Available online at www.sciencedirect.com
ScienceDirect
Journal of Functional Analysis 255 (2008) 2462–2487
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

Homogenization of periodic linear degenerate PDEs

Martin Hairer[a,*], Etienne Pardoux[b]

[a] *Department of Mathematics, The University of Warwick, Coventry CV4 7AL, United Kingdom*
[b] *Laboratoire d'Analyse, Topologie, Probabilités, Université de Provence, F-13453 Marseille, France*

Received 10 July 2007; accepted 23 April 2008

Communicated by C. Villani

Dedicated to Paul Malliavin

Abstract

It is well known under the name of 'periodic homogenization' that, under a centering condition of the drift, a periodic diffusion process on $\mathbf{R}^d$ converges, under diffusive rescaling, to a d-dimensional Brownian motion. Existing proofs of this result all rely on uniform ellipticity or hypoellipticity assumptions on the diffusion. In this paper, we considerably weaken these assumptions in order to allow for the diffusion coefficient to even vanish on an open set. As a consequence, it is no longer the case that the effective diffusivity matrix is necessarily non-degenerate. It turns out that, provided that some very weak regularity conditions are met, the range of the effective diffusivity matrix can be read off the shape of the support of the invariant measure for the periodic diffusion. In particular, this gives some easily verifiable conditions for the effective diffusivity matrix to be of full rank. We also discuss the application of our results to the homogenization of a class of elliptic and parabolic PDEs.

Keywords: Homogenization; Effective diffusivity; Malliavin calculus; Spectral gap; Degenerate

* Corresponding author.
E-mail addresses: M.Hairer@Warwick.ac.uk (M. Hairer), pardoux@cmi.univ-mrs.fr (E. Pardoux).

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2008.04.014

1. Introduction

Our goal is to study, by a probabilistic method, the limit as $\varepsilon \to 0$ of the solution $u^\varepsilon(t,x)$ of an elliptic PDE in the regular bounded domain $D \subset \mathbf{R}^d$

$$\begin{cases} \mathcal{L}_\varepsilon u^\varepsilon(x) + f\left(x, \dfrac{x}{\varepsilon}\right) u^\varepsilon(x) = 0, & x \in D, \\ u^\varepsilon(x) = g(x), \quad x \in \partial D, \end{cases} \tag{1.1}$$

where f is bounded from above, and g is continuous, as well as the limit of $u^\varepsilon(t,x)$, the solution of a parabolic PDE of the form

$$\begin{cases} \dfrac{\partial u^\varepsilon(t,x)}{\partial t} = \mathcal{L}_\varepsilon u^\varepsilon(t,x) + \left(\dfrac{1}{\varepsilon} e\left(\dfrac{x}{\varepsilon}\right) + f\left(x, \dfrac{x}{\varepsilon}\right)\right) u^\varepsilon(t,x), \\ u^\varepsilon(0,x) = g(x), \quad x \in \mathbf{R}^d. \end{cases} \tag{1.2}$$

In both cases, the linear operator $\mathcal{L}_\varepsilon$ is assumed to be a second order differential operator with rapidly oscillating coefficients given by

$$\mathcal{L}_\varepsilon = \frac{1}{2} \sum_{i,j=1}^d a_{ij}\left(\frac{x}{\varepsilon}\right) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d \left[\frac{1}{\varepsilon} b_i\left(\frac{x}{\varepsilon}\right) + c_i\left(\frac{x}{\varepsilon}\right)\right] \frac{\partial}{\partial x_i}.$$

The novelty of our result lies in the fact that we allow the matrix a to degenerate (and even possibly to vanish) in some open subset D of $\mathbf{R}^d$. There is by now quite a vast literature concerning the homogenization of second order elliptic and parabolic PDEs with a possibly degenerating matrix of second order coefficients a, see among others [2,3,6,7,19]. But, as far as we know, in all of these works, either the coefficient a is allowed to degenerate in certain directions only, or else it may vanish on sets of Lebesgue measure zero only. There is quite an extensive literature covering homogenisation of perforated domains [1,5] (which one could interpret as having a vanish inside the perforations), but it seems that the present work is the first to cover situations where the domain on which the matrix a is nondegenerate has no unbounded component. The main technical difficulty that we have to overcome is the lack of regularisation since we do not assume $\mathcal{L}_\varepsilon$ to be hypoelliptic (not even on a set of full measure). However, it turns out that it is possible to show nevertheless that under very weak assumptions, its resolvent maps $\mathcal{C}^1$ into $\mathcal{C}^1$ (see Lemma 2.6), which provides a C^1 solution to certain Poisson equations, and is sufficient to make an approximation argument work (see Lemma 3.2).

Because of the high degree of degeneracy allowed by our approach, it is no longer obvious that the effective diffusivity A of the homogenized operator

$$\mathcal{L}_0 = \frac{1}{2} \sum_{i,j=1}^d A_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d C_i \frac{\partial}{\partial x_i} \tag{1.3}$$

is non-degenerate. We shall therefore also seek to characterize the image of the homogenized diffusion matrix. It turns out that this can be done in terms of the support of the invariant measure of the diffusion process on the torus $\mathbf{T}^d$ with drift b and diffusion matrix a.

The paper is organized as follows. Section 2 contains our assumptions and several important preliminary results. Section 3 presents the homogenization result, in probabilistic terms. Section 4 contains our characterization of the image of the homogenized diffusion matrix A, and Sections 5 and 6 present the application to elliptic and parabolic PDEs. Finally, Section 7 contains a few concrete example that illustrate the scope of the results in this paper and highlight the differences with the existing literature.

2. Assumptions and preliminary results

Given $\varepsilon \geqslant 0$, $x \in \mathbf{R}^d$, let $\{X_t^{x,\varepsilon}\}$ denote the solution of the SDE

$$X_t^{x,\varepsilon} = x + \int_0^t \left[\frac{1}{\varepsilon} b\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) + c\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right)\right] ds + \sum_{j=1}^m \int_0^t \sigma_j\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) dW_s^j, \tag{2.1}$$

where b, c, and σ are periodic, of period one in each direction, and the process $\{W_t = (W_t^1, \dots, W_t^m),\ t \geqslant 0\}$ is a standard m-dimensional Brownian motion.

Define $\tilde{X}_t^{x,\varepsilon} = \frac{1}{\varepsilon} X_{\varepsilon^2 t}^{x,\varepsilon}$. Then there exists a standard m-dimensional Brownian motion $\{W_t\}$, depending on ε (but we forget that dependence since it has no incidence on the law of the process), such that

$$\tilde{X}_t^{x,\varepsilon} = \frac{x}{\varepsilon} + \int_0^t \left[b\left(\tilde{X}_s^{x,\varepsilon}\right) + \varepsilon c\left(\tilde{X}_s^{x,\varepsilon}\right)\right] ds + \sum_{j=1}^m \int_0^t \sigma_j\left(\tilde{X}_s^{x,\varepsilon}\right) dW_s^j. \tag{2.2}$$

In the sequel, we shall consider the solution of (2.2), as taking values in the torus $\mathbf{T}^d$. We will also consider the same equation starting from x, but without the term εc in the drift, namely

$$\tilde{X}_t^x = x + \int_0^t b\left(\tilde{X}_s^x\right) ds + \sum_{j=1}^m \int_0^t \sigma_j\left(\tilde{X}_s^x\right) dW_s^j. \tag{2.3}$$

We denote by J_t^x the Jacobian of the stochastic flow associated to $\tilde{X}_t^x$, that is $\{J_t^x,\ t \geqslant 0\}$, the $d \times d$-matrix valued stochastic process solving

$$dJ_t^x = Db\left(\tilde{X}_t^x\right) J_t^x\, dt + \sum_{j=1}^m D\sigma_j\left(\tilde{X}_t^x\right) J_t^x\, dW_t^j, \qquad J_0^x = I. \tag{2.4}$$

To the SDE satisfied by the process $\{\tilde{X}_\cdot^x\}$, we associate, inspired by Stroock–Varadhan's support theorem, the following controlled ODE (*from now on, we adopt the convention of summation over repeated indices*). For each $x \in \mathbf{T}^d$, $u \in L^2_{\mathrm{loc}}(\mathbf{R}_+; \mathbf{R}^m)$, let $\{z_u^x(t),\ t \geqslant 0\}$ denote the solution of

$$\begin{cases} \dfrac{dz_i}{dt}(t) = b_i\big(z(t)\big) - \dfrac{1}{2}\left[\dfrac{\partial \sigma_{ij}}{\partial x_k}\sigma_{kj}\right]\big(z(t)\big) + \sigma_{ij}\big(z(t)\big) u_j(t), \\ z(0) = x. \end{cases} \tag{2.5}$$

We shall also need the controlled ODE with b replaced by $b + \varepsilon c$, namely we shall denote by $\{z_u^{x,\varepsilon}(t),\ t \geqslant 0\}$ the solution of

$$\begin{cases} \dfrac{dz_i}{dt}(t) = [b_i + \varepsilon c_i]\big(z(t)\big) - \dfrac{1}{2}\left[\dfrac{\partial \sigma_{ij}}{\partial x_k}\sigma_{kj}\right]\big(z(t)\big) + \sigma_{ij}\big(z(t)\big)u_j(t), \\ z(0) = x. \end{cases} \tag{2.6}$$

We will throughout this paper make the following assumptions on the drift and the diffusion coefficient.

Assumption H.1. The functions σ, b, and c are of class $\mathcal{C}^\infty$ and periodic of period 1 in each direction.

Consider now b and the σ_j as vector fields on the torus $\mathbf{T}^d$. We say that the *strong Hörmander condition* holds at some point $x \in \mathbf{T}^d$ if the Lie algebra generated by $\{\sigma_j\}_{j=1}^m$ spans the whole tangent space of $\mathbf{T}^d$ at x. We furthermore say that the *parabolic Hörmander condition* holds at x if the Lie algebra generated by $\{\partial_t + b\} \cup \{\sigma_j\}_{j=1}^m$ spans the whole tangent space of $\mathbf{R} \times \mathbf{T}^d$ at $(0, x)$.

Assumption H.2. There exists a non-empty, open and connected subset U of $\mathbf{T}^d$ on which the strong Hörmander condition holds. Furthermore, there exist $t_0 > 0$ and ε_0 such that, for all $x \in \mathbf{T}^d$, $0 \leqslant \varepsilon \leqslant \varepsilon_0$, one has

$$\inf_{u \in L^2(0,t_0;\mathbf{R}^m)} \left\{ \|u\|_{L^2};\ z_u^{x,\varepsilon}(t_0) \in U \right\} < \infty.$$

Note that, by upper semicontinuity, the supremum over $x \in \mathbf{T}^d$ and $0 \leqslant \varepsilon \leqslant \varepsilon_0$ of the above infimum is bounded by a universal constant K.

Whenever X is a random variable and A an event, we shall use the notation

$$\mathbf{E}(X; A) = \mathbf{E}(X\chi_A) = \mathbf{E}(X \mid A)\mathbf{P}(A).$$

Assumption H.3. One has

$$\inf_{t>0} \sup_{x \in \mathbf{T}^d} \mathbf{E}\left(\left|J_t^x\right|; \left\{\tau_V^x \geqslant t\right\}\right) < 1,$$

where V denotes the subset of $\mathbf{T}^d$ where the parabolic Hörmander condition holds and τ_V^x is the first hitting time of V by the process $\{\tilde{X}_t^x\}$.

Remark 2.1. One simple criteria for this assumption to hold is the existence of a time t_2 such that $\mathbf{P}(\tau_V^x \leqslant t_2) = 1$ for all $x \in \mathbf{T}^d$.

Denote by $p_\varepsilon(t; x, A)$ the transition probabilities of the $\mathbf{T}^d$-valued Markov process $\{\tilde{X}_t^{\varepsilon,x}\}$. (Note that $\varepsilon = 0$ is allowed and corresponds to the process $\{\tilde{X}_t^x\}$ defined in (2.3).)

Lemma 2.2. *Under Assumptions* H.1 *and* H.2, *the following Doeblin condition is satisfied*: *there exist $t_1 > 0$, $0 < \varepsilon_1 \leqslant \varepsilon_0$, $\beta > 0$ and ν a probability measure on $\mathbf{T}^d$ which is absolutely continuous with respect to Lebesgue's measure, such that for all $0 \leqslant \varepsilon \leqslant \varepsilon_1$, $x \in \mathbf{T}^d$, A, Borel subset of $\mathbf{T}^d$,*

$$p_\varepsilon(t_1; x, A) \geqslant \beta \nu(A). \tag{2.7}$$

Proof. It is well known since the works of Malliavin, Bismut, Stroock et al. [4,12–15] that H.1 and H.2 imply that the transition probabilities starting from $x \in U$ have a $\mathcal{C}^\infty$ density with respect to the Lebesgue measure, and that this density is also smooth in the initial condition and in ε. In particular this implies that, for every $t > 0$, $(\varepsilon, x) \mapsto p_\varepsilon(t; x, \cdot)$ is continuous from $[0, 1] \times U$ into the set of probability measures on $\mathbf{T}^d$, equipped with the total variation distance. Let now $x_0 \in U$ be arbitrary. It follows that there exist an open neighbourhood $U_0 \subset U$ of x, $0 < \varepsilon_1 \leqslant \varepsilon_0$ and a probability measure ν such that $p_\varepsilon(t; x, \cdot) \geqslant \frac{1}{2}\nu$ for every $0 \leqslant \varepsilon \leqslant \varepsilon_1$, $x \in U_0$. Since the measures $p_\varepsilon(t; x, \cdot)$ are absolutely continuous, ν must also be absolutely continuous with respect to Lebesgue's measure.

Since $U_0 \subset U$ and since Assumption H.2 implies that (2.5) is locally controllable in U, it follows from the support theorem that, for every $s > t_0$ and every $x \in \mathbf{T}^d$, $0 \leqslant \varepsilon \leqslant \varepsilon_1$, one has $p_\varepsilon(s; x, U_0) > 0$. Since the Markov semigroup generated by the solutions of (2.3) is Feller and since U_0 is open, the function $(\varepsilon, x) \mapsto p_\varepsilon(s; x, U_0)$ is lower semicontinuous and therefore attains its lower bound β' on $[0, \varepsilon_1] \times \mathbf{T}^d$. The claim follows by taking $\beta = \beta'/2$ and $t_1 = s + t$. □

In the sequel we shall denote by $\{\mathcal{P}_{\varepsilon,t},\ t \geqslant 0\}$ the Markov semigroup associated to the $\mathbf{T}^d$-valued diffusion process $\{\tilde{X}_t^{\varepsilon,x},\ t \geqslant 0\}$. We shall also write $\mathcal{P}_t$ for the Markov semigroup associated to $\{\tilde{X}_t^x,\ t \geqslant 0\}$.

Corollary 2.3. *Under Assumptions* H.1 *and* H.2, *for each $0 \leqslant \varepsilon \leqslant \varepsilon_1$, $\{\tilde{X}_t^\varepsilon\}$ possesses a unique invariant probability measure μ_ε, which is absolutely continuous with respect to Lebesgue's measure. Moreover, there exist constants C and $\rho > 0$ such that*

$$\sup_{x \in \mathbf{T}^d,\ 0 \leqslant \varepsilon \leqslant \varepsilon_1} \left\| p_\varepsilon(t; x, \cdot) - \mu_\varepsilon \right\|_{\mathrm{TV}} \leqslant C e^{-\rho t}, \tag{2.8}$$

for every $t \geqslant 0$.

Proof. It follows from (2.7) that for all $0 \leqslant \varepsilon \leqslant \varepsilon_1$, one has

$$\left\| p_\varepsilon(t_1; x, \cdot) - p_\varepsilon(t_1; y, \cdot) \right\|_{\mathrm{TV}} \leqslant 2 - 2\beta,$$

uniformly over $x, y \in \mathbf{T}^d$. Given now any two probability measures μ_1 and μ_2 on $\mathbf{T}^d$, we recall from the definition of the total variation distance that there exists a positive measure ν with mass $\delta := 1 - \frac{1}{2}\|\mu_1 - \mu_2\|$ such that both $\mu_1 - \nu$ and $\mu_2 - \nu$ are positive measures of mass $1 - \delta$. One therefore has the bound

$$\begin{aligned}
\|\mathcal{P}_{\varepsilon,t_1}\mu_1 - \mathcal{P}_{\varepsilon,t_1}\mu_2\|_{\mathrm{TV}} &= \left\| \mathcal{P}_{\varepsilon,t_1}(\mu_1 - \nu) - \mathcal{P}_{\varepsilon,t_1}(\mu_2 - \nu) \right\|_{\mathrm{TV}} \\
&= \left\| \int_{\mathbf{T}^d} p_\varepsilon(t_1; x, \cdot)(\mu_1 - \nu)(dx) - \int_{\mathbf{T}^d} p_\varepsilon(t_1; y, \cdot)(\mu_2 - \nu)(dy) \right\|_{\mathrm{TV}}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{1-\delta}\bigg\| \int_{\mathbf{T}^d}\int_{\mathbf{T}^d} \big(p_\varepsilon(t_1;x,\cdot) - p_\varepsilon(t_1;y,\cdot)\big)(\mu_1-\nu)(dx)(\mu_2-\nu)(dy)\bigg\|_{\mathrm{TV}} \\
&\leqslant \frac{1}{1-\delta}\int_{\mathbf{T}^d} \big\|p_\varepsilon(t_1;x,\cdot) - p_\varepsilon(t_1;y,\cdot)\big\|_{\mathrm{TV}}(\mu_1-\nu)(dx)(\mu_2-\nu)(dy) \\
&\leqslant \frac{1}{1-\delta}(2-2\beta)(1-\delta)^2 = (1-\beta)(2-2\delta) \\
&\leqslant (1-\beta)\|\mu_1-\mu_2\|_{\mathrm{TV}}.
\end{aligned}
\tag{2.9}
$$

This immediately implies the existence and uniqueness of an invariant probability measure μ_ε for $\tilde{X}^\varepsilon_t$.

To see that μ_ε is absolutely continuous with respect to Lebesgue's measure, note that one can decompose the transition semigroup $\mathcal{P}_{\varepsilon,t}$ for $t = t_1$ as

$$
\mathcal{P}_{\varepsilon,t_1} = \beta\nu + (1-\beta)\bar{\mathcal{P}}_\varepsilon \tag{2.10}
$$

for some Markov operator $\bar{\mathcal{P}}_\varepsilon$. It follows from the fact that the solutions of (2.2) generate flows of diffeomorphisms [11], that $\bar{\mathcal{P}}^n_\varepsilon\nu$ is absolutely continuous with respect to the Lebesgue measure for every n. Fix now an arbitrary set A with Lebesgue measure 0. It follows from the invariance of μ_ε that $\mu_\varepsilon = (\mathcal{P}^*_{\varepsilon,t_1})^n\mu_\varepsilon$ for every $n > 0$. From this, (2.10), and the absolute continuity of ν, it follows that $\mu_\varepsilon(A) \leqslant (1-\beta)^n$ for every n, and therefore that $\mu_\varepsilon(A) = 0$.

Finally (2.8) follows from iterating (2.9) with $\mu_2 = \mu_\varepsilon$. $\square$

We shall need the

Lemma 2.4. *Denote by $\mu = \mu_0$ the invariant measure for* (2.3). *As $\varepsilon \to 0$, one has $\mu_\varepsilon \to \mu$ weakly.*

Proof. The tightness is obvious, since $\mathbf{T}^d$ is compact. Hence from any sequence ε_n converging to 0, we can extract a further subsequence, which we still denote by $\{\varepsilon_n\}$, such that

$$
\mu_{\varepsilon_n} \to \tilde{\mu}.
$$

Now if $f \in \mathcal{C}(\mathbf{T}^d)$, $t > 0$, clearly $\mathcal{P}_{\varepsilon_n,t} f(x) \to \mathcal{P}_t f(x)$ as $n \to \infty$, uniformly for $x \in \mathbf{T}^d$, hence

$$
\int_{\mathbf{T}^d} \mathcal{P}_{\varepsilon_n,t} f(x)\,\mu_{\varepsilon_n}(dx) \to \int_{\mathbf{T}^d} \mathcal{P}_t f(x)\,\tilde{\mu}(dx).
$$

But the left-hand side equals

$$
\int_{\mathbf{T}^d} f(x)\,\mu_{\varepsilon_n}(dx) \to \int_{\mathbf{T}^d} f(x)\,\tilde{\mu}(dx).
$$

Consequently $\tilde{\mu}$ is invariant under $\mathcal{P}^*_t$ for all $t > 0$, hence $\tilde{\mu} = \mu$, and $\mu_\varepsilon \to \mu$ weakly, as $\varepsilon \to 0$. $\square$

We can now deduce from (2.8), using also Lemma 2.4, exactly as in [17, Proposition 2.4], the following corollary.

Corollary 2.5. *Whenever $f \in L^\infty(\mathbf{T}^d)$, for any $t > 0$,*

$$\int_0^t f\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) ds \to t \int_{\mathbf{T}^d} f(y)\,\mu(dy)$$

in probability, as $\varepsilon \to 0$.

We finally assume that

Assumption H.4. The drift b satisfies the centering condition $\int_{\mathbf{T}^d} b(x)\,\mu(dx) = 0$.

Denoting by

$$L = \frac{1}{2} a_{ij}(x) \frac{\partial^2}{\partial x_i x_j} + b_i(x) \frac{\partial}{\partial x_i}$$

with $a(x) = \sigma\sigma^*(x)$ the infinitesimal generator of the $\mathbf{T}^d$-valued diffusion process $\{\tilde{X}_t^x, t \geqslant 0\}$, it follows from Lemma 2.6 below that under Assumptions H.1–H.4, there exists a unique $\mathcal{C}^1(\mathbf{T}^d; \mathbf{R}^d)$ solution of the Poisson equation

$$L\hat{b}(x) + b(x) = 0, \quad x \in \mathbf{T}^d.$$

This solution is given by [18]

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

It is not clear a priori that $\hat{b}$ is differentiable, however the following result shows that since b belongs to $\mathcal{C}^1(\mathbf{T}^d)$, so does $\hat{b}$.

Lemma 2.6. *Under Assumption* H.1, *the semigroup $\mathcal{P}_t$ generated by* (2.3) *maps $\mathcal{C}^1(\mathbf{T}^d)$ into itself. If furthermore* H.2 *and* H.3 *hold, then there exist positive constants C and γ such that*

$$\|\mathcal{P}_t f\|_{\mathcal{C}^1(\mathbf{T}^d)} \leqslant Ce^{-\gamma t} \|f\|_{\mathcal{C}^1(\mathbf{T}^d)}, \tag{2.11}$$

for every $f \in \mathcal{C}^1(\mathbf{T}^d)$ such that $\int f(x)\,\mu(dx) = 0$, and for every $t > 0$.

Proof. The proof relies on the techniques developed in [9]. Suppose that we can find a time t_2 such that there exist a constant C and $\delta > 0$ with

$$\|D\mathcal{P}_{t_2} f\|_{L^\infty} \leqslant C\|f\|_{L^\infty} + (1-\delta)\|Df\|_{L^\infty}, \tag{2.12}$$

for every test function $f \in \mathcal{C}^1(\mathbf{T}^d)$. It follows from [9, Section 2] that this, together with the Doeblin condition given by Lemma 2.2, implies that (2.11) holds.

It remains to show that (2.12) does indeed hold. The set V can be defined as

$$V = \left\{x;\ \exists \ell\colon\ Q_\ell(x) := \sum_{Y \in \mathcal{Y}_\ell} Y(x) \otimes Y(x) > 0\right\},$$

where $\mathcal{Y}_\ell$ denotes the set of vector fields consisting of the σ_j's, their brackets and their brackets with b, the number of those being bounded by ℓ. The inequality $Q_\ell(x) > 0$ means that the quadratic form $Q_\ell(x)$ is strictly positive definite. Writing V as a union of compact sets, we see that one can find a sequence ℓ_k such that $V = \bigcup_{k \geqslant 1} V_k$, where V_k is the set of points where $Q_{\ell_k}(x) > k^{-1} I$.

We now want to take advantage of the fact that as soon as the process $\tilde{X}^x$ hits V_k, its Malliavin matrix becomes invertible. However, we face the problem that the hitting time of V_k does not necessarily have a Malliavin derivative. This is the motivation for the next construction. Let Ω and $\bar{\Omega}$ be two independent copies of the m-dimensional Wiener space, and denote by B and $\bar{B}$ the corresponding canonical processes. We first consider the auxiliary SDE

$$dy = b(y)\,dt + \sum_{j=1}^m \sigma_j(y)\,dB_t^j, \qquad y(0) = x,$$

and define $\tilde{\tau}_k^x$ as the hitting time of V_k by the process y. Let $t_2 > 0$ be such that

$$\sup_{x \in \mathbf{T}^d} \mathbf{E}\left(\left|J_{t_2}^x\right|;\ \left\{\tau_V^x \geqslant t_2\right\}\right) < 1, \tag{2.13}$$

and

$$\tau_k^x = \begin{cases} \tilde{\tau}_k^x, & \text{when } \tilde{\tau}_k^x < t_2, \\ +\infty, & \text{otherwise.} \end{cases}$$

We now define a process W by

$$W_t = \begin{cases} B_t & \text{for } t \leqslant \tau_k^x, \\ B_{\tau_k^x} + \bar{B}_{t - \tau_k^x} & \text{for } t \geqslant \tau_k^x. \end{cases}$$

Since τ_k^x is a stopping time, W is again an m-dimensional Wiener process. We can (and will between now and the end of this proof) therefore consider the process $\tilde{X}_t^x$ to be driven by the process W that was just constructed.

Consider (see (2.4)) the $(d \times d)$-matrix valued SDE

$$dJ_t^x = Db\left(\tilde{X}_t^x\right) J_t^x\,dt + \sum_{j=1}^m \left(D\sigma_j\left(\tilde{X}_t^x\right) J_t^x\right) dW_t^j, \qquad J_0 = I.$$

We next define the (random) linear map $A_x : L^2([0,1]; \mathbf{R}^m) \to \mathbf{R}^d$ by

$$A_x u = \sum_{j=1}^{m} \int_{\tau_k^x}^{\tau_k^x+1} \left(J_s^x\right)^{-1} \sigma_j\left(X_s^x\right) u_j\left(s - \tau_k^x\right) ds,$$

with the convention that $A_x = 0$ on the set $\{\tau_k^x = +\infty\}$. Define the (random) map $C_x : \mathbf{R}^d \to \mathbf{R}^d$ by $C_x = A_x A_x^*$. It then follows from [16] and the strong Markov property that

Proposition 2.7. *For every $x \in \mathbf{T}^d$, the matrix C_x is almost surely invertible on the set $\{\tau_k^x < +\infty\}$. Furthermore, for every k and every p there exists a constant $K_{k,p}$ such that $\mathbf{E}(\|C_x^{-p}\| \mid \tau_k^x < \infty) \leqslant K_{k,p}$.*

Fix now an arbitrary vector $\xi \in \mathbf{R}^d$ and define a stochastic process $v : [0, 1] \to \mathbf{R}^m$ by

$$v = A_x^* C_x^{-1} \xi.$$

(We set as before $v = 0$ on the set $\{\tau_k^x = +\infty\}$.) The reason for performing this construction is that we are now going to use Malliavin calculus on the probability space $\Sigma = \Omega \times \bar{\Omega}$ equipped with the canonical Gaussian subspace generated by the Wiener process $\bar{B}$. It is then straightforward to check [16] that one has $v \in D^{1,2}$. Note however that since the stopping time τ_k^x is not Malliavin differentiable in general, it is *not* necessarily true that $v \in D^{1,2}$ if we equip Σ with the Gaussian structure inherited from W.

This construction yields the following equality, valid for every $f \in \mathcal{C}^1(\mathbf{T}^d)$ and for $t = t_2 + 1$:

$$\int_0^t \mathcal{D}_s\left(f\left(\tilde{X}_t^x\right)\right) v_s \, ds = \begin{cases} (Df)(\tilde{X}_t^x) J_t^x \xi & \text{if } \tau_k^x < \infty, \\ 0 & \text{if } \tau_k^x = \infty, \end{cases}$$

where $\mathcal{D}$ stands for the Malliavin derivative with respect to $\bar{B}$. It then follows from the integration by parts formula [16, Section 1.3] that

$$D\mathcal{P}_t f(x)\xi = \mathbf{E}\left(f\left(\tilde{X}_t^x\right) \int_0^1 v(s)\, d\bar{B}s \right) - \mathbf{E}\left((Df)\left(\tilde{X}_t^x\right) J_t^x \xi;\, \tau_k^x = \infty\right),$$

where the integral is to be understood in the Skorokhod sense. It follows from [16, p. 39] that

$$\|D\mathcal{P}_{t_2} f\|_\infty \leqslant \|Df\|_\infty \mathbf{E}\left(\left|J_{t_2}^x\right|; \left\{\tau_k^x = \infty\right\}\right) + \|f\|_\infty \mathbf{E}\left(\int_0^1 \|v_s\|^2\, ds + \int_0^1 \int_0^1 \|\mathcal{D}_s v_r\|^2\, dr\, ds \right).$$

Since, for every k, the second term is bounded uniformly in $x \in \mathbf{T}^d$, it remains to show that we can choose k such that

$$\sup_x \mathbf{E}\left(\left|J_{t_2}^x\right|; \left\{\tau_k^x = \infty\right\}\right) < 1.$$

Assume this is not the case. Then to each k, we can associate a point x_k such that

$$\mathbf{E}\left(\left|J_{t_2}^{x_k}\right|; \left\{\tau_k^{x_k} \geqslant t_2\right\}\right) \geqslant 1.$$

We can assume (up to extracting a subsequence) that $x_k \to x$. For all $\ell \in \mathbf{N}$,

$$1 \leqslant \limsup_{k\to\infty} \mathbf{E}\big(\big|J_{t_2}^{x_k}\big|; \big\{\tau_k^{x_k} \geqslant t_2\big\}\big) \leqslant \limsup_{k\to\infty} \mathbf{E}\big(\big|J_{t_2}^{x_k}\big|; \big\{\tau_\ell^{x_k} \geqslant t_2\big\}\big) \leqslant \mathbf{E}\big(\big|J_{t_2}^{x}\big|; \big\{\tau_\ell^{x} \geqslant t_2\big\}\big),$$

since the mapping $x \to \chi_{\{\tau_\ell^x \geqslant t_2\}}$ is a.s. upper semicontinuous, as the indicator function of a closed subset of trajectories. Since

$$\lim_{\ell\to\infty} \chi_{\{\tau_\ell^x \geqslant t_2\}} = \chi_{\{\tau_V^x \geqslant t_2\}} \quad \text{a.s.},$$

this contradicts (2.13). $\square$

Remark 2.8. We claim that Assumption H.3 is close to being sharp for the conclusion of Lemma 2.6 to hold. The simplest example (which is however not a diffusion on the torus) on which this can be seen is as follows. Let τ be an exponential random variable with parameter $b > 0$, and

$$X_t^x = \begin{cases} e^{at}x, & \text{if } t < \tau, \\ 0, & \text{otherwise.} \end{cases}$$

In this case

$$\mathbf{E}\big(\big|J_t^x\big|; \{\tau > t\}\big) = e^{(a-b)t},$$

and

$$\|D\mathcal{P}_t f\|_\infty = e^{(a-b)t}\|Df\|_\infty,$$

so that $\|\mathcal{P}_t f\|_{\mathcal{C}^1} \to 0$ if and only if H.3 holds.

3. The homogenization result

The goal of this section is to show the

Theorem 3.1. *We have that, in the sense of weak convergence on the space $\mathcal{C}(\mathbf{R}_+)$ equipped with the topology of uniform convergence on compact sets,*

$$X^{x,\varepsilon} \Rightarrow X^x, \quad \text{where } X_t^x = x + Ct + A^{1/2}W_t,$$

as $\varepsilon \to 0$. Here, $\{W_t, t \geqslant 0\}$ is a standard Brownian motion, and the homogenized coefficients C and A are given by

$$C = \int_{\mathbf{T}^d} (I + \nabla\hat{b})c(x)\,\mu(dx),$$

$$A = \int_{\mathbf{T}^d} (I + \nabla\hat{b})a(I + \nabla\hat{b})^*(x)\,\mu(dx).$$

In order to prove this theorem, we need to get rid of the term of order ε^{-1} in the SDE (2.1). The trick is to replace $X_t^{x,\varepsilon}$ by

$$\hat{X}_t^{x,\varepsilon} := X_t^{x,\varepsilon} + \varepsilon \hat{b}\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right),$$

and to take advantage of the

Lemma 3.2. *The following equality holds almost surely*:

$$\hat{X}_t^{x,\varepsilon} = x + \varepsilon \hat{b}\left(\frac{x}{\varepsilon}\right) + \int_0^t (I + \nabla \hat{b}) c\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) ds + \int_0^t (I + \nabla \hat{b}) \sigma\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) dW_s.$$

Proof. Let $\rho : \mathbf{R}^d \to \mathbf{R}_+$ be a smooth function with compact support, such that

$$\int_{\mathbf{R}^d} \rho(x)\, dx = 1,$$

and define $\rho_n(x) = n^d \rho(nx)$. We regularize $\hat{b}$ by convolution

$$\hat{b}_n = \hat{b} * \rho_n.$$

We now deduce from Itô's formula that

$$\begin{aligned}
\hat{X}_t^{\varepsilon,n} &:= X_t^{x,\varepsilon} + \varepsilon \hat{b}_n\left(\frac{X_t^{x,\varepsilon}}{\varepsilon}\right) \\
&= x + \varepsilon \hat{b}_n\left(\frac{x}{\varepsilon}\right) + \varepsilon^{-1} \int_0^t (L\hat{b}_n + b)\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) ds + \int_0^t (I + \nabla \hat{b}_n) c\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) ds \\
&\quad + \int_0^t (I + \nabla \hat{b}_n) \sigma\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) dW_s.
\end{aligned}$$

We want next to let $n \to \infty$. Clearly $\hat{b}_n \to \hat{b}$ and $\nabla \hat{b}_n \to \nabla \hat{b}$ pointwise, and the two sequences are bounded, uniformly with respect to n and $x \in \mathbf{T}^d$. It remains to treat the term containing the second order derivatives. One has

$$L\hat{b}_n = (L\hat{b}) * \rho_n + \varphi_n,$$

and since $\hat{b}$ is a weak solution of the Poisson equation,

$$(L\hat{b}) * \rho_n = -b * \rho_n \to -b,$$

the sequence being again uniformly bounded. It remains to study the sequence φ_n. Using again the convention of summation over repeated indices, we have

$$\varphi_n(x) = \frac{1}{2}\int_{\mathbf{R}^d} \big[a_{ij}(x) - a_{ij}(x-y)\big]\frac{\partial^2 \hat{b}}{\partial x_i \partial x_j}(x-y)n^d\rho(ny)\,dy$$
$$+ \int_{\mathbf{R}^d} \big[b_i(x) - b_i(x-y)\big]\frac{\partial \hat{b}}{\partial x_i}(x-y)n^d\rho(ny)\,dy. \tag{3.1}$$

The fact that the second integral in (3.1) converges to 0 and is uniformly bounded is easily established. We now integrate by parts the first integral, yielding

$$\int_{\mathbf{R}^d} \big[a_{ij}(x) - a_{ij}(x-y)\big]\frac{\partial^2 \hat{b}}{\partial x_i \partial x_j}(x-y)n^d\rho(ny)\,dy$$
$$= \int_{\mathbf{R}^d} \frac{\partial a_{ij}}{\partial x_i}(x-y)\frac{\partial \hat{b}}{\partial x_j}(x-y)n^d\rho(ny)\,dy$$
$$+ \int_{\mathbf{R}^d} \big[a_{ij}(x) - a_{ij}(x-y)\big]\frac{\partial \hat{b}}{\partial x_j}(x-y)n^{d+1}\rho_i'(ny)\,dy.$$

The first term on the right-hand side is uniformly bounded and converges to

$$\frac{\partial a_{ij}}{\partial x_i}(x)\frac{\partial \hat{b}}{\partial x_j}(x).$$

The second term is equal to

$$\int_{\mathbf{R}^d} y \cdot \nabla a_{ij}(x-y')\frac{\partial \hat{b}}{\partial x_j}(x-y)n^{d+1}\rho_i'(ny)\,dy,$$

where $|y'| \leqslant |y|$. This last quantity is equal to a bounded sequence converging to 0, plus

$$\frac{\partial a_{ij}}{\partial x_k}(x)\frac{\partial \hat{b}}{\partial x_j}(x)\int y_k n^{d+1}\rho_i'(ny)\,dy = -\frac{\partial a_{ij}}{\partial x_i}(x)\frac{\partial \hat{b}}{\partial x_j}(x).$$

The lemma is established. □

We now proceed with the

Proof of Theorem 3.1. Recall that, as in the statement of the theorem, we use $\Rightarrow$ to denote weak convergence for laws of processes on the space $\mathcal{C}(\mathbf{R}_+)$ equipped with the topology of uniform convergence on compact sets. We first note that since

$$\big|\hat{X}_t^{x,\varepsilon} - X_t^{x,\varepsilon}\big| \leqslant C\varepsilon,$$

the theorem will follow if we prove that, as $\varepsilon \to 0$,

$$\hat{X}^{x,\varepsilon} \Rightarrow X^x.$$

But since

$$\hat{X}_t^{x,\varepsilon} = x + \int_0^t (I + \nabla\hat{b})c\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) ds + M_t^\varepsilon,$$

one has

$$\langle\!\langle M^\varepsilon \rangle\!\rangle_t = \int_0^t (I + \nabla\hat{b})a(I + \nabla\hat{b})^*\left(\frac{X_s^{x,\varepsilon}}{\varepsilon}\right) ds,$$

and the result follows from Corollary 2.5 and the martingale central limit theorem, see e.g. [8, Theorem 7.1.4]. □

We conclude this section with the following result, which extends the averaging result of Corollary 2.5. It is going to be needed in the applications of Sections 5 and 6.

Proposition 3.3. *Let $f \in \mathcal{C}(\mathbf{R}^d \times \mathbf{T}^d)$. Then for every $t > 0$, the following convergence holds in law as $\varepsilon \to 0$:*

$$\int_0^t f\left(X_s^{x,\varepsilon}, \frac{X_s^{x,\varepsilon}}{\varepsilon}\right) ds \Rightarrow \int_0^t \bar{f}\left(X_s^x\right) ds,$$

where

$$\bar{f}(x) := \int_{\mathbf{T}^d} f(x,y)\,\mu(dy).$$

Proof. It is easily checked that $\bar{f}$ is continuous, hence

$$\int_0^t \bar{f}\left(X_s^{x,\varepsilon}\right) ds \Rightarrow \int_0^t \bar{f}\left(X_s^x\right) ds,$$

as $\varepsilon \to 0$. It then suffices to show that

$$\left| \int_0^t f\left(X_s^{x,\varepsilon}, \frac{X_s^{x,\varepsilon}}{\varepsilon}\right) ds - \int_0^t \bar{f}\left(X_s^{x,\varepsilon}\right) ds \right| \to 0$$

in probability, as $\varepsilon \to 0$. This is proved by exploiting the tightness in $\mathcal{C}(\mathbf{R}_+)$ of the collection of processes $\{X_\cdot^{x,\varepsilon},\ \varepsilon > 0\}$, and Corollary 2.5. For the details, see the proof of Lemma 4.2 in [17] (there is a misprint in the statement of that lemma). □

4. Characterization of the image of the homogenized diffusion matrix

The aim of this section is to give a characterisation of the range of A which can very easily be read off the support of the invariant measure μ. We first state all of our results and then prove them in the following subsection.

4.1. Statements of the results

It will be convenient in our statements to choose an arbitrary particular point $x_0 \in U$. We define the set Γ, consisting of loops on $\mathbf{T}^d$, starting and ending at x_0, which take the form

$$\gamma(s) = z_u^{x_0}(s), \quad 0 \leqslant s \leqslant t,$$

with arbitrary $t > 0$ and $u \in L^2(0, t; \mathbf{R}^m)$. We call loops in Γ *admissible*. It follows from H.2 that the set of admissible loops does not depend on the particular choice of x_0.

We next define a mapping $g : \Gamma \to \mathbf{Z}^d$ as follows. Lifting each loop γ, so as to transform it into a curve $\bar{\gamma}$ in $\mathbf{R}^d$, we define

$$g(\gamma) = \bar{\gamma}(t) - \bar{\gamma}(0),$$

where t is the end time of the loop γ. The first step in our characterization of the image of A is to show that

Theorem 4.1. *Under Assumptions* H.1–H.4*,* $g(\Gamma) \subset \operatorname{Im} A$.

The converse to this result takes the following form:

Theorem 4.2. *Under Assumptions* H.1–H.4*, for each* $e \in \operatorname{Im} A$*, with* $|e| = 1$*, and for every* $\delta > 0$*, there exists a loop* $\gamma \in \Gamma$ *such that*

$$\left| \frac{g(\gamma)}{|g(\gamma)|} - e \right| \leqslant \delta.$$

It follows immediately from the above theorems that under the same assumptions,

Corollary 4.3. $\operatorname{Im} A = \operatorname{span}\{g(\gamma), \gamma \in \Gamma\}$.

Corollary 4.4. *The matrix A is non-degenerate if and only if there exists a collection of admissible loops* $\{\gamma_1, \dots, \gamma_d\} \subset \Gamma$ *such that*

$$\operatorname{span}\bigl\{g(\gamma_1), \dots, g(\gamma_d)\bigr\} = \mathbf{R}^d.$$

The next corollary is slightly less obvious and will be shown in the next subsection. It will mainly be useful in the proof of Theorem 4.6 below, but since it is not a priori an obvious fact, we state it separately.

Corollary 4.5. *The set* $G = \{g(\gamma),\ \gamma \in \Gamma\}$ *is a subgroup of* $\mathbf{Z}^d$.

We denote by S the interior of the support of the invariant measure μ. It follows from H.2 that the support of μ is then the closure of S. It turns out that as far as the characterisation of $\operatorname{Im} A$ is concerned, we can replace admissible loops by arbitrary loops in S. This may sound somewhat surprising at first sight, since it is certainly not true in general that every loop in S is admissible. One has however

Theorem 4.6. *Every admissible loop γ satisfies $\gamma \subset S$. Conversely, for every loop $\rho \subset S$, one can find an admissible loop γ with $g(\rho) = g(\gamma)$.*

Remark 4.7. Combining Theorem 4.6 with Corollary 4.3, we find that $\operatorname{Im} A$ is completely characterised by the topology of the set S. This is however not necessarily true if we drop the regularity conditions H.2–H.3. In particular, it follows from Corollary 4.3 that one can find a basis for $\operatorname{Im} A$ which has rational coordinates. A 'counterexample' (with $d = 2$ and $m = 1$) to this claim is given by taking $b = c = 0$ and $\sigma_1 = (1, \sqrt{2})$, so that $\operatorname{Im} A$ is the vector space generated by $(1, \sqrt{2})$. Of course, this example satisfies neither H.2 nor H.3 and does therefore not contradict our results.

Remark 4.8. The argument in [10, p. 264] shows that A is non-degenerate whenever the set $\operatorname{span}\{g(\gamma),\ \gamma \in S_0\}$ is equal to $\mathbf{R}^d$, where S_0 is a connected subset of the support of the invariant measure μ where a is elliptic. Clearly the result presented here is stronger.

4.2. Proofs of the results

We start with the

Proof of Theorem 4.1. Let $\xi \in \mathbf{R}^d$ be such that $\langle A\xi, \xi\rangle = 0$. We will prove that for any $\gamma \in \Gamma$, $\langle g(\gamma), \xi\rangle = 0$. We shall make use of the notation $b_\xi(x) = \langle b(x), \xi\rangle$, and of the three following facts:

- $\hat{b}_\xi$ solves the Poisson equation
$$L\hat{b}_\xi(x) + b_\xi(x) = 0, \quad x \in \mathbf{T}^d; \tag{4.1}$$
- $\langle A\xi, \xi\rangle = 0$, which, when we explicit the matrix A, amounts to the fact that
$$\langle \nabla \hat{b}_\xi, \sigma_{\cdot j}\rangle + \langle \sigma_{\cdot j}, \xi\rangle = 0, \quad 1 \leqslant j \leqslant d, \quad \text{for } \mu \text{ a.e. } x. \tag{4.2}$$
Note that (4.2) holds true μ-almost everywhere. It can be noted that the trajectories of the solutions of (2.5) starting from the point x_0 remain in the interior of the support of μ, which is absolutely continuous with respect to the Lebesgue measure. Hence (4.2) holds true almost everywhere in the neighbourhood of such a trajectory, consequently everywhere there, due to continuity.
- We can now differentiate (4.2) with respect to x_k, and multiply the resulting identity by σ_{kj}, from which we deduce that
$$\sigma_{kj}(x)\frac{\partial}{\partial x_k}\left(\sigma_{ij}\frac{\partial \hat{b}_\xi}{\partial x_i}\right)(x) = -\xi_i\left(\sigma_{kj}\frac{\partial \sigma_{ij}}{\partial x_k}\right)(x), \quad x \in \mathbf{T}^d. \tag{4.3}$$

Choose a loop $\gamma \in \Gamma$, and denote by t its end time. We now have, for $0 < s < t$,

$$\begin{aligned}
\frac{d}{ds}\hat{b}_\xi\big(\gamma(s)\big) &= \nabla \hat{b}_\xi\big(\gamma(s)\big)\frac{d\gamma}{ds}(s) \\
&= \langle b, \nabla \hat{b}_\xi\rangle\big(\gamma(s)\big) - \frac{1}{2}\left[\frac{\partial \hat{b}_\xi}{\partial x_i}\frac{\partial \sigma_{ij}}{\partial x_k}\sigma_{kj}\right]\big(\gamma(s)\big) + \left[\frac{\partial \hat{b}_\xi}{\partial x_i}\sigma_{ij}\right]\big(\gamma(s)\big)u_j(s) \\
&= -b_\xi\big(\gamma(s)\big) - \frac{1}{2}\left[a_{ij}\frac{\partial^2 \hat{b}_\xi}{\partial x_i \partial x_j} + \sigma_{kj}\frac{\partial \sigma_{ij}}{\partial x_k}\frac{\partial \hat{b}_\xi}{\partial x_i}\right]\big(\gamma(s)\big) - \xi_i \sigma_{ij}\big(\gamma(s)\big)u_j(s) \\
&= -b_\xi\big(\gamma(s)\big) + \frac{1}{2}\xi_i\left[\sigma_{kj}\frac{\partial \sigma_{ij}}{\partial x_k}\right]\big(\gamma(s)\big) - \xi_i \sigma_{ij}\big(\gamma(s)\big)u_j(s) \\
&= -\left\langle \frac{d\gamma}{ds}(s), \xi\right\rangle,
\end{aligned}$$

where we have used (4.1) and (4.2) for the third equality, and (4.3) for the fourth equality. Integrating from $s = 0$ to $s = t$, we deduce that

$$\big\langle g(\gamma), \xi\big\rangle + \big\langle \hat{b}_\xi\big(\bar{\gamma}(t)\big) - \hat{b}_\xi\big(\bar{\gamma}(0)\big), \xi\big\rangle = 0,$$

from which it follows, since $\hat{b}$ is periodic, that $\langle g(\gamma), \xi\rangle = 0$. The result follows. □

Proof of Theorem 4.2. Let $e \in \operatorname{Im} A$ with $|e| = 1$ and $\alpha > 0$. We denote by $B(x, \alpha)$ the ball in $\mathbf{R}^d$ of radius α, centered at x. Note that

$$\mathbf{P}\big(A^{1/2}W_1 \in B(e, \alpha)\big) > 0.$$

It follows that for $\varepsilon > 0$ small enough (here X^ε is defined with $c = 0$),

$$\mathbf{P}\big(X_1^{\varepsilon, \varepsilon x_0} \in B(\varepsilon x_0 + e, \alpha)\big) > 0.$$

Consequently

$$\mathbf{P}\left(\tilde{X}^{x_0}_{1/\varepsilon^2} \in B\left(x_0 + \frac{e}{\varepsilon}, \frac{\alpha}{\varepsilon}\right)\right) > 0.$$

It then follows from Stroock–Varadhan's support theorem that there exists a control $u \in L^2(0, 1/\varepsilon^2; \mathbf{R}^m)$ such that the corresponding z-trajectory, lifted to $\mathbf{R}^d$, satisfies

$$\bar{z}(0) = x_0, \qquad \big|\bar{z}\big(1/\varepsilon^2\big) - (x_0 + e/\varepsilon)\big| \leqslant \alpha/\varepsilon.$$

Now from H.2, there exist $t \leqslant t_0$ and $u \in L^2(\frac{1}{\varepsilon^2}, \frac{1}{\varepsilon^2} + t; \mathbf{R}^m)$, such that $\bar{z}(\frac{1}{\varepsilon^2} + t) = x_0 + g(\gamma)$, where

$$\gamma = \left\{z(s),\ 0 \leqslant s \leqslant \frac{1}{\varepsilon^2} + t\right\} \in \Gamma.$$

We have constructed a loop γ such that

$$\left|g(\gamma) - \frac{e}{\varepsilon}\right| \leqslant \frac{\alpha}{\varepsilon} + C,$$

for some universal constant C (see the remark following Assumption H.2). Multiplying by ε, we get

$$\left|\varepsilon g(\gamma) - e\right| \leqslant \alpha + \varepsilon C,$$

and consequently

$$\left|\frac{g(\gamma)}{|g(\gamma)|} - e\right| \leqslant 2(\alpha + \varepsilon C),$$

from which the result follows, provided we let $\varepsilon \leqslant \delta/4C$ and choose $\alpha = \delta/4$. □

Proof of Corollary 4.5. It follows from Theorem 4.2 that linear combinations of elements of G with positive coefficients are dense in $\operatorname{Im} A$. Since they form a cone, they must therefore coincide with $\operatorname{Im} A$. Denoting the dimension of $\operatorname{Im} A$ by k, it follows from Theorem 4.1 that, for any $g \in G$, there exist $\alpha_1, \ldots, \alpha_k \in \mathbf{R}_+$ and linearly independent elements $g_1, \ldots, g_k \in G$ such that

$$-g = \sum_{i=1}^{k} \alpha_i g_i.$$

In fact, since g and all g_i have integer coordinates and since the inverse of a matrix with integer coefficients has rational coefficients, the α_i's must be rational. Consequently, there exists $n \in \mathbf{N}$ such that $-ng \in G$, from which it follows immediately that $-g \in G$. We have proved that G is a group. □

Proof of Theorem 4.6. The first statement follows from the fact that any admissible trajectory starting at x_0 stays in S.

Conversely, we first note that since S is connected, to each loop $\rho \subset S$, we can associate a loop $\rho' \subset S$ starting from x_0, with $g(\rho') = g(\rho)$. So we might as well assume that ρ starts from x_0. Denote by $\bar{S}$ the lift of S to $\mathbf{R}^d$. It now follows that x_0 and $x_0 + g(\rho)$ belong to the same connected component $\bar{S}'$ of $\bar{S}$. It remains to show that there exists an admissible loop γ such that $\bar{\gamma}$ joins x_0 and $x_0 + g(\rho)$. In other words, it remains to show that $g(\rho) \in G$.

For that sake, first notice that

$$\bar{S} = \bigcup_{k \in \mathbf{Z}^d} C(k),$$

where $C(k)$ denotes the set of accessible points from $x_0 + k$ by the solution of the controlled ODE (2.5) lifted to $\mathbf{R}^d$. Consequently we have that

$$\bar{S}' = \bigcup_{k \in \mathbf{Z}^d,\ x_0 + k \in \bar{S}'} C(k).$$

Recall that $x_0 + g(\rho) \in \bar{S}'$. Since each $C(k)$ is open, there exist $n > 0$ and a chain $f_1, \dots, f_n$ of points in $\mathbf{Z}^d$ such that, with $f_0 = 0$ and $f_{n+1} = g(\rho)$,

$$C(f_i) \cap C(f_{i+1}) \neq \emptyset, \quad 0 \leqslant i \leqslant n.$$

Pick a point $y \in C(f_i) \cap C(f_{i+1})$. There exist a control such that the solution of (2.5) (lifted to $\mathbf{R}^d$) starting from $x_0 + f_i$ reaches y in finite time, and another one such that the solution starting from $x_0 + f_{i+1}$ also reaches y in finite time. It follows from Assumption H.2 that there exists a control such that the solution starting from y reaches some $x_0 + f$ in finite time. Hence $f - f_i, f - f_{i+1} \in G$. Since G is a group, $f_{i+1} - f_i \in G$, for all $0 \leqslant i \leqslant n$. We have proved that $g(\rho) \in G$. $\square$

5. Homogenization of an elliptic PDE

In this section, we show how to apply the homogenization results obtained in the previous sections to the elliptic homogenization problem (1.1). Let D be a bounded domain in $\mathbf{R}^d$ with a $\mathcal{C}^1$ boundary, and define

$$\tau^\varepsilon = \inf\{t \geqslant 0, X_t^\varepsilon \notin \bar{D}\}.$$

Let $\alpha \geqslant 0$ be such that for all $x \in D$,

$$\sup_{\varepsilon>0} \mathbf{E}_x \exp(\alpha \tau^\varepsilon) < \infty. \tag{5.1}$$

We assume that $f \in \mathcal{C}(\mathbf{R}^{2d})$, that it is periodic with respect to its second variable, and that there exists $\delta > 0$ such that

$$f(x, y) \leqslant (\alpha - \delta)^+, \quad \forall x \in \mathbf{R}^d,\ y \in \mathbf{T}^d. \tag{5.2}$$

Remark 5.1. We will need conditions (5.1) and (5.2) in order to deduce some uniform integrability. Condition (5.1) must be checked in each particular example. It is always satisfied with $\alpha = 0$ (unless one is in the case $A = 0$), in which case (5.2) requires that $f(x, y) < 0$, $x \in \mathbf{R}^d$, $y \in \mathbf{T}^d$. This condition can be relaxed only if (5.1) is satisfied with some $\alpha > 0$.

The solution of the elliptic PDE

$$\begin{cases} \mathcal{L}_\varepsilon u^\varepsilon(x) + f\left(x, \dfrac{x}{\varepsilon}\right) u^\varepsilon(x) = 0, x \in D, \\ u^\varepsilon(x) = g(x), x \in \partial D, \end{cases}$$

is then given by the Feynman–Kac formula

$$u^\varepsilon(x) = \mathbf{E}_x\left[g\left(X_{\tau^\varepsilon}^\varepsilon\right) \exp\left(\int_0^{\tau^\varepsilon} f\left(X_s^\varepsilon, \frac{X_s^\varepsilon}{\varepsilon}\right) ds\right)\right].$$

We define as before

$$A = \int_{\mathbf{T}^d} (I + \nabla\hat{b})a(I + \nabla\hat{b})^*(x)\,\mu(dx),$$

$$C = \int_{\mathbf{T}^d} (I + \nabla\hat{b})c(x)\,\mu(dx),$$

$$D(x) = \int_{\mathbf{T}^d} f(x, y)\,\mu(dy),$$

$$\mathcal{L} = \frac{1}{2}A_{ij}\frac{\partial^2}{\partial x_i \partial x_j} + C_i\frac{\partial}{\partial x_i}.$$

From Theorem 3.1, $X^\varepsilon \Rightarrow X$ as $\varepsilon \to 0$, where

$$X_t = x + Ct + A^{1/2}W_t, \quad t \geqslant 0.$$

We assume here moreover that

$$A \text{ is strictly positive definite.} \tag{5.3}$$

Then

$$\tau = \inf\{t \geqslant 0, X_t \notin \bar{D}\}$$

is an a.s. continuous function of the limiting trajectory $\{X_t\}$. It follows from Proposition 3.3 that

$$\int_0^t f\left(X_s^\varepsilon, \frac{X_s^\varepsilon}{\varepsilon}\right) ds \Rightarrow \int_0^t D(X_s)\,ds.$$

From (5.1) and (5.2) we deduce the necessary uniform integrability in order to establish the

Theorem 5.2. *Under the conditions* H.1–H.4, (5.1)–(5.3)*,*

$$u^\varepsilon(x) \to \mathbf{E}_x\left[g(X_\tau)\exp\left(\int_0^\tau D(X_s)\,ds\right)\right], \quad \textit{as } \varepsilon \to 0,$$

where $u(x) := \mathbf{E}_x[g(X_\tau)\exp(\int_0^\tau D(X_s)\,ds)]$ *is the solution of the elliptic PDE*

$$\begin{cases} \mathcal{L}u(x) + D(x)u(x) = 0, & x \in D, \\ u(x) = g(x), & x \in \partial D, \end{cases}$$

at least in the viscosity sense.

Remark 5.3. The assumption that A be nondegenerate is not really necessary for our purpose. All we need is that for almost all (with respect to the law of X^x_τ) points $x \in \partial D$, $\langle An(x), n(x)\rangle > 0$, where $n(x)$ denotes the normal at x to ∂D. Depending on the geometry of D, this can be true even if $\dim(\operatorname{Im} A) = 1$ and $d > 1$.

6. Homogenization of a parabolic PDE

Assume that a, b, c satisfy the above assumptions. Let $e \in \mathcal{C}^1(\mathbf{R}^d, \mathbf{R})$, $f \in \mathcal{C}_\mathrm{b}(\mathbf{R}^d \times \mathbf{R}^d, \mathbf{R})$, e is periodic, f is periodic with respect to its second argument, and $g \in \mathcal{C}(\mathbf{R}^d)$ grows at most polynomially at infinity. For each $\varepsilon > 0$, we consider the PDE:

$$\begin{cases} \dfrac{\partial u^\varepsilon}{\partial t}(t,x) = \mathcal{L}_\varepsilon u^\varepsilon(t,x) + \left(\dfrac{1}{\varepsilon} e\left(\dfrac{x}{\varepsilon}\right) + f\left(x, \dfrac{x}{\varepsilon}\right)\right) u^\varepsilon(t,x), \\ u^\varepsilon(0,x) = g(x), \quad x \in \mathbf{R}^d. \end{cases} \tag{6.1}$$

We assume that

$$\int_{\mathbf{T}^d} e(x)\, \mu(dx) = 0. \tag{6.2}$$

Define

$$Y^\varepsilon_t = \int_0^t \left[\frac{1}{\varepsilon} e\left(\frac{X^\varepsilon_s}{\varepsilon}\right) + f\left(X^\varepsilon_s, \frac{X^\varepsilon_s}{\varepsilon}\right)\right] ds.$$

Then the solution of (6.1) is given by

$$u^\varepsilon(t,x) = \mathbf{E}\big[g\big(X^\varepsilon_t\big) \exp\big(Y^\varepsilon_t\big)\big], \tag{6.3}$$

where X^ε_t is the solution of the SDE (2.1).

Define

$$\hat{e}(x) = \int_0^\infty \mathbf{E}_x\big[e(\tilde{X}_t)\big]\, dt,$$

$$\hat{b}_i(x) = \int_0^\infty \mathbf{E}_x\big[b_i(\tilde{X}_t)\big]\, dt, \quad i = 1, \ldots, d,$$

the weak sense solutions of the Poisson equations

$$L\hat{e}(x) + e(x) = 0,$$

$$L\hat{b}_i(x) + b_i(x) = 0, \quad i = 1, \ldots, d.$$

We let

$$\begin{cases} A = \int\limits_{\mathbf{T}^d} (I + \nabla\hat{b})a(I + \nabla\hat{b})^*(x)\,\mu(dx), \\ C = \int\limits_{\mathbf{T}^d} (I + \nabla\hat{b})(c + a\nabla\hat{e})(x)\,\mu(dx), \\ D(x) = \int\limits_{\mathbf{T}^d} \left(\frac{1}{2}\nabla\hat{e}^* a\nabla\hat{e} + f(x,\cdot) + \nabla\hat{e}c\right)(y)\,\mu(dy). \end{cases} \tag{6.4}$$

Then, defining again $X_t^x = x + Ct + A^{\frac{1}{2}} W_t$,

$$u(t,x) = \mathbf{E}\big[g\big(X_t^x\big)e^{\int_0^t D(X_s^x)\,ds}\big]$$

is the solution (at least in the viscosity sense) of the parabolic PDE

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = \dfrac{1}{2}A_{ij}\dfrac{\partial^2 u}{\partial x_i \partial x_j}(t,x) + C_i\dfrac{\partial u}{\partial x_i}(t,x) + D(x)u(t,x), \\ u(0,x) = g(x), \quad x \in \mathbf{R}^d. \end{cases}$$

Theorem 6.1. *Under the conditions* H.1–H.4*, the smoothness of e, the boundedness and continuity of f, the continuity and growth condition on g and* (6.2)*, for all $t \geqslant 0$, $x \in \mathbf{R}^d$,*

$$u^\varepsilon(t,x) \to u(t,x)$$

as $\varepsilon \to 0$.

For the proof of Theorem 6.1, we need a result which is proved exactly as Lemma 3.2, namely

Lemma 6.2. *Define*

$$\hat{Y}_t^\varepsilon = Y_t^\varepsilon + \varepsilon\hat{e}\left(\frac{X_t^\varepsilon}{\varepsilon}\right).$$

Then the following holds for each $\varepsilon > 0$ and $t > 0$:

$$\hat{Y}_t^\varepsilon = \varepsilon\hat{e}\left(\frac{x}{\varepsilon}\right) + \int\limits_0^t [f + \nabla\hat{e}c]\left(X_s^\varepsilon, \frac{X_s^\varepsilon}{\varepsilon}\right)ds + \int\limits_0^t \nabla\hat{e}\sigma\left(\frac{X_s^\varepsilon}{\varepsilon}\right)dW_s.$$

We can now proceed with the

Proof of Theorem 6.1. We first define a new probability $\tilde{\mathbf{P}}$ by the formula

$$\left.\frac{d\tilde{\mathbf{P}}}{d\mathbf{P}}\right|_t = \exp\left(\int\limits_0^t \nabla\hat{e}\sigma\left(\frac{X_s^\varepsilon}{\varepsilon}\right)dW_s - \frac{1}{2}\int\limits_0^t \langle\nabla\hat{e}, a\nabla\hat{e}\rangle\left(\frac{X_s^\varepsilon}{\varepsilon}\right)ds\right).$$

We next remark that it follows from (6.3) that the behaviour as $\varepsilon \to 0$ of $u^\varepsilon(t,x)$ is the same as that of

$$\hat{u}^\varepsilon(t,x) := \mathbf{E}\big[g\big(\hat{X}_t^\varepsilon\big)\exp\big(\hat{Y}_t^\varepsilon - \varepsilon\hat{e}(x/\varepsilon)\big)\big].$$

The definition of $\tilde{\mathbf{P}}$ then yields

$$\hat{u}^\varepsilon(t,x) = \tilde{\mathbf{E}}\left[g\big(\hat{X}_t^\varepsilon\big)\exp\int_0^t \tilde{f}\left(X_s^\varepsilon, \frac{X_s^\varepsilon}{\varepsilon}\right)ds\right],$$

where we defined

$$\tilde{f} = f + \nabla\hat{e}c + \frac{1}{2}\langle\nabla\hat{e}, a\nabla\hat{e}\rangle.$$

On the other hand, it follows from Girsanov's theorem that

$$\hat{X}_t^\varepsilon = x + \varepsilon\hat{b}\left(\frac{x}{\varepsilon}\right) + \int_0^t (I+\nabla\hat{b})(c+a\nabla\hat{e})\left(\frac{X_s^\varepsilon}{\varepsilon}\right)ds + \int_0^t (I+\nabla\hat{b})\sigma\left(\frac{X_s^\varepsilon}{\varepsilon}\right)d\tilde{W}_s,$$

where $\tilde{W}_t = W_t - \int_0^t \sigma^t\nabla\hat{e}(\frac{X_s^\varepsilon}{\varepsilon})\,ds$ is a Brownian motion under $\tilde{\mathbf{P}}$.

An obvious adaptation of the proof of Theorem 3.1 shows that $\hat{X}^{x,\varepsilon} \Rightarrow X^x$, where $X_t^x = x + Ct + A^{1/2}\tilde{W}_t$, C is given by (6.4), and $\{\tilde{W}_t\}$ is a Brownian motion under $\tilde{\mathbf{P}}$. Since it follows from Proposition 3.3 that

$$\int_0^t \tilde{f}\left(X_s^\varepsilon, \frac{X_s^\varepsilon}{\varepsilon}\right)ds \Rightarrow \int_0^t D\big(X_s^x\big)\,ds$$

as $\varepsilon \to 0$, the result follows. □

7. Examples and counterexamples

This section provides several examples of degenerate diffusions for which our results apply. In each case, we furthermore give a characterisation of the range of the effective diffusion matrix A.

The figures should be interpreted as follows. The little black arrows show the vector field b. The shaded dark grey regions denote the points where the strong Hörmander condition holds (take for example a elliptic in the grey regions). We make *no assumption whatsoever* on the positivity of a in the white regions. In particular, a is allowed to vanish there.

Our first example (which is the only one that we are going to work out in some detail) is a typical example of the type of diffusions for which our conditions H.1–H.4 apply. Observe first that the diffusion $\{\tilde{X}_t, t \geqslant 0\}$ on $\mathbf{T}^d$ has the measure with density p as invariant measure, and H.4

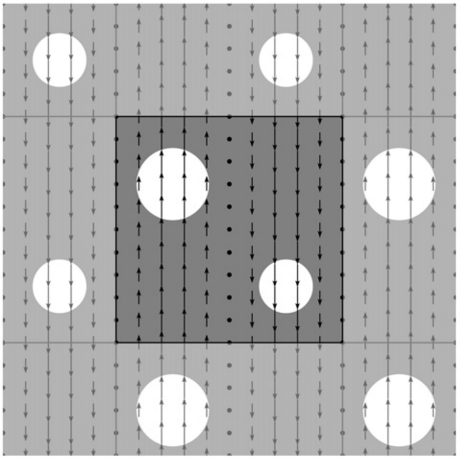

Fig. 1.

is satisfied if and only if there exists a mapping H from $\mathbf{T}^d$ into the set of $d \times d$ antisymmetric matrices, such that

$$b_i(x) = \frac{1}{2p(x)} \sum_j \frac{\partial}{\partial x_j}(pa_{ij} + H_{ij})(x), \quad 1 \leqslant i \leqslant d, x \in \mathbf{T}^d.$$

Let $\alpha : \mathbf{T}^d \to [0, 1]$ be a smooth function such that the set $\mathcal{A} = \{x \mid \alpha(x) = 0\}$ has a finite number of bounded connected components. For $\rho > 0$, we define the set $\mathcal{A}_\rho := \{x, d(x, \mathcal{A}) \leqslant \rho\}$ and we assume that there exists $\rho > 0$ such that $\mathcal{A}_\rho$ intersects neither $\{x_1 = 0\}$ nor $\{x_1 = 1/2\}$.

We choose our diffusion matrix σ to be given by $\sigma(x) = \alpha(x)I$, and we let

$$H = \begin{pmatrix} 0 & 0 & \cdots & 0 & (2\pi)^{-1}\cos(2\pi x_1) \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ -(2\pi)^{-1}\cos(2\pi x_1) & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Hence, if we choose

$$2b(x) = \nabla\big(\alpha^2\big)(x) + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sin(2\pi x_1) \end{pmatrix},$$

then the Lebesgue measure is invariant for the process $\tilde{X}_t^x$. Fig. 1 shows the vector field given by b, together with an example of a set $\mathcal{A}$ (the white disks) that satisfies our conditions.

Assumptions H.1 and H.4 are satisfied by construction. Assumption H.2 is satisfied since, by taking the trivial control $u = 0$, there exists a time t such that every solution to the controlled system reaches the complement of $\mathcal{A}$ before time t. The same argument shows that H.3 is satisfied as well.

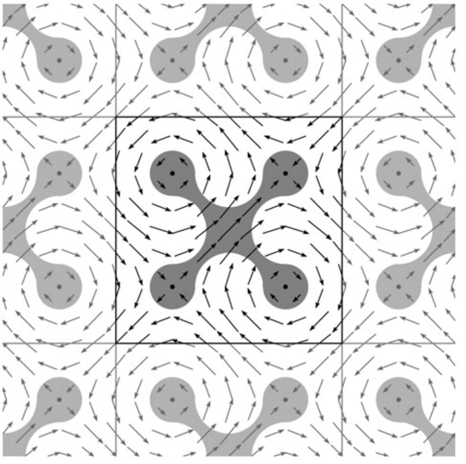

Fig. 2.

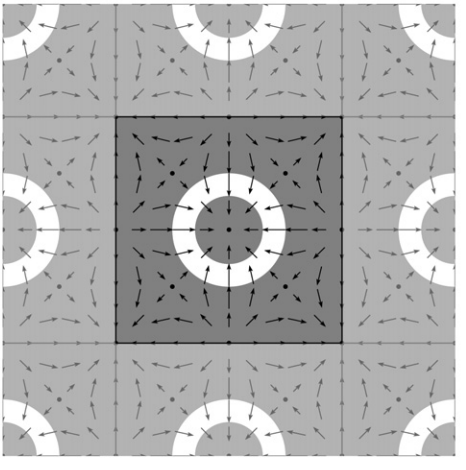

Fig. 3.

Our second example is depicted in Fig. 2 and provides a situation where the diffusion matrix of the limiting Brownian motion is non-degenerate even though the area in which the strong Hörmander condition holds does not intersect the boundaries of the fundamental domain. In particular, the argument of [10] mentioned in Remark 4.8 does not cover this situation.

In this case, it is easy to see that every point of the torus can be reached by the diffusion on the torus, so that μ has full support and therefore A has full rank. Indeed, take any point on the torus. If it is in the interior of the dark region, it can trivially be reached by the controlled ODE. If not, follow the drift flow in the reverse direction, starting from that point. The trajectory which we create in this way intersects the interior of the dark region. It is now clear that the controlled ODE can reach the given point in finite time.

Our third example, depicted in Fig. 3, is in a way opposite to the first one. It shows that it is possible for the limiting Brownian motion to have zero diffusion coefficient, even though the area in which the strong Hörmander condition holds stretches over the whole space.

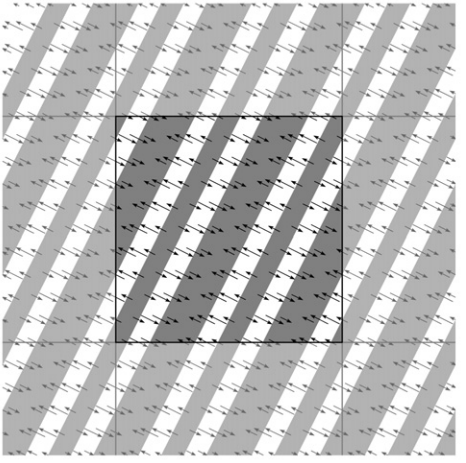

Fig. 4.

In this case, the support S of the invariant measure is given by the gray disk, so that $g(\gamma) = 0$ for every loop in S. This example again relies in a crucial way on the deterministic drift b. If we take the same example but change the sign of b, then the homogenized diffusion matrix A has full rank, since the support of the invariant measure will contain the unbounded shaded area.

In Fig. 4, we finally show an example where the effective diffusivity degenerates in one direction. In this particular case, the range of A is the one-dimensional subspace of $\mathbf{R}^2$ spanned by the vector $(1, 2)$.

Acknowledgments

We would like to thank the Centro de Giorgi in Pisa for their kind hospitality and for creating a stimulating research environment. M.H. was supported by an EPSRC Advanced Research Fellowship, grant number EP/D071593/1.

References

[1] E. Acerbi, V. Chiadò Piat, G. Dal Maso, D. Percivale, An extension theorem from connected sets, and homogenization in general periodic domains, Nonlinear Anal. 18 (5) (1992) 481–496.

[2] M. Bellieud, G. Bouchitté, Homogénéisation de problèmes elliptiques dégénérés, C. R. Acad. Sci. Paris Sér. I Math. 327 (8) (1998) 787–792.

[3] M. Biroli, U. Mosco, N.A. Tchou, Homogenization for degenerate operators with periodical coefficients with respect to the Heisenberg group, C. R. Acad. Sci. Paris Sér. I Math. 322 (5) (1996) 439–444.

[4] J.-M. Bismut, Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions, Z. Wahrsch. Verw. Geb. 56 (4) (1981) 469–505.

[5] D. Cioranescu, J.S.J. Paulin, Homogenization in open sets with holes, J. Math. Anal. Appl. 71 (2) (1979) 590–607.

[6] R. De Arcangelis, F. Serra Cassano, On the homogenization of degenerate elliptic equations in divergence form, J. Math. Pures Appl. (9) 71 (2) (1992) 119–138.

[7] J. Engström, L.-E. Persson, A. Piatnitski, P. Wall, Homogenization of random degenerated nonlinear monotone operators, Glasg. Mat. Ser. III 41(61) (1) (2006) 101–114.

[8] S.N. Ethier, T.G. Kurtz, Markov Processes, Wiley Series in Probab. Math. Statist., Wiley, New York, 1986.

[9] M. Hairer, J.C. Mattingly, Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations, Ann. Probab., in press.

[10] V.V. Jikov, S.M. Kozlov, O.A. Oleĭnik, Homogenization of Differential Operators and Integral Functionals, Springer, Berlin, 1994, translated from the Russian by G.A. Yosifian [G.A. Iosif'yan].
[11] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge Stud. Adv. Math., vol. 24, Cambridge Univ. Press, Cambridge, 1990.
[12] S. Kusuoka, D. Stroock, Applications of the Malliavin calculus. I, in: Stochastic Analysis, Katata/Kyoto, 1982, in: North-Holland Math. Library, vol. 32, North-Holland, Amsterdam, 1984, pp. 271–306.
[13] S. Kusuoka, D. Stroock, Applications of the Malliavin calculus. II, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 32 (1) (1985) 1–76.
[14] P. Malliavin, Stochastic calculus of variation and hypoelliptic operators, in: Proceedings of the International Symposium on Stochastic Differential Equations, Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976, Wiley, New York, 1978, pp. 195–263.
[15] J. Norris, Simplified Malliavin calculus, in: Séminaire de Probabilités, XX, 1984/85, in: Lecture Notes in Math., vol. 1204, Springer, Berlin, 1986, pp. 101–130.
[16] D. Nualart, The Malliavin Calculus and Related Topics, Probab. Appl. (N.Y.), Springer, New York, 1995.
[17] É. Pardoux, Homogenization of linear and semilinear second order parabolic PDEs with periodic coefficients: A probabilistic approach, J. Funct. Anal. 167 (2) (1999) 498–520.
[18] E. Pardoux, A.Y. Veretennikov, On the Poisson equation and diffusion approximation. III, Ann. Probab. 33 (3) (2005) 1111–1133.
[19] F. Paronetto, Homogenization of degenerate elliptic–parabolic equations, Asymptot. Anal. 37 (1) (2004) 21–56.