Some topics in homogenization

by

Charles Manson

Thesis

Submitted to the University of Warwick

in partial fulfilment of the requirements

for admission to the degree of

Doctor of Philosophy

Mathematics Institute

July 2010

THE UNIVERSITY OF
WARWICK

Contents

Acknowledgments

Thanks for to my mother and brother for all their love and support. Thanks to my supervisor, Martin Hairer for his help and encouragement. In addition I would like to thank Warwick University for financial support via a Warwick Postgraduate Fellowship.

Declarations

All work in this thesis unless otherwise stated is the work of the author.

Abstract

This thesis is mainly concerned with solving a new type of periodic homogenization problem. A solution of removing the Diophantine hypothesis on the homogenization problem where the interface sits at an irrational angle to the period is attempted but is not yet complete. As an aside an oscillator problem is analyzed using the corrector based approach of homogenization.

Abbreviations

We will use the following notation,

$\mathbb{R}$ the real numbers,

$\mathbb{N}$ the natural numbers,

$\mathbb{T}^d$ the d dimensional torus,

B a Brownian motion,

W a Brownian motion.

Chapter 1

Foreword

The aim of this thesis is to extend existing results on periodic homogenization. Fully periodic homogenization was initially investigated in the late 70s by Bensoussan, Lions and Papanicolaou [BLP78] and since then a large number of variations on the theme have emerged as have many applications in a diverse range of fields. There are a number of standard techniques which have led to the study of properties of solutions to certain PDEs, such as the solution to the Poisson equation, studied extensively in a series of papers. This particular equation is useful in the production of a ”corrector term”, which will be explained fully in due course.

Initially, the assumptions were quite rigid, such as the diffusion coefficient was assumed to be uniformly elliptic, and full periodicity was enforced, then ellipticity was relaxed to hypoellipticity. Then recently in a paper by my supervisor and Pardoux [HP08], the uniformly elliptic and hypoellipticity assumption was relaxed allowing the diffusion matrix to vanish even on an open set. There are myriad other ways in which the assumptions have been changed, or the techniques adapted to new fields such as materials science since the seminal book [BLP78].

There are a number of papers that relax the periodicity in one set of spatial variables and average over the periodic fast variables to produce a diffusion with ”homogenized” coefficients, that is, coefficients derived from the original coefficients, typically where the fast variables have been replaced by averaging a derived quantity with respect to the invariant measure (on the torus) of the fast process. Periodicity can also be relaxed with respect to the consideration of a reflected diffusion process (the solution to an SDE involving a local time term). These two approaches are even combined such as by Diakhaby and Ouknine in [DO06].

We allow two periodic regions where the drift is centered and periodic in each region with a thickened hyperplane interface region in between these two periodic regions. To show the weak convergence in one dimension to a rescaled skew Brownian motion with a different scaling factor on either side of the interface, which upon homogenization, has now reduced to a point (this only happens in one dimension, in general the interface becomes a hyperplane), we use a scheme first employed by Freidlin and Wentzell [FW93] to perform averaging on a graph.

The end result in homogenization is weak convergence in the space of probability distributions on a suitably chosen space, usually the space of continuous paths or cadlag paths (right continuous paths with left hand limits). This is usually achieved via a three or four step process of first showing "tightness" and then identifying the limit point, either by showing resolvent convergence or using the tool of the martingale problem by showing that the solution must then satisfy a particular martingale problem. If the martingale approach is utilized, then there is the extra step of showing that the solution of the martingale problem is unique. The final step, and a step not always taken or not relevant in some cases, is characterizing the diffusion process corresponding to a particular resolvent or particular martingale problem. This thesis does not deviate from the standard approach in taking these steps when performing homogenization and we use the martingale problem approach.

Chapter 2 is an introduction to the field in general, the main problem to be solved, and techniques required to solve it. The subsequent chapter 3 is the proof of the result in one dimension. Chapter 4 is the main result, the multidimensional result, which is followed by an attempt to homogenize in the reflected periodic setting where the interface sits at an angle to the period with an irrational tangent. In the final chapter 6 we study a homogenization inspired approach to a coupled oscillator chain problem.

Chapter 2

Introduction

2.1 The machinery of homogenization

There are several important concepts that perhaps should be explained before we proceed. At this point it is probably foremost to bear in mind that in homogenization, through the variation of a parameter (such as ε above), the object of study becomes a collection of probability distributions on an appropriate space, usually the space of continuous or cadlag functions from an interval of $\mathbb{R}$ to a suitably "nice" metric space. (In our case the space is that of continuous functions from $[0, \infty)$ to $\mathbb{R}^d$.) We will now embark on a quick review of the background theory to homogenization, the language of stochastic differential equations, semigroups and such like. The first formal concept to be introduced, therefore, is tightness of a family of probability distributions, establishment of which is usually the first step in a typical probabilistic homogenization proof.

Definition 2.1.1 (Tightness for a family of probability measures [Bil99]).
A family of probability measures, on a metric space (S, d), $\{\mathbb{P}_i\}_I \subseteq \mathscr{P}(S)$ ($\mathscr{P}(S)$ the set of probability measures on S), for an arbitrary (possibly uncountable) indexing set, is tight, if given $\varepsilon > 0$, there exists a compact set $K \subseteq S$ such that,

$$\inf_{i \in I} \mathbb{P}_i(K) \geq 1 - \varepsilon \ .$$

This is then combined, when (S, d) is complete and separable, with Prohorov's theorem. Prohorov's theorem gives the equivalence of tightness and relative compactness in the Prohorov metric on probability measures. The precise nature of the Prohorov metric is not important in this context, just that in the case S is separable, convergence in the Prohorv metric is equivalent to (in a general metric space this reduces to implies) weak convergence of probability measures [EK86, Theorem 3.1 Chapter 3]. It is weak convergence that is the sought after property in a lot of homogenization proofs and this is defined as follows,

Definition 2.1.2 (Weak convergence [Bil99])**.**
Given a metric space (S, d), let C_b be the space of bounded continuous functions with the supremum norm. A sequence $\{\mathbb{P}_n\}_n$ converges weakly to $\mathbb{P} \in \mathscr{P}(S)$, if

$$\lim_{n\to\infty} \int f \, d\mathbb{P}_n = \int f \, d\mathbb{P}, \quad \forall f \in C_b(S) \ .$$

Therefore, by showing tightness in a complete and separable metric space and applying the Prohorov theorem we have that any subsequence from our original collection of probability measures has a weakly convergent subsequence. The tightness of the family of measures is usually shown on the space $\mathbb{C}([0,\infty), \mathbb{R}^d)$ with the topology of uniform convergence on compact sets.

The second step is then to show that any weak limit of the family of probability distributions either satisfies a particular martingale problem or possesses a particular resolvent. It is probably prudent at this point to define both these quantities.

Definition 2.1.3 (Martingale problem)**.**
The definition of the solution to a martingale problem according to [EK86] is as follows.

Let $B(S)$ denote bounded measurable real-valued functions on the metric space S. Then given an operator $\mathcal{L} : B(S) \to B(S)$, a solution of the martingale problem for $\mathcal{L}$ is a measurable stochastic process X defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with values in S such that for $f, g \in B(S)$ with $f \in \mathscr{D}(\mathcal{L})$ (f in the domain of $\mathcal{L}$) and $g = \mathcal{L}f$, we have that

$$f(X(t)) - \int_0^t g(X(s)) \, ds \ ,$$

is a martingale with respect to the filtration

$$^{\star}\mathscr{F}_t^X = \mathscr{F}_t^X \vee \sigma\left(\int_0^s h(X(u)) : s \le t, h \in B(S) \right) ,$$

for $\mathscr{F}_t^X = \sigma(X(s) : s \le t)$, for a right continuous process ${}^\star\mathscr{F}_t^X = \mathscr{F}_t^X$ and in general the difference between the two consists only of null sets.

Since we are dealing with martingale problems in spaces of continuous paths we have ${}^\star\mathscr{F}_t^X = \mathscr{F}_t^X$ and hence the distinction between ${}^\star\mathscr{F}_t^X$ and $\mathscr{F}_t^X$is ignored from now on.

The concept of a resolvent belongs to the study of strongly continuous semigroups, so it is logical to define the resolvent in this setting and then to move onto how this definition relates to the particular situation of homogenization of the solutions to SDEs.

Definition 2.1.4 (Strongly continuous semigroup [EK86])**.**
A one parameter family $\{S(t) : t \ge 0\}$ of bounded linear operators on a Banach space B is called a semigroup if $S(0) = I$ and $S(t+s) = S(t)S(s)$ for all $s, t \ge 0$. A semigroup $\{S(t)\}$ on B is said to be strongly continuous if $\lim_{t\to 0} S(t)f = f$, for every $f \in B$.

The infinitesimal generator of a semigroup $\{S(t)\}$ on B is the linear operator A defined by

$$Af = \lim_{t\to 0} \frac{1}{t}\Big(S(t)f - f\Big)\,.$$

The domain $\mathscr{D}(A)$ of A is the subspace of all $f \in B$ for which the above limit exists.

It can be shown that the generator of a strongly continuous semigroup is closed, i.e. that the graph of the generator, considered as a subset of $B \times B$, is closed, and that $\mathscr{D}(A)$ is dense, see for instance [EK86, Corollary 1.6 Chapter 1].

Definition 2.1.5 (Resolvent [EK86])**.**
Given a closed operator A on the Banach space B, if, for a real λ, $\lambda I - A := \lambda - A$ is one-to-one, has a range equal to B and $(\lambda - A)^{-1}$ is a bounded linear operator on B, then λ is said to belong to the resolvent set of A and the bounded linear operator $(\lambda - A)^{-1}$ is called the resolvent of A.

In the particular case of an SDE with suitably nice coefficients (Lipschitz coefficients are sufficient), the semigroup is of course that of the solution considered as a Markov process. The semigroup on a closed subspace $M \subseteq B(S)$ corresponding to a Markov process is that given by,

$$\mathbb{E}\left[f(X(t+s))|\mathscr{F}_t^X\right] = \mathbb{E}_{X(t)}\left[f(X(s))\right] = \int_S P(s, X(t), dy) f(y) ,$$

for all $f \in M$, and where $P(s, X(t), dy)$ is the measure corresponding to the Markov transition kernel.

Where the Markov process is the solution of an SDE on $\mathbb{R}^d$, in order to obtain the generator we look to Itô's formula and take the expectation. For a real valued C^2 (twice continuously differentiable) function, f on $\mathbb{R}^n$, we have, from a repeated application of stochastic integration by parts followed by approximation argument [RY91], Itô's formula:

$$f(X(t)) = \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X(s))\, d\langle X_i, X_j\rangle(s) + \int_0^t \frac{\partial f}{\partial x_i}(X(s))\, dX_i(s) , \tag{2.1.1}$$

where the convention of summation over repeated indices (Einstein summation convention as it's otherwise known), is assumed to hold.

This finishes a brief introduction to the language of homogenization, in the next section we will move onto a review of some classic homogenization problems.

2.2 A Classic Problem

In this section we will begin to show how homogenization techniques are implemented in practice. We will begin with a classic problem that was solved in the early days in the book [BLP78].

Let B be a standard Wiener process and in the classic problem it is assumed that $b : \mathbb{R}^n \to \mathbb{R}^n$, smooth, and periodic. In the homogenization problems we will study later, $b : \mathbb{R}^n \to \mathbb{R}^n$, smooth, is periodic away from an 'interface' region $[-\eta, \eta]$, $\eta \geq 0$. More precisely, we will assume that there exist periodic functions $b_i : \mathbb{R} \to \mathbb{R}$, $i = \pm$, such that $b_i(x+1) = b_i(x)$ and such that $b(x) = b_+(x - \eta)$ for $x > \eta$ and $b(x) = b_-(x + \eta)$ for $x < -\eta$. We study the solution of the stochastic differential equation (SDE),

$$\frac{dX}{dt} = b(X(t)) + \frac{dB}{dt} ,$$

in particular the weak convergence of the solution under a diffusive rescaling. This SDE has a unique solution up to indistinguishability as the coefficients b and 1 are Lipschitz [RY91, Theorem 2.1 Chapter IX]. Two processes Y, Y' defined on the same probability space Ω are indistinguishable if for almost all $\omega \in \Omega$, $Y_t(\omega) = Y'_t(\omega)$ for all t [RY91].

A diffusive rescaling is given by $x \mapsto \varepsilon x$ and $t \mapsto \varepsilon^2 t$, exactly the ratio of powers that preserves the Brownian motion component of the motion. This rescaling gives the behavior of the process over large distances and large times. Applying the aforementioned rescaling, what is really the object of study is the family of solutions of the SDEs indexed by ε,

$$\frac{dX_x^\varepsilon}{dt} = \frac{1}{\varepsilon} b\left(\frac{X_x^\varepsilon}{\varepsilon}\right) + \frac{dB}{dt} ,$$
$$X_x^\varepsilon(0) = x . \tag{2.2.1}$$

The above rescaled equation is arrived at by ignoring the initial condition and considering the integrated form of the standard equation, given by, for,

$$X(t) = \int_0^t b(X_x(s))\, ds + B(t) ,$$

setting,

$$X^\varepsilon(t) = \varepsilon X\left(\frac{t}{\varepsilon^2}\right) .$$

We use the normal change of variables formula to exchange t for $\frac{t}{\varepsilon^2}$ in the integral with respect to dt on the RHS, i.e. set "new" s equals ε^2 times "old" s in the integrals and then end up with the rescaled equation except with $\varepsilon B\left(\frac{t}{\varepsilon^2}\right)$ as the second term on the RHS,

$$X^\varepsilon(t) = \int_0^t \frac{1}{\varepsilon} b\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) ds + \varepsilon B\left(\frac{t}{\varepsilon^2}\right) . \tag{2.2.2}$$

As mentioned in passing above though, this term is another Brownian motion courtesy of Levy's characterization theorem [RY91], which states,

Theorem 2.2.1. *[Levy's Characterization Theorem] For an $(\mathscr{F}_t)$-adapted continuous d-dimensional process X vanishing at 0, the following three conditions are equivalent: (i) X* *is an $\mathscr{F}_t$-Brownian motion; (ii) X is a continuous local martingale and $\langle X^i, X^j\rangle_t = \delta_{ij}t$ for every $1 \le i, j \le d$; (iii) X is a continuous local martingale and for every d-uple $f = (f_1, \ldots, f_d)$ of functions in $L^2(\mathbb{R}_+)$, the process*

$$\mathscr{E}_t^{ij} = \exp\left\{ i \sum_k \int_0^t f_k(s) dX_s^k + \frac{1}{2} \sum_k \int_0^t f_k^2(s)\, ds \right\} ,$$

is a complex martingale.

(Where $\langle \cdot, \cdot \rangle$ denotes quadratic variation.)

So using parts (i) and (ii) of Levy's characterization theorem, for the process $\varepsilon B(\varepsilon^{-2}t)$, this is, by definition of B, $\mathscr{F}_{\varepsilon^{-2}t}$ adapted (notice the change of filtration), vanishes at zero, continuous, and has $\langle \varepsilon B(\varepsilon^{-2}t), \varepsilon B(\varepsilon^{-2}t)\rangle = \varepsilon^2 \langle B(\varepsilon^{-2}t), B(\varepsilon^{-2}t)\rangle = \varepsilon^2 \varepsilon^{-2} t = t$. Therefore it is an $\mathscr{F}_{\varepsilon^{-2}t}$ Brownian motion.

In addition, in order for the scaling limit of the SDE (2.2.1) that we will consider, to exist, we require that b (or each of the two functions b_i, glued together to make b), satisfy a centering condition. The centering condition prevents the process from, intuitively speaking, "running off to infinity" by exhibiting an increasing overall drift in a particular direction as the parameter ε is reduced to zero. This condition is used in the classic, fully periodic case to ensure existence of a weak limit (the concept of a weak limit will be explained below, in the next section). The fact that a weak limit exists [BLP78] with the centering condition implies that violation of this condition results in a process that exhibits an increasing overall drift. Explicitly, the centering condition on b_i is given by,

$$\int_0^1 b_i(x) p_i(x)\, dx = 0 ,$$

where p_i denotes the density of the invariant measure for the solution to the SDE

$$\frac{dX}{dt} = b_i(X) + \frac{dB}{dt} , \tag{2.2.3}$$

for $i = \pm$, considered as a diffusion on the torus, which satisfies the Fokker Planck equation,

$$\mathcal{L}_i^\star p_i = 0 ,$$

for $\mathcal{L}$ the generator of (2.2.3) given by $b\partial_x + (1/2)\partial_x^2$ and $\mathcal{L}^\star$ its adjoint in a suitable space (the Sobolev space H^1 for instance) given by $\mathcal{L}^\star\phi = -\partial_x(b\phi) + (1/2)\partial_x^2\phi$. See [BLP78] for a proof of the existence of the density of the invariant measure as the unique solution of the Fokker Planck equation using the Fredholm alternative.

If $X(t)$ is the solution to the SDE on $\mathbb{R}^d$ given by the integral equations for $i = 1,\ldots, d$,

$$X^i(t) = X^i(0) + \sum_{j=1}^r \int_0^t g_{ij}\big(X(s)\big)\, dB^j(s) + \int_0^t h_i\big(X(s)\big)\, ds\,,$$

for g and h measurable functions taking values in the $d \times r$ matrices and d dimensional vectors respectively, and the B^j are r standard independent Brownian motions. Then applying Itô's formula gives, for $f \in C^2$, a real valued function,

$$\begin{aligned} f(X(t)) = & \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X(s)) g_{ik}\big(X(s)\big) g_{jk}\big(X(s)\big)\, ds \\ & + \int_0^t \frac{\partial f}{\partial x_i}(X(s)) h_i\big(X(s)\big)\, ds \\ & + \int_0^t \frac{\partial f}{\partial x_i}(X(s)) \left(\sum_j g_{ij}\big(X(s)\big)\, dB^j(s) \right). \end{aligned}$$

Looking at the RHS of the above expression, observe the semimartingale structure evident from the grouping of terms, the first two terms are of bounded variation and the last term, by the martingale property of the Itô integral, is a martingale. Hence, if the solution of the SDE were considered as a Markov process then the generator would be given by $g_{ik}g_{jk}\frac{\partial^2}{\partial x_i \partial x_j} + h_i \frac{\partial}{\partial x_i}$. Not all solutions to SDEs are Markov Processes but under basic assumptions, such as g and h Lipschitz, we have that the solution is a Markov Process, in fact with Lipschitz coefficients the solution is a Feller process [RY91, Theorem 2.5 Chapter IX].

Now we have introduced Itô's formula, we will show how it is used to define a corrector term that facilitates homogenization in a simple case, by considering the corrected process instead of the original process which is "close" to the original process but with a much nicer form of generator. The corrected process takes the form, $X^\varepsilon + \varepsilon g(\varepsilon^{-1}X^\varepsilon)$, and what we aim to do with the additional term is remove any terms of order ε^{-1} from the rescaled equation (2.2.2) where we assume that we are in the case of wholly periodic drift. In other words for the purposes of the following discussion we will assume that b is fully periodic and satisfies the centering condition. An application of Itô's formula then gives the appropriate g, in particular it shows that this is the case if g satisfies the Poisson equation for b, $\mathcal{L}g = -b$, componentwise. Note that we will frequently denote the generator of the semigroup corresponding to the solution of an SDE by $\mathcal{L}$, and in this case $\mathcal{L} = b.\bigtriangledown + \bigtriangledown.\bigtriangledown$. Applying Itô's formula to $X_x^\varepsilon + \varepsilon g(\varepsilon^{-1}X^\varepsilon)$ gives,

$$\begin{aligned} X_x^\varepsilon(t) &= x + \varepsilon g\left(\frac{x}{\varepsilon}\right) + \int_0^t \frac{1}{\varepsilon} b\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) ds + B(t) \\ &\quad + \int_0^t \frac{1}{\varepsilon} b_i\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) \frac{\partial g}{\partial x_i}\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) ds \\ &\quad + \int_0^t \frac{1}{2\varepsilon} \sum_i \frac{\partial^2 g}{\partial x_i \partial x_i}\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) ds \\ &\quad + \int_0^t \frac{\partial g}{\partial x_i}\left(\frac{X_x^\varepsilon}{\varepsilon}\right) dB_i(s) - \varepsilon g\left(\frac{X^\varepsilon}{\varepsilon}\right), \end{aligned}$$

where B_i denotes the components of the d dimensional Brownian motion B. Thus if $\mathcal{L}g = -b$, where $\mathcal{L} = b_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\triangle$ is the generator of the non-rescaled process, then the above equation becomes,

$$\begin{aligned} X_x^\varepsilon(t) + \varepsilon g\left(\frac{X^\varepsilon}{\varepsilon}\right) &= x + \varepsilon g\left(\frac{x}{\varepsilon}\right) + B(t) + \int_0^t \frac{\partial g}{\partial x_i}\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) dB_i(s) \\ &= x + \varepsilon g\left(\frac{x}{\varepsilon}\right) + \int_0^t \sum_i \left[1 + \frac{\partial g}{\partial x_i}\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right)\right] dB_i(s)\,. \end{aligned} \tag{2.2.4}$$

Notice that the term of order ε^{-1} is now gone. With one extra lemma, 2.2.3, we can finish the homogenization of this equation using the martingale central limit theorem [EK86] which is as follows,

Theorem 2.2.2 (Martingale Central Limit Theorem). *For $n = 1, 2, \ldots, n$, let $(\mathscr{F}_t^n)$ be a filtration and let M_n be an $\mathscr{F}_t^n$-local martingale with sample paths in $D_{\mathbb{R}^d}[0,\infty)$ (space of cadlag paths) and $M_n(0) = 0$. Let A_n^{ij} be a symmetric $d \times d$ matrix-valued process such that A_n^{ij} has sample paths in $D_{\mathbb{R}}[0,\infty)$ and $A_n(t) - A_n(s)$ is non-negative definite for $t > s \geq 0$. Assume one of the following conditions holds:*

(i) For each $T > 0$,

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\leq T} |M_n(t) - M_n(t-)|\right] = 0\,,$$

and

$$A_n^{ij} = \langle M_n^i, M_n^j \rangle \; .$$

(ii) For each $T > 0$ and $i, j = 1, 2, \ldots, d$,

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\le T} |A_n^{ij}(t) - A_n^{ij}(t-)| \right] = 0 \; ,$$
$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{t\le T} |M_n(t) - M_n(t-)|^2 \right] = 0 \; ,$$

and for $i, j = 1, 2, \ldots, d$,

$$M_n^i(t) M_n^j(t) - A^{ij}(t) \; ,$$

is an $\mathscr{F}_n$-local martingale.

Suppose that $C = c_{ij}$ is a continuous, symmetric, $d \times d$ matrix-valued function, defined on $[0, \infty)$, satisfying $C(0) = 0$ and

$$\sum_{ij} \big(c_{ij}(t) - c_{ij}(s)\big) \xi_i \xi_j \geq 0 \quad \xi \in \mathbb{R}^d, \quad t > s \geq 0 \; ,$$

and in addition, we have that for each $t \geq 0$ and $i, j = 1, 2, \ldots, d$

$$A_n^{ij}(t) \to c_{ij}(t) \; ,$$

in probability. Then $M_n \Rightarrow X$, where X is the process with independent Gaussian increments and covariance $c_{ij}(t)$. (The existence of which is ensured by [EK86, Theorem 1.1 Chapter 7].)

The lemma in question, remembering we are still in the fully periodic case for the purposes of this discussion, is (see for example [BMP07, Proposition 2.4]),

Lemma 2.2.3. *Let $f \in L^\infty(\mathbb{T}^d, \mathbb{R})$, then for any $t > 0$, we have the following convergence in probability,*

$$\int_0^t f\left(\frac{X^\varepsilon(s)}{\varepsilon} \right) ds \to t \int_{\mathbb{T}^d} f(x) \, \mu(dx) \; ,$$

where μ is the invariant measure of the non-rescaled process on the torus, as before.

The proof of this lemma is very similar to a lemma in the one dimensional case of the problem we will be studying, so to avoid repetition, will be omitted.

Looking at the form of the RHS of (2.2.4) it is quite clear how we will apply the martingale central limit theorem. Since the paths of the stochastic integral are continuous, both conditions (i) and (ii) from the martingale central limit theorem are satisfied and we will use the lemma above to show convergence in probability of the quadratic variation of the stochastic integral. The quadratic variation of the stochastic integral on the RHS of (2.2.4) is given by,

$$\begin{aligned}&\left\langle \int_0^t \sum_k \left[1+\frac{\partial g_i}{\partial x_k}\left(\frac{X^\varepsilon}{\varepsilon}\right)\right] dB_k(s), \int_0^t \sum_k \left[1+\frac{\partial g_j}{\partial x_k}\left(\frac{X^\varepsilon}{\varepsilon}\right)\right] dB_k(s) \right\rangle \\ &= t\int_{\mathbb{T}^d} \sum_k \left[1+\frac{\partial g_i}{\partial x_k}\left(\frac{X^\varepsilon(s)}{\varepsilon}\right)\right]\left[1+\frac{\partial g_j}{\partial x_k}\left(\frac{X^\varepsilon(s)}{\varepsilon}\right)\right] ds\ .\end{aligned}$$

Lemma 2.2.3 then gives the convergence in probability of the quadratic variation to,

$$\begin{aligned}&\int_0^t \sum_k \left[1+\frac{\partial g_i}{\partial x_k}\left(\frac{X^\varepsilon(s)}{\varepsilon}\right)\right]\left[1+\frac{\partial g_j}{\partial x_k}\left(\frac{X^\varepsilon(s)}{\varepsilon}\right)\right] ds \\ &\to t\int_{\mathbb{T}^d} \sum_k \left[1+\frac{\partial g_i}{\partial x_k}(x)\right]\left[1+\frac{\partial g_j}{\partial x_k}(x)\right] \mu(dx) = c^{ij}(t)\ ,\end{aligned}$$

and then the martingale central limit theorem gives the weak convergence of the stochastic integral term to the process with independent Gaussian increments given by $c^{ij}(t)$, which by an application of the Levy characterization theorem quoted above componentwise, is a Brownian motion with a homogenized diffusion coefficient in front of it. Then once it has been noted that addition of the term $\varepsilon g(\varepsilon^{-1}X_x^\varepsilon)$ (using the Wasserstein metric) does not affect the weak limit, we have the weak convergence of the process to the aforementioned limit.

To see that the addition of the term $\varepsilon g(\varepsilon^{-1}X^\varepsilon)$ and in fact any term that is bounded above by a term that tends to zero as $\varepsilon \to 0$ does not affect weak convergence we will use the Wasserstein metric [Vil03]. This metrizes weak convergence on a Polish space of bounded diameter, which is what $\mathbb{C}([0,\infty),\mathbb{R}^d)$ is; uniform convergence on compact sets can be metrized using the metric on $\mathbb{C}([0,\infty),\mathbb{R}^d)$, D, given by,

$$D(\omega, \omega') = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\sup_{0 \le t \le n} |\omega(t) - \omega'(t)|}{1 + \sup_{0 \le t \le n} |\omega(t) - \omega'(t)|} .$$

This then gives that the corrected process will tend to the same weak limit as the uncorrected process which is the rationale behind correcting in this fashion. The Wasserstein metric is defined as follows,

Definition 2.2.4 (Wasserstein Metric)**.**
The Wasserstein metric is defined by, for μ_1, μ_2 measures on a space Ω, the corresponding distance between them is given by,

$$\inf_{(\mu_1, \mu_2) \in \Gamma} \int |x_1 - x_2| \, d(\mu_1, \mu_2) ,$$

where Γ is the set of all couplings of the measures μ_1 and μ_2. That is measures on $\Omega \times \Omega$ that have projection onto the first component μ_1, and onto the second component, μ_2.

Consider the family of couplings indexed by ε given by the push forward onto $\mathbb{C}([0,\infty), \mathbb{R}^d) \times \mathbb{C}([0,\infty), \mathbb{R}^d)$ of μ_1, the law of X_x^ε, by the map $x \mapsto (x, x + \varepsilon g(\varepsilon^{-1} x))$. The convergence of the two families of measures given by the corrected and non-corrected processes to the same limit, considered as probability measures on the space $\mathbb{C}([0,\infty), \mathbb{R}^d)$, is seen by taking this family of couplings indexed by ε. Then the Wasserstein distance between the limiting point of the family of corrected processes and the distribution of the non-corrected process as $\varepsilon \to 0$ is less than,

$$\int D\left(x_1, x_1 + \varepsilon g\left(\frac{x_1}{\varepsilon}\right)\right) d\mu_1 \le \varepsilon \|g\|_\infty ,$$

which tends to zero as $\varepsilon \to 0$. Hence the measures given by the processes considered as probability measures on the space $\mathbb{C}([0,\infty), \mathbb{R}^d)$, converge weakly as $\varepsilon \to 0$.

This is known as periodic homogenization, albeit a simple case, and the manner of usage of the corrector term is similar to that used for the problems to be solved in this thesis. In the next subsection we will introduce the slightly more complicated equations that have in practice been homogenized since the very beginning of the theory, and this will be the starting point for a review of some of the research into this theory.

The "toy" example of the last subsection was a simplification of the problems investigated in fully periodic homogenization problems. The SDE to be homogenized usually has an extra drift term and the diffusion term is no longer a constant equal to the identity. Such an SDE would for example be given by, in integrated form,

$$X_x^\varepsilon(t) = x + \int_0^t \frac{1}{\varepsilon} b\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) ds + \int_0^t c\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) ds + \int_0^t \sigma\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) dB(s),$$

where $b, c : \mathbb{R}^d \to \mathbb{R}^d$ are smooth vector valued functions and $\sigma : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ is a smooth matrix valued function. Note that because of the extra term in the drift, this equation does not come from a rescaling as the more simple example of the previous chapter did. Without the extra drift term it would correspond again to a rescaling.

In addition to an extra term and additional complexity of the diffusion term, there is also in many homogenization problems the consideration of fast and slow variables. An example of this is known as locally periodic homogenization. In the case of locally periodic homogenization where there are both fast and slow variables, the drift is for example,

$$\varepsilon^{-1} b(X_x^\varepsilon(s), \varepsilon^{-1} X_x^\varepsilon(s)) \ .$$

The variable $X_x^\varepsilon(s)$ is termed the slow variable whereas the variable $\varepsilon^{-1} X_x^\varepsilon(s)$ is the fast variable. Obviously they are related, but the terminology comes from the general strategy used to tackle such problems, namely treat the slow and fast variables as independent variables, and through the use of correctors and bounding inequalities for the results, contrive a situation where averaging with respect to the fast variables can be conducted. The full SDE in a typical locally periodic homogenization problem is of the form,

$$\begin{aligned} X_x^\varepsilon(t) \ &= \ x + \int_0^t \frac{1}{\varepsilon} b\big(X_x^\varepsilon(s), \varepsilon^{-1} X_x^\varepsilon(s)\big) + c\big(X_x^\varepsilon(s), \varepsilon^{-1} X_x^\varepsilon(s)\big)\, ds \\ &\quad + \int_0^t \sigma\big(X_x^\varepsilon, \varepsilon^{-1} X_x^\varepsilon(s)\big)\, dB(s). \end{aligned} \tag{2.2.5}$$

Apart from the "standard" assumptions of boundedness and differentiability with bounded derivatives on the coefficients, the hypotheses on the drift in this case are that, for fixed x, $b(x,y)$ is periodic in y and centered, and $c(x,y)$, $\sigma(x,y)$ are periodic in y. The regularity assumptions are made to avoid too many technical details that detract from a presentation of the overall scheme. The concept of centering in this situation clearly differs slightly from that considered above since the process is no longer a diffusion on the torus. The operator $\mathcal{L}_{x,y}$ on $\mathbb{R}^{2d}$, obtained as the generator of the diffusion process given in (2.2.5), except pretending the fast and slow variables are now distinct spatial variables, for fixed x, is the generator of a diffusion process on $\mathbb{R}^d$ that can be considered as the generator of a diffusion process on the torus. For fixed x, the drift $b(x,y)$ is centered with respect to $m(x,dy)$ the invariant measure of the aforementioned diffusion on the torus. In what follows we assume smoothness to investigate the general principles and avoid a lot of the details that arise when regularity assumptions are much looser. For example in [BLP78] the coefficients are assumed to be, for the second variable fixed, twice continuously differentiable in the first variable but in [BMP07] this assumption is relaxed and although the basic structure of the proof is the same, the latter is significantly heavier in technical detail.

In both of these cases, the use of a corrector term does not completely remove the drift, but it does remove the term of order ε^{-1} from the equation and the remaining terms can be homogenized by a similar lemma to that used to show the convergence in probability of the quadratic variation term in the simple case. We will first deal with the simpler situation of periodic homogenization with the extra drift term c as in [BLP78], then move onto outlining how a locally periodic homogenization result would proceed. Applying Itô's formula to $X_x^\varepsilon(t) + \varepsilon g(\varepsilon^{-1} X_x^\varepsilon(t))$ (the corrector term which we refer to as g is also commonly referred to as $\hat{b}$),

$$\begin{aligned} X_x^\varepsilon(t) + \varepsilon g\left(\frac{X_x^\varepsilon(t)}{\varepsilon}\right) &= x + \varepsilon g\left(\frac{x}{\varepsilon}\right) + \int_0^t (I + \nabla g) c\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) ds \\ &\quad + \int_0^t (I + \nabla g)\sigma\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) dW(s)\ , \end{aligned}$$

since $\mathcal{L}g + b = 0$. Then instead of just the initial term plus a martingale term to demonstrate weak convergence of, we have an additional integral term of bounded variation. Remembering how convergence in probability of the quadratic variation of the stochastic integral term was achieved, using the lemma (2.2.3) that gave the convergence in probability of the quadratic variation term, we therefore also have convergence in probability of the integral with respect to ds. Convergence in probability implies weak convergence, hence this lemma gives,

$$\int_0^t (I+\nabla g)c\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) ds \Rightarrow t\int_{\mathbb{T}^d}(I+\nabla g)c(y)\,\mu(dy)\;.$$

Therefore putting all parts together, we have the weak convergence of $X_x^\varepsilon(t)$ to the limit,

$$\begin{aligned} &x + t\int_{\mathbb{T}^d}(I+\nabla g)c(y)\,\mu(dy)\\ &+\left[\int_{\mathbb{T}^d}(I+\nabla g)\sigma\sigma^T(I+\nabla g)^T(y)\,\mu(dy)\right]^{\frac{1}{2}} d\tilde{B}_s, \end{aligned}$$

where the exponent of the homogenized quadratic variation term denotes the unique symmetric $d\times d$ matrix that when composed with its transpose will give this quantity and $\tilde{B}$ denotes a d dimensional Brownian motion.

Moving on now to locally periodic homogenization as in [BLP78]. This is far more complicated than the cases considered previously due to the presence of the process to be homogenized as an order 1 input term in the drift and diffusion coefficients. Instead of just using a corrector and showing convergence in probability of the SDE termwise, the procedure of homogenization at least on a general level is more akin to the method we will employ. First tightness for the family of probability distributions on $\mathbb{C}([0,\infty),\mathbb{R}^d)$ is shown and then the limit point is identified. This is a departure from the method of homogenization employed in the previous simpler case. Originally in the proof of the simpler homogenization results outlined above in [BLP78], the proof did have a verification of tightness followed by an argument to identify the limit, following the same general method as we will outline in the case of locally periodic homogenization. Practically speaking, tightness of the family $\{\mathbb{P}^\varepsilon\}_\varepsilon$ is often shown by verifying that the following conditions hold, firstly that, given $T<\infty$, $\eta>0$,

$$\lim_{\delta\searrow 0}\sup_{\varepsilon}\mathbb{P}^\varepsilon\left[\sup_{t_2-t_1<\delta,\,0\le t_1<t_2\le T}|x(t_2)-x(t_1)|<\eta\right]=0\,,$$

where we denote the evaluation map by $x(\cdot)$. The second condition is that,

$$\lim_{M\to\infty} \sup_{\varepsilon} \mathbb{P}^{\varepsilon}\big[\, x(0) \geq M \,\big] = 0 \,.$$

The first of these is basically asserting the existence of a uniform modulus of continuity over the whole family in a probabilistic sense. The second condition is a measure of the uniform occupation by the initial conditions of a compact set, in a probabilistic sense. This second condition in much of the homogenization that will be outlined is trivial since we are seeking to demonstrate weak convergence of the family of solutions (viewed as probability distributions) starting from the same initial point. Therefore we only have to verify the first of these conditions with the supremum over ε replaced by $\limsup_{\varepsilon\to 0}$. In order to do this, a corrector term is used, in fact the same one as above but with an additional slow component, call the corrector g again, which is therefore a smooth periodic vector-valued function satisfying,

$$\mathcal{L}_{x,y} g(x,y) = -b(x,y), \quad \int_{\mathbb{T}^d} g(x,y)\, \mu(x,dy) = 0 \,.$$

Where $\mathcal{L}_{x,y}$ is the differential operator in y given by,

$$\mathcal{L}_{x,y} = b_i(x,y)\frac{\partial}{\partial y_i} + \frac{1}{2} a_{i,j}(x,y)\frac{\partial^2}{\partial y_i \partial y_j} \,,$$

and $m(x,dy)$ is the invariant measure of the process generated on the torus by $\mathcal{L}_{x,y}$.

We then apply Itô's formula as standard to $X^{\varepsilon}_x + \varepsilon g(X^{\varepsilon}_x, \varepsilon^{-1} X^{\varepsilon}_x)$, to give, for $a = \sigma\sigma^T$ as is common notation,

$$\begin{aligned}
X^{\varepsilon}_x(t) &+ \varepsilon g\Big(X^{\varepsilon}_x(t), \frac{X^{\varepsilon}_x(t)}{\varepsilon}\Big) \\
&= x + \varepsilon g\Big(x, \frac{x}{\varepsilon}\Big) + \int_0^t \frac{1}{\varepsilon} b + c \, ds \\
&\quad + \int_0^t \sigma \, dB(s) + \int_0^t (\nabla_x g + \frac{1}{\varepsilon}\nabla_y g) b \, ds \\
&\quad + \int_0^t (\varepsilon \nabla_x g + \nabla_y g) c \, ds + \int_0^t (\varepsilon \nabla_x g + \nabla_y g)\sigma \, dB(s) \\
&\quad + \int_0^t \Big[\varepsilon tr(a H_x(g)) + \frac{1}{\varepsilon} tr(a H_y(g)) \\
&\qquad\quad + tr\big(a(\nabla_x \nabla_y g + \nabla_y \nabla_x g)\big)\Big] \, ds \,.
\end{aligned}$$

Where $H_x(g)$ and $H_y(g)$ are the matrices of second order derivatives of g with respect to x and y respectively. Now, the terms of order $\frac{1}{\varepsilon}$ cancel since $\mathcal{L}_{x,y}g = -b$. In fact in [BLP78] they show tightness for the non-corrected process by adding the corrector term and its expansion from the Itô formula on the same side of the equality. In keeping with the what follows we have done the opposite. Now, to verify tightness we look at $\sup_{|t_1-t_2|<\delta} |X_x^\varepsilon(t_1) - X_x^\varepsilon(t_2)|$. Since the integrands in the expression above with respect to ds are bounded due to the assumption of bounded derivatives, for $|t_1 - t_2| < \delta$, and $\varepsilon < 1$, these integrals are bounded by $C_1\delta$. Since the only other non-fixed quantities are the stochastic integral terms. The integrand in the quadratic variation of the stochastic integral term (collecting them into one term) is bounded for $\varepsilon < 1$, by C say, hence the quadratic variation up to time T is bounded by CT. Then to produce the same type of inequality as that produced in [BLP78], we use Chebychev's inequality which is frequently used in bounding arguments on the stochastic integral term and note the bounds on the integrals of bounded variation. Chebychev's inequality is as follows,

Definition 2.2.5 (Chebychev's Inequality)**.**
For any positive random variable Y and $k > 0$, we have that,

$$\mathbb{P}[Y > k] \leq \frac{\mathbb{E}[Y]}{k} \,.$$

There are many forms of this inequality, but this is the one we will use most frequently and is very general in its formulation.

So using this inequality, and denoting the stochastic integral term by $S(t)$, we have that, for δ sufficiently small,

$$\begin{aligned}
&\mathbb{P}\Big[\sup_{|t_1-t_2|<\delta,\, t_1,\, t_2<T} |X_x^\varepsilon(t_1) - X_x^\varepsilon(t_2)| \geq \eta\Big] \\
\leq\,&\mathbb{P}\Big[\sup_{|t_1-t_2|<\delta,\, t_1,\, t_2<T} |S(t_1) - S(t_2)| \geq \eta - C_1\delta\Big] \\
=\,&\mathbb{P}\Big[\sup_{|t_1-t_2|<\delta,\, t_1,\, t_2<T} |S(t_1) - S(t_2)|^2 \geq (\eta - C_1\delta)^2\Big] \\
\leq\,&\frac{CT\delta}{(\eta - C_1\delta)^2} \,.
\end{aligned}$$

Where Chebychev's inequality is used in the last inequality. We then clearly have,

$$\lim_{\delta\searrow 0}\limsup_{\varepsilon\to 0}\mathbb{P}\Big[\sup_{t_2-t_1<\delta,\,0\le t_1<t_2\le T}|X^\varepsilon_x(t_2)-X^\varepsilon_x(t_1)|\ge\eta\Big]=0\ ,$$

and hence tightness.

So now the second (and last) stage in a large number of homogenization proofs, but usually the most difficult, is identifying any limit point of the family of probability measure as one particular probability measure for all convergent subsequences. The rationale behind this is as follows, by Prohorov's theorem, tightness is equivalent to weak relative compactness, i.e. for any subsequence of measures in a tight family of probability measures, there is a further subsequence that converges weakly to a probability measure. Thus, if it is possible to identify the limit of any convergent subsequence as a particular measure then this must be the limit point of any subsequence and therefore the unique limit point of the entire family. In a proof of locally periodic homogenization or in fact any homogenization, a common type of lemma to see is the following where L^2 convergence is commonly substituted for convergence in probability as in the corresponding result for fully periodic homogenization investigated earlier, or even convergence in a Sobolev space. This is because as the diffusion part of the limiting process is no longer Gaussian but a variable stochastic integral with respect to a Gaussian process (Brownian motion) so we are no longer using the Martingale Central Limit Theorem and instead just showing the convergence of the appropriate conditional expectations for which we only need L^1 convergence. The stochastic part of the limiting process is no longer a Gaussian since the dependence of the diffusion coefficient on the slow variable remains.

Lemma 2.2.6. *Let $h(x,y)$ be a function which is C^2 in x, C^1 in y, with all partial derivatives continuous and bounded globally in x,y. Suppose also that h is periodic in y for fixed x and satisfies,*

$$\int_{\mathbb{T}^d}h(x,y)\,\mu(x,dy)=0\ ,\quad\forall x\ .$$

Then we have that,

$$\mathbb{E}\bigg[\bigg(\int_s^t h\Big(X^\varepsilon(u),\frac{X^\varepsilon(u)}{\varepsilon}\Big)\,du\bigg)^2\bigg|\mathscr{F}_s\bigg]\to 0\ . \tag{2.2.6}$$

Proof. The following method of proof is quite common for showing such convergence. Take the periodic solution of

$$\mathcal{L}_{x,y}\tilde{h} = h \ .$$

Then we apply Itô's formula to,

$$\begin{aligned}
&\varepsilon^2\tilde{h}\left(X^\varepsilon(t), \frac{X^\varepsilon(t)}{\varepsilon}\right) - \varepsilon^2\tilde{h}\left(X^\varepsilon(s), \frac{X^\varepsilon(s)}{\varepsilon}\right) \\
&= \int_s^t (\varepsilon\nabla_x\tilde{h} + \nabla_y\tilde{h})b\, ds + \int_s^t (\varepsilon^2\nabla_x\tilde{h} + \varepsilon\nabla_y\tilde{h})c\, ds \\
&\quad + \int_s^t (\varepsilon^2\nabla_x\tilde{h} + \varepsilon\nabla_y\tilde{h})\sigma\, dB(s) \\
&\quad + \int_s^t \varepsilon^2 tr(aH_x(\tilde{h})) + tr(aH_y(\tilde{h})) + \varepsilon tr\big(a(\nabla_x\nabla_y\tilde{h} + \nabla_y\nabla_x\tilde{h})\big)\, ds \ .
\end{aligned}$$

Upon doing so, we then note that the only constant order terms are $\mathcal{L}_{x,y}\tilde{h} = h$. For bounded t, s, the rest of the integral terms multiplied by ε to some positive power are bounded thus we have,

$$\mathbb{E}\left[\left(\int_s^t h\left(X^\varepsilon(u), \frac{X^\varepsilon(u)}{\varepsilon}\right) du\right)^2 \middle| \mathscr{F}_s\right] = (O(\varepsilon) + O(\varepsilon^2))^2 = O(\varepsilon^2)$$

Thus showing the required L^2 convergence. □

The next step in identifying the limit is to adopt the same approach as [BLP78] and use the tool of the martingale problem mentioned above with the corrected process. Denote $X^\varepsilon_x(t) + \varepsilon g(X^\varepsilon_x(t), \varepsilon^{-1}X^\varepsilon_x(t))$ by $Z^\varepsilon(s)$ for ease of notation. Let $\phi \in \mathbb{C}^\infty_0(\mathbb{R}^d)$ (the space of smooth real valued functions on $\mathbb{R}^d$ with compact support), and using Itô's formula with the semimartingale expression for the corrected process derived by Itô's formula above gives for $\phi(Z^\varepsilon)$,

$$\begin{aligned}
&\phi\big(Z^\varepsilon(s)\big) + \int_s^t \Big[\nabla_x\phi\big(Z^\varepsilon(s)\big)\Big(c + \nabla_x g\, b + \varepsilon\nabla_x g\, c + \nabla_y g\, c \\
&\qquad\qquad + \varepsilon tr(aH_x(g)) + tr(a(\nabla_x\nabla_y g + \nabla_y\nabla_x g))\Big)\Big]\, du \\
&\quad + \int_s^t \nabla_x\phi(Z^\varepsilon(s))\Big(\sigma + \nabla_y g\sigma + \varepsilon\nabla_x g\sigma\Big)\, dB(u) \\
&\quad + \int_s^t tr\Big(H(\phi)\big((\sigma + \nabla_y g\sigma + \varepsilon\nabla_x g\sigma)(\sigma + \nabla_y g\sigma + \varepsilon\nabla_x g\sigma)^T\big)\, du \ .
\end{aligned}$$

Now we introduce the averaged coefficients (vector-valued and matrix-valued, respectively), which come from the constant order terms in the expression above, that will be coefficients of the generator of the weak limit point of the family to be homogenized,

$$\begin{aligned} r(x) &= \int_{\mathbb{T}^d} c(x,y) + \nabla_x g(x,y) b(x,y) \\ &\quad + \nabla_y g(x,y) c(x,y) + 2tr(a(x,y) \nabla_x \nabla_y g(x,y))\, \mu(dy) \\ &= \int_{\mathbb{T}^d} r(x,y)\, \mu(dy)\,, \\ q(x) &= \int_{\mathbb{T}^d} (I + \nabla_y g(x,y)) a(x,y) (I + \nabla_y g(x,y))^T\, \mu(dy) \\ &= \int_{\mathbb{T}^d} q(x,y)\, \mu(dy)\,, \end{aligned}$$

for $r(x,y)$ and $q(x,y)$ the two integrands. So what we do now is take $\phi(X_x^\varepsilon(s) + \varepsilon g(X_x^\varepsilon(s), \varepsilon^{-1} X_x^\varepsilon(s)))$ and subtract the generator of the process given by $r(x)$ and $q(x)$ above applied to ϕ and take conditional expectation of both sides, apply Lemma 2.2.6 and then take expectation of both sides. Since all derivatives are bounded we can make the substitution of $X^\varepsilon(t)$ for $Z^\varepsilon(t)$ in all the terms of finite variation at a cost of $K\varepsilon$ in total, for some constant K,

$$\begin{aligned} &\mathbb{E}_x \Bigg[\phi(X^\varepsilon(t)) - \phi(X^\varepsilon(s)) - \int_s^t \nabla\phi\, r(X^\varepsilon(u))\, du \\ &\qquad - \int_s^t H(\phi) q(X^\varepsilon(u))\, du \,\Bigg|\, \mathscr{F}_s \Bigg] \\ &= \mathbb{E}_x \Bigg[\int_s^t \nabla\phi \left(r(X^\varepsilon(u), \varepsilon^{-1} X^\varepsilon(u)) - r(X^\varepsilon(u)) \right) du \\ &\qquad + \int_s^t H(\phi).\left(q(X^\varepsilon(u), \varepsilon^{-1} X^\varepsilon(u)) - q(X^\varepsilon(u)) \right) du + K'\varepsilon \,\Bigg|\, \mathscr{F}_s \Bigg] \end{aligned} \tag{2.2.7}$$

since all the stochastic integral terms vanish after taking conditional expectation with respect to $\mathscr{F}_s$ (which is traditionally what stochastic integral terms of the form $\int_s^t \ldots\, dM(u)$, for M a martingale, do when expectation with respect to $\mathscr{F}_s$ is taken under appropriate bounding assumptions). Taking the expectation of the modulus of both sides of (2.2.7) and then using the Cauchy Schwarz inequality together with (2.2.6) and Jensen's inequality on the right hand side gives that the left hand side tends to zero in L^1 as $\varepsilon \to 0$. We have convergence of probability distributions weakly at least along a subsequence. Thus, taking the limit of (2.2.7) along such a subsequence, we can put the limit into the LHS of (2.2.7) and zero on the RHS, which gives that any limiting measure in the sense of weak convergence is the solution to the martingale problem given by the operator $\bar{\mathcal{L}}$,

$$\bar{\mathcal{L}} = r(x).\nabla + q(x).\nabla\nabla \tag{2.2.8}$$

It is at this point that things get slightly more complicated than the previous example since we no longer have the martingale central limit theorem to do all the work for us. Although we now know that any weak limit point of the family of measure indexed by ε must satisfy the martingale problem given by (2.2.8), we do not know that the solution to this martingale problem started from a particular point is unique. In other words we still do not know that the family of probability measures converges to any specific measure as yet. If we could show that there is only one solution to the martingale problem given a particular initial point then we would have done just that. Looking at the form of the limiting solution of the martingale problem given in (2.2.8) it is clear to see this is of the same form as martingale problems solved by the solutions to SDEs and a theorem from [SV69] is used to show uniqueness of the solutions to such martingale problems corresponding to the solution of SDEs.

2.3 Associated benefits of homogenizing SDEs

The probabilistic homogenization of an SDE leads to homogenization in an appropriate sense of the Dirichlet problem. For instance in conducting the locally periodic homogenization discussed above, Bensoussan, Lions and Papanicolaou [BLP78] were not just studying homogenization of SDEs as a problem in itself, once you have homogenized an SDE like the one just previously, you get for free the limiting solution to the set of Dirichlet problems, for continuous f,

$$-a_{ij}\Big(x,\frac{x}{\varepsilon}\Big)\frac{\partial u_\varepsilon}{\partial x_{ij}} - \frac{1}{\varepsilon}b_i\Big(x,\frac{x}{\varepsilon}\Big)\frac{\partial u_\varepsilon}{\partial x_i} - c\Big(x,\frac{x}{\varepsilon}\Big)\frac{\partial u_\varepsilon}{\partial x_i} + a_0\Big(x,\frac{x}{\varepsilon}\Big)u_\varepsilon = f(x),$$
$$f|_\Gamma \equiv 0\ ,$$

as $\varepsilon \to 0$. The Dirichlet problem is considered in a domain $\mathscr{O}$ with boundary Γ, and we have $u_\varepsilon \in W^{2,p}$ for $p > 2$. We also have $a_0 \geq \alpha_0 > 0$ for α_0 a constant. The connection between the SDE and the solution to the Dirichlet problem is given by the explicit formula for the solution to the Dirichlet problem,

$$\mathbb{E}_x\left[\int_0^{\tau^\varepsilon_{\mathscr{O}}} f(X^\varepsilon(t))\exp\left(-\int_0^t a_0\left(X^\varepsilon(s),\frac{X^\varepsilon(s)}{\varepsilon}\right)ds\right)dt\right], \tag{2.3.1}$$

where $\tau^\varepsilon_{\mathscr{O}}$ is the escape time from the domain $\mathscr{O}$ by the process X^ε. Then as a result of the weak convergence of the laws of the stochastic differential equations corresponding to X^ε we have $u^\varepsilon(x) \to u(x)$ for all $x \in \mathscr{O}$ as $\varepsilon \to 0$ where $u(x)$ is the solution to the Dirichlet problem,

$$-q_{ij}(x)\frac{\partial u}{\partial x_{ij}} - r_i(x)\frac{\partial u}{\partial x_i} + \bar{a}_0(x)u = f(x), \quad f|_\Gamma \equiv 0,$$

for,

$$\bar{a}_0(x) = \int_{\mathbb{T}^d} a(x,y)\mu(x,dy),$$

The problem term in the expression (2.3.1) is the order ε^{-1} argument of a_0 in the expression (2.3.1), since the most obvious route to the convergence is to apply the weak convergence that has just been outlined to get the pointwise convergence of solutions. We get round this by swapping a_0 for its averaged version and demonstrating that the cost of doing so tends to zero as $\varepsilon \to 0$. We have,

$$\begin{aligned}&\Bigg|\mathbb{E}_x\left[\int_0^{\tau^\varepsilon_{\mathscr{O}}} f(X^\varepsilon(t))\exp\left(-\int_0^t a_0\left(X^\varepsilon(s),\frac{X^\varepsilon(s)}{\varepsilon}\right)ds\right)dt\right]\\
&\quad - \mathbb{E}_x\left[\int_0^{\tau^\varepsilon_{\mathscr{O}}} f(X^\varepsilon(t))\exp\left(-\int_0^t \bar{a}_0(X^\varepsilon(s))\,ds\right)dt\right]\Bigg|\\
&\leq C\mathbb{E}_x\left[\int_0^\infty\left|\exp\left(-\int_0^t a_0\left(X^\varepsilon(s),\frac{X^\varepsilon(s)}{\varepsilon}\right)ds\right)\right.\right.\\
&\qquad\qquad \left.\left. - \exp\left(-\int_0^t \bar{a}_0(X^\varepsilon(s))\,ds\right)\right|dt\right],\end{aligned}$$

for C bounding the supremum norm of f on $\mathscr{O}$. With the hypothesis of strict positivity on a_0, we have the convergence of the RHS of this expression to zero by Lemma 2.2.6. By the positivity of a_0, the second term on the LHS of the above expression is a continuous bounded function on the space of continuous paths, hence tends to, by weak convergence of X^ε_x,

$$\mathbb{E}_x\left[\int_0^{\tau^\varepsilon_{\mathscr{O}}} f(\bar{X}(t))\exp\left(-\int_0^t \bar{a}_0(\bar{X}(s))\,ds\right)dt\right],$$

denoting the solution to the martingale problem corresponding to the differential operator $\bar{\mathcal{L}}$, by $\bar{X}$, i.e. $\bar{X}_x$ is the weak limit point of the family X^ε_x.

2.4 Homogenization in random media

The remaining large field in probabilistic homogenization is known as homogenization in random media. It can be briefly summarized as follows. A particularly accessible introduction can be found in [Oll94]. The original, seminal text in this field is by Papanicolaou and Varadhan [PV81].

For homogenization in random media the typical set up [Rho09a] is that we have a random medium given by a probability space $(\Omega, \mathcal{G}, \mu)$ and a group of measure preserving $\{\tau_x : x \in \mathbb{R}^d\}$ transformations acting ergodically on Ω:

- $\forall A \in \mathcal{G}, x \in \mathbb{R}^d, \mu(\tau_x(A)) = \mu(A)$,
- If for any $x \in \mathbb{R}^d$, $\tau_x(A) = A$, then $\mu(A) = 1$ or 0,
- For any measurable function g on $(\Omega, \mathcal{G}, \mu)$, the function $(x, \omega) \to g(\tau_x \omega)$ is a measurable function on $(\mathbb{R}^d \times \Omega, \mathcal{B}(\mathbb{R}^d) \times \mathcal{G})$.

Then there are a number of different avenues to explore in terms of how this translates into a homogenization problem, the classic case is explored in [PV81], given by the diffusion corresponding to the martingale problem with uniformly elliptic operator,

$$\frac{1}{2}\partial_i(a(\tau_{-\frac{x}{\varepsilon}}\omega)\partial_j) .$$

Although most of this paper is approaching the problem from the direction of the convergence of the corresponding PDE. An example of another quite classic problem can be found in [Oll94], where the problem under consideration is that of the homogenization of SDEs of the form,

$$X(t) = \int_0^t -DV(\tau_{-X(s)}(\omega))\, ds + W . \tag{2.4.1}$$

W is a Brownian motion independent of the random medium and $V : \Omega \to \mathbb{R}$ is a random potential, and $DV(\omega)$ is then defined as the differential in x of the function $x \mapsto V(\tau_{-x}(\omega))$. It is assumed that V and DV are bounded and in addition, stochastic continuity can be shown from the properties of the random medium above.

Remark 2.4.1. In private communication with Rémi Rhodes it was brought to my attention that the three properties of the random medium given above, by a proof constructed by himself and F. Delarue, together imply stochastic continuity,

Definition 2.4.2 (Stochastic continuity). For $T_h f(\omega) = f(\tau_{-h}\omega)$,

$$\lim_{h\to 0} \mu(|T_h f(\omega) - f(\omega)| \geq \delta) = 0, \quad \forall \delta > 0, f \in L^2(\mu)$$

Stochastic continuity is used to give density of the domain of the infinitesimal generator D, of T_x in $L^2(\mu)$ from [EK86, Corollary 1.1.6] as the generator of a strongly continuous semigroup. Ultimately the density of the domain is used to show convergence of the resolvent equation via the spectral theorem that will be our surrogate corrector in the random environment case. We will make this precise shortly.

Under these hypotheses, as one would expect, the question is that of the weak convergence of $\varepsilon X(\varepsilon^{-2}t)$ as $\varepsilon \to 0$.

An extension of the classic problem explored in this article [Oll94] is that of a particle moving in a free divergence random field,

$$X(t) = \int_0^t F(\tau_{X(s)}\omega)\, ds + W(t) \ ,$$

for W as above and $F = D.H$ for all anti-symmetric matrix H. In the case of a free divergence random field there is no longer reversibility to simplify the analysis slightly, as there is in the case of (2.4.1), as far as the weak convergence of $\varepsilon X(\varepsilon^{-2}t)$.

Extending this scenario in the manner in which periodic homogenization was extended, results in situations akin to those explored above for periodic homogenization. Such homogenization problems have been explored in a number of papers by Rhodes. In [Rho09a], the locally periodic homogenization mentioned previously is extended to the random medium case. We have a family of SDEs of the form,

$$\begin{aligned} X^\varepsilon(t) =& x + \int_0^t b\left(\omega, \frac{X^\varepsilon(s)}{\varepsilon}, X^\varepsilon(s)\right) ds + \int_0^t c\left(\omega, \frac{X^\varepsilon(s)}{\varepsilon}, X^\varepsilon(s)\right) ds \\ &+ \int_0^t \sigma\left(\omega, \frac{X^\varepsilon(s)}{\varepsilon}, X^\varepsilon(s)\right) dB(s) \ , \end{aligned}$$

for B as above and ω from the random medium again. In addition we have that $b(\omega, \cdot, y)$, $c(\omega, \cdot, y)$ and $\sigma(\omega, \cdot, y)$ are stationary random fields i.e. they are of the standard form $f(\omega, x, y) = f(\tau_y \omega, x)$. The assumption of uniform ellipticity ($\Lambda I \leq \sigma\sigma \leq \Lambda^{-1} I$ for some $\Lambda > 0$) of the diffusion coefficient is made together with ability of the generator of X to be written in divergence form (for the matrix $a + H$, a symmetric, H antisymmetric in this case) and regularity assumptions on the coefficients. The potentially degenerate case is dealt with in another paper by Rhodes, [Rho09b]. The uniform ellipticity is used to provide local ergodicity of the process and also to provide uniform (over the random environment) (Aronson type) estimates on the density to give tightness. Lack of uniform ellipticity in [Rho09b] with regard to local ergodicity is countered with weak local ergodicity assumptions which are enforced through a reference matrix which controls the degeneracies of the diffusion coefficient.

Then in [Rho09c], we have a corresponding homogenization problem for a reflected diffusion process,

$$X^{\varepsilon}(t) = \int_0^t b(\tau_{X^{\varepsilon}(s)}(\omega))\, ds + \int_0^t \sigma(\tau_{X^{\varepsilon}(s)}(\omega))\, dB(s) \\ + \int_0^t \gamma(\tau_{X^{\varepsilon}(s)}(\omega))\, dL(t)^{\varepsilon},$$

where the situation is as before except L^{ε} is the local time of $X^{\varepsilon}(t)$ and $\gamma_1 = 1$ as is required for a reflected diffusion process and the matrix that is used to give the generator of the process in divergence form is symmetric this time. The restrictions on the coefficients are so made to enable certain calculations to be carried out regarding the auxiliary problems mentioned below.

Reflecting for a minute on the generic method one must employ to successfully carry out the proof of homogenization, it becomes apparent that the traditional notion of a corrector is no longer applicable since we do not have an affirmative answer to the existence of a corrector with boundedness or sufficiently slow growth properties. In every classical type random medium homogenization problem we have a result of the form,

Proposition 2.4.3. *The solution to the resolvent equation (componentwise in $\mathbb{R}^d$),*

$$(\lambda I - L) u_{\lambda}^{i} = b^{i} \ ,$$

has the property, for some $\zeta \in L^2(\mu)$,

$$\lambda |u_{\lambda}^{i}|_2 + |Du_{\lambda}^{i} - \zeta|_2 \to 0, \tag{2.4.2}$$

where the subscript 2 denotes the L^2 norm with respect to μ.

This is used to compensate for the lack of a corrector in the traditional sense and is the corresponding auxiliary problem in this type of homogenization. It still retains the philosophy of a corrector though since it is used to cancel highly oscillatory terms at the expense of introducing convergent quantities. Although in one version of the proof presented in [PV81] this result is not taken to hold strongly.

Another possible issue that cannot be resolved as in the periodic case is compactness in the space of continuous paths. This is because in the random case generically we want tightness over the random medium as well as over the family of diffusion processes in some sense (tightness of the annealed law), this is either arrived at as in [PV81, Rho09a] where we obtain tightness on the space of continuous paths for the family of measures given by $\mathbb{P}^\varepsilon$, for all x in a compact set, ω,

$$\mathbb{P}^\varepsilon(F) = \int F(\zeta(\cdot))\, Q_x^\varepsilon(d\zeta, \omega(\cdot)) \ ,$$

for F a real valued function on path space and $Q_x^\varepsilon(\cdot, \omega)$ the measure on the space of paths induced by X_x^ε with ω as the realization of the random environment. The tightness with respect to the family of annealed laws $\bar{\mathbb{P}}^\varepsilon = \mu(d\omega) \times \lambda(dx)\mathbb{P}^\varepsilon$ for a suitable measure on $\mathbb{R}^d$, $\lambda(dx)$, can also be obtained directly, see for instance [Rho09c, Rho09b]. One method of achieving this is using the Garsia-Rodemich-Rumsey inequality via the Feynman Kac formula to show tightness directly as in [Oll94] and [Rho09c]. Another alternative is to make estimates on the density using Nash's estimates [PV81] or Aronson estimates [Rho09a] to obtain bounds above on the density that are uniform over the random environment to obtain tightness as in [PV81].

After obtaining tightness and a result akin to that of Proposition 2.4.3, the homogenization can then be achieved directly through bounding error terms (in the PDE framework) [PV81]. Alternatively it is achieved using an ergodic result (or existence of an ergodic invariant measure [Oll94]) to obtain convergence of the coefficients/quadratic variation to their averaged form, analogous to those used in the periodic case. For instance for $\Psi : \Omega \times \mathbb{R}^d \to \mathbb{R}$ sufficiently integrable (dependent on the situation in question), this type of result is usually something similar to,

$$\lim_{\varepsilon \to 0} \bar{\mathbb{E}}^{\varepsilon}_{x} \left[\sup_{0 \le t \le T} \left| \int_0^t \Psi(X^{\varepsilon}(s), \tau_{X^{\varepsilon}(s)/\varepsilon}(\omega)) - \bar{\Psi}(X^{\varepsilon}(s))\, ds \right| \right] = 0\,, \tag{2.4.3}$$

where $\bar{\mathbb{E}}^{\varepsilon}_{x}$ denotes expectation with respect to the measure $\bar{\mathbb{P}}^{\varepsilon}$ i.e. the annealed version of $\mathbb{P}^{\varepsilon}$ and $\bar{\Psi}(\cdot) = \int \Psi(\cdot, \omega)\mu(d\omega)$. Then proceeding as before except working in a suitable (annealed) L^2 space and using the limit of the "corrector" as $\lambda = \varepsilon^2$ tends to zero i.e. using the properties of both terms in (2.4.2). This is a slight oversimplification of the more complicated cases, for instance in [Rho09c] there is a local time term that must be dealt with. This means that also required are ergodic results for local time integrals of the form corresponding to that in (2.4.3), tightness of the local time terms L^{ε} and in addition after identifying the limit $\bar{L}$, it is necessary to show that $\bar{L}$ is the local time of the limiting process. This follows from the fact that $\bar{L}$ can be shown to be associated to the Skorohod problem of the limiting process and that we have uniqueness in law for the solution of the corresponding Skorohod problem [LS84] which has a local time solution. By saying $\bar{L}$ is associated to the Skorohod problem of the limiting process $\bar{X}$ we mean that it is an increasing process and its points of increase are contained within the set $\{\bar{X} = 0\}$. In [Rho09a] there is the problem of the dependence of the solution to the resolvent equation on the slow variable which emerges through the difficulty in dealing with the term $b\partial_y u_{\varepsilon^2}$ necessitating an addition ergodic result, Theorem 6.3.

Hopefully the above (brief) summary of homogenization in a random environment serves to illustrate the prevalence of corrector based approaches even when there is no existence of correctors in the tradition sense, and additionally to emphasize the breadth of homogenization literature.

This completes a review of the most accessible and classic cases of probabilistic homogenization. It is at this point that we take stock of non-probabilistic methods to analyze the convergence of solutions in the simplest cases such as the method of multiscale expansions [PS08].

2.5 Multiscale methods for the homogenization of SDEs

This section is a quick summary of the method of multiscale expansions which is often applied to similar homogenization problems to the ones we will be dealing with. For a more detailed exposition of this section see the book by Pavliotis and Stuart, Multiscale Methods [PS08].

Here instead of a corrector we have the solution to the cell problem that features prominently and occupies the corresponding position in the method. If we work through the problem we wish to solve given by (2.2.1) in the fully periodic case again from this perspective and see how the probabilistic methodology is mirrored in terms of the corrector function by the cell problem. There are a wide range of different forms of problems that the multiscale expansion method can be applied to; multiscale expansions is an analytical PDE method that can be applied to many different classes of PDEs not just those that correspond to SDEs. We will outline the breadth of the scope of the method of multiscale expansions in due course. The hypotheses are once again, that the drift b is fully periodic and centered over the whole of $\mathbb{R}^d$ and hence we can consider the process as a process on $\mathbb{T}^d$ for $\varepsilon = 1$. In the multiscale method setting the problem is recast from (2.2.1) to homogenization of the Cauchy problem,

$$\begin{aligned} \partial_t u^\varepsilon(x,t) &= \Big(\frac{\Delta}{2} + \frac{b(\varepsilon^{-1}x)}{\varepsilon}.\bigtriangledown\Big) u^\varepsilon(x,t), \\ u^\varepsilon(x,0) &= f(x), \quad x \in \mathbb{R}^d, \end{aligned} \tag{2.5.1}$$

for $f \in C_0^\infty(\mathbb{R}^d)$. This corresponds to the equation satisfied by $\mathbb{E}_x[f(X^\varepsilon(t))]$, for the SDE with drift $\varepsilon^{-1}b(\varepsilon^{-1}\cdot)$ and diffusion coefficient 1 on $\mathbb{R}^d$. Using the rigorous mathematical background in [PS08], we have the convergence of the solutions to (2.5.1) in $L^\infty(\mathbb{R}^d \times (0,T))$. This is sufficient after verifying tightness in $C([0,\infty),\mathbb{R}^d)$ for weak convergence since then convergence of $\mathbb{E}_x[f(X^\varepsilon(t))]$ (the finite dimensional distributions) in $L^\infty(\mathbb{R}^d \times (0,T))$ implies that the weak limit is unique.

The basic methodology behind a multiscale expansion is to split the process into two scales, a slow, macroscopic scale given by x and a fast, microscopic scale given by $y = x/\varepsilon$ and then treat x and y as independent variables. In doing so we are assuming an ansatz of the form,

$$u^{\varepsilon}(x) = u_0(x,y) + \varepsilon u_1(x,y) + \varepsilon^2 u_2(x,y) + \dots \tag{2.5.2}$$

for u_i periodic in y. u_0 is then the limit of the u^{ε} and hence the solution to the homogenized equation.

Of course this has to be accompanied by a rigorous mathematical justification, which is one of the main issues of homogenization via multiscale expansions.

The treatment of the variables as independent then means that we have,

$$\bigtriangledown \to \bigtriangledown_x + \frac{1}{\varepsilon}\bigtriangledown_y ,$$

i.e. $\bigtriangledown f(x, x/\varepsilon) = \bigtriangledown_x f(x,y)|_{y=x/\varepsilon} + \varepsilon^{-1} \bigtriangledown_y f(x,y)|_{y=x/\varepsilon}$. This means our generator $\mathcal{L}^{\varepsilon} = \Delta + \varepsilon^{-1} b.\bigtriangledown$ becomes $\varepsilon^{-2}\mathcal{L}_0 + \varepsilon^{-1}\mathcal{L}_1 + \mathcal{L}_2$ for,

$$\begin{aligned}
\mathcal{L}_0 &= \Delta_y/2 + b(y).\bigtriangledown_y , \\
\mathcal{L}_1 &= \bigtriangledown_x . \bigtriangledown_y + b(y).\bigtriangledown_x , \\
\mathcal{L}_2 &= \Delta_x/2 ,
\end{aligned}$$

we then have,

$$(\mathcal{L}_0 + \mathcal{L}_1 + \mathcal{L}_2)u^{\varepsilon} = \partial_t u^{\varepsilon} ,$$

which leads to the series of equations after equating powers of ε,

$$\mathcal{L}_0 u_0 = 0 , \tag{2.5.3}$$

$$\mathcal{L}_1 u_0 + \mathcal{L}_0 u_1 = 0 , \tag{2.5.4}$$

$$\mathcal{L}_2 u_0 + \mathcal{L}_1 u_1 + \mathcal{L}_0 u_2 = \partial_t u_0 . \tag{2.5.5}$$

By uniqueness of solution to the elliptic PDE (2.5.3) on the torus up to a constant we have that u_0 is independent of y and is purely a function of x. Hence we can write $u_0(x,y) = u(x)$.

(2.5.4) leads to,

$$(\Delta_y/2 + b.\bigtriangledown_y)u_1(x,y) = -b.\bigtriangledown_x (u(x)) . \tag{2.5.6}$$

Now we introduce a few theorems that we will need to progress this scheme further,

Theorem 2.5.1 (Fredhom Alternative). *Let H be a Hilbert space and let $K: H \to H$ be a compact operator. Then the following holds.*

- *Either the equations*

$$(I - K)u = f \,, \tag{2.5.7}$$

$$(I - K^{\star})U = F \,, \tag{2.5.8}$$

have unique solutions for every $f, F \in H$ or

- *the homogeneous equations*

$$(I - K)u_0 = 0 \,,$$

$$(I - K^{\star})U_0 = 0 \,,$$

have the same finite number of non trivial solutions i.e. the dimensions of the null spaces of $(I - K)$ and $(I - K^{\star})$ are both finite and equal. In this case (2.5.7) and (2.5.8) have a solution if and only if

$$(f, U_0) = 0, \quad \forall U_0 \in \mathcal{N}(I - K^{\star}) \,,$$

and

$$(f, u_0) = 0, \quad \forall u_0 \in \mathcal{N}(I - K) \,,$$

respectively.

A compact operator on a Hilbert space H is one for which the image of a bounded set is precompact (a set where any sequence of elements has a convergent subsequence). Second order periodic uniformly elliptic operators can be shown to possess the Fredholm alternative except we have to proceed via the resolvent which can be shown to be a compact operator $\mathcal{L}^2_{per} \to \mathcal{L}^2_{per}$. This done using the Lax Milgram Theorem to get the existence of weak solutions to the resolvent problem in the weak formulation on the Sobolev space $\mathcal{H}^1_{per}$. Then showing that the resolvent operator is bounded $\mathcal{L}^2_{per} \to \mathcal{H}^1_{per}$, followed by the Rellich compactness theorem which gives that the space $\mathcal{H}^1_{per}$ is compactly embedded in $\mathcal{L}^2_{per}$. The

Fredholm alternative is then deduced from the Fredholm alternative of the resolvent operator. Where the notation is as follows, $\mathcal{L}^2_{per}$ is the completion of $C^\infty(\mathbb{T}^d)$ under the L^2 norm and $\mathcal{H}^1_{per}$ is the completion of $C^\infty(\mathbb{T}^d)$ under the H^1 norm.

The Lax Milgram Theorem is a workhorse of PDE theory used to give existence of solutions and is as follows,

Theorem 2.5.2 (Lax Milgram). *Let H be a Hilbert space with norm $\|\cdot\|$ and inner product $(\cdot,\cdot)$. Let $\langle\cdot,\cdot,\rangle$ denote the pairing between $H^\star$ and H. Let $a: H \times H \to \mathbb{R}$ be a bilinear mapping which satisfies,*

- *Coercivity: there exists a constant $\alpha > 0$ such that,*

$$a[u,u] \geq \alpha\|u\|^2 \quad \forall u \in H\,.$$

- *Continuity: there exists a constant $\beta > 0$ such that,*

$$a[u,v] \leq \beta\|u\|\|v\| \quad \forall u,\, v \in H\,.$$

Let $f: H \to \mathbb{R}$ be a bounded linear functional on H. Then there exists a unique element $u \in H$ such that

$$a[u,v] = \langle f,v\rangle\,,$$

for all $v \in H$.

and the Rellich Compactness Theorem,

Theorem 2.5.3. *From every bounded sequence in $H^1(\mathbb{T}^d)$ we can extract a subsequence which is strongly convergent in $L^2(\mathbb{T}^d)$.*

Returning to the problem at hand, the Fredholm alternative implies that (2.5.6) has a solution if and only if the RHS is centered in y for every x with respect to the invariant measure of the process with drift b and diffusion coefficient 1. Call the density of this measure ρ. This is because the LHS of the above equation is a differential operator in y only. From the centering of b, this is indeed the case. So, noting the form of the RHS of the equation we assume a solution of the form,

$$\chi(y).\,\nabla_x u(x,t)\,. \tag{2.5.9}$$

χ in this context is known as the solution of the cell problem, but is none other than the corrector g from the probabilistic method shown earlier. Note that if we are seeking a higher order expansion (for instance we want convergence at some specified rate in ε) then we must take note of the fact that we potentially have a term in x, $\tilde{u}_1$ in addition to (2.5.9) which we then use higher order terms to derive an equation for. For now we will take $\tilde{u}_1 \equiv 0$.

Now we analyze (2.5.5). Solving for u_2, the solvability condition given by the Fredholm alternative is now;

$$\int_{\mathbb{T}^d} (\partial_t u - \mathcal{L}_2 u - \mathcal{L}_1 u_1)\rho \, dy = 0 \, .$$

The homogenized equation now falls out of this solvability condition,

$$\partial_t u = \Delta u/2 + \int_{\mathbb{T}^d} (\mathcal{L}_1 u_1)\rho \, dy \, , \tag{2.5.10}$$

and

$$\mathcal{L}_1 u_1 = (b \otimes \chi + \nabla_y \chi). \nabla_x \nabla_x u \, ,$$

which corresponds exactly to the answer obtained from the probabilistic approach, since,

$$\int_{\mathbb{T}^d} b \otimes \chi \rho \, dy. \nabla_x \nabla_x u = \frac{1}{2} \int_{\mathbb{T}^d} \nabla_{y_k} \chi_i \nabla_{y_k} \chi_j \rho \, dy \nabla_{x_i} \nabla_{x_j} u \, ,$$

since the invariant density ρ satisfies the stationary Fokker Planck equation (also known as the Kolmogorov forward equation), $\mathcal{L}^\star \rho = 0$.

In terms of the rigorous proof of convergence we end up with convergence in the space $L^\infty(\mathbb{R}^d \times (0,T))$ of solutions to the equations in (2.5.1) to the limit u given by (2.5.10). The basic methodology behind the rigorous proof is to obtain an equation for the remainder term R^ε of the form,

$$\begin{aligned} \partial_t R^\varepsilon &= \mathcal{L}^\varepsilon R^\varepsilon + \varepsilon F^\varepsilon(x,t), \quad (x,t) \in \mathbb{R}^d \times (0,T) \, , \\ R^\varepsilon &= \varepsilon r^\varepsilon(x), \quad (x,t) \in \mathbb{R}^d \times \{0\} \, , \end{aligned}$$

for functions F^ε, r^ε bounded over all ε in $L^\infty(\mathbb{R}^d \times (0,T))$. Then the convergence follows from estimates obtained using the maximum principle.

To give some idea of how this analytic approach actually works, below is the convergence proof for the multiscale expansions we have just produced for the parabolic PDE homogenization problem above as it appears in [PS08].

Theorem 2.5.4. *Let $u^\varepsilon(x,t)$, $u(x,t)$ be as above, for $b(y) \in C^\infty_{per}$ and $f \in C^\infty_b(\mathbb{R}^d)$. Then*

$$\|u^\varepsilon(x,t) - u(x,t)\|_{L^\infty(\mathbb{T}^d \times (0,T))} \le C\varepsilon \ ,$$

and hence $u^\varepsilon \to u$ in $L^\infty(\mathbb{T}^d \times (0,T))$.

Proof. First let us derive an expression for u_2 in (2.5.2). Solving the constant order equation (2.5.5), noting the equation solved by u we have,

$$u_2(x, \varepsilon^{-1}x, t) = \Theta(\varepsilon^{-1}x). \nabla_x \nabla_x u(x,t) \ ,$$

where Θ is the solution of

$$\mathcal{L}_0\Theta = -b(y) \otimes \chi(y) - \nabla_y \chi(y) + \int_{\mathbb{T}^d} b(y) \otimes \chi(y) + \nabla_y \chi(y)\rho(y)\, dy \ , \tag{2.5.11}$$

obviously with periodic boundary conditions. Both χ and Θ satisfy uniformly elliptic PDEs on the torus hence both functions and all their derivatives are bounded. In addition $f \in C^\infty_b(\mathbb{R}^d)$ implies that u is bounded in $L^\infty(\mathbb{R}^d \times (0,T))$ together with all its derivatives as the solution of a PDE with constant coefficients. Therefore noting the form of u_1, u_2, we have that $\|u_1(x,\varepsilon^{-1}x,t)\|_{L^\infty(\mathbb{R}^d\times(0,T))} \le C$ and $\|u_2(x,\varepsilon^{-1}x,t)\|_{L^\infty(\mathbb{R}^d\times(0,T))} \le C$ uniformly over ε. We are now halfway there, since we have good behavior of the corrector terms now up to sufficiently high order, now we just analyze the equation satisfied by the error term. Let R^ε be given by,

$$u^\varepsilon(x,t) = u(x,t) + \varepsilon u_1(x,\varepsilon^{-1}x,t) + \varepsilon^2 u_1(x,\varepsilon^{-1}x,t) + R^\varepsilon(x,t) \ . \tag{2.5.12}$$

Applying $\mathcal{L}^\varepsilon$ to (2.5.12) gives,

$$\mathcal{L}^\varepsilon u^\varepsilon = \partial_t u + \varepsilon(\mathcal{L}_1 u_2 + \mathcal{L}_2 u_1) + \varepsilon^2 \mathcal{L}_2 u_2 + \mathcal{L}^\varepsilon R^\varepsilon \ ,$$

by construction. Combining this with the partial differential with respect to t of (2.5.12) gives

$$\partial_t R^\varepsilon = \mathcal{L}^\varepsilon R^\varepsilon + \varepsilon(\mathcal{L}_1 u_2 + \mathcal{L}_2 u_1 - \partial_t u_1) + \varepsilon^2(\mathcal{L}_2 u_2 - \partial_t u_2)$$

$$= \mathcal{L}^\varepsilon R^\varepsilon + \varepsilon F^\varepsilon(x,t) ,$$

for F^ε chosen correctly. In addition,

$$f(x) = u^\varepsilon(x,0) = u(x,0) + \varepsilon u_1(x, \varepsilon^{-1}x, 0) + \varepsilon^2 u_2(x, \varepsilon^{-1}x, 0) + R^\varepsilon(x,0) ,$$

and since $u(x,0) = f(x)$ also we have,

$$R^\varepsilon(x,0) = -\varepsilon u_1(x, \varepsilon^{-1}x, 0) - \varepsilon^2 u_2(x, \varepsilon^{-1}x, 0) = \varepsilon r^\varepsilon(x) .$$

Hence we have a Cauchy problem for the remainder term R^ε. Explicit calculation of F^ε in terms of u, Θ, χ and their derivatives show that F^ε is bounded in $L^\infty(\mathbb{R}^d \times (0,T))$ uniformly over ε. Another explicit calculation expressing r^ε in terms of χ, Θ and the space derivatives of $u(\cdot,0)$ gives r^ε is also bounded in $L^\infty(\mathbb{T}^d \times (0,T))$ since $f \in C_b^\infty(\mathbb{R}^d)$. Using the maximum principle for parabolic PDEs, we have that,

$$\begin{aligned} \|R^\varepsilon\|_{L^\infty(\mathbb{R}^d \times (0,T))} &\leq \varepsilon \|r^\varepsilon\|_{L^\infty(\mathbb{R}^d)} + \varepsilon \int_0^T \|F^\varepsilon(\cdot,s)\|_{L^\infty(\mathbb{R}^d)} \, ds \\ &\leq \varepsilon C + \varepsilon C' T \leq C\varepsilon . \end{aligned}$$

Hence,

$$\begin{aligned} &\|u^\varepsilon - u\|_{L^\infty(\mathbb{T}^d \times (0,T))} \\ &= \|\varepsilon u_1 + \varepsilon^2 u_2 + R^\varepsilon\|_{L^\infty(\mathbb{T}^d \times (0,T))} \\ &\leq \varepsilon \|u_1\|_{L^\infty(\mathbb{T}^d \times (0,T))} + \varepsilon^2 \|u_2\|_{L^\infty(\mathbb{T}^d \times (0,T))} + \|R^\varepsilon\|_{L^\infty(\mathbb{T}^d \times (0,T))} \\ &\leq C\varepsilon , \end{aligned}$$

thus the proof is complete. □

Such expansions can even be used to deal with time dependent coefficients in the cases where we have an operator in a domain Ω, of the form,

$$\begin{aligned} \partial_t u^\varepsilon - \nabla_x . (a^\varepsilon(x,t) \nabla_x u^\varepsilon) = f , \\ u^\varepsilon(x,0) = u_0(x) , \end{aligned}$$

where a^ε is bounded in $L^\infty(\Omega \times [0,T])$ uniformly over ε, and uniformly elliptic over all ε. In addition, $a^\varepsilon(x,t) = a(x, x/\varepsilon, t, t/\varepsilon^k)$ for $k > 0$, a is a fixed matrix valued function with these properties. A similar asymptotic expansion is used but this time the trick of producing separate variables on different scales must be applied to time as well as space. Incidentally we have a different limit dependent on whether $k < 2$, $k = 2$ or $k > 2$. See [BLP78, Chapter 2] for a discussion of this problem.

2.6 An important extension to multiscale expansions: boundary layers

All the problems surveyed so far have been very classical in nature. However, multiscale expansions is more of a philosophy than an algorithm and as a result there are some useful extensions.

A very useful extension of the previous method of multiscale expansions is the homogenization of elliptic problems using boundary layers (which provide a more precise ansatz). There is as yet no corresponding construction for parabolic problems although Allaire is rumored to be working on such a construction. In order to underline the effectiveness of boundary layers we initially follow the same path as in [AA99] before studying in detail an example [MGV08]. Boundary layers are extra terms in the expansion, used to remove large oscillations in the gradient near the boundary which is important for convergence in H^1 (note that the previous convergence neglected the derivatives since we were using L^∞). For instance, to illustrate the suboptimality of a standard multiscale expansion based approach, we have the theorem,

Theorem 2.6.1 ([BLP78])**.** *Given the elliptic problem in a bounded open subset of $\mathbb{R}^d$, Ω,*

$$\begin{aligned} -divA_\varepsilon \bigtriangledown u^\varepsilon &= f, \quad in\,\Omega\,, \\ u^\varepsilon &= 0, \quad on\,\partial\Omega\,, \end{aligned} \tag{2.6.1}$$

where $A_\varepsilon(x) = A(\varepsilon^{-1}x)$ for a 1-periodic bounded uniformly elliptic matrix valued function A and $f \in L^2(\Omega)$. In addition, we have the solution to the homogenized problem,

$$-div\bar{A} \bigtriangledown u = f, \quad in\,\Omega\,,$$

$$u = 0, \quad on\ \partial\Omega\ ,$$

for the averaged diffusion matrix $\bar{A}$ *given by,*

$$\int_{[0,1]^d} A - A \bigtriangledown \chi \, dy\ .$$

Since the invariant measure in the absence of drift has density 1 with respect to Lebesgue measure.

Then we have weak convergence of u^ε *to* u *in* $H^1_0(\Omega)$ *(as previously stated). If* $u \in \mathcal{W}^{2,\infty}$ *(two weak derivatives in* L^∞*), then*

$$\|u^\varepsilon - u - \varepsilon u_1\|_{H^1(\Omega)} = O(\sqrt{\varepsilon})$$

for the usual multiscale expansion $u^\varepsilon = u + \varepsilon u_1 + \varepsilon^2 u_2$.

Generically this is considered to be an optimal bound [MGV08, AA99]. This encompasses an important class of homogenization problems as this can be used as a model for electrical and thermal conduction in composite materials where the microscopic structure is periodic [NR01].

Notice in the formal multiscale expansions conducted previously in this section there is no attempt made to deal with the fact that the boundary conditions of the error terms are basically just ignored. We will now be matching the boundary conditions over the multiscale expansion as a whole, i.e. removing the perturbations near the boundary introduced by the terms $u_1, u_2, \ldots$.

If we assume an ansatz of the form, $u^\varepsilon = u(x) + \varepsilon(u_1(x, \varepsilon^{-1}x) + u_1^{bl,\varepsilon}(x)) + \varepsilon^2(u_2(x, \varepsilon^{-1}x) + u_2^{bl,\varepsilon}(x)) + \ldots$. Where the boundary layer corrector functions satisfy the series equations,

$$\mathrm{div} A_\varepsilon \bigtriangledown u_i^{bl,\varepsilon} = 0, \quad \text{in } \Omega\ ,$$
$$u_i^{bl,\varepsilon}(x) = -u_i(x, \varepsilon^{-1}x), \text{ on } \partial\Omega\ .$$

So now taking $u_1 = \chi. \bigtriangledown u + \tilde{u}_1(x)$ for any reasonable $\tilde{u}_1$, we have the result,

Theorem 2.6.2 ([MV97])**.** *In the former situation, assume that* $u \in \mathcal{W}^{2,\infty}$ *again, then*

$$\|u^\varepsilon(x) - u(x) - \varepsilon u_1(x, \varepsilon^{-1}x) - \varepsilon u_1^{bl,\varepsilon}(x)\|_{H^1(\Omega)} = O(\varepsilon)\ .$$

Of course since we have now introduced the solution of a problem dependent on ε, we usually would like some form of good behavior of $u_1^{bl,\varepsilon}$ uniformly over ε.

Convergence results like that in Theorem 2.6.2 above are very similar to the standard case demonstrated previously except that we work in the space $H_0^1(\Omega)$ instead of L^∞.

Using the first three terms of the multiscale expansion, u, u_1 and $u_1^{bl,\varepsilon}$, we write an equation for the remainder again, $R^\varepsilon = u^\varepsilon - u - \varepsilon u_1 - \varepsilon u_1^{bl,\varepsilon}$,

$$
\begin{aligned}
-\bigtriangledown .A_\varepsilon \bigtriangledown R^\varepsilon = (f + \bigtriangledown .A_\varepsilon \bigtriangledown u) + \varepsilon \bigtriangledown .A_\varepsilon \bigtriangledown u_1, \quad \text{in}\,\Omega\,, \\
R^\varepsilon = 0, \quad \text{on}\,\partial\Omega\,.
\end{aligned}
$$

Since the remainder is zero on the boundary (thanks to the boundary layer term) using Poincaré's inequality all we have to show is that the gradient is $O(\varepsilon)$ in $L^2(\Omega)$, which by the uniform ellipticity of the matrix will follow if it can be shown that,

$$
\int_\Omega A_\varepsilon \bigtriangledown R^\varepsilon . \bigtriangledown \phi \, dx \le C\varepsilon \|\phi\|_{H_0^1(\Omega)}\,,
$$

for all $\phi \in H_0^1(\Omega)$, which is equivalent to,

$$
\begin{aligned}
-\int_\Omega \bigtriangledown .(A_\varepsilon \bigtriangledown R^\varepsilon).\phi \, dx &= \int_\Omega (f + \bigtriangledown .A_\varepsilon \bigtriangledown u + \varepsilon \bigtriangledown .A_\varepsilon \bigtriangledown u_1).\phi \, dx \\
&= \int_\Omega (-\bigtriangledown_y .A_\varepsilon \bigtriangledown_y u_2 + \varepsilon \bigtriangledown_x .A_\varepsilon \bigtriangledown_x u_1).\phi \, dx \\
&\le \left| \int_\Omega (-\varepsilon \bigtriangledown_x .A_\varepsilon \bigtriangledown_y u_2 - \bigtriangledown_y .A_\varepsilon \bigtriangledown_y u_2).\phi \, dx \right| \\
&\quad + \left| \int_\Omega (\varepsilon \bigtriangledown_x .A_\varepsilon \bigtriangledown_y u_2 + \varepsilon \bigtriangledown_x .A_\varepsilon \bigtriangledown_x u_1).\phi \, dx \right| \\
&\le 2\varepsilon \left| \int_\Omega A_\varepsilon \bigtriangledown_y u_2 \bigtriangledown \phi \, dx \right| + C\varepsilon \|\phi\|_{H_0^1(\Omega)} \\
&\le D\varepsilon \|\phi\|_{H_0^1(\Omega)}\,.
\end{aligned}
$$

Since by corresponding calculations as in the standard situation we have bounds on the relevant derivatives of u_1, u_2 above which have a similar form to the previous calculation. This last inequality implies the desired convergence.

In particular as in [BLP78], we can choose u_1 to be given by,

$$
\chi(y). \bigtriangledown_x u(x)
$$

where $\chi(y)$ satisfies the cell problem

$$-\bigtriangledown_y .A(y) \bigtriangledown_y \chi(y) = \bigtriangledown.A(y)$$

χ 1-periodic, $\int_{\mathbb{T}^d} \chi \, dy = 0$.

u_2 is given by $\Theta. \bigtriangledown_x \bigtriangledown_x u(x)$. Set $A(y) = a_{ij}(y)$, then Θ is,

$$\bigtriangledown_y.(A \bigtriangledown_y \Theta^{ij}(y)) = a_{ij}(y) - a_{kj}\frac{\partial \chi^i}{\partial y_k} - a_{ik}\frac{\partial \chi^j}{\partial y_k} - \frac{\partial a_{kj}}{\partial y_k}\chi^i - q_{ij} \; ,$$

for, q_{ij} constants given by,

$$q_{ij} = \int_{\mathbb{T}^d} a_{ij}(y) - a_{kj}\frac{\partial \chi^i}{\partial y_k} - a_{ik}\frac{\partial \chi^j}{\partial y_k} - \frac{\partial a_{kj}}{\partial y_k}\chi^i \, dy \; .$$

dy is the invariant measure of the system on $\mathbb{T}^d$ given by the generator $\bigtriangledown_y(A(y)\bigtriangledown_y)$.

Under certain slightly stronger assumptions [AA99], using the boundary layer terms in the proof then omitting them from the asymptotic expansion, we have results of the form,

Theorem 2.6.3. *Let u^ε, u_1 and u be given as in the elliptic problem immediately above. If we have $u \in \mathcal{W}^{3,\infty}(\Omega)$ (a slightly stricter hypothesis than before). Provided also we have some sensible regularity assumptions, e.g. either*

- *the solution of (2.6.1) is a real valued function as opposed to existing only in a Sobolev space,*
- *or the boundary is at least C^2 and we have that A is a holder continuous matrix valued function.*

Then for any compact set ω compactly embedded in Ω we have the convergence,

$$\|u^\varepsilon(x) - u(x) - \varepsilon u_1(x, \varepsilon^{-1}x)\|_{H^1(\omega)} \leq C\varepsilon \; ,$$

for C a constant depending on ω.

This is really an estimate away from the boundary, although including derivatives, and is reliant on good behavior of the order ε boundary layer term.

Proof of Theorem 2.6.3 [AA99]. The good behavior (bounding in $H^1(\omega)$ in this case) of $u_1^{bl,\varepsilon}$ is given by a technical lemma,

Lemma 2.6.4. *[AA99] For a sequence ϕ_ε in $H^1(\Omega)$, we define the sequence of solutions $z_\varepsilon \in H^1(\Omega)$ of,*

$$\bigtriangledown . A_\varepsilon \bigtriangledown z_\varepsilon = 0, \quad \text{in } \Omega,$$
$$z_\varepsilon = \phi_\varepsilon, \quad \text{on } \partial\Omega.$$

Assume that we have the solution to (2.6.1) is a real valued function and

$$\|\phi_\varepsilon\|_{L^\infty(\partial\Omega)} \leq C$$

(i.e. we will consider only the first set of hypotheses in Theorem 2.6.3). Then for any open set compactly embedded in Ω, there exists another positive constant C, such that,

$$\|z_\varepsilon\|_{H^1(\omega)} \leq C \, .$$

Proof of Lemma 2.6.4. Let ϕ be a smooth function with compact support in Ω and such that $\phi = 1$ in ω. Using the equation for z_ε, we obtain

$$\int_\Omega \phi^2 (A_\varepsilon \bigtriangledown z_\varepsilon). \bigtriangledown z_\varepsilon \, dx = -2 \int_\Omega \phi z_\varepsilon (A_\varepsilon \bigtriangledown z_\varepsilon). \bigtriangledown \phi \, dx \, ,$$

which implies, using the uniform ellipticity of A,

$$\| \bigtriangledown z_\varepsilon\|_{L^2(\omega)} \leq \|\phi \bigtriangledown z_\varepsilon\|_{L^2(\Omega)} \leq C\| \bigtriangledown \phi z_\varepsilon\|_{L^2(\Omega)} \leq C\|z_\varepsilon\|_{L^2(\Omega)} \, .$$

Using the maximum principle we then have that $\|z_\varepsilon\|_{L^2(\Omega)} \leq \|\phi_\varepsilon\|_{L^\infty(\partial\Omega)}$. Hence,

$$\|z_\varepsilon\|_{H^1(\omega)} \leq (C+1)\|z_\varepsilon\|_{L^2(\Omega)} \leq K \, ,$$

C, K constants. □

Remark 2.6.5. The same result follows in the case of regularity assumptions on the boundary of Ω except that this time we use the bound on ϕ_ε in $L^2(\partial\Omega)$ instead of $L^\infty(\partial\Omega)$. Then the result follows from compactness arguments instead of the maximum principle.

Then,

$$\|u^\varepsilon - u - \varepsilon u_1\|_{H^1(\omega)} \leq \|u^\varepsilon - u - \varepsilon u_1 - \varepsilon u_1^{bl,\varepsilon}\|_{H_0^1(\Omega)} + \varepsilon\|u_1^{bl,\varepsilon}\|_{H^1(\omega)} \, ,$$

by the bounds on u_1 in either $L^2(\partial\Omega)$ or $L^\infty(\Omega)$, an application of Lemma 2.6.4 completes the proof. □

Remark 2.6.6. Without $u_1^{bl,\varepsilon}$ present in the multiscale expansion i.e. for $\|u^\varepsilon - u - \varepsilon u_1\|_{H^1}$, then the best we can achieve is an interior estimate in the fashion of Theorem 2.6.3. This is because, in general, the $H^1(\Omega)$ norm of $u_1^{bl,\varepsilon}$ behaves like $\varepsilon^{-1/2}$ as $\varepsilon \to 0$ [AA99, Lemma 2.6]. Hence we would be stuck with an expansion only to accuracy $O(\varepsilon^{1/2})$ even though it would include terms of order ε.

Analyzing the behavior of the boundary layer corrector is often the most complicated part of producing an asymptotic expansion in a homogenization and can lead to the study of approximations to the boundary layer problem that are dependent on the shape of the domain in question [AA99, MGV08].

As an example of what this could entail, in [MGV08] (which closely mirrors the earlier work [AA99] on which it builds), the production of an $O(\varepsilon^2)$ expansion for exactly the elliptic problem above in a convex polygon is studied. This is under the assumption that the tangent of the angles of the sides are sufficiently irrational; the approximation is made by localizing the boundary layer corrector problem near the hyperplanes and being in possession of sufficient convergence properties of the solution to these problems that the presence of the other hyperplanes does not perturb the solution too much in the fast variable. Taking $n = 1$ (1.1) in [MGV08] to parallel the treatment in [AA99]. Assume that we are given a convex polygonal domain $P = \cap_{k=1}^N \{x : n^k.x > c^k\}$ bounded by N hyperplanes of $\mathbb{R}^d$ for $d = 2$ or 3 with inward unit normals n^k.

The normals n^k satisfy the following condition giving sufficient irrationality in the slope of the faces of the polygon relative to the periodic mesh grid of the coefficients.

Definition 2.6.7 (Small divisor assumption)**.** There is a $c,\, l > 0$ such that $\forall \xi \in \mathbb{Z}^d \backslash \{0\}$, $|n^k \times \xi| \geq c|\xi|^{-l}$ where $\times$ denotes the cross product in 3 dimensions and for $d = 2$ it denotes the 2 dimensional equivalent, $(x_1, x_2) \times (y_1, y_2) = x_2 y_1 - x_1 y_2$.

Then we have $u_1^{bl,\varepsilon} \approx \sum_{k=1}^N v_{bl,\varepsilon}^k$ for $v_{bl,\varepsilon}^k(x, y)$ defined as follows. Let the half space $\Omega^{\varepsilon,k}$ be the half space defined by,

$$\Omega^{\varepsilon,k} = \{y : n^k.y - c^k/\varepsilon > 0\} ,$$

then $v^k_{bl,\varepsilon}(x,y)$ is the solution of,

$$\begin{aligned} \nabla_y . A(y) \nabla_y v^k_{bl,\varepsilon} &= 0, \quad y \in \Omega^{\varepsilon,k}, \\ v^k_{bl,\varepsilon} &= -u_1(x,y), \quad y \in \partial\Omega^{\varepsilon,k} . \end{aligned} \tag{2.6.2}$$

Firstly, favorable behavior of the boundary layer terms $v^k_{bl,\varepsilon}(x,y)$ is shown and then the significance of this is explored.

Despite the value of $v^k_{bl,\varepsilon}$ changing subtly on the boundary $\partial\Omega^{\varepsilon,k}$, by looking at a related system that corresponds to this, it is shown that this effect does not affect the limit obtained. In fact we have that,

Lemma 2.6.8. *$v^k_{bl,\varepsilon}$ tends to a constant limit (in y, remember that the entire problem is a function of x and hence the limit is also a function of x) as $(y.n_k - c^k/\varepsilon) \to \infty$ with a convergence rate better than any power of $|y.n_k - c^k/\varepsilon|$.*

Usually this is achieved using an inequality related to the Poincaré inequality,

$$\int_{\mathcal{O}} |\phi(x)|^2 \, dx \leq C_{\mathcal{O}} \int_{\mathcal{O}} | \nabla \phi(x)|^2 \, dx ,$$

for $\mathcal{O}$ a bounded open set in $\mathbb{R}^n$ for any $n \in \mathbb{N}$ and $\phi \in \mathcal{W}^{1,2}$ with zero average. The constant $C_{\mathcal{O}}$ is dependent on the domain only. The Poincaré type equality is then applied to the cube in the first $d-1$ coordinates times successive unit slices of the d th coordinate after an appropriate orthogonal transformation to take the face to a hyperplane of the form $x_d = c$ for c some constant, see [AA99, Lemma 4.4]. Except that now, due to the irrationality of the tangent of the angle of the faces to the period, we no longer can consider the problem over a region compact in the first $d-1$ coordinates. This is where the sufficient irrationality assumption (diophantine condition) on all the normals comes to the rescue (and of course where full generality of the result is lost, despite density of such normals in $\mathbb{R}^d$). The Poincaré inequality is replaced by using the small divisor assumption 2.6.7 in conjunction with the Plancherel formula and Hölder inequality to produce a similar inequality. For all $1 < p < \infty$, ϕ smooth enough with zero average over $\mathbb{T}^d$,

$$\int_{\mathbb{T}^d} |\phi(x)|^2 \, dx \leq C \left(\int_{\mathbb{T}^d} |N^t \nabla_x \phi|^2 \, dx \right)^{1/p} \left(\|\phi\|_{H^{l/(p-1)}(\mathbb{T}^d)} \right)^{2-2/p} ,$$

(2.6.3)

for N^t a $(d-1)\times d$ matrix all $d-1$ rows of which are drawn in order from the first $d-1$ rows of the transpose of M in (2.6.4) below.

Consider the system related to that given in (2.6.2),

$$\begin{aligned} \nabla_z . B(Mz) \nabla_z v = 0, \quad z_d > a \ , \\ v(z) = v_0(Mz), \quad z_d = a \ , \end{aligned} \tag{2.6.4}$$

where B has the same properties as A, v_0 is a smooth 1-periodic function and M is a $d \times d$ orthogonal matrix (that is used to align the hyperplane $x_d = c$ with the relevant face). Then we have a general result that gives Lemma 2.6.8,

Lemma 2.6.9. *There exists a constant $v^\infty(B, M, v_0, a)$, such that the limit,*

$$v(z) \to v^\infty(B, M, v_0, a), \quad \text{as } z_d \to \infty, \text{ uniformly in } z' \ ,$$

exists, where $z = (z', z_d)$, z' is a $d-1$ dimensional vector of the first $d-1$ coordinates. Moreover,

$$\lim_{t\to\infty} \left| t^m \partial_{z'}^\alpha \partial_t^\beta (v - v^\infty(B, M, v_0, a)) \right| = 0 \ ,$$

$\alpha \in \mathbb{N}^d$, $\beta \in \mathbb{N}$, *uniformly in z'.*

Remark 2.6.10. The independence from a of the limit above one might expect from the previous discussion is shown in due course.

Remark 2.6.11. Note that we are using the torus on the z' coordinate in [MGV08] due to fact that the solution of the problem satisfied by v after transformation given by (2.6.4) is of the form $V(N\theta, t)$ where V is 1-periodic in its first argument. The second argument then behaves like the d th coordinate along the strip in [AA99].

Outline of proof of Lemma 2.6.9. Using Hardy's inequality, the norm of $\tilde{V} = V - \int_{\mathbb{T}^d} V$ and $\partial_t^k V$ are shown to be less than a constant in any L^2 Sobolev space H^s, $s \in \mathbb{N}$. Using the small divisor assumption, Plancherel's formula and the Hölder inequality to obtain (2.6.3), (2.6.3) is then applied to $\tilde{V}$ to analyze the behavior of the tails of the integrals of the first derivatives with respect to z' and t of V (equivalently v),

$$f(T) = \int_{\mathbb{T}^d} \int_T^{+\infty} \left(|N^t \nabla_{z'} V|^2 + |\partial_t V|^2 \right) dt \, dz' \ ,$$

upon doing so, it is obtained that the tails decay faster than any polynomial.

Then with this chain of inequalities in hand, the properties of $v^k_{bl,\varepsilon}$ are established by applying the same argument as applied to $\tilde{V}$ inductively after differentiating (2.6.4) and showing that the same convergence to 0 applies to the derivatives of v.

Finally convergence of the average $\bar{V}$, $V = \tilde{V} + \bar{V}$, is shown by showing that it is a Cauchy function. □

Remark 2.6.12. cf the proof of Lemma 4.4 in [AA99], where recursive estimates using the Poincaré inequality are used to prove a corresponding result.

It still remains to be proved that the limit above is independent of a. This follows by proving continuity of the limit in a which follows from the convergence of an error term and energy estimates using the small divisor assumption. Once continuity is obtained, we have from the system satisfied by $v(z + M^t\xi)$ for $\xi \in \mathbb{Z}^d$, and Lemma 2.6.9 that,

$$v^\infty(B, M, v_0, a) = v^\infty(B, M, v_0, a - \xi.Me_d) \ ,$$

for e_d the vector $(0, 0, \ldots, 1)$. We have that $Me_d \notin \alpha\mathbb{Q}^d$ for any $\alpha \in \mathbb{R}$ by the small divisor assumption (if the normal was proportion to a rational vector we would have $n \times \xi = 0$ for some $\xi \in \mathbb{Z}^d$) which implies that the set $\{\xi.Me_d,\ \xi \in \mathbb{Z}^d\}$ is dense in $\mathbb{R}$, and the result follows from continuity.

Now it is shown what the favorable behavior of the boundary layer terms derived above means in terms of the main asymptotic expansion problem.

By altering the value of the function of x term only, $\tilde{u}_1$, present in the u_1 term, it is possible to make the limit of $v^k_{bl,\varepsilon}$ zero. In particular by choosing the boundary conditions of $\tilde{u}_1$. Since the limit is linear in the boundary conditions of (2.6.2) all we do is subtract the limit.

In light of the convergence above we have, since $v^k_{bl,\varepsilon} = v(y) \bigtriangledown u(x)$ from the form of systems (2.6.2), (2.6.4),

$$v^k_{bl,\varepsilon} \to v^{k,\infty}_{bl}(x) = -G^{k,\alpha}. \bigtriangledown u \ ,$$

as $y.n_k - c^k/\varepsilon \to \infty$, $G^{k,\alpha}$ defined below.

To be precise, if the constants $G^{k,\alpha}$, $1 \leq \alpha \leq d$ are then defined by (dropping the dependence on a),

$$G^{k,\alpha} = -v^{\infty}(M^k A(M^k)^T, M^k, \chi^{\alpha}) ,$$

where χ^{α} are the solutions to the cell problems. Then the required result with regard to the limit of $v^k_{bl,\varepsilon}$ is achieved by setting,

$$\tilde{u}_1 = G^{k,\alpha}\partial_{x_\alpha} u, \quad x \in \partial\Omega \cap K^k, \quad 1 \leq k \leq N,$$

where K^k is the face of the polygon i.e. $K^k = \{x : n^k.x = c^k\}$. That this is the right definition of $\tilde{u}_1$ is obtained by consideration of the solution of (2.6.4) for $B = M^k A(M^k)^T$, $M = M^k$, $v_0 = \chi^{\alpha}\partial_{x_\alpha} u$ and noting that the function of x, $\partial_{x_\alpha} u$, is linear over the problem.

In this way the $v^k_{bl,\varepsilon}$ are almost a deconstruction of the problem of the boundary layer near that side of Ω with the overall behavior encoded in the choice of $\tilde{u}_1$ needed to make the limit of these zero.

By removing the boundary layers through a clever choice of $\tilde{u}_1$, this allows estimates on the size of,

$$\|u^{\varepsilon} - u(x) - \varepsilon u_1(x, \varepsilon^{-1}x) - \varepsilon^2 u_2(x, \varepsilon^{-1}x)\|_{H^1(\omega)} ,$$

to be made, showing that it is $O(\varepsilon^2)$.

Now we will observe how the behavior derived above of the boundary layer terms is instrumental in the convergence result.

This takes place with a global error estimate and a boundary error estimate (remember $u_1^{bl,\varepsilon}$ was only approximated by the sum of $v^k_{bl,\varepsilon}$). The global error estimate is given by,

Lemma 2.6.13. $\|e_{\varepsilon}\|_{H^1(\Omega)} = O(\varepsilon^2)$, *where,*

$$\begin{aligned} e^{\varepsilon} =& u^{\varepsilon} - u(x) - \varepsilon u_1(x, \varepsilon^{-1}x) - \varepsilon^2 u_2(x, \varepsilon^{-1}x) \\ &- \varepsilon u_1^{bl,\varepsilon}(x, \varepsilon^{-1}x) - \varepsilon^2 u_2^{bl,\varepsilon}(x, \varepsilon^{-1}x) , \end{aligned}$$

with the usual definition for $u_1^{bl,\varepsilon}$, $u_2^{bl,\varepsilon}$.

The boundary layer estimate is given by,

Lemma 2.6.14. $\|e^{bl,\varepsilon}\|_{L^2(\Omega)} = O(\varepsilon)$*, where,*

$$e^{bl,\varepsilon} = u_1^{bl,\varepsilon} - \sum_{k=1}^{N} v_{bl,\varepsilon}^k + \varepsilon u_2^{bl,\varepsilon} \,.$$

Then given these bounds, the result can be shown using the behavior of the boundary layer terms $v_{bl,\varepsilon}^k$ derived earlier. We have that, using the global error bound,

$$\begin{aligned} &\|u^\varepsilon - u(x) - \varepsilon u_1(x,\varepsilon^{-1}x) - \varepsilon^2 u_2(x,\varepsilon^{-1}x)\|_{H^1(\omega)} \\ \le& C\varepsilon^2 + \varepsilon \left\| \sum_{k=1}^{N} v_{bl,\varepsilon}^k \right\|_{H^1(\omega)} + \varepsilon \|e^{bl,\varepsilon}\|_{H^1(\omega)} \,. \end{aligned}$$

By the earlier clever choice of $\tilde{u}_1$, we have the convergence of the boundary layer terms $v_{bl,\varepsilon}^k$ and their derivatives to 0 faster than any polynomial power as $(y.n^k - c^k/\varepsilon) \to \infty$, uniformly in x, ε i.e.,

$$\left\| \sum_{k=1}^{N} v_{bl,\varepsilon}^k(x,\varepsilon^{-1}x) \right\|_{H^s(\omega')} = O(\varepsilon^m), \quad \forall s,\, m, \quad \forall \omega' \subset\subset \Omega \,.$$

Then, noting that $\bigtriangledown.A \bigtriangledown u_i^{bl,\varepsilon} = 0$, we have,

$$\bigtriangledown.A \bigtriangledown e^{bl,\varepsilon} = - \bigtriangledown.A \bigtriangledown \sum_{k=1}^{N} v_{bl,\varepsilon}^k(x,\varepsilon^{-1}x), \quad x \in \Omega \,.$$

Hence, if $0 \le \phi \le 1$ is compactly supported in Ω and $\phi = 1$ in ω, then we have,

$$\begin{aligned} \int_\Omega \phi^2 A(x/\varepsilon) \bigtriangledown e^{bl,\varepsilon}. \bigtriangledown e^{bl,\varepsilon}\, dx =& -2\int_\Omega \phi e^{bl,\varepsilon}.(A(x/\varepsilon) \bigtriangledown e^{bl,\varepsilon}. \bigtriangledown \phi)\, dx \\ &+ \int_\Omega \big(\bigtriangledown.A \bigtriangledown \sum_{k=1}^{N} v_{bl,\varepsilon}^k(x,\varepsilon^{-1}x)\big)\phi^2 e^{bl,\varepsilon}\, dx \,, \end{aligned} \tag{2.6.5}$$

using the decay properties of the derivatives of $v_{bl,\varepsilon}^k$ on the second term we have that the second term is $O(\varepsilon^m)$ for any $m \in \mathbb{N}$. Thus we have, for $m \in \mathbb{N}$,

$$\|e^{bl,\varepsilon}\|_{H^1(\omega)} \le C\|e^{bl,\varepsilon}\|_{L^2(\Omega)} + C_m \varepsilon^m \,.$$

Taking $m = 1$ we then have,

$$\|u^\varepsilon - u(x) - \varepsilon u_1(x,\varepsilon^{-1}x) - \varepsilon^2 u_2(x,\varepsilon^{-1}x)\|_{H^1(\omega)} \le C\big(\varepsilon^2 + \epsilon\|e^{bl,\varepsilon}\|_{L^2(\Omega)}\big) \,,$$

which when combined with the boundary estimate will give the required result.

Proof of Lemma 2.6.13. Like the previous multiscale expansion example, the main part of the verification of this estimate is verifying that $\bigtriangledown A \bigtriangledown e^{\varepsilon} = O(\varepsilon^2)$. We have,

$$\begin{aligned}\bigtriangledown.A \bigtriangledown e^{\varepsilon} =& \varepsilon \bigtriangledown_x .(A \bigtriangledown_x u_1 + A \bigtriangledown_y u_2)(x, \varepsilon^{-1}x) \\ &+ \varepsilon \bigtriangledown_y .(A \bigtriangledown_x u_2)(x, \varepsilon^{-1}x) + \varepsilon^2 \bigtriangledown_x .(A \bigtriangledown_x u_2)(x, \varepsilon^{-1}x) \ .\end{aligned} \tag{2.6.6}$$

Since by the usual multiscale method of comparing equations of order $1/\varepsilon^2$, $1/\varepsilon$, and 1, plus the aforementioned clever choice of $\tilde{u}_1(x)$, we have explicitly,

$$u_1(x, y) = -\chi(y) \bigtriangledown u(x) + \tilde{u}_1(x) \ ,$$

as before χ is the solution to the cell problem,

$$-\bigtriangledown .A \bigtriangledown \chi = \bigtriangledown.A, \quad \int_{\mathbb{T}^d} \chi \, dy = 0 \ .$$

If we let the family of 1-periodic matrices Y considered as a function of y satisfy,

$$\begin{aligned}-\bigtriangledown_y .A \bigtriangledown_y \mathrm{Y} =& A - A \bigtriangledown_y \chi - \bigtriangledown_y.(A \otimes \chi) \\ &- \int_{\mathbb{T}^d} A - A \bigtriangledown_y \chi - \bigtriangledown_y.(A \otimes \chi) \, dy \ ,\end{aligned}$$

then u_2 is given by,

$$u_2(x, y) = \mathrm{Y}^{\alpha\beta} \partial^2_{x_\alpha x_\beta} u - \chi. \bigtriangledown \tilde{u}_1 \ .$$

Note that $\mathrm{Y} = \Theta$ from 2.5.11 and this is to be expected since the formal calculation is identical for both cases. $\tilde{u}_1$ is given by,

$$-\bigtriangledown .\bar{A} \bigtriangledown \tilde{u}_1 = c^{\alpha\beta\gamma} \partial^3_{x_\alpha x_\beta x_\gamma} u, \quad c^{\alpha\beta\gamma} = \int_y A^{\gamma\eta} \partial_{y_\eta} \mathrm{Y}^{\alpha\beta} - A^{\alpha\beta} \chi^\gamma \ .$$

For $\bar{A}$ the averaged matrix,

$$\bar{A} = \int_{\mathbb{T}^d} A + A \bigtriangledown \chi \, dy \ .$$

Note the lack of boundary conditions on $\tilde{u}_1$, they are specified as above and not by formal considerations.

Then by explicit calculation, in (2.6.6) we have that $\tilde{w} = \bigtriangledown_x.(A \bigtriangledown_x u_1 + A \bigtriangledown_y u_2)(x, \varepsilon^{-1}x)$ has zero average with respect to y. To get order ε^2 bounds from the order ε bounds, a standard trick is deployed. There are a number of matrix fields $W(x,y)$ such that $\bigtriangledown_y . W(x,y) = \tilde{w}$. Then if we set,

$$V(x,y) = W(x,y) + A(y) \bigtriangledown_x u_2(x,y) \ ,$$

we have,

$$\begin{aligned} \bigtriangledown A \bigtriangledown e^{\varepsilon} &= \varepsilon \bigtriangledown_y W(x, \varepsilon^{-1}x) + \varepsilon \bigtriangledown_y .(A \bigtriangledown_x u_2)(x, \varepsilon^{-1}x) \\ &\quad + \varepsilon^2 \bigtriangledown_x .(A \bigtriangledown_x u_2)(x, \varepsilon^{-1}x) \\ &= \varepsilon \bigtriangledown_y .V(x, \varepsilon^{-1}x) + \varepsilon^2 \bigtriangledown_x .(A \bigtriangledown_x u_2)(x, \varepsilon^{-1}x) \\ &= \varepsilon^2 \bigtriangledown .V(\cdot, \varepsilon^{-1} \cdot)(x) - \varepsilon^2 \bigtriangledown_x .W(\cdot, \varepsilon^{-1} \cdot)(x) \ . \end{aligned}$$

Providing we have V, W in H^1 then the last line yields the required ε^2 bound. The required regularity of W is given as in [BLP78], by choosing $W = -\text{curl}_x \phi(x,y)$ (which is smooth in y) of an appropriate vector field ϕ to make $\bigtriangledown_x . W = 0$ if the assumptions made are not sufficient to give automatic regularity of this term.

Then the result follows by a simple energy estimate. □

Proof of Lemma 2.6.14. The boundary layer estimate is by far the most complicated of the estimates since it is affected by the geometry of the situation to a large extent.

More precisely than before, we expect $u^{bl,\varepsilon} = u_1^{bl,\varepsilon} + \varepsilon u_2^{bl,\varepsilon}$ to have an expansion of the type,

$$u^{bl,\varepsilon} \approx \sum_{k=1}^{N} \left(v_{bl,\varepsilon}^k (x, \varepsilon^{-1}x) + \varepsilon w_{bl,\varepsilon}^k (x, \varepsilon^{-1}x) \right) \ ,$$

where $v_{bl,\varepsilon}^k$, $w_{bl,\varepsilon}^k$ are defined as before on the half space $\Omega^{\varepsilon,k}$. As a result of the method of formal expansions the next order corrector $w_{bl,\varepsilon}^k$ satisfies,

$$\begin{aligned} - \bigtriangledown_y .A \bigtriangledown_y w_{bl,\varepsilon}^k = \bigtriangledown_x . A \bigtriangledown_y v_{bl,\varepsilon}^k + \bigtriangledown_y . A \bigtriangledown_x v_{bl,\varepsilon}^k, \quad y \in \Omega^{\varepsilon,k} \ , \\ w_{bl,\varepsilon}^k = -u_2(x,y), \quad y \in \partial\Omega^{\varepsilon,k} \ . \end{aligned}$$

Since we have

$$\| w_{bl,\varepsilon}^k (x, \varepsilon^{-1}x) \|_{L^2(\Omega)} = O(1) \ ,$$

from the decay properties of $v_{bl,\varepsilon}^k$ which give similar convergence properties of $w_{bl,\varepsilon}^k$. We can prove the boundary layer estimate if it can be shown that $\tilde{e}^{bl,\varepsilon} =$

$e^{bl,\varepsilon} - \varepsilon \sum_{k=1}^{N} w_{bl,\varepsilon}^k(x, \varepsilon^{-1}x)$ is $O(\varepsilon)$ in $L^2(\Omega)$. $\tilde{e}^{bl,\varepsilon}$ satisfies the elliptic problem,

$$\begin{aligned} -\bigtriangledown .A \bigtriangledown \tilde{e}^{bl,\varepsilon} := r^\varepsilon = & \sum_{k=1}^{N} \bigtriangledown_x .(A \bigtriangledown_x v_{bl,\varepsilon}^k + A \bigtriangledown_y w_{bl,\varepsilon}^k)(x, \varepsilon^{-1}x) \\ & + \varepsilon \sum_{k=1}^{N} \bigtriangledown .(A \bigtriangledown_x w_{bl,\varepsilon}^k)(x, \varepsilon^{-1}x), \quad x \in \Omega, \\ \tilde{e}^{bl,\varepsilon} := \phi^{bl,\varepsilon} = & - u_1(x, \varepsilon^{-1}x) - \varepsilon u_2(x, \varepsilon^{-1}x) \\ & - \sum (v_{bl,\varepsilon}^k + \varepsilon w_{bl,\varepsilon}^k)(x, \varepsilon^{-1}x), \quad x \in \partial\Omega \, . \end{aligned} \tag{2.6.7}$$

Note that although the boundary terms match in terms of u_1, u_2 and $v_{bl,\varepsilon}^k$, $w_{bl,\varepsilon}^k$ for the relevant k, the sum over k contributes extra terms, which of course given the decay of these terms, will pose most of a problem at the vertices/edges. The second term on the rhs of the condition in the domain Ω (2.6.7) is equal to $\varepsilon \bigtriangledown .R^\varepsilon$ where $\|R^\varepsilon\|_{L^2(\Omega)} \leq C$. The first term is $L^2(\Omega)$ in x, smooth in y and decays to zero as $(y.n^k - c^k/\varepsilon) \to \infty$ faster than any power of $(y.n^k - c^k/\varepsilon)$ uniformly in x, ε. Hence given $\psi \in H_0^1(\Omega)$, and denoting this first term $\sum_k r^k$,

$$\begin{aligned} \left| \int_\Omega r^k(x, \varepsilon^{-1}x).\psi(x)\, dx \right| & \leq \int_\Omega |r^k| d(x, \partial\Omega) \frac{|\psi|}{d(x, \partial\Omega)}\, dx \\ & \leq \int_\Omega |r^k| d(x, K^k) \frac{|\psi|}{d(x, \partial\Omega)}\, dx \\ & = \int_\Omega |r^k| |x.n^k - c^k| \frac{|\psi|}{d(x, \partial\Omega)}\, dx \\ & = \varepsilon \int_\Omega |r^k| |y.n^k - c^k/\varepsilon| \frac{|\psi|}{d(x, \partial\Omega)}\, dx \\ & \leq C\varepsilon \| \bigtriangledown \psi \|_{L^2(\Omega)} \, , \end{aligned}$$

using the Cauchy Schwarz inequality followed by Hardy's inequality with $p = 2$. Combining the bounds on R^ε, r^k gives that $\|r^\varepsilon\|_{H^{-1}(\Omega)} \leq C\varepsilon$ (H^{-1} the dual space of H^1).

Remark 2.6.15. In the case of a half space, the boundary term $\phi^{bl,\varepsilon}$ would be zero and all the following work to produce the bounds on the boundary term is redundant.

If it can be proven that,

$$\|\phi^{bl,\varepsilon}\|_{W^{1-1/p,p}(\partial\Omega)} \leq C(p)\varepsilon, \quad \forall p < 2 \, , \tag{2.6.8}$$

then there exists φ^ε such that, for $p < 2$,

$$\varphi^\varepsilon \in W^{1,p}(\Omega), \quad \|\varphi^\varepsilon\|_{W^{1,p}(\Omega)} = O(\varepsilon), \quad \varphi^\varepsilon|_{\partial\Omega} = \phi^{bl,\varepsilon} \, ,$$

hence for $e^{\varepsilon} = e^{bl,\varepsilon} - \varphi^{\varepsilon}$ satisfying,

$$\begin{aligned} \triangledown.(A \triangledown e^{\varepsilon}) &= r^{\varepsilon}_{bl} - \triangledown.(A \triangledown \varphi^{\varepsilon}), \quad \text{in } \Omega \\ e^{\varepsilon}|_{\partial\Omega} &= 0 \,, \end{aligned}$$

considering the RHS in $\mathcal{W}^{-1,p}$ in the manner above, using the general results of Meyers [Mey63] which relate the norms of solutions in Sobolev spaces with zero boundary conditions to the norm of the result, considered in $\mathcal{W}^{-1,p}$, of applying the elliptic operator to the solution in Ω, we have for some $p_m < 2$, that for all $p_m < p < 2$,

$$\|e^{\varepsilon}\|_{\mathcal{W}^{1,p}} \leq C(p)\|r^{\varepsilon}_{bl} - \triangledown.(A \triangledown \varphi^{\varepsilon})\|_{\mathcal{W}^{-1,p}} \leq C\varepsilon \,.$$

By the Sobolev embedding with $k = 1$ and suitable values chosen for l, $p_m < p' < 2$, to make $q' = 2$, we have the $\|e^{\varepsilon}\|_{L^2(\Omega)}$ bound. Where the Sobolev embedding theorem we are using is as follows,

Theorem 2.6.16 (Sobolev embedding [Kon45])**.** *Let k be a non-negative integer and $1 \leq p' \leq \infty$. If $k > l$ and $1 \leq p' \leq q' \leq \infty$ are such that*

$$\frac{1}{q'} = \frac{1}{p'} - \frac{k-l}{d} \,,$$

then,

$$\mathcal{W}^{k,p'}(\mathbb{R}^d) \subset \mathcal{W}^{\lfloor l \rfloor,q'}(\mathbb{R}^d) \,.$$

The last step is then to produce (2.6.8). This is where the geometry of Ω comes into play. We break $\phi^{bl,\varepsilon}$ into those terms with an $O(1)$ coefficient and those with an $O(\varepsilon)$ coefficient in front.

$$\begin{aligned} \phi^{bl,\varepsilon} = &\left[-u_1(x,\varepsilon^{-1}x) - \sum_{k=1}^{N} v^k_{bl,\varepsilon}(x,\varepsilon^{-1}x)\right] \\ &+ \varepsilon\left[-u_2(x,\varepsilon^{-1}x) - \sum_{k=1}^{N} w^k_{bl}(x,\varepsilon^{-1}x)\right] . \end{aligned}$$

Clearly in $W^{1-1/p,p}(\Omega)$ we would want an $O(\varepsilon)$ bound on the first bracketed term on the RHS and an $O(1)$ bound on the second bracketed term on the RHS.

Starting with the first term on the RHS, calling it ϕ_v, we have that this is given by,

$$\phi_v|_{\partial\Omega\cap K^k} = -\sum_{j\neq k} V^j(\varepsilon^{-1}x) \bigtriangledown u(x) \ ,$$

for suitable V^j, by construction of the $v^k_{bl,\varepsilon}$. Now the problem with showing the bounds we seek is in $d = 2$, the vertices, and $d = 3$, the edges and vertices (basically the sets where at least 2 faces meet), as is shown by the following calculation, if ψ is a function with compact support away from the bad sets outlined above, for $p < 2$,

$$\|\psi\phi_v\|_{\mathcal{W}^{1-1/p,p}(\partial\Omega)} \leq C\|\psi\phi_v\|_{H^{1/2}(\partial\Omega)} = C_m\varepsilon^m\| \bigtriangledown u\|_{H^{1/2}(\Omega)} \leq K_m\varepsilon^m \quad \forall m \ ,$$

where the equality follows from the better than polynomial power of convergence to 0 of $v^k_{bl,\varepsilon}$ and its derivatives.

In order to obtain an $O(\varepsilon)$ bound, the property that $\bigtriangledown u = 0$ at a vertex is used since we have the Dirichlet condition $u|_{\partial\Omega} = 0$. Consider $d = 2$, the case $d = 3$ follows from similar calculations. Let ψ be a smooth function with support centered on a vertex now. For $f, g \in L^\infty \cap \mathcal{W}^{1-1/p,p}(\partial\Omega)$,

$$\begin{aligned}\|fg\|_{\mathcal{W}^{1-1/p,p}(\partial\Omega)} \leq C\big(&\|f\|_{L^\infty(\partial\Omega)}\|g\|_{\mathcal{W}^{1-1/p,p}(\partial\Omega)} \\ &+ \|g\|_{L^\infty(\partial\Omega)}\|f\|_{\mathcal{W}^{1-1/p,p}(\partial\Omega)}\big) \ ,\end{aligned}$$

hence, making the appropriate definition of V to have $\phi_v = V(\varepsilon^{-1}x) \bigtriangledown u$,

$$\begin{aligned}\|\psi^2\phi_v\|_{\mathcal{W}^{1-1/p,p}(\partial\Omega)} \leq C\Bigg(&\left\|\psi\frac{\nabla u}{|x|}\right\|_{L^\infty(\partial\Omega)} \|\psi|x|V(\varepsilon^{-1}x)\|_{\mathcal{W}^{1-1/p,p}(\partial\Omega)} \\ &+ \|\psi|x|V(\varepsilon^{-1}x)\|_{L^\infty(\partial\Omega)}\left\|\psi\frac{\nabla u}{|x|}\right\|_{\mathcal{W}^{1-1/p,p}(\partial\Omega)}\Bigg) \ .\end{aligned}$$

We have, using the vanishing derivative at a vertex in the Taylor formula, that,

$$\psi\frac{\nabla u}{|x|} = \psi\frac{x}{|x|}.\int_0^1 \bigtriangledown\bigtriangledown u(tx)\, dt \in L^\infty(\partial\Omega)\cap\mathcal{W}^{1-1/p,p}(\partial\Omega), \quad \forall p < 2 \ .$$

Consider a vertex that is the intersection of $K^1 \cap \partial\Omega$ and $K^2 \cap \partial\Omega$. Hence for $j \neq 1$, we have from the decay properties of v^k_{bl},

$$\|\psi|x|V^j(\varepsilon^{-1}x)\|_{H^s(\partial\Omega\cap K^2)} = O(\varepsilon^m) \quad \forall m, s \ .$$

In addition,

$$\|\psi|x|V(\varepsilon^{-1}x)\|_{L^\infty(\partial\Omega\cap K^2)} \le C\sup_x \varepsilon|x||V^1(x)| \le C\varepsilon\,.$$

Similarly,

$$\begin{aligned}\|\psi|x|V(\varepsilon^{-1}x)\|_{L^p(\partial\Omega\cap K^2)} &\le C\varepsilon^{1+1/p}\sup_{x_1}\left(\int_0^\infty |V^1(x_1,x_2)|^p\,dx_2\right)^{1/p}\\ &\le C\varepsilon^{1+1/p}\,.\end{aligned}$$

We apply the same reasoning on K_1 to turn these into corresponding bounds over the whole boundary, then applying the same argument applied to the tangential derivatives gives,

$$\|\psi|x|V(\varepsilon^{-1}x)\|_{\mathcal{W}^{1,p}(\partial\Omega)} \le C\varepsilon^{1/p}\,.$$

Putting these together we have, by interpolation of the L^p and $\mathcal{W}^{1,p}$ bounds, [Ada75, Theroem 5.2],

$$\|\psi|x|V(\varepsilon^{-1}x)\|_{\mathcal{W}^{1-1/p,p}(\partial\Omega)} = O(\varepsilon^{2/p})\,,$$

which completes the bound for these terms when d=2.

When d=3, edges (meeting places of 2 faces) and vertices (meeting places of 3 faces) are considered differently. It is convenient to use locally cylindrical coordinates similar to $d=2$ on an edge far from a vertex and spherical type coordinates for a vertex. For instance for a vertex, look at $x = rs$, for r radius and $s \in \partial\tilde{\Omega}$, where the tilde denotes those of unit length, then we are studying $(\psi|rs|V)(\varepsilon^{-1}x)$. Bounds of identical orders are obtained in this fashion.

So we just need a constant order bound on the $\mathcal{W}^{1-1/p,p}(\partial\Omega)$ norm of,

$$\phi_w = -u_2(x,\varepsilon^{-1}x) - \sum_{k=1}^N w_{bl}^k(x,\varepsilon^{-1}x)\,.$$

Explicitly, ϕ_w is given by, by construction of the $w_{bl,\varepsilon}^k$ for suitable W, $\tilde{W}$,

$$\begin{aligned}&W(\varepsilon^{-1}x)\bigtriangledown\bigtriangledown u(x) + \tilde{W}(\varepsilon^{-1})\bigtriangledown\tilde{u}_1(x)\\ =&(-\Upsilon + \sum_{k=1}^N W^k(\varepsilon^{-1}x))\bigtriangledown\bigtriangledown u(x) + (-\chi(\varepsilon^{-1}x) + \sum_{k=1}^N \tilde{W}^k(\varepsilon^{-1}x))\bigtriangledown\tilde{u}_1(x)\,,\end{aligned}$$

and unlike the V^k, by the properties of $w^k_{bl,\varepsilon}$, W^k and $\tilde{W}^k$ converge to constant limits (not necessarily 0), given by $W^{k,\infty}$ and $\tilde{W}^{k,\infty}$ respectively. Hence,

$$\phi^\infty_w = \sum_{j\neq k} W^{j,\infty} \bigtriangledown \bigtriangledown u(x) + \sum_{j\neq k} \tilde{W}^{j,\infty}(\varepsilon^{-1}) \bigtriangledown \tilde{u}_1(x) \ ,$$

is introduced on K^k and then it is verified by the assumptions made that we have $\|\phi^\infty_w\|_{\mathcal{W}^{1-1/p,p}(\Omega)} < \infty$. Then we can consider $\phi_w - \phi^\infty_w$ which has fields which converge to 0 like ϕ_v except since only a constant order bound is required it is not necessary to use the vanishing of derivatives at a vertex and the prefactor of $|x|$. □

In fact using boundary layer terms to choose $\tilde{u}_1$, we can 'outperform' the degree of the expansion in the case of [MGV08] with regard to interior estimates and even on the entire open set Ω in L^2.

Theorem 2.6.17. *In the situation as above, noting the bounds on u_2 in $L^2(\Omega)$, we have,*

$$\|u^\varepsilon - u(x) - \varepsilon u_1(x, \varepsilon^{-1}x) - \varepsilon \sum_{k=1}^{N} v^k_{bl,\varepsilon}(x, \varepsilon^{-1}x)\|_{L^2(\Omega)} \leq C\varepsilon^2 \ ,$$

and hence noting the bounds on $v^k_{bl,\varepsilon}(x, \varepsilon^{-1}x)$ in the interior

$$\|u^\varepsilon - u(x) - \varepsilon u_1(x, \varepsilon^{-1}x)\|_{L^2(\omega)} \leq C\varepsilon^2 \ .$$

This completes this example. Note that at certain key points, assumptions on u and $\tilde{u}_1$ are brought to bear plus the approach (as far as bounds on the boundary layers) is influenced greatly by the shape of Ω.

In the chapter 5 we will attempt a probabilistic solution to a similar problem in order to show that via such an approach it might be possible to dispense with a lot of the technicality of the above result as well as relaxing the small divisor assumption on the gradients of the sides relative to the period.

2.7 Applications of Multiscale expansions

Aside from the purely mathematical applications of multiscale expansions, such expansions have found important applications in aiding the computational analysis of heterogeneous materials. Composite materials have found many applications in everything from aviation, defence (composite armor plated vehicle hulls), shipbuilding, the auto industry (weight reduction compared to steel/aluminium) to prosthetic limbs [Chu10]. The importance of composite materials is illustrated for instance in the aviation industry, uses of composite materials are not limited to lightweight wing materials but structures such as the keel beam for the now defunct Comanche helicopter program [MTSM98]. Although the macroscopic properties of such materials are well understood by standard engineering techniques, the computational methods produced with the aid of AEH (asymptotic expansion homogenization) as it is known, attempt to provide an understanding of how the microstructure of the material translates into macroscopic properties. This could be used to reduce cost in producing composite materials for novel applications by allowing prediction of material properties at a design stage. Hence the models used for the microstructure must exhibit sufficient complexity in order to model local dynamics whilst at the same time retaining enough simplicity to allow tractable numerical simulation over the macroscopic scale [CTN01].

A standard problem in the analysis of composite elastic materials is [LPW95], which illustrates that numerical solution of the cell problem is a problem to be studied also and this work exposes a couple of methods for the numerical solution of the cell problem. These are then implemented to obtain the effective stiffness tensor and local stress variations of a unidirectional elastic fibre composite.

The AEH method can also be applied to composite materials that are reinforced with a periodic grid of orthotropic (properties vary in different orthogonal directions) reinforcements, see for instance [KHGS09] where this problem is considered in 3 dimensions and different geometries of the reinforcements are explored. In a further extension to composite materials, in further work, [KG02], Kalamkarov studies the properties of smart materials using the same methods. Smart composite materials are composite materials that incorporate sensors and actuators. Dependent on whether they are passively or actively controlled, they either report information on their integrity, stress, etc, or report and make adjustments respectively.

Multiscale methods can even be used to derive non static properties of composite materials. For instance this is the case in [CTN00] (arising from the juxtaposition of dissimilar materials at a micro level) for heterogeneous woven-fabric layered media. The Kelvin-Voigt model of a heterogeneous material (a model commonly used to model creep) is employed in a constant stress creep (deformation under stress) simulation using a finite element numerical scheme applied to integral equations derived from AEH that are translated into variational equations to which computational techniques can be employed. The fact that deformation over time is analyzed means that there is not merely a single homogenization calculation to carry out.

I hope the above examples serve to illustrate to the reader the importance and flexibility of the asymptotic expansion method of homogenization when at first glance the method seems rather simplistic.

2.8 Weakness of Multiscale expansions

Although as illustrated above, multiscale methods are very powerful and flexible, with many applications, they have a couple of drawbacks. One of these is if there are not globally defined correctors on which to apply the formal side of the method which is then rigorously justified then there is no clear way to apply them to the problem in hand. This seemed to be the case in [HM10a] where it was not even known if there was a globally defined first order corrector.

The second presumption of multiscale expansions is that there is a clear separation between scales, which is not the case in a number of applications such as geology, differential effective medium theory and proving superdiffusivity for turblent diffusions. For more details see the references in [BAO03], where such a problem is analyzed in the context of proving the anomalous slow diffusion property of motion in a potential with many scales that have no clear separation which is accomplished via estimates on the effective diffusivity.

2.9 Degenerate Diffusion Coefficients

Up until this point, all the homogenization problems have been strictly elliptic. This makes things a lot easier but is not necessary, although it should be noted that this is not a particularly large field of homogenization since in all degenerate cases some compensatory assumptions must be made.

The difficulty with non uniformly elliptic differential operators is producing a corrector with sufficient regularity. This can be easily overcome with the assumption of hypoellipticity.

Definition 2.9.1 (Hypoellipticity of a differential operator [Hör61])**.** A differential operator $\mathcal{L}$ with coefficients in C^∞ is called hypoelliptic if

$$\mathcal{L}h = f \ ,$$

has only solutions (in an appropriate distributional sense) in C^∞ for $f \in C^\infty$.

But even the absence of hypoellipticity this can be overcome with certain (relatively non-restrictive) assumptions.

One approach to this problem is to proceed as in [HP08], where in context of the usual framework from [BLP78] there is one main problem. This is to show that we have the existence of a corrector $\mathcal{L}\hat{b} = -b$ for the centered function b (centered with respect to the weak limit of invariant probability measures of the SDEs satisfied by $\tilde{X}^\varepsilon$ below, μ) with sufficient regularity. The family of processes X_x^ε are the solutions to the following family of SDEs,

$$X_x^\varepsilon(t) = x + \int_0^t \frac{1}{\varepsilon} b\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) + c\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) ds + \int_0^t \sigma\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) dW(s) \ .$$

Then we let $\tilde{X}_x^\varepsilon(t) = \frac{1}{\varepsilon} X_x^\varepsilon(\varepsilon^2 t)$ so that,

$$\tilde{X}_x^\varepsilon(t) = \frac{x}{\varepsilon} + \int_0^t b\big(\tilde{X}_x^\varepsilon(s)\big) + \varepsilon c\big(\tilde{X}_x^\varepsilon(s)\big)\, ds + \int_0^t \sigma\big(\tilde{X}_x^\varepsilon(s)\big)\, dW(s) \ ,$$

for some Brownian motion W. We give $\tilde{X}_x^\varepsilon(t)$ the obvious meaning as a markov process on $\mathbb{T}^d$ and resultant semigroup for $\varepsilon \geq 0$. If $\tilde{X}_x^\varepsilon(t)$ has invariant measure μ^ε, we will denote the weak limit as $\mu = \mu^0$.

It is here that the hypoelliptic and degenerate cases within the general framework established in [BLP78] diverge. In both cases if we denote the semigroup of $\tilde{X}_x^0$ by $\mathcal{P}^t$, the prospective solution (this form is quite standard, [PV01, PV05]) is given by,

$$\hat{b}(x) = \int_0^\infty \mathcal{P}^t b(x)\, dt \ ,$$

since we have,

$$\mathbb{E}_x[f(\tilde{X}^0(t))] = \mathcal{L}\int_0^t \mathbb{E}_\cdot[f(\tilde{X}^0(s))]\,ds + f(x)\ ,$$

hence in a distributional sense, $\hat{b}$ is the solution; if we let $\phi \in C^\infty(\mathbb{T}^d)$ then, denoting the inner product in $L^2(\mathbb{T}^d)$ by $(\cdot,\cdot)$,

$$(\mathcal{L}^\star\phi, \int_0^t \mathbb{E}_\cdot[f(\tilde{X}^0(s))]\,ds) = (\phi, -f + \mathbb{E}_\cdot[f(\tilde{X}^0(t))])\ .$$

Taking the limit as $t \to \infty$, gives $\forall\phi \in C^\infty(\mathbb{T}^d)$,

$$(\mathcal{L}^\star\phi, \int_0^\infty \mathbb{E}_\cdot[f(\tilde{X}^0(s))]\,ds) = (\phi, -f)\ .$$

Both this limit and the existence of $\hat{b}$ is guaranteed by the exponential convergence in time of $\mathcal{P}_t f$ to $\mu(f)$ via the exponential convergence of the semigroup to the invariant regime in the supremum norm on $C^1(\mathbb{T}^d)$ [HP08, Lemma 2.6]. Then it just remains to show that this is indeed the solution. This involves showing that it has the required regularity properties. This is considerably more involved in the degenerate case. Since in the hypoelliptic case all we have to do now is appeal to ellipticity and our solution $\hat{b}$ in the distributional sense has to be a $C^\infty(\mathbb{T}^d)$ solution.

In the degenerate case, in [HP08], it follows from [HP08, Lemma 2.6] that $\hat{b} \in C^1(\mathbb{T}^d)$ which is shown by using Malliavin calculus together with the assumptions on the Hörmander condition holding on an accessible open set and good behavior of the Jacobian of the stochastic flow associated to the process with respect to the accessibility of this set. Then a regularization procedure (convolution with a smooth function with compact support which by means of increasing a parameter is made to converge in a distributional sense to a delta function) is used to show $\hat{b}$ can be used like a twice differentiable corrector.

From the exponential convergence of the semigroup of $\tilde{X}^\varepsilon_x(t)$ to μ^ε, we get,

Theorem 2.9.2. *We have for $f \in L^\infty(\mathbb{T}^d)$ for any $t > 0$,*

$$\int_0^t f\left(\frac{X^\varepsilon_x(s)}{\varepsilon}\right) ds = t\int_{\mathbb{T}^d} f(x)\,\mu(dx)\ , \tag{2.9.1}$$

in probability as $\varepsilon \to 0$.

This is shown from the exponential convergence in time to the invariant measure μ^ε of $\tilde{X}^\varepsilon_x$ in terms of total variation convergence of the semigroup [HP08, Corollary 2.3] together with the weak convergence of μ^ε to μ [HP08, Lemma 2.4] (cf [Par99, Proposition 2.4]).

With the existence of a corrector and the result of (2.9.1), the convergence result follows an identical pattern as before in the non-degenerate case.

This completes a brief summary of the main fields of homogenization. We briefly introduce the skew Brownian motion that we will use in our homogenization results before providing some background on the oscillator problem we will be solving in the final chapter.

2.10 A word on the skew Brownian motion

Since the skew Brownian motion is closely related to the limiting processes in the main results of this thesis, it seems that a few words on the topic of this process are in order.

The skew Brownian motion is related via a rescaling to the weak limit of the solutions of (2.2.1) in one dimension and the weak limit in multiple dimensions is a generalization of this. There exist multiple descriptions of skew Brownian motion, putting aside the intuitive interpretation initially introduced by Itô and McKean [IM65] of a skew Brownian motion of parameter p as a Wiener process reflected at 0 converted into a process on $\mathbb{R}$ by placing an excursion to the right of 0 with probability p and to the left with probability $(1-p)$, independently of the Wiener process. Commonly [HS81], skew Brownian motion is referred to as the solution to the SDE,

$$X(t) = W(t) + \beta L^0(t) \,, \qquad (2.10.1)$$

where W is a standard Wiener process, L^0 is the two sided local time at 0 of X and $\beta = 2p - 1$. The local time of a process is defined as follows,

Definition 2.10.1 (Local time [RY91])**.** By generalizing the standard Itô formula (2.1.1) to convex functions, we have the Tanaka formula. For a continuous semimartingale $X(t)$, we have a continuous increasing process such that, for $u \in \mathbb{R}$,

$$\begin{aligned}
|X(t)-u| =& |X_0-u| + \int_0^t \operatorname{sgn}(X(s)-u)\, dX(s) + L^u(t) \,, \\
(X(t)-u)^+ =& (X_0-u)^+ + \int_0^t 1_{\{X(s)>u\}}\, dX(s) \\
& + \frac{1}{2}\int_0^t 1_{\{X(s)=0\}}\, dX(s) + \frac{L^u(t)}{2} \,, \\
(X(t)-u)^- =& (X_0-u)^- - \int_0^t 1_{\{X(s)<u\}}\, dX(s) \\
& - \frac{1}{2}\int_0^t 1_{\{X(s)=0\}}\, dX(s) + \frac{L^u(t)}{2} \,,
\end{aligned}$$

for sgn $=1$ if $x>0$, sgn $=0$ if $x=0$ and -1 otherwise. $L^u(t)$ above is then termed the two sided local time of X at u at time t. There is also the concept of a left or right local time, for which the value of sgn at 0 is defined as -1, 1 respectively and there is a corresponding alteration in the subsequent equations also. We will use the two sided local time by default and from this point drop the preceding 'two sided'. Note that for a martingale all local times are equivalent but not for a semimartingale, since we have continuity of $L^u(t)$ in u $\forall t$ for a martingale but not for a semimartingale, although we still have left and right limits in u for all local times in this case. $L^u(t)$ is a continuous increasing process with $L^u(0)=0$ a.s. for all u. In fact the measure $dL^u(t)$ is supported entirely within the set $\{t: X(t)=u\}$ i.e. $L^u(t)$ increases only on this set. From the occupation times formula we have an alternative definition for the local time of a continuous semimartingale which is,

$$L^u(t) = \lim_{\varepsilon\to 0}\frac{1}{2\varepsilon}\int_0^t 1_{(u-\varepsilon,u+\varepsilon)}(X(s))\, d\langle X,X\rangle(s) \,,$$

and from this second definition it is easy to see why it is known as the local time at a point u. The right (left) local time is defined by a similar formula except we only divide by ε instead of 2ε and consider the occupation of the set $(u,u+\varepsilon)$ (respectively $(u-\varepsilon,u)$).

As alluded to above, the motivation for the second definition of the local time is provided by the very useful occupation times formula,

Theorem 2.10.2 (Occupation Times Formula [RY91]). *There is a $\mathbb{P}$ negligible set out-*

side of which,

$$\int_0^t \Phi(X(s))\, d\langle X, X\rangle = \int_{-\infty}^{\infty} \Phi(u) L^u(t)\, du\ ,$$

for every t and Borel function Φ.

Denote the left local time ($\mathrm{sgn}(0) = -1$) by $\hat{L}$. When combined with the following result which gives the continuity properties alluded to before, the occupation times formula (which follows from the first definition) is sufficient to easily derive the second definition of the local time from the first definition,

Theorem 2.10.3 ([RY91])**.** *For any continuous semimartingale X there exists a modification of the process $\{\hat{L}^u(t) : u \in \mathbb{R}, t \in [0, \infty)\}$ such that the map $(u, t) \to \hat{L}^u(t)$ is a.s. continuous in t and cadlag in u. Moreover if $X = V + M$ is the decomposition of X into its finite variation V, and martingale part M, then,*

$$\hat{L}^u(t) - \hat{L}^{u-}(t) = 2\int_0^t 1_{(X(s)=u)}\, dV(s) = 2\int_0^t 1_{(X(s)=u)}\, dX(s)\ . \tag{2.10.2}$$

Thus in particular if X is a local martingale there exists a bicontinuous modification of the family L^u of local times.

A modification of a stochastic process is defined as follows:

Definition 2.10.4. Two processes X, X' defined on the same probability space are defined as modifications of one another if for every $t \geq 0$, $X(t) = X'(t)$ a.s..

The modification above in Theorem 2.10.3 is usually taken to be the definition of the left local time and we will do the same. It is also clear from the occupation times formula, we have,

$$L^a(t) = \frac{1}{2}(\hat{L}^a(t) + \hat{L}^{a-}(t))\ ,$$

hence $\hat{L}^{a-}(t))$ is then the right local time.

Moving back to the skew Brownian motion, in [HS81] (2.10.1) is shown to have a unique solution (unique in law) for $|\beta| \leq 1$. $\beta = 1$ corresponds to a reflected Brownian motion and $|\beta| < 1$ is a true skew Brownian motion. In fact there is pathwise uniqueness of solution for $|\beta| < 1$ by [BC05, Theorem 2.2] and $|\beta| = 1$ by [BC05, Theorem 4.1].

Theorem 2.10.5 (Theorem 2.2 [BC05]). *Suppose a is a measurable function on $\mathbb{R}$ that is bounded above and below by positive constants and suppose there is a strictly increasing function f on $\mathbb{R}$ such that,*

$$|a(x) - a(y)|^2 \leq |f(x) - f(y)| \quad x, y \in \mathbb{R} . \tag{2.10.3}$$

Denote the two sided local time of x at u by L^u. Let μ be a finite signed measure on $\mathbb{R}$ such that $\mu(\{x\}) < 1 \; \forall x \in \mathbb{R}$. Then for each x_0, $u \in \mathbb{R}$, the SDE,

$$X(t) = x_0 + \int_0^t a(X(s))\, dW_s + \int L^u(t)\mu(du) \quad t \geq 0 , \tag{2.10.4}$$

has a continuous strong solution and the continuous solution is pathwise unique.

The following theorem gives existence and pathwise uniqueness in the case $|\beta = 1|$.

Theorem 2.10.6 (Theorem 4.1 [BC05]). *Suppose a is a measurable function on $\mathbb{R}$ that is bounded above and below by positive constants, and satisfies condition (2.10.3). Then for any x_0, $u \in \mathbb{R}$, the SDE*

$$X(t) = x_0 + \int_0^t a(X(s))\, dW(s) + L^u(t) , \tag{2.10.5}$$

has a strong continuous solution and the solution is pathwise unique. The same conclusion holds for the SDE

$$X(t) = x_0 + \int_0^t a(X(s))\, dW(s) - L^u(t) . \tag{2.10.6}$$

Strong solutions and pathwise uniqueness are defined as follows,

Definition 2.10.7 (Strong Solution [RY91]). A solution is known as a strong solution of an SDE driven by a Brownian motion W if it is adapted with respect to the filtration of W.

Definition 2.10.8 (Pathwise Uniqueness [RY91]). An SDE is said to have a pathwise unique solution if, for X, X' solutions with respect to the same driving Brownian motion and $X(0) = X'(0)$ a.s., then $X(t) = X'(t) \; \forall t \geq 0$ a.s., i.e. X and X' are indistinguishable.

This brief introduction to skew Brownian motion concludes with a look at the martingale problem satisfied by Brownian motion. We will need the result of the Itô-Tanaka formula,

Definition 2.10.9 (Itô-Tanaka formula [RY91])**.** If f is the difference of two convex functions and if X is a continuous semimartingale,

$$f(X(t)) = f(X_0) + \int_0^t \frac{1}{2}(f'_- + f'_+)(X_s)\, dX(s) + \frac{1}{2}\int_{\mathbb{R}} L^u(t)\, f''(du)\ .$$

In particular, $f(X)$ is a semimartingale.

Given $f \in C^2(\mathbb{R}\setminus\{0\}) \cap C^0(\mathbb{R})$, if we apply the Itô-Tanaka formula combined with the occupation times formula to $f(X(t))$, for $X(t)$ the solution of (2.10.1), we obtain that $f(X(t))$ is equal to,

$$\begin{aligned} & f(X_0) + \int_0^t f'(X(s))\, dW(s) + \frac{1}{2}\int_0^t f''(X(s))\, ds \\ & + \frac{\beta}{2}(f'_- + f'_+)L^0(t) + \frac{1}{2}(f'_+ - f'_-)L^0(t) \\ = & f(X_0) + \int_0^t f'(X(s))\, dW(s) + \frac{1}{2}\int_0^t f''(X(s))\, ds \\ & + \frac{1}{2}[f'_-(\beta - 1) + f'_+(\beta + 1)]L^0(t)\ . \end{aligned}$$

This implies that for any function f for which,

$$f'_-(\beta - 1) + f'_+(\beta + 1) = 0\ , \tag{2.10.7}$$

we have that,

$$f(X(t)) - \frac{1}{2}\int_0^t f''(X(s))\, ds - f(X(0))\ ,$$

is a martingale. Therefore the skew Brownian motion is the solution to the martingale problem given by the operator $(1/2)\partial_x^2$ on the subset of $C^2(\mathbb{R}\setminus\{0\}) \cap C^0(\mathbb{R})$ satisfying (2.10.7). The domain of the generator is in fact sufficiently large to characterize the skew Brownian motion [HM10b], i.e. the domain of definition is sufficiently large to guarantee uniqueness in law of the solution to the martingale problem.

2.11 The Intersection of Homogenization and Oscillator Problems: Correctors

At first glance the properties of chains of coupled oscillators and homogenization are two problems that would appear to have little in common. However, the use of correctors in some coupled oscillator chain problems provides common ground.

In homogenization, the purpose of a corrector term is to replace increasingly large oscillations ($O(\varepsilon^{-1})$) in the drift with an $O(\varepsilon)$ term as the parameter $\varepsilon \to 0$ in order that some form of averaged behavior can be derived in the limit. However, in certain circumstances in order to obtain a good description of the behavior of coupled oscillator chains in a dominating proportion of the regime canceling increasingly large, fast oscillations in the velocity variable can be useful. For instance in [Hai09], the following system of two coupled oscillators is considered,

$$\begin{aligned}
dq_i &= p_i \, dt, \quad i = \{0, 1\}, \\
dp_0 &= -V_1'(q_0)\, dt + \alpha(q_1 - q_0)\, dt - \gamma p_0 \, dt + \sqrt{2\gamma T}\, dW_0(t), \\
dp_1 &= -V_1'(q_1) + \alpha(q_0 - q_1)\, dt + \sqrt{2\gamma T_\infty}\, dW_1(t)
\end{aligned}$$

where W_0, W_1 are two independent Wiener processes and V_1 is given by,

$$V_1(q) = \frac{|q|^k}{2k}.$$

We have the existence of the corrector Φ from [HM09, Proposition 3.7], the unique centered solution of $\mathcal{L}\Phi = -q$. $\mathcal{L}$ is the generator for an isolated oscillator in potential V_1, given by,

$$\mathcal{L} = p\partial_q - q|q|^{2k-2}\partial_p$$

Where we call a function centered if it is centered in the sense of [HM09] in that the average over one period of the function in the case of an isolated oscillator is zero. If ψ scales like H^α and averages out to 0, then the solution to $\mathcal{L}\phi = \psi$, ϕ scales like $H^{\alpha+1/2k-1/2}$.

Hence in the case $k \geq 2$, since q scales like $H^{1/2k}$, Φ scales like $H^{1/k-1/2}$ and $1/k - 1/2 \leq 0$ for such k, the equations for the oscillators can be seen to exhibit approximate decoupling at high energies, by setting

$$\tilde{p}_0 = p_0 + \alpha\Phi(p_1, q_1)$$

in order to remove the q_1 (coupling) term from the SDE satisfied by p_0 at the expense of the addition of a constant order corrector term and a number of error terms that converge to zero in terms of absolute value, coefficient of 0 variables or quadratic variation as appropriate with increasing energy in the 1 (undamped) variables,

$$\begin{aligned} d\tilde{p}_0 = &- V_1'(q_0)\,dt + -\alpha q_0\,dt - \gamma p_0\,dt + \sqrt{2\gamma T}\,dw_0(t) \\ &+ \alpha\partial_p\Phi(q_0 - q_1)\,dt + \frac{1}{2}\partial_p^2\Phi\,dt + \partial_p\Phi\,dw_1(t). \end{aligned}$$

The aforementioned convergence of the second line of terms is given by [HM09, Proposition 3.5] which states that if ψ scales like H^α then $\partial_p\psi$ scales like $H^{\alpha-1/2}$ and $\partial_q\psi$ scales like $H^{\alpha-1/2k}$.

However the effect of approximate decoupling is not sufficiently strong to violate existence of an invariant measure in the case $k = 2$, $T < \alpha^2\langle\Phi^2\rangle$. The decoupling is effect is sufficiently strong for $k > 2$ or $k = 2$, $T > \alpha^2\langle\Phi^2\rangle$ and there is no invariant measure in this case. See [Hai09] for details.

In fact we take advantage greatly of this decoupling using the above correction together with a few others in the chapter 6 where we deal with the cases $k > 2$ and $k = 2$, $T > \alpha^2\langle\Phi^2\rangle$ and show that the undamped oscillator has energy that behaves similarly to the square of a Brownian motion. The reluctance of the undamped component to occupy a compact set together with the decoupling at high energies of the undamped oscillator is then used to show the convergence of the damped oscillator to an invariant distribution.

Chapter 3

Periodic Homogenization with an interface: the one dimensional case

This chapter is the solution of the one dimensional case of the problem given in the introduction by (6.0.1). It is a joint work with Martin Hairer due to appear in Stochastic Processes and Applications [HM10b].

We consider a one-dimensional diffusion process with coefficients that are periodic outside of a finite 'interface region'. The question investigated in this article is the limiting long time / large scale behaviour of such a process under diffusive rescaling. Our main result is that it converges weakly to a rescaled version of skew Brownian motion, with parameters that can be given explicitly in terms of the coefficients of the original diffusion.

Our method of proof relies on the framework provided by Freidlin and Wentzell [FW93] for diffusion processes on a graph in order to identify the generator of the limiting process. The graph in question consists of one vertex representing the interface region and two infinite segments corresponding to the regions on either side.

3.1 Introduction

Consider a diffusion process in $\mathbb{R}^d$ of the type

$$dX(t) = b(X)\,dt + dB(t)\;, \quad X(0) \in \mathbb{R}^d\;, \tag{3.1.1}$$

where B is a d-dimensional Wiener process and $b\colon \mathbb{R}^d \to \mathbb{R}^d$ is periodic and smooth, and define the diffusively rescaled process $X^\varepsilon(t) = \varepsilon X(t/\varepsilon^2)$. If b is periodic and satisfies a natural centering condition, then it is well-known that X^ε converges in law as $\varepsilon \to 0$ to a Wiener process with a 'diffusion tensor' that can be expressed in terms of the solution to a suitable Poisson equation, see for example [BLP78, PS08].

Similar types of homogenization results still hold true if b is not exactly periodic, but of the form $b(X) = \tilde{b}(X, \varepsilon X)$, for some smooth function $\tilde{b}$ that is periodic in its first argument. In other words, b consists of a slowly varying component, modulated by fast oscillations. In this case, the limiting process is not a Brownian motion anymore, but can be an arbitrary diffusion, whose coefficients can again be obtained by a suitable averaging procedure [BMP05]. The aim of this article is to consider a somewhat different situation where there is an abrupt change from one type of periodic behaviour to another, separated by an interface of size order one in the original 'microscopic' scale.

To the best of our knowledge, this situation has not been considered before, although a similar problem was studied in [ACP03]. In order to keep calculations simple, we restrict ourselves here to the one-dimensional situation. Building on this analysis, we are able to address the multidimensional case in [HM10a]. Restricting ourselves to the one-dimensional case considerably simplifies the analysis due to the following two facts:

- Any one-dimensional diffusion is reversible [IM65], so that its invariant measure can be given explicitly, enabling us to make a reasonable guess at the limiting process (see below).
- The 'interface' is a zero-dimensional object, so that it cannot exhibit any internal structure in the limit.

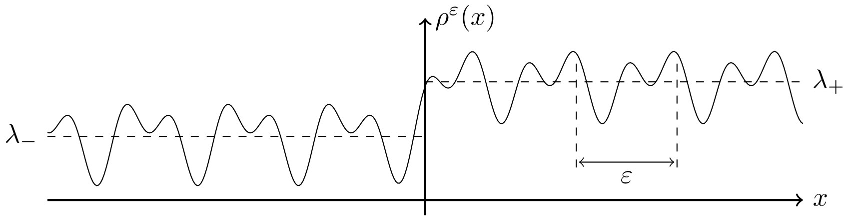
Before we give a more detailed description of our results, let us try to 'guess' what any limiting process X^0 should look like, if it exists. Away from the interface, we can apply the existing results on periodic homogenization, as in [BLP78, PS08]. We can therefore compute diffusion coefficients $C_\pm$ such that X^0 is expected to behave like $C_+ W(t)$ whenever $X^0 > 0$ and like $C_- W(t)$ whenever $X^0 < 0$, for some Wiener process W. One possible way of constructing a Markov process with this property is to take $X^0(t) = G(W(t))$, where $G\colon \mathbb{R} \to \mathbb{R}$ is given by

$$G(x) = \begin{cases} C_+ x & \text{if } x \geq 0, \\ C_- x & \text{otherwise.} \end{cases}$$

In turns out that processes of this form do not describe all the possible limiting processes that one can get in the presence of an interface. The reason why this is so can be seen by comparing the invariant measure of $G(W(t))$ to the invariant measure of X^ε. Since the invariant measure for W is Lebesgue measure (or multiples thereof), the invariant measure for $G(W(t))$ is given by

$$\mu(dx) = \begin{cases} \lambda_+ \, dx & \text{if } x > 0, \\ \lambda_- \, dx & \text{otherwise,} \end{cases} \qquad \text{with} \quad \frac{\lambda_+}{\lambda_-} = \frac{C_-}{C_+} \,. \tag{3.1.2}$$

On the other hand, if we denote the invariant measure for X^ε by $\mu^\varepsilon = \rho^\varepsilon(x)\, dx$, then ρ^ε will typically look as follows:



It follows that one does indeed have $\mu^\varepsilon \to \mu^0$ as $\varepsilon \to 0$, where μ^0 is of the type (3.1.2), but the ratio λ_+/λ_- depends on the behaviour of b, not only away from the interface, but also at the interface. This can be understood as the process X^0 picking up an additional drift, proportional to the local time spent at 0, that skews the proportion of time spent on either side of the interface. A Markov process with these properties can be constructed by applying the function G to a skew-Brownian motion (see for example [Lej06]) with parameter p for a suitable value of p.

An intuitive way of constructing this process goes as follows. First, draw the zeroes of a standard Wiener process on the real line. These form a Cantor set that partitions the line into countably many disjoint open intervals. Order them by decreasing length and denote by I_n the length of the nth interval. For each $n \geq 0$, toss an independent biased coin and draw an independent Brownian excursion. If the coin comes up heads (with probability p), fill the interval with the Brownian excursion, scaled horizontally by I_n and vertically by $C_+\sqrt{I_n}$. Otherwise (with probability $1-p$), fill the interval with the Brownian excursion, scaled horizontally by I_n and vertically by $-C_-\sqrt{I_n}$. One can check that the invariant measure for this process is given by (3.1.2), but with

$$\frac{\lambda_+}{\lambda_-} = \frac{p\, C_-}{(1-p)\, C_+} \;. \tag{3.1.3}$$

We denote the corresponding process by $B_{C_\pm,p}(t)$.

This should almost be sufficient to guess the main result of this article. To fix notations, we consider the process $X(t)$ as in (3.1.1) and its rescaled version X^ε, and we assume that the drift function b is smooth and periodic away from an 'interface' region $[-\eta,\eta]$. More precisely, we assume that there exist smooth periodic functions $b_i : \mathbb{R} \to \mathbb{R}$, $i \in \{+,-\}$, such that $b_i(x+1) = b_i(x)$ and such that $b(x) = b_+(x-\eta)$ for $x > \eta$ and $b(x) = b_-(x+\eta)$ for $x < -\eta$. Additionally, we assume that the functions b_i satisfy the centering condition

$$\int_0^1 b_i(x)\, dx = 0 \;.$$

We also set $V(x) = \int_0^x b(x)\, dx$ for $x \in \mathbb{R}$, so that $\exp(2V(x))\, dx$ is invariant for X, and similarly for V_i. Denote by C_i the effective diffusion coefficients for the periodic homogenization problems corresponding to b_i (see equation (3.3.7) below or [PS08] for a more explicit expression). Define furthermore $\lambda_\pm$ by

$$\lambda_+ = \int_\eta^{\eta+1} \exp(2V(x))\, dx \;, \qquad \lambda_- = \int_{-\eta-1}^{-\eta} \exp(2V(x))\, dx \;, \tag{3.1.4}$$

and let $p \in (0,1)$ be the unique solution to (3.1.3). With all these notations at hand, we have:

Theorem 3.1.1. *For any $t > 0$, the law of X^ε converges weakly to the law of $B_{C_\pm,p}$ in the space $\mathcal{C}([0,t],\mathbb{R})$.*

Remark 3.1.2. *In order to keep notations simple, we have assumed that the diffusion coefficient of X is constant and equal to 1. The case of a non-constant, but smooth and uniformly elliptic diffusion coefficient can be treated in exactly the same way, noting that it reduces to the case treated here after a time change that can easily be controlled.*

Remark 3.1.3. *A natural extension of the results presented in this article is case of a random potential, along the lines of the situation first considered in [PV82]. While the heuristic argument presented in the introduction still applies, the method of proof considered here does not seem to apply readily.*

An alternative method would be to consider an injective harmonic function H_ε *for* X^ε, *so that* $H_\varepsilon(X^\varepsilon(t))$ *is a martingale for every* ε. *In the situation at hand, one can take*

$$H_\varepsilon(x) = \int_0^x e^{2V(y/\varepsilon)}\,dy\;.$$

One might hope that in this case it is possible to show that the process $H^\varepsilon(t) = H_\varepsilon(X^\varepsilon(t))$ *then converges as* $\varepsilon \to 0$ *to the martingale* $H(t)$ *given by*

$$H(t) = \int_0^t A(H(s))\,dB(s)\;,$$

where $A(H)$ *takes one constant value for* $H > 0$ *and a different constant value for* $H < 0$ *(see also Section 3.4 below). This type of approach has been successfully applied to a number of multiscale problems, including homogenisation on fractals, see for example [Koz93, Zhi95, KK96, HK98, Owh03, Kum04].*

Remark 3.1.4. *Another natural extension of the results presented in this article is case of a random potential, along the lines of the situation first considered in [PV82].*

The proof of the weak convergence of the probability distributions on $\mathcal{C}[0,\infty)$ associated to X_x^ε presented in this article will rely heavily on the 1993 paper by Freidlin and Wentzell [FW93], in which the authors consider a 'fast' Hamiltonian system perturbed by a 'slow' diffusion. Theorems 2.1 and 4.1 from [FW93] provide a general framework for proving first the tightness and then the convergence of a family of probability distributions on $\mathcal{C}[0,\infty)$.

We will start by showing tightness of our family of processes in Section 3.2. Once tightness is established, we show in Section 3.3 that every limiting process X^0 solves the martingale problem associated to a certain generator. Finally, we show in Section 3.4 that this martingale problem has a unique solution which is precisely the rescaled skew-Brownian motion, thus concluding the proof.

3.2 Proving tightness

The main result of this section is the following:

Theorem 3.2.1. *The family of probability measures on* $\mathcal{C}([0,\infty),\mathbb{R})$ *given by the laws of* X_x^ε *for* $\varepsilon \in (0,1]$ *is tight.*

Proof. Given that the initial condition is kept fixed at one single point across the entire family of laws, tightness follows from uniformity in the modulus of continuity over $\varepsilon \in (0,1]$.

For $x \in \mathbb{R}$ and $\rho > 0$, denote by τ_ρ^x the first exit time of the canonical process from the interval $[x-\rho, x+\rho]$. We also denote by $\mathbb{P}_{x,\varepsilon}$ the law of X_x^ε. It then follows immediately from the proof of [SV79, Theorem 1.4.6] that a sufficient criterion for tightness is that, for every $\rho > 0$, there exists a constant A_ρ such that the bound

$$\mathbb{P}_{x,\varepsilon}(\,\tau_\rho^x \leq \delta) \leq \delta A_\rho\;, \tag{3.2.1}$$

holds uniformly over all $x \in \mathbb{R}$, $\varepsilon \in (0,1]$, and $\delta \in \mathbb{R}$. Before we proceed, we note the following two crucial facts:

1. It follows from the periodic case [BLP78, Section 3.4] that (3.2.1) holds uniformly for $x \notin (-\varepsilon\eta - \rho, \varepsilon\eta + \rho)$. We denote the corresponding constants by A_ρ^1.

2. A standard martingale argument as in [SV79, Section 1.4] shows that, for every $\varepsilon_0 > 0$, the bound (3.2.1) holds uniformly over all $x \in \mathbb{R}$, provided that we restrict ourselves to $\varepsilon \in [\varepsilon_0, 1]$. For the sequel of the proof it will be convenient to make the choice $\varepsilon_0 = \rho/(4\eta)$ and we denote the corresponding constants by A_ρ^2.

Combining these two facts, we see that it remains to find a family of constants A_ρ^3 such that (3.2.1) holds for $x \in (-\varepsilon\eta - \rho, \varepsilon\eta + \rho)$ and for $\varepsilon < \rho/(4\eta)$. At this stage we note that since ρ is greater than twice the width of the interface, for every $x \in \mathbb{R}$ there exist two points $\tilde{x}_\pm$ with the following two properties:

1. The process started at x has to hit either $\tilde{x}_+$ or $\tilde{x}_-$ before it can reach the boundary of the interval $[x-\rho, x+\rho]$.

2. The intervals $I_\pm = [\tilde{x}_\pm - \frac{\rho}{8}, \tilde{x}_\pm + \frac{\rho}{8}]$ satisfy $I_\pm \cap (-\eta\varepsilon, \eta\varepsilon) = \emptyset$ and $I_\pm \subset [x-\rho, x+\rho]$.

Restarting the process when it hits one of the $\tilde{x}_\pm$, it follows from the strong Markov property that we can choose $A_\rho^3 = A_{\rho/8}^1$, which concludes the proof by setting $A_\rho = \max\{A_\rho^i\}$. □

3.3 Convergence of the laws

As usual in the theory of homogenization, we do not show directly that the processes X^ε converge to a limit, but we first introducing a compensator $g\colon \mathbb{R} \to \mathbb{R}$ that 'kills' the strong drift of the rescaled process and consider instead the family of processes

$$Y^\varepsilon(t) = X^\varepsilon(t) + \varepsilon g\Big(\frac{X^\varepsilon(t)}{\varepsilon}\Big) . \tag{3.3.1}$$

Since we will choose g to be a bounded function, the weak convergence in the space of continuous functions of the laws of X^ε to some limiting process is equivalent to that of the Y^ε.

In order to construct g, let $\mathcal{L}_i$ denote the generator of the diffusion with drift b_i, that is $\mathcal{L}_i = \frac{1}{2}\partial_x^2 + b_i(x)\partial_x$, and denote by,

$$\mu_i(dx) = Z^{-1}\exp(2V_i(x))\,dx ,$$

the corresponding invariant probability measure on $[0,1]$. We then denote by g_i the unique smooth function solving

$$\mathcal{L}_i g_i = -b_i , \quad \int_0^1 g_i(x)\,\mu_i(dx) = 0 . \tag{3.3.2}$$

Since b is assumed to be centred on either side of the interface, such a function exists (and is unique) by the Fredholm alternative. We now choose *any* smooth function $g\colon \mathbb{R} \to \mathbb{R}$ such that $g(x) = g_-(x+\eta)$ for $x \in (-\infty, -\eta)$ and $g(x) = g_+(x-\eta)$ for $x \in (\eta, \infty)$, with a smooth joining region in between.

The main ingredient in our proof of convergence will be [FW93, Theorem 4.1], which is used in conjunction with the previous tightness result to identify the weak limit points of the family of probability distributions as the solutions to a martingale problem. The aim of this section is to explain how to fit our problem into the framework of [FW93] and to verify the assumptions of their main convergence theorem.

Before we proceed, let us recall what is understood by the "martingale problem" corresponding to some operator A (see for example [EK86]), and let us try to guess what the generator A for the limiting process is expected to be. Let X be a Polish (i.e. complete separable metric) space; $\mathcal{C}[0,\infty)$, the space of all continuous functions on $[0,\infty)$ with values in X. For any subset $I \subset [0,\infty)$, denote by $\mathscr{F}_I$ the σ-algebra of subsets of $\mathcal{C}[0,\infty)$ generated by the sets $\{x \in \mathcal{C}[0,\infty) : x(s) \in B\}$, where $s \in I$ and $B \subset X$ is an arbitrary Borel set. We also denote by $\mathcal{C}(X)$ the space of all continuous real-valued functions on X.

Let A be a linear operator on $\mathcal{C}(X)$, defined on a subspace $\mathscr{D}(A) \subseteq \mathcal{C}(X)$. We will say that a probability measure $\mathbb{P}$, on $\big(\mathcal{C}[0,\infty), \mathscr{F}_{[0,\infty)}\big)$, is a solution to the martingale problem corresponding to A, starting from a point $x_0 \in X$, if

$$\mathbb{P}\{x : x(0) = x_0\} = 1 \tag{3.3.3}$$

and, for any $f \in \mathscr{D}(A)$, the random function defined on the probability space $\big(\mathcal{C}[0,\infty), \mathscr{F}_{[0,\infty)}, \mathbb{P}\big)$ by

$$f\big(x(t)\big) - \int_0^t \big(Af\big)\big(x(s)\big)\, ds,\ t \in [0,\infty)\ , \tag{3.3.4}$$

is a martingale with respect to the filtration $\big\{\mathscr{F}_{[0,t]}\big\}_{t>0}$.

What do we expect the operator A to be given by in our case? On either side of the interface, we argued in the introduction that the limiting process should be given by Brownian motion, scaled by factors $C_\pm$ respectively. Therefore, one would expect A to be given by

$$\big(Af\big)(x) = \begin{cases} \frac{1}{2}C_-^2 \partial_x^2 f(x) & \text{if } x < 0, \\ \frac{1}{2}C_+^2 \partial_x^2 f(x) & \text{otherwise,} \end{cases} \tag{3.3.5}$$

and the domain $\mathscr{D}(A)$ to contain functions that are $\mathcal{C}^2$ away from the origin. This however does not take into account for the "skewing", which should be encoded in the behaviour of functions in $\mathscr{D}(A)$ at the origin.

Since the limiting process spends zero time at the origin (the invariant measure is continuous with respect to Lebesgue measure), it was shown in [Lej06] that the possible behaviours at the origin are given by matching conditions for the first derivatives of functions belonging to $\mathscr{D}(A)$. We know from the introduction that the invariant measure of the limiting process is proportional to Lebesgue measure on either side of the origin, with proportionality constants $\lambda_\pm$. We should therefore have the identity

$$\lambda_- \int_{-\infty}^0 Af(x)\,dx + \lambda_+ \int_0^\infty Af(x)\,dx = 0\;,$$

for every function $f \in \mathscr{D}(A)$. Using (3.3.5), we thus obtain

$$\lambda_- C_-^2 \int_{-\infty}^0 f''(x)\,dx + \lambda_+ C_+^2 \int_0^\infty f''(x)\,dx = 0\;.$$

Integrating by parts, this yields (for say compactly supported test functions f) the condition

$$\lambda_- C_-^2 f'(0^-) = \lambda_+ C_+^2 f'(0^+)\;. \tag{3.3.6}$$

This is exactly the general form of a generator produced by Theorem 4.1 in Freidlin and Wentzell [FW93], a differential operator on the regions away from some distinguished points termed nodes, combined with a restriction on the ratios of the limits of the derivatives at this point.

The main theorem of this section that is also very closely related to the main theorem of the article is as follows:

Theorem 3.3.1. *Let $C_\pm$ be given by*

$$C_\pm^2 = \int_0^1 (1 + g_\pm(x))^2\, \mu_\pm(dx)\;, \tag{3.3.7}$$

where $g_\pm$ and $\mu_\pm$ are as in (3.3.2). Let A be given by (3.3.5) and let $\mathscr{D}(A)$ be the set of continuous functions $f\colon \mathbb{R} \to \mathbb{R}$, vanishing at infinity, that are $\mathcal{C}^2$ away from 0 and that satisfy the condition (3.3.6) at the origin.

Then, every limit point of the family of processes Y_x^ε is solution to the martingale problem corresponding to A.

As already mentioned, our main ingredient is [FW93, Theorem 4.1] applied to the sequence of processes Y^ε as defined in (3.3.1). For completeness, we give a simplified statement of this result adapted to the situation at hand:

Theorem 3.3.2 (Freidlin & Wentzell). *Let $\mathcal{L}_i$, $i = \pm$, be elliptic second order differential operators with smooth coefficients on I_i, $I_+ = [0,\infty)$, $I_- = (-\infty, 0]$, and let Y^ε be a family of real-valued processes satisfying the strong Markov property. For some fixed $\tilde{\eta} > 0$, let τ^ε be the first hitting time of the set $(-\varepsilon\tilde{\eta}, \varepsilon\tilde{\eta})$ by Y^ε.*

Assume that there exists a function $k\colon \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{\varepsilon\to 0} k(\varepsilon) = 0$ such that, for any function $f \in \mathcal{C}_v^\infty(I_i)$ and for any $\lambda > 0$, one has the bound[1]

$$\begin{aligned}&\mathbb{E}_y\Big[e^{-\lambda\tau^\varepsilon} f\big(Y^\varepsilon(\tau^\varepsilon)\big) - f(y) + \int_0^{\tau^\varepsilon} e^{-\lambda t}\Big(\lambda f\big(Y^\varepsilon(t)\big) - \mathcal{L}_i f\big(Y^\varepsilon(t)\big)\Big)\,dt\Big] \\ &= \mathcal{O}\big(k(\varepsilon)\big) \end{aligned} \tag{3.3.8}$$

as $\varepsilon \to 0$, uniformly with respect to $y \in I_i$. Assume furthermore that there exists a function $\delta\colon \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{\varepsilon\to 0}\delta(\varepsilon) = 0$ and $\lim_{\varepsilon\to 0}\delta(\varepsilon)/k(\varepsilon) \to \infty$ such that, for any $\lambda > 0$,

$$\mathbb{E}_y\Big[\int_0^\infty e^{-\lambda t} 1_{(-\delta,\delta)}\big(Y^\varepsilon(t)\big)\,dt\Big] \to 0 \tag{3.3.9}$$

as $\varepsilon \to 0$, uniformly over all $y \in R$. Finally, writing σ^δ for the first hitting time of the set $(-\infty,-\delta)\cup(\delta,\infty)$ by Y^ε, assume that there exist $p_i \geq 0$ with $p_- + p_+ = 1$ such that

$$\mathbb{P}_y\big[Y^\varepsilon(\sigma^\delta) \in I_i\big] \to p_i\,, \quad i \in \{+,-\}\,, \tag{3.3.10}$$

uniformly for $y \in (-\varepsilon\tilde{\eta}, \varepsilon\tilde{\eta})$.

Let now A be the operator defined by $Af(x) = \mathcal{L}_i f(x)$ for $x \in I_i$ with domain $\mathscr{D}(A)$ consisting of functions f such that $f|_{I_i} \in \mathcal{C}_v^\infty(I_i)$ and such that the 'matching condition' $p_+ f'(0^+) = p_- f'(0^-)$ holds. Then for any fixed $t_0 \geq 0$, any $\lambda > 0$, and any $f \in \mathscr{D}(A)$, the bound

$$\begin{aligned}\operatorname{ess\,sup}\Big|&\int_{t_0}^\infty e^{-\lambda t}\mathbb{E}_y\Big[\lambda f\big(Y^\varepsilon(t)\big) - Af\big(Y^\varepsilon(t)\big)\Big|\mathscr{F}_{[0,t_0]}\Big]\,dt \\ &- e^{-\lambda t_0} f\big(Y^\varepsilon(t_0)\big)\Big| \to 0\end{aligned} \tag{3.3.11}$$

holds as $\varepsilon \to 0$, uniformly for all $y \in \mathbb{R}$.

Remark 3.3.3. *The version of Theorem 3.3.2 stated in [FW93] does actually treat more general diffusions on graphs, but assumes that the edges of the graph are finite. This is*

[1] We denote by $\mathcal{C}_v^\infty$ the space of smooth functions that vanish at infinity, together with all of their derivatives.

not really a restriction, since I_+ is in bijection with $[0,1)$ (and similarly for I_-) and we can simply add non-reachable vertices at ± 1 to turn our process into a process on a finite graph.

Remark 3.3.4. *As can be seen by combining (3.1.3) and (3.3.6), the probabilities $p_\pm$ appearing in the statement of Theorem 3.3.2 are not quite the same in general as the probabilities $\tilde p_\pm = \{p, 1-p\}$ appearing in the construction of skew Brownian motion in the introduction. The relation between them is given by $\frac{pC_+}{(1-p)C_-} = \frac{p_+}{p_-}$. The reason is that $p_\pm$ give the respective probabilities of hitting two points located at a fixed distance from the 'interface', whereas the non-trivial scaling of the Brownian bridges on either side of the interface means that $\tilde p_\pm$ give the probabilities of hitting two points whose distances from the interface have the ratio C_+/C_-.*

Most of the remainder of this section is devoted to the fact that:

Proposition 3.3.5. *The family of processes Y^ε given by (3.3.1) satisfies the assumptions of Theorem 3.3.2 with $\mathcal{L}_\pm = \frac{1}{2}C_\pm\partial_x^2$ and $p_\pm$ defined by the relations*

$$p_+ + p_- = 1\,, \qquad \frac{p_+}{p_-} = \frac{\lambda_+ C_+^2}{\lambda_- C_-^2}\,,$$

and $\lambda_\pm$ as in (3.1.4).

This yields the

Proof of Theorem 3.3.1. Before we start, let us remark that the initial condition y for the corrected process Y^ε and the initial condition x for the original process X^ε are related by $y = x + \varepsilon g(x/\varepsilon)$.

Note also that, thanks to the identity

$$\int_{t_0}^\infty e^{-\lambda s} F(s)\,ds = \int_{t_0}^\infty \lambda e^{-\lambda t}\int_{t_0}^t F(s)\,ds\,dt\,,$$

valid for any bounded measurable function F, the left hand side in (3.3.11) can be written as

$$\begin{aligned}\Delta(\varepsilon) &= \int_{t_0}^\infty \lambda e^{-\lambda t}\mathbb{E}_y\Big(f(Y^\varepsilon(t)) - f(Y^\varepsilon(t_0)) - \int_{t_0}^t Af(Y^\varepsilon(s))\,ds\,\Big|\,\mathscr{F}_{[0,t_0]}\Big)\,dt \\ &\stackrel{\text{def}}{=} \int_{t_0}^\infty \lambda e^{-\lambda t}\mathbb{E}_y\big(\mathcal{G}_f(Y^\varepsilon, t_0, t)\,\big|\,\mathscr{F}_{[0,t_0]}\big)\,dt\,.\end{aligned}$$

We have already established the weak precompactness of the family $\{\mathbb{P}_{x,\varepsilon},\ \varepsilon > 0\}$ in the space $\mathcal{C}(\mathbb{R}_+,\mathbb{R})$. The uniformity in x of the convergence of $\Delta(\varepsilon)$ to 0 then implies that for any n, any $0 \leq t_1 < \cdots < t_n \leq t_0$, and any bounded measurable function $G(x_1,\dots,x_n)$, $x_i \in \mathbb{R}$,

$$\left| \mathbb{E}_y\Big(G(Y^\varepsilon(t_1),\dots,Y^\varepsilon(t_n)) \cdot \int_{t_0}^\infty \lambda e^{-\lambda t} \mathcal{G}_f(Y^\varepsilon,t_0,t)\,dt\Big)\right| \leq \sup\left|G\right| \cdot \Delta(\varepsilon)\ . \quad (3.3.12)$$

If we furthermore assume that G is continuous, then the expression inside the expectation is a continuous function on $\mathcal{C}(\mathbb{R}_+,\mathbb{R})$, so that any accumulation point X^0 satisfies

$$\int_{t_0}^\infty \lambda e^{-\lambda t} \mathbb{E}\Big(G\big(X^0(t_1),\dots,X^0(t_n)\big)\mathcal{G}_f(X^0,t_0,t)\Big)\,dt = 0\ . \quad (3.3.13)$$

Since the integrand is a continuous function of t and a continuous function is determined uniquely by its Laplace transform, this implies that

$$\mathbb{E}\big(G(\dots)\mathcal{G}_f(X^0,t_0,t)\big) = 0\ ,$$

for all n and $0 \leq t_1 < \cdots < t_n \leq t_0$, so that in particular the random function $f(X^0(t)) - \int_0^t Af(X^0(s))\,ds$ is indeed a martingale in the filtration generated by the process X^0.

Since the laws of the starting points of X^ε are all equal to δ_x by construction, we conclude that the law of X^0 is indeed a solution of the martingale problem corresponding to A, starting from x_0. □

Proof of Proposition 3.3.5. The proofs of (3.3.8), (3.3.9) and (3.3.10) will be given as three separate propositions.

Proposition 3.3.6. *There exists $\tilde\eta > 0$ such that the process $Y^\varepsilon(t)$ satisfies (3.3.8) with $k(\varepsilon) = \varepsilon$, that is,*

$$\begin{aligned}\mathbb{E}_y\Big[& e^{-\lambda\tau^\varepsilon} f\big(Y^\varepsilon(\tau^\varepsilon)\big) - f(x) \\ & + \int_0^{\tau^\varepsilon} e^{-\lambda t}\Big(\lambda f\big(Y^\varepsilon(t)\big) - \frac{1}{2}C_+^2 f''\big(Y^\varepsilon(t)\big)\Big)\,dt\Big] = \mathcal{O}(\varepsilon)\ ,\end{aligned}$$

for every function $f \in \mathcal{C}_0^\infty(\mathbb{R}_+)$, uniformly in $y \in [\tilde\eta\varepsilon,\infty)$, and similarly for the left side of the interface.

Proof. The treatment of both sides of the interface is identical, so we restrict ourselves to $\mathbb{R}_+$. As before, the initial condition y for the corrected process Y^ε and the initial condition x for the original process X^ε are related by $y = x + \varepsilon g(x/\varepsilon)$. Note that one has $g'_\pm(x) \neq -1$ for any x since otherwise, by uniqueness of the solutions to the ODE $g'' = -2(1+g')b$, this would entail that $g'_\pm(x) = -1$ over the whole interval $[0,1]$, in contradiction with the periodic boundary conditions.

By possibly making η slightly larger, we can (and will from now on) therefore assume that $g'(x) > -1$ uniformly over $x \in \mathbb{R}$, so that the correspondence $x \leftrightarrow y$ is a bijection. Since g is bounded, this shows that one can find $\tilde\eta > 0$ so that $y \notin [-\varepsilon\tilde\eta, \varepsilon\tilde\eta]$ implies that $x \notin [-\varepsilon\eta, \varepsilon\eta]$. In particular, fixing such a value for $\tilde\eta$ from now on, we see that the drift vanishes in the SDE satisfied by Y^ε, provided that we consider the process only up to time τ^ε.

Using the integration by parts formula and Itô's formula for each Y^ε_y we get

$$\begin{aligned} e^{-\lambda t} f\big(Y^\varepsilon_y(\tau^\varepsilon)\big) = f(x) &+ \int_0^{\tau^\varepsilon} e^{-\lambda s}\big(1 + g'_+(\varepsilon^{-1}X^\varepsilon_x(s))\big) f'\big(Y^\varepsilon_y(s)\big)\, dB_s \\ &- \int_0^{\tau^\varepsilon} e^{-\lambda s}\Big[\lambda f\big(Y^\varepsilon_y(s)\big) \\ &\qquad + \frac{1}{2}\big(1 + g'_+(\varepsilon^{-1}X^\varepsilon_x(s))\big)^2 f''\big(Y^\varepsilon_y(s)\big)\Big]\, ds\;. \end{aligned} \tag{3.3.14}$$

Since the expectation of the stochastic integral vanishes (both f' and g'_+ are uniformly bounded), all that remains to be shown is that the last term in the above equation converges at rate ε to the same term with $(1+g')^2$ replaced by C_+^2.

This will be a consequence of the following result (variants of which are quite standard in the theory of periodic homogenization), which considers the fully periodic case. It is sufficient to consider this case in the situation at hand since we restrict ourselves to times before τ^ε, so that the process does not 'see' the interface.

Lemma 3.3.7. *Let $b\colon \mathbb{R}^n \to \mathbb{R}^n$ be smooth and periodic with fundamental domain $\Lambda \subset \mathbb{R}^n$, and denote by μ the (unique) probability measure on Λ invariant for the SDE*

$$dX(t) = b(X(t))\, dt + dB_t\;, \quad X(0) = x\;, \tag{3.3.15}$$

where B is a standard d-dimensional Wiener process. Assume furthermore that $\int_\Lambda b(x)\,\mu(dx) = 0$*. (This condition will be referred to in the sequel as b being centred.)*

Let $h\colon \mathbb{R}^d \to \mathbb{R}$ *be any smooth function that is periodic with fundamental domain* Λ *and centred. Let furthermore* $X^\varepsilon(t) = \varepsilon X(t/\varepsilon^2)$ *and let* τ_ε *be a family of (possibly infinite) stopping times with respect to the natural filtration of B. Then, for every* $F \in \mathcal{C}_0^4(\mathbb{R}^d, \mathbb{R})$ *there exists* $C > 0$ *independent of* τ_ε *such that the bound*

$$\mathbb{E}_x\left[\int_0^{\tau_\varepsilon} e^{-\lambda s} F\big(X^\varepsilon(s)\big) h\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\,ds\right] \le C\varepsilon\,,$$

holds for any $\varepsilon \in (0,1]$*, uniformly in x.*

Proof. Denote by $\mathcal{L} = \frac{1}{2}\Delta + \langle b(x), \nabla\rangle$ the generator of (3.3.15) and let g be the unique periodic centred solution to $\mathcal{L}g = h$. (Such a solution exists by the Fredholm alternative.) Applying Itô's formula to the process

$$\varepsilon^2 e^{-\lambda t} F\big(X^\varepsilon(t)\big) g\Big(\frac{X^\varepsilon(t)}{\varepsilon}\Big)\,,$$

we obtain the identity

$$\begin{aligned}
\varepsilon^2 e^{-\lambda\tau_\varepsilon} F\big(X^\varepsilon(\tau_\varepsilon)\big) g\Big(\frac{X^\varepsilon(\tau_\varepsilon)}{\varepsilon}\Big) &= \int_0^{\tau_\varepsilon} e^{-\lambda s} F\big(X^\varepsilon(s)\big) h\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\,ds \\
&\quad + \varepsilon^2 \int_0^{\tau_\varepsilon} -\lambda e^{-\lambda s} F\big(X(s)\big) g\Big(\frac{X(s)}{\varepsilon}\Big)\,ds \\
&\quad + \varepsilon \int_0^{\tau_\varepsilon} e^{-\lambda s} b\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big).\,\nabla F\big(X^\varepsilon(s)\big) g\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\,ds \\
&\quad + \frac{1}{2}\varepsilon^2 \int_0^{\tau_\varepsilon} e^{-\lambda s} \Delta F\big(X^\varepsilon(s)\big) g\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\,ds \\
&\quad + \varepsilon^2 \int_0^{\tau_\varepsilon} e^{-\lambda s} \nabla F\big(X^\varepsilon(s)\big) g\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\,dB_s \\
&\quad + \varepsilon \int_0^{\tau_\varepsilon} e^{-\lambda s} F\big(X^\varepsilon(s)\big) \nabla g\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\,dB_s \\
&\quad + \varepsilon \int_0^{\tau_\varepsilon} e^{-\lambda s}\Big(\nabla F\big(X^\varepsilon(s)\big).\,\nabla g\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\Big)\,ds\,.
\end{aligned}$$

The claim then follows by taking expectations and noting that all the functions of X^ε appearing in the various terms are uniformly bounded. □

Returning to the proof of Proposition 3.3.6, we first note that $f''(Y_y^\varepsilon) = f''(X_x^\varepsilon) + \mathcal{O}(\varepsilon)$, so that we can replace $f''(Y_y^\varepsilon)$ by $f''(X_x^\varepsilon)$ in (3.3.14), up to errors of $\mathcal{O}(\varepsilon)$. Applying Lemma 3.3.7 with $F = f''$ and $h = (1+g'_+)^2 - C_+^2$, the claim then follows at once. □

Proposition 3.3.8. *The convergence*

$$\mathbb{E}\Big[\int_0^\infty e^{-\lambda t} 1_{(-\sqrt{\varepsilon},\sqrt{\varepsilon})}\big(Y_y^\varepsilon(t)\big)\, dt\Big] \to 0 \tag{3.3.16}$$

takes place as $\varepsilon \to 0$, uniformly in the initial point $y \in \mathbb{R}$. In particular, (3.3.9) holds with $\delta(\varepsilon) = \sqrt{\varepsilon}$.

Proof. The main idea is to first perform a time-change that turns the diffusion coefficient of Y^ε into 1 and to then compare the resulting process to the process V^ε which is the solution to

$$dV^\varepsilon = b_V^\varepsilon(V^\varepsilon)\, dt + dB_t\ , \tag{3.3.17}$$

where the drift b_V^ε is given by

$$b_V^\varepsilon(x) \;=\; \begin{cases} -\frac{C_V}{\varepsilon} & \text{for } 0 \le x \le \hat{\eta}\varepsilon, \\ \frac{C_V}{\varepsilon} & \text{for } -\hat{\eta}\varepsilon \le x < 0, \\ 0 & \text{otherwise,} \end{cases} \tag{3.3.18}$$

for C_V and $\hat{\eta}$ some positive constants independent of ε to be determined below. An explicit resolvent equation then allows one to show that (3.3.16) with Y replaced by V tends to zero as $\varepsilon \to 0$, uniformly in the initial point.

First, let us start with the time change. As in the proof of Proposition 3.3.6, we assume that g is chosen in such a way that g' is bounded away (from below) from -1, so that there exists a constant $\phi > 0$ such that $g'(x) \ge \phi - 1$ for every $x \in \mathbb{R}$. In order to turn the diffusion coefficient of Y^ε into 1, we use the time change associated with the quadratic variation of the Y^ε,

$$\langle Y^\varepsilon, Y^\varepsilon\rangle(t) = \int_0^t \Big(1 + g'\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\Big)^2\, ds\ ,$$

thus setting

$$C_t^\varepsilon = \inf\big\{t' > 0 : \langle Y^\varepsilon, Y^\varepsilon\rangle(t') > t\big\}\ .$$

Defining the function $\hat{b} = b + \mathcal{L}g$, the process $Z^\varepsilon(t) = Y^\varepsilon(C_t^\varepsilon)$ then satisfies the equation

$$Z^\varepsilon(t) = y + \int_0^{C_t^\varepsilon} \frac{1}{\varepsilon}\,\hat{b}\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\,ds + \int_0^{C_t^\varepsilon} \Big(1 + g'\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\Big)\,dB_s\;.$$

Note that the time-change was defined precisely in such a way that the second term in this expression is equal to some Brownian motion $B^\varepsilon(t)$. Inserting the expression for the time change, the first term can be rewritten as

$$\int_0^{C_t^\varepsilon} \frac{1}{\varepsilon}\,\hat{b}\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\,ds = \int_0^t \frac{1}{\varepsilon}\,\hat{b}\Big(\frac{X^\varepsilon(C_s^\varepsilon)}{\varepsilon}\Big)\Big(1 + g'\Big(\frac{X^\varepsilon(C_s^\varepsilon)}{\varepsilon}\Big)\Big)^{-2}\,ds\;.$$

It follows that the drift term is non-zero only when the time-changed process occupies the region $(-\varepsilon\eta - \varepsilon\|g\|_\infty, \varepsilon\eta + \varepsilon\|g\|_\infty)$ just as for the non time-changed process. To summarise, there exists a drift $\tilde{b}$ bounded uniformly by $\frac{C_V}{\varepsilon}$ for some constant $C_V > 0$ and vanishing outside of $(-\tilde{\eta}\varepsilon, \tilde{\eta}\varepsilon)$ for $\tilde{\eta} = \eta + \|g\|_\infty$, as well as a Brownian motion B^ε, so that the process Z_x^ε satisfies the SDE

$$dZ_x^\varepsilon = \tilde{b}(Z_x^\varepsilon)\,dt + dB_t^\varepsilon\;, \qquad Z_x^\varepsilon(0) = y\;. \tag{3.3.19}$$

Now, look at how the time change affects the expression (3.3.16), where we set $G^\varepsilon = (-\sqrt{\varepsilon}, \sqrt{\varepsilon})$:

$$\begin{aligned}\int_0^\infty e^{-\lambda t}1_{G^\varepsilon}\big(Y^\varepsilon(t)\big)\,dt &= \int_0^{C_\infty^\varepsilon} e^{-\lambda t}1_{G^\varepsilon}\big(Y^\varepsilon(t)\big)\,dt \\ &= \int_0^\infty e^{-\lambda t}1_{G^\varepsilon}\big(Z^\varepsilon(t)\big)\Big(1 + g'\Big(\frac{X^\varepsilon(C_t^\varepsilon)}{\varepsilon}\Big)\Big)^{-2}\,dt \\ &\le \sup_{x\in\mathbb{R}}\big(1 + g'(x)\big)^{-2}\int_0^\infty e^{-\lambda t}1_{G^\varepsilon}\big(Z^\varepsilon(t)\big)\,dt\;.\end{aligned}$$

Hence if it can be shown that,

$$\mathbb{E}_y\Big[\int_0^\infty e^{-\lambda t}1_{G^\varepsilon}\big(Z^\varepsilon(t)\big)\,dt\Big] \to 0 \tag{3.3.20}$$

uniformly in the initial point x for the underlying process X_x^ε, as $\varepsilon \to 0$, then our claim follows. The idea is to bound (3.3.20) by the 'worst-case scenario' obtained by replacing the process Z^ε by the process V^ε described in (3.3.17).

One technical problem that arises is that it is tricky to get pathwise control on the behaviour of V due to the discontinuity of its drift. We therefore first compare Z^ε with the process U_x^ε solution to

$$dU^\varepsilon = b_U^\varepsilon(U^\varepsilon)\,dt + dB_U^\varepsilon(t)\;, \quad U_x^\varepsilon(0) = y\;, \tag{3.3.21}$$

where B_U^ε is a Brownian motion to be determined and b_U^ε is the Lipschitz continuous odd function defined on the positive real numbers by

$$b_U^\varepsilon(x) = \begin{cases} -\frac{C_V}{\varepsilon^2}\, x & \text{for } |x| \leq \varepsilon, \\ -\frac{C_V}{\varepsilon} & \text{for } \varepsilon < x \leq (2+\tilde{\eta})\varepsilon, \\ -\frac{C_V}{\varepsilon^2}\left((3+\tilde{\eta})\varepsilon - x\right) & \text{for } (2+\tilde{\eta})\varepsilon < x \leq (3+\tilde{\eta})\varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

The SDE (3.3.21) satisfies pathwise uniqueness, which is why we are using it as an intermediary between Z^ε and V^ε. Now, what we are going to do is, given a realisation of the Brownian motion B^ε driving Z^ε in (3.3.19), to choose the Brownian motion B_U^ε driving U_x^ε by changing the sign of the increments in such a way that the absolute value of U_x^ε is always less than or equal to $|Z_x^\varepsilon| + 2\varepsilon$. By pathwise uniqueness, we are indeed free to choose the Brownian motion in (3.3.21). The choice of the Brownian motion is the content of the following lemma:

Lemma 3.3.9. *For every initial condition x, there exists a map $B^\varepsilon \mapsto B_U^\varepsilon$ that preserves Wiener measure and such that $|U_x^\varepsilon| \leq |Z_x^\varepsilon| + 2\varepsilon$ almost surely. In particular, it follows that (3.3.20) is bounded by*

$$\mathbb{E}_x\Big(\int_0^\infty e^{-\lambda t} 1_{(-\delta',\delta')}\big(U_x^\varepsilon(t)\big)\, dt\Big)\;, \tag{3.3.22}$$

where $\delta' = \delta + 2\varepsilon$.

Proof. The construction works in the following way. Consider first the processes driven by the same realisation B^ε and define a stopping time τ_0 by $\tau_0 = \inf\{t > 0 : |U_x^\varepsilon(t)| = |Z_x^\varepsilon(t)| + 2\varepsilon\}$. This stopping time is strictly positive and one has $|U_x^\varepsilon(\tau_0)| \geq 2\varepsilon$. For times after τ_0, we determine B_U^ε by

$$B_U^\varepsilon(t) = B_U^\varepsilon(\tau_0) + sign\big(U_x^\varepsilon(\tau_0)\big) \int_{\tau_0}^t sign\big(Z_x^\varepsilon(s)\big)\, dB^\varepsilon(s)\;,$$

and we introduce the stopping time $\sigma_1 = \inf\{t > \tau_0 : |U_x^\varepsilon| = \varepsilon\}$. Since by construction U^ε does not change sign between τ_0 and σ_1, it then follows from the Itô-Tanaka formula that up to σ_1 one has

$$d|U^\varepsilon| = b_U^\varepsilon\big(|U^\varepsilon|\big)\, dt + sign\big(Z^\varepsilon\big)\, dB^\varepsilon(t)\;,$$

$$d|Z^\varepsilon| = sign(Z^\varepsilon)\tilde{b}^\varepsilon(Z^\varepsilon)\,dt + sign(Z^\varepsilon)\,dB^\varepsilon(t) + dL(t)\;,$$

for some local time term L. Since the local time term always yields positive contributions and since it follows from the definition that $b^\varepsilon_U(u) \le sign(z)\tilde{b}^\varepsilon(z)$ for $|u| \ge \varepsilon$ and $|u| \le |z| + 2\varepsilon$, we can apply a simple comparison result for SDEs to conclude that the inequality $|U^\varepsilon_x| \le |Z^\varepsilon_x| + 2\varepsilon$ holds almost surely between times τ_0 and σ_0.

We then drive again both processes by the same noise and define as before τ_1 by $\tau_1 = \inf\{t > \sigma_1 : |U^\varepsilon_x(t)| = |Z^\varepsilon_x(t)| + 2\varepsilon\}$. Note that $\tau_1 > \sigma_1$ almost surely since one has $|U^\varepsilon_x(\sigma_0)| \le |Z^\varepsilon_x(\sigma_0)| + \varepsilon$. We then apply the previous construction iteratively, so that, setting $\sigma_0 = 0$, we have constructed B^ε_U by

$$B^\varepsilon_U(t) = \int_0^t \Big(\sum_{n=0}^\infty 1_{[\sigma_n,\tau_n)}(s) + \sum_{n=0}^\infty 1_{[\tau_n,\sigma_{n+1})}(s)\; sign\big(U^\varepsilon_x(\tau_n)Y^\varepsilon_x(s)\big)\Big)\,dB^\varepsilon(s)\;.$$

Since the process U^ε has finite quadratic variation and has to move by at least ε between any two successive stopping times, our sequence of stopping times does converge to infinity, so that $B^\varepsilon_U(t)$ is indeed a Brownian motion with the required property. □

In our next step, we compare the process U^ε_x that we just constructed with the process V^ε_x defined in (3.3.17), where we set $\hat{\eta} = 5 + \tilde{\eta}$. Since the drift coefficient is bounded, it follows from an application of Girsanov's theorem like in [RY91, Corollary IX.1.12] that this SDE has a solution for some Brownian motion B^ε_V, say.

We now fix B^ε_V and use it to construct a Brownian motion B^ε_U driving (3.3.21) in such a way that the absolute value of V^ε_x always stays less than $|U^\varepsilon_x| + 2\varepsilon$:

Lemma 3.3.10. *There exists a map $B^\varepsilon_V \mapsto B^\varepsilon_U$ that preserves Wiener measure and such that $|V^\varepsilon_x| \le |U^\varepsilon_x| + 2\varepsilon$ for all times almost surely. In particular, (3.3.22) is bounded by*

$$\mathbb{E}_x\Big(\int_0^\infty e^{-\lambda t} 1_{(-\delta'',\delta'')}(V^\varepsilon_x(t))\,dt\Big)\;, \tag{3.3.23}$$

with $\delta'' = \delta' + 2\varepsilon$.

Proof. The argument is virtually identical to that of Lemma 3.3.9, so we do not reproduce it here. □

It now remains to show that:

Lemma 3.3.11. *The expression (3.3.23) converges to* 0 *uniformly in the initial point as* $\varepsilon \to 0$.

Proof. We write ε for $5\varepsilon + \varepsilon\eta$ and δ for $\delta + 4\varepsilon$ for ease of notation, but this has no bearing on the rates of convergence of the aforementioned quantities and hence on the calculation. We have the identity

$$\begin{aligned}\mathbb{E}_x\left[\int_0^\infty e^{-\lambda t} 1_{(-\delta,\delta)}\big(V_x^\varepsilon(t)\big)dt\right] &= \int_0^\infty e^{-\lambda t} P_t^\varepsilon\big(1_{(-\delta,\delta)}\big)(x)\,dt \\ &= (\lambda - \mathcal{L}_V^\varepsilon)^{-1} 1_{(-\delta,\delta)}(x) \qquad (3.3.24)\end{aligned}$$

by the resolvent equation (see for example [EK86, Chapter 1]), where $\mathcal{L}_V^\varepsilon$ is the generator of the Markov semigroup P_t^ε associated to V^ε.

We now proceed to computing this expression explicitly in order to show that its supremum tends to zero uniformly in x. In order to keep notations simple, we assume for the remainder of this proof that $\eta = C_V = 1$, which can always be achieved by rescaling space and redefining ε. In this case, the solution $f(x) = (\lambda - \mathcal{L}_V^\varepsilon)^{-1} 1_{(-\delta,\delta)}(x)$ to the resolvent equation can be assembled piecewise on the intervals $(-\infty, -\delta)$, $(-\delta, -\varepsilon)$, $(-\varepsilon, 0)$, $(0, \varepsilon)$, (ε, δ) and (δ, ∞) by making sure that it is $\mathcal{C}^1$ at each junction. Owing to the symmetry of the problem, the function f will be an even function of x, hence we only have to analyze it on one side of the origin.

The general solution on each interval can be written as

$$f(x) = \begin{cases} B_0 e^{-\sqrt{2\lambda}x} & \text{for } x \geq \delta, \\ \frac{1}{\lambda} + A_1 e^{\sqrt{2\lambda}x} + B_1 e^{-\sqrt{2\lambda}x} & \text{for } \varepsilon \leq x \leq \delta, \\ \frac{1}{\lambda} + \varepsilon^2 A_2 e^{\gamma_1 x} + B_2 e^{-\gamma_2 x} & \text{for } x \leq \varepsilon, \end{cases}$$

where

$$\begin{aligned}\gamma_1 &= \left(\frac{1}{\varepsilon^2} + 2\lambda\right)^{\frac{1}{2}} + \frac{1}{\varepsilon} = \frac{2}{\varepsilon} + \mathcal{O}(\varepsilon) \\ \gamma_2 &= \left(\frac{1}{\varepsilon^2} + 2\lambda\right)^{\frac{1}{2}} - \frac{1}{\varepsilon} = \lambda\varepsilon + \mathcal{O}(\varepsilon^2)\,.\end{aligned}$$

The reason for the somewhat strange choice of adding an explicit factor ε^2 in front of A_2 is justified *a posteriori* by noting that with this scaling, the matching conditions at ε and δ (as well as the fact that the derivative should vanish at the origin) yield the following linear system:

$$M \begin{pmatrix} B_0 \\ A_1 \\ B_1 \\ A_2 \\ B_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -\frac{1}{\lambda} \\ 0 \end{pmatrix} , \qquad M = \begin{pmatrix} 0 & 0 & 0 & 2 & -\lambda \\ 0 & -1 & -1 & 0 & 1 \\ 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 \end{pmatrix} + \mathcal{O}(\delta) \; .$$

To lowest order in ε and δ, this can easily be solved exactly, yielding

$$(B_0, A_1, B_1, A_2, B_2) = -\frac{1}{2\lambda}(0, 1, 1, \lambda, 2) + \mathcal{O}(\delta) \; .$$

Inserting this into the expression for f shows that $\sup_{x\in\mathbb{R}} |f(x)| = \mathcal{O}(\delta) = \mathcal{O}(\sqrt{\varepsilon})$, thus completing the proof. □

With the lemma regarding the resolvent calculation above, the proof of Proposition 3.3.8 is complete. □

We finally show that

Proposition 3.3.12. *For every $c > 0$, the exit probabilities from the interval $(-\delta, \delta)$ satisfy the bound*

$$\mathbb{P}_x\big[Y^\varepsilon(\sigma^\delta) \in I_i\big] = p_i + \mathcal{O}(\sqrt{\varepsilon}) \; ,$$

uniformly for $x \in [-c\varepsilon, c\varepsilon]$.

Proof. For the proof of this result, it turns out to be simpler to consider the original process $X^\varepsilon(t)$.

Whenever $Y^\varepsilon(t)$ exits the set $(-\delta, \delta)$, due to the deterministic relationship between the processes, $X^\varepsilon(t)$ exits a set $(-\delta', \delta'')$, where δ' and δ'' are contained in the interval $(\delta - \varepsilon\|g\|_\infty, \delta + \varepsilon\|g\|_\infty)$. Therefore, we just look at the exit of $X^\varepsilon(t)$ from an interval of this form as the computations are much easier to carry out. This is due to the simpler form of the scale function for $X^\varepsilon(t)$ compared with that of $Y^\varepsilon(t)$. It follows from [RY91, Exercise VII.3.20] that the scale function of the diffusion on $\mathbb{R}$ with generator $\mathcal{L} = \frac{1}{2}\sigma^2(x)\frac{d^2}{dx^2} + b(x)\frac{d}{dx}$ is given by,

$$s(x) = \int_c^x \exp\Big(-\int_c^y 2b(z)\sigma^{-2}(z)dz\Big)\, dy$$

where c is an arbitrary point in $\mathbb{R}$. Recall that the scale function of a real-valued process is a continuous, strictly increasing function such that for any $a < x < b$ in the set where the Markov process takes its values, one has

$$\mathbb{P}_x\big(T_b < T_a\big) = \frac{s(x) - s(a)}{s(b) - s(a)}\,,$$

where T_a, T_b are the first hitting times of the points a and b respectively. For $X^\varepsilon(t)$ we have that $\sigma = 1$ and the drift is equal to $\frac{1}{\varepsilon}\, b(x/\varepsilon)$. We are from now on going to use the notation $q_\varepsilon(y) = \big(1 + \varepsilon g(\cdot/\varepsilon)\big)^{-1}(y)$ for the transformation that allows recovery of X^ε from Y^ε. We also denote by T_a the first hitting time of the point $a \in \mathbb{R}$ by the process X^ε. We also use the shorthand notation

$$F_\varepsilon(u) = \exp\Big(-2\int_0^u \frac{1}{\varepsilon}\, b(z/\varepsilon)\, dz\Big)\,,$$

so that the scale function for X^ε is given by $s(z) = \int_0^z F_\varepsilon(y)\, dy$.

With this notation at hand, we have, for $x \in (-\delta, \delta)$, that, denoting the escape time of $Y^\varepsilon(t)$ from $(-\delta, \delta)$ by σ^δ,

$$\begin{aligned}\mathbb{P}_x\big(Y^\varepsilon(\sigma^\delta) \in I_+\big) = \mathbb{P}_x\big(T_{\delta''} < T_{-\delta'}\big) &= \frac{s\big(q_\varepsilon(x)\big) - s(-\delta')}{s(\delta'') - s(-\delta')}\\ &= \frac{\int_0^{q_\varepsilon(x)} F_\varepsilon(y)\, dy - \int_0^{-\delta'} F_\varepsilon(y)\, dy}{\int_0^{-\delta''} F_\varepsilon(y)\, dy - \int_0^{-\delta'} F_\varepsilon(y)\, dy} = \frac{\int_{-\delta'}^{q_\varepsilon(x)} F_\varepsilon(y)\, dy}{\int_{-\delta'}^{\delta''} F_\varepsilon(y)\, dy}\,.\end{aligned}$$

Noting that F_ε has the scaling property $F_\varepsilon(u) = F_1(u/\varepsilon)$, we thus obtain the identity

$$\mathbb{P}_x\big(Y^\varepsilon(\sigma^\delta) \in I_+\big) = \frac{\int_{-\delta'}^{q_\varepsilon(x)} F_1(y/\varepsilon)\, dy}{\int_{-\delta'}^{\delta''} F_1(y/\varepsilon)\, dy} = \frac{\int_{-\delta/\varepsilon}^{0} F_1(y)\, dy + \mathcal{O}(1)}{\int_{-\delta/\varepsilon}^{\delta/\varepsilon} F_1(y)\, dy + \mathcal{O}(1)}\,, \tag{3.3.25}$$

where we used the fact that $q_\varepsilon(x) = \mathcal{O}(\varepsilon)$ and F_1 is uniformly bounded, due to the fact that the functions $b_\pm$ are centred by assumption.

Note now that the effective diffusion coefficients $C_\pm$ can alternatively be expressed as [PS08, Sec 13.6]

$$C_+^2 = \Big[\int_\eta^{\eta+1} \exp\big(-2V(u)\big)\, du \int_\eta^{\eta+1} \exp\big(2V(u)\big)\, du\Big]^{-1}\,,$$

and similarly for C_-. Therefore, since F_1 is periodic away from $[-\eta, \eta]$, it follows immediately from the definitions of $\lambda_\pm$ and $C_\pm$ that

$$\int_{-N}^0 F_1(y)\, dy = \frac{1}{C_-^2\lambda_-} N + \mathcal{O}(1)\,, \qquad \int_0^N F_1(y)\, dy = \frac{1}{C_+^2\lambda_+} N + \mathcal{O}(1)\,.$$

Since $\delta = \sqrt{\varepsilon}$, combining these bounds with (3.3.25) implies that $\mathbb{P}_x\big(Y^\varepsilon(\sigma^\delta) \in I_+\big) = C_+^2\lambda_+/(C_-^2\lambda_+ + C_+^2\lambda_-) + \mathcal{O}(\sqrt{\varepsilon})$, from which the requested bound follows. □

Combining Propositions 3.3.6, 3.3.8, and 3.3.12 completes the proof of Proposition 3.3.5. □

3.4 Uniqueness and characterisation of the martingale problem

To conclude this article, we show that solution to the martingale problem corresponding to A is unique and is indeed given by the variant of skew Brownian motion constructed in the introduction.

The skew Brownian motion B_p of 'skewness' parameter p is known to have generator $\mathcal{L}_p = \frac{1}{2}\frac{d^2}{dx^2}$ on the set of functions that are continuous, twice continuously differentiable except at the origin where we have $pf'(0^+) = (1-p)f'(0^-)$. (See for example the review article [Lej06].) The process $B_{C_\pm,p}$ constructed in the introduction is given by

$$B_{C_\pm,p}(t) = G(B_p(t))\;, \qquad G(x) = \begin{cases} C_+x & \text{if } x \geq 0, \\ C_-x & \text{otherwise.} \end{cases}$$

Since G is a continuous bijection, this is again a strong Markov process and it has generator given by $Af = \big(\mathcal{L}_p(f\circ G)\big)\circ G^{-1}$ with $\mathscr{D}(A) = \{f \,:\, f\circ G \in \mathscr{D}(\mathcal{L}_p)\}$. Using the relation (3.1.3), it is now a straightforward calculation to show that $\mathscr{D}(A)$ consists precisely of those functions satisfying the derivative condition in (3.3.6), hence we will have

$$\frac{p_+}{p_-} = \frac{C_+p}{C_-(1-p)}\;.$$

This provides a clue as to how to show uniqueness easily. Let Y denote any solution of the martingale problem corresponding to A acting on $\mathscr{D}(A)$, Y also represents any possible limit point of the family Y^ε as $\varepsilon \to 0$. Defining g like G, but with constants p/C_+ and $(1-p)/C_-$ instead of C_+ and C_-, we note that g satisfies the derivative condition at 0 imposed for elements of $\mathscr{D}(A)$. Since g doesn't vanish at infinity, we have $g \notin \mathscr{D}(A)$, but we can approximate g by a sequence $g_n \in \mathscr{D}(A)$ such that $g = g_n$ on $[-n, n]$. Indeed, given $t > 0$, the probability of escaping from $[-n, n]$ before time t tends to zero as $n \to \infty$ uniformly for $\varepsilon \in (0, 1]$ by tightness. As a consequence, the process $V = g(Y)$ satisfies the SDE

$$V(t) = \int_0^t A_+ 1_{\{V(s)>0\}} + A_- 1_{\{V(s)<0\}} \, dB_s \, ,$$

where $A_\pm$ are some constants. It is known [BC05, Theorem 2.1] that this equation has a pathwise unique solution which in particular implies uniqueness in law. Since g is an invertible map, this immediately implies that Y is unique in law and one can check that Y is indeed the variant of skew Brownian motion constructed in the introduction.

Chapter 4

Periodic Homogenization with an interface: the multidimensional case

Further to the previous chapter, we now tackle the multidimensional version of the homogenization problem (6.0.1). This is a joint work with Martin Hairer to appear in the Annals of Probability.

We consider a diffusion process with coefficients that are periodic outside of an 'interface region' of finite thickness. The question investigated in this article is the limiting long time / large scale behaviour of such a process under diffusive rescaling. It is clear that outside of the interface, the limiting process must behave like Brownian motion, with diffusion matrices given by the standard theory of homogenization. The interesting behaviour therefore occurs on the interface. Our main result is that the limiting process is a semimartingale whose bounded variation part is proportional to the local time spent on the interface. The proportionality vector can have non-zero components parallel to the interface, so that the limiting diffusion is not necessarily reversible. We also exhibit an explicit way of identifying its parameters in terms of the coefficients of the original diffusion.

Similarly to the one-dimensional case, our method of proof relies on the framework provided by Freidlin and Wentzell [FW93] for diffusion processes on a graph in order to identify the generator of the limiting process.

4.1 Introduction

The theory of periodic homogenization is by now extremely well understood, see for example [BLP78, PS08]. Recall that the most basic result states that if X is a diffusion with smooth periodic coefficients, then the diffusively rescaled process $X^\varepsilon(t) = \varepsilon X(t/\varepsilon^2)$ converges in law to a Brownian motion with an explicitly computable diffusion matrix. If one considers diffusions that are 'locally periodic', but with slow modulations over spatial scales of order ε^{-1}, then it was shown in [BMP05] that the rescaled process converges in general to some diffusion process with a computable expression for both its drift and diffusion coefficients.

In this article, we will also consider the 'locally periodic' situation, but instead of considering slow modulations of the coefficients, we consider the case of a sharp (*i.e.* of size $\mathcal{O}(1)$) transition between two periodic structures. In the (much simpler) one-dimensional case, this model was previously studied in [HM10b], where we showed that the rescaled process converges in law to skew Brownian motion with an explicit expression for the skewness parameter. In higher dimensions, this model has not yet been studied to the best of our knowledge. The aim of this article is to clarify what is the behaviour of X^ε near the interface for very small values of ε. It is important to remark at this stage that we do *not* make the assumption that our diffusion is reversible i.e. we are not necessarily dealing with a drift given by a gradient vector field (extension of the results in [Vos97]). As we will see in Section 4.2 below, there are then situations in which the limiting process is not reversible either, contrary to the one-dimensional situation. The limiting process is not reversible when we have a non-zero local time term in those directions contained within the interface. This is due to the measures induced by the process and its time reversed counterpart being mutually singular on path space as a result of said non-zero local time terms.

If the diffusion we are homogenizing has drift given by a gradient vector field then it is reversible and as a consequence the limiting process is also reversible. This implies that the limiting process has no non-zero local time terms in the components the directions of which are contained within the interface. As a consequence, when conducting the identification of the limit stage of the homogenization scheme, in this case we do not need to explore those directions contained within the interface. The method of identification of the limit therefore becomes akin to that studied in the one dimensional case [HM10b] i.e., in this case, we only have to identify the coefficient of the local time term for the component perpendicular to the interface.

One feature of the problem at hand is that there is no *finite* invariant measure built into the framework of the problem. This is unlike most other homogenization problems, even those exhibiting rather 'bad' ergodic properties, such as the random environment case [PV81, Oll94] or the quenched convergence results for the Bouchaud trap model [BAČ07]. Since in our case the invariant measure μ of X is only σ-finite, this leads to two problems when trying to compute the effect of the behaviour of X near the interface in the limit $\varepsilon \to 0$. Indeed, one would 'naïvely' expect that an effective drift along the interface can be described by the quantity

$$\int b(x)\,\mu(dx)\;. \tag{4.1.1}$$

One problem with this expression is that there is no obvious natural normalisation for μ. Furthermore, since b is periodic away from the interface and the same is (approximately) true for μ, this integral certainly does not converge, even if we consider it as an integral over $\mathbb{R} \times \mathbb{T}^{d-1}$ by making use of the periodic structure in the directions parallel to the interface. See however (4.2.4) and Proposition 4.6.3 below for the correct way of interpreting (4.1.1) and our main result, Theorem 4.2.4 below, on how this quantity appears in the construction of the limiting process.

Another common feature of many homogenization results is the usage of a globally defined corrector function to compensate for the singular terms appearing in the problem. This is of course the case for standard periodic homogenization [BLP78], but also for a number of stochastic homogenization problems, as for example in [OS04, PV81, Oll94, Rho09a]. For the present problem however, it will be convenient to make use of corrector function that only cancels the singular terms away from the interface and to treat the behaviour of the limiting process at the interface by completely different means.

One very recent homogenization result where discontinuous coefficients appear in the limiting equation can be found in [BEP09] (which in turn generalises

[KK01]). However, their framework is quite different to the one considered here and doesn't seem to encompass our problem. Much more closely related problems are homogenization problems with the presence of a boundary [AA99, GVM08]. Those have been mostly studied by analytical tools so far. In our probabilistic language, what comes closest to the boundary layers studied in these articles is the σ-finite invariant measure of X, which is shown in Proposition 4.5.5 below to converge exponentially fast to a measure with periodic densities away from the interface.

For simplicity, we will consider the case of a constant diffusion matrix, but it is straightforward to adapt the proofs to cover the case of non-constant diffusivity as well. More precisely, we consider the family of processes X^ε taking values in $\mathbb{R}^d$, solutions to the stochastic differential equations

$$dX^\varepsilon = \frac{1}{\varepsilon} b\left(\frac{X^\varepsilon}{\varepsilon}\right) ds + dB(s) , \qquad X^\varepsilon(0) = x , \tag{4.1.2}$$

where B is a d-dimensional standard Wiener process. The drift b is assumed to be smooth and such that $b(x + e_i) = b(x)$ for the unit vectors e_i with $i = 2, \dots, d$ (but not for $i = 1$). Furthermore, we assume that there exist smooth vector fields $b_\pm$ with unit period in *every* direction and $\eta > 0$ such that

$$b(x) = b_+(x) , \quad x_1 > \eta , \qquad b(x) = b_-(x) , \quad x_1 < -\eta .$$

Figure 4.1 is a typical illustration of the type of vector fields that we have in mind.

If we denote by X the same process, but with $\varepsilon = 1$, then the process X^ε given by (4.1.2) is equal in law to the diffusive rescaling of X by a factor $\frac{1}{\varepsilon}$. In the sequel, we denote the generator of X by $\mathcal{L}$ and the generator of X^ε by $\mathcal{L}_\varepsilon$. We furthermore denote by $\mathcal{L}_\pm$ the generators for the diffusion processes on the torus given by

$$dX^\pm = b_\pm\big(X^\pm\big)\, ds + dB(s) , \tag{4.1.3}$$

and by $\mu_\pm$ the corresponding invariant probability measures. With this notation at hand, we impose the centering condition $\int_{\mathbb{T}^d} b_\pm(x)\, \mu_\pm(x) = 0$.

Under these conditions, our main result formulated in Theorem 4.2.4 below states that the family X^ε converges in law to a limiting process $\bar{X}$. Furthermore, we give an explicit characterisation of $\bar{X}$, both as the unique solution of a martingale

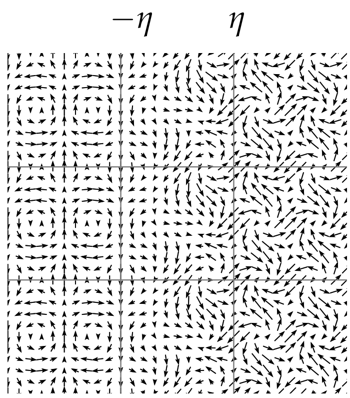


Figure 4.1: Example of a vector field b satisfying our conditions.

problem with some explicitly given generator and as the solution of a stochastic differential equation involving a local time term on the interface $\{x_1 = 0\}$. In addition to the homogenized diffusion coefficients on either side of the interface, this limiting process is characterised by a 'transmissivity coefficient', as well as by a 'drift vector' pointing along the interface.

The remainder of this article is structured as follows. After formulating our main results in Section 4.2, we show tightness of the family in Section 4.3. In Section 4.4, we then formulate the main tool used in the identification of the limiting process, namely a multidimensional analogue of the tool used by Freidlin and Wentzell in [FW93] to study homogenisation problems where the limiting process takes values in a graph. Section 4.5 is then devoted to the computation of the transmissivity coefficient, whereas Section 4.6 contains the computation of the drift vector. Finally, we show in Section 4.7 that the martingale problem is well-posed and we identify its solution with the solution to a stochastic differential equation.

4.1.1 Notation

We define the 'interface' of width K by

$$\mathscr{I}_K = \{x \in \mathbb{R}^d \,:\, x_1 \in [-K, K]\} \,.$$

We also denote by $\partial \mathscr{I}_K$ its boundary.

Frequently throughout the paper we will construct successive escape and subsequent reentry times particularly when constructing invariant measures in terms of the invariant measure of an embedded Markov chain as in [Has60]. We will denote such pairs of stopping times as σ, ϕ, which denote escape and reentry times respectively. Other stopping times not part of such a sequence will be denoted by τ.

4.2 The Main Result

Before stating the main result we will first define the various quantities involved and their relevance. It is clear that, in view of standard results from periodic homogenization [BLP78, PS08], any limiting process for X^ε should behave like Brownian motion on either side of the interface $\mathscr{I}_0 = \{x_1 = 0\}$, with effective diffusion tensors given by

$$D_{ij}^{\pm} = \int_{\mathbb{T}^d} (\delta_{ik} + \partial_k g_i^{\pm})(\delta_{kj} + \partial_k g_j^{\pm})\, d\mu_{\pm} \ .$$

(Summation of k is implied.) Here, the corrector functions $g_{\pm}\colon \mathbb{T}^d \to \mathbb{R}^d$ are the unique solutions to $\mathcal{L}_{\pm} g_{\pm} = -b_{\pm}$ such that

$$\int_{\mathbb{T}^d} g_{\pm}(x)\, \mu_{\pm}(dx) = 0 \ .$$

Since $b_{\pm}$ are centered with respect to $\mu_{\pm}$, such functions do indeed exist.

This justifies the introduction of a differential operator $\bar{\mathcal{L}}$ on $\mathbb{R}^d$ defined in two parts by $\bar{\mathcal{L}}_+$ on $I_+ = \{x_1 > 0\}$ and $\bar{\mathcal{L}}_-$ on $I_- = \{x_1 < 0\}$ with

$$\bar{\mathcal{L}}_{\pm} = \frac{D_{ij}^{\pm}}{2} \partial_i \partial_j \ , \tag{4.2.1}$$

then one would expect any limiting process to solve a martingale problem associated to $\bar{\mathcal{L}}$. However, the above definition of $\bar{\mathcal{L}}$ is not complete, since we did not specify any boundary condition at the interface $\mathscr{I}_0$.

One of the main ingredients in the analysis of the behaviour of the limiting process at the interface is the invariant measure μ for the (original, not rescaled) process X. It is not clear *a priori* that such an invariant measure exists, since X is not expected to be recurrent in general. However, if we identify points that differ by integer multiples of e_j for $j = 2, \ldots, d$, we can interpret X as a process with state space $\mathbb{R} \times \mathbb{T}^{d-1}$. It then follows from the results in [Has60] that this process admits a σ-finite invariant measure μ on $\mathbb{R} \times \mathbb{T}^{d-1}$.

Note that the invariant measure μ is *not* finite and can therefore not be normalised in a canonical way. However, if we define the 'unit cells' $C_j^\pm$ by

$$C_j^+ = [j, j+1] \times \mathbb{T}^{d-1} , \qquad C_j^- = [-j-1, -j] \times \mathbb{T}^{d-1} ,$$

then it is possible to make sense of the quantity $q_\pm = \lim_{j\to\infty} \mu(C_j^\pm)$ (we will show in Proposition 4.5.5 below that this limit actually exists).

Let now $p_\pm$ be given by

$$p_\pm = \frac{q_\pm D_{11}^\pm}{q_+ D_{11}^+ + q_- D_{11}^-} ,$$

which can also we rewritten in a more suggestive way as

$$\frac{p_+}{p_-} = \frac{q_+ D_{11}^+}{q_- D_{11}^-} . \tag{4.2.2}$$

This is the homogenized diffusion coefficient in the direction perpendicular to the interface, weighted by the invariant measure of a unit cell. Comparing with the one-dimensional case [HM10b], one would expect this to yield the likelihood for X^ε to exit a small (but still much larger than ε) neighborhood of the interface on a specific side.

Remark 4.2.1. *The ratio*

$$\frac{p_+ \sqrt{D_{11}^-}}{p_- \sqrt{D_{11}^+} + p_+ \sqrt{D_{11}^-}} . \tag{4.2.3}$$

gives the asymptotic probability of the process being located in the rhs (+) of the interface after a long time. This follows from the weak convergence of the first component to a skew Brownian motion with (possibly) different diffusion coefficients on either side of the interface. If we rescale this skew BM on either side of the interface by $\sqrt{D_{11}^\pm}$ to obtain a standard skew BM, we can use the scale function of BM to finish the verification of (4.2.3).

However, unlike in the one-dimensional case, these quantities are not yet sufficient to characterise the limiting process. The reason is that since X^ε is expected to spend time proportional to ε in the interface, but the drift is of order ε^{-1}

there, it is not impossible that the limiting process picks up a non-trivial drift along the interface. It turns out that this drift can be described by the coefficients α_j given by

$$\alpha_j = 2\Big(\frac{p_+}{D_{11}^+} + \frac{p_-}{D_{11}^-}\Big) \int_{\mathbb{R}\times\mathbb{T}^{d-1}} \big(b_j(x) + \mathcal{L}g_j(x)\big)\,\mu(dx)\;, \tag{4.2.4}$$

where μ is again normalised in such a way that $q_+ + q_- = 1$ and where g is any smooth function agreeing with $g_\pm$ on either side of the interface (see Section 4.3 below).

Remark 4.2.2. *Since $\int_{\mathbb{R}\times\mathbb{T}^{d-1}} \mathcal{L}\phi(x)\,\mu(dx) = 0$ for every smooth compactly supported function ϕ, one should interpret the integral on the right hand side of (4.2.4) as a 'renormalised' form of the intuitive more meaningful quantity (4.1.1).*

Remark 4.2.3. *The expression (4.2.4) is useful in order to generate examples with non-vanishing values for the coefficients α_i.*

Given all of these ingredients, we can construct an operator $\bar{\mathcal{L}}$ as follows. The domain $\mathscr{D}(\bar{\mathcal{L}})$ of $\bar{\mathcal{L}}$ consists of functions $f\colon \mathbb{R}^d \to \mathbb{R}$ such that

- The restrictions of f to I_+, I_-, and $\mathscr{I}_0$ are smooth.
- The partial derivatives $\partial_i f$ are continuous for $i \geq 2$.
- The partial derivative $\partial_1 f(x)$ has right and left limits $\partial_1 f|_{I_\pm}$ as $x \to \mathscr{I}_0$ and these limits satisfy the gluing condition

$$p_+\partial_1 f|_{I_+} - p_-\partial_1 f|_{I_-} + \sum_{j=2}^d \alpha_j \partial_j f = 0\;. \tag{4.2.5}$$

For any $f \in \mathscr{D}(\bar{\mathcal{L}})$, we then set $\bar{\mathcal{L}}f(x) = \mathcal{L}_\pm f(x)$ for $x \in I_\pm$. With these definitions at hand, we can state the main result of the article:

Theorem 4.2.4. *The family of processes X^ε converges in law to the unique solution $\bar{X}$ to the martingale problem given by the operator $\bar{\mathcal{L}}$. Furthermore, there exist matrices $M_\pm$ and a vector $K \in \mathbb{R}^d$ such that this solution solves the SDE*

$$d\bar{X}(t) = \mathbf{1}_{\bar{X}_1\leq 0} M_- dW(t) + \mathbf{1}_{\bar{X}_1>0} M_+ dW(t) + K\,dL(t)\;. \tag{4.2.6}$$

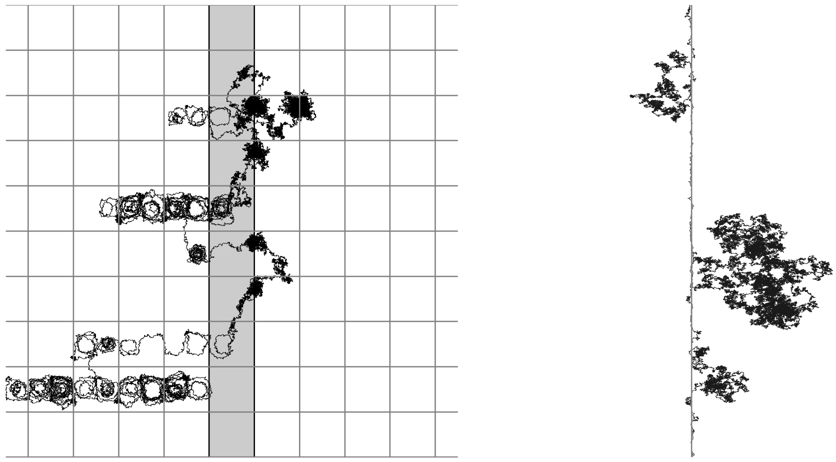


Figure 4.2: Sample paths at small (left) and large (right) scales.

where L denotes the symmetric local time of $\bar{X}_1$ at the origin and W is a standard d-dimensional Wiener process. The matrices $M_\pm$ and the vector K satisfy

$$M_\pm M_\pm^T = D^\pm \;, \quad K_1 = p_+ - p_- \;, \quad K_j = \alpha_j \;,$$

for $j = \{2, \ldots, d\}$.

In Figure 4.2, we show an example of a numerical simulation of the process studied in this article. The figure on the left shows the small-scale structure (the periodic structure of the drift is drawn as a grid). One can clearly see the periodic structure of the sample path, especially to the left of the interface. One can also see that the effective diffusivity is not necessarily proportional to the identity. In this case, to the left of the interface, the process diffuses much more easily horizontally than vertically.

The picture to the right shows a simulation of the process at a much larger scale. We used a slightly different vector field for the drift in order to obtain a simulation that shows clearly the strong drift experienced by the process when it hits the interface.

Remark 4.2.5. *Since the quadratic variation of $\bar{X}$ has a discontinuity at $\bar{X}_1 = 0$, we do have to specify which kind of local time L is. Using the symmetric local time yields nicer expressions. See for example [RY91, Lej06] for a definition of the symmetric local time.*

Analyzing what this means for a simple example, we consider the case of a two dimensional problem where we have $b_1 = 0$ and $b_2 = f(x_1)$ for f a smooth function that is zero outside of $\mathscr{I}_\eta$. Clearly $p_\pm = \frac{1}{2}$. In this case the invariant measure μ of the process X is given by $\frac{1}{2}$ times Lebesgue measure on $\mathbb{R} \times S^1$ and we can choose $g = 0$. This implies that we then simply have

$$\alpha_2 = \int_{\mathbb{R}} f(x)\, dx \;,$$

as one would expect.

4.3 Tightness of the family

The aim of this section is to prove the following tightness result:

Theorem 4.3.1. *Denote by $\mathbb{P}^\varepsilon$ the law of X_x^ε on $\mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$. Then the family $\{\mathbb{P}^\varepsilon\}_{\varepsilon \in (0,1]}$ is tight.*

Similar to what happens in the classical theory of periodic homogenization, it will be very convenient to construct a 'corrected process' Y, obtained by adding to X a corrector function that cancels out to first order the effect of the small oscillations. To this aim, we introduce a smooth function $g\colon \mathbb{R}^d \to \mathbb{R}^d$ which is periodic in the directions $2, \ldots, d$ and such that $g(x) = g_+(x)$ for $x_1 \geq \eta$ and similarly for $x_1 \leq -\eta$. (Recall that $g_\pm$ were defined in Section 4.2.) We do not specify the behaviour of g inside the interface $\mathscr{I}_\eta$, except that it has to be smooth in the whole space and periodic in the directions parallel to the interface. We fix such a function g once and for all from now on. We furthermore denote by Y^ε the process defined by $Y^\varepsilon = X^\varepsilon + \varepsilon g(\varepsilon^{-1} X^\varepsilon)$, as well as $y = x + \varepsilon g(x/\varepsilon)$ for its initial condition.

Defining the corrected drift $\tilde b(x) = \big(\mathcal{L} g + b\big)(x)$ and the corrected diffusion coefficient $\tilde\sigma_{ij}(x) = \delta_{ij} + \partial_j g_i(x)$, it follows from Itô's formula that the ith component of Y_y^ε satisfies

$$\big(Y_y^\varepsilon\big)_i(t) = y_i + \int_0^t \frac{1}{\varepsilon} \tilde b_i\Big(\frac{1}{\varepsilon} X_x^\varepsilon(s)\Big)\, ds + \int_0^t \tilde\sigma_{ij}\Big(\frac{1}{\varepsilon} X_x^\varepsilon(s)\Big)\, dW_j(s) \;. \tag{4.3.1}$$

It is very important to note that the corrected drift $\tilde b$ *vanishes* outside of $\mathscr{I}_\eta$, so that the process Y is subject to a large drift only when X is inside the interface.

Our main tool in the proof of Theorem 4.3.1 is the following result, which is very similar to [SV79, Thm 1.4.6]:

Proposition 4.3.2. *Let $\mathscr{P}$ be a family of probability measures on $\Omega = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ and denote by x the canonical process on Ω. Assume that*

$$\lim_{R \nearrow \infty} \sup_{\mathbb{P} \in \mathscr{P}} \mathbb{P}\big(|x(0)| \geq R \big) = 0 \; .$$

Furthermore, for any given $\rho > 0$, let $\tau_0 = 0$, and define recursively $\tau_{i+1} = \inf_{t > \tau_i} |x(t) - x(\tau_i)| > \rho$. Assume that the limit

$$\lim_{\delta \to 0} \operatorname{ess\,sup} \mathbb{P}\left[\tau_{n+1} - \tau_n \leq \delta | \mathscr{F}_{\tau_n} \right] \to 0, \qquad \mathbb{P} \text{ a.s., on } \{\tau_n < \infty\} \tag{4.3.2}$$

holds uniformly for every $\mathbb{P} \in \mathscr{P}$ and every $n \geq 0$. Then the family of probability measures $\mathscr{P}$ is tight on Ω.

Proof. The proof is similar to that of Theorem 1.4.6 in [SV79], except that their Lemma 1.4.4 is replaced by (4.3.2).

Fix an arbitrary final time $T > 0$. Furthermore, denote for $\omega \in \Omega$,

$$N_\rho = N_\rho(\omega) = \min\{n : \tau_{n+1} > T\} \; .$$

and the modulus of continuity by δ_ρ,

$$\delta_\rho = \delta_\rho(\omega) = \min\{\tau_n - \tau_{n-1} : 1 \leq n \leq N_\rho(\omega)\} \; .$$

Note that this expression depends on ρ via the definition of the stopping times τ_i.

With this notation at hand, tightness follows as in [SV79] if one can show that $\lim_{\delta \to 0} \sup_{\mathbb{P} \in \mathscr{P}} \mathbb{P}(\delta_\rho \leq \delta) = 0$ for every fixed $\rho > 0$. As in [SV79], one has for every $k > 0$ the bound

$$\mathbb{P}(\delta_\rho \leq \delta) \;\leq\; \sum_{i=1}^{k} \mathbb{E}\big[\mathbb{P}\left[\tau_{i+1} - \tau_i \leq \delta \,|\, \mathscr{F}_{\tau_i} \right] \big] + \mathbb{P}(N_\rho > k) \; .$$

For every fixed $k > 0$, the first term then converges uniformly to 0 by assumption. Since the second term is independent of δ, it remains to verify that converges to 0 as $k \to \infty$, uniformly over $\mathscr{P}$ (convergence for every fixed $\mathbb{P} \in \mathscr{P}$ is trivial but not sufficient for our needs).

This is a consequence of [SV79, Lemma 1.4.5], provided that one can find $\lambda < 1$ such that $\mathbb{E}\big[e^{-(\tau_{i+1} - \tau_i)} \,|\, \mathscr{F}_{\tau_i}\big] \leq \lambda$. This in turn follows from

$$\mathbb{E}\big[e^{-(\tau_{i+1} - \tau_i)} \,|\, \mathscr{F}_{\tau_i}\big] \;\leq\; \mathbb{P}\left[\tau_{i+1} - \tau_i \leq t_0 | \mathscr{F}_{\tau_i} \right] + e^{-t_0} \mathbb{P}\left[\tau_{i+1} - \tau_i > t_0 | \, \mathscr{F}_{\tau_i} \right]$$

$$\leq e^{-t_0} + (1 - e^{-t_0})\mathbb{P}\left[\tau_{i+1} - \tau_i \leq t_0 \middle| \mathscr{F}_{\tau_i}\right] .$$

Indeed, by choosing t_0 sufficiently small, this term can be made strictly less than 1, provided that $\mathbb{P}\left[\tau_{i+1} - \tau_i \leq t_0 \middle| \mathscr{F}_{\tau_i}\right]$ tends to zero uniformly (over the members of $\mathscr{P}$ and over i) as t_0 tends to zero, which is precisely our assumption. □

We now turn to the

Proof of Theorem 4.3.1. Recall that we defined the process

$$Y^\varepsilon = X^\varepsilon + \varepsilon g(\varepsilon^{-1} X^\varepsilon) ,$$

in Section 4.2. Note then that, just as in [HM10b, Prop. 2.5], the tightness of the laws of X^ε_x is equivalent to that of the laws of Y^ε_x. Therefore, all that remains to be shown is that we have the bound (4.3.2) for the law of Y^ε, uniformly over $\varepsilon \in (0, 1]$. The approach that we use is to consider separately the martingale part and the bounded variation part for Y^ε_y given by (4.3.1), and to show that the probability of either of these moving by at least $\frac{\rho}{2}$ during a time interval δ tends to zero uniformly over the initial condition.

Given any fixed $\rho, \gamma > 0$, we want to show that there exists a sufficiently small $\delta > 0$ such that $\mathbb{P}(\tau_{n+1} - \tau_n \leq \delta \,|\, \mathscr{F}_{\tau_n}) < \gamma$ uniformly over $\mathbb{P} \in \mathscr{P}$ (that is uniformly over the laws of Y^ε_x with $\varepsilon \in (0, 1]$) and n. We split the contributions from the martingale and the bounded variation parts in the following way:

$$\begin{aligned}
\mathbb{P}(\tau_{n+1} - \tau_n \leq \delta \,|\, \mathscr{F}_{\tau_n}) &= \mathbb{P}_{X(\tau_n)}\Big(\sup_{t<\delta} |Y(t) - Y(0)| > \rho \Big) \\
&\leq \sup_x \mathbb{P}_x \Big(\sup_{t<\delta} \Big| \frac{1}{\varepsilon} \int_0^t \tilde{b}_i(\varepsilon^{-1} X^\varepsilon_x(s))\, ds \Big| > \frac{\rho}{2} \Big) \\
&\quad + \sup_x \mathbb{P}_x \Big(\sup_{t<\delta} \Big| \int_0^t \tilde{\sigma}_{ij}(\varepsilon^{-1} X^\varepsilon_x(s))\, dW_j(s) \Big| > \frac{\rho}{2} \Big) \\
&\leq \frac{2}{\varepsilon\rho} \sup_x \mathbb{E}_x \int_0^t \big| \tilde{b}_i(\varepsilon^{-1} X^\varepsilon_x(s)) \big|\, ds \\
&\quad + \frac{2}{\rho} \sup_x \mathbb{E}_x \sup_{t\leq\delta} \Big| \int_0^t \tilde{\sigma}_{ij}(\varepsilon^{-1} X^\varepsilon_x(s))\, dW_j(s) \Big| . \qquad (4.3.3)
\end{aligned}$$

Here, we used the Chebychev's inequality to obtain the last bound. Since the functions $\tilde{\sigma}_{ij}$ are uniformly bounded, the stochastic integral appearing in the second term is easily bounded by $\mathcal{O}(\sqrt{\delta})$ by the Burkholder-Davis-Gundy inequalities.

Furthermore, by the definition of the corrector function g, there exists $\tilde{\eta} > 0$ such that $\tilde{b}(x) = 0$ for $x \notin \mathscr{I}_{\tilde{\eta}\varepsilon}$, so that there exists a constant C such that

$$\mathbb{P}(\tau_{n+1} - \tau_n \leq \delta \,|\, \mathscr{F}_{\tau_n}) \leq \frac{C}{\rho\varepsilon} \sup_x \mathbb{E}_x \Big(\int_0^\delta \mathbf{1}_{\mathscr{I}_{\tilde{\eta}\varepsilon}}\big(X_x^\varepsilon(s)\big)\, ds \Big) + \frac{C\sqrt{\delta}}{\rho} \;. \tag{4.3.4}$$

For fixed $\rho > 0$, the second term obviously goes to 0 as $\delta \to 0$, uniformly in ε, so it remains to consider the first term. As one would expect from the expression for the local time of a Brownian motion, it turns out that the expected time spent by the process in $\mathscr{I}_{\tilde{\eta}\varepsilon}$ scales like $\varepsilon\sqrt{\delta}$, thus showing that this term is also of order $\sqrt{\delta}/\rho$. Once we are able to show this, the proof is complete.

The occupation time of the interface appearing in the first term of (4.3.4) is bounded by the trivial estimate $C\delta/(\rho\varepsilon)$, which goes to 0 as $\delta \to 0$ provided that we consider $\varepsilon \geq \sqrt{\delta}$, say. We can therefore assume without any loss of generality in the sequel that we consider $\varepsilon < \sqrt{\delta}$.

The idea to bound the occupation time is the following. We decompose the trajectory for the process X^ε into excursions away from the interface, separated by pieces of trajectory inside the interface. We first show that if the process starts inside the interface, then the expected time spent in the interface before making a new excursion is of order ε^2. Then, we show that each excursion has a probability at least $\varepsilon/\sqrt{\delta}$ of being of length δ or more. This shows that in the time interval δ of interest, the process will perform at most of the order of $\sqrt{\delta}/\varepsilon$ excursions, so that the total time spent in the interface is indeed of the order $\varepsilon\sqrt{\delta}$, thus showing that the first term in (4.3.4) behaves like $\sqrt{\delta}/\rho$, as expected.

More precisely, we first choose two constants $K > 0$ and $\hat{K} > 0$ such that the chain of implications

$$\{X^\varepsilon \in \mathscr{I}_{\tilde{\eta}\varepsilon}\} \Rightarrow \{Y^\varepsilon \in \mathscr{I}_{\hat{K}\varepsilon}\} \Rightarrow \{X^\varepsilon \in \mathscr{I}_{(K-1)\varepsilon}\} \Rightarrow \{X^\varepsilon \in \mathscr{I}_{K\varepsilon}\} \;, \tag{4.3.5}$$

holds. We then set up a sequence of stopping times in the following way. We set $\phi_0 = 0$ and we set recursively

$$\begin{aligned} \sigma_n &= \inf\{t \geq \phi_n \;:\; X^\varepsilon(t) \notin \mathscr{I}_{K\varepsilon}\} \;, \\ \phi_n &= \inf\{t \geq \sigma_{n-1} \;:\; Y^\varepsilon(t) \in \mathscr{I}_{\hat{K}\varepsilon}\} \;. \end{aligned}$$

(Note that we can have $\sigma_0 = 0$ if the initial condition does not belong to $\mathscr{I}_{K\varepsilon}$. Apart from that, the second implication in (4.3.5) shows that increments from one stopping time to the next are always strictly positive.) This construction was chosen in such a way that the times when $X^\varepsilon \in \mathscr{I}_{\tilde{\eta}\varepsilon}$ always fall between ϕ_n and σ_n for some $n \geq 0$. In particular, if we set

$$N = \inf\{n \geq 0 \,:\, \phi_{n+1} - \sigma_n \geq \delta\}\;,$$

then we have the bound

$$\begin{aligned}
&\sup_x \mathbb{E}_x\Big(\int_0^\delta \mathbf{1}_{\mathscr{I}_{\tilde{\eta}\varepsilon}}\big(X_x^\varepsilon(s)\big)\,ds\Big)\\
&\leq \sup_x \mathbb{E}_x\Big(\sum_{n=0}^N (\sigma_n - \phi_n)\Big)\\
&= \sup_x \sum_{n=0}^\infty \mathbb{E}_x\big((\sigma_n - \phi_n)\mathbf{1}_{N\geq n}\big) = \sum_{n=0}^\infty \sup_x \mathbb{P}_x(N \geq n) \sup_x \mathbb{E}_x\big(\mathbb{E}_{X^\varepsilon(\phi_n)}\sigma_1\big)
\end{aligned}$$

where we used the strong Markov property and the fact that $\{N \geq n\}$ is $\mathscr{F}_{\phi_n}$-measurable in order to obtain the last identity. It follows from the definition of N that this expression is in turn bounded by

$$\sup_{x\in\mathbb{R}^d} \mathbb{E}_x\sigma_0 \sum_{n\geq 0}\big(\sup_{x\notin\mathscr{I}_{K\varepsilon}} \mathbb{P}_x(\phi_0 < \delta)\big)^n = \frac{\sup_{x\in\mathbb{R}^d}\mathbb{E}_x\sigma_0 \sup_{x\notin\mathscr{I}_{K\varepsilon}}\mathbb{P}_x(\phi_0 < \delta)}{\inf_{x\notin\mathscr{I}_{K\varepsilon}}\mathbb{P}_x(\phi_0 \geq \delta)}\;.$$

We now bound both terms appearing in this expression separately.

First, we turn to the expected escape time from the interface, $\mathbb{E}_x\sigma_0$. The idea is to use a comparison argument just like in [HM10b, Proposition 3.8]. We define a 'worst-case scenario' process V_x^ε, which is the solution to the SDE with initial condition x, diffusion coefficient 1 and drift coefficient given by b_V^ε, where

$$b_V^\varepsilon(x) \;=\; \begin{cases} \frac{-b_V}{\varepsilon} & \text{for } x \geq 0,\\ \frac{b_V}{\varepsilon} & \text{for } x < 0,\end{cases}$$

for some constant $b_V > 0$. We then have:

Lemma 4.3.3. *There exist $b_V > 0$ and $\tilde{K} > 0$ such that, if we define $\tau^{\tilde{K}} = \inf\{t \geq 0 \,:\, V_x^\varepsilon(t) \notin \mathscr{I}_{\tilde{K}\varepsilon}\}$, we have*

$$\mathbb{E}_x\sigma_0 \leq \mathbb{E}_x\tau^{\tilde{K}}\;,$$

for every $x \in \mathbb{R}^d$.

The proof of Lemma 4.3.3 is almost identical to that of [HM10b, Proposition 3.8], so we are going to omit it. A straightforward calculation using the particular form of the drift coefficient for V allows to check that there exists indeed a constant $C > 0$ such that the bound

$$\sup_x \mathbb{E}_x \tau^{\tilde{K}} \le C\varepsilon^2 ,$$

holds so that, combining this with Lemma 4.3.3, we have $\sup_{x\in\mathbb{R}^d} \mathbb{E}_x \sigma_0 \le C\varepsilon^2$.

Let us now turn to the bound on $\mathbb{P}_x(\phi_0 \ge \delta)$. The idea here is to look at the process Y^ε instead of X^ε and to time-change it in such a way that we can compare it to a standard Brownian motion. Note first that the last two implications in (4.3.5) show that if we start with X^ε anywhere outside of $\mathscr{I}_{K\varepsilon}$, then the first component of Y^ε has to travel by at least ε before the process Y^ε can hit $\mathscr{I}_{\hat{K}\varepsilon}$. Furthermore, it follows from (4.3.1) that the time change C_t such that $Y^\varepsilon(C_t)$ is a standard Brownian motion satisfies $C_t \ge ct$ for some $c > 0$. It therefore follows that, setting $H(z) = \inf_{t>0}\{B_t > z\}$, one has the lower bound

$$\inf_{x\notin\mathscr{I}_{K\varepsilon}} \mathbb{P}_x(\phi_0 \ge \delta) \ge \mathbb{P}(H(\varepsilon) \ge \delta/c) .$$

The explicit expression for the law of $H(z)$ given in [BS96, p. 163, eq. 2.02] yields in turn

$$\mathbb{P}(H(\varepsilon) \ge \delta/c) = \int_{\delta/(c\varepsilon^2)}^\infty \frac{e^{-1/(2t)}}{\sqrt{2\pi}\, t^{\frac{3}{2}}}\, dt .$$

It follows immediately that this in turn is bounded from below by $C\varepsilon/\sqrt{\delta}$ for some $C > 0$, provided that $\varepsilon \le \sqrt{\delta}$. Collecting these bounds completes the proof of Theorem 4.3.1. □

4.4 Main tool for identifying the limit process

Instead of considering a graph as before, we will consider a generalized multi-dimensional version different from that considered by Freidlin and Wentzell in [FW93, Section 6]. Note that the generalisation considered here is different (and actually simpler) than the one considered in [FW06]. We consider processes in $\mathbb{R}^d$

and we set $I_- = \{x \in \mathbb{R}^d \ : \ x_1 < 0\}$, and similarly for I_+. We consider a family of $\mathbb{R}^d$-valued processes X^ε and we denote by τ^ε the first hitting time of $\mathscr{I}_{\varepsilon\eta}$. Correspondingly, τ^δ is the first escape time of the set $\mathscr{I}_\delta$ by X^ε.

With this the main tool will be the following multidimensional analogue of [FW93, Thm 4.1]:

Theorem 4.4.1. *Let $\bar{\mathcal{L}}_i$ be second order differential operators on I_i with bounded coefficients and let D_i be some sets of test functions over I_i whose members are bounded and have bounded derivatives of all orders. Suppose that for $i \in \{+,-\}$, any function $f \in D_i$ and for any $\lambda > 0$, the bound*

$$\mathbb{E}_x\Big[e^{-\lambda\tau^\varepsilon} f\big(X^\varepsilon(\tau^\varepsilon)\big) - f\big(X^\varepsilon(0)\big) \\ + \int_0^{\tau^\varepsilon} e^{-\lambda t}\Big(\lambda f\big(X^\varepsilon(t)\big) - \bar{\mathcal{L}}_i f\big(X^\varepsilon(t)\big)\Big)\, dt \Big] = O\big(k(\varepsilon)\big)\,, \tag{4.4.1}$$

holds as $\varepsilon \to 0$, uniformly with respect to $x \in I_i$. Assume furthermore that the rate k is such that $\lim_{\varepsilon\to 0} k(\varepsilon) = 0$.

Assume that, for every $\lambda > 0$ and every $i \in \{+,-\}$, there exist functions $u_{i,\lambda} \in D_i$ such that $\bar{\mathcal{L}}_i u_{i,\lambda}(x) = \lambda u_{i,\lambda}(x)$ holds for $x \in I_i$ with $|x_1| \le 1$ and such that $u_{\pm,\lambda}(x) = 1$ for $x_1 = 0$ and $x_1 = \pm 1$.

Assume that there exists a rate $\delta = \delta(\varepsilon) \to 0$ such that $\delta(\varepsilon)/k(\varepsilon) \to \infty$ as $\varepsilon \to 0$ and such that for $\lambda > 0$,

$$\mathbb{E}_x^\varepsilon\Big[\int_0^\infty e^{-\lambda t}\mathbf{1}_{(-\delta,\delta)}\big(X_1^\varepsilon(t)\big)\, dt\Big] \to 0 \tag{4.4.2}$$

as $\varepsilon \to 0$, uniformly in the initial point. Assume the convergence

$$\mathbb{P}_x^\varepsilon[X^\varepsilon\big(\tau^\delta\big) \in I_i] \to p_i\,, \tag{4.4.3}$$

holds uniformly in x in the set $\mathscr{I}_{\varepsilon\eta}$ for some constants $p_\pm$ with $p_+ + p_- = 1$. Assume furthermore that there exist constants α_j and C such that

$$\frac{1}{\delta}\,\mathbb{E}_x^\varepsilon\Big[X_j^\varepsilon\big(\tau^\delta\big) - x_j\Big] \to \alpha_j\,, \quad \frac{1}{\delta^2}\,\mathbb{E}_x^\varepsilon\Big[\big(X_j^\varepsilon\big(\tau^\delta\big) - x_j\big)^2\Big] \le C\,, \tag{4.4.4}$$

for $j \ge 2$. Again, the limit is assumed to be uniform over $x \in \mathscr{I}_{\varepsilon\eta}$ as $\varepsilon \to 0$, and the inequality is assumed to be uniform over all $\varepsilon \in (0,1]$ and all $x \in \mathscr{I}_{\varepsilon\eta}$.

Let then D be the set of continuous functions $f\colon \mathbb{R}^d \to \mathbb{R}$ such that the restriction of f to I_i belongs to D_i and such that the gluing condition (4.2.5) holds. Then, for any fixed $f \in D$, $t_0 \geq 0$ and $\lambda > 0$,

$$\begin{aligned}\Delta(\varepsilon) \;=\; \operatorname{ess\,sup}\biggl|\,\mathbb{E}_x^\varepsilon\biggl[\int_{t_0}^\infty e^{-\lambda t}\Bigl[\lambda f\bigl(X^\varepsilon(t)\bigr) - \bar{\mathcal{L}} f\bigl(X^\varepsilon(t)\bigr)\Bigr]\,dt \\ -\; e^{-\lambda t_0} f\bigl(X^\varepsilon(t_0)\bigr)\,\Big|\, \mathscr{F}_{[0,t_0]}\biggr]\biggr| \to 0 \qquad (4.4.5)\end{aligned}$$

as $\varepsilon \to 0$, uniformly with respect to x. In particular, every weak limit of X^ε as $\varepsilon \to 0$ satisfies the martingale problem for $\bar{\mathcal{L}}$.

Remark 4.4.2. Note that we did not specify how 'large' the sets D_i of admissible test functions need to be. If these sets are too small then the theorem still holds, but the corresponding martingale problem might become ill-posed.

Proof. Since the proof is virtually identical to that of [FW93, Thm 4.1], we only sketch it here. The basic idea behind the proof given by Freidlin and Wentzell is to rewrite (4.4.5) using the Strong Markov Property of X^ε as a sum of terms between successive stopping times. To this effect, set for example $\sigma_0 = 0$ and then recursively $\phi_n = \inf\{t > \sigma_n : X_1^\varepsilon(t) \in \mathscr{I}_{\varepsilon\eta}\}$, $\sigma_{n+1} = \inf\{t > \phi_n : X_1^\varepsilon(t) \notin \mathscr{I}_\delta\}$. They then break up the term produced from (4.4.5) into two sums of analogous terms between times σ_n and ϕ_n and those between ϕ_n and σ_{n+1}.

The terms covering the time intervals $[\sigma_n, \phi_n]$ are bounded exactly as in [FW93], making use of (4.4.1), together with the bound $\sum_n \mathbb{E}_x e^{-\lambda\sigma_n} = \mathcal{O}(1/\delta)$ which follows from the existence of the functions $u_{i,\lambda}$ just as in [FW93].

Using assumption (4.4.2), the terms covering the time intervals $[\phi_n, \sigma_{n+1}]$ are then simplified to

$$\sum_n e^{-\lambda\phi_n}\Bigl(f\bigl(X^\varepsilon(\sigma_{n+1})\bigr) - f\bigl(X^\varepsilon(\phi_n)\bigr)\Bigr)\,,$$

modulo contributions that converge to 0 as $\varepsilon \to 0$. Since the expectation of this term is bounded by

$$\sup_{x\in\mathscr{I}_{\eta\varepsilon}} \mathbb{E}_x\bigl(f\bigl(X^\varepsilon(\tau^\delta)\bigr) - f(x)\bigr)\sum_n \mathbb{E}e^{-\lambda\phi_n}\,,$$

and since we already know that $\sum_n \mathbb{E}e^{-\lambda\phi_n} = \mathcal{O}(1/\delta)$, it remains to show that the supremum is of order $o(\delta)$. It follows from Taylor's expansion and the fact that

$f \in \mathcal{C}^2$ outside of the interface, that on the event $\Omega_+ \stackrel{\text{def}}{=} \{X_1^\varepsilon(\tau^\delta) > 0\}$, one has

$$
\begin{aligned}
f\big(X^\varepsilon(\tau^\delta)\big) - f(x) =& \delta\partial_1 f(x)|_{I_+} + \sum_{i=2}^d \partial_i f(x)\big(X_i^\varepsilon(\tau^\delta) - x_i\big) \\
&+ \mathcal{O}\big(\big|X_i^\varepsilon(\tau^\delta) - x_i\big|^2\big) ,
\end{aligned}
$$

and similarly on $\Omega_- = \{X_1^\varepsilon(\tau^\delta) < 0\}$. Combining this with (4.4.4), we thus have

$$
\begin{aligned}
\mathbb{E}_x\big(f\big(X^\varepsilon(\tau^\delta)\big) - f(x)\big) = \delta\partial_1 f(x)|_{I_+}\mathbb{P}_x(\Omega_+) + \delta\partial_1 f(x)|_{I_-}\mathbb{P}_x(\Omega_-) \\
+ \delta\sum_{i=2}^d \alpha_i \partial_i f(x) + o(\delta) .
\end{aligned}
$$

Since we assume that $\mathbb{P}_x(\Omega_\pm) \to p_\pm$ uniformly over $x \in \mathscr{I}_{\eta\varepsilon}$, the required bound now follows from the gluing condition. □

Most of the remainder of this article is devoted to the verification of the assumptions of Theorem 4.4.1. The bounds (4.4.1) and (4.4.2) will be relatively straightforward to verify and this will form the content of the remainder of this section. The convergence (4.4.3) is the one that is most difficult to obtain and will be the content of Section 4.5. Finally, we will show that (4.4.4) holds in Section 4.6. We start by the following result:

Lemma 4.4.3. *Let $\bar{\mathcal{L}}_\pm$ be as in (4.2.1) and let X^ε be the family of processes from Section 4.2. Then, the bound (4.4.1) holds with $k(\varepsilon) = \varepsilon$ for every $\lambda > 0$ and for every smooth bounded function $f\colon I_i \to \mathbb{R}$ that has bounded derivatives of all orders.*

Proof. It follows from [HM10b, Lemma 3.4] that, for any initial point x with $x_1 \neq 0$ and for ε sufficiently small so that $x \notin \mathscr{I}_{\varepsilon\eta}$,

$$
\mathbb{E}_x\left[\int_0^{\tau^\varepsilon} e^{-\lambda s} f\big(X^\varepsilon(s)\big) h\left(\frac{X^\varepsilon(s)}{\varepsilon}\right) ds\right] = \mathcal{O}(\varepsilon) , \tag{4.4.6}
$$

for h centered with respect to μ_+ (resp. μ_- if $x_1 < 0$). We assume that $x_1 > 0$ from now on, but the calculations are identical for the case $x_1 < 0$.

Note now that it suffices to obtain the bound (4.4.1) for the family of processes Y^ε, since $\|Y^\varepsilon(t) - X^\varepsilon(t)\| = \mathcal{O}(\varepsilon)$, uniformly. Applying Itô's formula to $e^{-\lambda\tau^\varepsilon} f\big(Y^\varepsilon(\tau^\varepsilon)\big)$, we obtain the identity

$$
e^{-\lambda\tau^\varepsilon} f\big(Y^\varepsilon(\tau^\varepsilon)\big) = f(y) + \int_0^{\tau^\varepsilon} -\lambda e^{-\lambda s} f\big(Y^\varepsilon(s)\big)\, ds
$$

$$+\frac{1}{2}\int_0^{\tau^\varepsilon} e^{-\lambda s}\big(\tilde\sigma_{ik}\tilde\sigma_{kj}\big)\Big(\frac{X^\varepsilon}{\varepsilon}\Big)\partial^2_{ij}f(Y^\varepsilon(s))\,ds$$
$$+\int_0^{\tau^\varepsilon} e^{-\lambda s}\tilde\sigma_{ik}\Big(\frac{X^\varepsilon(s)}{\varepsilon}\Big)\partial_i f(Y^\varepsilon(s))\,dW_k(s)\;.$$

Since $|Y^\varepsilon - X^\varepsilon| \le \mathcal{O}(\varepsilon)$ and since all derivatives of f are assumed to be bounded, it then follows from (4.4.6) that

$$\mathbb{E}\Big(e^{-\lambda\tau^\varepsilon} f\big(Y^\varepsilon(\tau^\varepsilon)\big)\Big) = f(y) - \lambda\mathbb{E}\int_0^{\tau^\varepsilon} e^{-\lambda s} f\big(Y^\varepsilon(s)\big)\,ds$$
$$+\frac{1}{2}\mathbb{E}\int_0^{\tau^\varepsilon} e^{-\lambda s} D^+_{ij}\partial^2_{ij}f(Y^\varepsilon(s))\,ds + \mathcal{O}(\varepsilon)\;,$$

which is precisely the required result. □

Additionally we have that the solution to $\bar{\mathcal{L}}_i u = \lambda u$ on I_i, $u = 1$ on $\{x_1 = 0\}$ and $\{x_1 = \pm 1\}$, is bounded and has bounded derivatives of all orders. This follows from the fact that u is given explicitly by $u(x) = C_1 e^{\sqrt{\lambda (D^\pm_{11})^{-1}}x_1} + C_2 e^{-\sqrt{\lambda (D^\pm_{11})^{-1}}x_1}$ for some constants C_i. We now show that the process Y^ε satisfies the bound (4.4.2), i.e. it does not spend too much time in the vicinity of the interface:

Lemma 4.4.4. *If we choose $\delta = \varepsilon^\alpha$ for any $\alpha \in (\frac{1}{2}, 1)$, then (4.4.2) holds for the family of processes X^ε from Section 4.2.*

Proof. Again, it suffices to show the bound for the process Y^ε since it differs from X^ε by $\mathcal{O}(\varepsilon)$. We would like to use an argument similar to what can be used in the one-dimensional case [HM10b], that is we time-change the corrected process Y^ε in such a way that it becomes a diffusion with diffusion coefficient 1. Its drift then vanishes outside of the interface and is bounded by K/ε for some $K > 0$. At this stage, one compares this process to the 'worst-case scenario' process Z^ε given by

$$dZ^\varepsilon = \hat b(Z^\varepsilon)\,dt + dB(t)\;,$$

where the drift $\hat b$ is given by

$$\hat b(z) = \begin{cases} -K\varepsilon^{-1} & \text{if } z \in [0, l\varepsilon), \\ K\varepsilon^{-1} & \text{if } z \in (-l\varepsilon, 0), \\ 0 & \text{otherwise.} \end{cases}$$

for some $l \in \mathbb{R}$. It can then be shown that Z^ε spends more time in the interface than Y^ε does, so that the requested bound can be obtained from a simple calculation.

The problem with this argument is that in the multi-dimensional case the time-change required to turn the first component of Y^ε into a diffusion with unit diffusion coefficient is given by

$$T_t = \inf\Big\{ s \in \mathbb{R}_+ : \int_0^s \sum_{i=1}^n \Big(\delta_{1i} + \partial_i g_1\big(\varepsilon^{-1} X^\varepsilon(u)\big)\Big)^2 du > t \Big\}\,. \tag{4.4.7}$$

We do not know of an argument giving a uniform bound *from below* on the quantity appearing under the integral in this expression. Therefore, an upper bound on the time spent by the process Z^ε in the interval $(-\delta, \delta)$ does not give us any control on the time spent by Y^ε (and therefore X^ε) in that interval.

Because of this, we modify our argument in the following way. We break up the integral in (4.4.2) as

$$\mathbb{E}_x\Big[\int_0^\infty e^{-\lambda t}\mathbf{1}_{(-\delta,\delta)}\big(Y_1^\varepsilon(t)\big)\,dt\Big] = \mathbb{E}_x\Big[\int_0^\infty e^{-\lambda t}\mathbf{1}_{(-c\varepsilon,c\varepsilon)}\big(Y_1^\varepsilon(t)\big)\,dt\Big] \tag{4.4.8}$$

$$+\,\mathbb{E}_x\Big[\int_0^\infty e^{-\lambda t}\mathbf{1}_{(-\delta,-c\varepsilon)}\big(Y_1^\varepsilon(t)\big)\,dt\Big] \tag{4.4.9}$$

$$+\,\mathbb{E}_x\Big[\int_0^\infty e^{-\lambda t}\mathbf{1}_{(c\varepsilon,\delta)}\big(Y_1^\varepsilon(t)\big)\,dt\Big]\,,$$

where Y_1^ε is the first component of Y^ε and c is a value to be determined. By symmetry, the last two terms are of the same order, so that it is sufficient to bound the first two terms. In order to bound the first term, we use the argument outlined above, but we replace Y^ε by the process $\tilde Y^\varepsilon$ given by $\tilde Y^\varepsilon(t) = X^\varepsilon(t) + \varepsilon\tilde g(\varepsilon^{-1}X^\varepsilon(t))$, where the corrector $\tilde g$ has the following properties:

1. The function $\tilde g(x)$ is smooth, periodic in the variables parallel to the interface, and equal to $g(x)$ for $x \notin \mathscr{I}_{c_1}$ for some c_1.

2. One has the implication $Y^\varepsilon \in \mathscr{I}_{c\varepsilon} \Rightarrow \tilde Y^\varepsilon \in \mathscr{I}_{c_2\varepsilon}$ for some $c_2 < c_1$.

3. If $\tilde Y^\varepsilon \in \mathscr{I}_{c_2\varepsilon}$, then $\tilde g(\varepsilon^{-1}X^\varepsilon) = 0$.

It is always possible to satisfy these properties by choosing c_1 sufficiently large and setting $g = 0$ in a sufficiently wide band around the interface. We now set $\tilde Z(t) = \tilde Y(\tilde T_t)$, where $\tilde T_t$ is defined as in (4.4.7), but with g replaced by $\tilde g$, so that it follows from the second property that one has the bound

$$\mathbb{E}_x \int_0^\infty e^{-\lambda t}\mathbf{1}_{(-c\varepsilon,c\varepsilon)}\big(Y_1^\varepsilon(t)\big)\,dt \le \mathbb{E}_x \int_0^\infty e^{-\lambda t}\mathbf{1}_{(-c_2\varepsilon,c_2\varepsilon)}\big(\tilde Y_1^\varepsilon(t)\big)\,dt$$

$$\le \mathbb{E}_x \int_0^\infty e^{-\lambda T_t} \mathbf{1}_{(-c_2\varepsilon, c_2\varepsilon)}\big(\tilde Z_1^\varepsilon(t)\big)\, dT_t \;.$$

At this stage, we remark that since the function $\tilde g$ has bounded derivatives, there exists a constant K_1 such that $T_t \ge K_1 t$ almost surely. On the other hand, it follows from the last property that one actually has $dT_t = dt$ whenever $\tilde Y^\varepsilon \in \mathscr{I}_{c_2\varepsilon}$, so that this expression is bounded by

$$\mathbb{E}_x \int_0^\infty e^{-K_1 t} \mathbf{1}_{(-c_2\varepsilon, c_2\varepsilon)}\big(\tilde Z_1^\varepsilon(t)\big)\, dt \;.$$

This expression in turn can be bounded by $\mathcal{O}(\varepsilon)$ just as in [HM10b].

We now proceed to bounding the term (4.4.9). For this, let us first introduce a constant $c_3 < c$ and make c from (4.4.8) sufficiently large such that

4. The implication $X^\varepsilon(t) \in \mathscr{I}_{c_3\varepsilon} \Rightarrow Y^\varepsilon(t) \in \mathscr{I}_{c\varepsilon}$ holds.

5. One has $c_3 > \eta + 1$.

Then, we define a series of stopping times $\{\phi_n'\}_n$ and $\{\sigma_n'\}_n$ recursively by $\phi_{-1}' = 0, \ldots, \sigma_n' = \inf\{t \ge \phi_{n-1}' : X_1^\varepsilon(t) \notin (-2\delta, -c_3\varepsilon + \varepsilon)\}$ and $\phi_n' = \inf\{t \ge \sigma_n' : X_1^\varepsilon(t) \in (-\delta, -c_3\varepsilon)\}$.

Now we can use the Strong Markov property as in [FW93, Lemma 4.1] with the stopping times ϕ_n' to obtain the bound $\mathbb{E}_x\big[\sum_{n=0}^\infty e^{-\lambda\sigma_n'(\varepsilon)}\big] = O\big(\frac{1}{\varepsilon}\big)$, uniformly in the initial point x for $x \in \{x : x_1 = -c_3\varepsilon + \varepsilon\} \cup \{x : x_1 = -2\delta\}$. This is a consequence of the fact that $\mathbb{E}_x[e^{-\lambda\sigma_0'}] = 1 - O(\varepsilon)$ uniformly. Furthermore, it follows from the definition of these stopping times, property 4., and the strong Markov property that (4.4.9) is bounded by

$$\begin{aligned}
\mathbb{E}_x \int_0^\infty e^{-\lambda t} \mathbf{1}_{(-\delta, -c_3\varepsilon)}\big(X_1^\varepsilon(t)\big)\, dt &\le \mathbb{E}_x \sum_{n\ge 0} \int_{\phi_{n-1}'}^{\sigma_n'} e^{-\lambda t}\, dt \\
&\le \lambda^{-1} \mathbb{E}_x \sum_{n \ge 0} e^{-\lambda \phi_{n-1}'} \big(\sigma_n' - \phi_{n-1}'\big) \qquad (4.4.10) \\
&\le \lambda^{-1} \Big(\mathbb{E}_x \sum_{n=0}^\infty e^{-\lambda\phi_n'(\varepsilon)}\Big) \sup_x \mathbb{E}_x \sigma_0' \le \frac{C}{\varepsilon\lambda} \sup_x \mathbb{E}_x \sigma_0' \;.
\end{aligned}$$

It follows that it suffices to be able to choose δ in such a way that $\mathbb{E}_x\sigma_0'$ is $o(\varepsilon)$ uniformly in the initial point. Specifically, we will show that (4.4.10) is $\mathcal{O}(\delta^2)$, so that the claim follows.

This will be a consequence of the following result:

Lemma 4.4.5. *Let X^- be as in (4.1.3) and define $X^{-,\varepsilon}(t) = \varepsilon X^-(\varepsilon^{-2}t)$. Let $\tau = \inf\{t > 0 : X_1^{-,\varepsilon}(t) \notin [-1,0]\}$. Then, there exists a constant C such that*

$$\mathbb{E}_x \tau \leq C \,,$$

independently of $\varepsilon \in (0,1]$ and independently of $x \in \mathbb{R}^d$.

Before we prove Lemma 4.4.5, we use it to complete the proof of Lemma 4.4.4. It follows from property 5. that up to time σ_0', the process X^ε is identical in law to the process $X^{-,\varepsilon}$. Furthermore, the stopping time σ_0' is certainly bounded from above by the first exit time of the first component of $X^{-,\varepsilon}$ from $(-2\delta, 0)$. Rescaling space by a factor 2δ and rescaling time correspondingly by $4\delta^2$, we deduce from Lemma 4.4.5 that $\mathbb{E}\sigma_0' \leq 4C\delta^2$, uniformly in the initial condition as required. □

We now turn to the

Proof of Lemma 4.4.5. Denote by U the region $\{x \in \mathbb{R}^d : x_1 \in [-1,0]\}$ and define f^ε by $f^\varepsilon(x) = \mathbb{E}_x \tau$. Then f^ε satisfies

$$\mathcal{L}^\varepsilon f^\varepsilon = -1, \quad f^\varepsilon(x) = 0 \text{ for } x \in \partial U,$$

where $\mathcal{L}^\varepsilon = \frac{1}{2}\Delta + \varepsilon^{-1} b_-(\varepsilon^{-1}\cdot)\nabla_x$. In order to obtain a bound on f, we will give a uniformly bounded (uniformly over ε) function g^ε such that it satisfies,

$$\mathcal{L}^\varepsilon g^\varepsilon = -1, \quad g^\varepsilon(x) \geq 0 \text{ for } x \in \partial U. \tag{4.4.11}$$

It then follows from the strong maximum principle (which we can apply since our diffusion is periodic in the directions in which U is unbounded) that $g^\varepsilon \geq f^\varepsilon$, so that the requested bound holds.

We use a standard multiscale expansion for g^ε of the form

$$g^\varepsilon = g_0 + \varepsilon g_1 + \varepsilon^2 g_2 \,.$$

Now to find such a g^ε. We proceed by starting off with a constant order term, that is, the typical term one would expect for the escape time if we were dealing with a Brownian motion, then removing the order $\frac{1}{\varepsilon}$ terms that arise when the operator $\mathcal{L}^\varepsilon$ acts on the constant order term by adding an order ε term. Then finally, we add an order ε^2 term to remove the constant order terms that are produced by the action of $\mathcal{L}^\varepsilon$ on the order ε term. Incidentally, this approach of correction works exactly with the maximum order term in ε being 2 and produces a series of terms that are known and have the right properties to provide a uniform bound.

Taking guidance from the fact that the homogenized process is given by Brownian motion, we make the ansatz $g_0(x) = C_2 - C_1 x_1(1 + x_1)$, for C_1 and C_2 two constants to be determined. Applying $\mathcal{L}^\varepsilon$ to g_0 yields

$$\mathcal{L}^\varepsilon g_0(x) = -C_1 - \frac{C_1}{\varepsilon} b_{-,1}\Big(\frac{x}{\varepsilon}\Big)(1 + 2x_1) .$$

for $b_{-,1}$ the first component of b_-. Our aim now is to choose g_1 in such a way that $\mathcal{L} g_1$ contains a term of order ε^{-1} that precisely cancels out the second term in this expression. Denote as in the introduction by g_- the unique centered solution to the Poisson equation

$$\mathcal{L} g_- = b_- , \tag{4.4.12}$$

where $\mathcal{L} = \frac{1}{2}\Delta + b_- \nabla_x$ is the generator for the non-rescaled process. We then set $g_1(x) = C_1(1 + 2x_1) g_{-,1}(\varepsilon^{-1}x)$, where $g_{-,1}$ is the first component of g_-, and we note that

$$\begin{aligned}
\varepsilon \mathcal{L}^\varepsilon g_1(x) &= \frac{C_1}{\varepsilon} b_{-,1}\Big(\frac{x}{\varepsilon}\Big)(1 + 2x_1) + 2C_1 b_{-,1}\Big(\frac{x}{\varepsilon}\Big) g_{-,1}\Big(\frac{x}{\varepsilon}\Big) + 2C_1 \frac{\partial g_{-,1}}{\partial x_1}\Big(\frac{x}{\varepsilon}\Big) \\
&= \frac{C_1}{\varepsilon} b_{-,1}\Big(\frac{x}{\varepsilon}\Big)(1 + 2x_1) + C_1 F\Big(\frac{x}{\varepsilon}\Big) ,
\end{aligned} \tag{4.4.13}$$

for some periodic function F independent of ε and of C_1. The term involving F appearing in this expression is still of order one, so we aim to compensate it by a judicious choice of g_2. It is not necessarily centred with respect to the invariant measure μ of our process, but there exists a periodic centred function h such that

$$\begin{aligned}
\mathcal{L} h = F - K , \quad K &= \int F(x)\, \mu(dx) \\
&= -\int |\nabla g_{-,1}(x)|^2\, \mu(dx) + 2\int \frac{\partial g_{-,1}}{\partial x_1}\, \mu(dx) .
\end{aligned}$$

Finally, setting $g_2(x) = -h(\varepsilon^{-1}x)$, we obtain

$$\mathcal{L}^\varepsilon g^\varepsilon = C_1(K - 1) = -C_1 \int |e_1 - \nabla g_{-,1}(x)|^2\, \mu(dx) . \tag{4.4.14}$$

Since the integral is strictly positive, the right hand side can be made to be equal to -1. Furthermore, since the corrector terms $\varepsilon g_1 + \varepsilon^2 g_2$ are uniformly bounded for $\varepsilon < 1$, it is straightforward to find a constant C_2 that ensures that $g(x) \geq 0$ for $x \in \partial U$, thus concluding the proof. □

4.5 Computation of the transmissivity coefficient

The aim of this section is to prove that

Proposition 4.5.1. *The identity (4.4.3) holds for the family of processes X^ε in Section 4.2 with $p_\pm$ given by (4.2.2).*

Let us first introduce some notation. Given a starting point $x \in \mathscr{I}_\eta$, we set $p_+^{x,k} = \mathbb{P}_x(X(\tau^{(k)}) > 0)$, and similarly for $p_-^{x,k}$, where $\tau^{(k)}$ is the first hitting time of $\partial\mathscr{I}_k$. We furthermore set

$$\bar{p}_+^k = \sup_{x\in\mathscr{I}_\eta} p_+^{x,k}\;, \qquad \underline{p}_+^k = \inf_{x\in\mathscr{I}_\eta} p_+^{x,k}\;, \qquad p_+^{(k)} = \frac{1}{2}\big(\bar{p}_+^k + \underline{p}_+^k\big)\;,$$

and similarly for p_-. It is clear that Proposition 4.5.1 follows if we can show that p_+^k converges to a limit satisfying (4.2.2) and $\bar{p}_+^k - \underline{p}_+^k \to 0$ as $k \to \infty$.

We will first show the latter, as it is relatively straightforward to show. In order to show the convergence of p_+^k, our main ingredient will be to show that the invariant measure $\mu(dx)$ for the process X looks more and more similar to $\mu_\pm(dx)$ as $x_1 \to \pm\infty$. Note that in this whole section, we will always consider X and $X_\pm$ as processes on $\mathbb{R}\times\mathbb{T}^{d-1}$, obtained by identifying points (x,y) such that $x_1 = y_1$ and $x_j - y_j \in \mathbb{Z}$ for $j \geq 2$. With this interpretation, the interface is compact and we will show that the processes are recurrent. If we were to consider them as processes in $\mathbb{R}^d$, they would *not* be recurrent for $d \geq 3$.

Before we show that indeed $\bar{p}_+^k - \underline{p}_+^k \to 0$, we obtain some recurrence properties of X and ensure that it visits any open set in $\mathscr{I}_\eta$ sufficiently often before the hitting time $\tau^{(k)}$.

Lemma 4.5.2. *Fix a neighborhood $\gamma \subset \mathscr{I}_\eta$. Then the probability for X to enter γ before hitting $\partial\mathscr{I}_k$, starting from an arbitrary initial point in $\mathscr{I}_\eta$ tends to 1 uniformly as $k \to \infty$. In particular, the process X is recurrent.*

Our first step in showing this result is to argue that if the process starts at distance $\mathcal{O}(1)$ of the interface, then it will return to the interface with overwhelming probability before exiting $\mathscr{I}_k$:

Lemma 4.5.3. *There exists $K > 0$ such that the probability, starting at x, for X to return to $\mathscr{I}_\eta$ before hitting $\partial\mathscr{I}_k$, is bounded from above by $1 - \frac{x-K}{k}$ and from below by $1 - \frac{x+K}{k}$.*

Proof. Denote by $f^k(x)$ the probability of hitting $\mathscr{I}_\eta$ before $\partial\mathscr{I}_k$, starting from x. We assume without loss of generality that $x_1 > 0$, since the case $x_1 < 0$ follows using the same argument. The function f^k then satisfies the equation $\mathcal{L}f^k = 0$, endowed with the boundary conditions $f^k(x) = 1$ if $x_1 = \eta$ and $f^k(x) = 0$ if $x_1 = k$. As in the proof of Lemma 4.4.5, we aim to construct a function g^k satisfying $\mathcal{L}g^k = 0$ and such that either $g^k(x) \le f^k(x)$ on the two boundaries or $g^k(x) \ge f^k(x)$ on the two boundaries. The claim then follows from the maximum principle.

Let g_+ be as in (4.4.12) and set

$$g^k(x) = 1 - k^{-1}\big(K + x_1 - g_{+,1}(x)\big) \;,$$

for some constant K to be determined. It is straightforward to check that g^k does indeed satisfy $\mathcal{L}g^k = 0$, as well as the required inequalities on the boundary, provided that K is either sufficiently large or sufficiently small. This concludes the proof. □

We now use the result of lemma 4.5.3 to prove lemma 4.5.2. This is done using the strong Markov property in conjunction with success/failure trials.

Proof of Lemma 4.5.2. Consider the two hyperplanes that delimit $\mathscr{I}_\eta$ and two further hyperplanes at distance m from $\mathscr{I}_\eta$, with m a sufficiently large constant to be determined later. We then break the process into excursions from $\partial\mathscr{I}_\eta$ to $\partial\mathscr{I}_{\eta+m}$ and back.

More precisely, we define two sets of stopping times $\{\sigma_n^m\}_n$ and $\{\phi_n^m\}_n$ recursively by $\sigma_1^m = \inf\{t \ge 0 : X(t) \in \partial\mathscr{I}_{\eta+m}\}, \ldots, \phi_n^m = \inf\{t > \sigma_n^m : X(t) \in \mathscr{I}_\eta\}$, $\sigma_{n+1}^m = \inf\{t > \phi_n^m : X(t) \in \partial\mathscr{I}_{\eta+m}\}$. We furthermore denote by $\mathscr{F}_n$ the σ-algebra generated by trajectories of X up to the time ϕ_n^m and by $\bar{\mathscr{F}}_n$ the σ-algebra generated by trajectories of X up to the time σ_{n+1}^m. We also denote by τ_γ the first hitting time of the set γ and by $\tau^{(k)}$ the first hitting time of the set $\partial\mathscr{I}_k$.

It follows from the ellipticity of X and the resulting smoothness of its transition probabilities that there exists some $p > 0$ such that

$$\inf_{x\in\partial\mathscr{I}_\eta} P_1(x,\gamma) = 2p > 0 \;.$$

Furthermore, it is straightforward, for instance using a comparison argument with a process with constant drift away from the interface and using the continuity of paths, to show that

$$\lim_{m\to\infty} \sup_{x\in\mathscr{I}_\eta} \mathbb{P}_x(\sigma_1^m \le 1) = 0 \;. \tag{4.5.1}$$

It follows that we can choose m large enough so that the probability appearing in (4.5.1) is bounded above by p. As a consequence, for such a choice of m, one has the almost sure bound

$$\mathbb{P}\big(\tau_\gamma < \sigma_{n+1}^m \,|\, \mathscr{F}_n\big) \ge p \;. \tag{4.5.2}$$

On the other hand, it follows from Lemma 4.5.3 that the probability that the process hits $\partial\mathscr{I}_k$ between σ_n^m and ϕ_n^m is bounded from *above* uniformly by $\beta_k = \mathcal{O}(k^{-1})$ so that, almost surely,

$$\mathbb{P}\big(\tau^{(k)} < \phi_{n+1}^m \,|\, \bar{\mathscr{F}}_n\big) \le \beta_k \;. \tag{4.5.3}$$

Note furthermore that by construction the event appearing in (4.5.2) is $\bar{\mathscr{F}}_n$-measurable.

Denote now by Y_n a Markov chain with states $\{-1, 0, 1\}$ such that $\{\pm 1\}$ are absorbing and such that $P(Y_{n+1} = -1 \,|\, Y_n = 0) = p$, $P(Y_{n+1} = 1 \,|\, Y_n = 0) = \beta_k$. As a consequence of (4.5.2) and (4.5.3), it is then possible to couple Y and X in such a way that the following two implications hold almost surely:

$$\begin{aligned}\{Y_n = 0 \text{ and } Y_{n+1} = -1\} &\Rightarrow \{\phi_n^m < \tau_\gamma < \sigma_{n+1}^m < \tau^{(k)}\}\\ \{\sigma_{n+1}^m < \tau^{(k)} < \phi_{n+1}^m < \tau_\gamma\} &\Rightarrow \{(Y_n = 0 \text{ and } Y_{n+1} = 1)\}\end{aligned}$$

It follows that the probability of entering γ before the hitting time $\tau^{(k)}$ is bounded from below by

$$\mathbb{P}(\tau_\gamma < \tau^{(k)}) \ge \mathbb{P}(\lim_{n\to\infty} Y_n = -1) = \frac{p}{p+\beta_k} \;.$$

Since p is fixed and $\beta_k = \mathcal{O}(k^{-1})$, this quantity can be made arbitrarily close to 1.

This shows that the set γ is recurrent for X. Since furthermore X has transition probabilities that have strictly positive densities with respect to Lebesgue measure (as a consequence of the ellipticity of the equations describing it), recurrence follows from [MT93, Theorem 8.0.1]. □

We now use this result to prove

Proposition 4.5.4. *$\bar{p}_{+}^{k} - \underline{p}_{+}^{k} \to 0$ as $k \to \infty$.*

Proof. The idea is to use the fact that, before the process exits $\mathscr{I}_k$, it has had sufficient amount of time to forget about its initial condition by visiting a small set on which a strong minorising condition holds for its transition probabilities.

Fix a value $\beta > 0$. Our aim is to show that there then exists $k_0 > 0$ such that

$$\underline{p}_{\pm}^{k} \geq p_{\pm}^{0,k} - \beta \;,$$

say, for every $k \geq k_0$. Since $p_{+}^{x,k} = 1 - p_{-}^{x,k}$, the claim then follows. We restrict ourselves to the bound for p_{+} since the other bound can be obtained in exactly the same way.

The argument is now the following. It follows from the smoothness of transition probabilities that there exists a neighbourhood γ of the origin such that the transition probabilities at time 1 for X, starting from γ satisfy the lower bound

$$\rho(y) = \inf_{x\in\gamma} P_1(x,y) \;,$$

with $\int_{\mathbb{R}\times\mathbb{T}^{d-1}} \rho(y)\,dy \geq 1 - \beta/2$. It then follows immediately that for $x \in \gamma$, one has $p_{+}^{x,k} \geq p_{+}^{0,k} - \beta/2 - \mathbb{P}_x\big(\exists t \leq 1 \,:\, X(t) \in \partial\mathscr{I}_k\big)$. For arbitrary x, it therefore follows from the strong Markov property that

$$p_{+}^{x,k} \geq p_{+}^{0,k} - \beta/2 - \sup_{y\in\gamma} \mathbb{P}_y\big(\exists t \leq 1 \,:\, X(t) \in \mathscr{I}_k\big) - \mathbb{P}_x\big(X \text{ hits } \partial\mathscr{I}_k \text{ before } \gamma\big) \;.$$

The last term can be made smaller than $\beta/4$ by Lemma 4.5.2. The remaining term $\mathbb{P}_y\big(\exists t \leq 1 \,:\, X(t) \in \mathscr{I}_k\big)$ on the other hand was already shown to be arbitrary small in (4.5.1). □

We next show that the invariant measure of the process converges to that of the relevant periodic process with increasing distance from the interface.

Proposition 4.5.5. *Let A denote a bounded measurable set and denote by μ the (unique up to scaling) invariant σ-finite measure of the process X. Denote furthermore by $\mu_\pm$ the invariant measure of the relevant periodic process, normalised in such a way that $\mu_\pm([k,k+1]\times\mathbb{T}^{d-1}) = 1$ for every $k \in \mathbb{Z}$. Then there exist normalisation constants $q_\pm$ such that,*

$$\lim_{k\to\infty} \big(|\mu(A+k) - q_+\mu_+(A)| + |\mu(A-k) - q_-\mu_-(A)|\big) = 0 \;. \tag{4.5.4}$$

(Here k is an integer.) Furthermore, this convergence is exponential, and uniform over the set A if we restrict its diameter.

Remark 4.5.6. *We used the shorthand notation $A + k$ for $\{x + k \,:\, x \in A\}$.*

Proof. We restrict ourselves to the estimate of $\mu(A+k)$, since the one on $\mu(A-k)$ is similar. For fixed $k \geq 0$, we introduce the sequence of stopping times given by $\phi_0^{(k)} = \inf\{t \geq 0 \,:\, X_1(t) = k\}$ and then recursively $\sigma_n^{(k)} = \inf\{t \geq \phi_n^{(k)} \,:\, |X_1(t) - k| = 1\}$, $\phi_{n+1}^{(k)} = \inf\{t \geq \sigma_n^{(k)} \,:\, X_1(t) = k\}$. This allows us to define an embedded Markov chain $Z^{(k)}$ on $\mathbb{T}^{d-1}$ by setting $Z_n^{(k)} = \Pi X(\phi_n^{(k)})$, where $\Pi(x,y) = y$ for $(x,y) \in \mathbb{R} \times \mathbb{T}^{d-1}$.

We similarly define an embedded Markov chain Z for the process X^+. (By periodicity of X^+, the choice of k is unimportant for the law of Z, so that we drop its dependence of k.) Denote by $\pi^{(k)}$ the invariant measure for $Z^{(k)}$ and by π the invariant measure for Z. We then define σ-finite measures μ_+ and $\mu^{(k)}$ on $\mathbb{R} \times \mathbb{T}^{d-1}$ through the identities

$$\mu^{(k)}(B) = \int_{\mathbb{T}^{d-1}} \mathbb{E}_{x+ke_1} \int_0^{\phi_1^{(k)}} \mathbf{1}_B(X(s))\, ds\, \pi^{(k)}(dx)\,, \tag{4.5.5}$$

$$\mu_+(B) = \int_{\mathbb{T}^{d-1}} \mathbb{E}_{x+ke_1} \int_0^{\phi_1^{(k)}} \mathbf{1}_B(X^+(s) - k)\, ds\, \pi(dx)\,. \tag{4.5.6}$$

(Here and below we make a slight abuse of notation and identify elements $x \in \mathbb{T}^{d-1}$ with the element $(0,x) \in \mathbb{R} \times \mathbb{T}^{d-1}$.) It follows from [Has60, Thm 2.1] that $\mu^{(k)}$ is invariant for the process X and μ_+ is invariant for X^+. Therefore, there exist constants $c_k > 0$ such that $\mu^{(k)} = c_k \mu$ since the invariant measure for X is unique up to normalisation. Note that by translation invariance of X^+, μ_+ does not depend on k.

Note that we can assume without any loss of generality that $A \subset \{x \,:\, x_1 > 0\}$ (it suffices to shift it by a finite number of steps to the right in (4.5.4)). In this case, we can rewrite (4.5.5) as

$$\mu^{(k)}(A+k) = \int_{\mathbb{T}^{d-1}} \mathbb{E}_{x+ke_1} \int_0^{\phi_1^{(k)}} \mathbf{1}_A(X^+(s) - k)\, ds\, \pi^{(k)}(dx)\,. \tag{4.5.7}$$

This is because $X(t) = X^+(t)$ for $t \leq \sigma_1^{(k)}$ and, if $X(\sigma_1^{(k)}) < k$, then

$$\int_{\sigma_1^{(k)}}^{\phi_1^{(k)}} \mathbf{1}_A(X(s) - k)\, ds = 0\,,$$

whereas if $X(\sigma_1^{(k)}) > k$, then $X(t) = X^+(t)$ for $t \le \phi_1^{(k)}$. This shows that the claim follows if we can show that $\|\pi - \pi^{(k)}\|_{\mathrm{TV}} \to 0$ as $k \to \infty$ and there exists a constant c_∞ such that $c_k \to c_\infty$.

Let us first show that the latter is a consequence of the former. Setting $B_k = [k, k+1] \times \mathbb{T}^{d-1}$, we have $c_{k+1}/c_k = \mu^{(k)}(B_{k+1})/\mu^{(k+1)}(B_{k+1})$. On the other hand a straightforward trial/error argument allows one to show that $\mathbb{E}_x \int_0^{\phi_1^{(0)}} \mathbf{1}_A(X^+(s))\,ds$ is bounded uniformly over $x \in \mathbb{T}^{d-1}$. It then follows immediately from (4.5.7) that there exists a constant C such that

$$|\mu^{(k)}(B_{k+1}) - \mu(B_0)| \le C\|\pi - \pi^{(k)}\|_{\mathrm{TV}}\;,$$

and similarly for $|\mu^{(k+1)}(B_{k+1}) - \mu(B_0)|$. It follows that provided that $\sum_{k\ge 0} \|\pi - \pi^{(k)}\|_{\mathrm{TV}} < \infty$, one does indeed have $c_k \to c_\infty$.

Denote now by P the transition probabilities for Z and by $P^{(k)}$ the transition probabilities for $Z^{(k)}$. Then, we can write $P = QR$, where R is the Markov kernel from $\mathbb{T}^{d-1}$ to $\{-1,1\} \times \mathbb{T}^{d-1}$ given by $R(x, A) = \mathbb{P}_x(X^+(\sigma_1) \in A)$ and Q is the Markov kernel from $\{-1,1\} \times \mathbb{T}^{d-1}$ to $\mathbb{T}^{d-1}$ given by $Q(x, A) = \mathbb{P}_x(X^+(\phi_0) \in A)$ for $X_1(0) = 0$, $\sigma_1 = \inf\{t > 0 : |X_1(t)| = 1\}$ and $\phi_1 = \inf\{t > \sigma_1 : X_1(t) = 0\}$. Since the diffusion X^+ is elliptic, both Q and R are strong Feller and irreducible. It follows from the Doeblin-Doob-Khasminskii theorem [DPZ96, Proposition 4.1.1] that $P(x,\cdot\,)$ and $P(y,\cdot\,)$ are mutually equivalent for any $x, y \in \mathbb{T}^{d-1}$. Furthermore, it follows from the Meyer-Mokobodzki theorem [DM83, Sei01, Hai07] that the map $x \mapsto P(x,\cdot\,)$ is continuous in the total variation topology. We conclude that the map $(x,y) \mapsto \|P(x,\cdot\,) - P(y,\cdot\,)\|_{\mathrm{TV}}$ reaches its maximum and that this is strictly less than 2, so that P satisfies Doeblin's condition. It follows that there exists a constant $\eta < 1$ such that P has the contraction property

$$\|P\nu_1 - P\nu_2\|_{\mathrm{TV}} \le \eta\|\nu_1 - \nu_2\|_{\mathrm{TV}}\;,$$

for any two probability measures ν_1, ν_2 on $\mathbb{T}^{d-1}$. Therefore, if we can find constants ε_k such that

$$\sup_{x\in\mathbb{T}^{d-1}} \|P(x,\cdot\,) - P^{(k)}(x,\cdot\,)\|_{\mathrm{TV}} \le \varepsilon_k\;, \tag{4.5.8}$$

then we have

$$\|\pi - \pi^{(k)}\|_{\mathrm{TV}} \le \|P\pi - P\pi^{(k)}\|_{\mathrm{TV}} + \|P\pi^{(k)} - P^{(k)}\pi^{(k)}\|_{\mathrm{TV}} \tag{4.5.9}$$

$$\leq \eta \|\pi - \pi^{(k)}\|_{\mathrm{TV}} + \varepsilon_k ,$$

so that $\|\pi - \pi^{(k)}\|_{\mathrm{TV}} \leq \varepsilon_k/(1-\eta)$. The problem thus boils down to obtaining (4.5.8) for an exponentially decaying sequence ε_k.

It follows from the same calculation as in Lemma 4.5.3 that the probability that X reaches the interface $\mathscr{I}_\eta$ before time $\phi_1^{(k)}$ when started on the hyperplane $\{x_1 = k\}$ is bounded from above by $\mathcal{O}(1/k)$. This yields the "trivial" bound $\varepsilon_k \leq \mathcal{O}(1/k)$, which unfortunately isn't even summable. However, a more refined analysis allows to obtain Proposition 4.5.7 below, thus concluding the proof. □

Proposition 4.5.7. *There exists a constant $\rho \in (0,1)$ such that $\varepsilon_k \leq \mathcal{O}(\rho^k)$.*

Proof. The intuitive idea behind the proof of Proposition 4.5.7 is that if the process goes all the way back to the interface then, by the time it reaches again the plane $\{x_1 = k\}$, its hitting distribution depends only very little on its behaviour near the interface. In order to formalise this, let us introduce the Markov transition kernel Q_+ from $\mathbb{T}^{d-1}$ to $\mathbb{T}^{d-1}$ which is such that $Q_+(x,\cdot\,)$ is the hitting distribution of the plane $\{1\} \times \mathbb{T}^{d-1}$ for the process X_+ started at $(0,x)$. Similarly, we denote by $Q^{\ell,k}(x,\cdot\,)$ the hitting distribution of the plane $\{k\} \times \mathbb{T}^{d-1}$ for the process X started at (ℓ, x).

For a fixed integer $\ell > \eta$, our aim is to show that $Q^{\ell,k}(x,\cdot\,)$ gets very close to $Q_+^{k-\ell}(x,\cdot\,)$. Here, we denote by Q_+^k the kth iteration of the Markov transition kernel Q_+. With these notations at hand, define the quantities

$$\alpha_k \equiv \sup_{x\in\mathbb{T}^{d-1}} \|Q^{\ell,k}(x,\cdot\,) - Q_+^{k-\ell}(x,\cdot\,)\|_{\mathrm{TV}} ,$$

Note now that since, for fixed ℓ, the probability that X reaches the interface $\mathscr{I}_\ell$ before time $\phi_1^{(k)}$ when started on the hyperplane $\{x_1 = k\}$ is bounded from above by $\mathcal{O}(1/k)$, we have

$$\begin{aligned} \varepsilon_k &\leq \sup_{x\in\mathbb{T}^{d-1}} \|Q^{k-1,k}(x,\cdot) - Q_+(x,\cdot)\|_{\mathrm{TV}} \\ &\leq \frac{C}{k} \sup_{x\in\mathbb{T}^{d-1}} \|Q^{\ell,k}(x,\cdot\,) - Q_+^{k-\ell}(x,\cdot\,)\|_{\mathrm{TV}} \leq \frac{C}{k}\alpha_k , \end{aligned} \tag{4.5.10}$$

so that it suffices to obtain an exponentially decaying bound on the α_k's.

We now look for a recursion relation on the α_k's which then yields the required bound. We have the identities $Q^{\ell,k} = Q^{k-1,k}Q^{\ell,k-1}$ and $Q_+^{k-\ell} = Q_+ Q_+^{k-\ell-1}$. It follows from the triangle inequality that one has the bound

$$\|Q^{\ell,k}\delta_x - Q_+^{k-\ell}\delta_x\|_{\mathrm{TV}} \le \|(Q^{k-1,k} - Q_+)Q^{\ell,k-1}\delta_x\|_{\mathrm{TV}} \\ + \|Q_+(Q^{\ell,k-1}\delta_x - Q_+^{k-\ell-1}\delta_x)\|_{\mathrm{TV}} . \tag{4.5.11}$$

At this stage, we note that by exactly the same reasoning as for P, the kernel Q_+ satisfies Doeblin's condition. Therefore, there exists a constant $\bar\eta < 1$ such that

$$\|Q_+\nu_1 - Q_+\nu_2\|_{\mathrm{TV}} \le \bar\eta \|\nu_1 - \nu_2\|_{\mathrm{TV}} ,$$

for any two probability measures ν_1, ν_2. This and the definition of α_k immediately implies that the second term in (4.5.11) is uniformly bounded by $\bar\eta\alpha_{k-1}$. On the other hand, it follows from (4.5.10) that the first term is bounded by $\frac{C}{k}\alpha_k$, so that

$$\alpha_k \le \frac{C}{k}\alpha_k + \bar\eta\alpha_{k-1} ,$$

for some fixed constant C. The claim now follows at once. □

Finally, the last estimate that we need is the following. Denote by τ the first hitting time of the interface $\partial\mathscr{I}_\eta$ and fix an arbitrary smooth positive function φ that is supported in the interval $[1,2]$. Set furthermore $\varphi_n^+(x) = n^{-2}\varphi(n^{-1}x_1)$ and $\varphi_n^-(x) = n^{-2}\varphi(-n^{-1}x_1)$. Then we have:

Lemma 4.5.8. *With the above notations, setting $\bar\varphi = \int_1^2 \varphi(x)\,dx$, we have*

$$\Big|\mathbb{E}_x \int_0^\tau \varphi_n^\pm(X_\pm(t))\,dt - \frac{2\bar\varphi}{D_{11}^\pm}\Big| \to 0 ,$$

uniformly for all $x \in \{\pm n\} \times \mathbb{T}^{d-1}$ as $n \to \infty$.

Proof. Again, we only consider the expression for X_+, the one for X_- follows in the same way. It follows from standard homogenization results [BLP78, PS08] that the law of $n^{-1}X_+(n^2t)$ converges weakly as $n \to \infty$ to the law of Brownian motion with diffusion coefficient D_{11}^+. It thus follows from [Bog07, Cor 8.4.2] that the law of $n^{-1}X_+(n^2t)$, where X_+ is stopped at the first hitting time of $\mathscr{I}_\eta$ converges weakly as $n \to \infty$ to the law of Brownian motion stopped when it reaches the hyperplane $\mathscr{I}_0$.

Denoting this limiting process by X_+^∞, an explicit calculation allows to check that $\mathbb{E}_x \int_0^\tau \varphi(X_+^\infty(t))\,dt = \frac{2\bar\varphi}{D_{11}^\pm}$ when $x_1 = 1$. Now, for any fixed $T > 0$, the map $\Phi_T \colon X \mapsto \int_0^{\tau\wedge T} \varphi_n^+(X(t))\,dt$ is continuous, so that $\mathbb{E}_x \int_0^{\tau\wedge T} \varphi_n^+(X_+(t))\,dt$ converges as $n \to \infty$ to $\mathbb{E}_x \int_0^{\tau\wedge T} \varphi(X_+^\infty(t))\,dt$. Letting $T \to \infty$ concludes the proof. □

We now have all the tools that we need to show that the exit probabilities from the interface converge to the desired limiting values.

Proof of Proposition 4.5.1. Similarly to the proof of Proposition 4.5.5 we use a representation of the invariant measure μ in terms of an embedded Markov chain. This time, we consider the stopping times $\tilde\phi_0^{(k)} = \inf\{t \geq 0 : |X_1(t)| = \eta\}$ and then $\tilde\sigma_n^{(k)} = \inf\{t \geq \tilde\phi_n^{(k)} : |X_1(t)| = k\}$, $\tilde\phi_{n+1}^{(k)} = \inf\{t \geq \tilde\sigma_n^{(k)} : |X_1(t)| = \eta\}$. Denoting as similar to before by $\tilde\pi^{(k)}$ the invariant measure of the embedded Markov chain $\tilde Z_n^{(k)} = X(\tilde\phi_n^{(k)})$ (which is now a Markov chain on $\partial\mathscr{I}_\eta$), we set

$$\tilde\mu^{(k)}(B) = \int_{\partial\mathscr{I}_\eta} \mathbb{E}_x \int_0^{\tilde\phi_1^{(k)}} \mathbf{1}_B(X(s))\,ds\,\tilde\pi^{(k)}(dx)\;. \tag{4.5.12}$$

Again, the measures $\tilde\mu^{(k)}$ differ from μ purely through a scaling factor, so that there are constants C_k such that $\tilde\mu^{(k)}(B) = \tilde C_k \mu(B)$ for every measurable set B.

The idea now is to evaluate $\tilde\mu^{(k)}(\varphi_k^\pm)$ in two different ways and to compare the resulting answers. First, we note from Proposition 4.5.5 that

$$\mu^{(k)}(\varphi_k^\pm) = \frac{C_k}{k}\big(q_\pm\bar\varphi + \mathcal{O}(k^{-1})\big)\;.$$

On the other hand, combining Proposition 4.5.4 and Lemma 4.5.8 with the definition (4.5.12), we see that

$$\mu^{(k)}(\varphi_k^\pm) = \frac{2p_\pm^{(k)}\bar\varphi}{D_{11}^\pm} + o(1) \tag{4.5.13}$$

as $k \to \infty$. Combining these two identities, we see that

$$\frac{p_+^{(k)}}{p_-^{(k)}} = \frac{D_{11}^+ q_+}{D_{11}^- q_-} + o(1)\;,$$

thus concluding the proof. □

4.6 Computation of the drift along the interface

This section is devoted to the computation of the drift coefficients α_j along the interface. Denote by τ^n the first hitting time of $\partial\mathscr{I}_n$ by the process X. With this notation, recall that, by (4.4.4), we have the identity

$$\alpha_j = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_x \int_0^{\tau^n} b_j(X_s)\, ds\;, \tag{4.6.1}$$

provided that this limit exists and is independent (and uniform) over starting points $x \in \mathscr{I}_\eta$.

Proposition 4.6.1. *The expression on the right hand side in (4.6.1) converges to the expression given by (4.2.4), uniformly in $x \in \mathscr{I}_\eta$.*

In order to show this, we will use the same construction as in the proof of Proposition 4.5.1. In particular, recall the definition (4.5.12) of the measures $\tilde\mu^{(k)}$, which are nothing but multiples of the invariant measure μ, as well as the sequence of stopping times $\tilde\phi_n^{(k)}$ and $\tilde\sigma_n^{(k)}$. Denote furthermore by $\tilde\pi_n^{(k)}$ the invariant measure for the process on $\partial\mathscr{I}_\eta$ with transition probabilities $P(x,A)$ given by

$$P(x,A) \stackrel{\text{def}}{=} \mathbb{P}_x\big(X(\tilde\phi_1^{(k)}) \in A \,|\, \tau^n > \tilde\phi_1^{(k)}\big)\;. \tag{4.6.2}$$

Our proof will proceed in two steps. First, we show that the limit (4.6.1) exists and is equal to the value (4.2.4) given in the interface, provided that we start the process X in the stationary measure $\tilde\pi_n^{(k)}$ and let $k\to\infty$. In the second step, we then show by a coupling argument similar to the proof of Proposition 4.5.4 that the expression in (4.6.1) depends only weakly on the initial condition as n gets large, thus concluding the proof.

Before we proceed with this programme, we perform the following preliminary calculation:

Lemma 4.6.2. *One has the normalisation*

$$\lim_{k\to\infty} k^{-2}\tilde\mu^{(k)}\big([-k,k]\times\mathbb{T}^{d-1}\big) = 2\Big(\frac{p_+}{D_{11}^+} + \frac{p_-}{D_{11}^-}\Big) \stackrel{\text{def}}{=} \beta\;,$$

where the coefficients $p_\pm$ are as in (4.2.2). In particular, if μ is normalised as in the introduction, then one has $k^{-1}\tilde\mu^{(k)} \approx \beta\mu$ for large values of k.

Proof. We know from Proposition 4.5.5 that $\mu(dx) \to \mu_\pm(dx)$ at exponential rate as $x_1 \to \pm\infty$, so that on large scales μ behaves like a multiple of Lebesgue measure on either side of the interface. Furthermore, we know from Proposition 4.5.1 that the corresponding normalisation constants satisfy the relation (4.2.2). Combining this with the fact that $\tilde{\mu}^{(k)}$ is just a multiple of μ, the result then follows from (4.5.13). □

Using this result, we obtain:

Proposition 4.6.3. *The limit*

$$\alpha_j = \lim_{k\to\infty}\lim_{n\to\infty}\frac{1}{n}\mathbb{E}_{\tilde{\pi}_n^{(k)}}\int_0^{\tau^n} b_j(X_s)\,ds\;,$$

exists and is equal to

$$\beta\int_{\mathbb{R}\times\mathbb{T}^{d-1}}\big(b_j(x)+\mathcal{L}g_j(x)\big)\,\mu(dx)\;, \tag{4.6.3}$$

where g is the function fixed in Section 4.3 and the constant β is as in Lemma 4.6.2.

Remark 4.6.4. *Note that if ϕ is any smooth compactly supported function, then the identity $\int \mathcal{L}\phi(x)\,\mu(dx) = 0$ holds. As a consequence, the expression (4.6.3) is independent of the choice of the compensator g.*

Proof. It follows from the definition of $\tilde{\pi}_n^{(k)}$ and the strong Markov property of X that one has the identity

$$\begin{aligned}\mathbb{E}_{\tilde{\pi}_n^{(k)}}\int_0^{\tau^n}\tilde{b}_j(X_s)\,ds &= \sum_{m\geq 0}\big(\mathbb{P}_{\tilde{\pi}_n^{(k)}}(\tilde{\phi}_1^{(k)}<\tau^n)\big)^m\,\mathbb{E}_{\tilde{\pi}_n^{(k)}}\int_0^{\tilde{\phi}_1^{(k)}\wedge\tau^n}\tilde{b}_j(X_s)\,ds\\ &= \frac{\mathbb{E}_{\tilde{\pi}_n^{(k)}}\int_0^{\tilde{\phi}_1^{(k)}\wedge\tau^n}\tilde{b}_j(X_s)\,ds}{\mathbb{P}(\tilde{\phi}_1^{(k)}>\tau^n)}\;.\end{aligned} \tag{4.6.4}$$

Note now that it follows from Lemma 4.5.3 that

$$\mathbb{P}(\tilde{\phi}_1^{(k)}>\tau^n) = k/n + \mathcal{O}(1/n)\;. \tag{4.6.5}$$

Since $\lim_{n\to\infty} g_j(X(\tau^n))/n = 0$ and furthermore, using the same argument as in 4.5.9, we have $\lim_{n\to\infty}\|\tilde{\pi}_n^{(k)}-\tilde{\pi}^{(k)}\|_{\mathrm{TV}} = 0$ for every $k>0$, so that

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}_{\tilde{\pi}_n^{(k)}}\int_0^{\tau^n} b_j(X_s)\,ds = \lim_{n\to\infty}\frac{1}{n}\mathbb{E}_{\tilde{\pi}_n^{(k)}}\Big[\int_0^{\tau^n} b_j(X_s)\,ds + g_j(X(\tau_n))\Big]$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{\tilde\pi_n^{(k)}} \int_0^{\tau^n} \tilde b_j(X_s)\,ds \\
&= \lim_{n\to\infty} \frac{1}{k}\mathbb{E}_{\tilde\pi^{(k)}} \int_0^{\tilde\phi_1^{(k)}\wedge\tau^n} \tilde b_j(X_s)\,ds \\
&= \frac{1}{k}\mathbb{E}_{\tilde\pi^{(k)}} \int_0^{\tilde\phi_1^{(k)}} \tilde b_j(X_s)\,ds \\
&= \frac{1}{k}\int_{\mathbb{R}\times\mathbb{T}^{d-1}} \tilde b_j(x)\,\tilde\mu^{(k)}(dx)\;. \qquad (4.6.6)
\end{aligned}
$$

Here, we used (4.6.4) and (4.6.5) to go from the second to the third line and we used the definition of the $\tilde\mu^{(k)}$ to obtain the last identity. The claim now follows from Lemma 4.6.2. □

We can now complete the

Proof of Proposition 4.6.1. In view of Proposition 4.6.3, it remains to show that

$$
\lim_{n\to\infty}\frac{1}{n}\Big|\mathbb{E}_x\int_0^{\tau^n} b(X_s)\,ds - \mathbb{E}_y\int_0^{\tau^n} b(X_s)\,ds\Big| = 0\;,
$$

uniformly over $x, y \in \mathscr{I}_\eta$. Fix an arbitrary value of $k > \eta$ and consider again the transition probabilities P given by (4.6.2). Since they arise as exit probabilities for an elliptic diffusion, we can show again by the same argument as in the proof of Proposition 4.5.5 that P satisfies the Doeblin condition for some constant η, namely $\|P\nu_1 - P\nu_2\|_{\mathrm{TV}} \le (1-\eta)\|\nu_1-\nu_2\|_{\mathrm{TV}}$, uniformly over probability measures ν_1 and ν_2 on $\partial\mathscr{I}_\eta$. Note now that one has the identity

$$
\begin{aligned}
\mathbb{E}_x\int_0^{\tau^n} b(X_s)\,ds &= \sum_{m\ge 0}\Big(\prod_{0\le\ell<m}\mathbb{P}_\ell^x(\tilde\phi_1^{(k)}<\tau^n)\Big)\mathbb{E}_m^x\int_0^{\tilde\phi_1^{(k)}\wedge\tau^n} b(X_s)\,ds \\
&= \sum_{m\ge0}\mathbb{P}_x(\tilde\phi_m^{(k)}<\tau^n)\,\mathbb{E}_m^x\int_0^{\tilde\phi_1^{(k)}\wedge\tau^n} b(X_s)\,ds\;, \qquad (4.6.7)
\end{aligned}
$$

where we denote by $\mathbb{P}_m$ (resp. $\mathbb{E}_m$) the probability (resp. expectation) for the process X started at $P^m(x,\cdot\,)$.

Note now that we have the identity

$$
\mathbb{P}_x(\tilde\phi_m^{(k)}<\tau^n) = \mathbb{P}_x\big(\tilde\phi_\ell^{(k)}<\tau^n\big) + \mathbb{P}_{P^\ell(x,\cdot)}\big(\tilde\phi_{m-\ell}^{(k)}<\tau^n\big)\;.
$$

Also, by choosing k sufficiently large (but independent of n), we can ensure that there exist constants $c, C > 0$ such that

$$
1-\frac{C}{n} \le \mathbb{P}_x(\tilde\phi_1^{(k)}<\tau^n) \le 1-\frac{c}{n}\;,
$$

uniformly for $x \in \mathscr{I}_\eta$ and for n sufficiently large. It also follows from the contraction properties of P that

$$\big|\mathbb{P}^x_m(\tilde\phi_1^{(k)} < \tau^n) - \mathbb{P}^y_m(\tilde\phi_1^{(k)} < \tau^n)\big| \le 2(1-\eta)^m ,$$

uniformly over $x, y \in \mathscr{I}_\eta$.

Combining these bounds, we obtain for every $\ell \le m \wedge n$ the estimate

$$\big|\mathbb{P}_x(\tilde\phi_m^{(k)} < \tau^n) - \mathbb{P}_y(\tilde\phi_m^{(k)} < \tau^n)\big| \le \frac{K\ell}{n} + 2(1-\eta)^\ell .$$

In particular, there exists a constant K, such that we have the uniform bound

$$\big|\mathbb{P}_x(\tilde\phi_m^{(k)} < \tau^n) - \mathbb{P}_y(\tilde\phi_m^{(k)} < \tau^n)\big| \le \frac{K}{\sqrt{n}} \wedge \frac{Km}{n} \wedge \Big(1-\frac{c}{n}\Big)^m ,$$

valid for every $m > 0$ and every n sufficiently large. Summing over m, it follows that

$$\sum_{m\ge 0} \big|\mathbb{P}_x(\tilde\phi_m^{(k)} < \tau^n) - \mathbb{P}_y(\tilde\phi_m^{(k)} < \tau^n)\big| \le K\sqrt{n} ,$$

for a possibly different constant K.

On the other hand, it is possible to check that there exists a constant C (depending on k) such that

$$\Big|\mathbb{E}_x \int_0^{\tilde\phi_1^{(k)}\wedge\tau^n} b(X_s)\,ds\Big| \le C ,$$

uniformly over $x \in \mathscr{I}_\eta$, so that

$$\Big|\mathbb{E}^x_m \int_0^{\tilde\phi_1^{(k)}\wedge\tau^n} b(X_s)\,ds - \mathbb{E}^y_m \int_0^{\tilde\phi_1^{(k)}\wedge\tau^n} b(X_s)\,ds\Big| \le 2C(1-\eta)^m .$$

Inserting these bounds into 4.6.7, we obtain

$$\Big|\mathbb{E}_x \int_0^{\tau^n} b(X_s)\,ds - \mathbb{E}_y \int_0^{\tau^n} b(X_s)\,ds\Big| \le 2C \sum_{m\ge 0} (1-\eta)^m + C\sqrt{n} ,$$

so that the requested bound follows at once. □

4.6.1 Bound on the second moment

In order to conclude the verification of the assumptions of Theorem 4.2.4, it remains to show that the second bound holds in (4.4.4). For the non-rescaled process, we can reformulate this as

Proposition 4.6.5. *For every $\bar\eta > 0$, there exists a constant $C > 0$ such that the bound*

$$\mathbb{E}_y \big\| Y(\tau^n) - y \big\|^2 \le C n^2 \;,$$

holds for every $n \ge 1$ and every initial condition $y \in \mathscr{I}_{\bar\eta}$.

Proof. It follows from (4.3.1) that

$$\mathbb{E}_y \big\| Y(\tau^n) - y \big\|^2 \le 2\mathbb{E}_y \Big\| \int_0^{\tau^n} \tilde b(X_s)\, ds \Big\|^2 + 2\mathbb{E}_y \Big\| \int_0^{\tau^n} \tilde\sigma(X_s)\, dW(s) \Big\|^2 . \tag{4.6.8}$$

It follows from Itô's isometry that the second term is bounded by $C\mathbb{E}\tau^n$, which in turn is bounded by $\mathcal{O}(n^2)$ by a calculation virtually identical to that of Lemma 4.4.5.

It remains to bound the first term, which we will do with the help of a decomposition similar to that used in the proof of Proposition 4.5.1. For two constants $c > 0$ and $a > 0$ to be determined, we set $\phi_0 = 0$, $\sigma_n = \inf\{t \ge \phi_n : |X_1(t)| = c + a\}$, and $\phi_n = \inf\{t \ge \sigma_{n-1} : |X_1(t)| = c\}$. Define furthermore

$$N = \inf\{k \ge 0 : \sigma_k \ge \tau^n\} \;.$$

Since $\tilde b$ is supported in a bounded strip around $\mathscr{I}_0$, we can make c sufficiently large so that the first term in (4.6.8) is bounded by some multiple of

$$\begin{aligned} \mathbb{E}_y \Big(\sum_{k=0}^{N} (\sigma_k - \phi_k) \Big)^2 &\le \sqrt{\mathbb{E}_y N^3 \mathbb{E}_y \sum_{k=0}^{N} (\sigma_k - \phi_k)^4} \\ &\le \sqrt{\mathbb{E}_y N^3 \sum_{k=0}^{\infty} \mathbb{E}_y \big((\sigma_k - \phi_k)^4 \,\big|\, N \ge k \big) \mathbb{P}_y(N \ge k)} \;. \end{aligned}$$

Note now that since σ_k is the exit time from a compact region for an elliptic diffusion, there exists a constant C such that $\mathbb{E}_y \big((\sigma_k - \phi_k)^4 \,\big|\, N \ge k \big) \le C$, uniformly in y. Furthermore, it follows from Lemma 4.5.3 that if a is sufficiently large, then

$$\mathbb{P}_y(N > 1) \le 1 - \frac{c}{n} \;,$$

for some constant $c > 0$, uniformly in y. The strong Markov property then immediately implies that $\mathbb{P}_y(N > k) \le \big(1 - \frac{c}{n}\big)^k$, so that N is stochastically bounded by a Poisson random variable with parameter $\mathcal{O}(n)$ and the claim follows. □

4.7 Well-posedness of the martingale problem and characterization of the limiting process

The aim of this section is to show that the martingale problem associated to the operator $\bar{\mathcal{L}}$ as defined in Theorem 4.2.4 is unique and to characterise the corresponding (strong) Markov process. Our main tool is the following general result by Ethier and Kurtz [EK86, Theorem 4.1]:

Theorem 4.7.1. *Let E be a separable metric space, and let $A\colon \mathscr{D}(A) \to \mathcal{B}_b(E)$ be linear and dissipative. Suppose there exists $\lambda > 0$ such that*

$$\mathcal{C} \stackrel{def}{=} \overline{\mathscr{R}(\lambda - A)} = \overline{\mathscr{D}(A)}\;, \tag{4.7.1}$$

and such that $\mathcal{C}$ is separating. Let $\mu \in \mathscr{P}(E)$ and suppose X is a solution of the martingale problem for (A, μ). Then X is a Markov process corresponding to the semigroup on $\mathcal{C}$ generated by the closure of A, and uniqueness holds for the martingale problem for (A, μ).

See also [BP87] for a more general result on the well-posedness of a martingale problem with discontinuous coefficients. This allows us go finally give the

Proof of Theorem 4.2.4. Since we already know from the results in the previous two sections that limit points of X^ε solve the martingale problem associated to $\bar{\mathcal{L}}$, it suffices to show that this martingale problem is well-posed and that its solutions are of the form (4.2.6).

For this, we somehow take the reverse approach: first, we construct a solution to (4.2.6) and we show that this is a Markov process solving the martingale problem associated to $\bar{\mathcal{L}}$. We then show that this Markov process generates a strongly continuous semigroup on $\mathcal{C}_0(\mathbb{R}^d)$, whose generator is the closure of $\bar{\mathcal{L}}$ in $\mathcal{C}_0$. Since $\mathcal{C}_0$ is separating and since generators of strongly continuous semigroups are dissipative and satisfy (4.7.1) by the Hille-Yosida theorem, the claim then follows.

In order to construct a solution to (4.2.6), let $M_\pm$ be matrices satisfying $M_\pm M_\pm^T = D^\pm$ and such that

$$M_\pm = \begin{pmatrix} \sqrt{D_{11}^\pm} & 0 \\ v_\pm & \tilde{M}_\pm \end{pmatrix}\;,$$

for some vectors $v_\pm \in \mathbb{R}^{d-1}$ and some $(d-1)\times(d-1)$ matrices $\tilde{M}_\pm$. (This is always possible by the QR decomposition.) We then first construct a Wiener process W_1 and a process $\bar{X}_1$ such that

$$d\bar{X}_1 = \big(\mathbf{1}_{\bar{X}_1\le 0}\sqrt{D_{11}^-} + \mathbf{1}_{\bar{X}_1>0}\sqrt{D_{11}^+}\big)\, dW(t) + (p_+ - p_-)\, dL(t)\;,$$

where L is the symmetric local time of $\bar{X}_1$ at the origin. This can be achieved for example by setting $\bar{X}_1 = g(Z)$, where

$$g(x) = \begin{cases} \sqrt{D_{11}^+} & \text{if } x > 0, \\ \sqrt{D_{11}^-} & \text{otherwise,} \end{cases}$$

Z is a skew-Brownian motion with parameter

$$p = \frac{p_+\sqrt{D_{11}^-}}{p_+\sqrt{D_{11}^-} + p_-\sqrt{D_{11}^+}}\;,$$

and W is the martingale part of Z. Given such a pair $(\bar{X}_1, W)$, we then let $\tilde{W}$ be an independent $d-1$-dimensional Wiener process and we define pathwise the $\mathbb{R}^{d-1}$-valued process $\tilde{X}$ by

$$\begin{aligned}\tilde{X}(t) = &\int_0^t \big(\mathbf{1}_{\bar{X}_1\le 0}\tilde{M}_- + \mathbf{1}_{\bar{X}_1>0}\tilde{M}_+\big)\, d\tilde{W}(t) + \int_0^t \big(\mathbf{1}_{\bar{X}_1\le 0}v_- + \mathbf{1}_{\bar{X}_1>0}v_+\big)\, dW(t) \\ &+ \tilde{\alpha}\int_0^t dL(t)\;,\end{aligned}$$

where $\tilde{\alpha}_j = \alpha_{j+1}$. Since we know that skew-Brownian motion enjoys the Markov property, it follows immediately that $\bar{X}_1$ is Markov, so that $\bar{X} = (\bar{X}_1, \tilde{X})$ is also a Markov process. Applying the symmetric Itô-Tanaka formula to $f(\bar{X})$ it is furthermore a straightforward exercise to check that $\bar{X}$ does indeed solve the Martingale problem for $\bar{\mathcal{L}}$.

The corresponding Markov semigroup $\{\mathcal{P}_t\}_{t\ge 0}$ maps $\mathcal{C}_0(\mathbb{R}^d)$ into itself as a consequence of the Feller property of skew-Brownian motion [Lej06]. Furthermore, as a consequence of the uniform stochastic continuity of $\bar{X}$, it is strongly continuous, so that its generator must be an extension of $\bar{\mathcal{L}}$. Since the range of $\bar{\mathcal{L}}$ contains $\mathcal{C}_0^\infty(\mathbb{R}^d)$, which is a dense subspace of $\mathcal{C}_0(\mathbb{R}^d)$, the claim follows. □

In this chapter we make an attempt to generalize the approach of the previous chapter to the situation of a reflected diffusion process where the interface reflecting the particle sits at an angle to the period with an irrational tangent. At present this remains a work in progress, in some sense it is 'close' to solution but so far ultimately the ergodic results presented are insufficient.

Chapter 5

Homogenization of a diffusion process reflected at an angle with irrational tangent

Consider the diffusion process in the half plane $\mathbb{R}^d_+ = \{x \in \mathbb{R}^d : x_1 > 0\}$ given by,

$$X(t) = X_0 + \int_0^t b(X(s))\,ds + W(t) + \int_0^t \gamma(X(s))\,dL^0_s\,, \tag{5.0.1}$$

where b, γ are d dimensional vectors with $\gamma_1 \equiv 1$, W is a d dimensional Brownian motion and $L^0(t)$ is the two sided local time of $X(t)$ on $x_1 = 0$, given by,

$$L^0(t) = \lim_{\delta \to 0} \frac{1}{2\delta} \int_0^t 1_{[0,\delta)}(X(s))\,ds\,.$$

For this SDE there is pathwise uniqueness of solution and hence uniqueness in law of solution from [LS84]. In addition, assume that γ is C^1 and periodic in every direction contained in the hyperplane $x_1 = 0$, and that b is periodic in the sense that there exists an orthonormal basis of vectors for $\mathbb{R}^d$, $\tilde{e}_i$, $i = 1,\ldots, d$, such that $b(x + \tilde{e}_i) = b(x)$ for all i, $x \in \mathbb{R}^d$. It is also assumed that the angle between $\tilde{e}_i$ and x_1 has an irrational tangent for at least one i since otherwise we would be in the same situation to that explored in [Tan84]. The usual centering condition is in force, i.e. we have $\int b\,d\mu = 0$, for μ the invariant measure of the diffusion given by

$$V(t) = V_0 + \int_0^t b(V(s))\,ds + W(t)\,.$$

We will denote the generator of V by $\mathcal{L}$.

Then under a diffusive rescaling $X_t^\varepsilon = \varepsilon X_{\varepsilon^{-2}t}$, the weak limit in continuous path space in $\mathbb{R}^d$, $C([0,\infty),\mathbb{R}^d)$ of the family of processes,

$$X(t) = X_0 + \frac{1}{\varepsilon}\int_0^t b\left(\frac{X(s)}{\varepsilon}\right) ds + W(t) + \int_0^t \gamma\left(\frac{X(s)}{\varepsilon}\right) dL^{0,\varepsilon}(s) ,$$

for $L^{0,\varepsilon}(t)$ the two sided local time of $X^\varepsilon(t)$ on $x_1 = 0$, as $\varepsilon \to 0$, is studied.

In order to carry out this homogenization the framework established in [FW93] will be used once again. This time however in one sense the situation is more simple in that it is no longer necessary to establish convergence of the probabilities of exiting on the left or right of a large neighborhood of the interface since there is only one direction to exit the multidimensional vertex this time. In order to attempt to obtain the expected point of exit in all other coordinate directions, inspiration is drawn from the method implemented in [HM10a], but this time we cannot 'roll the process up' to allow the state space to become a tube. Previously, an invariant measure of an embedded Markov chain (a process obtained from the returns of the full process) on the boundary of the interface $\partial\mathscr{I}_\eta$ was obtained with these identifications, to which there was convergence in total variation norm of this Markov chain. Instead one has to be clever about choosing a suitable state space on which to consider a similar embedded Markov chain. The Markov process in this case is formed on a subset of the $2d-1$ dimensional cube with the subspace topology. As in [HM10a] this is again as a result of forming the product of identifications with the position in the periodic cell that gives the drift times the position occupied in the interface. This part of the proof is incomplete at present though, since we cannot verify one of the hypotheses of the ergodic result that it is suspected can be used. The verification of tightness is quite simple apart from the local time term since there is an everywhere defined corrector.

5.1 Tightness where we have the angle of the interface at a angle with an irrational tangent

As in the multidimensional case, the main tool in the proof of Theorem 5.1.2 is the following result [HM10a], which is very similar to [SV79, Thm 1.4.6]:

Proposition 5.1.1. *Let $\mathscr{P}$ be a family of probability measures on $\Omega = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$ and denote by x the canonical process on Ω. Assume that*

$$\lim_{R \nearrow \infty} \sup_{\mathbb{P} \in \mathscr{P}} \mathbb{P}\big(|x(0)| \geq R \big) = 0 \, .$$

Furthermore, for any given $\rho > 0$, let $\tau_0 = 0$, and define recursively $\tau_{i+1} = \inf_{t > \tau_i} |x(t) - x(\tau_i)| > \rho$. Assume that the limit

$$\lim_{\delta \to 0} \operatorname{ess\,sup} \, \mathbb{P}\left[\tau_{n+1} - \tau_n \leq \delta | \mathscr{F}_{\tau_n}\right] \to 0, \quad \mathbb{P} \text{ a.s., on } \{\tau_n < \infty\}, \tag{5.1.1}$$

holds uniformly for every $\mathbb{P} \in \mathscr{P}$ and every $n \geq 0$. Then the family of probability measures $\mathscr{P}$ is tight on Ω.

Theorem 5.1.2. *The family of processes X_x^ε for fixed $x \in \mathbb{R}_+^d$ is tight in the space $C([0,\infty), \mathbb{R}_+^d)$.*

Proof. In order to use Theorem 5.1.1, first correct the process using the periodic corrector, the existence of which follows from the centering condition on b, i.e. there is a function g such that $\mathcal{L} g = -b$. So now considering $Y^\varepsilon(t) = X^\varepsilon(t) + \varepsilon g(\varepsilon^{-1} X(t))$, then by the Itô formula, $Y^\varepsilon(t)$ satisfies,

$$Y_x^\varepsilon(t) = x + \varepsilon g\left(\frac{x}{\varepsilon}\right) + \int_0^t \tilde{\sigma}\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) dW(s) + \int_0^t \tilde{\gamma}\left(\frac{X_x^\varepsilon(s)}{\varepsilon}\right) dL^{0,\varepsilon}(s) \, ,$$

for $\tilde{\sigma} = I + \bigtriangledown g$, $\tilde{\gamma} = \gamma + \bigtriangledown g \gamma$. Instead of being a reflected diffusion process in the half plane, $Y^\varepsilon(t)$ is a process in the set $D^\varepsilon = \{x : \; x = y + \varepsilon g(\varepsilon^{-1} y), \, y \in \mathbb{R}_+^d\}$, but convergence in the Wasserstein metric, in $\mathbb{R}^d$ for instance, implies that tightness of $Y_x^\varepsilon(t)$ is equivalent to tightness of $X_x^\varepsilon(t)$.

$$\begin{aligned}
\mathbb{P}_x(\tau_{n+1} - \tau_n \leq \delta \,|\, \mathscr{F}_{\tau_n}) &= \mathbb{P}_{X_x^\varepsilon(\tau_{n-1})}\Big(\sup_{t < \delta} |Y_x^\varepsilon(t) - Y_x^\varepsilon(0)| > \rho \Big) \\
&\leq \sup_x \mathbb{P}_x\Big(\sup_{0 \leq t \leq \delta} \Big| \int_0^t \tilde{\sigma}\big(\varepsilon^{-1} X_x^\varepsilon(s)\big) \, dW(s) \Big| > \frac{\rho}{2} \Big) \\
&\quad + \sup_x \mathbb{P}_x\Big(\sup_{0 \leq t \leq \delta} \Big| \int_0^t \tilde{\gamma}(\varepsilon^{-1} X_x(s)) \, dL_s^{0,\varepsilon} \Big| > \frac{\rho}{2} \Big) \\
&\leq \frac{2}{\rho} C_1 \sup_x \mathbb{E}_x \sup_{0 \leq t \leq \delta} \Big(\int_0^t (\tilde{\sigma}.\tilde{\sigma})\big(\varepsilon^{-1} X_x^\varepsilon(s)\big) \, ds \Big)^{\frac{1}{2}} \\
&\quad + \sup_\varepsilon \sup_x \mathbb{P}_x\Big(\|\tilde{\gamma}\|_\infty L_\delta^{0,\varepsilon} > \frac{\rho}{2} \Big) \, ,
\end{aligned} \tag{5.1.2}$$

using the Burkholder Davis Gundy inequality together with the Chebychev inequality. In addition,

$$\sup_x \mathbb{E}_x \sup_{0\le t\le\delta} \Big(\int_0^t (\tilde{\sigma}.\tilde{\sigma})\big(\varepsilon^{-1}X_x^{\varepsilon}(s)\big)\, ds \Big)^{\frac{1}{2}} \le \|\tilde{\sigma}.\tilde{\sigma}\|_\infty \sqrt{\delta}\ ,$$

which implies that the first term on the RHS of (5.1.2) satisfies the hypothesis of Theorem 5.1.1.

Consider the SDE satisfied by the local time of the first component,

$$X_1(t) = \frac{1}{\varepsilon}\int_0^t b_1(\varepsilon^{-1}X^{\varepsilon}(s))\, ds + e_1.W(t) + L^{0,\varepsilon}(t)\ . \tag{5.1.3}$$

In order to obtain tightness, using Proposition 5.1.1, it is sufficient to have a uniform (over ε, and initial point) modulus of continuity for $L^{0,\varepsilon}(t)$. Where we say that a family of processes $Y^{i\in I}$ have a uniform modulus of continuity (over the family and initial point), if given $\rho > 0$, $\xi > 0$, there exists a $\delta = \delta(\rho, \xi)$ such that,

$$\sup_x \sup_{i\in I} \sup_{0\le t\le\delta} \mathbb{P}(|Y_x^i(t) - x| > \rho) < \xi\ . \tag{5.1.4}$$

In order to do this we will obtain a uniform modulus of continuity for the terms other than the local time term in (5.1.3). Noting that everywhere away from the boundary we simply have the periodic situation, we can obtain a uniform modulus of continuity for X_1 exactly as in [HM10b]. For the term $e_1.W_t$ we proceed exactly as above for the stochastic integral term using the Chebychev inequality followed by the Burkholder Davis Gundy inequality. The only problematic term is therefore,

$$\frac{1}{\varepsilon}\int_0^t b_1(\varepsilon^{-1}X^{\varepsilon}(s))\, ds\ . \tag{5.1.5}$$

For $\varepsilon \le \xi\rho/4\|g_1\|_\infty$, we will add a term to this of the form $\varepsilon\phi(X^\varepsilon(t))g_1(\varepsilon^{-1}X^\varepsilon(t))$ where $0 \le \phi \le 1$ smooth, satisfies $\phi(x) = \phi(e_1.x)$ is 0 in $\mathscr{I}_{k\varepsilon}$ and 1 outside $\mathscr{I}_{(k+1)\varepsilon}$ for some $k > 0$. Using Itô's formula we have,

$$\begin{aligned} &\frac{1}{\varepsilon}\int_0^t b_1(\varepsilon^{-1}X^{\varepsilon}(s))\, ds + \varepsilon\phi(X^\varepsilon(t))g_1(\varepsilon^{-1}X^\varepsilon(t)) \\ =&\frac{1}{\varepsilon}\int_0^t \tilde{b}_1(\varepsilon^{-1}X^{\varepsilon}(s))\, ds \\ &+ \int_0^t \nabla(\phi g_1)(\varepsilon^{-1}X^\varepsilon(s))\, dW_s + \varepsilon\phi(X^\varepsilon(0))g_1(\varepsilon^{-1}X^\varepsilon(0))\ , \end{aligned}$$

where $\tilde{b}(\cdot/\varepsilon)$ has support contained within $\mathscr{I}_{(k+1)\varepsilon}$. By the same argument once again we have a uniform modulus of continuity of the stochastic term above. To show that the term of bounded variation has a bound above for $t = \delta$ in expectation of $C\sqrt{\delta}$, it can be shown that the expected time up to δ spent in $\mathscr{I}_{(k+1)\varepsilon}$ is bounded above by $C\varepsilon\sqrt{\delta}$ by reusing the calculation from [HM10a] (noting that we have an identical comparison result to that in [HM10b] for a reflected diffusion process). Hence taking

$$\delta(\rho,\xi) = \min\{(\xi\rho)^2/16\|g_1\|_\infty\|b_1\|_\infty, (\xi\rho)^2/16C^2(\|\tilde{b}_1\|_\infty)^2, \\ (\xi\rho)^2/16\|\bigtriangledown\phi\|_\infty\|\bigtriangledown g_1\|_\infty, (\xi\rho)^2/16\}$$

we have (5.1.4) for (5.1.5). With the verification of the uniform modulus of continuity of the local time term, the proof of tightness is complete. □

5.2 Main Theorem

The theorem that it is suspected we could use to identify the weak limit point of the family X^ε_x is adapted from [HM10a] and based on the famous theorem of [FW93],

Theorem 5.2.1. *Let $\bar{\mathcal{L}}$ be a second order differential operator on $I = \{x \in \mathbb{R}^d : x_1 > 0\}$ with bounded coefficients and let D' be a set of test functions over I whose members are bounded and have bounded derivatives of all orders. Denote the hitting time of the set $\{x_1 < \varepsilon\}$ by τ^ε. Suppose that for any function $f \in D'$ and for any $\lambda > 0$, the bound*

$$\mathbb{E}_x\Bigg[e^{-\lambda\tau^\varepsilon} f\big(X^\varepsilon(\tau^\varepsilon)\big) - f\big(X^\varepsilon(0)\big) \\ + \int_0^{\tau^\varepsilon} e^{-\lambda t}\Big(\lambda f\big(X^\varepsilon(t)\big) - \bar{\mathcal{L}}_i f\big(X^\varepsilon(t)\big)\Big)\,dt\Bigg] = O\big(k(\varepsilon)\big)\,, \tag{5.2.1}$$

holds as $\varepsilon \to 0$, uniformly with respect to $x \in I$. Assume furthermore that the rate k is such that $\lim_{\varepsilon\to 0} k(\varepsilon) = 0$.

Condition 1: Assume that, for every $\lambda > 0$ there exist functions $u_\lambda \in D'$ such that $\bar{\mathcal{L}}u_\lambda(x) = \lambda u_\lambda(x)$ holds for $x \in I$ with $0 \le x_1 \le 1$ and such that $u_\lambda(x) = 1$ for $x_1 = 0$ and $x_1 = 1$.

Assume that there exists a rate $\delta = \delta(\varepsilon) \to 0$ such that $\delta(\varepsilon)/k(\varepsilon) \to \infty$ as $\varepsilon \to 0$

and such that for $\lambda > 0$,

$$\mathbb{E}_x^\varepsilon\Big[\int_0^\infty e^{-\lambda t}\mathbf{1}_{(-\delta,\delta)}\big(X_1^\varepsilon(t)\big)\,dt\Big] \to 0 \tag{5.2.2}$$

as $\varepsilon \to 0$, *uniformly in the initial point. Assume furthermore that there exist constants* α_j *and* C *such that*

$$\frac{1}{\delta}\,\mathbb{E}_x^\varepsilon\Big[X_j^\varepsilon(\tau^\delta) - x_j\Big] \to \alpha_j\;, \quad \frac{1}{\delta^2}\,\mathbb{E}_x^\varepsilon\Big[\big(X_j^\varepsilon(\tau^\delta) - x_j\big)^2\Big] \le C\;, \tag{5.2.3}$$

for $j \ge 2$. *Again, the limit is assumed to be uniform over* $x \in \mathscr{I}_{\varepsilon\eta}$ *as* $\varepsilon \to 0$, *and the inequality is assumed to be uniform over all* $\varepsilon \in (0,1]$ *and all* $x \in \mathscr{I}_{\varepsilon\eta}$.

Let then D *be the set of continuous functions* $f\colon \mathbb{R}^d \to \mathbb{R}$ *such that the restriction of* f *to* I *belongs to* D' *and such that the gluing condition*

$$\partial_1 f(0+) + \sum_{j=2}^d \alpha_j \partial_j f(0) = 0\;, \tag{5.2.4}$$

holds. Then, for any fixed $f \in D$, $t_0 \ge 0$ *and* $\lambda > 0$,

$$\begin{aligned}\Delta(\varepsilon) \;=\; \operatorname{ess\,sup}\Big|\,\mathbb{E}_x^\varepsilon\Big[\int_{t_0}^\infty e^{-\lambda t}\Big[\,\lambda f\big(X^\varepsilon(t)\big) - \bar{\mathcal{L}} f\big(X^\varepsilon(t)\big)\Big]\,dt \\ -\; e^{-\lambda t_0} f\big(X^\varepsilon(t_0)\big)\,\Big|\,\mathscr{F}_{[0,t_0]}\Big]\Big| \to 0 \end{aligned} \tag{5.2.5}$$

as $\varepsilon \to 0$, *uniformly with respect to* $x \in \{x \in \mathbb{R}^d \,:\, x_1 \ge 0\}$. *In particular, every weak limit of* X^ε *as* $\varepsilon \to 0$ *satisfies the martingale problem for* $\bar{\mathcal{L}}$.

Compared to the corresponding theorem in [HM10a], this time we have only one edge to worry about and hence $p_- \equiv 0$, $p_+ \equiv 1$. Therefore the resulting gluing conditions on $x_1 = 0$ are dependent on the α_j only.

We will now demonstrate how, given a sufficiently good ergodic result, it would be possible to apply this theorem.

Note that by the calculations in [HM10a] we have (5.2.1) and the existence of the functions that satisfy Condition 1 immediately. In addition by identical arguments to those in [HM10a] on one side of the interface only this time, we also have (5.2.2) for $\delta = \varepsilon^\alpha$ for $1/2 < \alpha < 1$. It remains to verify (5.2.3). We will begin with the second expression,

Lemma 5.2.2. *Recasting the problem as in [HM10a], for τ^n the escape time of $X(t)$ from $\mathscr{I}_n$, we have,*

$$\mathbb{E}_x\left[\left(X_j(\tau^n) - x_j\right)^2\right] \leq Cn^2 \, ,$$

for $j = 1, \ldots, d$.

Proof. Assuming that we have a bound of the form,

$$\mathbb{E}_x[L_0(\tau^n)] \leq C'n^2 \tag{5.2.6}$$

then the bound will follow by an identical calculation to that in [HM10a]. Fix $0 < c_1 < c_2$, and a smooth function $\phi : \mathbb{R}_+ \to \mathbb{R}_+$, such that for $0 \leq x < c_1$, $\phi(x) = 0$, and for $x > c_2$, $\phi(x) = 1$. Studying the first component, inspired by the approach in [HM10a], we introduce a smooth corrector $\tilde{g}_1(x) = \phi(x_1)g_1(x)$. Then we have,

$$X_1(t) + \tilde{g}_1(X(t)) = \int_0^t \tilde{b}(X(s))\, ds + \int_0^t \tilde{\sigma}(X(s))\, dW(s) + L^0(t) \, ,$$

for $\tilde{b}_1 = b_1 + \mathcal{L}\tilde{g}_1$ with support inside of $\mathscr{I}_{c_2}$ and $\tilde{\sigma} = e_1^T + \nabla \tilde{g}_1$. Now we have that,

$$\begin{aligned}\mathbb{E}_x[L^0(\tau^n)]^2 \leq & 4\mathbb{E}_x[X_1(\tau^n) + \tilde{g}_1(X(\tau^n))]^2 + 4\mathbb{E}_x\left[\int_0^{\tau^n} \tilde{b}(X(s))\, ds\right]^2 \\ & + 4E_x\left[\int_0^{\tau^n} \tilde{\sigma}(X(s))\, ds\right]^2 .\end{aligned}$$

The first term on the RHS is bounded above by $4(n + \|g\|_1)^2$, the second and third terms are bounded above by Cn^2 by the arguments presented in [HM10a, Proposition 6.5] which implies (5.2.6) and hence completes the proof. $\square$

Now for the problematic first expression in (5.2.3),

Proposition 5.2.3. *We have that the limit,*

$$\alpha_j = \lim_{\varepsilon \to 0} \frac{1}{\delta}\, \mathbb{E}_x^\varepsilon\left[X_j^\varepsilon(\tau^\delta) - x_j\right]$$

exists uniformly over the initial point x.

Remark 5.2.4. We will attempt to verify the value of α_j as a limit involving the invariant measure of an auxiliary Markov process but this expression is not particularly instructive, even if it were shown to be true.

Let us now introduce the series of auxiliary discrete time Markov processes that we suspect could be used in the proof of Proposition 5.2.3.

First the state space, M, of the auxiliary Markov processes, which we will denote by $Z^{(k)}$. The state space is a subset of $\mathbb{R}^{2d-1}$ with the topology inherited from $\mathbb{R}^{2d-1}$ and then we make some identifications. The projection of M onto the last d coordinates is equal to the unit cube with periodic identifications, $\tilde{C} \cong \mathbb{T}^d$, delimited by the sides $\tilde{e}_i$. In other words to construct $\tilde{C}$, identify all points in $\mathbb{R}^d$ that differ by $\tilde{e}_i$, $i = 1, \ldots, d$, using the map $\pi : \mathbb{R}^d \to \tilde{C}$. The position in the last d coordinates will then represent the position with regard to the periodic drift of a point on $\mathscr{I}_0$ in a loose sense. Next, assume that the reflection term has periodicity given by the vectors $l_i e_i$, $l_i \in \mathbb{R}$, $i = 2, \ldots, d$ and identify all points in $\{x_1 = 0\}$ that differ by $l_i e_i$ under the map π_l to form $C_I \cong \mathbb{T}^{d-1}$, a rectangle with periodic identifications. The first $d-1$ coordinates will then be used to represent the position in the period of the interface reflection term for a point on $\mathscr{I}_0$ in a loose sense. For each point $x \in \tilde{C}$ form the set $A_x = \cap_n \{y \in \pi_l(\{x_1 = 0\}) : \exists z \in \{x_1 = 0\}, \pi_l(z) = y, |x - \pi(z)| < 1/n\}$ we then form the state space as $\cup_{x \in \tilde{C}} A_x \times x \subset C_I \times \tilde{C}$. For $x \in \pi(\{x_1 = 0\})$ we clearly have $A_x \neq \emptyset$, moreover for such points A_x simply differs by a translation. If $x \notin \pi(\{x_1 = 0\})$, then by the hypothesis on the angles subtended by the vectors $\tilde{e}_i$ we have that for all n, the set $\{y \in \pi_l(\{x_1 = 0\}) : \exists z \in \{x_1 = 0\}, \pi_l(z) = y, |x - \pi(z)| < 1/n\}$ is non-empty, it contains z_n, say. Moreover any limit point of the z_n is a member of A_x, hence A_x is also non-empty in this case. For x, $x' \in \tilde{C}$ we have that A_x and $A_{x'}$ differ by a translation and in addition, we have the property that for any x, A_x is invariant under any translation of $C_I = \pi_l(\{x_1 = 0\})$ that takes any member of A_x to another member of A_x. This implies it is the same set of translations of C_I that permutes A_x for all $x \in \tilde{C}$. We also have continuity in the following sense, there exists $z'_n \in A_{x_n}$, $x_n \to x$ (in $\tilde{C}$) and $z'_n \to z'$ (in C_I) only if $z' \in A_x$. This implies that $M = \cup_{x \in \tilde{C}} A_x \times x$ is a closed subspace of $C_I \times \tilde{C}$ in the subspace topology. As a closed subset of a compact space, M is compact.

Consider the series of stopping times $\phi_0^{(k)} = \inf\{t > 0 : X(t) \in \{x_1 = 0\}\}$, $\sigma_0^{(k)} = \inf\{t > \phi_0^{(k)} : X(t) \in \{x_1 = k\}\}, \ldots, \phi_n^{(k)} = \inf\{t > \sigma_{n-1}^{(k)} : X(t) \in \{x_1 = 0\}\}$, $\sigma_n^{(k)} = \inf\{t > \phi_n^{(k)} : X(t) \in \{x_1 = k\}\}$, and the embedded Markov process on $\{x_1 = 0\}$, $X(\phi_n^{(k)})$. We can extend this Markov process on $\{x_1 = 0\}$ to a Markov process on $C_I \times \tilde{C}$ by moving the period relative to the interface. Using the translation properties of the A_x, we can consider $X(\phi_n^{(k)})$ as a Markov process on M instead of on $C_I \times \tilde{C}$ by restricting the extension to M. This is the Markov process we will denote by $Z^{(k)}$ and explore the ergodic properties thereof.

Remark 5.2.5. In the case of oblique reflection we would have $M = p \times \tilde{C} \cong \tilde{C}$ and for any number of ergodic hypotheses on the periodicity l_i of the reflection terms in relation to the $\tilde{e}_i$, we would have $M = C_I \times \tilde{C}$. It might be that the complication introduced by using M instead of simply introducing ergodic hypothesis so that $M = C_I \times \tilde{C}$ is too much, but assuming a sufficiently good ergodic result this would be how to get maximum generality.

We then use this result in a similar fashion to [HM10a], replicating the stages. Fix any $\xi > 0$.

Using the optional stopping theorem, and the corrected process Y^ε

$$\left| \mathbb{E}_x[X_j^\varepsilon(\tau^\delta) - x_j] - \mathbb{E}_x\left[\int_0^{\tau^\delta} \tilde{\gamma}(\varepsilon^{-1}X(s))\, dL^{0,\varepsilon}(s)\right]\right| = O(\varepsilon) \; .$$

By pathwise uniqueness of solution to SDEs such as that satisfied by X^ε, [LS84], for X_x and $X^\varepsilon_{\varepsilon x}$, $L^{0,\varepsilon}(t)$ is equal to $\varepsilon L^0(\varepsilon^{-2}t)$ for $L^{0,\varepsilon}$ the local time of $X^\varepsilon_{\varepsilon x}$ and L^0 the local time of X_x.

We set up the discrete time Markov process on $C_I \times \tilde{C}$, $X(\phi_t^{(k)})$. Then we have, that, for $\phi_j^{(k)}$ the j th successive escape from $\mathscr{I}_k$ after making a return to $\mathscr{I}_0$,

$$\begin{aligned}
&\lim_{\varepsilon\to 0} \mathbb{E}^\varepsilon_{\varepsilon x}\left[\frac{1}{\delta}\int_0^{\tau^\delta} \tilde{\gamma}(\varepsilon^{-1}X^\varepsilon(s))dL^{0,\varepsilon}(s)\right] \\
&= \lim_{n\to\infty} \mathbb{E}_x\left[n^{-1}\int_0^{\tau^{\mathscr{I}_n}} \tilde{\gamma}(X(s))\, dL^0(s)\right] \\
&= \lim_{n\to\infty} \mathbb{E}_x\left[kn^{-1}\sum_j 1_{\{\phi_j^{(k)}<\tau^n\}} k^{-1}\int_{\phi_j^{(k)}}^{\sigma_j^{(k)}} \tilde{\gamma}(X(s))\, dL^0(s)\right] .
\end{aligned}$$

Applying the strong Markov property at times $\phi_j^{(k)}$, this becomes,

$$\lim_{n\to\infty} \mathbb{E}_x\left[kn^{-1}\sum_j 1_{\{\phi_j^{(k)}<\tau^n\}} \mathbb{E}_{X(\phi_j^{(k)})}\left[k^{-1}\int_0^{\tau^{\mathscr{I}_k}} \tilde{\gamma}(X(s))\, dL^0(s)\right]\right] \tag{5.2.7}$$

Denote $\tilde{X}^m$ as a process derived from X on $\mathbb{R}^d_+$. For $m = (m_1, m_2) \in M$ we take $\gamma(\cdot + \iota_1 m_1)$ and $b(\cdot + \iota_2 m_2)$ instead of the usual reflection and drift terms in (5.0.1) and denote the local time of $\tilde{X}^m_1$ as $L^{0,\tilde{X}}$, ι_1 as the inclusion of C_I in $\mathbb{R}^d$ and ι_2 as the inclusion of $\tilde{C}$ in $\mathbb{R}^d$. Basically $\tilde{X}$ is the process obtained when we translate the periodic reflection term within $\{x_1 = 0\}$ and the periodic drift term to make a point specified by the coordinate (m_1, m_2) the new origin and contained in $\{x_1 = 0\}$. More precisely $\tilde{X}^m$ is given by,

$$\tilde{X}^m(t) = \int_0^t b(\tilde{X}^m(s) + \iota_2 m_2)\, ds + W(t) + \int_0^t \gamma\big(\tilde{X}^m(s) + \iota_1 m_1\big)\, dL^{0,\tilde{X}}(s)\ .$$

Denote the escape time of $\tilde{X}^m$ from $\mathscr{I}_k$ as $\tilde{\tau}^{I_k}$. Assume that we have the results of the following two lemmas.

Lemma 5.2.6. *We have that*

$$\begin{aligned} l_{k,i}(m) :&= k^{-1}\mathbb{E}_0\Big[\tilde{X}^m_i(\sigma_0^{(k)}) + \tilde{g}^m_i\Big(\tilde{X}^m(\sigma_0^{(k)})\Big) - \tilde{g}^m_i(0)\Big] \\ &= \mathbb{E}_0\left[k^{-1}\int_0^{\tilde{\tau}^{I_k}} \tilde{\gamma}^m(\tilde{X}^m(s)) + \hat{\gamma}^m(\tilde{X}^m(s))\, dL^{0,\tilde{X}}(s)\right] \end{aligned} \tag{5.2.8}$$

is a continuous bounded function of $m \in M$ and is in fact bounded over k also, for all $i = 2, \ldots, n$. Where $\tilde{g}^m = g(\cdot + \iota_2 m_2)$, $\tilde{\gamma}^m = \gamma(\cdot + \iota_1 m_1)$ and $\hat{\gamma}^m = \bigtriangledown \tilde{g}^m \tilde{\gamma}^m$.

Lemma 5.2.7. *We have uniform weak convergence to an invariant measure $\mu^{(k)}$ on M of the discrete time process $X(\phi_j^{(k)})$ for all k. That is, given a continuous bounded function f on M and a $\xi > 0$, we have that there exists an $N(\xi)$ such that for all $n' > N$*

$$\Big|\mathbb{E}_x\Big[f\big(X(\phi_{n'}^{(k)})\big)\Big] - \mu^{(k)}(f)\Big| < \xi$$

for all $x \in M$.

If we take n, j, dependent on k, sufficiently large, we have the weak convergence

$$1_{\{\phi_j^{(k)} < \tau^n\}} X(\phi_j^{(k)}) \Rightarrow 1_{\{\phi_j^{(k)} < \tau^n\}} \pi^{(k)} \tag{5.2.9}$$

uniformly in the initial point. This follows from Lemma 5.2.7 since we only need a fixed number of steps for weak convergence irrespective of the initial distribution thanks to Lemma 5.2.7 so we have (5.2.9) by considering the distribution of $X(\phi^{(k)}_{j-l})$ conditioned on $\phi^{(k)}_{j-l} < \tau^n$ for l large. We will use this in an identical fashion to [HM10a].

More precisely, first, using the scale function of a Brownian motion of which the first component of X after correction away from the interface and subsequent time change, is, we have $\mathbb{P}_x[\phi^{(k)}_1 < \tau^n] = (1-(k+c_1)/(n+c_2))$. c_i are bounded above by a constant $C > 0$. Thus we choose k so that $C/k < \xi$. Now, for x such that $x_1 = 0$, if $\pi_l \times \pi(x) = m$ then we have, by definition of $\tilde{X}^m$,

$$\begin{aligned}
&\mathbb{E}_x\left[k^{-1}\int_0^{\tau^{\mathscr{I}_k}} \tilde{\gamma}_i(X_s)dL^0_s\right] \\
=&\mathbb{E}_0\left[k^{-1}\int_0^{\tau^{\mathscr{I}_k}(\tilde{X}^m)} \tilde{\gamma}^m_i(\tilde{X}^m_s) + \hat{\gamma}^m_i(\tilde{X}^m_s)dL^{0,\tilde{X}}_s\right] \\
:=&\mathbb{E}_0\left[k^{-1}\int_0^{\tau^{\mathscr{I}_k}(\tilde{X}^m_s)} \tilde{\tilde{\gamma}}_i(\tilde{X}^m_s)dL^{0,\tilde{X}}_s\right].
\end{aligned}$$

Then we take n sufficiently large compared to k so that for $N(\xi/2)$ in Lemma 5.2.7 for f given by the RHS of (5.2.8) $(1-(k+C)/(n-C))^{N(\xi/2)} > 1 - \xi/2\|f\|_\infty$, then we have, using the Strong Markov property at $\phi^{(k)}_j$,

$$\begin{aligned}
&\sum_{j\geq 0}\mathbb{E}_x\left[1_{\{\phi^{(k)}_{j+N(\xi/2)}<\tau^n\}} f\left(X(\phi^{(k)}_{j+N(\xi/2)})\right)\right] \\
=&\sum_{j\geq 0}\mathbb{E}_x\left[1_{\{\phi^{(k)}_j<\tau^n\}}(\mathbb{E}_{\pi^{(k)}}[f] + R(\xi))\right] \\
=&\sum_{j\geq 0}[(1-k/n)^{j+N(\xi/2)}\mathbb{E}_{\pi^{(k)}}[f] \\
&+ (1-k/n)^{j+N(\xi/2)}R'(\xi)] + R''(\xi,k,n)
\end{aligned}$$

for $|R(\xi)| < \xi$, $|R'(\xi)| < 2\xi$, $|R''(\xi,k,n)| < 3C(\|f\|_\infty + 1)n/k^2 < 3\xi(\|f\|_\infty + 1)n/k$ for ξ sufficiently small. Hence we have that (5.2.7) is equal to,

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{E}_x\left[kn^{-1}\sum_j(1-k/n)^j k^{-1}\mathbb{E}_{\pi^k}\left[\int_0^{\tau^{\mathscr{I}_k}} \tilde{\tilde{\gamma}}_i(X_s)dL^{0,\tilde{X}}_s\right]\right] + K(\xi) \\
=&k^{-1}\mathbb{E}_{\pi^k}\left[\int_0^{\tau^{\mathscr{I}_k}} \tilde{\tilde{\gamma}}_i(X_s)dL^{0,\tilde{X}}_s\right] + K(\xi)
\end{aligned}$$

for $|K(\xi)| < 5(\|f\|_\infty + 1)\xi$ for n (dependent on k, ξ), k (fixed, dependent on ξ) sufficiently large, making n larger if necessary so that $(N(\xi/2)\|f\|_\infty k/n) < \xi$ to make the error due to the first $N(\xi/2)$ terms less than ξ, all estimates uniform over the initial point x.

$\xi > 0$ is arbitrary so this completes the proof of convergence to a (as yet unidentified limit) on the assumption that the Lemmas 5.2.6 and 5.2.7 hold. We will now investigate these lemmas, beginning with the proof of Lemma 5.2.6,

Proof of Lemma 5.2.6. A word on notation before we begin, for the process $\tilde{X}^m$, the superscript $m \in M$ gives the position of the interface relative to the period and by definition of $\tilde{X}^m$, we center the process so that the position in the interface/period specified by m becomes the origin. $\tilde{X}^m$ without specification of initial point, in this case $\tilde{X}^m$, is assumed to start at 0 and under the measure $\mathbb{E}_x$ this is the process $\tilde{X}^m$ started at x.

We proceed as in [HM10a]. Fix $i \geq 2$. Set $k_1 = 4\|g\|_\infty + 1$. Then for $x \in \mathscr{I}_k$, $\tilde{\tau}^{\mathscr{I}_k}$ the escape time from $\mathscr{I}_k$ of $\tilde{X}^m$, we will consider,

$$k^{-1}\mathbb{E}_x\left[\tilde{X}_i^m(\tilde{\tau}^{\mathscr{I}_k})\right] , \tag{5.2.10}$$

by considering for $k > k_1 + 2\|g\|_\infty + 1$, for $x \in \mathscr{I}_0$,

$$k_1^{-1}\mathbb{E}_x\left[\tilde{X}_i^m(\tilde{\sigma}_0^{(k_1)})\right] ,$$

using the Strong Markov property at times $\tilde{\sigma}_n^{(k_1)}$ for n such that $\tilde{\sigma}_n^{(k_1)} < \tilde{\tau}^{\mathscr{I}_k}$, where we denote the corresponding series of stopping times for $\tilde{X}^m$ to $\phi_n^{(k)}$ and $\sigma_n^{(k)}$ with a tilde. First a uniform bound on (5.2.8) for all $1 < k < k_1 + 4\|g\|_\infty + 1$. This follows from the Itô formula since we have,

$$\begin{aligned} &k^{-1}\mathbb{E}_0\left[\tilde{X}_i^m(\tilde{\sigma}_0^{(k)})\right] \\ &= k^{-1}\mathbb{E}_0\left[\int_0^{\tilde{\sigma}_0^{(k)}} b_1(\tilde{X}^m(s) + \iota_2 m_2)\, ds\right] + k^{-1}\mathbb{E}_0\left[L^{0,\tilde{X}}(\tilde{\sigma}_0^{(k)})\right] , \end{aligned}$$

from this we obtain,

$$k = \mathbb{E}_0\left[\int_0^{\sigma_0^{(k)}} b_1(\tilde{X}_s^m + \iota_2 m_2)\, ds\right] + \mathbb{E}_0\left[L^{0,\tilde{X}}(\sigma_0^{(k)})\right] .$$

By using the comparison argument from [HM10b] we have that for k bounded, the expectation of the escape time from $\mathscr{I}_k$ for any initial point is uniformly bounded by K_{et} say, which implies that,

$$\mathbb{E}_0\left[L^{0,\tilde{X}}(\sigma_0^{(k)})\right] \leq k + K_{et}\|b_1\|_\infty . \tag{5.2.11}$$

This implies that for this restricted set of k, (5.2.8) is bounded above by,

$$\begin{aligned} &k^{-1}\|b_i\|_\infty K_{et} + k^{-1}\|\gamma_i\|_\infty \mathbb{E}_0\left[L^{0,\tilde{X}}\left(\sigma_0^{(k)}\right)\right] \\ \leq &\|b_i\|_\infty K_{et} + \|\gamma_i\|_\infty (1 + K_{et}\|b_1\|_\infty) \ , \end{aligned}$$

using (5.2.11). Denote the quantity on the RHS by K_{k_1}. Moreover, this also implies bounds in L^1. Now, for any $k > k_1 + 4\|g\|_\infty + 1$ we have,

$$\begin{aligned} k^{-1}\mathbb{E}_x\left[\tilde{X}_i^m\left(\tilde{\tau}^{\mathscr{I}_k}\right)\right] = k^{-1}\mathbb{E}_x\Bigg[&\sum_{\tilde{\phi}_n^{(k_1)} < \tilde{\tau}^{\mathscr{I}_k}} \tilde{X}_i^m\left(\tilde{\sigma}_n^{(k_1)}\right) - \tilde{X}_i^m\left(\tilde{\phi}_n^{(k_1)}\right) \\ &+ \sum_n \tilde{X}_i^m\left(\tilde{\phi}_{n+1}^{(k_1)} \wedge \tilde{\tau}^{\mathscr{I}_k}\right) - \tilde{X}_i^m\left(\tilde{\sigma}_n^{(k_1)} \wedge \tilde{\tau}^{\mathscr{I}_k}\right)\Bigg] \ . \end{aligned} \tag{5.2.12}$$

Applying the Strong Markov property at the times $\tilde{\phi}_n^{(k_1)}$ to the first term in (5.2.12) gives,

$$k^{-1}\mathbb{E}_x\left[\sum_{\tilde{\phi}_n^{(k_1)} < \tau^{\mathscr{I}_k}} \mathbb{E}_{\tilde{X}^m(\tilde{\phi}_n^{(k_1)})}\left[\tilde{X}_i^m\left(\tilde{\sigma}_0^{(k_1)}\right) - \tilde{X}_i^m(0)\right]\right] \ . \tag{5.2.13}$$

Now from above we have that for all x',

$$\left|\mathbb{E}_{x'}\left[\tilde{X}_i^m\left(\tilde{\sigma}_0^{(k_1)}\right) - \tilde{X}^m(0)\right]\right| \leq K_{k_1} \ ,$$

and that $\mathbb{P}[\tilde{\phi}_n^{(k_1)} < \tilde{\tau}^{\mathscr{I}_k}] < [(k - k_1 + 2\|g\|_\infty)/(k - 2\|g\|_\infty)]^n$, using the fact that the first component of the process under the map $x \mapsto x + \tilde{g}^m(x)$ is a time changed Brownian motion in conjunction with the scale function of a one dimensional diffusion as in [HM10a]. This gives a bound of $K_{k_1}(k - 2\|g\|_\infty)/k(k_1 - 4\|g\|_\infty) \leq K_{k_1}/(k_1 - 4\|g\|_\infty)$ on the modulus of (5.2.13).

Similarly, applying the Strong Markov property at times $\tilde{\sigma}_n^{(k_1)}$, the expectation of the second term is bounded above in modulus by,

$$\begin{aligned} &k^{-1}\sum_{n=1}^{\infty}\mathbb{P}\left[\tilde{\phi}_{n-1}^{(k_1)} < \tilde{\tau}^{\mathscr{I}_k}\right] \sup_{\{x:\, x_1 = k_1\}}\left|\mathbb{E}_x\left[\tilde{X}_i^m\left(\tilde{\phi}_0^{(k)} \wedge \tilde{\tau}^{\mathscr{I}_k}\right) - x_i\right]\right| \\ \leq &\frac{2\|g\|_\infty (k - 2\|g\|_\infty)}{k(k_1 - 4\|g\|_\infty)} \\ \leq &\frac{2\|g\|_\infty}{(k_1 - 4\|g\|_\infty)} \ , \end{aligned}$$

using the corrected process as in [HM10a] again.

We will denote the bound produced on (5.2.10) in this manner by $K - 1 - 2\|g\|_\infty$. This implies that (5.2.8) is bounded above by $K - 1$.

Moving onto the continuity of (5.2.8) in $m \in M$, first we will construct from the original (uncorrected) process X, an additional slightly modified process $\hat{X}^{m'}$, $m' \in M$, that we will compare with $\tilde{X}^m$, $m \in M$. The construction of these two processes makes comparison computationally easier. Note that in general, due to the computation we perform, we cannot express $\hat{X}^{m'}$ as $\tilde{X}^m$ with a different starting point $x \in \mathbb{R}^d_+$, as the periodic drift has been shifted relative to the interface once again between $\hat{X}^{m'}$ and $\tilde{X}^m$ as can be seen by the definition of $\hat{X}^{m'}$ below. Firstly, we will use the ideas of [LS84] to show proximity between $\hat{X}^{m'}$ and $\tilde{X}^m$ in the supremum norm up to time T.

First the definition of $\hat{X}^{m'}$. Consider two points $m, m' \in M$. m will form the initial point for $\tilde{X}^m$ as before and m' will be an initial point for $\hat{X}^{m'}$ in a sense we will now define. We define $\hat{X}^{m'}$ as,

$$\begin{aligned}\hat{X}^{m'}_t =& \iota_1(m'_1 - m_1) + \int_0^t b\big(\hat{X}^{m'}(s) + \iota_2 m'_2 - \iota_1(m'_1 - m_1)\big)\, ds + W(t) \\ &+ \int_0^t \gamma\big(\hat{X}^{m'}(s) + \iota_1 m_1\big)\, dL^{0,\hat{X}}(s)\ ,\end{aligned}$$

for $L^{0,\hat{X}}$ the local time of $\hat{X}^{m'}_1$ at 0. Note from the definition of $\tilde{X}^m$, $\hat{X}^{m'}$ we have that $\hat{X}^{m'} = \tilde{X}^{m'} + \iota_1(m'_1 - m_1)$.

Then for fixed $T > 0$, with the help of calculations similar to [LS84, proof of Proposition 4.1], we will analyze $\sup_{0\le t\le T} |\tilde{X}^m_t - \hat{X}^{m'}_t|$. For π_1 the projection onto the first coordinate, set $\phi(x) = \psi(\pi_1(x))\pi_1(x)$ for $\psi : \mathbb{R} \to \mathbb{R}$ a smooth positive function with $\psi(z) = 1$ for $|z| < 1/2$, $\psi(z) = 0$ for $|z| > 1$. Consider the real valued process Φ given by,

$$\begin{aligned}\Phi(t) =& e^{-\lambda(\phi(\tilde{X}^m(t))+\phi(\hat{X}^{m'}(t)))} \\ &\times (\tilde{X}^m(t) - \hat{X}^{m'}(t))^T\big(A(\tilde{X}^m(t)) + A(\hat{X}^{m'}(t))\big)(\tilde{X}^m(t) - \hat{X}^{m'}(t))\ ,\end{aligned} \tag{5.2.14}$$

for λ a constant we will choose later and $A = a_{ij}$, $A \ge \nu \mathrm{I}$, a periodic symmetric C^2 matrix field, the existence of which is assured by [LS84, Lemma 4.1] such that,

$$A(x)\gamma(x + \iota_1 m_1) = e_1\ , \tag{5.2.15}$$

for all $x \in \{\pi_1(x) = 0\}$, and since the domain we are considering is convex, for all $x \in \{x_1 = 0\}$, $x' \in \mathbb{R}^d_+$,

$$\big(A(x)(x - x'), \gamma(x + \iota_1 m_1)\big) \leq 0 \,.$$

In addition, since $\gamma \in C^2$ so is A by [LS84, Lemma 4.1]. There exists $C_0 \geq 0$ such that, for all $x, x' \in \{x_1 = 0\}$,

$$-C_0|x - x'|^2 + \big(A(x')(x - x'), \gamma(x + \iota_1 m_1)\big) \leq 0 \,. \tag{5.2.16}$$

For compactness of the following formulae we will denote $\tilde{X}^m(s) - \hat{X}^{m'}(s)$ as $\vec{X}_s$, $A(\tilde{X}^m(s)) + A(\hat{X}^{m'}(s))$ as A_{X_s} and $\lambda(\phi(\tilde{X}^m(s)) + \phi(\hat{X}^{m'}(s)))$ as $\Lambda_X(s)$. Then we apply the Itô formula to (5.2.14), which gives,

$$\begin{aligned}
\Phi(t) =& \iota_1(m_1 - m_1')^T(A(0) + A(\iota_1(m_1' - m_1))\iota_1(m_1 - m_1') \\
&- \lambda\Big[\int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T A_{X_s}\vec{X}_s E_1(s)\,ds \\
&+ \int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T A_{X_s}\vec{X}_s E_2(s)\,dW(s)\Big] \\
&+ \lambda^2\int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T A_{X_s}\vec{X}_s E_3(s)\,ds \\
&+ \int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T (A_1(\tilde{X}^m(s)) + A_1(\hat{X}^{m'}(s)))\vec{X}_s\,ds \\
&+ \int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T (A_2(\tilde{X}^m(s)) + A_2(\hat{X}^{m'}(s)))\vec{X}_s\,dW_s \\
&+ \int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T A_2(\tilde{X}^m(s))\vec{X}_s\gamma(\tilde{X}^m(s) + \iota_1 m_1)\,dL^{0,\tilde{X}}(s) \\
&+ \int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T A_2(\hat{X}_s^{m'})\vec{X}_s\gamma(\hat{X}^{m'}(s) + \iota_1 m_1)\,dL^{0,\hat{X}}(s) \\
&- \lambda\int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T A_{E,1}(s)\vec{X}_s E_{A,1}(s)\,ds \\
&+ 2\int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T A_{X_s}[b(\tilde{X}^m(s) + \iota_2 m_2) \\
&\qquad\qquad\qquad - b(\hat{X}^{m'}(s) + \iota_2 m_2' - \iota_1(m_1' - m_1))]\,ds \\
&+ 2\int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T A_{X_s}\gamma(\tilde{X}^m(s) + \iota_1 m_1)\,dL^{0,\tilde{X}}(s) \\
&- 2\int_0^t e^{-\Lambda_X(s)}\vec{X}_s^T A_{X_s}\gamma(\hat{X}^{m'}(s) + \iota_1 m_1)\,dL^{0,\hat{X}}(s) \\
&- \lambda\int_0^t \partial_{x_1}\phi(\tilde{X}^m(s))e^{-\Lambda_X(s)}\vec{X}_s^T A_{X_s}\vec{X}_s\gamma_1(\tilde{X}^m(s) + \iota_1 m_1)\,dL^{0,\tilde{X}}(s) \\
&- \lambda\int_0^t \partial_{x_1}\phi(\hat{X}^{m'}(s))e^{-\Lambda_X(s)}\vec{X}_s^T A_{X_s}\vec{X}_s\gamma_1(\hat{X}^{m'}(s) + \iota_1 m_1)\,dL^{0,\hat{X}}(s)\,,
\end{aligned}$$

where E_i denotes the terms derived solely from differentiation with respect to the drift and Brownian motion terms of the exponential term, A_i denotes the terms derived solely from the derivatives of the matrix field term and $A_{E,i}$, $E_{A,i}$ denote the cross terms from these two terms. All such terms are bounded. Denote the stopping time when $\sup_{0\le s\le t}|\tilde{X}^m_s - \hat{X}^{m'}_s| = 1$ by τ. Note that the C_i may change from line to line but always denote constants. Hence we have, for $s \le t$, $\mathbb{E}[\Phi(s\wedge\tau)]$ is bounded above by,

$$\begin{aligned}
&|\iota_1(m_1 - m_1')^T(A(0) + A(\iota_1(m_1' - m_1)))\iota_1(m_1 - m_1')| \\
&+ C_1(1+\lambda+\lambda^2)\mathbb{E}\left[\int_0^{s\wedge\tau} e^{-\lambda(\phi(\tilde{X}^m(s'))+\phi(\hat{X}^{m'}(s')))}|\vec{X}_s|^2\,ds'\right] \\
&+ C_2(|m_2 - m_2'| + |m_1 - m_1'|)\mathbb{E}[s\wedge\tau] \\
&+ 2\mathbb{E}\Bigg\{C_3\int_0^{s\wedge\tau} e^{-\lambda(\phi(\tilde{X}^m(s'))+\phi(\hat{X}^{m'}(s')))}|\vec{X}_{s'}|^2\|\gamma\|_\infty(dL^{0,\tilde{X}}(s') + dL^{0,\hat{X}}(s')) \\
&+ \int_0^{s\wedge\tau} e^{-\lambda(\phi(\tilde{X}^m(s'))+\phi(\hat{X}^{m'}(s')))}\vec{X}^T_{s'}A_{X_{s'}}\Big[\gamma(\tilde{X}^m(s') + \iota_1 m_1)\,dL^{0,\tilde{X}}(s') \\
&\qquad\qquad - \gamma(\hat{X}^{m'}(s') + \iota_1 m_1)\,dL^{0,\hat{X}}(s')\Big] \\
&- \lambda\int_0^{s\wedge\tau} e^{-\lambda(\phi(\tilde{X}^m(s'))+\phi(\hat{X}^{m'}(s')))}\vec{X}^T_{s'}A_{X_{s'}}\vec{X}_{s'}\Big[dL^{0,\tilde{X}}(s') + dL^{0,\hat{X}}(s')\Big]\Bigg\},
\end{aligned}$$

noting that we have $\gamma_1 \equiv 1$. The last term is at least as large in modulus as,

$$\lambda\nu\int_0^{s\wedge\tau} e^{-\lambda(\phi(\tilde{X}^m(s'))+\phi(\hat{X}^{m'}(s')))}|\vec{X}_{s'}|^2\Big[dL^{0,\tilde{X}}(s') + dL^{0,\hat{X}}(s')\Big],$$

from the properties of A. Provided we choose $\lambda\nu - 2C_3\|\gamma\|_\infty > 2C_0$, we then have, for $c = e^{-2\lambda\|\phi\|_\infty}\nu > 0$,

$$\begin{aligned}
c\mathbb{E}\left[|\vec{X}_{s\wedge\tau}|^2\right] \le&\, \mathbb{E}\left[\Phi(s)\right] \\
\le&\, \iota_1(m_1 - m_1')^T(A(0) + A(\iota_1(m_1' - m_1))\iota_1(m_1 - m_1') \\
&+ C_1\mathbb{E}\left[\int_0^s e^{-\lambda(\phi(\tilde{X}^m(s'))+\phi(\hat{X}^{m'}(s')))}|\vec{X}_{s'\wedge\tau}|^2\,ds'\right] \\
&+ C_2(|m_2 - m_2'| + |m_1 - m_1'|)s\,.
\end{aligned}$$

We can then conclude using Gronwall's inequality that,

$$\begin{aligned}
&\mathbb{E}\left[|\vec{X}_{s\wedge\tau}|^2\right] \\
\le&\, c^{-1}\iota_1(m_1 - m_1')^T(A(0) + A(\iota_1(m_1' - m_1)))\iota_1(m_1 - m_1') \\
&+ c^{-1}C_2(|m_2 - m_2'| + |m_1 - m_1'|)t
\end{aligned}$$

$$+ c^{-2}C_1 \int_0^t \big[\iota_1(m_1 - m_1')^T(A(0) + A(\iota_1(m_1' - m_1)))\iota_1(m_1 - m_1') + C_2(|m_2 - m_2'| + |m_1 - m_1'|)s\big] e^{c^{-1}C_1(t-s)}\, ds \,,$$

which can be made arbitrarily small by making $|m_1 - m_1'| + |m_2 - m_2'|$ sufficiently small. Denote this bound by $c(m)$.

We then study $\sup_{0\le s\le t} \Phi(s)$ using the bound we have on $\mathbb{E}\big[|\vec{X}_{s\wedge\tau}|^2\big]$ for all $0 \le s \le t$. Hence we have, $\mathbb{E}[\sup_{0\le s\le t\wedge\tau} \Phi(s)]$ is bounded above by,

$$\begin{aligned}
&|\iota_1(m_1 - m_1')^T(A(0) + A(\iota_1(m_1' - m_1)))\iota_1(m_1 - m_1')| \\
&+ C_1(\lambda + \lambda^2)\mathbb{E}\left[\int_0^{t\wedge\tau} e^{-\lambda(\phi(\tilde{X}^m(s)) + \phi(\hat{X}^{m'}(s)))}|\vec{X}_s|^2\, ds\right] + C_4\sqrt{tc(m)} \\
&+ C_2(|m_2 - m_2'| + |m_1 - m_1'|)t \\
&+ 2\mathbb{E}\sup_{0\le s\le t\wedge\tau}\left\{C_3\int_0^s e^{-\lambda(\phi(\tilde{X}^m(s')) + \phi(\hat{X}^{m'}(s')))}|\vec{X}_{s'}|^2\|\gamma\|_\infty(dL^{0,\tilde{X}}(s') + dL^{0,\hat{X}}(s'))\right. \\
&+ \int_0^s e^{-\lambda(\phi(\tilde{X}^m(s')) + \phi(\hat{X}^{m'}(s')))}\vec{X}_{s'}^T A_{X_{s'}}\Big[\gamma(\tilde{X}^m(s') + \iota_1 m_1)\, dL^{0,\tilde{X}}(s') \\
&\qquad\qquad - \gamma(\hat{X}^{m'}(s') + \iota_1 m_1)\, dL^{0,\hat{X}}(s')\Big] \\
&\left.- \lambda\int_0^s e^{-\lambda(\phi(\tilde{X}^m(s')) + \phi(\hat{X}^{m'}(s')))}\vec{X}_{s'}^T A_{X_{s'}}\vec{X}_{s'}\left[dL^{0,\tilde{X}}(s') + dL^{0,\hat{X}}(s')\right]\right\},
\end{aligned}$$

by applying the Burkholder Davis Gundy inequality to the stochastic integral term this time. We then have,

$$\begin{aligned}
c\mathbb{E}\left[\sup_{0\le s\le t}|\vec{X}_{s\wedge\tau}|^2\right] \le& \mathbb{E}\left[\sup_{0\le s\le t}\Phi(s\wedge\tau)\right] \\
\le& \iota_1(m_1 - m_1')^T(A(0) + A(\iota_1(m_1' - m_1))\iota_1(m_1 - m_1') \\
&+ C_1\mathbb{E}\left[\int_0^t e^{-\lambda(\phi(\tilde{X}^m(s)) + \phi(\hat{X}^{m'}(s)))}|\vec{X}_{s\wedge\tau}|^2\, ds\right] \\
&+ C_2(|m_2 - m_2'| + |m_1 - m_1'|)t + C_4\sqrt{tc(m)}\,.
\end{aligned}$$

We can then conclude that,

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\le s\le t}|\vec{X}_{s\wedge\tau}|^2\right] \le& c^{-1}\iota_1(m_1 - m_1')^T(A(0) + A(\iota_1(m_1' - m_1))\iota_1(m_1 - m_1') \\
&+ c^{-1}C_1 tc(m) + c^{-1}C_4\sqrt{tc(m)} \\
&+ c^{-1}C_2(|m_2 - m_2'| + |m_1 - m_1'|)t
\end{aligned}$$

which can be made arbitrarily small by making $|m_1 - m_1'| + |m_2 - m_2'|$ sufficiently small.

To re-iterate what we have shown up to this point is that given T, $\xi > 0$, there exists a $\delta(T,\xi) > 0$ such that up to time T for $|m_1 - m_1'| + |m_2 - m_2'| < \delta(T,\xi)$, $\mathbb{P}[\sup_{0\leq s\leq T} |\vec{X}_s| > \xi] < \xi$. We will now apply this to complete the verification of Lemma 5.2.6 by bounding the probability of being in 'bad' sets where the proximity of these two processes in the supremum norm is not sufficient to give the proximity of the quantity in the expectation in (5.2.8).

By a comparison argument similar to [HM10b] we can obtain a $T(\xi)$ such that, for any $m \in M$,

$$\mathbb{P}\left[\exists s \in [0, T(\xi)] : \tilde{X}_1^m(s) \notin [0, k+1]\right] > 1 - \xi/2K\,.$$

Note that this implies the same inequality for $\hat{X}^{m'}$ for any $m' \in M$ by definition of the two processes. Now what we do is given an index for $\tilde{X}^m$, $m = (m_1, m_2) \in M$, for $\hat{X}^{m'}$ restrict ourselves to indices $m' \in M$ such that $|m_1 - m_1'| + |m_2 - m_2'| < \delta(T(\xi), \eta) \wedge \xi$. Where $0 < \eta < \xi/2K$ is sufficiently small so that given any initial point x such that $x_1 < \eta$, $m' \in M$, we have a probability under $\mathbb{P}_x$, $1 - \xi/2K$, dependent only on the noise, for the cube of side ξ centered on the point $(k - \xi/2, x_2, \ldots, x_d)$ (denote this by C_ξ) that $\tau_{\mathscr{I}_k} = \tau^{C_\xi}$. $\tau_{\mathscr{I}_k}$ is the hitting time of $\mathscr{I}_k$ by $\hat{X}^{m'}$ and τ^{C_ξ} is the escape time of $\hat{X}^{m'}$ from the C_ξ. For instance this follows from choosing η sufficiently small so that there exists a $t' > 0$ (small) so that $\mathbb{P}_x[\{\sup_{0\leq s\leq t'} \pi_1(W(s)) > \eta\} \cap \{\sup_{0\leq s\leq t'} |W(s)| + \|b\|_\infty t' < \xi/2\}] > 1 - \xi/2K$. Choosing η in this fashion the same will hold for $\tilde{X}^m$.

Consider $\sigma_0^{(k),min} = \min\{\sigma_0^{(k)}(\tilde{X}^m), \sigma_0^{(k)}(\hat{X}^{m'})\}$ with the obvious notation that $\sigma_0^{(k)}(\tilde{X}^m)$ denotes $\sigma_0^{(k)}$ for $\tilde{X}^m$ and correspondingly for $\hat{X}^{m'}$. If $\sigma_0^{(k),min} < T$, $\tau_{\vec{X}}(\eta) > T$ for $\tau_{\vec{X}}(\eta) = \inf\{t > 0 : |\vec{X}_t| > \eta\}$, $\sigma_0^{(k)}(\tilde{X}^m) < \sigma_0^{(k)}(\hat{X}^m)$, and

$$\hat{X}^{m'}\left(\sigma_0^{(k)}(\tilde{X}^m) + \tau^{C_\xi}\left(\hat{X}^{m'}(\cdot + \sigma_0^{(k)}(\tilde{X}^m))\right)\right) \neq \hat{X}^{m'}(\sigma_0^{(k)}(\hat{X}^{m'}))\,,$$

denote this set union the corresponding set with $\tilde{X}^m$, $\hat{X}^{m'}$ reversed as $\tilde{E}$. We have $\mathbb{P}[\tilde{E}] < \xi/2K$. In addition, denote,

$$\sigma_0^{(k)}(\tilde{X}^m) + \tau^{C_\xi}\left(\hat{X}^{m'}(\cdot + \sigma_0^{(k)}(\tilde{X}^m))\right)\,,$$

or,

$$\sigma_0^{(k)}(\hat{X}^{m'}) + \tau^{C_\xi}\left(\tilde{X}^m(\cdot + \sigma_0^{(k)}(\hat{X}^{m'}))\right)\,,$$

as relevant by τ^e. Now on $\bar{E} := \tilde{E}^c \cap \{\sigma_0^{(k),min} < T\} \cap \{\tau_{\vec{X}}(\eta) > T\}$, we have,

$$\left|\hat{X}_i^{m'}(\sigma_0^{(k)}(\hat{X}^{m'})) - \tilde{X}_i^m(\sigma_0^{(k)}(\tilde{X}^m))\right| < \xi \; .$$

and $\mathbb{P}_0[\bar{E}] > 1 - \xi/2K - \xi/2K - \xi/2K > 1 - 2\xi/K$.

Hence if we denote $\pi_i + \tilde{g}_i^m$ as $\hat{g}_i^m$, for π_i projection onto the ith component, we have,

$$\begin{aligned}
&k^{-1}\left|\mathbb{E}\left[\hat{g}_i^m(\tilde{X}^m(\sigma_0^{(k)}(\tilde{X}^m))) - \hat{g}_i^m(\hat{X}^{m'}(\sigma_0^{(k)}(\hat{X}^{m'})))\right]\right| \\
\leq &k^{-1}\Big[\Big|\mathbb{E}\Big[1_{\{\sigma_0^{(k),min} \geq T\} \cap \{\tau_{\vec{X}}(\eta) > T\}}\Big(\hat{g}_i^m(\tilde{X}^m(\sigma_0^{(k)}(\tilde{X}^m))) \\
&\qquad\qquad\qquad\qquad\qquad\qquad - \hat{g}_i^m(\hat{X}^{m'}(\sigma_0^{(k)}(\hat{X}^{m'})))\Big)\Big]\Big| \\
&+ \left|\mathbb{E}\left[1_{\{\tau_{\vec{X}}(\eta) < T\}}\left(\hat{g}_i^m(\tilde{X}^m(\sigma_0^{(k)}(\tilde{X}^m))) - \hat{g}_i^m(\hat{X}^{m'}(\sigma_0^{(k)}(\hat{X}^{m'})))\right)\right]\right| \\
&+ \left|\mathbb{E}\left[1_{\tilde{E}}\left(\hat{g}_i^m(\tilde{X}^m(\sigma_0^{(k)}(\tilde{X}^m))) - \hat{g}_i^m(\hat{X}^{m'}(\sigma_0^{(k)}(\hat{X}^{m'})))\right)\right]\right| \\
&+ \mathbb{E}\left[(1 + \|Dg\|_\infty)1_{\bar{E}}\left|\tilde{X}^m(\sigma_0^{(k)}(\tilde{X}^m)) - \hat{X}^{m'}(\sigma_0^{(k)}(\hat{X}^{m'}))\right|\right]\Big] \\
\leq &4\xi + (1 + \|Dg\|_\infty)\xi \; ,
\end{aligned}$$

using the Strong Markov property at the times T, $\tau_{\vec{X}}(\eta)$ and τ^e respectively in the first three terms on the RHS together with the bounds of $K-1$ on $k^{-1}\mathbb{E}_x[\hat{g}_i^m(\tilde{X}_{\tau^{\mathscr{I}_k}(\tilde{X}^m)}) - \hat{g}_i^m(x)]$ derived above.

Now from the definition of $\tilde{X}^m$, $\hat{X}^{m'}$ we have that $\hat{X}^{m'} = \tilde{X}^{m'} + \iota_1(m_1' - m_1)$. Therefore if we have $|m_1 - m_1'| + |m_2 - m_2'| < \delta(T(\xi), \eta) \wedge \xi$ we have that,

$$\begin{aligned}
&k^{-1}\left|\mathbb{E}\left[\hat{g}_i^m(\tilde{X}^m(\sigma_0^{(k)}(\tilde{X}^m)))\right] - \mathbb{E}\left[\hat{g}_i^{m'}(\tilde{X}^{m'}(\sigma_0^{(k)}(\tilde{X}^{m'})))\right]\right| \\
\leq &4\xi + (1 + \|Dg\|_\infty)\xi + \xi + \|Dg\|_\infty\xi \\
\leq &4\xi + (1 + 2\|Dg\|_\infty)\xi \; ,
\end{aligned}$$

which from the definition of $\tilde{X}^m$ proves Lemma 5.2.6. □

In order to prove Lemma 5.2.7 it is suspected that the following result can be used,

Theorem 5.2.8. *Consider a discrete time Markov process Z on a compact state (metric) space M. Assume that we have the Feller property, i.e. that for any continuous f we have $\mathbb{E}_x[f(Z(1))]$ is a continuous function of x. We denote the oscillation of a function f as,*

$$\|f\|_{osc} = \sup_{x \in M} f(x) - \inf_{x \in M} f(x) \; .$$

Then Z admits an invariant measure on M, μ, moreover if and only if condition (C) holds this is the unique invariant measure of Z and we have uniform (over the initial point) weak convergence to the invariant measure, i.e. we have that given any continuous bounded f, $\xi > 0$, there exists $N(\xi) \in \mathbb{N}$ such that, for all $n > N(\xi)$,

$$\sup_{x \in M} \left| \mathbb{E}_x[f(Z(n))] - \mu \right| < \xi \, .$$

Condition (C): there exists $c > 0$ fixed (over δ, x, y) such that given $\delta > 0$, any x, $y \in M$, there exists a finite collection of disjoint subsets of M, $V^{x,y}_{\delta,n}$ such that $M = \cup_n V^{x,y}_{\delta,n}$, $\sup_n \sup_{u,v \in V^{x,y}_{\delta,n}} d(u,v) < \delta$ and there exists $a(\delta) \in \mathbb{N}$ (independent of x, y) so that $\sum_n \min\{\mathbb{P}_x[Z(a(\delta)) \in V^{x,y}_{\delta,n}], \mathbb{P}_y[Z(a(\delta)) \in V^{x,y}_{\delta,n}]\} > c$.

Proof of Theorem 5.2.8. We begin by proving condition (C) implies uniform weak convergence.

Firstly we have existence of an invariant measure from the proof of [LS06, Proposition 3.1]. This result gives the existence of the invariant measure of a continuous time Feller process that has some positive occupation property of a compact set but the proof generalizes without modification to discrete time.

Since we have now established the existence of an invariant measure, the proof of the result will be complete if we can show that there exists fixed $0 < \lambda < 1$, such that given any continuous bounded g, for any $x, y \in M$, there exists $m(g) \in \mathbb{N}$ such that,

$$\left| \mathbb{E}_x[g(Z(m))] - \mathbb{E}_y[g(Z(m))] \right| < \lambda \|g\|_{osc} \tag{5.2.17}$$

Now given any bounded continuous function f, we have, for any $j \in \mathbb{N}$, $j > m$, we have, for $x, y \in M$,

$$\begin{aligned} \left| \mathbb{E}_x[f(Z(j))] - \mathbb{E}_y[f(Z(j))] \right| =& \Big| \mathbb{E}_x[\mathbb{E}_{Z(j-m(f))}[f(Z(m(f)))]] \\ & - \mathbb{E}_y[\mathbb{E}_{Z(j-m(f))}[f(Z(m(f)))]] \Big| \\ <& \lambda \|f\|_{osc} \, , \end{aligned}$$

using the Markov property since we have $\|\mathbb{E}_\cdot[f(Z(m(f)))]\|_{osc} < \lambda \|f\|_{osc}$.

Now we simply iterate this procedure by taking $g = \mathbb{E}_\cdot[f(Z(m(f)))]$ in (5.2.17) followed by,

$$g = \mathbb{E}_\cdot \left[\mathbb{E}_{Z(m(\mathbb{E}_\cdot[f(Z(m(f)))]))} \left[f(Z(m(f))) \right] \right],$$

in (5.2.17) and so on using the Markov property. In other words we use the sequence of functions given by $f_0 = f$ with iterative step $f_n = \mathbb{E}_\cdot[f_{n-1}(Z(m(f_{n-1})))]$ in (5.2.17). Proceeding in this manner we obtain that,

$$\lim_{j\to\infty} \|\mathbb{E}_\cdot[f(Z(j))]\|_{osc} \le \lim_{n\to\infty} \lambda^n \|f\|_{osc} = 0 \,.$$

Thus we have that given $\xi > 0$ there exists $K(\xi) \in \mathbb{N}$ such that for any $x \in M$ that,

$$\left|\mathbb{E}_x[f(Z(j))] - \mu(f)\right| \le \int \left|\mathbb{E}_x[f(Z(j))] - \mathbb{E}_y[f(Z(j))]\right| \mu(dy) < \xi$$

for $j > K(\xi)$. This implies uniform convergence to the invariant measure.

So now it remains to verify (5.2.17). Given the continuous bounded function g choose δ small enough so that for $|u - v| < \delta$, $|g(u) - g(v)| < \|g\|_{osc}/2$. Then, given x, y, if we set $V^x = \cup_n V_n^x$ for $V_n^x \subset \{Z_x(a(\delta)) \in V_{\delta,n}^{x,y}\} \subset \Omega$, such that $\mathbb{P}[V_n^x] = \min\{\mathbb{P}[\{Z_x(a(\delta)) \in V_{\delta,n}^{x,y}\}], \mathbb{P}[\{Z_y(a(\delta)) \in V_{\delta,n}^{x,y}\}]$. Use the corresponding definition for V_n^y and V^y

$$\begin{aligned}
&\left|\mathbb{E}[g(Z_x(a(\delta)))] - \mathbb{E}[g(Z_y(a(\delta)))]\right| \\
=&\left|\mathbb{E}[1_{V^x} g(Z_x(a(\delta))) + 1_{(V^x)^c} g(Z_x(a(\delta)))]\right. \\
&\left. - \mathbb{E}[1_{V^y} g(Z_y(a(\delta))) + 1_{(V^y)^c} g(Z_y(a(\delta)))]\right| \\
\le&\left|\mathbb{E}[1_{V^x} g(Z(a(\delta)))] - \mathbb{E}[1_{V^y} g(Z_y(a(\delta)))]\right| \\
&+ \left|\mathbb{E}[1_{(V^x)^c} g(Z(a(\delta)))] - \mathbb{E}[1_{(V^y)^c} g(Z_y(a(\delta)))]\right| \\
\le&\sum_n \left|\mathbb{E}[1_{V_n^x} g(Z_x(a(\delta)))] - \mathbb{E}[1_{V_n^y} g(Z_y(a(\delta)))]\right| + (1-c)\|g\|_{osc} \\
<&\frac{c\|g\|_{osc}}{2} + (1-c)\|g\|_{osc} \\
=&(1 - c/2)\|g\|_{osc} \,.
\end{aligned}$$

Hence we have (5.2.17) with $\lambda = 1 - c/2$. This proves one direction of the theorem.

Now for the other implication. A word on notation, $cl(A)$ will denote the topological closure of the set A and $B(x,r)$ denotes the open ball of radius r centered at x. Assume we can construct a finite cover of M by closed $V^{k,n}$ such that $\sup_n \sup_{u,v\in V_{k,n}} d(u,v) < 2^{-k}$ and $\mu(\cup_{k,n}\partial V_{k,n}) < 1/2$. In addition, assume we have the existence of a set of continuous functions, indexed by k, $0 \le f_k \le 1$ such that $\mu(f_k) > 1/2$ and $f_k = 0$ on $\cup_n \partial V_{k,n}$. Studying $f_{k,i} = f_k|_{V_{k,i}}$ for $i = 1$ to n, and choosing $a(2^{-k}) = N$ such that $\mathbb{E}_x[f_{k,i}(Z(j))] > \mu(f_{k,i})/2$ for all x, i, $j > N$, the result is clear, we have verified condition (C) for $c = 1/4$. Then the $V_{k,n}$ are constructed as follows, given k we want $\mu(\cup_n \partial V_{k,n}) < 2^{-k-1}$ so we pick any point $x \in M$ and for some $2^{-k-2} < r < 2^{-k-1}$ we have that $\mu(\partial B(x,r)) < 2^{-k-2})$, we denote such a ball by B_1. Set $V_{k,1} = cl(B_1)$. Inductively we pick any $x_l \in M \setminus \cup_{j<l} cl(B_j)$ and then choose B_l, $B_l = B(x_l, r)$ such that $\mu(\partial B(x_l, r)) < 2^{-k-1-l}$ for $2^{-k-2} < r < 2^{-k-1}$, then we set $V_{k,l} = cl(B_l \cap M \setminus (\cup_{j<l} V_{k,j}))$. By compactness M is covered by a finite number of these $V_{k,l}$ hence this inductive process eventually comes to an end. Then we construct the function f_k by considering the series of continuous functions given by $f_k^j(x) = \min\{2^j d(x, \cup_n \partial V_{k,n}), 1\}$ as $j \to \infty$. By the dominated convergence theorem we have the existence of a j such that $\mu(f_k^j) > 1/2$ and we simply take $f_k = f_k^j$ for such a j. □

Remark 5.2.9. There is no attempt to produce a rate for this convergence and the rate produced by any such method is liable to be appalling.

Remark 5.2.10. As a simple example illustrating the application of this theorem consider the following situation. Take $\mathbb{T}^2$ and draw a line of irrational slope on $\mathbb{T}^2$, L, that without loss of generality we will assume passes through 0. L is a dense set in $\mathbb{T}^2$. If we consider a standard Brownian motion $\tilde{B}$ starting from 0 on the line L then we can use the above theorem to show that on $\mathbb{T}^2$, we have uniformly in the initial point $x \in L$,

$$\tilde{B} + x \Rightarrow \mu \ ,$$

for μ Lebesgue measure on the torus, with the obvious definition of vector addition on the torus. This is achieved by extending this process to a process on the torus, given by $\hat{B}_x = \tilde{B} + x$, $x \in \mathbb{T}^2$, and then applying Theorem 5.2.8. So we just have to verify the Feller property, condition (C) and that Lebesgue measure is indeed an invariant measure for the process $\hat{B}$.

First the Feller property, this is clear from the fact that a.s. for all $t \geq 0$, x, $x' \in \mathbb{T}^2$ we have,

$$|\hat{B}_x(t) - \hat{B}_{x'}(t)| = |x - x'| \ .$$

Secondly, we have to verify condition (C). For this, given $\delta > 0$, our δ-net in this case does not need to exhibit dependence on x, y and we can simply choose a uniform $n \times n$ grid for $n > \delta^{-1}$. To show condition (C) we will use the fundamental Weyl equidistribution theorem. Let $\psi : \mathbb{R}^2 \to \mathbb{R}_+$, $0 \le \psi(x) \le 1$ be such that for $(x_1, x_2) \in \mathbb{R}^2$, we have $\psi(x) = 1$ on the set $\{x : 1/4 \le x_1 \le 3/4, 1/4 \le x_2 \le 3/4\}$ and $\psi(x) = 0$ when $x \in \{x_1 \le 0\} \cup \{x_1 \ge 1\} \cup \{x_2 \le 0\} \cup \{x_2 \ge 1\}$. For (i, j), $1 \le i, j \le n$ an $n^{-1} \times n^{-1}$ cell in our δ-net denote the continuous function $\psi(n(x - (n^{-1}i, n^{-1}j)))$ by g_{ij} and the corresponding $n^{-1} \times n^{-1}$ cell on which it is non-zero as c_{ij}. We will apply the fundamental Weyl equidistribution theorem [Tao09] to g_{ij} using as the map T, translation by $\alpha\hat{l}$ for any fixed $\alpha \in \mathbb{R}$, where we denote the unit vector in the direction of L as $\hat{l}$. The fact that the two components of $\alpha\hat{l}$, (α_1, α_2) are independent over the rational numbers implies that there is no non-zero character (continuous homomorphism $\mathbb{T}^2 \to \mathbb{T}$, which therefore is of the form $(x_1, x_2) \mapsto n_1x_1 + n_2x_2$) of the 2 dimensional torus χ such that $\chi(\alpha_1, \alpha_2) = 0$. Therefore by the fundamental Weyl equidistribution theorem we have that for any $x \in \mathbb{T}^2$, $x + n\alpha\hat{l}$ is equidistributed on $\mathbb{T}^2$. Hence for all $x \in \mathbb{T}^2$ we have an $N_{ij}(x)$, such that for all $n > N_{ij}(x)$,

$$\frac{1}{n}\sum_{k=1}^{n} g_{ij}(T^n(x)) > \frac{1}{2}\mu(g_{ij}) \ . \tag{5.2.18}$$

Owing to the continuity of g_{ij} we have that,

$$\frac{1}{n}\sum_{k=1}^{n} g_{ij}(T^n(x)) \ .$$

is a family of functions over which we have uniform continuity in x. Therefore by compactness we have a single N_{ij}, such that for any $n > N_{ij}$ and any x we have (5.2.18). We use this to show that given any point $x \in \mathbb{T}^2$, if we take $\tilde{\mu}^x_n$ as normalized Lebesgue measure on the line segments $[x, x + n\alpha\hat{l}]$, then, there exists an N'_{ij} such that for any x we have, $n > N'_{ij}$,

$$\tilde{\mu}^x_n(g_{ij}) > \frac{1}{2}\mu(g_{ij}) \ . \tag{5.2.19}$$

This is clear by taking the normalized Lebesgue measure on the line segment of $[x, x + \alpha\hat{l})$ of both sides of (5.2.18) for $n > N_{ij}$ hence we set $N'_{ij} = N_{ij}$.

The standard Brownian motion has a probability density that satisfies the condition $p_t(0,y) > c_1 t^{-1/2}$ for $y \in [-c_2 t^{1/2}, c_2 t^{1/2}]$ for some fixed constants c_1, c_2 arising from the standard normal distribution. For $n \in \mathbb{N}$, we are going to consider the set of times $\{(n\alpha/c_2)^2\}$ and the density $c_1 c_2 n^{-1} \alpha^{-1}$ on the line segment from x to $x + n\alpha\hat{\imath}$. For each pair (i,j) we have an N_{ij} that gives a corresponding result to (5.2.19), so taking $N = \max_{(i,j)} N_{ij} + 1$, we have that at time $(N\alpha/c_2)^2$, for any x, $y \in \mathbb{T}^2$,

$$\sum_{ij} \min_{z=x,y} \mathbb{P}_z[\hat{B}_z \in c_{ij}] > \sum_{ij} \min_{z=x,y} \mathbb{E}_z[g_{ij}(\hat{B}_z)] > \frac{c_1 c_2}{2} \sum_{ij} \mu(g_{ij}) > \frac{c_1 c_2}{8}$$

which is precisely condition (C).

Lastly, the fact that normalized Lebesgue measure is the unique invariant probability measure follows from the fact that any unique invariant measure must be translation invariant since,

$$\hat{B}_{x+y}(t) = \hat{B}_x(t) + y ,$$

and that Lebesgue measure is the unique translation invariant measure on the torus.

Remark 5.2.11. Note that in the previous example since the process is translation invariant we have equicontinuity of the semigroup in time and hence can use a slight modification of [KPS, Theorem 1] to give uniform convergence to the invariant measure immediately.

Remark 5.2.12. This example provides the clue as to how Theorem 5.2.8 may be used to complete the proof. Under appropriate ergodic assumptions (maybe the framework is too general) if we have a lower bound on the embedded Markov process $X(\phi_t^{(k)})$ as it spreads out along the interface as above then we would be able to deploy Theorem 5.2.8 as above to complete the proof of the homogenization result. It is unknown to the author at this point how to proceed toward obtaining such a result, although it seems likely to the author that such a result would hold.

5.3 Well-Posedness of the martingale problem and characterization of the limiting process

By application of Theorem 5.2.1, contingent on the completion of Proposition 5.2.3 we would have the existence of a limiting process that satisfies the martingale problem given by the differential operator,

$$\bar{\mathcal{L}} = D_{ij}\partial_i\partial_j,$$

where D_{ij} is a $d \times d$ matrix given by,

$$D_{ij} = \int_{\mathbb{T}^d} (\delta_{ik} + \partial_k g_i)(\delta_{kj} + \partial_k g_j)\, d\mu \ ,$$

on the space of functions that are bounded, have bounded derivatives of all orders and in addition satisfy (5.2.4). Proceeding almost identically as in [HM10a], we construct a solution to the martingale problem corresponding to $\bar{\mathcal{L}}$ to show that the domain of definition of the martingale problem is sufficiently large and then apply the result by Ethier and Kurtz [EK86, Theorem 4.1]. The solution in this case is given as follows: let M be a matrix satisfying $MM^T = D$ and such that

$$M = \begin{pmatrix} \sqrt{D_{11}} & 0 \\ v & \tilde{M} \end{pmatrix} \ ,$$

for some vector $v \in \mathbb{R}^{d-1}$ and a $(d-1) \times (d-1)$ matrix $\tilde{M}$. (This is always possible by the QR decomposition.) We then first construct a Wiener process W_1 and a process $\bar{X}_1$ such that

$$d\bar{X}_1 = \sqrt{D_{11}^+}\, dW(t) + dL(t) \ ,$$

where L is the symmetric local time of $\bar{X}_1$ at the origin. This can be achieved for example by setting $\bar{X}_1 = h(Z)$, where

$$h(x) = \sqrt{D_{11}^+} x$$

Z is a reflected Brownian motion (also known as a skew-Brownian motion with parameter 1), and W is the martingale part of Z. Given such a pair $(\bar{X}_1, W)$, we then let $\tilde{W}$ be an independent $d-1$-dimensional Wiener process and we define pathwise the $\mathbb{R}_+^{d-1}$-valued process $\tilde{X}$ by

$$\tilde{X}(t) = \int_0^t \tilde{M}\, d\tilde{W}(t) + \int_0^t v\, dW(t) + \tilde{\alpha} \int_0^t dL(t) \ ,$$

where $\tilde{\alpha}_j = \alpha_{j+1}$. Thus a solution (which can be shown to be the only solution) to the martingale problem associated to $\bar{\mathcal{L}}$ has been characterized. The argument that gives uniqueness of solution to the martingale problem corresponding to $\bar{\mathcal{L}}$ then proceeds word for word as in [HM10a] and will not be repeated here.

Chapter 6

Convergence of Marginal Distributions for coupled oscillators

Consider the following system,

$$\begin{aligned} dq_0 &= p_0\,dt, \\ dp_0 &= \left[-q_0|q_0|^{2k-2} - \alpha(q_0 - q_1) - \gamma p_0\right] dt + \sqrt{2\gamma T_0}\,dW_0, \\ dq_1 &= p_1\,dt, \\ dp_1 &= \left[-q_1|q_1|^{2k-2} - \alpha(q_1 - q_0)\right] dt + \sqrt{2\gamma T_1}\,dW_1, \end{aligned} \tag{6.0.1}$$

where W_0, W_1 are independent standard Brownian motions. Where necessary for compactness of notation, we will refer to the process $(p_0(t), q_0(t), p_1(t), q_1(t))$ as $U(t) \in \mathbb{R}^4$ and the generator of U will be denoted by $\mathcal{L}$. We will analyze this system in the context of the results of [Hai09]; by a refinement of the methods introduced in [HM09], it was shown in [Hai09] that such a system admits an invariant measure if either $k < 2$, or, if $k = 2$ but the coupling constant is sufficiently large, $\alpha^2 > T_1\langle\Phi^2\rangle^{-1}$. Φ is a corrector, away from the origin it is the solution to the equation $\mathcal{L}_H\Phi = -q$, which is used to approximately decouple the two oscillators. We denote $\mathcal{L}_H$ for the generator of a free oscillator, $-q|q|^{2k-2}\partial_p + p\partial_q$. The aim of this article is to study the regime where no invariant measure exists. In this case, the 'cold' oscillator denoted by the 0 variables and the 'hot' oscillator denoted by the 1 variables asymptotically decouple as $t \to \infty$ due to the increasing energy of the hot oscillator like in [Hai09].

For $k \geq 2$, we will show that the cold oscillator still admits a limiting distribution and for $k > 2$ we have weak convergence to this distribution. Furthermore we can express it as the invariant measure, $\bar{\mu}$, of a related process given by the bar variables below,

$$\begin{aligned} d\bar{q}_0 &= \bar{p}_0 \, dt, \\ d\bar{p}_0 &= \left[-\bar{q}_0 |\bar{q}_0|^{2k-2} - \alpha \bar{q}_0 - \gamma \bar{p}_0 \right] dt + \sqrt{2\gamma T_0} \, d\bar{W}_0, \end{aligned} \tag{6.0.2}$$

where $\bar{W}_0$ is a standard Brownian motion. The existence of $\bar{\mu}$ follows from the Lyapunov functional method in [HM09] for instance. The weak convergence of the cold oscillator is the content of Theorem 6.1.1, the proof of which is the aim of section 6.1. In addition to weak convergence, we also have convergence in an ergodic sense and by this we mean that, for f continuous and bounded,

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(p_0(s), q_0(s)) \, ds = \int_{\mathbb{R}^2} f(u) \, d\bar{\mu}(u), \quad \text{a.s.} \, .$$

This is the content of Theorem 6.2.1, and section 6.2 is devoted to the proof of this statement.

For $k = 2$, $\alpha^2 < T_1 \langle \Phi^2 \rangle^{-1}$, the result is slightly different since the corrector Φ no longer vanishes with increasing energy. Weak convergence of the 0 variables in this case is expected but any such result would necessitate establishing whether the periods of the 1 variables become sufficiently mixed over time in order to eliminate the influence of Φ in a weak sense. We do however have the ergodic convergence, which is the content of Theorem 6.3.5 and section 6.3 contains the proof of this theorem.

We also show in section 6.4 that the energy of the hot oscillator H_1 grows linearly in time and that, after a suitable rescaling, its distribution for large times approaches that of the square of a Bessel process multiplied by a prefactor (or we could consider it as a time changed squared Bessel process), the parameter of which can be computed explicitly. This is the content of Theorem 6.4.1. For $k > 2$, consider,

$$X(t) = \int_0^t \gamma T_1 \, ds + \int_0^t (2\gamma T_1 \langle p^2 \rangle)^{\frac{1}{2}} \sqrt{X}(s) \, dW(s), \tag{6.0.3}$$

and for $k = 2$,

$$Y(t) = \int_0^t (\gamma T_1 - \gamma\alpha^2 \langle \Phi^2 \rangle)\, ds + \int_0^t (2\gamma T_1 \langle p^2 \rangle)^{\frac{1}{2}} \sqrt{Y}(s)\, dW(s), \qquad (6.0.4)$$

Then for $k > 2$ we have the weak convergence, for $H_1 = p_1^2/2 + q_1^{2k}/2k$,

$$\frac{H_1(T)}{T} \Rightarrow X(1),$$

and for $k = 2$, $T_1 > \alpha^2 \langle \Phi^2 \rangle$,

$$\frac{H_1(T)}{T} \Rightarrow Y(1).$$

The dimension of the squared Bessel process in the case $k > 2$ is then $2/\langle p^2 \rangle$ and the prefactor is $\gamma T_1 \langle p^2 \rangle / 2$. In the case $k = 2$, $T_1 > \alpha^2 \langle \Phi^2 \rangle$, the dimension of the squared Bessel process is then $2(\gamma T_1 - \alpha^2 \langle \Phi^2 \rangle)/\gamma T_1 \langle p^2 \rangle$ with the same prefactor as for the case $k > 2$.

6.1 $k > 2$: weak convergence of the 0 variables to $\bar{\mu}$

6.1.1 General Strategy

For (6.0.1), in the case $k > 2$, we will show for any initial point $(p_0(0), q_0(0), p_1(0), q_1(0))$, the zero variables $(p_0(t), q_0(t))$, exhibit weak convergence to the invariant regime for the system given above in (6.0.2). We have a different result for $k = 2$.

The general strategy is to show that given $\epsilon > 0$, and a bounded continuous function f of (p_0, q_0), the process at a large time $T(\epsilon)$ has a high probability, $1 - \epsilon$, of being at a sufficiently large energy in the 1 variables so the equations in the 1 and 0 variables will "decouple". When the equations decouple we will have convergence to within ϵ of the integral of f against the distribution of p_0, q_0, at the end of a fixed smaller time interval $t(\epsilon)$ providing $H_1 = \frac{p_1^2}{2} + \frac{q_1^{2k}}{2k}$ does not re-enter a compact set (dependent on ϵ) and we have a sufficiently good initial distribution at the beginning of the small fixed time interval in the 0 variables. We wait until the initial (before starting the fixed time interval on which we expect convergence of the 0 variables) energy in the 1 variables is large enough so that the probability of the process re-entering the compact set where we do not have sufficient decoupling within $t(\epsilon)$ is less than ϵ.

Upon completion of the details of this strategy, we then have,

Theorem 6.1.1. *For $k > 2$, we have the convergence, for bounded continuous f,*

$$\mathbb{E}_x[f(p_0(t), q_0(t))] \to \bar{\mu}(f),$$

as $t \to \infty$, for any initial point $x = (p_0(0), q_0(0), p_1(0), q_1(0))$. $\bar{\mu}$ is the invariant measure of the bar variables $\bar{p}_0$, $\bar{q}_0$ above. In other words we have weak convergence of the distribution of the cold oscillator, $(p_0(t), q_0(t))$, to $\bar{\mu}$.

The strategy of proof for Theorem 6.1.1 is the following. First, we show tightness of the 0 variables, then the convergence of the 0 variables to the regime given by $\bar{\mu}$ in the high energy regime for the 1 variables, then we show the existence of advantageous lower bounds on H_1, the energy of the 1 variables.

6.1.2 A priori bounds on the cold oscillator

In this subsection we will consider $k \geq 2$ not just $k > 2$. We copy the methodology of [Hai09] here. We introduce the tilde variables, $\tilde{p}_0$, $\tilde{q}_0$. $\tilde{q}_0$ is the integral with respect to time of $\tilde{p}_0 = p_0 + \alpha\Phi(q_1, p_1)$, for $\Phi(q_1, p_1)$ the unique centered solution (as defined in [HM09]) to $\mathcal{L}_H\Phi = -q + \mathcal{R}(p, q)$. $\mathcal{L}_H$ is the generator of the dynamics corresponding to the Hamiltonian $\frac{p^2}{2} + \frac{q^{2k}}{2k}$. $\mathcal{R}$ is a function that averages out to 0 over the level sets of H, is equal to 0 outside a ball containing the origin, and has all of its derivatives approaching those of q at the origin. The function $\mathcal{R}$ is required, since without it, lack of smoothness at the origin of the coefficients of $\mathcal{L}_H$ is a barrier to the smoothness of Φ [Hai09]. The corrected variables $\tilde{p}_0$, $\tilde{q}_0$ then satisfy

$$\begin{aligned} d\tilde{q}_0 &= \tilde{p}_0\, dt \\ d\tilde{p}_0 &= \left[-q_0|q_0|^{2k-2} - \alpha q_0 - \gamma p_0\right] dt + \sqrt{2\gamma T_0}\, dW_0 + R_1\, dt + R_2 dW_1, \end{aligned} \tag{6.1.1}$$

Where R_1, R_2 given by

$$R_1 = \alpha^2(q_0 - q_1)\partial_p\Phi + \alpha\gamma T_1\partial_p^2\Phi$$

$$R_2 = \alpha\sqrt{2\gamma T_1}\partial_p\Phi,$$

and are of order $H_1^{\frac{3}{2k}-1}(H_0^{\frac{1}{2k}}+1)$. Note that p_0 differs from $\tilde{p}_0$ by a term of order $H_1^{\frac{1}{k}-\frac{1}{2}}$ from [HM09].

Remark 6.1.2. For $k > 2$, it might seem at this point that we lose control of how $\tilde{q}_0$ relates to q_0, but in reality we will only be using this approximation over finite time intervals of prescribed length. We will then use lower bounds on the 1 variables to ensure that $\tilde{q}_0$ and q_0 are close over this finite time interval when they are considered as having corresponding initial values at the start of this finite time interval.

Then consider the test function, $E_0(\tilde{p}_0, q_0) = \frac{\tilde{p}_0^2}{2} + \frac{|q_0|^{2k}}{2k} + \alpha\frac{q_0^2}{2} + \theta\tilde{p}_0 q_0 + C$ for $0 < \theta < (\gamma/2) \wedge 1$ like in [Hai09] and $C > 0$ large enough to make E_0 positive.

$$\begin{aligned}\mathcal{L}E_0 = &-\gamma\tilde{p}_0^2 + \alpha\gamma\Phi\tilde{p}_0 + \gamma T_0 - \alpha\Phi q_0|q_0|^{2k-2} - \alpha\theta\Phi\tilde{p}_0 + \theta\tilde{p}_0^2 - \alpha^2\Phi q_0 \\ &- \theta q_0^{2k} - \theta\alpha q_0^2 - \theta\gamma\tilde{p}_0 q_0 + \alpha\theta\gamma\Phi q_0 + R_1\tilde{p}_0 + R_1\theta q_0 + \frac{1}{2}(R_2)^2.\end{aligned} \tag{6.1.2}$$

This is negative outside of a compact set in the 0 variables, since the coefficients of all leading order terms are negative. The leading order terms in (6.1.2) are of order E_0 whereas the stochastic terms of E_0 are of order $E_0^{\frac{1}{2}}$,

$$(\tilde{p}_0 + \theta q_0)(\sqrt{2\gamma T_0}\,dW_0 + R_2 dW_1).$$

Lemma 6.1.3. *Consider $k \geq 2$. For any initial condition $x = (p_0(0), q_0(0), p_1(0), q_1(0))$, given $\epsilon > 0$, there exists a time $T(\epsilon, x) > 0$ and a compact set $K(\epsilon) \subset \mathbb{R}^2$ independent of the initial condition, such that at all times $T > T(\epsilon, x)$ we have,*

$$\mathbb{P}_x\big[(p_0(T), q_0(T)) \in K(\epsilon)\big] > 1 - \epsilon.$$

Proof. We already have a Lyapunov function for the 0 variables, E_0. Hence by Chebychev we have the required result. □

Lemma 6.1.4. *Consider $k \geq 2$. Given $\epsilon > 0$ and a polynomial $p_0^{k_1} q_0^{k_2}$ for k_1, $k_2 \in \mathbb{Z}$, there exists a compact set $K_{p,q} \subset \mathbb{R}^2$ (dependent on the indices k_1, k_2) such that,*

$$\frac{1}{T}\int_0^T 1_{K_{p,q}^c}\big((p_0(s), q_0(s))\big)\, p_0^{k_1}(s) q_0^{k_2}(s)\, ds < \epsilon, \tag{6.1.3}$$

for T sufficiently large almost surely.

Proof. Basically what we are going to do is time change E_0 and then compare to a suitable one dimensional SDE where we will have this result by virtue of Birkhoff's ergodic theorem, then reverse the time change and this will give the required result. Initially, choose a large compact set $K' \subset \mathbb{R}$, such that outside of K', we have that the drift is less than $-2k(2\gamma T_0 + \|R_2\|_\infty^2)E_0$, for $k > 0$ a constant sufficiently small. Then the process we will make the comparison with is $|F|$ given by,

$$dF = -k\text{sgn}(F) + \mathcal{R}'\,dt + dW,$$

for W a standard Brownian motion, and $\mathcal{R}'$, $|\mathcal{R}'| \leq k$, a smoothing term that is zero outside a small neighborhood of the origin (smaller than K') such that $|-k\text{sgn}(F) + \mathcal{R}'| \leq k$ and $-k\text{sgn}(F) + \mathcal{R}'$ is smooth everywhere including the origin, which gives the existence of strong solutions for the SDE satisfied by F.

The time change we will use on E_0 is the obvious one,

$$T_t = \inf\left\{ s > 0 : \int_0^s (\tilde{p}_0(s') + \theta q_0(s'))^2 (2\gamma T_0 + (R_2(s'))^2)\,ds' > t \right\},$$

and by application of the time change we end up with a process with diffusion coefficient 1 and drift outside of K' of less than $-k$. Note that this transformation of the SDE is possible since the time change is continuous by the occupation times formula applied to $(\tilde{p}_0 + \theta q_0)^2$.

With an appropriate choice of Brownian motion for F as in [HM10b] and a matching choice of initial conditions for F relative to E_0, we have that $E_0(T_t) \leq |F(t)| + 2|K'|$, for $|K'|$ the largest element in K'.

By the continuous time Birkhoff ergodic theorem (for instance see [KS07]) applied to F, we have that, for compact sets K'' sufficiently large,

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T 1_{\mathbb{R}\setminus K''}(F(s))(|F(s)| + 2|K'|)^{\frac{k_1}{2}+\frac{k_2}{2k}}\,ds \leq (|K''|)^{-1}\epsilon \quad \text{a.s.}, \tag{6.1.4}$$

since by the existence of a Lyapunov function for F, we have the existence of an invariant measure for this process, and in addition from [Hai09, Theorem 3.4], we have the existence of a unique (hence ergodic) invariant measure. Therefore the application of the Birkhoff ergodic theorem is justified. Then we simply choose the set K'' such that the invariant measure of the complement $\mathbb{R} \setminus K''$ is less than $(|K''|)^{-1}\epsilon$, where the modulus denotes the size of the maximal element. This is possible since the invariant measure of the complement decays exponentially in the size of K'' due to F having $e^{k'|x|}$ as a Lyapunov functional for $0 < k' < k$.

To transfer this statement from one about F to one about E_0 there are two issues with the time change T_t that we must deal with when choosing our compact set $K_{p,q}$. The first is that we do not end up with the situation $T_t = \infty$ for some $t < \infty$ since then we may not explore enough of the path of F to have the convergence of the proportions assured along the paths of F. The second is that we have an upper bound on the inverse time change for the expansion of times outside of K'' relative to those inside.

The first point. This is not a problem since the result is immediate on this portion of the probability space when $K_{p,q} = \{(x,y) \in \mathbb{R}^2 : E_0(x,y) \in |K''| + 2K'\}$ is made larger if necessary so that it is larger than a compact set outside of which we have $(\tilde{p}_0(s') + \theta q_0(s'))^2(2\gamma T_0 + (R_2(s'))^2) > c > 0$ for some fixed c. This is easily possible if we choose $\mathcal{R}$ so that $\|R_2\|_\infty < 2\gamma T_0$, for instance $\mathcal{R}(p,q) = \zeta(H(p,q))q$ and $\zeta : \mathbb{R} \to \mathbb{R}$ is 1 on a sufficiently large neighborhood of 0. Then we expand $K_{p,q}$ if necessary so that $\{(x,y) : x^2 + |y| > (\|\Phi\|_\infty + 1)^2\} \subset K_{p,q}$. Explicit calculation then gives that (6.1.3) is bounded above by $(c|K''|)^{-1}\epsilon$ in this case.

The second point. Note that for E_0 outside of a sufficiently large compact set, $(\tilde{p}_0(s') + \theta q_0(s'))^2(2\gamma T_0 + (R_2(s'))^2)$ varies between $c_1 E_0^{\frac{1}{k}}$ and $c_2 E_0$, for $c_1, c_2 > 0$ constants, dependent on the precise definition of Φ. Hence we expand K'' if necessary to include this set and be such that $\sup\{(x + \Phi + \theta y)^2(2\gamma T_0 + \|R_2\|_\infty^2) : (x,y) \in \mathbb{R}^2,\ E_0(x,y) \in K''\} \le c_2|K''|$. Assume F spends less than $1/2$ of the time outside K'' a.s. as time tends to infinity which is clearly possible by choosing K'' sufficiently large from (6.1.4). Then we choose $K_{p,q} = \{(x,y) \in \mathbb{R}^2 : E_0(x,y) \in |K''| + 2K'\}$ and thus have almost surely, for t sufficiently large,

$$
\begin{aligned}
&\frac{1}{T_t}\int_0^{T_t} 1_{K_{p,q}^c}\big((p_0(s), q_0(s))\big)\, p_0^{k_1}(s) q_0^{k_2}(s)\, ds \\
\le &\frac{1}{T_t}\int_0^{t} 1_{K_{p,q}^c}\big((p_0(T_s), q_0(T_s))\big)\big(|F(s)| + 2|K'|\big)^{k_1/2 + k_2/2k}\, dT_s
\end{aligned}
$$

$$\leq \frac{2c_2(|K''|+2K')}{c_1(|K''|+2K')^{\frac{1}{k}}}(|K''|)^{-1}\epsilon$$

since then we have $T_t \geq t/2c_2(|K''|+2K')$ by explicit calculation using the time spent by F in K''. This proves the result. □

6.1.3 Stage 2: Convergence of the cold oscillator in distribution

When the hot oscillator is at high energy, we will show convergence of the 0 variables by showing proximity of the 0 variables and bar variables (using the tilde variables as an intermediary) over a finite time $t'(\epsilon)$ for given $\epsilon > 0$, then using the weak convergence of the bar variables to their invariant measure in that time. More precisely if we desire weak convergence of the 0 variables at the large time T we will study the 0 variables in the time period $[T - t'(\epsilon), T]$ and show weak convergence to $\bar{\mu}$.

What we do at this point is fix our $t'(\epsilon)$ so that for the bar variables, with bounded initial conditions, the integral of f with respect to the distribution of the variables $\bar{p}_0, \bar{q}_0$ is within ϵ of the integral against the invariant measure. We have exponential convergence of the bar variables to their invariant distribution with prefactor given by E_0 at the beginning of the time interval from the existence of the Lyapunov function $E_0(\bar{p}_0, \bar{q}_0)$ giving a spectral gap [Hai09, Theorem 3.4]. If we assume an exponential rate of convergence given by $\zeta > 0$, then denoting E_0 at the beginning of the time interval by $E_0(T - t'(\epsilon))$,

$$t'(\epsilon) = \zeta^{-1}(\log E_0(T - t'(\epsilon)) - \log \epsilon) \ .$$

Moving back to the 0 variables of which we seek convergence, since f is bounded we can ignore what happens on the rest of the probability space where we do not have good bounds on the 0 variables at $T - t'(\epsilon)$.

The difference $p_0 - \tilde{p}_0$ is quite clear and now we will study the error term $\tilde{p}_0 - \bar{p}_0$. Studying (6.1.1) we can see that the difference between these variables and the variables given in (6.0.2), when $W_0 = \bar{W}_0$, is given by the remainder terms R_1, R_2 plus the effect of the difference between $\bar{p}_0$, $\tilde{p}_0$ and $\bar{q}_0$, $\tilde{q}_0$.

Lemma 6.1.5. *For every compact set $K \subset \mathbb{R}^2$, every $\epsilon > 0$ and every $T > 0$, there exists a compact set $K' \subset \mathbb{R}^2$ so that if $(p_0(0), q_0(0)) \in K$ and $(p_1(0), q_1(0)) \notin K'$, we have*

$$\mathbb{P}\Big(\sup_{0 \le t \le T} \big(|\tilde{p}_0(t) - \bar{p}_0(t)| + |\tilde{q}_0(t) - \bar{q}_0(t)| \big) \ge \epsilon \Big) \le \epsilon \, ,$$

where $(\bar{p}_0, \bar{q}_0)$ solves (6.0.2) with respect to the same Brownian motion as that driving the 0 variables, W_0, and $\bar{p}_0(0) = \tilde{p}_0(0)$, $\bar{q}_0(0) = \tilde{q}_0(0)$.

Proof. Denote $e_0(t) = \sup_{0 \le s \le t} |\tilde{p}_0(s) - \bar{p}_0(s)|$. What we want is that within the time interval of length T, the processes (p_0, q_0), $(\tilde{p}_0, \tilde{q}_0)$ and $(\bar{p}_0, \bar{q}_0)$ remain with high probability $1 - \epsilon/3$, within an energy of $l(K, \epsilon, T)$. This is certainly the case when l is sufficiently large, at least as large as $(2k-1)^k (2k)^{2k-2}$, from a simple calculation using $E_0(\tilde{p}_0, q_0)$ for (p_0, q_0), $(\tilde{p}_0, \tilde{q}_0)$, and $E_0(\bar{p}_0, \bar{q}_0)$ for $(\bar{p}_0, \bar{q}_0)$, looking purely at the Brownian motion part (neglecting the negative drift). We can conduct a similar analysis using the bounds on the 0 variables with H_1 which changes on a scale of $H_1^{1/2}$ over finite times, hence by choosing K' sufficiently large (dependent on T) we can assume that the 1 variables are outside of $K'/2$ over the interval of time of length T. Therefore e_0 satisfies,

$$\begin{aligned} e_0 \le & (l + \alpha) \int_0^t \left(\int_0^{t'} e_0 \, ds \right) dt' + \int_0^t \gamma e_0 \, dt' \\ & + \sup_{s \le t} \left| \int_0^t R_2 \, dW_1(t') \right| + \int_0^t |R_1| \, dt' \\ & + (l + \alpha) \int_0^t \left(C_1 \Big[\sup_{(p_1, q_1) \notin \frac{K'}{2}} \left(\frac{p_1^2}{2} + \frac{q_1^{2k}}{2k} \right)^{\frac{1}{k} - \frac{1}{2}} \Big] t' \right) dt' \\ & + \int_0^t C_1 \gamma \sup_{(p_1, q_1) \notin \frac{K'}{2}} \left(\frac{p_1^2}{2} + \frac{q_1^{2k}}{2k} \right)^{\frac{1}{k} - \frac{1}{2}} dt' \, , \end{aligned}$$

for some constant C_1 bounding Φ over one period at energy 1. Simplifying, this is less than,

$$\begin{aligned} & \int_0^t (l + \alpha) t' e_0 \, dt' + \int_0^t \gamma e_0 \, dt' + 3\epsilon^{-1} C \big(\| R_2 |_{\mathbb{R}^2 \setminus \frac{K'}{2}} \|_\infty^2 T \big)^{\frac{1}{2}} \\ & + \int_0^t \| R_1 |_{\mathbb{R}^2 \setminus \frac{K'}{2}} \|_\infty \, dt' + \int_0^t (l t' + \alpha t' + \gamma) \left(C_1 \Big[\sup_{(p_1, q_1) \notin \frac{K'}{2}} H_1^{\frac{1}{k} - \frac{1}{2}} \Big] \right) dt' \end{aligned}$$

on $1 - \epsilon$ of the probability space, for C a constant sufficiently large which exists by the Burkholder Davis Gundy inequality combined with the Chebychev inequality.

By Gronwall's inequality e_0 is bounded above by,

$$
\begin{aligned}
&3\epsilon^{-1}C\|R_2|_{\mathbb{R}^2\setminus\frac{K'}{2}}\|_\infty T^{\frac{1}{2}} + \|R_1|_{\mathbb{R}^2\setminus\frac{K'}{2}}\|_\infty T \\
&+ \left(\frac{l}{2}T^2 + \frac{\alpha}{2}T^2 + \gamma T\right)\left(C_1\left[\sup_{(p_1,q_1)\notin\frac{K'}{2}} H_1^{\frac{1}{k}-\frac{1}{2}}\right]\right) \\
&+ \int_0^T (\beta t + \gamma)e^{\beta\frac{t^2}{2}+\gamma t}\Bigg(3\epsilon^{-1}C\|R_2|_{\mathbb{R}^2\setminus\frac{K'}{2}}\|_\infty T^{\frac{1}{2}} + \|R_1|_{\mathbb{R}^2\setminus\frac{K'}{2}}\|_\infty t \\
&\qquad\qquad + \left(\frac{l}{2}t^2 + \frac{\alpha}{2}t^2 + \gamma t\right)\left(C_1\left[\sup_{(p_1,q_1)\notin\frac{K'}{2}} H_1^{\frac{1}{k}-\frac{1}{2}}\right]\right)\Bigg)\,dt
\end{aligned}
$$

for $\beta = l + \alpha$. We therefore have a quantity with a prefactor of,

$$
O\left(\sup_{(p_1,q_1)\notin\frac{K'}{2}}\left(\frac{p_1^2}{2} + \frac{q_1^{2k}}{2k}\right)^{\frac{1}{k}-\frac{1}{2}}\right),
$$

noting the bounds on H_0 throughout this time interval with probability $1-\epsilon$. The prefactor tends to zero with increasing K'. This completes the proof of the bounds on $\sup_{0\le t\le T}|\tilde{p}_0(t) - \bar{p}_0(t)|$.

The part of the theorem involving $\sup_{0\le t\le T}|\tilde{q}_0(t) - \bar{q}_0(t)|$ follows since this is then bounded above by Te_0. □

For the bar variables we have the weak convergence to the stationary regime given by the invariant measure of the bar variables, $\bar{\mu}$. Given a bounded continuous function f, compact set $K \subset \mathbb{R}^2$, and $\epsilon > 0$, we simply take $t'(\epsilon)$ to be a bound above on the time taken from any initial condition in the tilde variables $(\tilde{p}_0(0), \tilde{q}_0(0)) = (p_0(0) + \Phi, q_0(0))$, $(p_0(0), q_0(0)) \in K$, until,

$$
\left|\mathbb{E}_{(\tilde{p}_0(0),\tilde{q}_0(0))}[f(\bar{p}_0(t'(\epsilon)), \bar{q}_0(t'(\epsilon)))] - \int f\,d\bar{\mu}\right| < \epsilon\,.
$$

Then from the Lemma 6.1.5 for H_1 sufficiently large at the start of the time interval, the error $|\mathbb{E}_{(\tilde{p}_0(0),\tilde{q}_0(0))}[f(\tilde{p}_0(t), \tilde{q}_0(t))] - \mathbb{E}_{(\tilde{p}_0(0),\tilde{q}_0(0))}[f(\bar{p}_0(t), \bar{q}_0(t'))]|$ is less than $\|f\|_\infty\epsilon + \|Df\|_\infty\epsilon$ within the time interval $[0, t'(\epsilon)]$. Hence we have

$$
\left|\mathbb{E}_{(\tilde{p}_0(0),\tilde{q}_0(0))}[f(\tilde{p}_0(t'(\epsilon)), \tilde{q}_0(t'(\epsilon)))] - \int f\,d\bar{\mu}\right| < \|f\|_\infty\epsilon + \|Df\|_\infty\epsilon + \epsilon,
$$

for the energy of the 1 variables, H_1, sufficiently large at time 0. Hence for the 1 variables sufficiently large at time 0 noting the bounds $\sup_{0\le s\le t'(\epsilon)}|p_0(s) - \tilde{p}_0(s)| = O(H_1^{\frac{1}{k}-\frac{1}{2}}(0))$ and $\sup_{0\le s\le t'(\epsilon)}|q_0(s) - \tilde{q}_0(s)| = O(H_1^{\frac{1}{k}-\frac{1}{2}}(0))$, we have $|\mathbb{E}_{(p_0(0),q_0(0))}[f(p_0(t'(\epsilon)),$ $\int f\,d\bar{\mu}| < \|f\|_\infty\epsilon + \|Df\|_\infty + 2\epsilon$. This is the content of

Proposition 6.1.6. *For every compact set $K \subset \mathbb{R}^2$, every $\epsilon > 0$, there exists a time $t'(\epsilon)$ and a compact set $K' \subset \mathbb{R}^2$ so that if $(p_0(0), q_0(0)) \in K$, $(p_1(0), q_1(0)) \notin K'$, we have for a bounded continuous function f,*

$$\left| \mathbb{E}_{(p_0(0),q_0(0))}[f(p_0(t'(\epsilon)), q_0(t'(\epsilon)))] - \int f \, d\bar{\mu} \right| < \epsilon$$

where $\bar{\mu}$ is the invariant measure of the system given in (6.0.2).

6.1.4 Stage 3: Advantageous Lowers Bounds on H_1

In this section we study the behavior of the 1 variables (p_1, q_1) with a view to completing the proof of Theorem 6.1.1, giving the weak convergence of the distribution of the 0 variables. It is sufficient to show that given any compact set K', containing 0 in $\mathbb{R}^2$, initial conditions for the full process $U(t)$, and time t', we can find a sufficiently large time $T(K', U(0), t')$ such that at time T the process $(p_1(t), q_1(t))$ is outside of K' and it does not re-enter K' in the time interval $(T, T + t']$ with arbitrarily large probability. We show this by looking at an appropriate test function. Note that Φ is again a function of the 1 variables. For $k > 2$, start with the function,

$$\tilde{H}_1 = \frac{p_1^2}{2} + \frac{q_1^{2k}}{2k} + \frac{\alpha(q_1 - q_0)^2}{2} - \alpha p_0 \Phi - \frac{\alpha^2 \Phi^2}{2} - \frac{\alpha q_0^2}{2} \,. \tag{6.1.5}$$

Then we have,

$$\begin{aligned} d\tilde{H}_1 =& \gamma T_1 + \alpha \Phi (q_0 |q_0|^{2k-2} + \alpha q_0 + \gamma p_0) \\ & - \alpha^2 p_0 (q_0 - q_1) \partial_p \Phi - \alpha^3 \Phi (q_0 - q_1) \partial_p \Phi \\ & - \gamma T_1 (\alpha p_0 \partial_p^2 \Phi + \alpha^2 (\partial_p \Phi)^2 + \alpha^2 \partial_p^2 \Phi \Phi) \, dt \\ & + p_1 \sqrt{2\gamma T_1} \, dW_1 - \alpha \Phi \sqrt{2\gamma T_0} \, dW_0 \\ & - \alpha p_0 \partial_p \Phi \sqrt{2\gamma T_1} \, dW_1 - \alpha^2 \Phi \partial_p \Phi \sqrt{2\gamma T_1} \, dW_1 \,. \end{aligned} \tag{6.1.6}$$

Proposition 6.1.7. *Consider $k \geq 2$, with $\alpha^2 < T_1 \langle \Phi^2 \rangle^{-1}$ for $k = 2$. Given any compact set $K_{H_1} \subset \mathbb{R}^2$, we have,*

$$\limsup_{T \to \infty} \frac{1}{T} \int_0^T 1_{K_{H_1}}(p_1(s), q_1(s)) \, ds = 0 \quad a.s. \,. \tag{6.1.7}$$

Proof. The proof of this follows from the more general statement,

Theorem 6.1.8. *Fix $\delta \in (0,1]$. Given the solution to an SDE in $\mathbb{R}^{n+m}$, n, $m \in \mathbb{N}$, generating a filtration $\mathcal{F}_t$, for $b \in \mathbb{R}^{n+m}$, σ an $(m+n) \times l$ dimensional matrix, b, σ smooth, W an l dimensional standard Brownian motion, $l \in \mathbb{N}$,*

$$dX(S) = b(X(s))\,ds + \sigma(X(s))\,dW(s)\ ,$$

where the first n variables are bounded in the sense that for every $\epsilon > 0$, $\exists K_n(\epsilon) \subset \mathbb{R}^n$, a compact set such that,

$$\liminf_{T\to\infty} \frac{1}{T}\int_0^T 1_{K_n(\epsilon)}\big((x_1(t),\ldots,x_n(t))\big)\,dt > 1-\epsilon \quad a.s.\ . \tag{6.1.8}$$

Assume that we have a smooth function $f : \mathbb{R}^{n+m} \to \mathbb{R}$ such that for every compact set K' in $\mathbb{R}^m$, there exists $a \in \mathbb{R}$ so that $f(x) > a$ for $x = (x_1,\ldots,x_{n+m})$ implies $(x_{n+1},\ldots,x_{n+m}) \notin K'$. Consider the SDE satisfied by $f(X)$. Assume that for $f(X)$ outside of an interval of the form $K_2 = (-\infty, k_2]$, $k_2 \in \mathbb{R}$, we have that there exists an $\mathcal{F}_{T_t}$ Brownian motion B, and a time change T_t continuous for $f(X(T_t)) \notin K_2$, with $T_t > ct$ for some $c > 0$, sufficiently large t a.s. (the size of t for this to occur not uniform over the probability space), such that if $\tau_1 < \tau_2$ are two $\mathcal{F}_{T_t}$ stopping times which satisfy $f(X(T_t)) \notin K_2$ for $\tau_1 \le t < \tau_2$, then,

$$f(X(T_t)) \ge f(X(\tau_1)) + (t-\tau_1)\delta + 2\int_{\tau_1}^t \sqrt{Z}\,dB(s)\ . \tag{6.1.9}$$

In other words, in some sense we have a favorable comparison with the square of a Bessel process of dimension δ.

Assume also that given any $K_1 = (-\infty, k_1]$, $k_1 \in \mathbb{R}$, if $(x_1(0),\ldots,x_n(0)) \in K'_n$, $K'_n \subset \mathbb{R}^n$ compact again, and $f(X(0)) \in K_1$, there is a time $t_{esc}(K_1, K'_n) > 0$, such that there is a non-zero probability $p(K_1, K'_n) > 0$, that after t_{esc}, $f(X) \notin K_1$ independent of the initial value of $(x_1,\ldots x_{n+m})$ and entirely dependent on the independent increment of the noise in time t_{esc}.

Then we have, for every compact set $K_m \subset \mathbb{R}^m$,

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T 1_{K_m}\big((x_{n+1}(t),\ldots,x_{n+m}(t))\big)\,dt = 0 \quad a.s.\ .$$

Remark 6.1.9. The condition (6.1.9) is a condition that results from a favorable comparison due to larger drift.

Proof of Theorem 6.1.8. Fix a compact set $K_m \subset \mathbb{R}^m$. If we can show that given $\epsilon > 0$,

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T 1_{K_m}\big((x_{n+1}(t),\ldots,x_{n+m}(t))\big)\,dt < 3\epsilon \quad \text{a.s.}\,,$$

then we are done.

There is an $a > 0$ such that $f(x) > a$ implies $(x_{n+1},\ldots,x_m) \notin K_m$. Set $k_2' \in \mathbb{R}$ to be such that, $K_2' = (-\infty, k_2'] := (-\infty, a) \cup K_2 \subset \mathbb{R}$ (K_2 from the statement of the theorem). Then set $\sigma_0 = \inf\{t > 0 : f(X(t)) \notin 2K_2'\}$, $\phi_0 = \inf\{t > \sigma_0 : f(X(t)) \in K_2'\}$,..., $\sigma_n = \inf\{t > \phi_{n-1} : f(X(t)) \notin 2K_2'\}$, $\phi_n = \inf\{t > \sigma_n : f(X(t)) \in K_2'\}$, with the obvious meaning for the multiplication of sets by constants.

What we are going to show is,

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{n=0}^{\infty}(\sigma_n \wedge T - \phi_{n-1}\wedge T) < 3\epsilon \quad \text{a.s.}\,, \tag{6.1.10}$$

with $\phi_{-1} \equiv 0$.

At any point in the probability space $\omega \in \Omega$ we build a new process $Y \in \mathbb{R}$ (no longer continuous, cadlag instead) by taking the portions of paths of $f(X)$ between the times $[\sigma_n, \phi_n)$ in order, but with an increased time between these intervals. Between these time intervals we construct the process as follows. We will denote the uniform bound below on the probability of escape from $2K_2'$ with the first n coordinates in $K_n(\epsilon)$, given in the statement of the theorem, as $p > 0$. Where $K_n(\epsilon)$ is given by (6.1.8).

- Step 1: first we ask if $(x_1(\phi_n),\ldots,x_n(\phi_n)) \in K_n(\epsilon)$, if it is not we wait until it is, at time $\phi_n + \phi$ say. Then if $f(X)$ is still in $2K_2'$ at time $\phi_n + \phi$ we take the path of $f(X)$ from ϕ_n to $(\phi_n + \phi + t_{esc}) \wedge \sigma_{n+1}$, but if $f(X) \notin 2K_2'$ at $\phi_n + \phi$ then we conduct independent success/failure trials with success probability p over time intervals of length t_{esc} to determine the time of escape (we take the path from ϕ_n to $(\phi_n + \phi) \wedge \sigma_{n+1} = \sigma_{n+1}$ and then leave Y stationary until these trials are over).

- Step 2: if it is the case that $(x_1(\phi_n,\ldots,x_n(\phi_n)) \in K_n(\epsilon)$ then we take the path of $f(X)$ from ϕ_n to $(\phi_n + t_{esc}) \wedge \sigma_{n+1}$.

- Step 3: repeat the two steps above until escape from $2K_2'$ has occurred.

It is clear from the above construction that Y spends longer than $f(X)$ in $2K_2'$ in any given time interval. Using a further process Z, we are now going to bound above almost surely the proportion of time spent in $2K_2'$ by Y. Firstly we are going to discount the portion of the paths of Y where we are waiting for the condition $(x_1(\phi_n, \ldots, x_n(\phi_n)) \in K_n(\epsilon)$ to become true in between escape attempts, then this is at most ϵ as a proportion of total time remembering that time progresses more slowly for Y than X in terms of how much of the path of X is seen by Y in any given time interval. We then only allow a successful escape if the noise of X is within the p of the probability space that guarantees an escape designated in the statement of the theorem and embark on extra trials if an escape occurs whilst the noise is not in this set. To make this precise, take the path from Y and perform modification of the sections of path between time intervals of path that are taken from $f(X)$ during $[\sigma_n, \phi_n)$.

- STEP 1: cut out all portions of path in Step 1 above where we are waiting for $(x_1(\phi_n + t), \ldots, x_n(\phi_n + t)) \in K_n(\epsilon)$ to become true.
- STEP 2: At every piece of path giving a re-entry to K_2' and subsequent escape of Y from $2K_2'$ perform one of the following two steps:
- STEP 2(i): If the escape was due to extra trials added in Step 1, retain this piece of path from Y.
- STEP 2(ii): If the above fails to hold then ask if the noise was in the designated proportion of size p of the probability space over the time t_{esc}. If so retain the path of Y up to escape with a waiting time at the end of this portion of path to allow completion of the last interval of time of length t_{esc}. If not then wait a length of time after ϕ_n given by independent success failure trials of length t_{esc} after an initial time t_{esc} before resuming the path of Y along the next excursion.

Z is again constructed from the portions of paths of $f(X)$ between the times $[\sigma_n, \phi_n)$ in order but the time in between is the time given by independent success/failure trials with success probability p (derived directly from the noise of the process X) over time intervals of length t_{esc} to determine the time of escape. (6.1.10) for Z is at least as large as (6.1.10) for Y minus 2ϵ pathwise, owing to the addition of extra trials for Z. For this process Z we have that (6.1.10) is almost surely bounded above by,

$$\frac{1}{T}\sum_{i=0}^{N_Z(T)} M_{Z,i} \; , \tag{6.1.11}$$

for $M_{Z,i}$ a series of independent geometric distributions with probability of success p, and for σ_n^Z, ϕ_n^Z the corresponding series of stopping times to σ_n, ϕ_n as above but for Z instead,

$$N_Z(T) = \sum_{n=0}^{\infty} 1_{\{\phi_{n-1}^Z < T\}} \; .$$

Using the strong law of large numbers applied to,

$$\frac{1}{N_Z(T)}\sum_{i=0}^{N_Z(T)} M_{Z,i} \; .$$

(6.1.11) is almost surely less than, for large enough T,

$$\frac{N_Z(T)(p^{-1}+\epsilon)}{T} \; , \tag{6.1.12}$$

since $N_Z(T) \to \infty$ a.s. as $T \to \infty$ since any geometric distribution is almost surely finite.

We will now use a lemma regarding the square of a Bessel process of dimension δ to show that (6.1.12) tends to 0 a.s..

Lemma 6.1.10. *Let Z be a squared Bessel process of dimension $\delta \in (0,2]$. Given a compact set $K' = [0,k'] \subset \mathbb{R}$ and a larger compact set $K'' = [0,k''] \subset \mathbb{R}$, $0 < k' < k''$, set $\phi_{-1}^Z \equiv 0$, $\sigma_0^Z = \inf\{t > 0 : Z(t) \notin K''\}$, $\phi_0^Z = \inf\{t > \sigma_0 : Z(t) \in K'\}, \ldots$, $\sigma_n^Z = \inf\{t > \phi_{n-1}^Z : Z(t) \notin K''\}$, $\phi_n^Z = \inf\{t > \sigma_n^Z : Z(t) \in K'\}$ and $N_Z(T) = \sum_{n=0}^{\infty} 1_{\{\phi_n^Z < T\}}$, then*

$$\limsup_{T\to\infty} \frac{N_Z(T)+1}{T - \sum_{n\geq 0} \sigma_n^Z \wedge T - \phi_{n-1}^Z \wedge T} = 0 \quad a.s. \; . \tag{6.1.13}$$

Proof of Lemma 6.1.10. For Z by the Strong Markov Property, we have that for $n \geq 1$,

$$\mathbb{E}_x[\sigma_n^Z - \phi_{n-1}^Z] = \mathbb{E}_{\phi_{n-1}^Z}[\sigma_0^Z] = \mathbb{E}_{k'}[\sigma_0^Z] \; ,$$

and σ_0^Z with the initial point k' is a series of independent identically distributed random variables. Hence by the strong law of large numbers we have that, since $\sigma_0^Z - \phi_{-1}^Z$ can easily be shown to have finite expectation,

$$\frac{1}{N}\sum_{n=0}^{N}(\sigma_n^Z - \phi_{n-1}^Z) \to \mathbb{E}_{k'}[\sigma_0^Z] > 0 \quad \text{a.s.}\,, \tag{6.1.14}$$

as $N \to \infty$. In addition for all N, we have that,

$$\sum_{n=0}^{N}(\sigma_n^Z \wedge T - \phi_{n-1}^Z \wedge T) \leq \int_0^T 1_{K''}(Z(s))\, ds\,.$$

We have that the RHS of the above equation tends to zero a.s. when divided by T which follows from the ratio ergodic theorem for recurrent continuous Markov processes in $\mathbb{R}$ of [Der54]. Therefore so does the LHS. Now if $N_Z(T) \to \infty$ as $T \to \infty$ for this point in the probability space, for T (dependent on the probability space) sufficiently large the LHS is a.s. greater than $(1-\epsilon)N_Z(T)\mathbb{E}_{k'}[\sigma_0^Z]$ from (6.1.14). If the assumption fails to hold then the result is clearly true. Hence we have $\frac{N_Z(T)}{T} \to 0$ which implies $\frac{N_Z(T)+1}{T-\sum_{n=0}^{N_Z(T)}(\sigma_n^Z-\phi_{n-1}^Z)} \to 0$ as $T \to \infty$ a.s. as required. □

Remark 6.1.11. The quantity (6.1.13) is the number of stopping times in a time interval of length $T - \sum_{n=0}\sigma_n^Z \wedge T - \phi_n^Z \wedge T$ where the process skips the intervals $[\phi_{n-1}^Z, \sigma_n^Z)$.

We construct a squared Bessel process of dimension δ as follows, we take the driving $\mathcal{F}_{T_t}$ Brownian motion from comparison in the intervals $T_t \in [\sigma_n^Z, \phi_n^Z]$ from the time σ_n^Z until the squared Bessel process hits k_2', and then fill in these independent excursions using paths of a squared Bessel process driven by a series of independent Brownian motions with initial points k_2' and terminal point $2k_2'$. The excursions are independent since the times σ_n^Z correspond to $\mathcal{F}_{T_t}$ stopping times cf [RY91, Exercise 3.21 Chapter IV].

Taking $K = K_2$ and $K' = 2K_2$ in the above lemma for the square of the Bessel process of dimension δ just constructed, and observing the bounds on the time change T_t from which we obtain $\sigma_n^Z \geq (1-\epsilon)c(\sigma_n^Z - \sum_{j<n}\sigma_j^Z - \phi_j^Z)$ for n sufficiently large. This gives the a.s. convergence of (6.1.12) to zero and completes the proof of the theorem. □

We are going to apply the above theorem to $U(t) = (p_0(t), q_0(t), p_1(t), q_1(t)) \in \mathbb{R}^4$ with $n = m = 2$. We will start with the case $k > 2$. Consider the equation satisfied by the test function $\tilde{\tilde{H}}_1 = \tilde{H}_1 - K_0 E_0(\tilde{p}_0, q_0)$ where $K_0 > 0$ is a constant and $\tilde{H}_1$ is as in (6.1.5). $\tilde{\tilde{H}}_1$ is less than $\tilde{H}_1$ and we claim that the one dimensional SDE satisfied by $\tilde{\tilde{H}}_1$ has positive drift (bounded away from 0) when $\tilde{\tilde{H}}_1$ is outside of an interval of the form $(-\infty, k_1)$, $k_1 > 0$ sufficiently large. This claim is verified as follows, studying the leading order terms in $\tilde{\tilde{H}}_1$, when $\tilde{\tilde{H}}_1$ is large, then H_1 is large. The only way of having a negative drift greater in magnitude than γT_1, noting that every term has a prefactor in the 1 variables that tends to zero as H_1 tends to infinity, is if the 0 variables are large. If the 0 variables are large the positive drift terms $\gamma p_0^2 + \theta q_0^{2k}$ arising from E_0 dominate all other drift terms in the 0 variables. Hence when $\tilde{\tilde{H}}_1$ is sufficiently large the drift term is strictly positive and bounded away from 0 which verifies the claim. We use the positive drift outside of such a set to demonstrate the reluctance of the 1 variables to occupy a compact set using Theorem 6.1.8 by setting $f = \tilde{\tilde{H}}_1$ and making a comparison after a time change.

Recall that the martingale part of $\tilde{\tilde{H}}_1$ is given by

$$\begin{aligned}
&K_0(-\tilde{p}_0 - \theta q_0)(\sqrt{2\gamma T_0}\, dW_0 + R_2\, dW_1) \\
&- \alpha\Phi\sqrt{2\gamma T_0}\, dW_0 + p_1\sqrt{2\gamma T_1}\, dW_1 \\
&- \alpha p_0 \partial_p \Phi \sqrt{2\gamma T_1}\, dW_1 - \alpha^2 \Phi \partial_p \Phi \sqrt{2\gamma T_1}\, dW_1\ .
\end{aligned}$$

We now verify the hypotheses of Theorem 6.1.8. The non-zero probability of escape from any compact set is given by the following lemma,

Lemma 6.1.12. *Consider $k \geq 2$. For any $t' > 0$, compact sets K_0, $K_1 \subset \mathbb{R}^2$ with nonempty interiors, there exists a small probability $p > 0$ (dependent on t', K_0, K_1) such that if $(p_0(0), q_0(0)) \in K_0$ and $(p_1(0), q_1(0)) \in K_1$, then*

$$\mathbb{P}\big((p_1(t'), q_1(t')) \notin K_1 \,\&\, (p_0(t'), q_0(t')) \in K_0\big) \geq p\ .$$

Proof. Fix a target point U' in the interior of $K_0 \times K_1^c$. Then, for every starting point U_0 in $K_0 \times K_1$, we can find an instance of (W_0, W_1) such that the corresponding solution to (6.0.1) satisfies $U(t') = U'$. The claim then follows from the continuity of the solution map with respect to (W_0, W_1), combined with the fact that Wiener measure has full support. □

Now for the second part. We have that outside a large set $K_1 \subset \mathbb{R}$ of the form $(-\infty, k_1)$, the process $\tilde{\tilde{H}}_1$ is equal to, for K_0 small enough so that $K_0 C_l \leq \gamma T_1/4$ for C_l the constant involved in the Lyapunov condition satisfied by E_0, $\mathcal{L}E_0 \leq C_l - cE_0$,

$$\begin{aligned} d\tilde{\tilde{H}}_1 = &\frac{\gamma T_1}{2}\,dt + b_1\,dt + K_0(\sqrt{2\gamma T_0})(-\tilde{p}_0 - \theta q_0)\,dW_0 - \alpha\Phi\sqrt{2\gamma T_0}\,dW_0 \\ &+ p_1\sqrt{2\gamma T_1}\,dW_1 - \alpha p_0 \partial_p \Phi \sqrt{2\gamma T_1}\,dW_1 \\ &- \alpha^2 \Phi \partial_p \Phi \sqrt{2\gamma T_1}\,dW_1 - K_0(\tilde{p}_0 + \theta q_0) R_2\,dW_1\ , \end{aligned}$$

where b_1 is positive and at least as large as $cK_0E_0/2$ for $c > 0$ the prefactor of E_0 involved in the Lyapunov condition on E_0. Let $K_1' \subset \mathbb{R}$ be such that,

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T (1 - 1_{K_1'})\left(p_0^2(t) + q_0^{2k}(t)\right) dt < 1 \quad \text{a.s.} \tag{6.1.15}$$

Being loose for a second with notation about the starting time. We will use the following time change, for comparison with a squared Bessel process, we want T_t to have the same rate as T_t' for $\tilde{\tilde{H}}_1$ outside the set K_1'' given by $K_1 \cup 2K_1' \subset K_1'' = (-\infty, k_1'') \subset \mathbb{R}$, differing by possibly a linear factor,

$$\begin{aligned} T_t' = \inf\Bigg\{ s : \int_0^s &\{[K_0(\sqrt{2\gamma T_0})(-\tilde{p}_0 - \theta q_0) - \alpha\Phi\sqrt{2\gamma T_0}]^2 \\ &+ [p_1\sqrt{2\gamma T_1} - \alpha p_0\partial_p\Phi\sqrt{2\gamma T_1} - \alpha^2\Phi\partial_p\Phi\sqrt{2\gamma T_1} \\ &- K_0(\tilde{p}_0 + \theta q_0)R_2]^2\}\{4\tilde{\tilde{H}}_1(s)\}^{-1}\,ds > t\Bigg\}\,. \end{aligned} \tag{6.1.16}$$

The LHS of the inequality is continuous in time and zero only on a set of Lebesgue measure zero. That this is zero only on a set of Lebesgue measure zero follows from the occupation times formula applied to the first integrand on the LHS. This time change results from studying the SDE satisfied by $\sqrt{\tilde{\tilde{H}}_1}$ and then making a time change to replace the integrand in the stochastic integral by 1. Since the integrand in (6.1.16) is zero only on a set of Lebesgue measure zero, the time change is continuous. On making this time change outside of the set K_1'' we end up with,

$$d\tilde{\tilde{H}}_1(T_t') = b_2'\,dt + \sqrt{\tilde{\tilde{H}}_1(T_s')}\,d\tilde{W}(s)\ , \tag{6.1.17}$$

for some constant $b_2' > b_2 > 0$, considering the two cases where $(2\gamma K_0^2 T_0 + 2(K_0\|R_2\|_\infty)^2 + 1)H_0$ is less than $\tilde{\tilde{H}}_1/2(K_0+1)$ and larger than $\tilde{\tilde{H}}_1/2(K_0+1)$ separately. In the process, making K_1'' larger if necessary to obtain these bounds using the leading order terms. Where the existence of such a Brownian motion is guaranteed from the description of the time change as a time change for $\sqrt{\tilde{\tilde{H}}_1}$, the expression for which we then square back again using Itô's formula to get back to an expression for $\tilde{\tilde{H}}_1$. The conversion of the Stieltjes integral with respect to dT_t' to one with respect to dt follows from the fact that Stieltjes measure can be shown to be absolutely continuous with respect to Lebesgue measure as P_1' is zero only on a set of zero Lebesgue measure. So we set T_t in between excursions from K'' as having the same rate as T_t'. Then we will use $f = \tilde{\tilde{H}}_1$ in Theorem 6.1.8 with a squared Bessel process of dimension $b_2/2$.

When $\tilde{\tilde{H}}_1$ is inside K_1'', the time change simply advances to the point in time at which $f = a\tilde{\tilde{H}}_1$ has made its escape from a slightly larger set. More specifically, for $j \in \mathbb{N}$ we have a series of stopping times ϕ_n and σ_n where we have $\phi_{-1} \equiv 0$, $\sigma_0 = \inf\{t > 0 : a\tilde{\tilde{H}}_1 \notin K_1'' + 1\}$, $\phi_0 = \inf\{t > \sigma_0 : a\tilde{\tilde{H}}_1 \in K_1''\}, \ldots, \sigma_n = \inf\{t > \phi_{n-1} : a\tilde{\tilde{H}}_1 \notin K_1'' + 1\}$, $\phi_n = \inf\{t > \sigma_n : a\tilde{\tilde{H}}_1 \in K_1''\}$. Then we choose the time change T_t as follows, we progress immediately to the terminal time σ_n in the intervals $[\phi_{n-1}, \sigma_n]$, $n \in \mathbb{N}_0$, i.e when $T_t = \phi_{n-1}$, and in the intervals $[\sigma_n, \phi_n)$, we progress at rate given by (6.1.16). Therefore by a simple comparison argument, for instance using the Itô formula applied to $\sqrt{\tilde{\tilde{H}}_1}$ (cf [HM10b]) we have (6.1.9), outside of $K_1'' + 1$, using the driving Brownian motion from (6.1.17) with a time change with a factor greater than a positive fixed amount except where H_0 is large. In other words the Brownian motion we use in Theorem 6.1.8 is the DDS Brownian motion of the appropriate martingale term with T_t as the DDS time change,

$$\begin{aligned}\int_0^t \sum_n 1_{[\sigma_n,\phi_n)}[K_0(\sqrt{2\gamma T_0})(-\tilde{p}_0 - \theta q_0) - \alpha\Phi\sqrt{2\gamma T_0}][2\sqrt{\tilde{\tilde{H}}_1}]^{-1}\,dW_0 \\ + \int_0^t \sum_n 1_{[\sigma_n,\phi_n)}[p_1\sqrt{2\gamma T_1} - \alpha p_0\partial_p\Phi\sqrt{2\gamma T_1} - \alpha^2\Phi\partial_p\Phi\sqrt{2\gamma T_1} \\ - K_0(\tilde{p}_0 + \theta q_0)R_2][2\sqrt{\tilde{\tilde{H}}_1}]^{-1}\,dW_1\,.\end{aligned}$$

This is not a problem with regard to the condition $T_t \geq ct$ for some $c > 0$, by (6.1.15), we have an average rate of at least, for A_t the inverse of the time change

T_t,

$$\frac{T_t}{t} = \frac{T_t}{A_{T_t}} \geq \frac{1}{\gamma T_1 + 1 + (2\gamma T_1 + 1) T_t^{-1} \int_0^{T_t} (1 - 1_{K_1'}) \left(\frac{p_0^2(s)}{2} + \frac{q_0^{2k}(s)}{2k} \right) ds}$$
$$\geq \frac{1}{\gamma T_1 + 1 + (2\gamma T_1 + 1)} \tag{6.1.18}$$

a.s. as $t \to \infty$, for $K_0 < (4\gamma T_0 + 2)^{-1/2}$, K_1'' sufficiently large.

Noting the bounds on the 0 variables given by Lemma 6.1.4 completes the verification of the hypotheses of Theorem 6.1.8, hence we can apply it to $\tilde{\tilde{H}}_1$. The form of the error terms in the context of Lemma 6.1.4 when comparing H_1 to $\tilde{\tilde{H}}_1$ proves the result.

Moving onto the case $k = 2$, $\alpha^2 < T_1 \langle \Phi^2 \rangle^{-1}$, we have to use additional correction terms in this case in order to get rid of terms that we previously had convergence to 0 of in the case $k > 2$, specifically those terms in the 0 variables that arise as a result of applying the generator to $-\alpha p_0 \Phi$. Consider $\tilde{H}_1$ given by,

$$\tilde{H}_1 = \frac{p_1^2}{2} + \frac{q_1^{2k}}{2k} + \alpha \frac{(q_1 - q_0)^2}{2} - \alpha p_0 \Phi - \frac{\alpha^2}{2} \Phi^2 - \alpha \frac{q_0^2}{2}$$
$$+ \alpha \Psi (q_0 |q_0|^{2k-2} + \alpha q_0 + \gamma p_0) + \alpha^2 \gamma \Psi \Phi - \alpha^2 \gamma \Xi \,,$$

where $\mathcal{L}\Psi = -\Phi$ and $\mathcal{L}\Xi = -\Phi^2 + \langle \Phi^2 \rangle$. We set $\tilde{\tilde{H}}_1 = \tilde{H}_1 - K_0 E_0$ again, for K_0 slightly modified so that $K_0 = [(\gamma T_1 - \alpha^2 \langle \Phi^2 \rangle)/4C_l] \wedge (4\gamma T_0 + 2)^{1/2}$. Applying Itô's formula we have almost the same form of stochastic terms and drift terms as before and hence an almost identical argument applies to $f = \tilde{\tilde{H}}_1$, where we use a corresponding time change to T_t. The only difference in the argument required is due to the term $\alpha \partial_p \Psi q_0 |q_0|^2$ which appears to thwart the bounding of the drift away from 0 after the time change T_t' outside of the set K_1'' by being of greater order in the 0 variables than H_0. Except that if we make the compact set K_1'' sufficiently large then for $\tilde{\tilde{H}}_1$ outside of K_1'', we have $H_1 \geq K_0 H_0 / 2$, thus implying (noting the order of $\partial_p \Psi$) that $\alpha \partial_p \Psi q_0 |q_0|^2$ is bounded above by a fixed constant (dependent on Ψ, K_0, α), so actually this is a constant order term and an identical argument then applies. □

Proposition 6.1.13. *Given fixed $t' > 0$, $\epsilon > 0$, compact $K \subset \mathbb{R}^2$, there exists $T' > 0$ sufficiently large such that for all $T > T'$,*

$$\mathbb{P}[\exists s,\, T < s < T + t', H_1(p_1(s), q_1(s)) \in K] < \epsilon$$

Proof. We will retain the notation from the previous proof. Note that by defining the set K_1'' in the previous proof large enough we can obtain that a similar time change to T_t, T_t^ϵ, is arbitrarily close to a constant factor time change with arbitrarily high probability for t sufficiently large, ϵ sufficiently small. In other words, if $k_1'' > 0$ is large enough, for large enough times, small ϵ, the factor of the time change, T_t^ϵ / t, that we can use to compare $\tilde{\tilde{H}}_1$ to a squared Bessel process is arbitrarily close to the constant,

$$\frac{4T_p(1)}{2\gamma T_1 A^p_{T_p(1)}} = \frac{2}{\gamma T_1 \langle p^2 \rangle},$$

over an arbitrarily large proportion of the probability space. In verifying this we will use the result of Lemma 6.3.1 regarding the proximity of the trajectories of the 1 variables at high energy and that of an isolated oscillator. $A^p_{T_p(1)}$ denotes the inverse of a time change similar to T_t' for an isolated oscillator at energy 1 at the end of one period $T_p(1)$ i.e. the inverse of the time change given by $T_t^p = \inf\{s : \int_0^s (p(s'))^2 / H(s')\, ds' > t\}$. $\langle p^2 \rangle$ denotes the ergodic average of p^2 over one time period for an isolated oscillator.

This fact regarding the time change T_t^ϵ, is the content of the following two lemmas,

Lemma 6.1.14. *Fix $t' > 0$, $\epsilon > 0$, and a compact set $K_0 \subset \mathbb{R}^2$. Assume that the 0 variables are within the compact set K_0 at time 0. Then if the energy H_1 is sufficiently large initially at time 0 (lower bounds on H_1 dependent on ϵ, K_0), we have, for A_t' the inverse of the time change T_t' above,*

$$\mathbb{P}\left[\sup_{0 \le s \le t'} \left| A_s' - \frac{1}{2}\gamma T_1 \langle p^2 \rangle s \right| > \epsilon \right] < \epsilon$$

This leads to the following lemma.

Lemma 6.1.15. *Fix $\epsilon > 0$, then for $T_\epsilon > 0$ sufficiently large we have,*

$$\mathbb{P}\left[\left| T^\epsilon_{T_\epsilon} - \frac{2}{\gamma T_1 \langle p^2 \rangle} T_\epsilon \right| > C_1 \epsilon T_\epsilon \right] < C_2 \epsilon$$

where C_1, C_2 constants and T_t^ϵ is defined exactly as T_t except that we choose K_1' such that,

$$\limsup_{T \to \infty} \frac{1}{T} \int_0^T (1 - 1_{K_1'}) \left(\frac{p_0^2(t)}{2} + \frac{q_0^{2k}(t)}{2k} \right) dt < \epsilon \quad a.s. \tag{6.1.19}$$

instead of 1.

Remark 6.1.16. The reason for this choice of K_1' dependent on ϵ is to limit the average rate of the inverse time change to T_t^ϵ.

Remark 6.1.17. It may seem counterintuitive to prove a statement about a time change through a local statement about the inverse of the time change, but the proof of Lemma 6.1.15 is conducted using the inverse to T_t^ϵ, A_t^ϵ, for reasons that will become apparent.

Taking the result of these two lemmas for granted, we have from a comparison argument similar to that in [HM10b], and the definition of the time change T_t^ϵ, that there exists a Brownian motion (the DDS Brownian motion mentioned previously) $\tilde{W}$ driving a squared Bessel process starting at 0 such that,

$$f(X(T_t^\epsilon)) \geq Z(t) \ . \tag{6.1.20}$$

Now if K is sufficiently large, we have

$$H_1(p_1(t), q_1(t)) \leq C_K = \sup_{(p_1,q_1)\in K} H_1(p_1, q_1) \ ,$$

which implies that

$$Z(A_t^\epsilon) \leq 2C_K + CH_0 + k_1'' + 1 \ , \tag{6.1.21}$$

where C is a constant, for A_t^ϵ the inverse of T_t^ϵ.

Given $\epsilon > 0$ small, we will use (6.1.21). We rescale time and space by $2/\gamma T_1 \langle p^2 \rangle T$ and look at $\tilde{Z}(1)$, the tilde denoting that it is the rescaled squared Bessel process that is also a squared Bessel process of the same dimension itself. Choose $k', k'' \in \mathbb{R}$ large enough so that $\mathbb{P}[2C_K + CH_0(T) + k_1'' + 1 > k''] < \epsilon$, which is possible by Lemma 6.1.3, and then k' such that the probability of $2C_K + CH_0 + k_1''$ exceeding k' in time t' from an initial condition in k'' is less than ϵ (this follows from the form of the SDE satisfied by E_0). Then we finish the result by bounding the oscillations in probability over a small time in the future or past $\sim \epsilon$ resulting from errors in the time change in order to bound the probability that $Z(A_\cdot^\epsilon) \leq 2C_K + CH_0 + k_1'' + 1$. We have, for T sufficiently large dependent on c such that $\mathbb{P}[|\tilde{Z}(1)| < c] < \epsilon$, at least as large as $\max\{4k''/\gamma T_1 \langle p^2 \rangle c, (\gamma T_1 \langle p^2 \rangle + 2C_1'\epsilon)t'/C_1'\gamma T_1 \langle p^2 \rangle \epsilon - 2(C_1'\epsilon)^2\}$, for $C_1' = \frac{1}{2}\gamma T_1 \langle p^2 \rangle C_1$,

$$\mathbb{P}\left[\inf_{T<s<T+t'} H_1(s) < C_K\right]$$

$$
\begin{aligned}
&\leq \mathbb{P}\left[|\tilde{B}(1)| < c\right] \\
&\quad + \mathbb{P}\Big[\sup_{T\leq s\leq T+t'} 2C_K + CH_0 + k_1'' + 1 > k'\Big] \\
&\quad + \mathbb{P}\left[\,\Big|\,A_T^\epsilon - \frac{1}{2}\gamma T_1\langle p^2\rangle T\,\Big| > C_1'\epsilon T\right] \\
&\quad + \mathbb{P}\left[\,\Big|\,A_{T+t'}^\epsilon - \frac{1}{2}\gamma T_1\langle p^2\rangle (T+t')\,\Big| > C_1'\epsilon (T+t')\right] \\
&\quad + \mathbb{P}\Big[\sup_{1-2C_1'\epsilon\leq s\leq 1+3C_1'\epsilon} |\tilde{Z}_s - \tilde{Z}_{1-2C_1'\epsilon}| > c\Big]
\end{aligned}
$$

The first term is bounded above by ϵ by construction, the second is bounded above by 2ϵ by construction, the third and fourth are bounded above by $2C_2\epsilon$ by Lemma 6.1.15 and the last one converges to zero as $\epsilon \to 0$, $T \to \infty$ by tightness of $\tilde{Z}$ on the space of continuous paths. □

Now the proof of the lemmas,

Proof of Lemma 6.1.14. The inverse to T_t', is given by, for appropriate P_1^2,

$$A_t' = \int_0^t \frac{\gamma T_1 P_1^2}{2\tilde{\tilde{H}}_1}(s)\, ds\ . \tag{6.1.22}$$

Now we have that, for all $\delta > 0$, for $H_1(0)$ sufficiently large,

$$\mathbb{P}\left[\sup_{0\leq s\leq t'}\left|\frac{H_1(s)}{H_1(0)} - 1\right| + \sup_{0\leq s\leq t'}\left|\frac{\tilde{\tilde{H}}_1(s)}{H_1(s)} - 1\right| < \delta\right] > 1-\delta,$$

which follow easily from the order of change of H_0 over a finite time being H_0 and the fact that H_1 changes by order $H_1^{1/2}$, therefore if $H_1(0)$ is sufficiently large dependent on $H_0(0)$ we have this result. We therefore have bounds on $\sup_{0\leq s\leq t'} H_0(s)/H_1^{1/4}(s)$ as small as we like with arbitrarily high probability by choosing $H_1(0)$ sufficiently large. This implies that P_1^2 behaves like p_1^2 when $H_1(0)$ is large and $H_0(0)$ is not.

If we let the series of stopping times ξ_n be given by $\xi_1 = T_p(H_1(0)),\ldots,\xi_n = \xi_{n-1} + T_p(H_1(\xi_{n-1}))$ as in Lemma 6.3.1, then from above, for $H_1(0)$ sufficiently large, we can exchange A_t' for $\hat{A}_t$ given by

$$\hat{A}_t = \sum_n \int_{\xi_n\wedge t'}^{\xi_{n+1}\wedge t'} \frac{\gamma T_1 p_1^2}{2H_1(\xi_n)}\, ds, \tag{6.1.23}$$

at an error of at most $\epsilon/3$ with probability at least $1-\epsilon/2$. By Lemma 6.3.1, if $H_1(0)$ is sufficiently large we can exchange the trajectory of p_1 within the times $[\xi_n, \xi_{n+1}]$

for p in the case of an isolated oscillator and the sum of the supremums of the errors in p^2 on the scale of $H_1(0)$ is less than $\epsilon/2\gamma T_1 t'$ with probability at least $1-\epsilon/2$. The supremum of the errors from the trajectory of an isolated oscillator is certainly less than the sum of the supremums within each interval. The intervals $[\xi_n, \xi_{n+1}]$ are complete periods for the isolated oscillator, and the corresponding inverse time change for the isolated oscillator has a fixed rate r (independent of the energy or starting position in the period) over each of them. The length of any period is at most $[\inf_{0\le s\le t'} H_1(s)]^{1/2k-1/2}$ and the rate of progression of (6.1.23) for an isolated oscillator is at most $\gamma T_1/2$.

By the above discussion we have the convergence of (6.1.22) for $t = t'$ to within ϵ of rt' on $1-\epsilon$ of the probability space where $H_1(0)$ is sufficiently large.

So now with an isolated oscillator at energy 1 we have that the rate of $A^p_{T_p(1)}$ at the end of one period $T_p(1)$ is equal to,

$$\frac{A^p_{T_p(1)}}{T_p(1)},$$

where T^p_t denotes the corresponding time change for an isolated oscillator at energy 1 and $A^p_{T_p(1)}$ denotes the inverse at the end of one period $T_p(1)$. We then have,

$$A^p_{T_p(1)} = \int_0^{T_p(1)} p^2(s)\, ds\ ,$$

Hence we have,

$$r = \frac{\gamma T_1 \int_0^{T_p(1)} p^2(s)\, ds}{2T_p(1)}\ ,$$

which completes the proof. □

Proof of Lemma 6.1.15. Given $\epsilon > 0$, fix any $t_c > 1$. To prove the lemma we will study the inverse time change.

Start with K_0 sufficiently large so that the proportion of time spent inside K_0 by the 0 variables is at least $1-\epsilon$ for T sufficiently large almost surely, from Lemma 6.1.4. Then choose K_1 sufficiently large (at least as large as $K_1''+1$ we use in the production of the time change T^ϵ_t) such that we have the result of Lemma 6.1.14 substituting ϵ^2 for ϵ over a time interval of length $t' = t_c$. For T sufficiently large we can choose a series of stopping times $t_1, \ldots, t_{n(T)}$, $n(T) = \lceil (1-2\epsilon)T/t_c \rceil$, such that

$$\mathbb{P}[H_1(t_n) \notin K_1 \text{ and } H_0(t_n) \in K_0 \, \forall n] > 1 - \epsilon^2 \, .$$

We then have that the intervals $(t_i, t_i + t_c)$ in which we have the proximity of the inverse, discarding those intervals where the inverse is not within ϵ of the constant time change, cover a proportion of T greater than $(1 - \epsilon - 2\epsilon) > 1 - 3\epsilon$ (for ϵ small) by Fubini's theorem, with probability at least $1 - \epsilon - \epsilon^2 > 1 - 2\epsilon$.

Looking at A^ϵ_t with regard to proximity to the constant time change with rate $\gamma T_1 \langle p^2 \rangle /2$, the only potential problem is where the inverse time change proceeds at a fast rate in between these intervals and on those intervals where the bound from Lemma 6.1.14 fails to hold. For K_1', K_1'' sufficiently large, the inverse time change is bounded in rate by $2\gamma T_1 + 2\gamma T_0$ except where H_0 is outside of K_1' and $\tilde{\tilde{H}}_1$ is outside of K_1''. Choose T sufficiently large so that with probability $1 - \epsilon$ we have the bounds of ϵ on (6.1.19). When $\tilde{\tilde{H}}_1$ is outside of K_1'' on the time interval $[\tau_0, t]$, the inverse A^ϵ_T is given by,

$$\int_{\tau_0}^{t} \frac{P_1^2(s)}{\tilde{\tilde{H}}_1(s)} \, ds + A^\epsilon_{\tau_0} \, ,$$

where τ_0 is a random time. Studying P_1, $\tilde{\tilde{H}}_1$, the LHS of (6.1.19) bounds $T^{-1}(2c\gamma T_1 + c\gamma T_0)^{-1} \in \mathbb{R}$ times the contribution to the average rate of the inverse time change where H_0 is outside of K_1', above the rate of $2\gamma T_1$. For c a constant less than $(K_1'')^{-1}$ for the choice of K_1'' dependent on ϵ, K_1'' increasing as $\epsilon \to 0$. Then noting the coverage of the intervals of length t_c and the definition of T^ϵ_t, we have immediately that for time T sufficiently large, with probability $1 - 2\epsilon - \epsilon$ the inverse $A^\epsilon_T = (T^\epsilon_T)^{-1}$ is within the range $(\gamma T_1 \langle p^2 \rangle (1 - 3\epsilon - \epsilon^2) T/2, \gamma T_1 \langle p^2 \rangle T/2 + (2c\gamma T_1 + c\gamma T_0)\epsilon T + 6\epsilon\gamma(T_0 + T_1)T)$. Converting this to a statement about T^ϵ_T completes the proof. □

6.1.5 Putting all the parts together: proof of Theorem 6.1.1

Consider given $\epsilon > 0$. Put ϵ into Lemma 6.1.3 to obtain a set $K(\epsilon)$ and a time $T(\epsilon)$, for the 0 variables to occupy after time $T(\epsilon)$. Next input ϵ, $K(\epsilon)$ into Proposition 6.1.6 to obtain a time $t'(\epsilon)$ and a compact set $K'(\epsilon)$ for the 1 variables not to occupy.

Then finally input ϵ, $t'(\epsilon)$, $K'(\epsilon)$ into Proposition 6.1.13 to obtain a large time $T'(\epsilon)$. Then for all times t larger than $T'(\epsilon) + t'(\epsilon) + T(\epsilon)$, we have using the Strong Markov property,

$$\left| \mathbb{E}_{(p_0(0),q_0(0),p_1(0),q_1(0))}[f(p_0(t), q_0(t))] - \int f\, d\bar{\mu} \right|$$
$$= \left| \mathbb{E}_{(p_0(\hat{t}),q_0(\hat{t}),p_1(\hat{t}),q_1(\hat{t}))}[f(p_0(t(\epsilon)), q_0(t(\epsilon)))] - \int f\, d\bar{\mu} \right| < \epsilon + 2\|f\|_\infty ,$$

where $\bar{\mu}$ is the invariant measure of (6.0.2) and $\hat{t} = t - t'(\epsilon)$. This completes the proof.

6.2 Ergodic Averaging for $k > 2$

The various stages remain the same as for weak convergence except the convergence section is somewhat modified and the requirement on the energy H_1 is also slightly different. For H_1, we need it to be large an increasingly large proportion of the time as opposed to at the end of the time interval with increasingly high probability. We can approximate any continuous and bounded f in the limit as time tends to infinity by functions of compact support from Lemma 6.1.4. By taking the positive and negative parts of f, we can assume the f we are dealing with is positive. Given an f of compact support we can approximate it uniformly by smooth functions of compact support, hence we will assume that f is positive, of compact support and smooth from now on.

6.2.1 Modification to the convergence stage, stage 2

Assume we are given $\epsilon > 0$. Then we modify this stage in a way that means in the next subsection we are looking for a compact set K_1 outside of which the error given by e_0 in Lemma 6.1.5 above is sufficiently small on a large proportion of the probability space, over a longer length of time, which we call t_c again. $t_c > 1$ is such that the proportion of time where the distribution of $(\bar{p}_0, \bar{q}_0)$ integrated against f differs from that against the invariant measure of the bar variables $\bar{\mu}$ by less than ϵ, is less than ϵ. This is taking into account the time taken for convergence from an initial distribution in the bar variables corresponding to the tilde variables initially in the compact set K' derived from K given by,

$$\limsup_{T\to\infty} T^{-1}\int_0^T 1_{K^c}(p_0(s),q_0(s))\,ds < \epsilon/2 \quad a.s.\,. \tag{6.2.1}$$

Set $K' = \{(x,y)\in\mathbb{R}^2 : \exists|k| \le \|\Phi\|_\infty,\ (x+k,y)\in K\}$. (6.2.1) follows from Lemma 6.1.4. Then in other words, by the uniform convergence to the invariant measure on compact sets [Hai09], t_c is such that for all $t > t_c$,

$$\sup_{(\bar p_0,\bar q_0)\in K'} t^{-1}\int_0^t 1_{|\mathbb{E}_{(\bar p_0,\bar q_0)}[f(\bar p_0(s),\bar q_0(s))]-\bar\mu(f)|>\epsilon}\,ds < \epsilon\,. \tag{6.2.2}$$

We want K_1 to be such that for the initial conditions in the 1 variables outside of K_1, e_0 over the time period t_c is greater than ϵ/t_c on at most ϵ of the probability space (see the last stage in this section). We are going to break time into a series of intervals of length t_c that cover a large proportion of the time.

6.2.2 Putting it all together

Theorem 6.2.1. *For $k > 2$, we have the following almost sure convergence for $f(p_0,q_0)$ continuous and bounded,*

$$\frac{1}{T}\int_0^T f(p_0(s),q_0(s))\,ds \to \int_{\mathbb{R}^2} f(u)\,d\bar\mu(u)\,. \tag{6.2.3}$$

Proof. Assume sufficiently nice f as usual, i.e. positive, smooth, compactly supported.

Given $\epsilon > 0$. Choose K, K', t_c as in the first subsection. In addition, choose K_1 such that on K_1^c we have $|\alpha\Phi| < \epsilon/(t_c+1)$ and over a time period of length t_c, for $e_0 = |\bar p_0 - \tilde p_0|$ where $\tilde p_0$ and $\bar p_0$ start at the same point as before, and initial conditions in the 1 variables outside of K_1,

$$\sup_{(p_0,q_0)\in K} \mathbb{P}_{(p_0,q_0)}[e_0 > \epsilon/t_c] < \epsilon\,,$$

since the calculation bounding e_0 (Lemma 6.1.5) is dependent on the initial conditions in the 0 variables only through the bounds on H_0 which can easily be seen to uniform over compact sets in the 0 variables. Designate the series of time intervals of length t_c by setting $t_1 = \inf\{t > 0 : (p_0(t),q_0(t))\in K_0,\ (p_1(t),q_1(t))\notin K_1\}, \ldots,$

$t_n = \inf\{t > t_{n-1} + t_c : (p_0(t), q_0(t)) \in K_0, (p_1(t), q_1(t)) \notin K_1\}$. By Lemmas 6.1.4 and 6.1.7, these intervals cover at least $1 - \epsilon$ of T for T sufficiently large and satisfy $\Pi_0 U(t_i) \in K_0$, $\Pi_1 U(t_i) \notin K_1$, for Π_0 the projection onto the 0 variables (the first two coordinates), with a corresponding definition of Π_1. On $1 - \epsilon$ of the probability space, artificially setting $(\bar{p}_0(t_i), \bar{q}_0(t_i)) = (\tilde{p}_0(t_i), \tilde{q}_0(t_i))$, $\tilde{q}_0(t_i) = q_0(t_i)$,

$$\left| \frac{1}{t_c} \int_{t_i}^{t_i+t_c} f(\bar{p}_0(s), \bar{q}_0(s))\, ds - \frac{1}{t_c} \int_{t_i}^{t_i+t_c} f(\tilde{p}_0(s), \tilde{q}_0(s))\, ds \right| < 2\|Df\|_\infty \epsilon \; .$$

The portions of the probability space where this does/does not hold are independent over the intervals $[t_i, t_i + t_c]$, as one can observe by studying the calculation bounding e_0, Lemma 6.1.5; it is entirely dependent on the noise in the interval $[t_i, t_i + t_c]$. Therefore by the law of large numbers for T sufficiently large at a cost of $\epsilon\|f\|_\infty + 4\|Df\|_\infty \epsilon + 2\epsilon\|f\|_\infty$ a.s., we can consider, for n the number of the intervals $[t_i, t_i + t_c]$,

$$\frac{nt_c}{T} \frac{1}{n} \sum_i \frac{1}{t_c} \int_{t_i}^{t_i+t_c} f(\bar{p}_0(s), \bar{q}_0(s))\, ds \; , \tag{6.2.4}$$

instead of the LHS of (6.2.3), where we artificially set

$$(\bar{p}_0(t_i), \bar{q}_0(t_i)) = (\tilde{p}_0(t_i), q_0(t_i)) \; ,$$

and the noise for the bar variables is exactly that from the tilde/original variables over the time period $[t_i, t_i + t_c]$, i.e. $\bar{W}_0 = W_0$. The bar variables obey (6.0.2) over the time period $[t_i, t_i + t_c]$.

We would prefer it if the distributions of the initial points in (6.2.4) were independent distributions given by the measure $\bar{\mu}$ instead of $(\tilde{p}_0(t_i), \tilde{q}_0(t_i))$ so that we could use the strong law of large numbers. We can do something similar though.

We will now use the well known method of proof of the law of large numbers [Ete81]. Consider a set of bounded positive random variables X_i (which means we do not have to bother with the truncation of the random variables), an increasing sequence of σ algebras $\mathcal{F}_i$, such that X_i is $\mathcal{F}_{i+1}$ measurable. Then we have,

Lemma 6.2.2. *With the set up as above,*

$$\liminf_{i\to\infty} \mathbb{E}[X_i|\mathcal{F}_i] \le \liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i \le \limsup_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$\leq \limsup_{i\to\infty} \mathbb{E}[X_i|\mathcal{F}_i] \quad a.s. \ .$$

Proof of lemma. Fix $\alpha > 1, \epsilon > 0$. Let $k_n = \lfloor \alpha^n \rfloor$. We prove the lemma as follows,

$$\sum_{n=1}^{\infty} \mathbb{P}\left[\frac{\left|\sum_{i=1}^{k_n} X_i - \sum_{i=1}^{k_n} \mathbb{E}[X_i|\mathcal{F}_i]\right|}{k_n} > \epsilon\right]$$

$$= \sum_{n=1}^{\infty} \mathbb{P}\left[\frac{\left|\sum_{i=1}^{k_n} X_i - \sum_{i=1}^{k_n} \mathbb{E}[X_i|\mathcal{F}_i]\right|^2}{k_n^2} > \epsilon^2\right] . \tag{6.2.5}$$

We have,

$$\begin{aligned}
&\mathbb{E}\left|\sum_{i=1}^{k_n} X_i - \sum_{i=1}^{k_n} \mathbb{E}[X_i|\mathcal{F}_i]\right|^2 \\
=&\mathbb{E}\sum_{i=1}^{k_n}(X_i - \mathbb{E}[X_i|\mathcal{F}_i])^2 + 2\sum_{i=1}^{k_n}\sum_{j>i}\mathbb{E}\Big[\big(X_i - \mathbb{E}[X_i|\mathcal{F}_i]\big)\big(X_j - \mathbb{E}[X_j|\mathcal{F}_j]\big)\Big] \\
=&\mathbb{E}\sum_{i=1}^{k_n}(X_i - \mathbb{E}[X_i|\mathcal{F}_i])^2 \\
&+ 2\sum_{i=1}^{k_n}\sum_{j>i}\mathbb{E}\Big[\mathbb{E}\Big[\big(X_i - \mathbb{E}[X_i|\mathcal{F}_i]\big)\big(X_j - \mathbb{E}[X_j|\mathcal{F}_j]\big)\big|\mathcal{F}_j\Big]\Big] .
\end{aligned}$$

We have for $j > i$,

$$\begin{aligned}
&\mathbb{E}\Big[\mathbb{E}\Big[\big(X_i - \mathbb{E}[X_i|\mathcal{F}_i]\big)\big(X_j - \mathbb{E}[X_j|\mathcal{F}_j]\big)\big|\mathcal{F}_j\Big]\Big] \\
=&\mathbb{E}\Big[\big(X_i - \mathbb{E}[X_i|\mathcal{F}_i]\big)\big(\mathbb{E}[X_j|\mathcal{F}_j] - \mathbb{E}[X_j|\mathcal{F}_j]\big)\Big] = 0 ,
\end{aligned}$$

and therefore using the Chebychev inequality the RHS of (6.2.5) is bounded above by,

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\sum_{i=1}^{k_n}\big(X_i - \mathbb{E}[X_i|\mathcal{F}_i]\big)^2}{(k_n\epsilon)^2} \leq \sum_{n=1}^{\infty} 4\|X\|_\infty \epsilon^{-2} k_n^{-1} < \infty .$$

Hence by the Borel Cantelli lemma for any $\epsilon > 0$, only finitely many of the events,

$$\frac{\left|\sum_{i=1}^{k_n} X_i - \sum_{i=1}^{k_n} \mathbb{E}[X_i|\mathcal{F}_i]\right|}{k_n} > \epsilon , \tag{6.2.6}$$

are true a.s.. We will now use the lacunary property of the sequence k_n. We have, from (6.2.6) and the positivity of the X_i,

$$\frac{1}{\alpha}(\liminf_{i\to\infty} \mathbb{E}[X_i|\mathcal{F}_i] - \varepsilon) \leq \frac{\sum_{i=1}^{n} X_i}{n} \leq \alpha(\limsup_{i\to\infty} \mathbb{E}[X_i|\mathcal{F}_i] + \varepsilon) ,$$

for any $\varepsilon > 0$, $\alpha > 1$, a.s. for n sufficiently large. This implies,

$$\liminf_{i\to\infty} \mathbb{E}[X_i|\mathcal{F}_i] \leq \liminf_{n\to\infty} \frac{\sum_{i=1}^n X_i}{n} \leq \limsup_{n\to\infty} \frac{\sum_{i=1}^n X_i}{n} \leq \limsup_{i\to\infty} \mathbb{E}[X_i|\mathcal{F}_i] \,,$$

a.s.. □

In (6.2.4) we are now in the situation of Lemma 6.2.2. The X_i are given by,

$$t_c^{-1} \int_{t_i}^{t_i+t_c} f(\bar{p}_0(s), \bar{q}_0(s))\, ds \,.$$

We set $\mathcal{F}_i = \mathcal{F}_{t_i}$. Then by (6.2.2), we have, by Fubini's theorem,

$$\left| \sup_{(\bar{p}_0,\bar{q}_0)\in K'} \mathbb{E}_{(\bar{p}_0,\bar{q}_0)} t_c^{-1} \int_0^{t_c} f(\bar{p}_0(s), \bar{q}_0(s))\, ds - \bar{\mu}(f) \right| < \|f\|_\infty \epsilon + \epsilon \,. \qquad (6.2.7)$$

Therefore we obtain using the Strong Markov property that the value of,

$$\mathbb{E}_{(\bar{p}_0(t_i),\bar{q}_0(t_i))} t_c^{-1} \left[\int_{t_i}^{t_i+t_c} f(\bar{p}_0(s), \bar{q}_0(s))\, ds \right] = \mathbb{E}[X_i|\mathcal{F}_i] \,,$$

is within $\|f\|_\infty \epsilon + \epsilon$ of $\bar{\mu}(f)$ for all i by (6.2.7). Therefore by applying Lemma 6.2.2 we find that (6.2.4) is within $\|f\|_\infty \epsilon + 2\epsilon + 2\|f\|_\infty \epsilon$ of $\bar{\mu}(f)$ a.s. for $\varepsilon < 1/2$ and hence the LHS of (6.2.3) is within $6\|f\|_\infty \epsilon + 2\epsilon + 4\|Df\|_\infty$ of $\bar{\mu}(f)$ a.s. for any $\epsilon > 0$. This proves (6.2.3).

□

6.3 Ergodic Averaging for $k = 2$, $T_1 > \alpha^2 \langle \Phi^2 \rangle$

The two stages involved in ergodic averaging from the case $k > 2$ remain the same. Specifically we show that when the energy of the first oscillator is high we have convergence with high probability of the ergodic average for the tilde variables over a time interval when H_1 is sufficiently large at the beginning of the time interval. In addition we show that a high proportion of the time is spent in the region where H_1 is large.

6.3.1 Ergodic averaging of the 0 variables when H_1 is large

We want to proceed in the same way as for the case $k > 2$ except we now have the problem that $q_0(t) - \tilde{q}_0(t) = O(t)$ and $p_0(t) - \tilde{p}_0(t) = O(1)$, i.e. since the value of $\Phi(p_1, q_1)$ is of the order H_1^0 we don't have the convergence to zero of the supremum of Φ over one period at high energy.

In order to get round this problem we use that the integral over one period of a free oscillator of Φ is zero, the decreasing period of the free oscillator with increasing energy ($H^{\frac{1}{2k}-\frac{1}{2}}$) for $k > 1$ and the proximity, in an appropriate sense at least, between the trajectories of the 1 variables and those of a free oscillator over a short period of time.

Lemma 6.3.1. *Take $k \geq 2$. Assume that we are given $\epsilon > 0$. Denote the trajectory of the free oscillator with initial configuration $(x, y) \in \mathbb{R}^2$ as $(p_{(x,y)}(t), q_{(x,y)}(t))$. The series of stopping times ξ_n are given by $\xi_1 = T_p(H_1(0)), \ldots, \xi_n = \xi_{n-1} + T_p(H_1(\xi_{n-1}))$. Given the trajectory of p_1, q_1, $\epsilon > 0$ started at $p_1(0)$, $q_1(0)$, the sum of errors of p_1, q_1 from the trajectory of a free oscillator p, q over each period, satisfies,*

$$\mathbb{P}_{(p_1(0),q_1(0))}\Bigg[\sum_{\xi_n<t'} \sup_{\xi_n\leq s\leq \xi_{n+1}\wedge t'} \frac{|p_1(s) - p_{(p_1(\xi_n),q_1(\xi_n))}(s)|}{H_1^{\frac{1}{2}}(0)} \\ + \sum_{\xi_n<t'} \sup_{\xi_n\leq s\leq \xi_{n+1}\wedge t'} \frac{|q_1(s) - q_{(p_1(\xi_n),q_1(\xi_n))}(s)|}{H_1^{\frac{1}{2k}}(0)} > 2E'(H_1(0), K, \epsilon)\Bigg] < \epsilon\,,$$

where $E'(y, K, \epsilon) \to 0$ as $y \to \infty$. Note that the probability above is entirely dependent on the noise in $[0, t']$.

Proof. Set $\tilde{p}_1 = p_1 + \alpha\Phi(p_0, q_0)$. Looking at the error in p over one time period on the scale of the energy of the 1 variables to the power $1/2$, $e_n(s) = H_1^{-1/2}(p_1(\xi_n), q_1(\xi_n)) \sup_{\xi_n\leq}$
$p_{(p_1(\xi_n),q_1(\xi_n))}(u)|$, $e_n(s)$ is
bounded above by,

$$\frac{1}{H_1^{1/2}(\xi_n)} \sup_{\xi_n\leq u\leq s\wedge\xi_{n+1}} |p_1(u) - \alpha\Phi(\tilde{p}_1(u), q_1(u)) + \alpha\Phi(\tilde{p}_0(u), q_0(u)) \\ - p_{(p_1(\xi_n),q_1(\xi_n))}(u)| + \frac{1}{H_1^{1/2}(\xi_n)}(2\alpha\|\Phi\|_\infty)\,. \tag{6.3.1}$$

For $\xi_n \leq s \leq \xi_{n+1}$, set $\hat{p}_1 = p_1(s) - \alpha\Phi(\tilde{p}_1(s), q_1(s)) + \alpha\Phi(\tilde{p}_0(s), q_0(s))$, $\hat{q}_1 = q_1(\xi_n) + \int_{\xi_n}^s \hat{p}_1\, dt$ and,

$$\hat{e}_n(s) = H_1^{-1/2}(\xi_n) \sup_{\xi_n\leq u\leq s\wedge\xi_{n+1}\wedge t'} |\hat{p}_1(u) - p_{(p_1(\xi_n),q_1(\xi_n))}(u)|\,.$$

Note that a simple calculation can be used to show that for $H_1(p_1(0), q_1(0))$ (which we will denote by $H_1(0)$) sufficiently large, the probability of H_1 doubling or halving over the time period t is less than $\epsilon/3$. Both of these probabilities dependent on the noise in $[0, t']$. Now using these facts, we will use Gronwall's inequality to bound the sum over n such that $\xi_n < t'$, of the $\hat{e}_n(\xi_{n+1} \wedge t')$. For $s \leq \xi_{n+1}$,

$$\begin{aligned}
\hat{e}_n(s) =& H_1^{-1/2}(\xi_n) \sup_{\xi_n \leq u \leq s \wedge \xi_{n+1} \wedge t'} |\hat{p}_1(u) - p_{(p_1(\xi_n), q_1(\xi_n))}(u)| \\
\leq& H_1^{-1/2}(\xi_n) \int_{\xi_n}^{s} \sup_{\xi_n \leq v \leq s'} |q_1^{2k-1}(v) - q_{(p_1(\xi_n), q_1(\xi_n))}^{2k-1}(v)| \, ds' \\
&+ H_1^{-1/2}(\xi_n) C + H_1^{-1/2}(\xi_n) \int_{\xi_n}^{\xi_n + 1} R \, dt \\
&+ H_1^{-1/2}(\xi_n) \sup_{\xi_n \leq t \leq \xi_{n+1}} \left| \int_{\xi_n}^{t} (R' + 2\gamma T_1) \, dW_1 \right| \\
&+ H_1^{-1/2}(\xi_n) \sup_{\xi_n \leq t \leq \xi_{n+1}} \left| \int_{\xi_n}^{t} R_0 \, dW_0 \right| \\
\leq& 2H_1^{-1/2}(0) \int_{\xi_n}^{s} \sup_{\xi_n \leq v \leq s'} |\hat{q}_1^{2k-1}(v) - q_{(p_1(\xi_n), q_1(\xi_n))}^{2k-1}(v)| \, ds' \\
&+ C' H_1^{1-1/k-1/2+1/2k-1/2+1/2k-1/2}(0) + 2H_1^{-1/2}(0) C \\
&+ 2H_1^{-1/2}(0) \int_{\xi_n}^{\xi_n + 1} R \, dt + 2H_1^{-1/2}(0) \sup_{\xi_n \leq t \leq \xi_n + 1} \left| \int_{\xi_n}^{t} (R' + 2\gamma T_1) \, dW_1 \right| \\
&+ 2H_1^{-1/2}(0) \sup_{\xi_n \leq t \leq \xi_n + 1} \left| \int_{\xi_n}^{t} R_0 \, dW_0 \right| ,
\end{aligned}$$

for appropriate bounded remainder terms R, R_0, R', C', with probability at least $1 - \epsilon/3$ since we have used the bounds in probability of $H_1(s)$ in terms of $H_1(0)$. Now we have that,

$$\begin{aligned}
&\int_{\xi_n}^{s} \sup_{\xi_n \leq v \leq s'} |\hat{q}_1^{2k-1}(v) - q_{(p_1(\xi_n), q_1(\xi_n))}^{2k-1}(v)| \, ds' \\
&\leq 2(2k-1)(2k)^{2-2/k} H_1^{1/2-1/2k}(0) \int_{\xi_n}^{s} \hat{e}_n(u) \, du \, .
\end{aligned}$$

Hence,

$$\begin{aligned}
\hat{e}_n(s) \leq& 4(2k-1)(2k)^{2-2/k} H_1^{-1/2k}(0) \int_{\xi_n}^{s} \hat{e}_n(u) \, du \\
&+ 2H_1^{-1/2}(0) C'' + 2H_1^{-1/2}(0) \int_{\xi_n}^{\xi_n + 1} R \, dt \\
&+ 2H_1^{-1/2}(0) \sup_{\xi_n \leq t \leq \xi_{n+1}} \left| \int_{\xi_n}^{t} (R' + 2\gamma T_1) \, dW_1 \right|
\end{aligned}$$

$$+ 2H_1^{-1/2}(0) \sup_{\xi_n \le t \le \xi_{n+1}} \left| \int_{\xi_n}^t R_0 \, dW_0 \right| .$$

We have,

$$\sum_n \sup_{\xi_n \le t \le \xi_{n+1}} \left| \int_{\xi_n}^t R_0 \, dW_0 \right| \le \sum_n \max_n \sup_{\xi_n \le t \le \xi_{n+1}} \left| \int_{\xi_n}^t R_0 \, dW_0 \right|$$
$$\le 2 \sum_n \sup_{0 \le t \le t' + O(H_1(0))^{1/2k-1/2}} \left| \int_0^t R_0 \, dW_0 \right| .$$

Therefore summing over n, dealing with the stochastic integral terms as indicated by the above, then using the Burkholder Davis Gundy inequality, gives, with probability at least $1 - \epsilon$,

$$\sum_n \hat{e}_n(s) \le O(H_1^{-1/2k}(0)) \int_0^{O(H_1^{1/2k-1/2}(0))} \sum_n \hat{e}_n(u) \, du + O(H_1^{-1/2k}(0)) ,$$

which, by Gronwall's inequality gives that

$$\sum_n \hat{e}_n(s) \le O(H_1^{-1/2k}(0)) + \int^{O(H_1^{1/2k-1/2}(0))} O(H_1^{-1/k}(0)) e^{O(H_1^{-1/2}(0))} \, ds$$
$$\le O(H_1^{-1/2k}(0)) ,$$

which completes the result for the velocity variable p noting that the sum of the second terms in (6.3.1) is $O(H_1^{1/2-1/2k-1/2}(0)) = O(H_1^{-1/2k}(0))$. The result for q follows from that for p by noting that we are simultaneously multiplying and dividing the error in p by $O(H_1^{1/2-1/2k})$. □

Then Lemma 6.3.1 can then be applied to give,

Lemma 6.3.2. *For H_1 sufficiently large at the beginning of a time interval of length t' we have,*

$$\sup_{0 \le s \le t'} \int_0^s \Phi(p_1(u), q_1(u)) \, du < \epsilon , \qquad (6.3.2)$$

with probability $1 - \epsilon$ dependent on the noise in $[0, t']$. Where how large $H_1(0)$ has to be is dependent on the length of the time period t'.

Proof. From Lemma 6.3.1 and the bounds on $\inf_{0 \le s \le t'} H_1(s)$ as discussed in the proof of that lemma, we have that on at least $1 - \epsilon$ of the probability space, for $H_1(0)$ sufficiently large, the LHS of (6.3.2) is bounded above in modulus by,

$$2^{1/2-1/2k} \|\Phi\|_\infty H_1^{1/2k-1/2}(0)$$

$$+\, 2^{1/2-1/2k} H_1^{1/2k-1/2}(0) \|D\Phi\|_\infty E'(H_1(0), K, \epsilon) \to 0\,,$$

as $H_1(0) \to \infty$ since the supremum of the errors over one time interval $[\xi_n, \xi_{n+1}]$ is certainly less than their sum. This proves the result. Denote the LHS of the above expression by $E_\Phi(H_1(0))$. □

Now we are in a position to show that the tilde variables can be made as close as desired to the bar variables when the energy is large.

Lemma 6.3.3. *Given $\epsilon > 0$, $t' > 0$. For sufficiently large $H_1(0)$ dependent on the bounds on $(\tilde{p}_0(0), \tilde{q}_0(0))$, we have that*

$$\mathbb{P}\left[\sup_{0\le s\le t'} |\tilde{p}_0 - \bar{p}_0| + \sup_{0\le s\le t'} |\tilde{q}_0 - \bar{q}_0| > \epsilon\right] < \epsilon\,, \tag{6.3.3}$$

where both sets of variables start from the same initial conditions and driving Brownian motion, i.e. $\bar{W}_0 = W_0$, in addition $\tilde{q}_0(0) = q_0(0)$. The probability in (6.3.3) is entirely dependent on the noise in the interval $[0, t']$.

Proof. We prove the statement about p, the one concerning q is simply the integrated version over a fixed time interval. Let $e_0'(s) = \sup_{0\le s'\le s} |\tilde{p}_0(s') - \bar{p}_0(s')|$. Then we have,

$$\sup_{0\le s\le t'} \max\{(2k-1)|\bar{q}_0(s)|^{2k-2}, (2k-1)|\tilde{q}_0(s)|^{2k-2}\} + \alpha + 1\,,$$

is bounded above by K say, with probability at least $1-\epsilon/3$, which combined with Lemma 6.3.2, means, with probability $1 - 2\epsilon/3$, we have,

$$\begin{aligned} e_0'(t) \le& K\int_0^t \int_0^s e_0'(s')\, ds'\, ds + \int_0^t \gamma e_0'\, ds + \left|\int_0^t R_1\, ds\right| + \sup_{0\le s\le t}\left|\int_0^s R_2\, dW_1\right| \\ &+ \alpha \int_0^t K E_\Phi(H_1(0))\, ds + \alpha\left|\int_0^t \gamma\Phi\, ds\right|\,. \end{aligned}$$

The calculation proceeds essentially as in section 1 except we end up with,

$$\alpha\int_0^t K E_\Phi(H_1(0))\, ds + \alpha\gamma E_\Phi(H_1(0))\,, \tag{6.3.4}$$

term both added to and as a prefactor in the integral against the exponential term in the previous calculation. We have that (6.3.4) tends to zero as $H_1(0)$ increases, which implies the result. □

The above lemma gives us proximity of the tilde and bar variables as we have for the case $k > 2$.

Now use this result by showing the ergodic averages involving the bar variables and the ordinary variables exhibit a familiar relationship (if you are familiar with the averaging over a fast variable paradigm of homogenization).

Lemma 6.3.4. *Given $\epsilon > 0$, $t' > 0$, initial condition $(p_0(0), q_0(0))$. Consider the bar variables with initial condition $(\tilde{p}_0(0), \tilde{q}_0(0))$ as in Lemma 6.3.3. We have, for continuous bounded f, with probability greater than $1 - \epsilon$ (dependent on the noise in the interval $[0, t']$), for $H_1(0)$ sufficiently large, dependent on the bounds on the initial condition $(p_0(0), q_0(0))$,*

$$\frac{1}{t'}\Bigg| \int_0^{t'} f(p_0(s), q_0(s))\, ds - \int_0^{t'} \frac{1}{T_p(1)} \int_0^{T_p(1)} f(\bar{p}_0(s) - \alpha\Phi(p(u), q(u)), \bar{q}_0(s))\, du\, ds\Bigg| < \epsilon\,, \quad (6.3.5)$$

where $T_p(1)$ is the time taken for one period of the free oscillator at energy 1 and p, q are the corresponding trajectories.

Proof. Again we approximate f by smooth compactly supported f as the energy of the 0 variables and tilde variables are less than K with high probability, $1 - \epsilon/5$ (K dependent on the bounds on the initial condition $(p_0(0), q_0(0))$, ϵ and t). With probability $1 - \epsilon/5$ dependent on the noise in the interval $[0, t']$ we have that $\inf_{0 \le s \le t'} H_1(s) > H_1(0)/2$ for $H_1(0)$ sufficiently large.

We have that,

$$\int_0^{t'} f(p_0(s), q_0(s))\, ds = \int_0^{t'} f(\tilde{p}_0(s) - \alpha\Phi(p_1(s), q_1(s)), q_0(s))\, ds\,, \quad (6.3.6)$$

by definition. We can swap q_0 for $\tilde{q}_0$ in (6.3.6) at a cost of $\|Df\|_\infty E_\Phi(H_1(0))$ with probability at least $1 - \epsilon/5$ dependent on the noise from Lemma 6.3.2. With probability $1 - \epsilon/5$ dependent on the noise in $[0, t']$, we have,

$$\Bigg| \int_0^{t'} f(\tilde{p}_0(s) - \alpha\Phi(p_1(s), q_1(s)), \tilde{q}_0(s))\, ds - \int_0^{t'} f(\bar{p}_0(s) - \alpha\Phi(p_1(s), q_1(s)), \bar{q}_0(s))\, ds\Bigg| < \frac{\epsilon}{4}\,,$$

from Lemma 6.3.3 for $H_1(0)$ sufficiently large. Therefore the expression we will be considering the value of is given by,

$$\frac{1}{t'}\int_0^{t'} f(\bar{p}_0(s) - \alpha\Phi(p_1(s), q_1(s)), \bar{q}_0(s))\, ds\,. \tag{6.3.7}$$

We reuse the stopping times ξ_n from Lemma 6.3.1 here once again. By studying the SDE satisfied by the bar variables, through the tightness of the one measure family W_0 on the space of continuous paths we have, with probability $1 - \epsilon/5$ dependent on W_0 during $[0, t']$, for $H_1(\xi_n)$ sufficiently large (which follows from the observation on $\inf_{0\le s\le t'} H_1(s)$ above),

$$\sup_{n:\,\xi_n<t'}\ \sup_{\xi_n\le s\le t\le \xi_{n+1}} |\bar{p}_0(s) - \bar{p}_0(t)| < \frac{\epsilon}{4\|Df\|_\infty}\,, \tag{6.3.8}$$

since by choosing $H_1(0)$ sufficiently large, we have made the largest interval $[\xi_n, \xi_{n+1})$ as small as required with probability at least $1 - \epsilon/5$. Therefore we can swap $\bar{p}_0$ for the cadlag process $\sum_n 1_{[\xi_n,\xi_{n+1})}(s)\bar{p}_0(\xi_n)$ in (6.3.7) at a cost of $\epsilon/4$. So we are now evaluating,

$$\frac{1}{t'}\int_0^{t'} f\Big(\sum_n 1_{[\xi_n,\xi_{n+1})}(s)\bar{p}_0(\xi_n) - \alpha\Phi(p_1(s), q_1(s)), \bar{q}_0(s)\Big)\, ds\,.$$

Using the result of Lemma 6.3.1 again, this differs from the second term on the LHS of (6.3.5) by,

$$\|Df\|_\infty E_\Phi(H_1(0))/t' + \epsilon/4\,,$$

with probability at least $1 - \epsilon/5$ (the same portion as in the application of Lemma 6.3.2) dependent on the noise in the interval $[0, t']$, the first term of which tends to 0 as $H_1 \to \infty$, thus proving the result. All of the bounds in probability are entirely dependent on the noise in $[0, t']$, hence so is the overall probability of the result holding. □

From Lemma 6.3.4 and the exponential convergence of the bar variables to the invariant measure, we then have the convergence as $H_1(0) \to \infty$, of

$$\frac{1}{T}\,\mathbb{E}_x\left[\int_0^T f(p_0(s), q_0(s))\, ds\right],$$

to

$$\int \frac{1}{T_p(1)} \int_0^{T_p(1)} f(\bar{p}_0 - \alpha\Phi(p(u), q(u)), \bar{q}_0)\, du\, d\bar{\mu} \ ,$$

for a sufficiently long time fixed interval T. Both T and the required size of $H_1(0)$ dependent on the bounds on the initial condition $(p_0(0), q_0(0))$.

Thus we have an equivalent result in this case as for the case $k > 2$.

6.3.2 Putting it all together

Theorem 6.3.5. *For $k = 2$ and $T_1 > \alpha\langle\Phi^2\rangle$, we have the following almost sure convergence for $f(p_0, q_0)$ continuous and bounded,*

$$\frac{1}{T}\int_0^T f(p_0(s), q_0(s))\, ds \to \int_{\mathbb{R}^2} \frac{1}{T_p(1)} \int_0^{T_p(1)} f(u - \alpha\Phi(t))\, dt\, d\mu(u) \ .$$

Proof. This proof is almost identical to that in the case $k > 2$ except that the ergodic average with respect to the normal variables compared to the bar variables involves a corrector term averaged over one period of the isolated oscillator, i.e. we use Lemma 6.3.4 to make the substitution $f(\cdot) \mapsto (T_p(1))^{-1} \int_0^{T_p(1)} f(\cdot - \alpha\Phi(t))\, dt$ in the previous argument. □

6.4 Convergence of H_1 to a squared Bessel process

In this section we will prove,

Theorem 6.4.1. *For $k > 2$, consider,*

$$X(t) = \int_0^t \gamma T_1\, ds + \int_0^t (2\gamma T_1 \langle p^2\rangle)^{\frac{1}{2}} \sqrt{X}(s)\, dW(s) \ , \tag{6.4.1}$$

and for $k = 2$,

$$Y(t) = \int_0^t (\gamma T_1 - \gamma\alpha^2\langle\Phi^2\rangle)\, ds + \int_0^t (2\gamma T_1\langle p^2\rangle)^{\frac{1}{2}} \sqrt{Y}(s)\, dW(s) \ . \tag{6.4.2}$$

Then for $k > 2$ we have the weak convergence,

$$\frac{H_1(T)}{T} \Rightarrow X(1) \ ,$$

and for $k = 2$,

$$\frac{H_1(T)}{T} \Rightarrow Y(1) \ .$$

We are going to prove this theorem by substituting $\tilde{H}_1$ for H_1 and then time changing $\tilde{H}_1$ in order to convert it to a process close to a fixed multiple of the appropriate squared Bessel process in between sets of stopping times where $\tilde{H}_1$ is outside of a compact set. Then we will build a new process close to the multiple of the appropriate squared Bessel process out of these excursions, $\tilde{B^2}$. We will then show that we have a map between $\tilde{H}_1$ and $\tilde{B^2}$ that becomes close to a constant factor time change for large times. Using the reluctance of $\tilde{H}_1$ to occupy a compact set, we will use a pair of comparisons to show that $\tilde{B^2}$ is trapped between the appropriate multiple of two squared Bessel processes with dimension close to the limiting squared Bessel process. At a fixed large time with high probability we find $\tilde{B^2}$ at a large value, well away from the compact set. Then we use the map between $\tilde{H}_1$ and $\tilde{B^2}$ when $\tilde{B^2}$ is outside of a compact set to show that we have a corresponding distribution for $\tilde{H}_1$.

Remark 6.4.2. Notice that we are considering $\tilde{H}_1$ and not $\tilde{\tilde{H}}_1$ now, the reason for this is we no longer need absolute bounds below on the quantity we are considering, we simply want the perturbation of the drift averaged over time to converge to 0.

Remark 6.4.3. As far as identifying the solution to (6.4.1), this is a $(2^{-1}\langle p^2\rangle)^{-1}$ dimensional squared Bessel process accelerated by a time change by factor $2^{-1}\gamma T_1\langle p^2\rangle$. Correspondingly (6.4.2) is a $(\gamma T_1 - \gamma\alpha^2\langle\Phi^2\rangle)(2^{-1}\gamma T_1\langle p^2\rangle)^{-1}$ dimensional squared Bessel process accelerated by a time change by factor $2^{-1}\gamma T_1\langle p^2\rangle$.

Proof of Theorem 6.4.1. Without loss of generality we assume a fixed (non-random) initial condition in $\mathbb{R}^4$. We consider a similar sequence of time changes to T^{ϵ}_t considered earlier, $T^{\epsilon,\tilde{H}_1}_t$. Let $K_2(\epsilon) = (-\infty, k_2(\epsilon)]$ be such that,

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T 1_{K^c_2(\epsilon)}(p_0^2(s) + q_0^2(s))\, ds < \epsilon$$

Consider the series of stopping times $\sigma^{\tilde{H}_1}_n, \phi^{\tilde{H}_1}_n$ given by $\phi^{\tilde{H}_1}_{-1} = 0$, $\sigma^{\tilde{H}_1}_0 = \inf\{t > 0 : \tilde{H}_1(t) > 2k'\}, \ldots, \phi^{\tilde{H}_1}_n = \inf\{t > \sigma^{\tilde{H}_1}_n : \tilde{H}_1 < k'\}$, $\sigma^{\tilde{H}_1}_{n+1} = \inf\{t > \phi^{\tilde{H}_1}_n : \tilde{H}_1 > 2k'\}$ for $k' = k_2(\epsilon) + C$ for $C > 0$, larger than $\tilde{H}_1(0)$, sufficiently large for Lemma 6.4.4 to hold. When $T^{\epsilon,\tilde{H}_1}_t \in [\sigma^{\tilde{H}_1}_n, \phi^{\tilde{H}_1}_n]$ we have similarly to (6.1.16), $T^{\epsilon,\tilde{H}_1,n}_t$ on the interval $[\sigma^{\tilde{H}_1}_n, \phi^{\tilde{H}_1}_n]$ as the solution of, for $k > 2$,

$$T^{\epsilon,n,\tilde{H}_1}_t = \inf\bigg\{ s : \int_{\sigma_n}^{s} \{2\gamma T_0[\alpha\Phi]^2 + 2\gamma T_1[p_1 - \alpha p_0 \partial_p \Phi$$

$$- \alpha^2\Phi\partial_p\Phi]^2\Big\}[\tilde{H}_1(s')]^{-1}\,ds' > t\Big\} \qquad (6.4.3)$$

and for $k = 2$,

$$T_t^{\epsilon,n,\tilde{H}_1} = \inf\Big\{s : \int_{\sigma_n}^{s} \big\{2\gamma T_0[\alpha\Phi + \alpha\gamma\Psi]^2 + 2\gamma T_1[p_1 - \alpha p_0\partial_p\Phi - \alpha^2\Phi\partial_p\Phi + \alpha\partial_p\Psi(q_0|q_0|^{2k-2} + \alpha q_0 - \gamma p_0) + \alpha^2\gamma(\partial_p\Psi\Phi_\Psi\partial_p\Phi) - \alpha^2\gamma\partial_p\Xi]^2\big\}[\tilde{H}_1(s')]^{-1}\,ds' > t\Big\} \qquad (6.4.4)$$

For $k > 2$, using the explicit expression for Φ in the modified polar coordinates from [HM09], we see that the set of zeroes of Φ in $\mathbb{R}^4$ has zero Lebesgue measure (it is only 2 points in the trajectory of the isolated oscillator at energy 1). From the existence of a density with respect to Lebesgue measure we have that the occupation time of the set of zeroes of Φ over any finite time [Hai09] is zero a.s., hence the integral of the quadratic variation of the numerator in (6.4.3) is non-zero over any finite time interval a.s.. In the case $k = 2$, we have an extra $\alpha\gamma\Psi$ for the quadratic variation of W_0 term. The zeroes of $[\Phi + \gamma\Psi]^2$ in $\mathbb{R}^4$ are also of null Lebesgue measure. This can be seen for instance using the polar coordinates of [HM09] and noting that since Φ is zero only on a set of θ of null Lebesgue measure the set of radii for which $\Phi = -\gamma\Psi$ is of non null Lebesgue measure in θ is countable due to the scaling of Ψ (meaning these non null sets Lebesgue sets in θ minus the zeroes of Φ are disjoint). The set of radii for which this occurs is therefore of null Lebesgue measure, which implies that the set of zeroes of $[\Phi + \gamma\Psi]^2$ is of null Lebesgue measure in $\mathbb{R}^4$. This implies that the times changes $T_t^{\epsilon,n,\tilde{H}_1}$ are continuous.

We are going to use the process time changed in these intervals to build a new process. Construct $\tilde{B^2}_\epsilon$ as follows, take the excursions of the process $\tilde{H}_1(T_t^{\epsilon,\tilde{H}_1})$ in the time intervals $[\sigma_n^{\tilde{H}_1}, \phi_n^{\tilde{H}_1}]$ in order, and place them end to end.

Now we need some properties of $\tilde{B^2}_\epsilon$ in order that we can analyze the effect of the drift perturbing it from the process X or Y (dependent on whether $k > 2$ or $k = 2$) using Lemma 6.4.7. When $\tilde{B^2}_\epsilon$ is on an excursion we have an obvious map between the time of the process $\tilde{B^2}_\epsilon$ and that of $\tilde{H}_1$. Define the sets of stopping times $\tilde{\sigma}_n$, $\tilde{\phi}_n = \tilde{\sigma}_{n+1}$ as above but for the process $\tilde{B^2}_\epsilon$. Now we have direct correspondence between $\tilde{H}_1(t)$ in the intervals $[\sigma_n^{\tilde{H}_1}, \phi_n^{\tilde{H}_1}]$ and $\tilde{B^2}_\epsilon$ in the intervals

$[\tilde{\sigma}_n, \tilde{\phi}_n]$ given by time the change $T_t^{\epsilon,n,\tilde{H}_1}$, i.e. to go from paths of $\tilde{H}_1(t)$ to those of $\tilde{B^2}_\epsilon$ we take, $\kappa^\epsilon : \mathbb{R} \to \mathbb{R}$, given by,

$$t \mapsto \tilde{\sigma}_n - \sigma_n^{\tilde{H}_1} + A^{\epsilon,n,\tilde{H}_1}_{t-\sigma_n^{\tilde{H}_1}} ,$$

then, for $t \in [\sigma_n^{\tilde{H}_1}, \phi_n^{\tilde{H}_1}]$, we have,

$$\tilde{H}_1(t) = \tilde{B^2}_\epsilon(\kappa^\epsilon(t)) . \tag{6.4.5}$$

We map all times in the intervals $[\phi_{n-1}^{\tilde{H}_1}, \sigma_n^{\tilde{H}_1}]$ to $\tilde{\sigma}_n$. We will denote the right continuous inverse which is a time change by $F^\epsilon : \mathbb{R} \to \mathbb{R}$. By the same argument as in Lemma 6.1.15 we have the stronger result that,

Lemma 6.4.4. *For T_ϵ sufficiently large, C_3, C_4, constants,*

$$\mathbb{P}\left[\sup_{0<T<T_\epsilon} \left| F^\epsilon(T) - \frac{T}{2\gamma T_1 \langle p^2 \rangle} \right| > C_3 \epsilon T_\epsilon \right] < C_4 \epsilon$$

Consider now the form of the SDE satisfied by $\tilde{B^2}_\epsilon$, for τ_1, $\tau_2 \in [\sigma_n, \sigma_{n+1}]$

$$\begin{aligned}
\tilde{B^2}_\epsilon(\tau_2) - \tilde{B^2}_\epsilon(\tau_1) =& c[F^\epsilon(\tau_2) - F^\epsilon(\tau_1)] \\
&+ \int_{F^\epsilon(\tau_1)}^{F^\epsilon(\tau_2)} \sum_n 1_{[\sigma_n^{\tilde{H}_1}, \phi_n^{\tilde{H}_1}]}(s) E(s)\, ds \\
&+ \int_{\tau_1}^{\tau_2} \sqrt{\tilde{B^2}_\epsilon(s)}\, dW(s) ,
\end{aligned}$$

where c is equal to $1/2\langle p^2 \rangle$ for $k > 2$ and $(\gamma T_1 - \gamma \alpha^2 \langle \Phi^2 \rangle)/2\gamma T_1 \langle p^2 \rangle$ for $k = 2$. $E(s)$ are the error terms in the drift given below.

It will be shown below in Lemma 6.4.6 that,

$$\limsup_{T \to \infty} \frac{1}{T} \int_0^T 1_{[\sigma_n^{\tilde{H}_1}, \phi_n^{\tilde{H}_1}]}(s) |E(s)|\, ds = 0 \quad \text{a.s.} \tag{6.4.6}$$

For sufficiently large t we therefore have, by Lemmas 6.4.4 and 6.4.6 that with probability at least $1 - 2C_4\epsilon$,

$$\frac{1}{t} \int_0^{F^\epsilon(t)} \sum_n 1_{[\sigma_n^{\tilde{H}_1}, \phi_n^{\tilde{H}_1}]}(s) |E(s)|\, ds < \epsilon . \tag{6.4.7}$$

We have, rescaling so that the situation is defined on the time interval $[0, 1]$,

$$\frac{\tilde{B^2}_\epsilon(\tau_2 t)}{t} - \frac{\tilde{B^2}_\epsilon(\tau_1 t)}{t} = c[F^\epsilon(t\tau_2) - F^\epsilon(t\tau_1)]$$

$$+ \int_{F^\epsilon(t\tau_1)}^{F^\epsilon(t\tau_2)} \sum_n 1_{[\sigma_n^{\tilde{H}_1}, \phi_n^{\tilde{H}_1}]}(s) E(s)\, ds$$

$$+ \int_{\tau_1}^{\tau_2} \sqrt{\frac{\tilde{B^2}_\epsilon(ts')}{t}}\, d\tilde{W}(s')$$

for $0 \leq \tau_1 < \tau_2 \leq 1$, $t\tau_1, t\tau_2 \in [\sigma_n, \sigma_{n+1}]$ for some n, $\tilde{W}$ a Brownian motion dependent on t. From Lemma 6.4.4 and (6.4.7) we have that the condition on the finite variation terms (6.4.9) is satisfied with probability $1 - 3C_4\epsilon$ for $\epsilon = \delta^2$, ϵ sufficiently small, t dependent on ϵ sufficiently large. In addition we have that (6.4.8) is satisfied for t sufficiently large dependent on ϵ and the given initial conditions, $\tilde{H}_1(0)$.

So it just remains to verify the lower bounding condition (6.4.10) in order to apply Lemma 6.4.7 to $\tilde{B^2}_\epsilon(st)/t$. So we set $c(\varsigma)$ to be such that for the squared Bessel process of the same dimension (dependent on $k > 2$ and α, $\langle \Phi^2 \rangle$ for $k = 2$) as Z in the comparison (6.1.20) we have (6.4.10) when we replace $c(\varsigma)$ with $2c(\varsigma)$. We will use the following strengthening of Lemma 6.1.15 which holds in the case $k = 2$ also,

Lemma 6.4.5. *Fix $\epsilon > 0$, then for $T_\epsilon > 0$ sufficiently large we have,*

$$\mathbb{P}\left[\sup_{0<T<T_\epsilon} \left| T_T^\epsilon - \frac{2}{\gamma T_1 \langle p^2 \rangle} T \right| > C_1 \epsilon T_\epsilon \right] < C_2' \epsilon$$

where C_1', C_2' constants.

which follows from the proof of Lemma 6.1.15 as it stands. We use this in conjunction with (6.1.20) which holds for $k = 2$ also and gives lower bounds on $\tilde{\tilde{H}}_1$ and therefore since $\tilde{\tilde{H}}_1 \leq \tilde{H}_1$, $\tilde{H}_1$ also. Consider composing the two maps $A^\epsilon_\cdot \circ F^\epsilon(\cdot)$ to map from paths of $\tilde{B^2}_\epsilon$ to those of $\tilde{\tilde{H}}_1(T^\epsilon_\cdot)$ up to a large time T, then this map is within $C_5 T\epsilon$ of $1/4$ times the identity map with probability $C_6\epsilon$ in the same sense as above. Given a sufficiently large time T for this to hold, at least as large as either compact set inside of which we have discontinuities for the time changes $T^\epsilon_\cdot$, F^ϵ divided by $c(\varsigma)/2$. Rescale time and space by $1/T$. Then we end up with another squared Bessel process of the same dimension defined on the time interval $[0, 1]$ (it is really approximately defined on $[0, 1/4]$), $\tilde{Z}$, from the scaling property of squared Bessel processes. For $0 \le t \le 1$ if,

$$\tilde{Z}\left(\frac{A^{\epsilon}_{\cdot} \circ F^{\epsilon}(tT)}{T}\right) > c(\varsigma)$$

then we have $\tilde{B^2}_{\epsilon}(tT)/T > c(\varsigma)$. Noting the error bounds on the map $A^{\epsilon}_{\cdot} \circ F^{\epsilon}(\cdot)$, we use the modulus of continuity of $\tilde{Z}$ on $[0,1]$ to complete the verification of (6.4.10). In particular, if we denote the modulus of continuity of $\tilde{Z}$ on $[0,1]$ as $\omega_{\tilde{Z}}$ and let C_ς be such that $\mathbb{P}[\omega_{\tilde{Z}}(C_\varsigma) < c(\varsigma)] < \varsigma/2$, hence if $\tilde{B^2}_{\epsilon}(tT)/T < c(\varsigma)$ for $C_5\epsilon < C_\varsigma$ we have that $\tilde{B^2}_{\epsilon}(tT)/T < c(\varsigma)$ implies $\tilde{Z}(t/4) < 2c(\varsigma)$. Integrating up with respect to t gives (6.4.10) for T sufficiently large provided $C_6\epsilon < (\varsigma/2)$ and $C_5\epsilon < 3/4$.

Therefore we can apply the result of Lemma 6.4.7 to $\tilde{B^2}_{\epsilon}(tT)/T$, which when combined with (6.4.5) and Lemma 6.4.4 gives the required weak convergence for T sufficiently large. This completes the proof of the result contingent on the proof of the two lemmas.

Now the two lemmas giving the missing results.

Collect together the bad drift terms, for $k > 2$, E is given by,

$$\begin{aligned}
&\alpha\Phi(q_0|q_0|^{2k-2} + \alpha q_0 + \gamma p_0) + \alpha^2 p_0(q_1 - q_0)\partial_p\Phi + \alpha^3\Phi(q_1 - q_0)\partial_p\Phi \\
&\quad - \gamma T_1(\alpha p_0\partial_p^2\Phi + \alpha^2(\partial_p\Phi)^2 + \alpha^2\partial_p^2\Phi\Phi)\ ,
\end{aligned}$$

and for $k = 2$, E is equal to,

$$\begin{aligned}
&\alpha^2 p_0(q_1 - q_0)\partial_p\Phi + \alpha^3(q_1 - q_0)\partial_p\Phi\Phi + \alpha(q_0 - q_1)\partial_p\Psi(q_0|q_0|^{2k-2} + \alpha q_0 + \gamma p_0) \\
&+ \alpha^3\gamma(q_0 - q_1)\partial_p\Phi\Psi + \alpha^3\gamma(q_0 - q_1)\Phi\partial_p\Psi + \alpha^3\gamma(q_1 - q_0)\partial_p\Xi \\
&+ \gamma T_1(-\alpha p_0\partial_p^2\Phi - \alpha^2(\partial_p\Phi)^2 - \alpha^2\partial_p^2\Phi\Phi + \alpha\partial_p^2\Psi(q_0|q_0|^{2k-2} + \alpha q_0 + \gamma p_0) \\
&\quad + \alpha^2\gamma\partial_p^2\Psi\Phi + \alpha^2\gamma\Psi\partial_p^2\Phi + 2\alpha^2\gamma\partial_p\Psi\partial_p\Phi - \alpha^2\gamma\partial_p^2\Xi) \\
&+ \alpha\Psi(p_0|q_0|^{2k-2} + \alpha p_0 - \gamma q_0|q_0|^{2k-2} - \gamma\alpha q_0 - \gamma^2 p_0)\ ,
\end{aligned}$$

these are of the form $f(p_1,q_1)g(p_0,q_0)$ where $f(p_1,q_1) \to 0$ as $|(p_1,q_1)| \to \infty$ and g is a polynomial.

Moving back to the drift terms of the form fg.

Lemma 6.4.6. *The drift terms of the form fg above satisfy,*

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T |fg|(s)\, ds = 0 \quad a.s.\ .$$

Proof of Lemma 6.4.6. Take arbitrary $\epsilon > 0$. First, by Lemma 6.1.4 there exists a compact set $K_0 \subset \mathbb{R}^2$ such that

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T 1_{K_0^c}((p_0(s),q_0(s)))|g|(p_0(s),q_0(s))\,ds < \epsilon \quad \text{a.s.}\ .$$

We designate a compact set $K_1 \subset \mathbb{R}^2$ such that when the 1 variables are outside of K_1, we have that the prefactor in the 1 variables is less than $\epsilon/\|g|_{K_0}\|_\infty$, and we choose T' large enough so that for all larger times, the proportion of time spent in K_1 is less than $\epsilon/\|g|_{K_0}\|_\infty$. We end up with the bound, for a generic term fg, for $T > T'$,

$$\begin{aligned}
&\limsup_{T\to\infty} \frac{1}{T}\int_0^T |f(p_1(s),q_1(s))g(p_0(s),q_0(s))|\,ds \\
\le &\limsup_{T\to\infty} \frac{1}{T}\int_0^T (1_{\mathbb{R}^2\times K_1} + 1_{\mathbb{R}^2\times K_1^c})|f(p_1(s),q_1(s))|1_{K_0\times\mathbb{R}^2}|g(p_0(s),q_0(s))|\,ds \\
&+ \limsup_{T\to\infty} \frac{1}{T}\int_0^T |f(p_1(s),q_1(s))|1_{K_0^c\times\mathbb{R}^2}|g(p_0(s),q_0(s))|\,ds \\
\le &(\|f\|_\infty\epsilon + \epsilon) + \|f\|_\infty\epsilon\ ,
\end{aligned}$$

a.s.. □

Hence we have shown almost surely that the time averages of the modulus of the non-constant drift terms tend to zero.

Now for a lemma which gives the robustness of the square of a Bessel process with respect to the appropriate form of perturbations.

Lemma 6.4.7. *Assume that the cadlag process Y^δ (continuous outside of $[0,\delta]$) satisfies the expression $Y^\delta(t) = a_\delta + S_t + \int_0^t 2\sqrt{|Y(s)|}\,dW_\delta(s)$. Where the starting point of Y^δ, a_δ, satisfies,*

$$a_\delta \le \delta\ . \tag{6.4.8}$$

S_t is a cadlag process of bounded variation, continuous when Y^δ is outside of $[0,\delta]$. S_t satisfies for any $0 \le \tau_1 < \tau_2 \le 1$, Y^δ outside of $[0,\delta]$ in the interval $[\tau_1,\tau_2]$,

$$k(\tau_2-\tau_1) - \delta < S_{\tau_2} - S_{\tau_1} < k(\tau_2-\tau_1) + \delta\ , \tag{6.4.9}$$

Assume also that we have uniform lower bounds over the perturbed processes, i.e. that given $\varsigma > 0$ there exists $\delta' > 0$ and a small set $c(\varsigma)$ such that with probability at least

$1-\varsigma$, for $\delta < \delta'$,

$$\int_0^2 1_{c(\varsigma)}(Y^\delta(s))\, ds < \varsigma \, . \tag{6.4.10}$$

Set $X_{\pm,\varsigma}(t) = k_{\pm,\varsigma}t + \int_0^t 2\sqrt{|X(s)|}\, dW_{\pm,\varsigma,\delta}(s)$. Then for $\delta < \varsigma$ sufficiently small, we have, with probability at least $1-\varsigma$, for $k-\varsigma \le k_{-,\varsigma} < k_{+,\varsigma} \le k+\varsigma$, $|t_\pm(\varsigma)| < 2\varsigma$, and $t(\delta)$ random with $|t(\delta)| < \delta$ a.s.,

$$X_{-,\varsigma}(1-t_-(\varsigma)) \le Y^\delta(1+t(\delta)) \le X^{+,\varsigma}(1+t_+(\varsigma)) \, .$$

which in turn implies the weak convergence as $\delta \to 0$,

$$Y^\delta(1+t(\delta)) \Rightarrow X(1) \, ,$$

for X given by $X(t) = kt + \int_0^t 2\sqrt{X(s)}\, ds$.

Proof. In order to prove this we will conduct a comparison when the larger process (in each comparison) is on an excursion from a small compact set.

We will conduct the comparison between Y^δ and $X_{+,\varsigma}$ in detail, the comparison between Y^δ and $X_{-,\varsigma}$ is similar.

Given $\varsigma > 0$, set $k_+ = k+\varsigma$, and let $c > 0$ be sufficiently small such that with probability at least $1-\varsigma/4$, the time spent by $X_{+,\varsigma}$ in the set $[0,2c]$ up to time 2, is less than ς (cf Lemma 6.1.10, for squared Bessel processes dimensions of larger than 2 the required part of this result holds since the process is then transient).

Construct a cadlag process $X^{\tilde{+},\varsigma}(t)$ as follows, begin with initial point $2c$ and let $\tilde{\phi}_1 = \inf\{t > 0 : X_0^{\tilde{+},\varsigma}(t) < c\}$ for $X_0^{\tilde{+},\varsigma}(t) = 2c + k_+t + \int_0^t \sqrt{X_0^{\tilde{+},\varsigma}(s)}\, dW_\delta(s)$. In the time interval $[0,\phi_1)$ we take the path of $X^{\tilde{+},\varsigma}(t)$ the same as $X_0^{\tilde{+},\varsigma}(t)$. The inductive step is then given by setting $\tilde{\phi}_n = \inf\{t > \tilde{\phi}_{n-1} : X_{n-1}^{\tilde{+},\varsigma}(t) < c\}$ for

$$X_{n-1}^{\tilde{+},\varsigma}(\tilde{\phi}_{n-1}+t) = 2c + k_+t + \int_0^t \sqrt{X_{n-1}^{\tilde{+},\varsigma}(\tilde{\phi}_{n-1}+s)}\, dW_\delta(\tilde{\phi}_{n-1}+s) \, ,$$

and in the time interval $[\tilde{\phi}_{n-1}, \tilde{\phi}_n)$ by taking the path of $X^{\tilde{+},\varsigma}(t)$ the same as that of $X_{n-1}^{\tilde{+},\varsigma}(t)$. By adding independent squared Bessel bridges between the independent excursions, we can extend the excursion process to a full squared Bessel process of dimension k_+, $X^{+,\varsigma}$. What we will show is that there exists a $\xi > 0$ such that $X^{\tilde{+},\varsigma}(t) > Y^\delta(t) + \xi$. First we choose δ sufficiently small so that the process Y^δ

starts at an initial point somewhere in the interval $[0, c/2]$. To make the comparison easier, we will compare $\sqrt{X^{\tilde{+},\varsigma}}$ and $\sqrt{Y^{\delta}}$ (away from 0). Remembering that at all times $X^{\tilde{+},\varsigma} \geq c$, by choosing ξ sufficiently small, $\xi < c/4$, we can assume that $Y^{\delta} > c/2$ on any portion of path that may be of interest (where Y^{δ} approaches within a distance of ξ of $X^{\tilde{+},\varsigma}$). In addition, until the path of Y^{δ} intersects that of $X^{\tilde{+},\varsigma}$, by bounds in probability $(1 - \varsigma/4$ say) above on squared Bessel processes we have bounds on both $X^{\tilde{+},\varsigma}$ and Y^{δ} above by $C > 0$ and below in a bounded region where the square root function is smooth with probability at least $1 - \varsigma/4$. Noting these bounds it is easy to see that for ξ sufficiently small if Y^{δ} is within 2ξ of $X^{\tilde{+},\varsigma}$, and assuming a drift term of kt instead of the bounded variation term then we would have $\sqrt{Y^{\delta}} - \sqrt{X^{\tilde{+},\varsigma}}$ has a negative drift of at least 2ξ. So we form the series of stopping times $\phi_0^{\xi} = \inf\{t > 0 : X^{\tilde{+},\varsigma}(t) - Y^{\delta}(t) < \xi\}$, $\sigma_0^{\xi} = \inf\{t > 0 : X^{\tilde{+},\varsigma}(t) - Y^{\delta}(t) > 2\xi\}, \ldots, \phi_n^{\xi} = \inf\{t > \sigma_{n-1}^{\xi} : X^{\tilde{+},\varsigma}(t) - Y^{\delta}(t) < \xi\}, \sigma_n^{\xi} = \inf\{t > \phi_n^{\xi} : X^{\tilde{+},\varsigma}(t) - Y^{\delta}(t) > 2\xi\}$. We will of course be studying the interior of the time intervals $[\phi_n^{\xi}, \sigma_n^{\xi}]$ and showing that the paths of Y^{δ} and $X^{\tilde{+},\varsigma}(t)$ never intersect.

If we relax the continuity assumption on the finite variation term, then it is clear that the largest drift is given by a finite variation term with the mass of the associated Stieltjes measure concentrated at the times when we have the largest drift term that it is allowable for the mass to be centered on by the restriction (6.4.9). We have that the situation differs from a mass of δ somewhere in each time segment of length δ/k in that allowable combinations are in bijection with those where the mass can move one time segment to the left or stay put, where moving to the site designated as -1 equates to removal and the mass at -1 is assumed to move to the left. This follows by a simple combinatoric argument after dividing time into segments of length δ/k using the condition (6.4.9) looking at the occupation of segments that are occupied by at least one mass of δ in between two unoccupied time segments. We must have all time segments in between with single occupation except for one time segment which must have a double occupation where we have a mass that stays put followed by one that moves one space to the left. Choose δ such that $2\delta/k < \omega(c^{3/2}\xi/\sqrt{2}, \varsigma/4)$ for ω bounds on the modulus of continuity of the family of squared Bessel processes with dimensions in the range $[k/2, 3k/2]$ with probability at least $1 - \varsigma/4$. Tightness of this family of squared Bessel processes follows for instance from Prohorov's theorem combined with pathwise convergence that occurs when we have convergence of dimension [RY91, Chapter IX section 3]. A simple contradiction argument shows that for δ sufficiently small in the time intervals $[\phi_i^{\xi} \wedge (1+\varsigma), \sigma_i^{\xi} \wedge (1+\varsigma)]$, $X^{\tilde{+},\varsigma}(t) - Y^{\delta}(t) \geq \xi - \sqrt{(2C/c)}\delta > 0$ for δ sufficiently small. This implies that the paths of $X^{+,\varsigma}(t)$ and $Y^{\delta}(t)$ do not cross in the time interval $[0, 1+\varsigma]$, which implies the first part of the inequality.

The second part of the inequality follows from an almost identical argument except that this time we are using the excursions of Y^{δ} from a small set, the required property of which are given by (6.4.10), and the Brownian motion driving the squared Bessel process is constructed by sticking together the Brownian motions that drive these excursions i.e. using an appropriate DDS Brownian motion as before.

The weak convergence follows for instance from the tightness in the space of continuous paths of the family of squared Bessel process with dimensions in the range $[k/2, 3k/2]$ together with bounds above on the probability of $Y^{\delta}(1+t(\delta)) \in [0, b]$ following from the distribution of $X^{+,\varsigma}(1)$ and bounds below following from the distribution of $X^{-,\varsigma}(1)$. □

With the completion of the proof of the lemmas, the proof is now complete.

□

Chapter 7

Concluding remarks

With regard to the homogenization problems, I would like to finish the proof of the homogenization result studied where the period sits at an angle with an irrational tangent to the interface from which it is reflected but it is at present unclear to me how to proceed forward in obtaining the necessary ergodic result to complete the scheme. Even once this has been completed there is the problem of an analogous homogenization problem to that studied in the second chapter where the interface is no longer reflecting but we have two periodic regions either side of an interface region that sits at an angle with an irrational tangent to the period. This would require yet more inventiveness to get convergence of the probabilities of exiting either side of a large neighborhood of the interface.

Originally, before attempting the coupled oscillator problem that became the final chapter of this thesis, I tried (and failed) to prove the conjecture of [Hai09] regarding the existence of the invariant measure of a chain of n oscillators. In particular that the borderline value of k is $2n/(2n-1)$ but there were no convincing ideas for a general scheme as to how to approach such a problem. The correctors as constructed in [HM09] can seemingly no longer be used nor is does there appear to be an obvious analogy to the correctors. It would be interesting to find an approach to resolve this conjecture.

On that note I would like to conclude this thesis.

Bibliography

[AA99] G. ALLAIRE and M. AMAR. Boundary layer tails in periodic homogenization. *ESAIM Control Optim. Calc. Var.* **4**, (1999), 209–243 (electronic).

[ACP03] G. ALLAIRE, Y. CAPDEBOSCQ, and A. PIATNITSKI. Homogenization and localization with an interface. *Indiana Univ. Math. J.* **52**, no. 6, (2003), 1413–1446.

[Ada75] R. A. ADAMS. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[BAČ07] G. BEN AROUS and J. ČERNÝ. Scaling limit for trap models on $\mathbb{Z}^d$. *Ann. Probab.* **35**, no. 6, (2007), 2356–2384.

[BAO03] G. BEN AROUS and H. OWHADI. Multiscale homogenization with bounded ratios and anomalous slow diffusion. *Comm. Pure Appl. Math.* **56**, no. 1, (2003), 80–113. URL `http://dx.doi.org/10.1002/cpa.10053`.

[BC05] R. F. BASS and Z.-Q. CHEN. One-dimensional stochastic differential equations with singular and degenerate coefficients. *Sankhyā* **67**, no. 1, (2005), 19–45.

[BEP09] K. BAHLALI, A. ELOUAFLIN, and E. PARDOUX. Homogenization of semilinear PDEs with discontinuous averaged coefficients. *Electron. J. Probab.* **14**, (2009), no. 18, 477–499.

[Bil99] P. BILLINGSLEY. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second ed., 1999. URL `http://dx.doi.org/10.1002/9780470316962`. A Wiley-Interscience Publication.

[BLP78] A. BENSOUSSAN, J.-L. LIONS, and G. PAPANICOLAOU. *Asymptotic analysis of periodic structures*. North-Holland, Amsterdam, 1978.

[BMP05] A. BENCHÉRIF-MADANI and É. PARDOUX. Homogenization of a diffusion with locally periodic coefficients. In *Séminaire de Probabilités XXXVIII*, vol. 1857 of *Lecture Notes in Math.*, 363–392. Springer, Berlin, 2005.

[BMP07] A. BENCHÉRIF-MADANI and É. PARDOUX. Homogenization of a semilinear parabolic PDE with locally periodic coefficients: a probabilistic approach. *ESAIM Probab. Stat.* **11**, (2007), 385–411 (electronic). URL `http://dx.doi.org/10.1051/ps:2007026`.

[Bog07] V. I. BOGACHEV. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.

[BP87] R. F. BASS and É. PARDOUX. Uniqueness for diffusions with piecewise constant coefficients. *Probab. Theory Related Fields* **76**, no. 4, (1987), 557–572.

[BS96] A. N. BORODIN and P. SALMINEN. *Handbook of Brownian motion—facts and formulae*. Probability and its Applications. Birkhäuser Verlag, Basel, 1996.

[Chu10] D. CHUNG. *Composite Materials: Functional Materials for Modern Technologies*. Science and Applications. Springer, New York, 2010. Second Addition.

[CTN00] P. CHUNG, K. TAMMA, and R. NAMBURU. A micro/macro homogenization approach for viscoelastic creep analysis with dissipative correctors for heterogeneous woven-fabric layered media. *Composites Science and Technology* **60**, no. 12-13, (2000), 2233–2253.

[CTN01] P. CHUNG, K. TAMMA, and R. NAMBURU. Asymptotic expansion homogenization for hetergeneous media: computational issues and applications. *Composites Part A: Applied Science and Manufacturing* **32**, no. 9, (2001), 1291–1301.

[Der54] C. DERMAN. Ergodic property of the Brownian motion process. *Proc. Nat. Acad. Sci. U. S. A.* **40**, (1954), 1155–1158.

[DM83] C. DELLACHERIE and P.-A. MEYER. *Probabilités et potentiel. Chapitres IX à XI*. Publications de l'Institut de Mathématiques de l'Université de Strasbourg [Publications of the Mathematical Institute of the University of Strasbourg], XVIII. Hermann, Paris, revised ed., 1983.

[DO06] A. DIAKHABY and Y. OUKNINE. Locally periodic homogenization of reflected diffusion. *J. Appl. Math. Stoch. Anal.* Art. ID 37643, 17.

[DPZ96] G. DA PRATO and J. ZABCZYK. *Ergodicity for infinite-dimensional systems*, vol. 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.

[EK86] S. N. ETHIER and T. G. KURTZ. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[Ete81] N. ETEMADI. An elementary proof of the strong law of large numbers. *Z. Wahrsch. Verw. Gebiete* **55**, no. 1, (1981), 119–122. URL `http://dx.doi.org/10.1007/BF01013465`.

[FW93] M. I. FREIDLIN and A. D. WENTZELL. Diffusion processes on graphs and the averaging principle. *Ann. Probab.* **21**, no. 4, (1993), 2215–2245.

[FW06] M. I. FREIDLIN and A. D. WENTZELL. Long-time behavior of weakly coupled oscillators. *J. Stat. Phys.* **123**, no. 6, (2006), 1311–1337.

[GVM08] D. GŔARD-VARET and N. MASMOUDI. Homogenization in polygonal domains, 2008. Preprint.

[Hai07] M. HAIRER. Ergodic properties for a class of non-Markovian processes, 2007. Lecture notes.

[Hai09] M. HAIRER. How hot can a heat bath get? *Comm. Math. Phys.* **292**, no. 1, (2009), 131–177. URL http://dx.doi.org/10.1007/s00220-009-0857-6.

[Has60] R. Z. HAS′MINSKIĬ. Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations. *Teor. Verojatnost. i Primenen.* **5**, (1960), 196–214.

[HK98] B. M. HAMBLY and T. KUMAGAI. Heat kernel estimates and homogenization for asymptotically lower-dimensional processes on some nested fractals. *Potential Anal.* **8**, no. 4, (1998), 359–397.

[HM09] M. HAIRER and J. C. MATTINGLY. Slow energy dissipation in anharmonic oscillator chains. *Comm. Pure Appl. Math.* **62**, no. 8, (2009), 999–1032. URL http://dx.doi.org/10.1002/cpa.20280.

[HM10a] M. HAIRER and C. MANSON. Periodic homogenization with an interface: the multi-dimensional case, 2010. To appear in Annals of Probability.

[HM10b] M. HAIRER and C. MANSON. Periodic homogenization with an interface: the one-dimensional case, 2010. To appear in Stoch. Proc. Appl.

[Hör61] L. HÖRMANDER. Hypoelliptic differential operators. *Ann. Inst. Fourier Grenoble* **11**, (1961), 477–492, XVI.

[HP08] M. HAIRER and E. PARDOUX. Homogenization of periodic linear degenerate PDEs. *J. Funct. Anal.* **255**, no. 9, (2008), 2462–2487. URL http://dx.doi.org/10.1016/j.jfa.2008.04.014.

[HS81] J. M. HARRISON and L. A. SHEPP. On skew Brownian motion. *Ann. Probab.* **9**, no. 2, (1981), 309–313.

[IM65] K. ITÔ and H. P. MCKEAN, JR. *Diffusion processes and their sample paths*. Die Grundlehren der Mathematischen Wissenschaften, Band 125. Academic Press Inc., Publishers, New York, 1965.

[KG02] A. KALAMKAROV and A. GEORGIADES. Micromechanical modelling of smart composite structures. *Journal of Smart Materials and Strcutures* **11**, no. 3, (2002), 423–434.

[KHGS09] A. KALAMKAROV, E. HASSAN, A. GEORGIADES, and M. SAVI. Asymptotic homogenization model for 3d grid-reinforced composite structures with generally orthotropic reinforcements. *Composite Structures* **89**, no. 2, (2009), 186–196.

[KK96] T. KUMAGAI and S. KUSUOKA. Homogenization on nested fractals. *Probab. Theory Related Fields* **104**, no. 3, (1996), 375–398.

[KK01] R. KHASMINSKII and N. KRYLOV. On averaging principle for diffusion processes with null-recurrent fast component. *Stochastic Process. Appl.* **93**, no. 2, (2001), 229–240.

[Kon45] W. KONDRACHOV. Sur certaines propriétés des fonctions dans l'espace. *C. R. (Doklady) Acad. Sci. URSS (N. S.)* **48**, (1945), 535–538.

[Koz93] S. M. KOZLOV. Harmonization and homogenization on fractals. *Comm. Math. Phys.* **153**, no. 2, (1993), 339–357.

[KPS] T. KOMOROWSKI, S. PESZAT, and T. SZAREK. On ergodicity of some markov processes. *Annals of Probability* .

[KS07] L. B. KORALOV and Y. G. SINAI. *Theory of probability and random processes*. Universitext. Springer, Berlin, second ed., 2007.

[Kum04] T. KUMAGAI. Homogenization on finitely ramified fractals. In *Stochastic analysis and related topics in Kyoto*, vol. 41 of *Adv. Stud. Pure Math.*, 189–207. Math. Soc. Japan, Tokyo, 2004.

[Lej06] A. LEJAY. On the constructions of the skew Brownian motion. *Probab. Surv.* **3**, (2006), 413–466 (electronic).

[LPW95] D. LUKKASSEN, L. PERSSON, and P. WALL. Some engineering and mathematical aspects on the homogenization method. *Composites Engineering* **5**, no. 5, (1995), 519–531.

[LS84] P.-L. LIONS and A.-S. SZNITMAN. Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.* **37**, no. 4, (1984), 511–537. URL http://dx.doi.org/10.1002/cpa.3160370408.

[LS06] A. LASOTA and T. SZAREK. Lower bound technique in the theory of a stochastic differential equation. *J. Differential Equations* **231**, no. 2, (2006), 513–533. URL http://dx.doi.org/10.1016/j.jde.2006.04.018.

[Mey63] N. G. MEYERS. An L^pe-estimate for the gradient of solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa (3)* **17**, (1963), 189–206.

[MGV08] N. MASMOUDI and D. GERARD-VARET. Homogenization in polygonal domains, 2008. URL http://www.math.jussieu.fr/ gerard-varet/GeMa2.pdf. To appear in J. Eur. Math.

[MT93] S. P. MEYN and R. L. TWEEDIE. *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1993.

[MTSM98] R. MOHAN, K. TAMMA, D. SHIRES, and A. MARK. Advanced manufacturing of large-scale composite structures: process modeling, manufacturing simulations and massively parallel computing platforms. *Advances in Engineering Software* **29**, no. 3-6, (1998), 249–263.

[MV97] S. MOSKOW and M. VOGELIUS. First-order corrections to the homogenised eigenvalues of a periodic composite medium. A convergence proof. *Proc. Roy. Soc. Edinburgh Sect. A* **127**, no. 6, (1997), 1263–1299.

[NR01] M. NEUSS-RADU. The boundary behavior of a composite material. *Mathematical Modelling and Numerical Analysis* **35**, no. 3, (2001), 406–435.

[Oll94] S. OLLA. Cours de l'école doctorale, 1994. École Polytechnique.

[OS04] S. OLLA and P. SIRI. Homogenization of a bond diffusion in a locally ergodic random environment. *Stochastic Process. Appl.* **109**, no. 2, (2004), 317–326.

[Owh03] H. OWHADI. Anomalous slow diffusion from perpetual homogenization. *Ann. Probab.* **31**, no. 4, (2003), 1935–1969.

[Par99] É. PARDOUX. Homogenization of linear and semilinear second order parabolic PDEs with periodic coefficients: a probabilistic approach. *J. Funct. Anal.* **167**, no. 2, (1999), 498–520. URL http://dx.doi.org/10.1006/jfan.1999.3441.

[PS08] G. A. PAVLIOTIS and A. M. STUART. *Multiscale methods*, vol. 53 of *Texts in Applied Mathematics*. Springer, New York, 2008. Averaging and homogenization.

[PV81] G. C. PAPANICOLAOU and S. R. S. VARADHAN. Boundary value problems with rapidly oscillating random coefficients. In *Random fields, Vol. I, II (Esztergom, 1979)*, vol. 27 of *Colloq. Math. Soc. János Bolyai*, 835–873. North-Holland, Amsterdam, 1981.

[PV82] G. C. PAPANICOLAOU and S. R. S. VARADHAN. Diffusions with random coefficients. In *Statistics and probability: essays in honor of C. R. Rao*, 547–552. North-Holland, Amsterdam, 1982.

[PV01] E. PARDOUX and A. Y. VERETENNIKOV. On the Poisson equation and diffusion approximation. I. *Ann. Probab.* **29**, no. 3, (2001), 1061–1085. URL http://dx.doi.org/10.1214/aop/1015345596.

[PV05] E. PARDOUX and A. Y. VERETENNIKOV. On the Poisson equation and diffusion approximation. III. *Ann. Probab.* **33**, no. 3, (2005), 1111–1133. URL http://dx.doi.org/10.1214/009117905000000062.

[Rho09a] R. RHODES. Diffusion in a locally stationary random environment. *Probab. Theory Related Fields* **143**, no. 3-4, (2009), 545–568.

[Rho09b] R. RHODES. Homogenization of locally stationary diffusions with possibly degenerate diffusion matrix. *Ann. Inst. Henri Poincaré Probab. Stat.* **45**, no. 4, (2009), 981–1001. URL http://dx.doi.org/10.1214/08-AIHP190.

[Rho09c] R. RHODES. Stochastic homogenization of reflected diffusion processes, 2009. URL http://hal.archives-ouvertes.fr/hal-00361800/en/. Online.

[RY91] D. REVUZ and M. YOR. *Continuous martingales and Brownian motion*, vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1991.

[Sei01] J. SEIDLER. A note on the strong Feller property, 2001. Unpublished lecture notes.

[SV69] D. W. STROOCK and S. R. S. VARADHAN. Diffusion processes with continuous coefficients. I. *Comm. Pure Appl. Math.* **22**, (1969), 345–400.

[SV79] D. W. STROOCK and S. R. S. VARADHAN. *Multidimensional diffusion processes*, vol. 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979.

[Tan84] H. TANAKA. Homogenization of diffusion processes with boundary conditions. In *Stochastic analysis and applications*, vol. 7 of *Adv. Probab. Related Topics*, 411–437. Dekker, New York, 1984.

[Tao09] T. TAO. *Poincaré's legacies, pages from year two of a mathematical blog. Part I*. American Mathematical Society, Providence, RI, 2009.

[Vil03] C. VILLANI. *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[Vos97] J. VOSS. Über die asymptotik des bayesrisikos bei diffusionsprozessen, 1997. Diplomarbeit Universität Kaiserslautern.

[Zhi95] V. V. ZHIKOV. Connectedness and homogenization. Examples of the fractal conductivity. In *Homogenization and applications to material sciences (Nice, 1995)*, vol. 9 of *GAKUTO Internat. Ser. Math. Sci. Appl.*, 421–430. Gakkōtosho, Tokyo, 1995.