MATHEMATICS OF COMPUTATION
VOLUME 42, NUMBER 166
APRIL 1984, PAGES 505-519

The Convergence Rate of a Multigrid Method with Gauss-Seidel Relaxation for the Poisson Equation

By Dietrich Braess

Abstract. The convergence rate of a multigrid method for the numerical solution of the Poisson equation on a uniform grid is estimated. The results are independent of the shape of the domain as long as it is convex and polygonal. On the other hand, pollution effects become apparent when the domain contains reentrant corners. To estimate the smoothing of the Gauss-Seidel relaxation, the smoothness is measured by comparing the energy norm with a (weaker) discrete seminorm.

1. Introduction. In this paper we will treat a multigrid method for the numerical solution of a discretization of the elliptic boundary value problem

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

Here Ω is assumed to be a polygonal domain such that its boundary matches the horizontal, vertical or diagonal lines of a uniform grid. The discretization will be done by the finite element approximation which leads to the standard 5-point formula.

The advantage of the multigrid method is the fact that the convergence rate can be bounded away from one while in the classical iterative procedures it tends to 1 as $h \to 0$. Here the convergence rate for a sequence (u^k), $u^k \to u$, is measured by the error damping factor $\delta = \sup_k \{\| u^{k+1} - u\| / \|u^k - u\|\}$ with an appropriately chosen norm.

There are two types of convergence proofs for multigrid methods in the literature. The results of the first type, e.g. [1], [8], [13], [19], refer to arbitrary convex domains or to domains with smooth boundary. Moreover, more general elliptic problems than (1.1) are treated. A convergence rate bounded away from 1 is established for sufficiently many smoothing steps under the condition that the problem (and the solution) are sufficiently regular. The investigations of the second type, e.g. [4], [7], [11], [16], establish explicit and promising bounds for the convergence rate by Fourier methods; but they are restricted to rectangular domains.

Because of the gap between the two theories there is still the unanswered question whether the good performance of multigrid methods depends on regularity properties [9, p. 157].

Received September 29, 1981; revised October 22, 1982.
1980 *Mathematics Subject Classification*. Primary 65N20, 65F10.

We will partially fill this gap and establish explicit bounds for arbitrary convex (polygonal) domains. A typical result from Table 1 in Section 6 shows that the convergence rate is 0.172 or better, when at least one Gauss-Seidel relaxation step per cycle is performed for the smoothing. Our result is not far from the number 0.125 which was recently determined as the asymptotic convergence rate for rectangular domains [16]. The given rate guarantees quick convergence of the multigrid iteration with W-cycles. Moreover, we show that the multigrid iteration with V-cycles has a convergence rate better than 0.5. Though our theoretical result might be still considered as a poor bound relative to the actually observed efficiency, no comparable rigorous result for V-cycles seems to exist in the literature.

On the other hand, in our theoretical study, pollution effects become apparent when the domain contains reentrant corners (*einspringende Ecken*). Thus, these corners may spoil the very fast convergence. Fortunately, their influence on the multigrid iteration may probably be considered as a pollution effect of local character. We will comment on this in the last section, though a rigorous treatment goes far beyond the framework of this paper. In any case, the convergence rate cannot become worse than $1/2$, since this bound holds for arbitrary domains and not merely for convex ones [3].

The linear equations from the discretization of (1.1) characterize the solution of a variational problem in a finite element space S_h. The central idea is the decomposition of S_h as a direct sum of two subspaces: $S_h = V \oplus W$, where $V = S_H$ is the finite element space for a coarser grid [2], [3], [11], cf. also [10], [12]. The alternate solution of the variational problem in $V \oplus W$ generates an iteration, for which the convergence rate may be estimated via a strengthened Cauchy inequality. In contrast to the previous investigations cited above, we will control the multiplying factor in this Cauchy inequality for the individual elements. By measuring smoothness with a discrete seminorm, which like the energy norm may be defined locally but which is weaker, we find quasi-orthogonality for the decomposition in just those cases where the Gauss-Seidel relaxation is not effective, and conversely. Finally a duality argument enables us to prove convergence of the algorithm with V-cycles.

Our analysis refers to multigrid methods where the mesh-size ratio is $\sqrt{2}$ [3], [7], [16]. Similar investigations for algorithms with mesh ratio 2 are found in [18].

2. Notation. Let $\Omega \subset \mathbf{R}^2$ be a bounded domain. Assume that there is a triangulation of Ω with rectangular isosceles triangles with sides of length h and $H = \sqrt{2}\,h$ (see Figure 1). The set of grid points $\{p_i\}$ which are contained in Ω is denoted by Ω_h, while Ω_H refers to the (rotated coarser) subgrid formed by triangles with sides of length H and $H\sqrt{2} = 2h$. Similarly $\Omega_{2h} \subset \Omega_H$ is defined. When we associate to each grid point one or two colors:

black	to the points of $\Omega_h \setminus \Omega_H$,
white/red	to the points of $\Omega_H \setminus \Omega_{2h}$,
white/green	to the points of Ω_{2h},

then the black-white coloring endows Ω_h with the structure of a checkered board, while the red-green one does the same for Ω_H.

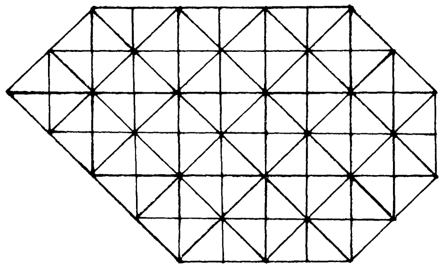

FIGURE 1

As usual, let S_h denote the space of those functions from $C(\Omega)$ which are linear on each triangle of Ω_h and which vanish on $\partial\Omega$. The discretization of (1.1) for the grid Ω_h is understood to be the (solution of the) variational problem in S_h:

$$J(u) := a(u, u) - 2(f, u)_0 \to \min! \tag{2.1}$$

Here,

$$a(u, v) = \int_\Omega (u_\xi v_\xi + u_\eta v_\eta)\, d\xi\, d\eta, \qquad (u, v)_0 = \int_\Omega uv\, d\xi\, d\eta. \tag{2.2}$$

When S_h is equipped with the energy norm $\|u\| = \sqrt{a(u, u)}$, the norm is easily given in terms of the values on the grid $u_i = u(p_i)$:

$$\|u\|^2 = \sum_{\substack{i, j \\ d(i, j) = h}} (u_i - u_j)^2. \tag{2.3}$$

The summation runs over all pairs of grid points with Euclidean distance h, and the points on the boundary are included. Here and in subsequent symmetrical sums, each pair is taken only once.

Next we endow S_h with a seminorm $|\cdot|$:

$$|u|^2 = \sum_{\substack{n, m \\ p_n, p_m \in \Omega_H \\ d(n, m) = 2h}} \frac{1}{2}(u_n - u_m)^2. \tag{2.4}$$

In this sum the terms related to points next to the boundary are to be understood as follows. Let (p_n, p_m) be a pair of points with distance $2h$, such that p_n is an interior point of Ω while $p_m \notin \bar{\Omega}$. (The point in the middle between p_n and p_m is located on $\partial\Omega$.) Then by convention

$$\tfrac{1}{2}(u_n - u_m)^2 \text{ is to be replaced by } (u_n - 0)^2. \tag{2.5}$$

We note that

$$|u| \leqslant \|u\| \quad \text{for each } u \in S_h, \tag{2.6}$$

because we have $(u_n - u_m)^2 \leqslant 2(u_n - u_{nm})^2 + 2(u_{nm} - u_m)^2$ whenever $p_n, p_m \in \Omega_{2h}$ and p_{nm} is the grid point in between.

We mention that in the analysis of the Gauss-Seidel iteration, seminorms of the more general form

$$|u|^2_{\delta, \tilde{\Omega}_h} := c(\delta) \sum_{\substack{p_k, p_l \in \tilde{\Omega}_h \\ d(p_k, p_l) = \delta}} (u_k - u_l)^2 \tag{2.7}$$

will enter, where $\tilde{\Omega}_h = \Omega_H$ or $\Omega_h \setminus \Omega_H$ and $\delta = H$, $2h$ or $2H$. They will be estimated by $\|u\|$ and $|u|$, which are also special cases of the general expression given in (2.7).

3. Decomposition of S_h and the Strengthened Cauchy Inequality. As in [3], we decompose S_h as a direct sum

$$S_h = S_H \oplus T_h, \tag{3.1}$$

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

$$T_h = \{ w \in S_h;\ w(p_i) = 0 \text{ for } p_i \in \Omega_H \}. \tag{3.2}$$

The letters u, v and w will always be understood in the spirit of the decomposition $u = v + w$ according to (3.1). In particular, we may write

$$|u| = |v| \tag{3.3}$$

because the seminorm $|\cdot|$ refers to points of Ω_H only. (On the other hand, $|\cdot|$ is a norm for the subspace S_H.)

In [2], [3], [11] *strengthened Cauchy inequalities* were used with constants which depended only on the choice of the subspace. Here we will derive improved bounds which depend on the smoothness of the individual elements.

LEMMA 3.1. *Assume that Ω_h has no concave corner at the points of $\Omega_h \setminus \Omega_H$. If $v \in S_H$ and $w \in T_h$, then*

$$|a(v,w)| \leqslant \sqrt{\frac{1}{2}\left(1 - \frac{|v|^2}{\|v\|^2}\right)}\, \|v\|\,\|w\|. \tag{3.4}$$

Proof. Let $v \in S_H$ and $w \in T_h$. In order to estimate $a(v, w)$, consider the integral over the triangle I in Figure 2a. We recall that v_ξ, v_η and w_η are constant while $|w_\xi| = |w_\eta|$ and w_ξ changes its sign at the broken line (cf., (3.5) in [3]). Now

$$\begin{aligned}
\int_{\mathrm{I}} (v_\xi w_\xi + v_\eta w_\eta) &= \int_{\mathrm{I}} v_\eta w_\eta \\
&= \frac{1}{2}\int_{\mathrm{I}} v_\eta^2 + \frac{1}{2}\int_{\mathrm{I}} w_\eta^2 - \frac{1}{2}\int_{\mathrm{I}} (v_\eta - w_\eta)^2 \\
&= \frac{1}{2}\int_{\mathrm{I}} (v_\xi^2 + v_\eta^2) + \frac{1}{4}\int_{\mathrm{I}} (w_\xi^2 + w_\eta^2) - \frac{1}{2}\int_{\mathrm{I}} v_\xi^2 - \frac{1}{2}\int_{\mathrm{I}} (v_\eta - w_\eta)^2 \\
&=: A_{\mathrm{I}} + B_{\mathrm{I}} - C_{\mathrm{I}} - D_{\mathrm{I}}.
\end{aligned} \tag{3.5}$$

For convenience we will skip the discussion of the influence of the boundary for a moment and first proceed as if there are only interior triangles.

Obviously, when summing over all triangles, we get

$$\sum_\nu C_\nu = \frac{1}{2} \sum_{\substack{n,\, m \text{ green} \\ d(n,m) = 2h}} \frac{1}{2}(v_n - v_m)^2. \tag{3.6}$$

The last term in (3.5) is also evaluated in connection with the corresponding term for triangle II.

$$\begin{aligned}
D_{\mathrm{I}} + D_{\mathrm{II}} &= \frac{1}{2}[v_1 - (v - w)_5]^2 + \frac{1}{2}[v_3 - (v - w)_5]^2 \\
&\geqslant \frac{1}{2} \min_{z \in \mathbf{R}} \{ (v_1 - z)^2 + (v_3 - z)^2 \} = \frac{1}{4}(v_1 - v_3)^2.
\end{aligned} \tag{3.7}$$

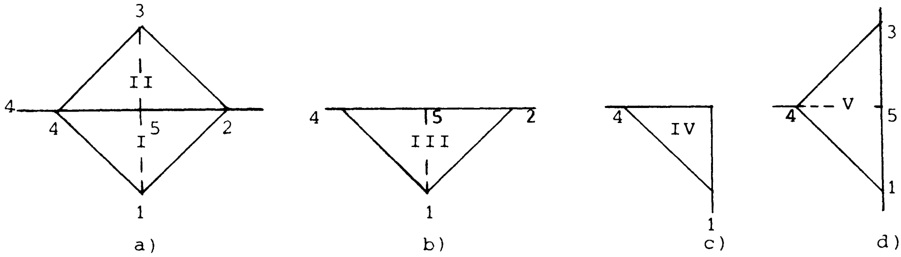

FIGURE 2

Specifications for the proof of Lemma 3.1
(*The points* p_2 *and* p_4 *belong to* Ω_{2h})
a: *triangles in the interior*; b, c, d: *boundary situations.*

Hence,

$$\sum_{\nu} D_\nu \geqslant \frac{1}{2} \sum_{\substack{n,\, m \text{ red} \\ d(n,m)=2h}} \frac{1}{2}(v_n - v_m)^2,$$

and the collection of terms yields

$$a(v, w) \leqslant \frac{1}{2}(\|v\|^2 - |v|^2) + \frac{1}{4}\|w\|^2. \tag{3.8}$$

Now we turn our attention to the triangles next to the boundary. Because of the zero boundary condition, w and v_ξ vanish in triangle III (see Figure 2b). Consequently,

$$\int_{\mathrm{III}} (v_\xi w_\xi + v_\eta w_\eta) = 0 = \frac{1}{2}\int_{\mathrm{III}} (v_\xi^2 + v_\eta^2) + 0 - 0 - \frac{1}{2}\int_{\mathrm{III}} v_\eta^2$$
$$= A_{\mathrm{III}} + B_{\mathrm{III}} - C_{\mathrm{III}} - D_{\mathrm{III}}.$$

This splitting is consistent with the summation for the regular triangles. Since also w vanishes on the triangles IV and V, the analysis there is analogous. We note that in this boundary formation triangle V belongs to the grid Ω_H; but this does not change the arguments.

Finally, if $\|w\| = \sqrt{2(\|v\|^2 - |v|^2)}$, it follows from (3.8) that

$$a(v, w) \leqslant \|v\|^2 - |v|^2 = \sqrt{\frac{1}{2}(\|v\|^2 - |v|^2)}\, \|w\|. \tag{3.9}$$

A simple homogeneity argument shows that (3.9) holds for all $w \in T_h$ and that the left-hand side may be replaced by its absolute value. $\square$

We note that given $v \in S_h$ one can find a $w \in T_h$, $w \neq 0$, such that equality holds in (3.4). To this end we need only choose w such that (3.7) becomes an equality. (Actually, this w is computed by a half step of the Gauss-Seidel iteration.)

As a consequence of Lemma 3.1 we have

$$|v| \leqslant \|v\|. \tag{3.10}$$

We will often put $\lambda = |v|/\|v\|$. In order to estimate $|u|/\|u\|$, we refer to Figure 3. Given $u = v + w$, we conclude that $\|u\| \geqslant \|v\| \sin\alpha$, where by Lemma 3.1 $\cos^2\alpha \leqslant \frac{1}{2}(1 - \lambda^2)$. Hence,

$$|u| = |v| = \lambda\|v\| \leqslant \frac{\lambda}{\sin\alpha}\|u\| \leqslant \frac{\lambda}{\sqrt{(1+\lambda^2)/2}}\|u\|. \tag{3.11}$$

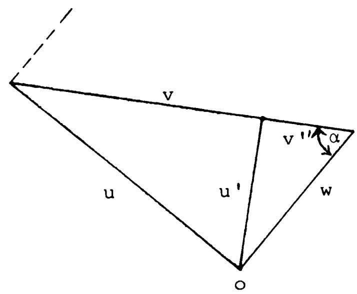

FIGURE 3

Decomposition $u = v + w$ after the optimization in W and $u' = v' + w$ after the optimization in V. Here $\cos\alpha = a(v, w)/(\|v\| \cdot \|w\|)$.

4. The Gauss-Seidel Relaxation. The solution of the variational problem (2.1) in S_h is determined by a linear system of the form

$$u_i = \frac{1}{4}\sum_j{}'_h u_j + b_i, \qquad p_i \in \Omega_h, \tag{4.1}$$

where Σ'_h refers to the summation over all neighbors in the grid Ω_h with distance h. The numerical solution of linear equations of the form (4.1) and the convergence rate of the multigrid methods for solving them is the central point of this paper.

A classical tool to solve (4.1) is the Gauss-Seidel iteration [17] which for convenience will be split into two steps:

$$(G_h^{\mathrm{I}}u)_i = \begin{cases} u_i & \text{if } p_i \in \Omega_H, \\ \dfrac{1}{4}\sum_j{}'_h u_j + b_i & \text{if } p_i \notin \Omega_H. \end{cases}$$

$$(G_h^{\mathrm{II}}u)_i = \begin{cases} \dfrac{1}{4}\sum_j{}'_h u_j + b_i & \text{if } p_i \in \Omega_H, \\ u_i & \text{if } p_i \notin \Omega_H. \end{cases}$$

In multigrid algorithms the point Gauss-Seidel relaxation $G_h^{\mathrm{I}} \cdot G_h^{\mathrm{II}}$ often occurs as a smoothing procedure. Here it is also applied for another purpose: The minimum of the variational functional $J(u)$ in the subset $u + T_h$ is just $G_h^{\mathrm{I}}u$. The following lemma will show that the error of an approximate solution will be reduced by the relaxation provided that it is not smooth. For another inequality which describes this effect, the reader is referred to [18].

We recall that interest in error estimates naturally leads to a study of the homogeneous equation. Therefore in the rest of this section we will assume $b = 0$.

LEMMA 4.1. *Assume that Ω is convex. If $u = G_h^{\mathrm{I}}u$, then*

$$\|G_h^{\mathrm{II}}u\| \leqslant |u|. \tag{4.2}$$

Proof. First we will ignore the boundary to simplify the analysis.

Put $\bar{u} = G^{\mathrm{II}}u$. Given $p_0 \in \Omega_H$, choose the (local) indices 1, 2, 3 and 4 for its neighbors in Ω_h. We will make use of the identity:

$$2\sum_{j=1}^{4} z_j^2 = (z_1 - z_2)^2 + (z_2 - z_3)^2 + (z_3 - z_4)^2 + (z_4 - z_1)^2 + \frac{1}{2}(z_1 + z_2 + z_3 + z_4)^2 - \frac{1}{2}(z_1 - z_2 + z_3 - z_4)^2.$$

From $\bar{u} = G_h^{\mathrm{II}}u$ it follows that

$$\bar{u}_0 = \frac{1}{4}\sum_{i=1}^{4} \bar{u}_i = \frac{1}{4}\sum_{i=1}^{4} u_i.$$

When we put $z_i = \bar{u}_i - \bar{u}_0 = u_i - \bar{u}_0$, we obtain from the identity above

$$(4.3)\quad \sum_{j=1}^{4} (\bar{u}_j - \bar{u}_0)^2 \leqslant \frac{1}{2}\{(u_1 - u_2)^2 + (u_2 - u_3)^2 + (u_3 - u_4)^2 + (u_4 - u_1)^2\}.$$

We rewrite (2.3):

$$(4.4)\quad \|\bar{u}\|^2 = \sum_{i \text{ white}} \sum_{\substack{j \text{ black} \\ d(i,j)=h}} (\bar{u}_i - \bar{u}_j)^2.$$

The inner sum may be estimated by using (4.3). The summation yields

$$(4.5)\quad \|G_h^{\mathrm{II}}u\|^2 \leqslant \sum_{\substack{k,l \text{ black} \\ d(k,l)=H}} (u_k - u_l)^2.$$

The factor $1/2$ in (4.3) compensates for the fact that each pair k, l occurs in two terms of the outer sum. For the situation shown in Figure 4 this happens for $i = 2$ and $i = 6$.

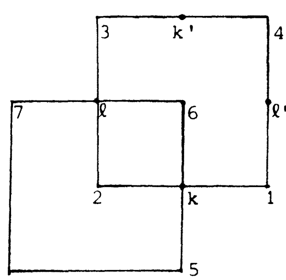

FIGURE 4

Illustration to the estimation of (4.5) (k, l, k', l' *are black*).

Next we consider one term from the sum in (4.5) and enumerate the neighbors of p_k and p_l as illustrated in Figure 4. Since $u = G^{\mathrm{I}}u$, it follows that u_k and u_l are the mean values of the numbers which u attains at the neighbors.

$$(4.6)\quad 16(u_k - u_l)^2 = [u_1 + u_5 - u_3 - u_7]^2 \leqslant 2(u_1 - u_3)^2 + 2(u_5 - u_7)^2.$$

The differences refer to pairs of points at a distance $2h\sqrt{2} = 2H$. When the sum is evaluated, each pair occurs twice, e.g., $(u_1 - u_3)^2$ is found in the bounds for $(u_k - u_l)^2$ and for $(u_{k'} - u_{l'})^2$. Hence,

$$(4.7)\quad \|G^{\mathrm{II}}u\|^2 \leqslant \frac{1}{4} \sum_{\substack{m,n \text{ white} \\ d(m,n)=2H}} (u_m - u_n)^2.$$

Finally consider an arbitrary square with length $2h$ and vertices in Ω_H, e.g. the upper one in Figure 4. Then we treat the terms in (4.7) corresponding to the diagonals according to

$$(u_1 - u_3)^2 + (u_2 - u_4)^2 = (u_1 - u_2)^2 + (u_2 - u_3)^2 + (u_3 - u_4)^2 + (u_4 - u_1)^2 - (u_1 - u_2 + u_3 - u_4)^2.$$

Since each side of length $2h$ separates two squares, we obtain by summation the estimate wanted:

$$\|G^{\mathrm{II}}u\|^2 \leqslant \frac{1}{2} \sum_{\substack{n,\, m \text{ white} \\ d(n,m)=2h}} (u_n - u_m)^2. \tag{4.8}$$

Now we turn our attention to the analysis of boundary terms. The case, where the white point p_0 lies on a (horizontal) boundary, is presented in Figure 5a. Then

$$\sum_{\substack{j \\ d(j,0)=h}} (u_j - \bar{u}_0)^2 = (u_1 - 0)^2 = \frac{1}{2}(u_1 - u_2)^2 + \frac{1}{2}(u_1 - u_3)^2 \tag{4.9}$$

is consistent with (4.3).

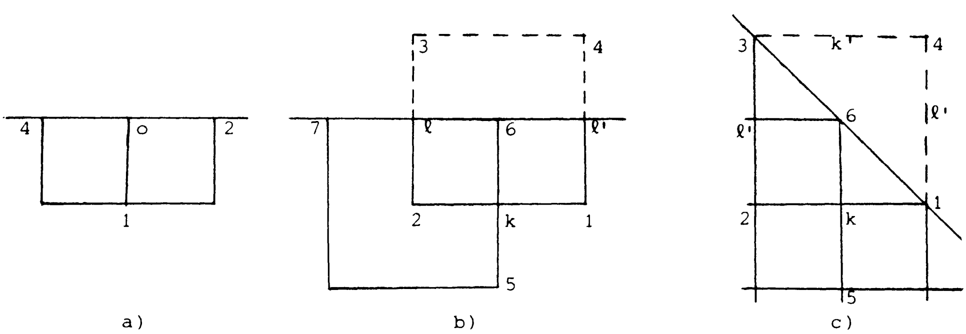

FIGURE 5
Boundary formations.

The boundary configuration which corresponds to that one from Figure 4 is shown in Figure 5b. Then only u_k is a mean value, while $u_l = 0$ is fixed, and

$$\begin{aligned} 16(u_k - u_l)^2 &= (u_1 + u_2 + u_5)^2 \\ &\leqslant 4(u_1 - 0)^2 + 4(u_2 - 0)^2 + 2(u_5 - u_7). \end{aligned}$$

The last term is standard, while the first ones replace $2(u_1 - u_3)^2$. Another copy of $(u_k - u_l)^2$ replaces $2(u_1 - u_5)^2$. Therefore we obtain immediately the expression comparable to (4.8) and the step leading to (4.7) is skipped. The doubling of the weights is consistent with (2.5).

When a white point lies on a diagonal boundary (see Figure 5c), the corresponding terms from (4.4) are directly expressed by those of (4.7). We have

$$\begin{aligned} 16(u_k - u_1)^2 = 16(u_k - u_6)^2 &= (u_2 + u_5)^2 \\ &\leqslant 2(u_1 - u_2)^2 + 2(u_5 - u_6)^2. \end{aligned}$$

We note that no other term of the form $(u_1 - u_2)^2$ from the square (p_1, p_2, p_3, p_4) enters into the right-hand side of (4.8). So we are done.

We abandon the discussion of convex corners because their treatment is simpler or is done by combining arguments from above. $\square$

The reader will observe in the preceding proof that $|u|_{\delta,\ldots}$ may always be estimated by some $|u|_{\delta',\ldots}$ whenever $\delta' < \delta$. The converse is possible only in connection with relaxation.

Since G_h^{I} acts only on the w-part and $|u|$ depends only on the v-part we have

COROLLARY 4.2. *If Ω is convex, then $\|G_h^{\mathrm{II}} \cdot G_h^{\mathrm{I}} u\| \leqslant |u|$ for all $u \in S_h$.*

We note that the constant 1 in (4.2) is the best possible because the convergence rate of the Gauss-Seidel iteration is close to 1 and $\|G_h^{\mathrm{II}} \cdot G_h^{\mathrm{I}}\| = 1 - O(h^2)$, see [17].

5. Remark on the Alternate Method. One element of the multigrid algorithms is the method of alternate minimization in two subspaces [2]. We find it—though not in its pure form—in two connections. The combination of a coarse-grid-correction with the relaxation (half-) step G_h^{I} is the alternate method associated with the decomposition $S_h = S_H \oplus T_h$. On the other hand, the classical Gauss-Seidel relaxation $G_h^{\mathrm{I}} \cdot G_h^{\mathrm{II}}$ is associated with the decomposition $S_h = T_h \oplus \hat{T}_h$, where $\hat{T}_h = \{ u \in S_h;\ u(p_i) = 0 \ \text{ for all } \ p_i \in \Omega_h \setminus \Omega_H\}$.

Let $H = V \oplus W$, where V and W are closed subspaces of the Hilbert space H. Denote the orthogonal projections onto V and W by P_V and P_W, resp. Moreover, put $Q_V = \mathrm{id} - P_V$ and $Q_W = \mathrm{id} - P_W$. When the alternate method is used for the homogeneous problem, the iteration is given by (5.1).

Remark 5.1. Given $u^0 \in H$, let

$$u^{\nu+1/2} = Q_V u^\nu, \qquad u^{\nu+1} = Q_W u^{\nu+1/2}, \qquad \nu = 0,1,2,\ldots. \tag{5.1}$$

Then after the first half-step the relative norm reduction becomes successively slower, i.e.,

$$\frac{\|u^{\nu+1}\|}{\|u^{\nu+1/2}\|} \geqslant \frac{\|u^{\nu+1/2}\|}{\|u^\nu\|} \tag{5.2}$$

holds for any integer or half-integer $\nu \geqslant 1/2$.

Proof. Assume that $u = Q_W u$. Recalling that Q_V and Q_W are projection operators, we get

$$\begin{aligned}\|Q_V u\|^2 &= (Q_V u, Q_V u) = (u, Q_V u) = (Q_W u, Q_V u)\\ &= (u, Q_W Q_V u) \leqslant \|u\| \|Q_W Q_V u\|.\end{aligned}$$

Since any u^ν, ν an integer $\nu \geqslant 1$, satisfies the assumption above, it follows that (5.2) is true for these ν's. By interchanging the role of Q_V and Q_W we get the statement for half-integer ν's. $\square$

Obviously (5.2) means that the function $\nu \mapsto \log\|u^\nu\|$ is convex.

6. Convergence Rate of the Multigrid Method. Now we are in position to analyze a multigrid method for the numerical solution of (4.1) on uniform meshes. Let h_q, $q = 0,1,\ldots, q_{\max}$, be a finite sequence of mesh-sizes with $h_{q-1} = \sqrt{2}\, h_q$, $q \geqslant 1$. The corresponding grids will be denoted by Ω^q instead of Ω^{h_q}. We will replace each suffix (or superscript) h_q by q, when adapting the notation from the previous sections.

If $q = 0$, the linear equations associated to Ω^q are solved directly. For $q > 1$ an iteration is defined as a recursive process. Each iteration loop contains one *correction* on the next coarser grid and *r smoothing steps*, $r \geqslant 1$.

As usual, the variables carry three superscripts; these refer to (1) the level, (2) an iteration count, and (3) a count of the steps within one iteration loop.

For convenience, we understand G_q^ν to be G_q^{I} if ν is odd and that it is G_q^{II} otherwise.

ALGORITHM 6.1 (kth loop at level q for the multigrid iteration with r smoothings).

0. *Start*. Given $u^{q,k,0}$. If $k = 0$, replace it by $G_q^{r+1}u^{q,k,0}$.

1. *Pre-smoothing*. For $\nu = 1, 2, \ldots, r - 1$, compute

$$u^{q,k,\nu} = G_q^{r+1-\nu}u^{q,k,\nu-1}.$$

2. *Transition step*. Put $u^{q,k,r} = G_q^{\mathrm{I}}u^{q,k,r-1}$.

3. *Coarse grid correction*. Let u^{q-1} denote the solution of the variational problem

$$J(u^{q,k,r} + u) \to \min!$$

when $u \in S_{q-1}$.

If $q - 1 = 0$, compute $v_1 = u^{q-1}$. If $q - 1 > 0$, compute an approximate solution v_1 to u^{q-1} by applying $\mu = 1$ or $\mu = 2$ iteration steps at the level $q - 1$, starting with $u^{q-1,0,0} = 0$. Put $u^{q,k,r+1} = u^{q,k,r} + v_1$.

4. *Post-smoothing*. For $\nu = 1, 2, \ldots, r + 1$, determine

$$u^{q,k,r+\nu+2} = G_q^\nu u^{q,k,r+\nu+1},$$

and proceed with $u^{q,k+1,0} = u^{q,k,2r+2}$.

Here r Gauss-Seidel relaxation sweeps are partly performed before the correction step and partly after it. Another half-step is added, because it simplifies the transfer to the coarse grid [3], [7]. We may abandon the discussion of the numerical realization of Algorithm 6.1, because this can be found in [3, Section 2] and in [16]. Other aspects for the preparation of efficient codes are given in [6].

The estimate of the convergence rate will be established for a generalization of Step 3.

3′. Determine an approximation v_1 to u^{q-1}, where u^{q-1} is defined as in Step 3, such that

$$\|v_1 - u^{q-1}\| < \delta\|u^{q-1}\|, \tag{6.1}$$

and put $u^{q,k,r+1} = u^{q,k,r} + v_1$.

In particular, we have:

$\delta = 0$, if $q = 1$ or more generally if the 2-level-iteration is performed,

$\delta = (\delta_{q-1})^\mu$, if μ iteration steps are performed at level $q - 1$.

Here δ_{q-1} is the convergence rate of the $(q - 1)$-level process. We notice that in this investigation convergence rates are measured by norms and not by spectral radii.

For ease of notation we will drop the superscripts q and k. Moreover, we will refer to the homogeneous equation again.

We decompose u^r in the sense of (3.1), $u^r = v^r + w^r$, $v^r \in S_{q-1}$, $w^r \in T_q$ and set $\lambda = |v^r|/\|v^r\|$, $\rho = 2\lambda^2/(\lambda^2 + 1)$. From (3.11) and Lemma 4.1 we have

$$\frac{\|G_q^{\mathrm{II}} u^r\|}{\|u^r\|} \leqslant \frac{|u^r|}{\|u^r\|} \leqslant \sqrt{\rho}\,. \tag{6.2}$$

Recalling that G_q^{I} and G_q^{II} constitute an alternate method, we conclude from Remark 5.1 that each application of half a Gauss-Seidel relaxation in the first two steps improves the approximate solution at least by the factor on the right-hand side of (6.2). Hence,

$$\|u^r\| \leqslant \rho^{r/2}\|u^0\|. \tag{6.3}$$

Let $u' := u^r - u^{q-1}$ denote the exact solution of the auxiliary problem in Step 3. From Lemma 3.1 in [3] we know that the execution of the exact coarse grid correction (with $\delta = 0$) yields an improvement by the factor in the strengthened Cauchy inequality (3.4):

$$\begin{aligned}\|u'\| &\leqslant \sqrt{\tfrac{1}{2}(1-\lambda^2)} \cdot \|u^r\| \leqslant \rho^{r/2}\sqrt{\tfrac{1}{2}(1-\lambda)^2} \cdot \|u^0\| \\ &= \rho^{r/2}\sqrt{\frac{1-\rho}{2-\rho}} \cdot \|u^0\|.\end{aligned} \tag{6.4}$$

Now we will control the perturbation in the correction step by applying the same trick as in [3]. Referring to (6.1) we may write u^{r+1} in the form

$$u^{r+1} = u' + \delta v_2$$

with some $v_2 \in S_{q-1}$, $\|v_2\| \leqslant \|u^{q-1}\|$. From the well-known characterization of closest points in subspaces of Hilbert spaces it follows that

$$\|u' + v_2\|^2 = \|u'\|^2 + \|v_2\|^2 \leqslant \|u'\|^2 + \|u^{q-1}\|^2 = \|u^r\|^2. \tag{6.5}$$

In order to estimate u^{2r+2} we use a duality argument $\|u^{2r+2}\| = \sup_{\hat{u}}\{(\hat{u}, u^{2r+2})/\|\hat{u}\|\}$. Let

$$\tilde{G} = \underbrace{G_q^{\mathrm{I}} \cdot G_q^{\mathrm{II}} \cdot G_q^{\mathrm{I}} \cdot \cdots \cdot G_q^{r+1}}_{r+1 \text{ alternating factors}}$$

and $Q = 1 - P_{S_{q-1}}$ be the projector associated to the exact coarse-grid-correction, then

$$u^{2r+2} = \tilde{G}^* u^{r+1}, \qquad u' = Qu^r,\ u^r = \tilde{G}u^0. \tag{6.6}$$

To understand this recall that $u^0 = G_q^{r+1}u^0$ and that $\tilde{G}^*$, the adjoint of $\tilde{G}$, is a product of $r + 1$ projectors such that G_q^{I} is to be applied first. Now (6.5), (6.6) and $Q^2 = Q = Q^*$ imply that

$$\begin{aligned}(\hat{u}, u^{2r+2}) &= (\hat{u}, \tilde{G}^* u^{r+1}) = \big(\tilde{G}\hat{u}, (1-\delta)u' + \delta[u' + v_2]\big) \\ &\leqslant (1-\delta)\big(\tilde{G}\hat{u}, Q^2\tilde{G}u^0\big) + \delta\|\tilde{G}\hat{u}\| \cdot \|u' + v_2\| \\ &= (1-\delta)\big(Q\tilde{G}\hat{u}, Q\tilde{G}u^0\big) + \delta\|\tilde{G}\hat{u}\| \cdot \|u^r\| \\ &\leqslant (1-\delta)\|Q\tilde{G}\hat{u}\| \cdot \|Q\tilde{G}u^0\| + \delta\|\tilde{G}\hat{u}\| \cdot \|\tilde{G}u^0\|.\end{aligned}$$

With this, the Cauchy-Schwarz inequality for 2-space yields

$$(6.7)\qquad (\hat{u}, u^{2r+2}) \leqslant \left[(1-\delta)\|Q\tilde{G}\hat{u}\|^2 + \delta\|\tilde{G}\hat{u}\|^2\right]^{1/2} \cdot \left[(1-\delta)\|Q\tilde{G}u^0\|^2 + \delta\|\tilde{G}u^0\|^2\right]^{1/2}.$$

From (6.3) and (6.4) we conclude that $(1-\delta)\|Q\tilde{G}u\|^2 + \delta\|\tilde{G}u\|^2 \leqslant \delta_q\|u\|^2$ for each $u \in S_q$ and not only for $u = u^0$, when putting

$$(6.8)\qquad \delta_q = \max_{0\leqslant\rho\leqslant 1} \rho^r\left[(1-\delta)\frac{1-\rho}{2-\rho} + \delta\right].$$

Consequently $(\hat{u}, u^{2r+2}) \leqslant \delta_q \cdot \|\hat{u}\| \cdot \|u^0\|$ and $\|u^{2r+2}\| \leqslant \delta_q\|u^0\|$. In particular, the two-grid convergence rate δ_{TG} is easily estimated by the evaluation of the maximum of $\rho^r(1-\rho)(2-\rho)^{-1}$, cf. Table 1.

TABLE 1

Convergence rates (error damping per cycle)

r	0	1	2	3
δ_{TG}	0.5	0.172	0.114	0.086
W-cycles:				
$\delta_{MG,\mu=2}$		0.187	0.120	0.089
$\delta_{MG,\mu=1,2}$		0.205	0.127	0.093
V-cycles:				
$\delta_{MG,\mu=1}$		$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$
if Ω is a square:				
δ_{TG}	0.5	0.125	0.053	0.042
ρ_{TG}	0.5	0.074	0.041	0.028

We also obtain by elementary calculations

$$(6.9)\qquad \delta_{TG} < \frac{1}{(r+1)e}.$$

A bound for the multigrid convergence rate $\delta_{MG,\mu} = \sup\{\delta_q, q > 1\}$ is obtained by solving the equation

$$(6.10)\qquad \delta_{MG,\mu} = \max_{0\leqslant\rho\leqslant 1} \rho^r \frac{1+\delta^\mu_{MG,\mu}-\rho}{2-\rho}.$$

In Table 1 the results for the iteration with a V-cycle, i.e., for $\mu = 1$, and for two cases with W-cycles are listed. In the case where in the recursive procedure $\mu = 1$ and $\mu = 2$ alternate (specifically $\mu = 1$ if $q_{\max} - q$ is even, and $\mu = 2$ otherwise) the computing effort for one cycle is roughly the same as the effort for two V-cycles. The effort for the pure W-cycle is larger by a factor $O(\log h^{-1}_{q_{\max}})$ since the number of unknowns per level grows only by a factor of 2.

In the last two rows of the table there are the error damping factors in terms of the energy norm and the spectral radius ρ for the iteration which were calculated in [16] for rectangular or square domains by the Fourier method.

In any case the application of one Gauss-Seidel relaxation sweep per cycle, i.e. the choice $r = 1$, turns out to be the most efficient. This is consistent with the results from [14]. On the other hand we cannot decide whether the V-cycle or a W-cycle is preferable (even if we exclude the pure W-cycle). If we compare "log δ_{MG}/effort per cycle" for different choices of the parameters, there are no large differences. From some extra calculations we find that the V-cycle is already substantially improved if merely on each fourth or fifth level the W-type is chosen. Nevertheless, our conservative bounds will possibly not tell us which parameters for r and μ yield the best code.

7. Concluding Remarks. We have emphasized that our proofs are no longer correct when Ω has got reentrant corners. Consider, e.g., the corner given in Figure 6, where p_0, p_3, p_4 lie on the boundary. If $u_1 = u_2 = 1$, then

$$(u_1 - \bar{u}_0)^2 + (u_2 - \bar{u}_0) = (u_1 - u_2)^2 + (u_2 - u_3)^2 + (u_4 - u_1)^2$$

in contrast to (4.3) and (4.9), where a factor $\frac{1}{2}$ enters into the inequality.

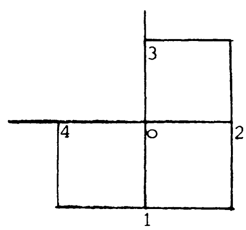

FIGURE 6
Reentrant corner

It is not surprising for an investigation in the framework of finite element theory that reentrant corners may produce *pollution* effects. On the other hand the situation is not quite the same as in the analysis of the discretization error and here the pollution effects should probably not be overestimated. Since our calculations are always done in a local way, we have to replace (4.1) by

$$\|G_h^{\mathrm{II}} u\|^2 \leqslant |u|^2 + \frac{1}{2}\|G_h^{\mathrm{II}} u\|_{\mathrm{rc}}^2,$$

where $\|\cdot\|_{\mathrm{rc}}$ refers to the energy norm for the restriction to a neighborhood of the reentrant corners. Only if the error from the viewpoint of the iteration (i.e., $u^{q,k,0} - u^q$) were concentrated in the neighborhood of the corners, the actual convergence rate could be substantially worse than the bounds determined in the preceding sections. (For a rigorous treatment, the techniques developed in [15] might be used.)

In this context a suggestion given by A. Brandt [5] is easily understood: If the computation shows large defects close to the boundaries, one should insert extra relaxation steps which operate in that region and which help to reduce the error there.

Another point seems to be worth mentioning. The result (6.9) shows an asymptotic behavior

$$\delta_{TG} \approx \text{const} \cdot r^{-1},$$

where r is the number of smoothing steps. It is well known (see e.g. [1], [9]) that the power of r depends on the difference of the indices of the involved Sobolev norms. We note, however, that the iteration with all smoothings performed before the coarse-grid-correction leads to a factor: const $\cdot\, r^{-1/2}$. The doubling of the power was gained by the duality argument.

Implicitly, we have referred to the H_1- and the H_2-norms only. This becomes apparent, when the convergence of the V-cycles is shown in the framework of the setting of Bank and Dupont [1] by applying our techniques. Let $u = \sum c_i \psi_i$ be a spectral decomposition of $u \in S_h$. Recall that $\|u\|_1^2 = |||u|||_1^2 = \sum_i \lambda_i c_i^2$ and put $|u|^2 := \sum_i \lambda_i (1 - \lambda_i/\lambda_{\max}) c_i^2$. Then one Jacobi relaxation

$$u \mapsto Ju = \sum (1 - \lambda_i/\lambda_{\max}) c_i \psi_i$$

implies a reduction of the $\|\cdot\|_1$-norm by a factor $\rho = |Ju|^2/\|Ju\|^2$. Moreover, if $r > 1$ Jacobi relaxations are performed, the error reduction in each step is at least as good as in the last step. Note that

$$\|u\|_1^2 - |u|^2 = \sum \left(\lambda_i^2/\lambda_{\max}\right) c_i^2 = \lambda_{\max}^{-1} |||u|||_2.$$

Consider the multigrid algorithm with r Jacobi relaxations before the (δ-perturbed) coarse-grid-correction and another r Jacobi relaxations after it. Then the damping factor for the $\|\cdot\|_1$-norm is estimated from above by

$$\max_{0 \leqslant \rho \leqslant 1} \rho^{2r}[\text{const}(1 - \rho)(1 - \delta) + \delta].$$

Consequently, the iteration with V-cycles has a convergence rate less than or equal to const$/(2r +$ const$) < 1$, provided that $r \geqslant$ const$/3$.

Institut für Mathematik
Ruhr-Universität Bochum
Postfach 150
4630 Bochum 1, Federal Republic of Germany

1. 1. R. E. Bank & T. Dupont, "An optimal order process for solving finite element equations," *Math. Comp.*, v. 36, 1981, pp. 35–51.

2. R. E. Bank & T. Dupont, *Analysis of a Two-Level Scheme for Solving Finite Element Equations*, Report CNA-159, Center for Numerical Analysis, Austin, 1980.

3. D. Braess, "The contraction number of a multigrid method for solving the Poisson equation," *Numer. Math.*, v. 37, 1981, pp. 387–404.

4. A. Brandt, "Multi-level adaptive solutions to boundary-value problems," *Math. Comp.*, v. 31, 1977, pp. 333–390.

5. A. Brandt, private communication.

6. A. Brandt & N. Dinar, "Multigrid solutions to elliptic flow problems," in *Numerical Methods for PDEs* (S. Parter, ed.), Academic Press, New York, 1979, pp. 53–107.

7. H. Foerster, K. Stüben & U. Trottenberg, "Non-standard multigrid techniques using checkered relaxation and intermediate grids," in *Elliptic Problem Solvers* (M. Schultz, ed.), Academic Press, New York, 1981.

8. W. Hackbusch, "On the convergence of multi-grid iterations," *Beiträge Numer. Math.*, v. 9, 1981, pp. 213–239.

9. W. Hackbusch, "Survey of convergence proofs for multi-grid iterations," in *Special Topics of Applied Mathematics* (J. Frehse, D. Pallaschke, U. Trottenberg, eds.), North-Holland, Amsterdam, 1980, pp. 152–164.

10. S. F. McCormick & J. W. Ruge, "Multigrid methods for variational problems,"*SIAM J. Numer. Anal.*, v. 19, 1982, pp. 924–929.

11. Th. Meis & H.-W. Branca, "Schnelle Lösung von Randwertaufgaben, " *ZAMM*, v.62, 1982, pp. T263–T270.

12. W. L. MIRANKER & V. YA. PAN, "Methods of aggregation," *Linear Algebra Appl.*, v. 29, 1980, pp. 231–257.

13. R. A. NICOLAIDES, "On the l^2-convergence of an algorithm for solving finite element equations," *Math. Comp.*, v. 31, 1977, pp. 892–906.

14. R. A. NICOLAIDES, "On some theoretical and practical aspects of multigrid methods," *Math. Comp.*, v. 33, 1979, pp. 933–952.

15. J. A. NITSCHE & A. H. SCHATZ, "Interior estimates for Ritz-Galerkin methods," *Math. Comp.*, v. 28, 1974, pp. 937–958.

16. M. RIES, U. TROTTENBERG & G. WINTER, "A note on MGR methods," *Linear Algebra Appl.*, v. 49, 1983, pp. 1–26, supplemented by private communication of the second author.

17. R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N. J., 1962.

18. R. VERFÜRTH, "The contraction number of a multigrid method with mesh ratio 2 for solving Poisson's equation," *Linear Algebra Appl.* (To appear.)

19. D. BRAESS & W. HACKBUSCH, "A new convergence proof for the multigrid method including the V-cycle," *SIAM J. Numer. Anal.*, v. 20, 1983, pp. 967–975.