Numer. Math. 37, 387–404 (1981)

Numerische Mathematik

The Contraction Number of a Multigrid Method for Solving the Poisson Equation

Dietrich Braess

Institut für Mathematik, Ruhr-Universität Bochum, Universitätsstr. 150, D-4630 Bochum 1, Germany (Fed. Rep.)

Summary. The treatment of a multigrid method in the framework of numerical analysis elucidates that regularity of the solution is not necessary for the convergence of the multigrid algorithm but only for fast convergence. For the linear equations which arise from the discretization of the Poisson equation, a convergence factor 0,5 is established independent of the shape of the domain and of the regularity of the solution.

Subject Classifications: AMS(MOS): 65N20; CR: 5.17.

1. Introduction

In recent years multigrid methods have been successfully applied to the numerical solution of the linear equations that arise from the discretization of elliptic problems. The convergence proofs in [5, 7, 10] and in the literature cited there make use of the approximation properties of finite element approximations. Therefore convergence is strongly related to the regularity of the solution of the differential equation.

In this paper we will treat multigrid methods for the discretization of the boundary value problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega \subset \mathbb{R}^2, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{1.1}$$

Our study of the standard 5-point-formula and piecewise quadratic elements will be done completely in the framework of *numerical linear algebra*, without using approximation properties.

The resulting system of linear equations will be approximately reduced to one of the same structure, but with fewer variables. This reduced system corresponds to the discretization of a Dirichlet problem in a coarser grid. The

Dedicated to Professor Dr.Dr.h.c. Lothar Collatz on the occasion of his 70th birthday

0029-599X/81/0037/0387/$03.60

multigrid iteration combines an approximative solution of the reduced system with one half step of the Gauß-Seidel algorithm or another appropriate smoothing procedure. The latter can partially compensate for the error caused by such a reduction.

The central idea for the study of the two-level procedure is the following: The linear equations characterize the solution of a variational problem in a finite element space S_h. We decompose S_h as the direct sum of two subspaces, $S_h = V \oplus W$, where $V = S_H$ is the finite element space for a coarser grid. If the variational problem is solved alternately in V and W, an iteration is obtained, for which the convergence rate may be estimated via a *strengthened Cauchy inequality*. The results for the two-level process are extended to the multigrid method by recursion.

After the first draft of this paper had been written the author was directed by H. Jarausch to similar investigations by Bank and Dupont [2]. In [2] the two-level iteration was analyzed via the decomposition of the finite element spaces for even a wider class of elliptic problems. But the multi-level case was again studied in a framework in which regularity assumptions are necessary. We will do this rigorously without such assumptions (though we are certain that some phenomena cannot be understood in the framework of numerical linear algebra). Of course, for contrast we will put more stress on these items for which our point of view is different from that of Bank and Dupont.

Our investigations aim at the following results and properties:

1. We establish an explicit bound

$$\frac{1}{2}$$

for the contraction number of the two-level iteration for the treatment of (1.1) independent of the regularity of the solution.

2. This bound is independent of the domain Ω. We assume only that the domain Ω is polygonal and that its corners belong to the finest grid.

3. The bound $\frac{1}{2}$ is established without the assumption that in each iteration sufficiently many relaxation steps are performed (cf. [7, 10]).

4. The recursion shows that the multi-level procedures for smooth solutions should be slightly different from those for less regular solutions, for which the rate of convergence is bounded from above by the number 0.62.

In Sect. 2 to 4 we study the standard 5-point-discretization in sequences of grids where the ratio of the mesh sizes is $\sqrt{2}$. Then the constant which enters into the strengthened Cauchy inequality is easily understood. In Sect. 5 the results are extended to multiple grids with mesh ratio 2. Since here the constants are computed numerically, this case is not as illustrative as the previous one. In Sect. 7, where piecewise quadratic elements are treated, the matrices of the linear systems are denser. Here it turns out that, on the highest level, the system may be reduced to the piecewise linear functions considered above with almost the same rate of convergence.

2. The Two-Level Process

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain. Assume that there is a triangulation of Ω which is generated by horizontal and vertical lines of distance h and by

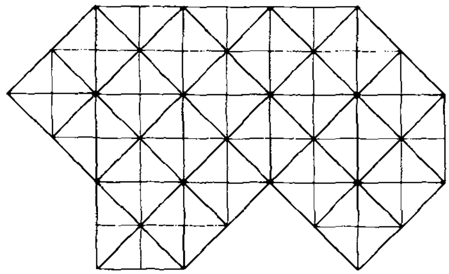

Fig. 1. Triangulation of a polygonal domain

diagonal lines of distance $H=h\sqrt{2}$ (see Fig. 1). The set of grid points $\{p_i\}_{i=1}^n$ which are contained in (the interior of) Ω is denoted as Ω_h, while Ω_H refers to the subset of points which also belong to the coarser grid formed by the diagonal lines.

Since there is no danger of confusion, we will also use the symbols Ω_h and Ω_H for the associated triangulations. The discretization of (1.1) with the standard 5-point-formula [3, p. 282] leads to a linear system of the form

$$A^h x = b^h \tag{2.1}$$

where

$$A_{ij}^h = \begin{cases} 1 & \text{if } i=j, \\ -1/4 & \text{if } i\neq j \text{ and } p_i, p_j \text{ are adjacent in } \Omega_h \\ 0 & \text{otherwise.} \end{cases} \tag{2.2}$$

We may rewrite (2.1) as

$$x_i = \tfrac{1}{4} \sum_j{}'_h x_j + b_i^h, \tag{2.3}$$

where Σ'_h refers to the summation over all neighbours in the grid Ω_h.

For convenience the Gauß-Seidel relaxation is split into two half-steps.

$$(G_h^{\mathrm{I}} x)_i = \begin{cases} x_i, & \text{if } p_i \in \Omega_H, \\ \tfrac{1}{4} \sum_j{}'_h x_j + b_i^h, & \text{if } p_i \notin \Omega_H. \end{cases}$$

$$(G_h^{\mathrm{II}} x)_i = \begin{cases} \tfrac{1}{4} \sum_j{}'_h x_j + b_i^h, & \text{if } p_i \in \Omega_H, \\ x_i, & \text{if } p_i \notin \Omega_H. \end{cases}$$

Obviously $G_h^{\mathrm{I}} x$ depends only on the components of x on Ω_H and $G_h^{\mathrm{II}} x$ only on the other ones.

As usual, the variables in multi-level algorithms carry three superscripts; they refer to (1) the level (or equivalently to the grid), (2) an iteration count and (3) a count of the steps within one iteration loop.

Algorithm 2.1. (k-th loop of the two-level iteration for Ω_h).

1. Given $x^{h,k,0}$ compute

$$x^{h,k,1} = (G_h^{\mathrm{II}} \circ G_h^{\mathrm{I}})^v x^{h,k,0},$$

where $v = 0, 1$ or 2.

2. Put

$$x^{h,k,2} = G_h^{\mathrm{I}} x^{h,k,1}. \tag{2.4}$$

3. Determine the residual $d = b^h - A^h x^{h,k,2}$ and solve in the coarser grid the linear equations

$$\begin{aligned} y_i &= \tfrac{1}{4} \sum_j{}'_H y_j + d_i, \qquad p_i \in \Omega_H, \\ &= \text{arbitrary}, \qquad p_i \in \Omega_h \setminus \Omega_H. \end{aligned} \tag{2.5}$$

(The iteration is independent of the choice of the y_i's, $p_i \in \Omega_h \setminus \Omega_H$. For the theoretical analysis the values should be chosen such that y may be interpolated by a function which is piecewise linear on Ω_H.)

4. Put

$$x^{h,k,3} = x^{h,k,2} + y.$$

5. Compute

$$x^{h,k,4} = G_h^{\mathrm{I}} x^{h,k,3}$$

and proceed with $x^{h,k+1,0} = x^{h,k,4}$.

Note that each iteration loop begins and ends with the execution of G_h^{I}. Therefore Step 5 need only be performed in the last iteration.

Furthermore we note that the residual vetor d has only non-zero coefficients d_i for $p_i \in \Omega_H$. This is caused by the execution of G_h^{I} in Step 2.

To understand the algorithm we interprete (2.1) as an Eulerian equation for the variational problem

$$J(u) := a(u, u) - 2(f, u)_0 \to \min, \tag{2.6}$$

where

$$a(u, v) = \int_\Omega (u_\xi v_\xi + u_\eta v_\eta)\, d\xi\, d\eta,$$

$$(u, v)_0 = \int_\Omega u v\, d\xi\, d\eta. \tag{2.7}$$

Specifically, the minimum of (2.6) is to be determined in S_h, the space of those continuous functions in $H_0^1(\Omega)$ which are linear on the triangles associated to Ω_h.

Each $u \in S_h$ may be written in the form

$$u = \sum_{i=1}^n x_i \phi_i^h, \qquad x_i = u(p_i). \tag{2.8}$$

Here $\{\phi_i^h;\ p_i \in \Omega_h\}$ is a basis of S_h such that

$$\phi_i(p_j) = \delta_{ij}, \qquad p_i, p_j \in \Omega_h.$$

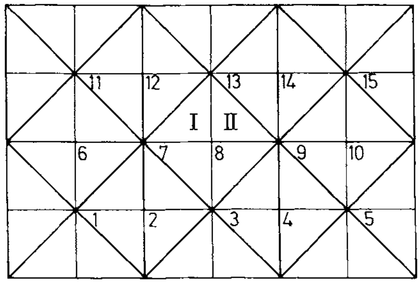

Fig. 2. Section from triangulation in Fig. 1. Triangulation of a polygonal domain

We will associate to the vectors $x^{h,k,\mu}$ from Algorithm 2.1 the functions $u^{h,k,\mu}$ via (2.8).

Note that the support of ϕ_i contains 8 triangles if $p_i \in \Omega_H$, and 4 triangles otherwise. The triangulation which is shown in more detail in Fig. 2 differs from Courant's choice [6, p. 218]. Nevertheless,

$$a(\phi_i^h, \phi_j^h) = 4\, A_{ij}^h. \tag{2.9}$$

Let $(f, \phi_i^h) = 4 b_i$, then the given linear equation (2.1) is equivalent to

$$a(u, \phi_i^h) = (f, \phi_i^h)_0, \qquad p_i \in \Omega_h, \tag{2.10}$$

or

$$a(u, \phi) = (f, \phi)_0, \qquad \phi \in S_h, \tag{2.11}$$

and u is the (weak) solution of the variational problem.

Next we decompose S_h as a direct sum

$$S_h = S_H \oplus T_h, \tag{2.12}$$

where S_H is the analogous finite element space for the coarser grid and

$$T_h = \operatorname{span}\{\phi_i^h; p_i \in \Omega_h \setminus \Omega_H\} = \{w \in S_h; w(p_i) = 0 \quad \text{for} \quad p_i \in \Omega_H\}.$$

The crucial point is the following observation:

Assertion 2.2. The functional J attains its minimum

$$\text{in } u^{h,k,1} + T_h \quad \text{at } u^{h,k,2}, \tag{2.13a}$$

$$\text{in } u^{h,k,2} + S_H \quad \text{at } u^{h,k,3}, \tag{2.13b}$$

and

$$\text{in } u^{h,k,3} + T_h \quad \text{at } u^{h,k,4}. \tag{2.13c}$$

To verify this assertion, let $\hat{u} \in S_h$, and consider the minimization of J in $\hat{u} + V$, V being a linear subspace of S_h. Then

$$J(\hat{u} + v) = a(v, v) - 2[(f, v)_0 - a(\hat{u}, v)] + \text{const}, \tag{2.14}$$

where const$=J(\hat{u})$. Since the expression $r(v):=(f,v)_0-a(\hat{u},v)$ is linear in v, the minimum v^* is given by

$$a(u^*,\phi)=(f,\phi)_0-a(\hat{u},\phi), \qquad \text{for all} \quad \phi\in V. \tag{2.15}$$

Now the restriction of the minimization of J to the subspace T_h is equivalent to keeping all values on the coarser grid Ω_H fixed. If we put $V=T_h$, then (2.15) reads

$$\left.\begin{array}{rl} & a(u,\phi_i^h)=(f,\phi_i^h), \\ \text{or} & (Ax)_i=b_i, \end{array}\right\} p_i\in\Omega_h\backslash\Omega_H. \tag{2.16}$$

Obviously (2.16) holds after each application of the Gauß-Seidel half-step G_h^1, and in particular for $x^{h,k,2}$ and $x^{h,k,4}$. This proves (2.13a) and (2.13c).

Before we analyse the restriction to S_H, we consider the transition from Ω_h and S_h to Ω_H and S_H in more detail. Each diagonal line of the fine grid, but only every second horizontal or vertical line of the fine grid are found in the coarser grid. The basis functions for the coarser grid are computed with two distinct formulas. Referring to Fig. 2 we have e.g.

$$\phi_1^H=\phi_1^h$$

and

$$\phi_7^H=\phi_7^h+\tfrac{1}{2}(\phi_2^h+\phi_6^h+\phi_8^h+\phi_{12}^h). \tag{2.17}$$

The basis functions for points on the 2nd, 4th, 6th, ... horizontal line are computed like ϕ_1^H, while the others are obtained like ϕ_7^H. In any case

$$\phi_i^H-\phi_i^h\in T_h. \tag{2.18}$$

The minimum of J in $u^{h,k,2}+S_H$ is characterized by

$$a(v,\phi_i^H)=r(\phi_i^h), \qquad \text{for} \quad p_i\in\Omega_H, \tag{2.19}$$

where $r(v)=(f,v)_0-a(u^{h,k,2},v)$. From (2.16) we know that $r(\phi_i^h)=0$ if $p_i\notin\Omega_H$. This and (2.18) imply

$$\begin{aligned} r(\phi_i^H)&=r(\phi_i^h)\\ &=4[b_i^h-(Ax^{h,k,2})_i]=4d_i. \end{aligned}$$

Moreover, the matrix with entries $a(\phi_i^H,\phi_j^H)$ has a structure which is analogous to (2.2). Hence, the solution of (2.19) is computed by Step 3 of the two-level iteration. □

3. Convergence Rate of the Two-Level Iteration

In determing a numerical bound for the convergence rate we use the following abstract lemma (cf. [2b, Theorem 1] and [4]). If V is any closed linear subspace of a Hilbert space, P_V will denote the orthogonal projector onto V. Note that $x-P_Vx$ is the element in $x+V$ with least norm.

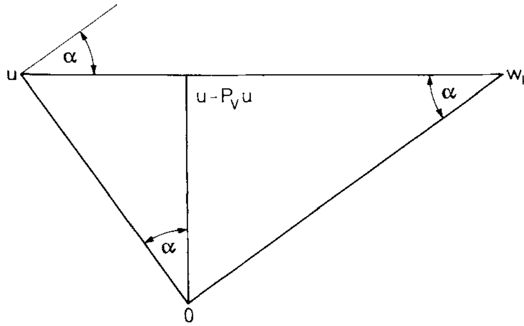

Fig. 3. Illustration of Lemma 3.1

Lemma 3.1. *Let the Hilbert-space U be a direct sum of its subspaces V and W. Assume that there is a* $\gamma<1$ *such that a strengthened Cauchy inequality holds:*

$$|(v, w)| \leqq \gamma \|v\| \cdot \|w\|, \qquad v \in V, \quad w \in W. \tag{3.1}$$

If u is optimal in $u+W$, *i.e., if* $P_W u=0$, *then*

$$\|u-P_V u\| \leqq \gamma \|u\|. \tag{3.2}$$

A simple proof of this lemma is given here with regard to its extension in Sect. 4. First we consider the special case $\dim V = \dim W = 1$, as illustrated in Fig. 3. Then $\|u-P_V u\| = \|u\| \cdot \cos\alpha \leqq \gamma \|u\|$, where α is an angle between vectors from V and W. In the general case we decompose $u=v_1+w_1$, $v_1 \in V$, $w_1 \in W$. By assumption we have $(u, w_1)=0$. Since we know that the lemma is true for $V_1 = \operatorname{span} v_1$, $W_1 = \operatorname{span} w_1$, and $u \in V_1 \oplus W_1$, we have

$$\min_{v \in V} \|u-v\| \leqq \min_{v \in V_1} \|u-v\| \leqq \gamma \|u\|.$$

This proves Lemma 3.1. □

Recall that $S_h \subset H_0^1(\Omega)$ is a Hilbert space when endowed with the inner-product

$$(u, v) = a(u, v). \tag{3.3}$$

From (2.8) and (2.9) we obtain the induced inner-product on the set of coefficient vectors in n-space:

$$(x, y) = 4 x^T A y.$$

The associated norm is the energy norm $\|u\| = \sqrt{a(u, u)}$.

Lemma 3.2. *If* $v \in S_H$ *and* $w \in T_h$, *then*

$$|a(v, w)| \leqq \frac{1}{\sqrt{2}} \|v\| \cdot \|w\|. \tag{3.4}$$

Proof. For the evaluation of $a(v, w)$ we consider first the integral on an arbitrary triangle of Ω_H, e.g., the triangle in Fig. 2 which consists of triangles I and II from the fine triangulation. Since the derivatives are piecewise constant and $w(p_7)=w(p_9)=w(p_{13})=0$, one has

$$\begin{aligned} v_\xi(\mathrm{I}) &= v_\xi(\mathrm{II}), \quad & w_\xi(\mathrm{I}) &= -w_\xi(\mathrm{II}), \\ v_\eta(\mathrm{I}) &= v_\eta(\mathrm{II}), \quad & w_\eta(\mathrm{I}) &= w_\eta(\mathrm{II}), \\ & & |w_\xi| &= |w_\eta| \end{aligned} \tag{3.5}$$

Consequently, the first term vanishes when we integrate over the triangles I and II,

$$\begin{aligned} \left|\int (v_\xi w_\xi + v_\eta w_\eta)\right| &= \left|\int v_\eta w_\eta\right| = \sqrt{\int |v_\eta|^2 \cdot \int |w_\eta|^2} \\ &= \sqrt{\int |v_\eta|^2 \cdot \int \tfrac{1}{2}(|w_\xi|^2 + |w_\eta|^2)} \\ &\leqq \frac{1}{\sqrt{2}} \|v\|_{\mathrm{I+II}} \cdot \|w\|_{\mathrm{I+II}} \\ &\leqq \frac{1}{\sqrt{2}} \cdot \{\tfrac{1}{2}\|v\|^2_{\mathrm{I+II}} + \tfrac{1}{2}\|w\|^2_{\mathrm{I+II}}\}. \end{aligned} \tag{3.6}$$

By summing over all triangles we get

$$|a(v, w)| \leqq \frac{1}{\sqrt{2}} \{\tfrac{1}{2}\|v\|^2 + \tfrac{1}{2}\|w\|^2\}.$$

Therefore (3.4) holds whenever $\|v\| = \|w\| = 1$. A simple homogenity argument shows that (3.4) is correct for all $v \in S_H$, $w \in T_h$. □

Let u^h denote the solution of (2.6) in S_h. Then

$$J(u) = \|u - u^h\|^2 + \text{const.} \qquad \text{for any} \quad u \in S_h. \tag{3.7}$$

By applying Lemma 3.1 to the optimization of $\|u - u^h\|$ instead of $\|u\|$ we obtain the first main result.

Theorem 3.3. *For the two-level iteration, independent of the number* ν *of smoothings, one has*

$$\|x^{h,k+1,0} - x^h\| \leqq \tfrac{1}{2}\|x^{h,k,0} - x^h\|. \tag{3.8}$$

Proof. Since for any $u \in S_h$

$$\begin{aligned} \|G_h^{\mathrm{I}} u - u^h\| &\leqq \|u - u^h\|, \\ \|G_h^{\mathrm{II}} u - u^h\| &\leqq \|u - u^h\|, \end{aligned}$$

it follows that $\|x^{h,k,2} - x^h\| \leqq \|x^{h,k,0} - x^h\|$. By applying Lemma 3.1 with $V = S_H$, $W = T_h$ we get

$$\|x^{h,k,3} - x^h\| \leqq \frac{1}{\sqrt{2}} \|x^{h,k,2} - x^h\|.$$

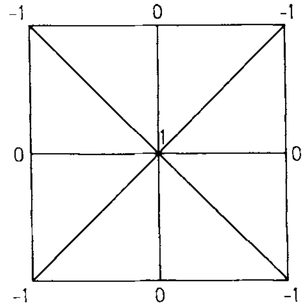

Fig. 4. Values of an approximately worst error function on a period of the domain

Another application with $V=T_h$, $W=S_H$ yields

$$\|x^{h,k,4}-x^h\| \leqq \frac{1}{\sqrt{2}} \|x^{h,k,3}-x^h\|.$$

Combining all estimates we obtain (3.8). □

The constant $\frac{1}{2}$ in (3.8) is the best possible constant for $v=0$. Indeed, let Ω_h be a domain with a very large number of interior points. Assume that on the points of some square the error attains the values given in Fig. 4 and that the error is extended periodically to all interior points of Ω_h. If the influence of the boundary is neglected, one iteration would cause the error to be multiplied by a factor of exactly 0.5.

We will proceed for a moment with the discussion of the optimality of the constants. Assume that the constant in (3.1) is sharp; i.e., we have $(v^*, w^*) = \gamma \|v^*\| \cdot w^*\| \neq 0$ for some $v^* \in V$, $w^* \in W$. Moreover let $u_0 = v + w^*$, $v \in V$. Then the vector in $u_0 + V$ with minimal norm has the form $\lambda v^* + w^*$, $\lambda \in \mathbb{R}$. This is a consequence of $\|\lambda v^* + w^*\| = \gamma \|w^*\|$ (see Fig. 3). Hence, the constant in (3.2) cannot be improved.

Thus the smoothing step, Step 1 in Algorithm 2.1, is reasonable. The half-step G_h^{II} just annihilates the asymptotically worst error function sketched in Fig. 4.

4. The Multi-Level Iteration

Generally, the reduced linear system (2.5) still has a large number of unknowns. Therefore it makes sense to solve it approximately by applying Algorithm 2.1 to the coarser grid. When this process is repeated, a recursive multigrid procedure is established.

Let h_q, $q=0, 1, \ldots, q_{\max}$ be a finite sequence of mesh sizes with $h_{q-1} = \sqrt{2}\, h_q$, $q \geqq 1$. The corresponding grids will be denoted by Ω^q instead of Ω^{h_q}. We will also replace each suffix (or superscript) h_q by q, when we adopt the notation from the previous sections.

Algorithm 4.1. (k-th loop of the iteration on the level q in the recursive algorithm).

1. and 2. Same as in the two-level iteration with $h=h_q$ (and $H=h_{q-1}$).
3. Determine the residual $d=b^q-A^q x^{q,k,2}$. Let y^{q-1} be a solution of

$$A^{q-1}y^{q-1}=d. \tag{4.1}$$

Compute an approximation y satisfying

$$\|y-y^{q-1}\| \leqq \tilde{\delta}_{q-1}\|y^{q-1}\|. \tag{4.2}$$

Specifically, if $q=1$ then (4.1) is solved exactly. If $q>1$ then μ iterations ($\mu=1,2$, or 3) of the level $q-1$ are performed for Eq. (4.1) with the starting vector $x^{q-1,0,0}=0$.

4. and 5. Same as in the two-level iteration with $h=h_q$ (and $H=h_{q-1}$).

Since the auxiliary equations in the coarser grids are now solved only approximately we have to modify the estimate given in Lemma 3.1.

Lemma 4.2. *Let the Hilbert space U be a direct sum of its subspaces V and W such that* (3.1) *holds. Let u be optimal in $u+W$ and let $v_1\in V$ satisfy*

$$\|v_1-P_V u\| \leqq \delta\|P_V u\|. \tag{4.3}$$

Then

$$\min_{w\in W}\|u-v_1-w\| \leqq [\gamma^2+\delta(1-\gamma^2)]\cdot\|u\|. \tag{4.4}$$

Proof. Put $v_0=P_V u$ and $w_0=P_W(u-v_0)$. From Lemma 3.1 we know that $\|u-v_0\|\leqq\gamma\|u\|$ and

$$\|u-v_0-w_0\| \leqq \gamma\|u-v_0\| \leqq \gamma^2\|u\|. \tag{4.5}$$

We may rewrite (4.3) as $v_1-v_0=\delta v$ where $v\in V$ and $\|v\|\leqq\|v_0\|$. Recalling the well-known characterization of closest points in a subspace of a Hilbert space we have

$$\begin{aligned}\|u-v_0-v\|^2 &= \|u-v_0\|^2+\|v\|^2\\ &\leqq \|u-v_0\|^2+\|v_0\|^2=\|u\|^2.\end{aligned} \tag{4.6}$$

Combining this with (4.5) we obtain

$$\begin{aligned}\|u-v_1-(1-\delta)\,w_0\| &= \|u-v_0-\delta v-(1-\delta)\,w_0\|\\ &\leqq \delta\,\|u-v_0-v\|+(1-\delta)\|u-v_0-w_0\|\\ &\leqq [\delta+(1-\delta)\,\gamma^2]\cdot\|u\|.\end{aligned}$$

From this (4.4) is immediate. □

Figure 5 illustrates that (4.4) cannot be improved unless additional information is available.

By modifying the proof of Theorem 3.3 in an obvious manner we obtain our main result:

Theorem 4.3. *The multi-level iteration converges independently of* v, *and*

$$\|x^{q,k+1,0}-x^q\| \leqq \delta_q\|x^{q,k,0}-x^q\|,$$

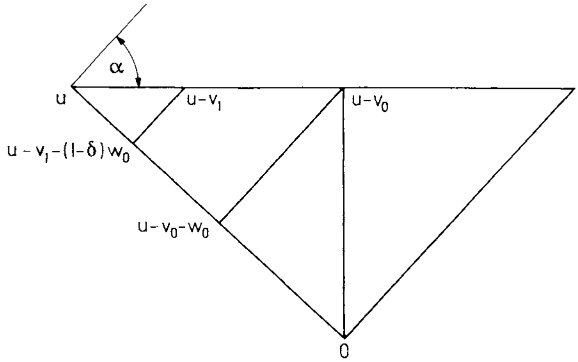

Fig. 5. Illustration of Lemma 4.2

Table 1. Contraction numbers δ_q from Theorem 4.3

μ \ q	0	1	2	3	4	5	6	7	8
1	0	0.5	0.750	0.875	0.938	0.969	0.985		
2	0	0.5	0.625	0.696	0.742	0.776	0.801	0.821	0.837
3	0	0.5	0.563	0.589	0.602	0.610	0.614	0.616	0.617

where δ_q is defined by the recursion

$$\begin{aligned}&\delta_0=0,\\&\delta_q=\tfrac{1}{2}(1+\delta_{q-1}^{\mu}).\end{aligned}\tag{4.7}$$

(The number μ of lower level iterations enters into (4.7) as a power $\tilde{\delta}_{q-1}=\delta_{q-1}^{\mu}$.)

The contraction numbers δ_q are listed for $\mu=1, 2$, and 3 in Table 1. Numerical results with $\mu=1$ reported in [5] show that our estimates are generally far too pessimistic. On the other hand, Hackbusch [5] reported that the observed convergence factor tends indeed to 1 for $q\to\infty$ if the domain is very irregular. It is obvious from Table 1 that in this case the choice $\mu=2$ gives a better performance, though the effort for each iteration loop is larger. We will return to this point in Sect. 6.

5. Non-Uniform Meshes

When variational problems on domains with corners are treated, often the meshes are refined close to the corners [1]. Theorem 4.2 applies to those cases as well. Before we discuss this situation, we turn our attention to a computational matter.

We have assumed that there is a triangulation of the domain Ω consistent with the finest grid. Then the coefficients of the linear system $A^h x=b$ have,

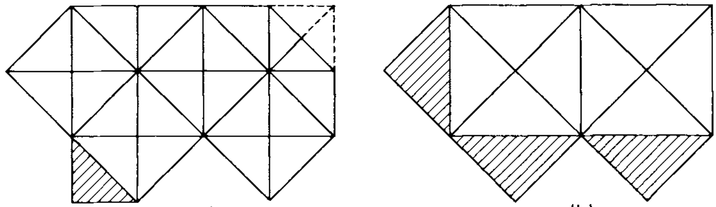

Fig. 6. Grid from Fig. 1 after 2 and 3 reductions

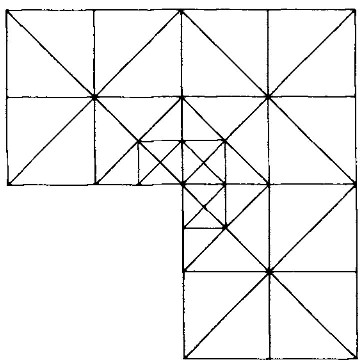

Fig. 7. Irregular grid for L-shaped region

uniformly, the structure given in (2.2). Though the boundary of Ω does not necessarily lie on the lines of the coarser meshes the reduced linear equations still have the standard form (2.2).

To illustrate this, we consider the reduced meshes of the example from Fig. 1. The hatched triangle in Fig. 6a may be eliminated because all its corners lie on the boundary. On the other hand one may complete the square on the upper right hand side, without changing the equations for the inner points. After these modifications the domain is adapted to the new grid. Analogously three triangles may be eliminated with the next coarsening as shown in Fig. 6b.

Consequently, there may be triangles in some Ω_q which belong to the support of a $w \in T_q$ but to no $v \in S_{q-1}$. Since this does not contradict Lemma 3.2, it is no drawback in the analysis of the method.

Now we turn to the point mentioned at the beginning of this section. Figure 7 shows an L-shaped domain, where the mesh is refined in the neighborhood of the corner in the center. Then the coefficients of the associated linear system differ from (2.2). But when the system is solved with a multigrid algorithm, it is not necessary to know (and to compute) the coefficients. The only difference from the standard case is the fact that the computation for the finer grids is restricted to a subdomain $\hat{\Omega}$. If we apply Lemma 3.2 to $v \in S_{q-1}$, $w \in T_q$ we get

$$|a(v, w)| \leq \frac{1}{\sqrt{2}} \|v\|_{\hat{\Omega}} \cdot \|w\|_{\hat{\Omega}}. \tag{5.1}$$

Since in general $\|v\|_{\hat{\Omega}} \ll \|v\|_{\Omega}$, the effective contraction constant will be substantially smaller than $2^{-1/2}$.

A refinement of the grid near the boundary of Ω in order to compensate for the loss of regularity seems reasonable too in this framework.

6. Multigrid Algorithms with $h_{q-1}=2h_q$

Usually, multigrid algorithms are used with sequences of meshes with $h_{q-1}/h_q = 2$ instead of $\sqrt{2}$. Then the decomposition in question is

$$S_h = S_{2h} \oplus T_h' \tag{6.1}$$

where $T_h' = \operatorname{span}\{\phi_i^h;\ p_i \in \Omega_{2h} \setminus \Omega_h\} = T_h \oplus T_{h\sqrt{2}}$. We will see that this decomposition is more advantageous then (2.12). There are only two complications, one is theoretical and the other is practical.

First we derive the constant for the strengthened Cauchy inequality for the decomposition (6.1). From Sect. 3 we know that there are elements $v \in S_{2h}$ and $w \in T_{h\sqrt{2}} \subset T_h'$, for which the ratio $a(v, w)/\|v\| \cdot \|w\|$ is close to $2^{-1/2}$. Hence we do not expect a smaller constant. But, surprisingly, the result is the same as for the previous decomposition. Unfortunately we can prove this only by numerical computations which give no insight into this phenomenon.

Lemma 6.1. *If $v \in S_{2h}$ and $w \in T_h'$, then*

$$|a(v, w)| \leq \frac{1}{\sqrt{2}} \|v\| \cdot \|w\|. \tag{6.2}$$

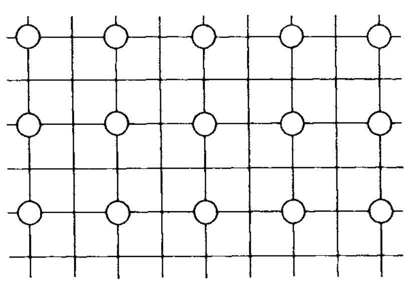

Fig. 8. Points of coarse grid for $h_{q-1}/h_q = 2$

Proof. It is sufficient to consider the functions on a triangle T of Ω_{2h}. A basis of $S_h|_T$ (mod constants) is specified in Table 2 in connection with Fig. 9. The basis functions are enumerated such that the first three ones are symmetrical to the line through p_3 and p_6, while ψ_4 and ψ_5 are antisymmetrical. Moreover $S_{2h|T} = \operatorname{span}\{\psi_1, \psi_4, \text{const.}\}$. A simple computation yields

Table 2. Values of basis function on the triangle in Fig. 9

	p_1	p_2	p_3	p_4	p_5	p_6
ψ_1	2	1	0	1	2	2
ψ_2	0	0	0	0	0	1
ψ_3	0	1	0	1	0	1
ψ_4	−2	−1	0	1	2	0
ψ_5	0	−1	0	1	0	0

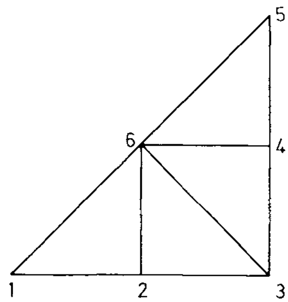

Fig. 9. Enumeration of points for the specification of a basis

$$a(\psi_i, \psi_j)_T = \left(\begin{array}{cc|c|cc} 4 & 2 & & & \\ 2 & 2 & & 0 & \\ \hline & & 2 & & \\ \hline & 0 & & 4 & 2 \\ & & & 2 & 4 \end{array}\right). \tag{6.3}$$

From (6.3) the estimate (6.2) is easily obtained. □

The solution of the variational problem in T_h in Sect. 2 was not problematic. Here the situation is different. The auxiliary problem is equivalent to the solution of a linear system of type (2.3), where all points of Ω_{2h} are extracted from Ω_h and are considered as boundary points. The solution by Gaussian Elimination is still very expensive. Fortunately, there are fast iterative procedures available. Point relaxation has a convergence rate [9; Sect. 3.2] of $-\log \frac{1}{2}$ while the convergence rate for a loop consisting of a horizontal line relaxation and a vertical one is $-\log \frac{1}{9}$. The extension of Lemma 4.2 to the case where both of the auxiliary problems are only solved approximately is left to the reader.

Algorithm 6.2 (k-th loop of the iteration on the level q in the recursive algorithm for $h_{q-1}/h_q = 2$).

1. Given $x^{h,k,0}$ determine $x^{h,k,1}$ satisfying $\|x^{h,k,1} - x^h\| \leqq \|x^{h,k,0} - x^h\|$, e.g. by the same smoothing procedure as in Algorithm 2.1 or line relaxation along the lines omitted in Step 2.

2. Perform line relaxations to $x^{h,k,1}$ first for all horizontal lines not meeting Ω_{2h} and then for all vertical lines with the same property. Denote the results as $x^{h,k,2}$.

3. Determine the residuals

$$d_i=(f,\phi_i^{2h})_0-a(u^{h,k,2},\phi_i^{2h}),\qquad p_i\in\Omega_{2h}.$$

Let y^{q-1} be a solution of $A^{q-1}y^{q-1}=d$. Compute an approximation y as in Algorithm 4.1.

4. Put

$$x^{h,k,3}=x^{h,k,2}+y.$$

5. Perform line relaxations to $x^{h,k,3}$ as in Step 2 above. Denote the results as $x^{h,k,4}$ and proceed with $x^{h,k+1,0}=x^{h,k,4}$.

The advantage of the multigrid algorithms with a mesh ratio $h_{q-1}/h_q=2$ is the strict reduction of the number of unknowns $\dim S_{2h}\approx\frac{1}{4}\dim S_h$. Consequently, the expense for the numerical calculations may be bounded independently of the number of levels, even if we choose $\mu=3$. Then the contraction numbers (see Table 1) are bounded by the solution of the equation $\delta=\frac{1}{2}(1+\delta^3)<1$;

$$\sup_{q\geqq 0}\delta_q\leqq 0.62. \tag{6.4}$$

This is the most pessimistic estimate if the auxiliary problem in T_h' is solved exactly. The rigorous value for Algorithm 6.2 will be slightly larger.

7. Quadratic Elements

When the Galerkin method is performed with piecewise quadratic functions in $C^0(\Omega)$, a straightforward application of the multigrid method would result in (approximate) reductions to functions which are polynomials on larger triangles. But here another multi-level procedure seems to be more advantageous. It is possible to decompose the finite element space such that the nontrivial subspace consists only of piecewise linear functions (cf. [2b, Sect. 3]). Then the resulting matrix has not only a reduced dimension but is also sparser. The algorithms and the numerical analysis derived for the simple case may be used for the steps on the lower levels.

Let Q_h denote the set of C^0-functions that are quadratic polynomials in each triangle T of the triangulation associated to Ω_{2h} (cf. Fig. 10). Consider the decomposition

$$Q_h=S_{2h}\oplus U_h' \tag{7.1}$$

where

$$U_h'=\{w\in Q_h;\ w(p_i)=0 \text{ for all corners } p_i \text{ of the triangles}\}.$$

Lemma 7.1. *If $v\in S_{2h}$ and $w\in U_h'$, then*

$$|a(v,w)|\leqq\sqrt{\tfrac{2}{3}}\,\|v\|\,\|w\|. \tag{7.2}$$

Proof. It is sufficient to consiser v and w on the triangle T given in Fig. 10 when $h=1$. We choose the basis functions

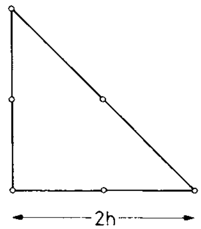

Fig. 10. Triangle whose corners and midpoints lie on a square grid with mesh size h

$$\begin{aligned} &\psi_0=1 && \psi_3=\xi(2-\xi),\\ &\psi_1=\xi, && \psi_4=\eta(2-\eta),\\ &\psi_2=\eta, && \psi_5=\xi\eta. \end{aligned}$$

Noting that

$$S_{2h|T}=\operatorname{span}\{\psi_0,\psi_1,\psi_2\},\qquad U'_{h|T}=\operatorname{span}\{\psi_3,\psi_4,\psi_5\},$$

we compute the matrix

$$a(\psi_i,\psi_j)_{i,j=1}^5=\begin{pmatrix} 2 & 0 & 4/3 & 0 & 4/3\\ 0 & 2 & 0 & 4/3 & 4/3\\ 4/3 & 0 & 8/3 & 0 & 4/3\\ 0 & 4/3 & 0 & 8/3 & 4/3\\ 4/3 & 4/3 & 4/3 & 4/3 & 8/3 \end{pmatrix}.$$

If we replace ψ_5 by $\psi'_5=\sqrt{2}(\psi_5-\frac{1}{2}\psi_3-\frac{1}{2}\psi_4)$ and leave the other functions unchanged, we obtain

$$\tfrac{3}{2}a(\psi'_i,\psi'_j)_{i,j=1}^5=\left(\begin{array}{cc|ccc} 3 & & 2 & & \sqrt{2}\\ & 3 & & 2 & \sqrt{2}\\ \hline 2 & & 4 & & \\ & 2 & & 4 & \\ \sqrt{2} & \sqrt{2} & & & 4 \end{array}\right)=:\begin{pmatrix} 3I & B\\ B^T & 4I \end{pmatrix}.$$

Now the constant γ for (7.2) is estimated by evaluating the spectral radius ρ of $B^T B$.

$$\gamma^2\leqq\tfrac{1}{3}\,\tfrac{1}{4}\,\rho(B^T B)=\tfrac{2}{3}.$$

This proves (7.2). □

The decomposition (7.1) has some properties and consequences which are very similar to those of (6.1). The solution of the variational problem in U'_h is not trivial but may be treated by line relaxation. Moreover, we may also interpret the decomposition (7.1) as the result of two steps. The intermediate space here contains those functions, which are piecewise linear on the horizontal and vertical lines but which are quadratic polynomials along the diagonals.

Finally we note that rectangular triangles are less interesting for practical computations than general triangles. For the extension of Lemma 7.1 to general regular triangulation refer to the proof of Lemma 1 in [2], from which the interpretation of the quotient $(1+\gamma)/(1-\gamma)$ as a condition number becomes also clear.

8. The Paradox of Smoothing. Further Remarks

In his paper [5] Hackbusch interprets the two-level iteration as a combination of a "*smoothing procedure*" and a "*correction by approximation*". From that point of view the Steps 2 and 5 in Algorithm 2.1 contain the smoothing procedures. We will show, however, that to the contrary (2.4) may produce a rough approximation from a smooth one. This paradox is not only of theoretical interest, but has some consequences for practical computations.

Generally, one chooses the minimal mesh size $h=h_{q_{\max}}$ so conservatively, that the finite element solutions of the elliptic equation for S_h and S_{2h} do not differ by much. If we decompose the solution $u^h=v^h+w^h$, $v^h\in S_{2h}$, $w^h\in T_h'$, we will therefore expect $\|w^h\|\ll\|v^h\|$. Assume that the algorithm is used with the parameter $\nu=0$. When we start it with $x^{h,0,0}=x^{h,0,1}=0$, it will produce a vector $x^{h,k,2}$, whose portion in T_h' is substantially greater.

Indeed, let there be a $w_1\in T_h'$, $\|w_1\|\neq 0$ with $|(v^h,w_1)|>\frac{1}{3}\|v^h\|\cdot\|w_1\|$. Recalling that Lemma 6.1 is stated with a close to optimal constant, this assumption seems reasonable. Then $\|x^{h,0,2}\|\leqq\sqrt{\frac{8}{9}}\,\|x^{h,0,0}\|$ and the part of $x^{h,0,2}$ in T_h' is at least $\frac{1}{3}\|x^{h,0,2}\|\gg\|w^h\|$.

In this case it is more appropriate to begin with the computation in the coarser grid instead of the *smoothing*. This argument agrees with the advice of some authors to start the multigrid iteration at the lowest level and not at the highest one.

At this point we might try to explain why the observed convergence rate is better than the theoretical one given in Sect. 3. The interference between the two auxiliary optimizations in V and W is worst, when the error is decomposed into two parts with the same order of magnitude. We do not meet that situation when looking for smooth solutions. On the other hand this explanation is a speculation because it is not known how to check it seriously.

However, we can explain another well-known phenomenon. Consider the Dirichlet problem

$$\begin{aligned} \varepsilon u_{\xi\xi}+u_{\eta\eta}&=f \quad \text{in } \Omega,\\ u&=0 \quad \text{on } \partial\Omega, \end{aligned} \tag{8.1}$$

where ε is a very small positive constant [8, p. 948]. Recalling (3.5) we have

$$\int|w_\eta|^2=\frac{1}{1+\varepsilon}\int(\varepsilon|w_\xi|^2+|w_\eta|^2).$$

The energy norm is now so anisotropic that instead of (3.4) we get only $\gamma<(1+\varepsilon)^{-1/2}$. Therefore, a multigrid procedure with point relaxation is not very effective (c.f. [8]).

References

1. Babuška, I., Rheinboldt, W.: Mathematical problems of computational decisions in the finite element method. Technical report TR-426, University of Maryland, 1975
2. Bank, R.E., Dupont, T.: An optimal order process for solving finite element equations. Math. Comput. **36**, 35–51 (1981)
   and: Analysis of a two-level scheme for solving finite element equations. Numer. Math. (1981, in press)
3. Collatz, L.: Numerische Behandlung von Differentialgleichungen. Berlin-Göttingen-Heidelberg: Springer 1951
4. Gunn, J.E.: The solution of difference equations by semi-explicit iterative techniques. SIAM J. Numer. Anal. **2**, 24–45 (1965)
5. Hackbusch, W.: On the multigrid method applied to difference equations. Computing **20**, 291–306 (1978)
6. Meis, T., Marcowitz, U.: Numerische Behandlung partieller Differentialgleichungen. Berlin-Heidelberg-New York: Springer 1978
7. Nicolaides, R.A.: On the l^2-convergence of an algorithm for solving finite element equations. Math. Comput. **31**, 892–906 (1977)
8. Nicolaides, R.A.: On some theoretical and practical aspects of multigrid methods. Math. Comput. **33**, 933–952 (1979)
9. Varga, R.S.: Matrix Iterative Analysis. Englewood Cliffs: Prentice-Hall 1962
10. Wesseling, P.: A convergence proof for a multiple grid method. In: Numerical analysis. Proceedings Dundee 1979 (G.A. Watson, ed.), pp. 164–183. Lecture Notes in Mathematics, Vol. 733. Berlin-Heidelberg-New York: Springer 1980

Received September 9, 1980