Numer. Math. (1996) 75: 135–152

Numerische Mathematik

Electronic Edition

The cascadic multigrid method for elliptic problems

Folkmar A. Bornemann, Peter Deuflhard

Konrad-Zuse-Zentrum Berlin, Heilbronner Strasse 10, D-10711 Berlin, Germany and Fachbereich Mathematik, Freie Universität Berlin, Germany; e-mail: bornemann@zib-berlin.de

Received November 12, 1994 / Revised version received October 12, 1995

Summary. The paper deals with certain adaptive multilevel methods at the confluence of nested multigrid methods and iterative methods based on the cascade principle of [10]. From the multigrid point of view, no correction cycles are needed; from the cascade principle view, a basic iteration method without any preconditioner is used at successive refinement levels. For a prescribed error tolerance on the final level, more iterations must be spent on coarser grids in order to allow for less iterations on finer grids. A first candidate of such a *cascadic multigrid method* was the recently suggested *cascadic conjugate gradient method* of [9], in short CCG method, which used the CG method as basic iteration method on each level. In [18] it has been proven, that the CCG method is accurate with optimal complexity for elliptic problems in 2D and quasi-uniform triangulations. The present paper simplifies that theory and extends it to more general basic iteration methods like the traditional multigrid smoothers. Moreover, an adaptive control strategy for the number of iterations on successive refinement levels for possibly highly non-uniform grids is worked out on the basis of a posteriori estimates. Numerical tests confirm the efficiency and robustness of the cascadic multigrid method.

Mathematics Subject Classification (1991): 65F10, 65N30, 65N55

0. Introduction

In this paper we consider linear scalar elliptic problems on general domains with space dimension d arbitrary, but practically up to 3. For the numerical solution of such problems with finite elements, *multigrid methods* are a both popular and efficient choice, cf. Hackbusch [13]. Such methods work on a sequence of grid levels $j = 0, 1, \ldots \ell$, where in our notation $j = 0$ denotes the coarse grid level and $j > 0$ the refinement levels. For the case of *adaptive* grids, the local multigrid

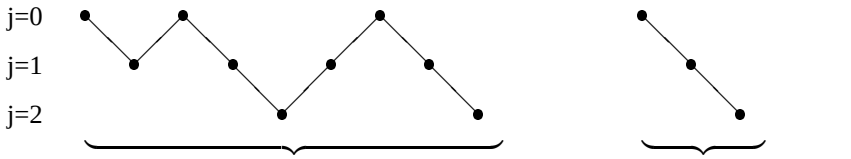
method suggested by [7] turned out to be the method of choice. In the adaptive setting it is favorable to consider nested iterations, wherein the computed solution on the previous level serves as starting point for the iteration on the new level. A typical distinction of different types of nested multigrid methods is made by the number p of correction cycles on each level: W-cycles are characterized by $p = 2$, V-cycles by $p = 1$. In view of convergence theory, W-cycles are much simpler to handle, whereas for actual implementation V-cycles are mostly preferred, since they require less computational effort per total iteration. During the eighties, further simplifications skipping the correction cycles at all ($p = 0$) seem to have been discussed; people called them "one-way multigrid" methods [11]. However, those simplifications have been abandoned in the sequel, since they were not regarded as robust enough in terms of optimal computational complexity.

The present paper deals with a kind of "one-way multigrid" method, which has evolved from the *cascade principle* due to Deuflhard, Leinen, Yserentant [10], the main idea of which was to terminate the iteration by an *adaptive control* as soon as the algebraic error is considerably below the discretization error. The first algorithmic realization [10] for $d = 2$ had used the conjugate gradient method with the hierarchical basis preconditioner of Yserentant [21], later realizations [3] included the case $d = 3$ and were based on the theoretically more satisfactory multilevel preconditioner of Bramble, Pasciak and Xu [8, 19, 17, 6], the so called BPX-preconditioner, which can be viewed as the additive Schwarz equivalent of a V-cycle [20].

Deuflhard [9] suggested, only quite recently, the use of an a posteriori algorithmic control of the "one-way multigrid" in combination with the conjugate gradient method. He called the method cascadic conjugate gradient method (CCG method). As a distinctive feature this method performs *more iterations on coarser levels* so as to obtain less iterations on finer levels. The first publication of the algorithm in [9] contained rather convincing numerical results, but no theoretical justification. Shaidurov [18] was the first to prove accuracy and optimal complexity of the CCG iteration for H^2-regular problems in dimension $d = 2$ and quasi-uniform triangulations using some a priori choice of the number of iterations, strictly decreasing for increasingly fine triangulations.

The purpose of the present paper is to both simplify Shaidurov's theoretical results and extend them to the case when the CG iteration on each refinement level is replaced by some general smoother like the traditional candidates symmetric Gauß-Seidel, SSOR or damped Jacobi iteration. We call such "one-way multigrid" methods *cascadic multigrid methods*.

In order to convey their basic structure, we give a schematic comparison of the cascadic multigrid method with the nested V-cycle multigrid method:



In Sect. 1 below, we first prove accuracy and optimal complexity of the cascadic multigrid method for $H^{1+\alpha}$-regular elliptic problems using linear energy reducing smoothers. In the next section, we show how Shaidurov's results for the (nonlinear) conjugate gradient method fit into our frame work. Both Sect. 1 and Sect. 2 require *quasi-uniform* grids. The main results are that

- for $d = 3$: the cascadic multigrid method is accurate with optimal complexity for any choice of an energy reducing smoother (symmetric Gauß-Seidel, SSOR, damped Jacobi or conjugate gradient iteration)
- for $d = 2$: the cascadic multigrid method has optimal complexity for the conjugate gradient method as a smoother and nearly optimal complexity (apart from a log-term) for the one-step iterative smoothers.

In the adaptive setting, however, we aim at the possibility of rather nonuniform grids. Therefore, in Sect. 3, we revisit the theory in view of an *adaptive* implementation and derive an a posteriori termination strategy, backed by theory, for the number of smoothing iterations on each refinement level. Finally, in Sect. 4, the derived strategy is illustrated by comparative numerical experiments for the CCG method, the cascadic multigrid method with the conjugate gradient method preconditioned by the BPX method, and the cascadic multigrid method with the symmetric Gauß-Seidel iteration as smoother.

1. Convergence analysis for general smoothers

In this section, we analyze the cascadic multigrid method with respect to accuracy and computational complexity using a general smoother as iterative method on each discretization level.

Let $\Omega \subset \mathbb{R}^d$ be a polygonal Lipschitz domain. We consider an elliptic Dirichlet problem on Ω in the weak formulation:

$$u \in H_0^1(\Omega): \qquad a(u,v) = (f,v)_{L^2} \quad \forall v \in H_0^1(\Omega).$$

Here $f \in H^{-1}(\Omega)$ and $a(\cdot,\cdot)$ is assumed to be a $H_0^1(\Omega)$-elliptic symmetric bilinear form. The induced energy-norm will be denoted by

$$\|u\|_a^2 = a(u,u) \qquad \forall u \in H_0^1(\Omega).$$

Given a nested family of triangulations $(\mathcal{T}_j)_{j=1}^{\ell}$, the spaces of linear finite elements are

$$X_j = \{u \in C(\bar{\Omega}) : u|_T \in P_1(T) \quad \forall T \in \mathcal{T}_j, \quad u|_{\partial\Omega} = 0\},$$

where $P_1(T)$ denotes the linear functions on the triangle T. We have

$$X_0 \subset X_1 \subset \ldots \subset X_\ell \subset H_0^1(\Omega).$$

The finite element approximations are given by

$$u_j \in X_j : \qquad a(u_j, v_j) = (f, v_j)_{L^2} \qquad \forall v_j \in X_j.$$

Denoting the *basic iterative procedure* on each level by $\mathcal{I}$, the cascadic multigrid method can be written as:

$$\text{(1)} \qquad \begin{aligned} &\text{(i)} \quad u_0^* = u_0 \\ &\text{(ii)} \quad j = 1, \ldots, \ell : \quad u_j^* = \mathcal{I}_{j,m_j} u_{j-1}^*. \end{aligned}$$

Here $\mathcal{I}_{j,m_j}$ denotes m_j steps of the basic iteration applied on level j.

We call a cascadic multigrid method *optimal* for level ℓ (with respect to the energy norm), if we obtain *both accuracy*

$$\|u_\ell - u_\ell^*\|_a \approx \|u - u_\ell\|_a,$$

which means that the iteration error is comparable to the approximation error, *and multigrid complexity*

$$\text{amount of work } = O(n_\ell),$$

where $n_\ell = \dim X_\ell$.

In this and the following section we consider *quasi-uniform* triangulations with meshsize parameter

$$\frac{1}{c} 2^{-j} \le h_j = \max_{T \in \mathcal{T}_j} \operatorname{diam} T \le c\, 2^{-j}.$$

For ease of notation, we will use the symbol c for any positive constant, that only depends on the bilinear form $a(\cdot,\cdot)$, on Ω and the shape regularity as well as the quasi-uniformity of the triangulations. All other dependencies will be stated explicitly.

A general assumption on the elliptic problem will be $H^{1+\alpha}$-regularity for some $0 < \alpha \le 1$, i.e.,

$$\|u\|_{H^{1+\alpha}} \le c\, \|f\|_{H^{\alpha-1}} \qquad \forall f \in H^{\alpha-1}(\Omega).$$

The approximation error of the finite element method is then given in energy norm as

$$\text{(2)} \qquad \|u - u_j\|_a \le c\, h_j^\alpha \|f\|_{H^{\alpha-1}} \qquad j = 0, \ldots, \ell,$$

cf. [14, Lemma 8.4.9]. By the well-known Aubin-Nitsche lemma and an interpolation argument one gets the *approximation property*

$$\text{(3)} \qquad \|u_j - u_{j-1}\|_{H^{1-\alpha}} \le c\, h_j^\alpha \|u_j - u_{j-1}\|_a \qquad j = 1, \ldots, \ell,$$

cf. [14, Theorem 8.4.14].

We consider the following type of basic iterations for the finite-element problem on level j started with $u_j^0 \in X_j$:

$$u_j - \mathcal{I}_{j,m_j} u_j^0 = \mathcal{S}_{j,m_j}(u_j - u_j^0)$$

with a linear mapping $\mathcal{S}_{j,m_j} : X_j \to X_j$ for the error propagation. We call the basic iteration an *energy reducing smoother*, if it obeys the *smoothing properties*

$$\text{(4)} \qquad \begin{aligned} &\text{(i)} \quad \|\mathcal{S}_{j,m_j} v_j\|_a \le c\, \frac{h_j^{-1}}{m_j^{\gamma}} \|v_j\|_{L^2} \qquad \forall v_j \in X_j, \\ &\text{(ii)} \quad \|\mathcal{S}_{j,m_j} v_j\|_a \le \|v_j\|_a \end{aligned}$$

with a parameter $0 < \gamma \le 1$. As is shown in [13] the symmetric Gauß-Seidel, the SSOR and the damped Jacobi iteration are smoothers in the sense of (4) with parameter

$$\gamma = 1/2.$$

Lemma 1.1. *A smoother in the sense of* (4) *fulfills*

$$\|\mathcal{S}_{j,m_j} v_j\|_a \le c\, \frac{h_j^{-\alpha}}{m_j^{\alpha\gamma}} \|v_j\|_{H^{1-\alpha}} \qquad \forall v_j \in X_j.$$

Proof. This can be shown be a usual interpolation argument using discrete Sobolev norms like those introduced by Bank and Dupont [1]. These discrete norms are equivalent to the fractional Sobolev norms in the considered range of smoothness. □

We are now able to state and prove the main convergence estimate for the cascadic multigrid method (1).

Theorem 1.2. *The error of the cascadic multigrid method with a smoother as basic iteration can be estimated by*

$$\|u_\ell - u_\ell^*\|_a \le c \sum_{j=1}^{\ell} \frac{1}{m_j^{\alpha\gamma}} \|u_j - u_{j-1}\|_a \le c \sum_{j=1}^{\ell} \frac{h_j^{\alpha}}{m_j^{\alpha\gamma}} \|f\|_{H^{\alpha-1}}.$$

Proof. For $j = 1, \ldots, \ell$ we get by the linearity of $\mathcal{S}_{j,m_j}$

$$\begin{aligned} \|u_j - u_j^*\|_a &= \|u_j - \mathcal{I}_{j,m_j} u_{j-1}^*\|_a = \|\mathcal{S}_{j,m_j}(u_j - u_{j-1}^*)\|_a \\ &\le \|\mathcal{S}_{j,m_j}(u_j - u_{j-1})\|_a + \|\mathcal{S}_{j,m_j}(u_{j-1} - u_{j-1}^*)\|_a. \end{aligned}$$

The first term can be estimated by Lemma 1.1 and the approximation property (3):

$$\begin{aligned} \|\mathscr{K}_{j,m_j}(u_j - u_{j-1})\|_a &\leq c\,\frac{h_j^{-\alpha}}{m_j^{\alpha\gamma}}\|u_j - u_{j-1}\|_{H^{1-\alpha}} \\ &\leq c\,\frac{1}{m_j^{\alpha\gamma}}\|u_j - u_{j-1}\|_a . \end{aligned}$$

If we estimate the second term by property (4(ii)) of a smoother, we get

$$\|u_j - u_j^*\|_a \leq \frac{c}{m_j^{\alpha\gamma}}\|u_j - u_{j-1}\|_a + \|u_{j-1} - u_{j-1}^*\|_a . \tag{5}$$

Using $u_0^* = u_0$ we get by induction

$$\|u_\ell - u_\ell^*\|_a \leq c \sum_{j=1}^{\ell} \frac{1}{m_j^{\alpha\gamma}}\|u_j - u_{j-1}\|_a .$$

Galerkin orthogonality gives

$$\|u_j - u_{j-1}\|_a \leq \|u - u_{j-1}\|_a ,$$

so that the error estimate (2) yields the second assertion. □

Since we have

$$2^{\ell-j} h_\ell / c \leq h_j \leq c\, 2^{\ell-j} h_\ell$$

Theorem 1.2 leads us to consider sequences $m_1, \ldots, m_\ell$ of the kind

$$m_j = \lceil \beta^{\ell-j} m_\ell \rceil , \tag{6}$$

for some fixed $\beta > 0$. *Accuracy* of the cascadic multigrid method is now given for sufficiently large β as the following Lemma shows.

Lemma 1.3. *Let the number m_j of iterations on level j be given by* (6)*. The cascadic multigrid method yields the error*

$$\|u_\ell - u_\ell^*\|_a \leq \begin{cases} c \cdot \dfrac{1}{1-(2/\beta^\gamma)^\alpha} \cdot \dfrac{h_\ell^\alpha}{m_\ell^{\alpha\gamma}}\|f\|_{H^{\alpha-1}}, & \text{for} \quad \beta > 2^{1/\gamma}, \\ c \cdot \ell \cdot \dfrac{h_\ell^\alpha}{m_\ell^{\alpha\gamma}}\|f\|_{H^{\alpha-1}}, & \text{for} \quad \beta = 2^{1/\gamma}. \end{cases}$$

Proof. By Theorem 1.2 we get

$$\sum_{j=1}^{\ell} \frac{h_j^\alpha}{m_j^{\alpha\gamma}} \leq c\,\frac{h_\ell^\alpha}{m_\ell^{\alpha\gamma}} \sum_{j=0}^{\ell-1} \left(\frac{2}{\beta^\gamma}\right)^{j\alpha} .$$

If $\beta > 2^{1/\gamma}$ the last sum is geometric and can be estimated by

$$\frac{1}{1-(2/\beta^\gamma)^\alpha} .$$

In the case $\beta = 2^{1/\gamma}$ the last sum is equal to ℓ. □

In contrast, the *optimal complexity* of the method is given for a sufficiently small β.

Lemma 1.4. *Let the number m_j of iterations on level j be given by* (6)*. The computational cost of the cascadic multigrid method is proportional to*

$$\sum_{j=1}^{\ell} m_j\, n_j \le \begin{cases} c \cdot \dfrac{1}{1-\beta/2^d} \cdot m_\ell\, n_\ell, & \text{for} \quad \beta < 2^d, \\ c \cdot \quad \ell \quad \cdot m_\ell\, n_\ell, & \text{for} \quad \beta = 2^d. \end{cases}$$

Proof. We have $2^{dj}/c \le n_j = \dim X_j \le c\, 2^{dj}$. Therefore we get

$$\sum_{j=1}^{\ell} m_j\, n_j \le c\, m_j\, n_j \sum_{j=0}^{\ell-1} \left(\frac{\beta}{2^d}\right)^j.$$

If $\beta < 2^d$ the last sum is geometric and can be estimated by

$$\frac{1}{1-\beta/2^d}.$$

In the case $\beta = 2^d$ the last sum is equal to ℓ. □

The last two Lemmas show that the two goals accuracy and multigrid complexity are not contradicting each other as long as

$$\gamma > 1/d,$$

since whenever this condition holds we are able to choose some β inbetween

$$2^{1/\gamma} < \beta < 2^d.$$

In the case $\gamma = 1/d$ either the accuracy or the complexity has to deteriorate logarithmically. We choose to fix accuracy and obtain as immediate consequence of our results the following Lemma.

Lemma 1.5. *Let $\gamma = 1/d$. If we choose $\beta = 2^d$ and the number of iterations on level ℓ as*

$$m_\ell = \lceil m_* \cdot \ell^{d/\alpha} \rceil$$

we get for the error of the cascadic multigrid method

$$\|u_\ell - u_\ell^*\|_a \le c\, \frac{h_\ell^\alpha}{m_*^{\alpha/d}} \|f\|_{H^{\alpha-1}},$$

and as complexity

$$\sum_{j=1}^{\ell} m_j\, n_j \le c\, m_* n_\ell (1 + \log n_\ell)^{1+d/\alpha}.$$

Our results show, that the cascadic multigrid method with a plain symmetric Gauß-Seidel, SSOR or damped Jacobi iteration (all with $\gamma = 1/2$) as basic iteration is

- optimal for $d = 3$,
- accurate with complexity $O(n_\ell\,(1 + \log n_\ell)^{1+2/\alpha})$ for $d = 2$.

2. Conjugate gradient method as smoother

When using the conjugate gradient method as the basic iteration we have to tackle with a problem: the result

$$\mathscr{T}_{j,m_j} u_j^0$$

of m_j steps of the cg-method is *not* linear in the starting value u_j^0. Thus, it seems that our frame work, developed so far, does not cover the cg-method. However, there is a remedy which uses results on the cg-method well known in the Russian literature [12, 15] or can be seen as a consequence of the theory of smoothing properties for semi-iterative methods [13, Sect. 6.2.5].

We have to fix some notation. Let $\langle \cdot, \cdot \rangle$ be the euclidean scalar product of the nodal basis in the finite element space X_j, the induced norm will be denoted by

$$|v_j|^2 = \langle v_j, v_j \rangle \qquad \forall v_j \in X_j .$$

We define the linear operator $A_j : X_j \to X_j$ by

$$\langle A v_j, w_j \rangle = a(v_j, w_j) \qquad \forall v_j, w_j \in X_j ,$$

which is represented in the nodal basis by the usual *stiffness matrix*. The error of the cg-method applied to the stiffness matrix can be expressed by

$$\|u_j - \mathscr{T}_{j,m_j} u_j^0\|_a = \min_{\substack{p \in P_{m_j} , \\ p(0)=1}} \|p(A_j)(u_j - u_j^0)\|_a .$$

Here P_{m_j} denotes the set of polynomials p with $\deg p \leq m_j$.

The idea is now, to find some polynomial $q_{j,m_j} \in P_{m_j}$ with $q_{j,m_j}(0) = 1$, such that

$$\mathscr{S}_{j,m_j} = q_{j,m_j}(A_j)$$

defines a linear smoother in the sense of (4). Since the error in the energy-norm of the CG method is then *majorized* by this smoother, the results of Sect. 1 are immediately valid for the CG method.

The choice of q_{j,m_j} depends on the following solution of a polynomial minimization problem, which can be traced back at least to a paper of Godunov and Prokopov [12].

Lemma 2.1. *Let $\lambda > 0$. The Chebyshev polynomial T_{2m+1} has the representation*

$$T_{2m+1}(x) = (-1)^m (2m+1)\, x\, \phi_{\lambda,m}(\lambda x^2)$$

with a unique $\phi_{\lambda,m} \in P_m$ and $\phi_{\lambda,m}(0) = 1$. The polynomial $\phi_{\lambda,m}$ solves the minimization problem

$$\max_{x \in [0,\lambda]} |\sqrt{x}\, p(x)| = \min!$$

over all polynomials $p \in P_m$ which are normalized by $p(0) = 1$. The minimal value is given by

$$\max_{x\in[0,\lambda]} |\sqrt{x}\phi_{\lambda,m}(x)| = \frac{\sqrt{\lambda}}{2m+1}.$$

Moreover we have

$$\max_{x\in[0,\lambda]} |\phi_{\lambda,m}(x)| = 1.$$

As a fairly easy consequence one obtains the following result.

Theorem 2.2. *The linear operator*

$$\mathscr{S}_{j,m_j} = \phi_{\lambda_j,m_j}(A_j), \qquad \lambda_j = \max\sigma(A_j)$$

satisfies

$$\text{(i)} \quad \|\mathscr{S}_{j,m_j} v_j\|_a \le \frac{\sqrt{\lambda_j}}{2m_j+1}|v_j| \qquad \forall v_j \in X_j.$$

$$\text{(ii)} \quad \|\mathscr{S}_{j,m_j} v_j\|_a \le \|v_j\|_a$$

Proof. During the proof we will suppress all indices which are used in the formulation of the Theorem. Let $\psi_1,\ldots,\psi_n$ be the orthonormal basis of X consisting of eigenvectors for A with

$$A\psi_\mu = \lambda_\mu\psi_\mu, \qquad \mu = 1,\ldots,n.$$

We choose the ordering to be

$$0 < \lambda_1 \le \ldots \le \lambda_n = \lambda.$$

Now we get for $v \in X$ by Lemma 2.1

$$\begin{aligned}\|\mathscr{S}v\|_a^2 &= \langle A\phi_{\lambda,m}^2(A)v, v\rangle = \sum_{\mu=1}^{n} \lambda_\mu \phi_{\lambda,m}^2(\lambda_\mu)\langle v,\psi_\mu\rangle^2 \\ &\le \begin{cases} \dfrac{\lambda}{(2m+1)^2}\sum_\mu \langle v,\psi_\mu\rangle^2 = \dfrac{\lambda}{(2m+1)^2}|v|^2, \\ \sum_\mu \lambda_\mu\langle v,\psi_\mu\rangle^2 = \langle Av, v\rangle = \|v\|_a^2. \quad \square \end{cases}\end{aligned}$$

Compared to the standard Chebyshev type convergence estimate, this estimate seems to be comparatively poor. However, it can be made a useful analytical tool, if we take care of the scaling of the linear system in a clever way. In fact, a little finite element theory shows that along this line we are able to define a majorizing linear smoother for the cg-method.

Corollary 2.3. *The linear operator*

$$\mathscr{S}_{j,m_j} = \phi_{\lambda_j,m_j}(A_j), \qquad \lambda_j = \max\sigma(A_j)$$

defines a smoother in the sense of (4) with parameter $\gamma = 1$.

Proof. The usual inverse inequality shows that the maximum eigenvalue of the stiffness matrix can be estimated by

$$\lambda_j \le c h_j^{d-2},$$

cf. [14, Theorem 8.8.6]. The euclidean norm with respect to the nodal basis is related to the L^2-norm by

$$\frac{1}{c} h_j^d |v_j|^2 \le \|v_j\|_{L^2}^2 \le c h_j^d |v_j|^2,$$

cf. [14, Theorem 8.8.1]. Thus Theorem 2.2 gives

$$\|\mathscr{S}_{j,m_j} v_j\|_a \le c \frac{h_j^{(d-2)/2}}{2m_j+1} |v_j| \le c \frac{h_j^{-1}}{2m_j+1} \|v_j\|_{L^2},$$

i.e., (4(i)) with $\gamma = 1$. Property (4(ii)) was already stated in Theorem 2.2. □

With the help of this majorizing smoother, it is immediately clear, that Theorem 1.2, Lemma 1.3, Lemma 1.4 and Lemma 1.5 remain valid for the cascadic multigrid method with the cg-method as basic iteration, which is the CCG-method of [9]. In particular, the CCG method is seen to be optimal for $d = 2, 3$.

3. Adaptive cascadic multigrid method

In this section, we develop an adaptive control of the cascadic multigrid method which is based on our theoretical considerations and some additional assumptions on the family of triangulations. For adaptive triangulations we drop the assumption of quasi-uniformity. All constants in the sequel will not depend on the quasi-uniformity, but only on the shape regularity of the triangulations.

In order to bound the maximum eigenvalue of all involved matrices simultaneously, we compute the stiffness matrix A_j with respect to the scaled nodal basis

$$h_{j,i}^{(2-d)/2} \psi_{j,i} \qquad i = 1, \ldots, n_j,$$

where $\{\psi_{j,i}\}_i$ is the usual nodal basis of X_j and

$$h_{j,i} = \operatorname{diam} \operatorname{supp} \psi_{j,i}.$$

Remark . This scaling is equivalent to a *diagonal preconditioning* of the stiffness matrix. We strongly recommend to use a diagonal scaling for all problems. For instance, unpleasant effects resulting from (jump-)discontinuous coefficients of the elliptic operator can be avoided, certainly in dimension $d = 2$ and for non-pathological situations in dimension $d = 3$.

In this section we denote by $\langle \cdot, \cdot \rangle$ the euclidean scalar product with respect to this *scaled* nodal basis. Xu [19] has shown the equivalence of norms

$$\frac{1}{c}|v_j|^2 \le \sum_{T \in \mathscr{T}_j} h_T^{-2} \|v_j\|^2_{L^2(T)} \le c\,|v_j|^2, \qquad h_T = \operatorname{diam} T,$$

and the bound

$$\lambda_j = \max \sigma(A_j) \le c$$

for the maximum eigenvalue of A_j. Hence, we get for the smoothing iterations considered

$$\|\cdot \mathscr{S}_{j,m_j} v_j\|_a \le \frac{c}{m_j^\gamma} \left(\sum_{T \in \mathscr{T}_j} h_T^{-2} \|v_j\|^2_{L^2(T)} \right)^{1/2},$$

where we have $\gamma = 1/2$ for the symmetric Gauß-Seidel, SSOR, damped Jacobi iteration and $\gamma = 1$ for the majorizing linear smoother of the conjugate gradient method as introduced in Theorem 2.2

In order to turn this into a starting point for a theorem like Theorem 1.2 we make the following two assumptions on the family of triangulations:

$$\begin{aligned} &\text{(i)} \quad h_T^{-2}\|u_j - u_{j-1}\|^2_{L^2(T)} \le c\|u_j - u_{j-1}\|^2_{H^1(T)}, \qquad \forall T \in \mathscr{T}_j \\ &\text{(ii)} \quad \|u - u_j\| \le c\, n_j^{-1/d} \|f\|_{L^2}, \end{aligned} \tag{7}$$

which are heuristically justified for *adaptive* triangulations. Assumption (i) means, that the finite element correction is locally of high frequency with respect to the finer triangulation. In other words, the refinement resolves changes *but not more*. Thus it is a statement of the *efficiency* of a triangulation. Assumption (ii) is a statement of optimal *accuracy*, which is justified by results of nonlinear approximation theory like [16]. Note that quasi-uniform triangulations do not satisfy assumption (ii) for problems which are not H^2-regular.

The same proof as for Theorem 1.2 now gives the following

Lemma 3.1. *Assumption (7) implies, for the error of the cascadic multigrid method,*

$$\|u_\ell - u_\ell^*\|_a \le c \sum_{j=1}^{\ell} \frac{1}{m_j^\gamma} \|u_j - u_{j-1}\|_a \le c \sum_{j=1}^{\ell} \frac{1}{m_j^\gamma\, n_j^{1/d}} \|f\|_{L^2}.$$

We can now extend Lemma 1.3 and Lemma 1.4 to the adaptive case. Here we have to use additionally, that the sequence of number of unknowns belongs to a geometric progression:

$$n_j < \sigma_0 n_j \le n_{j+1} \le \sigma_1 n_j \qquad j = 0, 1, \ldots.$$

With the choice of the iteration numbers m_j on level j as

$$m_j = \left\lceil m_\ell \left(\frac{n_\ell}{n_j} \right)^{(d+1)/2d\gamma} \right\rceil , \tag{8}$$

we get for $d > 1$ under assumption (7) the final error

$$\|u_\ell - u_\ell^*\|_a \leq \frac{c}{m_\ell^\gamma \, n_\ell^{1/d}} \|f\|_{L^2}$$

and for $\gamma = 1$ the complexity

$$\sum_{j=1}^{\ell} m_j \, n_j \leq c \, m_\ell \, n_\ell .$$

However, we will see in the numerical experiments, that optimal complexity can also be obtained for $\gamma = 1/2$, indicating that there should be sharper estimates than those used in our present analysis.

However, at the intermediate level j we do *not* know the number n_ℓ of nodal points at the final level ℓ, which means that so far our iteration is not yet implementable. To make it implementable, we *define* the final level ℓ as the first level on which the approximation error is below some user given tolerance TOL. Hence assumption (7) gives us the relation

$$\frac{\|u - u_j\|_a}{\mathrm{TOL}} \approx \left(\frac{n_\ell}{n_j} \right)^{1/d} , \tag{9}$$

which leads us to replace (8) by

$$m_j = \left\lceil m_\ell \left(\frac{\|u - u_j\|_a}{\mathrm{TOL}} \right)^{(d+1)/2\gamma} \right\rceil . \tag{10}$$

This algorithm is closest to the a priori choice of the parameters m_j. In actual computation, the basic iteration can be accurate enough much earlier than stated by theory. With these preparations, we may now go back to the crucial recursion (5), i.e.,

$$\|u_j - u_j^*\|_a \leq \frac{c}{m_j^\gamma} \|u - u_{j-1}\|_a + \|u_{j-1} - u_{j-1}^*\|_a ,$$

which we simply turn into a termination criterion for the basic iteration by inserting (10). We thus end up with the termination criterion

$$\|u_j - u_j^*\|_a \leq \rho \left(\frac{\mathrm{TOL}}{\|u - u_j\|_a} \right)^{(d+1)/2} \|u - u_{j-1}\|_a + \|u_{j-1} - u_{j-1}^*\|_a , \tag{11}$$

where $0 < \rho \leq 1$ is some safety factor. Note that the smoothing parameter γ dropped out since we stress the accuracy aspect of our analysis for the adaptive control. Herein the approximation error $\|u - u_j\|_a$ is not known, but can be replaced by the estimate

$$\|u - u_j\|_a \approx \|u - u_{j-1}\|_a \left(\frac{n_{j-1}}{n_j}\right)^{1/d} \approx \epsilon_{j-1} \left(\frac{n_{j-1}}{n_j}\right)^{1/d},$$

where ϵ_{j-1} denotes some estimate of the discretization error on the previous level, which is certainly provided by an adaptive algorithm; cf. [10, 4]. If we replace the iteration errors $\|u_j - u_j^*\|_a$, $\|u_{j-1} - u_{j-1}^*\|_a$ by appropriate estimates δ_j, δ_{j-1}, the design of an implementable control strategy for the adaptive cascadic multigrid method is completed:

$$\delta_j \le \rho \left(\frac{\mathrm{TOL}}{\epsilon_{j-1}} \left(\frac{n_j}{n_{j-1}}\right)^{1/d}\right)^{(d+1)/2} \epsilon_{j-1} + \delta_{j-1}. \tag{12}$$

Finally, we want to compare the termination criterion (11) with the criterion

$$\|u_j - u_j^*\|_a \le \rho_* \ \mathrm{TOL}$$

as suggested in [9] for the CCG method, wherein $0 < \rho_* \le 1$ was understood as some safety factor. By using the relation (9) and the assumption (7(ii)) we can bound the right hand side of (11) as follows

$$\|u_j - u_j^*\|_a \le c\,\rho\,\sigma_1^{1/d}\,\mathrm{TOL} \left(\frac{n_j}{n_\ell}\right)^{(d-1)/2d} + \|u_{j-1} - u_{j-1}^*\|_a.$$

This recursion yields

$$\begin{aligned}
\|u_j - u_j^*\|_a &\le c\,\rho\,\sigma_1^{1/d}\,\mathrm{TOL} \sum_{k=1}^{j} \left(\frac{n_k}{n_\ell}\right)^{(d-1)/2d} \\
&\le c\,\rho\,\sigma_1^{1/d}\,\mathrm{TOL} \sum_{k=1}^{j} \left(\frac{1}{\sigma_0}\right)^{(\ell-k)(d-1)/2d} \\
&\le \underbrace{\frac{c\,\sigma_1^{1/d}}{1 - \sigma_0^{(1-d)/2d}}\rho}_{=\rho_*} \cdot \mathrm{TOL}\,.
\end{aligned}$$

At first glance, we seem to have confirmed the above termination criterion. However, the derivation leads to a factor ρ_*, which is no longer understood to be restricted by $\rho_* \le 1$. In complicated examples, a reasonable choice of ρ_* will be rather delicate – as has been confirmed by numerical experiments especially with reentrant corners. In contrast to this undesirable sensitivity, the new termination criterion (12) does not depend too critically on the choice of the safety parameter ρ.

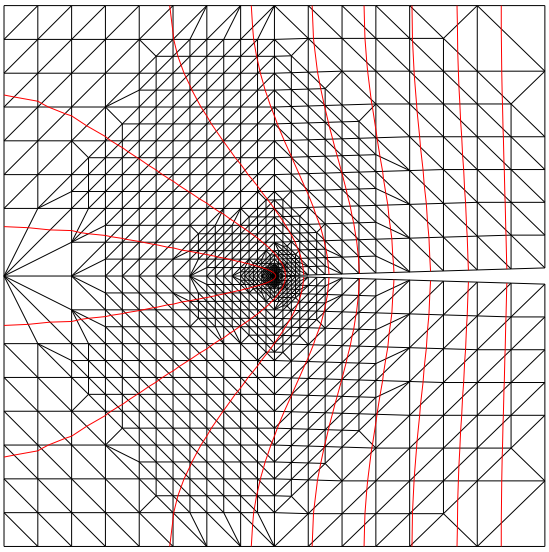


Fig. 1. Typical adaptive triangulation with isolines of solution

4. Numerical experiments

The above designed adaptive cascadic multigrid method has been implemented in three different variants:

- *CCG*: the cascadic multigrid method with the conjugate gradient method as the basic iteration (cascadic conjugate gradient method),
- *CSGS*: the cascadic multigrid method with the symmetric Gauß-Seidel iteration as the basic iteration,
- *adaptive V-cycle*: the cascadic multigrid method with a multigrid V-cycle as the basic iteration using a Jacobi smoother, local (new nodal points and neighbors only) in the case of locally refined triangulations. This implementation can be viewed as a robust and efficient *automatic choice* of the number m_j of V-cycles on level j for the usual nested multigrid method.

All variants used the termination criterion (12), where the algebraic error estimate δ_j was realized as the diagonally scaled residual and the discretization error estimate ϵ_j was taken to be the usual edge oriented estimator due to [10] for $d = 2$ and the edge and face oriented estimator from [4] for $d = 3$.

The three algorithms were tested in a variety of problems including

- spatial dimension $d = 2, 3$
- uniform and adaptive grid refinement
- selfadjoint singular perturbation problems arising in the transient phase of parabolic problems due to discretization in time, cf. [2]
- problems with material jumps in the coefficients

The new methods CCG and CSGS worked efficiently in all these cases, when *diagonal* preconditioning was used, as indicated by the theoretical considerations of Sect. 3. In particular, the new termination strategy (12) turned out to be clearly more robust than the previous one suggested in [9]. Moreover, there was no difference of the behavior in two and three space dimensions. On this basis, only one single two-dimensional example with a reentrant corner was regarded to be enough to demonstrate the features of the cascadic multigrid method.

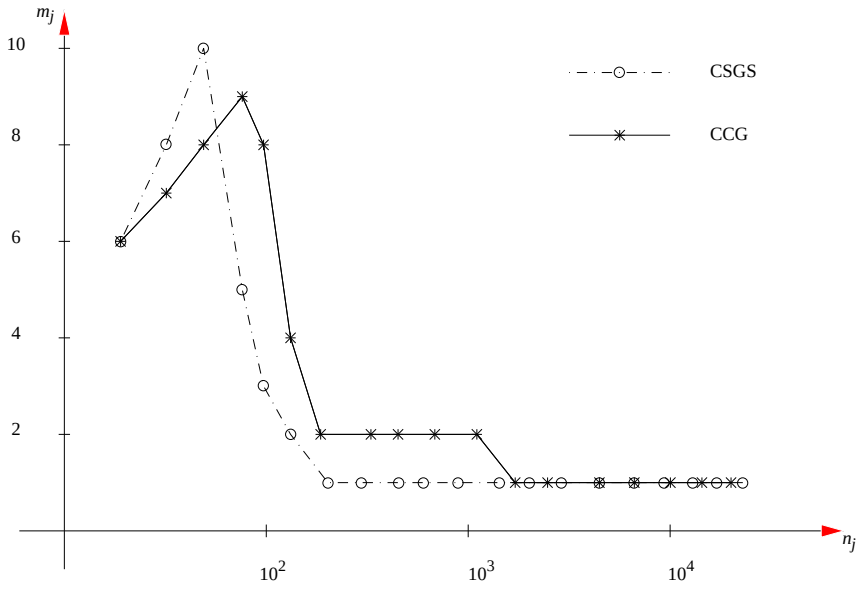


Fig. 2. Comparative number of iterations vs. number of unknowns

Example . Consider the elliptic problem

$$-\Delta u = 0, \qquad u|_{\Gamma_1} = 10^3, \quad u|_{\Gamma_2} = 0, \quad \partial_n u|_{\Gamma_3} = 0$$

on a domain Ω which is a unit square with slit

$$\Omega = \{x \in \mathbb{R}^2 : |x|_\infty \leq 1\} \cap \{x \in \mathbb{R}^2 : |x_2| \geq 0.03 x_1\}.$$

The boundary pieces are

$$\Gamma_1 = \{x \in \Omega : x_1 = 1, x_2 \geq 0.03\}, \quad \Gamma_2 = \{x \in \Omega : x_1 = 1, x_2 \leq -0.03\},$$

and $\Gamma_3 = \partial\Omega \setminus (\Gamma_1 \cup \Gamma_2)$.

The accuracy parameter was set to a relative $\mathrm{TOL}_{\mathrm{rel}} = 10^{-2}$, i.e.,

$$\mathrm{TOL} = \|u\|_a \cdot \mathrm{TOL}_{\mathrm{rel}} = 412.52 \cdot \mathrm{TOL}_{\mathrm{rel}} .$$

Throughout the safety factor was set to $\rho = 0.4$.

Figure 1 shows a typical adaptive triangulation, wherein a geometric increase of nodes with $\sigma_0 = 1.1$ was imposed, and the corresponding isolines of the solution. In order to illustrate the iterative behavior of the cascadic multigrid variants, Fig. 2 shows the number m_j of iterations. The V-cycle has been excluded here since global smoothing of CCG and CSGS is not comparable to local smoothing of the V-cycle. However, all methods can be compared with respect to computing time. Therefore, Fig. 3 shows the actual accumulated computing time (SUN sparc IPX) spent in the iteration versus the number n_j of nodes. For comparison reasons we included as a further method:

– *plain V-cycle*: nested V-cycle multigrid with local Jacobi smoothing, using *only one* V-cycle per level. This minimal variant serves as the *lower bound* for all possible strategies to control the nested V-cycle multigrid method. It cannot be viewed as a generally robust choice, but leads to reasonable results in this example.

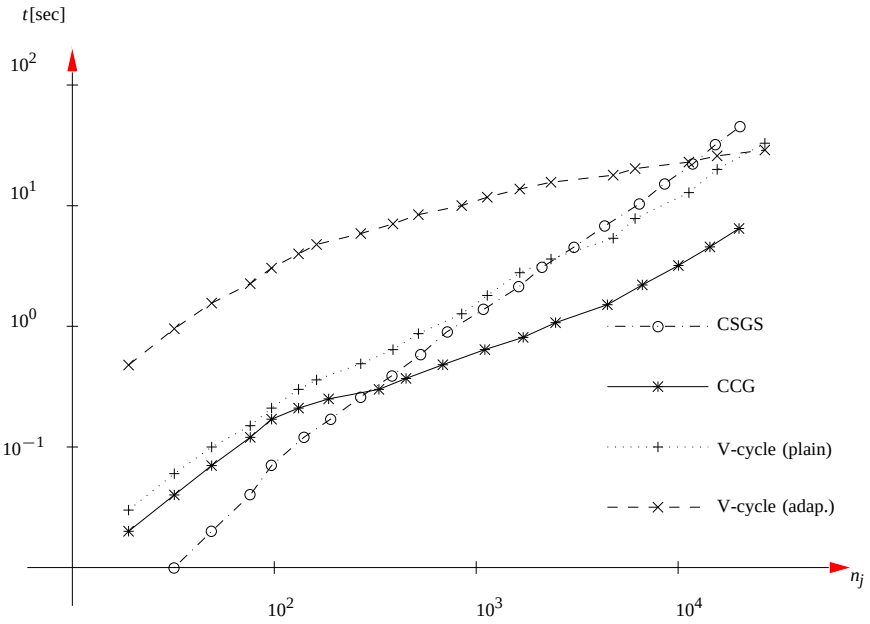


Fig. 3. Accumulated iteration time vs. number of unknowns

On the final level the CCG variant is roughly a factor of two faster than the other variants. This effect is due to the fact, that we have only one iteration on the final levels which dominate the total iteration cost. But *one* symmetric Gauß-Seidel step is twice as costly as one conjugate gradient step. The same observation holds for a V-cycle with local Jacobi smoothing, since it touches roughly twice as many nodes as contained in the final level; cf. [2, Lemma 2.6].

For the sake of clarity, however, it should be mentioned that the process of *solving* the problem required only a minor portion of the total computing time

compared to the effort spent in *establishing* the problem, i.e., computing the stiffness matrices.

Table 1. Algebraic errors of the variants for $\mathrm{TOL}_{\mathrm{rel}} = 2.24 \cdot 10^{-2}$

variant	$\frac{\|u_{15}-u_{15}^*\|_a}{\|u_{15}\|_a}$	$\frac{\|u_{15}-u_{15}^*\|_{L^2}}{\|u_{15}\|_{L^2}}$	$\frac{\sum \text{CPU-time}}{n_{15}}$
CCG	$6 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	2.14 ms
CSGS	$6 \cdot 10^{-2}$	$1 \cdot 10^{-2}$	2.70 ms
V-cycle (plain)	$2 \cdot 10^{-2}$	$1 \cdot 10^{-3}$	3.61 ms
V-cycle (adap.)	$2 \cdot 10^{-3}$	$1 \cdot 10^{-4}$	4.27 ms

Table 1 shows the behavior of the (purely) *algebraic error* for the different variants. The relative tolerance $\mathrm{TOL}_{\mathrm{rel}} = 2.24 \cdot 10^{-2}$ is used yielding 15 refinement levels with a final triangulation of roughly 4400 nodal points. This allows one to estimate the *discretization error* using two further *uniform* refinements:

$$\frac{\|u - u_{15}\|_a}{\|u\|_a} \approx 2 \cdot 10^{-2}, \qquad \frac{\|u - u_{15}\|_{L^2}}{\|u\|_{L^2}} \approx 1 \cdot 10^{-4},$$

where

$$\|u\|_a = 412.52, \qquad \|u\|_{L^2} = 1121.36.$$

The algebraic errors in *energy norm* are comparable to the required accuracy. However, the variants without coarse grid corrections are slightly less accurate by a factor of three, whereas the nested multigrid variants stay below the discretization error.

With respect to the L^2*-norm* only the *adaptive* V-cycle variant gives satisfactory results. This fact points out that the given termination criterion (12) can be viewed as a robust tool to control the number of iterations in nested multigrid methods.

Remark . The reader should note, that the term *optimality* has been used in this paper with respect to the *energy norm*. As Table 1 already indicates things are totally different for the L^2- or L^∞-norm. By means of simple examples one can *prove* that the cascadic multigrid method *cannot* be optimal with respect to the L^2-norm. There is a rather precise theoretical understanding of this phenomenon, which will be subject of a forthcoming publication [5].

Acknowledgments. The authors gratefully acknowledge Ralf Kornhuber, who had pointed out to them the relation of the termination criterion (11) to the one used in [9]. Moreover, they want to thank Rudi Beck and Rolf Krause for computational assistance.

References

1. Bank, R.E., Dupont, T.F. (1981): An optimal order process for solving elliptic finite element equations. Math. Comp. **36**, 967–975
2. Bornemann, F.A. (1992): An adaptive multilevel approach to parabolic equations III. 2D error estimation and multilevel preconditioning. IMPACT Comput. Sci. Engrg. **4**, 1–45
3. Bornemann, F.A., Erdmann, B., Kornhuber, R. (1993): Adaptive multilevel methods in three space dimensions. Inter. J. Numer. Meth. Engrg. **36**, 3187–3202
4. Bornemann, F.A., Erdmann, B., Kornhuber, R. (1996): A posteriori error estimates for elliptic problems in two and three space dimensions. SIAM J. Numer. Anal. (to appear)
5. Bornemann, F.A., Krause, R. (1995): manuscript in preparation
6. Bornemann, F.A., Yserentant, H. (1993): A basic norm equivalence for the theory of multilevel methods. Numer. Math. **64**, 455–476
7. Bramble, J.H., Pasciak, J.E. Wang, J., Xu, J. (1991): Convergence estimates for multigrid algorithms without regularity assumptions. Math. Comp. **57**, 23–45
8. Bramble, J.H., Pasciak, J.E., Xu, J. (1990): Parallel multilevel preconditioners. Math. Comp. **55**, 1–22
9. Deuflhard, P. (1994): Cascadic conjugate gradient methods for elliptic partial differential equations. Algorithm and numerical results. In: Keyes, D., Xu, J. (eds) Proceedings of the 7th International Conference on Domain Decomposition Methods 1993, pp. 29–42. AMS, Providence
10. Deuflhard, P., Leinen, P., Yserentant, H. (1989): Concepts of an adaptive hierarchical finite element code. IMPACT Compt. Sci. Engrg. **1**, 3–35
11. Douglas, C.C. (1994): personal communication
12. Godunov, S.K., Prokopov, G.P. (1969): On the solution of the Laplace difference equation (Russian). Zh. vychisl. Mat. mat. Fiz. **9**, 462–468
13. Hackbusch, W. (1985): Multi-Grid Methods and Applications. Springer-Verlag, Berlin, Heidelberg, New York
14. Hackbusch, W. (1992): Theory and Numerical Treatment of Elliptic Differential Equations. Springer-Verlag, Berlin, Heidelberg, New York
15. Il'in **V.P. (1976)**, Some estimates for conjugate gradient methods. USSR Comput. Math. and Math. Phys. **16**, 22–30
16. Oswald, P. (1990): On function spaces related to finite element approximation theory. Z. Anal. Anwend. **9**, 43–64
17. Oswald, P. (1991): On discrete norm estimates related to multilevel preconditioners in the finite element method. In Proceedings International Conference on the Contructive Theory of Functions, Varna
18. Shaidurov, V.V. (1994): Some estimates of the rate of convergence for the cascadic conjugate-gradient method. Otto-von-Guericke-Universität, Magdeburg (preprint)
19. Xu, J. (1989): Theory of Multilevel Methods. Department of Mathematics, Pennsylvania State University, University Park Report No. AM 48
20. Xu, J. (1992): Iterative methods by space decomposition and subspace correction. SIAM Review **34**, 581–613
21. Yserentant, H. (1986): On the multi-level splitting of finite element spaces. Numer. Math. **49**, 379–412