IN-64

48119

P. 19

NASA Contractor Report 187633

ICASE Report No. 91-70

ICASE

THE IMPROVED ROBUSTNESS OF MULTIGRID ELLIPTIC
SOLVERS BASED ON MULTIPLE SEMICOARSENED GRIDS

Naomi H. Naik

John Van Rosendale

NASA

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665-5225

N92-10348

Unclas
0048119

G3/64

(NASA-CR-187633) THE IMPROVED ROBUSTNESS OF
MULTIGRID ELLIPTIC SOLVERS BASED ON MULTIPLE
SEMICOARSENED GRIDS Final Report (ICASE)
19 p   CSCL 12A

The Improved Robustness of Multigrid Elliptic Solvers Based on Multiple Semicoarsened Grids*

Naomi H. Naik
Mathematics Department
Vassar College
Poughkeepsie, NY 12601

John Van Rosendale
ICASE, NASA Langley Research Center
Mailstop 132C
Hampton, VA 23665

Abstract

Multigrid convergence rates degenerate on problems with stretched grids or anisotropic operators, unless one uses line or plane relaxation. For three dimensional problems, only plane relaxation suffices, in general. While line and plane relaxation algorithms are efficient on sequential machines, they are quite awkward and inefficient on parallel machines. This paper presents a new multigrid algorithm, based on the use of multiple coarse grids, that eliminates the need for line or plane relaxation in anisotropic problems. We develop this algorithm, and extend the standard multigrid theory to establish rapid convergence for this class of algorithms. The new algorithm uses only point relaxation, allowing easy and efficient parallel implementation, yet achieves robustness and convergence rates comparable to line and plane relaxation multigrid algorithms.

The algorithm described here is a variant of Mulder's multigrid algorithm [5] for hyperbolic problems. The latter uses multiple coarse grids to achieve robustness, but is unsuitable for elliptic problems, since its V-cycle convergence rate goes to one as the number of levels increases. The new algorithm combines the contributions from the multiple coarse grids via a local "switch," based on the strength of the discrete operator in each coordinate direction. This improvement allows us to show that the V-cycle convergence rate is uniformly bounded away from one, on model anisotropic problems. Moreover, the new algorithm can be combined with the idea of concurrent iteration on all multigrid levels to yield a highly parallel algorithm for strongly anisotropic problems.

1 Introduction

As is well known, the convergence rate of multigrid algorithms based on point relaxation smoothers degenerates on problems exhibiting strong anisotropies. Thus line or plane relaxation in each of the coordinate directions is often needed to obtain good multigrid convergence rates. Anisotropic discrete operators arise in problems in which the differential operator exhibits stronger coupling in some coordinate directions than in others, or when the discretization is based on highly stretched grids having mesh aspect ratios far from unity. For such problems, line relaxation typically suffices in two dimensional problems, while for three dimensional problems plane relaxation is often required. With standard (full coarsening) multigrid in three dimensions, plane relaxation in each of the coordinate directions is required in general [6].

Plane relaxation is very expensive, especially for systems of equations. Given this, interest has focused recently on "semicoarsening" algorithms, in which the plane relaxation is carried out in only one direction, while the grid is "coarsened" only in the direction orthogonal to these planes [1]. The required plane solves can then be done recursively, via an analogous two dimensional algorithm, based on line relaxation in one direction and coarsening in the orthogonal direction.

While such "semicoarsening" algorithms are fast and effective on sequential architectures, they have limited and awkward parallelism. The recursive solution of two dimensional problems, required in the plane relaxation, takes $O(\log^2 N)$ parallel operations, so that it takes $O(\log^3 N)$ parallel operations per three dimensional V-cycle, where N is the number of mesh points. Thus it takes time at least $O(\log^4 N)$ to converge to truncation error.

An alternative way of achieving robustness, which avoids line and plane relaxations altogether, is to use multiple coarse grids formed by semicoarsening in each of the coordinate directions, as in Mulder's hyperbolic algorithms [5]. In this paper, a modification of Mulder's method is proposed, which substantially improves the convergence properties of his method, when applied to elliptic problems. This enables one to design effective V-cycle elliptic solvers for anisotropic problems, using only point relaxation smoothers. This new class of methods is shown to be simple, robust, and effective.

One can also construct a highly parallel algorithm for anisotropic problems by combining the new algorithm with the idea of concurrent iteration on all multigrid levels [2]. This paper gives numerical experiments suggesting the efficacy of this approach, though we have yet to explore the use of this algorithm on parallel machines.

2 Algorithm Design

Strong coupling of the operator in a particular direction can easily degrade the performance of a multigrid method. There are several ways of accelerating convergence in the case of such anisotropy. Line relaxation and semicoarsening methods can be used to correct, respectively, the inability of the relaxation method to solve for some high frequencies, and the inability of the standard coarse grid to represent high frequencies. In the line relaxation method, the ineffective point relaxation is replaced by line relaxation in the direction of strong coupling.

By removing the residual components due to the strong coupling, the remaining residual due to weak coupling in the other direction can be effectively smoothed by the relaxation. Thus line relaxation together with standard coarsening is sufficient to uniformly reduce all Fourier components, in two dimensional problems. In the second approach, instead of using line relaxations, the grid is coarsened only in the direction of strongest coupling. In this case, point relaxation together with semicoarsening suffices to uniformly reduce all Fourier components.

In addition to line relaxation and semicoarsening, other methods have been proposed that more aggressively solve for the difficult frequencies. Hackbusch's Robust Parallel Multigrid [3] uses 'forced aliasing' to represent high frequency components on standard coarse grids. The high frequency components of the residual are aliased to low frequencies, solved for on a coarse grid, and then the coarse grid correction is "de-aliased" back to the high frequency. Although this method uses point relaxations and standard coarsening, it requires the use of multiple coarse grids, each with a different discrete operator, and is thus quite complex.

In this paper we will look at a natural extension of the second approach, use of semicoarsening. This approach was originally proposed by Mulder [5] for overcoming the problem of alignment in fluid flow computations. The simple technique of semicoarsening simultaneously in all coordinate directions, and properly weighting the contributions from each of the coarse grids, yields an efficient, robust, and easily parallelizable multigrid method for general tensor product grid.

3   The Algorithm

The multiple semicoarse grid (MSG) correction scheme (for linear problems) is similar to the standard multigrid correction scheme, except that there are now extra grids involved. In two dimensions every grid is simultaneously coarsened in two directions.

We first suppose, for simplicity, that the domain of the model boundary problem is the unit square, and that this problem is to be solved on an $N \times N$ uniform grid given by

$$\Omega_h = \{(ih, jh) \mid i = 0, 1, \ldots, N - 1; \quad j = 0, 1, \ldots, N - 1\},$$

where $h = 1/N$ and N is a power of two. Let the subgrid, $\Omega^{m,n}$, obtained by successively semicoarsening Ω_h, be the grid with $N/2^m$ grid points in the x direction and $N/2^n$ grid points in the y direction.

Notice that the notation does not distinguish between a grid obtained by semicoarsening first in the y direction and then in the x direction and a grid obtained by semicoarsening first in the x direction and then in the y direction. As shown in Mulder, in order to construct reasonable algorithms in three or more dimensions, the problems on equivalent grids must be combined. Figure 1 shows the interrelations between the various grids for a two dimensional problem with an 8×8 fine grid. With coarse grids combined as in this diagram, one has a only 16 grids altogether, while without combining the full binary tree of grids would contain 69 grids and have no real numerical advantage.

Now introducing more notation, the discrete equations on grid $\Omega^{m,n}$ are written as:

$$A^{m,n} U^{m,n} = F^{m,n} \tag{1}$$

The operators $A^{m,n}$ can be thought of as either discretizations of the differential operator, L, on the grid $\Omega^{m,n}$, or as operators obtained variationally from the fine grid and intergrid transfer operators.

A k grid ($N = 2^k$) V-cycle for this method is performed in three parts. In part a the information is propagated from the fine grids to the coarse grids, in part b the equations are solved on the coarsest grid, and in part c the information is propagated back from the coarse grids to the fine grids.
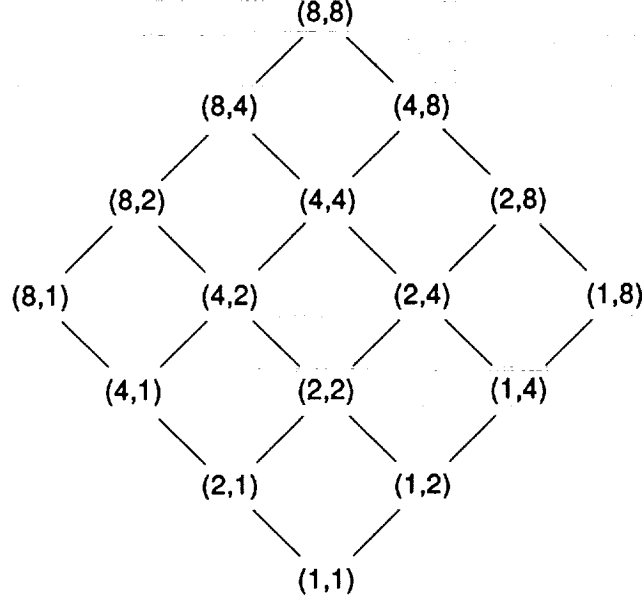
MSG algorithm

a. For $l = 0, 1, 2, \ldots, 2k - 1$:
  For all $m \geq 0$, $n \geq 0$ such that $m + n = l$:
   1. If $l > 0$, combine restricted residuals on $\Omega^{m,n}$
   2. Relax ν_1 times on the $\Omega^{m,n}$-grid equations
   3. If $m < k$, transfer (restrict) residual from $\Omega^{m,n}$ to $\Omega^{m+1,n}$
   4. If $n < k$, transfer (restrict) residual from $\Omega^{m,n}$ to $\Omega^{m,n+1}$

b. For $l = 2k$:
   1. Combine restricted residuals on $\Omega^{k,k}$
   2. Solve (1) by any direct or iterative method on $\Omega^{k,k}$
   3. Transfer (interpolate) correction from $\Omega^{k,k}$ to $\Omega^{k-1,k}$ and $\Omega^{k,k-1}$

c. For $l = 2k - 1, 2k - 2, \ldots, 1, 0$:
  For all $m \geq 0$, $n \geq 0$ such that $m + n = l$:
   1. If $l > 0$, combine interpolated corrections on $\Omega^{m,n}$
   2. Relax ν_2 times on the $\Omega^{m,n}$-grid equations
   3. If $m > 0$, transfer (interpolate) correction from $\Omega^{m,n}$ to $\Omega^{m-1,n}$
   4. If $n > 0$, transfer (interpolate) correction from $\Omega^{m,n}$ to $\Omega^{m,n-1}$

Any point relaxation on equation (1) can be used in steps $a2$ and $c2$. For steps $a3$, $a4$, $c3$, and $c4$ we consider intergrid transfer operators which are one dimensional. For example, the residual restriction operators could be the usual three point averaging formulas, given by the stencils:

$$I_{m,n}^{m+1,n} = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}_{m,n}^{m+1,n}$$

$$I_{m,n}^{m,n+1} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}_{m,n}^{m,n+1}$$

Figure 1: Semicoarsening of an 8×8 grid

```
                          (8,8)
                         /     \
                  (8,4)           (4,8)
                 /     \         /     \
            (8,2)        (4,4)        (2,8)
           /     \      /     \      /     \
       (8,1)       (4,2)        (2,4)        (1,8)
           \      /     \      /     \      /
            (4,1)        (2,2)        (1,4)
                 \      /     \      /
                  (2,1)          (1,2)
                       \        /
                          (1,1)
```

Similarly, one dimensional linear interpolation in the direction of grid refinement could be used to bring the corrections from coarse to fine grids.

We now look at the details of step $a1$, in which the residuals on $\Omega^{m,n}$ must be combined, and step $c1$, in which the corrections on $\Omega^{m,n}$ must be combined. The restricted residuals are simply averaged. Specifically,

$$
r^{m,n} = \begin{cases}
\frac{1}{2}(I^{m,n}_{m-1,n} r^{m-1,n} + I^{m,n}_{m,n-1} r^{m,n-1}) & \text{if } m > 0, n > 0, \\[2mm]
I^{m,n}_{m-1,n} r^{m-1,n} & \text{if } n = 0, \\[2mm]
I^{m,n}_{m,n-1} r^{m,n-1} & \text{if } m = 0.
\end{cases}
\tag{2}
$$

A weighted average of the interpolated corrections is used, so that

$$
u^{m,n} = \begin{cases}
\omega_1^{m,n} I^{m,n}_{m+1,n} u^{m+1,n} + \omega_2^{m,n} I^{m,n}_{m,n+1} u^{m,n+1} & \text{if } m < k, n < k, \\[2mm]
I^{m,n}_{m+1,n} u^{m+1,n} & \text{if } n = k, \\[2mm]
I^{m,n}_{m,n+1} u^{m,n+1} & \text{if } m = k.
\end{cases}
$$

The MSG algorithm can lead to multigrid convergence rates independent of mesh size, provided the weights ω_1 and ω_2 are chosen properly.

4

4   Convergence Theory

In this section, we give our convergence proof for the MSG method for a constant coefficient model problem in two dimensions, and derive sufficient conditions on the weights. These conditions are then used in the next section to motivate the choice of weights for the variable coefficient problem.

In order to show that the convergence rate of the MSG method is independent of mesh size for linear, constant coefficient model problems, we make the following assumptions:

A1. The coarse grid operators are Galerkin, or 'variational'.

A2. The restriction and projection operators are adjoints of each other and are one dimensional.

A3. The discretized operator is symmetric positive definite.

A4. The linear part of the smoother and the discrete operator commute.

We model both our analysis and our notation after that in [4], although some notational changes are needed in order to keep track of the multiple grids on each level. In particular, it is more convenient here to label grid levels in the reverse order, so that the grid level increases as the grid becomes coarser, contrary to the standard convention. If the two grid directions are to be coarsened a maximum of $\bar{m}$ times in the first direction and $\bar{n}$ times in the second direction ($0 \le m \le \bar{m}$ and $0 \le n \le \bar{n}$) then the coarsest level will be given by $\bar{l} = \bar{m} + \bar{n}$ and the coarsest grid will be given by the indices $\bar{m}, \bar{n}$.

We are looking for the solution of

$$A^{0,0} u^{0,0} = f^{0,0}$$

where $A^{0,0}$ is symmetric positive definite. For each coarser grid level, $l = 1, \ldots, \bar{l}$, we recursively define each of the operators $A^{m,n}$, for $m + n = l$,

$$A^{m,n} = \begin{cases} I^{m,n}_{m-1,n} A^{m-1,n} I^{m-1,n}_{m,n} & \text{if } m > 0 \\[2mm] I^{m,n}_{m,n-1} A^{m,n-1} I^{m,n-1}_{m,n} & \text{if } m = 0 \end{cases}$$

Note that if m and n are both positive then there are two ways to construct the coarse grid operators from the fine grid operator. However, since the intergrid operators are one-dimensional,

$$I^{m,n}_{m,n-1} I^{m,n-1}_{m-1,n-1} = I^{m,n}_{m-1,n} I^{m-1,n}_{m-1,n-1}, \tag{3}$$

and therefore either way gives the same result.

Our notation is as follows. Each of the $A^{m,n}$ are operators on a finite dimensional space, $H^{m,n}$. Since we will only be looking at two grid levels at a time, we simplify the grid indices

by the shorthand notation:

$$\begin{aligned} k &= m, n \\ k_1 &= m+1, n \\ k_2 &= m, n+1 \end{aligned}$$

We also define the inner products

$$\langle u^k, v^k \rangle_k = \sum_i u_i^k v_i^k$$

$$[u^k, v^k]_k = \langle A^k u^k, v^k \rangle_k$$

for any u^k, v^k in H^k. The second, 'energy', inner product induces a norm on H^k, which we denote by $\| \cdot \|_k$, thus

$$\|v^k\|_k^2 = [v^k, v^k]_k = \langle A^k v^k, v^k \rangle_k.$$

We also define four subspaces of each space H^k as follows. For $i = 1, 2$:

$$\begin{aligned} \hat{S}_i^k &= R(I_{k_i}^k) \\ \hat{T}_i^k &= \{v^k \in H^k | \ [v^k, w^k]_k = 0 \quad \text{for all } w^k \in \hat{S}_i^k\} \end{aligned}$$

Corresponding to these subspaces, we define projection operators, T_i^k and S_i^k such that

$$\begin{aligned} R(T_i^k) &= \hat{T}_i^k \\ ker(T_i^k) &= \hat{S}_i^k \\ S_i^k &= I - T_i^k. \end{aligned}$$

for $i = 1, 2$.

With the above notation, we are now ready to discuss the V-cycle convergence analysis. The approximation to the solution of the kth grid equations is updated three times per V-cycle; once after the ν_1 relaxation sweeps in step A2, once after the coarse grid correction in step C1 and once after the ν_2 relaxation sweeps in step c2. We label the initial approximation as $u_{(0)}^k$, and the approximations after each of the three updates as $u_{(i)}^k$ for $i = 1, 2$ and 3. If $\mathcal{G}$ denotes the relaxation operator, then the updates are given by

$$u_{(1)}^k = \mathcal{G}^{\nu_1}(u_{(0)}^k, f^k)$$

$$u_{(2)}^k = u_{(1)}^k + \omega_1^k I_{k_1}^k u_{(3)}^{k_1} + \omega_2^k I_{k_2}^k u_{(3)}^{k_2}$$

$$u_{(3)}^k = \mathcal{G}^{\nu_2}(u_{(2)}^k, f^k)$$

where

$$u_{(3)}^{k_j} = MG(u_{(0)}^{k_j}, f^{k_j}), \qquad j = 1, 2.$$

We make two additional assumptions in order to simplify the case when two grids on the same grid level are semicoarsened in opposite directions, yielding coarse grids of the same dimensions. Recall that the coarse grid problems on these coarse grids are combined to form a single problem.

6

A5. The initial approximation is equal to zero on all except the finest level:

$$u_{(0)}^k = 0, \qquad \text{for all } k = m, n \neq 0, 0$$

A6. There is no smoothing in the fine-to-coarse part of the V-cycle:

$$\nu_1 = 0$$

These assumptions guarantee that the residuals from both grids are identical, as shown in the following lemma.

Lemma 1 *For m, n with $0 < m \leq \bar{m}$, $0 < n \leq \bar{n}$:*

$$f^{m,n} = I_{m-1,n}^{m,n} f^{m-1,n} = I_{m,n-1}^{m,n} f^{m,n-1}$$

Proof:
The lemma is proved by a simple induction argument on the grid level l, using the additional assumptions A5 and A6, together with Equations (2) and (3). ∎

A standard multigrid V-cycle convergence result can be based on a single assumption, which combines both the smoothing and the approximation properties of the problem, namely

$$\frac{\|v\|_k^2}{\|F^k v\|_k^2} \geq 1 + \beta \left(\frac{\|T^k F^k v\|_k^2}{\|F^k v\|_k^2} \right)^{1/\alpha}$$

for all $v \in H^k$ and for all grid levels k. See [4]. Here F^k is the linear part of the smoothing operator $\mathcal{G}^{\nu_2}$. We assume sufficient regularity and take $\alpha = 1$. Then if we define the multigrid convergence rate on the kth grid level as

$$\varepsilon^k = \inf_{\varepsilon} \{ \|u^k - MG^k(v^k, f^k)\|_k \leq \varepsilon \|u^k - v^k\|_k \text{ for all } v^k \in H^k \},$$

the V-cycle convergence theorem for standard multigrid algorithms is

Theorem 1 (Standard) *Let $k \geq 2$ and suppose we have already bounded ε^{k-1} by $\dfrac{1}{1 + \beta}$. Then*

$$\varepsilon^k \leq \frac{1}{1 + \beta}.$$

In our case, in which we have multiple coarse grids on each grid level, we can prove a similar result. In fact, the convergence of the MSG V-cycle algorithm can be as good as the convergence of a standard V-cycle multigrid in which every grid is semicoarsened only in the optimal direction.

7

Theorem 2 (MSG) *Suppose that*

$$\|v\|_k^2 \geq \|F^k v\|_k^2 + \beta_i \|T_i^k F^k v\|_k^2, \qquad i = 1, 2, \tag{4}$$

for all $v \in H^k$ and for all grid levels k, and choose the weights, ω_1^k and ω_2^k, so that

$$\omega_i^k = \frac{\beta_i}{\beta_1 + \beta_2}.$$

If

$$max(\varepsilon^{m+1,n}, \varepsilon^{m,n+1}) \leq min\left(\frac{1}{1+\beta_1}, \frac{1}{1+\beta_2}\right)$$

then

$$\varepsilon^k \leq min\left(\frac{1}{1+\beta_1}, \frac{1}{1+\beta_2}\right).$$

Proof:

First we note that

$$\omega_1 + \omega_2 = 1. \tag{5}$$

We denote the errors corresponding to the updates by $e_{(i)}^k$ for $i = 0, 1, 2$ and 3, where

$$e_{(i)}^k = u^k - u_{(i)}^k$$

and where u^k is the exact solution of the kth grid equations. Using Lemma 1 we see that, for both $i = 1$ and $i = 2$,

$$e_{(1)}^k - I_{k_i}^k u^{k_i} = T_i^k e_{(1)}^k.$$

By our assumptions we also have

$$e_{(1)}^k = e_{(0)}^k. \tag{6}$$

Combining these results, we can write

$$e_{(2)}^k = (\omega_1 T_1^k + \omega_2 T_2^k)e_{(0)}^k + \omega_1 I_{k_1}^k e_{(3)}^{k_1} + \omega_2 I_{k_2}^k e_{(3)}^{k_2}.$$

Then the error before and after the ν_2 post-relaxation sweeps is

$$e_{(3)}^k = F^k e_{(2)}^k.$$

Since we need to look at only two grid levels at a time, we will temporarily suppress the notational references to the current grid, k. Thus, we define, for $i = 1, 2$ and $j = 0, 1, 2, 3$,

$$\begin{aligned}
u^{k_i} &= u^i, \\
e^{k_i} &= e^i, \\
e_{(j)}^k &= e_{(j)},
\end{aligned}$$

and so on.

8

Consider an arbitrary element, v, of H^k, and let $y = Fv$. Then

$$[e_{(3)}, v]_k = [e_{(2)}, y]$$

$$= [(\omega_1 T_1 + \omega_2 T_2)e_{(0)} + \omega_1 I_{k_1}^k e_{(3)}^1 + \omega_2 I_{k_2}^k e_{(3)}^2, y]$$

$$= \omega_1\left([T_1 e_{(1)}, T_1 y] + [I_{k_1}^k e_{(3)}^1, S_1 y]\right) + \omega_2\left([T_2 e_{(1)}, T_2 y] + [I_{k_2}^k e_{(3)}^2, S_2 y]\right)$$

By the Cauchy-Schwartz inequality,

$$|[e_{(3)}, v]| \le \omega_1\left(\|T_1 e_{(1)}\|\,\|T_1 y\| + \varepsilon_1\|S_1 e_{(1)}\|\,\|S_1 y\|\right) + \omega_2\left(\|T_2 e_{(1)}\|\,\|T_2 y\| + \varepsilon_2\|S_2 e_{(1)}\|\,\|S_2 y\|\right),$$

since, for $i = 1, 2$,

$$\|I_{k_i}^k e_{(3)}^i\| = \|e_{(3)}^i\| \le \varepsilon_i\|e_{(0)}^i\| = \varepsilon_i\|u^i\| = \varepsilon_i\|I_{k_i}^k u^i\| = \varepsilon_i\|S_i e_{(1)}\|.$$

Using the Cauchy-Schwartz inequality one more time gives us

$$|[e_{(3)}, v]|^2 \le \left(\omega_1(\|T_1 e_{(1)}\|^2 + \|S_1 e_{(1)}\|^2) + \omega_2(\|T_2 e_{(1)}\|^2 + \|S_2 e_{(1)}\|^2)\right)$$
$$\cdot \left(\omega_1(\|T_1 y\|^2 + \varepsilon_1\|S_1 y\|^2) + \omega_2(\|T_2 y\|^2 + \varepsilon_2\|S_2 y\|^2)\right).$$

Dividing through by the square norms of the initial error and v and using Equations (5) and (6) this simplifies to

$$\frac{|[e_{(3)}, v]|^2}{\|e_{(0)}\|^2 \|v\|^2} \le \frac{\|y\|^2}{\|v\|^2}\left(\frac{\omega_1(\|T_1 y\|^2 + \varepsilon_1\|S_1 y\|^2) + \omega_2(\|T_2 y\|^2 + \varepsilon_2\|S_2 y\|^2)}{\|y\|^2}\right).$$

It is convenient to define two variables, t_1 and t_2, such that

$$t_1 = \frac{\|T_1 y\|^2}{\|y\|^2}, \qquad t_2 = \frac{\|T_2 y\|^2}{\|y\|^2}$$

and rewrite our inequality as

$$\frac{|[e_{(3)}, v]|^2}{\|e_{(0)}\|^2 \|v\|^2} \le \frac{\|Fv\|^2}{\|v\|^2}\left(\omega_1(t_1 + \varepsilon_1(1 - t_1)) + (\omega_2(t_2 + \varepsilon_2(1 - t_2))\right). \tag{7}$$

The smoothing and approximation hypotheses given by Equation (4) of the theorem can then be rewritten in terms of the new variables as,

$$\frac{\|v\|^2}{\|Fv\|^2} \ge 1 + \beta_1 t_1,$$

$$\frac{\|v\|^2}{\|Fv\|^2} \ge 1 + \beta_2 t_2.$$

9

Figure 2: Limits on ω_1
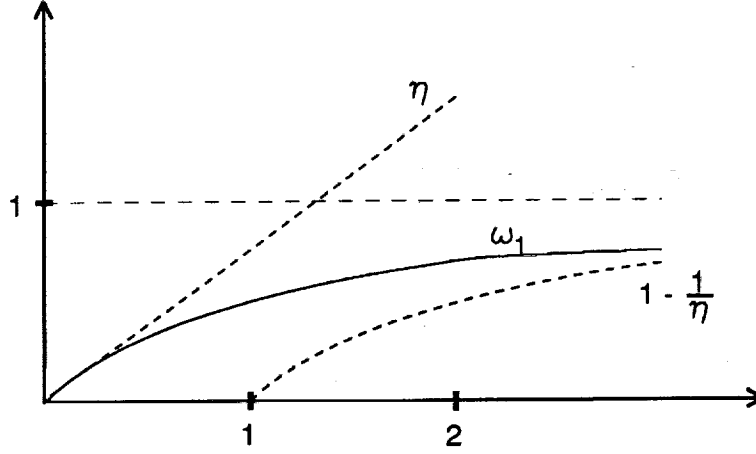
We can therefore write inequality (7) in terms of t_1 and t_2 with either of these upper bounds on $\frac{\|Fv\|^2}{\|v\|^2}$, so we are free to use whichever is smaller. Thus,

$$\frac{|[e_{(3)}, v]|^2}{\|e_{(0)}\|^2 \|v\|^2} \leq \min\left(\frac{1}{1+\beta_1 t_1}, \frac{1}{1+\beta_2 t_2}\right)\left((\omega_1(t_1 + \varepsilon_1(1-t_1))) + (\omega_2(t_2 + \varepsilon_2(1-t_2)))\right). \quad (8)$$

Taking the maximum over all values of t_1 and t_2 we arrive at the following bound on the convergence rate,

$$\varepsilon^k \leq \max_{\substack{0 \leq t_1 \leq 1 \\ 0 \leq t_2 \leq 1}} \left(\min\left(\frac{1}{1+\beta_1 t_1}, \frac{1}{1+\beta_2 t_2}\right)\left((\omega_1(t_1 + \varepsilon_1(1-t_1))) + (\omega_2(t_2 + \varepsilon_2(1-t_2)))\right)\right).$$

Using the definitions of the weights ω_1 and ω_2 and the conditions on the ε_i's, it follows that

$$\varepsilon^k \leq min\left(\frac{1}{1+\beta_1}, \frac{1}{1+\beta_2}\right). \qquad \blacksquare$$

Note that the weights in the hypothesis of the theorem, while convenient, are not the only choice. All we really needed in the proof was that the weights lie within some bounds, determined by the β_i's. If we define the ratio, $\eta = \frac{\beta_1}{\beta_2}$, then the theorem will be proved provided $\omega_1 \leq \eta$ and $\omega_2 \geq 1 - \eta$ whenever $\eta \leq 1$ and $\omega_1 \geq 1 - 1/\eta$ and $\omega_2 \leq 1/\eta$ whenever $\eta \geq 1$. For the statement of the theorem, we have chosen $\omega_1 = \eta/(\eta + 1)$. See Figure 2.

5 Practical choice of the weights

Suppose that, as in Mulder, the corrections from the two coarse grids are averaged. Thus, the weights are given by

$$\omega_1^{m,n} = \frac{1}{2}, \qquad \omega_2^{m,n} = \frac{1}{2}.$$

These weights give a two-grid convergence rate of approximately 1/2 in the case of strong alignment, because the appropriate grid gets only half of the needed information. In this case, our convergence result does not guarantee good convergence. Thus we should look for ways, based on our theorem, to improve the convergence of this method.

Recall that only semicoarsening in the direction of strongest coupling can significantly reduce the frequencies which cannot be reduced by the point relaxation. The weights provide a way of "switching" to the appropriate coarse grid in the cases of strong alignment in one of the coordinate directions. For our model problem, $\alpha u_{xx} + \gamma u_{yy} = f$, we have some degree of freedom in the choice of the weights. We could take, for instance,

$$\omega_1 = \frac{\alpha^2}{\alpha^2 + \gamma^2}, \qquad \omega_2 = \frac{\gamma^2}{\alpha^2 + \gamma^2}.$$

Since the appropriate grid gets all of the needed information in the case of strong alignment, these weights can lead to convergence rates which can be made arbitrarily small by increasing the number of relaxation sweeps.

In general, ω_1 and ω_2 will vary over the domain and we will not know the relative strengths α and γ explicitly. Suppose we know that, locally, all modes which cannot be efficiently reduced by point relaxation can be well approximated on the same semicoarsened coarse grid. That is, suppose semicoarsening can be used locally to accelerate the convergence. In this case, we would like to determine the most efficient direction of semicoarsening and choose our weights accordingly. One way to do this is to test the operator, at the given grid point, on two different high frequency Fourier modes, one oscillatory only in the x-direction, the other oscillatory only in the y-direction. The two modes, call them u and v, which are most natural look locally like:

```
1 -1  1 -1  1 -1           -1 -1 -1 -1 -1 -1
1 -1  1 -1  1 -1            1  1  1  1  1  1
1 -1  1 -1  1 -1           -1 -1 -1 -1 -1 -1
1 -1  1 -1  1 -1            1  1  1  1  1  1
1 -1  1 -1  1 -1           -1 -1 -1 -1 -1 -1
1 -1  1 -1  1 -1            1  1  1  1  1  1
         u                          v
```

Appropriate weights at the grid point (i,j) can be determined by applying the operator,

11

$A^{m,n}$ to u and v. We define

$$(A^{m,n}u)(i,j) = \lambda_u(i,j) \quad \text{and} \quad (A^{m,n}v)(i,j) = \lambda_v(i,j).$$

Then a reasonable choice for the weights is:

$$\omega_1^{m,n}(i,j) = \frac{\lambda_u^2}{\lambda_u^2 + \lambda_v^2}, \qquad \omega_2^{m,n}(i,j) = \frac{\lambda_v^2}{\lambda_u^2 + \lambda_v^2}.$$

Thus, if there is a direction in which semicoarsening can give an acceptable convergence rate, this method should find that direction.

## 6 Theoretical Complexity - Sequential and Parallel

As shown in Mulder [5], the cost of a sequential MSG V-cycle is proportional to the total number of points on all grids. In two dimensions, where the fine grid consists of $M \times N$ points, there is a total of $(2M - 1)(2N - 1)$ grid points on all of the grids combined, as can be easily seen by arranging all of the grids as in Figure 3. Thus, there are approximately four times as many points on all the grids as there are on the finest grid. A similar arrangement of all of the grids obtained by semicoarsening a three dimensional $L \times M \times N$ grid is also shown in Figure 3 giving a total of $(2L - 1)(2M - 1)(2N - 1)$ grid points on all of the grids, approximately 8 times the number of points on the finest grid. The cost of the d dimensional algorithm, will be roughly proportional to 2^d times the number of points on the finest grid. A parallel implementation of the MSG algorithm is relatively straightforward since the computational work on a given level is local and can be performed simultaneously at many grids points. For a modest number of processors, most of the computation time is spent on the fine grid levels since, on each fine grid level, l, there are

$$\binom{l + d - 1}{d - 1} \frac{MN}{2^l} \approx \frac{l^{d-1}}{(d-1)! 2^l} MN$$

grid points per level. On coarser levels this is an upper bound on the number of grid points. Therefore the number of grid points decreases like a polynomial divided by an exponential as the levels become coarser. For a large number of processors, approximately equal to the number of grid points on the finest grid level N, an equal amount of parallel computation time is spent on all grid levels, resulting in a computational cost per V-cycle on the order of $log(N)$.

On message passing machines, the communication between grid levels can become a problem as the number of processors is increased. Consider what happens in the extreme case where we have as many processors available as we have grid points on the finest grid. If we assign one grid point to each processor, then some processors have more work on the coarser levels and some processors will have no work, simply because the multiple semicoarse grids have some grid points in common. Thus, some sort of re-distribution must occur in order to keep the load balanced. We propose two different schemes, a simple scheme involving

Figure 3: Successive semicoarsening, total number of points
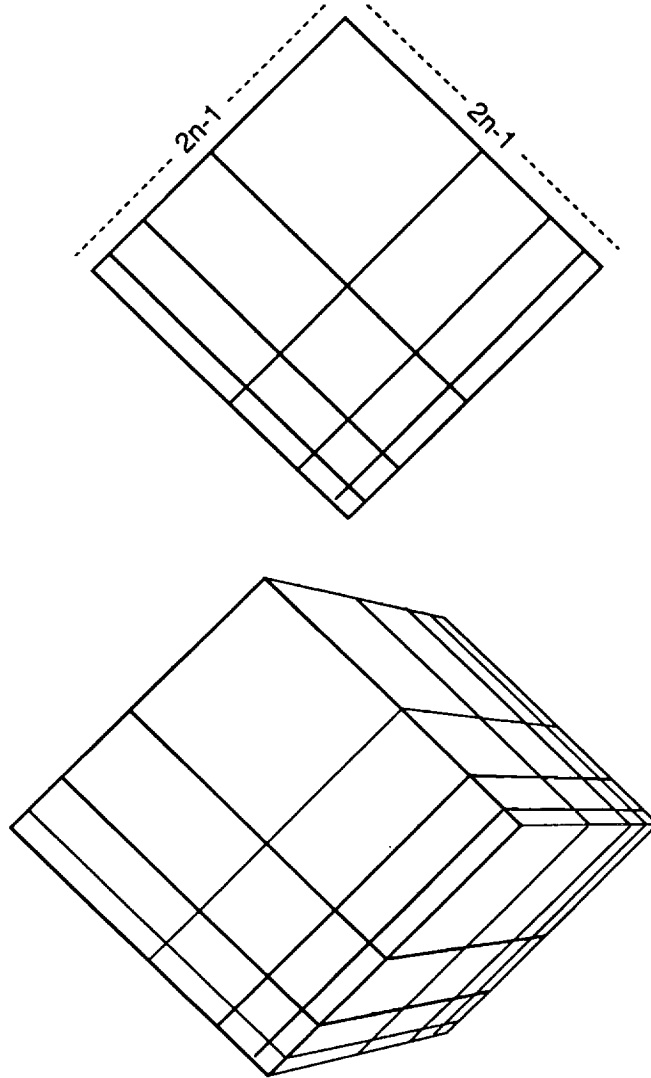
transposes, which works only in two dimensions, and a second scheme which offsets the grids in order to reduce the communication. Both schemes preserve the computational complexity, but have differing communication requirements. The offsetting scheme is ideal for hypercube communication networks, since all communication is between nearest neighboring processors.

The transpose scheme is based on the observation that in two dimensions, even if the grids of the same dimensions are not combined, the total number of points on each level does not increase. For example, if we start with an 8×8 grid on level zero, we get a 4×8 and an 8×4 grid on level one, still having 64 mesh points. On level two, we get a 2×8, two 4×4, and an 8×2 grid. By carefully transposing and packing these grids, we can fit all 2^l grids on level l into a 8×8 array, as shown in Figure 4. This mapping does keep all calculations on a particular grid local, but involves packing and shifting between grid levels.

In higher dimensions, this type of packing of the coarse grids does not work. Moreover,

Figure 4: Transpose scheme
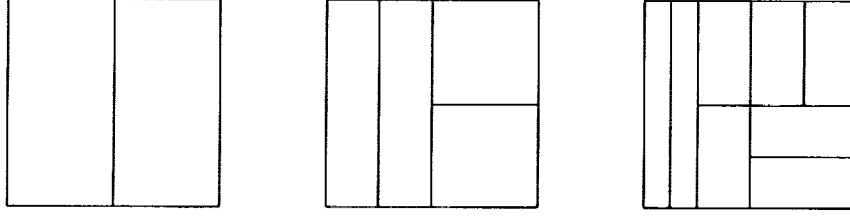
this scheme involves intense communication in the packing phase which could make it prohibitively expensive on some machines. For hypercube communication networks, there are alternate schemes which keep all of the grid data local. One possibility is to offset the various coarse grids to redistribute the load. For example, in two dimensions, using N^2 points and $(2N - 1)^2$ processors, there is a simple one-to-one mapping from all of the points on all of the grids to the $(2N - 1)^2$ processors. For every grid level l $(0 \le l \le k)$, and for every grid $\Omega^{m,n}$ on level l $(m \ge 0, n \ge 0, m + n = l)$, let the (i,j)th grid point on grid $\Omega^{m,n}$ be assigned to the $(\hat{i}, \hat{j})$ processor where

$$\hat{i} = 2^{m+1}i + 2^m - 1$$
$$\hat{j} = 2^{n+1}j + 2^n - 1,$$

where $0 \le i < N/2^m$ and $0 \le j < N/2^m$. The quantities $2^m - 1$ and $2^n - 1$ in the above expressions are the horizontal and vertical offsets for the $\Omega^{m,n}$ grid. Since information is only transferred to grids differing by one in either m or n, then the relative offsets are always by a power of 2 in one direction. This scheme works equally well in three or more dimensions and the extension is obvious.

Finally, we note that the offsetting of the various grids can easily be incorporated into the interpolation and restriction operations. During a restriction from $\Omega^{m,n}$ to $\Omega^{m+1,n}$, for example, an averaging of three values of the residual in the x direction is immediately followed by a shift of all values to the processor which is 2^m to the right. Similarly for the coarsening in the y direction. During the interpolation, this process is reversed. Interpolated values are shifted to the left. This method automatically maintains the correct offset for all of the grids and increases the communication by only a constant.

Note that, when relaxations are only performed on one grid level at a time, it is sufficient to offset the grids in only one of the two coordinate directions, using only $N(2N - 1)$ processors.

7 Experimental Results

Experimentally, the MSG algorithm converges extremely well for the model problem $\alpha u_{xx} + \gamma u_{yy} = f$, using the weights suggested in Section 5. Asymptotic convergence rates are given in Table 1. These were obtained using a random initial approximation to the solution, rescaling after each iteration and observing the limit of the ratio of subsequent errors (l_2).

14

All available coarse grid levels were used, with two red/black SOR relaxation sweeps per grid level and an exact solve on the coarsest level. The convergence rates are seen to be uniformly small for all ratios, α/γ.

MSG was also used on Poisson's equation on non-uniformly stretched grids. Chebyshev grids were used in both directions, with convergence comparable to the uniformly stretched grids. See Table 1. The convergence rates for grids which have Chebyshev stretching in only one direction are also given and can be seen to be in the same range as for the model problem.

The last entry in Table 1 is for exponential stretching of the grid in one of the coordinate directions. The exponential stretching is done so that the ratio of the lengths of the first and the last cell is 10,000. The convergence rates are slightly worse, but still appear to be bounded independently of grid size.

	8×8	16×16	32×32	64×64
Uniform Grid				
$\alpha/\gamma = 1$	0.07	0.09	0.10	0.10
$\alpha/\gamma = 10$	0.13	0.15	0.15	0.15
$\alpha/\gamma = 100$	0.16	0.19	0.19	0.19
$\alpha/\gamma = 1000$	0.16	0.19	0.21	0.21
Chebyschev Grid	0.14	0.15	0.16	0.16
Uniform/Chebyshev	0.09	0.13	0.15	0.16
Exponential Stretching	0.15	0.18	0.19	0.20

Table 1: Asymptotic convergence rates of MSG on various types of grids

On massively parallel architectures, the relaxation sweeps in the MSG algorithm can be performed concurrently on all grids on all grid levels using the CMG algorithm of Gannon and Van Rosendale [2]. The combined CMG/MSG algorithm proceeds in two phases. In the first phase, the relaxation is performed on all grids, on all levels. The second phase is the intergrid transfer phase, in which residuals and corrections from each grid are transferred to neighboring coarse and fine grids, respectively. Experimental results indicate that the robustness properties of the MSG algorithm are retained. In Table 2, the observed convergence rates are given for the model problem. Note that we again observe that strong alignment does not seriously degrade the convergence. The convergence rates per concurrent iteration are mostly in the 0.4-0.6 range.

	8×8	16×16	32×32	64×64
Uniform Grid				
$\alpha/\gamma = 1$	0.42	0.52	0.59	0.60
$\alpha/\gamma = 10$	0.46	0.48	0.59	0.61
$\alpha/\gamma = 100$	0.40	0.43	0.52	0.57
$\alpha/\gamma = 1000$	0.41	0.43	0.51	0.55
Chebyschev Grid	0.37	0.44	0.50	0.53
Uniform/Chebyshev	0.44	0.53	0.56	0.57
Exponential Stretching	0.48	0.59	0.63	0.63

Table 2: Observed convergence rates of MSG/CMG on various types of grids

References

[1] J.E. Dendy, Jr., M.P. Ida and J.M. Rutledge. "A Semi-coarsening Multigrid Algorithm for SIMD Machines," Los Alamos Technical Report LA-UR 91-894.

[2] D. Gannon and J. Van Rosendale. "Highly Parallel Multigrid Solvers for elliptic PDE's: an experimental analysis," ICASE Report 82-36.

[3] W. Hackbusch. "The Frequency Decomposition Multigrid Method, Part I: Application to Anisotropic Equations," Numer. Math. 56 (1989) pp. 229-245.

[4] J. Mandel, S. McCormick and R. Bank. "Variational Multigrid Theory," in Multigrid Methods, S. McCormick, Ed., SIAM, 1987.

[5] W. Mulder. "A New Multigrid Approach to Convection Problems," CAM Report 88-04, March 1988.

[6] K. Stüben and U. Trottenberg. "On the construction of Fast Solvers for elliptic equations," IMA-Report Mr. 82.0201, GMD MBH Bonn, 1982.

Report Documentation Page

1. Report No. NASA CR-187633 ICASE Report No. 91-70	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle THE IMPROVED ROBUSTNESS OF MULTIGRID ELLIPTIC SOLVERS BASED ON MULTIPLE SEMICOARSENED GRIDS		5. Report Date September 1991
		6. Performing Organization Code
7. Author(s) Naomi H. Naik John Van Rosendale		8. Performing Organization Report No. 91-70
		10. Work Unit No. 505-90-52-01
9. Performing Organization Name and Address Institute for Computer Applications in Science and Engineering Mail Stop 132C, NASA Langley Research Center Hampton, VA 23665-5225		11. Contract or Grant No. NAS1-18605
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Langley Research Center Hampton, VA 23665-5225		13. Type of Report and Period Covered Contractor Report
		14. Sponsoring Agency Code

15. Supplementary Notes
Langley Technical Monitor: Michael F. Card Submitted to SIAM Journal of Numerical Analysis Final Report

16. Abstract

Multigrid convergence rates degenerate on problems with stretched grids or anisotropic operators, unless one uses line or plane relaxation. For three dimensional problems, only plane relaxation suffices, in general. While line and plane relaxation algorithms are efficient on sequential machines, they are quite awkward and inefficient on parallel machines. This paper presents a new multigrid algorithm, based on the use of multiple coarse grids, that eliminates the need for line or plane relaxation in anisotropic problems. We develop this algorithm, and extend the standard multigrid theory to establish rapid convergence for this class of algorithms. The new algorithm uses only point relaxation, allowing easy and efficient parallel implementation, yet achieves robustness and convergence rates comparable to line and plane relaxation multigrid algorithms.

The algorithm described here is a variant of Mulder's multigrid algorithm [5] for hyperbolic problems. The latter uses multiple coarse grids to achieve robustness, but is unsuitable for elliptic problems, since its V-cycle convergence rate goes to one as the number of levels increases. The new algorithm combines the contributions from the multiple coarse grids via a local "switch," based on the strength of the discrete operator in each coordinate direction. This improvement allows us to show that the V-cycle convergence rate is uniformly bounded away from one, on model anisotropic problems. Moreover, the new algorithm can be combined with the idea of concurrent iteration on all multigrid levels to yield a highly parallel algorithm for strongly anisotropic problems.

17. Key Words (Suggested by Author(s)) spectral methods, parallel multigrid	18. Distribution Statement 64 - Numerical Analysis Unclassified - Unlimited
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages 18	22. Price A03