On Multigrid Convergence in the Indefinite Case

By R. A. Nicolaides*

Abstract. Previous results of the author on the convergence of the multigrid method
for positive definite elliptic problems are generalized to cover the indefinite case.

I. Introduction. In [8] (see also [7]) we suggested an algorithm for solving
discrete finite element systems arising from positive definite problems of elliptic type
and showed that it had very favorable properties. This type of algorithm has come to
be known as being of multigrid type and so we shall use this term here. A considerable
amount of work has been done by others since submission of [8] for publication, and
we wish here to indicate some of this work. In the finite difference area there is com-
prehensive (although not rigorous) work by Brandt [1], [2] covering a large number
of topics, including adaptive computation schemes. The basic tool used in these
investigations is Fourier analysis; and it is used to analyze not only the smoothing,
but also the fine to coarse and coarse to fine grid transfer operations of the multigrid
method. In the finite element area, in addition to work by Bank and Dupont [3],
papers by Hackbusch [4], [5], [6] extend the basic multigrid theorems to nonself-
adjoint and other problems. This latter work is of an exceedingly theoretical character,
largely because it addresses the difficult question of the effect of the regularity of the
given problem on the multigrid algorithm. In [8] the question was avoided by assum-
ing triangulations such that the L^2 estimate (see below for notations)

$$\|u - u^h\|_{L^2(\Omega)} \leqslant Kh^2 \|f\|_{L^2(\Omega)}$$

holds on each, where h is a mesh size parameter. The methods of [4] are thus of a
different character from those of [8].

The purpose of the present note is to extend to a wider class of problems the
$O(N)$ type convergence results of [8]. The extension is to the indefinite case, i.e.,
nonsingular problems for which the operator has not only positive, but also negative
eigenvalues (the Helmholtz equation is a second order example). The usual iterative
methods will fail when applied to such problems. True, they can in principle be solved
by iterating with the discrete equations $K_p^T K_p u^p = K_p^T f^p$ or $K_p K_p^T V^p = f^p$, where
$u^p = K_p^T V^p$ (continuing to use the notations of [8] as we shall do throughout),
where K_p is the system, or stiffness matrix of the finite element method. This is done,

however, at the expense of squaring the condition numbers of the algebraic problem, with a consequent loss of convergence speed. Furthermore, such a least squares formulation while superfically attractive for use with Gauss-Seidel smoothing, may have rather poor smoothing properties. As it turns out, it is not necessary to use any new smoothing methods. The method of [8] for which the iteration matrix is $(I - \alpha K_p)$ may be used, as may other schemes. The corrections from the coarse grids are sufficiently able to deal with the amplified low frequency error components, provided the algorithm is operated sensibly. Our principal goal is to extend the two main results of [8] so that they also apply to the new class of problems. As a matter of fact, it is necessary only to alter one hypothesis and somewhat rework the proof of one theorem in [8] in order to make the generalizations sought. Once this is done the two main results referred to above continue to hold verbatim. To repeat all the preliminaries, definitions and lemmas which lead to the theorem to be modified would mean an overly large amount of redundancy in the present note which we wish to avoid. Therefore, extensive references to [8] cannot be avoided. For each reference, we shall use exactly the notations of [8]. The more important ones are given again below, but for the others, especially a number of constants occurring in the proofs, we must refer the reader to the original source. It will also be assumed that the reader is familiar with the basic multigrid algorithm as given in [8, Section 4].

   II. Results. Let us now recall some of the notations of [8]. K_p is the discrete system matrix on the grid where the solution to the continuous problem is sought, while K_q, $q = 1, 2, \ldots, p - 1$, are the system matrices on the coarser grids. The vectors of unknowns on the subgrids are denoted by u^q or x^q possibly with additional superscripts denoting position in a hierachy of multigrid approximations. Thus, for example $x^{q,k,i}$ denotes a vector defined on a grid G_q in the ith smoothing sweep of the kth iteration of the multigrid algorithm. Norm symbols used are $\| \ \|$, and sometimes $\| \ \|_{l^2}$ for the l^2 norm and $\| \ \|_m$ for the norm on $H^m(\Omega)$. $(\cdot, \cdot)$ is the usual l^2 inner product. h_i denotes the mesh size parameters on G_i, respectively, and N_i the number of unknowns in the various linear systems. K_q has dimensions $N_q \times N_q$. We also write $R(q)$ for the space of N_q tuples. All of these conventions are fully explained in [8].
   We turn now to the necessary modification which has to be made in the hypotheses of [8]. It is the positive definiteness which has to be altered. We require here that the principal part of the form $a(u, u) - 2(u, f)$ should satisfy the Gårding inequality

(2.1) $$a(u, u) \geqslant b_1 \|u\|_m^2 - \bar{c_0}^2 \|u\|_0^2, \quad b_1 > 0.$$

The earlier hypothesis had $\bar{c_0}^2 \equiv 0$ and ensured the positive definiteness of the form $a(u, u)$. The form is now indefinite and is to be rendered stationary, rather than minimized. In addition to (2.1) it is necessary to assume that this extremization problem has a unique solution, as, unlike the definite case, this cannot be deduced from (2.1). It then follows from (2.1) and the basic eigenvalue estimates of the finite

element method [9, Chapter 6] that for all h_q sufficiently small, K_q is nonsingular. The pure Neumann problem is not included in our formulation. In addition to (2.1) the hypotheses H1 and H2 of [8, §3] specifying the L^2 error estimate and the uniformity properties of the basis are assumed to hold. From the uniformity assumption H2 and (2.1) it follows immediately that

$$(2.2) \qquad (K_q u^q, u^q) \geqslant - \lambda_0 \bar{c}_0^2 \|u^q\|_{l^2}, \qquad u^q \in R(q),$$

where $\lambda_0 > 0$ is a lower bound, independent of h_q, to the eigenvalues of the mass matrix (see [8, (6.1)]).

We may now restate the first main result of [8] and prove it under the current conditions. Recall that α' denotes the parameter in the smoothing algorithm, n' the number of iterations of the smoothing algorithm and δ' the accuracy parameter for the subsystem solution.

THEOREM 2.1 [CF. [8, THEOREM 5.1]]. *There exist numbers δ_0 and n_0 not dependent on q and a number α_0 such that for any fixed $\nu \geqslant 2$ with $\alpha' = \alpha_0$, $\delta' = \delta_0$ and $n' = n_0$*

$$\|x^q - x^{q,\nu,0}\| \leqslant \delta_0 \|x^q - x^{q,0,0}\|, \qquad 0 < \delta_0 < 1.$$

Proof. The proof is a verbatim repetition of that in [8] up to and including [8, (7.6)] which we repeat here:

$$\|\epsilon^{q,1,0}\|^2 \leqslant 4[\mu^2 (B_2')^2 \gamma_{1,q,\mu}^{2n'} \|\epsilon_{1,\mu}^{q,0,0}\|^2 + (B_2')^2 \gamma_{2,q,\mu}^{2n'} \|\epsilon_{2,\mu}^{q,0,\,0}\|^2]$$

$$+ 2(B_3')^2 (\delta')^2 \|\epsilon^{q,0,n'}\|^2.$$

($\gamma_{i,q,\mu}^{2n'}$, $i = 1, 2$, are bounds on the smoothing operator in a pair of subspaces dependent on the parameter μ.) Define α_0 by the equation

$$\alpha_0 = \alpha_0(\mu) = \frac{2}{\mu B_1^{-1} h_q^{-2m} + B_2 h_q^{-2m}} = \frac{2 h_q^{2m}}{\mu B_1^{-1} + B_2}.$$

In this case, with $\bar{c}_0^2 \lambda_0 \equiv c_0^2$,

$$\gamma_{1,q,\mu} \leqslant 1 + \frac{2 c_0^2 h_q^{2m}}{\mu B_1^{-1} + B_2} \leqslant 1 + \frac{2 c_0^2 h_q^{2m}}{B_2}.$$

Denote the right member of this inequality by M_1, dependent on h but not on μ. Also

$$\gamma_{2,q,\mu} \leqslant \frac{B_2 - \mu B_1^{-1}}{B_2 + \mu B_1^{-1}} = \theta(\mu),$$

where θ does not depend on h. Then

$$2(B_3')^2 (\delta')^2 \|\epsilon^{q,0,n'}\|^2 \leqslant 2(B_3')^2 (\delta')^2 \gamma_{1,q,\mu}^{2n'} \|\epsilon^{q,0,0}\|^2,$$

and

$$\|\epsilon^{q,1,0}\|^2 \leqslant 4(B_2')^2\mu^2 M_1^{2n'}\|\epsilon_{1,\mu}^{q,0,0}\|^2 + 4(B_2')^2\theta^{2n}\|\epsilon_{2,\mu}^{q,0,0}\|^2$$

(2.3)

$$+ 2(B_3')^2(\delta')^2 M_1^{2n'}\|\epsilon^{q,0,0}\|^2.$$

Now, let $L > 1$ be any given fixed number. Select $\delta' = \delta_0$ where δ_0 is some solution of

$$2(B_3')^2 L \leqslant \tfrac{1}{2}\delta_0^{2/\nu-2}, \qquad \delta_0 \neq \delta_0(h),$$

and $\mu = \mu_0$ such that

$$4(B_2')^2\mu_0^2 L \leqslant \tfrac{1}{2}\delta_0^{2/\nu}.$$

μ_0 and δ_0 are independent of h. Additionally, let n_0 be so large that

$$4(B_2')^2\theta(\mu_0)^{2n_0} \leqslant \tfrac{1}{2}\delta^{2/\nu}$$

and, finally, $h \leqslant h_0$ where h_0 is chosen so that

$$M_1(h_0)^{2n_0} \leqslant L.$$

Substituting these values into (2.2) gives

$$\|\epsilon^{q,1,0}\|^2 \leqslant \tfrac{1}{2}\delta_0^{2/\nu}(\|\epsilon_{1,\mu}^{q,0,0}\|^2 + \|\epsilon_{2,\mu}^{q,0,0}\|^2) + \tfrac{1}{2}\delta_0^{2/\nu}\|\epsilon^{q,0,0}\|^2$$

$$= \delta_0^{2/\nu}\|\epsilon^{q,0,0}\|^2.$$

Then ν repetitions of the algorithm gives

$$\|\epsilon^{q,1,0}\| \leqslant \delta_0\|\epsilon^{q,0,0}\|,$$

and the result is proved.

All the subsequent results of [8] now go through unchanged. In particular the $O(N)$ result, [8, Theorem 8.1], is valid with the new hypothesis, and thus the main theorems hold for the indefinite case.

There is one point to which we must draw attention. In the definite case it is safe (though not efficient) to do as many smoothing sweeps as desired. The algorithms cannot diverge on this account. In the indefinite case this is no longer true. Excessive applications of the smoothing algorithm can drive up the low frequency error components to the point where the coarse grid corrections cannot control them. This feature makes the algorithm more amenable to divergence and somewhat harder to operate than in the definite case, which is almost foolproof. The problem is most acute when c_0^2 is large in some relative sense. Nevertheless it seems that the multigrid method will furnish us with a powerful tool for solving indefinite problems and make possible the solution of problems which cannot be solved by the elimination methods in present use due to the large storage and execution inefficiencies inherent in these latter algorithms. Work on implementation is in progress.

With this, we conclude the present analysis of the indefinite problem. It seems reasonable to the author that similar $O(N)$ type results can be proved for any properly

posed elliptic system, not merely the scalar symmetric case we have considered here and in [8].

Acknowledgment. My best thanks are due to James Ortega for continuous help and advice.

Department of Mathematics
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

1. A. BRANDT, "Multi-level adaptive solution to boundary value problems," *Math. Comp.,* v. 31, 1977, pp. 333–391.

2. A. BRANDT, *Multi-level Adaptive Techniques (MLAT): Ideas and Software,* Proc. Conf. Math. Software, Math. Res. Center, Wisconsin, 1977.

3. RANDOLPH A. BANK & T. DUPONT, *An Optimal Order Process for Solving Elliptic Finite Element Equations,* Department of Mathematics, University of Chicago, 1978. (Manuscript.)

4. W. HACKBUSCH, "On the convergence of a multigrid iteration applied to finite element equations," *Numer. Math.* (To appear.)

5. W. HACKBUSCH, "On the computation of eigenvalues and eigenfunctions of elliptic operator by means of a multigrid method," *Numer. Math.* (To appear.)

6. W. HACKBUSCH, *A Fast Numerical Method for Elliptic Boundary Value Problems with Variable Coefficients,* 2nd GAMM Conference on Numer. Math. in Fluid Mechanics, DFVLR, Köln, October 1977.

7. R. A. NICOLAIDES, *On a Proposal for solving Discrete Finite Element Equations,* Proc. Conf. Linear Algebra and Finite Elements, Brunel University, June 1975.

8. R. A. NICOLAIDES, "On the l^2 convergence of an algorithm for solving finite element equations," *Math. Comp.,* v. 31, 1977, pp. 892–906.

9. G. STRANG AND G. FIX, *An Analysis of the Finite Element Method,* Prentice-Hall, Englewood Cliffs, N. J., 1973.