A STUDY OF SOME MULTI-GRID IDEAS*

by

David Kamowitz
and
Seymour V. Parter

Computer Sciences Technical Report #545

June 1984

A STUDY OF SOME MULTI-GRID IDEAS*

by

David Kamowitz[1]
and
Seymour V. Parter[1]

(1) Computer Sciences Department, University of Wisconsin-Madison, Madison, WI.

*Supported by the U.S. Air Force Office of Scientific Research under Contract No. AFOSR-82-0275.

ABSTRACT

In an effort to understand certain ideas and concepts associated with multi-grid iterations we give an in-depth study of a particular simple problem. We consider a standard finite-difference system associated with the two-point boundary value problem

$$-(pu')' + bu' + qu = 0, \quad u(0) = u(1) = 0 \ .$$

The operators I_h^{2h}, I_{2h}^h are "operator" based interpolation and projection operators while the smoothers are the damped Jacobi iterations with parameter $a > 0$.

We determine the exact rates of convergence for the "two-grid" scheme and upper bounds (<1 !) for the multi-grid schemes. Experimental results are discussed.

1. Introduction

The multi-grid approach for the numerical solution of boundary value problems for elliptic partial differential equation is proving itself as one of the fastest and most efficient methods - see [1], [3], [4], [5], [6], [17]. Moreover, there are a large number of theoretical papers on this subject - see [2], [3], [6], [7], [8], [9], [10], [11], [13], [16]. Nevertheless, it seems (at least it seems so to these authors) that we are just beginning to understand this powerful idea. In particular, there are questions of: how do we choose the interpolation and projection operators?, how do we choose the smoothing operators?, what do we mean when we say smoothing? and ...?

This report is a reflection of our efforts to understand and appreciate the theoretical insights of Frederickson [7], McCormick and Ruge [13], McCormick [14], [15] and Greenbaum [8] and apply those ideas to extend the explicit convergence rates given by Hackbusch [12, (2.21)], [11] for the very simplest problem

$$u'' = f, \quad u(0) = u(1) = 0 .$$

Specifically, we consider the two-point boundary-value problem

$$Lu := -(pu')' + b(x)u' + qu = f, \quad 0 < x < 1 \tag{1.1}$$

$$u(0) = u(1) = 0 \tag{1.2}$$

where $p(x)$, $b(x)$, $q(x)$ are smooth functions and

$$p(x) \geq p_0 > 0, \quad q(x) \geq 0 . \tag{1.3}$$

In section 2 we describe a basic approach to multi-grid which is based on the ideas of Frederickson [7], McCormick and Ruge [13] and Greenbaum [8].

In section 3 we describe a discretization (finite-difference) of the problem (1.1), (1.2) and a specific two-grid iterative procedure for its solution. In section 4 we describe the class of damped Jacobi "smoothers" and use our knowledge of these schemes and their eigenvectors to describe the basic spaces: *Range* $I_{2h}^{h} = \mathcal{R}$ and *Nullspace* $I_{h}^{2h} L_{h} = \eta$. In section 5 we obtain estimates for the norm decay of a single step in a two grid scheme for two different norms. In addition we obtain a better estimate for the norm decay for all iterative steps beyond the first in both these norms. Interestingly enough, this estimate is the same in both norms. This latter result is an improvement over the estimates of Hackbusch [11], [12].

It is well-known that the problem (1.1), (1.2) is equivalent to a self-adjoint problem. Moreover, the discretization (3.4) is also equivalent to a symmetric problem. In fact, our multi-grid treatment of this problem is equivalent to the "same" multi-grid treatment of this symmetric problem. This equivalence is not needed for the discussion in sections 1-5. However, as we turn to the extension of a two-grid scheme to a true multi-grid scheme, we require this information. In section 6 we demonstrate this equivalence. In section 7 we describe the n-grid "saw-tooth" multi-grid schemes and give a theory (closely related to a theorem of McCormick [14], [15]) which describes the rates of convergence of this scheme. In addition section 7 contains some experimental results.

2. A Basic Theory

The theory presented in this section is based on the work of Frederickson [7], McCormick and Ruge [13] and Greenbaum [8]. We consider a finite-dimensional linear space S_h and a problem

$$(2.1) \qquad L_h U = f; \quad u,\ f \in S_h$$

where

$$L_h: S_h \to S_h$$

is a linear, nonsingular operator.

Multi-grid is an iterative method for the solution of this problem. The basic idea is to utilize another finite dimensional space S_{2h} with

$$(2.2) \qquad \dim S_{2h} < \dim S_h .$$

Hence we require operators I_h^{2h}, I_{2h}^h which enable us to effect communication between these spaces. In particular, we have

$$(2.3a) \qquad I_h^{2h}: S_h \to S_{2h} \qquad \text{(projection)}$$

$$(2.3b) \qquad I_{2h}^h: S_{2h} \to S_h \qquad \text{(Interpolation)}$$

where I_h^{2h}, I_{2h}^h are linear operators. We also require a "smoothing" operator $\mathcal{S}_h$ and a "coarse grid" operator L_{2h}. The smoothing operator $\mathcal{S}_h$ is an affine operator which has U, the unique solution of (2.1) as its only fixed point. That is

$$(2.4a) \qquad \mathcal{S}_h v = G_h v + \hat{f}$$

where $G_h\colon S_h \to S_h$ is a linear operator *and* if U is the solution of (2.1), then

(2.4b) $$\mathcal{S}_h U = U .$$

Finally, the "coarse grid" operator

(2.5) $$L_{2h}\colon S_{2h} \to S_{2h}$$

is a linear, nonsingular operator taking S_{2h} onto itself.

Let $U^0 \in S_h$ be a guess for the solution U of (2.1). Set

(2.6a) $$\varepsilon^0 = U - U^0 ,$$

(2.6b) $$\tilde{U} = \mathcal{S}_h U^0 ,$$

(2.6c) $$\tilde{\varepsilon} = U - \tilde{U} = G_h(U-U^0) = G_h\varepsilon^0 ,$$

(2.6d) $$r = f - L_h\tilde{U} = L_h(U-\tilde{U}) = L_h\tilde{\varepsilon} ,$$

(2.6e) $$R = I_h^{2h} r .$$

Solve

(2.7) $$L_{2h}\hat{\varepsilon} = R , \quad \text{i.e.,} \quad \hat{\varepsilon} = L_{2h}^{-1}R .$$

Set

(2.8) $$U^1 = \tilde{U} + I_{2h}^h\hat{\varepsilon} , \quad \varepsilon^1 = \tilde{\varepsilon} - I_{2h}^h\hat{\varepsilon}$$

$$U^0 := U^1$$

and return to (2.6a).

Remark: While one might say that we have merely described a two-grid scheme, the iterative scheme described above does, in fact, describe "multi-grid" schemes. The point is that the "coarse grid correction" operator L_{2h} may be a complicated procedure involving more grids.

The work of Frederickson [7], McCormick and Ruge [13] and Greenbaum [8] suggests we study two fundamental subspaces. These are

(2.9a) $$\mathcal{R} = \text{range of } I_{2h}^{h},$$

(2.9b) $$\eta = \text{nullspace of } I_h^{2h} L_h .$$

In addition we consider a special two grid "coarse grid" operator

(2.10) $$\hat{L}_{2h} = I_h^{2h} L_h I_{2h}^{h} .$$

This particular operator is the "Galerkin choice" and is "optimal" in a certain sense. This fact is emphasized by the following

Lemma 2.1: Consider one iteration as described above by (2.6a)-(2.8) with

(2.11) $$L_{2h} = \hat{L}_{2h} .$$

Suppose $\tilde{\varepsilon} \in \mathcal{R}$. That is, suppose there is a $w(2h) \in S_{2h}$ and

(2.12) $$\tilde{\varepsilon} = \tilde{\varepsilon}(h) = I_{2h}^{h} w(2h) .$$

Then

(2.13a) $$\hat{\varepsilon} = w(2h)$$

and

$$U^1 = U .$$

Hence, the problem is solved.

Proof: From (2.6d) and (2.12) we have

$$r = L_h\tilde{\varepsilon} = L_h I_{2h}^h w(2h) .$$

Thus

$$R = I_h^{2h} r = (I_h^{2h} L_h I_{2h}^h) w(2h) .$$

That is,

$$R = \hat{L}_{2h} w(2h) . \tag{2.14}$$

Hence, from (2.7) we have (2.13a). Finally, (2.13b) follows from (2.8) and (2.12). ■

Now suppose we can write S_h as the direct sum (not necessarily orthogonal) of *Range* (I_{2h}^h) and *Nullspace* $(I_h^{2h} L_h)$. That is, every grid function $w(h) \in S_h$ can be uniquely written as

$$w(h) = I_{2h}^h w^1(2h) + w^2(h) \tag{2.15a}$$

where

$$I_h^{2h} L_h w^2(h) = 0 . \tag{2.15b}$$

Let us apply this decomposition to the intermediate error $\tilde{\varepsilon} = \tilde{\varepsilon}(h)$. Then

$$\tilde{\epsilon}(h) = I_{2h}^{h} w^{1}(2h) + \epsilon^{1}(h) \ . \tag{2.16}$$

Thus

$$r = L_h I_{2h}^{h} w^{1}(2h) + L_h \epsilon^{1}(h)$$

and

$$R = I_h^{2h} r = [I_h^{2h} L_h I_{2h}^{h}] w^{1}(2h) \ .$$

Using lemma 2.1 we see that

$$U^{1} = \tilde{U} + I_{2h}^{h} w^{1}(2h) \tag{2.17}$$

and hence

$$\epsilon^{1} = U - U^{1} = U - \tilde{U} - I_{2h}^{h} w^{1}(2h) \ .$$

That is

$$\epsilon^{1} = \epsilon^{1}(h) \ .$$

Thus, the convergence of the process can be tested by

$$\| \epsilon^{1}(h) \| / \| \epsilon^{0} \|$$

where

$$\epsilon^{1}(h) \in \mathit{Nullspace}(I_h^{2h} L_h) = \eta$$

and the norm involved is any norm. The authors mentioned above all dealt with self-adjoint L_h and assume that

$$I_h^{2h} = c(I_{2h}^h)^T .$$

In this case it is an easy matter to see that the *Range* (I_{2h}^h) and *Nullspace* $(I_h^{2h}L_h)$ are L_h-orthogonal complements of S_h . In the general case we must assume the decomposition (2.15a). However, it is a simple counting argument to see that (2.15a) is valid if L_{2h} is non-singular and

i) rank $I_{2h}^h = \dim S_{2h}$

ii) dim *Nullspace* $I_h^{2h}L_h \geq \dim S_h - \dim S_{2h}$.

Recall that Lemma 2.1 implies that the zero vector is the only vector common to both subspaces.

3. The Discrete Problem

Let an integer $N > 1$ be chosen and set

$$h = \frac{1}{2(N+1)} = \frac{1}{(2N+1)+1} \tag{3.1}$$

and let the fine "grid points" be given by

$$x_j(h) = jh, \quad j = 0,1,2,\ldots,2N+2 . \tag{3.2}$$

We define a difference operator L_h by

$$[L_h U]_k = -\alpha_k U_{k-1} + \beta_k U_k - \gamma_k U_{k+1} \tag{3.3a}$$

where

$$\begin{cases} \alpha_k = \left[\dfrac{p_{k-\frac{1}{2}}}{h^2} + \dfrac{b_k}{2h}\right] , \\[2ex] \beta_k = \left[\dfrac{p_{k+\frac{1}{2}} + p_{k-\frac{1}{2}}}{h^2} + q_k\right] , \\[2ex] \gamma_k = \left[\dfrac{p_{k+\frac{1}{2}}}{h^2} - \dfrac{b_k}{2h}\right] . \end{cases} \tag{3.3b}$$

Then L_h is a consistent approximation to the operator L described in (1.1).

We assume h is so small that $\alpha_k > 0$, $\gamma_k > 0$ for all k.

We are concerned with the system of linear equations

$$\begin{cases} [L_h U]_k = f_k, \quad k = 1,2,\ldots,2N+1 , \\ U_0 = U_{2N+2} = 0 \end{cases} \tag{3.4}$$

We shall solve this system with a particular multi-grid iterative scheme. Consider a course grid on which we have a mesh spacing of $2h$ and the course grid points are given by

$$X_k(2h) = 2hk, \quad k = 0,1,2,\dots,N+1 .$$

We have a space S_h of grid functions $\{U_k(h); k = 0,1,\dots,2N+2\}$ defined on the fine grid, and we have a space S_{2h} of grid functions $\{U_k(2h); k = 0,1,\dots,N+1\}$ defined on the coarse grid. Our first step is to construct the interpolation operator I_{2h}^h which maps S_{2h} into S_h, i.e.

$$I_{2h}^h : S_{2h} \to S_h .$$

Following the experience of Dendy [4], [5] we choose the following mapping

$$\text{(3.5a)} \qquad [I_{2h}^h U(2h)]_{2k} = U_k(2h) \qquad \text{(common points)}$$

and

$$\text{(3.5b)} \qquad [I_{2h}^h U(2h)]_{2k-1} = \frac{1}{\beta_{2k-1}} [\alpha_{2k-1} U_{k-1}(2h) + \gamma_{2k-1} U_k(2h)] \qquad \text{(new points).}$$

This choice of "operator" interpolation may be described in the following way: If the physical point $x_j(h)$ of the fine grid is also a physical point of the coarse grid, i.e. if j is even, say $j = 2k$, we set $U_j(h) = U_k(2h)$ to be the same value as the coarse grid function assumed at that point; that is, (3.5a) holds. If the physical point $x_j(h)$ of the fine grid is not a point of the coarse grid, i.e. j is odd, say $j = 2k - 1$, we require that

$$\text{(3.5c)} \qquad \{L_h[I_{2h}^h U(2h)]\}_{2k-1} = 0 .$$

We formalize this remark as

Lemma 3.1: Let this interpolation operator I_{2h}^{h} be defined by (3.5a), (3.5b). Then a function $U(h) \in S_h$ is in the range of I_{2h}^{h} if and only if

$$[L_h U(h)]_{2k-1} = 0, \quad k = 1,2,\ldots,N+1 \ . \qquad \blacksquare$$

We now turn to the construction of a projection operator $I_h^{2h}: S_h \to S_{2h}$. We define

$$(3.6) \qquad [I_h^{2h} U(h)]_k = 1/2 \left[\frac{\alpha_{2k}}{\beta_{2k-1}} U_{2k-1}(h) + U_{2k}(h) + \frac{\gamma_{2k}}{\beta_{2k+1}} U_{2k+1}(h) \right] .$$

Remark: if $b(x) = 0$ and the operator L is self-adjoint then

$$\alpha_k = \gamma_{k-1}$$

and we see that

$$(3.7) \qquad I_h^{2h} = c[I_{2h}^{h}]^T .$$

As we have said in section 2, the relationship (3.7) between I_h^{2h} and I_{2h}^{h} is the "variational choice" and is frequently recommended for self-adjoint problems - see [13], [7], [4].

For the purposes of this exposition we take the optimal choice of "coarse grid correction", i.e.

$$L_{2h} := \hat{L}_{2h} = I_h^{2h} L_h I_{2h}^{h} .$$

Remark: A direct computation shows that

$$[L_{2h}U(2h)]_k = -\alpha_k^{(2)}U_{k-1} + \beta_k^{(2)}U_k - \gamma_k^{(2)}U_{k+1} \tag{3.8a}$$

where

$$\alpha_k^{(2)} = \frac{1}{2}\left[\frac{\alpha_{2k}\alpha_{2k-1}}{\beta_{2k-1}}\right] \tag{3.8b}$$

$$\beta_k^{(2)} = \frac{1}{2}\left[\beta_{2k} - \frac{\alpha_{2k}\gamma_{2k-1}}{\beta_{2k-1}} - \frac{\gamma_{2k}\alpha_{2k+1}}{\beta_{2k+1}}\right] \tag{3.8c}$$

$$\gamma_k^{(2)} = \frac{1}{2}\left[\frac{\gamma_{2k}\gamma_{2k+1}}{\beta_{2k+1}}\right] \tag{3.8d}$$

Hence, L_{2h} is again a diagonally dominant three term operator of the form (3.3a).

Now we need only describe the smoothing operator which we do in the next section.

4. Jacobi Iterative Schemes

A direct iterative method for the solution of (3.4) is described by a splitting of the operator L_h. We set

$$L_h = M - N .$$

Then, given a first guess U^0 we define successive iterates by the formula

$$MU^{j+1} = NU^j + f . \tag{4.1}$$

The convergence of this scheme is determined by the eigenvalue problem

$$\lambda MU = NU . \tag{4.2}$$

As is well known, the scheme (4.1) is convergent iff

$$\max|\lambda| < 1 .$$

It is easy to verify that: if $\langle \lambda, \phi \rangle$ are an eigenpair of (4.2) then

$$L_h\phi = (1-\lambda)M\phi . \tag{4.3}$$

In this section we are concerned with a particularly simple class of such iterative methods, the Jacobi methods. We set

$$M = (1+a)B , \quad a \geq 0 , \tag{4.4a}$$

where

$$B = \mathit{diag}(\beta_k) . \tag{4.4b}$$

In this case we may rewrite (4.1) as

$$U^{j+1} = U^j + \frac{1}{1+a} B^{-1}(f-L_h U^j) .$$

When $a = 0$ we call this scheme the Jacobi method. When $a > 0$ we call this scheme a damped Jacobi method. In these cases we are able to give a relatively complete discussion of the eigenvalue problem (4.2). We have the following facts.

i) The method is convergent for $a \geq 0$.

ii) Let $a = 0$. Let $\langle \mu,\phi \rangle$ be an eigenpair, i.e.

$$\mu B\phi = (B-L_h)\phi .$$

Let $\hat{\phi}$ be defined by

$$\hat{\phi}_k = (-1)^{k+1}\phi_k . \tag{4.6}$$

Then $\langle -\mu,\hat{\phi} \rangle$ is also an eigenpair. The eigenvalues μ are real and distinct, furthermore: as $h \to 0$ the $\{\mu\}$ fill out the interval $[-1,1]$. For completeness we repeat the basic relationship between ϕ and $\hat{\phi}$. Namely,

(4.7a) if k is odd: $\phi_k = \hat{\phi}_k$

(4.7b) if k is even: $\phi_k = -\hat{\phi}_k$.

Since $\dim S_h = 2N+1$ there is a single eigenvector $\tilde{\phi}$ associated with the eigenvalue $\mu = 0$. This eigenvector satisfies

$$\tilde{\phi}_{2k} = 0 . \tag{4.8}$$

For $a > 0$ there are corresponding eigenpairs $\langle \lambda,\phi \rangle$, $\langle \hat{\lambda},\hat{\phi} \rangle$ where ϕ and $\hat{\phi}$ are the eigenfunctions described in (2) and

$$\lambda = \frac{\mu+a}{1+a}, \quad \hat{\lambda} = \frac{a-\mu}{1+a} \tag{4.9}$$

the eigenpair $\langle 0,\tilde{\phi} \rangle$ corresponds to an eigenpair $\langle \lambda,\tilde{\phi} \rangle$ with

$$\lambda = \frac{a}{1+a} .$$

We now turn to the determination of the three important subspaces

$$\textit{Range } I^h_{2h}, \ \textit{Nullspace } I^{2h}_h, \ \textit{Nullspace } I^{2h}_h L_h .$$

where the operators I^h_{2h}, I^{2h}_h are given by (3.5a), (3.5b) and (3.6).

<u>Lemma 4.1</u>: Let $\langle \mu,\phi \rangle$, $\langle -\mu,\hat{\phi} \rangle$ be the two eigenpairs described in (ii) above with $\mu \neq 0$. Let

$$\Phi = [(1+\mu)\phi - (1-\mu)\hat{\phi}] \tag{4.10}$$

then $\Phi \in \textit{Range } I^h_{2h}$. Further, since the vectors Φ corresponding to different pairs $\langle \mu,\phi \rangle$, $\langle -\mu,\hat{\phi} \rangle$ are linearly independent this construction provides N linearly independent elements of the *Range* I^h_{2h}.

<u>Proof</u>: Using (4.3) we have

$$L_h\Phi = [(1+\mu)(1-\mu)B\phi - (1-\mu)(1+\mu)B\hat{\phi}]$$

$$= (1+\mu)(1-\mu)B[\phi-\hat{\phi}] .$$

Thus, the lemma follows from Lemma 3.1 which characterizes *Range* (I^h_{2h}) and from (4.7a), (4.7b). ∎

<u>Lemma 4.2</u>:

i) For all choices of $a \geq 0$, the vector $B\tilde{\phi}$ (associated with $\mu=0$) is an element of *Nullspace* I_h^{2h} and hence (4.3) implies that $\tilde{\phi}$ is an element of *Nullspace* $I_h^{2h}L_h$.

ii) Let $\langle \mu,\phi \rangle$, $\langle -\mu,\hat{\phi} \rangle$ with $\mu \neq 0$ be the two eigenpairs described above. Let

$$\Psi = B[(1-\mu)\phi + (1+\mu)\hat{\phi}] \ . \tag{4.11a}$$

Then ψ is an element of *Nullspace* of I_h^{2h} and (4.3) implies that

$$(\phi+\hat{\phi}) \in \textit{Nullspace} \text{ of } I_h^{2h}L_h \ . \tag{4.11b}$$

Further, since the vectors Ψ associated with different pairs $\langle \mu,\phi \rangle$, $\langle -\mu,\hat{\phi} \rangle$ are linearly independent, we have N linearly independent elements of this nullspace and N linearly independent vectors of *Nullspace* $I_h^{2h}L_h$. The vectors $B\tilde{\phi}$, $\tilde{\phi}$ provide one more independent vector of each of these subspaces.

<u>Proof</u>: The result follows from a direct computation using (3.6) and the defining eigenvalue problem. ∎

<u>Corollary</u>: There is a unique decomposition of S_h into *Range* I_{2h}^h and *Nullspace* $(I_h^{2h}L_h.)$

<u>Proof</u>: Since L_h is non-singular we have shown that

$$\textit{dim Range } I_{2h}^h \geq N \ ,$$

$$\textit{dim Nullspace } (I_h^{2h}L_h) = \textit{dim Nullspace } (I_h^{2h}) \geq N + 1 \ ,$$

$$\textit{dim } S_h = 2N + 1 \ .$$

Thus the corollary follows from the observation that Lemma 3.1 implies that these two subspaces have only the zero vector in common. ∎

5. Some Estimates

Let $a \geq 0$. We take as our smoother m applications of the corresponding Jacobi iteration. That is, given $U^0 = U^0(h)$ we obtain $\tilde{U}$ (as in 2.6b) from the formula

$$MU^{j+1} = NU^j + f\ , \quad j = 0,1,\ldots,m-1\ , \tag{5.1a}$$

$$\tilde{U} = U^m\ . \tag{5.1b}$$

First let us consider the special vector $\tilde{\phi}$ with its associated eigenvalue

$$\lambda = \frac{a}{1+a}\ .$$

Suppose

$$\epsilon^0 = U - U^0 = C\tilde{\phi}$$

then

$$\tilde{\epsilon} = C\lambda^m\tilde{\phi}$$

$$L_h\tilde{\epsilon} = C\lambda^m B\tilde{\phi}$$

and, using Lemma 4.2 we see that

$$I_h^{2h} L_h \tilde{\epsilon} = 0\ .$$

Hence, in this case, for any norm

(5.2)
$$\frac{\|\epsilon^1\|}{\|\epsilon^0\|} = \left(\frac{a}{1+a}\right)^m .$$

We now consider two norms defined on S_h. Let $w, v \in S_h$ be of the form

(5.3a)
$$w = \sum_{j=1}^{N} A_j\phi_j + \tilde{A}\tilde{\phi} + \sum_{j=1}^{N} \hat{A}_j\hat{\phi}_j$$

(5.3b)
$$v = \sum_{j=1}^{N} C_j\phi_j + \tilde{C}\tilde{\phi} + \sum_{j=1}^{N} \hat{C}_j\hat{\phi}_j .$$

Define

$$\langle w,v \rangle_0 = \sum_{j=1}^{N} A_jC_j + \tilde{A}\tilde{C} + \sum_{j=1}^{N} \hat{A}_j\hat{C}_j$$

$$\langle w,v \rangle_1 = \sum_{j=1}^{N} A_jC_j(1-\mu_j) + \tilde{A}\tilde{C} + \sum_{j=1}^{\tilde{}} \hat{A}_j\hat{C}_j(1+\mu_j)$$

$$\|w\|_0^2 = \langle w,w \rangle_0$$

$$\|w\|_1^2 = \langle w,w \rangle_1 = \sum_{j=1}^{N} A_j^2(1-\mu_j) + \tilde{A}^2 + \sum_{j=1}^{N} \hat{A}_j^2(1+\mu_j) .$$

<u>Lemma 5.1</u>: Suppose

(5.4)
$$\epsilon^0 = U - U^0 = c\phi + d\hat{\phi}$$

and let ϵ^1 be defined by the two-grid iteration scheme. Then

(5.5a) $$\max_{c,d} \frac{\|\epsilon^1\|_0^2}{\|\epsilon^0\|_0^2} = 1/2\left[\left(\frac{\mu+a}{1+a}\right)^{2m}(1-\mu)^2 + \left(\frac{a-\mu}{1+a}\right)^{2m}(1+\mu)^2\right]$$

(5.5b) $$\max_{c,d} \frac{\|\epsilon^1\|_1^2}{\|\epsilon^0\|_1^2} = 1/2\left[\left(\frac{\mu+a}{1+a}\right)^{2m}(1-\mu) + \left(\frac{a-\mu}{1+a}\right)^{2m}(1+\mu)\right] .$$

<u>Proof</u>: From (5.4) we see that

$$\tilde{\epsilon} = c\lambda^m\phi + d\hat{\lambda}^m\hat{\phi} .$$

Following the theory of section 2, we write (see 2.16)

(5.6) $$\tilde{\epsilon} = I_{2h}^h w^1(2h) + \epsilon^1(h) ,$$

where

$$I_h^{2h} L_h \epsilon^1(h) = 0 .$$

We claim that

(5.7a) $$I_{2h}^h w^1(2h) = \frac{c\lambda^m - d\hat{\lambda}^m}{2} [(1+\mu)\phi - (1-\mu)\hat{\phi}]$$

(5.7b) $$\epsilon^1 = \frac{c\lambda^m(1-\mu) + d\hat{\lambda}^m(1+\mu)}{2} [\phi + \hat{\phi}] .$$

To verify this we need merely verify that the sum of the right-hand-sides of (5.7a) and (5.7b) is $\tilde{\epsilon}$, and use (4.10) and (4.11b).

Having verified (5.6), (5.7a), (5.7b) we proceed as follows

(5.8a) $$\| \epsilon^1 \|_0^2 = 1/2[c\lambda^m(1-\mu) + d\hat{\lambda}^m(1+\mu)]^2$$

(5.8b) $$\| \epsilon^0 \|_0^2 = c^2 + d^2$$

(5.9a) $$\| \epsilon^1 \|_1^2 = 1/2[c\lambda^m(1-\mu) + d\hat{\lambda}^m(1+\mu)]^2$$

(5.9b) $$\| \epsilon^0 \|_1^2 = c^2(1-\mu) + d^2(1+\mu)$$

Thus, a simple argument shows that

(5.10a) $$\sup \frac{\| \epsilon^1 \|_0^2}{\| \epsilon^0 \|_0^2} = 1/2[\lambda^{2m}(1-\mu)^2 + \hat{\lambda}^{2m}(1+\mu)^2] ,$$

(5.10b) $$\sup \frac{\| \epsilon^1 \|_1^2}{\| \epsilon^0 \|_1^2} = 1/2[\lambda^{2m}(1-\mu) + \hat{\lambda}^{2m}(1+\mu)] .$$

Using the basic formulae (4.9) we obtain

$$\sup \frac{\| \epsilon^1 \|_0^2}{\| \epsilon^0 \|_0^2} = 1/2\left[\left(\frac{\mu+a}{1+a}\right)^{2m}(1-\mu)^2 + \left(\frac{a-\mu}{1+a}\right)^{2m}(1+\mu)^2\right]$$

$$\sup \frac{\| \epsilon^1 \|_1^2}{\| \epsilon^0 \|_1^2} = 1/2\left[\left(\frac{\mu+a}{1+a}\right)^{2m}(1-\mu) + \left(\frac{a-\mu}{1+a}\right)^{2m}(1+\mu)\right]$$

Thus, the lemma is proven. ∎

Theorem 5.1: In the general case we have

$$(5.11a)\qquad \left(\frac{\|\epsilon^1\|_0}{\|\epsilon^0\|_0}\right)^2 \le \sup_{-1\le\mu\le1}\left\{1/2\left[\left(\frac{\mu+a}{1+a}\right)^{2m}(1-\mu)^2+\left(\frac{a-\mu}{1+a}\right)^{2m}(1+\mu)^2\right]\right\}$$

$$(5.11b)\qquad \left(\frac{\|\epsilon^1\|_1}{\|\epsilon^0\|_1}\right)^2 \le \sup_{-1\le\mu\le1}\left\{1/2\left[\left(\frac{\mu+a}{1+a}\right)^{2m}(1-\mu)+\left(\frac{a-\mu}{1+a}\right)^{2m}(1+\mu)\right]\right\}$$

Proof: The Theorem follows immediately from the previous lemma. ∎

We observe that (5.11a) is precisely the formula obtained by Hackbusch [12, (2.21)] in the special case $p(x) = 1$, $b(x) = q(x) = 0$ and $a = 1$. To make a complete identification we merely set

$$(5.12)\qquad \sigma = \frac{1-\mu}{2}\,,\quad (1-\sigma) = \frac{1+\mu}{2}\,.$$

However, while (5.11a), (5.11b) describe the worst case decay in one multigrid iteration in a 2-grid scheme, it does not give the estimate of real interest. From the discussion in the proof of Lemma 5.1 - and (5.7) in particular - we realize that, even though the constants c and d of (5.4) may be arbitrary for ε^0, that is not true for ε^k, $k \ge 1$. We have (from (5.7b))

$$(5.13a)\qquad \varepsilon^1 = \sigma[\phi+\hat{\phi}]$$

with

$$\sigma = \frac{c\lambda^m(1-\mu) + d\hat{\lambda}^m(1+\mu)}{2} .$$

Therefore, following the argument of Lemma 5.1,

$$\varepsilon^{(2)} = \frac{\sigma}{2}[\lambda^m(1-\mu) + \hat{\lambda}^m(1+\mu)][\phi+\hat{\phi}] .$$

Hence

$$\| \varepsilon^{(1)} \|_0^2 = \| \varepsilon^{(1)} \|_1^2 = 2\sigma^2$$

$$\| \varepsilon^{(2)} \|_0^2 = \| \varepsilon^{(2)} \|_1^2 = \frac{\sigma^2}{2}[\lambda^m(1-\mu) + \hat{\lambda}^m(1+\mu)]^2$$

and for $j = 0,1$ we have

(5.14) $$\frac{\| \varepsilon^{(2)} \|_j^2}{\| \varepsilon^{(1)} \|_j^2} = \frac{1}{4}[\lambda^m(1-\mu) + \hat{\lambda}^m(1+\mu)]^2 .$$

Thus we have proven

Theorem 5.2: In the general case, for $j = 0,1$ and all $k \geq 1$ we have

(5.15) $$\frac{\| \varepsilon^{(k+1)} \|_j^2}{\| \varepsilon^{(k)} \|_j^2} \leq \sup_{-1\leq\mu\leq 1} \left\{\frac{1}{4}[\lambda^m(1-\mu) + \hat{\lambda}^m(1+\mu)]^2\right\} .$$

Remark: The distinction between Hackbusch's result (5.11a) and (5.15) is non-trivial for large m. We have, as $m \to \infty$

(5.16a) $$\frac{\| e^1 \|_0}{\| e^0 \|_0} \sim \frac{(1+a)}{\sqrt{2}} \; \frac{1}{em}$$ (5.11a) while

$$\text{(5.16b)} \qquad \frac{\| e^{(k+1)} \|_j}{\| e^k \|_j} \sim \frac{1+a}{2} \frac{1}{em} \qquad \text{(5.14)}$$

Thus for $k \geq 1$ we have

$$\text{(5.17)} \qquad \frac{(\| e^{k+1} \| / \| e^k \|)}{(\| e^1 \|_0 / \| e^0 \|_0)} \sim \frac{\sqrt{2}}{2} = .7071$$

6. Symmetrization

Consider the difference equation (3.4) described by (3.3a), (3.3b). Let

(6.1) $$U_k = d_k V_k$$

where the coefficients d_k are computed recursively by

(6.2a) $$d_0 = 1$$

(6.2b) $$d_{k+1} = d_k \sqrt{\frac{\alpha_{k+1}}{\gamma_k}}\ , \quad k=0,1,\ldots$$

then (3.4) becomes

$$-\alpha_k d_{k-1} V_{k-1} + \beta_k d_k V_k - \gamma_k d_{k+1} V_{k+1} = f_k$$

or

(6.3) $$-\alpha_k \left(\frac{d_{k-1}}{d_k}\right) V_{k-1} + \beta_k V_k - \gamma_k \left(\frac{d_{k+1}}{d_k}\right) V_{k+1} = \frac{1}{d_k} f_k\ ,$$

that is

(6.3a) $$\hat{L}_h V = -\hat{\alpha}_k V_{k-1} + \beta_k V_k - \hat{\gamma}_k V_{k+1} = \hat{f}_k\ ,$$

where

$$(6.4a)\qquad \hat{\alpha}_k = \alpha_k \frac{d_{k-1}}{d_k} = \alpha_k \sqrt{\frac{\gamma_{k-1}}{\alpha_k}} = \sqrt{\alpha_k \gamma_{k-1}} ,$$

$$(6.4b)\qquad \hat{\gamma}_k = \gamma_k \frac{d_{k+1}}{d_k} = \gamma_k \sqrt{\frac{\alpha_{k+1}}{\gamma_k}} = \sqrt{\gamma_k \alpha_{k+1}} .$$

Hence

$$(6.5)\qquad \hat{\alpha}_k = \hat{\gamma}_{k-1}$$

and $\hat{L}_h$ is given by a symmetric operator.

We now turn our attention to the Jacobi iterative schemes of section 4. We have

$$(6.6)\qquad U^{j+1} = U^j + \frac{1}{1+a} B^{-1}(f-L_h U^j) .$$

The change of variables (6.1) is conveniently described by

$$(6.7a)\qquad U = DV$$

where

$$(6.7b)\qquad D = \mathrm{diag}(d_1,\ldots,d_{2N+1}) .$$

With this change of variables the iteration (6.6) becomes

$$V^{j+1} = V^j + \frac{1}{1+a} D^{-1} B^{-1} (f-L_h D V^j) .$$

But, since D and B are both diagonal matrices, we have

$$(6.8a)\qquad V^{j+1} = V^j + \frac{1}{1+a} B^{-1}(D^{-1}f-\hat{L}_h V^j)$$

where

$$D^{-1}f = (\hat{f}_1, \hat{f}_2, \ldots, \hat{f}_{2N+1})^T \; . \tag{6.8b}$$

Since B is also the diagonal of $\hat{L}_h$, (6.8) is precisely the same Jacobi iterative scheme for modified symmetric equations.

Now let us study the effect of this transformation on the two-grid iterative scheme. We compute

$$D_k^{-1}[I_{2h}^h U]_k \; .$$

Imagine $U(2h)_k = d_{2k}V(2h)_k$ is given in S_{2h}. Then, from (3.5a)

$$d_{2k}^{-1}[I_{2h}^h U(2h)]_{2k} = V(2h)_k \; , \tag{6.9a}$$

$$d_{2k-2}^{-1}[I_{2h}^h U(2h)]_{2k-2} = V(2h)_{k-1} \; . \tag{6.9b}$$

Thus (3.5b) yields

$$d_{2k-1}^{-1}[I_{2h}^h U(2h)]_{2k-1} = \frac{1}{\beta_{2k-1}}[\alpha_{2k-1}d_{2k-1}^{-1}d_{2k-2}V_{k-1}^{(2h)} + \gamma_{2k-1}d_{2k-1}^{-1}d_{2k}V_k^{(2h)}]$$

$$= \frac{1}{\beta_{2k-1}}[\hat{\alpha}_{2k-1}V(2h)_{k-1} + \hat{\gamma}_{2k-1}V(2h)_k] \; .$$

Thus, with this change of variables the mapping I_{2h}^h of our original unsymmetric problem becomes $\hat{I}_{2h}^h$, the appropriate mapping associated with the new symmetric problem.

Finally, let us consider

$$d_{2k}^{-1}[I_h^{2h}U(h)]_k \ .$$

A straight forward calculation verifies that

$$d_{2k}^{-1}[I_h^{2h}U]_k = \frac{1}{2}[\frac{\hat{\alpha}_{2h}}{\beta_{2k-1}} V_{2k-1}(h) + V_{2k(h)} + \frac{\hat{\gamma}_{2h}}{\beta_{2k+1}} V_{2k+1}]$$

Thus, following the remarks of section 3 [see (3.7)] we see that

$$[d_{2k}^{-1}I_h^{2h}] \sim \hat{I}_h^{2h} \ ,$$

the appropriate projection operator.

For our purposes, the major significance of these calculations is that the "1" norm introduced in section 5 is the "operator norm" for the symmetric problem. Hence, we have a norm which is well-defined on all spaces S_{h_j} .

7. Multi-Grid and Experimental Results

The results of the previous sections, and Theorem 5.1 in particular, provide exact estimates of the decay of the error (in two norms) in one iteration of a 2-grid scheme - in the worst case.

Since L_{2h} is again a three term (diagonally dominant) operator of the form (3.3a) - and given specifically by (3.8) - we may apply our multi-grid approach inductively as follows: Assume that the n-grid multi-grid scheme based on "smoothing" with m applications of the damped Jacobi iteration with parameter a is defined. Suppose

(7.1a) $$h = \frac{H}{2^n}, \quad n \geq 2,$$

where H is of the form

(7.1b) $$H = \frac{1}{p+1}, \quad p \geq 1.$$

We wish to solve (3.4) on the h-grid. The iterative scheme is given by the following inductive description.

(1.) On the h-grid $(h = 2^{-n}H)$:

(a) Let U^0 be chosen.

(b) Apply the damped Jacobi (with parameter a) iteration m times to obtain $\tilde{U}$.

(c) Form $r(h) = f - L_h\tilde{U}$.

(2.) Transfer Information:

(a) Set $r(2h) = I_h^{2h} r(h)$

(3.) On the 2h grid:

(a) Consider the problem

$$\hat{L}_{2h}\tilde{U}(2h) = r(2h).$$

(b) if $2h = H$, solve exactly.

(c) if $2h < H$, set $\tilde{U}^0(2h) = 0$ and apply the n-grid iterative scheme (based on smoothing with m applications of the damped Jacobi iteration with parameter a). Let $U^1(2h)$ be the result of this step.

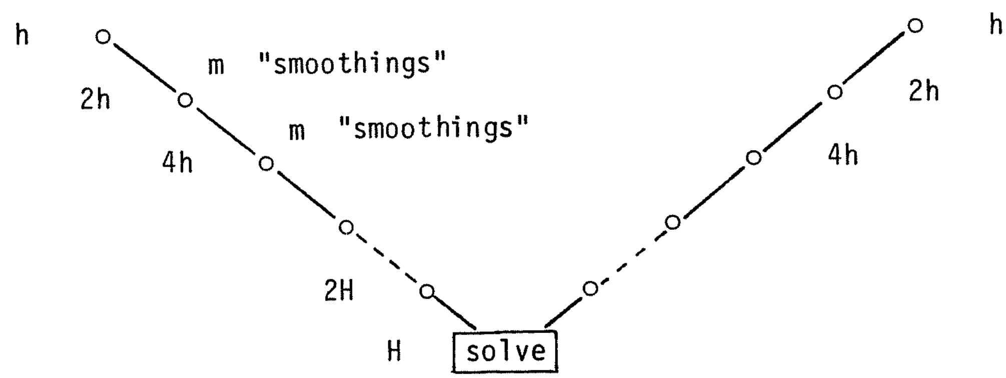
(4.) Transfer Information:

(a) $U^1 = \tilde{U} + I_{2h}^{h} U^1(2h)$.

(b) $U^1 \rightarrow U^0$.

Return to 1(b).

In the multi-grid jargon this is the so-called slash or sawtooth cycle which we indicate schematically as:



Note: There are no smoothing steps during the transfers from coarse to fine grids.

McCormick [14], [15], calls such a multigrid cycle a $M_{\backslash h}$ cycle. When the smoothing occurs only on the way "up" the cycle and the errors are merely restricted on the fine to coarse transfers, he calls the cycle a $M_{/h}$

cycle. For the symmetric case, using Richardson's iteration, see [14], he shows that

(7.2) $$\|M_{/h}\|_1 = \|M_{\backslash h}\|_1 .$$

In discussing the symmetric $M_{/h}$ cycle he obtains the following estimate. Let

(7.3a) $$\varepsilon^0 = \eta^0 + I_{2h}^h \omega^0 .$$

Suppose α, $0 < \alpha < 1$ satisfies

(7.3b) $$\|G\varepsilon^0\|_1^2 \leq \alpha\|\eta^0\|_1^2 + \|I_{2h}^h\omega^0\|_1^2 , \quad \forall\varepsilon^0 ,$$

then

(7.4a) $$\|M_{/h}\|_1 \leq \alpha^{\frac{1}{2}} ,$$

that is:

(7.4b) $$\|\varepsilon^1\|_1 = \|u-u^1\|_1 \leq \alpha^{\frac{1}{2}} \|\varepsilon^0\|_1 .$$

Since our Jacobi iteration is not all that different than Richardson's iteration it is not surprising that a similar result holds in our case. Indeed, if one applies his argument to our multigrid cycle, i.e. the $M_{\backslash h}$ cycle, one gets the following result.

<u>Lemma 7.1</u>: Consider the symmetric case and suppose

(7.5) $$\|G\|_1 \leq 1 .$$

Let

(7.6a) $$G\varepsilon^0 = \tilde{\varepsilon} = \eta + I_{2h}^h\omega .$$

Suppose $\hat{\alpha}$, $0 < \hat{\alpha} < 1$ satisfies

(7.6b) $$\| \eta \|_1^2 + \hat{\alpha} \| I_{2h}^h \omega \|_1^2 \leq \hat{\alpha} \| \varepsilon^0 \|_1^2 .$$

Then

(7.7) $$\| M_{\backslash h} \|_1 \leq \hat{\alpha}^{\frac{1}{2}} .$$

Proof: The proof proceeds by induction. Since we use an exact solver on the coarsest grid,

(7.8a) $$\| M_{\backslash h} \|_1 = 0 < \hat{\alpha}^{\frac{1}{2}} .$$

Assume that

(7.8b) $$\| M_{\backslash 2h} \|_1 \leq \hat{\alpha}^{\frac{1}{2}} ,$$

that is:

(7.8c) $$\| M_{\backslash 2h} \omega(2h) - \omega(2h) \|_1 \leq \hat{\alpha}^{\frac{1}{2}} \| \omega(2h) \|_1 .$$

Then

(7.9) $$\| \varepsilon^1 \|_1^2 = \| \eta \|_1^2 + \| I_{2h}^h (M_{\backslash 2h} \omega - \omega) \|_1^2 .$$

Because this is a symmetric problem we know that

(7.10a) $$\| v \|_1^2 = \langle v, Lv \rangle ,$$

and that

(7.10b) $$\| I_{2h}^{h} v \|_1^2 = \langle I_{2h}^{h} v, L_h I_{2h}^{h} v \rangle = \langle v, \gamma I_h^{2h} L_h I_{2h}^{h} v \rangle$$

$$= \gamma \langle v, \hat{L}_{2h} v \rangle = \gamma \| v \|_1^2 .$$

Therefore,

$$\| \varepsilon^1 \|_1^2 = \| \eta \|_1^2 + \gamma \| M_{\backslash 2h} \omega - \omega \|_1^2 .$$

By the inductive hypotheses (7.8b) we have

(7.11) $$\| \varepsilon^1 \|_1^2 \leq \| \eta \|_1^2 + \gamma \hat{\alpha} \| \omega \|_1^2 = \| \eta \|_1^2 + \hat{\alpha} \| I_{2h}^{h} \omega \|_1^2 .$$

Note: In (7.10b) and in this calculation the symbols $\| \omega \|_1$ and $\| I_{2h}^{h} \omega \|_1$ refer to the designated norms on the spaces S_{2h} and S_h .

By the basic inequality (7.6b) we have

$$\| \varepsilon^1 \|_1^2 \leq \| \eta \|_1^2 + \hat{\alpha} \| I_{2h}^{h} \omega \|_1^2 \leq \hat{\alpha} \| \varepsilon^0 \|_1^2$$

which proves the Lemma.

Since the proof of this lemma is immediate once one understands the proof of McCormick's lemma 2.2 of [15] one would expect that

(7.12) $$\alpha = \hat{\alpha} .$$

Indeed, this is the case. Direct but messy calculations based on the results of section 5 yield

(7.13) $$\alpha = \hat{\alpha} = \sup_{-1 \leq \mu \leq 1} \left\{ \frac{\frac{1}{2}\left(\frac{\mu-a}{1+a}\right)^{2m}(1+\mu) + \frac{1}{2}\left(\frac{\mu+a}{1+a}\right)^{2m}(1-\mu) - \left(\frac{\mu+a}{1+a}\right)^{2m}\left(\frac{\mu-a}{1+a}\right)^{2m}}{1 - \frac{1}{2}\left(\frac{\mu-a}{1+a}\right)^{2m}(1-\mu) - \frac{1}{2}\left(\frac{\mu+a}{1+a}\right)^{2m}(1+\mu)} \right\}$$

Moreover, for all choices of a and m, the supremum is attained at $\mu = 1$. The corresponding values of $\alpha^{\frac{1}{2}}$ are displayed in the following table:

Bounds on the Convergence Rate

m\a	.333	.5	.667	.75	1	1.333
1	.633	.577	.561	.561	.577	.614
2	.435	.408	.417	.424	.447	.475
3	.336	.335	.349	.357	.378	.403
4	.283	.293	.307	.314	.333	.357

In view of the results of section 6 which demonstrate the complete equivalence of our problem to a related symmetric problem, these upper bounds apply in our case.

However the estimate $\alpha^{\frac{1}{2}} = \hat{\alpha}^{\frac{1}{2}}$ is only an upper bound for the rate of convergence of the multigrid iterative scheme. In order to complete our investigation we have undertaken an experimental project.

A computer program was written with the following capabilities: The user supplies

$$p(x),\ b(x),\ q(x),\ f(x),\ m,\ a,\ n, \text{ and}$$

$$M = \left(\frac{1}{h}\right) - 1 = (\text{number of points on the finest grid}),$$

where $p(x)$, $b(x)$, $q(x)$, $f(x)$ are the coefficients of the problems (1.1), (1.2) and

m = number of applications of the damped Jacobi iteration,

a = parameter of the damped Jacobi iteration

n = # of grid levels.

The user also supplies an initial guess U^0 and a tolerance E.

The program then executes multi-grid iterations until the ℓ_1 norm [see (7.14a)] of the residual is below the indicated tolerance E. The program is run in an interactive fashion which allows the user to change the parameters M, m, a and n.

The experiments reported here were run on the VAX 780 in both single and double precision arithmetic (approximately sixteen decimal digits of accuracy). The single precision results were qualitatively similar to the double precision results, however, for increased accuracy, the double precision results are reported here.

For our present purposes the basic program was modified to enable us to estimate the "rates of convergence" of the multi-grid iteration. For each test problem we used a known solution $u(x)$ of the boundary value problem (1.1), (1.2). Then we computed the exact solution $u(x,h)$ of the algebraic system (2.4). Then using two norms

$$\|u\|_{\ell_1} = h\sum |u_j| \tag{7.14a}$$

$$\|u\|_1 = \langle u, L_h u \rangle^{\frac{1}{2}} \tag{7.14b}$$

we computed the norms of the error $(u-u^i)$ at each iteration. The rate of convergence was measured by computing

$$\frac{\|\varepsilon^i\|}{\|\varepsilon^{i-1}\|} = \rho^i \tag{7.15}$$

at each iteration $i = 1,2,3....$

To check that the program was working correctly a number of measures were taken. The most simplistic was to carry out some of the iterations by hand and to compare the hand computed calculations to the iterates generated by the machine. In addition, since the discretization error is $O(h^2)$, it is not unreasonable to expect that halving the step size should reduce the final error in u by a factor of four. This property was checked and found to be true. One of the requirements for $I_{2^{\ell}h}^{2^{\ell+1}h}$ is that

$$\left[L_{2^{\ell}h} I_{2^{\ell+1}h}^{2^{\ell}h} u\right]_{2k-1} = 0 \tag{7.16}$$

(by lemma 2.1).

After each coarse to fine grid transfer, formula (7.16) was computed and checked. Finally, from (5.7b) one sees that the error, $\varepsilon_{2k}(h)$, on the even points of the coarsest grid should be zero. This requirement was also verified after each iteration.

The test problems are best described by giving the choices of $p(x)$, $b(x)$, $q(x)$ and $u(x)$, the true solution of the differential equation (1.1), (1.2) (which determines $f(x)$).

As a basic case we took

$$p(x) = 1, \quad q(x) = b(x) = 0 \quad \text{and} \quad u(x) = 0\ . \tag{7.17a}$$

This test was merely to be sure the program worked on this simple case. In addition there were six other problems based on two additional sets of coefficients $p(x)$, $b(x)$, $q(x)$ and three "solutions" $u(x)$. These are

$$p(x) = 1 + \frac{1}{2}\sin 4\pi x, \quad b(x) = 1 + x, \quad q(x) = (\sin 5\pi x)^2 \tag{7.17b}$$

$$p(x) = e^x, \quad b(x) = 1 + x^2, \quad q(x) = (1-x)e^{x/2}\ . \tag{7.17c}$$

The "solutions" were

(7.18a) $$u_1(x) = x(e-e^x) ,$$

(7.18b) $$u_2(x) = x^{5/2}(1-x) ,$$

(7.18c) $$u_3(x) = \sin(14\pi x) .$$

For each problem, test runs were made with a variety of initial guesses. After all, the point was to obtain the worst rate of convergence. Each initial guess consisted of a smooth component

$$u_k^s = 20 \sin \frac{k\pi}{M+1}$$ where M is the number of points on the finest grid

and a rough component. The rough component was chosen in various ways in order to have different compositions on the coarser grids. The rough components of the initial guesses are best described schematically, by setting

$$u_k = u_k^s + 40\delta_k$$

where $|\delta_k| = 1$, and the sign of δ_k follows the following patterns:

	Initial Guess	Pattern for δ_k
	A	+ - + - + - + - + - + - + - + .
	B	+ + - - + + - - + + - - + + - .
(7.19)	C	+ + + - - - + + + - - - + + + .
	D	+ + + + - - - - + + + + - - - .
	E	+ - - + + + - - - - + + + + + .

Runs were made with a, the damped Jacobi parameter, equal to .333, .5, .667, .75, 1.0, 1.333, while m, the number of smoothing iterations, ran from one to four and the number of grid layers varied from two to five. For each test problem, the program stopped when the discrete ℓ_1 norm of the residual vector was less than .00005. The most recently computed rate of convergence.

$$\frac{\| \varepsilon_{final} \|_{\ell_1}}{\| \varepsilon_{final-1} \|_{\ell_1}}$$

was computed and recorded in Tables III-VI.

The theoretical rate for a two grid iteration scheme was computed from Theorem 5.1 and Theorem 5.2 by solving for the maximum of

$$F(\mu) = \frac{1}{2}\left[\left(\frac{\mu+a}{1+a}\right)^{2m}(1-\mu) + \left(\frac{a-\mu}{1+a}\right)^{2m}(1+\mu)\right], \quad -1 \le \mu \le 1$$

and

$$F_1(\mu) = \frac{1}{4}\left[\left(\frac{\mu+a}{1+a}\right)^{m}(1-\mu) + \left(\frac{a-\mu}{1+a}\right)^{m}(1+\mu)\right]^2, \quad -1 \le \mu \le 1$$

using Newton's method. Table I exhibits $(\max F(\mu))^{1/2}$ (a predicted rate of convergence) as a function of m and a. The value of μ at which the maximum of $F(\mu)$ occurred can be found in Table I'

Table I

Predicted Rate Based on $F(\mu)$

m\a	.333	.500	.667	.750	1.000	1.333
1	.500	.333	.400	.429	.500	.571
2	.260	.248	.261	.268	.289	.331
3	.200	.206	.217	.223	.238	.258
4	.171	.180	.190	.195	.208	.225

Table I'

Damped Jacobi Parameter-μ

m\a	.333	.500	.667	.750	1.000	1.333
1	1	0	0	0	0	0
2	.883	.707	.666	.650	.577	.370
3	.833	.786	.762	.750	.714	.661
4	.857	.833	.814	.811	.777	.741

Table II exhibits $(\max F_1(\mu))^{\frac{1}{2}}$ (a better rate of convergence) as a function of m and a. The value of μ at which the maximum of $F_1(\mu)$ occurred can be found in Table II'.

Table II

Predicted Rate Based on $F_1(\mu)$

m\a	.333	.500	.667	.750	1.000	1.333
1	.500	.333	.400	.429	.500	.572
2	.250	.111	.160	.184	.250	.326
3	.125	.078	.088	.093	.125	.187
4	.068	.062	.068	.072	.083	.109

Table II'

Damped Jacobi Parameter-μ

m\a	.333	.500	.667	.750	1.000	1.333
1	1	0	0	0	0	0
2	1	0	0	0	0	0
3	1	.612	.577	.530	0	0
4	.883	.707	.667	.650	.577	.370

Tables III through VI contain the worst rate of convergence found experimentally as measured in the ℓ_1 norm.

Table III

Worst case, 2-grids

m\a	.333	.5	.667	.75	1.0	1.333
1	.499[a]	.333[b]	.400[b]	.429[b]	.500[b]	.571[b]
2	.250[c]	.111[b]	.160[b]	.184[b]	.250[b]	.327[b]
3	.124[c]	.075[b]	.087[d]	.093[e]	.125[b]	.187[b]
4	.063[a]	.062[f]	.068[g]	.071[h]	.082[e]	.107[i]

Table IV

Worst case, 3-grids

m\a	.333	.5	.667	.75	1.0	1.333
1	.499[c]	.333[b]	.400[b]	.429[b]	.500[b]	.571[b]
2	.250[c]	.165[g]	.192[g]	.210[g]	.267[g]	.337[g]
3	.124[c]	.099[k]	.116[k]	.124[k]	.170[g]	.210[g]
4	.087[j]	.079[j]	.087[j]	.092[j]	.130[k]	.141[g]

Table V

Worst case, 4-grids

m\a	.333	.5	.667	.75	1.0	1.333
1	.499[c]	.333[b]	.400[b]	.429[b]	.500[b]	.571[b]
2	.250[c]	.190[m]	.201[ℓ]	.212[j]	.267[q]	.337[q]
3	.124[c]	.121[m]	.136[j]	.145[j]	.170[j]	.214[j]
4	.095[i]	.095[i]	.104[i]	.111[i]	.130[i]	.157[j]

Table VI

Worst case, 5-grids

m\a	.333	.5	.667	.75	1.0	1.333
1	.499(c)	.333(b)	.400(b)	.429(b)	.500(b)	.571(b)
2	.250(c)	.209(o)	.221(o)	.222(ℓ)	.268(k)	.337(g)
3	.124(c)	.134(o)	.148(o)	.148(q)	.175(k)	.219(k)
4	.098(n)	.098(p)	.104(k)	.111(k)	.131(k)	.160(k)

The letters in the above tables correspond to the choices of coefficients, "solutions", and patterns for rough components in the initial guess [see (7.17), (7.18), (7.19)] displayed in Table VII.

Concluding Remarks

As can be seen from the computational results, no particular choice of problem or initial guess always resulted in giving the worst case. Moreover, it appears that $\alpha^{\frac{1}{2}} = \hat{\alpha}^{\frac{1}{2}}$ is an upper bound on the rate of convergence of the multigrid scheme but does not yield the exact rate of convergence. Notice that there seems to be no degradation for $m = 1$. However, as m increases we find some degradation in the rate of convergence. But, it appears to be quite less than $\alpha^{\frac{1}{2}} = \hat{\alpha}^{\frac{1}{2}}$.

Table VII

Worst Case Problems

Problem	Coefficients	"Solution"	Pattern for α_k in initial guess
a	7.17b	$x(e-e^x)$	B
b	7.17c	$\sin(14\pi x)$	B
c	7.17b	$x^{5/2}(1-x)$	C
d	7.17c	$\sin(14\pi x)$	C
e	7.17b	$x(e-e^x)$	C
f	7.17c	$x(e-e^x)$	D
g	7.17a	0	D
h	7.17c	$x(e-e^x)$	D
i	7.17c	$\sin(14\pi x)$	E
j	7.17a	0	E
k	7.17b	$x(e-e^x)$	E
ℓ	7.17c	$x^{5/2}(1-x)$	A
m	7.17a	0	E
n	7.17c	$x^{5/2}(1-x)$	C
o	7.17b	$x(e-e^x)$	A
p	7.17c	$x(e-e^x)$	A
q	7.17b	$\sin(14\pi x)$	A

REFERENCES

[1] R. E. Alcouffe, A. Brandt, J. E. Dendy Jr. and J. W. Painter: The multigrid methods for the diffusion equation with strongly discontinuous coefficients. SIAM S. Sci. Stat. Comput. 2, 430-454 (1981).

[2] N. S. Bahvalov: On the convergence of a relaxation method with natural constraints on the elliptic operator. Z. vycisl. Mat. Mat. Fiz. 6, 861-883 (1966) U.S.S.R. Comp. Math. Math. Phys. 6, 101-135.

[3] A. Brandt: Multilevel adaptive solutions to boundary-value problems, Math. Comp. 31, 333-390 (1977)

[4] J. E. Dendy, Jr.: Black Box Multigrid, J. Comp. Phys. 48, 366-386 (1982).

[5] J. E. Dendy, Jr.: Black Box Multigrid for nonsymmetric problems, Los Alamos Report LA-UR-83-625 presented at International Multigrid Conference, Copper Mountain, Co. April 6-8, 1983.

[6] R. P. Fedorenko: A relaxation method for solving elliptic difference equations. Z. vycisl. Mat. Mat. Fiz. 1, 922-927 (1961) U.S.S.R. Comp. Math. Math. Phys. 1 1092-1096 (1962).

[7] P. O. Frederickson: Fast approximate inversion of large sparse linear systems, Math. Report 7-75, Lakehead University, Ontario, Canada, 1975.

[8] A. Greenbaum: Analysis of a multigrid method as an iterative technique for solving linear systems, to appear: SIAM J. Num. Anal.

[9] W. Hackbusch: On the convergence of multi-grid iterations. Beiträge Numer. Math. 9 213-239 (1981)

[10] W. Hackbusch: Convergence of multi-grid iterations applied to difference equations. Math. Comp. 34 425-440, (1980).

[11] W. Hackbusch: On the multi-grid method applied to difference equations. Computing 20, 291-306 (1978).

[12] W. Hackbusch: Introduction to multi-grid methods for the solution of boundary value problems. SEMINAIRE D'ANALYSE NUMERIQUE No. 358, Institut National Polytechnique de Grenoble, France (1981).

[13] S. McCormick and J. Ruge: Multigrid methods for variational problems. SIAM J. Num. Anal. 19, 924-929 (1982).

[14] S. McCormick, Multigrid methods for variational problems: further results. SIAM J. Numer. Anal. 21, 255-263 (1984).

[15] S. McCormick, Multigrid methods for variational problems: general theory. To Appear

[16] R. A. Nicolaides: On the ℓ^2 convergence of an algorithm for solving finite-element equation. Math. of Comp. 31, 892-906, (1977).

[17] J. C. South (Jr) and A. Brandt: Application of a multi-level grid method to transonic flow calculations. Transonic Flow Problems in Turbomachinery. T. C. Adamson and M. F. Platzer Editors, Hemisphere, Washington D.C.