Andreas Rieder

A domain embedding method for Dirichlet problems in arbitrary space dimension

A DOMAIN EMBEDDING METHOD FOR DIRICHLET PROBLEMS IN ARBITRARY SPACE DIMENSION (*)

Andreas RIEDER (†)

Abstract — *An embedding method for the discretization of Dirichlet boundary value problems over general domains in arbitrary space dimension is proposed The main advantage of the method lies in the use of Cartesian coordinates independent of the underlying domain Error estimates and aspects of the numerical realization are considered To obtain an efficient solver for the resulting linear system of equations an easy-to-use preconditioning is recommended and analyzed A variety of numerical experiments illustrate and confirm the theoretical results* © Elsevier, Paris

Key words boundary value problem, fictitious domain, Galerkin scheme, scaling functions, biorthogonal wavelets, conjugate gradients, preconditioning

AMS subject classification. 65N12, 65N30, 65F10

Résumé — *On présente une méthode de plongement pour la discrétisation des problèmes aux limites de Dirichlet dans les domaines généraux en dimension quelconque L'avantage principal de cette méthode se trouve dans l'utilisation des coordonnées cartésiennes indépendantes du domaine Des estimations d'erreur et les aspects de la réalisation numérique sont considérés Pour obtenir un solveur efficace pour le système d'équations linéaires, on recommande et analyse un préconditionneur facile à réaliser Une multiplicité d'expériences numériques confirme les résultats théoriques* © Elsevier, Paris

1. INTRODUCTION

We will be concerned with the numerical solution of the Dirichlet boundary value problem

$$(1.1a) \qquad - \operatorname{div} (A \nabla u) + \alpha u = f \quad \text{in} \quad \Omega \subset \mathbb{R}^d,$$

$$(1.1b) \qquad u = g \quad \text{on} \quad \partial \Omega.$$

Discretizing the above Dirichlet problem by finite elements requires a triangulation of the domain Ω. If the boundary $\partial \Omega$ has a complicated structure, the generation of a finite element grid aligned with $\partial \Omega$ may be a delicate and time-consuming task.

Furthermore, the topology of a complicated grid is reflected in the data structures. To the net run-time for the arithmetic operations the memory access time has therefore to be added on a considerable scale.

Methods are consequently asked for which can easily be adapted to different domains and which lead to simple data structures. Especially users, which like to solve complicated 3D-problems from real-life applications, appreciate algorithms allowing a clear and fast coding. We refer, e.g., to the popular article [7] by Cipra.

In this paper we present the analysis and the realization of an algorithm for solving (1.1) which has the following three main advantages. Our algorithm

(i) is largely independent of the geometry of the domain Ω ;

(ii) allows Cartesian grids (coordinates) resulting in simple data structures and fast memory access times. Thus, the overhead due to enlarging the computational region is compensated by far;

(iii) requires only little geometric information, namely, a digitalized version of the characteristic function of Ω.

Of course, we have a price to pay for the advantages of the proposed method. Since the Cartesian grid cannot be adapted accurately enough to the boundary of Ω the accuracy of the numerical solution deteriorates slightly near $\partial\Omega$.

The basic idea is not a new one. We embed the domain Ω in a larger parallelepiped $\square$. Then we extend the differential equation (1.1) to a boundary value problem over $\square$ with periodic boundary conditions. The essential boundary condition (1.1b) will be forced approximately. The embedding domain $\square$ is also called *fictitious domain*.

From an abstract point of view our approach can be considered as a Galerkin scheme. The chosen approximation spaces have to satisfy the periodic boundary constraints. Canonical candidates are the periodized *scaling function spaces* generated by translated and scaled versions of one single *refinable* or *scalable* function.

The concept of a scaling function has several advantages. First of all, it allows a unified treatment of a very general class of functions. For instance, the Daubechies scaling functions, see Daubechies [13], the B-Splines and — more general — certain kind of box splines are covered, see, e.g., Chui [6] and de Boor, Hölling and Riemenschneider [14]. Even the classical Lagrange $\mathscr{C}^0$-finite element in $\mathbb{R}^2$, considered as a bivariate three-directional box spline, is a scaling function. Hence, the construction and application of test functions with an arbitrary high degree of smoothness can be realized without difficulties. This raises hopes to combine the localness of finite elements with the high order of approximation of spectral methods.

The scaling function spaces possess an intrinsic multilevel structure which is a main ingredient for the efficient multilevel solvers of the corresponding linear system of equations. Last but not least, the fast assembling of the stiffness matrix, see Dahmen and Micchelli [11], argues in favor of scaling functions as test functions.

This paper is organized as follows. We start in the next section with the weak formulation of the Dirichlet problem (1.1). Then we introduce and explain the fictitious domain method.

As mentioned above scaling functions are the building blocks of the approximation spaces in our Galerkin scheme. In Section 3.1 we therefore recall briefly the properties of scaling functions which we will need in our later analysis. We are now in a position to present the Galerkin discretization in Section 3.2.

A considerable part of the paper deals with error estimates. In Section 4 we obtain H^1- as well as L^2-error estimates. Our error analysis applies to those scaling functions permitting the construction of biorthogonal wavelets. The biorthogonal wavelets do not enter explicitly into our numerical scheme but their existence implies the validity of so called Jackson and Bernstein estimates, see, e.g., Dahmen and Kunoth [10], which we will rely on heavily.

We have already mentioned above that the accuracy of the numerical solution suffers slightly under the rough boundary approximation. Incorporating a-prior knowledge of the analytic solution we are still able to yield optimal H^1-error estimates. If this a-prior knowledge is not available then we will prove optimal error estimates in the interior of Ω at least.

The leading idea underlying our method is its easy implementation for irregular domains in arbitrary space dimensions. So we discuss aspects concerning the numerical realization in Section 5. We support our statements by numerical examples in three space dimensions.

The final numerical task is the solution of a sparse linear system of equations. For that we favor Krylov space iterations, like the conjugate gradient method, which can easily be modified to our needs. Since the condition number of the stiffness matrix grows like one over the square of the discretization step size, a preconditioning of the system is imperative. There is a straightforward way of preconditioning. It will be analyzed in Section 6. Though this way of preconditioning is not optimal, which we will prove by analytical statements as well as numerical experiments, it impresses by efficiency and simple implementation.

Our results in Section 6 finally solve an open problem addressed by Glowinski, Rieder, Wells and Zhou [21]. We comment on this in Remark 6.7.

The paper ends with a discussion of the results in Section 7.

A multitude of articles deals with fictitious domain methods. The papers by Glowinski et al. [19, 20, 21] and Wells and Zhou [35] are akin to the present one. Ficititious domain methods using finite element grids *aligned with* $\partial\Omega$ have been investigated, e.g., by Börgers and Widlund [2], Kuznetsov, Finogenov and Supalov [27] and Nepomnyaschikh [30]. An abstract theory of fictitious domain techniques has been developed by Nepomnyaschikh [29], see also Oswald [33] and Xu [37].

2. FICTITIOUS DOMAIN FORMULATION OF THE DIRICHLET PROBLEM

Let $\Omega \subset \mathbb{R}^d$, $d \geq 2$, be a bounded domain with finite perimeter and Lipschitz continuous boundary $\partial\Omega$ where Ω is located on one side of the boundary. For the details see the $N^{0,1}$-domains of Wloka [36]. These will be our assumptions on the geometry of Ω throughout the paper unless stronger assumptions are explicitly required.

We associate the bilinear form a (2.1) with the boundary value problem (1.1),

$$(2.1) \qquad a(u,v) = \int_\Omega ((A\nabla u) \cdot \nabla v + \alpha uv)\, dx.$$

Then, the *weak* formulation of the Dirichlet problem (1.1) is given by

$$(2.2) \qquad \begin{cases} \text{find } u \in H^1(\Omega) \text{ with } \gamma_{\partial\Omega}\, u = g \text{ such that} \\[2mm] a(u,v) = \int_\Omega fv\, dx \text{ holds for all } v \in H_0^1(\Omega) \end{cases}$$

where $\gamma_{\partial\Omega} : H^1(\Omega) \to H^{1/2}(\partial\Omega)$ represents the *trace operator* satisfying

$$(2.3) \qquad \|\gamma_{\partial\Omega}\, v\|_{H^{1/2}(\partial\Omega)} \leq C_\gamma \|v\|_{H^1(\Omega)}$$

with a positive constant C_γ. We refer, e.g., to Wloka [36] for a detailed description of the trace operator as well as the L^2-Sobolev spaces $H^s(\Omega) = W_2^s(\Omega)$ and $H^s(\partial\Omega) = W_2^s(\partial\Omega)$. By $H_0^1(\Omega)$ we denote the space $H_0^1(\Omega) := \{v \in H^1(\Omega) \mid \gamma_{\partial\Omega}\, v = 0\}$.

A proof of the following well-known existence and uniqueness result can be found, for instance, in the books of Glowinski [18] and Hackbusch [22].

THEOREM 2.1: *Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with a Lipschitz continuous boundary. Assume that the matrix $A = \{a_{ij} \mid 1 \leq i, j \leq d\}$ with entries $a_{ij} \in L^\infty(\Omega)$ satisfies*

$$(2.4) \qquad \langle A(\,\cdot\,)\xi, \xi\rangle_{\mathbb{R}^d} \geq \beta\|\xi\|_{\mathbb{R}^d}^2 \text{ a.e. in } \Omega \text{ for all } \xi \in \mathbb{R}^d$$

for some $\beta > 0$. Further, let α be in $L^\infty(\Omega)$ with $\alpha(x) \geq 0$ a.e. in Ω. Then, the variational problem (2.2) with a as in (2.1) has a unique solution $u \in H^1(\Omega)$ provided $f \in L^2(\Omega)$ and $g \in H^{1/2}(\partial\Omega)$.

Let $\square \subset \mathbb{R}^d$ be an open rectangle (Cartesian product of open intervals) which covers $\bar\Omega$. We now define the periodic Sobolev space $H_p^1(\square)$ by

$$(2.5) \qquad H_p^1(\square) := \{v \in H^1(\square) \mid v \text{ is periodic } w.r.t.\ \partial\square\}.$$

Suppose that $\square = (0, \ell)^d$, $\ell > 0$. Then the periodicity in (2.5) has to be understood by $v(\,\cdot + \ell n) = v(\,\cdot\,)$ for all $n \in \mathbb{Z}^d$.

Let $\tilde A : \square \to \mathbb{R}^{d \times d}$ as well as $\tilde\alpha : \square \to \mathbb{R}$ be extensions of $A : \Omega \to \mathbb{R}^{d \times d}$ and $\alpha : \Omega \to \mathbb{R}$, respectively, that is, $\tilde A|_\Omega = A$ and $\tilde\alpha|_\Omega = \alpha$. With those we define an extension $\tilde a$ of the bilinear form a (2.1) by

$$(2.6) \qquad \tilde a(u,v) := \int_\square ((\tilde A\nabla u) \cdot \nabla v + \tilde\alpha uv)\, dx.$$

Now, we consider the variational problem (2.7),

$$(2.7) \quad \begin{cases} \text{find } \tilde{u} \in H_p^1(\square) \text{ with } \gamma_{\partial\Omega} \tilde{u} = g \text{ such that} \\[2mm] \tilde{a}(\tilde{u}, v) = \int_\square \tilde{f} v \, dx \text{ holds for all } v \in H_p^1(\square) \text{ with } \gamma_{\partial\Omega} v = 0 \,, \end{cases}$$

which is the *fictitious domain formulation* of the Dirichlet problem (2.2). Here, $\tilde{f} : \square \to \mathbb{R}$ is an extension of $f : \Omega \to \mathbb{R}$.

In Lemma 2.2 we analyze the solvability of (2.7) and link (2.2) to (2.7).

LEMMA 2.2: *Let the open rectangle $\square$ cover $\bar{\Omega}$. Further, let the above defined extensions $\tilde{A} = \{\tilde{a}_{ij} \mid 1 \leq i, j \leq d\}$, $\tilde{\alpha}$ and $\tilde{f}$ fulfill: $\tilde{a}_{ij} \in L^\infty(\square)$, $\tilde{\alpha} \in L^\infty(\square)$ with $\tilde{\alpha}(x) \geq \alpha_0 > 0$ a.e. in $\square$ and $\tilde{f} \in L^2(\square)$. Suppose that $\tilde{A}$ satisfies (2.4) where A is replaced by $\tilde{A}$ and Ω by $\square$.*

If the boundary value g is in $H^{1/2}(\partial\Omega)$ then the fictitious domain formulation (2.7) has a unique solution $\tilde{u} \in H_p^1(\square)$ which coincides with the solution u of (2.2) in Ω.

Proof: The key is the existence of an extension $\tilde{g} \in H_0^1(\square)$ of g with $\gamma_{\partial\Omega} \tilde{g} = g$, see, e.g., Wloka [36]. The unique solvability of (2.7) follows now by standard techniques which can be found, e.g., in the books of Glowinski [18] and Hackbusch [22].

Finally, $w := u - \tilde{u}$ vanishes on $\partial\Omega$, i.e., $w \in H_0^1(\Omega)$. Further $a(w, v) = 0$ for all $v \in H_0^1(\Omega)$. Setting $v = w$ yields $0 = a(w, w) \geq \min\{\beta, \alpha_0\} \|w\|_{H^1(\Omega)}^2$ which implies that $u = \tilde{u}$ in Ω. $\square$

3. GALERKIN DISCRETIZATION OF THE FICTITIOUS DOMAIN FORMULATION

The fictitious domain formulation (2.7) is well suited for a Galerkin-type discretization. To that end we will replace $H_p^1(\square)$ in (2.7) by a finite dimensional approximation space. Also we will need a numerical realizable approximation to the trace operator $\gamma_{\partial\Omega}$.

For several reasons mentioned in the introduction we favor scaling function spaces as approximation spaces in our Galerkin scheme. We therefore sketch the concept of scaling functions and some of its implications needed throughout the paper.

3.1. Scaling functions

A function $\varphi \in L^2(\mathbb{R}^d)$ is called *scaling function* if it satisfies the following scaling or refinement equation

$$(3.1) \quad \varphi(x) = 2^{d/2} \sum_{k \in \mathbb{Z}^d} h_k \varphi(2x - k) \,.$$

In the sequel we will only consider scaling functions with compact support. Then the sequence $h = \{h_k\}_{k \in \mathbb{Z}^d}$ of real numbers is finite.

Taking the Fourier transform of both sides of (3.1) we realize that any non-trivial scaling function has a non-vanishing mean value. Thus, we assume the normalization $\int_{\mathbb{R}^d} \varphi(x) \, dx = 1$. Further, we require that the integer translates of φ generate a Riesz system in $L^2(\mathbb{R}^d)$, that is, we have the norm equivalence

$$(3.2) \quad \left\| \sum_{k \in \mathbb{Z}^d} c_k \varphi(\cdot - k) \right\|_{L^2(\mathbb{R}^d)} \sim \|c\|_{\ell^2(\mathbb{Z}^d)} \quad \text{for all} \quad c \in \ell^2(\mathbb{Z}^d) \,.$$

A scaling function φ is of *order* N if the polynomials up to degree $N - 1$ can be expressed by linear combinations of the integer translates of φ. Scaling functions of order 1 satisfy, see, e.g., Fix and Strang [16],

$$(3.3) \qquad 1 = \sum_{k \in \mathbb{Z}^d} \varphi(x - k) .$$

Typical examples for scaling functions with the above requirements are B-splines, several kinds of box splines and the Daubechies scaling functions whose integer translates are even orthonormal.

We now turn to an appropriate periodic setting. Essential properties of scaling functions carry over to their periodized versions.

Let f be in $L_0^2(\mathbb{R}^d)$, the space of compactly supported square integrable functions. We define the periodization $[f]$ of f by $[f](\, . \,) := \sum_{r \in \mathbb{Z}^d} f(\, . + r)$. The operator $[\, . \,]$ maps $L_0^2(\mathbb{R}^d)$ into $L_p^2(\square) = \{v \in L^2(\square) \mid v$ is periodic w.r.t. $\partial \square\}$ where $\square = [0, 1]^d$. For convenience we set

$$(3.4) \qquad f_k^l := [f_{l, k}] \quad \text{where} \quad f_{l, k}(\, . \,) := 2^{dl/2} f(2^l . - k) .$$

With a scaling function φ we associate the spaces V_l^p (3.5), $l \in \mathbb{N}_0$, of dimension 2^{dl},

$$(3.5) \qquad V_l^p := \operatorname{span} \{\varphi_k^l \mid k \in \mathbb{Z}^{d, l}\} \subset L_p^2(\square) ,$$

where $\mathbb{Z}^{d, l} := \mathbb{Z}^d /(2^l \mathbb{Z}^d)$. The refinement equation (3.1) is inherited by φ_k^l. Consequently, the spaces V_l^p are nested, i.e., $V_l^p \subset V_{l + 1}^p$.

3.2. Galerkin discretization of (2.7)

We choose the finite dimensional spaces V_l^p (3.5) as approximation spaces in our Galerkin scheme. The underlying scaling function φ is supposed to be in $H^1(\mathbb{R}^d)$. Without loss of generality we restrict our considerations to the fictitious domain $\square = (0, 1)^d$. So, V_l^p is a subspace of $H_p^1(\square)$.

It will prove convenient to use the following notation. We define the index set

$$(3.6) \qquad \mathscr{B}_l := \{m \in \mathbb{Z}^{d, l} \mid (\operatorname{supp} \varphi_m^l \cap \square)^0 \cap \partial \Omega \neq \varnothing \}$$

which contains the indices of those basis functions having supports whose interiors intersect the boundary of Ω ($\mathscr{B}$ for *boundary*). Next we introduce the approximation $\gamma^l : V_l^p \to V_l^p$ to the trace operator $\gamma_{\partial \Omega}$:

$$(3.7) \qquad \gamma^l(v_l) := \sum_{k \in \mathscr{B}l} v_{l, k} \varphi_k^l, \quad \text{if} \quad v_l = \sum_{k \in \mathbb{Z}^{d, l}} v_{l, k} \varphi_k^l, \quad v_{l, k} \in \mathbb{R} .$$

Now we have all ingredients to discretize (2.7) by the variational problem (3.8),

$$(3.8) \qquad \begin{cases} \text{find } \tilde{u}_l \in V_l^p \text{ with } \gamma^l(\tilde{u}_l) = \gamma^l(\tilde{g}_l) \text{ such that} \\ \tilde{a}(\tilde{u}_l, v_l) = \int_\square \tilde{f} v_l \, dx \text{ holds for all } v_l \in V_l^p \text{ with } \gamma^l(v_l) = 0 . \end{cases}$$

In (3.8), $\tilde{g}_l \in V_l^p$ is a function whose trace $\gamma_{\partial \Omega} \tilde{g}_l$ approximates the boundary value g, cf. (1.1b). A detailed definition of $\tilde{g}_l$ will be given later.

Under the hypotheses of Lemma 2.2 problem (3.8) has a unique solution $\tilde{u}_l$.

4. CONVERGENCE AND ERROR ESTIMATES

We start with some preparing considerations. We introduce two more index sets, the "interior" and "exterior" indices:

$$\mathscr{I}_l := \left\{ m \in \mathbb{Z}^{d,\,l} \mid \operatorname{supp} \varphi_m^l \cap \square \subset \Omega \right\} \quad \text{and} \quad \mathscr{E}_l := \mathbb{Z}^{d,\,l} \setminus (\mathscr{I}_l \cup \mathscr{B}_l) \,.$$

Any $v_l \in V_l^p$ can be split according to $v_l = v_l^{\mathscr{I}} + v_l^{\mathscr{E}} + \gamma^l(v^l)$ where $v_l^{\mathscr{I}} := \sum_{k \in \mathscr{I}_l} v_{l,k}\, \varphi_k^l$ and $v_l^{\mathscr{E}} := \sum_{k \in \mathscr{E}_l} v_{l,k}\, \varphi_k^l$. Please note that $v_l^{\mathscr{I}}$ and $v_l^{\mathscr{E}}$ have disjoint supports. Therefore, both parts $\tilde{u}_l^{\mathscr{I}}$ and $\tilde{u}_l^{\mathscr{E}}$ of the solution $\tilde{u}_l$ of (3.8) are decoupled. The interior part $\tilde{u}_l^{\mathscr{I}}$ is uniquely determined as the solution of the homogeneous variational problem (4.1),

$$(4.1) \qquad \begin{cases} \text{find } \tilde{u}_l^{\mathscr{I}} \in V_l^{p,\,\mathscr{I}} \text{ such that} \\[2mm] a(\tilde{u}_l^{\mathscr{I}}, v_l) = \displaystyle\int_\Omega f v_l \, dx - a(\gamma^l(\tilde{g}_l), v_l) \text{ holds for all } v_l \in V_l^{p,\,\mathscr{I}} \,. \end{cases}$$

Here, $V_l^{p,\,\mathscr{I}} := \left\{ v_l \in V_l^p \mid v_l = v_l^{\mathscr{I}} \right\} \subset H_0^1(\Omega)$.

THEOREM 4.1: *Adopt the hypotheses of Lemma 2.2. Let u and $\tilde{u}_l$ be the solutions of (2.2) and (3.8), respectively. Then, there is a positive constant C_S so that*

$$(4.2) \qquad \| u - \tilde{u}_l \|_{H^1(\Omega)} \le C_S \inf \left\{ \| u - w_l \|_{H^1(\Omega)} \mid w_l \in V_l^p \text{ with } \gamma^l(w_l) = \gamma^l(\tilde{g}_l) \right\} .$$

Proof: We have that

$$(4.3) \qquad a(u, v_l) = \langle f, v_l \rangle_{L^2(\Omega)} \quad \text{and} \quad a(\tilde{u}_l, v_l) = \langle f, v_l \rangle_{L^2(\Omega)}$$

for all $v_l \in V_l^{p,\,\mathscr{I}}$. The right relation comes from (4.1) since $\tilde{u}_l$ and $\tilde{u}_l^{\mathscr{I}} + \gamma^l(\tilde{g}_l)$ coincide in Ω. Subtracting the right from the left equality yields $a(u - \tilde{u}_l, v_l) = 0$ for all $v_l \in V_l^{p,\,\mathscr{I}}$. Actually, we have that $a(u - \tilde{u}_l, v_l) = 0$ for all $v_l \in V_l^p$ with $\gamma^l(v_l) = 0$. We can now proceed as in the proof of Cea's lemma, see, e.g., Glowinski [18, p. 327]. $\square$

4.1. H^1-error in Ω

In this section we analyze the infimum on the right-hand side of (4.2). The following investigations are related to the space

$$V_l(\Omega) := \operatorname{span} \left\{ \varphi_m^l \mid m \in \mathscr{B}_l \cup \mathscr{I}_l \right\}$$

which is spanned by those basis functions whose supports intersect $\bar{\Omega}$. Since $\bar{\Omega} \subset \square$, $V_l(\Omega)$ does not contain periodized basis functions for l sufficiently large. Thus, we henceforth assume the representation

$$(4.4) \qquad V_l(\Omega) = \operatorname{span} \left\{ \varphi_{l,\,m} \mid m \in \mathscr{B}_l \cup \mathscr{I}_l \right\}$$

for all l. The space $V_l(\Omega)$ is a subspace of $V_l := \text{span}\{\varphi_{l,m} \mid m \in \mathbb{Z}^d\}$ as soon as we identify the equivalence classes in $\mathbb{Z}^{d,l}$ with their representers in $[0, 2^l - 1]^d$. We associate the discretization step size $\delta_l = 2^{-l}$ with V_l (resp. with $V_l(\Omega)$ and V_l^p).

We require additional properties of the underlying scaling function φ. In the sequel we will employ the *Jackson* or *direct* estimate (4.5): there is a sequence of projection operators $P_l : L^2(\mathbb{R}^d) \to V_l$, $l \in \mathbb{N}_0$, two positive numbers v, N and a non-negative number q such that

$$(4.5) \qquad \|f - P_l f\|_{H^s(\mathbb{R}^d)} \leqslant C_J \delta_l^{t-s} \|f\|_{H^t(\mathbb{R}^d)} \quad \text{for all } f \in H^t(\mathbb{R}^d)$$

where $s \leqslant t$, $0 \leqslant s < q + v$ and $0 \leqslant t \leqslant N$. The positive constant C_J does not depend on l or f. Further we will rely on the *Bernstein* or *inverse* estimate (4.6),

$$(4.6) \qquad \|v_l\|_{H^t(\mathbb{R}^d)} \leqslant C_B \delta_l^{s-t} \|v_l\|_{H^s(\mathbb{R}^d)} \quad \text{for all } v_l \in V_l,$$

where $0 \leqslant s \leqslant t < q + v$ and $C_B > 0$ is independent of v_l and l.

REMARK 4.2: *Let φ be a scaling function of order N which permits the construction of biorthogonal wavelets and which generates a Riesz system (3.2) in $L^2(\mathbb{R}^d)$. Then, both estimates (4.5) and (4.6) are valid where $v > 0$ is the Hölder exponent of the q-th order derivatives of φ. Furthermore, the operators P_l are explicitly given by*

$$(4.7) \qquad P_l f := \sum_{m \in \mathbb{Z}^d} \langle f, \tilde{\varphi}_{l,m} \rangle_{L^2(\mathbb{R}^d)} \varphi_{l,m}$$

where $\tilde{\varphi}$ is the dual scaling function to φ.

These results originate in the work of Dahmen and Kunoth [10].

Tensor products of B-splines are scaling functions with the required properties. Likewise, certain kinds of box splines belong to this class of functions as well, see Dahlke, Latour and Gröchenig [9].

In view of the above remark we restrict our attention to scaling functions belonging to biorthogonal wavelets, see, e.g., Cohen, Daubechies and Feauveau [8] for an introduction to biorthogonal wavelets.

Let $\tilde{\varphi}$ be a dual scaling function to φ so that

$$(4.8) \qquad \text{supp } \varphi \subset \text{supp } \tilde{\varphi} .$$

The examples in [8] show that (4.8) is not a restrictive assumption. We set $\tilde{S} := \text{supp } \tilde{\varphi}$. Then, $\tilde{S}_{l,m} := \delta_l(m + \tilde{S})$ is the support of $\tilde{\varphi}_{l,m}$.

Up to now we did not specify the approximation $\tilde{g}_l \in V_l^p$ of the boundary value g, cf. (3.8). From now on let $\tilde{g}$ be a square integrable extension of g to $\square$. We define

$$(4.9) \qquad \tilde{g}_l := \sum_{m \in \mathscr{B}_l \cup \mathscr{I}_l} \langle \tilde{g}, \tilde{\varphi}_{l,m} \rangle_{L^2(\mathbb{R}^d)} \varphi_{l,m} .$$

We would like to apply the global Jackson and Bernstein estimates to achieve a local error estimate (over Ω). For that reason we recall the existence of a bounded linear extension operator $E_B : H^r(B) \to H^r(\mathbb{R}^d)$ which fulfills

$$(4.10) \qquad \|E_B f\|_{H^r(\mathbb{R}^d)} \leqslant C_E \|f\|_{H^r(B)}, \, 0 \leqslant r \leqslant s, \quad \text{and} \quad E_B f = f \quad \text{a.e. in } B$$

provided the bounded domain $B \subset \mathbb{R}^d$ has a sufficiently smooth boundary. If B enjoys our standard assumptions on Ω, see Section 2, then (4.10) holdstrue for $s, r \in \mathbb{N}$ and the positive constant C_E depends only on s and B, see, e.g., Wloka [36].

To guide the reader we give a brief outlook on what follows in the remainder of this subsection. We start with a fundamental inequality in Lemma 4.3. Based on this result we present an optimal H^1-error estimate in Theorem 4.4 as far as an *a priori* knowledge of the solution near the boundary is available. Theorem 4.5 contains a non-optimal convergence result under comparatively weak requirements.

We denote the diameter of a set $M \subset \mathbb{R}^d$ and the distance of a point $x \in \mathbb{R}^d$ to M in the Euclidean norm by diam M and dist(x, M), respectively.

LEMMA 4.3: *Let φ be a scaling function satisfying (4.5) and (4.6) with P_l from (4.7) for $1 < q + v \le N$. Further, let there exist a domain Ω_E with an arbitrary smooth boundary such that $\bar{\Omega} \subset \Omega_E$ and $\bar{\Omega}_E \subset \square$.*

If the solution $\tilde{u}$ of the Dirichlet problem (2.7) is in $H^t(\Omega_E)$ for some $t \in [1, N]$ and $\tilde{u}_l$ solves (3.8) with the boundary constraint (4.9) then

$$(4.11) \qquad \| u - \tilde{u}_l \|_{H^1(\Omega)} \le C_F \left\{ \delta_l^{t-1} \| \tilde{u} \|_{H^t(\Omega_E)} + \delta_l^{-1} \left(\sum_{m \in \mathcal{B}_l} \| \tilde{u} - \tilde{g} \|_{L^2(\tilde{S}_{l,m})}^2 \right)^{1/2} \right\}$$

for l sufficiently large. The positive constant C_F in (4.11) depends on Ω_E, N and $\tilde{\varphi}$.

Proof: The stated estimate follows immediately by (4.2) when choosing w_l in a special way, namely as $w_l = \tilde{w}_l^{\tilde{g}} + \gamma^l(\tilde{g}_l)$ where

$$\tilde{w}_l := \sum_{m \in \mathbb{Z}^d \setminus \mathcal{B}_l} \langle E_{\Omega_E} \tilde{u}, \tilde{\varphi}_{l,m} \rangle_{L^2(\mathbb{R}^d)} \varphi_{l,m} + \sum_{m \in \mathcal{B}_l} \langle \tilde{g}, \tilde{\varphi}_{l,m} \rangle_{L^2(\mathbb{R}^d)} \varphi_{l,m}.$$

Obviously, $w_l \in V_l(\Omega)$, $\gamma^l(w_l) = \gamma^l(\tilde{g}_l)$, and $w_l|_\Omega = \tilde{w}_l|_\Omega$. Now,

$$\| u - \tilde{u}_l \|_{H^1(\Omega)} \overset{(4.2)}{\le} C_S \| \tilde{u} - \tilde{w}_l \|_{H^1(\Omega_E)} \le C_S \| E_{\Omega_E} \tilde{u} - \tilde{w}_l \|_{H^1(\mathbb{R}^d)}$$

$$\le C_S (\| E_{\Omega_E} \tilde{u} - P_l E_{\Omega_E} \tilde{u} \|_{H^1(\mathbb{R}^d)} + \| P_l E_{\Omega_E} \tilde{u} - \tilde{w}_l \|_{H^1(\mathbb{R}^d)}).$$

Next we apply the Jackson and Bernstein inequalities, that is,

$$\| u - \tilde{u}_l \|_{H^1(\Omega)} \le C_S C_J C_E \delta_l^{t-1} \| \tilde{u} \|_{H^t(\Omega_E)} + C_S C_B \delta_l^{-1} \| P_l E_{\Omega_E} \tilde{u} - \tilde{w}_l \|_{L^2(\mathbb{R}^d)}.$$

where we also used (4.10). It remains to estimate the L^2-norm of

$$P_l E_{\Omega_E} \tilde{u} - \tilde{w}_l = \sum_{m \in \mathcal{B}_l} \langle E_{\Omega_E} \tilde{u} - \tilde{g}, \tilde{\varphi}_{l,m} \rangle_{L^2(\tilde{S}_{l,m})} \varphi_{l,m}.$$

Let m an arbitrary element of $\mathcal{B}_l$. The support $\tilde{S}_{l,m}$ intersects $\partial\Omega$, cf. (4.8). Therefore the distance of a point $x \in \tilde{S}_{l,m}$ to $\partial\Omega$ is bounded by dist $(x, \partial\Omega) \le$ diam $\tilde{S}_{l,m} = \delta_l \cdot$ diam $\tilde{S}$. Consequently, the support $\tilde{S}_{l,m}$ lies completely inside Ω_E if l is sufficiently large. Then $E_{\Omega_E} \tilde{u}|_{\tilde{S}_{l,m}} = \tilde{u}|_{\tilde{S}_{l,m}}$, see (4.10), and we finally end with

$$(4.12) \qquad \| P_l E_{\Omega_E} \tilde{u} - \tilde{w}_l \|_{L^2(\mathbb{R}^d)} \le C_R \| \tilde{\varphi} \|_{L^2(\tilde{S})} \left(\sum_{m \in \mathcal{B}_l} \| \tilde{u} - \tilde{g} \|_{L^2(\tilde{S}_{l,m})}^2 \right)^{1/2}$$

where we have first used (3.2) and then the Cauchy inequality. $\square$

In the following theorem we require the regularity $\tilde{u} \in H^2(\Omega_E)$. This regularity applies under suitable smoothness assumptions (in Ω_E) on the coefficients of the bilinear form $\tilde{a}$ (2.6) and the right-hand side $\tilde{f}$ of (2.7), see, e.g., Hackbusch [22, Chapter 9].

THEOREM 4.4: *Let the hypotheses of Lemma 4.3 be satisfied and let the solution $\tilde{u}$ of (2.7) be in $H^2(\Omega_E)$. Furthermore, we require that*

$$\tilde{u} - \tilde{g} \in \mathscr{C}^{1, 1/2}(T_\tau) \quad \text{as well as} \quad \nabla \tilde{u}|_{\partial\Omega} = \nabla \tilde{g}\,|_{\partial\Omega} \; a.e.$$

where $T_\tau = \{y \in \square \mid \operatorname{dist}(y, \partial\Omega) < \tau\}$, $\tau > 0$, *satisfying* $T_\tau \subset \Omega_E$. *Then,*

$$\|u - \tilde{u}_l\|_{H^1(\Omega)} \leq C_e \delta_l (\|\tilde{u}\|_{H^2(\Omega_E)} + \|\tilde{u} - \tilde{g}\|_{\mathscr{C}^{1, 1/2}(T_\tau)})$$

when l is sufficiently large and where C_e is a positive constant.

Proof: In a first step we provide an auxiliary estimate. Let $f \in \mathscr{C}^{1, 1/2}(D)$, $D \subset \mathbb{R}^d$ open, and let $x, y \in D$ be two points whose connecting line segment $[x, y]$ lies in D. The mean value theorem in integral form implies that

$$(4.13) \qquad |f(x) - f(y) - \langle \nabla f(y), y - x \rangle_{\mathbb{R}^d}| \leq \frac{2\sqrt{d}}{3} \|f\|_{\mathscr{C}^{1, 1/2}(D)} \|x - y\|_{\mathbb{R}^d}^{3/2}.$$

Now we set $f := \tilde{u} - \tilde{g} \in \mathscr{C}^{1, 1/2}(T_\tau)$. Both, f and ∇f vanish on $\partial\Omega$. Let $x \in T_\tau$ be arbitrary and let $z \in \partial\Omega$ be given by $\operatorname{dist}(x, \partial\Omega) = \|x - z\|_{\mathbb{R}^d}$. Clearly, $[x, z] \subset T_\tau$. By (4.13) and $f(x) = f(x) - f(z) - \langle \nabla f(z), z - x \rangle_{\mathbb{R}^d}$ we get

$$|\tilde{u}(x) - \tilde{g}(x)| \leq \frac{2\sqrt{d}}{3} \|\tilde{u} - \tilde{g}\|_{\mathscr{C}^{1, 1/2}(T_\tau)} \operatorname{dist}(x, \partial\Omega)^{3/2} \quad \text{for all} \quad x \in T_\tau.$$

The latter inequality yields

$$\|\tilde{u} - \tilde{g}\|_{L^2(\tilde{S}_{l,m})}^2 \leq C \|\tilde{u} - \tilde{g}\|_{\mathscr{C}^{1, 1/2}(T_\tau)}^2 \delta_l^{d+3}$$

for l such large that $\delta_l . \operatorname{diam} \tilde{S} \leq \tau$. The constant C depends only on $\tilde{S}$ and d. Since the cardinality of $\mathscr{B}_l$ grows like δ_l^{1-d}, the proof ends by applying (4.11). $\square$

The essential assumption $\tilde{u} - \tilde{g} \in \mathscr{C}^{1, 1/2}(T_\tau)$ of Theorem 4.4 might be satisfied, though neither $\tilde{u}$ nor $\tilde{g}$ are in $\mathscr{C}^{1, 1/2}(T_\tau)$. It suffices that $\tilde{g}$ reproduces the $\mathscr{C}^{1, 1/2}$-singular behavior of $\tilde{u}$ in a vicinity of $\partial\Omega$. If $\tilde{u} \in \mathscr{C}^{1, 1/2}(T_\tau)$ (see, e.g., Gilbarg and Trudinger [17] for sufficient conditions) then any $\mathscr{C}^{1, 1/2}$-extension of g locally around $\partial\Omega$ will do the job whenever its first order derivates agree with the ones of $\tilde{u}$ on $\partial\Omega$.

Of course, non-optimal H^1-convergence holds under weaker assumptions.

THEOREM 4.5: *Let φ be a scaling function satisfying (4.5) and (4.6) with P_l from (4.7) for $1 < q + v \leq N$. Let Ω_E be as in Lemma 4.3.*

Suppose that the solution u of the Dirichlet problem (2.2) is in $H^{1+t}(\Omega)$ and that $\tilde{g}$ is in $H^{1+t}(\Omega_E)$ for some $t \in [0, N-1]$.

If the extension operator E_Ω meets (4.10) with $s = t + 1$ then

$$(4.14) \qquad \|u - \tilde{u}_l\|_{H^1(\Omega)} \leq C_e(\delta_l^t \|u\|_{H^{1+t}(\Omega)} + \delta_l^\lambda \|E_\Omega u - \tilde{g}\|_{H^{1+\lambda}(\Omega_E)})$$

as $l \to \infty$ where $\lambda = \min\{1/2 - p, t\}$ for any $p > 0$.

Proof: The first part of the proof is completely analogous to the proof of Lemma 4.3. In the definition of $\tilde{w}_l$ we replace $E_{\Omega_E}\tilde{u}$ by $E_\Omega u$ and obtain

$$\|u - \tilde{u}_l\|_{H^1(\Omega)} \leq C(\delta_l^t \|u\|_{H^{1+t}(\Omega)} + \delta_l^{-1} \|P_l E_\Omega u - \tilde{w}_l\|_{L^2(\mathbb{R}^d)}).$$

Estimating the sum in (4.12) we realize that

$$\| P_l E_\Omega u - \tilde{w}_l \|_{L^2(\mathbb{R}^d)} \leq C_T \| E_\Omega u - \tilde{g} \|_{L^2(T)}$$

where C_T is a positive constant and where $T = \{ y \in \square \mid \text{dist}\,(y, \partial\Omega) < \vartheta_l \}$ with $\vartheta_l = \delta_l \cdot \text{diam}\,\tilde{S}$. Let l be such large that $T \subset \Omega_E$. Since $\gamma_{\partial\Omega}(E_\Omega u - \tilde{g}) = 0$ both inequalities of Lemma 6.1 of Bramble and Pasciak [4] may be used to estimate

$$\| E_\Omega u - \tilde{g} \|_{L^2(T)} \leq \bar{C}_{BP}\, \vartheta_l^{1 + \lambda} \| E_\Omega u - \tilde{g} \|_{H^{1 + \lambda}(\Omega_E)}$$

wich ends the proof of (4.14). $\square$

Remark 4.6: The error in (4.14) has the non-optimal decay $\mathcal{O}(\delta_l^{1/2 - \rho})$, $\rho > 0$, as $l \to \infty$ provided $u \in H^{1 + t}(\Omega)$ and $\tilde{g} \in H^{1 + t}(\Omega_E)$, $t \geq 1/2$. This order of decay comes arbitrarily close to the order achieved in using the Lagrange $\mathscr{C}^0$-finite element for a *polygonal* boundary approximation of non-convexe domains, see Hackbusch [22, Chapter 8.6] as well as Strang and Fix [34, Chapter 4.4].

4.2. L^2-error in Ω

We present a modification of the duality argument by Nitsche [31]. An estimate by Braess [3, Chapter 3.1] inspired the analysis of this section.

We will obtain a non-optimal L^2-error estimate. In contrary to the H^1-setting, the L^2-non-optimality cannot be overcome by *a priori* information about the solution near the boundary. The reason for that is solely the rough boundary approximation which has its effects here in full force.

In the sequel we adopt the hypotheses of Theorem 2.1 and additionally assume that $\partial\Omega$ is $\mathscr{C}^2$ and that the entries of the coefficient matrix A are uniformly Lipschitz continuous in Ω.

With the error $e := u - \tilde{u}_l \in H^1(\Omega)$ we define the homogeneous problem (4.15),

$$(4.15) \qquad \begin{cases} \text{find } w \in H_0^1(\Omega) \text{ such that} \\[2mm] b(w, v) = \displaystyle\int_\Omega ev\, dx \text{ holds for all } v \in H_0^1(\Omega)\,. \end{cases}$$

In (4.15), b is the bilinear form adjoint to a: $b(w, v) := a(v, w)$. The ellipticity, the continuity as wells as the regularity carry over from a to b. Especially, (4.15) has a unique solution w fulfilling

$$(4.16) \qquad\qquad \| w \|_{H^2(\Omega)} \leq C_H \| e \|_{L^2(\Omega)}\,,$$

see, e.g., Gilbarg and Trudinger [17, Theorem 8.12]. Since the trace $\gamma_{\partial\Omega} e = g - \gamma_{\partial\Omega}\tilde{g}_l$ of e does not vanish we have that $b(w, e) = \| e \|_{L^2(\Omega)}^2 + \displaystyle\int_{\partial\Omega} \lambda(g - \tilde{g}_l)\, ds$ as a consequence of Green's formula where $\lambda := (A^t \nabla w) \cdot \vec{n}$ and $\vec{n}$ is the outward normal vector to $\partial\Omega$. The equations in (4.3) give $b(z_l, e) = a(u - \tilde{u}_l, z_l) = 0$ for all $z_l \in V_l^{p,\mathscr{S}}$. We get

$$\| e \|_{L^2(\Omega)}^2 = b(w - z_l, e) - \int_{\partial\Omega} \lambda(g - \tilde{g}_l)\, ds \quad \text{for all } z_l \in V_l^{p,\mathscr{S}}$$

and further

$$(4.17) \qquad \| e \|_{L^2(\Omega)}^2 \leq M \| w - z_l \|_{H^1(\Omega)} \| e \|_{H^1(\Omega)} + \| \lambda \|_{L^2(\partial\Omega)} \| g - \tilde{g}_l \|_{L^2(\partial\Omega)}$$

for all $z_l \in V_l^{p,\mathcal{F}}$ where $M = \max \left\{ \sup_{x \in \Omega} \| A(x) \|_{\mathbb{R}^d}, \| \alpha \|_{L^\infty(\Omega)} \right\}$. We estimate the L^2-norm of λ employing $|\lambda(x)| \leq M \| \nabla w(x) \|_{\mathbb{R}^d}$, (2.3) and (4.16):

$$(4.18) \qquad \| \lambda \|_{L^2(\partial\Omega)}^2 \leq M^2 \sum_{i=1}^d \left\| \frac{\partial w}{\partial x_i} \right\|_{H^{1/2}(\partial\Omega)}^2 \leq M^2 C_\gamma^2 C_H^2 \| e \|_{L^2(\Omega)}^2 .$$

LEMMA 4.7: *Let $\tilde{g}$ be in $H^2(\Omega_E)$ with Ω_E as in Lemma 4.3. Adopt the assumptions from above and the hypotheses on φ of Lemma 4.3. Then,*

$$(4.19) \qquad \| g - \tilde{g}_l \|_{L^2(\partial\Omega)} \leq C_g \delta_l^{3/2 - \epsilon} \| \tilde{g} \|_{H^2(\Omega_E)}$$

for l sufficiently large and any $\epsilon > 0$. In (4.19), C_g is a positive constant.

Proof: For l large enough we have that $\tilde{g}_l |_\Omega = P_l E_{\Omega_E} \tilde{g} |_\Omega$. Further,

$$\| g - \tilde{g}_l \|_{L^2(\partial\Omega)} \leq \| \gamma_{\partial\Omega} (\tilde{g} - \tilde{g}_l) \|_{H^\epsilon(\partial\Omega)} \leq C_\gamma \| \tilde{g} - \tilde{g}_l \|_{H^{\epsilon + 1/2}(\Omega)},$$

$0 < \epsilon \leq 1/2$, which follows from the trace theorem, see, e.g., Wloka [36]. Finally,

$$\| g - \tilde{g}_l \|_{L^2(\partial\Omega)} \leq C_\gamma \| \tilde{g} - P_l E_{\Omega_E} \tilde{g} \|_{H^{\epsilon + 1/2}(\Omega)} \leq C_\gamma \| E_{\Omega_E} \tilde{g} - P_l E_{\Omega_E} \tilde{g} \|_{H^{\epsilon + 1/2}(\mathbb{R}^d)}$$

which implies (4.19) by (4.5) and (4.10). $\square$

We plug both inequalities (4.19) and (4.18) into (4.17) and get ($0 < \epsilon \leq 1/2$)

$$(4.20) \qquad \| e \|_{L^2(\Omega)}^2 \leq C_L (\| e \|_{H^1(\Omega)} \mathscr{A}(w) + \delta_l^{3/2 - \epsilon} \| \tilde{g} \|_{H^2(\Omega_E)} \| e \|_{L^2(\Omega)})$$

where $\mathscr{A}(w) := \inf \left\{ \| w - z_l \|_{H^1(\Omega)} \mid z_l \in V_l^{p,\mathcal{F}} \right\}$ and C_L is a suitable positive constant. In Theorem 4.5 we have already analyzed the infimum $\mathscr{A}(w)$ implicitly.

LEMMA 4.8: *Let $w \in H^2(\Omega)$ be the solution of (4.15) fulfilling (4.16). Let φ be as in Theorem 4.5. Then there is a positive constant C_e such that*

$$(4.21) \qquad \mathscr{A}(w) = \inf \left\{ \| w - z_l \|_{H^1(\Omega)} \mid z_l \in V_l^{p,\mathcal{F}} \right\} \leq C_e \delta_l^{1/2 - \beta} \| e \|_{L^2(\Omega)}$$

for any $\beta > 0$.

Proof: The infimum can be bounded by the right-hand side of (4.14) where we have to replace $u = w$, $\tilde{g} = 0$ and $t = 1$: $\mathscr{A}(w) \leq C \delta_l^{1/2 - \beta} \| w \|_{H^2(\Omega)}$. The proof ends by (4.16). $\square$

Plugging (4.21) into (4.20) finally yields ($\epsilon, \beta > 0$)

$$\| u - \tilde{u}_l \|_{L^2(\Omega)} \leq C_L (\delta_l^{1/2 - \beta} \| u - \tilde{u}_l \|_{H^1(\Omega)} + \delta_l^{3/2 - \epsilon} \| \tilde{g} \|_{H^2(\Omega_E)})$$

for l sufficiently large and where C_L is again a suitable positive constant. Applying the results of the Theorems 4.4 and 4.5 we find the following L^2-estimates.

THEOREM 4.9: *Let φ and Ω_E be as in Lemma 4.3. Let $\tilde{g}$ be in $H^2(\Omega_E)$. Adopt the assumptions on $\partial\Omega$ and A from above. Then,*

$$\| u - \tilde{u}_l \|_{L^2(\Omega)} \leq C_L \delta_l^{1 - \beta} (\| u \|_{H^2(\Omega)} + \| E_\Omega u - \tilde{g} \|_{H^{3/2}(\Omega_E)} + \| \tilde{g} \|_{H^2(\Omega_E)})$$

for l sufficiently large and $\beta > 0$. Under the hypotheses of Theorem 4.4 we have

$$\| u - \tilde{u}_l \|_{L^2(\Omega)} \leqslant C_L \delta_l^{3/2 - \beta} (\| \tilde{u} \|_{H^2(\Omega_E)} + \| \tilde{u} - \tilde{g} \|_{\mathscr{C}^{1,1/2}(T_z)} + \| \tilde{g} \|_{H^2(\Omega_E)})$$

for l sufficiently large and $\beta > 0$. In any of the above estimates C_L denotes another positive constant.

The non-optimal L^2-estimates are caused by the approximation error (4.21). This error is essential by the chosen kind of boundary approximation via γ^l. The constraint $\gamma^l(v_l) = 0$ forces zero not only on $\partial\Omega$ but also on the strip

(4.22)
$$\partial\Omega^l := \bigcup_{m \in \mathscr{B}_l} \operatorname{supp} \varphi_{l,m} .$$

This strip has width diam $(\operatorname{supp} \varphi) \cdot \delta_l$ in general. Consequently, one cannot expect a better order of convergence than $1/2$. In this sense the estimate (4.21) is optimal.

In the finite element theory this corresponds to the use of rectangular elements for an approximation of curvilinear boundaries leading to the same error decay as in (4.21), see, e.g., Strang and Fix [34, Chapter 4.4].

4.3. H^1-error in the interior of Ω

The effects caused by an inadequate extension of g are locally in nature. So we expect the validity of optimal error estimates in the interior of Ω independent of the behavior of $\tilde{g}$ in the vicinity of $\partial\Omega$.

The interior estimates we need are due to Nitsche and Schatz [32]. Bertoluzza [1] showed that our approximation spaces $V_l(\Omega)$ (4.4) possess the properties to apply the theory of Nitsche and Schatz as far as the scaling function φ satisfies the hypotheses of Lemma 4.3. Therefore we have

(4.23)
$$\| u - \tilde{u}_l \|_{H^s(\Omega_0)} \leqslant C_{NS} (\delta_l^{t-s} \| u \|_{H^t(\Omega_1)} + \| u - \tilde{u}_l \|_{H^{-p}(\Omega_1)})$$

for l sufficiently large and $s \in \{0,1\}$, $1 \leqslant t \leqslant N$, $p \in \mathbb{N}_0$. The domains Ω_0, Ω_1 and Ω are nested: $\bar{\Omega}_0 \subset \Omega_1$, $\bar{\Omega}_1 \subset \Omega$.

THEOREM 4.10: *Let $u \in H^2(\Omega)$ and $\tilde{g} \in H^2(\Omega_E)$ with Ω_E from Lemma 4.3. Further, we adopt the remaining assumptions of Theorem 4.5. The domains Ω_0, Ω_1 and Ω are nested as explained above. Then there is a positive constant C_{IN} such that*

$$\| u - \tilde{u}_l \|_{H^1(\Omega_0)} \leqslant C_{IN} \delta_l^{1-\beta} (\| u \|_{H^2(\Omega)} + \| \tilde{g} \|_{H^2(\Omega_E)})$$

for l sufficiently large and $\beta > 0$.

Proof: We apply (4.23) with $s = 1$, $t = 2$ and $p = 0$. Since $\| u - \tilde{u}_l \|_{L^2(\Omega_1)} \leqslant \| u - \tilde{u}_l \|_{L^2(\Omega)}$ we may apply the first inequality of Theorem 4.9. $\square$

In the interior of Ω we achieve virtually the optimal H^1-error estimate, however, under negligible assumptions on the extension $\tilde{g}$ of the boundary value g.

5. ASPECTS CONCERNING THE NUMERICAL REALIZATION

In this section we investigate the structure of the linear system of equations being equivalent to the variational problem (3.8). We will concentrate on the generation of the stiffness matrix and some of its properties. In that we will take care to manage without an explicit parametrization of the boundary of Ω. The minimal geometric information we allow will be a digitalized version of the characteristic function χ_Ω of Ω. In this respect we will be able to develop a program code which is *independent* of the shape of the domain. A complicated and time-consuming grid generation adapted to the domain is canceled.

5.1. The linear system

Let $\mathbf{A}_l \in \mathbb{R}^{\mathbb{Z}^{d,l} \times \mathbb{Z}^{d,l}}$ be the stiffness matrix with entries

$$(5.1) \qquad (\mathbf{A}_l)_{k,r} := \tilde{a}(\varphi_k^l, \varphi_r^l), \quad r, k \in \mathbb{Z}^{d,l},$$

and let $\mathbf{f}_l \leqslant \mathbb{R}^{\mathbb{Z}^{d,l}}$ be the vector with components $(\mathbf{f}_l)_k := \int_\square \tilde{f}\varphi_k^l \, dx$, $k \in \mathbb{Z}^{d,l}$. Furthermore we define a diagonal matrix $M_l \in \mathbb{R}^{\mathbb{Z}^{d,l} \times \mathbb{Z}^{d,l}}$ representing $\partial\Omega$,

$$(5.2) \qquad (M_l)_{k,r} := \begin{cases} 1 : & k = r \text{ and } k, r \in \mathscr{B}_l \\ 0 : & \text{otherwise} \end{cases},$$

and a vector $\mathbf{g}_l \in \mathbb{R}^{\mathbb{Z}^{d,l}}$ containing the boundary data: $(\mathbf{g}_l)_k := \int_\square \tilde{g} \, \tilde{\varphi}_k^l \, dx$, $k \in \mathbb{R}^{\mathbb{Z}^{d,l}}$, see (4.9). We consider the linear system

$$(5.3) \qquad \begin{cases} \text{find } u_l \in \mathsf{R}(I - M_l) \text{ as the unique solution of} \\ (I - M_l) \, \mathbf{A}_l \, (I - M_l) \, \mathbf{u}_l = (I - M_l) \, (\mathbf{f}_l - \mathbf{A}_l M_l \mathbf{g}_l) \, . \end{cases}$$

Here, $\mathsf{R}(I - M_l)$ denotes the range of $I - M_l$. The solution $\tilde{u}_l$ of (3.8) is now given by

$$\tilde{u}_l := \sum_{k \in \mathscr{E}_l \cup \mathscr{I}_l} (\mathbf{u}_l)_k \varphi_k^l + \sum_{k \in \mathscr{B}_l} (\mathbf{g}_l)_k \varphi_k^l \, .$$

Krylov space methods are well suited for the numerical solution of the system (5.3), for instance, the conjugate gradient method and its modifications for non-symmetric problems. These iterative schemes have only to be restricted to the subspace $\mathsf{R}(I - M_l)$ of $\mathbb{R}^{\mathbb{Z}^{d,l}}$ which can be realized easily enough. Thus, we can rely on the simple ordering of the (Cartesian) index set $\mathbb{Z}^{d,l}$. This reduces the coding effort as well as the run time considerably since the simple data structures allow fast memory access times.

The code can be designed independently of the geometry of the domain Ω as far as M_l is considered as an input to the program. This is a further advantage of our fictitious domain approach. In Section 5.2 we will see how to extract the matrix M_l from a discrete version of the characteristic function of Ω.

In the sequel we identify the equivalence classes in $\mathbb{Z}^{d,l}$ with their representers in $\{0, ..., 2^l - 1\}^d$ and vice versa.

Next we will resolve the structure of $\mathbf{A}_l$. For convenience we assume that the coefficients A and α of the bilinear form (2.1) are extended periodically w.r.t. $\partial\square$ by $\tilde{A}$ and $\tilde{\alpha}$, cf. Section 2.

In computing $(\mathbf{A}_l)_{k,r}$ we typically have to deal with integrals like

$$c_{k,r}^{i,j} := \int_\square c(x) \, D^{e_i} \varphi_k^l(x) \, D^{e_j} \varphi_r^l(x) \, dx$$

where c is periodic due to our assumption from above, i.e., $c(\, . + n) = c(\, . \,)$ for all $n \in \mathbb{Z}^d$. By e_i we denote the i-th canonical unit vector in $\mathbb{R}^d$. We rewrite the integrals $c_{k,r}^{i,j}$ using the definition (3.4) of the φ_k^l's, the periodicity of c and letting $\kappa := k - r \in \mathbb{Z}^d$:

$$c_{k,r}^{i,j} = \delta_l^{-2} \sum_{p \in \mathbb{Z}^d} \underbrace{\int_{\mathbb{R}^d} c(\delta_l(z + r)) \, D^{e_i} \varphi(z + 2^l p - \kappa) \, D^{e_j} \varphi(z) \, dz}_{:= \eta(p, \kappa)} \, .$$

Let $\text{diam}_\infty \mathcal{M} := \max\left\{ \|y - z\|_\infty \mid y, z \in \mathcal{M} \right\}$ be the diameter of the set $\mathcal{M} \subset \mathbb{R}^d$ measured w.r.t. the maximum norm $\| \cdot \|_\infty$. We set $S = \text{supp}\,\varphi$.

Since $(\overset{\circ}{S} + q) \cap \overset{\circ}{S} = \emptyset$ if $q \in \mathbb{R}^d$ with $\|q\|_\infty \geq \text{diam}_\infty S$ we have that $\eta(p, k - r) = 0$ for $\|k - r - 2^l p\|_\infty \geq \text{diam}_\infty S$. Therefore we are able to restrict the range of p in the above representation of $c_{k,r}^{l,J}$:

$$(5.4) \qquad\qquad c_{k,r}^{l,J} = \delta_l^{-2} \sum_{p \in \mathcal{P}(k-r,l)} \eta(p, k - r)$$

where $\mathcal{P}(\kappa, l) := \{ p \in \mathbb{Z}^d \mid \|\kappa - 2^l p\|_\infty < \text{diam}_\infty S \}$.

LEMMA 5.1: *Let $l \in \mathbb{N}$ be such large that $2^l \geq 2\,\text{diam}_\infty S$. Then, the set $\mathcal{P}(\kappa, l)$, $\kappa \in \mathbb{Z}^d$, contains one element at most.*

Proof: Let p_1 and p_2 be in $\mathcal{P}(\kappa, l)$. The estimate

$$2^l \|p_1 - p_2\|_\infty \leq \|2^l p_1 - \kappa\|_\infty + \|2^l p_2 - \kappa\|_\infty < 2\,\text{diam}_\infty S$$

implies $\|p_1 - p_2\|_\infty < 2^{-l}\,2\,\text{diam}_\infty S \leq 1$. Hence, $p_1 = p_2$. $\qquad\square$

COROLLARY 5.2: *Adopt the assumptions of Lemma 5.1. Let $\text{diam}_\infty S \geq 1$. Further, let k and r be in $\mathbb{Z}^{d,l}$ such that for any $m \in \{1, ..., d\}$ either $|k_m - r_m| < \text{diam}_\infty S$ or $2^l - \text{diam}_\infty S < |k_m - r_m| \leq 2^l - 1$ holds true. Then the set $\mathcal{P}(k - r, l)$ contains only the element p with components*

$$(5.5) \qquad\qquad p_m = \begin{cases} 0 & : \quad |k_m - r_m| < \text{diam}_\infty S \\ -\text{sgn}\,(k_m - r_m) & : \quad 2^l - \text{diam}_\infty S < |k_m - r_m| \leq 2^l - 1 \end{cases}.$$

Proof: Let $\kappa = k - r$. We claim that $|\kappa_m - 2^l p_m| < \text{diam}_\infty S$ for all $m \in \{1, ..., d\}$. For $|\kappa_m| < \text{diam}_\infty S$ our claim follows by $p_m = 0$. Now, let $2^l - \text{diam}_\infty S < |\kappa_m| \leq 2^l - 1$ and let κ_m be positive, i.e., $p_m = -1$. This implies $-\text{diam}_\infty S < \kappa_m + 2^l p_m \leq -1$ which gives $|\kappa_m + 2^l p_m| < \text{diam}_\infty S$. In the same manner one deals with $\kappa_m < 0$. The set $\mathcal{P}(k - r, l)$ contains p (5.5) and p is its only element due to Lemma 5.1. $\qquad\square$

Because the basis functions have a local support, the stiffness matrix $\mathbf{A}_l$ is sparse, of course. The indices of the zero entries are known *a-priori*.

LEMMA 5.3: *Let k and r be in $\mathbb{Z}^{d,l}$ with $\text{diam}_\infty S \leq |k_m - r_m| \leq 2^l - \text{diam}_\infty S$ for some $m \in \{1, ..., d\}$. Then, $c_{k,r}^{l,J} = 0$ as well as $(\mathbf{A}_l)_{k,r} = 0$.*

Proof: We will show that $\mathcal{P}(\kappa, l)$, $\kappa = k - r$, is the empty set. For this we assume there is some $p \in \mathcal{P}(\kappa, l)$. Since $\text{diam}_\infty S \leq |\kappa_m|$ we have that $|p_m| \geq 1$. However,

$$|\kappa_m - 2^l p_m| \geq |\,|\kappa_m| - 2^l |p_m|\,| = 2^l |p_m| - |\kappa_m| \geq \text{diam}_\infty S$$

tells us that $p \notin \mathcal{P}(\kappa, l)$ which contradicts $p \in \mathcal{P}(\kappa, l)$. Thus, $\mathcal{P}(\kappa, l) = \emptyset$. $\qquad\square$

In assembling the stiffness matrix $\mathbf{A}_l$, integrals of the kind (5.4) have to be evaluated. The efficient computation of such integrals has been studied by Dahmen and Micchelli [11] and by Latto, Resnikoff and Tenenbaum [28]. A code realizing the approach of Dahmen and Micchelli has been written by Kunoth [26].

5.2. Classification of $\mathscr{B}_l$

We concentrate on the generation of the matrix M_l, that is, we classify the indices in $\mathscr{B}_l$ (3.6). We will fail in a correct classification which is numerical realizable. However, we present a good approximation to $\mathscr{B}_l$ which becomes better as the discretization step size decreases.

Let K be the ball w.r.t. the maximum norm in $\mathbb{R}^d$ with smallest possible radius containing the support S of the scaling function φ. Let ζ denote the center of K. We define the approximation

$$(5.6) \qquad \chi_\Omega^l := \sum_{k \in \mathbb{Z}^{d,l}} \chi_\Omega(\zeta_{l,k}) \, \varphi_k^l$$

to the characteristic function χ_Ω of the domain Ω. Note that $\zeta_{l,k} := \delta_l(k + \zeta)$ is the center of $K_{l,k} := \delta_l(k + K)$. We have the periodization

$$(5.7) \qquad \zeta_{l,k} = \zeta_{l,k+2^l q}, \quad q \in \mathbb{Z}^d, k \in \mathbb{Z}^{d,l}.$$

The boundary of Ω coincides with support of the gradient $\nabla \chi_\Omega$ of χ_Ω. This motivate us to define the approximation $\widehat{\mathscr{B}}_l$ to $\mathscr{B}_l$ by

$$\widehat{\mathscr{B}}_l := \left\{ r \in \mathbb{Z}^{d,l} \mid \text{there is some } i \in \{1, ..., d\} \text{ with } \langle D^{e_i} \chi_\Omega^l, \varphi_r^l \rangle_{L^2(\square)} \neq 0 \right\}.$$

The elements in $\widehat{\mathscr{B}}_l$ can be determined easily. We have

$$(\mathbf{d}_l^i)_r := \langle D^{e_i} \chi_\Omega^l, \varphi_r^l \rangle_{L^2(\square)} = \sum_{k \in \mathbb{Z}^{d,l}} \chi_\Omega(\zeta_{l,k}) \langle D^{e_i} \varphi_k^l, \varphi_r^l \rangle_{L^2(\square)}.$$

In Section 5.1 we have already studied integrals like $\langle D^{e_i} \varphi_k^l, \varphi_r^l \rangle_{L^2(\square)}$. Analogously,

$$(5.8) \qquad \langle D^{e_i} \varphi_k^l, \varphi_r^l \rangle_{L^2(\square)} = \delta_l^{-1} \int_S D^{e_i} \varphi(z + 2^l p - (k - r)) \, \varphi(z) \, dz =: \delta_l^{-1} \Gamma_{k-r}^l,$$

where $p = p(k, r)$ is as in (5.5). With the sparse matrix $(\Gamma_l^i)_{k,r} = \delta_l^{-1} \Gamma_{k-r}^l$ and the vector $(\chi_l)_k = \chi_\Omega(\zeta_{l,k})$ we get $\mathbf{d}_l^i = \Gamma_l^i \chi_l$, $i \in \{1, ..., d\}$. Fast matrix-vector multiplications enable us to compute the vectors $\mathbf{d}_l^i$ which, in turn, give the set $\widehat{\mathscr{B}}_l$.

So far the connection between $\mathscr{B}_l$ and $\widehat{\mathscr{B}}_l$ remained unsettled. We formulate a first result in Lemma 5.4. Let $\mathrm{dist}_\infty(\mathscr{N}, \mathscr{M})$ be the distance of the sets $\mathscr{N}, \mathscr{M} \subset \mathbb{R}^d$ measured in the maximum norm.

LEMMA 5.4: *Let the order of the scaling function be 1 at least. Let $l \in \mathbb{N}$ be such large that $2^l > \max\{2, 1/\mathrm{dist}(\partial\square, \partial\Omega)\} \mathrm{diam}_\infty S$. If $\mathrm{dist}_\infty(\zeta_{l,r}, \partial\Omega) \geq \delta_l \mathrm{diam}_\infty S$ for $r \in \mathbb{Z}^{d,l}$, then r is neither in $\mathscr{B}_l$ nor in $\widehat{\mathscr{B}}_l$.*

Proof: The hypothesis $2^l > \max\{2, 1/\mathrm{dist}(\partial\square, \partial\Omega)\} \mathrm{diam}_\infty S$ allows, on one side, the application of Corollary 5.2, that is, the representation (5.8) is valid with p from (5.5). On the other side, it also guarantees that the part of the support of φ_r^l, which is in $\square$, does not intersect the boundaries of $\square$ and Ω simultaneously.

First, we will show that $r \notin \mathscr{B}_l$. We assume both: $r \in \mathscr{B}_l$ as well as $\mathrm{dist}_\infty(\zeta_{l,r}, \partial\Omega) \geq \delta_l \mathrm{diam}_\infty S$. That part of the support of φ_r^l, which lies in $\square$, coincides with $\overset{\circ}{S}_{l,r}$. The boundary of Ω intersects $\overset{\circ}{S}_{l,r}$. Therefore, $\mathrm{dist}(\zeta_{l,r}, \partial\Omega) < \delta_l(\mathrm{diam}_\infty S)/2$ which contradicts the assumption.

Finally we show that $r \notin \widehat{\mathscr{B}}_l$. To do so, we define $U_r := \{k \in \mathbb{Z}^{d,l} \mid \overset{\circ}{S}_{l,k-2^l p} \cap \overset{\circ}{S}_{l,r} \neq \emptyset\}$ with $p = p(k, r)$ from (5.5). We have that

$$(5.9) \qquad \|\zeta_{l,k-2^l p} - \zeta_{l,r}\|_\infty < \delta_l \mathrm{diam}_\infty S \quad \text{for} \quad k \in U_r.$$

(Otherwise, both open balls $\overset{\circ}{K}_{l,r}$ and $\overset{\circ}{K}_{l,k-2^l p}$ intersect, which implies that $\overset{\circ}{S}_{l,k-2^l p} \cap \overset{\circ}{S}_{l,r} = \Phi$. However this contradicts $k \in U_r$). It follows from inequality (5.9) that the centers $\zeta_{l,k-2^l p}$, $k \in U_r$, are located on the same side of the boundary of Ω, i.e., $\chi_\Omega(\zeta_{l,r}) = \chi_\Omega(\zeta_{l,k-2^l p})$ for all $k \in U_r$. Now,

$$(\mathbf{d}_l^t)_r \overset{(5.7)}{=} \delta_l^{-1} \sum_{k \in U_r} \Gamma_{k-r}^l \chi_\Omega(\zeta_{l,k-2^l p}) = \delta_l^{-1} \chi_\Omega(\zeta_{l,r}) \sum_{k \in U_r} \Gamma_{k-r}^l.$$

Since $\sum_{k \in U_r} \Gamma_{k-r}^l = 0$, which follows by (3.3), we obtain $r \notin \widehat{\mathscr{B}}_l$. $\square$
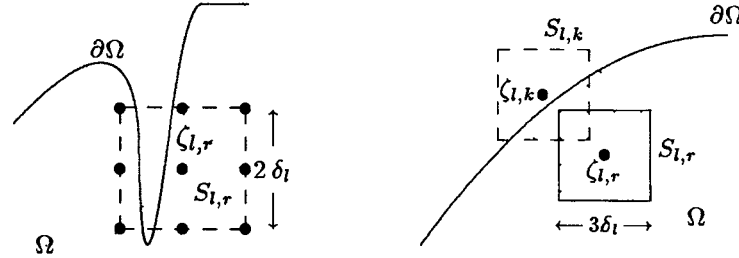

Figure 5.1. — Left: sketch of a situation in which r belongs to $\mathscr{B}_l$ but not to $\widehat{\mathscr{B}}_l$. The dashed line is the boundary of $S_{l,r}$ Marked are all centers $\zeta_{l,k}$ for $k \in U_r$ (here $p = 0$). Since all centers lie in Ω we have that $d_i^l = 0$ for all i, i.e., $r \notin \widehat{\mathscr{B}}_l$
Right: sketch of a situation in which r does not belong to $\mathscr{B}_l$ but to $\widehat{\mathscr{B}}_l$. The support $S_{l,r}$ is marked (solid line). Also we have drawn the support $S_{l,k}$ (dashed line) which intersects $S_{l,r}$ and the center of which lies outside Ω All other centers $\zeta_{l,m}$, $m \in U_r \backslash\{k\}$, lie in the interior of Ω Thus, there is one $(d_i^l)_r \neq 0$ which implies $r \in \widehat{\mathscr{B}}_l$.

Lemma 5.4 has to be interpreted in the following way. The indices r in $\widehat{\mathscr{B}}_l$ belong to basis functions φ_r^l whose supports are located near the boundary $\partial\Omega$. Analytically, this means that dist $(\zeta_{l,r}, \partial\Omega) < \delta_l \text{diam}_\infty S$ for $r \in \widehat{\mathscr{B}}_l$. Although dist $(\zeta_{l,r}, \partial\Omega) < \delta_l(\text{diam}_\infty S)/2$ for $r \in \mathscr{B}_l$, $\widehat{\mathscr{B}}_l$ is *not* a superset of $\mathscr{B}_l$. There are situations in which indices of $\mathscr{B}_l$ are not in $\widehat{\mathscr{B}}_l$ and vice versa. Such situations are sketched in *figure* 5.1. For convenience we considered the tensor product case. The scaling function φ is a d-fold tensor product of a univariate scaling function. Hence, $S = K$ (K is the smallest box containing S)

From a numerical point of view, the pathological situations of *figure* 5.1 pose no difficulties. The error caused by the wrong classification according to the left constellation has the same order of magnitude as the discretization error. Structures of the boundary which are smaller than the discretization step size cannot be resolved in principal. In constellations depicted on the right we get some boundary indices too many. In the worst case the width of the strip, where the boundary values are forced, is doubled from $\delta_l(\text{diam}_\infty S)/2$ to $\delta_l \text{diam}_\infty S$.

All things considered we have seen that $\widehat{\mathscr{B}}_l$ is a numerical realizable approximation to $\mathscr{B}_l$ which produces an "approximation error" having the same order of magnitude than the discretization error.

5.3. Numerical experiments

On the basis of two 3D-examples we give an impression on the mode of action of the proposed algorithm. We consider

(5.10a) $$-\alpha \Delta u + u = f \quad \text{in} \quad \Omega,$$

(5.10b) $$u = g \quad \text{on} \quad \partial\Omega,$$

where α is a positive constant and $\Delta = \sum_{i=1}^{d} D^{2e_i}$ is the Laplace operator. The underlying domain is

$$\Omega := \left\{ x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 < 1/16, x_3 < \sqrt{x_1^2 + x_2^2} \right\}.$$

The domain Ω has a re-entrant corner. It can be expressed by $\Omega = B \backslash \mathcal{K}$ where B is the ball with center in the origin and with radius 0.25. By $\mathcal{K}$ we denote the right circular cone whose vertex is the origin and which opens in the direction of the positive x_3-axis with angle $\pi/2$. We use the cube $\Box =] - 0.3, 0.3[^3$ as fictitious domain.

Let Φ be a univariate and weak differentiable scaling function with compact support. We get the scaling function φ by

$$\varphi(x) = \varphi(x_1, x_2, x_3) := \Phi(x_1)\, \Phi(x_2)\, \Phi(x_3)\,.$$

The periodization φ_k^l of $\varphi_{l,k}$ is 1-periodic. By the re-scaling $\check{\varphi}_k^l(x) := \varphi_k^l(x/0.6)$, $\check{\varphi}_k^l$ becomes 0.6-periodic. As approximation space we accordingly choose

$$(5.11) \qquad V_l := span\{\check{\varphi}_k^l \mid k \in \mathbb{Z}^{3,l}\} \subset H_p^1(]-0.3, 0.3[^3)\,.$$

In all our experiments Φ is the linear B-spline and the discretization level is $l = 7$. Hence, $0.6 \cdot \delta_7 = 0.6/128$ is the discretization step size. The extensions of the different right-hand sides, the boundary values and of the coefficients in (5.10) will be obvious because they will be defined everywhere in $\mathbb{R}^3$.
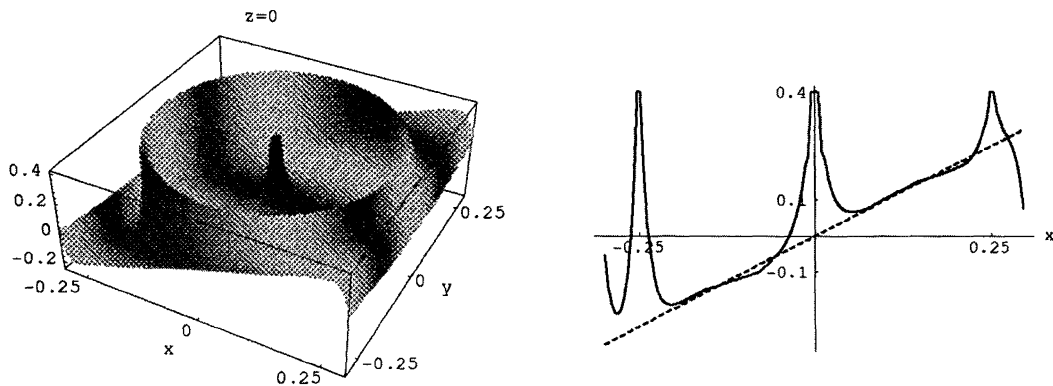

Figure 5.2. — Cross sections of the numerical solution $\tilde{u}_7$ of (5.10) with $a = 10^{-4}$, $f(x) = x_1 + x_2^2 + x_3^2$, and $g = 0.4$. **Left:** $\tilde{u}_7(\,.\,,.\,,0\,)$, **right:** $\tilde{u}_7(\,.\,,0,0\,)$ (solid line) and $f(\,.\,,0,0\,)$ (dashed line).

Firstly, we solve (5.10) with $\alpha = 10^{-4}$, $f(x) = x_1 + x_2^2 + x_3^2$, and $g = 0.4$. *Figure* 5.2 displays the cross sections $\tilde{u}_7(\,.\,,.\,,0\,)$ and $\tilde{u}_7(\,.\,,0,0\,)$. Note that the intersection of Ω with the plane $x_3 = 0$ is a circular disk with radius 0.25 punctured at the origin. Since α is very small, u and hence $\tilde{u}_7$ approximate f inside Ω, see *figure* 5.2 (right).

The two diagrams in *figure* 5.2 demonstrate clearly that the boundary constraints are forced not only on $\partial\Omega$ but also on the strip $\partial\Omega^l$ (4.22).

Figures 5.3 and 5.4 show different cross sections of the numerical approximation $\tilde{u}_7$ to the exact solution of (5.10) with $\alpha = 1$, $f = 1$, and $g = 0$.

Both graphs in *figure* 5.3 belong to the cross sections $\tilde{u}_7(\,.\,,.\,,-0.12\,)$ and $\tilde{u}_7(\,.\,,.\,,0.12\,)$. The intersection of Ω with the plane $x_3 = 0.12$ is a circular ring centered in the origin with inner radius 0.12 and outer radius 0.22. The approximation $\tilde{u}_7(\,.\,,.\,,-0.12\,)$ has to be zero on both boundaries of that circular ring.
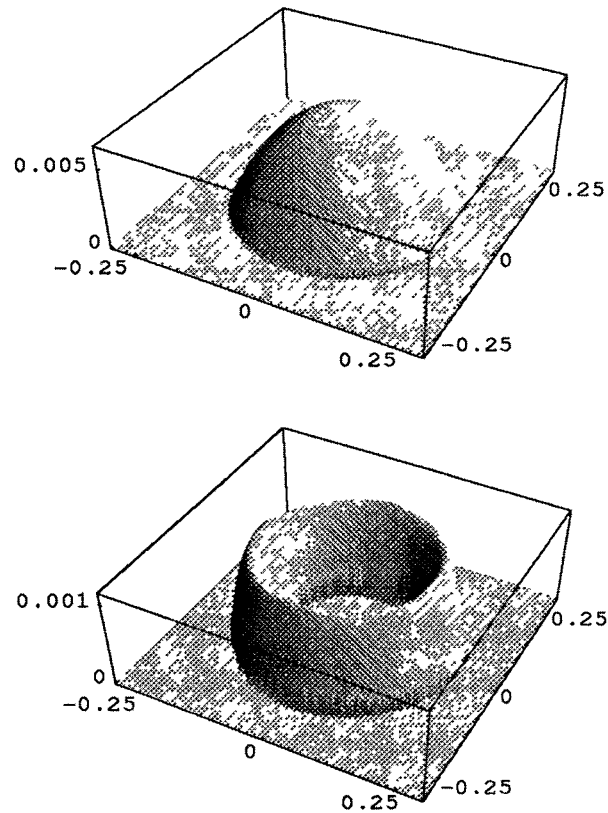
Figure 5.3. — Cross sections of the numerical solution $\bar{u}_7$ of (5.10) with $a = 1$, $f = 1$, and $g = 0$ Top: $\bar{u}_7(\,.\,,.\,,-\,0\,12\,)$, bottom: $\bar{u}_7(\,.\,,.\,,0\,12\,)$.

The diagrams of *figure* 5 4 show the graphs of $\bar{u}_7(0,\,.\,,.\,)$ and of $\bar{u}_7(\pm\,0\,12,\,.\,,.\,)$ Here we do not have "inner" boundaries The right circular cone $\mathscr{K}$ becomes noticeable in the missing radial symmetry of the domains of intersection

6. PRECONDITIONING

As already explained in the former section Krylov space methods are well suited for the iterative solution of the linear system (5.3) In this section we propose a preconditioner for the method of conjugate gradients (cg) applied to (5 3) We refer to, e g , Deuflhard and Hohmann [15] and Hackbusch [23] for an introduction to the cg-iteration and the concept of preconditioning

6.1. Theoretical studies

In the sequel we will assume the positive-definiteness of $\mathbf{A}_l$, that is, $\mathbf{A}_l$ is symmetric and has only positive eigenvalues For instance, this assumptions holds true when $\bar{A}$ and $\tilde{\alpha}$ are as in Lemma 2 2 and when the coefficient matrix $\bar{A}$ of the form $\tilde{a}$ (2 6) is additionally symmetric

We identify $\mathbb{R}^{\mathbb{Z}^{d,\,l}}$ with $\mathbb{R}^{n_l}$ where $n_l = 2^{dl}$. Thus, $\mathbf{A}_l \in \mathbb{R}^{n_l \times n_l}$. For writing convenience we set $\hat{\mathbf{A}}_l := (I - M_l)\,\mathbf{A}_l\,(I - M_l)$. This matrix, restricted to the subspace $\mathsf{R}(I - M_l)$ of $\mathbb{R}^{n_l}$, is also positive definite. So we may apply the cg-iteration to (5.3). It is well known that the performance of the cg-iteration deteriorates when the spectral condition number of $\hat{\mathbf{A}}_l$ gets larger.
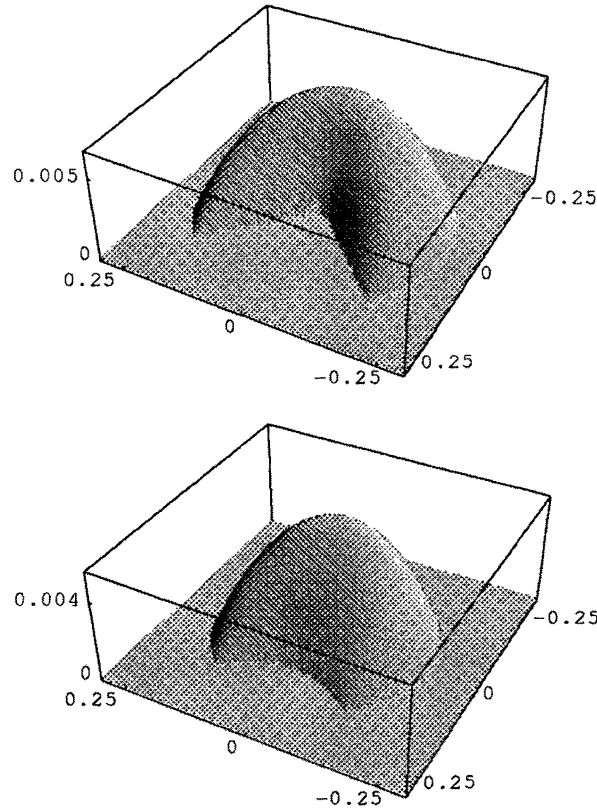

Figure 5.4. — Cross sections of the numerical solution $\bar{u}_7$ of (5.10) with $a = 1$, $f = 1$, and $g = 0$. Top: $\bar{u}_7(0,.,.)$, bottom: $\bar{u}_7(\pm\,0.12,.,.)$.

By standard techniques used for instance in the finite element theory, see, e.g., Hackbusch [22], one can show that $\mathrm{cond}_{\mathsf{R}(I - M_l)}(\hat{\mathbf{A}}_l) = \mathcal{O}(\delta_l^{-2})$ as $l \to \infty$. (We use the notation $\mathrm{cond}_X(T)$ for the spectral condition number of the automorphism T on the finite-dimensional vector space X.) Instead of the ill-conditioned system (5.3) we rather choose to solve the equivalent symmetric system

$$(6.1) \qquad \hat{\mathbf{W}}_l^{1/2}\hat{\mathbf{A}}_l\hat{\mathbf{W}}_l^{1/2}\,z_l = \hat{\mathbf{W}}_l^{1/2}(\mathbf{f}_l - \mathbf{A}_l M_l\,\mathbf{g}_l)\,.$$

In (6.1), $\hat{\mathbf{W}}_l := (I - M_l)\,\mathbf{W}_l\,(I - M_l)$ where the positive definite matrix $\mathbf{W}_l \in \mathbb{R}^{n_l \times n_l}$ is an *optimal* preconditioner for $\mathbf{A}_l$, that is,

$$(6.2) \qquad \mathrm{cond}_{\mathbb{R}^{n_l}}(\mathbf{W}_l\mathbf{A}_l) = \mathcal{O}(1) \quad \text{as} \quad l \to \infty\,.$$

Interpreting $\mathbf{A}_l$ as stiffness matrix of a Galerkin scheme with approximation space V_l^p applied to the periodic variational problem

$$\text{find } v \in H_p^1(\square) \text{ such that } \tilde{a}(v, w) = \int_\square \tilde{f}w \, dx \text{ holds true for all } w \in H_p^1(\square),$$

we immediately find a variety of optimal preconditioners for $\mathbf{A}_l$. For instance, we have the *BPX*-preconditioner, see Bramble, Pasciak and Xu [5], wavelet preconditioners, see Jaffard [25] and Dahmen and Kunoth [10], as well as multi-grid preconditioners, see, e.g., Hackbusch [23].

In the remainder of this section we will verify that the condition number of $\hat{\mathbf{W}}_l^{1/2}\hat{\mathbf{A}}_l\hat{\mathbf{W}}_l^{1/2}$ behaves nicer than the one of $\hat{\mathbf{A}}_l$ as $l \to \infty$. To this end we provide positive numbers γ and Γ such that

$$(6.3) \qquad\qquad \gamma\hat{\mathbf{W}}_l^{-1} \leq \hat{\mathbf{A}}_l \leq \Gamma\hat{\mathbf{W}}_l^{-1}.$$

Then, $\text{cond}_{R(I - M_l)}(\hat{\mathbf{W}}_l^{1/2}\hat{\mathbf{A}}_l\hat{\mathbf{W}}_l^{1/2}) \leq \Gamma/\gamma$. The notation $C \leq D(C < D)$ signifies that the matrix $D - C$ is positive semi definite (positive definite).

The following lemma connects the quality of the preconditioner $\mathbf{W}_l$ for $\mathbf{A}_l$ with the quality of the induced preconditioner $\hat{\mathbf{W}}_l$ for $\hat{\mathbf{A}}_l$. Its obvious proof is omitted.

LEMMA 6.1: *Suppose there are positive numbers θ, Θ, λ and Λ such that*

$$(6.4) \qquad\qquad \theta\mathbf{W}_l^{-1} \leq \mathbf{A}_l \leq \Theta\mathbf{W}_l^{-1} \quad and \quad \lambda(\widehat{\mathbf{A}_l^{-1}})^{-1} \leq \hat{\mathbf{A}}_l \leq \Lambda(\widehat{\mathbf{A}_l^{-1}})^{-1}.$$

Then, the inequality (6.3) is satisfied with $\gamma = \theta \cdot \lambda$ and $\Gamma = \Theta \cdot \Lambda$.

Before we are able to judge the quality of the preconditioner $\hat{\mathbf{W}}_l$ via Lemma 6.1 we have to supply positive numbers λ and Λ for the two-sided inequality on the right side in (6.4). This will take the rest of the section.

For our further investigations we order $\mathbb{Z}^{d,l}$ according to its disjoint splitting $\mathbb{Z}^{d,l} = \mathscr{I}_l \cup \mathscr{E}_l \cup \mathscr{B}_l$, see Section 4. We arrange the indices belonging to $\mathscr{I}_l \cup \mathscr{E}_l$ in the first place and then those indices belonging to $\mathscr{B}_l$. The matrices $\mathbf{A}_l$ and $\mathbf{A}_l^{-1}$ now have the block structure

$$(6.5) \qquad\qquad \mathbf{A}_l = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^t & \mathbf{A}_{22} \end{pmatrix}, \quad \mathbf{A}_l^{-1} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{12}^t & \mathbf{B}_{22} \end{pmatrix},$$

where the diagonal blocks are square matrices of dimensions $m_1 = |\mathscr{I}_l \cup \mathscr{E}_l|$ and $m_2 = |\mathscr{B}_l|$, respectively. Both matrices $\hat{\mathbf{A}}_l$ and $\widehat{\mathbf{A}_l^{-1}}$ can be expressed by $\hat{\mathbf{A}}_l = \mathbf{A}_{11}$ and $\widehat{\mathbf{A}_l^{-1}} = \mathbf{B}_{11}$. Writing $\mathbf{A}_l^{-1}\mathbf{A}_l = I$ in the above block form we see that

$$\mathbf{B}_{11}\mathbf{A}_{11} + \mathbf{B}_{12}\mathbf{A}_{12}^t = I \quad \text{as well as} \quad \mathbf{B}_{12}^t\mathbf{A}_{11} + \mathbf{B}_{22}\mathbf{A}_{12}^t = 0.$$

These relations firstly imply $\mathbf{B}_{11}\mathbf{A}_{11} = I + \mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}\mathbf{A}_{12}^t$ and then

$$(6.6) \qquad\qquad \mathbf{A}_{11}^{1/2}\mathbf{B}_{11}\mathbf{A}_{11}^{1/2} = I + \mathbf{A}_{11}^{-1/2}\mathbf{A}_{12}\mathbf{B}_{22}\mathbf{A}_{12}^t\mathbf{A}_{11}^{-1/2} > I$$

which yields $\lambda = 1$ in (6.4). We are now going to estimate Λ. For that we rely on (6.6) again. Let Y be the largest eigenvalue of the product $\mathbf{A}_{22}^{1/2}\mathbf{B}_{22}\mathbf{A}_{22}^{1/2}$, that is, $\mathbf{B}_{22} \leq Y\mathbf{A}_{22}^{-1}$. Plugging the latter inequality into (6.6) yields $\mathbf{A}_{11}^{1/2}\mathbf{B}_{11}\mathbf{A}_{11}^{1/2} = I + Y\mathbf{A}_{11}^{-1/2}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{12}^t\mathbf{A}_{11}^{-1/2}$.

LEMMA 6.2: *We have that $\mathbf{A}_{11}^{-1/2}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{12}^t\mathbf{A}_{11}^{-1/2} < I$.*

Proof: The assertion follows if the Schur complement $S := \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{12}^t$ is positive definite. Since $S = \mathbf{B}_{11}^{-1}$ and $\mathbf{B}_{11}^{-1} > 0$ we are done. $\square$

We summarize our provisional result in the next lemma.

LEMMA 6.3: *Let Y be the largest Eigenvalue of $\mathbf{A}_{22}^{1/2}\,\mathbf{B}_{22}\,\mathbf{A}_{22}^{1/2}$, then*

$$(\widehat{\mathbf{A}_l^{-1}})^{-1} \leqslant \hat{\mathbf{A}}_l < (1+Y)\,(\widehat{\mathbf{A}_l^{-1}})^{-1}\,.$$

So far we managed with purely algebraic techniques. The following estimate of Y is based on a representation which allows the application of analytical tools. Properties of the underlying bilinear form $\tilde{a}$ (2.6) enter our investigations crucially.

The eigenvalue Y is the maximal value of the corresponding Rayleigh quotient:

$$Y = \max_{0 \neq \xi \in \mathbb{R}^{m_2}} \langle \mathbf{A}_{22}\,\xi, \xi \rangle_{\mathbb{R}^{m_2}} / \langle \mathbf{B}_{22}^{-1}\,\xi, \xi \rangle_{\mathbb{R}^{m_2}}$$

We have that

$$(6.7) \qquad \langle \mathbf{A}_{22}\,\xi, \xi \rangle_{\mathbb{R}^{m_2}} = \tilde{a}(\gamma^l(v_l), \gamma^l(v_l))$$

for all $v_l \in V_l^p$ with $\gamma^l(v_l) = \sum_{k \in \mathcal{B}_l} \xi_{\gamma(k)}\,\varphi_k^l$. Here, γ^l is the discrete trace operator defined in (3.7) and $\gamma : \mathcal{B}_l \to \{1, ..., m_2\}$ is the chosen ordering of $\mathcal{B}_l$.

Let $\xi \in \mathbb{R}^{m_2}$. We introduce the discrete variational problem (6.8),

$$(6.8) \qquad \begin{cases} \text{find } v_l \in V_l^p \text{ with } \gamma^l(v_l) = \sum_{k \in \mathcal{B}_l} \xi_{\gamma(k)}\,\varphi_k^l \text{ such that} \\ \tilde{a}(v_l, w_l) = 0 \text{ holds true for all } w_l \in V_l^p \text{ with } \gamma^l(w_l) = 0\,. \end{cases}$$

LEMMA 6.4: *Let $v_l \in V_l^p$ be the unique solution of (6.8) w.r.t. $\xi \in \mathbb{R}^{m_2}$. Then,*

$$(6.9) \qquad \langle B_{22}^{-1}\,\xi, \xi \rangle_{\mathbb{R}^{m_2}} = \tilde{a}(v_l, v_l)\,.$$

Proof: We expand v_l by $v_l = \sum_{k \in \mathcal{I}_l \cup \mathcal{E}_l} \eta_{\iota(k)}\,\varphi_k^l + \sum_{k \in \mathcal{B}_l} \xi_{\gamma(k)}\,\varphi_k^l$ with one $\eta \in \mathbb{R}^{m_1}$. Here, $\iota : \mathcal{I}_l \cup \mathcal{E}_l \to \{1, ..., m_1\}$ is the chosen ordering of $\mathcal{I}_l \cup \mathcal{E}_l$. We have

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{12}^t & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \eta \\ \xi \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$$

for one $b \in \mathbb{R}^{m_2}$. This b will now be determined. Because $\mathbf{A}_{11}\,\eta = -\mathbf{A}_{12}\,\xi$ we get $b = \mathbf{A}_{12}^t\,\eta + \mathbf{A}_{22}\,\xi = (\mathbf{A}_{22} - \mathbf{A}_{12}^t\,\mathbf{A}_{11}^{-1}\,\mathbf{A}_{12})\,\xi$. Further, $\mathbf{B}_{22}^{-1} = \mathbf{A}_{22} - \mathbf{A}_{12}^t\,\mathbf{A}_{11}^{-1}\,\mathbf{A}_{12}$ which can be seen easily. Finally,

$$\tilde{a}(v_l, v_l) = \left\langle \begin{pmatrix} 0 \\ b \end{pmatrix}, \begin{pmatrix} \eta \\ \xi \end{pmatrix} \right\rangle_{\mathbb{R}^{n_l}} = \langle \mathbf{B}_{22}^{-1}\,\xi, \xi \rangle_{\mathbb{R}^{m_2}}$$

which ends the proof of Lemma 6.4. $\square$

Taking (6.7) and (6.9) into account we obtain

$$(6.10) \qquad Y = \max \{ \tilde{a}(\gamma^l(v_l), \gamma^l(v_l))/\tilde{a}(v_l, v_l) \mid 0 \neq v_l \in V_l^p, \tilde{a}(v_l, w_l) = 0 \text{ for all } w_l \in V_l^p \text{ with } \gamma^l(w_l) = 0 \}\,.$$

Under the assumptions of Lemma 2.2 the bilinear form $\tilde{a}$ is $H_p^1(\square)$-elliptic and continuous. Thus, there exists a positive constant $C_{\tilde{a}}$ such that

$$Y \leqslant C_{\tilde{a}} \max_{0 \neq v_l \in V_l^p} \frac{\|\gamma^l(v_l)\|_{H^1(\square)}^2}{\|v_l\|_{H^1(\square)}^2} \leqslant C_{\tilde{a}} C_B^2 \delta_l^{-2} \max_{0 \neq v_l \in V_l^p} \frac{\|\gamma^l(v_l)\|_{L^2(\partial\Omega^l)}^2}{\|v_l\|_{H^1(\square)}^2}$$

where the latter inequality is due to the Bernstein estimate (4.6) which holds correspondingly for the periodic spaces V_l^p, see Dahmen, Prößdorf and Schneider [12, Theorem 5.1]. The further procedure is based on the assumption

$$(6.11) \qquad \|\gamma^l(v_l)\|_{L^2(\partial\Omega^l)} \leqslant C_V \|v_l\|_{L^2(\partial\Omega^l)} \quad \text{for all} \quad v_l \in V_l^p$$

where the positive constant C_V does not depend on l or v_L. In Section 6.2 we will verify (6.11) for the tensor product approach.

In the next step we estimate $\|v_l\|_{L^2(\partial\Omega^l)}$. For that we require some properties of Ω which are of technical nature. Essentially, the boundary of Ω should be the zero set of a continuously differentiable function. Ellipsoidal domains, especially ball-shaped domains, satisfy the hypotheses of Lemma 6.5.

LEMMA 6.5: *Suppose there is a function $F : \square \to \mathbb{R}$ which is in $\mathscr{C}^1(\square)$ and whose gradien thas no zeros in $\square$. Furthermore, let there exist a $\beta > 0$ such that the sets $\Omega_\alpha := \{x \in \square \mid F(x) < \alpha\}$ are bounded domains with $\overline{\Omega_\alpha} \subset \square$ for $\alpha \in [-\beta, \beta]$. The domain Ω of the boundary value problem (1.1) coincide with $\Omega_0: \Omega = \Omega_0$.*

If l is sufficiently large then there is a positive constant C_Ω so that

$$\|v_l\|_{L^2(\partial\Omega^l)} \leqslant C_\Omega \delta_l^{1/2} \|v_l\|_{H^1(\square)} \quad \text{for all} \quad v_l \in V_l^p .$$

Proof: Let l be such large that $\partial\Omega^l \subset \Omega_\beta \backslash \overline{\Omega}_{-\beta}$. We define $\alpha_l^+ := \max \mathscr{N}_l$ and $\alpha_l^- := \min \mathscr{N}_l$ where $\mathscr{N}_l = \{\alpha \in [-\beta, \beta] \mid$ there is some $k \in \mathscr{B}_l$ with $S_{l,k} \cap \partial\Omega_\alpha \neq \phi\}$. Hence, $\partial\Omega^l \subset \Omega_{\alpha_l^+} \backslash \overline{\Omega}_{\alpha_l^-}$. The transformation rule for Lebesgue integrals yields

$$\|v_l\|_{L^2(\partial\Omega^l)}^2 \leqslant \int_{\alpha_l^-}^{\alpha_l^+} \int_{\partial\Omega_\alpha} \frac{|v_l|^2}{\|\nabla F\|_{\mathbb{R}^d}} ds_\alpha \, d\alpha \leqslant \rho^{-1} \int_{\alpha_l^-}^{\alpha_l^+} \|v_l\|_{L^2(\partial\Omega_\alpha)}^2 \, d\alpha$$

where ds_α denotes the surface measure on $\partial\Omega_\alpha$ and where $\rho := \min_{x \in \overline{\Omega}_\beta} \|\nabla F(x)\|_{\mathbb{R}^d} > 0$. Now we apply (2.3) yielding

$$\|v_l\|_{L^2(\partial\Omega^l)}^2 \leqslant \rho^{-1} C_\gamma^2 (\alpha_l^+ - \alpha_l^-) \|v_l\|_{H^1(\square)}^2 .$$

Since $\alpha_l^+ - \alpha_l^- = F(x) - F(y) \leqslant \max_{\xi \in D} \|\nabla F(\xi)\|_{\mathbb{R}^d} \|x - y\|_{\mathbb{R}^d}$ (D denotes the closure of $\square$) holds true for all $x \in \partial\Omega_{\alpha_l^+}$ and all $y \in \partial\Omega_{\alpha_l^-}$, the proof ends by

$$\alpha_l^+ - \alpha_l^- \leqslant \max_{\xi \in D} \|\nabla F(\xi)\|_{\mathbb{R}^d} \, \mathrm{dist}(\partial\Omega_{\alpha_l^+}, \partial\Omega_{\alpha_l^-})$$

and by $\mathrm{dist}(\partial\Omega_{\alpha_l^+}, \partial\Omega_{\alpha_l^-}) \leqslant 2\delta_l \mathrm{diam}\, S.$ $\square$

Theorem 6.6 is now an easy consequence of our former results.

THEOREM 6.6: *Adopt the assumptions of Lemma 6.5 and suppose that the estimate (6.11) holds. Let $\mathbf{W}_l$ be an optimal preconditioner for $\mathbf{A}_l$, see (6.2).*

Then there exists a positive constant $C_{\hat{W}}$ independent of l such that

$$(6.12) \qquad \text{cond}_{R(I - M_l)} (\hat{W}_l \hat{A}_l) \leq C_{\hat{W}} \delta_l^{-1} \quad \text{as} \quad l \to \infty .$$

Let $\tilde{a}$ be the H^1-inner product. In view of (6.10) we see that Y has to grow like δ_l^{-1} as $l \to \infty$. So the estimate (6.12) is not too pessimistic but qualitatively correct.

Unfortunately, the optimality of W_l does not carry over to $\hat{W}_l$. However, the cg-iteration applied to (6.1) is more efficient than might be expected, not at least because of its simple implementation (W_l has only to be realized on $\square$) and its high flexibility (once implemented it can be applied to any $\Omega \subset \square$).

In a finite element setting Börgers and Widlund [2] proposed a preconditioner which is akin to our method and enjoys the same non-optimality. Besides the restriction to two space dimensions and to linear elements, their approach has the further disadvantage that a triangulation adapted to the boundary of the domain is needed.

Not surprisingly, our preconditioner fits neatly into the abstract framework of the *fictitious space lemma* due to Nepomnyaschikh [29], see also Oswald [33] and Xu [37].

Remark 6.7: Glowinski et al. [21] proposed a preconditioned cg-iteration for the numerical solution of a penalty formulation of (1.1). However, no analytical estimate of the condition number was given. Since our more general approach covers the method of Glowinski et al. the missing estimate is finally supplied by Theorem 6.6.

6.2. Verification of the estimate (6.11)

Here we provide a proof of (6.11) when φ is the d-fold tensor product $\varphi(x) = \prod_{i=1}^{d} \Phi(x_i)$ of the univariate scaling function Φ with compact support in $[0, T]$, $T \in \mathbb{N}$. Let the integer translates of Φ be *locally linear independent* w.r.t. the interval $[0, 1]$, that is,

$$(6.13) \qquad \sum_{k \in \mathbb{Z}} c_k \Phi(\, . - k) = 0 \quad \text{in} \quad [0, 1] \Rightarrow c_k = 0 \quad \text{for all} \quad k \in \mathscr{A},$$

where $\mathscr{A} = \{ k \in \mathbb{Z} \mid \text{supp } \Phi(\, . - k) \cap \,]0, 1[\, \neq \varnothing \}$. For instance, all B-splines share the property (6.13). Local linear independence guarantees the existence of a positive constant C_u such that

$$(6.14) \qquad C_u \sum_{k \in \mathscr{A}} |c_k|^2 \leq \left\| \sum_{k \in \mathscr{A}} c_k \Phi(\, . - k) \right\|_{L^2(0, 1)}^2$$

It will be convenient to use the following notation. Let $E_k := [0, 1]^d + k$, $k \in \mathbb{Z}^d$, be the translated unit cube. We introduced E_k since the intersection $S_{l, m} \cap S_{l, r}$ of the supports of $\varphi_{l, m}$ and $\varphi_{l, r}$ can be written as disjoint union of the dilated cubes $\delta_l E_k$. To a set $B \subset \mathbb{R}^d$ we associate the index set

$$\mathscr{F}(B) := \{ k \in \mathbb{Z}^d \mid \overset{\circ}{S}_{0, k} \cap B \neq \varnothing \} .$$

LEMMA 6.8: *Let Φ be a univariate scaling function satisfying (6.13) and let the multivariate scaling function φ be defined by $\varphi(x) = \prod_{i=1}^{d} \Phi(x_i)$.*

Then, the integer translates of φ are locally linear independent w.r.t. any set $B \subset \mathbb{R}^d$ of the kind $B = \bigcup_{r \in \mathscr{M}} \text{supp } \varphi_{0, r}$ with arbitrary $\mathscr{M} \subset \mathbb{Z}^d$. More precisely we have

$$(6.15) \qquad C_u^d \sum_{k \in \mathscr{F}(B)} |c_k|^2 \leq \left\| \sum_{k \in \mathscr{F}(B)} c_k \varphi(\, . - k) \right\|_{L^2(B)}^2$$

for all $c \in \ell^2(\mathbb{Z}^d)$ where C_u is the constant from (6.14).

Proof: We split B into disjoint cubes, i.e., $B = \bigcup_{m \in \mathscr{X}} E_m$ with a suitable $\mathscr{X} \subset \mathbb{Z}^d$. Let $v := \sum_{k \in \mathscr{F}(B)} c_k \varphi_{0,k}$. Now,

$$\int_B |v(x)|^2 \, dx = \sum_{m \in \mathscr{X}} \int_{E_m} |v(x)|^2 \, dx. \quad \text{We have}$$

$$\int_{E_m} |v(x)|^2 \, dx = \sum_{k \in \mathscr{F}(E_m)} \sum_{r \in \mathscr{F}(E_m)} c_k \bar{c}_r \prod_{i=1}^d \langle \Phi(\,.\, - k_i), \Phi(\,.\, - r_i) \rangle_{L^2(m_i, m_i + 1)}$$

$$\geq C_u^d \sum_{r \in \mathscr{F}(E_m)} |c_r|^2$$

as an immediate consequence from (6.14). Hence,

$$\int_B |v(x)|^2 \, dx \geq C_u^d \sum_{m \in \mathscr{X}} \sum_{r \in \mathscr{F}(E_m)} |c_r|^2 .$$

Since $\mathscr{F}(B) = \bigcup_{m \in \mathscr{X}} \mathscr{F}(E_m)$ we get $\sum_{m \in \mathscr{X}} \sum_{r \in \mathscr{F}(E_m)} |c_r|^2 \geq \sum_{k \in \mathscr{F}(B)} |c_k|^2.$ $\quad\square$

Let $v_l = \sum_{k \in \mathbb{Z}^{d,l}} v_{l,k} \varphi_k^l \in V_l^p$. For l sufficiently large we have that

(6.16)
$$v_l|_{\partial \Omega^l} = \sum_{k \in \mathbb{Z}^d} v_{l,k} \varphi_{l,k}|_{\partial \Omega^l}.$$

A straightforward calculation shows that

$$\| v_l \|_{L^2(\partial \Omega^l)}^2 = \int_B \left| \sum_{k \in \mathscr{F}(B)} v_{l,k} \varphi(x - k) \right|^2 dx \quad \text{where} \quad B = \bigcup_{r \in \mathscr{B}_l} \mathrm{supp}\, \varphi_{0,r} .$$

Applying (6.15) we estimate the right-hand side of the above equality from below by

(6.17)
$$\| v_l \|_{L^2(\partial \Omega^l)}^2 \geq C_u^d \sum_{k \in \mathscr{F}(B)} |v_{l,k}|^2 .$$

THEOREM 6.9: *Let the scaling functions Φ and φ be as in Lemma 6.8. Then, the estimate (6.11) holds true for l sufficiently large.*

Proof: Let l such large that (6.16) applies. The local linear independence of Φ implies the norm equivalence (3.2) for φ. Consequently,

$$\| \gamma^l(v_l) \|_{L^2(\partial \Omega^l)}^2 = \| \gamma^l(v_l) \|_{L^2(\mathbb{R}^d)}^2 \leq C_\varphi \sum_{k \in \mathscr{B}_l} |v_{l,k}|^2$$

with $C_\varphi > 0$. Because $\mathscr{B}_l \subset \mathscr{F}(B)$ where $B = \bigcup_{r \in \mathscr{B}_l} \mathrm{supp}\, \varphi_{0,k}$, the stated estimate (6.11) with $C_V = C_\varphi / C_u^d$ results from (6.17). $\quad\square$

6.3. Numerical examples

We demonstrate the performance of our preconditioned cg-iteration by a variety of numerical examples.

All computations are based on the boundary value problem (5.10) in dimensions $d = 2$ and $d = 3$ where $\alpha = 1, f = 1$, and $g = 0$. The domain Ω is always the ball $\Omega = \{x \in \mathbb{R}^d \mid \|x\|_{\mathbb{R}^d} < 0.25\}$. The fictitious domain

is the cube $\square = [-0.3, 0.3]^d$. The approximation spaces V_l^p are realized as in (5.11) where the underlying scaling function is a tensor product. We considered four univariate scaling functions, namely the linear (B_2), the quadratic (B_3) and the cubic (B_4) B-splines as well as the Daubechies scaling function (DS_3) of order 3.

Tables 6.1 and 6.2 display the number of iteration steps S needed by the cg-method to yield a relative residue smaller than 0.01, that is,

$$(6.18) \qquad \| \hat{A}_l v^S - (I - M_l) \mathbf{f}_l \|_{\mathbb{R}^{n_l}} < 0.01 \, . \, \| (I - M_l) \mathbf{f}_l \|_{\mathbb{R}^{n_l}} \, .$$

(Note that $\mathbf{g}_l = 0$ in our examples). In (6.18), v^S is the S-th iterate of the cg-iteration applied either to (5.3) or (6.1) started with the initial guess $v^0 = 0$.

Table 6.1. — 2D-example. Number of iteration steps S to satisfy (6.18). Left table (top): cg-iteration applied to (5.3), right table (top): cg-iteration applied to (6.1) where W_l is the BPX-preconditioner. Table on the bottom: cg-iteration applied to (6.1) where W_l is the wavelet preconditioner.

l	5	6	7	8	9	10
B_2	12	25	52	105	349	860
B_3	7	15	31	64	191	525
B_4	6	13	26	54	144	448
DS_3	24	64	134	413	1 098	2 986

l	5	6	7	8	9	10
B_2	7	10	14	20	29	53
B_3	6	8	12	17	24	37
B_4	7	12	18	27	41	64
DS_3	12	18	26	38	54	81

l	5	6	7	8	9	10
$B_{2, \, {}_2\tilde{2}\tilde{\varphi}}$	7	9	15	22	35	58
DS_3	13	19	31	45	71	111

Table 6.2. — 3D-example. Number of iteration steps S to satisfy (6.18). Left table: cg-iteration applied to (5.3), (n.c. = not computed). Right table: cg-iteration applied to (6.1) where W_l is the BPX-preconditioner.

l	5	6	7	8
B_2	11	24	50	175
B_3	6	14	30	100
B_4	8	12	25	84
DS_3	28	70	212	n.c.

l	5	6	7	8
B_2	6	9	12	20
B_3	7	10	15	23
B_4	10	17	29	51
DS_3	12	18	27	49

All our numerical experiments confirm the theory. Without a preconditioning the number of iteration steps doubles at least as the discretization step size is halved. The application of an optimal preconditioner for the periodic problem, i.e., for A_l, reduces the number of necessary iteration steps significantly. Moreover, S grows only with about the factor 1.5 when the discretization step size is divided by 2. As predicted by the theory this behavior is independent of the space dimension d.

Table 6.1 shows results w.r.t. two optimal preconditioners for A_l. These are the BPX- and a certain wavelet preconditioner. The wavelet preconditioner depends on the chosen dual (biorthogonal) scaling function. In our computations with the linear B-spline B_2 we used its dual ${}_{2,2}\tilde{\varphi}$ constructed by Cohen, Daubechies and Feauveau [8, Section 6.A]. We also tested the wavelet preconditioner in case of the Daubechies scaling function DS_3. Here DS_3 coincides with its dual.

At a first glance the *BPX*- and the wavelet preconditioner seem to be of comparable efficiency. As soon as we consider the computational effort this impression changes. The wavelet preconditioner can be interpreted as a "sum" of 2^d *BPX*-preconditioners. Hence, the computational effort for the wavelet preconditioner is about 2^d times the effort of the *BPX*-preconditioner. This factor even increases when the support of the dual scaling function is much larger than the one of the scaling function.

It is worthwhile to mention that the quality of the preconditioner $\hat{\mathbf{W}}_l$ is not affected when the underlying boundary value problem has less than full elliptic regularity. Indeed, the statement of Theorem 6.6 holds without regularity assumptions. For numerical experiments in this situation we refer to Glowinski et al. [21].

7. DISCUSSION AND CONCLUSION

In this paper we proposed an embedding method for Dirichlet problems in $\mathbb{R}^d$, $d \geq 2$. The underlying leading idea was its convenient implementation for domains with complicated boundaries. Its further advantages have been reported in detail on the previous pages. Therefore, we address here two aspects of the algorithm which shall be improved in future research.

First of all, the accuracy of the computed solution gets worse near the boundary. The reasons have been explained. One way to overcome this dilemma could be as follows. We discretize the boundary value problem based on a low order scaling function. The resulting approximate solution is used to define an improved boundary approximation $\tilde{g}_l$ w.r.t. a scaling function of higher order. This could prove to be a practical approach realizing the assumptions of Theorem 4.4.

Iterative solvers with a convergence rate independent of the mesh size would additionally increase the attractiveness of our scheme. We think of multigrid methods whose prolongation and restriction operators take special care of the boundary components. Such local techniques have been employed successfully for free boundary value problems, see, e.g., Hoppe [4].

Our presented algorithm has very attractive features. Its potential deserves further exploration.

REFERENCES

[1] S. BERTOLUZZA, *Interior estimates for the wavelet Galerkin method,* Numer. Math. 78 (1997), pp. 1-20.

[2] C. BÖRGERS and O. B. WIDLUND, *On finite element domain imbedding methods,* SIAM J. Numer. Anal., 27 (1990), pp. 963-978.

[3] D. BRAESS, *Finite-Elemente,* Springer-Lehrbuch, Springer-Verlag, Berlin, 1992.

[4] J. H. BRAMBLE and J. E. PASCIAK, *New estimates for multilevel algorithms including the V-cycle,* Math. Comp., 60 (1993), pp. 447-471.

[5] J. H. BRAMBLE, J. E. PASCIAK and J. XU, *Parallel multilevel preconditioners,* Math. Comp., 55 (1990), pp. 1-22.

[6] C. K. CHUI, *Multivariate Splines,* vol. 54 of CBMS-NSF Series in Applied Mathematics, SIAM, Philadelphia, 1988.

[7] B. A. CIPRA, *A rapid-deployment force for CFD: Cartesian grids,* Siam News (Newsjournal of the Society for Industrial and Applied Mathematics), 25 (1995).

[8] A. COHEN, I. DAUBECHIES and J.-C. FEAUVEAU, *Biorthogonal bases of compactly supported wavelets,* Comm. Pure Appl. Math., 45 (1992), pp. 485-560.

[9] S. DAHLKE, V. LATOUR and K. GRÖCHENIG, *Biorthogonal box spline wavelet bases,* Bericht 122, Institut für Geometrie und Praktische Mathematik, RWTH Aachen, 1995.

[10] W. DAHMEN and A. KUNOTH, *Multilevel preconditioning,* Numer. Math., 63 (1992), pp. 315-344.

[11] W. DAHMEN and C. A. MICCHELLI, *Using the refinement equation for evaluating integrals of wavelets,* SIAM J. Numer. Anal., 30 (1993), pp. 507-537.

[12] W. DAHMEN, S. PRÖSSDORF and R. SCHNEIDER, *Wavelet approximation methods for pseudodifferential equations I: Stability and convergence,* Math. Z., 215 (1994), pp. 583-620.

[13] I. DAUBECHIES, *Orthonormal bases of compactly supported wavelets,* Comm. Pure Appl. Math., 41 (1988), pp. 906-966.

[14] C. DE BOOR, K. HÖLLING and S. RIEMENSCHNEIDER, *Box Splines*, vol. 98 of Applied Mathematical Sciences, Springer-Verlag, Berlin, 1993.

[15] P. DEUFLHARD and A. HOHMANN, *Numerical Analysis: A First Course in Scientific Computation*, de Gruyter Textbook, de Gruyter, Berlin, New York, 1994.

[16] G. J. FIX and G. STRANG, *A Fourier analysis of the finite element method in Ritz-Galerkin theory*, in Constructive Aspects of Functional Analysis, Rome, 1973, Edizioni Cremonese, pp. 265-273.

[17] D. GILBARG and N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, vol. 224 of Grundlehren der mathematischen Wissenschaften, Springer-Verlag, Berlin, 1983.

[18] R. GLOWINSKI, *Numerical Methods for Nonlinear Variational Problems*, Springer Series in Computational Physics, Springer-Verlag, New York, 1984.

[19] R. GLOWINSKI and T.-W. PAN, *Error estimates for fictitious domain/penalty/finite element methods*, Calcolo, 19 (1992), pp. 125-141.

[20] R. GLOWINSKI, T.-W. PAN, R. O. WELLS Jr. and X. ZHOU, *Wavelet and finite element solutions for the Neumann problem using fictitious domains*, J. Comp. Phys., 126 (1996), pp. 40-51.

[21] R. GLOWINSKI, A. RIEDER, R. O. WELLS Jr. and X. ZHOU, *A wavelet multilevel method for Dirichlet boundary value problems in general domains*, Modélisation Mathématique et Analyse Numérique (M^{2}AN), 30 (1996), pp. 711-729.

[22] W. HACKBUSCH, *Elliptic Differential Equations: Theory and Numerical Treatment*, vol. 18 of Springer Series in Computational Mathematics, Springer-Verlag, Heidelberg, 1992.

[23] —, *Iterative Solution of Large Sparse Systems of Equations*, Applied Mathematical Sciences, Springer-Verlag, New York, 1994.

[24] R. H. W. HOPPE, *Une méthode multigrille pour la solution des problèmes d'obstacle*, Modélisation Mathématique et Analyse Numérique (M^{2}AN), 24 (1990), pp. 711-736.

[25] S. JAFFARD, *Wavelet methods for fast resolution of elliptic problems*, SIAM J. Numer. Anal., 29 (1992), pp. 965-986.

[26] A. KUNOTH, *Computing refinable integrals: documentation of the program*, Manual Institut für Geometrie und Praktische Mathemtik, RWTH Aachen, 1995.

[27] Y. A. KUZNETSOV, S. A. FINOGENOV and A. V. SUPALOV, *Fictitiuos domain methods for 3D elliptic problems: algorithms and software within a parallel environment*, Arbeitspapiere der GMD 726, GMD, D-53754 St. Augustin, Germany, 1993.

[28] A. LATTO, H. L. RESNIKOFF and E. TENENBAUM, *The evaluation of connection coefficients of compactly supported wavelets*, in Proceedings of the USA-French Workshop on Wavelets and Turbulence, Princeton University, 1991.

[29] S. V. NEPOMNYASCHIKH, *Mesh theorems of traces, normalization of function traces and their inversion*, Sov. J. Numer. Anal. Math. Model., 6 (1991), pp. 223-242.

[30] —, *Fictitious space method on unstructured grids*, East-West J. Numer. Math., 3 (1995), pp. 71-79.

[31] J. A. NITSCHE, *Ein Kriterium für die Quasi-Optimalität des Ritzschen Verfahrens*, Numer. Math., 11 (1968), pp. 346-348.

[32] J. A. NITSCHE and A. H. SCHATZ, *Interior estimates for Ritz-Galerkin methods*, Math. Comp., 28 (1974), pp. 937-958.

[33] P. OSWALD, *Multilevel Finite Element Approximation: Theory and Applications*, Teubner Skripten zur Numerik, B. G. Teubner, Stuttgart, Germany, 1994.

[34] G. STRANG and G. J. FIX, *An Analysis of the Finite Element Method*, Prentice-Hall Series in Automatic Computation, Prentice-Hall, Englewood Cliffs, N.J., 1973.

[35] R. O. WELLS Jr. and X. ZHOU, *Wavelet-Galerkin solutions for the Dirichlet problem*, Numer. Math., 70 (1995), pp. 379-396.

[36] J. WLOKA, *Partial Differential Equations*, Cambridge Univesrity Press, Cambridge, U.K., 1987.

[37] J. XU, *The auxiliary space method and optimal multigrid preconditioning techniques for unstructured grids*, Computing, 56 (1996), pp. 215-235.