Contemporary Mathematics
Volume 157, 1994

ON THE ROBUSTNESS AND EFFICIENCY OF THE FULLY ADAPTIVE MULTIGRID METHOD

U. RÜDE

August 20, 1992

ABSTRACT. The *fully adaptive multigrid method* (FAMe) is a finite element based elliptic solver integrating self-adaptivity, error estimation and efficient iterative solution. Refined elements are not restricted to predetermined regions and need not be grouped in patches. Instead, whether an element is refined, is decided individually for each element using an integrated error indicator. The refinement process induces a multilevel structure and therefore a natural decomposition of the solution space into a nested sequence. This can be exploited to define an efficient solver and error estimator.

1. INTRODUCTION

This paper presents a description of the *virtual global grid refinement* technique and of the *multilevel adaptive iteration* as the basic concepts of a fully adaptive multigrid method. This idea is related to multigrid, hierarchical bases, and multilevel Schwarz methods, see e.g. Dryja and Widlund [2], Bramble, Pasciak and Xu [1], and Yserentant [11].

Virtual global grids are a *recursive infinite* data structure suitable for adaptive mesh refinement algorithms. The infinite structure can be handled using the *multilevel adaptive relaxation*, a variant of relaxation techniques employing a so-called *lazy evaluation*. Mathematically this leads to an incremental construction of a finite element approximation space and the corresponding finite element solution.

In the core of the paper we demonstrate that the resulting algorithm is fast and robust. The analysis is based on the appropriate multilevel additive Schwarz operator (BPX-operator). It can be shown that the condition number of the multilevel operator is bounded uniformly, independent of the number of levels, and this in turn leads to fast convergence of iterative solvers. Furthermore, it can be used to construct simple error estimates that are based on the multilevel structure. These results show that the algorithm is robust, and motivate the virtual global grid mesh refinement strategy.

1991 *Mathematics Subject Classification.* 65N55, 65N15, 65N50.

Key words and phrases. Fully adaptive multigrid, mesh refinement, virtual global grids, error estimates, multilevel additive Schwarz method.

Supported by DFG Stipend Ru 422/3-1

This paper is in final form and no version of it will be submitted for publication elsewhere.

0271-4132/94 $1.00 + $.25 per page

```
proc SequentialAdaptiveRelaxation( θ, x, S̃ )
        assert( S̃ ⊃ S(θ, x) )
        while( S̃ ≠ ∅ )
                pick i ∈ S̃
                S̃ ← S̃ \ {i}
                if |θ_i(x)| > θ then
                        x ← x + r_i
                        S̃ ← S̃ ∪ Neigh(i)
                end if
                assert( S̃ ⊃ S(θ, x) )
        end while
        assert( S̃ = ∅ )
end proc
```

FIGURE 1. Sequential adaptive relaxation

2. VIRTUAL GLOBAL GRIDS AND MULTILEVEL ADAPTIVE ITERATION

For a polygonal bounded domain $\Omega \subset \mathbf{R}^2$ we assume that a suitable initial triangulation $\mathcal{T}_0$ is given. A regular refinement $\mathcal{T}_1$ of $\mathcal{T}_0$ is defined by connecting the midpoints of each of the triangles in $\mathcal{T}_0$, thus partitioning each triangle into four congruent subtriangles. If this process is repeated, it generates an infinite sequence of nested triangulations $\{\mathcal{T}_i\}_{i \in \mathbf{N}_0}$. This infinite, recursive construction of meshes is the *virtual global grid* structure. Of course it cannot be represented directly on any finite machine. Therefore, we adopt the notion of *lazy evaluation*, where portions of the mesh are only allocated, when required by the solution process.

We distinguish between *live nodes*, that is nodes requiring individual values, and *ghost nodes* that have default values only. Ghosts values need not be stored explicitly, but can be calculated from nodes on coarser grids, whenever they are needed. It is natural to define the value for a ghost node by (recursive) interpolation. In hierarchical basis representation, this structure becomes particularly elegant, because ghost nodes are simply characterized by having a vanishing value. A more detailed description of this mesh adaption concept is given by Rüde [6, 7].

The virtual global grid data structure must be complemented with an effective solution algorithm and error indicator. We propose a multilevel iteration as introduced in [6, 7]. The core of this method consists of *adaptive relaxation*, as shown in the program fragment of Fig. 1.

The adaptive relaxation is based on the concept of an *active set* $\tilde{S}$. The idea is that relaxation needs only be performed on nodes (equations) with large residuals. The active set is the means to find these nodes efficiently.

In each elementary relaxation step, an arbitrary node is selected from the current active set. If relaxation leads to a large change from the current value for the node, the value is updated. Then, however, all neighboring nodes must be set active.

If the relaxation produces only a small change, then the update is skipped, and all neighbors keep their status. In any case, after the inspection, a node can be deleted

from the active set, because a new relaxation would reproduce the same result. Thus the active set grows and shrinks depending on the solution characteristics. The algorithm naturally terminates, if the active set is depleted. This in turn guarantees that all residuals are small.

This single level relaxation must be extended to the multilevel structure, leading to a rather intricate algorithm, described in Rüde [6, 7]. Here we only remark that all manipulations can be organized such that the overhead remains small compared to the numerical work involved.

The adaptive relaxation is an almost ideal supplement of the virtual global grid structure. If a ghost becomes *active* it must be set *live*. This simple concept produces a very effective mesh adaption technique, because each such change can be interpreted as an enlargement of the approximation space.

3. Theoretical background

In this section we now outline the theoretical background for the fully adaptive multigrid concept. We focus on the prototype case of a symmetric, elliptic, second order partial differential equation subject to homogeneous Dirichlet boundary conditions. In weak formulation this can be written as: Find $u \in H_0^1(\Omega)$, such that

$$a(u,v) = \Phi(v) \text{ for all } v \in H_0^1(\Omega), \tag{1}$$

where $a(\cdot,\cdot)$ is an H^1-elliptic bilinear form and Φ is a continuous, linear functional. The hierarchical mesh structure leads to a nested sequence of finite element spaces

$$V_0 \subset V_1 \subset \cdots \subset V = H_0^1. \tag{2}$$

We are interested in solving a finite element problem

$$a(u_K, v_K) = \Phi(v_K) \text{ for all } v_K \in V_K. \tag{3}$$

By introducing the nodal basis $\mathcal{B}_K$ of V_K this reduces to an algebraic system

$$A_K u_K = f_K, \tag{4}$$

where we interchangeably use u_K to stand for the function in V_K and its coefficient vector in the representation with respect to $\mathcal{B}_K$.

Multilevel algorithms can be introduced on the basis of projection operators $P_j : V \longrightarrow V_j$ that are defined by

$$(P_j u, v_j)_j = a(u, v_j) \text{ for all } v_j \in V_j, \tag{5}$$

where $(\cdot,\cdot)_j$ is a "natural" bilinear form in V_j. Algebraically, for $u_K \in V_K$, $K \geq j$, that is the projection P_j restricted to V_K. can be written as

$$P_j = I_j^K D_j^{-1} I_K^j A_K, \tag{6}$$

where $I_j^K = (I_K^j)^T$ is the prolongation defined by the natural embedding of V_j in V_K, $j \leq K$, and D_j is a matrix that defines the scalar product $(\cdot,\cdot)_j$ with respect

to the basis $\mathcal{B}_j$. The *additive Schwarz operator* of V_K with respect to the subspace system $\{V_j\}_{j=0,\ldots,K}$ is defined by

$$P = \sum_{j=0}^{K} P_j. \tag{7}$$

P depends on the the scalar products $(\cdot,\cdot)_j$, that is the choice of the D_j. Typically, D_j is chosen as an operator that is simple to invert but such that $(\cdot,\cdot)_j$ is equivalent to the L_2-inner-product in V_j, scaled by a constant. Possible choices are $D_j = \operatorname{diag}(A_j)$, that is the diagonal part of the stiffness matrix on level j. An alternative is the choice as $D_j = \operatorname{diag}(4^j)$, that is a diagonal matrix with constant entries. These two alternatives lead to the *multiple level diagonal scaling* (MDS) and the *BPX*-preconditioner, respectively. For more details see Zhang [12] and Bramble, Pasciak, and Xu [1].

The initial level plays a special role. To make the theory independent of V_0, we choose $D_0 = A_0$. Algorithmically this corresponds to solving the coarsest level problem in each application of P. Analogous to P_j and P, we finally define ϕ_j and ϕ by

$$(\phi_j, v_j)_j = \Phi(v_j) \text{ for all } v_j \in V_j, \tag{8}$$

$$\phi = \sum_{j=0}^{K} \phi_j. \tag{9}$$

Under the assumptions made in this paper, the following theorem holds.

Theorem 1 (Oswald). *The variational problem (1) is equivalent to the operator equation*

$$Pu = \phi. \tag{10}$$

There exist constants $0 < c_1 \le c_2 < \infty$ independent of the number of levels K, such that the spectrum of P_V can be estimated by

$$c_1 \le \lambda_{\min}(P_V) \le \lambda_{\max}(P_V) \le c_2. \tag{11}$$

Proof. See Oswald [3] □

Similar structures have been studied by Xu [10], Bramble, Pasciak and Xu [1], Yserentant [11], Zhang [12], and Dryja and Widlund [2].

We now study applications of this theorem for iterative solvers and error estimators. Clearly, the explicit construction of P is too expensive, because P is not sparse. However, P can be applied to a vector according to its definition as a sum. Thus iterative techniques, like the conjugate gradient method can be applied, and will yield a solver whose rate of convergence is independent of the number of levels. Thus the multilevel structure leads to asymptotically optimal *efficient* solvers.

The application of the operator P seems to be parallelizable. However, typically the individual terms in the sum should not be calculated independently. The calculation of the coarsest level term produces all finer level contributions as a side effect, so that computing them in parallel makes no sense. For domain decomposition based additive Schwarz methods this is different. There the $P_j u$ may be

computed in parallel. However, to have mesh-independent convergence rate without the subproblems becoming too large, the domain decomposition method must also be augmented with a coarse mesh.

Assume that $u_K, u_K^* \in V_K$ and $u_K^* = A_K^{-1} f_K$ is the correct solution of the level K equations. We introduce the *scaled residual of level j* by

$$\bar{r}_j = D_j^{-1} I_K^j A_K (u_K^* - u_K). \tag{12}$$

Note that

$$P(u_K^* - u_K) = \sum_{j=0}^{K} I_j^K \bar{r}_j. \tag{13}$$

We have thus related known quantities (the residuals) to unknown ones (the error). Using the properties of the additive Schwarz operator we derive error estimates in the L_2- and the H^1-norm.

Theorem 2. *There exist constants* $0 < c_0 \le c_1 < \infty$, *such that*

$$c_0 \left\| \sum_{j=1}^{K} I_{V_j}^V \bar{r}_j \right\|_{L_2} \le \|u - u^*\|_2 \le c_1 \left\| \sum_{j=1}^{K} I_{V_j}^V \bar{r}_j \right\|_{L_2} \tag{14}$$

and

$$c_0 \sum_{j=1}^{K} \bar{r}_j^T D_j \bar{r}_j \le \|u - u^*\|_{H^1}^2 \le c_1 \sum_{j=1}^{K} \bar{r}_k^T D_j \bar{r}_j. \tag{15}$$

Proof. See Rüde [8]. For the energy norm estimate see also Oswald [4]. □

These two error estimates show that in the multilevel context residuals are directly related to the *algebraic*[1] error. Note that this is not true for a single level, where residuals must be multiplied with the condition number to give error bounds.

Theorem 2 shows that in the multilevel context it suffices to eliminate all *large* residuals. Conversely, when no large residuals remain, the iteration will have produced an approximation with small *algebraic* error. Thus the multilevel adaptive relaxation is a *robust* algorithm, because it has been constructed such that the error estimate is satisfied at termination.

Beyond this, equations (14, 15) have the additional feature that they are independent of the number of levels. They formally also produce estimates for the *discretization* error, if we let the number of levels K tend to infinity. In the context of hierarchical basis and virtual global grids, this can be exploited, if we use regularity assumptions to bound ghost node residuals by those available for live nodes on coarser levels.

Furthermore, the relaxation of an individual equation automatically produces change in the error bound, and thus estimates the local contribution of the node to the error. In the case of a live node, a relaxation leads to a reduction of the algebraic error, for a ghost node also to a reduction of the discretization error. As

[1] We distinguish between *algebraic* errors, that is errors of the iterative linear system solver with respect to the true solution of the algebraic equations, and the *discretization* error, that is the error of the (correct) algebraic solution with respect to the true solution of the differential equation.

both kinds of errors are estimated by the same mechanism, we also have criteria when to switch between (algebraic) iteration and mesh refinement. In fact, the multilevel adaptive relaxation and virtual global grids can switch between iteration and refinement on a node by node basis.

The adaptive relaxation can be shown to produce an efficient algorithm by its property to relax the large residuals only. It can be interpreted as a more efficient variant of the Gauss-Southwell relaxation (see Southwell [9]).

The fully adaptive multigrid method terminates naturally, when the large residuals have been removed on all levels. Sophisticated data structures (see Rüde [5]) can be used to trace all the changes and dependencies within and between levels. This process by definition provides guaranteed accuracy and robustness of the algorithm.

4. Conclusions

In this paper we have outlined how recent results from the theory of additive Schwarz methods and the multilevel splitting of finite element spaces can be exploited to construct effective multilevel adaptive algorithms that are efficient and robust.

References

1. James Bramble, Joseph Pasciak, and Jinchao Xu, *Parallel multilevel preconditioners*, Math. Comp. **31** (1990), 333–390.
2. M. Dryja and O.B. Widlund, *Multilevel additive methods for elliptic finite element problems*, Parallel Algorithms for Partial Differential Equations, Proceedings of the Sixth GAMM-Seminar, Kiel, January 19-21, 1990 (Braunschweig) (W. Hackbusch, ed.), Vieweg-Verlag, 1991.
3. P. Oswald, *Norm equivalencies and multilevel schwarz preconditioning for variational problems*, Forschungsergebnisse Math/92/01, Friedrich Schiller Universität, Jena, January 1992.
4. ———, *Stable splittings of Sobolev spaces and fast solution of variational problems*, Forschungsergebnisse Math/92/05, Friedrich Schiller Universität, Jena, May 1992.
5. U. Rüde, *Data structures for multilevel adaptive methods and iterative solvers*, Bericht I-9217, Institut für Informatik, TU München, May 1992.
6. ———, *On the multilevel adaptive iterative method*, Preliminary proceedings of the 2nd Copper Mountain Conference on Iterative Methods, April 9–14, 1992 (T. Manteuffel, ed.), University of Colorado at Denver, 1992, Accepted for publication in SIAM J. Sci. Stat. Comput. and also available as TU-Bericht I-9216.
7. ———, *Fully adaptive multigrid methods*, SIAM J. Numer. Analysis **30** (1993), no. 1, 230–248.
8. ———, *Mathematical and computational adaptive techniques for elliptic problems*, Frontiers in Applied Mathematics, vol. 13, SIAM, Philadelphia, 1993, to be published.
9. R.V. Southwell, *Relaxation methods in engineering science – a treatise in approximate computation*, Oxford University Press, 1940.
10. J. Xu, *Theory of multilevel methods*, Report AM48, Department of Mathematics, Pennsylvania State University, July 1989.
11. H. Yserentant, *Two preconditioners based on the multi-level splitting of finite element spaces*, Numer. Math. **58** (1990), 163–184.
12. Xuejun Zhang, *Multilevel additive Schwarz methods*, Tech. Report 582, New York University, Courant Institute, Department of Computer Science, 1991.

Institut für Informatik, Technische Universität München, Arcisstr. 21, D-8000 München 2, Germany

E-mail address: ruede@informatik.tu-muenchen.de