LA-UR 97-2551

CONF-9707106--

Approved for public release; distribution is unlimited

Title: A Multi-Level Method for Sparse Linear Systems

Author(s): Yair Shapira

RECEIVED
AUG 27 1997
OSTI

Submitted to: SIAM's 45th Anniversary Meeting
July 14-18, 1997
Stanford, CA

MASTER

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Los Alamos
National Laboratory

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. The Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Form 836 (10/96)

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

A Multi-Level Method for Sparse Linear Systems

Yair Shapira*

Abstract

A multi-level method for the solution of sparse linear systems is introduced. The definition of the method is based on data from the coefficient matrix alone. An upper bound for the condition number is available for certain SPD problems. Numerical experiments confirm the analysis and illustrate the efficiency of the method for diffusion problems with discontinuous coefficients with discontinuities which are not aligned with the coarse meshes.

1 Introduction

Multi-level and hierarchical basis methods for the numerical solution of elliptic boundary value problems with possibly discontinuous coefficients are studied extensively in the literature (e.g., [2]–[8], [17]–[18]). Although upper bounds for the condition numbers of these methods are available, they are often derived under the assumption that the discontinuities in the coefficients of the PDE are aligned with the coarse mesh. This might limit significantly the number of levels that may be used. Furthermore, some hierarchical basis methods deteriorate when turning to 3-d problems.

Multigrid methods with matrix-dependent transfer and coarse-grid operators (such as those of [20] [22] [26] [28]) are considered robust for diffusion problems with variable coefficients. The Black Box Multigrid method of [1] [11] [12] [13] is the ancestor of these methods and is efficient for many examples. However, it is limited to structured grids, that is, coefficient matrices representing 3^d-coefficient stencils (where d is the dimension of the domain of the underlying PDE). For this kind of grids, although no analysis is available, Black Box Multigrid was applied successfully to problems with discontinuous coefficients (with no outer acceleration) even when the discontinuities are not aligned with the coarse grids ([11] [13]). A Black-Box Multigrid version which is related to the method of [16] is introduced and analyzed in [23]. It is further generalized to discretizations based on local mesh refinement in [24], where it is also shown that the condition number of the corresponding two-level method is mesh and jump-independent.

*Los Alamos National Laboratory, Mail Stop B-256, Los Alamos, NM 87545, e-mail: yairs@lanl.gov. This work was performed under the auspices of the US Department of Energy.

The above mentioned methods assume that the linear system of equations arises from an elliptic boundary value problem. They inherently use the geometrical properties of the domain on which the PDE is defined for defining the coarse grids. Although the definitions of transfer operators among grids and coarse-grid coefficient matrices in the methods mentioned in the latter paragraph are based on data contained in the coefficient matrix alone, the construction of the coarse grids themselves still depends on the underlying continuous problem. This is a disadvantage when elliptic boundary value problems on completely unstructured grids are considered, where no natural coarse grids are available. The Algebraic Multigrid method of [21] avoids this problem by defining the coarse grids using information contained in the coefficient matrix alone. Algebraic coarsening methods for unstructured meshes are introduced in [19]. A coarsening procedure based on graph theory is described in [14].

In this work a multi-level method for the solution of sparse linear systems is introduced. The definition of the method is based on data contained in the coefficient matrix alone; no underlying continuous problem is assumed. The method is designed so that the two-level analysis of [23] [24] is valid. Furthermore, this analysis is extended here to the multi-level case. (This extension may apply, with slight changes, also for the methods of [23] [24].) Since the method is purely algebraic, the assumptions made for the analysis are algebraic in nature. In particular, it is assumed in the analysis that the coefficient matrices at all levels are diagonally dominant. Although this is not guaranteed even for diagonally dominant problems, it turns out that it holds for the examples tested in [23], [24] and here. The present numerical experiments illustrate the efficiency of the method for diffusion problems with discontinuous coefficients with discontinuities which are not aligned with the coarse meshes.

The contents of the paper are as follows. In Section 2 the multi-level method is introduced. In Section 3 an upper bound for the condition number is given in the symmetric positive definite (SPD) case. In Section 4 numerical experiments are presented. In Section 5 concluding remarks are made.

2 The Multigrid Method

Consider the linear system of equations

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

where N is a positive integer, A is a nonsingular sparse matrix of order N with nonzero main-diagonal elements, b is a given N-dimensional vector and x is the N-dimensional vector of unknowns.

Next, we introduce the multi-level method. Here a 'level' refers to a subset of indices

of variables for which a reduced linear system of equations is considered in the multi-level algorithm.

Let $0 \leq \alpha < 1$ be a small threshold. (The purpose of introducing α is for anisotropic problems, where the standard coarsening fails ([9]); an appropriate choice for α results in coarsening in the strongly coupled directions only, which is suitable for this kind of problems.) Let $A^{(\alpha)}$ be a matrix of the same order as A with

$$A_{i,j}^{(\alpha)} = \begin{cases} 1 & \text{if } \max\{|A_{i,j}/A_{i,i}|, |A_{j,i}/A_{j,j}|\} \geq \alpha, \\ 0 & \text{otherwise.} \end{cases}$$

The coarse level c should be a maximal set of 'weakly-coupled' variable indices; that is, for every $i, j \in c$, $A_{i,j}^{(\alpha)}$ vanish and c cannot be enlarged without violating this requirement. This can be achieved as follows. Let $S \subset \{1, 2, \ldots, N\}$ be an index set. Define the following procedure:

<u>coarsen(S, A, c)</u>:

1. Set $f = F = \emptyset$.
2. Pick some $i \in S$ and set $c = C = \{i\}$.
3. Until $c \cup f = S$, do:
   (a) $F \leftarrow \cup_{i \in C}\{j \in S \mid A_{i,j}^{(\alpha)} \neq 0\} \setminus c \setminus f$.
   (b) $f \leftarrow f \cup F$.
   (c) coarsen$(\cup_{i \in F}\{j \in S \mid A_{i,j}^{(\alpha)} \neq 0\} \setminus c \setminus f, A, C)$.
   (d) If $C = \emptyset$ then pick an $i \in S \setminus c \setminus f$, preferably $i \in \cup_{i \in F}\{j \in S \mid A_{i,j}^{(0)} \neq 0\}$, and set $C = \{i\}$.
   (e) $c \leftarrow c \cup C$.

Of course, the application of this procedure is with $S = \{1, 2, \ldots, N\}$. By induction on $|S|$ (the number of elements in S) one can show that c is indeed a maximal subset of S in the above sense. After completing this application, c and f are modified as follows. f is scanned and if, for some $i \in f$,

$$\frac{\max\left\{A_{i,i} - \Sigma_{j \in f \setminus \{i\} \text{ or } A_{i,j}^{(\alpha)}=0} |A_{i,j}|, \Sigma_{j \in c,\ A_{i,j}^{(\alpha)} \neq 0} |A_{i,j}|\right\}}{|A_{i,i}|} \leq \bar{\alpha} \tag{2}$$

(where $0 \leq \bar{\alpha} < 1$ is a small parameter), then substitute

$$c \leftarrow c \cup \{i\} \quad \text{and} \quad f \leftarrow f \setminus \{i\}. \tag{3}$$

This modification is aimed to guarantee boundedness of the l_2 norm of the submatrix P_{ff} defined below, which is important in the analysis. In the sequel, we refer to the construction of c and f as the 'coarsening procedure'.

Consider, for example, a $(2d+1)$-coefficient stencil arising from a discretization of a second order elliptic boundary value problem on a d-dimensional domain. Then, for sufficiently small α and $\bar{\alpha}$, the coarsening procedure yields the usual red-black coloring of grid points, with c corresponding to red points and f to black ones. As a matter of fact, for any matrix of property-A (with sufficiently small α and $\bar{\alpha}$), the partitioning $c \cup f$ resulting from the above procedure is equivalent to the property-A partitioning.

The above coarsening procedure is inherently sequential. (Parallel coarsening methods were discussed recently in [15].) However, since it involves mostly integer comparisons, its time consumption is negligible.

The partitioning $\{1, 2, \ldots, N\} = c \cup f$ resulting from the above coarsening procedure induces the following block partitioning of A:

$$A = \begin{pmatrix} A_{ff} & A_{fc} \\ A_{cf} & A_{cc} \end{pmatrix} \tag{4}$$

and similarly for other matrices of the same order.

The subset f may be further partitioned in the form $f = f_1 \cup f_2$. Reasonable possible criteria for this might be

$$f_2 = \{i \in f \mid \sum_{j \in c,\ A_{i,j}^{(\alpha)} \neq 0} |A_{i,j}/A_{i,i}| \leq \alpha_1$$

or

$$f_2 = \{i \in f \mid \max_{j \in c,\ A_{i,j}^{(\alpha)} \neq 0} |A_{i,j}/A_{i,i}| \leq \alpha_1\}$$

(where $0 \leq \alpha_1 < 1$ is some parameter), with $f_1 = f \setminus f_2$. Then one may define restriction and prolongation matrices using this partitioning as in [24]. Similarly, one may further split f_1 using a parameter α_2 greater than α_1. By repeating this, one can obtain a decomposition of f as the union of d disjoint subsets (where d is some positive integer – usually the dimension of the domain of the underlying PDE). The restriction and prolongation matrices corresponding to this decomposition are defined in [23]. For simplicity, however, we do not consider these implementations here. (The analysis in Section 3 applies to these versions as well, provided that it is modified in the spirit of [23] [24].)

For any set $g \subset \{1, 2, \ldots, N\}$, let $J_g : l_2(\{1, 2, \ldots, N\}) \to l_2(g)$ denote the injection operator defined by

$$(J_g v)_j = v_j, \ v \in l_2(\{1, 2, \ldots, N\}), \ j \in g.$$

For any matrix M, $M = (m_{i,j})_{1 \leq i \leq K,\ 1 \leq j \leq L}$, let the absolute value of M be defined by $|M| = (|m_{i,j}|)_{1 \leq i \leq K,\ 1 \leq j \leq L}$ and the diagonal matrix of row-sums of M by

$$rs(M) = diag(\sum_{j=1}^{L} m_{i,j})_{1 \leq i \leq K}.$$

Let $\acute{A}$ be a matrix of the same order as A with

$$\acute{A}_{i,j} = \begin{cases} A_{i,j} & \text{if } A_{i,j}^{(\alpha)} = 0 \text{ and either } i \in f \text{ and } j \in c \text{ or } i \in c \text{ and } j \in f \\ 0 & \text{otherwise.} \end{cases}$$

Define

$$\tilde{A} = \begin{pmatrix} diag(A_{ff}) - rs(|A_{ff} - diag(A_{ff})|) - rs(|\acute{A}_{fc}|) & A_{fc} - \acute{A}_{fc} \\ A_{cf} - \acute{A}_{cf} & rs(|A_{cf} - \acute{A}_{cf}|) \end{pmatrix}.$$

If, for some $i \in f$, $(\tilde{A}_{ff})_{i,i}/|A_{i,i}| \leq \bar{\alpha}$, then substitute $(\tilde{A}_{ff})_{i,i} \leftarrow rs(|A_{fc} - \acute{A}_{fc}|)_{i,i}$. From (2)–(3) it follows that, after such a substitution, $(\tilde{A}_{ff})_{i,i}/|A_{i,i}| > \bar{\alpha}$. This guarantees that $\tilde{A}_{ff}$ is nonsingular.

Define the matrices P ('prolongation') and R ('restriction') by

$$P^{-1} = \begin{pmatrix} \tilde{A}_{ff} & \tilde{A}_{fc} \\ 0 & I \end{pmatrix} \quad \text{and} \quad R^{-1} = \begin{pmatrix} \tilde{A}_{ff} & 0 \\ \tilde{A}_{cf} & I \end{pmatrix}, \tag{5}$$

where I denotes the identity matrix of a suitable order, or

$$P = \begin{pmatrix} \tilde{A}_{ff}^{-1} & -\tilde{A}_{ff}^{-1}\tilde{A}_{fc} \\ 0 & I \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} \tilde{A}_{ff}^{-1} & 0 \\ -\tilde{A}_{cf}\tilde{A}_{ff}^{-1} & I \end{pmatrix}. \tag{6}$$

(In the nonsymmetric case, one may also consider deriving P from the symmetric part of A rather than A, as in [12].)

Define the coarse-level coefficient matrix

$$Q = J_c RAPJ_c^t$$

and the matrices

$$W = R_{ff} diag(A_{ff}) P_{ff} \quad \text{and} \quad F = blockdiag(W, Q).$$

Note that R, P and F are of the same order as A; however, the complexity of inverting F is much smaller than that of inverting A. It is assumed here that Q is nonsingular. This holds, e.g., when A is SPD. The two-level iteration for the solution of (1) is defined by

$$\text{TL}(x_{in}, A, b, x_{out}): \quad x_{out} = x_{in} + PF^{-1}R(b - Ax_{in}). \tag{7}$$

Note that, when A is of property-A and α and $\bar{\alpha}$ are sufficiently small, the TL method is a direct solver. Indeed, in this case $\tilde{A}_{fc} = A_{fc}$, $\tilde{A}_{cf} = A_{cf}$ and A_{ff} is diagonal, which implies $\tilde{A}_{ff} = A_{ff}$. Therefore, $A = R^{-1}FP^{-1}$ is just the block-LDU decomposition of A..

Let D be a nonsingular diagonal matrix of order N. If the TL method is applied to the system

$$D^{-1}Ax = D^{-1}b \tag{8}$$

rather than (1), then it is mathematically unchanged (see [24], Lemma 2).

The TL method may be supplemented with relaxations before and after it in the spirit of multigrid. However, this option is not considered here. Instead, we turn to the multi-level method.

Suppose that A represents the $(2d+1)$-coefficient stencil mentioned above. Naturally, one would like to use the TL method recursively for the solution of the coarse level problem

$$Qe = J_c R(b - Ax_{in}). \tag{9}$$

Unfortunately, Q is not of a $(2d+1)$-coefficient stencil but rather an oblique 3^d-coefficient one and, therefore, the TL method is not a direct solver for (9). In [10] it is suggested to modify Q by 'throwing' onto its main diagonal the elements located outside an oblique $(2d+1)$-coefficient stencil, which results in a $(2d+1)$-coefficient stencil for the coarse-grid problem as well. In our experience, however, modifying the Galerkin scheme for the coarse level might result in a non-robust method. Therefore, we choose another approach here.

We describe first the three-level method. Denote $f_0 = f$, $A_0 = A$, $\tilde{A}_0 = \tilde{A}$, $P_1 = P$, $R_1 = R$, $Q_1 = Q$, $W_1 = W$ and $A_1 = F$. We consider the case in which (9) is solved approximately by one iteration of the TL method (with a zero initial guess). This is equivalent to replacing the lower right block submatrix Q in F in (7) by the corresponding preconditioner of the TL method applied to (9). In fact, if $\tilde{Q}$ is generated from Q the same way $\tilde{A}$ was generated from A, $\check{P}$ and $\check{R}$ are generated from it the same way P and R were generated from $\tilde{A}$ and Q_2 is the corresponding coarse-level coefficient matrix of the TL method applied to (9) (using $F_2 = blockdiag(W_2, Q_2)$), then the three-level method for the solution of (1) obtains the form

$$x_{out} = x_{in} + P_1 P_2 A_2^{-1} R_2 R_1 (b - Ax_{in}),$$

where $P_2 = blockdiag(I, \check{P})$, $R_2 = blockdiag(I, \check{R})$ and $A_2 = blockdiag(W_1, W_2, Q_2)$ are of order N. (In the following, we also use the notation $\tilde{A}_1 = \tilde{Q}$.)

Let L be a positive integer denoting the number of levels used minus 1. The above procedure may then be repeated until level L, with f_i denoting the set of variable indices f resulting from the coarsening procedure applied to the ith level $(0 \le i < L)$ and f_L the set of variable indices on the coarsest level and $\tilde{A}_{i-1}$, P_i, R_i, Q_i, W_i and $A_i = blockdiag(W_1, W_2, \ldots, W_i, Q_i)$ denoting the matrices constructed for the ith level $(0 < i \le L)$. (P_i, R_i and A_i are of order N.) The multi-level method obtains then the form

$$x_{out} = x_{in} + P_{L,1} A_L^{-1} R_{L,1} (b - Ax_{in}), \tag{10}$$

where, for $i > k$, $P_{k,i} = R_{k,i} = I$, for $i = k$ $P_{k,i} = P_i$ and $R_{k,i} = R_i$ and, for $i < k$, $P_{k,i} = P_i P_{i+1} \cdots P_k$ is the 'prolongation' from the kth level to the $i-1$st one and $R_{k,i} = R_k R_{k-1} \cdots R_i$ is the 'restriction' from the $i-1$st level to the kth one.

The multi-level method can also be supplemented with relaxations before and after each coarse-level visit. Avoiding these relaxations is attractive for parallel implementation, where

it reduces the time consumption of an iteration from $O(L)$ (the time required for marching across levels) to $O(\log L)$ (the time required for summing across levels).

Using induction and Lemma 2 in [24], one may show that if the multi-level method is applied to (8) rather than (1) then it is mathematically unchanged. The obvious choice $D = diag(A)$ in (8) may be useful for improving the upper bound for the condition number in Theorem 1 below (see [24]).

3 Analysis in the SPD Case

Here we give an upper bound for the condition number of the multi-level method in the SPD case.

Recall that f_i denotes the set of variable indices f resulting from applying the coarsening procedure to the ith level ($0 \le i < L$). For $0 \le k \le i \le L$, define $f_i^k = \cup_{j=k}^{i} f_k$. With this notation, the coarsening procedure applied at the ith level ($0 \le i < L$) results in the partitioning $f_L^i = f_L^{i+1} \cup f_i$. Also, we have $f_L^0 = \{1, 2, \ldots, N\}$ and

$$J_{f_{i-2}^0} P_i J_{f_{i-2}^0}^t = J_{f_{i-2}^0} P_i J_{f_{i-2}^0}^t = I, \quad 2 \le i \le L.$$

Note also that, for $0 \le k \le j \le i < L$,

$$J_{f_j}^t J_{f_j} P_{j+1}^{-1} = J_{f_j}^t J_{f_j} P_{i+1,k+1}^{-1}. \tag{11}$$

In the sequel, $(\cdot,\cdot)_2$ denotes the usual inner product in $l_2(f_L^0)$ and $\|\cdot\|_2$ denotes the corresponding vector and matrix norms. Let D be a diagonal SPD matrix of order N. Let

- $(\cdot,\cdot) = (D\cdot,\cdot)_2$ denote the inner product induced by D;
- $\|\cdot\| = \sqrt{(\cdot,\cdot)}$ denote the corresponding vector norm;
- and $\|\cdot\| = \|D^{1/2} \cdot D^{-1/2}\|_2$ denote the matrix norm induced by D.

Define also the 'multi-level' inner product

$$[\cdot,\cdot] = \sum_{n=0}^{L} \|P_{n,1}\|^2 (P_{n,1} J_{f_L^n}^t J_{f_L^n} \cdot, P_{n,1} J_{f_L^n}^t J_{f_L^n} \cdot).$$

For $0 \le n \le L$, define the matrices

$$G_n = J_{f_L^n}^t J_{f_L^n} A_n J_{f_L^n}^t J_{f_L^n} \quad \text{and} \quad \tilde{G}_n = J_{f_L^n}^t \tilde{A}_n J_{f_L^n}.$$

(This means that G_n and $\tilde{G}_n$ are of order N and that $(G_n)_{i,j} = (A_n)_{i,j}$ and $(\tilde{G}_n)_{i,j} = (\tilde{A}_n)_{i,j}$ if $i, j \in f_L^n$ $(G_n)_{i,j} = (\tilde{G}_n)_{i,j} = 0$ otherwise.) This implies

$$[\cdot, A\cdot] = \sum_{n=0}^{L} \|P_{n,1}\|^2 (\cdot, G_n \cdot).$$

For any matrix M which is SPD with respect to some inner product, define its condition number by

$$\kappa(M) = \rho(M)\rho(M^{-1}),$$

where ρ denotes the spectral radius of a matrix. (This definition arises from the Preconditioned Conjugate Gradients algorithm, which uses the inner product for which the preconditioned coefficient matrix is SPD. With the matrix norm induced by this inner product, the norm of the preconditioned coefficient matrix is equal to its spectral radius.)

Define the constants

$$\begin{aligned}
g &= \max_{0\le k<L} \|G_k\|, \\
p &= \max_{0\le k<L} \|J_{f_k}^t J_{f_k} P_{k+1} J_{f_k}^t J_{f_k}\|, \\
\bar{p} &= \max_{0\le n\le k<L} \|P_{k,n+1} J_{f_k}^t J_{f_k}\| \|P_{n,1}\| / \|P_{k,1}\|, \\
\hat{p} &= \max_{0\le n\le k<L} \|P_{k,n+1}^{-1}\|, \\
w &= \max_{0\le k<L} \|J_{f_k}^t W_{k+1} J_{f_k}\| \\
\text{and } a &= \max_{0\le k<L} \rho\left(J_{f_k} diag(A_k)^{-1} G_k J_{f_k}^t\right).
\end{aligned}$$

Theorem 1 *Assume that A is symmetric with respect to $(\cdot,\cdot)$ and that the matrices G_i $(0 \le i < L)$ are diagonally dominant. Then*

$$\kappa(P_{L,1}A_L^{-1}R_{L,1}A) \le (1 + 2gp\bar{p}L^2 + g^2p^2\bar{p}^2L^3 + wg\hat{p}^2L)2L\max\{a,1\}. \tag{12}$$

Proof: Since A is symmetric with respect to $(\cdot,\cdot)$, it follows by induction that the P_i^{-1}'s are the conjugates of the corresponding R_i^{-1}'s with respect to $(\cdot,\cdot)$ and that the A_i's are also symmetric with respect to $(\cdot,\cdot)$. Consequently, the $P_{i,j}^{-1}$'s are the conjugates of the corresponding $R_{i,j}^{-1}$'s with respect to $(\cdot,\cdot)$ and A and $P_{L,1}A_L^{-1}R_{L,1}$ are symmetric with respect to $[\cdot,\cdot]$ as well.

Since G_n, $\tilde{G}_n$ and $G_n - \tilde{G}_n$ are symmetric with respect to $(\cdot,\cdot)$ and diagonally dominant, it follows from Gershgorin's theorem that they are positive semidefinite. Consequently, for any nonzero vector $x \in l_2(f_L^0)$ and $0 \le n < L$,

$$\|J_{f_n}^t J_{f_n} P_{n+1}^{-1} x\|^2 = \|J_{f_n}^t J_{f_n} \tilde{G}_n x\|^2 \le \|\tilde{G}_n x\|^2 \le \|\tilde{G}_n\|(x, \tilde{G}_n x) \le \|G_n\|(x, G_n x).$$

Using this and (11), we have

$$\begin{aligned}
&\left(P_{n,1}J_{f_L^n}^t J_{f_L^n} x, R_{L,1}^{-1}A_L P_{L,1}^{-1} P_{n,1} J_{f_L^n}^t J_{f_L^n} x\right) \\
&= \left(P_{L,n+1}^{-1} J_{f_L^n}^t J_{f_L^n} x, A_L P_{L,n+1}^{-1} J_{f_L^n}^t J_{f_L^n} x\right) \\
&= \left(J_{f_L}^t J_{f_L} x, A_L J_{f_L}^t J_{f_L} x\right) + \sum_{k=n}^{L-1} \left(J_{f_k}^t J_{f_k} P_{k+1}^{-1} x, J_{f_k}^t W_{k+1} J_{f_k} P_{k+1}^{-1} x\right)
\end{aligned}$$

$$
\begin{aligned}
&\leq \left(J_{f_L}^t J_{f_L} x, A_L J_{f_L}^t J_{f_L} x\right) + \sum_{k=n}^{L-1} \|J_{f_k}^t W_{k+1} J_{f_k}\| \|G_k\|(x, G_k x) \\
&= \left(\left(P_{L,n+1}^{-1} - J_{f_{L-1}^n}^t J_{f_{L-1}^n} P_{L,n+1}^{-1}\right) J_{f_L^n}^t J_{f_L^n} x, R_{L,n+1} A_n P_{L,n+1} \left(P_{L,n+1}^{-1} - J_{f_{L-1}^n}^t J_{f_{L-1}^n} P_{L,n+1}^{-1}\right) J_{f_L^n}^t J_{f_L^n} x\right) \\
&+ \sum_{k=n}^{L-1} \|J_{f_k}^t W_{k+1} J_{f_k}\| \|G_k\|(x, G_k x) \\
&= \left(P_{L,n+1}\left(P_{L,n+1}^{-1} - \sum_{k=n}^{L-1} J_{f_k}^t J_{f_k} P_{k+1}^{-1}\right) J_{f_L^n}^t J_{f_L^n} x, A_n P_{L,n+1}\left(P_{L,n+1}^{-1} - \sum_{k=n}^{L-1} J_{f_k}^t J_{f_k} P_{k+1}^{-1}\right) J_{f_L^n}^t J_{f_L^n} x\right) \\
&+ \sum_{k=n}^{L-1} \|J_{f_k}^t W_{k+1} J_{f_k}\| \|G_k\|(x, G_k x) \\
&= \left(\left(I - \sum_{k=n}^{L-1} P_{k,n+1} P_{k+1} J_{f_k}^t J_{f_k} P_{k+1}^{-1}\right) J_{f_L^n}^t J_{f_L^n} x, A_n \left(I - \sum_{k=n}^{L-1} P_{k,n+1} P_{k+1} J_{f_k}^t J_{f_k} P_{k+1}^{-1}\right) J_{f_L^n}^t J_{f_L^n} x\right) \\
&+ \sum_{k=n}^{L-1} \|J_{f_k}^t W_{k+1} J_{f_k}\| \|G_k\|(x, G_k x) \\
&\leq (x, G_n x) + 2\sqrt{\|G_n\|(x, G_n x)} \sum_{k=n}^{L-1} \|P_{k,n+1} J_{f_k}^t J_{f_k}\| \|J_{f_k}^t J_{f_k} P_{k+1} J_{f_k}^t J_{f_k}\| \sqrt{\|G_k\|(x, G_k x)} \\
&+ \|G_n\| \left(\sum_{k=n}^{L-1} \|P_{k,n+1} J_{f_k}^t J_{f_k}\| \|J_{f_k}^t J_{f_k} P_{k+1} J_{f_k}^t J_{f_k}\| \sqrt{\|G_k\|(x, G_k x)}\right)^2 \\
&+ \sum_{k=n}^{L-1} \|J_{f_k}^t W_{k+1} J_{f_k}\| \|G_k\|(x, G_k x).
\end{aligned}
$$

By multiplying both sides of this inequality by $\|P_{n,1}\|^2$, one obtains

$$
\begin{aligned}
&\|P_{n,1}\|^2 \left(P_{n,1} J_{f_L^n}^t J_{f_L^n} x, R_{L,1}^{-1} A_L P_{L,1}^{-1} P_{n,1} J_{f_L^n}^t J_{f_L^n} x\right) \\
&\leq \|P_{n,1}\|^2 (x, G_n x) + 2\sqrt{\|P_{n,1}\|^2 \|G_n\|(x, G_n x)} \sum_{k=n}^{L-1} \|J_{f_k}^t J_{f_k} P_{k+1} J_{f_k}^t J_{f_k}\| \\
&\cdot \sqrt{\|P_{n,1}\|^2 \|P_{k,n+1} J_{f_k}^t J_{f_k}\|^2 \|G_k\|(x, G_k x)} \\
&+ \|G_n\| \left(\sum_{k=n}^{L-1} \|J_{f_k}^t J_{f_k} P_{k+1} J_{f_k}^t J_{f_k}\| \sqrt{\|P_{n,1}\|^2 \|P_{k,n+1} J_{f_k}^t J_{f_k}\|^2 \|G_k\|(x, G_k x)}\right)^2 \\
&+ \sum_{k=n}^{L-1} \|P_{n,1}\|^2 \|J_{f_k}^t W_{k+1} J_{f_k}\| \|G_k\|(x, G_k x) \\
&\leq \|P_{n,1}\|^2 (x, G_n x) + 2gp\bar{p}\sqrt{\|P_{n,1}\|^2 (x, G_n x)} \sum_{k=n}^{L-1} \sqrt{\|P_{k,1}\|^2 (x, G_k x)} \\
&+ g^2 p^2 \bar{p}^2 \left(\sum_{k=n}^{L-1} \sqrt{\|P_{k,1}\|^2 (x, G_k x)}\right)^2 \\
&+ wg \sum_{k=n}^{L-1} \|P_{k,1}\|^2 \|P_{k,n+1}^{-1}\|^2 (x, G_k x) \\
&\leq \|P_{n,1}\|^2 (x, G_n x) + 2gp\bar{p} L[x, Ax] + g^2 p^2 \bar{p}^2 L^2 [x, Ax] + wg\hat{p}^2 [x, Ax].
\end{aligned}
$$

For $n = L$, we have the trivial equation

$$\|P_{n,1}\|^2 \left(P_{n,1} J^t_{f^n_L} J_{f^n_L} x, R^{-1}_{L,1} A_L P^{-1}_{L,1} P_{n,1} J^t_{f^n_L} J_{f^n_L} x\right) = \|P_{n,1}\|^2 (x, G_n x).$$

By summing over $0 \le n \le L$, we obtain

$$[x, R^{-1}_{L,1} A_L P^{-1}_{L,1} x] \le (1 + 2gp\bar{p}L^2 + g^2 p^2 \bar{p}^2 L^3 + wg\hat{p}^2 L)[x, Ax].$$

On the other hand, we have by repeated applications of Lemma 1 in [23] [24] [25] that, for any nonzero vector $x \in l_2(f^0_L)$ and $0 \le n < L$ (using the notation $y = P^{-1}_{L,n+1} J^t_{f^n_L} J_{f^n_L} x$)

$$\begin{aligned}
&(x, G_n x) \\
&\quad = \quad (y, R_{L,1} A P_{L,1} y) \\
&\quad \le \quad 2\left(y, \left(\sum_{j=0}^{1} J^t_{\cup_{k=j \ mod \ 2} \ f_k} J_{\cup_{k=j \ mod \ 2} \ f_k} R_{L,1} A P_{L,1} J^t_{\cup_{k=j \ mod \ 2} \ f_k} J_{\cup_{k=j \ mod \ 2} \ f_k}\right) y\right) \\
&\quad \le \quad 4\left(y, \left(\sum_{j=0}^{3} J^t_{\cup_{k=j \ mod \ 4} \ f_k} J_{\cup_{k=j \ mod \ 4} \ f_k} R_{L,1} A P_{L,1} J^t_{\cup_{k=j \ mod \ 4} \ f_k} J_{\cup_{k=j \ mod \ 4} \ f_k}\right) y\right) \\
&\le \cdots \le \quad 2L\left(y, \left(\sum_{k=n}^{L} J^t_{f_k} J_{f_k} R_{L,1} A P_{L,1} J^t_{f_k} J_{f_k}\right) y\right) \\
&\quad = \quad 2L\left(y, \left(\sum_{k=n}^{L-1} J^t_{f_k} J_{f_k} R_{k+1} G_k P_{k+1} J^t_{f_k} J_{f_k} + G_L\right) y\right) \\
&\quad \le \quad 2L \max\{a, 1\} (y, A_L y) \\
&\quad = \quad 2L \max\{a, 1\} (P_{n,1} J^t_{f^n_L} J_{f^n_L} x, R^{-1}_{L,1} A_L P^{-1}_{L,1} P_{n,1} J^t_{f^n_L} J_{f^n_L} x).
\end{aligned}$$

For $n = L$, we have the trivial equation

$$(x, G_n x) = \left(P_{n,1} J^t_{f^n_L} J_{f^n_L} x, R^{-1}_{L,1} A_L P^{-1}_{L,1} P_{n,1} J^t_{f^n_L} J_{f^n_L} . x\right).$$

By multiplying the above inequalities and equation on both sides by $\|P_{n,1}\|^2$ and summing over $0 \le n \le L$, one obtains

$$[x, Ax] \le 2L \max\{a, 1\} [x, R^{-1}_{L,1} A_L P^{-1}_{L,1} x]. \quad \square$$

The constants used in (12) are not yet estimated. The following inequalities (derived under the assumptions in Theorem 1) are helpful in evaluating these constants for the problems considered here and in [23] [24].

Define

$$\bar{g} = \max_{0 \le k < L} \|diag(G_k)\| \quad \text{and} \quad \hat{g} = \max_{0 \le k < L} \|J^t_{f_k} J_{f_k} diag(A_k)^{-1} J^t_{f_k} J_{f_k}\|. \tag{13}$$

From Gershgorin's theorem and the diagonal dominance and symmetry (with respect to $(\cdot, \cdot)$) of the G_k's, it follows that

$$a \le 2 \quad \text{and} \quad g \le 2\bar{g}. \tag{14}$$

For estimating p, we use (2)–(3). which imply

$$p \leq \bar{\alpha}^{-1}\hat{g}.$$

This also implies

$$w \leq p^2\bar{g} \leq \bar{\alpha}^{-2}\hat{g}^2\bar{g}. \tag{15}$$

Let us estimate $\hat{p}$. Assume that the $\tilde{G}_k$'s $(0 \leq k < L)$ are sparse in the sense that

$$\max_{i \in f_L^{k+1}} |\{j \in f_k \mid (\tilde{G}_k)_{i,j} \neq 0\}| \leq N_0$$

and

$$\max_{j \in f_L^{k+1}} |\{i \in f_k \mid (\tilde{G}_k)_{i,j} \neq 0\}| \leq N_0,$$

where N_0 is a constant independent of N and L. (For the examples in Section 4 and in [11] [12] [23] [24], this holds with $N_0 = 9$; for the examples in [13], it holds with $N_0 = 27$.) From (11) and the diagonal dominance of the G_k's it follows that, for $0 \leq n \leq k < L$ and $1 \leq i \leq N$,

$$rs\left(|R_{k,n+1}^{-1}P_{k,n+1}^{-1}|\right)_{i,i} \leq \max\left\{N_0^2 L\bar{g}^2, 1\right\}.$$

From the symmetry (with respect to $(\cdot,\cdot)$) of the G_k's and Gershgorin's theorem, it follows that

$$\hat{p} \leq \max\left\{N_0 L^{1/2}\bar{g}, 1\right\}. \tag{16}$$

In Tables 1 and 2 below we show that, for the present numerical examples (with the multi-level method applied to (8) with $D = duag(A)$) $\bar{g} \leq 1$ and $\hat{g} \leq 2$. Furthermore, for such diffusion problems in 2-d we have, for $0 \leq n \leq k < L$, $\|P_{k,n+1}\| = O(4^{k-n})$, which implies that $\bar{p}$ is bounded independent of N and L. (Similarly, $\bar{p}$ can be bounded for more general diffusion problems in 3-dimensional domains.) In this case, one can also see from the proof of Theorem 1 that the upper bound in (12) can be improved by replacing $\hat{p}$ by

$$\max_{0 \leq n \leq k < L} \|P_{n,1}\|/\|P_{k,1}\| = O(1).$$

4 Numerical Experiments

Here we apply the multi-level method to diffusion problems with discontinuous coefficients. We test the 'staircase' problem (Example IV in [1], also tested in [23]). This is an isotropic diffusion equation of the form

$$-(\mathcal{D}u_x)_x - (\mathcal{D}u_y)_y = f \tag{17}$$

the diffusion coefficient of which, $\mathcal{D}$, is a scalar function which is large ($\mathcal{D} = \mathcal{D}_s$) inside the staircase and small ($\mathcal{D} = 1$) outside it. The domain is a square and the cell-size (on the finest grid of size 65×65) is equal to 1. Mixed (vacuum) boundary conditions of the form

$$\mathcal{D}u_n + 0.5u = f_b \tag{18}$$

(where $\vec{n}$ is the outer normal vector and f_b is some function defined on the boundary of the domain) are imposed on the upper and right edges of the square and Neumann boundary conditions are imposed on the other two edges. A finite volume discretization is used (see [1]); therefore, the coefficient matrix is of property-A. Since no anisotropy is presented, the multi-level method is implemented as in Section 2 with $\alpha = 0$. As discussed there, the first coarse level consists of variables corresponding to red grid points in a red-black coloring of the finest grid. The stencil of the coefficient matrix of the first coarse level is of size 3×3, aligned with the oblique grid of the red points. The further coarser grids are obtained by standard coarsening of this oblique grid, that is, taking every other point in the south-west–north-east direction and every other point in the south-east–north-west direction, so that the number of coarse grid points is about 1/4 of that of the next finer grid. The stencil on all the coarse grids is of the same structure as that of the first coarse grid.

It is observed that the coefficient matrices G_k, $0 \leq k \leq 4$, are diagonally dominant. Hence, Theorem 1 is valid with (8) and $D = diag(A)$, which yields a jump-independent upper bound in (12) (see [24]). In the following we display quantities which are important in estimating the constants used in (12) (see (13)–(16)). These quantities are calculated as if the multi-level method were applied to (8) with $D = diag(A)$ instead of (1). As discussed in Section 2, this implementation is mathematically equivalent to the actual one and, hence, may be used for deriving jump-independent upper bounds. In Table 1 $\mathcal{D}_s = 10^3$ is used and in Table 2 $\mathcal{D}_s = 10^4$ is used. It is seen from the tables that using $L > 4$ yields a large p in (12) and, therefore, is not recommended. Of course, in realistic applications this can be avoided by choosing a suitable $\bar{\alpha}$ in (2), e.g., $\bar{\alpha} = .2$; here, however, we choose for simplicity to use standard coarsening as described above, which amounts to $\bar{\alpha} = .2$ if $L \leq 4$ and $\bar{\alpha} = 0$ if $L > 4$ (this way, $\bar{\alpha}$ is smaller than the numbers in the last columns in the tables, hence (2) does not occur and, hence, (3) is not used.)

Table 1: Quantities used in the upper bound in Theorem 1. $\mathcal{D}_s = 10^3$ is used.

k	$\|diag(G_k)\|$	$\|J_{f_k}^t J_{f_k} diag(A_k)^{-1} J_{f_k}^t J_{f_k}\|$	$\|J_{f_k}^t \left(J_{f_k} \tilde{G}_k J_{f_k}^t\right)^{-1} J_{f_k}\|$	$\|diag(G_k) J_{f_k}^t \left(J_{f_k} \tilde{G}_k J_{f_k}^t\right)^{-1} J_{f_k}\|^{-1}$
0	1	1	1	1
1	0.9	1.6	4	0.286
2	0.879	1.882	5.333	0.222
3	0.872	1.969	5.818	0.206
4	0.871	1.990	1492.8	10^{-3}

Next, we report convergence rates for the multi-level method applied to the above examples. The basic multi-level method (10) is accelerated by the Conjugate Gradient Squared (CGS) method of [27]. (Although this acceleration method is inferior to the Preconditioned Conjugate Gradients method for SPD problems with SPD preconditioners, we are interested here in the investigation of the preconditioner and not in optimizing the acceleration technique.) The method is iterated until the l_2 norm of the residual is reduced by 9 orders of

Table 2: Quantities used in the upper bound in Theorem 1. $\mathcal{D}_s = 10^4$ is used.

k	$\|diag(G_k)\|$	$\|J_{f_k}^t J_{f_k} diag(A_k)^{-1} J_{f_k}^t J_{f_k}\|$	$\|J_{f_k}^t \left(J_{f_k}\tilde{G}_k J_{f_k}^t\right)^{-1} J_{f_k}\|$	$\|diag(G_k) J_{f_k}^t \left(J_{f_k}\tilde{G}_k J_{f_k}^t\right)^{-1} J_{f_k}\|^{-1}$
0	1	1	1	1
1	0.9	1.6	4	0.286
2	0.879	1.882	5.333	0.222
3	0.872	1.969	5.818	0.206
4	0.872	1.992	14888.2	10^{-4}

magnitude. We define the preconditioned convergence factor by

$$\left(\frac{\|\mathcal{P}^{-1}(Ax_{last} - b)\|_2}{\|\mathcal{P}^{-1}(Ax_0 - b)\|_2}\right)^{last^{-1}},$$

where $\mathcal{P}$ is the multi-level preconditioner and *last* is the total number of multi-level iterations of the form (10). (The l_2 norm of the preconditioned residual is available within CGS and, as discussed in [23], reflects the error better than that of the residual itself since the preconditioned system is better-conditioned than the original one.)

The results in Table 3 are as expected from Theorem 1 in the sense that the preconditioned convergence factor increases slowly with L so long as $L \leq 4$. The deterioration of the convergence rate when $L = 5$ is used is due to the large p obtained in this case, which implies a large upper bound in (12). As discussed above, this can be fixed by using $\bar{\alpha} = .2$ in (2); however, here standard coarsening is used, which amounts to $\bar{\alpha} = .2$ if $L \leq 4$ and $\bar{\alpha} = 0$ if $L > 4$. Although this is not an optimal implementation, it is sufficient for our purpose, namely, illustrating the usefulness of the theory in constructing the multi-level method.

Table 3: Preconditioned convergence factors for the multi-level method (accelerated by CGS) applied to the 'staircase problem.

L	$\mathcal{D}_s = 10^3$	$\mathcal{D}_s = 10^4$
1	10^{-6}	10^{-6}
2	.439	.392
3	.530	.533
4	.562	.570
5	.793	$> .9$

5 Discussion

The multi-level method introduced here is general in the sense that it is well-defined for general sparse linear systems. Of course, pathological examples may be constructed for which the coarsening procedure is trivial and, therefore, the number of levels that can be used is limited. For linear systems arising from the discretization of elliptic boundary value

problems, however, the method has the potential of being general in the sense that it is applicable independent of the regular finite element mesh used.

The theory is limited to SPD problems for which the coefficient matrices at all levels are diagonally dominant. Although this is rarely the case in real applications, it holds for model diffusion problems with discontinuous coefficients. For this kind of problems, the condition number grows at most polynomially with the number of levels used, as is also illustrated numerically. The property that the discontinuities need not be aligned with the coarse meshes enables the use of a large number of levels.

The multi-level method is constructed such that it is applicable also for anisotropic and nonsymmetric problems. Verifying this is left for future research.

References

[1] Alcouffe, R.; Brandt, A.; Dendy, J. E.; and Painter J.: The Multigrid Method for the Diffusion Equation with Strongly Discontinuous Coefficients. *SIAM J. Sci. Stat. Comput.,* vol. 2, 1981, pp. 430-454.

[2] R.E. Bank and T. Dupont: An Optimal Order Process for Solving Elliptic Finite Element Equations. *Math. Comp.* 36, 1981, pp 35–51.

[3] Bank, R. E.; Dupont, T. F.; and Yserentant, Y.: The Hierarchical Basis Multigrid Method. *Numer. Math.* 52 (1988), 427-458.

[4] Bank, R. E.; and Gutsch, S.: Hierarchical Basis for the Convection-Diffusion Equation on Unstructured Meshes. In Ninth International Conference on Domain Decomposition Methods, Bjorstad, P., Espedal, M. and Keyes, D. (eds.), Willey and sons, Sussex, UK, 1997.

[5] J.H. Bramble and R.E. Ewing and J.E. Pasciak and A.H. Schatz: A Preconditioning Technique for the Efficient Solution of Problems with Local Grid Refinement. *Comp. Meth. Appl. Mech. Engng.* 67, 1988, pp. 149–159.

[6] Bramble J. H.; Pasciak J. E.; and Schatz A. H.: The Constructuring of Preconditioners for Elliptic Problems on Regions Partitioned into Substructures *I. Math. Comp.* 46 (1986), 361-369.

[7] Bramble J. H.; Pasciak J. E.; and Schatz A. H.: The Constructuring of Preconditioners for Elliptic Problems on Regions Partitioned into Substructures *II. Math. Comp.* 47 (1986) 103-134.

[8] J.H. Bramble and J.E. Pasciak and J. Xu: Parallel Multilevel Preconditioners. *Math. Comp.* 55, 1990, pp. 1–22.

[9] Brandt, A.: Guide to Multigrid Development. Multigrid Methods. Hackbusch, W.; and Trottenberg, U., eds.: *Lecture Notes in Mathematics* 960, Springer-Verlag, Berlin, Heidelberg 1982.

[10] Brand, C. W.: An Incomplete Factorization Preconditioning Using Repeated Red Black Ordering. *Numer. Math.* 61 (1992), pp. 433-454.

[11] Dendy, J. E.: Black Box Multigrid. *J. Comp. Phys.* 48 (1982), pp. 366-386.

[12] Dendy, J. E.: Black Box Multigrid for Nonsymmetric Problems. *Appl. Math. Comp.,* vol. 13, 1983, pp. 261-283.

[13] Dendy, J.E.: Two Multigrid Methods for the Three-Dimensional Problems with Discontinuous and Anisotropic Coefficients. *SIAM J. Sci. Stat. Comput.,* 8 (1987), pp. 673-685.

[14] Dutto, L. C.; and Habashi, W. G.: An Algebraic Multilevel Parallelizable Preconditioner for Large-Scale CFD Problems. In the MGnet virtual proceedings of the 7th Copper Mountain Conference on Multigrid Methods, April 6-11, 1997.

[15] Henson, V.H.: Towards a Fully-Parallelizable Algebraic Multigrid. Presented in the 7th Copper Mountain Conference on Multigrid Methods, April 6-11, 1997.

[16] Kettler, R.; and Meijerink, J. A.: A Multigrid Method and a Combined Multigrid-Conjugate Gradient Method for Elliptic Problems with Strongly Discontinuous Coefficients in General Domains. Shell Publ. 604, KSELP, Rijswijk, The Netherlands, 1981.

[17] Maitre, J.F.; and Musy, F.: Multigrid Methods: Convergence Theory in a Variational Framework. *SIAM J. Numer. Anal.,* vol. 21, 1984, pp. 657-671.

[18] Maitre, J.F.; and Musy, F.: Algebraic Formulation of the Multigrid Method in the Symmetric and Positive Definite Case–A Convergence Estimation for the V-Cycle. In Multigrid Methods for Integral and Differential Equations, D.J. Paddon and H. Holstein (eds.), Clarendon Press, Oxford, 1985.

[19] Mavriplis, D. J.: Directional Coarsening and Smoothing for Anisotropic Navier-Stokes Problems. In the MGnet virtual proceedings of the 7th Copper Mountain Conference on Multigrid Methods, April 6-11, 1997.

[20] Reusken, A.: A Multigrid Method Based on Incomplete Gaussian Elimination. *Numer. Linear Algebra Appl.* 3 (1996), pp. 369–390.

[21] Ruge, J. W.; and Stuben, K.: Algebraic Multigrid. In McCormick, S.F., ed.: *Multigrid Methods*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1987.

[22] Shapira, Y.: Multigrid Methods for 3-D Definite and Indefinite Problems. *Appl. Numer. Math.* (accepted).

[23] Shapira, Y.: Analysis of Matrix-Dependent Mulrigrid Algorithms. submitted to *Numer. Linear Algebra Appl.*.

[24] Shapira, Y.: Multigrid for Locally Refined Meshes. In the MGnet virtual proceedings of the 7th Copper Mountain Conference on Multigrid Methods, April 6-11, 1997. Submitted to *SIAM J. Sci. Comput.*.

[25] Shapira, Y.: Algebraic Domain Decomposition Method for Unstructured Grids. In Ninth International Conference on Domain Decomposition Methods, Bjorstad, P., Espedal, M. and Keyes, D. (eds.), Willey and sons, Sussex, UK, 1997 (accepted).

[26] Shapira, Y.; Israeli, M.; and Sidi, A.: Towards Automatic Multigrid Algorithms for SPD, Nonsymmetric and Indefinite Problems. *SIAM J. Sci. Comput.* 17 (1996), 439–453.

[27] Sonneveld, P.: CGS, a Fast Lanczos-Type Solver for Nonsymmetric Linear Systems. *SIAM J. Sci. Statist. Comput.* 10 (1989), pp. 36-52.

[28] de Zeeuw, P. M.: Matrix-dependent Prolongations and Restrictions in a Blackbox Multigrid Solver. *J. Comput. Appl. Math.* 33 (1990), pp. 1-27.