MATHEMATICS OF COMPUTATION
Volume 69, Number 230, Pages 501–520
S 0025-5718(99)01138-2
Article electronically published on March 18, 1999

THE CONVERGENCE OF THE CASCADIC CONJUGATE-GRADIENT METHOD APPLIED TO ELLIPTIC PROBLEMS IN DOMAINS WITH RE-ENTRANT CORNERS

VLADIMIR SHAIDUROV AND LUTZ TOBISKA

ABSTRACT. We study the convergence properties of the cascadic conjugate-gradient method (CCG-method), which can be considered as a multilevel method without coarse-grid correction. Nevertheless, the CCG-method converges with a rate that is independent of the number of unknowns and the number of grid levels. We prove this property for two-dimensional elliptic second-order Dirichlet problems in a polygonal domain with an interior angle greater than π. For piecewise linear finite elements we construct special nested triangulations that satisfy the conditions of a "triangulation of type (h, γ, L)" in the sense of I. Babuška, R. B. Kellogg and J. Pitkäranta. In this way we can guarantee both the same order of accuracy in the energy norm of the discrete solution and the same convergence rate of the CCG-method as in the case of quasiuniform triangulations of a convex polygonal domain.

1. INTRODUCTION

In this paper, we consider a cascadic conjugate-gradient method (CCG-method) for solving discretized elliptic equations that yield discrete symmetric positive definite problems. This algorithm can be considered as a multigrid or multilevel method, but without coarse grid correction, i.e., if a certain grid level is attained, we do not return to coarser grid levels but proceed only at the same or on higher grid levels. The CCG-method can be recursively defined as follows. On the coarsest grid, the linear system is solved directly. On finer grids, the system is solved iteratively by the conjugate-gradient method. These iterations are started by an interpolation of the approximate solution from the previous coarser grid. On each fixed grid level we do not use any preconditioning based on coarser grids nor any restrictions onto coarser grid levels. Nevertheless, the CCG-algorithm as a multilevel method has optimal arithmetic complexity, and its convergence rate is independent of the number of unknowns and of the number of grid levels.

A CCG-algorithm has been recently presented by P. Deuflhard in [4] and [5], where the excellent convergence properties of this algorithm were demonstrated by numerical test examples. Its optimal arithmetic complexity with respect to the number of unknowns was proved for H^2-regular elliptic problems in [12]. Then for

Received by the editor November 11, 1997 and, in revised form, July 10, 1998.
1991 *Mathematics Subject Classification.* Primary 65F10; Secondary 65N30.
Key words and phrases. Multigrid, cascadic algorithm, conjugate-gradient method, finite element method.
The research was supported by the Deutsche Forschungsgemeinschaft.

quasiuniform meshes this result was extended in [11] and [2] to elliptic problems with reduced regularity caused by interior angles greater than π. Nevertheless, the use of piecewise linear finite elements on quasiuniform triangulations reduces the convergence order of the Galerkin solution. Furthermore, F.A. Bornemann [2] studied the replacement of the CG-method by other iterative methods (damped Jacobi, Gauß-Seidel, SSOR, etc.) and gave sufficient conditions for optimal complexity of the cascadic algorithm. In the three-dimensional case these conditions are satisfied by many known iterative schemes, but in two dimensions the amount of work is suboptimal unless the CG-method is used.

Here we use piecewise linear finite elements on triangles in order to discretize a second-order elliptic problem in a polygonal domain with an interior angle greater than π. We construct special nested triangulations that are refined towards this angular point as in [13] and satisfy the conditions of a "triangulation of type (h, γ, L)"; these were defined in [1] and used for the classical multigrid method in [14]. We prove in detail that one obtains the same order of accuracy of the approximate solution and the same convergence rate of the CCG-method in the energy norm as in the H^2-regular case.

2. The cascadic algorithm

We denote by $M_0, M_1, \cdots, M_l$ finite-dimensional vector spaces of increasing dimension equipped with inner products $(\cdot,\cdot)_i$, for $i = 0, 1, \cdots, l$. Moreover, let linear prolongation operators

$$I_i : \; M_i \to M_{i+1}, \quad \text{for } i = 0, 1, \cdots, l-1, \tag{2.1}$$

and linear invertible operators

$$L_i : \; M_i \to M_i, \quad \text{for } \; i = 0, \cdots, l, \tag{2.2}$$

be given. Then the cascadic algorithm is an iterative method for solving the following problem:

For a given $f_l \in M_l$, find $u_l \in M_l$ such that

$$L_l u_l = f_l, \tag{2.3}$$

by using approximations of the solutions of the following problems:

For a given $f_i \in M_i$, find $u_i \in M_i$ such that

$$L_i u_i = f_i \tag{2.4}$$

on lower levels $i = 0, \cdots, l-1$. The idea is to start with the exact solution

$$v_0 = u_0$$

on the lowest level $i = 0$ and to prolong each approximate solution

$$v_i \in M_i$$

of (2.4) to the next higher level in order to find an initial guess for an iterative method that approximates the solution u_{i+1}. Applying the conjugate-gradient algorithm (CG-algoritm) on each level, we obtain the cascadic conjugate-gradient algorithm (CCG-algorithm), which can be formulated in the following way:

CCG-algorithm

{1. Set $v_0 = L_0^{-1} f_0$.
2. For $i = 1, 2, \cdots, l$ and given v_{i-1} do :
{ 2.1. Set $w_i = I_{i-1} v_{i-1}$;
2.2. Perform m_i iterations of the conjugate-gradient method :

$y_0 = w_i$;
$p_0 = r_0 = f_i - L_i y_0$;
$\sigma_0 = (r_0, r_0)_i$;
for $k = 1, 2, \cdots, m_i$ do :
{ $\alpha_{k-1} = \sigma_{k-1}/(p_{k-1}, L_i p_{k-1})$;
$y_k = y_{k-1} + \alpha_{k-1} p_{k-1}$;
$r_k = r_{k-1} - \alpha_{k-1} L_i p_{k-1}$;
$\sigma_k = (r_k, r_k)_i$;
if $\sigma_k = 0$ then { $y_{m_i} = y_k$; goto 2.3 };
$\beta_k = \sigma_k / \sigma_{k-1}$;
$p_k = r_k + \beta_k p_{k-1}$;
} the end of the iteration;

2.3. Set $v_i = y_{m_i}$;
} the end of the level i;
} the end of the algorithm.

We shall study the convergence properties of the CCG-algorithm under the assumption that the operators L_i, for $i = 0, 1, \cdots, l$, are self-adjoint and positive definite, i.e., for $i = 0, 1, \cdots, l$ we have

$$(L_i u, v)_i = (u, L_i v)_i, \qquad (L_i u, u)_i \geq \alpha_i (u, u)_i, \quad \alpha_i > 0, \qquad \forall u, v \in M_i \ . \tag{2.5}$$

Moreover, we assume that the operator $L_{i-1} : M_{i-1} \to M_{i-1}$ on the lower level can be represented by means of the operator $L_i : M_i \to M_i$ and the transfer operators $I_{i-1} : M_{i-1} \to M_i$, $I_{i-1}^* : M_i \to M_{i-1}$ in the form

$$L_{i-1} = I_{i-1}^* L_i I_{i-1} \ . \tag{2.6}$$

Note that the adjoint operator $I_{i-1}^* : M_i \to M_{i-1}$, for $i = 1, 2, \cdots, l$, is defined by

$$(I_{i-1}^* v, w)_{i-1} = (v, I_{i-1} w)_i \qquad \forall v \in M_i, \ w \in M_{i-1} \ .$$

We introduce a scale of norms on M_i by

$$|||u|||_i^{(\alpha)} := \sqrt{(L_i^\alpha u, u)_i}, \quad u \in M_i \ ,$$

with $\alpha \in (-\infty, \infty)$. In order to simplify the notation, we write

$$|||u|||_i := |||u|||_i^{(1)}, \quad \|u\|_i := |||u|||_i^{(0)}, \quad u \in M_i.$$

The operator norm induced by $||| \cdot |||_i$ for an operator $B : M_i \to M_i$ is given by

$$|||B|||_i = \sup_{u \in M_i \setminus \{0\}} \frac{|||Bu|||_i}{|||u|||_i}. \tag{2.7}$$

On a fixed level $i \in \{1, \cdots, l\}$ we apply the CG-algorithm to reduce the error $u_i - w_i$ of the initial guess w_i for the exact solution u_i of problem (2.4). After m_i steps we get the error

$$u_i - v_i := B_i(u_i - w_i)$$

of the final approximation v_i on level i. In this way we define the operator $B_i : M_i \to M_i$ of error reduction on level i. This operator can be represented as a polynomial in L_i:

$$B_i = P_i(L_i) = I + \sum_{k=1}^{m_i} a_k L_i^k, \tag{2.8}$$

with coefficients which depend on the parameters $\sigma_0, \cdots, \sigma_{m_i}, \alpha_0, \cdots, \alpha_{m_{i-1}}$ given in the definition of the CCG algorithm (see [10]). Here and in the following we denote by I the identity in the corresponding space. From [10] we recall the well-known optimality property of the CG-algorithm:

Lemma 2.1. *Among all polynomials of the form* (2.8) *with arbitrary coefficients a_k the conjugate-gradient method minimizes the error $u_i - v_i$ of the final approximation v_i in the norm $|||\cdot|||_i$ for a fixed given initial guess w_i.*

3. Optimal polynomials

For estimating the norm of the error-reduction operator B_i on the level i, we consider polynomials q_m of degree m with $q_m(0) = 1$. These polynomials can be written in the form

$$q_m(x) = \prod_{k=1}^{m} (1 - \mu_k x) , \tag{3.1}$$

with parameters $\mu_k \neq 0$ for $k = 1, \ldots, m$. Note that the polynomial P_i that defines the error-reduction operator B_i on the level i has the same structure. We shall show that on a given compact set $[0, d]$ the parameters μ_k, for $k = 1, \ldots, m$, can be chosen in such a way that the resulting polynomial satisfies certain optimality properties.

Lemma 3.1. *For any $\gamma > 0$ and any $d > 0$, there exist parameters μ_k, for $k = 1, \ldots, m$, such that the polynomial defined by* (3.1) *satisfies*

$$\max_{0 \le x \le d} |q_m(x)| \le 1 \tag{3.2}$$

and

$$\max_{0 \le x \le d} |x^{\gamma/2} q_m(x)| \le \eta_\gamma(m) d^{\gamma/2}, \tag{3.3}$$

where $\eta_\gamma(m)$ is independent of d and tends to 0 if m tends to infinity.

Proof. We consider the minimization problem

Find parameters $\mu_1, \cdots, \mu_m$ such that

$$M(\mu_1, \cdots, \mu_m) := \max_{0 \le x \le d} |\sqrt{x}\, q_m(x)| \tag{3.4}$$

becomes minimal.

In [13, §4.1], it has been proved that the solution of this problem defines the polynomial $\bar{q}_m$ with

$$\sqrt{x}\,\bar{q}_m(x) = (-1)^m p_m \cos((2m+1)\arccos\sqrt{x/d}), \quad p_m = \sqrt{d}/(2m+1). \tag{3.5}$$

One can check that $\bar{q}_m$ is a polynomial in x of degree m with zeros at $\bar{x}_k$ and whose parameters μ_k are given by

$$\bar{\mu}_k = \frac{1}{\bar{x}_k} = \frac{1}{d}\cos^{-2}\frac{\pi(2k+1)}{2(2m+1)}, \quad \text{for } k = 0, \cdots, m-1. \tag{3.6}$$

In [13, §4.1] it has also been proved that

$$|\bar{q}_m(x)| \leq 1 \quad \forall x \in [0, d]. \tag{3.7}$$

It follows directly from (3.5) that

$$|\sqrt{x}\,\bar{q}_m(x)| \leq p_m \quad \forall x \in [0, d]. \tag{3.8}$$

Now let us first consider the case $\gamma \in (0, 1]$. Using (3.7) and (3.8), we get

$$|x^{\gamma/2}\bar{q}_m(x)| \leq |\sqrt{x}\,\bar{q}_m(x)|^\gamma \leq d^{\gamma/2}/(2m+1)^\gamma \quad \forall x \in [0, d].$$

Thus we have shown that there are parameters μ_k in (3.1) such that the estimates (3.2) and (3.3) hold with $\eta_\gamma(m)$ defined by

$$\eta_\gamma(m) = \frac{1}{(2m+1)^\gamma}, \quad \text{for } \gamma \in (0, 1]. \tag{3.9}$$

Next we consider the case $\gamma > 1$ and put

$$r = -[-\gamma] = \begin{cases} \gamma & \text{if } \gamma \text{ is an integer,} \\ [\gamma] + 1 & \text{otherwise.} \end{cases}$$

Any integer m can be decomposed in the form

$$m = tr + s\ ,$$

where $t = [m/r]$ and $0 \leq s \leq r-1$. Then the polynomial

$$\bar{\bar{q}}_m\,(x) = \bar{q}_t^{\,r-s}(x)\bar{q}_{t+1}^{\,s}(x) \tag{3.10}$$

is of the form (3.1), with t parameters

$$\bar{\bar{\mu}}_i = \frac{1}{d}\cos^{-2}\frac{\pi(2i+1)}{2(2t+1)}, \quad \text{for } i = 0, \cdots, t-1, \tag{3.11}$$

of multiplicity $r - s$ and $t + 1$ parameters

$$\bar{\bar{\bar{\mu}}}_j = \frac{1}{d}\cos^{-2}\frac{\pi(2j+1)}{2(2t+3)}, \quad \text{for } j = 0, \cdots, t, \tag{3.12}$$

of multiplicity s. Applying (3.7) to $\bar{q}_t$ and $\bar{q}_{t+1}$, we get

$$|\,\bar{\bar{q}}_m\,(x)| = |\bar{q}_t(x)|^{r-s}\,|\bar{q}_{t+1}|^s \leq 1 \quad \forall x \in [0, d]. \tag{3.13}$$

In order to show (3.3) we use (3.8) for $\bar{q}_t$ and $\bar{q}_{t+1}$, and obtain

$$|x^{\gamma/2r}\bar{q}_t(x)| \leq |\sqrt{x}\,\bar{q}_t(x)|^{\gamma/r} \leq \left(\frac{\sqrt{d}}{2t+1}\right)^{\gamma/r} \quad \forall x \in [0, d]$$

and

$$|x^{\gamma/2r}\bar{q}_{t+1}(x)| \leq |\sqrt{x}\,\bar{q}_{t+1}(x)|^{\gamma/r} \leq \left(\frac{\sqrt{d}}{2t+3}\right)^{\gamma/r} \quad \forall x \in [0, d].$$

Consequently,

$$|x^{\gamma/2}\,\bar{\bar{q}}_m\,(x)| \le |x^{\gamma/2r}\bar{q}_t(x)|^{r-s}|x^{\gamma/2r}\bar{q}_{t+1}(x)|^s \le d^{\gamma/2}\eta_\gamma(m) \quad \forall x \in [0,d], \tag{3.14}$$

where

$$\eta_\gamma(m) = \frac{1}{(2t+1)^{\gamma(r-s)/r}(2t+3)^{\gamma s/r}} \quad \forall x \in [0,d]. \tag{3.15}$$

Note that the function $\eta_\gamma(\cdot)$ is monotonically decreasing with respect to $m = tr+s$. Moreover, if $s=0$ (i.e., if m is a multiple of r) we have

$$\eta_\gamma(m) = \frac{1}{(2m/r+1)^\gamma} = O(m^{-\gamma}). \tag{3.16}$$

Therefore $\eta_\gamma(m)$ tends to 0 when $m \to \infty$. Thus, setting μ_k in (3.1) equal to the values in (3.11) and (3.12) with their corresponding multiplicities, we get (3.3) for the case $\gamma > 1$ also. □

We set $d = \lambda_i^*$, where λ_i^* denotes the largest eigenvalue of the operator L_i in the space M_i, and choose the parameters μ_k as in the proof of Lemma 3.1. Then the optimal polynomial q_m defines the auxiliary operator $S_{i,m}$ by

$$S_{i,m} = q_m(L_i) = \prod_{k=1}^{m}(I - \mu_k L_i),$$

which majorizes the error-reduction operator B_i on level i owing to Lemma 2.1.

Lemma 3.2. *Let the operator L_i be self-adjoint and positive definite. Then for any $\gamma > 0$, we have the inequalities*

$$|||S_{i,m}w|||_i \le (\lambda_i^*)^{\gamma/2}\eta_\gamma(m)|||w|||_i^{(1-\gamma)} \quad \forall w \in M_i \tag{3.17}$$

and

$$|||S_{i,m}w|||_i \le |||w|||_i \quad \forall w \in M_i, \tag{3.18}$$

where the function η_γ is independent of d and tends to 0 if m tends to infinity.

Proof. We can assume that the set of eigenvectors $\{\varphi_j\}_{j=1}^{n_i}$ of the eigenvalue problem

$$L_i\varphi_j = \lambda_j\varphi_j, \quad \text{for } j = 1,\cdots,n_i, \quad \text{where } n_i = \dim M_i, \tag{3.19}$$

is orthonormal with respect to the inner product $(\cdot,\cdot)_i$, i.e.,

$$(\varphi_j,\varphi_k)_i = \delta_{jk}, \quad \text{for } j,k = 1,\ldots,n_i,$$

where δ_{jk} is Kronecker's symbol. Then, using the basis representation

$$w = \sum_{j=1}^{n_i}\alpha_j\varphi_j \tag{3.20}$$

of $w \in M_i$, we get

$$\left(|||w|||_i^{(1-\gamma)}\right)^2 = \sum_{j=1}^{n_i}\lambda_j^{1-\gamma}\alpha_j^2 \tag{3.21}$$

and

$$|||S_{i,m}w|||_i^2 = \sum_{j=1}^{n_i}\lambda_j q_m^2(\lambda_j)\alpha_j^2. \tag{3.22}$$

From (3.3) we obtain

$$\sum_{j=1}^{n_i} \lambda_j q_m^2(\lambda_j)\alpha_j^2 \le \eta_\gamma^2(m)(\lambda_i^*)^\gamma \sum_{j=1}^{n_i} \lambda_j^{1-\gamma}\alpha_j^2 \tag{3.23}$$
$$= \eta_\gamma^2(m)(\lambda_i^*)^\gamma (|||w|||_i^{(1-\gamma)})^2,$$

which implies (3.17). Using (3.2), we get immediately

$$\sum_{j=1}^{n_i} \lambda_j q_m^2(\lambda_j)\alpha_j^2 \le \sum_{j=1}^{n_i} \lambda_j \alpha_j^2 = |||w|||_i^2,$$

which implies (3.18). □

4. The algebraic convergence theorem

In order to formulate our abstract convergence result, we assume that the following criterion is satisfied:

There exist constants $c^* > 0$ and $\gamma > 0$ such that for $i = 1, \cdots, l$ we have the following relation between two neighbouring solutions u_{i-1} and u_i of the problems (2.4):

$$|||u_i - I_{i-1}u_{i-1}|||_i^{(1-\gamma)} \le c^*(\lambda_i^*)^{-\gamma/2}\, |||u_i - I_{i-1}u_{i-1}|||_i. \tag{4.1}$$

Note that this inequality can be proved not only in the case of $H^{1+\lambda}$-regularity with $\lambda \in (0,1]$, but also for $\lambda > 1$ when for example second-order finite elements are used.

Theorem 4.1. *Let the operators L_i, for $i = 0, \dots, l$, be self-adjoint, positive definite and satisfy*

$$L_{i-1} = I_{i-1}^* L_i I_{i-1}.$$

We assume that the convergence criterion (4.1) *holds for some $\gamma > 0$. Then, for each level i, where $i = 1, \dots, l$, the approximate solution v_i of the CCG-algorithm satisfies the inequality*

$$|||u_i - v_i|||_i \le c^* \sum_{j=1}^{i} \eta_\gamma(m_j)\, |||u_j - I_{j-1}u_{j-1}|||_j\ , \tag{4.2}$$

where the constant c^* and the function η_γ are independent of i and u_j for $j = 1, \dots, i$.

Proof. Let us denote the iteration error of the CCG-algorithm at level i after m_i steps by

$$\varepsilon_i = u_i - v_i \quad \forall i = 0, 1, \cdots, l.$$

Using the definition of the error-reduction operator B_i, we have

$$\varepsilon_i = B_i(u_i - w_i)\ ,$$

where w_i is the initial guess for the CG-algorithm on level i. The polynomial $q_m(\cdot)$ that defines $S_{i,m_i}(\cdot)$ has the form (2.8) with some coefficients altered. The minimization property of the CG-algorithm (Lemma 2.1) implies that

$$\begin{aligned} |||\varepsilon_i|||_i &= |||B_i(u_i - w_i)|||_i \\ &\le |||S_{i,m_i}(u_i - w_i)|||_i \\ &\le |||S_{i,m_i}(u_i - I_{i-1}u_{i-1})|||_i + |||S_{i,m_i}I_{i-1}\varepsilon_{i-1}|||_i. \end{aligned} \tag{4.3}$$

Taking into consideration (3.17) and (4.1), we can estimate the first term in the right-hand side of (4.3):

$$\begin{aligned} |||S_{i,m_i}(u_i - I_{i-1}u_{i-1})|||_i &\le (\lambda_i^*)^{\gamma/2}\eta_\gamma(m_i)\,|||u_i - I_{i-1}u_{i-1}|||_i^{(1-\gamma)} \\ &\le c^*\eta_\gamma(m_i)\,|||u_i - I_{i-1}u_{i-1}|||_i. \end{aligned} \tag{4.4}$$

To estimate the second term, we use (3.18):

$$|||S_{i,m_i}I_{i-1}\varepsilon_{i-1}|||_i \le |||I_{i-1}\varepsilon_{i-1}|||_i\ ,$$

and by (2.6) we have

$$\begin{aligned} |||I_{i-1}\varepsilon_{i-1}|||_i^2 &= (I_{i-1}^*L_iI_{i-1}\varepsilon_{i-1}, \varepsilon_{i-1})_{i-1} \\ &= (L_{i-1}\varepsilon_{i-1}, \varepsilon_{i-1})_{i-1} \\ &= |||\varepsilon_{i-1}|||_{i-1}^2. \end{aligned}$$

Thus we obtain

$$|||\varepsilon_i|||_i \le c^*\eta_\gamma(m_i)\,|||u_i - I_{i-1}u_{i-1}|||_i + |||\varepsilon_{i-1}|||_{i-1}, \tag{4.5}$$

from which the statement of the theorem follows by induction on i. □

Remark 4.2. Each term in the sum (4.2),

$$\frac{c^*}{2m_j+1}|||u_j - I_{j-1}u_{j-1}|||_j,$$

can be considered as the error contribution at the corresponding level j of the CCG-algorithm. Since c^* is independent of j and m_j, we can reduce this error contribution by taking a suffiently large number of smoothing steps m_j. Later we shall see that asymptotically the size of the term $|||u_j - I_{j-1}u_{j-1}|||_j$ also decreases as the level j increases. This is important, since the complexity of the CG-iteration increases with j.

5. The boundary value problem

Let us consider the following Dirichlet problem in an open bounded polygon $\Omega \subset \mathbf{R}^2$ with boundary $\Gamma = \partial\Omega$:

$$-\sum_{i,j=1}^{2} \partial_i(a_{ij}\partial_j u) + bu = f \quad \text{in} \quad \Omega, \tag{5.1}$$

$$u = 0 \quad \text{on} \quad \Gamma, \tag{5.2}$$

where the coefficients and the right-hand side of (5.1) satisfy the conditions

$$\begin{cases} \partial_k a_{ij} \in L_q(\Omega), \quad q > 2, \quad i,j,k = 1,2; \qquad a_{12} = a_{21} \quad \text{on} \quad \bar\Omega; \\ b \in L_2(\Omega); \qquad b \ge 0 \quad \text{on} \quad \bar\Omega; \quad \text{there are } \nu_2 \ge \nu_1 > 0 \text{ such that} \\ \nu_1 \sum_{i=1}^2 \xi_i^2 \le \sum_{i,j=1}^2 a_{ij}(\cdot)\xi_i\xi_j \le \nu_2 \sum_{i=1}^2 \xi_i^2 \quad \forall \xi_i \in \mathbf{R} \quad \text{on} \quad \bar\Omega. \end{cases} \tag{5.3}$$

We shall use the standard notation of Sobolev spaces $H^s(\Omega)$ equipped with the norm $\|\cdot\|_{s,\Omega}$ for any integer $s \geq 0$. The space $H^0(\Omega)$ coincides with the space $L_2(\Omega)$, and $H_0^1(\Omega)$ is the subspace of $H^1(\Omega)$ that is the closure of the set $C_0^\infty(\Omega)$ of infinite differentiable functions with compact support in Ω. With this notation the problem (5.1)–(5.2) can be formulated in its weak form:

Find $u \in H_0^1(\Omega)$ *such that*

$$a(u,v) = (f,v)_\Omega \quad \forall v \in H_0^1(\Omega), \tag{5.4}$$

where the bilinear form $a(\cdot,\cdot)$ and the linear form $(f,\cdot)$ are given by

$$a(u,v) = \int_\Omega \Big(\sum_{i,j=1}^{2} a_{ij}\, \partial_j u\, \partial_i v + buv\Big)\, dx, \quad (f,v)_\Omega = \int_\Omega fv\, dx. \tag{5.5}$$

It is known that under the assumptions (5.3) the problem (5.4), (5.5) has a unique solution [8]. Under the assumptions that f belongs to $L_2(\Omega)$ and that Ω is convex the solution is H^2-regular. This regular case has been studied in detail in [12]. Here we are interested in the more general case of a non-convex polygon. To simplify the presentation we consider only the case of one re–entrant corner with inner angle $\theta > \pi$ at the origin $(0,0)$.

In order to describe the type of regularity loss, let us take a positive r_0 in such a way that the circumference of the circle with center $(0,0)$ and radius r_0 cuts only a sector ω from the domain Ω. Then we introduce polar coordinates (r,φ), where $x_1 = r\cos\varphi$ and $x_2 = r\sin\varphi$, such that sector ω is described by $0 < r < r_0$ and $0 < \varphi < \theta$. The singular behaviour of the solution in these coordinates is characterized by the following function:

$$\tilde{w}(r,\varphi) = r^\mu \tilde{\xi}(r) \sin \pi\varphi/\theta, \tag{5.6}$$

where the constant $\mu \in (1/2, 1)$ can be given in explicit form [9, 6]; for example in the case of Poisson's equation we have $\mu = \pi/\theta$. The cutoff function $\tilde{\xi}(r) \in C^\infty[0,\infty)$ is given by

$$\tilde{\xi}(r) = \begin{cases} 1 \quad \text{if} \quad r \in [0, r_0/2], \\ \text{monotone if} \quad r \in [r_0/2, r_0], \\ 0 \quad \text{if} \quad r \in [r_0, \infty). \end{cases} \tag{5.7}$$

Using the singular function (5.6), we can represent the solution of (5.1)–(5.2) in the form

$$u(x) = v(x) + \sigma w(x), \tag{5.8}$$

where $w(x_1,x_2) = w(r\cos\varphi, r\sin\varphi) = \tilde{w}(r,\varphi)$, and σ and v denote a constant and the regular part of the solution, respectively, that satisfy

$$|\sigma| + \|v\|_{2,\Omega} \leq c\,\|f\|_{0,\Omega}. \tag{5.9}$$

The regularity properties of the solution u can also be described by special spaces with weighted norms [1]. For this purpose, we introduce the weighting function

$$\Phi_\beta(x_1,x_2) = \Phi_\beta(r\cos\varphi, r\sin\varphi) = \tilde{\Phi}_\beta(r,\varphi) = r^\beta \quad \text{for } \beta \in (-\infty,\infty) \tag{5.10}$$

and denote by $H^{m,\beta}(\Omega)$, for $m \geq 0$ and $0 \leq \beta < 1$, the closure of $C^\infty(\Omega)$ in the norm $\|\cdot\|_{m,\beta}$ defined by

$$\|u\|^2_{m,\beta} = \|u\|^2_{m-1} + \int_\Omega \Phi^2_\beta \sum_{m_1+m_2=m} (\partial_1^{m_1}\partial_2^{m_2} u)^2\, dx,$$
$$m \geq 1, \quad m_1 \geq 0, \quad m_2 \geq 0,$$
$$\|u\|^2_{0,\beta} = \int_\Omega \Phi^2_\beta u^2\, dx.$$

Then we have the following regularity result [7, Theorem 1.1]:

Theorem 5.1. *Suppose that* $1 - \mu < \beta < 1$ *and*

$$f \in H^{0,\beta}(\Omega). \tag{5.11}$$

Then the solution u *of* (5.4) *belongs to* $H^{2,\beta}(\Omega)$, *and there is a constant* $c_1 > 0$, *which is independent of* f, *such that*

$$\|u\|_{2,\beta} \leq c_1 \|f\|_{0,\beta}. \tag{5.12}$$

Remark 5.2. Note that $L_2(\Omega) \subset H^{0,\beta}(\Omega)$ for $\beta \geq 0$. Therefore the condition

$$f \in L_2(\Omega) \tag{5.13}$$

is sufficient for (5.11), and a positive constant c_2 exists such that

$$\|f\|_{0,\beta} \leq c_2 \|f\|_0. \tag{5.14}$$

6. The mesh refinement strategy

Standard finite element triangulations result in optimal convergence rates, provided that the solution is sufficiently regular. A reduction of the convergence rate can be observed both theoretically and numerically when the solution is not H^2-regular. Consequently special mesh refinement strategies have been developed to guarantee optimal convergence rates [9, 1, 14]. Standard techniques for adapting the grid in the neighbourhood of singular points lead to a family of meshes with optimal order of convergence, but not necessarily to a family of nested finite element spaces (see, e.g., [9]). In this section we derive a special mesh refinement technique which guarantees both optimal order of convergence and a nested family of finite element spaces.

We shall follow a technique described in [13]. Let us start with an initial *admissible* triangulation $\mathcal{F}_0$ of Ω into closed triangles, i.e., each pair of triangles has either no common points or a common vertex or a common edge. In order to define the refinement near the singular point $A = (0,0)$, we introduce the refinement index $\rho \geq 1$. Then every initial triangle is divided into 4 finer ones. If a triangle of refinement level i, where $i = 0, 1, \cdots, l-1$, does not contain the origin $(0,0)$, it is subdivided into 4 triangles by connecting the midpoints of its edges. This type of refinement is called a regular partition of the domain. Now let $\triangle AB_iC_i$ be a triangle of refinement level i and have a vertex $A = (0,0)$ (see Figure 1). We denote by b_0 the length of the perpendicular from A to the opposite side B_0C_0 of the initial triangle $\triangle AB_0C_0$. Then we construct the straight segment $B_{i+1}C_{i+1}$ parallel to B_0C_0 at a distance $b_0 2^{-(i+1)\rho}$ from A, with ends B_{i+1}, C_{i+1} on the edges AB_0, AC_0 respectively. The midpoint of B_iC_i is denoted by A_{i+1}. Connecting the points

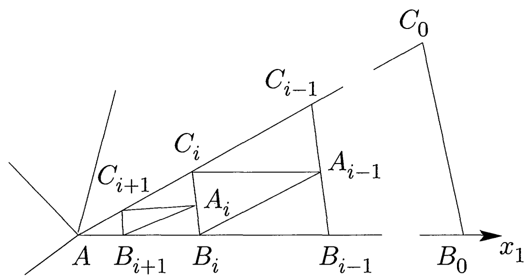

FIGURE 1. Irregular partition of triangle.

A_{i+1}, B_{i+1}, C_{i+1}, we get 4 finer triangles of the triangulation $\mathcal{F}_{i+1}$ at refinement level $(i+1)$. This type of refinement is called an irregular partition.

Thus, starting with the initial admissible triangulation $\mathcal{F}_0$ of Ω, we end up with a family of admissible triangulations $\mathcal{F}_i$ for any i, where $i = 1, 2, \cdots$. Before we show that $\mathcal{F}$ is a triangulation of the type considered in [1], we note an important property of the locally refined meshes constructed above. If we define the maximal length of all edges of the triangulation $\mathcal{F}_i$ by h_i, then the relation

$$h_i = h_1 2^{-i+1} \tag{6.1}$$

holds true for $i = 1, 2, \cdots$. This can be seen in the following way. On the triangulation $\mathcal{F}_i$, for $i = 0, 1, \cdots$, all edges of triangles sharing A as a vertex have length less than or equal to h_i. Now let us consider the next triangulation $\mathcal{F}_{i+1}$. Since the distance from each parallel line B_iC_i to A is $b_0 2^{-(i+1)\rho}$, for $i = 0, 1, \cdots$, we can show by an elementary argument that

$$|C_{i+1}C_i| \leq |C_iC_{i-1}|/2, \quad |C_{i+1}A_i| \leq |C_iA_{i-1}|/2 \,.$$
$$|B_{i+1}A_i| \leq |B_iA_{i-1}|/2, \quad |B_{i+1}B_i| \leq |B_iB_{i-1}|/2 \,.$$

Therefore we have

$$h_{i+1} = h_i/2 \,.$$

Thus by induction we get (6.1). To classify the properties of the triangulation we shall use the definition given in [1]:

A triangulation $\mathcal{F}$ is said to be of type (h, γ, L) if it satisfies the following three properties:

(i): for any triangle $\triangle \in \mathcal{F}$ and for any angle α of $\triangle$, we have

$$\alpha \geq L^{-1}; \tag{6.2}$$

(ii): if $\Phi_\gamma \neq 0$ on $\triangle$, then

$$L^{-1}\Phi_\gamma(x) \leq d_\triangle/h \leq L\Phi_\gamma(x), \tag{6.3}$$

where $d_\triangle = \sup\{|x-y| : x, y \in \triangle\}$ is the diameter of $\triangle$;

(iii): if $\Phi_\gamma = 0$ at some point of $\triangle$, then

$$L^{-1} \sup_{x\in\triangle} \Phi_\gamma(x) \leq d_\triangle/h \leq L \sup_{x\in\triangle} \Phi_\gamma(x). \tag{6.4}$$

Lemma 6.1. *Let ρ be the refinement index and $\gamma = (\rho-1)/\rho$. Then each triangulation $\mathcal{F}_i$ is of type (h_i, γ, L) with a constant L independent of i.*

Proof. First, we note that in each initial triangle there are only 4 different types of geometrically similar triangles at any level of triangulation. They are similar to the triangles of $\mathcal{F}_1$. Therefore their angles are independent of the level number i. Thus (6.2) holds true with

$$L \geq L_0 = \max_{\triangle \in \mathcal{F}_1} \alpha^{-1}. \tag{6.5}$$

Now let $x \in \triangle_j \in \mathcal{F}_j$ and let $\Phi_\gamma \neq 0$ on $\triangle_j$, i.e., the point A does not belong to the $\triangle_j$. We consider all triangles $\triangle_k$ on lower levels that contain this x:

$$x \in \triangle_j \subset \cdots \subset \triangle_{i+1} \subset \triangle_i \subset \cdots \subset \triangle_0 \in \mathcal{F}_0. \tag{6.6}$$

There are two cases. First, $\triangle_0$ does not contain A. This means that $\Phi_\gamma \geq c$ on $\triangle_0$ (and in particular on $\triangle_j$) for some constant c that is independent of j. Moreover, we have $d_{\triangle_j} = d_{\triangle_1} h_j / h_1$ for $j = 1, 2, \cdots$. Both these facts follow from (6.3) for this type of triangle. Second, $\triangle_0$ contains A. Then there is an index i such that

$$A \notin \triangle_{i+1} \quad \text{but} \quad A \in \triangle_i. \tag{6.7}$$

The notation used in the following is given in Figure 1. The statement (6.7) implies that x belongs to the trapezium $C_{i+1}B_{i+1}B_iC_i$. Therefore the inequality $|x - A| \geq b_0 2^{-(i+1)\rho}$ holds true. Hence, we have

$$b_0^\gamma 2^{-(i+1)(\rho-1)} \leq \Phi_\gamma(x) \leq d_{\triangle_i}^\gamma \tag{6.8}$$

by definition of Φ_γ. On the one hand, $d_{\triangle_{i+1}}$ can be estimated from above by

$$d_{\triangle_{i+1}} \leq d_{\triangle_i} = d_{\triangle_0} 2^{-i\rho}. \tag{6.9}$$

Taking

$$L \geq L_1 = d_{\triangle_0} h_1^{-1} b_0^{-\gamma} 2^{\rho-1},$$

we get

$$d_{\triangle_{i+1}} / h_{i+1} \leq L_1 \Phi_\gamma(x) \leq L \Phi_\gamma(x). \tag{6.10}$$

On the other hand, $d_{\triangle_{i+1}}$ can be estimated from below by

$$d_{\triangle_{i+1}} \geq d_{\triangle_i} / 2^\rho = 2^{-(i+1)\rho} d_{\triangle_0}.$$

Taking

$$L^{-1} \leq L_2^{-1} = d_{\triangle_0}^{1-\gamma} h_1^{-1} 2^{-\rho},$$

we get the inequalities

$$d_{\triangle_{i+1}} / h_{i+1} \geq L_2^{-1} \Phi_\gamma(x) \geq L^{-1} \Phi_\gamma(x). \tag{6.11}$$

Thus (6.3) is proved at the level $i+1$. At the higher levels $j > i+1$, we have

$$d_{\triangle_j} / h_j = d_{\triangle_{i+1}} / h_{i+1}$$

by construction. Therefore (6.10) and (6.11) give us

$$L^{-1} \Phi_\gamma(x) \leq L_2^{-1} \Phi_\gamma(x) \leq d_{\triangle_j} / h_j \leq L_1 \Phi_\gamma(x) \leq L \Phi_\gamma(x) \tag{6.12}$$

for all j when $\Phi_\gamma \neq 0$ on $\triangle_j$.

Finally, let $\Phi_\gamma(x) = 0$ at some point $x \in \triangle$. This means that $x = A$. For instance, let $x = A \in \triangle_{i+1} = \triangle A B_{i+1} C_{i+1}$ in Figure 1. By analogy with (6.8), we have

$$b_0^\gamma 2^{-(i+1)(\rho-1)} \leq \sup_{\triangle_{i+1}} \Phi_\gamma \leq d_{\triangle_{i+1}}^\gamma. \tag{6.13}$$

Let us choose

$$L \geq L_3 = d_{\triangle_0}/(2h_1 b_0^{\gamma}) .$$

Since

$$d_{\triangle_{i+1}} = d_{\triangle_0} 2^{-(i+1)\rho}, \tag{6.14}$$

we get

$$d_{\triangle_{i+1}}/h_{i+1} \leq L_3 \sup_{\triangle_{i+1}} \Phi_\gamma \leq L \sup_{\triangle_{i+1}} \Phi_\gamma. \tag{6.15}$$

Then, for L satisfying

$$L^{-1} \leq L_4^{-1} = d_{\triangle_0}^{1-\gamma}/2h_1 ,$$

and taking into consideration (6.13) and (6.14), we have

$$d_{\triangle_{i+1}}/h_{i+1} \geq L_4^{-1} \sup_{\triangle_{i+1}} \Phi_\gamma \geq L \sup_{\triangle_{i+1}} \Phi_\gamma. \tag{6.16}$$

Therefore, (6.2)–(6.4) are proved if L is sufficiently large; more precisely, the choice

$$L = \max\{L_0, \cdots, L_4\}$$

is sufficient to guarantee (6.2)–(6.4). □

Thus conditions (i)–(iii) of [1] are satisfied and we can use its results. For example, there is a constant c_3 independent of h_i such that the number n_i of interior vertices in $\mathcal{F}_i$ can be estimated in the following way:

$$n_i \leq c_3 h_i^{-2}. \tag{6.17}$$

Now we derive the Galerkin approximation based on the triangulation described above. We denote by $\bar{\Omega}_i$ the set of all nodes of the triangulation $\mathcal{F}_i$, and by Ω_i the set of all interior nodes. For each node $y \in \Omega_i$, we define the basis function $\varphi_y^i \in H_0^1(\Omega)$ by requiring it to be linear on each triangle of the triangulation $\mathcal{F}_i$, to equal 1 at the node y and to equal 0 at every other node $z \in \bar{\Omega}_i$. We denote the linear span of these functions by

$$\mathcal{H}^i = \text{span}\,\{\varphi_y^i : \ y \in \Omega_i\}.$$

Considering (5.4) on the subspace $\mathcal{H}^i \subset H_0^1(\Omega)$, we get the discrete problem:

Find $\tilde{u}_i \in \mathcal{H}^i$ such that

$$a(\tilde{u}_i, v) = (f, v)_\Omega \quad \forall v \in \mathcal{H}^i. \tag{6.18}$$

Let M_i be the n_i-dimensional vector space of all vectors $w = (w(x) : x \in \Omega_i)$.

Then the formulation (6.18) is equivalent to the linear system of algebraic equations

$$L_i u_i = f_i , \tag{6.19}$$

where $u_i \in M_i$ is the vector of unknowns with components $u_i(y)$, $y \in \Omega_i$; $f_i \in M_i$ is defined by $f_i(x) = (f, \varphi_x^i)_\Omega$ for all $x \in \Omega_i$; L_i is the matrix whose elements are

$$L_i(x, y) = a(\varphi_y^i, \varphi_x^i), \quad \forall x, y \in \Omega_i. \tag{6.20}$$

Let us define the usual isomorphism $\mathcal{J}_i$ between vectors $v \in M_i$ and functions $\tilde{v} \in \mathcal{H}^i$ that are their prolongations, i.e.,

$$\tilde{v} = \mathcal{J}_i v \quad \textit{means} \quad \tilde{v}(x) = \sum_{y \in \Omega_i} v(y)\varphi_y^i(x), \quad \forall x \in \bar{\Omega}, \tag{6.21}$$

and, vice versa,

$$v = \mathcal{J}_i^{-1}\tilde{v} \quad \textit{is defined by} \quad v(y) = \tilde{v}(y), \quad \forall y \in \Omega_i. \tag{6.22}$$

Now we introduce the energy norm for functions belonging to $H_0^1(\Omega)$:

$$|||\tilde{v}|||_\Omega = \mathcal{L}(\tilde{v}, \tilde{v})^{1/2},$$

and specify the inner product and the norm for vectors in M_i:

$$(v, w)_i = \sum_{x \in \Omega_i} v(x)w(x),$$

$$\|v\|_i = (v, v)_i^{1/2}, \quad \forall v, w \in M_i.$$

From (6.20) and (6.21), we have *for an isomorphic pair* $v \in M_i$ *and* $\tilde{v} = \mathcal{J}_i v \in \mathcal{H}^i$ the relationship

$$|||v|||_i = |||\tilde{v}|||_\Omega. \tag{6.23}$$

Now let us introduce the interpolation operator $I_i : M_i \to M_{i+1}$. Let $v \in M_i$. Since its prolongation $\tilde{v}$ belongs to $\mathcal{H}^{i+1}$, the isomorphism associates with $\tilde{v}$ a vector $w \in M_{i+1}$. In such a way, we have uniquely defined $I_i : v \to w$.

The convergence of the Bubnov-Galerkin solution to the exact solution was studied in a number of papers (e.g., [3]). A standard analysis on a quasi-uniform triangulation gives a non-optimal convergence rate because of a loss of regularity in the exact solution u. In our case we can however use the special nested triangulations that are refined towards the singular point and hence obtain the optimal first-order convergence.

Lemma 6.2. *Let the triangulations* $\mathcal{F}_i$ *at the levels* $i = 1, \cdots, l$ *be generated with a refinement index*

$$\rho > 1/\mu \tag{6.24}$$

and let $\beta = (\rho - 1)/\rho$. *Suppose that the assumptions* (5.3) *are satisfied and that* $f \in H^{0,\beta}(\Omega)$. *Then the solution* $\tilde{u}_i$ *of the Galerkin problem* (6.18) *satisfies the error estimate*

$$|||u - \tilde{u}_i|||_\Omega \le c_4 h_i \|f\|_{0,\beta} \,. \tag{6.25}$$

Proof. Let us note that β satisfies $1 - \mu < \beta < 1$. This means that we can apply Theorem 5.1 and prove existence of a solution of the continuous problem in $H^{2,\beta}(\Omega)$. Using Lemma 4.5 from [1], we can estimate the interpolation error by

$$\|u - \tilde{v}_i\|_{1,\Omega} \le c h_i \|u\|_{2,\beta} \,. \tag{6.26}$$

Here $\tilde{v}_i$ denotes the piecewise linear interpolant of u on $\mathcal{F}_i$. Next, applying (5.12) and using the equivalence of the norms $|||\cdot|||_\Omega$ and $\|\cdot\|_{1,\Omega}$, we get

$$|||u - \tilde{v}_i|||_\Omega \le c_4 h_i \|f\|_{0,\beta}. \tag{6.27}$$

Finally, the optimality of the Galerkin solution $\tilde{u}_i$ implies that

$$|||u - \tilde{u}_i|||_\Omega \le |||u - \tilde{v}_i|||_\Omega \le c_4 h_i \|f\|_{0,\beta}.$$

□

In the following we need

Lemma 6.3. *Under the assumptions of Lemma 6.2, the solutions* $\tilde{u}_{i-1}$, $\tilde{u}_i$ *of the Galerkin problem* (6.18) *in* $\mathcal{H}^{i-1}$ *and* $\mathcal{H}^i$ *satisfy the inequalities*

$$\|\tilde{u}_i - \tilde{u}_{i-1}\|_{0,-\beta} \le c_5 h_i |||\tilde{u}_i - \tilde{u}_{i-1}|||_\Omega, \tag{6.28}$$

$$|||\tilde{u}_i - \tilde{u}_{i-1}|||_\Omega \le |||u - \tilde{u}_{i-1}|||_\Omega. \tag{6.29}$$

Proof. We shall use a duality argument as in [1]. Let z denote the solution of the problem

$$a(z, v) = (g, v)_\Omega \quad \forall v \in H_0^1(\Omega) \tag{6.30}$$

with the right-hand side

$$g = \Phi_\beta^{-2}(\tilde{u}_i - \tilde{u}_{i-1}).$$

We show that $g \in H^{0,\beta}$. From [1] we know that there is a constant c such that for all $v \in H^1(\Omega)$ we have

$$\|v\|_{0,-\beta} \le c\,\|v\|_1.$$

Setting $v = \tilde{u}_i - \tilde{u}_{i-1}$, we get

$$\|g\|_{0,\beta} = \|\tilde{u}_i - \tilde{u}_{i-1}\|_{0,-\beta} \le \|\tilde{u}_i - \tilde{u}_{i-1}\|_1 < \infty. \tag{6.31}$$

Let $\tilde{z}_{i-1}$ be the Galerkin solution of the problem

$$a(\tilde{z}_{i-1}, v) = (g, v)_\Omega \quad \forall v \in \mathcal{H}^{i-1}. \tag{6.32}$$

Applying Lemma 6.2, we obtain

$$|||z - \tilde{z}_{i-1}|||_\Omega \le c_4 h_{i-1} \|\tilde{u}_i - \tilde{u}_{i-1}\|_{0,-\beta}. \tag{6.33}$$

From (6.30) and (6.18) we get the representation

$$\|\tilde{u}_i - \tilde{u}_{i-1}\|_{0,-\beta}^2 = a(z - \tilde{z}_{i-1}, \tilde{u}_i - \tilde{u}_{i-1}).$$

By means of the Cauchy-Bunjakovski inequality we obtain

$$\begin{aligned} \|\tilde{u}_i - \tilde{u}_{i-1}\|_{0,-\beta}^2 &\le |||z - \tilde{z}_{i-1}|||_\Omega |||\tilde{u}_i - \tilde{u}_{i-1}|||_\Omega \\ &\le c_4 h_{i-1} \|\tilde{u}_i - \tilde{u}_{i-1}\|_{0,-\beta} |||\tilde{u}_i - \tilde{u}_{i-1}|||_\Omega, \end{aligned}$$

which yields (6.28) with $c_5 = 2c_4$.

In order to prove (6.29), we set $v = \tilde{u}_i - \tilde{u}_{i-1}$ in (5.4) and (6.18); this gives

$$a(u, \tilde{u}_i - \tilde{u}_{i-1}) = a(\tilde{u}_i, \tilde{u}_i - \tilde{u}_{i-1}),$$

and taking into consideration

$$a(\tilde{u}_{i-1}, \tilde{u}_i - \tilde{u}_{i-1}) = 0,$$

we obtain the representation

$$a(u - \tilde{u}_i, \tilde{u}_i - \tilde{u}_{i-1}) = a(\tilde{u}_i - \tilde{u}_i, \tilde{u}_i - \tilde{u}_{i-1}).$$

Estimating the left-hand side by the Cauchy-Bunjakovski inequality,

$$\begin{aligned} |||\tilde{u}_i - \tilde{u}_{i-1}|||_\Omega^2 &= a(u - \tilde{u}_{i-1}, \tilde{u}_i - \tilde{u}_{i-1}) \\ &\le |||u - \tilde{u}_{i-1}|||_\Omega |||\tilde{u}_i - \tilde{u}_{i-1}|||_\Omega, \end{aligned}$$

we finally get (6.29). □

In order to check the convergence criterion (4.1) in the next section, we need a result on the equivalence of norms.

Lemma 6.4. *Let $v \in M_i$, $\tilde{v} = \mathcal{J}_i v \in \mathcal{H}^i$ be an isomorphic pair, let $\rho \geq 1$ be the refinement index of the triangulation, and $\beta = (\rho - 1)/\rho$. Then there are constants $c_7, c_8 > 0$, which are independent of i, v, and $\tilde{v}$, such that*

$$c_7 \|\tilde{v}\|_{0,-\beta} \leq h_i \|v\|_i \leq c_8 \|\tilde{v}\|_{0,-\beta}. \tag{6.34}$$

Proof. First we consider a triangle $\triangle a_1 a_2 a_3 \in \mathcal{F}_i$ that does not contain $A = (0,0)$. Then we have (6.3), which gives

$$L^{-2} \frac{h_i^2}{d_\triangle^2} \int_\triangle \tilde{v}^2 \, dx \leq \int_\triangle \tilde{v}^2 \Phi_\beta^{-2} \, dx \leq L^2 \frac{h_i^2}{d_\triangle^2} \int_\triangle \tilde{v}^2 \, dx \, .$$

From (6.2), we have the affine regularity of $\triangle$ and get, analogously to Theorem 3.13 of [13],

$$\begin{aligned} c_9^{-1} d_\triangle^2 (v^2(a_1) + v^2(a_2) + v^2(a_3)) &\leq \int_\triangle \tilde{v}^2 \, dx \\ &\leq c_9 d_\triangle^2 (v^2(a_1) + v^2(a_2) + v^2(a_3)), \end{aligned} \tag{6.35}$$

with a constant c_9 that is independent of $\triangle$. Thus

$$\begin{aligned} c_9^{-1} L^{-2} h_i^2 (v^2(a_1) + v^2(a_2) + v^2(a_3)) &\leq \int_\triangle \tilde{v}^2 \Phi_\beta^{-2} \, dx \\ &\leq c_9 L^2 h_i^2 (v^2(a_1) + v^2(a_2) + v^2(a_3)). \end{aligned} \tag{6.36}$$

Second, we consider a triangle $\triangle \in \mathcal{F}_i$ that contains $A = (0,0)$, e.g., $\triangle A B_i C_i$. From the left-hand side of (6.4), we get

$$L^{-1} \Phi_\beta(x) \leq d_\triangle / h_i \quad \forall x \in \triangle.$$

Therefore we obtain as above the left-hand side of (6.36) for this $\triangle$. It remains to show the second inequality of (6.36). For this we introduce barycentric coordinates $\lambda_1(x)$, $\lambda_2(x)$, $\lambda_3(x)$ corresponding to the vertices B_i, C_i, A respectively. Since $v(A) = 0$, we have

$$\tilde{v}(x) = v(B_i)\lambda_1(x) + v(C_i)\lambda_2(x) \quad \forall x \in \triangle. \tag{6.37}$$

Let $C_i = (b_1, b_2)$. Then

$$|\lambda_1(x)| = |x_1 b_2 - x_2 b_1| / (2 \, \text{meas} \triangle) \, .$$

From the Cauchy-Bunjakovski inequality, we obtain

$$|\lambda_1(x)| \leq r(x) d_\triangle / (2 \text{meas} \triangle) \quad \forall x \in \triangle.$$

The Law of Sines and (6.2) imply that

$$2 \, \text{meas} \triangle \geq d_\triangle^2 \sin^2(1/L).$$

Consequently we have

$$|\lambda_1(x)| \leq r(x) / (d_\triangle \sin^2(1/L)). \tag{6.38}$$

Analogously, we get

$$|\lambda_2(x)| \leq r(x) / (d_\triangle \sin^2(1/L)). \tag{6.39}$$

From the second inequality of (6.4), we have

$$d_\triangle / h_i \leq L \sup_\triangle r^\beta(x) \leq L d_\triangle^\beta, \quad \text{i.e.,} \quad d_\triangle^{1-\beta} \leq L h_i. \tag{6.40}$$

Using (6.37)–(6.39), we get

$$
\begin{aligned}
\int_{\triangle} \tilde{v}^2(x) r^{-2\beta}(x)\,dx &= \int_{\triangle} r^{-2\beta}(x)(v(B_i)\lambda_1(x) + v(C_i)\lambda_2(x))^2\,dx \\
&\le c_{10} d_{\triangle}^{-2}(|v(B_i)| + |v(C_i)|)^2 \int_{\triangle} r^{2-2\beta}(x)\,dx \\
&\le 2c_{10} d_{\triangle}^{-2}(v^2(B_i) + v^2(C_i)) \int_0^{\theta}\int_0^{d_{\triangle}} r^{3-2\beta}\,dr d\varphi \\
&\le c_{10}\theta d_{\triangle}^{2-2\beta}(v^2(B_i) + v^2(C_i)) \\
&\le c_{10}\theta L^2 h_i^2 (v^2(B_i) + v^2(C_i)) ,
\end{aligned}
\tag{6.41}
$$

where $c_{10} = \sin^{-2}(1/L)$. Now we are able to prove the second inequality of (6.34) by summing the left-hand side of (6.36) over all triangles $\triangle \in \mathcal{F}_i$:

$$
c_9^{-1} L^{-2} h_i^2 \|v\|_i^2 = c_9^{-1} L^{-2} h_i^2 \sum_{y\in\Omega_i} v^2(y) \le \int_{\Omega} \tilde{v}^2 \Phi_{\beta}^{-2}\,dx = \|\tilde{v}\|_{0,-\beta}^2.
$$

Thus we can take $c_8 = \sqrt{c_9}\,L$ in (6.34). To prove the first inequality of (6.34), we sum the second inequality of (6.36) over all triangles $\triangle \in \mathcal{F}_i$ that do not contain A and (6.41) over all triangles $\triangle \in \mathcal{F}_i$ that do contain A. This gives

$$
\|\tilde{v}\|_{0,-\beta}^2 = \int_{\Omega} \tilde{v}^2(x) r^{-2\beta}(x)\,dx \le c_{11} k L^2 h_i^2 \sum_{y\in\Omega_i} v^2(y) = c_{11} k L^2 h_i^2 \|v\|_i^2 ,
$$

where $c_{11} = \max\{c_9, c_{10}\theta\}$ and k is the maximum number of triangles having a common vertex. Therefore we can take $c_7 = c_{11}^{-1/2} k^{-1/2} L^{-1}$ in (6.34). □

7. The main convergence result

Now we are ready to apply the abstract convergence result of Theorem 4.1 to the boundary value problem.

Theorem 7.1. *Let the assumptions* (5.3) *and* (5.13) *on the data be satisfied and let the refinement index* ρ *of the triangulation be greater than* $1/\mu$. *Then for the CCG-algorithm with* m_j *iterations on each level* j *for* $j = 1, \cdots, l$, *we have the error estimate*

$$
|||u_l - v_l|||_l = |||\tilde{u}_l - \tilde{v}_l|||_{\Omega} \le c_{12} \sum_{j=1}^{l} \frac{h_j}{2m_j + 1} \|f\|_0, \tag{7.1}
$$

where $\tilde{v}_l = \mathcal{J}_l v_l$ *and* v_l *denotes the final approximation of the CCG-algorithm.*

Proof. Let us note that for $i = 1, \cdots, l$ we have the usual estimate

$$
0 < \lambda_i^* \le c_6 \tag{7.2}
$$

for the largest eigenvalue λ_i^* of the matrix L_i (see, e.g., [13, Theorem 3.14]). Taking into consideration (6.28) and (6.23), we get

$$
\|\tilde{u}_i - \tilde{u}_{i-1}\|_{0,-\beta} \le c_5 h_i |||u_i - I_{i-1}u_{i-1}|||_i. \tag{7.3}
$$

The norm equivalence (6.34) gives

$$
h_i \|u_i - I_{i-1}u_{i-1}\|_i \le c_8 \|\tilde{u}_i - \tilde{u}_{i-1}\|_{0,-\beta}. \tag{7.4}
$$

Combining these inequalities with (7.2), we obtain

$$\|u_i - I_{i-1}u_{i-1}\|_i \le c_5 c_8 \sqrt{\frac{c_6}{\lambda_i^*}} |||u_i - I_{i-1}u_{i-1}|||_i. \tag{7.5}$$

This proves (4.1) for $\gamma = 1$ and $c^* = c_5 c_8 \sqrt{c_6}$. Since all the assumptions of Theorem 4.1 are satisfied, we get

$$|||u_l - v_l|||_l \le c^* \sum_{j=1}^{l} \frac{1}{2m_j + 1} |||\tilde{u}_j - \tilde{u}_{j-1}|||_\Omega$$

from (3.9) and (6.23). Owing to (6.29) and (6.25), we have (7.1) with $c_{12} = c^* c_2 c_4$. □

Now we count the number N_{CCG} of arithmetic operations for the full CCG-algorithm and try to choose m_j in an optimal way. Upper bounds for the arithmetic operations have been considered, e.g., in [12], [13], and are given by

$$N_{CCG} \le N(m_1, \cdots, m_l) := d_1 \sum_{j=1}^{l} (m_j + d_2) n_j + d_3 \tag{7.6}$$

with constants d_1, d_2 and d_3 that are independent of m_j, n_j and h_j. On the other hand, the accuracy of the CCG-algorithm is characterized by the value $\varepsilon = \varepsilon(m_1, \dots, m_l)$:

$$|||\tilde{u}_l - \tilde{v}_l|||_\Omega \le \varepsilon(m_1, \cdots, m_l) := d_4 \sum_{j=1}^{l} \frac{h_j}{2m_j + 1} \tag{7.7}$$

with the constant $d_4 = c_{12}\|f\|_0$. This leads us to solve the optimization problem:

Find $m_1, \cdots, m_l$ such that the value $\varepsilon(m_1, \cdots, m_l)$ is minimal under the constraint

$$N(m_1, \cdots, m_l) \le d_5 \tag{7.8}$$

for an appropriate constant $d_5 > 0$.

Applying the method of Lagrange multipliers, we obtain

$$(2m_j + 1)^2 = d_6 h_j / n_j, \qquad \text{for } j = 1, \dots, l, \tag{7.9}$$

with a constant d_6 that depends on $d_1, \cdots, d_5$ but is independent of m_j. We eliminate d_6 by using (7.9) for $j = l$, and get

$$2m_j + 1 = (2m_l + 1)\sqrt{h_j n_l / h_l n_j}. \tag{7.10}$$

Since $h_j = 2^{l-j} h_l$ and $n_l \approx 4^{l-j} n_j$, we obtain

$$2m_j + 1 \approx (2m_l + 1) 2^{3(l-j)/2}. \tag{7.11}$$

We use this relation in the following way. Fixing the number of iterations at the highest level l by setting $m_l = m$, we choose m_j at the lower levels as the smallest integer m_j such that

$$2m_j + 1 \ge (2m_l + 1) 2^{3(l-j)/2}, \qquad \text{for } j = 1, \cdots, l-1. \tag{7.12}$$

Theorem 7.2. *In addition to the assumptions of Theorem 7.1, let the number m_j of iterations in the CCG-algorithm be defined by* (7.12). *Then the error of the final approximation v_l in the CCG-algorithm satisfies the estimates*

$$|||u_l - v_l|||_l = |||\tilde{u}_l - \tilde{v}_l|||_\Omega \leq c_{13} \frac{h_l}{2m+1} \|f\|_0 \tag{7.13}$$

and

$$|||u - \tilde{v}_l|||_\Omega \leq h_l(c_2 c_4 + \frac{c_{13}}{2m+1}) \|f\|_0 \,, \tag{7.14}$$

where $\tilde{v}_l = \mathcal{J}_l v_l$ and the constants c_2 and c_4 are given in (5.14) *and* (6.25). *The number of arithmetic operations is bounded by*

$$N_{CCG} \leq (c_{14} m + c_{15}) n_l; \tag{7.15}$$

the constants c_{13}–c_{15} are independent of the number of levels, the number of CG-iterations on the highest level m, and the number of unknowns on the highest level n_l.

Proof. Using (7.12) and the relation $h_j = 2^{l-j} h_l$ in (7.1), we get

$$|||\tilde{u}_l - \tilde{v}_l|||_\Omega \leq c_{12} \frac{h_l}{2m+1} \sum_{j=1}^{l} 2^{-(l-j)/2} \|f\|_0 \,.$$

Summing, we get (7.13) with the constant $c_{13} = c_{12}/(1 - 2^{-1/2})$. From the Euler formula for a polyhedron, we have

$$n_l \geq 4^{l-j} n_j. \tag{7.16}$$

The definition of m_j implies that

$$2m_j + 1 \leq (2m+1) 2^{3(l-j)/2} + 2.$$

Using these inequalities in (7.6), we obtain the estimate

$$N_{CCG} \leq d_1 n_l \sum_{j=1}^{l} ((2m+1) 2^{3(l-j)/2} - 1/2 + d_2) 2^{-2(l-j)} + d_3.$$

Calculating the sum, we finally get (7.15) with constants

$$c_{14} = 2/(1 - 2^{-1/2}) \quad \text{and} \quad c_{15} = 1/(1 - 2^{-1/2}) + (4d_2 - 2)/3 + d_3. \quad \square$$

Remark 7.3. Now choosing m in order to get $c_2 c_4 \sim c_{13}/(2m+1)$, we see that the amount of arithmetic operations is proportional only to n_l.

Remark 7.4. The inequality (7.14) shows that the final numerical solution produced by the CCG-method is, in the energy norm, of the same order of magnitude as the discretization error of the finite element method. Nevertheless, this approximate solution $\tilde{v}_l$ is not the finite element solution and may not have a higher-order L^2 norm error nor exhibit superconvergence. In this sense, the CCG-method is not as good as multigrid V-cycle iterations or the CG-method with a V-cycle preconditioner [13], [14].

Remark 7.5. The inequality (7.12) gives an upper bound for the estimated number of iterations on each level. For the coarse grids this bound may be too pessimistic (i.e., much larger than the number of unknowns).

Acknowledgments

The authors acknowledge Prof. J. Xu for the discussion that motivated these investigations. Moreover, the authors wish to thank the German Research Association (DFG) for supporting the research presented in this paper.

References

1. I. Babuška, R.B. Kellogg, and J. Pitkäranta, *Direct and inverse error estimates for finite elements with mesh refinements*, Numer. Math. **33** (1979), 447–471. MR **81c**:65054
2. F.A. Bornemann, *On the convergence of cascadic iterations for elliptic problems*, Tech. Report SC 94-8, Konrad-Zuse-Zentrum Berlin (ZIB), 1994.
3. P. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978. MR **58**:25001
4. P. Deuflhard, *Cascadic Conjugate Gradient Methods for Elliptic Partial Differential Equations I. Algorithm and Numerical Results*, Tech. Report SC 93–23, Konrad-Zuse-Zentrum Berlin (ZIB), 1993.
5. ______, *Cascadic conjugate gradient methods for elliptic partial differential equations. Algorithm and numerical results*, Domain decomposition methods in scientific and engineering computing. Proceedings of the 7th international conference on domain decomposition, October 27-30, 1993, Pennsylvania State University, (D. Keyes and J. Xu, eds.), Contemp. Math., vol. 180, Providence, RI: American Mathematical Society, 1994, pp. 29–42. MR **95i**:65008
6. P. Grisvard, *Singularities in Boundary Value Problems*, Springer-Verlag, Berlin, and Masson, Paris, 1992. MR **93h**:35004
7. V.A. Kondrat'ev, *Boundary problems for elliptic equations in domains with conical or angular points*, Trans. Mosc. Math. Soc. **16** (1967), 227–313. MR **37**:1777
8. O.A. Ladyzhenskaya and N.N. Uraltseva, *Linear and quasilinear elliptic equations*, Academic Press, New York, 1968. MR **39**:5941
9. L.A. Oganesjan and L.A. Rukhovets, *Variational Difference Methods of Solving the Elliptic Equations (in Russian)*, Acad. Sci. of the Armenian SSR, Erevan, 1979. MR **92m**:65101
10. A.A. Samarskii and E.S. Nikolaev, *Numerical Methods for Grid Equations.* Vol. II, *Iterative Methods*, Birkhäuser, Basel, 1989. MR **90m**:65003
11. V.V. Shaidurov, *The convergence of the cascadic conjugate-gradient method under a deficient regularity*, Problems and Methods in Mathematical Physics (L.Jentsch and F.Tröltzsch, eds.), Teubner, Stuttgart, 1994, 185–194. MR **95f**:65217
12. ______, *Some estimates of the rate of convergence for the cascadic conjugate-gradient method*, Comput. Math. Appl. **31** (1996), No. 4-5, 161–171. MR **96j**:65003
13. ______, *Multigrid methods for finite elements*, Kluwer Academic Publishers, Dordrecht, 1995. MR **97e**:65142
14. H. Yserentant, *The convergence of multi-level methods for solving finite-element equations in the presence of singularities*, Math. Comp. **47** (1986), 399–409. MR **88d**:65149

Institute of Computational Modelling, Siberian Branch of the Russian Academy of Sciences, Krasnoyarsk 660036, Russia

E-mail address: shidurov@cckr.krasnoyarsk.su

Otto-von-Guericke-Universität Magdeburg, Postfach 4120, D-39016 Magdeburg, Germany

E-mail address: tobiska@mathematik.uni-magdeburg.de