SIAM J. APPL. MATH.
Vol. 42, No. 6, December 1982

0036-1399/82/4206-0011 $1.00/0

A PROJECTION METHOD APPLIED TO DIFFUSION IN A PERIODIC STRUCTURE*

MICHAEL VOGELIUS† AND GEORGE PAPANICOLAOU†

Abstract. In this paper we analyze a method designed to replace PDE's with rapidly varying coefficients by PDE's with constant coefficients. This method is based on a combination of an asymptotic expansion and a variational principle. We show that for smooth data the method has the same approximation properties as the more standard approach based only on an asymptotic expansion. More importantly, we show that the present method works well also for nonsmooth data. This may be seen as a result of the optimal way in which the method treats boundary layers.

1. Introduction. We shall analyze a method for constructing effective constant coefficient equations corresponding to PDE's with rapidly varying coefficients. We consider the following model problem

$$
(1)\qquad \begin{aligned} &-\nabla_x \cdot \left(\mathbf{A}\left(\frac{x}{\varepsilon}\right)\nabla_x u^\varepsilon(x)\right) + b\left(\frac{x}{\varepsilon}\right)u^\varepsilon(x) = f(x) \quad \text{in } \Omega,\\ &u^\varepsilon(x) = 0 \quad \text{on } \partial\Omega, \end{aligned}
$$

where ε is a small positive parameter characterizing the length scale of variation of the coefficients relative to the size of Ω. Ω itself is a fixed bounded domain in $\mathbb{R}^n$ with a Lipschitz boundary. The elements of the coefficient matrix $\mathbf{A}(y) = \{a_{ij}(y)\}_{i,j=1}^n$ and the function $b(y)$ are bounded periodic functions of $y \in \mathbb{R}^n$ with period one in each direction, $\mathbf{A}(y)$ is assumed symmetric and uniformly positive definite, i.e., there is a constant $a_0 > 0$ such that

$$
a_0\xi^2 \leqq \langle \mathbf{A}(y)\xi, \xi\rangle \quad \forall \xi \in \mathbb{R}^n, \quad y \in \mathbb{R}^n
$$

and b is nonnegative.

We may think of $u^\varepsilon(x)$ as denoting the temperature of a conductor occupying Ω and having conductivity $\mathbf{A}(x/\varepsilon)$ which is rapidly varying.

It has been shown in some detail (cf. [1], [3], [5], [13], [16], [19]) that when ε is small and all the data is smooth $u^\varepsilon(x)$ is well approximated by the solution to the homogenized, constant coefficient, problem

$$
(2)\qquad \begin{aligned} &-\nabla_x \cdot (\mathscr{A}\nabla_x u) + \bar{b}u(x) = f(x) \quad \text{in } \Omega,\\ &u(x) = 0 \quad \text{on } \partial\Omega. \end{aligned}
$$

Here $\bar{}$ denotes averaging over a period cell, i.e.,

$$
\bar{b} = \iint_{[0,1]^n} b(y)\, dy
$$

and the matrix $\mathscr{A}$ is defined below.

Let χ_j, $1 \leqq j \leqq n$, be periodic functions satisfying

$$
(3)\qquad -\nabla_y \cdot (\mathbf{A}(y)\nabla_y \chi_j(y)) = \nabla_y \cdot (\mathbf{A}(y)e_j),
$$

where e_j is the standard jth basis vector in $\mathbb{R}^n$, and let χ denote the vector $(\chi_1, \chi_2, \cdots, \chi_n)$.

---

* Received by the editors June 3, 1981, and in revised form October 26, 1981. This research was supported by the U.S. Army Research Office under grant DAAG29-78-G-0177.

† Courant Institute of Mathematical Sciences, New York University, New York, New York 10012.

The matrix $\mathcal{A}$ is then given by

$$\mathcal{A} = \overline{\mathbf{A}(y)(\nabla_y \boldsymbol{\chi} + I)}.$$

It is clear that $\mathcal{A}$ is uniquely defined even though χ_j is determined only up to a constant from (3).

Let H^s be the standard Sobolev space of order s based on L^2 and let $W^{s,p}$ denote the corresponding space based on L^p, $p \neq 2$. The norms will be denoted by $\|\cdot\|_s$ and $\|\cdot\|_{s,p}$ respectively. $\mathring{H}^1(\Omega)$ refers to the closure of $C_0^\infty(\Omega)$ in $H^1(\Omega)$, i.e., the elements that vanish on $\partial\Omega$. We note that (3) is considered as an equation in variational form in $H^1_{\text{per}}([0, 1]^n)$ (the periodic elements of H^1).

The correspondence between $u^\varepsilon(x)$ of (1) and $u(x)$ of (2) for small ε is this:

$$u^\varepsilon \to u \quad \text{weakly in } H^1(\Omega) \quad \text{as } \varepsilon \to 0$$

and by compactness therefore $u^\varepsilon \to u$ strongly in $H^s(\Omega)$ for any $s < 1$. But one cannot show that $u^\varepsilon \to u$ strongly in $H^1(\Omega)$, which is convergence in the energy norm, and indeed this is false.

The big advantage of (2) over (1) is clearly that the highly oscillatory coefficients have been replaced by *constant* coefficients.

To understand why (1) can be approximated by (2) and how to construct approximations to u^ε that are close in the energy norm it is convenient to perform a multiple scales asymptotic expansion (cf. [3], [8], [12], [14]). We look for $u^\varepsilon(x)$ in the form

$$u^\varepsilon(x) = u(x) + \varepsilon u_1\left(x, \frac{x}{\varepsilon}\right) + \varepsilon^2 u_2\left(x, \frac{x}{\varepsilon}\right) + \cdots$$

We treat x and $y = x/\varepsilon$ as independent variables, insert the above expression into (1) and match equal powers of ε. This leads to a sequence of problems the solution of the first of which requires that $u_1(x, x/\varepsilon)$ has the form

$$u_1\left(x, \frac{x}{\varepsilon}\right) = \sum_{j=1}^{n} \chi_j\left(\frac{x}{\varepsilon}\right) \frac{\partial}{\partial x_j} u(x),$$

with χ_j, $1 \leqq j \leqq n$, as determined by (3). The second problem is similar to the first, but with a different right-hand side. It is the solvability requirement for this problem that leads to the effective equation (2) for u.

An interesting variation of this argument is found in [10]. There the effective equations are found entirely from considerations involving convergence of the energy, i.e., only a very weak form of matching.

In the references given above several other examples are treated: problems in fluid mechanics [3], [14], probabilistic problems [3], [12], etc. Connections with the averaging method for ordinary differential equations and other asymptotic problems are presented in detail in [12], while dynamic and high frequency problems are analyzed in [3].

In [3] it is shown by many methods, the most efficient and elegant being the one given by Tartar [16], that

$$\left\| u^\varepsilon(x) - \left[u(x) + \varepsilon \sum_{j=1}^{n} \chi_j\left(\frac{x}{\varepsilon}\right) m^\varepsilon(x) \frac{\partial}{\partial x_j} u(x) \right] \right\|_1 \leqq C\varepsilon^{1/2}, \tag{4}$$

provided $u \in H^2(\Omega)$, $(\partial/\partial x_j)u \in L^\infty(\Omega)$ and $\chi_j \in W^{1,\infty}([0, 1]^n)$, $1 \leqq j \leqq n$. $m^\varepsilon(x)$ is a function of the form

$$m^\varepsilon(x) = m\left(\frac{\text{dist}(x, \partial\Omega)}{\varepsilon}\right),$$

where $m \in W^{1,\infty}(\mathbb{R}_+)$ is such that $m(0)=0$ and $m(s)=1$ for $s \geqq 1$. (It is here implicitly assumed that $\partial\Omega$ is piecewise C^1.) The presence of the cutoff function m in (4) is crucial since we want the approximation to satisfy the homogeneous boundary condition.

Some clear disadvantages of (2) and (4) are

(a) The cutoff function m^ε can be chosen rather arbitrarily. This leaves open the very difficult question of how to select it optimally in order to obtain the smallest error.

(b) If the right-hand side f is not sufficiently regular and u therefore is not in $H^2(\Omega)$, then in general the expression

$$u(x)+\varepsilon \sum_{j=0}^{n} \chi_j\left(\frac{x}{\varepsilon}\right) m^\varepsilon(x) \frac{\partial}{\partial x_j} u(x) \tag{5}$$

does not make sense in $\mathring{H}^1(\Omega)$. Standard homogenization therefore fails to produce a family s^ε such that $u^\varepsilon - s^\varepsilon \to 0$ strongly in $\mathring{H}^1(\Omega)$.

(c) If the boundary $\partial\Omega$ is not sufficiently smooth (e.g., has corners) and u therefore is not in $H^2(\Omega)$ then the results as contained in [3] do not apply.

$$u(x)+\varepsilon \sum_{j=0}^{n} \chi_j\left(\frac{x}{\varepsilon}\right) m^\varepsilon(x) \frac{\partial}{\partial x_j} u(x)$$

may still make sense in $\mathring{H}^1(\Omega)$ provided f is at least in $L^2(\Omega)$ and $\chi_j \in W^{1,\infty}([0,1]^n)$, $1 \leqq j \leqq n$ (and for an appropriate cutoff $m^\varepsilon(x)$), but a separate analysis is needed to see if we recover the convergence in $\mathring{H}^1(\Omega)$ and at what rate. This analysis is to the authors' best knowledge not found in the literature.

(d) From a computational point of view

$$u(x)+\varepsilon \sum_{j=1}^{n} \chi_j\left(\frac{x}{\varepsilon}\right) m^\varepsilon(x) \frac{\partial}{\partial x_j} u(x)$$

is not very natural. If u is approximated by, say, a finite element method as u^Δ, then for the above expression to be in $H^1(\Omega)$ requires that $(\partial/\partial x_j)u^\Delta \in H^1(\Omega)$, $1 \leqq j \leqq n$. This means we have to use C^1-elements or alternatively treat (2) by a socalled mixed method.

The purpose of this paper is to introduce a new second order constant coefficient elliptic system that preserves the homogeneous boundary condition, leads directly to an approximation result of the same type as (4) and furthermore does not exhibit the disadvantages mentioned above. To be more specific:

(a) It will be shown that the solution to this system has a boundary-layer behavior identical to the optimal cutoff (to be defined) of

$$u(x)+\varepsilon \sum_{j=1}^{n} \chi_j\left(\frac{x}{\varepsilon}\right) \frac{\partial}{\partial x_j} u(x)$$

relative to the energy.

(b)–(c) In the case of a sufficiently smooth domain it is shown that we have strong $\mathring{H}^1(\Omega)$-convergence provided only $f \in H^s(\Omega)$ for some $-1<s$. Regarding nonsmooth domains we restrict ourselves to the case $n=2$ and show that if the boundary $\partial\Omega$ is piecewise smooth with a finite number of corners, then this system will provide a family s^ε such that $u^\varepsilon - s^\varepsilon \to 0$ strongly in $\mathring{H}^1(\Omega)$ as $\varepsilon \to 0$.

(d) Since the new system that we introduce is of second order it has a natural discretization by the finite element method using only C^0-elements. The analog of (5) in this case does not involve differentiation so we have no need for more regular elements.

For some numerical experiments, comparing this alternate method to the more standard technique, we refer to [17].

2. The nondegeneracy assumptions. The system we shall derive is essentially based on taking higher order terms in an asymptotic expansion and then incorporating them in the equations via a variational principle. It is therefore very natural that we have to make some kind of assumption to the effect that these terms are really present in the particular problem.

At various points in this paper we shall need either of the following two assumptions regarding the functions χ_j defined by (3).

$$\text{(N)}\qquad \begin{aligned}&\text{if } \beta_j \in \mathbb{R},\ 1 \leqq j \leqq n, \text{ are such that}\\ &\sum_{j=1}^{n} \beta_j \nabla_y \chi_j(y) = 0 \quad \forall y,\\ &\text{then } \beta_j = 0,\ 1 \leqq j \leqq n.\end{aligned}$$

$$\text{(NN)}\qquad \begin{aligned}&\text{if } \boldsymbol{\alpha}_j \in \mathbb{R}^n,\ 0 \leqq j \leqq n, \text{ and } \beta_j \in \mathbb{R},\ 1 \leqq j \leqq n \text{ are such that}\\ &\boldsymbol{\alpha}_0 + \sum_{j=1}^{n} \chi_j(y)\boldsymbol{\alpha}_j + \sum_{j=1}^{n} \beta_j \nabla_y \chi_j(y) = 0 \quad \forall y,\\ &\text{then } \boldsymbol{\alpha}_j = 0,\ \beta_j = 0,\ \forall j.\end{aligned}$$

Remark 2.1. If $\mathbf{A}(y) = a(y)I$, $a \in L^\infty(\mathbb{R}^n)$ and $a(y) \geqq a_0$, then the assumption (N) is clearly satisfied if and only if

$$\boldsymbol{\beta} \cdot \nabla_y a(y) = 0 \quad \forall y \Rightarrow \boldsymbol{\beta} = 0,$$

i.e., if and only if a is not totally independent of one particular direction.

Remark 2.2. The assumption (NN) is generally stronger than (N). If $n = 1$ the two are equivalent; this is because in that case, there is only one function χ and at the same time $\int_0^1 \chi(y)(d/dy)\chi(y)dy = 0$. For, $n = 1$ both assumptions are therefore equivalent to the requirement that the function a is not a constant.

In § 6 we shall take a closer look at what happens near or in the degenerate case. We show that although our elliptic system becomes singular the estimate corresponding to (4) is uniformly valid. This analysis is only carried out for $n = 1$.

3. The derivation of the system. Introduce the new independent variable $y = x/\varepsilon$ in (1). It is quite clear that if U^ε denotes the solution to

$$\text{(6)}\qquad \begin{aligned}&-(\nabla_x + \varepsilon^{-1}\nabla_y) \cdot (\mathbf{A}(y)(\nabla_x + \varepsilon^{-1}\nabla_y)U^\varepsilon(x, y)) + b(y)U^\varepsilon(x, y) = f(x),\\ &U^\varepsilon \text{ is periodic in } y \text{ with period 1, and}\\ &U^\varepsilon(x, y) = 0 \quad \forall x \in \partial\Omega, \forall y,\end{aligned}$$

then $u^\varepsilon(x) = U^\varepsilon(x, x/\varepsilon)$ solves (1).

The function $U^\varepsilon(x, y)$ itself has a physical meaning. For a fixed $y_0 \in [0, 1]^n$, $V^\varepsilon(x, y_0) = U^\varepsilon(x, \varepsilon^{-1}x + y_0)$ is the solution to the boundary value problem

$$\begin{aligned}&-\nabla_x \cdot (\mathbf{A}(\varepsilon^{-1}x + y_0)\nabla_x V^\varepsilon(x, y_0)) + b(\varepsilon^{-1}x + y_0)V^\varepsilon(x, y_0) = f(x) \quad \text{in } \Omega,\\ &V^\varepsilon(x, y_0) = 0 \quad \text{on } \partial\Omega.\end{aligned}$$

This problem is identical to (1) except that the periodic microstructure is no longer centered at the origin but at y_0. For the treatment of equations with rapidly oscillating random coefficients this formulation is essential. In that case y_0 is taken to be a uniformly distributed random variable with values in $[0, 1]^n$ and $V^\varepsilon(x, y_0)$ is interpreted as a function on Ω with values in $L^2(\mathcal{P}, d\mu)$ where $(\mathcal{P}, d\mu)$ is the probability space associated with y_0 (for more details see [13]).

Let $\mathcal{V}$ denote the set of functions

$$\left\{ v_0(x) + \varepsilon \sum_{j=1}^n \chi_j(y) v_j(x) \,\middle|\, v_j \in \mathring{H}^1(\Omega),\ 0 \leqq j \leqq n \right\},$$

where χ_j, $1 \leqq j \leqq n$, are as defined in (3). We can then "project" U^ε onto $\mathcal{V}$ in the semi-inner product associated with (6). By setting $y = x/\varepsilon$ in the result of this projection we obtain an approximation to $u^\varepsilon(x)$ of the form

$$u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) u_j^\varepsilon(x), \tag{7}$$

where u_0^ε, $\{u_j^\varepsilon\}_{j=1}^n \in \mathring{H}^1(\Omega)$ is the solution to

$$\begin{aligned} &-\nabla_x \cdot \begin{bmatrix} \bar{\mathbf{A}} & \varepsilon \overline{\mathbf{A}\chi_j} \\ \varepsilon \overline{\mathbf{A}\chi_k} & \varepsilon^2 \overline{\mathbf{A}\chi_k\chi_j} \end{bmatrix} \nabla_x \begin{bmatrix} u_0^\varepsilon \\ \{u_j^\varepsilon\} \end{bmatrix} + \begin{bmatrix} \bar{b} & -\nabla_x \cdot \overline{\mathbf{A}\nabla_y \chi_j} \\ \overline{\mathbf{A}\nabla_y\chi_k} \cdot \nabla_x & \overline{\mathbf{A}\nabla_y\chi_k\nabla_y\chi_j} \end{bmatrix} \begin{bmatrix} u_0^\varepsilon \\ \{u_j^\varepsilon\} \end{bmatrix} \\ &+ \begin{bmatrix} 0 & \varepsilon\overline{b\chi_j} \\ \varepsilon\overline{b\chi_k} & \varepsilon^2\overline{b\chi_k\chi_j} + \varepsilon\overline{\mathbf{A}(\nabla_y\chi_k\chi_j - \chi_k\nabla_y\chi_j)}\cdot\nabla_x \end{bmatrix} \begin{bmatrix} u_0^\varepsilon \\ \{u_j^\varepsilon\} \end{bmatrix} = \begin{bmatrix} f \\ \{\varepsilon\bar{\chi}_k f\} \end{bmatrix}. \end{aligned} \tag{8}$$

The system (8) for u_j^ε, $0 \leqq j \leqq n$, may also be obtained by averaging the variational principle associated with (1) as in Whitham's approach to modulation theory (cf. [18]). That is, in the expression

$$\frac{1}{2}\int_\Omega\int \left[\mathbf{A}\left(\frac{x}{\varepsilon}\right)\nabla_x V(x) \cdot \nabla_x V(x) + b\left(\frac{x}{\varepsilon}\right) V^2(x)\right] dx - \int_\Omega\int f(x)V(x)\,dx$$

we insert for V a function of the form $v_0(x) + \varepsilon \sum_{j=1}^n \chi_j(x/\varepsilon) v_j(x)$, keep x fixed and average the integrand with respect to $y = x/\varepsilon$ and then calculate the variation of the integral with respect to v_j, $0 \leqq j \leqq n$. This leads to the system (8).

We now prove two results about the structure of this system and its solutions.

LEMMA 3.1. *If the condition* (N) *is satisfied then* (8) *is a positive definite uniformly elliptic system. If*

$$B_\varepsilon\left(\begin{bmatrix} v_0 \\ \{v_j\} \end{bmatrix}, \begin{bmatrix} w_0 \\ \{w_j\} \end{bmatrix}\right)$$

denotes the energy inner product associated with the system (8) *and the condition* (NN) *is satisfied, then there exists* C (*independent of* ε) *such that*

$$\|v_0\|_1^2 + \varepsilon^2 \sum_{j=1}^n \|v_j\|_1^2 + \sum_{j=1}^n \|v_j\|_0^2 \leqq C B_\varepsilon\left(\begin{bmatrix} v_0 \\ \{v_j\} \end{bmatrix}, \begin{bmatrix} v_0 \\ \{v_j\} \end{bmatrix}\right)$$

for all $v_j \in \mathring{H}^1(\Omega)$, $0 \leqq j \leqq n$.

Proof. First let us prove that the system is positive definite, uniformly elliptic provided (N) holds. Let ξ_0, $\{\xi_j\}_{j=1}^n$ be $n+1$ vectors in $\mathbb{R}^n$. Then

$$\begin{aligned} &\left\langle \begin{bmatrix} \bar{\mathbf{A}} & \varepsilon\overline{\mathbf{A}\chi_j} \\ \varepsilon\overline{\mathbf{A}\chi_k} & \varepsilon^2\overline{\mathbf{A}\chi_k\chi_j} \end{bmatrix} \begin{bmatrix} \xi_0 \\ \{\xi_j\} \end{bmatrix}, \begin{bmatrix} \xi_0 \\ \{\xi_k\} \end{bmatrix} \right\rangle \\ &\qquad = \iint_{[0,1]^n} \left\langle \mathbf{A}(y)\left(\xi_0 + \varepsilon\sum_{j=1}^n \chi_j(y)\xi_j\right), \left(\xi_0 + \varepsilon\sum_{j=1}^n \chi_j(y)\xi_j\right)\right\rangle dy \end{aligned}$$

$$\geqq a_0 \iint_{[0,1]^n} \left(\xi_0+\varepsilon \sum_{j=1}^n \chi_j(y)\xi_j\right)^2 dy.$$

Now the condition (N) is clearly equivalent to the statement that 1, $\{\chi_j\}_{j=1}^n$ is a set of $n+1$ linearly independent functions. It therefore immediately follows that

$$\iint_{[0,1]^n} \left(\xi_0+\varepsilon \sum_{j=1}^n \chi_j(y)\xi_j\right)^2 dy \geqq C\left(\xi_0^2+\varepsilon^2 \sum_{j=1}^n \xi_j^2\right),$$

and this proves the first assertion.

Concerning the second part of the lemma it suffices to prove

$$\begin{aligned}&\|v_0\|_1^2+\varepsilon^2 \sum_{j=1}^n \|v_j\|_1^2+\sum_{j=1}^n \|v_j\|_0^2 \\ \text{(9)} \qquad &\leqq C \iint_{[0,1]^n} \int_\Omega \left[(\nabla_x+\varepsilon^{-1}\nabla_y)(v_0(x)+\varepsilon \sum_{j=1}^n \chi_j(y)v_j(x))\right]^2 dx\,dy,\end{aligned}$$

due to the fact that $\mathbf{A}(y)$ is uniformly positive definite.

With the assumption (NN), and since

$$(\nabla_x+\varepsilon^{-1}\nabla_y)\left(v_0(x)+\varepsilon \sum_{j=1}^n \chi_j(y)v_j(x)\right)=\nabla_x v_0(x)+\varepsilon \sum_{j=1}^n \chi_j(y)\nabla_x v_j(x)+\sum_{j=1}^n \nabla_y\chi_j(y)v_j(x),$$

(9) easily follows. □

Remark 3.1. It is also easy to see that, provided only (N) is satisfied, 0 is not an eigenvalue for (8), i.e. (8) has a unique solution. For the particular coercivity estimate, however, we need the stronger assumption (NN).

LEMMA 3.2. *Assume that the boundary of the domain Ω is locally given as the graph of a piecewise C^1 function (with bounded derivatives). Assume that the condition* (NN) *is satisfied and that $\chi_j \in W^{1,\infty}([0,1]^n)\forall 1\leqq j\leqq n$. Let u denote the solution to* (2) *and u_0^ε, $\{u_j^\varepsilon\}_{j=1}^n$ the solution to* (8).

There exists a constant C (independent of ε) such that

$$\|u_0^\varepsilon-u\|_1 \leqq C\varepsilon^{1/2}(\|u\|_2+\|u\|_{1,\infty}),$$

$$\left\|u_j^\varepsilon-\frac{\partial}{\partial x_j}u\right\|_0 \leqq C\varepsilon^{1/2}(\|u\|_2+\|u\|_{1,\infty}), \qquad 1\leqq j\leqq n,$$

and

$$\|\nabla_x u_j^\varepsilon\|_0 \leqq C\varepsilon^{-1/2}(\|u\|_2+\|u\|_{1,\infty}), \qquad 1\leqq j\leqq n,$$

provided $u\in H^2(\Omega)\cap W^{1,\infty}(\Omega)$.

Proof. The formal "limit" equations, for $\varepsilon=0$, corresponding to (8) are

$$\begin{aligned}&-\nabla_x\cdot(\bar{\mathbf{A}}\nabla_x u_0)-\sum_{j=1}^n \overline{\mathbf{A}\nabla_y\chi_j}\cdot\nabla_x u_j+\bar{b}u_0=f,\\ \text{(10)}\qquad & \\ &\overline{\mathbf{A}\nabla_y\chi_k}\cdot\nabla_x u_0-e_k\cdot\sum_{j=1}^n \overline{\mathbf{A}\nabla_y\chi_j}u_j=0, \qquad 1\leqq k\leqq n,\end{aligned}$$

with $u_0=0$ on $\partial\Omega$. (Here we used the identity $\overline{\mathbf{A}\nabla_y\chi_k\nabla_y\chi_j}=-e_k\cdot\overline{\mathbf{A}\nabla_y\chi_j}$.) Simple manipulations give that (10) is equivalent to

$$-\nabla_x\cdot(\mathscr{A}\nabla_x u_0)+\bar{b}u_0=f,$$

$$u_0=0\quad\text{on }\partial\Omega,$$

$$u_j=\frac{\partial}{\partial x_j}u_0,\qquad 1\leqq j\leqq n.$$

Let $m\in W^{1,\infty}(\mathbb{R}_+)$ be such that $m(0)=0$ and $m\equiv 1$ on $[1,\infty[$. Define

$$\tilde{u}_j^\varepsilon(x)=m\left(\frac{\text{dist}(x,\partial\Omega)}{\varepsilon}\right)u_j(x).$$

Because of Lemma 3.1 we know that

$$\text{(11)}\qquad \|u_0^\varepsilon-u_0\|_1+\varepsilon\sum_{j=1}^n\|u_j^\varepsilon-\tilde{u}_j^\varepsilon\|_1+\sum_{j=1}^n\|u_j^\varepsilon-\tilde{u}_j^\varepsilon\|_0\leqq C\sup_{\left[\begin{smallmatrix}v_0\\\{v_j\}\end{smallmatrix}\right]\in M_\varepsilon}B_\varepsilon\left(\begin{bmatrix}u_0^\varepsilon-u_0\\\{u_j^\varepsilon-\tilde{u}_j^\varepsilon\}\end{bmatrix},\begin{bmatrix}v_0\\\{v_j\}\end{bmatrix}\right),$$

where

$$M_\varepsilon=\left\{\begin{bmatrix}v_0\\\{v_j\}\end{bmatrix}\,\middle|\,v_j\in\mathring{H}^1(\Omega),\,0\leqq j\leqq n,\quad\text{and }\|v_0\|_1+\varepsilon\sum_{j=1}^n\|v_j\|_1+\sum_{j=1}^n\|v_j\|_0\leqq 1\right\}.$$

We easily get that

$$\begin{aligned}&B_\varepsilon\left(\begin{bmatrix}u_0^\varepsilon-u_0\\\{u_j^\varepsilon-\tilde{u}_j^\varepsilon\}\end{bmatrix},\begin{bmatrix}v_0\\\{v_j\}\end{bmatrix}\right)\\&\quad=\left\langle\begin{bmatrix}0\\\{\varepsilon\bar{\chi}_jf\}\end{bmatrix},\begin{bmatrix}v_0\\\{v_j\}\end{bmatrix}\right\rangle-\left\langle\begin{bmatrix}0&\varepsilon\overline{\mathbf{A}\chi_j}\\\varepsilon\overline{\mathbf{A}\chi_k}&\varepsilon^2\overline{\mathbf{A}\chi_k\chi_j}\end{bmatrix}\nabla_x\begin{bmatrix}u_0\\\{u_j\}\end{bmatrix},\nabla_x\begin{bmatrix}v_0\\\{v_j\}\end{bmatrix}\right\rangle\\&\qquad-\left\langle\begin{bmatrix}0&\varepsilon\overline{b\chi_j}\\\varepsilon\overline{b\chi_k}&\varepsilon^2\overline{b\chi_k\chi_j}+\varepsilon\overline{\mathbf{A}(\nabla_y\chi_k\chi_j-\chi_k\nabla_y\chi_j)}\cdot\nabla_x\end{bmatrix}\begin{bmatrix}u_0\\\{u_j\}\end{bmatrix},\begin{bmatrix}v_0\\\{v_j\}\end{bmatrix}\right\rangle\\&\qquad+B_\varepsilon\left(\begin{bmatrix}0\\\{u_j-\tilde{u}_j^\varepsilon\}\end{bmatrix},\begin{bmatrix}v_0\\\{v_j\}\end{bmatrix}\right).\end{aligned}$$

From this identity it follows that

$$\text{(12)}\qquad \sup_{\left[\begin{smallmatrix}v_0\\\{v_j\}\end{smallmatrix}\right]\in M_\varepsilon}B_\varepsilon\left(\begin{bmatrix}u_0^\varepsilon-u_0\\u_j^\varepsilon-\tilde{u}_j^\varepsilon\}\end{bmatrix},\begin{bmatrix}v_0\\\{v_j\}\end{bmatrix}\right)\leqq C\varepsilon^{1/2}(\|u_0\|_2+\|u_0\|_{1,\infty}).$$

Since we also have

$$\varepsilon\sum_{j=1}^n\|\tilde{u}_j^\varepsilon\|_1\leqq C\varepsilon^{1/2}(\|u_0\|_2+\|u_0\|_{1,\infty})$$

and

$$\sum_{j=1}^n\left\|\frac{\partial}{\partial x_j}u_0-\tilde{u}_j^\varepsilon\right\|_0\leqq C\varepsilon^{1/2}\|u_0\|_{1,\infty},$$

(11) and (12) immediately lead to

$$\|u_0^\varepsilon-u_0\|_1+\varepsilon\sum_{i=1}^n\|u_i^\varepsilon\|_1+\sum_{i=1}^n\left\|u_i^\varepsilon-\frac{\partial}{\partial x_i}u_0\right\|_0\leqq C\varepsilon^{1/2}(\|u_0\|_2+\|u_0\|_{1,\infty}).$$

We note that $u = u_0$, and hence this is exactly the desired result. $\square$

Remark 3.2. If $\partial\Omega$ is sufficiently smooth, then we have the following two regularity results:

$$\exists C \quad \text{such that } \|u\|_2 \leqq C\|f\|_0$$

and

$$\forall q > n, \quad \exists C_q \quad \text{such that } \|u\|_{1,\infty} \leqq C_q\|f\|_{0,q}.$$

For a proof of the last one see, e.g., [9]. This combined with Lemma 3.2 implies that for all $q > \max(n, 2)$ there exists C_q such that

$$\begin{aligned} &\|u_0^\varepsilon - u\|_1 \leqq C_q \varepsilon^{1/2} \|f\|_{0,q}, \\ (13) \qquad &\left\|u_j^\varepsilon - \frac{\partial}{\partial x_j} u\right\|_0 \leqq C_q \varepsilon^{1/2} \|f\|_{0,q}, \qquad 1 \leqq j \leqq n, \\ &\|\nabla_x u_j^\varepsilon\|_0 \leqq C_q \varepsilon^{-1/2} \|f\|_{0,q}, \qquad 1 \leqq j \leqq n. \end{aligned}$$

4. The approximation results for smooth domains. We are now in a position to prove that this alternate procedure for homogenization gives results that are close to u^ε in energy as $\varepsilon \to 0$. Whenever we require that the domain Ω be "sufficiently smooth" this means that we assume as much smoothness as needed for the regularity results of Remark 3.2 to hold (and at least a piecewise C^1 boundary).

THEOREM 4.1. *Assume that the condition* (NN) *is satisfied and that* $\chi_j \in W^{1,\infty}([0, 1]^n) \forall 1 \leqq j \leqq n$. *Let* u_0^ε, $\{u_j^\varepsilon\}_{j=1}^n$ *be the solution to* (8). *Provided* Ω *is sufficiently smooth, then for all* $q > \max(n, 2)$ *there exists* C_q *(independent of* ε*) such that*

$$\left\|u^\varepsilon(x) - \left(u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) u_j^\varepsilon(x)\right)\right\|_1 \leqq C_q \varepsilon^{1/2} \|f\|_{0,q}.$$

Proof. Using the same technique as in the proof of [3, Thm. 1.5.1] one easily gets that

$$\left\|u^\varepsilon(x) - \left(u(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) \frac{\partial}{\partial x_j} u(x)\right)\right\|_1 \leqq C \varepsilon^{1/2} (\|u\|_2 + \|u\|_{1,\infty}) \leqq C_q \varepsilon^{1/2} \|f\|_{0,q},$$

where u is the solution to (2). Because of the inequalities of Remark 3.2 it follows that

$$\left\|u(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) \frac{\partial}{\partial x_j} u(x) - \left(u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) u_j^\varepsilon(x)\right)\right\|_1 \leqq C_q \varepsilon^{1/2} \|f\|_{0,q}.$$

For any $q > \max(n, 2)$ we therefore obtain the estimate

$$\left\|u^\varepsilon(x) - \left(u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) u_j^\varepsilon(x)\right)\right\|_1 \leqq C_q \varepsilon^{1/2} \|f\|_{0,q},$$

which is exactly as desired. $\square$

Theorem 4.1 ensures that the procedure developed for homogenization here is as powerful as standard homogenization and formation of the sum

$$u(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) m^\varepsilon(x) \frac{\partial}{\partial x_j} u(x).$$

But, more important, we are also in a position to show that the alternate procedure guarantees strong H^1-convergence in cases where this sum is not well defined as an H^1-function, due to lack of regularity of f.

THEOREM 4.2. *Assume that the condition* (NN) *is satisfied and that* $\chi_j \in W^{1,\infty}([0,1]^n)$, *for all* $1 \leqq j \leqq n$. *Let* u_0^ε, $\{u_j^\varepsilon\}_{j=1}^n$ *be the solution to* (8), *with* $f \in H^s(\Omega)$ *for some* $-1 < s$. *Provided* Ω *is sufficiently smooth we then have*

$$u^\varepsilon(x) - \left(u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) u_j^\varepsilon(x)\right) \to 0 \quad \text{as } \varepsilon \to 0,$$

strongly in $\mathring{H}^1(\Omega)$.

Proof. Both $u^\varepsilon(x)$ and $u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j(x/\varepsilon) u_j^\varepsilon(x)$ are uniformly bounded in $H^1(\Omega)$ provided only $f \in H^{-1}(\Omega)$, i.e.

$$\|u^\varepsilon(x)\|_1 \leqq C\|f\|_{-1}$$

and

$$\left\|u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) u_j^\varepsilon(x)\right\|_1 \leqq C\|f\|_{-1}.$$

We therefore have

$$\left\|u^\varepsilon(x) - \left(u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) u_j^\varepsilon(x)\right)\right\|_1 \leqq C\|f\|_{-1}. \tag{14}$$

From the previous theorem we also know that

$$\left\|u^\varepsilon(x) - \left(u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) u_j^\varepsilon(x)\right)\right\|_1 \leqq C_q \varepsilon^{1/2}\|f\|_{0,q}, \tag{15}$$

for any $q > \max(n, 2)$.

Applying interpolation by the socalled K-method (cf. [4]) to (14) and (15) we obtain for $0 < \theta < 1$ and $q > \max(n, 2)$ that

$$\left\|u^\varepsilon(x) - \left(u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j\left(\frac{x}{\varepsilon}\right) u_j^\varepsilon(x)\right)\right\|_1 \leqq C_q \varepsilon^{\theta/2} |||f|||_\theta \tag{16}$$

where $|||\cdot|||_\theta$ is the norm on $(H^{-1}(\Omega), L^q(\Omega))_{\theta,\infty}$.

The Sobolev embedding theorem says that

$$H^t(\Omega) \subset L^q(\Omega)$$

for $t = n/2 - n/q$ and $2 \leqq q < \infty$.

Let us in the following assume that $n \geqq 2$. The case $n = 1$ may be handled by a similar argument. By taking $q = n + \delta$ and δ sufficiently small we thus obtain

$$H^{n/2-1+\sigma}(\Omega) \subset L^q(\Omega) \tag{17}$$

for a given but arbitrarily small $\sigma > 0$. The inclusion (17) immediately implies that

$$H^{-1+\theta(n/2+\sigma)}(\Omega) \subset (H^{-1}(\Omega), L^q(\Omega))_{\theta,\infty}. \tag{18}$$

Here we have used the well known fact (cf. [4], [11]) that $H^{-1+\theta(n/2+\sigma)}(\Omega) \subset (H^{-1}(\Omega), H^{n/2-1+\sigma}(\Omega))_{\theta,\infty}$.

Now choose $0 < \theta < 1$ such that

$$-1 + \theta\left(\left(\frac{n}{2}\right) + \sigma\right) < s.$$

(This can be done since $-1 < s$), then by a combination of (16) and (18) we get the desired result. $\square$

Remark 4.1. $f \in H^{-1}(\Omega)$ is necessary and sufficient in order for (8) to have a solution with components in $H^1(\Omega)$. If we only know about f that $f \in H^s(\Omega)$ for some $-1 < s < 0$, then standard homogenization and formation of the sum (5) will not give any H^1-convergence, but the alternate procedure does.

Remark 4.2. We actually proved a little more than just strong convergence in the previous proof. Specifically we got that $f \in H^s(\Omega)$, $-1 < s < n/2 - 1$ implies a convergence rate for $u^\varepsilon(x) - (u_0^\varepsilon(x) + \varepsilon \sum_{j=1}^n \chi_j(x/\varepsilon) u_j^\varepsilon(x))$ of order ε^ν, where ν is smaller than but arbitrarily close to $(1+s)/n$. This particular statement is only valid for $n \geqq 2$, but a similar result holds in the case $n = 1$.

5. An approximation result for domains with corners. Throughout this section we assume that $\Omega \subseteq \mathbb{R}^2$ has only a *piecewise* smooth boundary (for simplicity let us take this to mean piecewise C^∞) with a finite number of corners C_i and corresponding angles $0 < \alpha_i < 2\pi$, $\alpha_i \neq \pi$. As our main result we shall prove that the contents of Theorem 4.2 remain valid also in this case. The proof of this is based on Lemma 3.2 and interpolation. The interpolation result we need is one between spaces that satisfy certain boundary conditions, but because of the corners of the domain Ω we cannot rely on the results of this type proven e.g. by Lions and Magenes (cf. [11]). Instead we use a slight variation of a result due to Babuska, Kellogg and Pitkäranta. Since we are here only interested in the case of ordinary Sobolev spaces, which somewhat simplifies the proof, we shall give a sketch following the main ideas in [2].

LEMMA 5.1. *Let $\Omega \subset \mathbb{R}^2$ be piecewise smooth with a finite number of corners. The following result then holds for any $0 < \theta < 1$ and $1 \leqq t$:*

$$(\mathring{H}^1(\Omega), H^t(\Omega) \cap \mathring{H}^1(\Omega))_{\theta,\infty} = (H^1(\Omega), H^t(\Omega))_{\theta,\infty} \cap \mathring{H}^1(\Omega).$$

Proof. The inclusion $(\mathring{H}^1(\Omega), H^t(\Omega) \cap \mathring{H}^1(\Omega))_{\theta,\infty} \subset (H^1(\Omega), H^t(\Omega))_{\theta,\infty} \cap \mathring{H}^1(\Omega)$ is trivial. To prove the converse it obviously suffices, by means of a localization argument, to consider a domain Ω with only one corner of angle $0 < \alpha < 2\pi$, $\alpha \neq \pi$, the sides of which are locally linear.

Let us first take the case $\alpha < \pi$. By a linear transformation it is possible to map Ω onto a new domain $\tilde{\Omega}$ with a corner of angle $\tilde{\alpha} < \min(\pi/(t-1), \pi)$. This induces a mapping L of functions on $\tilde{\Omega}$ to functions on Ω.

We define a bounded operator P: $H^1(\tilde{\Omega}) \to \mathring{H}^1(\tilde{\Omega})$ the following way,

$$-\Delta Pu + Pu = -\Delta u + u \quad \text{in } \tilde{\Omega},$$
$$Pu = 0 \quad \text{on } \partial\tilde{\Omega}.$$

Because of the size of the angle $\tilde{\alpha}$ it is clear (cf. [6]) that P is also a bounded operator

$$H^t(\tilde{\Omega}) \to H^t(\tilde{\Omega}) \cap \mathring{H}^1(\tilde{\Omega}).$$

Define

$$P_\Omega := L \circ P \circ L^{-1}$$

as an operator $H^1(\Omega) \to \mathring{H}^1(\Omega)$. It is clear that P_Ω also takes $H^t(\Omega)$ continuously into $H^t(\Omega) \cap \mathring{H}^1(\Omega)$. By interpolation P_Ω is therefore a bounded operator,

$$P_\Omega: (H^1(\Omega), H^t(\Omega))_{\theta,\infty} \to (\mathring{H}^1(\Omega), H^t(\Omega) \cap \mathring{H}^1(\Omega))_{\theta,\infty}.$$

On $\mathring{H}^1(\Omega)$ P_Ω is by definition the identity, which proves that

$$(H^1(\Omega), H^t(\Omega))_{\theta,\infty} \cap \mathring{H}^1(\Omega) \subseteq (\mathring{H}^1(\Omega), H^t(\Omega) \cap \mathring{H}^1(\Omega))_{\theta,\infty}.$$

The case $\pi<\alpha<2\pi$ is a bit more difficult because we can no longer map the domain Ω by a linear transformation onto one with a convex corner. Let B be a ball centered at the corner and containing Ω. By E_Ω we denote the Stein extension (cf. [15, Chap. 6, § 3]) of functions on Ω to functions on B vanishing at ∂B. In the following Ω^c denotes the complement of Ω with respect to B, i.e. $\Omega^c=B\backslash\Omega$, and E_{Ω^c} the Stein extension of functions on Ω^c to all of B. Define

$$P_\Omega := E_{\Omega^c}\circ P_{\Omega^c}\circ E_\Omega+(I-E_{\Omega^c}\circ E_\Omega).$$

P_{Ω^c} is as defined before, since now the angle of Ω^c is $<\pi$. It is easily seen that P_Ω is a bounded operator

$$(H^1(\Omega), H^t(\Omega))_{\theta,\infty}\to(\mathring{H}^1(\Omega), H^t(\Omega)\cap\mathring{H}^1(\Omega))_{\theta,\infty}$$

such that $P_\Omega=I$ (the identity) on $\mathring{H}^1(\Omega)$. This concludes the proof of the lemma. □

With this lemma we are now in a position to prove the main result of this section.

THEOREM 5.1. *Let $\Omega\subset\mathbb{R}^2$ be piecewise smooth with a finite number of corners. Assume that condition* (NN) *is satisfied, and that $\chi_j\in W^{1,\infty}([0,1]^2)$, $j=1,2$. Let u_0^ε, $\{u_j^\varepsilon\}_{j=1}^2$ be the solution to* (8), *with $f\in H^s(\Omega)$ for some $-1<s$. We then have that*

$$u^\varepsilon(x)-\left(u_0^\varepsilon(x)+\varepsilon\sum_{j=1}^2\chi_j\left(\frac{x}{\varepsilon}\right)u_j^\varepsilon(x)\right)\to 0\quad\text{as } \varepsilon\to 0,$$

strongly in $\mathring{H}^1(\Omega)$.

Proof. It follows directly from Lemma 3.2 and the proof of Theorem 4.1 that there exists C such that

$$\left\|u^\varepsilon(x)-\left(u_0^\varepsilon(x)+\varepsilon\sum_{j=1}^2\chi_j\left(\frac{x}{\varepsilon}\right)u_j^\varepsilon(x)\right)\right\|_1\leqq C\varepsilon^{1/2}(\|u\|_2+\|u\|_{1,\infty})\tag{19}$$

provided $u\in H^2(\Omega)$, $(\partial/\partial x_j)u\in L^\infty(\Omega)$, and the boundary is piecewise C^1 (u here denotes the solution to (2)). Since $n=2$ we know from the Sobolev embedding theorem that

$$H^t(\Omega)\subset W^{1,\infty}(\Omega)\quad\text{for any } t>2,$$

i.e., for all $t>2$ there exists C_t such that

$$\left\|u^\varepsilon(x)-\left(u_0^\varepsilon(x)+\varepsilon\sum_{j=1}^2\chi_j\left(\frac{x}{\varepsilon}\right)u_j^\varepsilon(x)\right)\right\|_1\leqq C_t\varepsilon^{1/2}\|u\|_t,\tag{20}$$

provided $u\in H^t(\Omega)$.

We also have the estimate

$$\left\|u^\varepsilon(x)-\left(u_0^\varepsilon(x)+\varepsilon\sum_{j=1}^2\chi_j\left(\frac{x}{\varepsilon}\right)u_j^\varepsilon(x)\right)\right\|_1\leqq C\|f\|_{-1}\leqq C\|u\|_1.\tag{21}$$

Through (2) f is given as a function of $u\in\mathring{H}^1(\Omega)$. Since u^ε and u_0^ε, $\{u_j^\varepsilon\}_{j=1}^2$ are defined in terms of f these can also be regarded as functions of $u\in\mathring{H}^1(\Omega)$. With this interpretation in mind we can interpolate between (20) and (21) and obtain for any $0<\theta<1$ that

$$\left\|u^\varepsilon(x)-\left(u_0^\varepsilon(x)+\varepsilon\sum_{j=1}^2\chi_j\left(\frac{x}{\varepsilon}\right)u_j^\varepsilon(x)\right)\right\|_1\leqq C_t\varepsilon^{\theta/2}|||u|||_\theta,$$

provided $u\in(\mathring{H}^1(\Omega), H^t(\Omega)\cap\mathring{H}^1(\Omega))_{\theta,\infty}$. ($|||\cdot|||_\theta$ denotes a norm on this space.)

Using Lemma 5.1 we therefore get

$$\left\| u^{\varepsilon}(x) - \left(u_0^{\varepsilon}(x) + \varepsilon \sum_{j=1}^{2} \chi_j\left(\frac{x}{\varepsilon}\right) u_j^{\varepsilon}(x) \right) \right\|_1 \leqq C_t \varepsilon^{\theta/2} |||u|||_{\theta}, \tag{22}$$

provided $u \in (H^1(\Omega), H^t(\Omega))_{\theta,\infty}$ and $|||\cdot|||_\theta$ denotes a norm on this space (remember that u by definition is always in $\mathring{H}^1(\Omega)$).

It is well known that

$$H^{1+(t-1)\theta}(\Omega) \subset (H^1(\Omega), H^t(\Omega))_{\theta,\infty}.$$

A combination of this and (22) gives

$$\left\| u^{\varepsilon}(x) - \left(u_0^{\varepsilon}(x) + \varepsilon \sum_{j=1}^{2} \chi_j\left(\frac{x}{\varepsilon}\right) u_j^{\varepsilon}(x) \right) \right\|_1 \leqq C_{r,\nu} \varepsilon^{\nu} \|u\|_r, \tag{23}$$

where ν can take any value $<\min((r-1)/2, 1/2)$, provided $u \in H^r(\Omega)$. From, e.g., [6] we know that if $f \in H^s(\Omega)$ for some $-1<s$ then $u \in H^r(\Omega)$ for some $1<r$ (how large r can be taken depends both on s and the size of maximal angle). Formula (23) thus immediately guarantees convergence. □

Remark 5.1. From the proof of Theorem 5.1 we also gain some information about approximation rates.

It is always possible to perform a linear change of variables such that (2) transforms into

$$\begin{aligned} -\Delta u + \bar{b} u &= f \quad \text{in } \tilde{\Omega}, \\ u &= 0 \quad \text{on } \partial\tilde{\Omega}, \end{aligned}$$

If $f \in L^2(\Omega)$ and $\tilde{\alpha}$ denotes the size of the maximal corner in $\tilde{\Omega}$ then we have (cf. [6]) that $u \in H^r(\Omega)$ for any $r < \min((\pi/\tilde{\alpha})+1, 2)$. Based on the previous proof, this implies that

$$\left\| u^{\varepsilon}(x) - \left(u_0^{\varepsilon}(x) + \varepsilon \sum_{j=1}^{2} \chi_j(x) u_j^{\varepsilon}(x) \right) \right\|_1$$

converges like ε^{ν} where ν is smaller than, but arbitrarily close to $\min(\pi/2\tilde{\alpha}, 1/2)$.

6. The "nearly" degenerate case. We now study more closely what happens to this procedure for homogenization when $\mathbf{A}$ gets very near a point of degeneracy.

We limit our analysis to the case $n=1$, i.e., $\mathbf{A}(y) = a(y)$, and shall always implicitly assume that f is sufficiently smooth for our arguments to be valid. Our main result, Theorem 6.1, shows that although the system (8) becomes singular, the estimate $\|u^{\varepsilon}(x) - (u_0^{\varepsilon}(x) + \varepsilon\chi(x/\varepsilon)u_1^{\varepsilon}(x))\|_1 \leqq C\varepsilon^{1/2}$ holds uniformly and independently of how close we get to a point of degeneracy.

In the following the notation $'$ shall be used in place of d/dy. The absolute terms involving b have no bearing on the problem addressed here, so for convenience we set $b \equiv 0$. The domain Ω is taken to be $[-1, 1]$.

Due to the well-known formulas $(\overline{a^{-1}})^{-1} = \bar{a} + \overline{a\chi'}$ and $-\overline{a\chi'} = \overline{a(\chi')^2}$, equations (8) for u_0^{ε} and u_1^{ε} take the form

$$\begin{aligned} -(\overline{a^{-1}})^{-1}\left(\frac{d}{dx}\right)^2 u_0^{\varepsilon} + \overline{a(\chi')^2}\frac{d}{dx}\left(u_1^{\varepsilon} - \frac{d}{dx}u_0^{\varepsilon}\right) &= f, \\ -\varepsilon^2\overline{a\chi^2}\left(\frac{d}{dx}\right)^2 u_1^{\varepsilon} + \overline{a(\chi')^2}\left(u_1^{\varepsilon} - \frac{d}{dx}u_0^{\varepsilon}\right) &= \varepsilon\bar{\chi} f, \end{aligned} \tag{24}$$

Here the undetermined constant of χ has been selected such that $\overline{a\chi}=0$.

Multiply the first equation by $\varepsilon^2\overline{a^{-1}}\,\overline{a\chi^2}$, differentiate and subtract it from the second to obtain

$$(25)\qquad -\varepsilon^2 c\left(\frac{d}{dx}\right)^2\left(u_1^\varepsilon-\frac{d}{dx}u_0^\varepsilon\right)+\overline{a(\chi')^2}\left(u_1^\varepsilon-\frac{d}{dx}u_0^\varepsilon\right)=\varepsilon\bar{\chi}f-\varepsilon^2\overline{a^{-1}}\,\overline{a\chi^2}\frac{d}{dx}f,$$

with $c=\overline{a\chi^2}(1+\overline{a^{-1}}\,\overline{a(\chi')^2})$. Let us now estimate the boundary conditions for $u_1^\varepsilon-(d/dx)u_0^\varepsilon$. Integration of the first equation in (24) gives

$$(26)\qquad -\bar{a}\left(\frac{d}{dx}u_0^\varepsilon(1)-\frac{d}{dx}u_0^\varepsilon(-1)\right)=\int_{-1}^{1}f(x)\,dx.$$

Multiplication of the same equation by x and integration leads to

$$(27)\qquad -\bar{a}\left(\frac{d}{dx}u_0^\varepsilon(1)+\frac{d}{dx}u_0^\varepsilon(-1)\right)-\overline{a(\chi')^2}\int_{-1}^{1}u_1^\varepsilon(x)\,dx=\int_{-1}^{1}f(x)x\,dx.$$

LEMMA 6.1. *Assume that $0<a_0\leqq a(y)\leqq a_1<\infty$ and a is nonconstant. There exists C (independent of a and ε, but dependent on a_0 and a_1) such that*

$$\left|\frac{d}{dx}u_0^\varepsilon(x_0)\right|<C\|f\|_{0,1}$$

for $x_0=\pm1$ and all ε.

Proof. From (24) it follows through multiplication of the first equation by u_0^ε, the second by u_1^ε, addition and integration that

$$(28)\qquad \begin{aligned}&(\overline{a^{-1}})^{-1}\int_{-1}^{1}\left(\frac{d}{dx}u_0^\varepsilon\right)^2dx+\overline{a(\chi')^2}\int_{-1}^{1}\left(u_1^\varepsilon-\frac{d}{dx}u_0^\varepsilon\right)^2dx\\&+\varepsilon^2\overline{a\chi^2}\int_{-1}^{1}\left(\frac{d}{dx}u_1^\varepsilon\right)^2dx=\int_{-1}^{1}fu_0^\varepsilon\,dx+\varepsilon\bar{\chi}\int_{-1}^{1}fu_1^\varepsilon\,dx.\end{aligned}$$

Since $0<a_0\leqq a(y)$ it immediately follows that

$$|\bar{\chi}|\leqq C(\overline{a\chi^2})^{1/2},$$

and therefore due to (28) that

$$(29)\qquad (\overline{a(\chi')^2})^{1/2}\|u_1^\varepsilon\|_0\leqq C\|f\|_{-1}.$$

Here we have used the inequality

$$(30)\qquad \overline{a(\chi')^2}\leqq C,$$

which can be derived from the identity

$$a\chi'=-a+(\overline{a^{-1}})^{-1}$$

and the assumption $0<a_0\leqq a(y)\leqq a_1<\infty$. Based on (29) and (30) we immediately get

$$\overline{a(\chi')^2}\left|\int_{-1}^{1}u_1^\varepsilon(x)\,dx\right|\leqq C\|f\|_{-1}\leqq C\int_{-1}^{1}|f(x)|\,dx.$$

This inequality combined with (26) and (27) leads to the desired result. □

Since $u_1^\varepsilon(x_0)=0$ at $x_0=\pm1$ the conclusion in Lemma 6.1 may also be written

$$(31)\qquad \left|u_1^\varepsilon(x_0)-\frac{d}{dx}u_0^\varepsilon(x_0)\right|\leqq C\|f\|_{0,1}$$

for $x_0=\pm1$ and all ε.

LEMMA 6.2. *Assume that* $0 < a_0 \leqq a(y) \leqq a_1 < \infty$ *and* a *is nonconstant. There exist constants* α, c_ε, d_ε *and a function* R^ε *such that*

$$u_1^\varepsilon(x) - \frac{d}{dx} u_0^\varepsilon(x) = c_\varepsilon e^{-\alpha(x+1)/\varepsilon} + d_\varepsilon e^{-\alpha(1-x)/\varepsilon} + R^\varepsilon.$$

Specifically,

$$\alpha = [\overline{a(\chi')^2}[\overline{a\chi}^2(1 + \overline{a^{-1}} \cdot \overline{a(\chi')^2})]^{-1}]^{1/2},$$

c_ε *and* d_ε *are bounded independently of* a, ε *and*

$$\|R^\varepsilon\|_0 \leqq C\varepsilon, \qquad \left\|\frac{d}{dx} R^\varepsilon\right\|_0 \leqq C$$

for some C *independent of* a, ε *(but dependent on* a_0, a_1*).*

Proof. Dividing (25) by $\overline{a(\chi')^2}$ we get the equation

$$-\left(\frac{\varepsilon}{\alpha}\right)^2\left(\frac{d}{dx}\right)^2\left(u_1^\varepsilon - \frac{d}{dx}u_0^\varepsilon\right) + \left(u_1^\varepsilon - \frac{d}{dx}u_0^\varepsilon\right) = \frac{\varepsilon}{\alpha}\beta_1 f - \left(\frac{\varepsilon}{\alpha}\right)^2 \beta_2 \frac{d}{dx} f,$$

where

$$\beta_1 = \bar{\chi}\alpha(\overline{a(\chi')^2})^{-1} \quad \text{and} \quad \beta_2 = \overline{a^{-1}}\,\overline{a\chi}^2\alpha^2(\overline{a(\chi')^2})^{-1}.$$

It is easy to see that $|\beta_2| \leqq C$ independently of a. The similar result for β_1 is not quite as obvious. We have that

$$|\bar{\chi}| = \frac{1}{|C|}|\overline{\chi(y)(C - a(y))}| \leqq \frac{1}{|C|}\|\chi\|_0\|a(y) - C\|_0,$$

for any constant C. Since $0 < a_0 \leqq a(y)$ this leads to the existence of a constant D independent of a such that

$$|\bar{\chi}| \leqq D(\overline{a\chi^2})^{1/2} \inf_C \|a(y) - C\|_0, \tag{32}$$

where the infimum is taken over all constants C. Based on the identity $a\chi' = -a + (\overline{a^{-1}})^{-1}$ and the assumption $a(y) \leqq a_1 < \infty$ we conclude that there exists $D > 0$ (independent of a) with

$$\overline{a(\chi')^2} \geqq D \inf_C \|a(y) - C\|_0^2. \tag{33}$$

Formulas (32) and (33) now lead to the existence of a constant C (independent of a) such that $|\beta_1| \leqq C$.

By a very similar argument it is possible to prove that $\overline{a(\chi)^2} \cdot (\overline{a(\chi')^2})^{-1} \leqq C$ independently of a. This implies that α is bounded from below independently of a.

The fact that $|\beta_i| \leqq C$, $i = 1, 2$, and $0 < c \leqq \alpha$ independently of a now combine with (31) to give the result of this lemma. □

Based on Lemma 6.2 and the first equation of (24) we immediately get

LEMMA 6.3. *Assume that* $0 < a_0 \leqq a(y) \leqq a_1 < \infty$ *and* a *is nonconstant. Let* α, c_ε *and* d_ε *be the same constants as in the previous lemma. Then*

$$\frac{d}{dx} u_1^\varepsilon(x) = \overline{a^{-1}} \cdot \bar{a}\frac{\alpha}{\varepsilon}(-c_\varepsilon e^{-\alpha(x+1)/\varepsilon} + d_\varepsilon e^{-\alpha(1-x)/\varepsilon}) + \tilde{R}^\varepsilon,$$

where $\tilde{R}^\varepsilon$ *satisfies* $\|\tilde{R}^\varepsilon\|_0 \leqq C$ *for some constant independent of* a *and* ε *(but dependent on* a_0 *and* a_1*).*

We are now in a position to prove the main result of this section. It shows that although the alternate procedure for homogenization at first glance may seem singular at points $a(y)=$ constant, this is not really so.

THEOREM 6.1. *Assume that $0<a_0\leqq a(y)\leqq a_1<\infty$, and that f is sufficiently smooth. Let χ be adjusted so that $\overline{a\chi}=0$. Let u_0^ε, u_1^ε be the solution to* (24) *with homogeneous Dirichlet boundary conditions and u^ε the solution to the present version of* (1). *There exists a constant C independent of a and ε (but dependent on a_0, a_1 and f) such that*

$$\left\| u^\varepsilon(x)-\left(u_0^\varepsilon(x)+\varepsilon\chi\left(\frac{x}{\varepsilon}\right)u_1^\varepsilon(x)\right)\right\|_1\leqq C\varepsilon^{1/2}.$$

Proof. It is not difficult to see that

$$\begin{aligned}\int_{-1}^{1} a\left(\frac{x}{\varepsilon}\right)\frac{d}{dx}\left(u^\varepsilon(x)-u_0^\varepsilon(x)-\varepsilon\chi\left(\frac{x}{\varepsilon}\right)u_1^\varepsilon(x)\right)\frac{d}{dx}v(x)\,dx\\ =\int_{-1}^{1}\left(\bar a-a\left(\frac{x}{\varepsilon}\right)\right)\left(\frac{d}{dx}u_0^\varepsilon(x)-u_1^\varepsilon(x)\right)\frac{d}{dx}v(x)\,dx\\ -\varepsilon\int_{-1}^{1}a\left(\frac{x}{\varepsilon}\right)\chi\left(\frac{x}{\varepsilon}\right)\frac{d}{dx}u_1^\varepsilon(x)\frac{d}{dx}v(x)\,dx\quad \forall v\in \mathring{H}^1([-1,1]).\end{aligned}\tag{34}$$

By use of Lemma 6.2 the first term on the right-hand side of (34) is easily bounded by

$$C\varepsilon^{1/2}\|v\|_1,\quad \text{with } C \text{ independent of } a \text{ and } \varepsilon.$$

From [7, p. 194], we have the inequality

$$\|\phi\|_{0,\infty}\leqq C\|\phi\|_0^{1/2}\|\phi\|_1^{1/2}$$

for any $\phi\in H^1([0,1])$.

This immediately gives that

$$\|a\chi\|_{0,\infty}\leqq C\|\chi\|_0^{1/2}\|\chi\|_1^{1/2}\leqq C(\overline{a\chi^2})^{1/4},$$

with C independent of a (but of course dependent on a_0 and a_1).

Combining this estimate with the result of Lemma 6.3 we obtain that the second term in the right-hand side of (34) can be bounded by

$$C(\overline{a\chi^2})^{1/4}(\alpha\varepsilon)^{1/2}\|v\|_1\leqq C\varepsilon^{1/2}\|v\|_1$$

with C independent of a, ε. This finishes the proof of the theorem. □

Remark 6.1. Consider the case that *a is* a constant. The term u_0^ε is still uniquely determined by (24) (actually u_0^ε is also the exact solution to (1)). The term u_1^ε is completely undetermined, but due to the fact that $\chi=0$ this does not matter to the sum $u_0^\varepsilon(x)+\varepsilon\chi(x/\varepsilon)u_1^\varepsilon(x)$. As it should do, the alternate procedure thus produces the true solution.

7. A boundary-layer analysis for the cutoff. In this and the following section we shall justify that the boundary-layers associated with the singularly perturbed system (8) are identical to the optimal cutoff (near $\partial\Omega$) of the first two terms

$$u(x)+\varepsilon\sum_{j=1}^{n}\chi_j\left(\frac{x}{\varepsilon}\right)\frac{\partial}{\partial x_j}u(x)$$

of the formal asymptotic expansion. The optimality is defined in terms of energy.

This is the basis for our claim that not only does the system (8) preserve homogeneous Dirichlet boundary conditions, it automatically provides a solution which is at least as good as the optimal one can obtain based on $u(x)$, $\nabla_x u(x)$ and a cutoff. The optimality of boundary-layers also provides heuristic insight as to why this method is very good with problems with nonsmooth data.

The analysis presented here could be carried out in any dimension n (along the normal direction to the boundary $\partial\Omega$), but for simplicity of exposition we only consider the case $n=1$, $\Omega=[-1,1]$.

It is quite clear that the absolute terms involving b in the equations (1), (2) and (8) do not influence the optimal cutoff, nor the boundary layers of (8), so for convenience we set $b=0$. Since (8) is already an averaged system it is very natural that we have to define the optimal cutoff for $u(x)+\varepsilon\chi(x/\varepsilon)(d/dx)u(x)$ in an averaged sense in order to compare. To this end let us assume that the coefficient $\mathbf{A}$ in (1) is not only a function of x/ε but also of a random grid location. To be specific the coefficient is

$$a\left(\frac{x}{\varepsilon}+y_0(\mu)\right),$$

where $0\leqq y_0\leqq 1$ is a uniformly distributed random variable on the probability space $(\mathcal{P}, d\mu)$. (Remember that the matrix $\mathbf{A}$ is just a single function a in this case.)

We now proceed to find the equations that characterize the optimal cutoff.

Let $m_i(\cdot)$, $i=0,1$, be two $W^{1,\infty}$-functions on $[0,\infty[$ with the following properties:

$$m_1(0)=0, \qquad m_i(t)=1 \quad \forall t\in[1,\infty[, \quad i=0,1.$$

We shall refer to (m_0, m_1) as a set of cutoff functions, although only m_1 is required to be 0 at $t=0$.

Let u denote the solution to (2), i.e.

$$-\frac{d}{dx}\left(\mathcal{A}\left(\frac{d}{dx}\right)\right)u(x)=f(x) \quad \text{in } [-1,1],$$

$$u(x)=0 \quad \text{for } x=\pm 1,$$

and $\mathcal{A}=(\overline{a^{-1}})^{-1}$.

Define $s^\varepsilon(x,\mu)$ to be

$$s^\varepsilon(x,\mu)=m_0\left(\frac{d}{\varepsilon}\right)u(x)+\varepsilon\chi\left(\frac{x}{\varepsilon}+y_0(\mu)\right)m_1\left(\frac{d}{\varepsilon}\right)\frac{d}{dx}u(x),$$

with $d=\min\{1+x, 1-x\}$.

Note that we allow for adjustment of both u and $(d/dx)\,u$ near the boundary.

As before, $u^\varepsilon(x,\mu)$ denotes the solution to (1). (Remember that it depends on $\mu\in\mathcal{P}$ through the random shift y_0.)

We shall first derive an expression for the average energy of $u^\varepsilon(x,\mu)-s^\varepsilon(x,\mu)$ on $[-1,0]$.

LEMMA 7.1. *Assume that $f\in L^2([-1,1])$ and that χ has been selected so that $\overline{a\chi}=0$. If $E(\mu, m_0, m_1)$ denotes*

$$\int_{-1}^{0} a\left(\frac{x}{\varepsilon}+y_0(\mu)\right)\left[\frac{d}{dx}(u^\varepsilon-s^\varepsilon)\right]^2 dx,$$

then

$$\int_{\mathcal{P}} E(\mu, m_0, m_1)\, d\mu$$

$$= \left[\frac{d}{dx}u(-1)\right]^2 \cdot \left[\varepsilon^2 \bar{a} \int_0^1 \left(\frac{d}{dt}\left(\frac{t}{\varepsilon} M_0\left(\frac{t}{\varepsilon}\right)\right)\right)^2 dt + \varepsilon^2 \overline{a\chi^2} \int_0^1 \left(\frac{d}{dt} M_1\left(\frac{t}{\varepsilon}\right)\right)^2 dt \right.$$

$$+ 2\varepsilon \overline{a\chi'} \int_0^1 \frac{d}{dt}\left(\frac{t}{\varepsilon} M_0\left(\frac{t}{\varepsilon}\right)\right) M_1\left(\frac{t}{\varepsilon}\right) dt$$

$$\left. + \overline{a(\chi')^2} \int_0^1 \left(M_1\left(\frac{t}{\varepsilon}\right)\right)^2 dt - \varepsilon \overline{a\chi\chi'}\right] + O(\varepsilon^{3/2}),$$

where $M_i = 1 - m_i$, $i = 0, 1$.

Proof. From the definition of u^ε and s^ε it follows that

$$\int_{-1}^1 a\left(\frac{x}{\varepsilon} + y_0(\mu)\right) \frac{d}{dx}(u^\varepsilon - s^\varepsilon) \frac{d}{dx} v\, dx$$

$$= \int_{-1}^1 a\left(\frac{x}{\varepsilon} + y_0(\mu)\right) M_0\left(\frac{d}{\varepsilon}\right) \frac{d}{dx} u\left(\frac{d}{dx}\right) v\, dx$$

$$+ \int_{-1}^1 a\left(\frac{x}{\varepsilon} + y_0(\mu)\right)(\chi')\left(\frac{x}{\varepsilon} + y_0(\mu)\right) M_1\left(\frac{d}{\varepsilon}\right) \frac{d}{dx} u \frac{d}{dx} v\, dx \tag{35}$$

$$+ \int_{-1}^1 a\left(\frac{x}{\varepsilon} + y_0(\mu)\right) \frac{d}{dx} M_0\left(\frac{d}{\varepsilon}\right) u \frac{d}{dx} v\, dx$$

$$+ \varepsilon \int_{-1}^1 a\left(\frac{x}{\varepsilon} + y_0(\mu)\right) \chi\left(\frac{x}{\varepsilon} + y_0(\mu)\right) \frac{d}{dx} M_1\left(\frac{d}{\varepsilon}\right) \frac{d}{dx} u \frac{d}{dx} v\, dx + O(\varepsilon)\|v\|_1$$

for any $v \in \mathring{H}^1([-1, 1])$.

Based on (35) we conclude that

$$a\left(\frac{x}{\varepsilon} + y_0(\mu)\right) \frac{d}{dx}(u^\varepsilon - s^\varepsilon)$$

$$= a\left(\frac{x}{\varepsilon} + y_0(\mu)\right) M_0\left(\frac{d}{\varepsilon}\right) \frac{d}{dx} u$$

$$+ a\left(\frac{x}{\varepsilon} + y_0(\mu)\right)(\chi')\left(\frac{x}{\varepsilon} + y_0(\mu)\right) M_1\left(\frac{d}{\varepsilon}\right) \frac{d}{dx} u + a\left(\frac{x}{\varepsilon} + y_0(\mu)\right) \frac{d}{dx} M_0\left(\frac{d}{\varepsilon}\right) u \tag{36}$$

$$+ \varepsilon a\left(\frac{x}{\varepsilon} + y_0(\mu)\right) \chi\left(\frac{x}{\varepsilon} + y_0(\mu)\right) \frac{d}{dx} M_1\left(\frac{d}{\varepsilon}\right) \frac{d}{dx} u + O_{L^2}(\varepsilon).$$

The term $O_{L^2}(\varepsilon)$ denotes a quantity with L^2-norm bounded by $C\varepsilon$. For the derivation of this identity we have used the fact that $\int_0^1 M_i(t/\varepsilon)\, dt \leqq C\varepsilon$, $i = 0, 1$, and $\int_0^1 (d/dt) M_1(t/\varepsilon)\, dt \leqq C$ to bound the constant of integration undetermined by (35). Formula (36) immediately leads to

$$
\begin{aligned}
a\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)\frac{d}{dx}(u^\varepsilon-s^\varepsilon)
&=\frac{d}{dx}u(-1)\Big[a\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)M_0\Big(\frac{1+x}{\varepsilon}\Big)\\
&\quad+a\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)(\chi')\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)M_1\Big(\frac{1+x}{\varepsilon}\Big)\\
&\quad+\varepsilon a\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)\frac{d}{dx}M_0\Big(\frac{1+x}{\varepsilon}\Big)\Big(\frac{1+x}{\varepsilon}\Big)\\
&\quad+\varepsilon a\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)\chi\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)\frac{d}{dx}M_1\Big(\frac{1+x}{\varepsilon}\Big)\Big]+O_{L^2}(\varepsilon)\quad\text{near } x=-1,
\end{aligned}
$$

and

$$a\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)\frac{d}{dx}(u^\varepsilon-s^\varepsilon)=O_{L^2}(\varepsilon)$$

in the interior of $[-1, 1]$ (a distance ε away from the boundary).

We therefore get

$$
\begin{aligned}
&\int_{\mathcal{P}}\Big(\int_{-1}^{0}a\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)\Big[\frac{d}{dx}(u^\varepsilon-s^\varepsilon)\Big]^2dx\Big)\,d\mu\\
&=\Big[\frac{d}{dx}u(-1)\Big]^2\cdot\Big[\varepsilon^2\bar a\int_0^1\Big(\frac{d}{dt}\Big(\frac{t}{\varepsilon}M_0\Big(\frac{t}{\varepsilon}\Big)\Big)\Big)^2dt+\varepsilon^2\overline{a\chi^2}\int_0^1\Big(\frac{d}{dt}M_1\Big(\frac{t}{\varepsilon}\Big)\Big)^2dt\\
&\qquad+2\varepsilon\overline{a\chi'}\int_0^1\frac{d}{dt}\Big(\frac{t}{\varepsilon}M_0\Big(\frac{t}{\varepsilon}\Big)\Big)M_1\Big(\frac{t}{\varepsilon}\Big)\,dt\\
&\qquad+2\varepsilon\overline{a\chi\chi'}\int_0^1\frac{d}{dt}M_1\Big(\frac{t}{\varepsilon}\Big)M_1\Big(\frac{t}{\varepsilon}\Big)\,dt\\
&\qquad+\overline{a(\chi')^2}\int_0^1\Big(M_1\Big(\frac{t}{\varepsilon}\Big)\Big)^2dt\Big]+O(\varepsilon^{3/2}).
\end{aligned}
$$

Since

$$2\int_0^1\frac{d}{dt}M_1\Big(\frac{t}{\varepsilon}\Big)M_1\Big(\frac{t}{\varepsilon}\Big)\,dt=-1,$$

we have proven the result of this lemma. □

Remark 7.1. The same argument, but performed on $[0, 1]$ would lead to

$$
\begin{aligned}
&\int_{\mathcal{P}}\Big(\int_{0}^{1}a\Big(\frac{x}{\varepsilon}+y_0(\mu)\Big)\Big[\frac{d}{dx}(u^\varepsilon-s^\varepsilon)\Big]^2dx\Big)\,d\mu\\
&=\Big[\frac{d}{dx}u(1)\Big]^2\cdot\Big[\varepsilon^2\bar a\int_0^1\Big(\frac{d}{dt}\Big(\frac{t}{\varepsilon}M_0\Big(\frac{t}{\varepsilon}\Big)\Big)\Big)^2dt+\varepsilon^2\overline{a\chi^2}\int_0^1\Big(\frac{d}{dt}M_1\Big(\frac{t}{\varepsilon}\Big)\Big)^2dt\\
&\qquad+2\varepsilon\overline{a\chi'}\int_0^1\frac{d}{dt}\Big(\frac{t}{\varepsilon}M_0\Big(\frac{t}{\varepsilon}\Big)\Big)M_1\Big(\frac{t}{\varepsilon}\Big)\,dt\\
&\qquad+\overline{a(\chi')^2}\int_0^1\Big(M_1\Big(\frac{t}{\varepsilon}\Big)\Big)^2dt+\varepsilon\overline{a\chi\chi'}\Big]+O(\varepsilon^{3/2}).
\end{aligned}
$$

By adding the expression in Lemma 7.1 and the one corresponding to the interval $[0, 1]$ we get an expression for the total averaged energy of $u^\varepsilon(x,\mu)-s^\varepsilon(x,\mu)$. In order to find M_0 and M_1 (i.e. m_0 and m_1) that minimize this expression (assuming $((d/dx)u(-1))^2+((d/dx)u(1))^2\neq 0$), it is asymptotically in ε necessary and sufficient to minimize

$$\varepsilon^2\bar{a}\int_0^1\left(\frac{d}{dt}\left(\frac{t}{\varepsilon}M_0\left(\frac{t}{\varepsilon}\right)\right)\right)^2 dt+\varepsilon^2\overline{a\chi^2}\int_0^1\left(\frac{d}{dt}M_1\left(\frac{t}{\varepsilon}\right)\right)^2 dt$$
$$+2\varepsilon\overline{a\chi'}\int_0^1\left(\frac{d}{dt}\right)\left(\frac{t}{\varepsilon}M_0\left(\frac{t}{\varepsilon}\right)\right)M_1\left(\frac{t}{\varepsilon}\right)dt+\overline{a(\chi')^2}\int_0^1\left(M_1\left(\frac{t}{\varepsilon}\right)\right)^2 dt.$$

Performing the change of variable $t\leftarrow t/\varepsilon$ and taking variations we are finally led to the following set of equations for the optimal $\tilde{M}_0(t)=tM_0(t)$ and $M_1(t)$:

$$\begin{aligned}&-\bar{a}\left(\frac{d}{dt}\right)^2\tilde{M}_0-\overline{a\chi'}\frac{d}{dt}M_1=0,\\ &-\overline{a\chi^2}\left(\frac{d}{dt}\right)^2M_1+\overline{a\chi'}\frac{d}{dt}\tilde{M}_0+\overline{a(\chi')^2}M_1=0,\end{aligned}\tag{37}$$

with $\tilde{M}_0(0)=\tilde{M}_0(1)=M_1(1)=0$, $M_1(0)=1$.

8. A boundary-layer analysis for the elliptic system. In this section we shall, based on the analysis of the cutoff in the previous section and a result about the boundary-layer behavior of our elliptic system, justify the earlier claims made concerning optimality. If we select $\overline{a\chi}=0$, the system (8) in this case takes the form

$$-\bar{a}\left(\frac{d}{dx}\right)^2u_0^\varepsilon-\overline{a\chi'}\frac{d}{dx}u_1^\varepsilon=f,$$
$$-\varepsilon^2\overline{a\chi^2}\left(\frac{d}{dx}\right)^2u_1^\varepsilon+\overline{a\chi'}\frac{d}{dx}u_0^\varepsilon+\overline{a(\chi')^2}u_1^\varepsilon=\varepsilon\bar{\chi}f,\quad\text{with } u_0^\varepsilon=u_1^\varepsilon=0\quad\text{for } x=\pm 1.$$

Let $\tilde{M}_0$ and M_1 be the solution to (37) of the previous section, and let us also introduce

$$M_0(t)=t^{-1}\tilde{M}_0(t),$$
$$U_0^\varepsilon(x)=\left(1-M_0\left(\frac{d}{\varepsilon}\right)\right)u(x),$$
$$U_1^\varepsilon(x)=\left(1-M_1\left(\frac{d}{\varepsilon}\right)\right)\frac{d}{dx}u(x),$$

where u as before denotes the solution to (2) and $d=\min\{1+x, 1-x\}$.

LEMMA 8.1. *Assume that a is nonconstant. With notation as explained above we have the following* 3 *estimates,*

$$\|u_0^\varepsilon-U_0^\varepsilon\|_1\leqq C\varepsilon,\quad \|u_1^\varepsilon-U_1^\varepsilon\|_0\leqq C\varepsilon,\quad \|u_1^\varepsilon-U_1^\varepsilon\|_1\leqq C,$$

and therefore also

$$\left\|u_0^\varepsilon(x)+\varepsilon\chi\left(\frac{x}{\varepsilon}\right)u_1^\varepsilon(x)-\left(U_0^\varepsilon(x)+\varepsilon\chi\left(\frac{x}{\varepsilon}\right)U_1^\varepsilon(x)\right)\right\|_1\leqq C\varepsilon,$$

all provided $f\in L^2([-1,1])$.

Proof. This is a standard singular perturbation argument and shall for reasons of brevity not be given here. □

If we recall the results of Lemma 3.2, which in this case state that

$$\|u_0^\varepsilon - u\|_1 \leqq C\varepsilon^{1/2}, \quad \left\|u_1^\varepsilon - \frac{d}{dx}u\right\|_0 \leqq C\varepsilon^{1/2}, \quad \left\|u_1^\varepsilon - \frac{d}{dx}u\right\|_1 \leqq C\varepsilon^{-1/2}$$

(and which for general f are optimal), then Lemma 8.1 simply expresses the fact that U_0^ε, U_1^ε and $U_0^\varepsilon(x)+\varepsilon\chi(x/\varepsilon)U_1^\varepsilon(x)$ correctly reproduce the first boundary layers of u_0^ε, u_1^ε and $u_0^\varepsilon(x)+\varepsilon\chi(x/\varepsilon)u_1^\varepsilon(x)$ respectively.

Now, because of the way U_0^ε and U_1^ε were defined and the way we derived the functions M_0 and M_1, we have finally arrived at the following result.

THEOREM 8.2. *The first boundary-layer of* $u_0^\varepsilon(x)+\varepsilon\chi(x/\varepsilon)u_1^\varepsilon(x)$ *is identical to that which appears when an optimal set of cutoff functions* (m_0, m_1) *is applied to the sum*

$$u(x)+\varepsilon\chi\left(\frac{x}{\varepsilon}\right)\frac{d}{dx}u(x).$$

The optimality is defined in terms of smallest energy-error.

9. Concluding remarks.

Remark 9.1. In order to solve the system (8) numerically we can e.g. use a finite element discretization. This shall be based on a mesh $\mathcal{M}$. As test and trial functions for the u_j^ε's we can choose piecewise polynomials of degree $\leqq r$. The test and trial functions only have to be continuous.

In the engineering literature (see, e.g., [1] and references therein) the following numerical method to deal with problems involving composite materials can be found:

Introduce the set

$$W^r = \left\{ v_0(x)+\varepsilon \sum_{j=1}^{n} \chi_j\left(\frac{x}{\varepsilon}\right)v_j(x) \,\middle|\, v_j \in P_{\mathcal{M}}^r,\ 0 \leqq j \leqq n \right\},$$

where $P_{\mathcal{M}}^r$ denotes the set of continuous piecewise polynomials of degree $\leqq r$, with respect to the mesh $\mathcal{M}$. (Elements of W^r are sometimes referred to as "super elements" because of their similarity to splines in the "fast" variable x/ε.)

Select the set W^r as test and trial functions in the weak formulation

$$\int_\Omega\int \left(\mathbf{A}\left(\frac{x}{\varepsilon}\right)\nabla_x u(x)\nabla_x v(x)+b\left(\frac{x}{\varepsilon}\right)u(x)v(x)\right)dx = \int_\Omega\int f(x)v(x)\,dx\text{''}.$$

Since the coefficients of the bilinear form are nonconstant (highly oscillatory) and the test and trial functions are also highly oscillatory there is a considerable amount of work associated with computing the matrices for this finite dimensional problem exactly.

The finite dimensional equations resulting from the discretization of the system (8), as outlined before, represent the natural homogenization or "lumping" of the equations derived from the "super elements".

Remark 9.2. The numerical experiments that have been carried out, and some of which were presented in [17], indicate an additional benefit of the system (8). It seems to give good results even for moderate to large size ε, a range in which standard homogenization in general fails. This feature is not totally surprising, since after all (8) is very closely related to an energy projection.

Remark 9.3. The system (8) can be viewed as finding a stationary point for an appropriate energy expression within a certain class of functions. This stationary point is more specifically a minimum. Based on (8) we may therefore obtain reliable upper bounds for the energy. For more details and numerical experiments see [17].

A system similar to (8) can also be derived from the complementary energy expression and the adjoint asymptotic analysis as carried out in [3]. Since the energy principle is now one of maximizing an energy expression, this approach will naturally lead to lower bounds for the energy.

REFERENCES

[1] I. BABUSKA, *Homogenization approach in engineering*, in Computing Methods in Applied Sciences and Engineering, Lecture Notes in Economics and Mathematical Systems 134, Springer, New York, 1976, pp. 137–153.

[2] I. BABUSKA, R. B. KELLOGG AND J. PITKÄRANTA, *Direct and inverse error estimates for finite elements with mesh refinements*, Numer. Math., 33 (1979), pp. 447–471.

[3] A. BENSOUSSAN, J.-L. LIONS AND G. PAPANICOLAOU, *Asymptotic Analysis for Periodic Structures*, North-Holland, Amsterdam, 1978.

[4] J. BERGH AND J. LÖFSTRÖM, *Interpolation Spaces*, Springer, New York, 1976.

[5] E. DE GIORGI AND S. SPAGNOLO, *Sulla convergenza degli integrali dell'energia per operatori ellittici del secundo ordine*, Boll. U.M.I., 8 (1973), pp. 391–411.

[6] P. GRISVARD, *Boundary value problems in nonsmooth domains*, Lecture Notes 19, Univ. of Maryland, College Park, 1980.

[7] G. H. HARDY, J. E. LITTLEWOOD AND G. PÓLYA, *Inequalities*, Cambridge Univ. Press, Cambridge, 1934.

[8] J. B. KELLER, *Effective behavior of heterogeneous media*, in Statistical Mechanics and Statistical Methods in Theory and Application, U. Landman, ed., Plenum, New York, 1977, pp. 631–644.

[9] O. A. LADYZHENSKAYA AND N. N. URAL'TSEVA, *Linear and Quasilinear Elliptic Equations*, Academic Press, New York, 1968.

[10] J.-L. LIONS, *Asymptotic calculus of variations*, in Singular Perturbations and Asymptotics, R. Meyer and S. Parter, eds., Academic Press, New York, 1980, pp. 277–296.

[11] J.-L. LIONS AND E. MAGENES, *Nonhomogeneous Boundary Value Problems and Applications, I*, Springer, New York, 1972.

[12] G. PAPANICOLAOU, *Asymptotic analysis of stochastic equations*, in Studies in Probability Theory, MAA Studies 18, M. Rosenblatt, ed., Mathematics Association of America, New York, 1978, pp. 111–179.

[13] G. PAPANICOLAOU AND S. VARADHAN, *Boundary value problems with rapidly oscillating random coefficients*, Coll. Math. Societatis János Bólyai #27, Esztergom, Hungary, 1979, North-Holland, Amsterdam, 1981, pp. 835–873.

[14] E. SANCHEZ-PALENCIA, *Nonhomogeneous media and vibration theory*, Lecture Notes in Physics 127, Springer, New York, 1980.

[15] E. STEIN, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, Princeton, NJ, 1970.

[16] L. TARTAR, *Cours Peccot*, Collège de France, 1977.

[17] M. VOGELIUS, *A variational method to find effective coefficients for periodic media. A comparison with standard homogenization*, in Macroscopic Properties of Disordered Media, Lecture Notes in Physics 154, Springer, New York, 1982, pp. 285–296.

[18] G. WHITHAM, *Linear and Nonlinear Waves*, John Wiley, New York, 1974.

[19] V. V. ZHIKOV, S. M. KOZLOV, O. A. OLEINIK AND KHA T'EN NGOAN, *Averaging and G-convergence of differential operators*, Russian Math. Surveys 34, no. 5 (1979), pp. 69–147.