JOURNAL OF COMPUTATIONAL PHYSICS 19, 418–431 (1975)

On Multiple Grid and Related Techniques for Solving Discrete Elliptic Systems

R. A. NICOLAIDES

Department of Mathematics, Imperial College, London S. W. 7, United Kingdom

Received July 31, 1974; revised May 16, 1975

A descriptive account of a type of iterative method for solving elliptic equations is given. This type of method involves the use of more than one finite difference or element grid. A new rapidly convergent algorithm of this type is proposed and numerical results demonstrating its usefulness are presented.

1. INTRODUCTION

Some recent proposals for solving the finite difference equations arising from second-order elliptic differential equations have involved the use of more than one finite difference grid or finite difference operator. For brevity we refer to such methods as being of multiple grid type. The general principle underlying this type of method was understood by pencil and paper relaxation users, and the method used by them and based on this principle was called "block relaxation." The proposed machine implementations, however, owe little or nothing to this earlier work.

The class of algorithms under discussion is of importance for the following reasons: For second-order elliptic equations in the plane it contains members with the property that in order to reduce a norm of an initial error by a factor of 10^{-p}, requires an amount of work proportional to the number of gridpoints n and p. By this we mean that there is a constant K such that the amount of work is bounded by Kpn machine operations. It seems clear that no algorithm can improve on this figure, except for providing a lesser value for K. Many general algorithms involve corresponding work in Kpn^{α} where $\alpha > 1$. For SOR with optimal ω, $\alpha = 3/2$, moreover, even the "fast direct methods" with their limited usefulness, have a work count that is greater than $O(n)$ operations. All that this means is that multiple grid methods deserve investigation. We shall say more about where the above results can be found later. It must be said here that the implementation of a multiple grid method involves a high strategic component. One reason for this is the need to keep the machine overheads down to a reasonable level. Moreover,

although it may not be considered important from the point of view of the "optimal" use of the computing machinery, programming a multiple grid method is a rather complex operation. This situation can be compared with SOR, for which the programming is often rather simple. There appears to be a real need for more suggestions as to how a multiple grid method might be implemented efficiently, with supporting numerical evidence.

The present paper is descriptive in character. We hope to provide some rigorous analysis later. In Section 2 there is an explanation of the underlying principle of the methods and a brief account of what has been done so far. In Section 3 we show a general principle for constructing a kind of "coarse grid" finite difference operator. In Section 4 another multiple grid type of algorithm of general utility is proposed and its workings are explained, and in Section 5 numerical results are given to show the implementability of the algorithm (which is rapidly, but not optimally convergent).

The remainder of this section is devoted to a very brief note on elimination methods for solving elliptic difference equations. Although not strictly relevant to what follows, we shall discuss these here. There has been a tendency to assume that the demise of iterative methods is imminent (e.g., see [11, p. 33]). A look at the work count of the elimination process should dispel these notions. Well-known work [5, 8] shows that at least for second-order equations it seems unlikely that the work count can be obtained below $O(n^{3/2})$ operations. (This figure is for a five-point difference formula or the simplest finite element system.) If we agree that $O(n)$ operations is the optimal figure, then the implication is that simple elimination can never be optimal. This is a serious matter, but it is only to be expected since the elimination process takes no real account of the ellipticity, while the multiple grid methods, for example, make essential use of it. The situation with higher-order equations is less clear, though even here, without the use of iterative improvement, elimination processes might be prone to stability problems, while rapidly convergent iterative methods are relatively free of this difficulty because their stability properties are determined essentially by the condition of the iteration operator and not that of the coefficient matrix. Multidimensional problems can be expected to cause further difficulties. Probably the best kind of algorithm would use both iteration and elimination.

2. General Principles

To illustrate the essence of the multiple grid idea, we consider the following model problem: Let Ω denote the square

$$\Omega = \{P \equiv (x, y) \mid 0 < x, y < \pi\},$$

and $\partial\Omega$ its boundary. Find u such that

$$\begin{aligned} -\Delta u = -\left(\frac{\partial^2 u}{\partial u^2} + \frac{\partial^2 u}{\partial y^2}\right) = f(P) \qquad & P \in \Omega \\ u(Q) = \varphi(Q) \qquad & Q \in \partial\Omega, \end{aligned} \tag{2.1}$$

φ being given.

To solve this problem, we proceed in the usual way, namely, by writing down the discrete problem

$$\begin{aligned} -\Delta_h u_{ij}^h = f_{ij}^h \qquad & P_{ij} \in \Omega_h \\ u_{ij}^h = \varphi(Q_{ij}) \qquad & Q_{ij} \in \partial\Omega_h \end{aligned} \tag{2.2}$$

where u^h is a grid function defined at the points P_{ij} of the grid

$$\bar{\Omega}_h = \{(x_i, y_j) = P_{ij} \mid x_i = ih, y_j = jh, i, j = 0, 1, \ldots, (n+1)\},$$

and where Δ_h is the difference operator such that

$$\Delta_h u_{ij} = h^{-2}(u_{i+1,j} - 2u_{ij} + u_{i-1,j}) + h^{-2}(u_{i,j+1} - 2u_{ij} + u_{i,j-1}).$$

Choosing an arbitrary, fixed listing of the interior gridpoints $P_{ij} \in \Omega_h$, and applying (2.2) to each point in order gives the system of linear equations

$$A_h u^h = b^h \tag{2.3}$$

where A_h is a positive definite matrix with diagonal elements $4h^{-2}$ and b^h is defined in the obvious way.

The eigenvectors $w^{h;k,l}$ of A_h are known to be (up to a common constant multiplier)

$$w_{ij}^{h;k,l} = w^{h;k,l}(P_{ij}) = \sin kih \sin ljh; \qquad i, j = 1, 2, \ldots, n; \qquad k, l = 1, 2, \ldots, n \tag{2.4}$$

with corresponding eigenvalues

$$\lambda_{k,l}^h = \frac{4}{h^2}\sin^2\frac{kh}{2} + \frac{4}{h^2}\sin^2\frac{lh}{2}. \tag{2.5}$$

With this notation the general principle may be explained thus: Let $h_1 = \pi/(n_1 + 1)$ and $h_2 = \pi/(n_2 + 1)$ $n_1 > n_2$. The terms, h_1, defines the "fine" grid and h_2 defines the "coarse" grid; h_2 is assumed to be chosen in such a way that the work involved in reducing the error in a given initial approximation to the solution of

$$A_{h_2} u^{h_2} = b^{h_2} \tag{2.6}$$

by a given factor 10^{-p} is small compared to the work involved for a similar operation on (2.3). Let $\bar{u}^{h_2}$ be the approximate solution to (2.6) thereby obtained. Now suppose that $\bar{u}^{h_2}$ is extended by some (local) interpolation process onto $\bar{\Omega}_{h_1}$ and let $\tilde{u}^{h_2}$ denote the interpolated grid function, defined on $\bar{\Omega}_{h_1}$.

The important fact about $\tilde{u}^{h_2}$ is that provided h_1 and h_2 are small and sufficiently close to each other (how close will depend, among other things, on how rapidly b^{h_1} oscillates), the difference $u^{h_1} - \tilde{u}^{h_2}$ will have only small projections along the eigenvectors, $w^{h_1;k,l}$, $k, l = 1, 2,..., n_2$, of A_{h_1}. (The eigenvectors $w^{h_1;k,l}$ are assumed normalized in some way independent of i, $i = 1, 2$.) Loosely speaking, this comes about because $w^{h_1;k,l}$, $w^{h_2;k,l}$, $k, l = 1, 2,..., n_2$ for h_1 near to h_2 have the common "representation" (2.4). Evidently, certain facts about the interpolation process are implicit in this argument. Hence, there are grounds for supposing that most of the difference $\tilde{u}^{h_2} - u^{h_1}$ lies in the subspace of $R_{n_1^2}$ spanned by the more oscillatory eigenvectors, $w^{h_1;k,l}$, $k, l = n_2 + 1,..., n_1$. Oscillating error vectors are easy to reduce by conventional relaxation techniques, as we indicate later. The essence of a multiple grid method is the elimination of long wavelength error components by the use of coarse grid operations, and the elimination of short wavelength errors by other (relaxation) techniques. For various reasons the above suggestions must not be taken too literally and indeed must be modified to construct a working algorithm, but the principle is a valid one.

For the purposes of illustration, we now briefly consider a specific proposal for a multiple grid algorithm made by Fedorenko [3, 4]. This work seems to be the first systematic multiple grid method proposed and analyzed. The basic algorithm has three steps and will be described for the model problem of Section 2 with $\varphi = 0$. A few notations are needed. Let $\bar{\Omega}_{h'}$ be a grid, and with $h' \geqslant h$, suppose that $\bar{\Omega}_{h'} \subseteq \bar{\Omega}_h$ so that $\bar{\Omega}_{h'}$ is a subgrid of $\bar{\Omega}_h$. If v^h is defined on $\bar{\Omega}_h$ then by $v^{h'}$ we mean the restriction of v^h to $\bar{\Omega}_{h'}$. Below we choose

$$h = \pi/2^s \qquad \text{and} \qquad h' = \pi/2^{s-1}, \qquad \bar{\Omega}_{h'} \subset \bar{\Omega}_h .$$

For any h, we use the norm for grid functions

$$\| v^h \|_h^2 = h^2 \sum_{i,j \in \Omega_h} (v_{ij}^h)^2$$

and lastly, define the residuals

$$r^{h,k} = A_h u^{h,k} - b^h$$

for grid functions $u^{h,k}$, so that if $\epsilon^{h,k} = u^{h,k} - u^h$, then

$$A_h \epsilon^{h,k} = r^{h,k}.$$

To carry out the algorithm we need an initial approximation $u^{h,0}$, integers p and m_0, and a parameter $\alpha > 0$. Then

(1) Do m_0 iterations of the relaxation algorithm

$$u^{h,k} = u^{h,k-1} - \alpha(A_h u^{h,k-1} - b^h) \qquad k = 1, 2, \ldots, m_0 .$$

(2) Compute $v^{h'}$ so that

$$\| A_{h'} v^{h'} - r^{h',m_0} \|_{h'} \leqslant 10^{-p} \| r^{h',m_0} \|_{h'} .$$

(3) Interpolate $v^{h'}$ (defined on $\Omega_{h'}$) onto Ω_h to get $\bar{v}^{h'}$ and then set $\bar{u}^{h,k} = u^{h,k} - \bar{v}^{h'}$.

For completeness, we give the details of step 3. They are as follows: to interpolate a given grid function $v^{h'}$ from $\bar{\Omega}_{h'}$ onto $\bar{\Omega}_h$ where $h = h'/2$, if $\bar{v}^{h'}$ is the interpolate, then

$$\begin{aligned}
\bar{v}^{h'}_{ij} &= v^{h'}_{ij} \qquad i, j \text{ even} \\
&= \tfrac{1}{2}(v^{h'}_{i-1,j} + v^{h'}_{i+1,j}) \qquad i \text{ odd}, j \text{ even} \\
&= \tfrac{1}{2}(v^{h'}_{i,j-1} + v^{h'}_{i,j+1}) \qquad i \text{ even}, j \text{ odd} \\
&= \tfrac{1}{4}(v^{h'}_{i+1,j+1} + v^{h'}_{i-1,j-1} + v^{h'}_{i+1,j-1} + v^{h'}_{i-1,j+1}) \qquad i \text{ odd}, j \text{ odd}.
\end{aligned}$$

Returning now to the algorithm itself we may add the following remarks. Regarding step 1, we have that

$$\epsilon^{h,m_0} = \sum_{i,j=1}^{2^s-1} (1 - \alpha\lambda^h_{ij})^{m_0} e^{h,0}_{i,j} w^{h;i,j}$$

where $e^{h,0}_{i,j}$ are the Fourier components of the initial error. For all those indices i, j for which $|(1 - \alpha\lambda^h_{ij})|$ is sufficiently small, the corresponding error components will be substantially reduced by step 1. For example, if $\alpha = \frac{1}{5}$, a simple computation with (2.5) shows that for $i, j \geqslant 2^{s-1} + 1$, $|(1 - \alpha\lambda^h_{ij})| < 0.7$. Notice that if $\alpha = \frac{1}{4}$, so that the iterations in step 1 are Jacobi iterations, the eigencomponents associated with the center of the spectrum are the ones first eliminated. The suboptimal choice of $\alpha = \frac{1}{5}$ ensures that the eigencomponents associated with most rapidly varying eigenfunctions are reduced fastest. Notice also that no choice of α leading to a convergent relaxation algorithm will give "rapid" reduction of the components of the slower varying eigenfunctions. This phenomenon seems to be a characteristic of all "one" parameter relaxation methods, including SOR. If it were possible to discover an algorithm free of this difficulty (multiple grid algorithms are in this category), then obviously, rapidly convergent general methods could be constructed.

Step 2 of the algorithm may be described as the computation of an approximate solution to the equation

$$A_{h'}\epsilon^{h',m_0} = r^{h',m_0}.$$

Thus, if h' was equal to h, and p was $-\infty$, then steps 2 and 3 would give the exact solution to the original problem. To summarize, the purpose of step 1 is to reduce the oscillating part of the initial error; that of step 2 is to find its less rapidly oscillating part to sufficient accuracy; while in step 3 we find a new approximate solution by subtracting from the old, u^{h,m_0}, its (slowly oscillating) error. Certain of the steps must be repeated (on an inductive basis) to get the proper definition of the algorithm, but we shall not go into this any further here.

Fedorenko proved that a certain variant of the above algorithm applied to the model problem requires a number of machine operations proportional only to the number of gridpoints to reduce the initial error by a given factor, provided that h is sufficiently small. This result was generalized by Bakhvalov [1] to a wide class of second-order equations, but still with Ω a square. In [2], some other multiple grid algorithms are proposed. These latter algorithms are different in principle from the algorithm analyzed by Bakhvalov, and possibly more natural. The difference is basically this: In Federenko's algorithm it is pressupposed that a discrete problem is to be solved and the method adopted is to incorporate successively a sequence of grids, proceeding from the finest grid to the coarsest one. Brandt's approach starts from a very coarse grid and then works up to finer and yet finer grids. This approach has the great advantage of not requiring the user to specify the grid in advance. Indeed, it is possible to construct an "adaptive" algorithm where the grids are constructed precisely to match the behavior of the solution of the differential equation. The latter information is available from the calculations done on coarser grids. This idea is far removed from the algebraic approach traditionally used in the solution of discrete elliptic problems. It is worth noting that both of the approaches we are considering make essential use of the residual equation (3.2).

3. Calculation of Residual Corrections

We consider now a general principle for constructing algorithms making use of the multiple grid idea. We shall suppose that we have an approximation $u^{h,0}$, obtained by any method whatever, to the solution u^h of the system of n (slight notational change) equations:

$$A_h u^h = b^h. \tag{3.1}$$

Define the error $\epsilon^{h,0} = u^{h,0} - u^h$ and the residual $r^{h,0}$ by

$$A_h \epsilon^{h,0} = r^{h,0}. \tag{3.2}$$

Evidently, if we could solve (3.2), then (3.1) would be solved too, for $u^h = u^{h,0} - \epsilon^{h,0}$. To get a simpler problem than (3.2) one may look for a solution in a subspace, which will be taken to be a p-dimensional subspace of R_n spanned by the columns of the $n \times p$ matrix of rank p, E. Thus, the approximate solution to (3.2) will have the form Ec where c is a p-dimensional vector. Thus,

$$A_h E c = r^{h,0}, \tag{3.3}$$

and we shall determine c by the familiar technique of requiring the residual

$$A_h E c - r^{h,0}$$

to be orthogonal to another subspace, also of dimension p, and spanned by the columns of a rank p, $n \times p$ matrix W. This leads to the p linear equations in p unknowns for c,

$$W' A_h E c = W' r^{h,0}. \tag{3.4}$$

Clearly $W' A_h E$ is not singular and c is determined uniquely; therefore,

$$Ec = E(W' A_h E)^{-1} W' r^{h,0}$$

and for sensible choices of E and W there is reason to expect that

$$\tilde{u}^{h,0} = u^{h,0} - E(W' A_h E)^{-1} W' r^{h,0} \tag{3.5}$$

is a "better" approximation to u^h than $u^{h,0}$. From (3.5) it follows that

$$\tilde{r}^{h,0} = (I - A_h E(W' A_h E)^{-1} W')\, r^{h,0}$$

or

$$\tilde{\epsilon}^{h,0} = (I - E(W' A_h E)^{-1} W' A_h)\, \epsilon^{h,0}.$$

A simple choice of W is just $W = E$. In this case c is just the Galerkin solution to (3.2) and therefore Ec solves the problem

$$\min_{y \in \text{span}(E)} \| \epsilon^{h,0} - y \|_{A_h}$$

where span(E) means the subspace spanned by the columns of E and $\| x \|^2_{A_h} \equiv (A_h x, x)$. Henceforth, only the case $W \equiv E$ will be considered. We shall choose E so that $E' A_h E$ is a "coarse grid analogue" of A_h, and so that c will be a coarse grid solution of the discrete Poisson problem (3.2). Ec is a grid function defined on the fine grid and in this sense E may be thought of as an operator interpolating c onto Ω_h.

We shall consider a particular choice of E, made in the following way: Let the region Ω be divided up into p nonoverlapping subregions $\Omega^{(i)}$ $i = 1, 2, \ldots, p$ with

boundaries $\partial\Omega^{(i)}$ $i = 1, 2, \ldots, p$ in such a way that every point P of $\bar{\Omega}$ satisfies $P \in \bigcup_{i=1}^{p} \bar{\Omega}^{(i)}$. Assume that $\Omega^{(i)}$ are chosen in such a way that $P_i \in \Omega_h$ $i = 1, 2, \ldots, n \Rightarrow P_i \in \Omega^{(k)}$ for some $1 \leqslant k \leqslant p$, and let $| \Omega^{(k)} |$ $k = 1, 2, \ldots, p$ be the number of gridpoints P_j in $\Omega_h \cap \Omega^{(k)}$.

Let $E_{\cdot i}$ be the n-dimensional vector whose kth component is $| \Omega^{(i)} |^{-1/2}$ if $P_k \in \Omega^{(i)}$ and zero otherwise. Having formed the vectors $E_{\cdot i}$, $i = 1, 2, \ldots, p$ construct from them the matrix E with $E_{\cdot i}$ for its ith column; the hypotheses ensure that rank $E = p$, provided the intersections $\Omega^{(i)} \cap \Omega_h$, $i = 1, 2, \ldots, p$ are all nonempty. Furthermore, in this case (which we shall assume) $E'E$ is a diagonal nonsingular matrix and $E(E'E)^{-1}E'$ is the orthogonal projector onto span(E). In some cases it is desirable to select other values for the nonzero elements of E. There are many possible ideas one could use. In the "block relaxation" of Southwell [9], the numbers are as we have given them (in essence), but the regions $\Omega^{(i)}$ were chosen according to the progress of the ordinary relaxation process, which was "interrupted" for the purpose of carrying out the block relaxation.

A method in which the present method, in a sense, can be viewed as a special case, was proposed by Wachspress in [15]. This work seems to contain the first systematic proposals for the interruption of iterations in order to carry out variationally based accelerations. Furthermore, the formulation in [15] is applicable even when nonpositive definite operators are involved. A recent review paper of Wachspress contains an account of further progress, chiefly, though not entirely, in the area of nuclear reactor design.

In the general literature, there are various sporadic references to the type of idea we are considering; see, e.g., [10, 12, 13].

4. An Algorithm

For the purpose of testing out the various properties of the previously mentioned methods, the following algorithm was devised and implemented. We are not necessarily suggesting that it should be used to solve real problems (although it could be). Our main purpose was to see a working algorithm of the type we are considering. First of all, we shall state the algorithm and explain how it works. Some numerical results will then be given for the model problem. The algorithm is superficially similar to one proposed in [13], but the resemblance is not too much more than superficial.

In the algorithm, each iterate depends on two previous iterates. Suppose $u^0, u^1, \ldots, u^k$ are known. The general structure is as follows. From u^k we do in turn a relaxation cycle and then a coarse grid computation. Denote the result of these two by $\bar{u}^k$. Then u^{k-1} and $\bar{u}^k$ are combined to give u^{k+1} by means of a second-degree method. The latter requires us to have available two parameters ω and q. The

relaxation sweep requires another parameter α and the coarse grid calculation requires a matrix E. The choice of these data is discussed below and as will be seen, relative to, say, a simple second-degree algorithm, there is no problem associated with the choice of the parameters ω and q. Here is a formal statement of the algorithm for solving $Au = b$, A positive definite.

ALGORITHM. Given E, α, u^0, u^1, ω, and q, construct the sequence $\{u^k\}$ according to

(a) $\tilde{u}^k = u^k - \alpha(Au^k - b)$

(b) $\bar{u}^k = \tilde{u}^k - E(E'AE)^{-1}E'(A\tilde{u}^k - b)$

(c) $u^{k+1} = (1 - \omega)u^{k-1} + \omega[q\bar{u}^k + (1 - q)u^k]$ for $k = 1, 2, \ldots$.

Part (a) is the ordinary relaxation sweep; (b) is the coarse grid calculation and would, as usual in iteration methods, be done by computing a "correction" c from $E'AEc = r^k$. Part (c) is the second-degree calculation. In the usual notation, (a) and (b) may be combined into a stationary linear method with operator

$$(I - E(E'AE)^{-1}E'A)(I - \alpha A). \tag{4.1}$$

The role of the second-degree part is to accelerate the convergence of the iterates produced by this operator. The subregions $\Omega^{(k)}$ in the calculations done for the model problem were chosen in the following way: h was chosen to have the form $h = \pi/(N^2 + 1)$ for integers $N = 5(1)10$, giving problems with 5^4 to 10^4 unknowns. Region Ω was divided up, in the natural way, into N^2 rectangular subregions each containing N^2 gridpoints. The coarse grid operator $E'AE$ is then proportional to the discrete Laplacian for a regular $N \times N$ grid, as is easy to verify.

We need to recall several facts about the second-degree algorithms. Consider the algorithm with φ given, $\varphi: D \subset R_n \to R_n$, $x^{k+1} = \varphi(x^k)$ where $x^* = \varphi(x^*)$. Let φ' denote the Frechét derivative of φ evaluated at x^*, and let $\mu_n' \leqslant \mu_{n-1}' \leqslant \cdots \leqslant \mu_1' < 1$ be the eigenvalues, assumed real, of the operator φ'. Let $\mu_i = 1 - \mu_i'$, $i = 1, 2, \ldots, n$. The parameters q and ω in the second-degree method interact in the following way: Associated with each q is a certain number $\bar{\mu}(q)$; for each q there exists a unique, best ω, determined by the formula familiar from the SOR theory,

$$\omega(q) = \frac{2}{1 + (1 - \bar{\mu}(q)^2)^{1/2}}$$

and the effective spectral radius for this ω is

$$(\omega(q) - 1)^{1/2} = \frac{1 - (1 - \bar{\mu}(q)^2)^{1/2}}{1 + (1 - \bar{\mu}(q)^2)^{1/2}}.$$

It follows that the most rapid convergence occurs when $\bar{\mu}(q)^2$ is as small as possible Now it is easy to see that the best value of q and the corresponding value of $\bar{\mu}(q)$ in the above sense is

$$\hat{q} = \frac{2}{\mu_1 + \mu_n} \qquad \bar{\mu}(\hat{q}) = \frac{\mu_n - \mu_1}{\mu_n + \mu_1}.$$

This situation can be expressed by saying that the effect of using $\hat{q}$ is to translate and compress the spectrum $\{\mu_1, ..., \mu_n\}$ so that it is symmetric about the origin, the optimum value of ω then being determined by the extreme points of this translated spectrum. In particular, if the spectrum $\{\mu_1, ..., \mu_n\}$ is initially not symmetric about the origin, then improvements in the rate of convergence are obtained by using a suitable value for q.

In the particular algorithm defined above, the role of the matrix φ' is played by (4.1). We choose $\alpha = h^2/4$ and carry out a similarity transformation on (4.1) with the matrix S, where $S^2 = A$ and S is positive definite. Eq. (4.1) then takes the form

$$(I - SE(E'AE)^{-1} E'S)(I - (h^2/4)A). \tag{4.2}$$

The second member is the orthogonal projector onto span(SE) and is therefore equal to its own square. Consequently, the eigenvalues of (4.2) are those of the symmetric matrix

$$(I - SE(E'AE)^{-1} E'S)(I - (h^2/4)A)(I - SE(E'AE)^{-1} E'S). \tag{4.3}$$

By forming the Rayleigh quotient and using the well-known theorem on constrained Rayleigh quotients, we infer that the spectrum of (4.3) lies inside that of $(I - (h^2/4)A)$. It follows from our previous discussion that the optimal second-degree method for (4.2) converges no slower than that for $(I - (h^2/4)A)$. It may be expected to converge considerably faster as we now explain (somewhat heuristically): Let w_1 be the eigenfunction of A associated with the smallest eigenvalue λ_1. Then the matrix

$$(I - w_1(w_1'w_1)^{-1} w_1')(I - (h^2/4)A) \tag{4.4}$$

has the same eigenvectors as A and with the eigenvalue λ_1 "replaced" by zero; the other eigenvalues are those of $(I - (h^2/4)A)$. Note that $(I - w_1(w_1'w_1)^{-1} w_1')$ is an orthogonal projector. A similar principle is evident in (4.2). The difference is that SE is not now an invariant subspace of A. But it "nearly" is because of the choice of E and because of the fact that S and A have the same eigenvectors. In a sense, the subspace span(E) is more closely related to the invariant subspace of A, span($w_1, w_2, ..., w_p$), than to any other of the invariant subspaces. In this sense we can expect that the primary effect of the coarse grid operations is to deflate out

of the operator $(I - (h^2/4)A)$ some eigenvalues from the upper part (closest to $+1$) of its spectrum. The finer the coarse grid the more of the spectrum that will be deflated. Parameter q in the second-degree algorithm is chosen to centralize the (deflated) spectrum of the composite operator in (4.2), while use of the corresponding optimum ω enables an order of magnitude increase in the rate of convergence to be obtained over the algorithm with $\omega = 1$.

Obviously, there are serious shortcomings to the previous argument, but it is felt that the essentials of the working of the algorithm are contained in it.

Fuller details of the second-degree and related methods may be found in [6, 14, 17]. The formulation given is taken from [7].

5. Numerical Results

As already indicated, we used the algorithm of Section 4 to solve the model problem, with boundary function $\varphi \equiv 1$, and the right-hand side $f \equiv 0$. The range of values of h was $h = \pi/(N^2 + 1)$ for integers $N = 5(1)10$. The subregions were rectangular, N^2 of them each containing N^2 points. The idea of this subdivision and various other ideas were borrowed from the interesting paper [13]. We set out to discover the rate of convergence of the algorithm and the sensitivity of this rate relative to the choice of parameters. The algorithm chosen should be rather insensitive to the choice of parameters for the kind of h one might encounter in practice. In every case, the initial approximation was a random vector with entries uniformly distributed in [0, 2]. The second "initial" approximation was generated by use of the iteration with operator (4.1). This choice results in a significant saving of machine time because, as is shown in [6], the second-degree algorithm then has the property of actually reducing the length of the error vector at every iteration. In general, the length of the error vector will increase before it starts decreasing. Parameter α was chosen as $h^2/4$ (of course the calculations were not literally arranged in this way). Concerning the solving of the coarse grid equations, $E'AEc = r$, we did either N or $2N$ SOR iterations, starting from zero and using the theoretically and a priori computable optimum value of ω. The choice makes little difference to the overall machine time since the amount of work for the coarse grid calculation in either case is negligible compared to the time for one relaxation sweep over the fine grid. The greater part of the time for the coarse grid calculation, in the present implementation, was spent on the computation of the residual for the right-hand side. The entire coarse grid calculation is equivalent in time to one second-degree sweep (roughly). Although $2N$ iterations were used for the figures given below, the results are not changed in any essential way by doing just N iterations. These tactics might have to change if a finer coarse grid was used. All calculations were done in double precision on the California Institute

of Technology IBM 370/158. The use of double-precision arithmetic was found to be essential to the coarse grid calculations. A calculation was terminated when, for some k, $\| u^{k+1} - u^{k-1} \|_\infty \leqslant 10^{-6}$. This criterion is useful only because the exact solution is $u \equiv 1$.

The general nature of the calculations performed for each N was that for a range of values of q, the problem was "solved" in the sense of the termination criterion for a range of values of ω near to the optimum for that q. From our earlier discussion it follows that $\hat{q} \to 1$ as $h \to 0$, and from the second-degree theory of choosing q, it follows that there is little damage done from making rather drastic underestimates of $\hat{q}$. This has little to do with the multiple grid part of the algorithm, but it does concern the fact that the graph of $\bar{\mu}(q)$ as a function of q is linear and has a slope converging to zero as $h \to 0$ for $q \in (0, \hat{q})$.

In Table I we give some results for the case $N = 8$, $h = \pi/65$ corresponding to 4096 unknowns. These results are typical of all the calculations. The entries are the number of iterations used. The behavior in the last line is explained by the observation that $\hat{q} \in (0.98, 0.99)$. Note also the insensitivity of the number of iterations with respect to ω. The reason for this is the smallness of the number $\bar{\mu}(q)$.

TABLE I

($N = 8, h = \pi/65$)

q/ω	1.55	1.56	1.57	1.58	1.59	1.60	1.61
0.96	64	60	57	56	58	60	62
0.97	63	59	56	56	59	60	62
0.98	62	58	55	57	58	60	63
0.99	105	101	98	94	—	—	—

It is as if we were solving by SOR a problem on a grid of size $h' \sim O(h^{1/2})$. There is no point in looking for more than two decimal places in ω. A similar table was constructed for each value of N. In Table II we give a summary of the best results

TABLE II

N	q	ω	I	ω_r	I_r
5	0.95	1.41	36	1.39–1.43	40–37
6	0.97	1.48	42	1.45–1.50	49–45
7	0.97	1.53	49	1.50–1.55	56–51
8	0.98	1.57	55	1.55–1.60	62–60
9	0.98	1.61	62	1.60–1.65	64–71
10	0.98	1.64	70	1.62–1.65	77–71

obtained for each N (the smallest number of iterations in each case), along with the corresponding q and ω. No attempt was ever made to work with more than two decimal places in the parameters. The last two columns show the variation in the required number of iterations for ω varying in the indicated range ω_r and with the value of q shown. Whenever different parameter values gave the same smallest number of iterations, the values shown are the largest $q \leqslant \hat{q}$ and then the smallest ω for which the tie occurred. The noticeable things in Table II are the slow growth in the number of iterations required and the absence of critical behavior in the number of iterations as a function of ω.

Finally, we shall say a few words about the rate of convergence. At the conclusion of each computation an average effective spectral radius was computed from the formula

$$a = \left(\frac{\text{final residual}}{\text{initial residual}}\right)^{1/I}$$

where I is the number of iterations.

For a given value of q, not necessarily optimal, the second-degree theory predicts that provided ω is chosen to be the optimal ω for this q, $\omega(q)$, then as $I \to \infty$, $a \to (\omega(q) - 1)^{1/2}$. Although the empirically detremined ω's are not optimal, we give in Table III the numbers a and $a' = (\omega'(q) - 1)^{1/2}$ where $w'(q)$ is the best ω found in each case, as indicated in Table I. There is tolerably good agreement to two decimal places indicating that within the limitations of the experiment, the best ω's are near to optimal. The q's are at most 0.01 smaller than their optimal values. In the last row of Table III we compute the quantity

TABLE III

N	5	6	7	8	9	10
a_N	0.638	0.694	0.726	0.752	0.778	0.797
a_N'	0.640	0.693	0.728	0.755	0.781	0.800
$1 - 2(\pi^{1/2})^{-1}h^{1/2}$	0.608	0.671	0.717	0.752	0.779	0.800

$1 - 2(\pi^{1/2})^{-1} h^{1/2}$. This formula was obtained from heuristic reasoning based on the explanations in Section 4 and on other empirical considerations. The agreement does lend support to the conjecture that the rate of convergence is not better than $O(h^{1/2})$. The corresponding work count is $O(n^{5/4})$ operations where n is the number of equations.

We should also say that the model problem that was considered probably shows the algorithm at its best for the particular choice of E that we made.

6. Concluding Remarks

It is clear that the multiple grid idea can be used in many ways. Obvious variations are the use of different matrices E and the use of different relaxation algorithms such as the Gauss–Seidel method. The variation we looked at has the merit of theoretical simplicity. Although this paper has been largely descriptive, we hope that it has demonstrated that multiple grid methods are of some usefulness and interest.

Acknowledgments

Part of this work was done while I was working at the California Institute of Technology and at Stanford University during 1973–1974. I am very grateful indeed to Profs. H. B. Keller and G. H. Golub for making these visits possible and for all their hospitality during this period.

In connection with the work reported herein, I wish to express my gratitude to Prof. R. S. Varga for his encouragement, advice, and many enjoyable conversations, and also for his comments on an earlier draft of the manuscript. Lastly, I would like to thank the referees for their comments.

References

1. N. Bakhvalov, *Zh. Vych. Mat.* **6** (1966), 861–885.
2. A. Brandt, *in* "Proceedings of the 3rd Conference on Numerical Methods in Fluid Mechanics," Lecture Notes in Physics, Vol. 18, pp. 82–89, Springer Verlag, Berlin, 1972.
3. R. P. Fedorenko, *Zh. Vych. Mat.* **1** (1961), 922–927.
4. R. P. Fedorenko, *Zh. Vych. Mat.* **4** (1964), 559–564.
5. A. George, *SIAM J. Num. Anal.* **10** (1973), 2.
6. G. H. Golub and R. S. Varga, *Num. Math.* **3** (1961), 147–156.
7. R. A. Nicolaides, *Num. Math.*, **24** (1975), 95–101.
8. D. J. Rose, "Graph Theory and Computing" (R. Read, Ed.), Academic Press, New York, 1972.
9. R. Southwell, "Relaxation methods in Theoretical Physics," Clarendon Press, Oxford, 1946.
10. E. Steifel, *Z. Angew. Math. Phys.* **3**, 1–33.
11. G. Strang and G. Fix, "An analysis of the Finite Element Method," Prentice-Hall, New Jersey, 1973.
12. J. L. Synge, "The Hypercircle in Mathematical Physics," Cambridge University Press, Cambridge, 1957.
13. F. de la Vallee Poussin, *SIAM J. Num. Anal.* **5** (1968), 340–351.
14. R. S. Varga, "Matrix Iterative Analysis," Prentice–Hall, New Jersey, 1962.
15. E. L. Wachspress, "Iterative Solution of Elliptic Systems," Prentice-Hall, New Jersey, 1966.
16. E. L. Wachspress, Variational acceleration of linear iteration. Proc. Army Workshop, Watervliet Arsenal, 1974.
17. D. M. Young, "Iterative Solution of Large Linear Systems," Academic Press, New York, 1971.