Computing 66, 1–26 (2001)

Computing

Homogenization and Multigrid

N. Neuss, **W. Jäger**, and **G. Wittum**, Heidelberg

Received August 10, 1999; revised July 28, 2000

Abstract

For elliptic partial differential equations with periodically oscillating coefficients which may have large jumps, we prove robust convergence of a two-grid algorithm using a prolongation motivated by the theory of homogenization. The corresponding Galerkin operator on the coarse grid turns out to be a discretization of a diffusion operator with homogenized coefficients obtained by solving *discrete* cell problems. This two-grid method is then embedded inside a multi-grid cycle extending over both the fine and the coarse scale.

AMS Subject Classifications: 35B27, 65F10, 65N22, 65N55.

Key Words: Homogenization, partial differential equations, oscillating coefficients, multilevel methods, multigrid, block smoothers, robustness.

1. Introduction

When an elliptic partial differential equation with rapidly oscillating coefficients is discretized, the following problems arise:

(a) The period of the oscillations is often so small that direct simulation with sufficient accuracy is simply impossible.

(b) Even if direct numerical simulation is possible, the arising system of equations may be very hard to solve.

In the first case, one way out is to find auxiliary problems, whose solutions are good approximations of the solution of the original problem. This passage from a "microscopic" to a "macroscopic" equation is called *homogenization*, see [4], [21], [2], [27], [19]. Let us also mention the approach due to [23], [22] which, in a model case, improves the finite element discretization of a diffusion-reaction problem with periodically oscillating coefficients, the related work from [18], [12], and the articles [15], [20] which try to recover the homogenized equation by algebraic multigrid.

In the second case, the discretization of an elliptic boundary value problem usually ends up in having to solve systems of linear equations for which direct decompositions are too expensive, both in terms of memory and computing time. Instead, one should use iterative methods, especially hierarchical ones like multigrid, which can often be shown to have optimal complexity. Yet in many interesting cases, optimal performance of multigrid is not easily achieved. Usually, this deterioration of

multigrid performance occurs when one considers so-called "singularly perturbed" problems like convection-diffusion equations with dominant convection or large anisotropies, where the original problem is a small elliptic perturbation of some reduced equation, possibly of non-elliptic type. Indeed, the diffusion problem with oscillating coefficients may also be considered a perturbation of another diffusion law which is called the "homogenized problem".

Methods for improving on multigrid convergence for these singularly perturbed problems already exist, but up to now there has been no complete answer for any of the three kinds of singular perturbations mentioned above. The case of a diffusion problem with small-scale oscillations can in many cases be handled by using so-called "matrix-dependent" prolongation/restriction operators, see for example [1], [10], [28], [32], or by immediately applying "algebraic multigrid" (AMG), which relies on the matrix information even more heavily in setting up the coarse grid space; among the many references, we want to mention [29], [31], [5] (aimed at problems with geometric irregularities are [3], [17]). Unfortunately, given a diffusion problem that is more than one-dimensional and whose coefficient has small-scale oscillations with a large-scale variation, there is no theory proving robust convergence independent of the size of the problem and the variation of the coefficients. This is not particularly surprising, since the class of these problems is very large, and there is also no method supposed to work in all cases.

Therefore, one first has to restrict the range of problems in a suitable way. As in [13], [14] and [25], we choose the case of periodic oscillations. Here, a lot is known about the limiting problem, and one may hope that this information can be used to design a suitable multigrid algorithm.

A first result can be found in [13], [14], where the continuous homogenized problem was used to define the coarse-grid problem. Yet, if the problem with oscillating coefficients itself is not resolved exactly, this may not be a good choice, see Remark 3.6. In [25], a (discretely) homogenized problem was defined by solving *discretized* local problems. For the resulting algorithm, robust convergence could be shown both theoretically and practically.

This work is based on [25], but it contains several improvements. First, the central result is sharper (especially allowing arbitrarily fine triangulation of each cell) and makes the relation to algebraic multigrid clearer. Second, we are able to handle the case of "stiff" inclusions theoretically; this was only examined numerically in [25].

Finally, some words on the notation used in this paper: we often use Einstein's summation convention, and denote the relations $a \leq Cb$, $a \geq Cb$, $ca \leq b \leq Ca$ for generic constants $c, C > 0$ by $a \lesssim b$, $a \gtrsim b$, $a \sim b$, respectively. We also use the abbreviations $\partial_k = \frac{\partial}{\partial x_k}$, $\partial_{y_k} = \frac{\partial}{\partial y_k}$.

2. Setting of the Problem

2.1. Basic Setting

Let $\mathscr{A} : Y \to \mathbb{R}^{n^2}$ be a matrix-valued, 1-periodic function and assume that positive constants μ_0, μ_1 exist with

$$\mu_0|\eta|^2 \leq \mathscr{A}_{ij}(y)\eta_i\eta_j \leq \mu_1|\eta|^2, \quad \forall\, \eta \in \mathbb{R}^n, \quad \forall\, y \in Y. \tag{1}$$

Next, we define for each $\varepsilon > 0$ an ε-periodic coefficient function

$$\mathscr{A}^\varepsilon : \mathbb{R}^n \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n), \quad x \mapsto \mathscr{A}\left(\frac{x}{\varepsilon}\right). \tag{2}$$

Given then a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$ and $f \in L^2(\Omega)$, we want to solve the following problem: Find $u^\varepsilon \in H_0^1(\Omega)$ such that

$$-\partial_i\left(\mathscr{A}^\varepsilon_{ij}(x)\partial_j u^\varepsilon(x)\right)dx = f(x) \quad x \in \Omega, \tag{3}$$

for which we will mainly use the variational formulation: find $u^\varepsilon \in H_0^1(\Omega)$ satisfying for all $\varphi \in H_0^1(\Omega)$ the equation

$$a^\varepsilon(u^\varepsilon, \varphi) := \int_\Omega \mathscr{A}^\varepsilon_{ij}(x)\partial_j u^\varepsilon(x)\partial_i\varphi(x)dx = \int_\Omega f(x)\varphi(x)dx. \tag{4}$$

It is well-known from homogenization theory (see e.g. [4], [30], [2], [19]) that the solution u^ε of (4) is approximated by the solution $u^0 \in H_0^1(\Omega)$ satisfying the homogenized equation

$$a^0(u^0, \varphi) := \int_\Omega \mathscr{A}^0_{ij}(x)\partial_j u^0(x)\partial_i\varphi(x)dx = \int_\Omega f(x)\varphi(x)dx. \tag{5}$$

for all $\varphi \in H_0^1(\Omega)$. The coefficients $\mathscr{A}^0$ appearing in (5) can be computed as the cell averages ($\langle f \rangle_Y := \int_Y f dy$)

$$\mathscr{A}^0_{ik} := \langle \mathscr{A}_{ij}(\delta_{jk} + \partial_{y_j} w_k)\rangle_Y = \langle (\delta_{il} + \partial_{y_l} w_i)\mathscr{A}_{lj}(\delta_{jk} + \partial_{y_j} w_k)\rangle_Y, \tag{6}$$

where the $w_k, k = 1, \ldots, n$ are solutions to the "cell problems": find $w_k \in H^1_{per}(Y)$ such that

$$\int_Y \mathscr{A}_{ij}(y)(\partial_{y_j} w_k(y) + \delta_{jk})\partial_{y_i}\varphi(y)dy = 0 \quad \forall\, \varphi \in H^1_{per}(Y). \tag{7}$$

Here, $H^1_{per}(Y)$ is defined as the space of 1-periodic functions on $\mathbb{R}^n$ restricted to Y. Again we set

$$w^\varepsilon_k(x) := w_k\left(\frac{x}{\varepsilon}\right), \quad \text{for } x \in \mathbb{R}^n. \tag{8}$$

Usual estimates show weak convergence of u^ε to u^0 for $\varepsilon \to 0$. For our application, we will need error estimates of the following type (but in a discrete setting):

Theorem 2.1. *Define an approximation u^ε_{app} to u^ε by*

$$u^\varepsilon_{app} := u^0 + \varepsilon\partial_k u^0 w^\varepsilon_k \in H^1(\Omega) \tag{9}$$

together with a boundary correction term $v^\varepsilon \in H^1(\Omega)$, given as the solution of

$$-\partial_i\Big(\mathscr{A}^\varepsilon_{ij}(x)\partial_j v^\varepsilon(x)\Big) = 0 \quad \text{for } x \in \Omega,\ v^\varepsilon(x) = u^\varepsilon_{app}(x) \quad \text{for } x \in \partial\Omega. \tag{10}$$

Now assume that the solution u^0 *to the homogenized problem* (5) *lies in* $H^2(\Omega)$, *and that the solutions* w_k *of the cell problems* (7) *are in* $W^{1,p}(Y)$ *for some* $p > n$. *With* a^ε *being the bilinear form defined in* (4), *the estimate*

$$a^\varepsilon(u^\varepsilon - u^\varepsilon_{app}, \varphi) \le C\varepsilon\big\|D^2u^0\big\|_{L^2(\Omega)}\|\nabla\varphi\|_{L^2(\Omega)} \tag{11}$$

is valid for all $\varphi \in H^1_0(\Omega)$, *from which one immediately obtains the energy error estimate*

$$\Big\|u^\varepsilon - u^\varepsilon_{app} + v^\varepsilon\Big\|_{\mathrm{a}^\varepsilon} \le C\varepsilon\big\|D^2u^0\big\|_{L^2(\Omega)} \tag{12}$$

with $\|\cdot\|_{a^\varepsilon}$ *denoting the energy norm*

$$\|\varphi\|_{\mathrm{a}^\varepsilon} := a^\varepsilon(\varphi,\varphi)^{1/2}. \tag{13}$$

Proof: See [19], Section 1.4, [24] or [25]. □

2.2. Symmetric Cells

We observe that Theorem 2.1 does not yet give a computationally feasible approximation, because, to obtain the boundary-layer correction v^ε, one still has to solve a problem of the same complexity as the original one. The study of this boundary term is deep, and not yet completely solved. One can approximate it by simply cutting off the wrong boundary values taken by u^ε_{app}, but, unfortunately, this leaves the energy-norm approximation error at a size of $O(\varepsilon^{1/2})$. This is much worse than the $O(\varepsilon)$-estimate from (12), which is essentially what we need for our theory below. In this paper, we will therefore restrict ourselves to special cases where this boundary layer vanishes.

This happens, for example, if the domain $\Omega \subset \mathbb{R}^n$ is rectangular with sides lying in the hyper-planes $x_k = 0$ or $x_k = N_k\varepsilon, N_k \in \mathbb{N}$ for $k = 1,\dots,n$, and if one chooses periodic instead of Dirichlet boundary conditions for the original problem (3) (of course, one then has to require the solvability condition $\int_\Omega f(x)dx = 0$ for existence, and a normalization, e.g. $\int_\Omega u^\varepsilon(x)dx = 0$, for uniqueness). But periodic boundary conditions are seldom in practical applications, so we want to examine another situation.

The following definition can be found in [2], §6.3.

Definition 2.2. We say that the function $\mathscr{A} : Y \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ is symmetric with respect to the k-coordinate or k-symmetric, if it fulfills

$$(-1)^{\delta_{ik}}\mathscr{A}_{ij}(S^{(k)}y)(-1)^{\delta_{jk}} = \mathscr{A}_{ij}(y) \ (\textit{no summation}), \tag{14}$$

for the reflection

$$S^{(k)} : (y_1, \ldots, y_n) \mapsto (y_1, \ldots, y_{k-1}, -y_k, y_{k+1}, \ldots, y_n). \tag{15}$$

We call $\mathscr{A}$ symmetric, if (14) is fulfilled for all $k = 1, \ldots, n$.

If the coefficient $\mathscr{A}$ is k-symmetric, one can deduce some (anti-)symmetry properties for the solutions of the cell problems (7) as well:

Theorem 2.3. *Let $\mathscr{A} : Y \to \mathscr{L}(\mathbb{R}^n, \mathbb{R}^n)$ be symmetric with respect to the k-coordinate. Assume that the w_l from (7) are normalized to $\langle w_l \rangle_Y = 0$. Then*

1. We have

$$w_l(S^{(k)} y) = (-1)^{\delta_{kl}} w_l(y), \quad \forall y \in Y. \tag{16}$$

2. $w_k(y) = 0$ on $\{y \in Y | y_k = 0\}$.

3. The tensor $\alpha_{ij} := \mathscr{A}_{ij} + \mathscr{A}_{il} \partial_{y_l} w_j$ satisfies

$$(-1)^{\delta_{ik}} \alpha_{ij}(S^{(k)} y)(-1)^{\delta_{jk}} = \alpha_{ij}(y) \ (\textit{no summation}). \tag{17}$$

4. The homogenized coefficients $\mathscr{A}^0$ satisfy $\mathscr{A}^0_{kj} = \mathscr{A}^0_{jk} = 0$ for $j \neq k$. Especially, $\mathscr{A}^0$ is diagonal for (fully) symmetric $\mathscr{A}$.

Proof: See [2], §6.3 or [25]. □

Corollary 2.4. *Let $\varepsilon > 0$ and $\Omega = \prod_{k=1}^n [0, d_k] \subset \mathbb{R}^n$ where, for $k = 1, \ldots, n$, $d_k = n_k \varepsilon$ for some $n_k \in \mathbb{N}$. Assume further that $\mathscr{A}$ is symmetric with respect to the coordinate indices $\emptyset \neq S \subset \{1, \ldots, n$ and that Dirichlet zero boundary conditions are prescribed on*

$$\partial\Omega_0 := \{x \in \partial\Omega \mid \exists i \in S : x_i = 0 \vee x_i = d_i\} \tag{18}$$

while periodic boundary conditions are prescribed on $\partial\Omega - \partial\Omega_0$. Then u^ε_{app} vanishes on $\partial\Omega_0$, and therefore $v^\varepsilon \equiv 0$ in Theorem 2.1.

Proof: The tangential component of ∇u^0 vanishes on $\partial\Omega_0$ due to the homogeneous boundary conditions posed for u^0. Also the normal component of the vector of cell solutions vanishes there because of Theorem 2.3. Thus, the product $\nabla u^0 \cdot \vec{w}$ vanishes on $\partial\Omega_0$. □

3. Discrete Approximation Results

3.1. Basic Setting

Let Ω be of the form $\Omega = \prod_{k=1}^n [0, d_k] \subset \mathbb{R}^n$ where, for $k = 1, \ldots, n$, $d_k = n_k \varepsilon$ for some $n_k \in \mathbb{N}$. Let $\mathscr{T}_Y$ be a triangulation of the cell Y (consisting of triangles,

quadrangles or their higher-dimensional analogues), and let $\mathcal{S}_Y$ denote the space of continuous, periodic functions on Y that are piecewise (multi)linear on elements of $\mathcal{T}_Y$. Scaling $\mathcal{T}_Y$ with factor ε and repeating it periodically, we obtain a triangulation $\mathcal{T}_h$ of the whole domain Ω. We denote by $\mathcal{S}_h$ the space of continuous functions that are (multi) linear on elements of $\mathcal{T}_h$ and vanish on $\partial\Omega$.

Problem (4) now induces the discrete problem: Find $u_h^\varepsilon \in \mathcal{S}_h$ such that

$$\int_\Omega \mathcal{A}_{ij}^\varepsilon(x)\partial_j u_h^\varepsilon(x)\partial_i\varphi_h(x)dx = \int_\Omega f(x)\varphi_h(x)dx \quad \forall\varphi \in \mathcal{S}_h, \tag{19}$$

and, in the same way, the cell problem (7) induces a discrete cell problem: Find $w_{k,\mathcal{S}_Y} \in \mathcal{S}_Y$ such that

$$\int_Y \mathcal{A}_{ij}(y)(\partial_{y_j} w_{k,\mathcal{S}_Y}(y) + \delta_{jk})\partial_{y_i}\varphi_{\mathcal{S}_Y}\, dy = 0 \quad \forall\varphi_{\mathcal{S}_Y} \in \mathcal{S}_Y. \tag{20}$$

Since we assume throughout this paper that $\mathcal{A}$ is symmetric in the sense of Definition 2.2, it is natural to assume that the triangulation $\mathcal{T}_Y$ is also symmetric with respect to reflections on all coordinate planes. This situation is depicted in Fig. 1.

Let now $\mathcal{T}_H$ be a triangulation of Ω made of cubes aligned with the cell structure, i.e. $H = \varepsilon$. Denote by $\mathcal{S}_H$ the space consisting of multilinear finite elements on $\mathcal{T}_H$ vanishing on $\partial\Omega$ and define a prolongation $p_{h\leftarrow H} : \mathcal{S}_H \to \mathcal{S}_h$ as

$$p_{h\leftarrow H} : \mathcal{S}_H \to \mathcal{S}_h,\ \varphi_H \mapsto \varphi_h = I_h \circ (\varphi_H + \varepsilon\partial_k\varphi_H w_{k,h}^\varepsilon), \tag{21}$$

where $w_{k,h}^\varepsilon(x) := w_{k,\mathcal{S}_Y}\left(\frac{x}{\varepsilon}\right)$, and $I_h : C^0(\Omega) \to \mathcal{S}_h$ is the interpolation operator given by evaluation at nodes of $\mathcal{T}_h$.

Remark 3.1. Due to the above symmetry assumptions on coefficients and cell triangulation, the assertion of Theorem 2.3 still holds for the discrete cell solutions $w_{k,\mathcal{S}_Y}$ (by the same argument). Together with the observation that, for $\varepsilon = H$, the tangential component of $\nabla\varphi_H$ is continuous across element boundaries

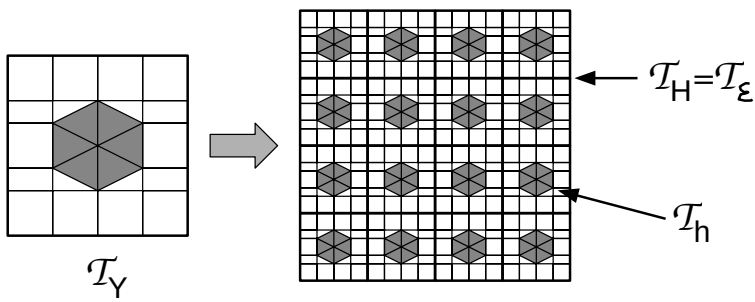

Fig. 1. Symmetric triangulations of cell and domain

of $\mathscr{T}_H$, it follows that the prolongation $p_{h\leftarrow H}$ is well-defined despite the discontinuities of $\nabla\varphi_H$.

With the help of the discrete cell solutions $w_{k,\mathscr{S}_Y}$, we can further define discretely homogenized coefficients $\tilde{\mathscr{A}}^0$ by

$$\tilde{\mathscr{A}}^0_{ik} := \langle \mathscr{A}_{ij}(y)(\delta_{jk} + \partial_{y_j} w_{k,\mathscr{S}_Y}(y))\rangle_Y, \tag{22}$$

for which we assume that constants $\mu_0', \mu_1' > 0$ exist such that

$$\mu_0'|\eta|^2 \le \tilde{\mathscr{A}}^0_{ij}\eta_i\eta_j \le \mu_1'|\eta|^2 \quad \forall\ \eta \in \mathbb{R}^n. \tag{23}$$

Remark 3.2. We observe that $\mathscr{A}^0 < \tilde{\mathscr{A}}^0$ in general, because the cell solutions w_k, resp. $w_{k,\mathscr{S}_Y}$, can be interpreted as orthogonal projections with respect to the energy scalar product of the functions $-y_k$ (not having periodic boundary conditions!) on $H^1_{per}(Y)$, resp. $\mathscr{S}_Y$, and $\mathscr{A}^0$, resp. $\tilde{\mathscr{A}}^0$, measures the energy of the remaining part.

With the coefficients $\tilde{\mathscr{A}}^0$, we define a discrete bilinear form on $H^1_0(\Omega)$ by

$$\tilde{a}^0(u,v) := \int_\Omega \tilde{\mathscr{A}}^0_{ij}\partial_j u(x)\partial_i v(x)dx. \tag{24}$$

which gives rise to the discretely homogenized problem

$$\text{Find } u^0 \in H^1_0(\Omega): \quad \tilde{a}^0(u^0,\varphi) = f(\varphi) \quad \forall\ \varphi \in H^1_0(\Omega). \tag{25}$$

3.2. The Uniformly Elliptic Case

We have the following basic theorem:

Theorem 3.3. *We assume the setting of Subsection* 3.1. *Especially, we have a symmetric diffusion coefficient and the coarse grid $\mathscr{T}_H$ is aligned with the cell structure*[1]. *Then let u^0_H be the a^0-orthogonal projection of u^0 from* (25) *on $\mathscr{S}_H$, which is defined by*

$$\textit{Find } u^0_H \in \mathscr{S}_H: \quad \tilde{a}^0(u^0_H,\varphi_H) = f(\varphi_H) \quad \forall \varphi_H \in \mathscr{S}_H. \tag{26}$$

Then a constant $C = C(n,\Omega,\mathscr{T}_Y)$ exists such that

$$a^\varepsilon(u^\varepsilon_h - p_{h\leftarrow H}u^0_H, \varphi_h) \le C(\mu_1' + \mu_1\|w_{k,\mathscr{S}_Y}\|_{1,\infty})\varepsilon\|D^2u^0\|_{L^2(\Omega)}\|\nabla\varphi_h\|_{L^2(\Omega)}. \tag{27}$$

Setting $\varphi_h := u^\varepsilon_h - p_{h\leftarrow H}u^0_H$, we immediately obtain the energy-norm error estimate

$$\|\nabla(u^\varepsilon_h - p_{h\leftarrow H}u^0_H)\|_{a^\varepsilon} \le C\frac{\mu_1' + \mu_1\|w_{k,\mathscr{S}_Y}\|_{1,\infty}}{\sqrt{\mu_0}}\varepsilon\|D^2u^0\|_{L^2(\Omega)}. \tag{28}$$

[1] This assumption of alignment with the cell structure can be weakened, see Theorem 3.9 below.

Since we are in a H^2-regular setting for the homogenized problem (25), *we may further replace* $\|D^2u^0\|_{L^2(\Omega)}$ *by* $\frac{1}{\mu_0'}\|f\|_{L^2(\Omega)}$ *in* (28).

Proof: 1. The idea is to decompose the test function φ_h into a "smooth" and an "oscillating" part. To achieve this, let $\Pi_H : H^1(\Omega) \to \mathscr{S}_H$ be a suitable quasi-interpolation operator (see [8], [7], [25]) which satisfies for $0 \le k < 2$

$$\|u - \Pi_h u\|_{H^k(\Omega)} \le Ch^{2-k}|u|_{H^2(\Omega)}.$$

Then the decomposition

$$\varphi_h^{(1)} := I_h \Pi_H \varphi_h, \ \varphi_h^{(2)} := \varphi_h - \varphi_h^{(1)}$$

fulfills (note that $H = \varepsilon$)

$$\left\|\varphi_h^{(2)}\right\|_{L^2(\Omega)} \le C\varepsilon\|\nabla\varphi_h\|_{L^2(\Omega)}, \quad \left\|\nabla\varphi_h^{(2)}\right\|_{L^2(\Omega)} \le C\|\nabla\varphi_h\|_{L^2(\Omega)}. \tag{29}$$

We now have to estimate

$$a^\varepsilon(u_h^\varepsilon - p_{h\leftarrow H}u_H^0, \varphi_h) = a^\varepsilon(u_h^\varepsilon - p_{h\leftarrow H}u_H^0, \varphi_h^{(1)}) + a^\varepsilon(u_h^\varepsilon - p_{h\leftarrow H}u_H^0, \varphi_h^{(2)}) \tag{30}$$

which is done in the following.

2. The second term on the right-hand side of (30) can be estimated as follows:

$$a^\varepsilon(u_h^\varepsilon - p_{h\leftarrow H}u_H^0, \varphi_h^{(2)}) = (f, \varphi_h^{(2)})_{L^2(\Omega)} + a^\varepsilon(p_{h\leftarrow H}u_H^0, \varphi_h^{(2)}).$$

Here, the first term on the right-hand side is of the right order because

$$\begin{aligned}(f, \varphi_h^{(2)})_{L^2(\Omega)} &\le C\varepsilon\|f\|_{L^2(\Omega)}\|\nabla\varphi_h\|_{L^2(\Omega)} \\ &\le C\mu_1'\varepsilon\left\|D^2u^0\right\|_{L^2(\Omega)}\|\nabla\varphi_h\|_{L^2(\Omega)}\end{aligned}$$

because of (29). To handle the second term, we first replace u_H^0 by u^0, which may be done since, according to standard finite element estimates,

$$\left\|\nabla(u_H^0 - u^0)\right\|_{L^2(\Omega)} \le CH\left\|D^2u^0\right\|_{L^2(\Omega)}.$$

Second, we have for every cell $e \in \mathscr{T}_H$ (D_k^e denoting a suitable projection of ∇u^0 on constant vectors)

$$\begin{aligned}\nabla\left(I_h \circ (\varepsilon\partial_k u_H^0 w_{k,h}^\varepsilon) - \varepsilon\partial_k u^0 w_{k,h}^\varepsilon\right) &\sim \nabla\left(I_h \circ (\varepsilon\partial_k u_H^0 w_{k,h}^\varepsilon) - \varepsilon D_k^e w_{k,h}^\varepsilon\right) \\ &\sim \varepsilon\nabla I_h \circ ((\partial_k u_H^0 - D_k^e)w_{k,h}^\varepsilon) \\ &\sim 0\end{aligned}$$

up to error terms of size

$$C\varepsilon \left\| w_{k,\mathscr{S}_Y} \right\|_{1,\infty} \left\| D^2 u^0 \right\|_{L^2(e)}. \tag{31}$$

We therefore obtain

$$\begin{aligned} a^\varepsilon(p_{h\leftarrow H} u_H^0, \varphi_h^{(2)}) &\sim \int_\Omega \left(\mathscr{A}_{ij}^\varepsilon \partial_j u_H^0 + \mathscr{A}_{ij}^\varepsilon \partial_j (\varepsilon \partial_k u_H^0 w_{k,h}^\varepsilon) \right) \partial_i \varphi_h^{(2)} dx \\ &\sim \int_\Omega \left(\mathscr{A}_{ij}^\varepsilon \partial_j u^0 + \mathscr{A}_{ij}^\varepsilon \partial_k u^0 \partial_{y_j} w_{k,\mathscr{S}_Y} \right) \partial_i \varphi_h^{(2)} dx \\ &\sim \int_\Omega \left(\mathscr{A}_{ij}^\varepsilon + \mathscr{A}_{ik}^\varepsilon \partial_{y_k} w_{j,\mathscr{S}_Y} \right) \partial_j u^0 \partial_i \varphi_h^{(2)} dx \end{aligned}$$

up to error terms of size

$$C\mu_1 \varepsilon \left\| w_{k,\mathscr{S}_Y} \right\|_{1,\infty} \left\| D^2 u^0 \right\|_{L^2(\Omega)} \left\| \nabla \varphi_h^{(2)} \right\|_{L^2(\Omega)}. \tag{32}$$

Next, we expand $\varphi_h^{(2)}$ with respect to the nodal basis $\{\psi_\lambda\}_{\lambda \in \Lambda}$ of $\mathscr{S}_h$:

$$\int_\Omega \left(\mathscr{A}_{ij}^\varepsilon + \mathscr{A}_{ik}^\varepsilon \partial_{y_k} w_{j,\mathscr{S}_Y} \right) \partial_j u^0 \partial_i \varphi_h^{(2)} = \sum_{\lambda \in \Lambda} \varphi_h^{(2)}(x_\lambda) \int_{\mathrm{supp}(\psi^\lambda)} \left(\mathscr{A}_{ij}^\varepsilon + \mathscr{A}_{ik}^\varepsilon \partial_k w_{j,\mathscr{S}_Y} \right) \partial_j u^0 \partial_i \psi_\lambda$$

Again, ∇u^0 may be replaced on $\mathrm{supp}(\psi_\lambda)$ by constant vectors D_λ up to an error of the form

$$\begin{aligned} &C\mu_1 \| w_{k,\mathscr{S}_Y} \|_{1,\infty} \sum_{\lambda \in \Lambda} h \left\| D^2 u^0 \right\|_{L^2(\mathrm{supp}(\psi_\lambda))} h^{-1} \varphi_h^{(2)}(x_\lambda) \mathrm{vol}(\mathrm{supp}(\psi_\lambda))^{\frac{1}{2}} \\ &\quad \le C\mu_1 \left\| w_{k,\mathscr{S}_Y} \right\|_{1,\infty} \left\| D^2 u^0 \right\|_{L^2(\Omega)} \left\| \varphi_h^{(2)} \right\|_{L^2(\Omega)} \\ &\quad \le C\mu_1 \varepsilon \left\| w_{k,\mathscr{S}_Y} \right\|_{1,\infty} \left\| D^2 u^0 \right\|_{L^2(\Omega)} \left\| \nabla \varphi_h^{(2)} \right\|_{L^2(\Omega)} \end{aligned}$$

because of (29). Since the remaining terms with the constant vectors vanish due to (20), the assertion is proved.

3. For the first term on the right-hand side of (30), we obtain

$$\begin{aligned} a^\varepsilon(u^\varepsilon - p_{h\leftarrow H} u_H^0, \varphi_h^{(1)}) &= (f, \varphi_h^{(1)})_{L^2(\Omega)} + a^\varepsilon(p_{h\leftarrow H} u_H^0, \varphi_h^{(1)}) \\ &= \tilde{a}^0(u^0, \varphi_h^{(1)}) + a^\varepsilon(p_{h\leftarrow H} u_H^0, \varphi_h^{(1)}) \\ &\sim \int_\Omega \left(\tilde{\mathscr{A}}_{ij}^0 - \mathscr{A}_{ij}^\varepsilon - \mathscr{A}_{ik}^\varepsilon \partial_{y_k} w_{j,\mathscr{S}_Y} \right) \partial_j u^0 \partial_i \varphi_h^{(1)} dx \end{aligned}$$

up to an error of size

$$C\mu_1 \left\| w_{k,\mathscr{S}_Y} \right\|_{1,\infty} \varepsilon \left\| D^2 u^0 \right\|_{L^2(\Omega)} \left\| \nabla \varphi_h^{(1)} \right\|_{L^2(\Omega)}.$$

Next, we decompose the integral into a sum over all elements $e \in \mathscr{T}_H$ (i.e. over all cells), and replace ∇u^0 by constant vectors D^e on each cell, which can be done up to errors of the form.

$$C(\mu_1' + \mu_1 \|w_{k,\mathscr{S}_Y}\|_{1,\infty})\varepsilon \|D^2 u^0\|_{L^2(\Omega)} \left\|\nabla \varphi_h^{(1)}\right\|_{L^2(\Omega)}. \tag{33}$$

The remaining term then is

$$\sum_{e \in \mathscr{T}_H} \int_e \left(\tilde{\mathscr{A}}^0_{ij} - \mathscr{A}^\varepsilon_{ij} - \mathscr{A}^\varepsilon_{ik} \partial_{y_k} w_{j,\mathscr{S}_Y}\right) D^e_j \partial_i \varphi_h^{(1)} dx. \tag{34}$$

Now, $\varphi_h^{(1)}$ was constructed by pointwise interpolation of some function which was multilinear on each cell. Therefore, since nodal evaluation interpolates linear functions exactly and anti-symmetric functions to anti-symmetric functions (because $\mathscr{T}_Y$ was assumed to be symmetric), $\varphi_h^{(1)}$ can be written as a (L^2-orthogonal) sum of a linear function plus functions that are anti-symmetric in at least two variables. Now, the integral over the linear function vanishes because of the definition of the homogenized coefficients. On the other hand, using a discrete analogue of (17) (which can be proved in precisely the same way), one easily checks that each of the integrals in (34) vanishes for those parts of $\varphi_h^{(1)}$ which are anti-symmetric in at least two variables (apply a reflection $S^{(l)} : y_l \mapsto 1 - y_l$ such that $l \neq j$ and $\varphi_h^{(1)}$ is anti-symmetric in y_l). Thus, the estimate follows. Note that at this point, symmetry is needed crucially except for the one-dimensional case where $\varphi_h^{(1)}$ is linear on each cell and no anti-symmetric parts appear. □

The following theorem shows that our discrete homogenization can be interpreted as a Galerkin approximation for the prolongation (21).

Theorem 3.4. 1. *In* (26), *instead of* $\tilde{a}^0$, *we may use the Galerkin form*

$$\tilde{a}^0_G(u_H, v_H) := a^\varepsilon(p_{h \leftarrow H} u_H, p_{h \leftarrow H} v_H). \tag{35}$$

2. *For practical computations, it is convenient to use the right-hand sides*

$$f_H(\varphi_H) := f(I_h \varphi_H) \tag{36}$$

or the Galerkin choice

$$f^G_H(\varphi_H) := f(p_{h \leftarrow H} \varphi_H) \tag{37}$$

in (26). *The choice* (36) *allows the same estimate as* (26). *The choice* (37) *even allows the improved* $L^2(\Omega)$*-approximation property*

$$\left\|u^\varepsilon_h - p_{h \leftarrow H} u^0_H\right\|_{L^2(\Omega)} \leq C\varepsilon^2 \|f\|_{L^2(\Omega)}. \tag{38}$$

Proof: 1. Since the bilinear forms $\tilde{a}^0_G$ and $\tilde{a}^0$ coincide on linear functions, we can apply Theorem A.1. The argument is as follows: one easily checks that the values

D_i from (71) coincide. According to Remark A.2, we only have to ensure that no degeneracy point $\xi \in [0,4]^n - 0$ with $\sigma(\xi) = 0$ exists. But if such a point existed, one could construct functions, oscillating with period $O(H)$ in at least one coordinate direction, having an L_2-norm of order 1 and an energy norm of order $O(H)$. Applying Poincaré's inequality on patches having the size of the oscillations, one obtains a contradiction for $H = \varepsilon \to 0$. Therefore, Theorem A.1 is applicable and shows that the stiffness matrix corresponding to $\tilde{a}_G^0$ is also a discretization of the homogenized problem (25), and that the corresponding solution of (26) approximates u^0 up to an energy error of size $O(H)$ (equivalent to $O(\varepsilon)$, since $\varepsilon = H$). Thus, estimate (27) remains valid.

2. First, because of Theorem 3.5, we have

$$(f - f_H)(\varphi_H) = \int_\Omega f(\varphi_H - I_h\varphi_H)dx \le CH\|f\|_{L^2(\Omega)}\|\nabla\varphi_H\|_{L^2(\Omega)}. \tag{39}$$

Consequently, the energy norm of the difference of the corresponding solutions is also of the same order, which shows that f_H may be used instead of f in (26).

Next, we observe that

$$(f_H^G - f_H)(\varphi_H) = \varepsilon \int_\Omega f\,\partial_k\varphi_H w_{k,h}^\varepsilon\, dx \le C\varepsilon\|f\|_{L^2(\Omega)}\|\nabla\varphi_H\|_{L^2(\Omega)}, \tag{40}$$

such that estimate (27) also follows if we use f_H^G instead of f in (26).

Then, in the Galerkin situation (i.e. if using both $\tilde{a}_G^0$ and f_H^G), the improved L^2-estimate follows with the usual duality argument: let v_h be the solution of the problem

$$a^\varepsilon(v_h, \varphi_h) = (u_h^\varepsilon - p_{h\leftarrow H}u_H^0, \varphi_h)_{L^2(\Omega)} \quad \forall\ \varphi_h \in \mathscr{S}_h, \tag{41}$$

and let v_H be the solution to

$$\begin{aligned}\tilde{a}_G^0(v_H, \varphi_H) &= a^\varepsilon(v_h, p_{h\leftarrow H}\varphi_H)\\ &= (u_h^\varepsilon - p_{h\leftarrow H}u_H^0, p_{h\leftarrow H}\varphi_H)_{L^2(\Omega)}, \quad \forall\ \varphi_H \in \mathscr{S}_H.\end{aligned} \tag{42}$$

Because of (35), (37), the Galerkin orthogonality

$$a^\varepsilon(u_h^\varepsilon - p_{h\leftarrow H}u_H^0, p_{h\leftarrow H}\varphi_H) = a^\varepsilon(u_h^\varepsilon, p_{h\leftarrow H}\varphi_H) - \tilde{a}_G^0(u_H^0, \varphi_H) = 0 \tag{43}$$

holds and yields

$$\begin{aligned}\left\|u_h^\varepsilon - p_{h\leftarrow H}u_H^0\right\|_{L^2(\Omega)}^2 &= a^\varepsilon(u_h^\varepsilon - p_{h\leftarrow H}u_H^0, v_h)\\ &= a^\varepsilon(u_h^\varepsilon - p_{h\leftarrow H}u_H^0, v_h - p_{h\leftarrow H}v_H)\\ &\le C\varepsilon^2\|f\|_{L^2(\Omega)}\left\|u_h^\varepsilon - p_{h\leftarrow H}u_H^0\right\|_{L^2(\Omega)}.\end{aligned}$$

The assertion then follows. Finally, due to Theorem A.1, this improved estimate remains valid if one uses $\tilde{a}^0$ instead of $\tilde{a}^0_G$. □

We used the following lemma in the proof:

Lemma 3.5. *Let $\mathcal{T}_h$ and $\mathcal{T}_H$ be two triangulations of the polygonal domain Ω. Assume that $h \lesssim H$ and denote by $\mathcal{S}_h$ and $\mathcal{S}_H$ the corresponding spaces consisting of (multi) linear finite elements. Let $I_h : C^0(\Omega) \to \mathcal{S}_h$ be the nodal evaluation operator. Assume that u_H is an approximation to $u \in H^2(\Omega)$ satisfying*

$$\|u - u_H\|_{L^2(\Omega)} \lesssim h^2 \left\|D^2 u\right\|_{L^2(\Omega)}. \tag{44}$$

Then, for $0 \leq k \leq 1$, there holds

$$|I_h u_H - u_H|_{H^k(\Omega)} \leq CH^{2-k}|u|_{H^2}. \tag{45}$$

Proof: We only sketch the proof, which can be done similarly to the proof of Theorem 1.5.22 of [25]: first, one replaces u_H by the function u. This in turn is replaced locally by linear polynomials. Since I_h does not change linear polynomials, one then can estimate the error by an inverse inequality. □

Remark 3.6. Note that the estimate (28) implies a certain optimality for the discretely homogenized coefficients given by (20) and (22), because any other choice of coefficients for the $\mathcal{S}_H$-problem (for example the choice (7) and (6) which was suggested in [14]) yields an $O(1)$ approximation error independent of ε.

Remark 3.7. Since h was allowed to be arbitrarily small compared with ε, the previous theorem yields an $O(\varepsilon^2)$ L^2-approximation for the continuous case as well, provided that, for the homogenized problem, the right-hand side is modified according to (37). Numerical tests confirm this observation, see Subsection 4.1.

Remark 3.8. It is possible to prove a result similar to Theorem 3.3 using less than H^2-regularity for domains with reentrant corners. We do not want to do this here, since it is technically more difficult, and one does not gain more insight. For a result in this direction, the reader is referred to [25].

Finally, we note that we can generalize the choice of the coarse grid:

Theorem 3.9. *If $\mathcal{T}_H$ is chosen different to the grid induced by the cell structure such that $\varepsilon \lesssim H$, then for the prolongation*

$$p_{h \leftarrow H} := p_{h \leftarrow \varepsilon} \circ I_\varepsilon, \tag{46}$$

estimates (27) *and* (38) *remain valid with H replacing ε.*

Proof: This follows immediately, since $I_\varepsilon u_H^0$ approximates the solution u^0 up to $O(H)$ in the energy norm according to Lemma 3.5. □

3.3. The Case of Soft Inclusions

While the results shown up to now are already quite satisfactory, they are not really applicable if the coefficients $\mathcal{A}$ fulfill (1) with $\frac{\mu_1}{\mu_0} \gg 1$. In this as well as in the following subsection, we consider special cases, where the condition of uniform ellipticity (with moderate constants) is violated.

Definition 3.10. *Let the unit cell Y be the union $\overline{Y_1} \cup \overline{Y_2}$ of two open, disjoint subsets Y_1, Y_2, and assume that the periodic repetition of Y_1 over the whole $\mathbb{R}^n$ is connected (see Fig. 2 where the light region could be Y_1 and the dark region Y_2). Assume that on Y_1 the matrix-valued function $\mathcal{A} : Y \to \mathbb{R}^n$ fulfills the uniform ellipticity condition (1) with constants μ_0, μ_1 and on Y_2 with constants κ_0, κ_1 with $\kappa_1 \leq \mu_0$. Then we say that $\mathcal{A}$ describes a soft inclusion.*

We can then prove the following analogue to Theorem 3.3:

Theorem 3.11. *Under the assumptions of Theorem* 3.3, *assume that the boundary $\partial Y_1 = \partial Y_0$ is aligned with edges of the cell triangulation $\mathcal{T}_Y$ (see again Fig. 2 for an example). Further assume that the operator $E : \mathcal{S}_Y \to \mathcal{S}_Y, u_h \mapsto \tilde{u}_h$ defined by*

$$\tilde{u}_h = u_h \text{ on } Y_1, \;\; a^\varepsilon(\tilde{u}_h, \varphi_h) = 0 \quad \forall\, \varphi_h \in \mathcal{S}_Y \text{ with } \varphi_h|_{Y_1} \equiv 0 \tag{47}$$

fulfills

$$\|E\tilde{u}_h\|_{H^1_{per}(Y)} \leq C_E \|u_h\|_{H^1_{per}(Y_1)}. \tag{48}$$

Define a weighted L^2-scalar product as

$$(u, v)_{L^2_\omega(\Omega)} = \int_\Omega u(x)v(x)\omega(x)dx \tag{49}$$

with the weight

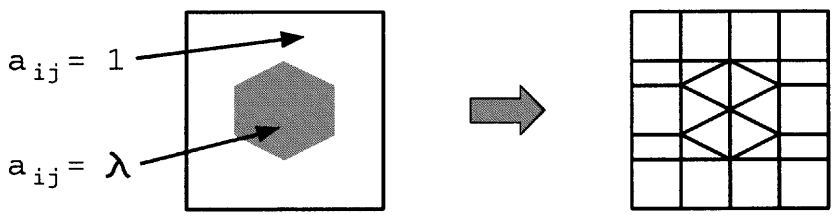

Fig. 2. Diffusion coefficient and coarse cell triangulation

$$\omega(x) := \left(\frac{1}{n}\sum_{i=1}^{n} \mathscr{A}^{\varepsilon}_{ii}(x)\right)^{-1}. \tag{50}$$

Then we have the error estimate

$$\left\|u^{\varepsilon}_h - p_{h\leftarrow H} u^0_H\right\|_{a^{\varepsilon}} \leq C\varepsilon \|f\|_{L^2_{\omega}(\Omega)} \tag{51}$$

with constant $C = C(n, \Omega, C_E, \mathscr{T}_Y, \mu_0, \mu_1, \mu_0', \mu_1', \left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty}, \frac{\kappa_1}{\kappa_0})$.

Proof: 1. For $i = 1, 2$ let

$$\Omega_i = \Omega \cap \varepsilon(\mathbb{Z}^n + Y_i)$$

Denoting $e_h := u^{\varepsilon}_h - p_{h\leftarrow H} u^0_H$, let $\tilde{e}_h = Ee_h$ be defined as in (47). Then

$$\begin{aligned} a^{\varepsilon}(e_h, e_h) &= a^{\varepsilon}(e_h, \tilde{e}_h) + a^{\varepsilon}(e_h, e_h - \tilde{e}_h) \\ &= a^{\varepsilon}(e_h, \tilde{e}_h) + a^{\varepsilon}(e_h - \tilde{e}_h, e_h - \tilde{e}_h). \end{aligned}$$

2. For the first part, we can apply (27) of Theorem 3.3 to obtain

$$\begin{aligned} a^{\varepsilon}(e_h, \tilde{e}_h) &\lesssim (\mu_1' + \mu_1 \left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty})\varepsilon \left\|D^2 u^0\right\|_{L^2(\Omega)} \|\nabla \tilde{e}_h\|_{L^2(\Omega)} \\ &\lesssim C_E \frac{\mu_1' + \mu_1 \left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty}}{\mu_0'} \varepsilon \|f\|_{L^2(\Omega)} \|\nabla e_h\|_{L^2(\Omega_1)} \\ &\lesssim C_E \frac{(\mu_1' + \mu_1 \left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty})}{\mu_0'} \sqrt{\frac{\mu_1}{\mu_0}} \varepsilon \|f\|_{L^2_{\omega}(\Omega)} \|\nabla e_h\|_{a^{\varepsilon}}. \end{aligned}$$

3. For the second part, setting $\varphi_h := e_h - \tilde{e}_h$, we have to estimate

$$a^{\varepsilon}(\varphi_h, \varphi_h) = a^{\varepsilon}(e_h, \varphi_h) = a^{\varepsilon}(u^{\varepsilon}_h, \varphi_h) - a^{\varepsilon}(p_{h\leftarrow H} u^0_H, \varphi_h).$$

For $a^{\varepsilon}(u^{\varepsilon}_h, \varphi_h)$, we have

$$a^{\varepsilon}(u^{\varepsilon}_h, \varphi_h) = (f, \varphi_h)_{L^2(\Omega_2)} \leq \|f\|_{L^2(\Omega_2)} \|\varphi_h\|_{L^2(\Omega_2)} \lesssim \sqrt{\frac{\kappa_1}{\kappa_0}} \varepsilon \|f\|_{L^2_{\omega}(\Omega)} \|\varphi_h\|_{a^{\varepsilon}}$$

where we used Poincaré's inequality (note that φ_h vanishes on parts of the cell).

For $a^{\varepsilon}(p_{h\leftarrow H} u^0_H, \varphi_h)$, we proceed exactly as in step 2 of the proof of Theorem 3.3: first, we expand $\varphi_h := e_h - \tilde{e}_h$ with respect to the nodal basis $\{\psi^{(\lambda)}\}_{\lambda\in\Lambda}$ of $\mathscr{S}_h$, then we replace ∇u^0 locally by constant vectors D_λ. This can be done up to an error of size

$$
\begin{aligned}
&C\left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty}\kappa_1\left\|D^2u^0\right\|_{L^2(\Omega_2)}\|\varphi_h\|_{L^2(\Omega_2)}\\
&\quad\le C\left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty}\kappa_1\frac{1}{\mu_0'}\|f\|_{L^2(\Omega)}\varepsilon\|\nabla\varphi_h\|_{L^2(\Omega_2)}\\
&\quad\le C\left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty}\frac{\kappa_1}{\mu_0'}\sqrt{\mu_1}\|f\|_{L^2_\omega(\Omega)}\varepsilon\frac{1}{\sqrt{\kappa_0}}\|\varphi_h\|_{a^\varepsilon}\\
&\quad\le C\left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty}\sqrt{\frac{\kappa_1}{\kappa_0}}\frac{\sqrt{\mu_0\mu_1}}{\mu_0'}\varepsilon\|f\|_{L^2_\omega(\Omega)}\|\varphi_h\|_{a^\varepsilon},
\end{aligned}
$$

since we assumed $\kappa_1 \le \mu_0$. The remaining term again vanishes, since the $w_{k,\mathscr{S}_Y}$ are solutions to the discrete cell problems (20). □

Remark 3.12. The constant in the assertion of Theorem 3.11 still depends on $\left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty}$ (which can be slightly relaxed to $\left\|w_{k,\mathscr{S}_Y}\right\|_{1,p}$). Thus, one has to ensure that this quantity does not depend too heavily on $\frac{\mu_0}{\kappa_1}$. This may be done in several situations. For example, it is shown in [25] for cell triangulations containing only a small number of unknowns. On the other hand, for fine cell triangulations, the regularity of the solutions $w_{k,\mathscr{S}_Y}$ of the discrete cell problems (20) will be determined essentially by the regularity of the solutions w_k of the continuous cell problems (7). Here, good behavior of $\left\|w_{k,\mathscr{S}_Y}\right\|_{1,\infty}$ can be expected for smooth coefficients and/or smooth interfaces between regions where jumps of the coefficients occur.

3.4. The Case of Stiff Inclusions

We now consider the case of an inlay with a large diffusion coefficient surrounded by a matrix where the diffusion coefficient is relatively small.

Definition 3.13. *Let the unit cell Y be the union $\overline{Y_1} \cup \overline{Y_2}$ of two open, disjoint subsets Y_1, Y_2, and assume that the periodic repetition of Y_2 over the whole $\mathbb{R}^n$ is not connected. Assume that on Y_1 the matrix-valued function $\mathscr{A} : Y \to \mathbb{R}^n$ fulfills the uniform ellipticity condition* (1) *with constants μ_0, μ_1 and on Y_2 with constants κ_0, κ_1 with $\kappa_0 \ge \mu_1$. Then we say that $\mathscr{A}$ describes a stiff inclusion.*

Throughout this subsection we only handle the following special case:

Definition 3.14. *We say that $\mathscr{A}$ describes a simple stiff inclusion, if $\mathscr{A}$ describes a stiff inclusion and $\overline{Y_2}$ does not touch the boundary $\partial Y = \partial]0,1[^n$.*

We now prove a form of Theorem 3.3 that does not depend on κ_1. To achieve this, we have to modify the prolongation $p_{h\leftarrow H}$ such that

$$
\left\|p_{h\leftarrow H}u_H^0\right\|_{a^\varepsilon} \sim \left\|u_H^0\right\|_{a^0}. \tag{52}
$$

This is done by using, instead of $\mathscr{S}_H$, a modified space of finite element functions which is generated by functions on the reference cube that are linear in a neigh-

borhood of Y_2. Such a space $\tilde{\mathscr{S}}_H$ can be constructed in the following way: as in the proof of Theorem 3.3, a multilinear function can be written as an L^2-orthogonal sum of basis functions that are either linear or multilinear in two or more variables. Now one can modify this basis by replacing the multilinear functions by some which take the same boundary values on ∂Y and vanish on Y_2. Since $\overline{Y_2}$ does not touch ∂Y, the resulting basis functions can be smoothed. The resulting finite element space still contains the linear functions which guarantees that the usual error estimates are valid. Of course, the constants in the estimates will depend on $d(\overline{Y_2}, \partial Y)$.

Replacing $\mathscr{S}_H$ by $\tilde{\mathscr{S}}_H$, we can prove the following theorem:

Theorem 3.15. *Assume again the setting of Theorem* 3.3, *but, instead of* $\mathscr{S}_H$, *use the space* $\tilde{\mathscr{S}}_H$ *to define the approximate solution* u_H^0. *Then the estimate*

$$a^\varepsilon(u_h^\varepsilon - p_{h\leftarrow H} u_H^0, \varphi_h) \leq C(\mu_1' + \mu_1 \|w_{k,\mathscr{S}_Y}\|_{1,\infty})\varepsilon \|D^2 u^0\|_{L^2(\Omega)} \|\nabla \varphi_h\|_{L^2(\Omega)} \tag{53}$$

is valid with $C = C(n, \Omega, \mathscr{T}_Y, \operatorname{dist}(\overline{Y_2}, \partial Y))$. *Setting* $\varphi_h := u_h^\varepsilon - p_{h\leftarrow H} u_H^0$, *we immediately obtain the energy-norm error estimate*

$$\|\nabla(u_h^\varepsilon - p_{h\leftarrow H} u_H)\|_{a^\varepsilon} \leq C \frac{\mu_1' + \mu_1 \|w_{k,\mathscr{S}_Y}\|_{1,\infty}}{\sqrt{\mu_0}} \varepsilon \|D^2 u^0\|_{L^2(\Omega)}. \tag{54}$$

Again we are in an H^2*-regular setting for the homogenized problem* (25), *such that we can replace* $\|D^2 u^0\|_{L^2(\Omega)}$ by $\frac{1}{\mu_0'}\|f\|_{L^2(\Omega)}$.

Proof: The proof of Theorem 3.3 can be transferred: First, step 1 can be carried out as before (note that also here the estimates (29) are valid for $\tilde{\mathscr{S}}_H$ since the linear functions are contained in the ansatz space).

Step 2 needs a little change: instead of replacing u_H^0 immediately by u^0, we have to replace it by a function $\tilde{u}^0 \in H^2(\Omega)$ with

$$\|\nabla(\tilde{u}^0 - u^0)\|_{L^2(\Omega)} \leq C\varepsilon \|D^2 u^0\|_{L^2(\Omega)}$$

which is equal to u_H^0 on the inlay part Ω_2. This can be done in a standard way by smoothing u_H^0 outside from some neighborhood of Ω_2. It is clear that the resulting function remains a good approximation of u^0, albeit with constants depending on $d(\overline{Y_2}, \partial Y)$. After this change, the argument runs as before: since no replacement is done on Ω_2, the size of the diffusion coefficient there does not enter into the error term (31), and since D_λ may be chosen to be equal to $\nabla \tilde{u}^0$ on the inlays, the error term (32) also keeps the same form.

For step 3, we note that

$$a^0(u^0, \varphi_h^{(1)}) \sim a^0(\tilde{u}^0, \varphi_h^{(1)})$$

up to an error term of the form

$$C\varepsilon\mu_1' \left\|D^2 u^0\right\|_{L^2(\Omega)} \left\|\nabla \varphi_h^{(1)}\right\|_{L^2(\Omega)},$$

which appears in (33) anyhow. Therefore, (33) also remains unchanged, and step 4 can be performed using $\tilde{u}^0$ instead of u^0. □

Corollary 3.16. *Using the Galerkin modification*

$$f_H^G(\varphi_H) := f(p_{h\leftarrow H}\varphi_H) \tag{55}$$

we obtain again by a duality argument the estimate

$$\left\|u_h^\varepsilon - p_{h\leftarrow H} u_H^0\right\|_{L^2(\Omega)} \le C\varepsilon \left\|u_h^\varepsilon - p_{h\leftarrow H} u_H^0\right\|_{a^\varepsilon}, \tag{56}$$

and combining this with (54), *we get*

$$\left\|u_h^\varepsilon - p_{h\leftarrow H} u_H^0\right\|_{L^2(\Omega)} \le C\varepsilon^2 \|f\|_{L^2(\Omega)}. \tag{57}$$

4. Numerical Applications

4.1. An Improved L^2-Estimate

We first want to establish the improvement of L^2-convergence numerically by using the modified right-hand side (37) for the homogenized problem. This improvement is valid only for symmetric cells, but symmetric cells are often found in industrial applications, so this method may be very helpful in practice.

To check this estimate numerically, we consider a two-dimensional example with the diffusion coefficient

$$\mathscr{A}(y_1, y_2) = 1 + (\sin(\pi y_1)\sin(\pi y_2))^2. \tag{58}$$

We note that $\mathscr{A}$ fulfills (1) with $\mu_0 = 1, \mu_1 = 2$, and is smooth, which ensures smooth cell solutions w_k and a finite element approximation of u^ε satisfying

$$\left\|D^2 u^\varepsilon\right\|_{L^2(\Omega)} \le C\frac{1}{\varepsilon}\|f\|_{L^2(\Omega)}. \tag{59}$$

We now choose $f \equiv 10^4$, and compute u_h^ε on a uniform grid $\mathscr{T}_h$ of meshsize $h = \varepsilon^2$. Because of (59), this ensures that u^ε is approximated up to $O(\varepsilon)$ in the energy norm and $O(\varepsilon^2)$ in the L^2-norm. The cell solutions are computed on a mesh $\mathscr{T}_Y$ of meshsize ε, and u_H^0 on a mesh $\mathscr{T}_H$ with meshsize $H = \varepsilon$ as well. Using f_H^G from (37) as the right-hand side and the bilinear form $\tilde{a}^0$ (using $\tilde{a}_G^0$ makes almost no difference) for computing u_H^0, we obtain the values from Table 1, which indeed show an $O(\varepsilon)$-approximation error in the energy norm as well as an $O(\varepsilon^2)$-approximation error in the L^2-norm.

Table 1. Energy error and L^2-error for $f \equiv 10^4$

$1/\varepsilon$	4	6	8	10	12	14	16
Energy error	503.7	349.9	266.0	214.2	179.1	153.8	134.8
L^2 error	34.10	15.62	8.881	5.713	3.979	2.929	2.245

Table 2. Energy and L^2-error for rough f

$1/\varepsilon$	4	6	8	10	12	14	16
Energy error (with f_H)	767.7	544.1	418.5	338.8	284.2	244.6	214.5
L^2 error (with f_H)	32.36	15.54	9.771	7.043	5.498	4.515	3.836
Energy error (with f_H^G)	765.2	542.0	416.8	337.4	283.0	243.6	213.7
L^2 error (with f_H^G)	31.18	13.75	7.711	4.917	3.405	2.495	1.907

Since one easily checks that we have $f_H = f_H^G$ for constant f, this example does not show that the Galerkin modification is necessary. To see this, we have to look at rougher right-hand sides, preferably oscillating with period ε. For such f, it is easy to give counter-examples already in the one-dimensional case[2]. In the setting above, the choice

$$f = 10^5 \cdot \begin{cases} \frac{x}{\varepsilon} - \left[\frac{x}{\varepsilon}\right] - \frac{1}{2} & \text{if } x \le \frac{1}{2} \\ -\frac{x}{\varepsilon} + \left[\frac{x}{\varepsilon}\right] - \frac{1}{2} & \text{if } x > \frac{1}{2} \end{cases} \tag{60}$$

leads to the values of Table 2. It is obvious that the L^2-error does not behave like $O(\varepsilon^2)$, unless we use the Galerkin restriction.

4.2. A Robust Two-Grid Scheme

In this subsection, we apply Theorem 3.3, Theorem 3.11, and Theorem 3.15 to show the convergence of a suitable two-grid algorithm for solving (19) efficiently. We assume that we are in the setting described in Section 3. For our numerical tests, we further assume the special case of coefficients $\mathscr{A}$ shown in Fig. 2 which depend on the parameter λ.

As ingredients for the two-grid algorithm we require the following: on $\mathscr{S}_h$, we need a smoothing algorithm which we assume to be given as some approximation W_h to the stiffness matrix A_h such that

$$A_h \le \omega \frac{W_h + W_h^*}{2}. \tag{61}$$

As coarse grid space we take $\mathscr{S}_H$, the space of multilinear finite elements on the cube decomposition with meshsize $H = \varepsilon$, and as prolongation we use $p_{h \leftarrow H}$ de-

[2] For example, take a coefficient $a = 1$ in $Y_1 = [0, 1/3] \cup [2/3, 1]$ and $a = 2$ in $Y_2 = [1/3, 2/3]$. On $\Omega = [0, 1]$, one then can take the right-hand side $f = w$ on $[0, 1/2]$ and $f = -w$ on $[1/2, 1]$. A computation then shows that the peaks of u^ε and u^0 are both of order ε, but differ by a multiplicative factor unless one uses the modified right-hand side for the homogenized problem.

fined in (21). The coarse-grid matrix is chosen as $A_H = p^t_{h\leftarrow H} A_h p_{h\leftarrow H}$ (which means that we are in a Galerkin setting, so that convergence can be shown for any number of smoothing steps). We also assume that an (approximate) solver for the $\mathscr{S}_H$-problem is given by an A_H-approximation W_H. Such a solver is easy to obtain, because we have a constant coefficient problem on $\mathscr{S}_H$. For our numerical tests, we have used a multigrid V-cycle with one Gauss-Seidel post-smoothing step, but other methods also work well. The two-grid algorithm then reads

Algorithm 4.1. (*Two-grid cycle with post-smoothing*) $\mathrm{TG}(f_h) :=$

begin
$u_h := 0$
coarse-grid correction:
defect computation: $d_h := f_h - A_h u_h$
$u_h := u_h + p_{h\leftarrow H} A_H^{-1} p^t_{h\leftarrow H} d_h$
post-smoothing: $u_h := u_h - W_h^{-1}(A_h u_h - f_h)$
return u_h.
end

This two-grid cycle converges robustly provided the $\mathscr{S}_h$-smoother is chosen reasonably. For example, we have

Theorem 4.2. *In the uniformly elliptic case, the two-grid cycle with Gauss-Seidel smoothing converges with a convergence rate independent of* $\varepsilon \sim h$.

Proof: We show that conditions $(V_0) - (V_2)$ of [26] are satisfied. First, condition (V_0) corresponds to (61) which is satisfied for Gauss-Seidel smoothing. Next, the approximation property (V_1) for unsymmetric smoothers can be infered from Theorem 3.3 using [26], Section 3.1. Finally, condition (V_2) is trivial for the two-grid situation where we can choose the subspace decomposition orthogonal with respect to the energy scalar product. Thus, Theorem 2.5 of [26] is applicable and yields the result. □

This can be verified also numerically. Using the coefficient and the cell triangulation shown in Fig. 2 and fixing $\lambda = 0.1$, we indeed observe robust convergence with respect to variations in ε, see Table 3. Here, as well as in the following tests, the termination criterion was a defect reduction by a factor of 10^{-12}, and the convergence rate is a geometric average.

Table 3. Averaged convergence rate for varying ε

$1/\varepsilon$	4	8	16	32	64
Convergence rate	0.670	0.703	0.715	0.714	0.712

Next, we look at robustness with respect to soft inclusions. Here, Theorem 3.11 yields the approximation estimate needed for the application of the theory of [26] with constants independent of λ. Consequently, Theorem 4.2 still holds in the limit case of arbitrarily small λ.

We check this assertion numerically by fixing $\varepsilon = 1/16$ and varying λ which gives the results from Table 4. We see that our algorithm is indeed robust with respect to soft inclusions. The values are even slightly better for small λ than for the constant coefficient case $\lambda = 1.0$.

We now turn to the case of stiff inclusions, i.e. $\lambda \gg 1$. We cannot expect the above cycle with Gauss-Seidel smoothing to be robust in this case, and indeed, numerical experiments show that it converges very poorly, even if we use the modified prolongation from Subsection 3.4. The problem is the following: errors constant on the inlays are damped neither by the Gauss-Seidel smoother (since that smoother does not damp such errors robustly for $\lambda \gg 1$) nor by the coarse-grid correction (because these errors are not "smooth"). Fortunately, a cure is found quite easily and follows a pattern which Brandt in [6] calls the "golden multigrid rule": points connected strongly should be treated as a block. This remedy is possible here, since for $\lambda \gg 1$ the patches consisting of strongly connected points are isolated from each other and consist (on $\mathscr{S}_h$) only of very few points (which allows effective application of block-wise smoothers).

And indeed, if we use block Gauss-Seidel smoothing on the block structure given by strong coupling (thus, blocks are all nodes which do not lie on an inlay together with blocks consisting of all nodes lying in or at the boundary of the same inlay), we obtain the values from Table 5. We see that our cycle is again perfectly robust.

Also here, robustness is an easy consequence of the above theory:

Theorem 4.3. *In the above setting, let the diffusion coefficient describe a stiff inclusion. If we use the modified prolongation described in Subsection* 3.4 *together with a block Gauss-Seidel smoother for a block decomposition where all nodes touching an inlay are inside a block for $\mathscr{S}_h$-smoothing, the corresponding two-grid cycle converges with a convergence rate independent of the number of cells.*

Table 4. Averaged convergence rate for soft inclusions

λ	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
Convergence rate	0.755	0.715	0.711	0.710	0.710	0.710	0.710

Table 5. Averaged convergence rate for stiff inclusions

λ	10^1	10^2	10^3	10^4	10^5	10^6
Convergence rate	0.621	0.619	0.618	0.618	0.618	0.617

Proof: The argument is very similar to that of Section 3.2 in [26]. Again, conditions (V_0) and (V_2) of [26] are easily seen to be satisfied. To show (V_1), we observe that

$$W_{h,symm} := W_h(W_h + W_h^* - A_h)^{-1}W_h^* = W_h(2\widetilde{W}_{h,symm} - A_h)^{-1}W_h^*$$

is spectrally equivalent to the block-diagonal D_h of the stiffness matrix A_h with respect to the above block structure (strongly coupled unknowns of the inlay forming a block). Therefore, (V_1) follows from

$$\|Qu_h\|_{D_h} \leq C\|u_h\|_{A_h} \tag{62}$$

with Q denoting the coarse grid correction operator. But on $\mathcal{S}_h$ we have

$$(\cdot,\cdot)_{D_h} \sim (\cdot,\cdot)_{a^\varepsilon,\Omega_2} + Ch^{-2}(\cdot,\cdot)_{L^2(\Omega)},$$

such that (62) follows by using

$$\|Qu_h\|_{a^\varepsilon,\Omega_2} \leq \|Qu_h\|_{A_h} \leq \|u_h\|_{A_h}$$

together with (56). □

Some concluding remarks are called for. It usually pays off to use more than one smoothing step on $\mathcal{S}_h$, which can also be shown theoretically, at least for the case of symmetric smoothing operations by a well-known argument, see e.g. [16]. Further, it can be useful to choose the meshsize H smaller than ε to decrease the "frequency gap" between $\mathcal{S}_h$ and $\mathcal{S}_H$. The above convergence results remain valid because of Theorem 3.9.

4.3. A Robust Multigrid Scheme

We now use the results of the previous subsection to construct robust multigrid algorithms for diffusion problems with periodically oscillating coefficients which are symmetric in the sense of Definition 2.2. Thus, for a model situation, we solve the problem mentioned in the introduction: for the above kind of diffusion equations, our algorithm converges robustly with respect to period and amplitude of the oscillation.

Adopting the setting of Subsection 4.2 and setting $\mathcal{T}_{h_0} := \mathcal{T}_h$, we obtain further triangulations $\mathcal{T}_{h_1}, \ldots, \mathcal{T}_{h_J}$ by uniform refinement. We denote the corresponding spaces of multilinear finite elements by $\mathcal{S}_{h_0}, \ldots, \mathcal{S}_{h_J}$. For this range of levels, the construction of a suitable multigrid algorithm would be straightforward if we could afford the exact solving of the $\mathcal{S}_{h_0}$-problem (see [33] or [26]). But again, if $\mathcal{S}_{h_0}$ is already very large, this may not be possible. As a remedy, we propose to use the two-grid algorithm described in Subsection 4.2 above. Given suitable smoothers W_k on $\mathcal{S}_{h_k}$ and defining prolongation/restriction by the finite-element

inclusion operator $I_{k\leftarrow k-1}$ and its adjoint, we obtain the following algorithm which executes the two-grid algorithm from Subsection 4.2 for solving the coarse-grid problem. We call it a multi-scale cycle because of Fig. 3, which depicts a situation where the homogenized problem is also solved with multigrid.

Algorithm 4.4. (*Multi-scale cycle*) $\mathrm{MSC}(k, f_k) :=$

begin
  if $k = 0$
    return $u_0 := \mathrm{TG}(f_0)$.
  else
    $u_k := 0$
    coarse-grid correction:
      defect computation: $d_k := f_k - A_k u_k$
      $u_k := u_k + I_{k\leftarrow k-1}\mathrm{MSC}(k-1, I^*_{k\leftarrow k-1} d_k)$
    post-smoothing: $u_k := u_k - W_k^{-1}(A_k u_k - f_k)$
    return u_k.
  endif
end

Robust convergence of this cycle can also be verified numerically: in the setting of Subsection 4.2, we fix $\varepsilon = 1/16$ and refine $\mathscr{T}_h$ three times (i.e. $J = 3$). We use Gauss-Seidel smoothing again, but since a single Gauss-Seidel step usually yields a convergence rate of about 0.5, it is sensible to improve the convergence of the two-grid algorithm by performing an additional pre-smoothing step (otherwise the convergence of the whole cycle would be determined by the coarse-grid correction). Since this pre-smoothing step does not increase the energy error, the

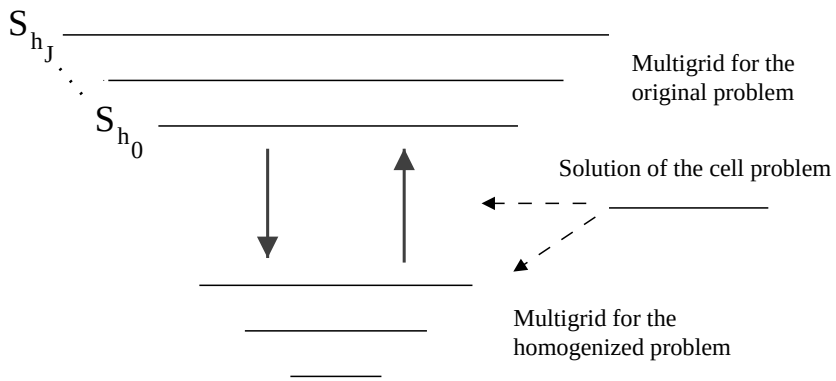

Fig. 3. Multi-scale cycle

Table 6. Convergence rates for the whole multi-scale cycle

λ	10^{-6}	10^{-4}	10^{-2}	10^{0}	10^{2}	10^{4}	10^{6}
Convergence rate	0.498	0.498	0.497	0.514	0.482	0.498	0.498

theoretical results can be carried over easily. The results for different values of λ (using the block Gauss-Seidel smoother for large values of $\lambda > 1$) are given in Table 6 and show that the whole cycle is perfectly robust.

A. An Approximation Estimate for Structured Grids

Theorem A.1. *On a structured grid with $(N+1)^n$ grid points*

$$x_{\vec{i}} = (i_1 h, \ldots, i_n h) \tag{63}$$

with $\vec{i} = (i_1, \ldots, i_n), i_j \in \{0, \ldots, N\}$ and meshsize $h = 1/N$, we consider a difference equation of the form

$$\frac{1}{h^2} \sum_{\vec{\alpha} \in \{-1,0,1\}^n} a_{\vec{\alpha}} x_{\vec{i}+\vec{\alpha}} = f_{\vec{i}} \quad \forall \vec{i} = (i_1, \ldots, i_n), i_j = 1, \ldots, N-1 \tag{64}$$

$$x_{\vec{i}+\vec{\alpha}} = 0 \quad \forall \vec{i} : \exists j : i_j = 0 \vee i_j = N. \tag{65}$$

If $a_{\vec{\alpha}}$ is symmetric, i.e.

$$a_{\vec{\alpha}} = a_{-\vec{\alpha}}, \tag{66}$$

one easily checks that the eigenvectors $\psi^{\vec{k}}, \vec{k} \in \{1, \ldots, N-1\}^n$ of the difference operator are given by

$$\psi^{\vec{k}}_{\vec{i}} = \prod_{j=1}^{n} \sin(\pi h i_j k_j), \quad \vec{i} \in \{0, \ldots, N\}^n \tag{67}$$

with

$$\lambda_{\vec{k}} = \frac{1}{h^2} \sum_{\vec{\alpha} \in \{-1,0,1\}^n} a_{\vec{\alpha}} \prod_{i=1}^{n} (\cos(\pi h k_i))^{|\alpha_i|}. \tag{68}$$

Setting

$$\sigma(\xi) := \sum_{\vec{\alpha} \in \{-1,0,1\}^n} a_{\vec{\alpha}} \prod_{i=1}^{n} (1 - \frac{1}{2}\xi_i)^{|\alpha_i|}, \tag{69}$$

we further assume

$$\sigma(0) = 0, \tag{70}$$

$$D_i := \frac{\partial \sigma}{\partial \xi_i}|_{\xi=0} = -\frac{1}{2} \sum_{\{a_{\vec{\alpha}} \in \{-1,0,1\}^n, \alpha_i \neq 0\}} a_{\vec{\alpha}} > 0, \quad \forall\, i = 1, \ldots, n, \tag{71}$$

$$\left| \frac{1}{\sigma(\xi)} - \frac{1}{\vec{D} \cdot \xi} \right| \leq C_\sigma, \quad \forall \xi \in [0,4]^n - 0. \tag{72}$$

Then (64), (65) *is a second-order discretization of*

$$-\sum_{i=1}^{n} D_i \partial_i^2 u(x) = f(x), \quad x \in \Omega := [0,1]^n, \tag{73}$$

$$u(x) = 0, \quad x \in \partial\Omega, \tag{74}$$

where f may be defined as $f = \sum_{\vec{\imath}} f_{\vec{\imath}} \varphi_{\vec{\imath}}$ with the usual multilinear finite-element basis functions of the cube decomposition associated with the grid. Especially, given $a_{\vec{\alpha}}$ and $a'_{\vec{\alpha}}$ where the associated functions σ and σ' fulfill (66), (70), (71) *with the same values $D_i, i = 1, \ldots n$, and* (72) *with constants C_σ and $C_{\sigma'}$, the difference of the corresponding solutions x and x' can be estimated as*

$$\left(h^n \sum_{\vec{\imath}} |x_{\vec{\imath}} - x'_{\vec{\imath}}|^2 \right)^{\frac{1}{2}} \leq (C_\sigma + C_{\sigma'}) h^2 \left(h^n \sum_{\vec{\imath}} f_{\vec{\imath}}^2 \right)^{\frac{1}{2}}. \tag{75}$$

Proof: This is done by eigenmode analysis: if $x = \sum_{\vec{k}} \hat{x}_{\vec{k}} \psi_{\vec{k}}$ and $f = \sum_{\vec{k}} \hat{f}_{\vec{k}} \psi_{\vec{k}}$, we have

$$\hat{x}_{\vec{k}} - \hat{x'}_{\vec{k}} = \left(\frac{1}{\lambda_{\vec{k}}} - \frac{1}{\lambda'_{\vec{k}}} \right) \hat{f}_{\vec{k}}, \tag{76}$$

and the assertion follows from

$$\begin{aligned} \left| \frac{1}{\lambda_{\vec{k}}} - \frac{1}{\lambda'_{\vec{k}}} \right| &\leq h^2 \sup_{\xi \in [0,4]^n - 0} \left| \frac{1}{\sigma(\xi)} - \frac{1}{\sigma'(\xi)} \right| \\ &\leq h^2 \sup_{\xi \in [0,4]^n - 0} \left| \frac{1}{\sigma(\xi)} - \frac{1}{\vec{D} \cdot \xi} \right| + \left| \frac{1}{\sigma'(\xi)} - \frac{1}{\vec{D} \cdot \xi} \right| \\ &\leq (C_\sigma + C_{\sigma'}) h^2 \end{aligned}$$

and the fact that the eigenmodes are orthogonal with respect to the Euclidean scalar product. □

Remark A.2. For (72), it is necessary and sufficient that σ from (69) does not vanish in $[0,4]^n - 0$.

Proof: It is obvious that the condition is necessary, and we will proceed with showing that it is also sufficient:

Since σ is obviously two times continuously differentiable, a neighborhood $U_\delta(0)$ exists in which $\sigma(\xi) > \frac{1}{2}\left|\vec{D}\cdot\xi\right|$ and $\left|\sigma(\xi) - \vec{D}\cdot\xi\right| \le C|\xi|^2$. Therefore

$$\left|\frac{1}{\sigma(\xi)} - \frac{1}{\vec{D}\cdot\xi}\right| \le \frac{\left|\vec{D}\cdot\xi - \sigma(\xi)\right|}{\sigma(\xi)\left|\vec{D}\cdot\xi\right|} \le \frac{2C}{\vec{D}^2}. \tag{77}$$

The complementary set $[0,4]^n - U_\delta(0)$ is compact, such that the function σ takes a positive minimum σ_{min} there. Therefore

$$\left|\frac{1}{\sigma(\xi)} - \frac{1}{\vec{D}\cdot\xi}\right| \le \frac{1}{\sigma_{min}} + \frac{1}{\delta\left|\vec{D}\right|}. \tag{78}$$

□

Acknowledgement

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) through the Sonderforschungsbereich 359 (SFB 359) at the University of Heidelberg.

References

[1] Alcouffe, R. E., Brandt, A., Dendy, J. E., Painter, J. W.: The multigrid method for the diffusion equation with strongly discontinuous coefficients. SIAM J. Sci. Stat. Comput. *2*, 430–454 (1981).
[2] Bakhvalov, N., Panasenko, G.: Homogenization: averaging processes in periodic media. Dordrecht: Kluwer, 1989.
[3] Bank, R., Xu, J.: An algorithm for coarsening unstructured meshes. Numer. Math. *73*, 1–36 (1996).
[4] Bensoussan, A., Lions, J.-L., Papanicolaou, G.: Asymptotic analysis for periodic structures. Amsterdam: North-Holland, 1978.
[5] Braess, D.: Towards algebraic multigrid for elliptic problems of second order. Computing *55*, 379–393 (1995).
[6] Brandt, A.: Multigrid techniques: 1984 guide with applications to fluid dynamics, GMD–Studien Nr. 85, Gesellschaft für Mathematik und Datenverarbeitung, St. Augustin, 1984.
[7] Brenner, S. C., Scott, L. R.: The mathematical theory of finite element methods. Texts in Applied Mathematics. New York: Springer, 1994.
[8] Clément, P.: Approximation by finite element functions using local regularization, RAIRO, 9, R-2 (1975), pp. 77–84.
[9] Dendy, J. E., Hyman, J. M., Moulton, J. D.: The black box multigrid numerical homogenization algorithm. J. Comput. Phys. *141*, 1–29 (1998).
[10] deZeeuw, P. M.: Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. J. Comput. Appl. Math. *33*, 1–27 (1990).
[11] Edwards, M. G., Rogers, C. F.: Multigrid and renormalization for reservoir simulation. In: Multigrid methods IV. Proceedings of the Fourth European Multigrid Conference, Amsterdam, July 6–9, 1993, pp. 189–200. ISNM, Vol. 116. Basel: Birkhäuser, 1994.

[12] Efendiev, Y. R., Hou, T. Y., Wu, X. H.: Convergence of a nonconformal multiscale finite element method. SIAM J. Numer. Anal. *37*, 888–910 (2000).
[13] Engquist, B., Luo, E.: Multigrid methods for differential equations with highly oscillatory coefficients. Presented at the Copper Mountain Conference, 1993.
[14] Engquist, B., Luo, E.: Multigrid methods for differential equations with highly oscillatory coefficients. SIAM J. Numer. Anal. *34*, 2254–2273 (1997).
[15] Griebel, M., Knapek, S.: A multigrid-homogenization method, In: Adaptive methods – algorithms, theory and applications, pp. 142–157. Notes on Numerical Fluid Mechanics, Vol. 59. Vieweg: Braunschweig, 1994.
[16] Hackbusch, W.: Multigrid methods and applications, Computational Mathematics, Vol. 4. Berlin, Hiedelberg, New York: Springer, 1985.
[17] Hackbusch, W., Sauter, S.: Composite finite elements for the approximation of partial differential equations on domains with complicated microstructure. Numer. Math. *55*, 447–472 (1997).
[18] Hou, T. Y., Wu, X. H., Cai, Z.: Convergence of a nonconformal multiscale finite element method for elliptic problems with rapidly oscillating coefficients. Math. Comp. *68*, 913–943 (1999).
[19] Jikov, V. V., Kozlov, S. M., Oleinik, O. A.: Homogenization of differential operators and integral functionals. Berlin, Heidelberg, New York, Tokyo: Springer, 1994.
[20] Knapek, S.: Matrix-dependent multigrid homogenization for diffusion problems. SIAM J. Sci. Comput. *20*, 515–533 (1998).
[21] Lions, J.-L.: Some methods in the mathematical analysis of systems and their control. New York: Gordon and Breach, 1981.
[22] Matache, A. M., Babuska, I., Schwab, C.: Generalized p-FEM in homogenization, SAM, ETH Zürich, CH-8092 Zürich, Switzerland, Research report 1999-01 (1999).
[23] Morgan, R. C., Babuska, I.: An approach for constructing families of homogenized equations for periodic media, I + II. Math. Anal. *22*, 1–33 (1991).
[24] Moskow, S., Vogelius, M.: First order corrections to the homogenised eigenvalues of a periodic composite medium, a convergence proof. Proc. R. Soc. Edinb., Sect. *A 127*, 1263–1299 (1997).
[25] Neuss, N.: Homogenisierung und Mehrgitter. PhD thesis, Universität Heidelberg, Heidelberg, Germany, 1995. Also as ICA-Report N96/7, ICA Stuttgart, Germany, 1996.
[26] Neuss, N.: V-cycle convergence with unsymmetric smoothers and application to an anisotropic model problem. SIAM J. Num. Anal. *35*, 1201–1212 (1998).
[27] Oleinik, O. A., Shamaev, A. S., Yosifian, G. A.: Mathematical problems in elasticity and homogenization. Amsterdam: North-Holland, 1992.
[28] Reusken, A.: Multigrid with matrix-dependent transfer operators for convection–diffusion problems. In: Multigrid Methods IV, Proceedings of the Fourth European Multigrid Conference, Amsterdam, July 6–9, 1993, pp. 269–280. ISNM, Vol. 116. Basel: Birkhäuser, 1994.
[29] Ruge, J. W., Stüben, K.: Algebraic multigrid (AMG). In: Multigrid Methods (McCormick, S. F. ed.), pp. 73–130. Frontiers in Applied Mathematics, Vol. 3. Philadelphia: SIAM, 1987.
[30] Sanchez-Palencia, E.: Non-homogenous media and vibration theory. Lecture Notes in Physics, Vol. 127. Berlin, Heidelberg, New York: Springer, 1980.
[31] Vaněk, P., Mandel, J., Brezina, M.: Algebraic multigrid based on smoothed aggregation for second and fourth order problems. Computing *56*, 179–196 (1996).
[32] Wagner, C., Kinzelbach, W., Wittum, G.: Schur-complement multigrid. A robust method for groundwater flow and transport problems. Numer. Math. *75*, 523–545 (1997).
[33] Zhang, X.: Multilevel Schwarz methods. Numer. Math. *63*, 521–539 (1992).

Nicolas Neuss
e-mail: Nicolas.Neuss@iwr.uni-heidelberg.de
WWW: http://www.iwr.uni-heidelberg.de/~Nicolas.Neuss

Willi Jäger
e-mail: jaeger@iwr.uni-heidelberg.de
WWW: http://www.iwr.uni-heidelberg.de/~Willi.Jäger

Gabriel Wittum
e-mail: wittum@iwr.uni-heidelberg.de

Interdisziplinäres Zentrum
für Wissenschaftliches Rechnen
Im Neuenheimer Feld 368
D-69120 Heidelberg
Germany