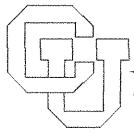
Recent Developments for the PSMG Multiscale Method

Paul O. Frederickson
Oliver A. McBryan

CU-CS-524-91



University of Colorado at Boulder
DEPARTMENT OF COMPUTER SCIENCE

Recent Developments for the PSMG Multiscale Method

Paul O. Frederickson, Oliver A. McBryan

CU-CS-524-91 March 1991

Department of Computer Science
University of Colorado at Boulder
Campus Box 430
Boulder, Colorado 80309-0430 USA

(303) 492-7514
(303) 492-2844 Fax
mcbryan@cs.colorado.edu

RECENT DEVELOPMENTS FOR THE PSMG MULTISCALE METHOD

Paul O. Frederickson [1]
RIACS,
NASA-Ames Research Center
Moffet Field, CA 94035

Oliver A. McBryan [2]
Dept. of Computer Science
University of Colorado
Boulder, CO 80309-0430

In this paper we discuss new developments for the PSMG multiscale method, which we have introduced previously as an efficient PDE solver for massively parallel architectures.

After an overview of the algorithm we introduce the fundamental multiscale recursion relation, as well as appropriate Fourier space notation. We derive the multiscale recursion as a single functional equation without reference to grids. We prove a sequence of rigorous convergence rate bounds which provide increasingly accurate estimates of the convergence rate for translation invariant problems. We show that in constant coefficient situations the convergence rates for the method may be derived to arbitrary precision, and we develop an efficient numerical scheme for computing such rates. Convergence rates are shown to be faster than reported previously. We present estimates for the normalized work involved in PSMG solution: the number of parallel arithmetic and communication operations required per digit of error reduction. The work estimates show that the algorithm is highly efficient.

---

1. Work supported by the NAS Systems Division of NASA and the Defense Advanced Research Project Agency via Cooperative Agreement NCC 2-387 between NASA and the University Space Research Association (USRA).
2. Work supported by the Air Force Office of Scientific Research, under grant AFOSR-89-0422

1. OVERVIEW

In many situations the most efficient algorithms for the numerical solution of large sparse elliptic problems are the various multigrid algorithms[1-3]. Usually these methods are able to compute a solution with N unknowns in $O(N)$ operations, asymptotically faster than most other algorithms. Efficient parallel implementations of multigrid algorithms have been reported on both SIMD and MIMD parallel computers[4-14]. Most of these methods are designed for moderately parallel systems, where each processor will hold many degrees of freedom.

The PSMG method, introduced in[15], provides a parallel algorithm appropriate to the case where the number of processors is comparable to the number of degrees of freedom. Related ideas have been introduced by Hackbusch[16]. As remarked in [17], PSMG is an example of an *intrinsically parallel algorithm.* It is highly efficient if sufficient processors are available, but is extremely inefficient on serial or low-parallelism computers. In situations where there are substantially more fine grid points than processors, an efficient approach might use a hybrid algorithm---using standard multigrid on the finest grids, but switching to PSMG on grid levels where the number of processors approximates or exceeds the number of grid points.

A brief but complete description of the PSMG algorithm is presented in Sections 2 and 3 below. Section 4 introduces the underlying multiscale recurrence relation which is used throughout the paper. Section 5 specializes to translation invariant problems and introduces notation in Fourier transform space. In section 6 we show that a single functional equation on the unit square is equivalent to the PSMG recursion formula across grid levels, greatly simplifying the analysis of PSMG convergence. Section 7 presents proofs of basic convergence rate bounds which show that the PSMG method actually converges, uniformly in the grid size. Section 8 discusses the numerical evaluation of convergence rates and develops an $O(N)$ algorithm for such evaluations. In section 9 we discuss methods for estimating the parallel work required by PSMG. Section 10 discusses actual PSMG performance, comparing the work per digit of convergence achieved across several different algorithms.

2. THE BASIC IDEA

Consider a simple discretization problem on a 1-dimensional grid. Standard multigrid techniques work with a series of coarser grids, each obtained by eliminating the odd numbered points of the previous grid. The error equation for the fine grid is then projected to the coarse grid at even numbered points, the coarse grid equation is solved approximately, and the error is interpolated back to the fine grid and added to the solution there. Finally a smoothing operation is performed on the fine grid. Recursive application of this procedure defines the complete multigrid procedure[1, 3].

The basic idea behind PSMG is the observation that for each fine grid there are two natural coarse grids - the even and odd points of the fine grid. (For simplicity we assume that periodic boundary conditions are enforced). Either of these coarse grids could be used at any point to construct the coarse grid solution, and both would presumably provide approximately equivalent quality solutions. Thus it ought to be possible to find a combination of the two solutions that is significantly better than either separately. It would follow immediately that such a scheme would converge faster (fewer iterations) than the corresponding standard multigrid scheme. Note that on a massively parallel machine the two coarse grid solutions may be solved simultaneously, in the same time as one of them would take - we assume here that the number of processors is comparable to the number of fine grid points. Both coarse grid problems are solved using the same set of machine instructions. Consequently the algorithm is well suited to SIMD parallel computers, as well as to MIMD machines.

The idea outlined above extends naturally to multi-dimensional problems. In d dimensions, 2^d coarse grids are obtained from a fine grid by selecting either the even or the odd points in each of the d coordinate directions. The fine grid solution is then defined by performing a suitable linear interpolation of all 2^d coarse grid points.

3. THE PSMG ALGORITHM

The PSMG algorithm works with a single grid of points $G^{(L)}$ of size $n = n(L) = 2^L$ in each dimension (called the level L grid, or the *fine* grid), but utilizes operators with different scales $l \leq L$ on that grid. Thus the algorithm is strictly speaking multiscale rather than multigrid. There are three basic operators: a finite difference operator A, an interpolation operator

Q and a smoothing operator $S = I - ZA$. All operators are periodic on the grid in each coordinate direction. The PSMG algorithm extends naturally to both Neumann and Dirichlet boundary conditions, with no increase in convergence rate. The simplest approach to implementing Neumann or Dirichlet boundary conditions is to use reflection or anti-reflection boundary conditions and an extended grid.

The operators at scale level l, denoted $A^{(l)}$, $Q^{(l)}$, and $Z^{(l)}$, couple points at a distance $d_l \equiv 2^{L-l}$. Each level l operator is defined at all points of the grid $G^{(L)}$. The basic steps involved at level l, $0<l\leq L$, for the solution of $A^{(L)}U = f$, starting with an initial guess u, with error e and residual r, are described by:

Algorithm PSMG(l,u,f):

1. Compute residual: $r = A^{(l)}e = f - A^{(l)}u$
2. Project residual to coarse grid: $r = r$ (trivial injection).
3. Solve coarse grid residual equation using PSMG: $e' = \mathbf{PSMG}(l-1,0,r)$
4. Interpolate to fine grid: $e'' = Q^{(l)}e'$
5. Apply a relaxation: $e''' = (I - Z^{(l)}A^{(l)})e'' + Z^{(l)}r$
6. Compute and return the new solution: $u''' = u + e'''$

An exact solver is utilized instead of **PSMG** on the coarsest grid in step 3. As indicated in step 3, the initial guess for all coarse grids is taken to be 0. Consequently steps 1 and 6 are required only on the finest grid. The process as described depends on an explicit choice for $Q^{(l)}$ and $Z^{(l)}$. The PSMG strategy is to choose $Q^{(l)}$ and $Z^{(l)}$ as functions of $A^{(l)}$ in such a way as to optimize the convergence rate of the above algorithm. Explicit choices for $Q^{(l)}$ and $Z^{(l)}$ are given in[15, 18], and in Section 10, for the cases where $A^{(l)}$ represents either the standard 5-point or Mehrstellen discretizations of the Laplacian.

4. THE PSMG RECURRENCE EQUATION

We note that a two-grid PSMG algorithm may be described in the form: $r^{(T)} = \mathbf{T}^{(L)} r$, where the two-grid residual reduction operator $\mathbf{T}^{(L)}$ is determined by:

$$\mathbf{T}^{(L)} = \mathbf{S}^{(L)}(I - A^{(L)} Q^{(L)} A^{(L-1)-1}) \quad .$$

We define the *two-grid convergence rate* τ of this iteration procedure as the quantity: $\tau = \sup_L ||\mathbf{T}^{(L)}||$. Clearly τ provides a bound on the convergence rate of the two-grid method on any grid.

We obtain an equation for the residual reduction operator of the full PSMG algorithm by recursive application of the two-grid algorithm described above. The corresponding residual reduction then takes the form: $r^{(M)} = \mathbf{M}^{(l)} r$, where the multi-grid iteration operator $\mathbf{M}^{(l)}$ is determined by:

$$\mathbf{M}^{(l)} = \mathbf{T}^{(l)} + (\mathbf{S}^{(l)} - \mathbf{T}^{(l)}) \mathbf{M}^{(l-1)} \quad , \quad l = L, \cdots, 1 \quad ,$$

with $\mathbf{M}^{(0)} \equiv 0$. We define the *multigrid convergence rate* of this procedure as the quantity:

$$\mu = \sup_L \sup_{l \le L} ||\mathbf{M}^{(l)}|| \quad .$$

Clearly μ provides a bound on the convergence rate of PSMG on any grid $G^{(L)}$.

5. FOURIER MODE ANALYSIS

In order to complete the description of the PSMG algorithm it is essential to define the operators $Q^{(l)}$ and $Z^{(l)}$ used for interpolation and smoothing. In this section, we describe suitable classes of $Q^{(l)}$ and $Z^{(l)}$ for the special case of an operator which has translation invariant coefficients. We will illustrate the ideas for the Poisson equation discretized on a periodic rectangular grid $G^{(L)}$ of $N = n \times n$ points, $n = 2^L$, which we label with the index $i = (i_1, i_2)$, $0 \le i_1, i_2 < n$. We will use two discretizations of the negative Laplacian $-\Delta$. The first of these is the standard five-point discretization defined by

$$(A_5^{(l)} u)_i = h_l^{-2} (4u_i - u_{i-e_1^l} - u_{i+e_1^l} - u_{i-e_2^l} - u_{i+e_2^l}) \quad ,$$

where e_i^l are integer vectors of length $d_l \equiv 2^{L-l}$ in the coordinate directions in index space, or alternatively by the familiar five-point star notation:

$$A_5^{(l)} = h_l^{-2} \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix} .$$

The second discretization we will study is the more accurate *Mehrstellen* discretization represented by the nine-point star

$$A_9^{(l)} = (6h_l^2)^{-1} \begin{bmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{bmatrix} .$$

Similarly, we will choose the operators $Q^{(l)}$ and $Z^{(l)}$ to be defined by simple symmetric three parameter nine-point star operators (with appropriate scale length depending on l):

$$Q^{(l)} = \begin{bmatrix} q_{11} & q_1 & q_{11} \\ q_1 & q_0 & q_1 \\ q_{11} & q_1 & q_{11} \end{bmatrix} , \quad Z^{(l)} = h_l^2 \begin{bmatrix} z_{11} & z_1 & z_{11} \\ z_1 & z_0 & z_1 \\ z_{11} & z_1 & z_{11} \end{bmatrix} ,$$

or equivalently in operator notation:

$$(Q^{(l)}u)_i = q_0 u_i + q_1 (u_{i+e_1^l} + u_{i-e_1^l} + u_{i+e_2^l} + u_{i-e_2^l}) +$$

$$q_{11} (u_{i+e_1^l+e_2^l} + u_{i-e_1^l+e_2^l} + u_{i+e_1^l-e_2^l} + u_{i-e_1^l-e_2^l}) ,$$

with a similar expression for $Z^{(l)}u$. For simplicity, we take the parameters q_i and z_i to be independent of the scale parameter l.

Since all of these operators are translation invariant, they are diagonalized by the discrete Fourier transform. The analysis of the PSMG algorithm then becomes particularly convenient. In the following we work entirely in Fourier transform space, where each of the operators $A^{(l)}$, $Q^{(l)}$ and $Z^{(l)}$ will reduce to multiplication by a trigonometric function. In terms of the two-dimensional discrete Fourier transform on $G^{(L)}$:

$$\hat{u_k} = n^{-1} \sum_{j_1,j_2=0}^{n-1} e^{i\,2\pi j\cdot k/n}\, u_j \ , \quad 0 \le k_1, k_2 < n \ ,$$

and the notation $x_i^{(l)} = \cos(2\pi k_i d_l/n)$, $i = 1,2$, the operators $A^{(l)}$ and $Q^{(l)}$ reduce to multiplication by the trigonometric polynomials:

$$A_5^{(l)} = h_l^{-2} (4 - 2(x_1^{(l)} + x_2^{(l)})) ,$$

$$A_9^{(l)} = (6h_l^2)^{-1} (20 - 8(x_1^{(l)} + x_2^{(l)}) - 4x_1^{(l)}x_2^{(l)}) ,$$

$$Q^{(l)} = q_0 + 2q_1(x_1^{(l)} + x_2^{(l)}) + 4q_{11}x_1^{(l)}x_2^{(l)} \ ,$$

with a similar expression for $Z^{(l)}$ in terms of parameters z_0, z_1 and z_{11}. Application of the two-grid iteration operator $\mathbf{T}^{(l)}$ over the grid $G^{(L)}$ reduces to multiplication by the function:

$$\mathbf{T}^{(l)}{}_k = \mathbf{S}^{(l)}{}_k \ \mathbf{C}^{(l)}{}_k \ ,$$

where $\mathbf{S}^{(l)}{}_k \equiv 1 - Z^{(l)}{}_k A^{(l)}{}_k$ and $\mathbf{C}^{(l)}{}_k \equiv 1 - Q^{(l)}{}_k A^{(l-1)}{}_k{}^{-1} A^{(l)}{}_k$ are the Fourier representations of the smoothing and coarse-scale correction operators, respectively.

For both the Laplace and Mehrstellen case, $\mathbf{C}^{(l)}{}_k$ has apparent poles at the four points $|x_1^{(l)}| = |x_2^{(l)}| = 1$, which are the zeroes of the coarse grid difference operator. The pole at $x_1^{(l)} = x_2^{(l)} = 1$ is canceled by a corresponding zero in the numerator, but this is not so for the other three poles. We cancel the remaining zeroes in the denominator by carefully choosing the three parameters q_i. It is easily checked that the two conditions:

$$q_1 = q_0/2 \ , \quad q_{11} = q_0/4 \ ,$$

suffice, leaving one free parameter q_0 in the Q operator to be chosen later.

A further constraint is necessary if the multigrid iteration operator $\mathbf{M}^{(l)}$ is to have a bound independent of l, namely the constraint that the coarse grid correction operator $\mathbf{C}^{(l)}$ vanish at the origin in frequency space. This constraint leads to the condition q_0=.25 on the interpolation operator $Q^{(l)}$ which is therefore uniquely determined. The resulting form of $\mathbf{C}^{(l)}$ is then:

$$\mathbf{C}^{(l)}{}_k = 1 - .5q_0 \, (1+x_1^{(l)})(1+x_2^{(l)}) \, \frac{2-x_1^{(l)}-x_2^{(l)}}{2-x_1^{(l)2}-x_2^{(l)2}} \ ,$$

for the five-point operator, with a similar expression for the nine-point operator. Note that the same restrictions on Q are required for the five and nine-point operators.

To get an improved convergence rate we have also used a 25-point star operator for Q:

$$Q = \begin{bmatrix} q_{22} & q_{12} & q_2 & q_{12} & q_{22} \\ q_{12} & q_{11} & q_1 & q_{11} & q_{12} \\ q_2 & q_1 & q_0 & q_1 & q_2 \\ q_{12} & q_{11} & q_1 & q_{11} & q_{12} \\ q_{22} & q_{12} & q_2 & q_{12} & q_{22} \end{bmatrix} \ .$$

Again we compute an explicit rational function expression for the coarse-scale operator $\mathbf{C}_k^{(l)}$ as

a function of the trigonometric variables $x_i^{(l)}$. As before, there are three poles of the denominator in $\mathbf{C}_k^{(l)}$ which we cancel by careful choice of Q, leading to the two constraints:

$$0 = q_0 - 4q_1 + 4q_{11} + 4q_2 + 4q_{22} - 8q_{12} \ ,$$

$$0 = q_0 - 4q_{11} + 4q_2 + 4q_{22} \ .$$

We must add the constraint

$$1 = q_0 + 4q_1 + 4q_{11} + 4q_2 + 4q_{22} + 8q_{12} \ ,$$

required to ensure that the coarse grid correction operator $\mathbf{C}^{(l)}$ vanishes at the origin in frequency space. It follows that there are 3 independent coefficients of Q, along with the 3 parameters of Z, that are available to minimize μ in the multigrid analysis.

In the translation invariant case all of the operators in the recurrence equations for the multigrid residual reduction operator $\mathbf{M}^{(l)}$ introduced in the previous section commute. Furthermore $\mathbf{M}^{(l)}$ is also then the error reduction operator. $\mathbf{M}_k^{(l)}$ is therefore determined recursively as a sum of products of low-degree rational functions in the $x_i^{(l)}$. Since $\mathbf{M}^{(l)}$ is a multiplication operator in the translation invariant case, its norm is the maximum value of $|\mathbf{M}_k^{(l)}|$ evaluated over all relevant frequencies k_i, or equivalently, evaluated over those discrete points in the square $-1 \le x_1^{(l)}, x_2^{(l)} \le 1$ that correspond to Fourier frequencies on the grid $G^{(L)}$.

Translation invariant problems will have an inherent singularity if constant functions are in the null space of the differential operator A. In Fourier space this results in bad behavior at the origin k=(0,0). The singularity must be dealt with. One solution is to omit constant functions and zero frequency from the domain of interest. Alternatively one may include these functions by using the Moore-Penrose pseudo-inverse[19, 20] to extend operators from the space of non-constant data. This latter choice is equivalent to defining $\mathbf{M}^{(l)}(0,0) = 1$ in the basic recurrence relation.

6. A FUNCTIONAL EQUATION FOR *M*

A remarkable feature of PSMG for translation invariant PDE without lower order terms (such as the Poisson equation) is that we can phrase the defining equation for the iteration in a form completely independent of grids or grid levels. The resulting equation is a simple

functional equation on the space of trigonometric polynomials. The key point in deriving this relationship is to make a transition from unbounded frequency variables (k) to uniformly bounded angle (θ) or cosine ($x = \cos(\theta)$) variables.

In previous sections, all operators have been explicitly shown as dependent on the grid $G^{(L)}$ and the level l, as well as on frequency k_i. We begin with a lemma which shows that the kernels of these operators are actually independent of L and need not be computed at even frequencies k.

Lemma 1: Each of the operators $O = \mathbf{S}^{(l)}, \mathbf{T}^{(l)}, \mathbf{M}^{(l)}$ defined on $G^{(L)}$ has a kernel $O_k^{(l)}$ independent of L, and satisfying $O_{2k}^{(l)} = O_k^{(l-1)}$.

Proof: From section 5 we note that the only explicit l or L dependence in the operators **S**, **T** is in the trigonometric functions $x_k^{(l)} = \cos(2\pi k/2^l)$. Factors of h_l involved in the operators A and Z always cancel out provided there are no first order or constant terms in the difference operator. Thus the operator kernels depend explicitly on l but not on L. Furthermore we note that $x_{2k}^{(l)} = \cos(2\pi\, 2k/2^l) = x_k^{(l-1)}$, and consequently we have the result of the lemma for **S** and **T**.

We prove the result for **M** by induction on l. For l=1 we apply the recurrence along with $\mathbf{M}^{(0)} \equiv 1$, obtaining $\mathbf{M}^{(1)} = \mathbf{S}^{(1)}$, and so the lemma applies for l=1. Now assume the lemma applies to $\mathbf{M}^{(l-1)}$. Then

$$\begin{aligned} \mathbf{M}_{2k}^{(l)} &= \mathbf{T}_{2k}^{(l)} + (\mathbf{S}_{2k}^{(l)} - \mathbf{T}_{2k}^{(l)})\mathbf{M}_{2k}^{(l-1)} \\ &= \mathbf{T}_k^{(l-1)} + (\mathbf{S}_k^{(l-1)} - \mathbf{T}_k^{(l-1)})\mathbf{M}_k^{(l-2)} \\ &= \mathbf{M}_k^{(l-1)} \ . \end{aligned}$$

This completes the induction proof.

As a consequence of the lemma, we may write the recurrence as a single-level functional equation:

$$\mathbf{M}_k^{(l)} = \mathbf{T}_k^{(l)} + (\mathbf{S}_k^{(l)} - \mathbf{T}_k^{(l)})\mathbf{M}_{2k}^{(l)} \ .$$

We now regard $\mathbf{M}^{(l)}$ as a function of the normalized variables $x_i = \cos(2\pi k_i/n)$ rather than of k_i. These variables span the range $[-1,1]$ as k_i and n vary. Alternatively we may introduce the variables $\theta_i \equiv 2\pi k_i/n$ and regard $\mathbf{M}$ as a function $\mathbf{M}(\theta)$ on $[0,2\pi]\times[0,2\pi]$. We denote by D_θ the set of values (θ_1,θ_2) where $0 \le k_i < n$, for any n of the form 2^l, and where the origin, $k_1=k_2=0$, is omitted. We denote the corresponding set of values of x by D_x. Thus all of the values of $\mathbf{M}_k^{(l)}$, for all grids $G^{(L)}$, are values of a single function $\mathbf{M}(\theta)$, $\theta=(\theta_1,\theta_2)$ in D_θ, or equivalently of a single function $\mathbf{M}(x)$, $x=(x_1,x_2)$ in D_x. Similarly the smoothing and two grid kernels extend into functions $\mathbf{S}(\theta)$ and $\mathbf{T}(\theta)$ defined over D_θ, or $\mathbf{S}(x)$ and $\mathbf{T}(x)$ defined over D_x, and the multiscale recursion relation reduces to the form:

$$\mathbf{M}(\theta) = \mathbf{T}(\theta) + (\mathbf{S}(\theta)-\mathbf{T}(\theta))\,\mathbf{M}(2\theta)\,, \quad \mathbf{M}(0) = 1\,,$$

or equivalently

$$\mathbf{M}(x) = \mathbf{T}(x) + (\mathbf{S}(x)-\mathbf{T}(x))\,\mathbf{M}(2x^2-1)\,, \quad \mathbf{M}(1) = 1\,,$$

where we have used the fact that $\cos(2\theta) = 2\cos^2(\theta) - 1$. We will use these functional equations in the following section. The choice of $\mathbf{M}(1) = 1$ was explained at the end of the previous section, and effectively defines the Moore-Penrose pseudo-inverse on constant functions.

7. CONVERGENCE RATE BOUNDS: STATEMENT AND PROOFS

In this section we will derive rigorous upper bounds on the multigrid convergence rate of the PSMG algorithm in the translation invariant case described in the previous section.

Theorem 1: Define $\mu_1 \equiv \sup_x |\mathbf{T}(x)| \,/\, (1 - |\mathbf{S}(x) - \mathbf{T}(x)|)$. Assume that the smoother satisfies $|\mathbf{S}| \le \mu_1$ at the three high frequency points $(-1,1)$, $(1,-1)$ and $(-1,-1)$. Then the multigrid convergence rate μ satisfies the bound $\mu \le \mu_1$.

Proof: Let P_3 denote the set $\{(1,-1), (-1,1), (-1,-1)\}$. We will use a mapping $g: R^2 \to R^2$ defined by $g(x_1,x_2) = (2x_1^2-1, 2x_2^2-1)$. We decompose D_x into a union of disjoint subsets D_n, $n \geq 0$, where we define recursively:

$$D_n = \{x \in D_x : x \text{ not } \in D_{n-1} \text{ and } g(x) \in D_{n-1}\} \ , \quad D_0 \equiv P_3 \ .$$

D_n consists of the points of D_x which correspond to angles in D_0 of the form $2\pi k/2^n$ with k odd. We will prove Theorem 1 holds for the sets D_n by induction on n.

The induction hypothesis holds for $D_0 \equiv P_3$. To see this we note that for each point $p \in P_3$, $g(p) = (1,1)$ and since $M(1,1) = 1$, the recursion relation reduces at these points to

$$\mathbf{M}(p) \ = \ \mathbf{T}(p) + (\mathbf{S}(p) - \mathbf{T}(p)) \cdot 1 \ = \ \mathbf{S}(p) \ , \ p \in P_3 \ .$$

Thus by the assumption of the theorem that $|\mathbf{S}(p)| \leq \mu_1$ for $p \in P_3$, we conclude that $|\mathbf{M}(p)| \leq \mu_1$, and thus that the induction hypothesis holds for D_0.

Now assume the result is true for D_{n-1} and let $x \in D_n$. Since $g(x) \in D_{n-1}$, it follows that:

$$\begin{aligned} |\mathbf{M}(x)| &\leq |\mathbf{T}(x)| + |\mathbf{S}(x) - \mathbf{T}(x)| \, |\mathbf{M}(2x^2-1)| \\ &\leq (1 - |\mathbf{S}(x) - \mathbf{T}(x)|)\mu_1 + |\mathbf{S}(x) - \mathbf{T}(x)| \mu_1 \\ &= \mu_1 \ . \end{aligned}$$

which completes the proof of the induction hypothesis for D_n.

Remarks: The assumption on **S** in the statement of Theorem 1 is a requirement that **S** be small at high frequencies, which is essential anyway if **S** is to be a good smoother. Thus it is not a major restriction. However the assumption can be avoided by terminating the PSMG iteration at the level 1 grid rather than at level 0. In that case an exact solution is performed on level 1, so that **M** is 0 on the set P_3. We also remark that the theorem does not require that the differential operator be homogeneous, as required to derive the grid independent recursion relation. The theorem is easily extended by the same proof to the case where $\mathbf{M}(x)$ is dependent on the level l. The theorem does however require translation invariance.

The bound described above was used in[15] to design good choices for Q and Z. In fact the bound μ_1 requires only that we find the suprenum of a rational function on the unit square. One can estimate μ_1 by evaluation of the suprenum over a large grid of values in the unit square. The resulting quantity is then used as an objective function and minimized with respect to the parameters defining Q and Z.

Sharper convergence criteria result if one first applies the iteration formula for $\mathbf{M}(x)$ several times. It is notationally convenient to write $\mathbf{M}$ as a function of the angles $0 \le \theta_i \equiv 2\pi k_i/n < 2\pi$. We also define $\mathbf{R}(\theta) \equiv \mathbf{S}(\theta)-\mathbf{T}(\theta)$. The recurrence relation then reduces to:

$$\mathbf{M}(\theta) = \mathbf{T}(\theta) + \mathbf{R}(\theta)\mathbf{M}(2\theta)\,, \quad \mathbf{M}(\mathbf{0}) = 1\ .$$

Iterating $\nu-1$ times we arrive at:

$$\begin{aligned}\mathbf{M}(\theta) \;=\; \mathbf{T}(\theta) \;+\; \mathbf{R}(\theta)\mathbf{T}(2\theta) \;+\; \mathbf{R}(\theta)\mathbf{R}(2\theta)\mathbf{T}(4\theta) \;+\; \cdots \\ +\; \mathbf{R}(\theta)\mathbf{R}(2\theta)\cdots\mathbf{R}(2^{\nu-1}\theta)\mathbf{T}(2^{\nu}\theta) \;+\; \mathbf{R}(\theta)\mathbf{R}(2\theta)\cdots\mathbf{R}(2^{\nu}\theta)\mathbf{M}(2^{\nu+1}\theta) \\ \equiv\; \mathbf{T}_\nu(\theta) \;+\; \mathbf{R}_\nu(\theta)\mathbf{M}(2^{\nu+1}\theta)\ .\end{aligned}$$

This again has a similar form to the original recurrence equation, and we prove exactly as in the previous theorem:

Theorem 2: Define $\mu_\nu \equiv \sup_\theta |\mathbf{T}_\nu(\theta)| \,/\, (1 - |\mathbf{R}_\nu(\theta)|)$. Assume that the smoother satisfies $|\mathbf{S}| \le \mu_\nu$ at the three high frequency points $(-1,1)$, $(1,-1)$ and $(-1,-1)$. Then the convergence rate satisfies the bound $\mu \le \mu_\nu$.

It is important to take care in evaluating the bounds μ_ν. Even if $\mathbf{R}_\nu(\theta)$ approaches 1 at some points, the bound may still be useful provided the numerator $\mathbf{T}_\nu(\theta)$ also vanishes at the same points. Indeed this situation occurs at the origin even for μ_1. In evaluating $|\mathbf{T}_\nu|$ it is also important to take the absolute value only after the various terms have been added together.

These convergence results may be extended to provide various bounds on derivatives of $\mathbf{M}(\theta)$, and in particular to show that with appropriate conditions imposed, μ may be approximated arbitrarily closely by taking the supremum of $|\mathbf{M}(\theta)|$ on a suitably fine mesh of points θ. For further details we refer to[21].

For illustration we provide here results obtained on grids with edges of size $n(L)$ from 4 through 1024 using each of the bounds μ_1,.., μ_5. The final column is the exact norm as computed using the full iteration formula. The results are for the case of a Mehrstellen operator with 9-point Q and Z stencils - the coefficients of Q and Z are given in the last section of the paper under PSMG9-9.

TABLE 1: CONVERGENCE BOUNDS μ_ν

n(L)	μ_1	μ_2	μ_3	μ_4	μ_5	μ
4	.0009	.0049	.0049	.0049	.0049	.0049
8	.0248	.0174	.0174	.0174	.0174	.0174
16	.0264	.0216	.0216	.0216	.0216	.0217
32	.0291	.0230	.0216	.0216	.0216	.0217
64	.0291	.0241	.0216	.0216	.0216	.0217
128	.0292	.0241	.0225	.0216	.0216	.0217
256	.0292	.0243	.0224	.0220	.0217	.0217
512	.0292	.0243	.0226	.0220	.0219	.0217
1024	.0292	.0243	.0226	.0221	.0219	.0217

8. COMPUTATION OF CONVERGENCE RATES

In our paper[15] we have used the first theorem of the previous section to provide upper bounds on PSMG convergence rates and to guide thereby, optimization processes. Basing an optimization process for a rate on optimization of a possibly crude upper bound for that rate is certainly not ideal. More recently[18], we have improved our knowledge of PSMG convergence rates by *exactly* computing the spectral radius $\mu^{(L)}$ of the self-adjoint PSMG multigrid

error reduction operator $\mathbf{M}^{(L)}$ for all grids $G^{(L)}$ with L ranging from 0 to 11. Here $G^{(L)}$ is a square grid with $n \equiv 2^L$ points on a side. The verification is based on the iterative formula:

$$\mathbf{M}^{(l)} = \mathbf{T}^{(l)} + (\mathbf{S}^{(l)} - \mathbf{T}^{(l)})\, \mathbf{M}^{(l-1)}, \quad 1 \le l \le L,$$

for the multigrid operator $\mathbf{M}^{(l)}$. The correct domain for the translation invariant Poisson equation is the set of grid functions orthogonal to constants. In Fourier space this means that points aliased to (0,0) are omitted from the computation. Equivalently one defines $\mathbf{M}^{(0)} \equiv 1$ in (1).

We compute $\mu^{(L)}$ by evaluating the recurrence above from l=0 to l=L in Fourier space for every frequency pair k_1, k_2 appropriate to $G^{(L)}$ - i.e. for $0 \le k_i < n$. The only approximation in this procedure is that the kernels are evaluated in double precision rather than infinite precision arithmetic.

At first sight it would appear that the evaluation of $\mu^{(L)}$ would require $O(n^2 \log L)$ operations - the number of grid points for level L times the cost $O(\log L)$ of an evaluation of the recurrence. Remarkably, in Fourier space only $O(n^2)$ operations are required. The idea is to use the previously derived fact that $\mathbf{M}_{2k}^{(l)} = \mathbf{M}_k^{(l-1)}$. The value of $\mathbf{M}_k^{(l)}$ is independent of which grid level l it is evaluated at (assuming frequency k exists at that level). Therefore we develop an inverse process where we first evaluate $\mathbf{M}$ at all coarse grid points, then evaluate $\mathbf{M}$ on the next finer grid but only at those points not in the coarse grid. Furthermore one can use symmetry in k_1, k_2 and between $k_i \le n/2$ and $k_i \ge n/2$ to reduce the work by a further factor of 8. We store all evaluated $\mathbf{M}_k$ on a given level, so that the recursion on the following level can be terminated after one step. It is this last fact that reduces the overall cost by $\log L$. Given that the result of a norm evaluation is a single data point in the optimization of the choice of Q and Z, the ability to have such a fast evaluation of the norm of $\mathbf{M}$ is the key to obtaining highly efficient PSMG processes in both 2D and 3D.

9. PARALLEL OPERATION COUNTS

In this section we discuss methods for estimating the work performed in a PSMG iteration. We use a model of parallel computation introduced in[22]. Thus we assume only nearest neighbor communication with four neighbors, purely SIMD communication and computation, and we consider separately each grid level. The operation counts will be those for intermediate

grids and the communication unit is grid level dependent - the cost for communication between "nearest neighbors" at that grid level.

The PSMG algorithm was described as 6 separate steps in section 3 above to which we now refer. For intermediate grids ($0<l<L$) the initial guess is taken to be 0 so that step 1 is not needed. Similarly step 6 is relevant only to the top level grid L. Step 2 is free because PSMG uses injection, while step 3 is counted at level $l-1$ or lower. Thus only steps 4-5 are to be counted at level l. Apart from the 3 operators involved, we note that no communication is required in these steps while two computations are required in step 5. Therefore we conclude that at intermediate grid levels ($0<l<L$), the parallel communication and computation costs for the PSMG algorithm are given by:

$$comm\,(PSMG\,(l)) = comm\,(Q) + comm\,(A) + comm\,(Z) \quad ,$$
$$comp\,(PSMG\,(l)) = comp\,(Q) + comp\,(A) + comp\,(Z) + 2 \quad ,$$

where $comm\,(O)$ and $comp\,(O)$ are the number of parallel communication and computation operations required to execute operator O.

In Table 2 below we present the operation counts for each of the individual operators encountered in the four PSMG algorithms corresponding to choice of 5 or 9 point A and to 9 or 25 point Q stencils (often referred to as the PSMG5-9, PSMG9-9, PSMG5-25 and PSMG9-25 methods). For full details of the derivation of these operation counts we refer to[18].

TABLE 2: OPERATION COUNTS FOR A, Q, Z

Operator	Comp. Steps	Comm. Steps
5-pt A (Laplacian)	3	4
9-pt A (Laplacian)	5	4
9-pt Q (Interpolation)	4	4
25-pt Q (Interpolation)	12	8
9-pt Z (Relaxation)	5	4

10. PSMG PERFORMANCE

An effective measure of algorithm performance depends on several factors besides convergence rate. For parallel environments, measures of parallel work and of communication need to be incorporated. Furthermore one needs to be precise about the underlying model machine specification. As a model of parallel computation we have followed[22], as discussed in the previous section.

If the asymptotic convergence rate of a method is ρ and the method requires w parallel operations per iteration, the normalized operation count is defined as $w/\log_{10}\rho$, and measures the parallel work required per grid level to reduce the error by a factor of 10. We consider parallel communication and computation operations separately.

For several PSMG methods we present asymptotic convergence rates, the number of parallel arithmetic and communication operations required on each grid per iteration, and also the normalized operation count for arithmetic and communication. We summarize the results for several simple cases in Table 3. For each cycle we use a single relaxation with a 9-point Z stencil. The cases are labeled PSMGa-q where a and q are the stencil size (5 or 9) of the difference operator a and of the q matrix (9 or 25) respectively.

TABLE 3: PSMG CONVERGENCE RATES

	Convergence	Steps per Level		Normalized Steps	
Method	**Rate**	**Comp.**	**Comm.**	**Comp.**	**Comm.**
PSMG 5-9	.08867	14	12	13.31	11.40
PSMG 5-25	.02504	22	16	13.74	9.99
PSMG 9-9	.02165	16	12	9.61	7.21
PSMG 9-25	.00165	24	16	8.62	5.75

The corresponding coefficients for the interpolation operator Q and the smoothing operator Z are (in the notation of[18]):

PSMG5-9:	q_0=.25	q_1=.125	q_{11}=.0625
	z_0=.278079	z_1=.0534577	z_{11}=.0125615
PSMG5-25:	q_0=.361017	q_1=.11458	q_{11}=.0625
	q_2=-.0309162	q_{12}=.00521024	q_{22}=.00316188
	z_0=.361452	z_1=.0891718	z_{11}=.0293793
PSMG9-9:	q_0=.25	q_1=.125	q_{11}=.0625
	z_0=.300589	z_1=.0432465	z_{11}=.0139994
PSMG9-25:	q_0=.34152	q_1=.0995677	q_{11}=.0625
	q_2=-.0199225	q_{12}=.0127161	q_{22}=-.00295755
	z_0=.283286	z_1=.0323815	z_{11}=.00835795

The convergence rates given in Table 3 are the maximum values of $\mu^{(L)}$ for $0 \le L \le 11$ and therefore bound the exact convergence rates for all grids up to size 4 million points. In practice we find that the convergence rates $\mu^{(L)}$ are unchanged to several digits of precision beyond about level L=6. We therefore are confident, that the convergence rates in Table 3 extend to arbitrary numbers of grid levels. As a final check we have solved the Poisson equation using PSMG, with zero right hand side and a random initial guess, and in each case verified the convergence rates of Table 3.

As a comparison point we note that the PSMG9-25 algorithm uses about half as much computation and close to one fifth as much communication as the best standard red-black algorithms[18, 22], assuming of course that there are about as many processors as fine grid points.

References

1. A. Brandt, "Multi-level adaptive solutions to boundary-value problems," *Math. Comp.*, vol. 31, pp. 333-390, 1977.
2. W. Hackbusch, "Convergence of multi-grid iterations applied to difference equations," *Math. Comp.*, vol. 34, pp. 425-440, 1980.
3. K. Stüben and U. Trottenberg, "Multigrid Methods: Fundamental algorithms, model problem analysis and applications," in *Multigrid Methods*, ed. U. Trottenberg, Lecture Notes in Mathematics 960., pp. 1-176, Springer-Verlag, Berlin, 1981,.
4. A. Brandt, "Multi-Grid Solvers on Parallel Computers," ICASE Technical Report 80-23, NASA Langley Research Center, Hampton Va., 1980.
5. D. Gannon and J. van Rosendale, "Highly Parallel Multi-Grid Solvers for Elliptic PDEs: An Experimental Analysis," ICASE Technical Report 82-36, NASA Langley Research Center, Hampton Va., 1982.
6. O. McBryan and E. Van de Velde, "Parallel Algorithms for Elliptic Equation Solution on the HEP Computer," *Proceedings of the Conference on Parallel Processing using the Heterogeneous Element Processor, March 1985*, University of Oklahoma, March 1985.
7. O. McBryan and E. Van de Velde, "Parallel Algorithms for Elliptic Equations," *Commun. Pure and Appl. Math.*, vol. 38, pp. 769-795, 1985.
8. C. Thole, "Experiments with Multigrid Methods on the Caltech Hypercube," GMD-Studien Nr. 103, Gesellschaft fur Mathematik und Datenvararbeitung, St. Augustin, West Germany, 1986.
9. O. McBryan and E. Van de Velde, "The Multigrid Method on Parallel Processors," in *Multigrid Methods II*, ed. W. Hackbusch and U. Trottenberg, Lecture Notes in Mathematics, vol. 1228, Springer-Verlag, Berlin, Heidelberg, 1986.
10. O. McBryan and E. Van de Velde, *Hypercube Algorithms and Implementations*, SIAM J. Sci. Stat. Comput., 8, pp. 227-287, 1987.
11. T. F. Chan, Y. Saad, and M. H. Schultz, "Solving elliptic partial differential equations on hypercubes," *Hypercube Multiprocessors 1986*, SIAM, Philadelphia, 1986.
12. P. O. Frederickson and M. W. Benson, "Fast parallel algorithms for the Moore-Penrose pseudo-inverse," *Hypercube Multiprocessors 1986*, SIAM, Philadelphia, 1986.
13. O. McBryan, "The Connection Machine: PDE Solution on 65536 Processors," *Parallel Computing*, vol. 9, pp. 1-24, North-Holland, 1988.
14. O. McBryan, P. Frederickson, J. Linden, A. Schuller, K. Solchenbach, K. Stüben, C.-A. Thole, and U. Trottenberg, "Multigrid Methods on Parallel Computers - a Survey on Recent Developments," *University of Colorado Tech. Report CU-CS-504-90, and*

IMPACT of Supercomputing, 1991, to appear.

15. P. O. Frederickson and O. McBryan, "Parallel Superconvergent Multigrid," in *Multigrid Methods: Theory, Applications and Supercomputing*, ed. S. McCormick, Math Applications Series, vol. 110, pp. 195-210, Marcel-Dekker Inc., New York, 1988.
16. W. Hackbusch, "A New Approach to Robust Multigrid Methods," Proceedings of the 1987 ICIAM Conference, 1987.
17. P. O. Frederickson and O. McBryan, "Intrinsically Parallel Multiscale Algorithms for Hypercubes," in *Proceedings of 3rd Conference on Hypercube Concurrent Computers and Applications, Vol II.*, ed. G. Fox, pp. 1726-1734, ACM Press, 1988.
18. P. Frederickson and O. McBryan, "Normalized Convergence Rates for the PSMG Method," *and SISSC*, vol. 12, no. 1, pp. 221-229, 1991.
19. E. H. Moore, "On the reciprocal of the general algebraic matrix," *Bull. Amer. Math. Soc.*, vol. 26, pp. 394-395, 1919.
20. R. Penrose, "A Generalized Inverse for Matrices," *Proc. Cambridge Philos. Soc.*, vol. 51, pp. 406-413, 1955.
21. P. Frederickson and O. McBryan, "Convergence Bounds for the PSMG Method," University of Colorado CS Dept Technical Report, 1991, to appear.
22. N. Decker, "A Note on the Parallel Efficiency of the Frederickson-McBryan Multigrid Algorithm," *SISSC*, vol. 12, no. 1, pp. 208-220, 1991.