52

On the Use of Multigrid as a Preconditioner

C.W. Oosterlee and T. Washio

1 Introduction

In the search for robust and efficient Krylov subspace methods, multigrid is being considered as a preconditioner. With preconditioners based on multigrid it is expected that robust convergence can be achieved for a large class of problems. GMRES [SS86] is used as the Krylov subspace solver. Singularly perturbed 2D problems of both convection-diffusion and of jumping coefficients type are considered, for which the design of optimal standard multigrid is not easy. For these problems the multigrid method is being compared as a solver and as a preconditioner. Eigenvalue spectra of the multigrid iteration matrix are analyzed to understand the convergence of the algorithms.

In the domain decomposition context we can think of the method as a robust subdomain solver. Also, the parallel multiblock method can be seen as an alternative for domain decomposition techniques on regular grids: The method is parallelized with *grid partitioning* [MFL$^+$91].

The purpose of this work is not to derive optimal multigrid methods for specific problems, but to construct a robust well-parallelizable solver, in which the smoother as well as the coarse grid correction is fixed. Another robust multigrid variant for solving scalar partial differential equations is Algebraic Multigrid (AMG) by Ruge and Stüben [RS87], in which the smoother is fixed but the transfer operators depend on the connections in a matrix. Efficient parallelization of AMG is, however, not trivial. Matrix-dependent transfer operators are employed so that problems with convection dominance, as well as problems with jumping coefficients, can be solved efficiently. The operators have been designed so that problems on grids with arbitrary mesh sizes, not just powers of two (+1), can be solved with similar efficiency. The algorithm uses the prolongation operators introduced by de Zeeuw [Zee90]. The multigrid algorithm employs Galerkin coarsening [Hac85], [Wes92] for building the matrices on coarser grids. The alternating zebra line Gauss-Seidel relaxation method is used as the smoother, since it is a robust smoother for anisotropic problems and it is efficiently parallelizable. In [Ket82] an early comparison of multigrid and multigrid

Ninth International Conference on Domain Decomposition Methods
Editor Petter E. Bjørstad, Magne S. Espedal and David E. Keyes

preconditioned CG for symmetric model equations showed the promising robustness of the latter method.

The solution method is analyzed in order to understand the convergence behavior of multigrid used as a solver and as a preconditioner. The eigenvalue spectrum of the multigrid iteration matrix for the singularly perturbed problems is calculated in Section 3. Interesting subjects for the convergence behavior are the spectral radius and the eigenvalue distribution. Numerical results are also presented in Section 3. The benefits of the constructed multigrid preconditioned Krylov methods are shown for a convection-diffusion problem and a problem with jumping coefficients on fine grids solved on the NEC Cenju–3 MIMD machine [HCH+96]. The message-passing is done with MPI.

2 The Multigrid Preconditioned Krylov Methods

We concentrate on linear matrix systems with nine diagonals,

$$A\phi \quad = \quad b \ . \tag{2.1}$$

Matrix A has right preconditioning as follows:

$$AK^{-1}(K\phi) = b \ . \tag{2.2}$$

The Krylov subspace method that is used for solving (2.2) is GMRES(m) [SS86]. Matrix K^{-1} in (2.2) is approximated by one iteration of the multigrid method.

Using a preconditioner as solver. A preconditioner, like the multigrid preconditioner, is a candidate for use as a solver. An iteration of a multigrid solver is equivalent to a Richardson iteration on the preconditioned matrix. With K being the iteration matrix, multigrid can be written as follows:

$$K\phi^{(k+1)} + (A-K)\phi^{(k)} = b \ . \tag{2.3}$$

This formulation is equivalent to,

$$\phi^{(k+1)} = \phi^{(k)} + K^{-1}(b - A\phi^{(k)}) = \phi^{(k)} + K^{-1}r^{(k)}; \ \ r^{(k+1)} = (I - AK^{-1})r^{(k)} \ . \tag{2.4}$$

The multigrid solver is implemented as a Richardson iteration with a left multigrid preconditioner for A. The convergence of (2.4) can be investigated by analyzing the spectrum of the iteration matrix. The spectral radius of $I - AK^{-1}$ determines the convergence. This spectrum is analyzed in Section 3 for the multigrid method for the problems tested on 33^2 and 65^2 grids. With this spectrum we can also investigate the convergence of GMRES, since the spectra of left and right preconditioned matrices are the same.

The multigrid preconditioner. The multigrid preconditioner implemented is now discussed in some more detail. The multigrid correction scheme [Hac85, Wes92] is used for solving the linear equations. Here, the robustness and efficiency of the F-cycle is evaluated. The multigrid F-cycle is a hybrid between the cheap V-cycle and the expensive W-cycle. The smoother is the alternating zebra line Gauss-Seidel smoother.

First, all odd (white) lines are processed in one direction, after which all even (black) lines are processed. This procedure takes place in the x- and y-directions. Fourier smoothing analysis for model equations, presented in [Wes92], indicates the robustness of this smoother. The algorithm evaluated adopts the "upwind" prolongation operator by de Zeeuw [Zee90]. This operator is specially designed for problems with jumping coefficients and for second-order differential equations with a dominating first-order derivative. As already indicated in [Den83] it is appropriate for the construction of transfer operators for unsymmetric matrices to split a matrix A into a symmetric and an antisymmetric part:

$$S = \frac{1}{2}(A + A^T), \;\; T = A - S = \frac{1}{2}(A - A^T) \; . \tag{2.5}$$

The investigated transfer operators are also based on this splitting. Analysis of this operator and the numerical experiments in [Zee90] shows the very interesting behavior of these operators. Restriction R^L is defined as the transpose of the prolongation operator. The coarse grid matrices A^L are defined with Galerkin coarsening [Hac85], [Wes92],

$$A^M = A \; , \tag{2.6}$$
$$A^L = R^L A^{L+1} P^{L+1}, \; 1 \leq L \leq M-1 \; . \tag{2.7}$$

M represents the finest grid level.

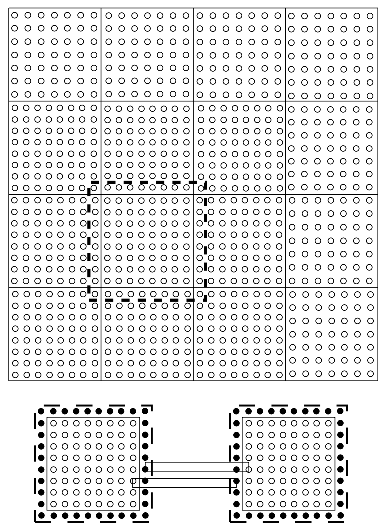
Grid partitioning. If grid applications are to be implemented on parallel computers, a straightforward approach is to use *grid partitioning*. This means that the original domain Ω is split into P parts (subdomains) in such a way that, with respect to the finest grid, each subdomain consists of (roughly) the same number of grid points. Because of the only local dependencies of grid points, each process needs foreign data only from boundary areas of neighbor subdomains. After a smoothing step is performed, data have to be communicated along the artificial boundaries (see Figure 1). The extension of the single grid case to parallel multigrid is obvious: On the finest grid level, all communication is a strictly local one. Similarly, also on all coarser grids necessary communication is "local" relatively to the corresponding grid level.

Parallelism is straightforward in Krylov methods, except for the multigrid preconditioner, the matrix-vector and inner products, which need communication among neighboring processors for the problems under consideration. We would like to point out that in our approach all parallel algorithms are algorithmically equivalent to their non–partitioned versions: the results of the partitioned and the non–partitioned versions are identical.

3 Numerical Results

The equations investigated are two 2D singularly perturbed problems. A convection-diffusion equation with a dominating convection term and an interface problem are solved. We concentrate on "difficult" problems for multigrid. As the initial guess $\phi^{(0)} = 0$ is used for all problems. Restart parameter m is set to 20 here. After some efficiency tests, we choose no pre-smoothing and 2 post-smoothing iterations;

Figure 1 A regular grid partitioned into 16 subgrids. To each subgrid an overlap area is assigned needed for data exchange in the exchange phase.



on the coarsest grid 2 smoothing iterations are performed ($\nu_3 = 2$) in order to keep the parallel method as cheap as possible. For the problems investigated this did not influence convergence negatively, since the coarsest grid is always a 3^2 grid. The results presented are the number of iterations (n) to reduce the L_2-norm of the initial residual with 8 orders of magnitude ($||r^{(n)}||_2/||r^{(0)}||_2 \leq 10^{-8}$). Furthermore, the elapsed time for this number of iterations obtained on the NEC Cenju–3 MIMD machine [HCH+96] is presented. For all problems 32 processors are used in a 4×8 configuration.

Rotating convection-diffusion equation. The first problem is a rotating convection-diffusion problem,

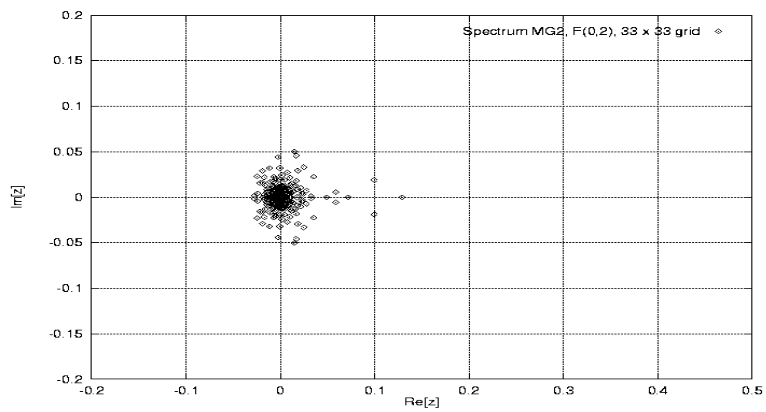
$$-\epsilon\Delta\phi + a(x,y)\frac{\partial\phi}{\partial x} + b(x,y)\frac{\partial\phi}{\partial y} = 1 \;\text{ on } \Omega = (0,1)\times(0,1) \;. \tag{3.8}$$

Here, $a(x,y) = -\sin(\pi x).\cos(\pi y)$, $b(x,y) = \sin(\pi y).\cos(\pi x)$

Dirichlet boundary conditions are prescribed: $\phi|_\Gamma = \sin(\pi x) + \sin(13\pi x) + \sin(\pi y) + \sin(13\pi y)$.

A convection dominated test case is chosen: $\epsilon = 10^{-5}$. The convection terms are discretized with a standard upwind discretization. A first order upwind discretization is chosen, since this is still a linear discretization, which can be tested and evaluated. The final target is a second order (nonlinear) discretization with a limiter, for which the components chosen here in multigrid (and the linear GMRES solver) are not appropriate. However, it is a useful discretization for understanding the behavior of our multigrid as a preconditioner and as a solver. We investigate the eigenvalue spectrum of the Richardson iteration matrix (2.4) on a 33^2 and a 65^2 grid. The spectra are presented in Figure 2. As can be seen, most eigenvalues are clustered around 0, only the largest eigenvalues are outside the clustering. The spectral radius determines the convergence of multigrid as a solver, as is well-known. This spectral radius increases on finer grids as can be seen in Figure 2. However, it is found that the eigenvectors belonging to the larger eigenvalues are very soon captured into the Krylov subspace when multigrid is used as a preconditioner, and that therefore the convergence is accelerated considerably.

Figure 2 *The eigenvalue spectra for the rotating convection-diffusion problem on two consecutive grid sizes.*



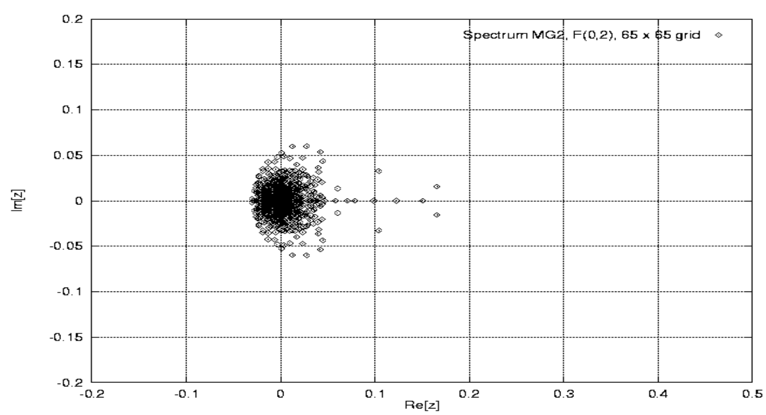


Table 1 presents the convergence results of multigrid as a solver and as a preconditioner on three different grid sizes, 129^2, 257^2 and 513^2. It can be seen that multigrid used as preconditioner handles this test case with dominating convection very well. Very satisfactory convergence associated with small elapsed times is found in most cases with the F-cycle. (With the V-cycle used as a preconditioner the best elapsed times are found, but the number of iterations is increasing with a higher degree than with the F-cycle for large grid sizes.)

An interface problem. Next, an interface problem is considered. This type of problems has been investigated with multigrid, for example in [Zee90]. The problem to be solved

Table 1 *Iterations (n) and elapsed time in seconds for the rotating convection-diffusion equation.*

grid: / method:	129^2	257^2	513^2
multigrid as solver	(15) 5.1	(20) 10.1	(29) 24.8
multigrid as preconditioner	(10) 3.5	(12) 6.1	(16) 14.3

looks as follows:

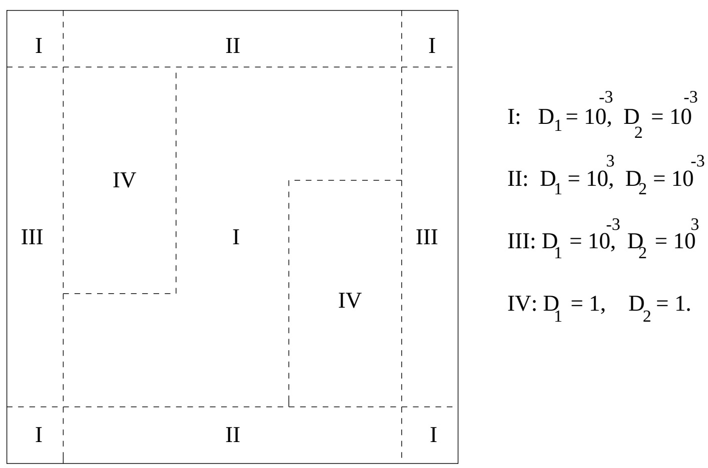
$$\frac{\partial}{\partial x} D_1 \frac{\partial \phi}{\partial x} + \frac{\partial}{\partial y} D_2 \frac{\partial \phi}{\partial y} = 1 \text{ on } \Omega = (0,1) \times (0,1) \ . \tag{3.9}$$

Dirichlet conditions are used:

$$\phi = 1 \quad \text{on } \{x \leq \frac{1}{2} \wedge y = 0\} \text{ and on } \{x = 0 \wedge y \leq \frac{1}{2}\}; \text{ elsewhere } \phi = 0. \tag{3.10}$$

The computational domain with the values of the jumping diffusion coefficients D_1 and D_2 is presented in Figure 3. The discretization is vertex-centered and all diffusion

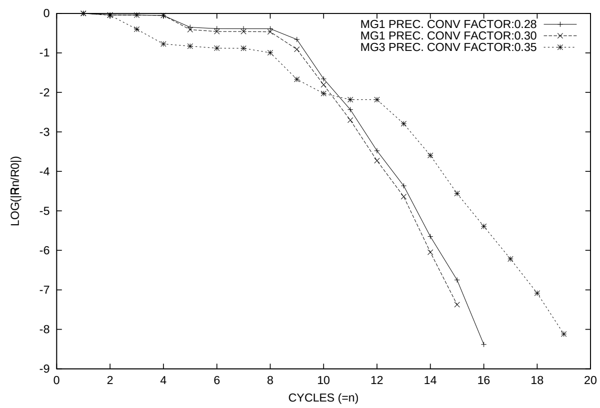
Figure 3 The domain for the interface problem



coefficients are assumed in the vertices. For the discretized diffusion coefficient between two vertices the harmonic average of the neighboring coefficients is taken.

Clearly multigrid can solve many interface problems, presented for example in [Zee90]. Here we constructed a difficult problem, where the robust components of our multigrid method are not satisfactory. The Krylov acceleration is really needed for convergence. The eigenvalue spectrum obtained with multigrid is presented in Figure 5. In Figure 5 we see two eigenvalues close to 1, so multigrid convergence is already very slow on this coarse grid. The convergence of GMRES(20) with multigrid as a preconditioner on the 33^2 grid is shown in Figure 4. Multigrid as a preconditioner is

Figure 4 The convergence of GMRES(30) with multigrid preconditioner, 33^2 grid.



converging well. In Figure 4 it can be seen that for the problem on a 33^2 grid the first 9 GMRES iterations do not reduce the residual very much, but after iteration 9 fast convergence is obtained. In our analysis of the evolution of the Krylov subspace it is found that the vector belonging to a second eigenvalue of $I - AK^{-1}$ around 1 is obtained in the Krylov space in the 9th iteration, and then GMRES starts converging very rapidly. For this test problem the convergence of the preconditioned Krylov methods with the multigrid preconditioner on three very fine grids is presented. The number of GMRES(20) iterations (n) and the elapsed time are presented in Table 2. The GMRES convergence is influenced by the fact that the restart parameter is 20; a larger parameter results in faster convergence. Again the F-cycle is preferred for its robustness and efficiency.

Figure 5 The eigenvalue spectrum for the interface problem on a 33^2 grid, F(0,2) cycle.

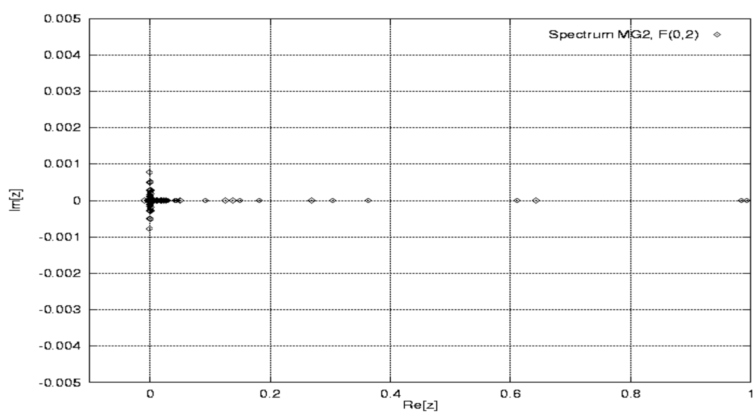


Table 2 *GMRES(20) iterations (n) and elapsed time in seconds for the interface problem.*

grid:	257^2	513^2	769^2
cycle:	GMRES	GMRES	GMRES
F	(34) 19.0	(33) 30.6	(36) 52.9

4 Conclusion

In the present work a multigrid method has been evaluated as a solver and as a preconditioner for GMRES. The problems investigated were singularly perturbed. The behavior of the multigrid method is much more robust when it is used as a preconditioner, since then the convergence is not sensitive to parameter changes. For the test problems many of the eigenvalues of a multigrid iteration matrix are clustered around the origin. In some cases there are some isolated large eigenvalues which limit the multigrid convergence, but are captured nicely by a Krylov acceleration technique. The most efficient results are obtained when the method is used as a preconditioner. The multigrid F-cycle is used, since it is robust and efficient. The convergence behavior can be well understood by investigating the eigenvalue spectrum of the iteration matrix.

REFERENCES

[Den83] Dendy Jr. J. (1983) Black box multigrid for nonsymmetric problems. *Appl. Math. Comp.* 13: 261–283.

[Hac85] Hackbusch W. (1985) *Multi-grid methods and applications.* Springer, Berlin.

[HCH+96] Hempel R., Calkin R., Hess R., Joppich W., Keller U., Koike N., Oosterlee C., Ritzdorf H., Washio T., Wypior P., and Ziegler W. (1996) Real applications on the new parallel system NEC Cenju–3. *Par. Comput.* 22: 131–148.

[Ket82] Kettler R. (1982) Analysis and comparison of relaxation schemes in robust multigrid and preconditioned conjugate gradient methods. In Hackbusch W. and Trottenberg U. (eds) *Multigrid Methods*, volume 960 of *Lect. Notes in Math.*, pages 502–534. Springer, Berlin.

[MFL+91] McBryan O., Frederickson P., Linden J., Schüller A., Solchenbach K., Stüben K., Thole C., and Trottenberg U. (1991) Multigrid methods on parallel computers - a survey of recent developments. *IMPACT Comp. Science Eng.* 3: 1–75.

[RS87] Ruge J. and Stüben K. (1987) Algebraic multigrid (AMG). In McCormick S. (ed) *Multigrid Methods*, volume 5 of *Frontiers in Appl. Math.*, pages 73–130. SIAM Press, Philadelphia.

[SS86] Saad Y. and Schultz M. (1986) GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Comput.* 7: 856–869.

[Wes92] Wesseling P. (1992) *An introduction to multigrid methods.* John Wiley, Chichester.

[Zee90] Zeeuw de P. (1990) Matrix-dependent prolongations and restrictions in a blackbox multigrid solver. *J. Comp. Appl. Math.* 33: 1–27.