Two-Scale FEM for Homogenization Problems

Two-Scale FEM for Homogenization Problems*

A.-M. Matache and Ch. Schwab

Research Report No. 2001-06
July 2001

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

*Research supported by the Swiss National Science Foundation under Project "Hierarchic FE-Models for periodic lattice and honeycomb materials" with Number BBW 21-58754.99 and under IHP Network 'Homogenization and Multiple Scales' "HMS2000" of the EC.

Two-Scale FEM for Homogenization Problems*

A.-M. Matache and Ch. Schwab

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 2001-06 July 2001

Abstract

The convergence of a two-scale FEM for elliptic problems in divergence form with coefficients and geometries oscillating at length scale $\varepsilon \ll 1$ is analyzed. Full elliptic regularity independent of ε is shown when the solution is viewed as mapping from the slow into the fast scale. Two-scale FE spaces which are able to resolve the ε scale of the solution with work independent of ε and without analytical homogenization are introduced. Robust in ε error estimates for the two-scale FE spaces are proved. Numerical experiments confirm the theoretical analysis.

*Research supported by the Swiss National Science Foundation under Project "Hierarchic FE-Models for periodic lattice and honeycomb materials" with Number BBW 21-58754.99 and under IHP Network 'Homogenization and Multiple Scales' "HMS2000" of the EC.

1 Introduction

1.1 Homogenization problem

We investigate finite element methods (FEM) for the numerical solution of elliptic homogenization problems in divergence form, i.e.

$$L^{\varepsilon}\left(\frac{x}{\varepsilon},\partial_x\right)u^{\varepsilon} := -\nabla\cdot\left(A\left(\frac{x}{\varepsilon}\right)\nabla u^{\varepsilon}\right) + a_0\left(\frac{x}{\varepsilon}\right)u^{\varepsilon} = f(x), \tag{1.1}$$

where ε is a small parameter and we assume that $A(y)$, $a_0(y)$ are 1-periodic in each variable and that

$$A(\cdot)\in L^{\infty}_{\text{per}}(\widehat{Q})^{n\times n}_{\text{symm}}, \quad a_0(\cdot)\in L^{\infty}_{\text{per}}(\widehat{Q}) \tag{1.2}$$

satisfy, for some $\gamma > 0$,

$$\xi^{\top}A(y)\xi \geq \gamma|\xi|^2, \quad a_0(y)\geq\gamma \quad \forall\xi\in\mathbb{R}^n, \text{ a.e. } y\in\widehat{Q}\subset[0,1]^n. \tag{1.3}$$

Here $\widehat{Q}\subset[0,1]^n$ is referred to as unit-cell domain and we assume that $\widehat{Q}$ has Lipschitz boundary $\partial\widehat{Q} = \widehat{\Gamma}_{\text{per}}\cup\widehat{\Gamma}_N$ with $\widehat{\Gamma}_{\text{per}} = \partial\widehat{Q}\cap\partial[0,1]^n$, and $\widehat{\Gamma}_N = \partial\widehat{Q}\backslash\widehat{\Gamma}_{\text{per}}$ (possibly empty). We assume further that $\widehat{\Gamma}_N$ is smooth to avoid regularity issues, but emphasize that this does not constitute an essential limitation.

We consider (1.1) in a bounded Lipschitz domain Ω covered by a pavement of cells of the form $\varepsilon(k+\widehat{Q})$, with $k\in\mathbb{Z}^n$ and $\varepsilon/\text{diam}\,(\Omega)\ll 1$. We set $\Omega_\varepsilon = \Omega_\varepsilon^\infty\cap\Omega$, where

$$\Omega_\varepsilon^\infty = \bigcup_{\mathbb{Z}^n}\varepsilon(k+\widehat{Q}), \quad \Gamma^\infty_{N,\varepsilon} := \bigcup_{\mathbb{Z}^n}\varepsilon(k+\widehat{\Gamma}_N). \tag{1.4}$$

We complete (1.1) in Ω_ε by Dirichlet boundary conditions on $\partial\Omega$, i.e.,

$$u^\varepsilon = 0 \quad \text{on } \partial\Omega_\varepsilon\cap\partial\Omega, \tag{1.5}$$

and, if $\widehat{\Gamma}_N\neq\emptyset$, by Neumann boundary conditions elsewhere

$$\gamma_1 u^\varepsilon := n\cdot A\left(\frac{x}{\varepsilon}\right)\nabla u^\varepsilon = 0 \quad \text{on } \partial\Omega_\varepsilon\backslash\partial\Omega = \partial\Omega_\varepsilon\cap\Gamma^\infty_{N,\varepsilon}. \tag{1.6}$$

Problems of type (1.1) have been thoroughly analyzed by asymptotic analysis as $\varepsilon\to 0$; we mention only [3, 11] and the references there. In this analytical approach to homogenization, the limiting problem as $\varepsilon\to 0$ of (1.1) is identified first and then solved numerically. Since the limiting problem does not depend on ε, no scale resolution is required. However, fine scale information on u^ε has been lost in the analytic homogenization process and numerical determination of correctors is as costly as solving the original problem.

Here we propose and analyze a two-scale FEM for (1.1), (1.5), (1.6) with $H > \varepsilon$ which does not require analytic homogenization as e.g. in [3, 11] **and** which is able to resolve the ε-scale of $u^\varepsilon(x)$ with $N \ll O(\varepsilon^{-n})$ degrees of freedom. In addition, we will establish rigorous error bounds for the h- , p- and the hp- versions of the two-scale FEM.

1.2 Finite Element Approximation

The FEM is based on the variational form of (1.1), (1.5), (1.6)

$$\text{Find } u^\varepsilon \in H^1_D(\Omega_\varepsilon) \; : \; B^\varepsilon(u^\varepsilon, v) = (f, v) \quad \forall v \in H^1_D(\Omega_\varepsilon), \tag{1.7}$$

where $H^1_D(\Omega_\varepsilon) := \{u \in H^1(\Omega_\varepsilon) \; : \; (1.5) \text{ holds for } u\}$ and the bilinear form $B^\varepsilon : H^1_D(\Omega_\varepsilon) \times H^1_D(\Omega_\varepsilon) \to \mathbb{R}$ is given by

$$B^\varepsilon(u, v) = \int\limits_{\Omega_\varepsilon} \left(A\left(\frac{x}{\varepsilon}\right) \nabla u(x) \right) \cdot \nabla v(x) + a_0\left(\frac{x}{\varepsilon}\right) u(x) v(x) \, dx.$$

By (1.3), (1.7) admits a unique solution $u^\varepsilon \in H^1_D(\Omega_\varepsilon)$ for every $\varepsilon > 0$ and every $f \in L^2(\Omega)$.

Let $V^\varepsilon_N \subset H^1_D(\Omega_\varepsilon)$ be any subspace of dimension $N = \dim(V^\varepsilon_N) < \infty$. Then

$$u^\varepsilon_N \in V^\varepsilon_N \; : \; B^\varepsilon(u^\varepsilon_N, v) = (f, v) \quad \forall v \in V^\varepsilon_N \tag{1.8}$$

defines a unique FE solution and there exists $C > 0$ independent of ε such that

$$\|u^\varepsilon - u^\varepsilon_N\|_{H^1(\Omega_\varepsilon)} \le C \min_{v \in V^\varepsilon_N} \|u^\varepsilon - v\|_{H^1(\Omega_\varepsilon)}. \tag{1.9}$$

Even if the right hand side f, the domain Ω and the coefficients A and a_0 are smooth (i.e., C^∞), if $\varepsilon/\text{diam}(\Omega) \ll 1$ the solution u^ε exhibits oscillations on the ε-scale obstructing FE convergence. More specifically, assume that $\widehat{Q} = [0,1]^n$ and that f and $\partial\Omega$ are smooth. Then $\Omega_\varepsilon = \Omega$ and there exist positive constants $C = C(\Omega)$ and $C(\alpha) = C(\alpha, \Omega)$, $\alpha \in \mathbb{N}^n$, such that

$$\|u\|_{L^2(\Omega)} \le C, \quad \|D^\alpha u\|_{L^2(\Omega)} \le C(\alpha)\varepsilon^{1-|\alpha|}, \quad \forall \alpha \in \mathbb{N}^n, \, |\alpha| > 0. \tag{1.10}$$

Denoting by $V^\varepsilon_N = V_N = S^{p,1}(\Omega, \mathcal{T}_H) \subset H^1(\Omega)$ the FE space of piecewise polynomials of degree $p \ge 1$ on a quasiuniform mesh $\mathcal{T}_H$ of meshwidth H, it holds

$$\min_{v \in S^{p,1}(\Omega, \mathcal{T}_H)} \|u^\varepsilon - v\|_{H^1(\Omega_\varepsilon)} \quad \le \quad C H^p \|D^{p+1} u\|_{L^2(\Omega)} \le C(H/\varepsilon)^p.$$

We have also that

$$\min_{v \in S^{p,1}(\Omega, \mathcal{T}_H)} \|u^\varepsilon - v\|_{H^1(\Omega^\varepsilon)} \le \|u^\varepsilon\|_{H^1(\Omega_\varepsilon)} \le C\|f\|_{L^2(\Omega)}.$$

Therefore the FE error with respect to the usual FE space $V_N = S^{p,1}(\Omega, \mathcal{T}_H)$ satisfies the following *a-priori* bounds

$$\|u^\varepsilon - u^\varepsilon_N\|_{H^1(\Omega_\varepsilon)} \le C \min(1, (H/\varepsilon)^p),$$

with $C = C(p, \Omega, f, A, a_0) > 0$ a constant independent of ε and H. Standard FEM, as e.g., piecewise linears on a quasiuniform mesh $\mathcal{T}_H$ of size H, thus converge only if $H < \varepsilon$, i.e., if $N = \dim V^\varepsilon_N = O(\varepsilon^{-n})$. This *scale resolution requirement* is often prohibitive, especially if $n \ge 3$.

In view of (1.9), the key to a robust discretization of (1.1) is the design of V^ε_N. Rather than incorporating e.g., the asymptotics of u^ε (which is not always defined, see [11] and the references there) into V^ε_N, we design V^ε_N based on a two-scale regularity theory of u^ε.

1.3 Scale Separation for u^ε

Ignoring boundary conditions (1.5), we consider (1.1) on the unbounded domain $\Omega_\varepsilon^\infty$ in (1.4). For any $f \in L^2(\mathbb{R}^n)$, (1.1), (1.6) admits a unique solution $u^\varepsilon \in H^1(\Omega_\varepsilon^\infty)$. We will exploit that u^ε admits the representation [7, 6, 5]

$$u^\varepsilon(x) = \frac{1}{(2\pi)^{n/2}} \int\limits_{t\in\mathbb{R}^n} \hat{f}(t)\psi(x,\varepsilon,t)\, dt, \quad x \in \Omega_\varepsilon^\infty, \tag{1.11}$$

where the kernel $\psi(x,\varepsilon,t)$ is the distributional solution of

$$L^\varepsilon \psi = e^{it\cdot x} \text{ on } \Omega_\varepsilon^\infty, \quad n \cdot A(x/\varepsilon)\nabla\psi = 0 \text{ on } \Gamma_{N,\varepsilon}^\infty. \tag{1.12}$$

To characterize precisely the solution of (1.12) in $\Omega_\varepsilon^\infty$, we introduce weighted Sobolev spaces $H_\nu^j(\Omega_\varepsilon^\infty)$ of complex-valued functions with exponential weights depending on a real parameter ν.

Definition 1.1 *For $j = 0, 1$ and for any $\nu \in \mathbb{R}$ the weighted Sobolev spaces $H_\nu^j(\Omega_\varepsilon^\infty)$ equipped with the norm $\|\cdot\|_{j,\nu}$ are defined to be*

$$H_\nu^j(\Omega_\varepsilon^\infty) = \overline{C_0^\infty(\mathbb{R}^n;\mathbb{C})\Big|_{\Omega_\varepsilon^\infty}}^{\|\cdot\|_{j,\nu}}, \tag{1.13}$$

where

$$\|u\|_{j,\nu}^2 = \int\limits_{\Omega_\varepsilon^\infty} \left(\sum_{|\alpha|\le j} |D_x^\alpha u|^2\right) e^{2\nu\|x\|}\, dx \quad (D_x^\alpha u = \partial_{x_1}^{\alpha_1}\ldots\partial_{x_n}^{\alpha_n} u,\ \forall\, \alpha \in \mathbb{N}_0^n). \tag{1.14}$$

Note that for $\nu > 0$ holds $H_\nu^1 \subset H_0^1 = H^1 \subset H_{-\nu}^1$. To specify the meaning of (1.12), we generalize (1.7) for right hand sides $f(x)$ which are not decaying at ∞. To do so, let us introduce the following sesquilinear form $\Psi(\varepsilon)[\cdot,\cdot] : H_{-\nu}^1(\Omega_\varepsilon^\infty) \times H_\nu^1(\Omega_\varepsilon^\infty) \to \mathbb{C}$:

$$\Psi(\varepsilon)[u,v] = \int\limits_{\Omega_\varepsilon^\infty} \left\{ \left(A\left(\frac{x}{\varepsilon}\right) \nabla_x u(x) \right) \cdot \overline{\nabla_x v(x)} + a_0\left(\frac{x}{\varepsilon}\right) u(x)\overline{v(x)} \right\} dx. \tag{1.15}$$

For all $\varepsilon > 0$ and for $\nu > 0$ sufficiently small, $\Psi(\varepsilon)$ is bounded and 'coercive' with respect to $H_{-\nu}^1(\Omega_\varepsilon^\infty) \times H_\nu^1(\Omega_\varepsilon^\infty)$, in the sense that the *inf-sup* stability condition holds (note that for $\nu = 0$, $\Psi(\varepsilon)$ coincides with B^ε in (1.7)). There holds:

Proposition 1.2 *There exist positive constants ν_0, C and γ such that for all $\nu \in (0, \nu_0)$ and all $\varepsilon > 0$*

1. $|\Psi(\varepsilon)[u,v]| \le C\|u\|_{1,-\nu}\|v\|_{1,\nu}$,
2. $\inf\limits_{\|u\|_{1,-\nu}=1} \sup\limits_{\|v\|_{1,\nu}=1} |\Psi(\varepsilon)[u,v]| \ge \gamma > 0$
3. $\sup\limits_{u\in H_{-\nu}^1(\Omega_\infty^\varepsilon)} |\Psi(\varepsilon)[u,v]| > 0$ *for all $v \in H_\nu^1(\Omega_\infty^\varepsilon)$ and $v \neq 0$.*

The continuity of the sesquilinear form $\Psi(\varepsilon)$ stated in *1.* is obvious. The *inf-sup* condition *2.* and the injectivity property *3.* can be verified in the following way: based on the coercivity of the bilinear form B^ε in (1.7) for the case $\nu = 0$ a perturbation argument can be employed to prove the existence of a positive $\nu_0 > 0$ such that *2.* and *3.* hold. We emphasize that ν_0 is independent of ε and depends only on the upper and lower bounds of the matrix A and of the zero order coefficient a_0. The next Proposition follows by standard elliptic regularity [1, 8] as a corollary of Proposition 1.2. Representation (1.11) is proved in [5].

Proposition 1.3 *The properties 1, 2 and 3 of $\Psi(\varepsilon)$ imply that the variational problem*

$$\begin{aligned} &\text{Given } f \in \left(H^1_\nu(\Omega^\infty_\varepsilon)\right)^*, \text{ find} \\ &u^\varepsilon \in H^1_{-\nu}(\Omega^\infty_\varepsilon) \;:\; \Psi(\varepsilon)[u^\varepsilon, v] = \langle f, v\rangle_{(H^1_\nu(\Omega^\infty_\varepsilon))^* \times H^1_\nu(\Omega^\infty_\varepsilon)}, \qquad \forall v \in H^1_\nu(\Omega^\infty_\varepsilon), \end{aligned} \tag{1.16}$$

admits a unique weak solution $u^\varepsilon \in H^1_{-\nu}(\Omega^\infty_\varepsilon)$ and the a-priori estimate

$$\|u^\varepsilon\|_{H^1_{-\nu}(\Omega^\infty_\varepsilon)} \le (1/\gamma)\|f\|_{(H^1_\nu(\Omega^\infty_\varepsilon))^*}$$

holds. Moreover, u^ε admits the representation (1.11) where the integral is understood as Bochner integral of $H^1_{-\nu}$-valued functions.

$\psi(x, \varepsilon, t)$ is the weak solution of (1.16) with respect to the functional $f = e^{it\cdot x} \in \left(H^1_\nu(\Omega^\infty_\varepsilon)\right)^*$. By Proposition 1.3 we know that

$$\left\|\psi(x, \varepsilon, t)\right\|_{H^1_{-\nu}(\Omega^\infty_\varepsilon)} \le 1/\gamma \|e^{it\cdot x}\|_{(H^1_\nu(\Omega^\infty_\varepsilon))^*}.$$

It is now not difficult to see that

$$\|e^{it\cdot x}\|_{(H^1_\nu(\Omega^\infty_\varepsilon))^*} \le 1/\nu^{n/2}.$$

Therefore, $\|\psi(\cdot, \varepsilon, t)\|_{H^1_{-\nu}(\Omega^\infty_\varepsilon)} \le 1/(\gamma\nu^{n/2})$.

Problem (1.1) has separated scales, a slow variable x and a fast variable $y = x/\varepsilon$, in the following sense: the kernel ψ in (1.12) (which is, in a sense, the fine scale response to the coarse scale excitation $e^{it\cdot x}$) can be written in separated form $\psi(x, \varepsilon, t) = e^{it\cdot x}\phi(x/\varepsilon, \varepsilon, t)$ where $\phi(y, \varepsilon, t)$ is the solution of the so-called *unit-cell problem*: $\phi \in H^1_{\mathrm{per}}(\widehat{Q})$

$$\begin{aligned} \mathcal{L}(\varepsilon, t, y; \partial_y)\phi &:= e^{-i\varepsilon t\cdot y} L^\varepsilon(y, \varepsilon^{-1}\partial_y) e^{i\varepsilon t\cdot y}\phi = 1 \text{ in } \widehat{Q}, \\ \mathcal{B}(\varepsilon, t, y; \partial_y)\phi &:= e^{-i\varepsilon t\cdot y} n \cdot A(y)\nabla_y(e^{i\varepsilon t\cdot y}\phi) = 0 \text{ on } \widehat{\Gamma}_N. \end{aligned} \tag{1.17}$$

Unlike ψ, the kernel ϕ is computable by solving the unit-cell problem (1.17) numerically, for example (but not necessary) with finite elements.

1.4 Two-Scale FEM and Outline of the Paper

Based on the representation (1.11), we see that on $\Omega^\infty_\varepsilon$ (i.e., in the absence of boundary layers) the solution $u^\varepsilon(x)$ can be viewed as a map from the 'slow' variable x into the 'fast' variable x/ε: $u^\varepsilon(x) = U^\varepsilon(x, x/\varepsilon)$, where $U^\varepsilon(x, y)$ depends smoothly on ε. In Section 2 we derive new, two-scale regularity results on $u^\varepsilon(x)$ by analyzing $U^\varepsilon(x, y)$. The two-scale point of view of regularity gives rise to a 'natural' FE discretization of (1.1) by means of a non-standard two-scale FE-space V_N^ε in Ω_ε constructed as follows: Let $\mathcal{T}_H$ be a quasiuniform mesh in Ω (**not** in Ω_ε, i.e., the fine structure of the coefficients is ignored) of meshwidth $H > \varepsilon$ and $S^p(\Omega, \mathcal{T}_H)$ the space of continuous, piecewise polynomials of degree p on $\mathcal{T}_H$ (we assume that $\mathcal{T}_H$ is aligned with the periodic pattern in Ω_ε even if this is not essential for our analysis). Next, we resolve the fast scale by a FEM in $\widehat{Q}$, based on the mesh $\widehat{\mathcal{T}}_h$ (for simplicity also quasiuniform of width h), and the space $S^\mu_{\text{per}}(\widehat{Q}, \widehat{\mathcal{T}}_h)$. The FE space V_N^ε in (1.8) is then the Bochner space

$$V_N^\varepsilon = S^p(\Omega, \mathcal{T}_H;\ S^\mu_{\text{per}}(\varepsilon\widehat{Q}, \varepsilon\widehat{\mathcal{T}}_h)). \tag{1.18}$$

Since $\{1\} \subset S^\mu_{\text{per}}(\widehat{Q}, \widehat{\mathcal{T}}_h)$, $S^p(\Omega, \mathcal{T}_H)|_{\Omega_\varepsilon} \subset V_N^\varepsilon$ and V_N^ε is a generalized FE-space. With V_N^ε robust convergence rates as $h, H \to 0$ can be achieved for u_N^ε as we shall show in Section 3. These two-scale approximation results are quite general and applicable whenever the solution has the two-scale regularity; in particular, the representation (1.11) which is valid only in the linear setting is not necessary. In contrast, in [4, 5, 6] a different (in general smaller) space V_N^ε than (1.18) is proposed. In that approach the kernel $\phi(y, \varepsilon, t)$ in (1.17) is incorporated directly in the FE-space via shape functions $\phi(y, \varepsilon, t)$ sampled at suitable points t_j in the Fourier space.

Since the understanding of the design and the properties of the two-scale FEM depend crucially on the two-scale regularity of $u^\varepsilon(x)$, we investigate it first in Section 2. Section 3 is then devoted to the definition and error analysis of the two-scale FEM. In Section 4 we address computational aspects of the two-scale FEM and present numerical results which support our error estimates.

We remark that we consider here only smooth Γ_N^∞, and smooth (i.e., C^∞) coefficients $A(\cdot), a_0(\cdot)$ in (1.1). The nonsmooth case (i.e., discontinuous $A(\cdot), a_0(\cdot)$, Lipschitz $\widehat{\Gamma}_N$) shall be treated elsewhere.

2 Two scale regularity

Uniform control of the kernel $\phi(y, \varepsilon, t)$ in (1.17) in terms of ε and t implies two-scale regularity results on $u^\varepsilon(x)$. The key is to interpret $u^\varepsilon(x)$ as a map from a Sobolev-space in the "slow variable" x into the "fast variable" $y = x/\varepsilon$. More precisely, $u^\varepsilon(x) = U^\varepsilon(x, x/\varepsilon)$, where $U^\varepsilon(x, y)$ is an element of the Bochner-space $H^r(\Omega, H^s_{\text{per}}(\widehat{Q}))$ for $r, s \geq 0$ depending on the regularity of the coefficients and on the data f and, more importantly, where the ε-dependence of $U^\varepsilon(x, y)$ is smooth. We establish the two-scale regularity by uniform (in t) estimates of the kernel ϕ obtained in Propositions 2.3, 2.4 below. To keep technicalities minimal, we consider here only the case when the unit cell problem admits maximal elliptic regularity. If this is not so, all assertions below have natural analogs in a scale of weighted spaces.

2.1 Two scale shift theorem

Theorem 2.1 *Assume that $A(\cdot)$, $a_0(\cdot)$ are smooth and 1-periodic in $y = x/\varepsilon \in \widehat{Q}$. Then, for $f \in H^k_{\text{comp}}(\mathbb{R}^n)$ $(k \geq 0)$, the solution $u^\varepsilon(x)$ of (1.1) on $\Omega^\varepsilon_\infty$ can be written as $u^\varepsilon(x) = U^\varepsilon(x, y)|_{y=x/\varepsilon}$, $x \in \Omega^\varepsilon_\infty$, where $U^\varepsilon(x, y)$ satisfies in $\Omega = \mathbb{R}^n$ the two-scale regularity estimate*

$$\|U^\varepsilon\|_{H^r(\Omega, H^s_{\text{per}}(\widehat{Q}))} \leq C(k)\ \|f\|_{H^{r+s-1}(\Omega)} \tag{2.1}$$

provided $r + s \leq k + 1$, $r, s \geq 0$, and

$$\|\varepsilon^{-1}\nabla_y U^\varepsilon\|_{H^r(\Omega, H^{s-1}_{\text{per}}(\widehat{Q}))} \leq C(k)\ \|f\|_{H^{r+s-1}(\Omega)} \tag{2.2}$$

provided $r+s \le k+1$, $r, s-1 \ge 0$. Here, $C(k)$ is independent of ε, but depends on $r+s$ (see Remark 3.15 ahead for this dependence).

Proof. The proof is based on the Fourier-Bochner integral representation (1.11) of the solution $u^\varepsilon(x) = U^\varepsilon(x, x/\varepsilon)$ and on the two-scale regularity estimates on the Fourier-Bochner integral kernel in Propositions 2.3, 2.4 below which are uniform in ε and t. For multiindices α, β with $|\alpha| \le r$, $|\beta| \le s$, the mixed derivative (in the sense of distributions) $D_x^\alpha D_y^\beta U^\varepsilon(x,y)$ can be interpreted as mapping $L^2_{\rm per}(\widehat{Q})$ into $L^2(\mathbb{R}^n)$. More precisely, for arbitrary $\varphi \in L^2_{\rm per}(\widehat{Q})$, $\langle D_x^\alpha D_y^\beta U^\varepsilon(x,\cdot), \varphi\rangle_{L^2_{\rm per}(\widehat{Q})\times L^2_{\rm per}(\widehat{Q})}$ is the inverse Fourier transform of a $L^2(\mathbb{R}^n)$ function

$$\langle D_x^\alpha D_y^\beta U^\varepsilon(x,y), \varphi\rangle_{L^2_{\rm per}(\widehat{Q})\times L^2_{\rm per}(\widehat{Q})} = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{it\cdot x} \hat{f}(t) (it)^\alpha \langle D_y^\beta \phi(y,\varepsilon,t), \varphi(y)\rangle_{L^2_{\rm per}(\widehat{Q})\times L^2_{\rm per}(\widehat{Q})}\, dt.$$

By Parseval's relation then, the $L^2(\mathbb{R}^n)$-norm of $\langle D_x^\alpha D_y^\beta U^\varepsilon(x,y), \varphi\rangle_{L^2_{\rm per}(\widehat{Q})\times L^2_{\rm per}(\widehat{Q})}$ is equal to

$$\left\| \langle D_x^\alpha D_y^\beta U^\varepsilon(x,y), \varphi\rangle_{L^2_{\rm per}(\widehat{Q})\times L^2_{\rm per}(\widehat{Q})} \right\|_{L^2(\mathbb{R}^n)} = \left\| (it)^\alpha \hat{f}(t) \langle D_y^\beta \phi(y,\varepsilon,t), \varphi(y)\rangle_{L^2_{\rm per}(\widehat{Q})\times L^2_{\rm per}(\widehat{Q})} \right\|_{L^2(\mathbb{R}^n)}.$$

By (2.7) ahead in Propositions 2.3, 2.4, respectively, there exists a positive constant $C > 0$ independent of ε, t and of the test function φ, such that for all $t \in \mathbb{R}^n$

$$\left| \langle D_y^\beta \phi(y,\varepsilon,t), \varphi(y)\rangle_{L^2_{\rm per}(\widehat{Q})\times L^2_{\rm per}(\widehat{Q})} \right| \le C(1+|t|)^{|\beta|-1} \|\varphi\|_{L^2(\widehat{Q})}. \tag{2.3}$$

Hence, by Parseval's identity again,

$$\left\| \langle D_x^\alpha D_y^\beta U^\varepsilon(x,y), \varphi\rangle_{L^2_{\rm per}(\widehat{Q})\times L^2_{\rm per}(\widehat{Q})} \right\|_{L^2(\mathbb{R}^n)} \le C \|f\|_{H^{r+s-1}(\mathbb{R}^n)} \|\varphi\|_{L^2(\widehat{Q})},$$

which proves (2.1). Proceeding in a similar fashion, by the uniform bounds on the derivatives of the integral kernel in (2.17) one can prove the two scale regularity estimate on the gradient of the solution in (2.2).

□

It therefore remains to analyze the kernel ϕ, i.e., to prove (2.3).

2.2 Regularity of the Fourier-Bochner integral kernel

Lemma 2.2 *Under the assumptions on $A(\cdot)$ and $a_0(\cdot)$ in (1.2), (1.3), there exists a positive constant $C > 0$ independent of ε and t such that for $\varepsilon = 1/M$, $M \in \mathbb{N}$:*

$$\|\phi\|_{0,\widehat{Q}} \le C(1+|t|)^{-1} \quad \textit{and} \quad \|\varepsilon^{-1}\nabla_y \phi\|_{0,\widehat{Q}} \le C. \tag{2.4}$$

Proof. The key observation is that $\chi(x,\varepsilon,t) := it\,\psi(x,\varepsilon,t) \in \left(H^1_{-\nu}(\Omega_\varepsilon^\infty)\right)^n$ solves

$$\begin{aligned} -\nabla_x \cdot \left(A\left(\frac{x}{\varepsilon}\right) \nabla_x \chi(x,\varepsilon,t)\right) + a_0\left(\frac{x}{\varepsilon}\right) \chi(x,\varepsilon,t) &= ite^{it\cdot x} \quad \text{in } \Omega_\varepsilon^\infty, \\ n \cdot A\left(\frac{x}{\varepsilon}\right) \nabla_x \chi(x,\varepsilon,t) &= 0 \quad \text{on } \Gamma_\varepsilon^\infty \text{ (if nonempty)}. \end{aligned}$$

and for $\nu \in (0,\nu_0)$ the $\|\cdot\|_{(H^1_\nu(\Omega_\varepsilon^\infty))^*}$ norm of the right hand side is uniformly bounded with respect to t and ε

$$\|ite^{it\cdot x}\|_{(H^1_\nu(\Omega_\varepsilon^\infty))^*} \le C(\nu, n).$$

Therefore, $\|\chi(x,\varepsilon,t)\|_{1,-\nu} \le C\|ite^{it\cdot x}\|_{(H^1_\nu(\Omega^\infty_\varepsilon))^*} \le C(\nu,n)$ and

$$\|\psi(x,\varepsilon,t)\|_{1,-\nu} \le C(1+|t|)^{-1},$$

with $C>0$ depending only on γ and ν. Then

$$\begin{aligned}\|\phi\|^2_{0,\widehat{Q}} &= \int_{\widehat{Q}} |\phi(y,\varepsilon,t)|^2 dy = \varepsilon^{-n}\int_{\varepsilon\widehat{Q}} \left|\phi\left(\frac{x}{\varepsilon},\varepsilon,t\right)\right|^2 dx \\ &= \sum_{k\in\{0,1,\dots,M-1\}^n} \int_{\varepsilon(\widehat{Q}+k)} \left|\phi\left(\frac{x}{\varepsilon},\varepsilon,t\right)\right|^2 e^{-2\nu|x|}e^{2\nu|x|}dx \\ &\le C\|\psi(x,\varepsilon,t)\|^2_{0,-\nu} \\ &\le C(1+|t|)^{-2}.\end{aligned} \tag{2.5}$$

We obtain estimates for $\|\varepsilon^{-1}\nabla_y\phi\|_{0,\widehat{Q}}$ by a similar argument:

$$\begin{aligned}\int_{\widehat{Q}} \left|\varepsilon^{-1}\nabla_y\phi(y,\varepsilon,t)\right|^2 dy &= \int_{\varepsilon\widehat{Q}} \left|\varepsilon^{-1}\nabla_y\phi\left(\frac{x}{\varepsilon},\varepsilon,t\right)\right|^2 \varepsilon^{-n}dx = \int_{\varepsilon\widehat{Q}} \left|\nabla_x\phi\left(\frac{x}{\varepsilon},\varepsilon,t\right)\right|^2 \varepsilon^{-n}dx \\ &= \sum_{k\in\{0,1,\dots,M-1\}^n} \int_{\varepsilon(\widehat{Q}+k)} \left|\nabla_x\phi\left(\frac{x}{\varepsilon},\varepsilon,t\right)\right|^2 e^{-2\nu|x|}e^{2\nu|x|}dx \\ &\le C\left(1+|t|^2\right)\|\psi(\cdot,\varepsilon,t)\|^2_{1,-\nu} \le C.\end{aligned}$$

□

We now bound higher order norms of ϕ and ψ. To this end, we discuss two cases separately: $\widehat{\Gamma}_N=\emptyset$, $\widehat{Q}=[0,1]^n$ and $\widehat{\Gamma}_N\neq\emptyset$, $\widehat{\Gamma}_N\in C^\infty$, $\widehat{Q}\subset[0,1]^n$.

2.2.1 Case $\widehat{Q}=[0,1]^n$ and $\widehat{\Gamma}_N=\emptyset$.

Recall that $\phi=\phi(y,\varepsilon,t)\in H^1_{\text{per}}(\widehat{Q})$ solves the unit-cell problem

$$\mathcal{L}(\varepsilon,t,y;\,\partial_y)\phi := -(it+\varepsilon^{-1}\nabla_y)^\top(A(y)(it+\varepsilon^{-1}\nabla_y)\phi)+a_0(y)\phi = 1 \quad \text{in } \widehat{Q}=[0,1]^n, \tag{2.6}$$

with periodic, uniformly bounded coefficients $A\in L^\infty(\widehat{Q})^{n\times n}_{\text{sym}}$, $a_0\in L^\infty(\widehat{Q})$.

Proposition 2.3 *Assume that $a_0\in W^{k-1,\infty}_{\text{per}}(\widehat{Q})$ and $A\in W^{k-1,\infty}_{\text{per}}(\widehat{Q})^{n\times n}_{\text{sym}}$ ($k\ge 1$). Then $\phi\in H^k_{\text{per}}(\widehat{Q})$ and the following (uniform in t and ε) estimates hold*

$$\|\phi\|_{L^2(\widehat{Q})} \le C(1+|t|)^{-1}, \quad |\phi|_{H^k(\widehat{Q})} \le C\varepsilon(1+\varepsilon|t|)^{k-1}, \ \textit{for } k\ge 1 \tag{2.7}$$

with a constant $C=C(k)>0$ depending only on k, $\|A\|_{W^{k-1,\infty}(\widehat{Q})}$ and $\|a_0\|_{W^{k-1,\infty}(\widehat{Q})}$.

Proof. (2.7) for $k=1$ is just the statement of Lemma 2.2. For $k\ge 2$ the proof is done by induction with respect to k. We use the notation $\psi^{(k)}(y)=(it\varepsilon+\nabla_y)^k\phi(y,\varepsilon,t)$ and we mean by this any k-th order derivative of ϕ of type $\prod_{j=1}^n(it_j\varepsilon+\partial_{y_j})^{k_j}\phi$, with $k_1+\dots+k_n=k$. From (2.6) we obtain bounds for $\psi^{(0)}$ and $\psi^{(1)}$ in terms of ε and the (lower) coefficient bounds (A, $a_0\ge\gamma>0$)

$$\|\psi^{(0)}\|^2_{0,\widehat{Q}} \le \frac{1}{\gamma}, \quad \|\psi^{(1)}\|^2_{0,\widehat{Q}} \le \frac{\varepsilon^2}{\gamma}. \tag{2.8}$$

By Lemma 2.2, the $L^2(\widehat{Q})$-estimate for $\psi^{(0)}$ in (2.8) can be improved to

$$\|\psi^{(0)}\|_{0,\widehat{Q}} \le C(1+|t|)^{-1}. \tag{2.9}$$

To establish the assertion for $k=2$ assume that $a_0 \in W^{1,\infty}_{\text{per}}(\widehat{Q})$ and $A \in W^{1,\infty}_{\text{per}}(\widehat{Q})^{n\times n}_{\text{sym}}$. Since (2.6) holds in $\mathbb{R}^n$ and the coefficients $A(\cdot)$, $a_0(\cdot)$ are in $W^{1,\infty}_{\text{per}}(\widehat{Q})$, we may apply the interior regularity theory and deduce that $\phi \in H^2_{loc}(\mathbb{R}^n)$. Apply $(it_r\varepsilon + \partial_r)$ to (2.6), multiply it by $v \in H^1(\widehat{Q})$ and integrate by parts. It follows that $\psi_r := (it_r\varepsilon + \partial_r)\phi \in H^1_{loc}(\mathbb{R}^n)$ satisfies

$$\Phi(\varepsilon,t)[\psi_r, v] = L_r(v) + \mathcal{I}_{\widehat{\Gamma}_{\text{per}}}, \tag{2.10}$$

where the bilinear form $\Phi(\varepsilon,t)[\cdot,\cdot]$ and the linear functional L_r in (2.10) are given by

$$\Phi(\varepsilon,t)[\psi,v] = \int_{\widehat{Q}} \left\{ A(y)(it\varepsilon + \nabla_y)\psi \cdot \overline{(it\varepsilon + \nabla_y)v} + \varepsilon^2 a_0(y)\psi\overline{v} \right\} dy$$

and

$$\begin{aligned} L_r(v) &= it_r\varepsilon^3 \int_{\widehat{Q}} \overline{v}(y)\,dy - \varepsilon^2 \int_{\widehat{Q}} \partial_r a_0(y)\phi\overline{v}(y)\,dy \\ &\quad - \int_{\widehat{Q}} \partial_r A(y)(it\varepsilon + \nabla_y)\phi \cdot \overline{(it\varepsilon + \nabla_y)v}\,dy, \end{aligned}$$

respectively, and the boundary integral $\mathcal{I}_{\widehat{\Gamma}_{\text{per}}}$ is given by

$$\begin{aligned} \mathcal{I}_{\widehat{\Gamma}_{\text{per}}} &= \int_{\widehat{\Gamma}_{\text{per}} = \partial\widehat{Q}} n \cdot (it_r\varepsilon + \partial_r)[A(y)(it\varepsilon + \nabla_y)\phi(y)]\overline{v}(y)\,ds_y \\ &= \int_{\widehat{\Gamma}_{\text{per}}} n \cdot \partial_r \left[A(y)(it\varepsilon + \nabla_y)\phi(y)\right] \overline{v}(y)\,ds_y \\ &= \int_{\widehat{\Gamma}_{\text{per}}} n_r \left(A(y)(it\varepsilon + \nabla_y)\phi(y)\right) \cdot \overline{(it\varepsilon + \nabla_y)v}(y)\,ds_y. \end{aligned}$$

If $v \in C^\infty_{\text{per}}(\widehat{Q})$, then $\mathcal{I}_{\widehat{\Gamma}_{\text{per}}}$ vanishes. In fact, it turns out that $\psi_r \in H^1_{\text{per}}(\widehat{Q})$ (i.e., $\phi \in H^2_{\text{per}}(\widehat{Q})$) is the unique weak solution of (2.10) for all $v \in H^1_{\text{per}}(\widehat{Q})$. Next, we take $v = \psi_r$ in (2.10). Then, it follows that

$$\begin{aligned} \gamma\,\|(it\varepsilon + \nabla_y)\psi_r\|^2_{0,\widehat{Q}} &\le \varepsilon^3|t|\,\|\psi_r\|_{0,\widehat{Q}} + \varepsilon^2\|\partial_r a_0\|_{L^\infty(\widehat{Q})}\|\phi\|_{0,\widehat{Q}}\|\psi_r\|_{0,\widehat{Q}} \\ &+ \|\partial_r A\|_{L^\infty(\widehat{Q})}\|(it\varepsilon + \nabla_y)\phi\|_{0,\widehat{Q}}\,\|(it\varepsilon + \nabla_y)\psi_r\|_{0,\widehat{Q}}\,. \end{aligned}$$

Hence,

$$\begin{aligned} \left\|\psi^{(2)}\right\|_{0,\widehat{Q}} &\le \frac{1}{\gamma}\,\|A\|_{W^{1,\infty}(\widehat{Q})} \left\|\psi^{(1)}\right\|_{0,\widehat{Q}} \\ &+ \left[\frac{1}{\gamma}\left(\varepsilon^3|t|\,\left\|\psi^{(1)}\right\| + \varepsilon^2\,\|a_0\|_{W^{1,\infty}(\widehat{Q})} \left\|\psi^{(0)}\right\|_{0,\widehat{Q}} \left\|\psi^{(1)}\right\|_{0,\widehat{Q}}\right)\right]^{1/2}. \end{aligned} \tag{2.11}$$

By (2.8), (2.9) and (2.11) it follows that

$$\|\psi^{(2)}\|_{0,\widehat{Q}} \leq C\varepsilon(1+\varepsilon|t|), \tag{2.12}$$

with a positive constant $C>0$ depending only on γ, $\|A\|_{W^{1,\infty}(\widehat{Q})}$ and $\|a_0\|_{W^{1,\infty}(\widehat{Q})}$.

The assertion for $k=3$ can be proved by assuming further that $A(\cdot) \in W^{2,\infty}_{\text{per}}(\widehat{Q})^{n\times n}_{\text{sym}}$ and $a_0(\cdot) \in W^{2,\infty}_{\text{per}}(\widehat{Q})$. By standard interior regularity theory it follows that $\phi \in H^3_{loc}(\widehat{Q})$. Apply the differential operator $(it_r\varepsilon+\partial_r)(it_s\varepsilon+\partial_s)$ to (2.6), multiply the resulting equation by $v \in H^1(\widehat{Q})$ and integrate by parts over $\widehat{Q}$. It follows that $\psi_{rs} := (it_r\varepsilon+\partial_r)(it_s\varepsilon+\partial_s)\phi \in H^1_{loc}(\mathbb{R}^n)$ satisfies

$$\Phi(\varepsilon,t)[\psi_{rs},v] = L_{rs}(v) + \mathcal{I}_{\widehat{\Gamma}_{\text{per}}}, \tag{2.13}$$

where

$$\begin{aligned}
L_{rs}(v) &= \varepsilon^2(it_r\varepsilon)(it_s\varepsilon)\int_{\widehat{Q}} \overline{v}(y)\,dy \\
&- \int_{\widehat{Q}} \left[\partial_r A(y)(it\varepsilon+\nabla_y)(it_s\varepsilon+\partial_s)\phi + \partial_s A(y)(it\varepsilon+\nabla_y)(it_r\varepsilon+\partial_r)\phi\right]\cdot\overline{(it\varepsilon+\nabla_y)v}\,dy \\
&- \varepsilon^2\int_{\widehat{Q}} \left[\partial_r a_0(y)(it_s\varepsilon+\partial_s) + \partial_s a_0(y)(it_r\varepsilon+\partial_r)\right]\phi\overline{v}\,dy \\
&- \int_{\widehat{Q}} \partial^2_{rs}A(y)(it\varepsilon+\nabla_y)\phi\cdot\overline{(it\varepsilon+\nabla_y)v} - \varepsilon^2\int_{\widehat{Q}} \partial^2_{rs}a_0(y)\phi\overline{v}\,dy,
\end{aligned} \tag{2.14}$$

and the boundary integral $\mathcal{I}_{\widehat{\Gamma}_{\text{per}}}$ in (2.13) is given by

$$\begin{aligned}
\mathcal{I}_{\widehat{\Gamma}_{\text{per}}} &= \int_{\widehat{\Gamma}_{\text{per}}} n\cdot\Big[(it_r\varepsilon+\partial_r)(it_s\varepsilon+\partial_s)\left(A(y)(it\varepsilon+\nabla_y)\phi\right)\Big]\overline{v}\,ds_y \\
&= \int_{\widehat{\Gamma}_{\text{per}}} n\cdot\Big[\partial_r\partial_s\left(A(y)(it\varepsilon+\nabla_y)\phi\right)\Big]\overline{v}\,ds_y
\end{aligned}$$

If $v \in C^\infty_{\text{per}}(\widehat{Q})$, then the boundary integral $\mathcal{I}_{\widehat{\Gamma}_{\text{per}}}$ in (2.13) vanishes and it turns out that $\psi_{rs} \in H^1_{\text{per}}(\widehat{Q})$, i.e., $\phi \in H^3_{\text{per}}(\widehat{Q})$. Take $v = (it_r\varepsilon+\partial_r)(it_s\varepsilon+\partial_s)\phi$ in (2.13). It follows that

$$\begin{aligned}
\gamma\left\|\psi^{(3)}\right\|^2_{0,\widehat{Q}} &\leq \varepsilon^4|t|^2\left\|\psi^{(2)}\right\|_{0,\widehat{Q}} + 2\|A\|_{W^{1,\infty}(\widehat{Q})}\left\|\psi^{(2)}\right\|_{0,\widehat{Q}}\left\|\psi^{(3)}\right\|_{0,\widehat{Q}} \\
&+ 2\varepsilon^2\|a_0\|_{W^{1,\infty}(\widehat{Q})}\left\|\psi^{(1)}\right\|_{0,\widehat{Q}}\left\|\psi^{(2)}\right\|_{0,\widehat{Q}} + \|A\|_{W^{2,\infty}(\widehat{Q})}\left\|\psi^{(1)}\right\|_{0,\widehat{Q}}\left\|\psi^{(3)}\right\|_{0,\widehat{Q}} \\
&+ \varepsilon^2\|a_0\|_{W^{2,\infty}(\widehat{Q})}\left\|\psi^{(0)}\right\|_{0,\widehat{Q}}\left\|\psi^{(2)}\right\|_{0,\widehat{Q}}.
\end{aligned}$$

Hence,

$$
\begin{aligned}
\left\|\psi^{(3)}\right\|_{0,\widehat{Q}} &\le \frac{1}{\gamma}\left(2\left\|A\right\|_{W^{1,\infty}(\widehat{Q})}\left\|\psi^{(2)}\right\|_{0,\widehat{Q}} + \left\|A\right\|_{W^{2,\infty}(\widehat{Q})}\left\|\psi^{(1)}\right\|_{0,\widehat{Q}}\right)\\
&+ \left[\frac{1}{\gamma}\left(\varepsilon^4|t|^2\left\|\psi^{(2)}\right\|_{0,\widehat{Q}} + 2\varepsilon^2\left\|a_0\right\|_{W^{1,\infty}(\widehat{Q})}\left\|\psi^{(1)}\right\|_{0,\widehat{Q}}\left\|\psi^{(2)}\right\|_{0,\widehat{Q}}\right)\right.\\
&+ \left.\varepsilon^2\left\|a_0\right\|_{W^{2,\infty}(\widehat{Q})}\left\|\psi^{(0)}\right\|_{0,\widehat{Q}}\left\|\psi^{(2)}\right\|_{0,\widehat{Q}}\right]^{1/2}.
\end{aligned}
$$

By (2.8), (2.9) and (2.12) it follows immediately that

$$\|\psi^{(3)}\|_{0,\widehat{Q}} \le C\varepsilon(1+\varepsilon|t|)^2,$$

with $C>0$ being a constant which depends only on γ, $\|A\|_{W^{2,\infty}(\widehat{Q})}$, $\|a_0\|_{W^{2,\infty}(\widehat{Q})}$.

For any $k \ge 3$, one can easily see that $\varepsilon^{-1}\|\psi^{(k)}\|_{0,\widehat{Q}}$ can be estimated analogously in terms of ε and t, the dominant term coming from $(|t|^{k-1}\|\psi^{(k-1)}\|_{0,\widehat{Q}})^{1/2}$. Therefore,

$$\varepsilon^{-1}\|\psi^{(k)}\|_{0,\widehat{Q}} \le C(\|A\|_{W^{k-1,\infty}(\widehat{Q})}, \|a_0\|_{W^{k-1,\infty}(\widehat{Q})}, k)(1+\varepsilon|t|)^{k-1}, \quad \text{for all } k = 1,2,\ldots.$$

Recall that $\psi^{(k)} = (it\varepsilon + \nabla_y)^k\phi$. It follows that

$$\varepsilon^{-1}\|D^k\phi\|_{0,\widehat{Q}} \le C(1+\varepsilon|t|)^{k-1}, \quad \text{for all } k=1,2,\ldots. \tag{2.15}$$

□

2.2.2 Case $\widehat{Q} \subset [0,1]^n$, $\widehat{\Gamma}_N \neq \emptyset$, $\widehat{\Gamma}_N \in C^\infty$.

By (1.17), $\phi \in H^1_{\text{per}}(\widehat{Q})$ solves

$$
\begin{aligned}
\mathcal{L}(\varepsilon,t,y;\partial_y)\phi := -(it+\varepsilon^{-1}\nabla_y)^\top(A(y)(it+\varepsilon^{-1}\nabla_y)\phi) + a_0(y)\phi &= 1 \quad \text{in } \widehat{Q}, && (2.16a)\\
n\cdot A(y)(it+\varepsilon^{-1}\nabla_y)\phi &= 0 \quad \text{on } \widehat{\Gamma}_N. && (2.16b)
\end{aligned}
$$

Proposition 2.4 *Assume that $a_0 \in W^{k-1,\infty}_{\text{per}}(\widehat{Q})$ and $A \in W^{k-1,\infty}_{\text{per}}(\widehat{Q})^{n\times n}_{\text{sym}}$ $(k \ge 1)$. Then $\phi \in H^k_{\text{per}}(\widehat{Q})$ and the following estimates hold*

$$\|\phi\|_{L^2(\widehat{Q})} \le C(1+|t|)^{-1}, \quad \|\varepsilon^{-1}\nabla_y\phi\|_{H^{k-1}(\widehat{Q})} \le C(1+\varepsilon|t|)^{k-1}, \tag{2.17}$$

with a constant $C>0$ depending only on k, $\|A\|_{W^{k-1,\infty}(\widehat{Q})}$ and $\|a_0\|_{W^{k-1,\infty}(\widehat{Q})}$, but independent of ε and t.

Proof. The inequality (2.17) for $k=1$ is just the statement of Lemma 2.2. We prove (2.17) in the general case by induction. Assume that $k \ge 2$ and that (2.17) holds for $k-1$. We write first (2.16a)–(2.16b) in the following form: $\phi \in H^1_{\text{per}}(\widehat{Q})$ solves

$$
\begin{aligned}
-\nabla_y\cdot(A(y)\nabla_y\phi) &= f(y) \quad \text{in } \widehat{Q}\\
n\cdot A(y)\nabla_y\phi &= g(y) \quad \text{on } \widehat{\Gamma}_N,
\end{aligned}
$$

with

$$
\begin{aligned}
f(y) &:= \varepsilon^2 - \varepsilon^2 a_0(y)\phi - \varepsilon^2 t^\top A(y) t\phi + 2i\varepsilon t^\top A(y)\nabla_y\phi + i\varepsilon t^\top\left(\nabla_y\cdot A(y)\right)\phi \\
g(y) &:= -i\varepsilon n\cdot A(y)t\phi.
\end{aligned}
$$

By standard elliptic regularity for Neumann problem, for all $k\geq 0$, it holds

$$
\|\nabla_y^{k+1}\phi\|_{L^2(\widehat{Q})} \leq C(k)\left(\|f\|_{H^{k-1}(\widehat{Q})} + \|g\|_{H^{k-1/2}(\widehat{\Gamma}_N)} + \|\nabla_y\phi\|_{H^{k-1}(\widehat{Q})}\right), \tag{2.18}
$$

with some constant $C(k)>0$. Let us estimate now the $H^k(\widehat{Q})$ norm of $\nabla_y\phi$

$$
\|\nabla_y\phi\|_{H^k(\widehat{Q})} \leq C(k)\left(\|\nabla_y\phi\|_{H^{k-1}(\widehat{Q})} + \|\nabla_y^{k+1}\phi\|_{L^2(\widehat{Q})}\right).
$$

We use the induction assumption and (2.18) to conclude

$$
\begin{aligned}
\|\nabla_y\phi\|_{H^k(\widehat{Q})} &\leq C\varepsilon(1+\varepsilon|t|)^{k-1} + \|\varepsilon^2 - \varepsilon^2 a_0\phi - \varepsilon^2 t^\top A(y)t\phi + 2i\varepsilon t^\top A(y)\nabla_y\phi \\
&\qquad + i\varepsilon t^\top(\nabla_y\cdot A(y))\phi\|_{H^{k-1}(\widehat{Q})} + \varepsilon\|n\cdot A(y)t\phi\|_{H^{k-1/2}(\widehat{\Gamma}_N)} \\
&\leq C(k)\varepsilon(1+\varepsilon|t|)^k.
\end{aligned}
$$

□

2.3 Sharpness of the two-scale regularity.

We consider 1-d problems, i.e., $n=1$. Let us assume that $f\in L^2_{\text{per}}(0,1)$ has the Fourier expansion $f(x)=\sum_{k\in\mathbb{Z}} f_k e^{2\pi ikx}$. Assume further that $a(\cdot)$ is a 1-periodic, L^∞ function and $\varepsilon = 1/M$, with $M\in\mathbb{N}^*$. Let $u^\varepsilon(x)\in H^1_0(0,1)$ be the solution of the following boundary value problem

$$
-\frac{d}{dx}\left(a\left(\frac{x}{\varepsilon}\right)\frac{du^\varepsilon}{dx}\right) = f(x) \quad \text{in } \Omega=(0,1), \quad u^\varepsilon\Big|_{\partial\Omega} = 0.
$$

Then, the solution $u^\varepsilon(x)$ exhibits a 'two-scale' behavior, in the sense that $u^\varepsilon(x) = U^\varepsilon(x, x/\varepsilon)$ and $U^\varepsilon(x,y)$ is 1-periodic in y and has certain regularity properties if seen as a mapping from $x\in\Omega$ into $L^2_{\text{per}}(\widehat{Q})$.
We use the notation $\langle g\rangle$ to denote the mean value of the function g

$$
\langle g\rangle = \int_{\widehat{Q}} g(t)\,dt, \quad \widehat{Q} = (0,1).
$$

We introduce two further periodic functions

$$
A(t) = \int_0^t \left(\frac{1}{a(s)} - \langle 1/a(\cdot)\rangle\right) ds, \quad \tilde{A}(t) = \int_0^t \left(A(s) - \langle A\rangle\right) ds. \tag{2.20}
$$

Let $A_p,\ p \in \mathbb{Z}$ be the Fourier coefficients of $A(\cdot) \in H^1_{\mathrm{per}}(0,1)$. With these definitions, we may write now explicitly the expression for $U^\varepsilon(x,y)$

$$\begin{aligned} U^\varepsilon(x,y) &= -f_0\Big[x\varepsilon A(y) - \varepsilon^2\tilde{A}(y) + \frac{1}{2}x^2\langle 1/a(\cdot)\rangle - \varepsilon x\Big(\frac{1}{2}\langle 1/a(\cdot)\rangle - \langle \cdot/a(\cdot)\rangle\Big)\Big] \\ &+ D\Big(\varepsilon A(y) + x\langle 1/a(\cdot)\rangle\Big) + \sum_{k\in\mathbb{Z}^*}\frac{f_k}{2\pi i k}\Big(\varepsilon A(y) + x\langle 1/a(\cdot)\rangle\Big) \\ &- \sum_{k\in\mathbb{Z}^*}\frac{f_k}{2\pi i k}\Big\{e^{2\pi i k x}\varepsilon A(y) + \frac{\langle 1/a(\cdot)\rangle}{2\pi i k}(e^{2\pi i k x}-1) \\ &- 2\pi i k\varepsilon^2\Big[\sum_{\substack{p\in\mathbb{Z}\\ p\neq -k\varepsilon}} A_p\frac{e^{2\pi i(kx+py)}-1}{2\pi i(p+k\varepsilon)}\dot{+}A_{-k\varepsilon}\frac{x}{\varepsilon}\Big]\Big\}, \end{aligned}$$

and $\dot{+}$ here means that the summation is done only over $k \in (1/\varepsilon)\mathbb{Z}^*$. The constant D is given by

$$D = \frac{1}{\langle 1/a(\cdot)\rangle}\int_0^1 \frac{F(t)}{a(t/\varepsilon)}\,dt, \quad F(t) = \int_0^t f(s)\,ds.$$

From this representation, we see from the terms in the sum $\sum_{k\in\mathbb{Z}^*}$ for large k that the regularity $\|f\|_{H^{r+s-1}}$ in (2.1), (2.2) is optimal. For simplicity, we assume $s=0$. Then, it holds

Proposition 2.5 *Assume that $a(\cdot)$ is smooth and 1-periodic in $y = x/\varepsilon \in \widehat{Q}$. Then, for $f \in H^{r-1}_{\mathrm{per}}(\Omega)$ $(r \geq 0)$, the solution $u^\varepsilon(x)$ of (2.19) on Ω satisfies the two-scale regularity estimate (2.1), (2.2)*

$$\|U^\varepsilon\|_{H^r(\Omega,\, L^2_{\mathrm{per}}(\widehat{Q}))} \leq C(r)\|f\|_{H^{r-1}(\Omega)}. \tag{2.21}$$

Moreover, for all $r \geq 1$

$$\|\varepsilon^{-1}\nabla_y U^\varepsilon\|_{H^{r-1}(\Omega,\, L^2_{\mathrm{per}}(\widehat{Q}))} \leq C(r)\|f\|_{H^{r-1}(\Omega)}. \tag{2.22}$$

The estimates (2.21)–(2.22) are sharp, in the sense that for ε sufficiently small, there exists a constant $c = c(r) > 0$, which does not depend on ε, such that for all $r \geq 1$

$$c(r)\|f\|_{H^{r-2}(\Omega)} \leq \|U^\varepsilon\|_{H^r(\Omega,\, L^2_{\mathrm{per}}(\widehat{Q}))} + \|\varepsilon^{-1}\nabla_y U^\varepsilon\|_{H^{r-1}(\Omega,\, L^2_{\mathrm{per}}(\widehat{Q}))}.$$

Proof. For the proof we refer to Appendix A.

Remark 2.6 Careful inspection of the proof reveals that for $r \geq 1$ the upper bounds in (2.21)–(2.22) have the form $C(r)\|f\|_{H^{r-2}(\Omega)} + C(\varepsilon,r)\|f\|_{H^{r-1}(\Omega)}$ with $C(\varepsilon,r) > 0$ depending on ε and vanishing with $\varepsilon \to 0$. We have a slightly different situation as, e.g., for

$$-\Delta u = f \quad \text{in } \Omega, \quad f \in H^{r-2}(\Omega), \quad u|_{\partial\Omega} \text{ smooth}$$

where $\partial\Omega$ is smooth and we have the sharp shift result: there exists $C(r,\Omega) > 0$

$$\|u\|_{H^r(\Omega)} \leq C(r,\Omega)\|f\|_{H^{r-2}(\Omega)}, \quad r \geq 1,$$

in the sense that for generic data, $\|u\|_{H^r(\Omega)}$ has a lower bound of the same type ($c(r,\Omega) > 0$)

$$\|u\|_{H^r(\Omega)} \geq c(r,\Omega)\|f\|_{H^{r-2}(\Omega)}.$$

In our case, however, the gap $C(\varepsilon,r)\|f\|_{H^{r-1}(\Omega)}$ cannot be removed.

3 Rate of convergence of the two-scale FEM

In the previous section we saw that uniform regularity of $u^\varepsilon(x)$ in dependence on ε could be properly expressed in terms of the two-scale Sobolev spaces $H^r(\mathbb{R}^n, H^s_{\mathrm{per}}(\widehat{Q}))$. The two-scale Finite-Element spaces in the Introduction are therefore natural for the discretization of homogenization problems. In the present section we prove robust convergence estimates for these two-scale FEM under two-scale regularity hypothesis on $u^\varepsilon(x)$.

3.1 Preliminaries

Let $\Omega \subset \mathbb{R}^n$, $\Omega' \subset \mathbb{R}^n$ be two Lipschitz domains. For $\alpha, \beta \in \mathbb{N}^n$ two multiindices we define the Sobolev spaces $\mathcal{H}^{\alpha,\beta}(\Omega \times \Omega')$ of mixed order on the product domain $\Omega \times \Omega'$ as

$$\mathcal{H}^{\alpha,\beta}(\Omega \times \Omega') := \{u \in L^2(\Omega \times \Omega') \ : \ D_x^\gamma D_z^\delta u \in L^2(\Omega \times \Omega'),\ \forall \gamma \le \alpha, \delta \le \beta\},$$

where $\gamma \le \alpha$ is understood componentwise. These are Hilbert spaces with respect to the norm

$$\|u\|^2_{\mathcal{H}^{\alpha,\beta}(\Omega\times\Omega')} := \sum_{\substack{0\le\gamma\le\alpha \\ 0\le\delta\le\beta}} \|D_x^\gamma D_z^\delta u\|^2_{L^2(\Omega\times\Omega')}.$$

3.1.1 Traces in Sobolev spaces of mixed order

For a function $f(\cdot,\cdot) : \Omega \times \Omega \to \mathbb{C}$, we denote by $(\mathcal{R}f)(x) = f(x,y)|_{y=x} : \Omega \to \mathbb{C}$ its restriction to the diagonal $\{(x,y) \in \Omega \times \Omega \mid x = y\}$.

Lemma 3.1 *Let $\Omega = \Omega' := (0,1)^n$ and denote by $\mathbf{1} \in \mathbb{N}^n$ the multiindex $(1,\dots,1)$. Then for any fixed pair of multiindices $\alpha, \beta \in \mathbb{N}^n$ with $\alpha + \beta = \mathbf{1}$ the restriction operator $\mathcal{R}$: $\mathcal{H}^{\alpha,\beta}(\Omega \times \Omega) \to L^2(\Omega)$ is continuous, i.e., there exists a constant $C = C(n) > 0$ such that*

$$\|\mathcal{R}f\|_{L^2(\Omega)} \le C(n)\|f\|_{\mathcal{H}^{\alpha,\beta}(\Omega\times\Omega)}, \quad \forall f \in \mathcal{H}^{\alpha,\beta}(\Omega \times \Omega).$$

Proof. Without loss of generality we present the proof for the case $n = 2$ only. By making eventually a change of coordinates we may assume that $\alpha = \mathbf{1}$ and $\beta = \mathbf{0}$, where we denote by $\mathbf{0}$ the multiindex $(0,\dots,0) \in \mathbb{N}^n$. We may restrict the proof to the case when $f \in C^\infty(\overline{\Omega} \times \overline{\Omega})$, we can use then a density argument to conclude for general $f \in \mathcal{H}^{\alpha,\beta}(\Omega \times \Omega)$. Let $\varphi \in C_0^\infty(\mathbb{R})$ be such that

$$0 \le \varphi(s) \le 1, \forall s \in \mathbb{R}, \quad \varphi(0) = 1, \quad \operatorname{supp}\varphi \subset (-1,1).$$

Then,

$$\begin{aligned}
f(x,x) &= \int_0^{x_1} \partial_{x_1} f(t,x_2,x)\,dt + f(0,x_2,x) \\
&= \int_0^{x_1}\int_0^{x_2} \partial^2_{x_1x_2} f(t,s,x)\,ds\,dt + \int_0^{x_1} \left(\partial_{x_1} f(t,0,x)\varphi(0) - \partial_{x_1} f(t,1,x)\varphi(1)\right)\,dt + f(0,x_2,x) \\
&= \int_0^{x_1}\int_0^{x_2} \partial^2_{x_1x_2} f(t,s,x)\,ds\,dt + \int_0^{x_1}\int_0^{1} \left(\partial^2_{x_1x_2} f(t,s,x)\varphi(s) + \partial_{x_1} f(t,s,x)\varphi'(s)\right)\,ds\,dt \\
&\quad + f(0,x_2,x).
\end{aligned}$$

Therefore,

$$\|\mathcal{R}f(\cdot)\|_{0,\Omega}^2 \le C\left(\sum_{0\le\alpha_j\le 1}\|D_x^\alpha f(\cdot,\cdot)\|_{0,\Omega\times\Omega}^2 + \int_\Omega |f(0,x_2,x)|^2 dx\right).$$

It remains to estimate $\int_\Omega |f(0,x_2,x)|^2 dx$. To this end, we proceed as before

$$\begin{aligned} f(0,x_2,x) &= -\int_0^{x_2}\int_0^1 \left(\partial^2_{x_1x_2} f(t,s,x)\varphi(t) + \partial_{x_2} f(t,s,x)\varphi'(t)\right)\,dt\,ds \\ &\quad -\int_0^1 \left(\partial_{x_1} f(s,0,x)\varphi(s) + f(s,0,x)\varphi'(s)\right)\,ds. \end{aligned}$$

The second integral term can be written as

$$\begin{aligned} &-\int_0^1 \left(\partial_{x_1} f(s,0,x)\varphi(s) + f(s,0,x)\varphi'(s)\right)\,ds = \\ &= \int_0^1 \Bigg[\varphi(s)\int_0^1 \left(\partial^2_{x_1x_2} f(s,t,x)\varphi(t) + \partial_{x_1} f(s,t,x)\varphi'(t)\right)\,dt \\ &+\ \varphi'(s)\int_0^1 \left(\partial_{x_2} f(s,t,x)\varphi(t) + f(s,t,x)\varphi'(t)\right)\,dt\Bigg]\,ds. \end{aligned}$$

Summing up, we obtain that

$$\int_\Omega |f(0,x_2,x)|^2 dx \le C(n)\sum_{0\le\alpha_j\le 1}\|D_x^\alpha f(\cdot,\cdot)\|_{0,\Omega\times\Omega}\,.$$

□

Remark 3.2 The trace result of Lemma 3.1 remains true if the domain Ω is replaced with any subdomain $\widehat{Q}$ satisfying the following condition. For all $x \in \widehat{Q}$ there is a $x^* \in \widehat{Q}$ such that $\lfloor x_1^*, x_1\rfloor \times \lfloor x_2^*, x_2\rfloor \subset \widehat{Q}$ and $|x_1^* - x_1|, |x_2^* - x_2| \ge c$ for some positive constant c independent of x, x^*. We denoted here by $\lfloor a,b\rfloor$ the set of all points between a and b

$$\lfloor a,b\rfloor = \begin{cases} [a,b] & \text{if } a\le b \\ [b,a] & \text{else.} \end{cases}$$

This condition is satisfied if e.g. we assume that $\widehat{Q}$ is given as a finite union of patches $\overline{\widehat{Q}} = \cup_{i=1}^I \overline{\widehat{Q}_i}$, I finite, such that $\widehat{Q}_i \cap \widehat{Q}_j = \emptyset$ for all $i \ne j$ and each subdomain $\widehat{Q}_i = F_i((0,1)^n)$ is the image of the unit n-simplex $(0,1)^n$ through a C^∞ diffeomorphism $F_i : [0,1]^n \to \overline{\widehat{Q}_i}$.

Then

$$\begin{aligned}
\mathcal{R}f(x) &= f(x_1,x_2,x)\varphi(1) - f(x_1^*,x_2,x)\varphi(0) \\
&= \int_0^1 \Big\{ (\partial_{x_1} f)(x_1^* + s(x_1 - x_1^*), x_2, x)(x_1 - x_1^*)\varphi(s) + f(x_1^* + s(x_1 - x_1^*), x_2, x)\varphi'(s) \Big\} ds \\
&= \int_{x_1^*}^{x_1} \Big\{ (\partial_{x_1} f)(s, x_2, x)\varphi\left(\frac{s - x_1^*}{x_1 - x_1^*}\right) + f(s,x_2,x)\varphi'\left(\frac{s - x_1^*}{x_1 - x_1^*}\right)\frac{1}{x_1 - x_1^*} \Big\} ds \\
&= \int_{x_1^*}^{x_1}\int_{x_2^*}^{x_2} \Big\{ (\partial^2_{x_1 x_2} f)(s,t,x)\varphi\left(\frac{s - x_1^*}{x_1 - x_1^*}\right)\varphi\left(\frac{t - x_2^*}{x_2 - x_2^*}\right) \\
&+ (\partial_{x_1} f)(s,t,x)\varphi\left(\frac{s - x_1^*}{x_1 - x_1^*}\right)\varphi'\left(\frac{t - x_2^*}{x_2 - x_2^*}\right)\frac{1}{x_2 - x_2^*} \\
&+ (\partial_{x_2} f)(s,t,x)\varphi'\left(\frac{s - x_1^*}{x_1 - x_1^*}\right)\varphi\left(\frac{t - x_2^*}{x_2 - x_2^*}\right)\frac{1}{x_1 - x_1^*} \\
&+ f(s,t,x)\varphi'\left(\frac{s - x_1^*}{x_1 - x_1^*}\right)\varphi'\left(\frac{t - x_2^*}{x_2 - x_2^*}\right)\frac{1}{x_1 - x_1^*}\frac{1}{x_2 - x_2^*} \Big\} ds dt.
\end{aligned}$$

It follows now that

$$\|\mathcal{R}f\|_{0,\widehat{Q}} \le C\frac{1}{c^2}\sum_{0\le\alpha_j\le 1}\|D_x^\alpha f(\cdot,\cdot)\|_{0,\widehat{Q}\times\widehat{Q}}.$$

3.1.2 Polynomial approximation results

In the two-scale error estimates below, we shall require the following error bounds for the tensor product interpolant (unlike standard H^1-estimates, here also **mixed** first derivatives are bound).

Let $|\cdot|_{H^k(\widehat{\Omega})}$ denote the Sobolev seminorm of order k on $\widehat{\Omega} = (-1,1)$ given by

$$|\hat{u}|_{H^k(\widehat{\Omega})} := \|\hat{u}^{(k)}\|_{L^2(\widehat{\Omega})}, \quad \forall\, \hat{u} \in H^k(\widehat{\Omega}).$$

Lemma 3.3 *Let $\hat{u} \in H^{k+1}(\widehat{\Omega})$ for some $k \ge 0$. Then, for each $p \ge 1$, $0 \le k \le p$, there exists a polynomial interpolant $\pi_p\hat{u} \in S^p(\widehat{\Omega})$, with $S^p(\widehat{\Omega})$ denoting the space of polynomials of degree at most p on $\widehat{\Omega}$, such that it holds*

$$\begin{aligned}
\|\hat{u}' - (\pi_p\hat{u})'\|^2_{L^2(\widehat{\Omega})} &\le \frac{(p-k)!}{(p+k)!}|\hat{u}|^2_{H^{k+1}(\widehat{\Omega})} \\
\|\hat{u} - \pi_p\hat{u}\|^2_{L^2(\widehat{\Omega})} &\le \frac{1}{p(p+1)}\frac{(p-k)!}{(p+k)!}|\hat{u}|^2_{H^{k+1}(\widehat{\Omega})}.
\end{aligned}$$

Proof. One takes $\pi_p\hat{u}$ such that $\pi_p\hat{u}(\pm 1) = \hat{u}(\pm 1)$ and $(\pi_p\hat{u})'$ is the Legendre series of $\hat{u}'$ truncated after the Legendre polynomial L_{p-1}. □

In the multi-dimensional case, we denote by $\widehat{\Pi}_p := \pi_p^{(x_1)} \otimes \cdots \otimes \pi_p^{(x_n)}$ $(n = 2,3)$ the tensor product polynomial interpolant of degree p in the reference element $\widehat{K} := (-1,1)^n$.

Lemma 3.4 *Let $n = 2, 3$ and let $\widehat{\Pi}_p = \pi_p^{(x_1)} \otimes \cdots \otimes \pi_p^{(x_n)}$ denote the tensor product polynomial interpolant of degree p ($p \geq 1$) in each variable in $\widehat{K} = (-1,1)^n$. Then, for all $\hat{u} \in H^{k+1}(\widehat{K})$, $n-1 \leq k \leq p$, it holds*

$$\sum_{0 \leq \alpha_j \leq 1} \|D^\alpha(\hat{u} - \widehat{\Pi}_p \hat{u})\|_{L^2(\widehat{K})} \leq C\Phi_n(p,k)\|D^{k+1}\hat{u}\|_{L^2(\widehat{K})}, \tag{3.1}$$

where $\Phi_n(p,k)$ is given by

$$\Phi_n(p,k) \quad = \quad \sqrt{\frac{(p-k+n-1)!}{(p+k-n+1)!}} \leq Cp^{-(k-n+1)}, \quad \text{for } p \to \infty$$

and $C > 0$ is a constant independent of p and k.

The proof of this lemma by a tensor product argument is given in Appendix B. The loss of p^n in (3.1) is due to the control of mixed first derivatives of the error. This control is needed for the application of the trace operator $\mathcal{R}$ below.

Remark 3.5 In the case when $n = 3$, $p = 1$ and $u \in \mathcal{H}^{\mathbf{1}}(\widehat{K})$, with $\mathcal{H}^{\mathbf{1}}(\widehat{K})$ being given by

$$H^1(\widehat{K}) \subset \mathcal{H}^{\mathbf{1}}(\widehat{K}) := \{\hat{u} \mid D^\alpha \hat{u} \in L^2(\widehat{K})\ \forall \alpha \text{ such that } \max_j \alpha_j \leq 1\}, \quad \mathbf{1} = (1,1,1)$$

it holds that

$$\sum_{0 \leq \alpha_j \leq 1} \|D^\alpha(\hat{u} - \widehat{\Pi}_1 \hat{u})\|_{L^2(\widehat{K})} \leq C \sum_{\substack{0 \leq \alpha_j \leq 1 \\ |\alpha| \geq 1}} \|D^\alpha \hat{u}\|_{L^2(\widehat{K})}. \tag{3.2}$$

3.1.3 Definition of the FE spaces

We assume that the domain Ω is axiparallel and we take $\mathcal{T}_H$ to be a quasiuniform triangulation of Ω of affine quadrilateral elements of size H. We take as macro FE space in $\Omega = (0,1)^n$ the standard affine FE space $S^p(\Omega, \mathcal{T}_H)$ defined as $S^p(\Omega, \mathcal{T}_H) = \{u \in H^1(\Omega) \mid u|_K \circ F_K^{-1} \in S^p(\widehat{K})\ \forall K \in \mathcal{T}_H\}$, where we denoted by $F_K : \widehat{K} \to K$ the affine element map associated to the element K.

We introduce next the micro FE space in $\widehat{Q}$, i.e. the FE space with respect to the fast variable in the unit-cell. If $\widehat{Q} = (0,1)^n$, we also take $\widehat{\mathcal{T}}_h$ as a quasiuniform mesh in $\widehat{Q}$ of axiparallel quadrilaterals. For the case when the unit-cell domain $\widehat{Q}$ has e.g. interior holes the 'micro' triangulation $\widehat{\mathcal{T}}_h$ is obtained as follows. First one assumes the existence of a partition $\overline{\widehat{Q}} = \cup_{i=1}^I \overline{\widehat{Q}_i}$ ($I < \infty$ fixed) of $\widehat{Q}$ in a finite number of quadrilateral patches $\widehat{Q}_i$. Each patch $\widehat{Q}_i = F_i((0,1)^n)$ is image of the reference domain $(0,1)^n$ via the C^∞ diffeomorphism $F_i : (0,1)^n \to \widehat{Q}_i$. These mappings satisfy also a compatibility condition along the common interfaces in the sense that $F_i \circ F_{i'}^{-1} = Id$ on $\overline{\widehat{Q}_i} \cap \overline{\widehat{Q}_{i'}}$ for all $i, i' = 1, \ldots, I$ (such F_i can be constructed by blending, see e.g. [10]). The mesh $\widehat{\mathcal{T}}_h$ in $\widehat{Q}$ has to be periodic and is given as union of patch meshes

$$\widehat{\mathcal{T}}_h = \cup_{i=1}^I \widehat{\mathcal{T}}_{h,i}, \quad \widehat{\mathcal{T}}_{h,i} = F_i(\widehat{\widehat{\mathcal{T}}}_h)$$

with $\widehat{\widehat{\mathcal{T}}}_h$ being a uniform, affine quadrilateral mesh in the reference domain $(0,1)^n$. Then $S^\mu_{\text{per}}(\widehat{Q}, \widehat{\mathcal{T}}_h, \mathbf{F})$ is the finite element space of all piecewise mapped polynomials of degree μ:

$$S^\mu_{\text{per}}(\widehat{Q}, \widehat{\mathcal{T}}_h, \mathbf{F}) = \{u \in H^1_{\text{per}}(\widehat{Q}) \mid \left(u|_{\widehat{Q}_i} \circ F_i\right)|_{\widehat{\widehat{K}}} \in S^\mu(\widehat{\widehat{K}})\ \forall \widehat{\widehat{K}} \in \widehat{\widehat{\mathcal{T}}}_h\}.$$

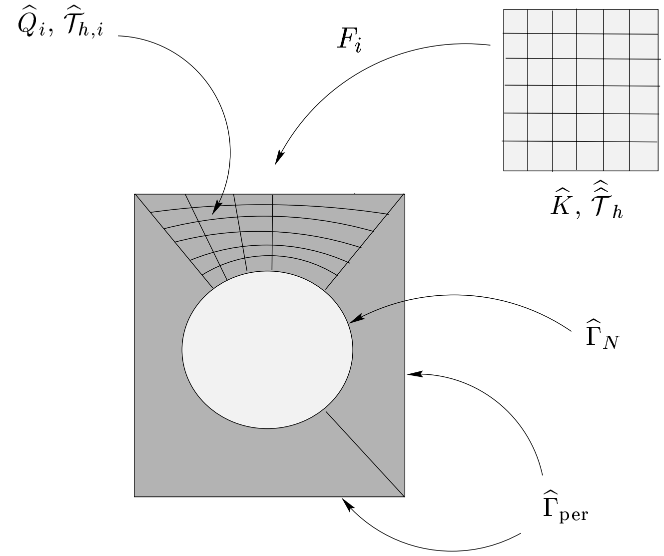

Figure 1: Unit cell domain $\widehat{Q}$ as union of four patches $\widehat{Q}_i$, $i = 1, \ldots, 4$, and the boundaries $\widehat{\Gamma}_N$, $\widehat{\Gamma}_{\rm per}$.

We take as two-scale FE space $\mathcal{V}_N^\varepsilon$ the space of traces of the two-scale Bochner space

$$S^p(\Omega, \mathcal{T}_H; S^\mu_{\rm per}(\widehat{Q}, \widehat{\mathcal{T}}_h, \mathbf{F})) =$$

$$= \Big\{ U(x,y) \mid \forall K \in \mathcal{T} : U(F_K(\hat{x}), y) \text{ is polynomial of degree } p \text{ w.r. to } \hat{x}$$

$$\text{in } \widehat{K} \text{ and continuous, periodic p.w. polynomial with respect to } \widehat{\widehat{\mathcal{T}}}_h \text{ in } y \in \widehat{Q} \Big\}. \tag{3.3}$$

More specifically,

$$\mathcal{V}_N^\varepsilon = \mathcal{R}^\varepsilon S^p(\Omega, \mathcal{T}_H; S^\mu_{\rm per}(\widehat{Q}, \widehat{\mathcal{T}}_h, \mathbf{F})), \tag{3.4}$$

where the trace operator $\mathcal{R}^\varepsilon$ is given by $(\mathcal{R}^\varepsilon U)(x) = U(x,y)|_{y=\frac{x}{\varepsilon}}$. Note that the elements of the FE space $\mathcal{V}_N^\varepsilon$ have the form

$$u^\varepsilon_{FE}(x) = \sum_{i,I} N_i(x) \phi_I\left(\frac{x}{\varepsilon}\right) \quad \forall\, x \in \Omega_\varepsilon$$

with shape functions $N_i(\cdot) \in S^p(\Omega, \mathcal{T}_H)$, and $\phi_I(\cdot) \in S^\mu_{\rm per}(\widehat{Q}, \widehat{\mathcal{T}}_h, \mathbf{F})$. Note also that $N_i(\cdot)$ are defined everywhere in Ω, whereas $\phi_I(\frac{\cdot}{\varepsilon})$ are defined only in Ω_ε.

3.1.4 Finite Element approximation results

We start with finite dimensional approximations with respect to the macro FE space $S^p(\Omega, \mathcal{T}_H)$ in $\Omega = (0,1)^n$, $n = 2, 3$. Let $\Pi_{p,\mathcal{T}_H}$ denote the piecewise polynomial interpolant of degree $p \geq 1$ given by $\Pi_{p,\mathcal{T}_H} u|_K = \widehat{\Pi}_p(u|_K \circ F_K) \circ F_K^{-1}$ in each element $K \in \mathcal{T}_H$ with F_K being the associated affine element mapping. Affine transformations of the elements in addition to the local estimates from Lemma 3.4 give

Lemma 3.6 *For any $u \in H^2(\Omega)$, $\Omega \subset \mathbb{R}^n$, $n = 2, 3$,*

$$\|u - \Pi_{p,\mathcal{T}_H} u\|_{L^2(\Omega)} + H|u - \Pi_{p,\mathcal{T}_H} u|_{H^1(\Omega)} \leq C \sum_{K \in \mathcal{T}_H} H^{s_K+1} \Phi_n(p, s_K) |u|_{H^{s_K+1}(K)} \tag{3.5}$$

for $n - 1 \leq s_K \leq p$ such that the right hand side in (3.5) is finite. The constant $C > 0$ is independent of p, s_K and H.

Remark 3.7 *If $n = 3$ and $p = 1$, then for all $u \in \mathcal{H}^1(\Omega)$*

$$\|u - \Pi_{1,\mathcal{T}_H} u\|_{L^2(\Omega)} + H|u - \Pi_{1,\mathcal{T}_H} u|_{H^1(\Omega)} \leq CH^2 |u|_{\mathcal{H}^1(\Omega)}, \tag{3.6}$$

where $C > 0$ is independent of h and we denoted by $|u|^2_{\mathcal{H}^1(\Omega)} := \sum_{\substack{0 \leq \alpha_j \leq 1 \\ |\alpha| > 0}} \|D^\alpha u\|^2_{L^2(\Omega)}$.

Remark 3.8 Estimates (3.5) for the interpolant $\Pi_{p,\mathcal{T}_H}$ are explicit in H, p and s_K. If only H-dependence is of interest, other interpolants, e.g., of Clément type, could be used.

In order to obtain similar FE approximation results with respect to the FE space $S^\mu_{\text{per}}(\widehat{Q}, \widehat{\mathcal{T}}, \mathbf{F})$ in $\widehat{Q}$ we define the piecewise polynomial interpolant $\mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} \in S^\mu_{\text{per}}(\widehat{Q}, \widehat{\mathcal{T}}, \mathbf{F})$ as given by

$$\mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} u|_{\widehat{Q}_i} = \left(\Pi_{\mu,\widehat{\widehat{\mathcal{T}}}_h} \left(u|_{\widehat{Q}_i} \circ F_i\right)\right) \circ F_i^{-1}.$$

Then a similar estimate as in Lemma 3.6 for the interpolation error $u - \mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} u$ holds.

Lemma 3.9 *For $n = 2, 3$ and for $u \in H^2_{\text{per}}(\widehat{Q})$ there exists a $\widehat{Q}$ periodic interpolant $\mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} u \in S^\mu_{\text{per}}(\widehat{Q}, \widehat{\mathcal{T}}, \mathbf{F})$ such that*

$$\|u - \mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} u\|_{L^2(\widehat{Q})} + h|u - \mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} u|_{H^1(\widehat{Q})} \leq Ch^{\min(\mu,s)+1} \Phi_n(\mu, s) \sum_{i=1}^{I} \|\hat{u}_i\|_{H^{s+1}(\widehat{Q}_i)}, \quad \hat{u}_i = u|_{\widehat{Q}_i} \tag{3.7}$$

for all $n - 2 \leq s, \mu$ such that the right hand side in (3.7) is finite. The constant $C > 0$ is independent of μ, s and h. If $n = 3$ and $\mu = 1$, then for all $u \in \mathcal{H}^1(\Omega)$ there holds

$$\|u - \mathcal{I}_{1,\widehat{\mathcal{T}}_h} u\|_{L^2(\Omega)} + h|u - \mathcal{I}_{1,\widehat{\mathcal{T}}_h} u|_{H^1(\Omega)} \leq Ch^2 |u|_{\mathcal{H}^1(\Omega)}. \tag{3.8}$$

Proof. The result is a direct consequence of the definition of the interpolation operator $\mathcal{I}_{\mu,\widehat{\mathcal{T}}_h}$ with respect to $S^\mu_{\text{per}}(\widehat{Q}, \widehat{\mathcal{T}}_h, \mathbf{F})$ and Lemma 3.6. We prove only the estimate (3.7) since (3.8) can be obtained similarly.

$$\begin{aligned}
\|u - \mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} u\|^2_{L^2(\widehat{Q})} + h|u - \mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} u|^2_{H^1(\widehat{Q})} &= \sum_{i=1}^{I} \|u - \mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} u\|^2_{L^2(\widehat{Q}_i)} + h|u - \mathcal{I}_{\mu,\widehat{\mathcal{T}}_h} u|^2_{H^1(\widehat{Q}_i)} \\
&\leq C \sum_i \|u \circ F_i - \mathcal{I}_{\mu,\widehat{\mathcal{T}}_h}(u \circ F_i)\|^2_{L^2([0,1]^n)} + h|u \circ F_i - \mathcal{I}_{\mu,\widehat{\mathcal{T}}_h}(u \circ F_i)|^2_{H^1([0,1]^n)} \\
&= C \sum_i \|u \circ F_i - \Pi_{\mu,\widehat{\widehat{\mathcal{T}}}_h}(u \circ F_i)\|^2_{L^2([0,1]^n)} + h|u \circ F_i - \Pi_{\mu,\widehat{\widehat{\mathcal{T}}}_h}(u \circ F_i)|^2_{H^1([0,1]^n)} \\
&\leq Ch^{2\min(s,\mu)+2} \Phi_n^2(\mu, s) \sum_i \|u|_{\widehat{Q}_i}\|^2_{H^{s+1}(\widehat{Q}_i)}.
\end{aligned}$$

□

3.2 Two-Scale Finite Element Convergence

By (1.9), it holds

$$\|u^\varepsilon - u^\varepsilon_{FE}\|_{H^1(\Omega_\varepsilon)} \leq C \inf_{v \in \mathcal{V}^\varepsilon_N} \|u^\varepsilon - v\|_{H^1(\Omega_\varepsilon)}.$$

The goal of this section is to estimate the approximation error for the two-scale FE space $\mathcal{V}^\varepsilon_N$ (3.3)–(3.4) and to obtain robust estimates with respect to ε for $H/\varepsilon \geq 1$.

We have seen that the solution u^ε may be interpreted as $u^\varepsilon(\cdot) = \mathcal{R}U^\varepsilon(\cdot, \cdot/\varepsilon)$, where $U^\varepsilon(\cdot,\cdot)$ is defined on $\Omega \times \widehat{Q}$. This suggests to use hp-interpolants in Ω (not Ω_ε!) and $\widehat{Q}$ to approximate U^ε in $\Omega \times \widehat{Q}$ and take traces.

For each element $K \in \mathcal{T}_H$ of the 'macro' triangulation, define $U^{\varepsilon,K}(\hat{x}, y) := U^\varepsilon(F_K(\hat{x}), y)$, with $F_K : \widehat{K} \to K$ being the affine element map of K to the reference element $\widehat{K} = [0,1]^n$. Then, the interpolation error $e^\varepsilon_{\mathcal{I}}$ has the form $e^\varepsilon_{\mathcal{I}}(\cdot) := \mathcal{R}E^\varepsilon_{\mathcal{I}}(\cdot, \cdot/\varepsilon)$, in which $E^\varepsilon_{\mathcal{I}}(x,y) = U^\varepsilon(x,y) - U^\varepsilon_{\mathcal{I}}(x,y)$, $x \in \Omega$, $y \in \Omega^\infty_\varepsilon$, is given by

$$E^\varepsilon_{\mathcal{I}}(F_K(\hat{x}), y) := U^{\varepsilon,K}(\hat{x}, y) - \Pi^{\hat{x}}_p \otimes \mathcal{I}^y_{\mu,\widehat{\mathcal{T}}_h} U^{\varepsilon,K}(\hat{x}, y), \quad \forall K \in \mathcal{T}_H,$$

with $\widehat{\Pi}^{\hat{x}}_p$ being the p interpolant in each reference element $\widehat{K}$ and $\mathcal{I}^y_{\mu,\widehat{\mathcal{T}}_h}$ the $S^\mu_{\mathrm{per}}(\widehat{Q}, \widehat{\mathcal{T}}_h, \mathbf{F})$ interpolant in $H^1_{\mathrm{per}}(\widehat{Q})$. If H denotes the mesh size of the quasiuniform 'macroscopic' triangulation on Ω and h is the mesh size of the quasiuniform 'micro' triangulation on the unit cell $\widehat{Q}$, we obtain

Proposition 3.10 *Assume that $n = 2$. For $p, \mu, k, s \geq 1$ and $H/\varepsilon \in \mathbb{N}$ in (3.3)–(3.4) it holds*

$$\|e^\varepsilon_{\mathcal{I}}\|_{L^2(\Omega_\varepsilon)} \leq C\left(H^{\min(p,k)+1}\Phi_n(p,k)\|U^\varepsilon\|_{H^{k+1}(\Omega;\, L^2_{\mathrm{per}}(\widehat{Q}))} + h^{\min(\mu,s)+1}\Phi_n(\mu,s)\|U^\varepsilon\|_{H^n(\Omega;\, H^{s+1}_{\mathrm{per}}(\widehat{Q}))}\right),$$

where $C > 0$ is a positive constant independent of p, μ, k, s and ε.

Proof. Let $K = F_K(\widehat{K}) \in \mathcal{T}_H$ be an element of the 'macro' triangulation, affine image of the reference element $\widehat{K}$ under the element mapping F_K. We split the interpolation error as

$$E^\varepsilon_{\mathcal{I}}(F_K(\hat{x}), y) := U^{\varepsilon,K}(\hat{x}, y) - \widehat{\Pi}^{\hat{x}}_p U^{\varepsilon,K}(\hat{x}, y) + \widehat{\Pi}^{\hat{x}}_p U^{\varepsilon,K}(\hat{x}, y) - (\widehat{\Pi}^{\hat{x}}_p \otimes \mathcal{I}^y_{\mu,\widehat{\mathcal{T}}_h}) U^{\varepsilon,K}(\hat{x}, y).$$

We estimate first the L^2 norm of the error on K and apply the trace result in Lemma 3.1 to move on full two scale interpolation error estimates

$$\begin{aligned}
\int_{K\cap\Omega_\varepsilon} |e^\varepsilon_{\mathcal{I}}(x)|^2\, dx &= \varepsilon^n \sum_{m\in\mathbb{Z}^n:\varepsilon(\widehat{Q}+m)\subset K} \int_{\widehat{Q}} \left| E^\varepsilon_{\mathcal{I}}(\varepsilon(z+m), y)\Big|_{z=y} \right|^2 dz \\
&\leq C\varepsilon^n \sum_{m\in\mathbb{Z}^n:\varepsilon(\widehat{Q}+m)\subset K} \sum_{0\leq\alpha_j\leq 1} \varepsilon^{2|\alpha|} \int_{\widehat{Q}\times\widehat{Q}} |(D^\alpha_x E^\varepsilon_{\mathcal{I}})(\varepsilon(z+m), y)|^2\, dz dy \\
&= C \sum_{0\leq\alpha_j\leq 1} \varepsilon^{2|\alpha|} \int_{(K\cap\Omega_\varepsilon)\times\widehat{Q}} |D^\alpha_x E^\varepsilon_{\mathcal{I}}(x,y)|^2\, dx dy \\
&\leq C \sum_{0\leq\alpha_j\leq 1} \varepsilon^{2|\alpha|} \int_{K\times\widehat{Q}} |D^\alpha_x E^\varepsilon_{\mathcal{I}}(x,y)|^2\, dx dy \\
&\leq CH^n \sum_{0\leq\alpha_j\leq 1} \left(\frac{\varepsilon}{H}\right)^{2|\alpha|} \int_{\widehat{K}\times\widehat{Q}} |D^\alpha_{\hat{x}} E^\varepsilon_{\mathcal{I}}(F_K(\hat{x}), y)|^2\, d\hat{x} dy \\
&\leq CH^n\left(\mathrm{I}_K + \mathrm{II}_K\right),
\end{aligned}$$

where

$$\mathrm{I}_K = \int\limits_{\widehat{K}\times\widehat{Q}} \sum_{0\le\alpha_j\le 1} \left| D^{\alpha}_{\hat{x}} \left(U^{\varepsilon,K}(\hat{x},y) - \widehat{\Pi}^{\hat{x}}_p U^{\varepsilon,K}(\hat{x},y) \right) \right|^2 \, d\hat{x}dy$$

$$\mathrm{II}_K = \int\limits_{\widehat{K}\times\widehat{Q}} \sum_{0\le\alpha_j\le 1} \left| D^{\alpha}_{\hat{x}} \left(\widehat{\Pi}^{\hat{x}}_p U^{\varepsilon,K}(\hat{x},y) - (\widehat{\Pi}^{\hat{x}}_p \otimes \mathcal{I}^y_{\mu,\widehat{\mathcal{T}}_h}) U^{\varepsilon,K}(\hat{x},y) \right) \right|^2 \, d\hat{x}dy.$$

By Lemma 3.4, the 'macro' error I_K can be estimated as follows

$$\begin{aligned}
\mathrm{I}_K &= \int\limits_{\widehat{K}\times\widehat{Q}} \sum_{0\le\alpha_j\le 1} \left| D^{\alpha}_{\hat{x}} \left(U^{\varepsilon,K}(\hat{x},y) - \widehat{\Pi}^{\hat{x}}_p U^{\varepsilon,K}(\hat{x},y) \right) \right|^2 \, d\hat{x}dy \\
&\le C\Phi^2_n(p,k) \int\limits_{\widehat{K}\times\widehat{Q}} \left| D^{k+1}_{\hat{x}} U^{\varepsilon,K}(\hat{x},y) \right|^2 \, d\hat{x}dy \\
&\le CH^{2(k+1)}\Phi^2_n(p,k) \int\limits_{\widehat{K}\times\widehat{Q}} \left| \left(D^{k+1}_{x} U^{\varepsilon} \right)(F_K(\hat{x}),y) \right|^2 \, d\hat{x}dy \\
&\le CH^{2(k+1)-n}\Phi^2_n(p,k) \int\limits_{K\times\widehat{Q}} \left| \left(D^{k+1}_{x} U^{\varepsilon} \right)(x,y) \right|^2 \, dxdy.
\end{aligned}$$

Applying now the error estimates in Lemma 3.9 for the interpolation error in the 'micro' FE space $S^{\mu}_{\mathrm{per}}(\widehat{Q},\widehat{\mathcal{T}}_h,\mathbf{F})$, the error II_K in the fast variable can be estimated as follows

$$\begin{aligned}
\mathrm{II}_K &= \int\limits_{\widehat{K}\times\widehat{Q}} \sum_{0\le\alpha_j\le 1} \left| D^{\alpha}_{\hat{x}} \left(\widehat{\Pi}^{\hat{x}}_p U^{\varepsilon,K}(\hat{x},y) - (\widehat{\Pi}^{\hat{x}}_p \otimes \mathcal{I}^y_{\mu,\widehat{\mathcal{T}}_h}) U^{\varepsilon,K}(\hat{x},y) \right) \right|^2 \, d\hat{x}dy \\
&\le C \sum_{0\le\alpha_j\le 1} \int\limits_{\widehat{K}\times\widehat{Q}} \left| D^{|\alpha|}_{\hat{x}} \left(U^{\varepsilon,K}(\hat{x},y) - \mathcal{I}^y_{\mu,\widehat{\mathcal{T}}_h} U^{\varepsilon,K}(\hat{x},y) \right) \right|^2 \, d\hat{x}dy \\
&\le CH^{-n} \sum_{0\le\alpha_j\le 1} \int\limits_{K\times\widehat{Q}} \left| D^{|\alpha|}_{x} \left(U^{\varepsilon}(x,y) - \mathcal{I}^y_{\mu,\widehat{\mathcal{T}}_h} U^{\varepsilon}(x,y) \right) \right|^2 \, dxdy \\
&\le CH^{-n}h^{2\min(\mu,s)+2}\Phi^2_2(\mu,s)\|U^{\varepsilon}\|^2_{H^n(K;\,H^{s+1}_{\mathrm{per}}(\widehat{Q}))}.
\end{aligned}$$

Summing up over all elements $K \in \mathcal{T}_H$ we obtain that

$$\|e^{\varepsilon}_{\mathcal{I}}\|_{L^2(\Omega_\varepsilon)} \le C\left(H^{\min(p,k)+1}\Phi_n(p,k)\|U^{\varepsilon}\|_{H^{k+1}(\Omega;\,L^2_{\mathrm{per}}(\widehat{Q}))} + h^{\min(\mu,s)+1}\Phi_n(\mu,s)\|U^{\varepsilon}\|_{H^n(\Omega;\,H^{s+1}_{\mathrm{per}}(\widehat{Q}))} \right).$$

□

A similar result can be derived now for the energy norm of the two scale interpolation error. To this end, we estimate the $L^2(\Omega)$-norm of $\nabla_x e^{\varepsilon}_{\mathcal{I}}$ in terms of the regularity of the data and of the 'macro', resp. 'micro' triangulations .

Proposition 3.11 *Assume that $n = 2$, $k,s \ge 1$ and $H/\varepsilon \in \mathbb{N}$. Then it holds for any $p,\mu \ge 1$*

$$\begin{aligned}
\|\nabla_x e^{\varepsilon}_{\mathcal{I}}(x)\|_{L^2(\Omega_\varepsilon)} &\le CH^{\min(p,k)}\Phi_n(p,k)\left(\|\varepsilon^{-1}\nabla_y U^{\varepsilon}\|_{H^k(\Omega;\,L^2_{\mathrm{per}}(\widehat{Q}))} + \|U^{\varepsilon}\|_{H^{k+1}(\Omega;\,L^2_{\mathrm{per}}(\widehat{Q}))} \right) \\
&+ Ch^{\min(\mu,s)}\Phi_n(\mu,s)\left(\|\varepsilon^{-1}\nabla_y U^{\varepsilon}\|_{H^n(\Omega;\,H^s_{\mathrm{per}}(\widehat{Q}))} + \|U^{\varepsilon}\|_{H^n(\Omega;H^{s+1}_{\mathrm{per}}(\widehat{Q})} \right).
\end{aligned} \tag{3.9}$$

Proof. Let $K \in \mathcal{T}_H$ be an arbitrary element of the 'macro' triangulation and consider the $H^1(K \cap \Omega_\varepsilon)$ seminorm of the interpolation error $e^\varepsilon_{\mathcal{I}}$ on $K \cap \Omega_\varepsilon$. Then it holds

$$\begin{aligned}\|\nabla_x e^\varepsilon_{\mathcal{I}}(x)\|^2_{L^2(K\cap\Omega_\varepsilon)} &= \int\limits_{K\cap\Omega_\varepsilon} |\nabla_x e^\varepsilon_{\mathcal{I}}(x)|^2\,dx = \int\limits_{K\cap\Omega_\varepsilon} \left| \left((\nabla_x + \varepsilon^{-1}\nabla_y)E^\varepsilon_{\mathcal{I}}\right)(x,y)\Big|_{y=\frac{x}{\varepsilon}} \right|^2 dx \\ &\le \mathrm{I}_K + \mathrm{II}_K,\end{aligned}$$

where

$$\mathrm{I}_K = \int\limits_{K\cap\Omega_\varepsilon} \left| \left(\nabla_x E^\varepsilon_{\mathcal{I}}\right)(x,y)\Big|_{y=\frac{x}{\varepsilon}} \right|^2 dx, \quad \mathrm{II}_K = \int\limits_{K\cap\Omega_\varepsilon} \left| \left(\varepsilon^{-1}\nabla_y E^\varepsilon_{\mathcal{I}}\right)(x,y)\Big|_{y=\frac{x}{\varepsilon}} \right|^2 dx.$$

By the trace result in Lemma 3.1 we obtain that

$$\begin{aligned}\mathrm{I}_K &= \varepsilon^n \sum_{r=1}^n \sum_{m:\varepsilon(m+\widehat{Q})\subset K} \int\limits_{\widehat{Q}} \left| (\partial_{x_r} E^\varepsilon_{\mathcal{I}})(\varepsilon(z+m),y)\Big|_{y=z} \right|^2 dz \\ &\le C\varepsilon^n \sum_{r=1}^n \sum_{m:\varepsilon(m+\widehat{Q})\subset K} \sum_{\substack{0\le\alpha_j\le 1\\ \alpha_r=1}} \varepsilon^{2|\alpha|-2} \int\limits_{\widehat{Q}}\int\limits_{\widehat{Q}} \Big[|(D^\alpha_x \partial_{y_r} E^\varepsilon_{\mathcal{I}})(\varepsilon(z+m),y)|^2 \\ &\qquad + |(D^\alpha_x E^\varepsilon_{\mathcal{I}})(\varepsilon(z+m),y)|^2 \Big] dz dy\end{aligned}$$

$$\begin{aligned} &= C\sum_{r=1}^n \sum_{\substack{0\le\alpha_j\le 1\\ \alpha_r=1}} \varepsilon^{2|\alpha|-2} \int\limits_{K\cap\Omega_\varepsilon}\int\limits_{\widehat{Q}} \Big[|(D^\alpha_x \partial_{y_r} E^\varepsilon_{\mathcal{I}})(x,y)|^2 + |(D^\alpha_x E^\varepsilon_{\mathcal{I}})(x,y)|^2 \Big] dxdy \\ &\le C\sum_{r=1}^n \sum_{\substack{0\le\alpha_j\le 1\\ \alpha_r=1}} \varepsilon^{2|\alpha|-2} \int\limits_{K}\int\limits_{\widehat{Q}} \Big[|(D^\alpha_x \partial_{y_r} E^\varepsilon_{\mathcal{I}})(x,y)|^2 + |(D^\alpha_x E^\varepsilon_{\mathcal{I}})(x,y)|^2 \Big] dxdy \\ &\le CH^n \sum_{r=1}^n \sum_{\substack{0\le\alpha_j\le 1\\ \alpha_r=1}} \frac{\varepsilon^{2|\alpha|}}{H^{2|\alpha|}} \int\limits_{\widehat{K}}\int\limits_{\widehat{Q}} \left| D^\alpha_{\hat{x}}(\varepsilon^{-1}\partial_{y_r} E^\varepsilon_{\mathcal{I}})(F_K(\hat{x}),y)\right|^2 d\hat{x}dy \\ &+ CH^n \sum_{r=1}^n \sum_{\substack{0\le\alpha_j\le 1\\ \alpha_r=1}} \frac{\varepsilon^{2|\alpha|-2}}{H^{2|\alpha|-2}} H^{-2} \int\limits_{\widehat{K}}\int\limits_{\widehat{Q}} |D^\alpha_{\hat{x}} E^\varepsilon_{\mathcal{I}}(F_K(\hat{x}),y)|^2\, d\hat{x}dy \\ &\le CH^n \sum_{r=1}^n \sum_{\substack{0\le\alpha_j\le 1\\ \alpha_r=1}} \int\limits_{\widehat{K}}\int\limits_{\widehat{Q}} \left| D^\alpha_{\hat{x}}(\varepsilon^{-1}\partial_{y_r} E^\varepsilon_{\mathcal{I}})(F_K(\hat{x}),y)\right|^2 + H^{-2}\,|D^\alpha_{\hat{x}} E^\varepsilon_{\mathcal{I}}(F_K(\hat{x}),y)|^2\, d\hat{x}dy.\end{aligned}$$

By the same arguments as in the proof of Proposition 3.10 we obtain that

$$\begin{aligned}\mathrm{I}_K &\le C_1 H^{2\min(p,k)} \Phi^2_n(p,k) \Bigg(\int\limits_{K\times\widehat{Q}} \left| D^k_x(\varepsilon^{-1}\nabla_y U^\varepsilon)(x,y)\right|^2\, dxdy + \int\limits_{K\times\widehat{Q}} \left| D^{k+1}_x U^\varepsilon(x,y)\right|^2\, dxdy \Bigg) \\ &+ C_2 h^{2\min(\mu,s)} \Phi^2_n(\mu,s) \Big(\|\varepsilon^{-1}D_y U^\varepsilon\|^2_{H^n(K;H^s_{\mathrm{per}}(\widehat{Q}))} + \|U^\varepsilon\|^2_{H^n(K;H^{s+1}_{\mathrm{per}}(\widehat{Q}))} \Big).\end{aligned}$$

Similar considerations for II_K lead to the following estimate

$$\begin{aligned}
\mathrm{II}_K &= \varepsilon^n \sum_{m\in\mathbb{Z}^n:\varepsilon(\widehat{Q}+m)\subset K} \int_{\widehat{Q}} \left|(\varepsilon^{-1}\nabla_y E^\varepsilon_{\mathcal{I}})(\varepsilon(z+m),y)\right|_{z=y}\Big|^2 dz \\
&\le C\varepsilon^n \sum_{m\in\mathbb{Z}^n:\varepsilon(\widehat{Q}+m)\subset K} \sum_{0\le\alpha_j\le 1} \varepsilon^{2|\alpha|} \int_{\widehat{Q}\times\widehat{Q}} |(D^\alpha_x E^\varepsilon_{\mathcal{I}})(\varepsilon(z+m),y)|^2\, dzdy \\
&= C \sum_{0\le\alpha_j\le 1} \varepsilon^{2|\alpha|} \int_{(K\cap\Omega_\varepsilon)\times\widehat{Q}} \left|D^\alpha_x(\varepsilon^{-1}\nabla_y E^\varepsilon_{\mathcal{I}})(x,y)\right|^2 dxdy \\
&\le C \sum_{0\le\alpha_j\le 1} \varepsilon^{2|\alpha|} \int_{K\times\widehat{Q}} \left|D^\alpha_x(\varepsilon^{-1}\nabla_y E^\varepsilon_{\mathcal{I}})(x,y)\right|^2 dxdy \\
&\le CH^n \sum_{0\le\alpha_j\le 1} \left(\frac{\varepsilon}{H}\right)^{2|\alpha|} \int_{\widehat{K}\times\widehat{Q}} \left|D^\alpha_{\hat{x}}(\varepsilon^{-1}\nabla_y E^\varepsilon_{\mathcal{I}})(F_K(\hat{x}),y)\right|^2 d\hat{x}dy \\
&\le C_1 H^{2\min(p,k)}\Phi_n^2(p,k) \int_{K\times\widehat{Q}} \left|\left(D^k_x(\varepsilon^{-1}\nabla_y U^\varepsilon)\right)(x,y)\right|^2\, dxdy \\
&+ C_2 h^{2\min(\mu,s)}\Phi_n^2(\mu,s)\|\varepsilon^{-1}\nabla_y U^\varepsilon\|^2_{H^n(K;\,H^s_{\mathrm{per}}(\widehat{Q}))}.
\end{aligned}$$

Summing up over all elements K of the 'macro' triangulation we obtain (3.9).

□

Remark 3.12 A careful inspection of the proofs of Propositions 3.10, 3.11 reveals that $\|U^\varepsilon\|_{H^n(\Omega;\,H^{s+1}_{\mathrm{per}}(\widehat{Q}))}$ and $\|\varepsilon^{-1}\nabla_y U^\varepsilon\|_{H^n(\Omega;\,H^s_{\mathrm{per}}(\widehat{Q}))}$ can be replaced by the weaker norms $\|U^\varepsilon\|_{\mathcal{H}^1(\Omega;\,H^{s+1}_{\mathrm{per}}(\widehat{Q}))}$ and $\|\varepsilon^{-1}\nabla_y U^\varepsilon\|_{\mathcal{H}^1(\Omega;\,H^s_{\mathrm{per}}(\widehat{Q}))}$ where for a Banach space X we denote

$$\mathcal{H}^1(\Omega;\,X) := \{u:\Omega\to X \;:\; D^\alpha_x u\in L^2(\Omega;\,X) \quad \forall\,\alpha \text{ s.t. } \max_j \alpha_j\le 1\}.$$

Remark 3.13 Similar error estimates for the interpolation error as in Propositions 3.10, 3.11 can be obtained in the case $n=3$ by using the appropriate regularity assumptions on $U^\varepsilon(x,y)$. We restrict here ourselves to the two-dimensional setting.

Theorem 3.14 *Assume for the solution u^ε of (1.7) the two-scale regularity (2.1)–(2.2) in Ω_ε. Then, for $H/\varepsilon\in\mathbb{N}$, the error in the two-scale FEM based on the space (3.3)–(3.4) can be estimated as follows:*

$$\|u^\varepsilon - u^\varepsilon_{FE}\|_{H^1(\Omega_\varepsilon)} \le C_1(k)H^{\min(p,k)}\Phi_n(p,k)\|f\|_{H^k(\Omega)} + C_2(s)h^{\min(\mu,s)}\Phi_n(\mu,s)\|f\|_{H^{n+s}(\Omega)}.$$

Proof. The proof is a direct consequence of Theorem 2.1 and Propositions 3.10, 3.11.

□

Remark 3.15 Suppose that the solution $U^\varepsilon(x,y)$ is patch-wise analytic on the 'macro' level and analytic on the 'micro' scale. Then, there exist $C>0$, $d_K>0$ independent of ε such that for all k

$$\begin{aligned}
\|D^k_x U^\varepsilon(x,y)\|_{L^2(K;\,L^2(\widehat{Q}))} &\le C(d_K)^k k!\,|K|^{1/2} \\
\|\varepsilon^{-1}\nabla_y D^k_x U^\varepsilon(x,y)\|_{L^2(K;\,L^2(\widehat{Q}))} &\le C(d_K)^k k!\,|K|^{1/2}.
\end{aligned}$$

In this case the estimates in Propositions 3.10, 3.11 lead to exponential convergence. Applying Stirling's formula after taking $k-1=\alpha p$ ($0<\alpha<1$ will be selected below) we obtain that

$$\begin{aligned}\Phi_2(p,k)\|D_x^k U^\varepsilon(x,y)\|_{L^2(K;L^2(\widehat{Q}))} &\leq CH(d_K)^k k!\left(\frac{(p-k+1)!}{(p+k-1)!}\right)^{1/2}\\ &\leq CpH(d_K)^{\alpha p}(\alpha p)^{\alpha p}\left(\frac{((1-\alpha)p)^{(1-\alpha)p}}{((1+\alpha)p)^{(1+\alpha)p}}\right)^{1/2}\\ &= CpH(F(\alpha,d_K))^p,\end{aligned}$$

where

$$F(\alpha,d_K)=(\alpha d_K)^{\alpha}\left(\frac{(1-\alpha)^{1-\alpha}}{(1+\alpha)^{1+\alpha}}\right)^{1/2}.$$

Now, for $d>1$, there exists $\alpha_{\min}(d)$ such that

$$\min_{0<\alpha<1}F(d,\alpha)=F(d,\alpha_{\min})<1,\quad \alpha_{\min}=\frac{1}{\sqrt{1+d^2}}.$$

Then, taking $\tilde{b}_K=|\log F(d_K,\alpha_{\min})|$, it follows that

$$\Phi_2(p,k)\|D_x^k U^\varepsilon(x,y)\|_{L^2(K;L^2(\widehat{Q}))}\leq CpHe^{-\tilde{b}_K p}\leq CHe^{-b_K p}\quad 0<b_K<\tilde{b}_K.$$

Summing up now the contribution in the energy error from the terms of the form

$$H^{2\min(p,k)}\Phi_2^2(p,k)\|D_x^k U^\varepsilon(x,y)\|^2_{L^2(K;L^2(\widehat{Q}))}$$

gives us

$$\sum_{K\in\mathcal{T}_H}H^{2k+2}e^{-2b_K p}\sim H^{2k}e^{-2bp},\quad b=\min_{K\in\mathcal{T}}b_K$$

showing that the interpolation error in the energy from the 'macro' interpolant decays exponentially in p, uniformly in ε. Analogous analysis can be carried out with respect to the 'micro' scale interpolation error. This shows that the contribution to the error from the terms

$$\sum_{K\in\mathcal{T}}\|U^\varepsilon(x,y)-\mathcal{I}_{h,y}U^\varepsilon(x,y)\|^2_{H^n(K;L^2(\widehat{Q}))}+\|\varepsilon^{-1}\nabla_y U^\varepsilon(x,y)-\mathcal{I}_{h,y}\varepsilon^{-1}\nabla_y U^\varepsilon(x,y)\|^2_{H^n(K;L^2(\widehat{Q}))}$$

is also exponentially decaying in μ and has the form $h^{2s}e^{-2a\mu}$, with $a>0$.

Remark 3.16 So far we have only discussed the preasymptotic case when $H\geq\varepsilon$ and we obtained the robust (in ε) error estimate $\|e_N^\varepsilon\|_{H^1(\Omega_\varepsilon)}\leq C(H^p+h^\mu)$.
Let us choose $p=\mu$. Then, for $h\cong H$ we obtain $\|e_N^\varepsilon\|_{H^1(\Omega_\varepsilon)}\leq Ch^p$ and $N=\dim(\mathcal{V}_N^\varepsilon)=O(H^{-n}h^{-n})=O(h^{-2n})$. Hence, in terms of number of degrees of freedom, the two-scale FE error estimate is qualitatively of the form

$$\|e_N^\varepsilon\|_{H^1(\Omega_\varepsilon)}\leq CN^{-\frac{p}{2n}}\quad\text{for } N\preceq\varepsilon^{-2n},$$

since the total number of degrees of freedom at the critical value $H\cong\varepsilon$, $h\cong\varepsilon$ is $N=O(\varepsilon^{-2n})$. At this transition point the fine scale is resolved and we switch from the two scale FE space to full discretization with mesh size $H=\varepsilon h$, $h\leq\varepsilon$. This is achieved by breaking the periodicity

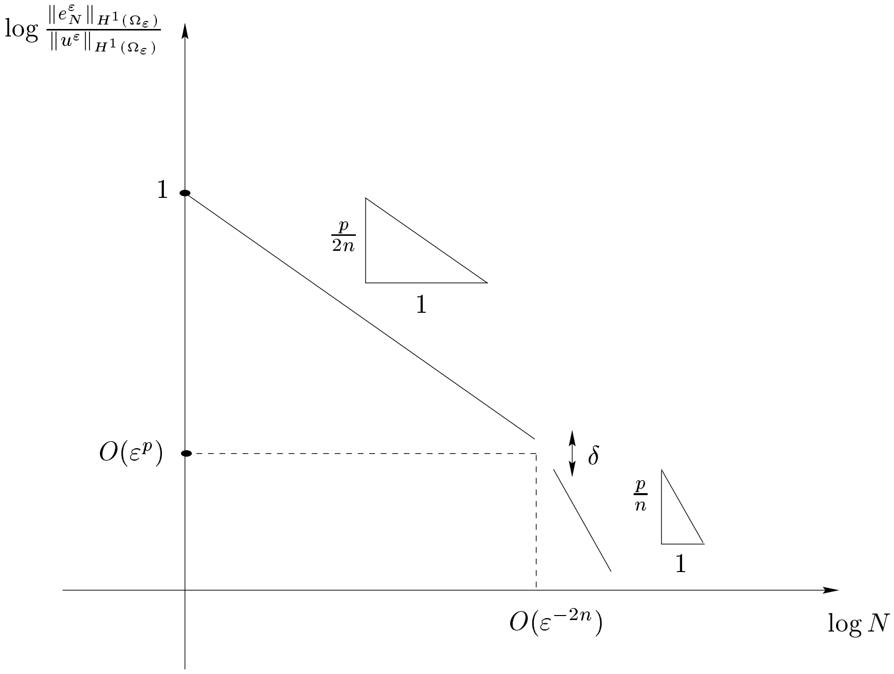

Figure 2: Qualitative picture on two-scale convergence as $H = h \to 0$ for fixed $p = \mu \geq 1$: relative error versus # dof in double logarithmic scale. At the critical point $N = O(\varepsilon^{-2n})$ the error slopes switch from $\frac{p}{2n}$ to $\frac{p}{n}$, the 'jump' δ is independent of ε and is due to possibly different constants in the *a-priori* estimates.

to get the full space with mesh width $H = \varepsilon h$, $h \leq \varepsilon$. The dimension of the FE space in this asymptotic regime is $N = O(H^{-n}) = O(\varepsilon^{-n} h^{-n})$. Using standard error estimates, i.e.,

$$\|e_N^\varepsilon\|_{H^1(\Omega_\varepsilon)} \leq CH^p \|u^\varepsilon\|_{H^{p+1}(\Omega_\varepsilon)}$$

and the *a-priori* estimate (1.10) for u^ε we obtain

$$\|e_N^\varepsilon\|_{H^1(\Omega_\varepsilon)} \leq C \left(\frac{H}{\varepsilon}\right)^p = Ch^p \leq C\varepsilon^{-p} N^{-\frac{p}{n}} \leq CN^{-\frac{p}{2n}} \quad \text{for } N \succeq \varepsilon^{-2n}.$$

We see that we obtain a robust convergence rate of $O(N^{-\frac{p}{2n}})$, as compared to the (non-robust) rate of $O(N^{-\frac{p}{n}})$ of standard FEM. The robustness of the two-scale FEM was achieved by an increase in dimension and the use of tensor product approximations in $\Omega \times \widehat{Q}$.

4 Implementation of the Two-Scale FEM

We address now the implementation of the 2-scale FEM. In order to obtain an efficient algorithm it is essential that the element stiffness and mass matrices can be computed in a complexity independent of ε and to an accuracy which will not compromise the asymptotic convergence rates in Theorem 3.14. Due to the rapid oscillations of the coefficients and of the micro-shapefunctions, the elemental stiffness matrices on the macro mesh can *not* be evaluated robustly by standard quadratures. The macro stiffness and mass matrices can be developed from *moments*, i.e., integrals in the fast variable corresponding to discretization of the unit-cell problem with monomial weighted coefficients, combined with certain lattice summation formulas. To explain this is the object of Section 4.1. In Section 4.2 we then present numerical experiments which confirm our error analysis.

Proposition 4.1 *For any $\varepsilon > 0$ and for any finite dimensional subspace $\mathcal{M}^{\mu}_{\text{per}}(\widehat{Q})$ of $H^1_{\text{per}}(\widehat{Q})$, with $\mathcal{M}^{\mu}_{\text{per}}(\widehat{Q}) = \text{Span}\,\{\Phi_i(y)\}_{i=1}^{\mu}$ of dimension μ independent of ε, the two-scale FEM with respect to the two-scale discretization $S^p(\Omega, \mathcal{T}_H; \mathcal{M}^{\mu}_{\varepsilon,\text{per}}(\varepsilon\widehat{Q}))$ ($\mathcal{M}^{\mu}_{\varepsilon,\text{per}}(\varepsilon\widehat{Q}) = \text{Span}\,\{\Phi_i(x/\varepsilon)\}_{i=1}^{\mu}$) can be implemented with a computational work independent of ε.*

4.1 Macroelement Stiffness Matrix

We start from the discrete variational formulation: Find $u \in S^p(\Omega, \mathcal{T}_H; \mathcal{M}^{\mu}_{\varepsilon,\text{per}}(\varepsilon\widehat{Q}))$ such that

$$B^{\varepsilon}(u, v) = \int_{\Omega} f v \, dx \quad \forall\, v \in S^p(\Omega, \mathcal{T}_H; \mathcal{M}^{\mu}_{\varepsilon}(\varepsilon\widehat{Q})),$$

where $\mathcal{M}^{\mu}_{\text{per}}(\widehat{Q}) = \text{Span}\,\{\Phi_i\}$ is any conforming FE discretization of $H^1_{\text{per}}(\widehat{Q})$. For $u, v \in S^p(\Omega, \mathcal{T}_H; \mathcal{M}^{\mu}_{\varepsilon}(\varepsilon\widehat{Q}))$ the bilinear form can be split in a sum of elemental bilinear forms B^{ε}_K

$$B^{\varepsilon}(u, v) = \sum_{K \in \mathcal{T}_H} B^{\varepsilon}_K(u, v).$$

For each macro element element $K \in \mathcal{T}_H$ with 'macroscopic' polynomial space $S^p(K) = \text{Span}\,\{\nu_I^{[K]}\}_I$, the elemental bilinear form B_K can be written in terms of the reference element matrix

$$B^{\varepsilon}_K(u, v) = \underline{v}^{\top} \underline{\underline{\mathbf{K}}}^{[K]} \underline{u}, \quad \underline{u} = \{u_{Ii}\}, \quad \underline{v} = \{v_{Ii}\},$$

where $u(x)\Big|_K = \sum_{I,i} u_{Ii} \nu_I^{[K]}(x) \Phi_i(x/\varepsilon)$ and $v(x)\Big|_K = \sum_{J,j} v_{Jj} \nu_J^{[K]}(x) \Phi_j(x/\varepsilon)$. The entries of the element stiffness matrix $\underline{\underline{\mathbf{K}}}^{[K]}$ are given by

$$\begin{aligned} \underline{\underline{\mathbf{K}}}^{[K]}_{(Ii)\,(Jj)} &= \int_K A\left(\frac{x}{\varepsilon}\right) \left(\nu_I^{[K]}(x) \Phi_i\left(\frac{x}{\varepsilon}\right)\right)' \left(\nu_J^{[K]}(x) \Phi_j\left(\frac{x}{\varepsilon}\right)\right)' dx \\ &+ \int_K a_0\left(\frac{x}{\varepsilon}\right) \nu_I^{[K]}(x) \Phi_i\left(\frac{x}{\varepsilon}\right) \nu_J^{[K]}(x) \Phi_j\left(\frac{x}{\varepsilon}\right) dx, \end{aligned} \tag{4.1}$$

where a prime denotes $\frac{d}{dx}$. Without loss of generality we assume now that $K = (0, H)$, with $M := H/\varepsilon \in \mathbb{N}$. For simplicity, we consider only the first integral term in (4.1). Since $K = \cup_{m=0}^{M-1} K_m$, with $K_m = \varepsilon(m + \widehat{Q})$ we obtain that

$$\begin{aligned} \underline{\underline{A}}^{[K]}_{(Ii)\,(Jj)} &= \int_K A\left(\frac{x}{\varepsilon}\right) \left(\nu_I^{[K]}(x) \Phi_i\left(\frac{x}{\varepsilon}\right)\right)' \left(\nu_J^{[K]}(x) \Phi_j\left(\frac{x}{\varepsilon}\right)\right)' dx \\ &= \sum_{m=0}^{M-1} \int_{K_m} A\left(\frac{x}{\varepsilon}\right) \left(\nu_I^{[K]}(x) \Phi_i\left(\frac{x}{\varepsilon}\right)\right)' \left(\nu_J^{[K]}(x) \Phi_j\left(\frac{x}{\varepsilon}\right)\right)' dx \\ &= \sum_{\gamma,\delta \le 1} \sum_{\alpha} c^{IJ}_{\gamma\delta\alpha}\, \varepsilon^{-(\gamma+\delta)} \sum_{m=0}^{M-1} \int_{K_m} A\left(\frac{x}{\varepsilon}\right) \Phi_i^{(\gamma)}\left(\frac{x}{\varepsilon}\right) \Phi_j^{(\delta)}\left(\frac{x}{\varepsilon}\right) x^{\alpha}\, dx \\ &= \sum_{\gamma,\delta \le 1} \sum_{\alpha} c^{IJ}_{\gamma\delta\alpha}\, \varepsilon^{n-(\gamma+\delta)+\alpha} \sum_{m=0}^{M-1} \int_{\widehat{Q}} A(\hat{y}) \Phi_i^{(\gamma)}(\hat{y}) \Phi_j^{(\delta)}(\hat{y})\, (\hat{y} + m)^{\alpha}\, d\hat{y}, \end{aligned}$$

with suitable constants $c^{IJ}_{\gamma\delta\alpha} = c^{IJ}_{\gamma\delta\alpha}(K)$ depending only on $I, J, \alpha, \gamma, \delta$ and on the element K. We see that for the calculation of the two-scale element stiffness matrices the basic integrals

$$\widehat{\underline{\underline{K}}}_{\mu}^{\gamma\delta\tau} = \left(\int_{\widehat{Q}} A(\hat{y}) \Phi_i^{(\gamma)}(\hat{y}) \Phi_j^{(\delta)}(\hat{y}) \hat{y}^{\tau} \, d\hat{y} \right)_{i,j=1,\dots,\mu} \tag{4.2}$$

are needed. Let us remark that (4.2) when $\tau = 0$ and $\delta = \gamma = 1$, corresponds to the global stiffness matrix of the unit cell problem discretized with $\mathcal{M}^{\mu}_{\text{per}} = \text{Span}\{\Phi_i \,|\, i = 1, \dots, \mu\}$. When $\tau > 0$ we obtain a scale interaction stiffness matrix and a discretization of the unit cell problem with monomial weight functions is generally needed. This procedure is suited for parallel implementation, since the computation for various values of τ can be done in parallel. The entries $\underline{\underline{A}}^{[K]}_{(Ii)\,(Jj)}$ of the element stiffness matrix are ultimately given by

$$\begin{aligned}
&\sum_{\gamma,\delta\le 1} \sum_{\alpha} c^{IJ}_{\gamma\delta\alpha} \varepsilon^{n-(\gamma+\delta)+\alpha} \sum_{\tau\le\alpha} \left(\widehat{\underline{\underline{K}}}_{\mu}^{\gamma\delta\tau}\right)_{ij} \binom{\alpha}{\tau} \sum_{m=0}^{M-1} m^{\alpha-\tau} \\
&= \sum_{\gamma,\delta\le 1} \sum_{\alpha} \sum_{\tau\le\alpha} \left(\widehat{\underline{\underline{K}}}_{\mu}^{\gamma\delta\tau}\right)_{ij} \sum_{m=0}^{M-1} S^{IJ}_{\gamma\delta\alpha\tau}(m, H, \varepsilon),
\end{aligned}$$

with $\sum_{m=0}^{M-1} S^{IJ}_{\gamma\delta\alpha\tau}(m, H, \varepsilon)$ being directly computable. The idea is to compute sums of powers of natural numbers appearing in $\sum_{m=0}^{M-1} S^{IJ}_{\gamma\delta\alpha\tau}(m, H, \varepsilon)$ in terms of Bernoulli numbers B_i. These can be easily tabulated and the sums can be computed with a computational work independent of M. More precisely, one exploits the fact that for $N \in \mathbb{N}$, $\sum_{k=1}^{N} k^q$ is given by

$$\sum_{k=1}^{N} k^q = \frac{N^{q+1}}{q+1} + \frac{N^q}{2} + \frac{1}{2}\binom{q}{1} B_2 N^{q-1} + \frac{1}{4}\binom{q}{3} B_4 N^{q-3} + \frac{1}{6}\binom{q}{5} B_6 N^{q-5} + \dots, \tag{4.3}$$

the last term containing either N or N^2.

Remark 4.2 It should be remarked that

$$\frac{1}{N^{q+1}} \sum_{k=1}^{N} k^q = O\left(\frac{1}{q+1}\right), \quad \text{as } N \to \infty,$$

so that (4.3) could also be used as asymptotic expansion for very small $\varepsilon = H/M$. The amount of work for computing the element stiffness matrix in the two-scale FEM is therefore independent of ε. If $n > 1$, the same arguments apply if all indices are changed to suitable multiindices.

4.2 Numerical results

We illustrate our error estimates for the two-scale FEM for the one dimensional model problem

$$\begin{aligned}
-\frac{d}{dx}\left(a\left(\frac{x}{\varepsilon}\right) \frac{du^{\varepsilon}}{dx}(x) \right) &= f(x) \quad \text{in } \Omega = (0,1), \\
u^{\varepsilon}|_{\partial\Omega} &= 0,
\end{aligned} \tag{4.4}$$

where $f(x) = e^x$ and

$$a(y) = 2 + \cos(2\pi y).$$

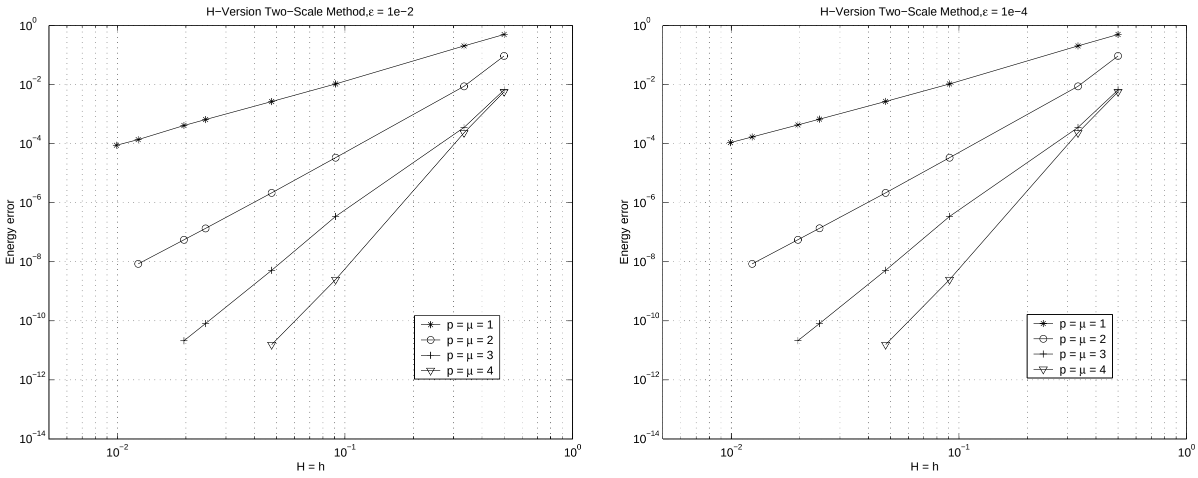

Figure 3: Energy error in the H-Version of the two-scale FEM

The shift Theorem 2.1 applies on Ω and the solution does not exhibit boundary layers, since $u^\varepsilon(x) = U^\varepsilon(x, x/\varepsilon)$, with $U^\varepsilon(x,y)$ smooth on $\Omega \times \widehat{Q}$ and 1-periodic in y.
In Figure 3 we plot the energy error versus $H = h$ and for different $p = \mu \in \{1,2,3,4\}$. Computations were performed for two different ε-scales, 10^{-2} and 10^{-4}, respectively. We see that the rate of convergence of $\|u^\varepsilon - u^\varepsilon_{FE}\|^2_{H^1(\Omega)}$ is proportional to H^{2p} as expected from the error estimates in Theorem 3.14. Moreover, we observe robustness of the convergence rates with respect to the parameter ε.
The next set of numerical experiments shows that simultaneous refinement on both scales is indeed necessary. To that end, calculations for $\varepsilon = 10^{-4}$, $\mu = 1$ and fixed h, p were performed. In Figure 4 we plot the error in energy versus H (for several fixed p). In agreement with our *a-priori* estimates $O(H^{2p} + h^{2\mu})$ we observe a saturation effect.

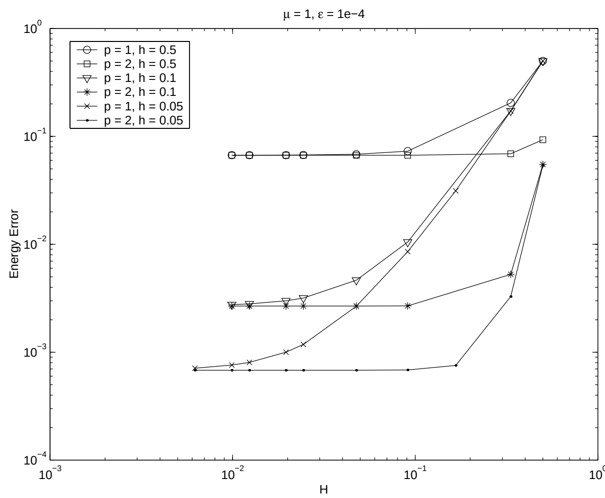

Figure 4: Energy error versus H for fixed fast scale resolution h ($\mu = 1$, $\varepsilon = 10^{-4}$)

Since the solution $U^\varepsilon(x,y)$ corresponding to (4.4) is analytic, according to Remark 3.15 we expect exponential rates of convergence of the p-version of the two-scale method, i.e., keeping

H, h fixed and increasing $p = \mu$. The exponential convergence is observed in Figure 5. We also note robustness as the error curves for $\varepsilon = 10^{-2}$, 10^{-4} are practically on top of each other.

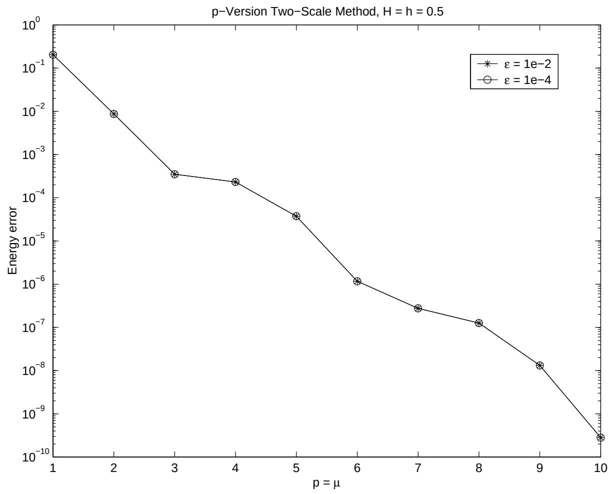

Figure 5: Robust exponential convergence of the p-Version of the two-scale FEM: energy error versus polynomial degree $p = \mu$ at fixed $H = h$, analytic solution.

A Proof of Proposition 2.5

Proof. We write first $U^\varepsilon = U_1^\varepsilon(x,y) + U_2^\varepsilon(x,y) + U_3^\varepsilon(x,y) + U_4^\varepsilon(x,y) + K$, with

$$\begin{aligned}
U_1^\varepsilon(x,y) &= -f_0\Big[\varepsilon A(y)\left(x + \frac{\varepsilon}{\langle 1/a(\cdot)\rangle}\left(\frac{1}{2}\langle 1/a(\cdot)\rangle - \langle \cdot/a(\cdot)\rangle\right) - \frac{1}{2}\right) \\
&\quad - \varepsilon^2 \tilde{A}(y) + \frac{1}{2}\langle 1/a(\cdot)\rangle x(x-1)\Big] \\
U_2^\varepsilon(x,y) &= -\varepsilon^2 \frac{1}{\langle 1/a(\cdot)\rangle}\Big(\sum_{k\varepsilon\in\mathbb{Z}^*} f_k A_{-k\varepsilon}\Big) A(y) \\
U_3^\varepsilon(x,y) &= -\sum_{k\in\mathbb{Z}^*} \frac{f_k}{2\pi i k}\left(\varepsilon A(y) + \frac{\langle 1/a(\cdot)\rangle}{2\pi i k}\right) e^{2\pi i k x} \\
U_4^\varepsilon(x,y) &= \sum_{k\in\mathbb{Z}^*} \varepsilon f_k \Big[\sum_{p\neq -k\varepsilon} A_p \frac{1}{2\pi i(pM+k)} e^{2\pi i(kx+py)}\Big] \\
K &= -\sum_{k\in\mathbb{Z}^*} \varepsilon f_k \Big[\sum_{p\neq -k\varepsilon} A_p \frac{1}{2\pi i(pM+k)}\Big] + \sum_{k\in\mathbb{Z}^*} \frac{f_k\langle 1/a(\cdot)\rangle}{(2\pi i k)^2}.
\end{aligned} \tag{A.1}$$

$A(\cdot)$ and $\tilde{A}(\cdot)$ are defined in (2.20) and $\{A_p\}_{p\in\mathbb{Z}}$ are the Fourier coefficients of $A(\cdot) \in H^1_{\mathrm{per}}(0,1)$. We prove here (2.21)for $r = 0$ and (2.22) for $r = 1$ since the estimates with respect to higher order norms can be obtained analogously.

The $\|\cdot\|_0$, $\|\cdot\|_{-1}$ norms of f are given in terms of Fourier coefficients by

$$\|f\|_0^2 \cong \sum_{k\in\mathbb{Z}} |f_k|^2, \quad \|f\|_{-1} \cong |f_0|^2 + \sum_{k\in\mathbb{Z}^*} \left|\frac{f_k}{k}\right|^2.$$

Let us estimate first $U_4^\varepsilon(x,y)$. The $L^2(\Omega,\, L^2(\widehat{Q}))$ norm of $U_4^\varepsilon(x,y)$ has the explicit form

$$\|U_4^\varepsilon(x,y)\|^2_{L^2(\Omega,\, L^2(\widehat{Q}))} = \sum_{k\in\mathbb{Z}^*} \varepsilon^2 |f_k|^2 \left[\sum_{p\neq -k\varepsilon} |A_p|^2 \, \frac{1}{(2\pi(pM+k))^2} \right].$$

Since

$$|A_0|^2 + \sum_{p\neq 0} |A_p|^2 p^2 \le C\|A\|^2_{H^1_{\rm per}(\widehat{Q})} \le C\|1/a\|_0^2$$

we have the following estimate

$$\begin{aligned}
\sum_{p\neq -k\varepsilon} |A_p|^2 \frac{1}{(pM+k)^2} &= \frac{|A_0|^2}{k^2} + \sum_{\substack{p\neq -k\varepsilon\\ p\neq 0}} |A_p|^2 p^2 \frac{1}{p^2(pM+k)^2} \\
&= \frac{|A_0|^2}{k^2} + \varepsilon^{-2}\frac{1}{k^2} \sum_{\substack{p\neq -k\varepsilon\\ p\neq 0}} |A_p|^2 p^2 \left(\frac{1}{pM} - \frac{1}{pM+k}\right)^2 \\
&\le C\varepsilon^{-2}\frac{1}{k^2}\left(|A_0|^2 + \sum_{p\neq 0} |A_p|^2 p^2\right) \le C\varepsilon^{-2}\frac{1}{k^2},
\end{aligned}$$

with $C>0$ independent of k and ε. With this observation the estimate for $\|U_4^\varepsilon(x,y)\|^2_{L^2(\Omega,\, L^2(\widehat{Q}))}$ becomes

$$\|U_4^\varepsilon(x,y)\|^2_{L^2(\Omega,\, L^2(\widehat{Q}))} \le C\sum_{k\neq 0} \frac{|f_k|^2}{k^2} \le C\|f\|^2_{-1},$$

with dominant terms coming from $\sum_{l\in\mathbb{Z}^*} \varepsilon^2 |f_{Ml\pm 1}|^2 |A_{-l}|^2$. One can even prove that $\|U_4^\varepsilon(x,y)\|_{L^2(\Omega, L^2(\widehat{Q}))} \to 0$ as $\varepsilon\to 0$ provided that $\|f\|_{-1} < \infty$. Since the ideas of this proof are relevant for our analysis we will give here the details. First observe that

$$\begin{aligned}
\|U_4^\varepsilon(x,y)\|^2_{L^2(\Omega, L^2(\widehat{Q}))} &= \varepsilon^2 \sum_{k\neq 0} \frac{|f_k|^2}{(2\pi)^2 k^2} |A_0|^2 + \sum_{k\neq 0} \frac{|f_k|^2}{k^2} \sum_{\substack{p\neq -k\varepsilon\\ p\neq 0}} \frac{1}{(2\pi)^2} |A_p|^2 p^2 \left(\frac{1}{pM} - \frac{1}{pM+k}\right)^2 \\
&\le C\varepsilon^2 \|f\|^2_{-1} + C\sum_{k\neq 0} \frac{|f_k|^2}{k^2} \sum_{\substack{p\neq -k\varepsilon\\ p\neq 0}} |A_p|^2 p^2 \left(\frac{1}{pM} - \frac{1}{pM+k}\right)^2.
\end{aligned}$$

We claim now that

$$\sum_{k\neq 0} \frac{|f_k|^2}{k^2} \sum_{\substack{p\neq -k\varepsilon\\ p\neq 0}} |A_p|^2 p^2 \left(\frac{1}{pM} - \frac{1}{pM+k}\right)^2 \le C(\varepsilon)\|f\|^2_{-1}$$

with $C(\varepsilon)\to 0$ as $\varepsilon\to 0$. We observe that

$$\sum_{|k|\ge M} \frac{|f_k|^2}{k^2} \sum_{\substack{p\neq -k\varepsilon\\ p\neq 0}} |A_p|^2 p^2 \left(\frac{1}{pM} - \frac{1}{pM+k}\right)^2 \le C \sum_{|k|\ge M} \frac{|f_k|^2}{k^2} \le C(\varepsilon)\|f\|^2_{-1}$$

and $C(\varepsilon)\to 0$ as $\varepsilon\to 0$. It remains to show that e.g.

$$\sum_{r=1}^{M-1} \frac{|f_r|^2}{r^2} \sum_{p\neq 0} |A_p|^2 p^2 \left(\frac{1}{pM} - \frac{1}{pM+r}\right)^2 \le C(\varepsilon)\|f\|^2_{-1}$$

with $C(\varepsilon) \to 0$ as $\varepsilon \to 0$. We write this in the following form

$$\begin{aligned}
&\sum_{r=1}^{M-1} \frac{|f_r|^2}{r^2} \sum_{\substack{p\geq 1\\ p\leq -2}} |A_p|^2 p^2 \left(\frac{1}{pM} - \frac{1}{pM+r}\right)^2 + \sum_{r=1}^{M-1} |A_{-1}|^2 \frac{|f_r|^2}{M^2(M-r)^2} \\
&\leq C\varepsilon^2 \|f\|_{-1}^2 + C \sum_{r=1}^{M-1} \frac{|f_r|^2}{M^2(M-r)^2} \\
&\leq C\varepsilon^2 \|f\|_{-1}^2 + C \sum_{r=1}^{[M/2]} \frac{|f_r|^2}{M^2(M-r)^2} + C \sum_{r=[M/2]+1}^{M} \frac{|f_r|^2}{M^2(M-r)^2} \\
&\leq C\varepsilon^2 \|f\|_{-1}^2 + C \sum_{r=[M/2]+1}^{M} \frac{|f_r|^2}{r^2} \\
&\leq C(\varepsilon) \|f\|_{-1}^2,
\end{aligned}$$

with $C(\varepsilon) \to 0$ as $\varepsilon \to 0$. Furthermore,

$$\|\varepsilon^{-1} \nabla_y U_4^\varepsilon(x,y)\|^2_{L^2(\Omega,\, L^2_{\text{per}}(\widehat{Q}))} = \sum_{k\in\mathbb{Z}^*} |f_k|^2 \sum_{p\neq -k\varepsilon} |A_p|^2 \frac{p^2}{(pM+k)^2} \leq C\|f\|_0^2.$$

The same type of arguments used for the estimate of the $L^2(\Omega, L^2(\widehat{Q}))$-norm of $U_4^\varepsilon(x,y)$ imply that

$$\|\varepsilon^{-1} \nabla_y U_4^\varepsilon(x,y)\|_{L^2(\Omega, L^2(\widehat{Q}))} + \|\nabla_x U_4^\varepsilon(x,y)\|_{L^2(\Omega, L^2(\widehat{Q}))} \leq C(\varepsilon)\|f\|_0^2$$

with $C(\varepsilon) > 0$ such that $C(\varepsilon) \to 0$ as $\varepsilon \to 0$.

We see that $U_1^\varepsilon(x,y)$ is polynomial in x and $U_2^\varepsilon(x,y) = U_2^\varepsilon(y)$ is constant with respect to x. The bounds for $U_1^\varepsilon(x,y)$ are straightforward

$$\|U_1^\varepsilon\|_{L^2(\Omega, L^2_{\text{per}}(\widehat{Q}))} \leq C|f_0| \leq C\|f\|_{-1}, \quad \|\varepsilon^{-1} \nabla_y U_1^\varepsilon\|_{L^2(\Omega, L^2_{\text{per}}(\widehat{Q}))} \leq C|f_0| \leq C\|f\|_{-1}.$$

$U_2^\varepsilon(x,y)$ has the form $-\varepsilon \left(\sum_{l\in\mathbb{Z}^*} f_{Ml} A_{-l}\right) 1/\langle 1/a(\cdot)\rangle \varepsilon A(y)$ and its $L^2(\Omega,\, L^2(\widehat{Q}))$-norm is bounded by $\varepsilon\|f\|_{-1}$. Further,

$$\|U_3^\varepsilon\|_{L^2(\Omega,\, L^2_{\text{per}}(\widehat{Q}))} \leq C(\varepsilon\|f\|_{-1} + \|f\|_{-2}), \quad \|\varepsilon^{-1} \nabla_y U_3^\varepsilon\|_{L^2(\Omega,\, L^2_{\text{per}}(\widehat{Q}))} \leq C\|f\|_{-1}.$$

It remains to analyze the convergence of the first sum in K in (A.1) and to obtain sharp estimates on K with respect to ε and the regularity of f. We split K as

$$K = S_1 + S_2,$$

where

$$S_1 = -\sum_{k\in\mathbb{Z}^*} \varepsilon f_k \left[\sum_{p\neq -k\varepsilon} A_p \frac{1}{2\pi i(pM+k)}\right]$$

and

$$S_2 = \omega \frac{1}{(2\pi i)^2} \sum_{k\in\mathbb{Z}^*} \frac{f_k}{k^2}.$$

We recall the definition (2.20) of $A(y)$. This implies that the Fourier coefficients A_p, $p \in \mathbb{Z}$, are, in terms of the Fourier coefficients b_q of $1/a(\cdot)$, given by

$$A_0 = -\frac{1}{2\pi i} \sum_{q \neq 0} \frac{b_q}{q}, \quad A_p = \frac{1}{2\pi i} \frac{b_p}{p} \quad \forall\, p \neq 0.$$

With this observation S_1 can be written in the following form

$$\begin{aligned} S_1 &= \frac{1}{(2\pi i)^2} \sum_{k \neq 0} \varepsilon f_k \Big[\sum_{\substack{q \neq 0 \\ qM+k \neq 0}} \frac{b_q}{q} \left(\frac{1}{k} - \frac{1}{qM+k} \right) \Big] - \frac{1}{(2\pi i)^2} \sum_{l \neq 0} \varepsilon \frac{f_{Ml}}{Ml} \frac{b_{-l}}{l} \\ &= \frac{1}{(2\pi i)^2} \sum_{k \neq 0} \frac{f_k}{k} \Big(\sum_{\substack{q \neq 0 \\ qM+k \neq 0}} \frac{b_q}{qM+k} \Big) - \frac{1}{(2\pi i)^2} \sum_{l \neq 0} \varepsilon \frac{f_{Ml}}{Ml} \frac{b_{-l}}{l}. \end{aligned}$$

Now we use that for all $k \in \mathbb{Z}^*$, the sequence $s^k = \{(s^k)_p\}_{p \in \mathbb{Z}}$ given by

$$(s^k)_p = \begin{cases} \dfrac{1}{p+k} & \text{if } p \neq -k \\ 0 & \text{if } p = -k \end{cases}$$

is an element of l^2 and $(s^k, s^{k'})_{l^2} = c\delta_k^{k'}$, i.e., $\{s^k\}_k$ is an orthogonal system in l^2 and all s^k have the same l^2-norm. Indeed, let $k \neq k'$. Let us show that $(s^k, s^{k'})_{l^2} = 0$. Without loss of generality we may assume that $k < k'$. Then, it holds

$$\begin{aligned} (s^k, s^{k'})_{l^2} &= \sum_{\substack{q \neq -k \\ q \neq -k'}} \frac{1}{q+k} \frac{1}{q+k'} \\ &= \frac{1}{k'-k} \sum_{\substack{q \neq -k \\ q \neq -k'}} \left(\frac{1}{q+k} - \frac{1}{q+k'} \right) \\ &= \frac{1}{k'-k} \lim_{N \to \infty} \sum_{\substack{|q| \leq N \\ q \neq -k, q \neq -k'}} \left(\frac{1}{q+k} - \frac{1}{q+k'} \right) \\ &= \frac{1}{k'-k} \lim_{N \to \infty} \left(\sum_{\substack{q=-N+k \\ q \neq 0}}^{N+k} \frac{1}{q} - \sum_{\substack{q=-N+k' \\ q \neq 0}}^{N+k'} \frac{1}{q} \right) \\ &= \frac{1}{k'-k} \lim_{N \to \infty} \left(\sum_{q=-N+k}^{-N+k'} \frac{1}{q} - \sum_{q=N+k}^{N+k'} \frac{1}{q} \right). \end{aligned}$$

The last sums are done after a finite, fixed number of q with $q \to \infty$ when $N \to \infty$. Therefore, the limit is zero. With this observation we can write K as

$$K = S^1 + S^2 = \frac{1}{(2\pi i)^2} \sum_{k \neq 0} \frac{f_k}{k} \sum_{\substack{q' \in M\mathbb{Z} \\ q'+k \neq 0}} \frac{b_{q'/M}}{q'+k} - \frac{1}{(2\pi i)^2} \sum_{l \neq 0} \varepsilon \frac{f_{Ml}}{Ml} \frac{b_{-l}}{l}.$$

The first sum in the above representation can be interpreted as the l^2 scalar product between the sequences $\{f_k/k\}_{k \in \mathbb{Z}^*} \in l^2$ and $\{c_k\}_k$, where $c_k = (\tilde{b}, s^k)_{l^2}$, with $\tilde{b}_q = 0$ if $q \notin M\mathbb{Z}$ and $\tilde{b}_q = b_{q/M}$ else. Therefore, this sum is bounded by $c\|\tilde{b}\|_{l^2}\|f\|_{-1} \leq c\|f\|_{-1}$. The second sum can be estimated by $C\varepsilon\|f\|_{-1}$. Summing up all these estimates we see that $|K| \leq C\|f\|_{-1}$.

It remains to check if the estimates in (2.21) are sharp. We will analyse this only for the case $r = 1$. We assume therefore that $\|f\|_0 < \infty$ and we show that for $\varepsilon \ll 1$ sufficiently small $\|U^\varepsilon\|_{H^1(\Omega,L^2_{\rm per}(\widehat{Q}))} \geq C\|f\|_{-1}$ with $C > 0$ a constant independent of ε. It is not difficult to see that for $\varepsilon \ll 1$ sufficiently small

$$\|U^\varepsilon\|_{H^1(\Omega,L^2_{\rm per}(\widehat{Q}))} \geq C\|1/2\omega f_0(x-x^2) - \sum_{k\neq 0} \frac{f_k}{(2\pi ik)^2}\omega e^{2\pi ikx} + K\|_{H^1(\Omega,L^2_{\rm per}(\widehat{Q}))},$$

in which we introduced the notation $\omega = \langle 1/a(\cdot)\rangle$. To this end, we used that $\|U_4^\varepsilon(x,y)_{H^1(\Omega,L^2_{\rm per}(\widehat{Q}))} \leq C(\varepsilon)\|f\|_0$ with $C(\varepsilon) \to 0$ as $\varepsilon \to 0$. Now,

$$\begin{aligned}
&\|1/2\omega f_0(x-x^2) - \sum_{k\neq 0} \frac{f_k}{(2\pi ik)^2}\omega e^{2\pi ikx} + K\|^2_{L^2(\Omega,L^2_{\rm per}(\widehat{Q}))} \\
&= \frac{1}{120}\omega^2|f_0|^2 + \omega^2\sum_{k\neq 0}\frac{|f_k|^2}{(2\pi k)^4} + |K|^2 + \frac{1}{6}\omega f_0\overline{K} - 2\omega^2 f_0 \sum_{k\neq 0}\frac{\overline{f_k}}{(2\pi k)^4} \\
&\geq C\omega^2|f_0|^2
\end{aligned}$$

and

$$\begin{aligned}
&\|1/2\omega f_0(1-2x) - \sum_{k\neq 0}\frac{f_k}{2\pi ik}\omega e^{2\pi ikx}\|^2_{L^2(\Omega,L^2_{\rm per}(\widehat{Q}))} \\
&= \frac{1}{12}\omega^2|f_0|^2 + \omega^2\sum_{k\neq 0}\frac{|f_k|^2}{(2\pi k)^2} - 2\omega^2 f_0\frac{\overline{f_k}}{(2\pi k)^2} \\
&\geq \frac{1}{12}\omega^2|f_0|^2 + \omega^2\sum_{k\neq 0}\frac{|f_k|^2}{(2\pi k)^2} - 2\alpha\omega^2|f_0|^2\frac{1}{(2\pi)^2}\sum_{k\geq 1}\frac{1}{k^2} - \frac{1}{\alpha}\omega^2\sum_{k\neq 0}\frac{|f_k|^2}{(2\pi k)^2} \\
&= \frac{1}{12}(1-\alpha)\omega^2|f_0|^2 + \omega^2\left(1-\frac{1}{\alpha}\right)\sum_{k\neq 0}\frac{|f_k|^2}{(2\pi k)^2}
\end{aligned}$$

for some arbitrary $\alpha > 0$. Summing up,

$$\begin{aligned}
&\|1/2\omega f_0(x-x^2) - \sum_{k\neq 0}\frac{f_k}{(2\pi ik)^2}\omega e^{2\pi ikx} + K\|_{H^1(\Omega,L^2_{\rm per}(\widehat{Q}))} \\
&\quad \geq C\omega^2|f_0|^2 + \frac{1}{12}(1-\alpha)\omega^2|f_0|^2 + \omega^2\left(1-\frac{1}{\alpha}\right)\sum_{k\neq 0}\frac{|f_k|^2}{(2\pi k)^2}.
\end{aligned}$$

Choosing $1 < \alpha$ such that $C + \frac{1}{12}(1-\alpha) > 0$ we conclude that

$$\|1/2\omega f_0(x-x^2) - \sum_{k\neq 0}\frac{f_k}{(2\pi ik)^2}\omega e^{2\pi ikx} + K\|_{H^1(\Omega,L^2_{\rm per}(\widehat{Q}))} \geq C\|f\|^2_{-1}.$$

B Proof of Lemma 3.4

Proof. We start with the two-dimensional case ($n = 2$). Let us investigate first the mixed derivatives of the interpolation error. Since ∂_{x_j} and $\pi_p^{(x_i)}$ commute for all $i \neq j$ it holds

$$\partial_{x_1}\partial_{x_2}(u - \widehat{\Pi}_p u) = \partial_{x_1}(\partial_{x_2}u - \pi_p^{(x_1)}(\partial_{x_2}u)) + \partial_{x_1}(\pi_p^{(x_1)}\partial_{x_2}(u - \pi_p^{(x_2)}u)).$$

For $0 \le t \le k-1$ we obtain that

$$
\begin{aligned}
\|\partial_{x_1}\partial_{x_2}(u - \widehat{\Pi}_p u)\|^2_{L^2(\widehat{K})} \quad &\le \\
&\le\; 2\,\|\partial_{x_1}(\partial_{x_2}u - \pi_p^{(x_1)}(\partial_{x_2}u))\|^2_{L^2(\widehat{K})} + 2\,\|\partial_{x_1}(\pi_p^{(x_1)}\partial_{x_2}(u - \pi_p^{(x_2)}u))\|^2_{L^2(\widehat{K})} \\
&\le\; 2\,\|\partial_{x_1}(\partial_{x_2}u - \pi_p^{(x_1)}(\partial_{x_2}u))\|^2_{L^2(\widehat{K})} + 2\,\|\partial_{x_2}(\partial_{x_1}u - \pi_p^{(x_2)}(\partial_{x_1}u))\|^2_{L^2(\widehat{K})} \\
&\le\; 2\,\frac{(p-t)!}{(p+t)!}\int_{\widehat{K}} \left((\partial_{x_1}^{t+1}\partial_{x_2}u(x_1,x_2))^2 + (\partial_{x_1}\partial_{x_2}^{t+1}u(x_1,x_2))^2\right)dx_1dx_2.
\end{aligned}
$$

Reasoning in a similar way for the lower order derivatives, for $0 \le s, t \le k$ we obtain that

$$
\begin{aligned}
\|\partial_{x_1}(u - \widehat{\Pi}_p u)\|^2_{L^2(\widehat{K})} \quad &\le\; 2\,\frac{(p-s)!}{(p+s)!}\int_{\widehat{K}} (\partial_{x_1}^{s+1}u(x_1,x_2))^2 dx_1dx_2 \\
&\quad + 2\,\frac{1}{p(p+1)}\,\frac{(p-t)!}{(p+t)!}\int_{\widehat{K}} (\partial_{x_1}\partial_{x_2}^{t+1}u(x_1,x_2))^2 dx_1dx_2
\end{aligned}
$$

and

$$
\begin{aligned}
\|u - \widehat{\Pi}_p u\|^2_{L^2(\widehat{K})} \quad &\le\; \frac{2}{p(p+1)}\left\{\frac{(p-s)!}{(p+s)!}\int_{\widehat{K}} (\partial_{x_1}^{s+1}u)^2 dx_1dx_2 \right.\\
&\quad \left. + \frac{(p-t)!}{(p+t)!}\int_{\widehat{K}} (\partial_{x_1}\partial_{x_2}^{t+1}u)^2 dx_1dx_2\right\}.
\end{aligned}
$$

Taking now $t = s-1$, $s = k$ and summing up all the estimates we get the result.

For the case $n = 3$ we start again from the mixed derivatives of the interpolation error

$$
\begin{aligned}
\|\partial_{x_1}\partial_{x_2}\partial_{x_3}(u - \widehat{\Pi}_p u)\|^2_{L^2(\widehat{K})} \quad &= \\
&= \|\partial_{x_1}(\partial_{x_2}\partial_{x_3}u - \pi_p^{(x_1)}\partial_{x_2}\partial_{x_3}u) + \partial_{x_1}\pi_p^{(x_1)}\partial_{x_2}(\partial_{x_3}u - \pi_p^{(x_2)}\partial_{x_3}u) \\
&\quad + \partial_{x_1}\pi_p^{(x_1)}\partial_{x_2}\pi_p^{(x_2)}\partial_{x_3}(u - \pi_p^{(x_3)}u)\|^2_{L^2(\widehat{K})} \\
&\le\; 3\|\partial_{x_1}(\partial_{x_2}\partial_{x_3}u - \pi_p^{(x_1)}\partial_{x_2}\partial_{x_3}u)\|^2_{L^2(\widehat{K})} + 3\|\partial_{x_1}\pi_p^{(x_1)}\partial_{x_2}(\partial_{x_3}u - \pi_p^{(x_2)}\partial_{x_3}u)\|^2_{L^2(\widehat{K})} \\
&\quad + 3\|\partial_{x_1}\pi_p^{(x_1)}\partial_{x_2}\pi_p^{(x_2)}\partial_{x_3}(u - \pi_p^{(x_3)}u)\|^2_{L^2(\widehat{K})} \\
&\le\; 3\|\partial_{x_1}(\partial_{x_2}\partial_{x_3}u - \pi_p^{(x_1)}\partial_{x_2}\partial_{x_3}u)\|^2_{L^2(\widehat{K})} + 3\|\partial_{x_2}(\partial_{x_1}\partial_{x_3}u - \pi_p^{(x_2)}\partial_{x_1}\partial_{x_3}u)\|^2_{L^2(\widehat{K})} \\
&\quad + 3\|\partial_{x_3}(\partial_{x_1}\partial_{x_2}u - \pi_p^{(x_3)}\partial_{x_1}\partial_{x_2}u)\|^2_{L^2(\widehat{K})} \\
&\le\; 3\frac{(p-t)!}{(p+t)!}\int_{\widehat{K}} \left\{\left|\partial_{x_1}^{t+1}\partial_{x_2}\partial_{x_3}u\right|^2 + \left|\partial_{x_1}\partial_{x_2}^{t+1}\partial_{x_3}u\right|^2 + \left|\partial_{x_1}\partial_{x_2}\partial_{x_3}^{t+1}u\right|^2\right\}dx.
\end{aligned}
$$

Next, we investigate the lower order derivatives

$$
\begin{aligned}
\|\partial_{x_1}\partial_{x_2}(u - \widehat{\Pi}_p u)\|^2_{L^2(\widehat{K})} \quad &= \\
&=\; \|\partial_{x_1}(\partial_{x_2}u - \pi_p^{(x_1)}\partial_{x_2}u) + \partial_{x_1}\pi_p^{(x_1)}\partial_{x_2}(u - \pi_p^{(x_2)u}) \\
&\quad +\; \partial_{x_1}\pi_p^{(x_1)}\partial_{x_2}\pi_p^{(x_2)}(u - \pi_p^{(x_3)}u)\|^2_{L^2(\widehat{K})}
\end{aligned}
$$

$$
\begin{aligned}
&\leq\ 3\|\partial_{x_1}(\partial_{x_2}u-\pi_p^{(x_1)}\partial_{x_2}u)\|^2_{L^2(\widehat{K})}\\
&+\ 3\|\partial_{x_2}(\partial_{x_1}u-\pi_p^{(x_2)}\partial_{x_1}u)\|^2_{L^2(\widehat{K})}+3\|\partial_{x_1}\partial_{x_2}u-\pi_p^{(x_3)}\partial_{x_1}\partial_{x_2}u\|^2_{L^2(\widehat{K})}\\
&\leq\ 3\frac{(p-s)!}{(p+s)!}\int_{\widehat{K}}\left\{\left|\partial_{x_1}^{s+1}\partial_{x_2}u\right|^2+\left|\partial_{x_1}\partial_{x_2}^{s+1}u\right|^2\right\}dx\\
&+\ 3\frac{1}{p(p+1)}\frac{(p-t)!}{(p+t)!}\int_{\widehat{K}}\left|\partial_{x_1}\partial_{x_2}\partial_{x_3}^{t+1}u\right|^2dx.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\|\partial_{x_1}(u-\widehat{\Pi}_pu)\|^2_{L^2(\widehat{K})}\ &=\ \|\partial_{x_1}(u-\pi_p^{(x_1)}u)+\partial_{x_1}\pi_p^{(x_1)}(u-\pi_p^{(x_2)}u)\\
&+\ \partial_{x_1}\pi_p^{(x_1)}\pi_p^{(x_2)}(u-\pi_p^{(x_3)}u)\|^2_{L^2(\widehat{K})}\\
&\leq\ 3\|\partial_{x_1}(u-\pi_p^{(x_1)}u)\|^2_{L^2(\widehat{K})}+3\|\partial_{x_1}u-\pi_p^{(x_2)}\partial_{x_1}u\|^2_{L^2(\widehat{K})}\\
&+\ 3\|\pi_p^{(x_2)}(\partial_{x_1}u-\pi_p^{(x_3)}\partial_{x_1}u)\|^2_{L^2(\widehat{K})}\\
&\leq\ 3\frac{(p-l)!}{(p+l)!}\int_{\widehat{K}}\left|\partial_{x_1}^{l+1}u\right|^2dx+3\frac{1}{p(p+1)}\frac{(p-s)!}{(p+s)!}\int_{\widehat{K}}\left\{\left|\partial_{x_1}\partial_{x_2}^{s+1}u\right|^2+2\left|\partial_{x_1}\partial_{x_3}^{s+1}u\right|^2\right\}dx\\
&+\ 6\frac{1}{p^2(p+1)^2}\frac{(p-t)!}{(p+t)!}\int_{\widehat{K}}\left|\partial_{x_1}\partial_{x_2}\partial_{x_3}^{t+1}u\right|^2dx.
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\|u-\widehat{\Pi}_pu\|^2_{L^2(\widehat{K})}\ &=\ \|u-\pi_p^{(x_1)}u+\pi_p^{(x_1)}(u-\pi_p^{(x_2)}u)+\pi_p^{(x_1)}\pi_p^{(x_2)}(u-\pi_p^{(x_3)}u)\|^2_{L^2(\widehat{K})}\\
&\leq\ 3\frac{1}{p(p+1)}\frac{(p-l)!}{(p+l)!}\int_{\widehat{K}}\left\{\left|\partial_{x_1}^{l+1}u\right|^2+2\left|\partial_{x_2}^{l+1}u\right|^2+4\left|\partial_{x_3}^{l+1}u\right|^2\right\}dx\\
&+6\frac{1}{p^2(p+1)^2}\frac{(p-s)!}{(p+s)!}\int_{\widehat{K}}\left\{\left|\partial_{x_1}\partial_{x_2}^{s+1}\right|^2+2\left|\partial_{x_2}\partial_{x_3}^{s+1}\right|^2+2\left|\partial_{x_1}\partial_{x_3}^{s+1}\right|^2\right\}dx\\
&+12\frac{1}{p^3(p+1)^3}\frac{(p-t)!}{(p+t)!}\int_{\widehat{K}}\left|\partial_{x_1}\partial_{x_2}\partial_{x_3}^{t+1}u\right|^2dx.
\end{aligned}
$$

Assume first that $p\geq k\geq 2$. We take then $t=k-2$, $s=k-1$ and $l=k$ in the above estimates and we obtain (3.1). In case of linear interpolation, i.e. $p=1$, under the assumption that u is in the space $\mathcal{H}^{\mathbf{1}}(\widehat{K})$, we obtain (3.2) by taking $t=s=l=0$ in the previous estimates. □

References

[1] I. Babuška, A.K. Aziz, *Survey lectures on the mathematical foundation of the finite element method*, in The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A. K. Aziz, ed., Academic Press, New York (1973) 5–359.

[2] A. Bensoussan, J.L. Lions, G. Papanicolau, *Asymptotic Analysis for Periodic Structures*, North Holland, Amsterdam, 1978.

[3] D. Cioranescu and Jeanine Saint Jean Paulin, "Homogenization of Reticulated Structures", Springer Applied Mathematical Sciences (1999).

[4] T. Y. Hou and X. Xin, "Homogenization of Linear Transport Equations with Oscillatory Vector Fields", SIAM J. Appl. Math. **52** (1992), 34–45.

[5] A. M. Matache, Spectral- and p-Finite Elements for problems with microstructure, Ph.D. Thesis, ETH Zürich 2000.

[6] A.M. Matache, I. Babuška and Ch. Schwab, "Generalized p-FEM in Homogenization", Numerische Mathematik, **86** Issue 2 (2000), 319–375.

[7] R.C. Morgan and I. Babuška, "An approach for constructing families of homogenized solutions for periodic media, I: An integral representation and its consequences, II: Properties of the kernel", SIAM J. Math. Anal. **22** (1991) 1–33.

[8] Ch. Schwab, "p- and hp- Finite Element Methods", Oxford Science Publications, 1998.

[9] Ch. Schwab, A.-M. Matache, "High order generalized FEM for lattice materials", Proceedings of the 3rd European Conference on Numerical Mathematics and Advanced Applications, Finland, 1999, ed. by P. Neittaanmäki, T. Tiihonen and P. Tarvainen, World Scientific, Singapore, 2000.

[10] B. Szabó, I. Babuška, "Finite Element Analysis", John Wiley & Sons, Inc. 1991.

[11] O. A. Oleinik, A.S. Shamaev, G. A. Yosifian, "Mathematical Problems in Elasticity and Homogenization", North-Holland, 1992.