

Wavelet Techniques Applied to Multigrid Methods

Doreen De Leon
Department of Mathematics
University of California
Box 951555
Los Angeles, CA 90095-1555
e-mail: ddeleon@math.ucla.edu
<http://www.math.ucla.edu/~ddeleon>

November 28, 2001

Abstract

The standard multigrid procedure performs poorly or may break down when used to solve certain problems, like elliptic problems with discontinuous or highly oscillatory coefficients. Here, we take the approach of using a wavelet transform and Schur complements to obtain coarse grid, interpolation, and restriction operators. ILU(0) and compression are used to improve the efficiency of the resulting method. Numerical examples are presented.

1 Introduction

The multigrid method is very useful in increasing the efficiency of iterative methods used to solve systems of algebraic equations approximating partial differential equations. However, when confronted by certain problems, for example problems with discontinuous or highly oscillatory coefficients as well as advection-dominated problems, the standard multigrid procedure converges slowly, with a rate dependent on mesh size, or may break down.

One method to correct for this for elliptic problems with periodic coefficients is through use of homogenization (e.g., [11, 10, 15]). This approach is taken because the homogenized operator provides a very good approximation of the important properties (eigenvalues, eigenfunctions) of the original fine grid operator. A problem with these homogenization techniques is that they are only applicable to periodic problems and can only be given closed form solutions in certain cases. Also, no natural definition of the restriction and interpolation operators follows from the homogenized coarse grid operator. More recently, Andreas Rieder [16] has used wavelet decompositions to obtain a multilevel method. His approach uses a choice of the filter operators obtained from wavelets for the restriction and interpolation operators.

Another problem analyzed in this paper is the advection-diffusion equation with dominant advection term. Many multigrid solutions to this problem appearing in the literature involve numbering and solving for the unknowns in a certain order (e.g., [1, 22]). M. Griebel and S. Knapek use matrix-dependent interpolations in [12], where the coarse grid operator, using the Galerkin approach, is determined to be a Schur complement.

In [9], the wavelet transform, which involves both high- and low-pass filter operators, is used to derive a new approach, under the assumption that the matrix on the fine grid is symmetric. Also,

some one-dimensional examples are examined. The goal of the present work is to extend the results of this approach to two dimensions, dropping the assumption of a symmetric fine grid operator. We consider this approach for several reasons.

First, we see from, say, [8] that the wavelet coarse grid operator provides a good approximation to the homogenized coarse grid operator, and it has a natural connection to the interpolation and restriction operators. Furthermore, wavelets can be applied to problems with periodic as well as non-periodic coefficients. Finally, the application of wavelet operators to vectors and matrices maintains the properties of the original problem.

The initial procedure followed is much the same as that in [9]. Removing the restriction of a symmetric fine grid operator has allowed us to apply the wavelet multigrid method to more, and different types of, problems than were examined in [9]. In Section 2, we discuss some multigrid background. Section 3 discusses wavelets, in both one and two dimensions, as background for later sections. Section 4 discusses the application of the wavelet transform, defined in Section 3, to multigrid methods. A theorem demonstrating the compressibility of the resulting coarse grid, interpolation, and restriction operators is also presented. In Section 5, we discuss the use of ILU(0) to obtain the LU decomposition of the component to be inverted followed by truncation to obtain sparser versions of the inverse component. Section 6 presents some numerical results of applying the wavelet multigrid method to a variety of problems, including advection-dominated equations and the checkerboard problem. We demonstrate the rapid convergence, independent of mesh size, of the wavelet multigrid method for these problems. We also examine the application of the wavelet multigrid method to anisotropic diffusion problems. In this case, semicoarsening is used instead of full coarsening, giving a convergence rate that is rapid, with a rate comparable to the algebraic multigrid method (AMG1R5, an earlier version of which is described in [17]). Application of the wavelet multigrid method to a reformulated version of the Stokes equations also yields good convergence results. For all of the numerical results in this paper, the two-level V-cycle method is used with one Gauss-Seidel iteration for both the coarsening and the correction phases, unless otherwise specified.

2 Multigrid

The problem we are concerned with solving is the system of linear equations

$$Au = b, \tag{1}$$

where A and b arise from discretization of a differential equation on some grid Ω^h , where h represents the step size.

For notational purposes, we briefly describe the V-cycle method used in this paper. Given some interpolation operator, I_{2h}^h , where the superscript refers to the fine grid and the subscript refers to the coarse grid, and a restriction operator, I_h^{2h} , we can define a multigrid method recursively. For the two-level method, first relax a few (usually one or two) steps on the fine grid Ω^h to get an initial guess u^h . Then, compute the residual $r^h = b^h - A^h u^h$, restrict the residual to the coarse grid $\Omega^{2h} : r^{2h} = I_h^{2h} r^h$, and solve the residual equation

$$A^{2h} e^{2h} = r^{2h}$$

on the coarse grid. Then, set $u^h = u^h + I_{2h}^h e^{2h}$ and relax again a few steps on the fine grid (usually one or two steps). This describes the two-level method. Based on this, we define the V-cycle multigrid scheme recursively. Some good references for multigrid methods are [3, 13, 21].

Another type of multigrid scheme is algebraic multigrid, which only uses the structure of the matrix in the problem to determine the coarsening process (choice of coarse grid and definition of interpolation/restriction operators). This process is performed in order to ensure that the range of interpolation approximates the errors not sufficiently reduced via relaxation. For a more detailed description of algebraic methods, see, e.g., [17, 4, 14, 19, 21]. Algebraic multigrid methods are of particular interest to us, in that they are the nearest methods to the approach taken in this paper. Note that in [4] the relation between algebraic multigrid and Schur complements is discussed. An updated version of the algebraic multigrid method discussed in [17] is one of the bases of comparison used in this paper.

It is expected that V-cycle schemes should converge at a rate independent of the mesh size. However, for certain problems, including problems with discontinuous or highly oscillatory coefficients or for advection-dominated problems, the above is no longer true. One difficulty is that the small eigenvalues of A may not be associated with smooth eigenfunctions, a key assumption for the standard multigrid method. For such problems, it is not as simple to approximate the smooth eigenfunctions on the coarse grids. New methods for restriction and prolongation (interpolation) or for treating the entire problem must be found.

3 Wavelets

For notational purposes, a brief description of wavelets follows. For more details, please refer to [5, 6]. Wavelets basically separate data (or functions or operators) into different frequency components and analyze them by scaling. We can choose the wavelets to form a complete orthonormal basis of $L^2(\mathbb{R})$. And, due to the scaling of the wavelet functions, they have time- or space-widths that are related to their frequency – at high frequencies, they are narrow, and at low frequencies, they are broader. Therefore, they provide good localization of functions in both the frequency domain and physical space, and representation by wavelets seems natural to apply to analysis of fine and coarse scales.

Basically, a multiresolution analysis (MRA) consists of a sequence of closed subspaces V_j of $L^2(\mathbb{R})$, the scaling spaces, that satisfy certain conditions. For every $j \in \mathbb{Z}$, let W_j be the orthogonal complement of V_j in V_{j-1} . W_j are called the wavelet spaces. Define H_j and G_j to be the operators that transform the basis of the space V_j to the bases of the spaces V_{j+1} and W_{j+1} , respectively. Now, define $\mathcal{W}_j : V_j \rightarrow V_{j+1} \oplus W_{j+1}$. Then, by definition,

$$\mathcal{W}_j = \begin{pmatrix} H_j \\ G_j \end{pmatrix}.$$

Note that \mathcal{W}_j is orthogonal due to the properties of H_j and G_j .

We would like to point out that the discrete wavelet operators are computationally efficient. With respect to the Haar multiresolution analysis, application of the low-frequency operator (H_j) to an $n \times n$ matrix involves only $2n$ operations. The same holds for the high-frequency operator (G_j). So, the application of the wavelet transform requires only $4n$ operations. In general, application of the wavelet transform requires $\mathcal{O}(n)$ operations, assuming a finite number of coefficients for the low- and high-frequency operators.

In two dimensions, we use the tensor product of one-dimensional multiresolution analyses. Define \mathbf{V}_j by $\mathbf{V}_j = V_j \otimes V_j$ (tensor product). Then, the \mathbf{V}_j 's form an MRA in $L^2(\mathbb{R}^2)$. Now, for

each $j \in \mathbb{Z}$, we define \mathbf{W}_j to be the orthogonal complement of \mathbf{V}_j in \mathbf{V}_{j-1} . So,

$$\begin{aligned}\mathbf{V}_{j-1} &= V_{j-1} \otimes V_{j-1} \\ &= (V_j \otimes V_j) \oplus [(W_j \otimes V_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes W_j)] \\ &= \mathbf{V}_j \oplus \mathbf{W}_j.\end{aligned}\tag{2}$$

Then, analogous to the one dimensional case, we define \mathbf{H}_j to be

$$\mathbf{H}_j = H_j^y H_j^x$$

and \mathbf{G}_j to be

$$\mathbf{G}_j = \begin{pmatrix} G_j^y H_j^x \\ H_j^y G_j^x \\ G_j^y G_j^x \end{pmatrix}.$$

\mathbf{H}_j and \mathbf{G}_j have the same properties as H_j and G_j . Define \mathbf{W}_j by

$$\mathbf{W}_j = \begin{pmatrix} \mathbf{H}_j \\ \mathbf{G}_j \end{pmatrix}.$$

Then, $\mathbf{W}_j : \mathbf{V}_j \rightarrow \mathbf{V}_{j+1} \oplus \mathbf{W}_{j+1}$ and \mathbf{W}_j is orthogonal.

4 Applying the Wavelet Transform to Multigrid

4.1 The One-Dimensional Case

The one-dimensional discussion of this method appears in [9]. In that paper, the fine grid operator is assumed to be symmetric. A brief re-cap of the results follows, after which we present a theorem related to the compression of the inverse component in the resulting operators.

Given the problem

$$L_j U = F,$$

where L_j represents the operator on the fine grid obtained by discretization of a one-dimensional boundary value problem, the wavelet transform is applied to both sides of the equation, yielding

$$(\mathcal{W}_j L_j \mathcal{W}_j^T) \mathcal{W}_j U = \mathcal{W}_j F$$

Define \tilde{L}_j as follows:

$$\tilde{L}_j = \mathcal{W}_j L_j \mathcal{W}_j^T = \begin{pmatrix} T_j & B_j \\ B_j^T & D_j \end{pmatrix}.$$

Taking the block UDL decomposition of \tilde{L}_j , where U is block upper triangular with unit diagonal, D is block diagonal, and L is block lower triangular with unit diagonal; calculating the inverse

of the factorization; and solving for U , it is clear that the interpolation and restriction operators should be defined by

$$I_{2h}^h = \sqrt{2}(H_j^T - G_j^T D_j^{-1} B_j^T)$$

and

$$I_h^{2h} = \frac{1}{2}(I_{2h}^h)^T,$$

respectively, and the coarse grid operator should be

$$L_{j+1} = T_j - B_j D_j^{-1} B_j^T, \quad (4)$$

which is the Schur complement of D_j in \tilde{L}_j . Now, although the matrices T_j , B_j , and D_j are as sparse as the original operator L_j , D_j^{-1} is not. But, we also observe that the fill-in that results from inversion decays exponentially as we move away from the original tridiagonal structure. This is evident in the structure of the coarse grid operator L_{j+1} (compare [2]).

The above procedure may be repeatedly applied until the desired coarseness is reached. Although the level of fill-in in the operator D_j^{-1} increases, D_j^{-1} has the property of exponential decay away from the diagonal (proved below), which makes compressibility possible. This property is due to the characteristics of wavelets (again, refer to [2]). One thing to keep in mind is that the number of gridpoints used must be an even number, and, in two dimensions, the next to coarsest grid can have no fewer than four gridpoints.

4.2 Investigation into the Exponential Decay in D_j^{-1}

Here, we investigate the exponential decay of the values of the elements of D_j^{-1} on terms that are not on the main diagonal. First, we will show that, using Haar wavelets, the matrix D_j has the same tridiagonal structure as the operator L_j , where L_j represents the operator formed by discretization of the one-dimensional differential equation

$$\begin{aligned} -\frac{d}{dx}(a(x)\frac{d}{dx}u(x)) + b(x)\frac{d}{dx}u(x) &= f(x), \quad x \text{ in } \Omega \\ u(x) &= g, \quad x \text{ on } \partial\Omega, \end{aligned} \quad (5)$$

where a is positive. The discretization of (5) takes the form

$$\frac{-a_{i+\frac{1}{2}}u_{i+1} + (a_{i+\frac{1}{2}} + a_{i-\frac{1}{2}})u_i - a_{i-\frac{1}{2}}u_{i-1}}{h^2} + \frac{b_i^- u_{i+1} + |b_i|u_i - b_i^+ u_{i-1}}{h} = f_{ij}, \quad (6)$$

where

$$b_i^- = \frac{1}{2}(b_i - |b_i|), \quad (7)$$

$$b_i^+ = \frac{1}{2}(b_i + |b_i|). \quad (8)$$

Lemma 1. *Given L_j , the operator obtained from discretizing (5), assuming constant coefficients, using the three-point discretization with upwinding ((6), (7), and (8)), the matrix $D_j = G_j L_j G_j^T$, where G_j is the Haar wavelet operator, has the same tridiagonal structure as L_j .*

Proof. Given the discretization from (6), (7), and (8), L_j takes the form

$$L_j = \begin{pmatrix} \frac{a_{i+\frac{1}{2}} + a_{i-\frac{1}{2}}}{h^2} + \frac{|b_i|}{h} & -\frac{a_{i+\frac{1}{2}}}{h^2} + \frac{b_i^-}{h} & & \\ -\frac{a_{i-\frac{1}{2}}}{h^2} - \frac{b_i^+}{h} & \frac{a_{i+\frac{1}{2}} + a_{i-\frac{1}{2}}}{h^2} + \frac{|b_i|}{h} & -\frac{a_{i+\frac{1}{2}}}{h^2} + \frac{b_i^-}{h} & \\ & \ddots & \ddots & \ddots \\ & & -\frac{a_{i-\frac{1}{2}}}{h^2} - \frac{b_i^+}{h} & \frac{a_{i+\frac{1}{2}} + a_{i-\frac{1}{2}}}{h^2} + \frac{|b_i|}{h} \end{pmatrix}.$$

For simplicity, define

$$\alpha = \frac{a_{i+\frac{1}{2}} + a_{i-\frac{1}{2}}}{h^2} + \frac{|b_i|}{h}, \quad (9)$$

$$\beta = -\frac{a_{i+\frac{1}{2}}}{h^2} + \frac{b_i^-}{h}, \quad (10)$$

$$\gamma = -\frac{a_{i-\frac{1}{2}}}{h^2} - \frac{b_i^+}{h}. \quad (11)$$

Assuming constant coefficients, $\alpha = -\beta - \gamma$. Then, a straightforward multiplication of $G_j L_{j+1} G_j^T$ proves the desired result. \square

Theorem 1. Define $D_j = G_j L_j G_j^T$ as in Lemma 1. Then, there exist constants $C > 0$ and $0 < \rho < 1$ such that

$$|(D_j^{-1})_{ij}| \leq C\rho^{|j-i|}. \quad (12)$$

Proof. The proof is done by using a power series expansion to calculate D_j^{-1} . It is straightforward to show that $\frac{|\beta|}{\alpha} < 1$ and $\frac{|\gamma|}{\alpha} < 1$. Also, note that $\beta, \gamma \leq 0$ and $\alpha > 0$. Rewriting D_j , we see that

$$D_j^{-1} = \frac{2}{3\alpha} \left(\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} + \begin{pmatrix} 0 & -\frac{\beta}{3\alpha} & & \\ -\frac{\gamma}{3\alpha} & 0 & -\frac{\beta}{3\alpha} & \\ & \ddots & \ddots & \ddots \\ & & -\frac{\gamma}{3\alpha} & 0 \end{pmatrix} \right)^{-1}.$$

Since $\frac{|\beta|}{\alpha} < 1$ and $\frac{|\gamma|}{\alpha} < 1$, a power series expansion may be used to represent the inverse. Let

$$A = \begin{pmatrix} 0 & -\frac{\beta}{3\alpha} & & \\ -\frac{\gamma}{3\alpha} & 0 & -\frac{\beta}{3\alpha} & \\ & \ddots & \ddots & \ddots \\ & & -\frac{\gamma}{3\alpha} & 0 \end{pmatrix}. \quad (13)$$

Then, we have

$$\begin{aligned} \left\| D_j^{-1} - \frac{2}{3\alpha} (I - A + A^2 + \dots + (-1)^{n-1} A^{n-1}) \right\| &\leq \frac{2}{3\alpha} \sum_{i=n}^{\infty} \|A^i\| \\ &\leq \frac{2}{3\alpha} \sum_{i=n}^{\infty} \left(\frac{2}{3}\right)^i = \frac{2}{\alpha} \left(\frac{2}{3}\right)^n \end{aligned}$$

The norm of A can be bounded:

$$\|A\|_\infty = \|A\|_1 = -\frac{\beta + \gamma}{3\alpha},$$

so that $0 < \|A\| = \rho < 1$, where

$$\rho = -\frac{\beta + \gamma}{3\alpha}. \quad (14)$$

The first appearance of the $(ij)^{\text{th}}$ entry in the estimate for D_j^{-1} occurs in the $|j - i|^{\text{th}}$ term in the power series approximation. Due to the structure of A , it is apparent from straightforward multiplication that further appearances occur in alternating terms of the expansion. This means that

$$\begin{aligned} |(D_j^{-1})_{ij}| &\leq \|A\|^{|j-i|} + \|A\|^{|j-i|+2} + \dots \\ &= \rho^{|j-i|} \frac{1}{1 - \rho^2} \\ &= C\rho^{|j-i|}, \end{aligned}$$

where

$$C = (1 - \rho^2)^{-1}. \quad (15)$$

Therefore, the estimate (12) holds, with C given by (15) and ρ given by (14). \square

4.3 The Two-Dimensional Case

Here, we will briefly describe the two-dimensional application of the wavelet multigrid method. Full details may be found in [7].

Given the problem

$$L_j U = F, \quad (16)$$

where L_j represents the operator on the fine grid obtained by discretizing a two-dimensional partial differential equation, we apply the wavelet transform to both sides of the equation. Denoting \mathcal{W}_j by \mathcal{W}_j for simplicity, we obtain

$$\begin{aligned} (\mathcal{W}_j L_j \mathcal{W}_j^T) \mathcal{W}_j U &= \mathcal{W}_j F \\ \implies (\mathcal{W}_j L_j \mathcal{W}_j^T) \begin{pmatrix} U_L \\ U_{LH} \\ U_{HL} \\ U_{HH} \end{pmatrix} &= \begin{pmatrix} F_L \\ F_{LH} \\ F_{HL} \\ F_{HH} \end{pmatrix}, \end{aligned} \quad (17)$$

where $U_L, F_L \in \mathbf{V}_j$ and $(U_{LH}, U_{HL}, U_{HH})^T, (F_{LH}, F_{HL}, F_{HH})^T \in \mathbf{W}_j$. Note that for simplicity, we will also let H_j denote \mathbf{H}_j and G_j denote \mathbf{G}_j . We also observe that in two-dimensions, application of the wavelet transform only requires $\mathcal{O}(n)$ operations. The reader is referred to [2] for more details regarding the fast wavelet transform.

Performing the multiplication on the left hand side, $\mathcal{W}_j L_j \mathcal{W}_j^T$, we then partition the resulting matrix, which we will denote by \tilde{L}_j :

$$\tilde{L}_j = \mathcal{W}_j L_j \mathcal{W}_j^T = \begin{pmatrix} T_j & B_j \\ C_j & D_j \end{pmatrix}. \quad (18)$$

Then, we determine the block UDL decomposition of \tilde{L}_j , where U is block upper triangular with unit diagonal, D is block diagonal, and L is block lower triangular with unit diagonal. Using this to find \tilde{L}_j^{-1} , defining

$$\begin{pmatrix} U_L \\ U_H \end{pmatrix} = \begin{pmatrix} U_L \\ U_{LH} \\ U_{HL} \\ U_{HH} \end{pmatrix}$$

and similarly for $\begin{pmatrix} F_L \\ F_H \end{pmatrix}$, and solving for $\begin{pmatrix} U_L \\ U_H \end{pmatrix}$, we see that

$$\begin{pmatrix} U_L \\ U_H \end{pmatrix} = \begin{pmatrix} (T_j - B_j D_j^{-1} C_j)^{-1} (H_j - B_j D_j^{-1} G_j) \\ -D_j^{-1} C_j (T_j - B_j D_j^{-1} C_j)^{-1} (H_j - B_j D_j^{-1} G_j) + D_j^{-1} G_j \end{pmatrix} F. \quad (19)$$

So,

$$U = \begin{pmatrix} H_j^T - G_j^T D_j^{-1} C_j & G_j^T \end{pmatrix} \begin{pmatrix} (T_j - B_j D_j^{-1} C_j)^{-1} & 0 \\ 0 & D_j^{-1} \end{pmatrix} \begin{pmatrix} H_j - B_j D_j^{-1} G_j \\ G_j \end{pmatrix} F.$$

Denote

$$I_{2h}^h = \sqrt{2} (H_j^T - G_j^T D_j^{-1} C_j) \quad (20)$$

and

$$I_h^{2h} = \frac{\sqrt{2}}{2} (H_j - B_j D_j^{-1} G_j) \quad (21)$$

as our interpolation and restriction operators, respectively. Note that if the fine grid operator L_j is symmetric, then $C_j = B_j^T$ and $I_h^{2h} = \frac{1}{2} (I_{2h}^h)^T$. Using the interpolation and restriction operators defined in (20) and (21), we have

$$U = I_{2h}^h (T_j - B_j D_j^{-1} C_j)^{-1} I_h^{2h} F + G_j^T D_j^{-1} G_j F. \quad (22)$$

We also note that in multigrid, we are working on the residual equation, i.e.,

$$e = I_{2h}^h (T_j - B_j D_j^{-1} C_j)^{-1} I_h^{2h} r + G_j^T D_j^{-1} G_j r.$$

If we assume that $G_j^T D_j^{-1} G_j r$ is small, i.e., r is almost in $\text{Range}(H_j^T)$, then we can approximate the error by

$$e = I_{2h}^h (T_j - B_j D_j^{-1} C_j)^{-1} I_h^{2h} r.$$

So,

$$(T_j - B_j D_j^{-1} C_j) e_{2h} = I_h^{2h} r_h.$$

The above assumption is good for most of the classical iterative methods, like Jacobi and Gauss-Seidel. Therefore, our coarse grid operator is

$$L_{j+1} = T_j - B_j D_j^{-1} C_j, \tag{23}$$

which is the Schur complement of D_j in \tilde{L}_j .

Observe that this operator is the same as the one we obtain if we solve for U_L in (19). Solving for U_L yields

$$\begin{aligned} U_L &= (T_j - B_j D_j^{-1} C_j)^{-1} F_L - (T_j - B_j D_j^{-1} C_j)^{-1} B_j D_j^{-1} F_H \\ &= (T_j - B_j D_j^{-1} C_j)^{-1} (F_L - B_j D_j^{-1} F_H). \end{aligned}$$

Also, if the fine grid operator is symmetric, then the coarse grid operator is $T_j - B_j D_j^{-1} B_j^T$.

We will denote the multigrid method thus formed as the wavelet multigrid method. Notice that although the wavelet and scaling operators are periodic, this method is applicable to any problem, even to those which are nonperiodic.

5 Improving Efficiency of the Wavelet Multigrid Method

We would like to make this procedure more efficient, so as to be practically useful. Although D_j is not dense (it is, in fact, essentially a banded matrix), its inverse is dense due to fill-in. But, we observe a significant amount of decay of the values on certain diagonals, indicating that it is possible to increase the efficiency of the method in this area, thereby improving the efficiency of the overall algorithm.

One step towards achieving this goal is to avoid computing the inverse exactly. To this end, we use ILU(0) to compute the incomplete LU factorization, and then use a sequence of forward and backward substitutions to compute the inverse. Although using ILU(0) reduces the computational complexity of calculating D_j^{-1} without compromising the convergence of the method, the resulting inverse is still dense. This results in representations for the coarse grid, restriction, and interpolation operators that are much denser than the fine grid operator. To improve further on these results, we use a truncation procedure – any values that appear in the computed inverse in locations that hold zero values in D_j are set to zero, thus eliminating any fill-in over the original matrix, D_j . We will call this method the truncated wavelet multigrid method, and we will refer to the original method as the dense or full wavelet multigrid method. The term wavelet multigrid method will refer to a generality applying to both versions.

We will briefly discuss the complexity of the coarse grid operator for the truncated wavelet multigrid method. After calculating D_j^{-1} using the method of ILU(0) followed by truncation, for the examples involving full coarsening and no coupling discussed in the following section, we find that the inverse is representative of a stencil that contains between seven and twenty-three elements. This leads to a coarse-grid operator $T_j - B_j D_j^{-1} C_j$ that corresponds to a stencil having between twenty-one and twenty-five elements. We observe that the matrix T_j has the same structure as the fine grid operator, i.e., it corresponds to a stencil with five elements. The additional elements, then, come solely from the product $B_j D_j^{-1} C_j$.

6 Numerical Applications

In this section, we display the numerical results of applying the truncated wavelet multigrid method to various problems. For more examples of applications, see [7]. We compare the convergence of the truncated wavelet multigrid method with the full wavelet multigrid method; the algebraic multigrid method (AMG1R5), the method designed by John Ruge, Klaus Stüben, and Rolf Hempel, an earlier version of which is described in [17]; the standard linear multigrid method, using nine-point interpolation, full-weighting restriction (a constant multiple of the adjoint of the nine-point interpolation, scaled so that the sum of the weights is one (see [3] and [13])), and a coarse grid operator defined by discretizing the equation on a grid with the appropriate step size; the homogenization method (where appropriate), using the homogenized coarse grid operator described in [10] and [11] and the nine-point interpolation and full-weighting restriction operators. For all problems, unless otherwise specified, numerical results are analyzed using, for the fine grid in the interior, a 16×16 grid, leading to a 256×256 matrix, a 32×32 grid, leading to a 1024×1024 matrix, and a 64×64 grid, leading to a 4096×4096 matrix. With respect to standard multigrid, multigrid with homogenization, and algebraic multigrid, an odd number of gridpoints is required, so that we have a 15×15 grid, a 31×31 grid, and a 63×63 grid in the interior.

6.1 Elliptic Problems

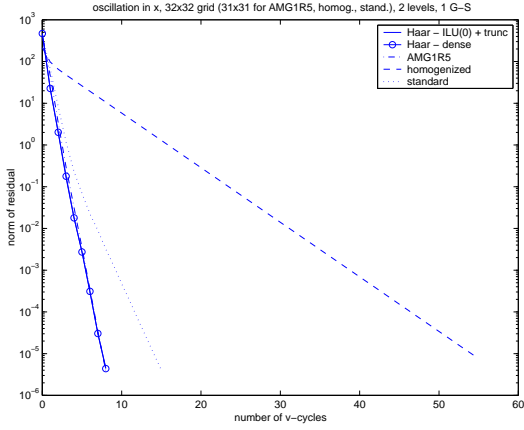
First, we look at the elliptic problem,

$$\begin{aligned} -\nabla \cdot (a(x, y) \nabla u(x, y)) &= 0, \text{ in } \Omega \\ u(x, y) &= 0, \text{ on } \partial\Omega, \end{aligned} \tag{24}$$

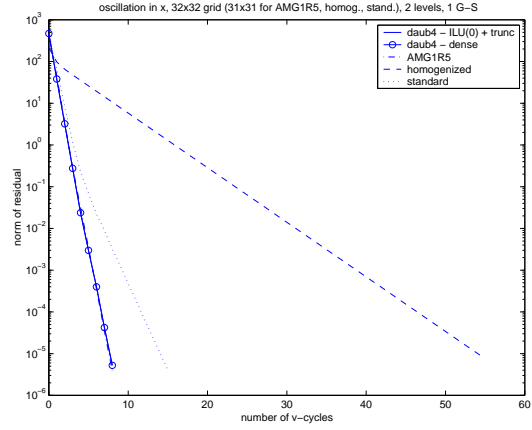
where Ω is the unit square and $a > 0$. We look at two cases: oscillation in the x -direction, $a(x, y) = 1 + 0.8 \sin(10\sqrt{2}\pi x)$ and oscillation along diagonals, $a(x, y) = 1 + 0.8 \sin(10\sqrt{2}\pi(x - y))$.

For both cases, we see that the full wavelet multigrid method actually has a convergence rate that is comparable to the algebraic multigrid method. In the case of oscillation in the x -direction, the truncated wavelet multigrid method using Haar wavelets applied on a 16×16 grid has a convergence rate that is worse than the algebraic multigrid method. Using Daubechies wavelets, however, we see a convergence rate that is very close to that of algebraic multigrid. For a 32×32 grid, both types of wavelets yield approximately the same convergence rate as algebraic multigrid. For the case of oscillation along diagonals, using either type of wavelets produces convergence rates that are almost identical to those achieved by the algebraic multigrid method, regardless of the mesh size of the fine grid. The convergence of the standard multigrid method is dependent on the mesh size, having very poor convergence for a 16×16 grid and somewhat better convergence for a 32×32 or 64×64 grid. For the two-level method, the improvement occurs as a result of the finer mesh size on the coarser grid. For multigrid using more levels, however, the same problem is encountered for the standard multigrid method, again due to the lack of detail in the coarse grid operator. The wavelet multigrid method does not suffer from this problem. For the 64×64 grid, we analyzed only the problem with oscillation along diagonals. For this problem, we see that the convergence rate of the truncated wavelet multigrid method using Haar wavelets is about three times as fast as for the standard and homogenized multigrid method and about the same as for the algebraic multigrid method. Figures 1 and 2 demonstrate the results on a 32×32 grid.

In Table 1, we compare the average convergence factor of the truncated wavelet multigrid method (using Haar wavelets) with the algebraic multigrid method, homogenized multigrid, and

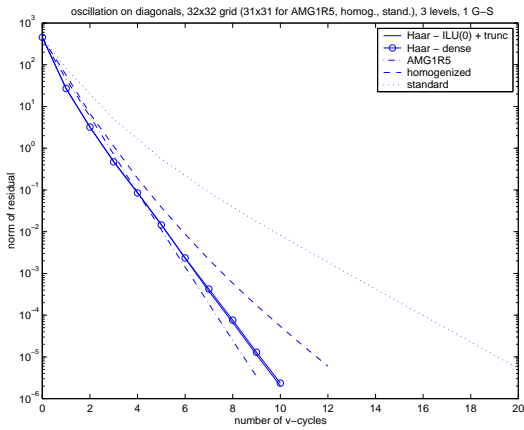


(a)

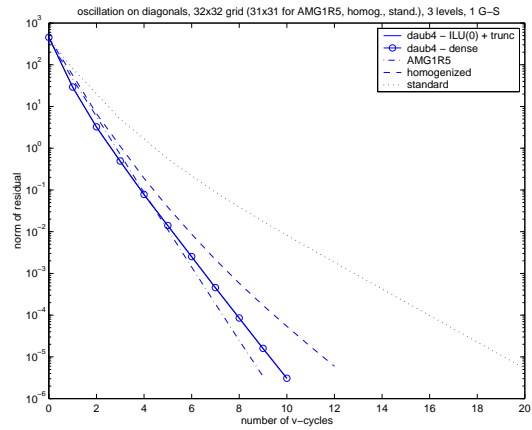


(b)

Figure 1: Oscillation in x -direction, 32×32 grid. Compare Haar and Daubechies wavelet multigrid with AMG1R5, homogenized, and standard methods. (a) uses Haar wavelets and (b) uses Daubechies wavelets.



(a)



(b)

Figure 2: Oscillation on diagonals, 32×32 grid, 3 levels. Compare Haar and Daubechies wavelet multigrid with AMG1R5, homogenized, and standard methods. (a) uses Haar wavelets and (b) uses Daubechies wavelets.

standard multigrid. The table demonstrates results with a fixed coarsest grid of 8×8 for each fine mesh size. These results show that the truncated wavelet multigrid method appears to have a convergence rate essentially independent of mesh size.

Table 1: Average convergence factor per cycle for diffusion problem with oscillation along diagonals

fine grid size	truncated Haar wavelet	AMG1R5	homogenized	standard
16×16 (2 levels)	0.1212	0.1189	0.2740	0.5286
32×32 (3 levels)	0.1461	0.1266	0.2211	0.4019
64×64 (4 levels)	0.1161	0.1260	0.4167	0.3546

Next, we look at the checkerboard problem, which is defined by (24) with

$$a = \begin{cases} 10^5 & \text{if } 0 < x, y < 0.5 \text{ or } 0.5 < x, y < 1, \\ 1 & \text{otherwise.} \end{cases} \quad (25)$$

The results for the checkerboard problem are quite good. The full Haar wavelet multigrid method has a convergence rate that is as good or better than the algebraic multigrid method. The truncated wavelet multigrid method also performs as well or better than the algebraic multigrid, except in the case where two iterations of Gauss-Seidel are used. In that case, the convergence rate of the algebraic multigrid method is slightly better. Clearly, the convergence rate of the wavelet multigrid method is essentially independent of the fine grid size. The standard multigrid method diverges for this problem. The results for the 16×16 and 32×32 grids are shown in Figure 3.

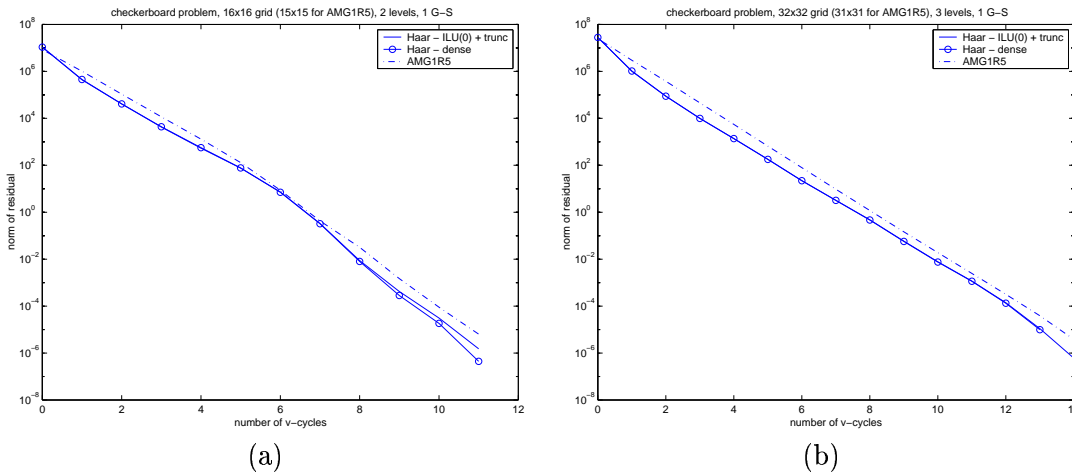


Figure 3: Checkerboard problem. Comparison of Haar wavelet multigrid method with AMG1R5 (standard multigrid fails to converge). (a) 16×16 grid, 2 levels; (b) 32×32 grid, 3 levels, both with 1 Gauss-Seidel.

In Table 2, we compare the average convergence factor of the truncated wavelet multigrid method (using Haar wavelets) with the algebraic multigrid method. The table demonstrates results with a fixed coarsest grid of 8×8 for each fine mesh size. These results show that the truncated wavelet multigrid method appears to have a convergence rate essentially independent of mesh size and also comparable to the algebraic multigrid method.

Table 2: Average convergence factor per cycle for the checkerboard problem

fine grid size	truncated Haar wavelet	AMG1R5
16×16 (2 levels)	0.0678	0.0791
32×32 (3 levels)	0.1057	0.1216
64×64 (4 levels)	0.1178	0.1271

6.2 The Advection-Diffusion Problem

Here, we are investigating the problem

$$\begin{aligned} -\epsilon\Delta u + b \cdot u &= 0, \text{ in } \Omega \\ u &= f(x), \text{ on } \partial\Omega, \end{aligned} \tag{26}$$

where Ω is the unit square and $\|b\| \gg \epsilon > 0$. In this problem, we encounter difficulties with multigrid methods due to the fact that some of the components of the solution oscillate along characteristics [22, 23]. So, moving to the coarse grid with the standard multigrid approach does not represent a good approximation to the problem on the coarse grid. We apply the wavelet multigrid method to these problems to overcome this difficulty, since application of the wavelet operator keeps the characteristics of the original problem.

To discretize, we use the usual five-point centered discretization for the diffusion term and a first order upwind scheme for the advection part of the equation. Although using first order upwind introduces artificial diffusion into the solution of the order of the mesh size squared, it provides a convenient test of the effectiveness of the different multigrid methods. Symmetric Gauss-Seidel is used as the smoother in order to ensure that we perform sweeps in the direction of the characteristics over the entire flow field.

The wavelet multigrid method will not be compared to AMG1R5 in this subsection, because AMG1R5 does not have the option of using symmetric Gauss-Seidel as the smoother. The standard multigrid, however, has been rewritten to allow for symmetric Gauss-Seidel as the smoother. Of course, as was mentioned in Section 1, standard multigrid may be modified to produce better convergence results (see e.g., [1, 22, 12]). First, we have a comparison of the methods for (26), where $b = ((2y - 1)(1 - x^2), 2xy(y - 1))$ and $f(x)$ is defined by

$$f(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{27}$$

Note that the discontinuous boundary condition will give rise to a boundary layer near the left-hand boundary. Also, the characteristics are parabolic, resulting in flow entering and exiting through the left-hand boundary. We set $\epsilon = 10^{-5}$ for the experiments. Convergence, as compared with multigrid only employing standard point Gauss-Seidel, has improved for both the standard multigrid method and the wavelet multigrid method. In fact, for both methods, convergence appears independent of mesh size, as can be seen in Figure 4.

Finally, we use the boundary conditions given by (27), but we change the advection component so that the characteristics are closed and the flow is skewed, so that it does not line up with the

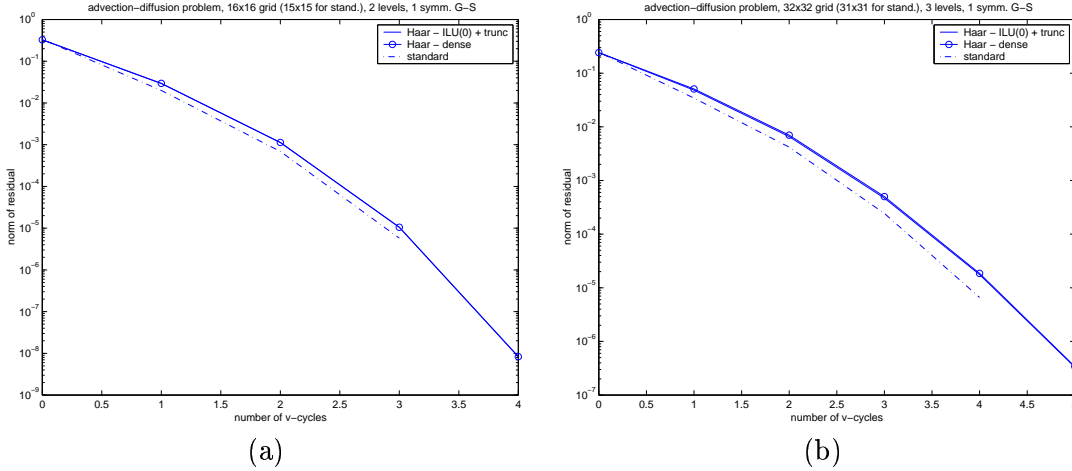


Figure 4: Comparison of wavelet multigrid method with standard multigrid method, using symmetric Gauss-Seidel as the smoother. $\epsilon = 10^{-5}$ and flow is parabolic, entering and exiting at the left-hand boundary. Boundary conditions are discontinuous, resulting in a boundary layer. (a) 16×16 grid, 2 levels; (b) 32×32 grid, 3 levels.

grid. Here,

$$b = (\sin(\pi y_1) \cos(\pi x_1) + \sin(\pi y_2) \cos(\pi x_2), \\ -\cos(\pi y_1) \sin(\pi x_1) - \cos(\pi y_2) \sin(\pi x_2)),$$

where

$$x_1 = x^2 + 0.5, \quad x_2 = (x - 1)^2 + 0.5, \quad y_1 = y^2 + 0.5, \quad y_2 = (y - 1)^2 + 0.5.$$

In this problem, the standard multigrid fails to converge, but the wavelet multigrid method performs very well (see Figure 5). Convergence is rapid and the convergence rate is essentially independent of the mesh size. The contour plot of the solution, which shows the boundary layer, is given in Figure 6.

Table 3 displays the average convergence factor per cycle of the truncated Haar wavelet multigrid method applied to the advection-diffusion problem with skewed flow and discontinuous boundary conditions. Here, the coarsest grid for all trials is 8×8 .

Table 3: Average convergence factor per cycle for advection-diffusion with skewed flow and discontinuous boundary conditions.

Fine grid size	Avg. Convergence/Cycle
16×16 (2 levels)	.0963
32×32 (3 levels)	.2003
64×64 (4 levels)	.2558

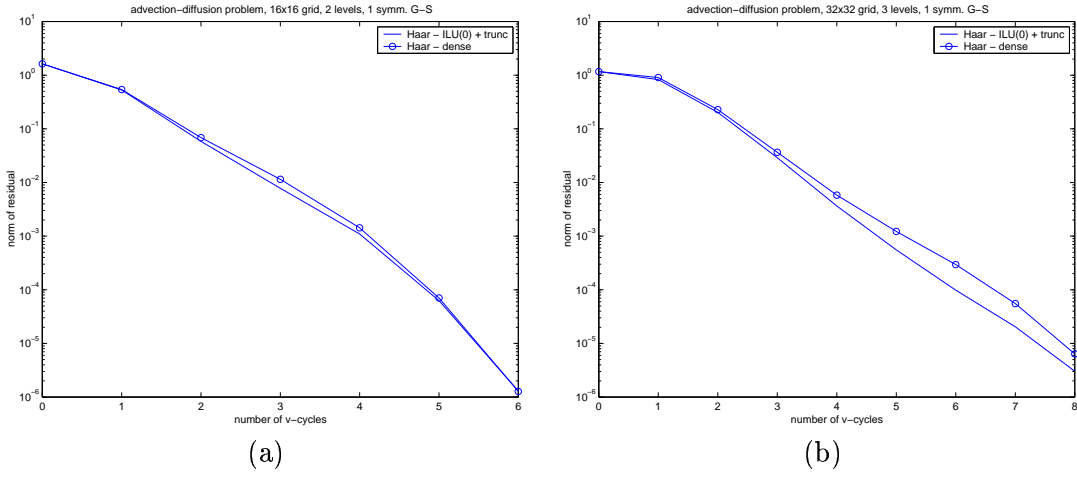


Figure 5: Wavelet multigrid method, using symmetric Gauss-Seidel as the smoother. $\epsilon = 10^{-5}$ and flow is skewed. Boundary conditions are discontinuous, resulting in a boundary layer. Standard multigrid fails to converge. (a) 16×16 grid, 2 levels; (b) 32×32 grid, 3 levels.

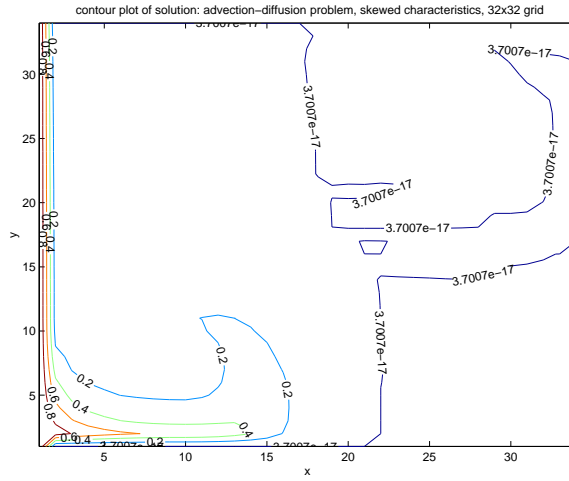


Figure 6: Contour plot of the solution of the advection-diffusion problem with skewed flow and discontinuous boundary conditions, resulting in a boundary layer. Results are shown for the 32×32 grid, 3 level multigrid.

6.3 The Anisotropic Diffusion Problem

Here we look at the problem

$$\begin{aligned} -u_{xx} - \epsilon u_{yy} &= 0, \text{ in } \Omega \\ u &= 0, \text{ on } \partial\Omega, \end{aligned} \tag{28}$$

where Ω is the unit square and $\epsilon > 0$.

Semicoarsening is used for this problem, so that the smoothed error will be approximated well on the coarser grid. Coarsening is done in the direction of the anisotropy, which, for the case of (28), is the x-direction. Figure 7 demonstrates that this approach is extremely successful. Not only is convergence greatly improved (over the wavelet multigrid method using full coarsening), but the convergence appears to be independent of the mesh size, as desired.

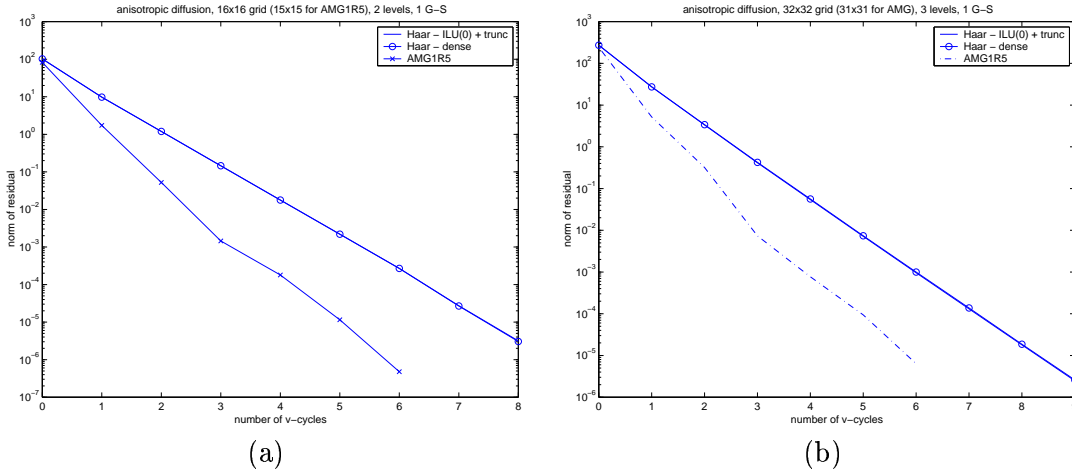


Figure 7: The anisotropic diffusion problem. Comparison of truncated wavelet multigrid, full wavelet multigrid, and AMG1R5 on a (a) 16x16 grid (2 levels) and (b) 32x32 grid (3 levels).

We will briefly discuss the complexity of the coarse grid operator for the truncated wavelet multigrid method in this special case. After calculating D_j^{-1} using ILU(0) and truncation, we find that D_j^{-1} corresponds to a stencil having only five elements. This leads to a coarse-grid operator $T_j - B_j D_j^{-1} C_j$ that corresponds to a stencil having approximately thirteen elements. The increase in density is due to the product $B_j D_j^{-1} C_j$.

6.4 Stokes Problem

Here, we consider the Stokes equations

$$\begin{aligned} -\Delta u + p_x &= f^u \text{ in } \Omega \\ -\Delta v + p_y &= f^v \text{ in } \Omega \\ u_x + v_y &= 0 \text{ on } \bar{\Omega}. \end{aligned}$$

As is done in [18], summing the first two equations and using the continuity condition ($u_x + v_y = 0$), we obtain the following system of equations:

$$-\Delta u + p_x = f^u \text{ in } \Omega \quad (29)$$

$$-\Delta v + p_y = f^v \text{ in } \Omega \quad (30)$$

$$\Delta p = f_x^u + f_y^v \text{ in } \Omega \quad (31)$$

$$u_x + v_y = 0 \text{ on } \partial\Omega. \quad (32)$$

Values for u and v are given on the boundary of Ω .

These equations are discretized using centered differencing for the first order terms (p_x, p_y, u_x, v_y) and the usual five-point discretization for the second order terms ($\Delta u, \Delta v, \Delta p$). Both f^u and f^v are assumed zero for the numerical calculations and Ω is the unit square. We use forward differencing (left-hand side and bottom of the square) and backward differencing (right-hand side and top of the square) to discretize the continuity equation (32). This discretization is used to obtain equations for the outermost interior values on the left and right boundaries of the square (for u) and on the top and bottom (for v). This can easily be done, because u and v are specified on the boundary of the unit square. So, we calculate v_y for $x = 0$ and $x = 1$, and then subtract this term from the right-hand side in our matrix equation and put the discretization for u_x into the matrix. The same holds true for the top and bottom, with u_x being calculated when $y = 0$ and $y = 1$ and v_y being discretized. The boundary values of p are obtained by using (29) for $x = 0$ and $x = 1$ and (30) for $y = 0$ and $y = 1$. Since the corner values of p do not appear directly in the discretization, we assign to them the value of the average of the adjacent boundary values for p . The V-cycle multigrid algorithm is followed, with Gauss-Seidel as the smoother, but fifty iterations are performed on the coarsest grid in lieu of an exact solve (since the matrix is singular). Fewer iterations are necessary on the coarsest grid if more than two levels are used (so that the coarse grid is actually sufficiently coarse).

One note must be made regarding this discretization. In order to have a solution of the Stokes problem, the velocity field on the boundary, $\mathbf{u} = (u, v)^T$ must satisfy the condition $\int \mathbf{u} \cdot \mathbf{n} dS = 0$, where $S = \partial\Omega$ (see, e.g., [20]). For the discretization employed here, this condition is insufficient to guarantee convergence. The velocity field on the boundary must also be chosen such that the condition $\int \frac{\partial p}{\partial n} dS = 0$ is met.

For the Stokes problem we have three unknowns, u , v , and p . Therefore, the wavelet multigrid method must be modified. For example, let H_u be the scaling operator that will be applied to the discrete values for u . Similarly, define H_v and H_p . We also define the wavelet operators G_u , G_v , and G_p accordingly. Let

$$\mathbf{H}_j = \begin{pmatrix} H_u & 0 & 0 \\ 0 & H_v & 0 \\ 0 & 0 & H_p \end{pmatrix}.$$

Define \mathbf{G}_j in the same manner. Then, the wavelet transform \mathbf{W}_j may be defined as

$$\mathbf{W}_j = \begin{pmatrix} \mathbf{H}_j \\ \mathbf{G}_j \end{pmatrix}.$$

The wavelet transform thus defined is orthogonal, and the new scaling and wavelet operators satisfy the conditions for the two-dimensional analogues. The wavelet multigrid method still follows as in

Section 4, letting

$$U = \begin{pmatrix} u \\ v \\ p \end{pmatrix}, F = \begin{pmatrix} f^u \\ f^v \\ f_x^u + f_y^v \end{pmatrix} + \text{boundary terms.}$$

In Figure 8, we demonstrate the effectiveness of the wavelet multigrid method in solving the Stokes equations with boundary conditions of

$$\begin{aligned} u &= \begin{cases} 1 & \text{for } y = 1, \\ 0 & \text{otherwise,} \end{cases} \\ v &= 0. \end{aligned}$$

The dense wavelet multigrid method clearly has rapid convergence for this problem. Similar convergence results are achieved for other boundary conditions tested that satisfy the restrictions given above, such as $u = v = 0$ everywhere on the boundary, except at $y = 1$, where $u = \sin(\pi x)$.

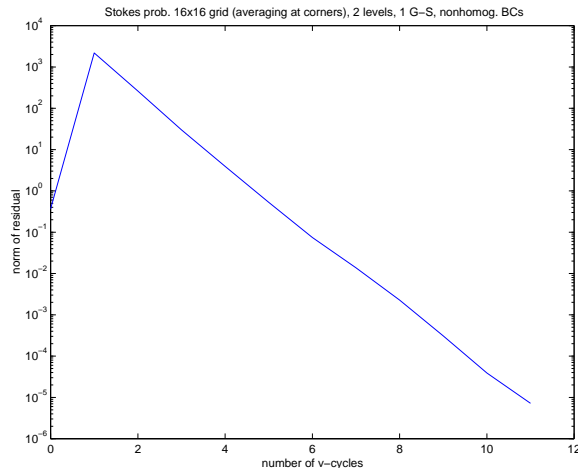


Figure 8: The Stokes equations on a 16×16 grid, nonhomogeneous boundary conditions.

7 Conclusion

The new multigrid method described here, called the wavelet multigrid method, has proven to be very useful in a wide variety of problems. In many of those problems where standard multigrid methods fail to converge independently of mesh size, the wavelet multigrid method does ensure such convergence. Also, due to the properties of wavelets, in many cases we can efficiently apply the wavelet multigrid method through use of compression. With respect to solving the anisotropic diffusion problem, semicoarsening ensures good convergence results. These results have shown that it is worthwhile to further explore the usefulness of this method.

8 Acknowledgements

My thanks go to John Ruge for providing a copy of the AMG1R5 program and for his willingness to answer questions regarding the program. I also wish to thank my mentor, Bjorn Engquist, whose guidance has been invaluable.

References

- [1] Jürgen Bey and Gabriel Wittum, "Downwind Numbering: A Robust Multigrid Method for Convection Diffusion Problems on Unstructured Grids," in *Fast Solvers for Flow Problems (Kiel, 1994)*, volume 49 of *Notes on Numerical Fluid Mechanics*, pp. 63–73, Vieweg, Braunschweig, 1995.
- [2] G. Beylkin, R. Coifman, and V. Rokhlin, "Fast Wavelet Transforms and Numerical Algorithms I," *Communications on Pure and Applied Mathematics*, 44:141–183, 1991.
- [3] William L. Briggs, Stephen F. McCormick, and Van Emden Henson, *A Multigrid Tutorial, Second Edition*, SIAM, Philadelphia, PA, 2000.
- [4] Wolfgang Dahmen and Ludwig Elsner, "Algebraic Multigrid Methods and the Schur Complement," in *Robust Multi-Grid Methods (Kiel, 1988)*, volume 23 of *Notes on Numerical Fluid Mechanics*, pp. 58–68, Vieweg, Braunschweig, 1989.
- [5] Ingrid Daubechies, "Orthonormal Bases of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, XLI(7):909–996, 1988.
- [6] Ingrid Daubechies, *Ten Lectures on Wavelets*, volume 61 of *CBMS-NSF Series in Applied Mathematics*, SIAM, Philadelphia, PA, 1992.
- [7] Doreen De Leon, "Wavelet Operators Applied to Multigrid Methods" (Ph. D. thesis), UCLA Mathematics Department CAM Report 00-22, June 2000.
- [8] Mihai Dorobantu and Bjorn Engquist, "Wavelet-Based Numerical Homogenization," *SIAM Journal on Numerical Analysis*, 35(2):540–559, April 1998.
- [9] Bjorn Engquist and Erding Luo, "The Multigrid Method Based on a Wavelet Transformation and Schur Complement," Unpublished.
- [10] Bjorn Engquist and Erding Luo, "New Coarse Grid Operators for Highly Oscillatory Coefficient Elliptic Problems," *Journal of Computational Physics*, 129:296–306, 1996.
- [11] Bjorn Engquist and Erding Luo, "Convergence of a Multigrid Method for Elliptic Equations with Highly Oscillatory Coefficients," *SIAM Journal on Numerical Analysis*, 34(6):2254–2273, December 1997.
- [12] M. Griebel and S. Knapek, "A Multigrid Homogenization Method," in *Modeling and Computation in Environmental Sciences (Stuttgart, 1995)*, volume 59 of *Notes on Numerical Fluid Mechanics*, pp. 187–202, Vieweg, Braunschweig, 1995.
- [13] Stephen McCormick, editor, *Multigrid Methods*, Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1987.

- [14] J. David Moulton, Joel E. Dendy, Jr., and James M. Hyman, "The Black Box Multigrid Numerical Homogenization Algorithm," *Journal of Computational Physics*, 142:80–108, 1998.
- [15] Nicolas Neuss, Willi Jäger, and Gabriel Wittum, "Homogenization and Multigrid," Preprint 98-04 (SFB 359), January 1998.
- [16] Andreas Rieder, "Multilevel Methods Based on Wavelet Decomposition," *East-West Journal of Numerical Mathematics*, 2(4):313–330, 1994.
- [17] J. W. Ruge and K. Stüben, "Algebraic Multigrid," in *Multigrid Methods*, Frontiers in Applied Mathematics, pp. 73-130, SIAM, Philadelphia, PA, 1987.
- [18] A. Schüller, "A Multigrid Algorithm for the Incompressible Navier-Stokes Equations," in *Numerical Methods for Advection-Diffusion Problems (Kiel, 1989)*, volume 30 of *Note on Numerical Fluid Mechanics*, pp. 124–133, Vieweg, Braunschweig, 1990.
- [19] K. Stüben, "Algebraic Multigrid (AMG): An Introduction with Applications," Technical Report 70, GMD, November 1999.
- [20] Roger Temam, *Navier-Stokes Equations: Theory and Numerical Analysis*, 3rd ed., volume 2 of *Studies in Mathematics and its Applications*, North-Holland, New York, 1984.
- [21] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, London, 2001.
- [22] Irad Yavneh, Cornelis H. Venner, and Achi Brandt, "Fast Multigrid Solution of the Advection Problem with Closed Characteristics," *SIAM Journal on Scientific Computing*, 19(1):111–125, January 1998.
- [23] Irad Yavneh, "Coarse-Grid Correction for Nonelliptic and Singular Perturbation Problems", *SIAM Journal on Scientific Computing*, 19(5):1682–1699, September 1998.