On the Rate of Convergence of an Alternating Direction Implicit Method in a Noncommutative Case*

By Olof B. Widlund

1. Introduction. It is well known that alternating direction implicit methods (ADI-methods) often provide very powerful techniques for computing numerical solutions to partial differential equations of elliptic type. The theoretical explanation for this is incomplete. The theory is in fact so unsatisfactory that it is often not possible to predict whether a suggested ADI-scheme is likely to converge or not.

Consider the following Dirichlet problem

$$(1 1) \qquad -\frac{\partial}{\partial x_1}\left(a_1(x)\,\frac{\partial u}{\partial x_1}\right) - \frac{\partial}{\partial x_2}\left(a_2(x)\,\frac{\partial u}{\partial x_2}\right) + c(x)u = f(x), \qquad x \in R;$$

$$u(x) = g(x), \qquad x \text{ on the boundary of } R,$$

where $x = (x_1, x_2)$ and R is the interior of a compact region. The coefficients a_1, a_2 and c are real valued sufficiently smooth functions. a_1 and a_2 are strictly positive and c is bounded from below by a sufficiently large lower bound c_0. Let us introduce a rectangular lattice of mesh points on the region. We approximate the boundary of the region by nearby mesh points at which the solution is approximately known. If we set up the standard five point difference approximation to (1.1) on this lattice, we get a system of linear equations with as many equations as unknowns. It can be shown that this system has a unique solution if c_0 is sufficiently large. It is well known that if the number of unknowns is very large it is often best to solve such a system with some iterative method. In this paper we will consider an ADI-method which for the case $a_1 = a_2 = 1$ and $c = 0$ is identical with the original ADI-method of Peaceman and Rachford [3]. It is known both from existing theory and numerical experiments that this ADI-method, using a sequence of different iteration parameters, is often superior to other iterative methods. Cf. Birkhoff, Varga and Young [2]. In practice one usually chooses the parameters in a cyclic way. In this paper we will always suppose that this is done and we will call the number of parameters m the cycle length. How to make the best choice for this number of steps, m, will not be discussed in this paper.

Before we describe our new result we will give a short survey of certain aspects of the theory for ADI-methods. For details we refer to Birkhoff, Varga and Young [2], Varga [5] and Wachspress [6].

Let us thus first consider Laplace's equation with Dirichlet data given on the boundary of a unit square. Suppose we use a uniform lattice of mesh points with

mesh size h. It is then possible, for any given m, to construct a set of iteration parameters for which the discrete L^2-norm of the error function is reduced by at least a factor

$$(1.2) \qquad p_m = \left(\frac{1 - (\tan (\pi h/2))^{1/m}}{1 + (\tan (\pi h/2))^{1/m}} \right)^2$$

for each full cycle of iterations. For the same problem the spectral radius of the point successive over relaxation operator is equal to p_1. The number of arithmetic operations used to complete a step in the ADI-scheme is of the same order as for the successive over relaxation method. These facts explain why ADI-methods can be much more efficient than other iterative methods when h is small and $m \geqq 2$. It has been shown that essentially the same rapid convergence is achieved in more general cases provided certain operators commute. Birkhoff and Varga [1] have however shown that these commutativity conditions impose very strong restrictions on both the coefficients and the region, which must then be rectangular.

Very little is known for noncommutative cases except when $m = 1$. In this one-parameter case there exists a simple convergence proof which at the same time gives an adequate estimate for the asymptotic rate of convergence. However, as we have seen earlier, $m = 1$ is the least interesting case. Some weak results have also been proved for the multi-parameter case. Thus Pearcy [4] has shown that we can always get convergence for any given problem and any given mesh if we choose m sufficiently large, have a mild restriction on the size of the iteration parameters and use them in a certain order. Pearcy's technique is however not refined enough to give more than a very weak estimate for the rate of convergence. But it is known from numerical experiments that the same rapid convergence as in commutative cases is possible even in noncommutative cases.

In this paper we will consider equations of type (1.1) in rectangular regions. Our operators will then in general not commute. For any given m we will construct iteration parameters which will make our scheme converge rapidly if the mesh size is sufficiently small. We will also be able to show that our estimate for the spectral radius of the operator which corresponds to a full cycle of iterations differs from that in a comparable commutative case only by a lower order term. Our technique would enable us to exhibit an estimate for this difference in terms of the mesh size, the iteration parameters, the coefficients and certain of their derivatives. Our iteration parameters will in general depend on the space variables. This dependence is however quite simple and it therefore adds little to the programming effort. Furthermore, a numerical experiment indicates that our theorems are not true if we restrict our choice to parameters which do not vary with the space variables. The result of this and other numerical experiments will be reported in a separate paper.

2. Preliminaries and Statement of Results.

To simplify the writing up of our proofs we will suppose that our region is a unit square and that our lattice of mesh points is uniform, i.e. $\Delta x_1 = \Delta x_2 = h$. The extension of our results to rectangular regions and rectangular lattices is immediate. We suppose that $h = 1/N$ for some integer N and that $h \leqq h_0$ where h_0 is a given constant. The interior of the lattice is defined by the cartesian product

$$R(h) = R_1(h) \times R_2(h),$$

where

$$R_i(h) = \{x_i \mid x_i = h, 2h, \cdots, 1 - h\}.$$

The boundary of the lattice is defined by

$$\partial R(h) = \partial R_1(h) \times R_2(h) \cup R_1(h) \times \partial R_2(h),$$

where

$$\partial R_i(h) = \{x_i \mid x_i = 0, 1\}.$$

Our difference approximation to (1.1) is the standard five point approximation

$$
\begin{aligned}
& -D_{-x_1}(a_1(x + (h/2)e_1)D_{+x_1}u(x)) \\
(2.1) \quad & -D_{-x_2}(a_2(x + (h/2)e_2)D_{+x_2}u(x)) + c(x)u(x) = f(x), \quad x \in R(h), \\
& \hspace{6.5cm} u(x) = g(x), \quad x \in \partial R(h),
\end{aligned}
$$

where e_i is the unit vector in the direction of the positive ith coordinate axis and

$$hD_{\pm x_i}u(x) = \pm(u(x \pm he_i) - u(x)).$$

Introduce the operators Q_i defined by

$$
\begin{aligned}
Q_i u(x) &= -D_{-x_i}(a_i(x + (h/2)e_i)D_{+x_i}u(x)) + c_i(x)u(x), \quad & x \in R(h) \\
&= 0, & x \in \partial R(h),
\end{aligned}
$$

where $c_1(x)$ and $c_2(x)$ are real valued sufficiently smooth functions which fulfill $c_1(x) + c_2(x) = c(x)$. If we consider $u(x)$ to be a vector with a component corresponding to each point of $R(h)$, we see easily that the centered difference structure implies that the operators Q_i, $i = 1, 2$, correspond to symmetric matrices and that they thus only have real eigenvalues. We will always assume that these eigenvalues are bounded from below by a strictly positive constant for all $h \leqq h_0$. We will refer to this condition as the *eigenvalue condition*.

We will discuss the eigenvalues of Q_i, $i = 1, 2$, in Section 3 and give a sufficient condition in terms of the coefficients of Q_i to ensure that the eigenvalue condition is fulfilled. In the same section we will also show that (2.1) has a unique solution and that the operators on the left-hand sides of our ADI-scheme

$$
\begin{aligned}
(\rho_n(x, h) + h^2 Q_1)u_{n+1/2} &= (\rho_n(x, h) - h^2 Q_2)u_n + h^2 f, \\
(\rho_n(x, h) + h^2 Q_2)u_{n+1} &= (\rho_n(x, h) - h^2 Q_1)u_{n+1/2} + h^2 f, \quad x \in R(h), \\
(2.2) \quad u_1(x) = \cdots = u_{n+1/2}(x) &= u_{n+1}(x) = g(x), \quad x \in \partial R(h), \\
n = 1, 2, \cdots, \quad u_1(x) \text{ given,} &
\end{aligned}
$$

have well defined inverses if $\rho_n(x, h) > 0$. Here $u_1(x)$ is the initial approximation to the solution of (2.1) and $\rho_n(x, h)$ are the iteration parameters. We will assume that these parameters satisfy

$$\rho_1(x, h) > \rho_2(x, h) > \cdots > \rho_m(x, h),$$

and

$$\rho_{n_1}(x, h) = \rho_n(x, h) \quad \text{if} \quad n_1 = n \bmod (m).$$

We will also assume that the parameters $\rho_n(x, h)$ can be written as a product of a fixed function of x and a function which depends on h and n. To simplify the notations later, we let $\rho_n(x, h)$ have the form

$$(2.3) \qquad \rho_n(x, h) = b^2(x)h^2(k_n(h))^{-1}$$

where $k_n(h) > 0$ and $b(x)$ is a sufficiently smooth, strictly positive function.

Now suppose that $u_1, u_2, \cdots$ converge to a limit. This limit must then be equal to the solution of (2.1). This fact enables us to reduce our problem to the study of the convergence to zero of the solution of the following error equation,

$$(2.4) \quad \begin{aligned} (\rho_n(x, h) + h^2 Q_1)\epsilon_{n+1/2} &= (\rho_n(x, h) - h^2 Q_2)\epsilon_n, \\ (\rho_n(x, h) + h^2 Q_2)\epsilon_{n+1} &= (\rho_n(x, h) - h^2 Q_1)\epsilon_{n+1/2}, \end{aligned} \qquad n = 1, 2, \cdots,$$

with

$$\epsilon_\nu = u_\nu - u,$$

where u is the solution of (2.1).

In the following we shall work almost exclusively with mesh functions which are zero for $x \in \partial R(h)$. We therefore introduce the space

$$U_0 = \{u \mid u(x) = 0, x \in \partial R(h)\},$$

with an inner product and a norm defined by

$$(u, v) = \sum_{x \in R(h)} h^2 \bar{u}(x) v(x), \qquad \| u \|^2 = (u, u).$$

By $\Omega(A)$ we denote the spectral radius of the operator A. It is well known that

$$(2.5) \qquad \Omega(A) \leqq \| A \|.$$

For $p \geqq q$ we will write $\prod_q^p A_\nu$ for $A_p A_{p-1} \cdots A_q$. We also define $\prod_q^{q-1} A_\nu = I =$ the identity operator.

Consider (2.4). To decide the convergence of our method we want to estimate

$$\Omega_m = \Omega \left\{ \prod_1^m (\rho_\nu + h^2 Q_2)^{-1}(\rho_\nu - h^2 Q_1)(\rho_\nu + h^2 Q_1)^{-1}(\rho_\nu - h^2 Q_2) \right\}.$$

Using (2.3) we see that

$$(2.6) \quad \begin{aligned} \rho_\nu - h^2 Q_i &= \frac{h^2}{k_\nu(h)} b(x)(1 - k_\nu Q_i')b(x) \quad \text{and} \\ (\rho_\nu + h^2 Q_i)^{-1} &= \frac{k_\nu(h)}{h^2} (b(x))^{-1}(1 + k_\nu Q_i')^{-1}(b(x))^{-1}, \end{aligned}$$

where

$$Q_i' = (b(x))^{-1}Q_i(b(x))^{-1}, \qquad i = 1, 2.$$

(2.6) tells us that

$$\Omega_m = \Omega\left\{(b(x))^{-1}\left(\prod_1^m (1 + k_\nu Q_2')^{-1}(1 - k_\nu Q_1')(1 + k_\nu Q_1')^{-1}\right.\right.$$

$$\left.\left.\times (1 - k_\nu Q_2')\right)b(x)\right\}.$$

By a similarity transformation we get

$$(2.7) \qquad \Omega_m = \Omega\left\{\prod_1^m (1 + k_\nu Q_2')^{-1}(1 - k_\nu Q_1')(1 + k_\nu Q_1')^{-1}(1 - k_\nu Q_2')\right\}.$$

The spectral radius Ω_m will eventually be estimated by using (2.5). Before this is done we will however carry out still another similarity transformation. Thus let us introduce

$$S_\nu = (1 + k_m Q_2')(1 + k_\nu Q_2')^{-1}(1 - k_\nu Q_1')(1 + k_\nu Q_1')^{-1}$$

$$\times (1 - k_\nu Q_2')(1 + k_m Q_2')^{-1}.$$

It is easily seen that

$$(2.8) \qquad \Omega_m = \Omega\left\{\prod_1^m S_\nu\right\} \leqq \left\|\prod_1^m S_\nu\right\|.$$

We will write $\prod_1^m S_\nu$ as the sum of two terms,

$$(2.9) \qquad \prod_1^m S_\nu = \prod_1^m T_{\nu 1} \prod_1^m T_{\nu 2} + R$$

where

$$T_{\nu i} = (1 + k_\nu Q_i')^{-1}(1 - k_\nu Q_i')$$

and the remainder R is defined by (2.9). In commutative cases $R = 0$. From (2.8) and (2.9) we obtain,

$$(2.10) \qquad \Omega_m \leqq \left\|\prod_1^m T_{\nu 1}\right\| \cdot \left\|\prod_1^m T_{\nu 2}\right\| + \|R\|.$$

We will show in Section 4 how to choose the k_ν to make the first of these terms small and in Sections 5 and 6 how to choose $b(x)$ to make $\|R\|$ very small. The first term in (2.10) can be treated just as if our problem were commutative.

As will be shown in Section 3, Q_i' as well as Q_i are selfadjoint operators and they therefore only have real eigenvalues. The bounds which we can get for $\|\prod_1^m T_{\nu i}\|$ depend on lower and upper bounds for these eigenvalues. Suppose that we know that the eigenvalues of Q_i', $i = 1, 2$, lie between α and β where $0 < \alpha < \beta$. We are then able to state our main result.

THEOREM 1. *If* $b^2(x) = a_1(x)$ *and we choose* k_ν *conveniently we have*

$$(2.11) \qquad \Omega_m \leqq \left(\frac{1 - (\alpha/\beta)^{1/2m}}{1 + (\alpha/\beta)^{1/2m}}\right)^2 + O(h^{3/2m}).$$

We give estimates for α and β in Section 3. There we will also show that

$$(2.12) \qquad \alpha/\beta = O(h^2)$$

if the eigenvalue condition is fulfilled. From (2.11) and (2.12) follows

$$\Omega_m \leqq \left(\frac{1 - q(h)}{1 + q(h)}\right)^2 + O(h^{3/2m}),$$

where $0 < q(h) = O(h^{1/m})$. Compare (1.2).

The case $a_1 = a_2$ is of course of particular interest in many applications. In that case we can prove a somewhat stronger theorem.

THEOREM 2. *Suppose that* $a_1(x) = a_2(x)$. *Then if* $b^2(x) = a_1(x)$ *and* k_ν *are chosen conveniently we have*

$$\| b\epsilon_{m+1} \| \leqq \left\{\left(\frac{1 - (\alpha/\beta)^{1/2m}}{1 + (\alpha/\beta)^{1/2m}}\right)^2 + O(h^{3/2m})\right\} \| b\epsilon_1 \|.$$

Remark. If we consider equations of type (1.1) for which the coefficients can be extended to be periodic (of period 2 in x_1 and x_2), even, sufficiently smooth functions for all values of x it can be shown that we do not necessarily need to choose $b^2(x) = a_1(x)$ to achieve rapid convergence. In fact it suffices to make max $| \partial/\partial x_2(a_1/b^2) |$ sufficiently small. The proof of this result is quite involved and we do not give it in this paper. Numerical experiments indicate that an upper bound for $| \partial/\partial x_2(a_1/b^2) |$ is needed to guarantee convergence for small values of m.

3. The Eigenvalues of Q_i and $Q_i{}'$.

We begin this section by proving

PROPOSITION 3.1. Q_i *and* $Q_i{}'$, $i = 1, 2$, *are selfadjoint operators, i.e.*

$$(Q_i u, v) = (u, Q_i v), \qquad (Q_i{}'u, v) = (u, Q_i{}'v),$$

for all $u, v \in U_0$. Q_i *and* $Q_i{}'$ *thus have only real eigenvalues.*

Proof. That $(Q_i u, v) = (u, Q_i v)$ follows immediately from the definitions of the inner product and Q_i, the fact that $u, v \in U_0$ and a summation by parts. The fact that $(Q_i{}'u, v) = (Q_i b^{-1}u, b^{-1}v)$ and $(u, Q_i{}'v) = (b^{-1}u, Q_i b^{-1}v) = (Q_i b^{-1}u, b^{-1}v)$ proves the second statement of the theorem.

We next give an estimate of the eigenvalues of Q_i.

PROPOSITION 3.2. *Suppose that* λ *is an eigenvalue of* Q_i *for a fixed value of* $x_j \in R_j$, $j \neq i$. *Then*

$$\min_{0 \leqq x_i \leqq 1} a_i(x)4 \sin^2 (\pi h/2)/h^2 + \min_{0 \leqq x_i \leqq 1} c_i(x) \leqq \lambda$$

$$\leqq \max_{0 \leqq x_i \leqq 1} a_i(x)4 \cos^2 (\pi h/2)/h^2 + \max_{0 \leqq x_i \leqq 1} c_i(x).$$

Before we give a proof we state a corollary to Proposition 3.2.

COROLLARY. *The eigenvalue condition is fulfilled if*

$$c_i(x) > -\min a_i(x)4 \sin^2 (\pi h_0/2)/h_0{}^2, \qquad\qquad i = 1, 2.$$

In this inequality the min are taken for all values of x in the closed unit square.

Proof of the proposition. We will use the following notation

$$(u, v)_{i:r,s} = \sum_{rh \leqq x_i \leqq sh} h\bar{u}(x)v(x),$$

$$\| u \|_{i:r,s}^2 = (u, u)_{i:r,s},$$

where $0 \leqq r < s \leqq N$ and u and v are mesh functions defined on $R_i(h) \cup \partial R_i(h)$. Let $\phi(x_i) \in U_0$ be an eigenfunction of Q_i for which $\| \phi \|_{i:1,N-1} = 1$ and

$$Q_i\phi = \lambda\phi, \qquad\qquad x_j \in R_j \text{ fixed}, \quad j \neq i.$$

Then

$$\lambda = (\lambda\phi, \phi)_{i:1,N-1} = (-D_{-x_i}a_i(x + (h/2)e_i)D_{+x_i}\phi, \phi)_{i:1,N-1} + (c_i\phi, \phi)_{i:1,N-1}.$$

Writing out the sums it is easy to see that

$$(-D_{-x_i}a_i(x + (h/2)e_i)D_{+x_i}\phi, \phi)_{i:1,N-1} = (a_i(x + (h/2)e_i)D_{+x_i}\phi, D_{+x_i}\phi)_{i:0,N-1}.$$

Thus

$$\min_{0 \leqq x_i \leqq 1} a_i \| D_{+x_i}\phi \|^2_{i:0,N-1} + \min_{0 \leqq x_i \leqq 1} c_i \leqq \lambda \leqq \max_{0 \leqq x_i \leqq 1} a_i \| D_{+x_i}\phi \|^2_{i:0,N-1} + \max_{0 \leqq x_i \leqq 1} c_i.$$

To finish our proof, we show that

$$(3.1) \qquad 4 \sin^2(\pi h/2)/h^2 \leqq \| D_{+x_i}\phi \|^2_{i:0,N-1} \leqq 4 \cos^2(\pi h/2)/h^2.$$

To do this we develop $\phi(x_i)$ in a discrete Fourier series

$$\phi(x_i) = \sum_{\nu=1}^{N-1} \alpha_\nu \sin \pi x_i \nu.$$

It is easy to show that $(\sin \pi x_i \nu, \sin \pi x_i \mu)_{i:1,N-1} = \frac{1}{2}\delta_{\nu\mu}$ if $1 \leqq \nu, \mu \leqq N - 1$. We know that

$$\| \phi \|^2_{i:1,N-1} = \frac{1}{2}\sum_{\nu=1}^{N-1} | \alpha_\nu |^2 = 1.$$

Furthermore

$$\| D_{+x_i}\phi \|^2_{i:0,N-1} = -(\phi, D_{-x_i}D_{+x_i}\phi)_{i:1,N-1}$$

$$= \frac{1}{2}\sum_{\nu=1}^{N-1} 4 \sin^2(\pi\nu h/2)/h^2 | \alpha_\nu |^2,$$

which proves (3.1).

The eigenvalues of Q_i' can be estimated in terms of the eigenvalues of Q_i.

PROPOSITION 3.3. *Suppose that for a given* x_j, $j \neq i$, *the eigenvalues of* Q_i *fulfill*

$$0 < \alpha(x_j) \leqq \lambda \leqq \beta(x_j).$$

Then the eigenvalues λ' *of* Q_i' *satisfy*

$$0 < \alpha(x_j)/\max_{0 \leqq x_i \leqq 1} b^2(x) \leqq \lambda' \leqq \beta(x_j)/\min_{0 \leqq x_i \leqq 1} b^2(x).$$

Proof. Let

$$Q_i'\phi = \lambda'\phi$$

where $\| \phi \|_{i:1,N-1} = 1$. Since Q_i is a selfadjoint operator $\alpha(x_j)\| \Psi \|^2_{i:1,N-1} \leqq (Q_i\Psi, \Psi)_{i:1,N-1} \leqq \beta(x_j)\| \Psi \|^2_{i:1,N-1}$, for all $\Psi \in U_0$. By the definition of Q_i' we get

$$\lambda' = (\lambda'\phi, \phi)_{i:1,N-1} = (Q_i'\phi, \phi)_{i:1,N-1}$$

$$= (Q_i(b)^{-1}\phi, (b)^{-1}\phi)_{i:1,N-1}.$$

Thus

$$\left(\alpha(x_j)/\max_{0 \le x_i \le 1} b^2(x)\right) \le \alpha(x_j)\| b^{-1}\phi \|^2_{i:1,N-1}$$

$$\le (Q_i b^{-1}\phi, b^{-1}\phi)_{i:1,N-1} \le \beta(x_j)\| b^{-1}\phi \|^2_{i:1,N-1}$$

$$\le \left(\beta(x_j)/\min_{0 \le x_i \le 1} b^2(x)\right),$$

which completes the proof.

We may use Propositions 3.2 and 3.3 to estimate the quantities α and β which were introduced in Section 2. These propositions also tell us that if the eigenvalue condition is fulfilled we always have $\alpha/\beta = O(h^2)$. It should however be pointed out that the values of α and β which can be computed from Propositions 3.2 and 3.3 often are rather pessimistic and that it is sometimes advisable to try to get better bounds by some other method since the error estimates are then improved. It is also known from experience that the efficiency of the ADI-scheme depends very much on the choice of α and β.

We next show some simple consequences of the eigenvalue condition.

PROPOSITION 3.4. *Equation* (2.1) *has a unique solution.*

Proof. (2.1) corresponds to a system of linear equations with as many equations as unknowns. We therefore only have to show that $u \equiv 0$ is the only solution to (2.1) for $f(x) = g(x) \equiv 0$. Suppose that $v(x) \ne 0$ and

$$(Q_1 + Q_2)v = 0.$$

This is not possible, since the Q_i are selfadjoint operators with strictly positive eigenvalues. That is, there is a $c > 0$ such that

$$((Q_1 + Q_2)v, v) \ge c\| v \|^2,$$

for all $v \in U_0$.

This contradiction establishes the proposition.

PROPOSITION 3.5.

$$(\rho_n(x, h) + h^2 Q_i)u(x_i) = f(x_i), \qquad x_i \in R_i(h),$$

$$u(0) = u_0, \qquad u(1) = u_1$$

has a unique solution.

Proof. Observe that by (2.6)

$$\rho_n(x, h) + h^2 Q_i = \frac{h^2}{k_n(h)} b(x)(1 + k_n Q_i')b(x).$$

By Proposition 3.3 the eigenvalues of Q_i' are positive, hence this operator has a unique inverse.

Proposition 3.5 shows that $u_{3/2}, u_2, \cdots$, defined by (2.2) are all well defined. The following lemmas will be used later.

LEMMA 3.1. *For any* $k > 0$

$$\|(1 + kQ_i')^{-1} \| \le \frac{1}{1 + k\alpha} < 1,$$

$$\|(1 + kQ_i')^{-1}kQ_i' \| \le \frac{k\beta}{1 + k\beta} < 1,$$

$$\|(1 + kQ_i')^{-1}(1 - kQ_i')\| \le \max\left(\frac{1 - k\alpha}{1 + k\alpha}, \frac{k\beta - 1}{k\beta + 1}\right) < 1.$$

Proof. Q_i' are selfadjoint operators and they therefore have complete sets of orthonormal eigenfunctions. Develop $u \in U_0$ in series of eigenfunctions of Q_i'.

$$u = \sum_{\nu=1}^{N-1} \alpha_\nu(x_j)\phi_\nu(x_i), \quad j \neq i, \qquad (\phi_\nu, \phi_\mu)_{i:1,N-1} = \delta_{\nu\mu}.$$

Now

$$(1 + kQ_i')^{-1}u = \sum_{\nu=1}^{N-1} \frac{\alpha_\nu(x_j)}{1 + k\lambda_\nu(x_j)} \cdot \phi_\nu(x_i)$$

where $\lambda_\nu(x_j)$ is the eigenvalue corresponding to $\phi_\nu(x_i)$. Thus

$$\| (1 + kQ_i')^{-1}u \|^2 = \sum_{\nu=1}^{N-1} \| \alpha_\nu/(1 + k\lambda_\nu) \|^2_{j:1,N-1}$$

$$\leq \left(\frac{1}{1 + k\alpha}\right)^2 \sum_{\nu=1}^{N-1} \| \alpha_\nu \|^2_{j:1,N-1} = \left(\frac{1}{1 + k\alpha}\right)^2 \| u \|^2.$$

The other inequalities follow in the same way.

LEMMA 3.2.

$$\left\| \prod_1^m T_{\nu i} \right\| \leq \max_{\alpha \leq \lambda \leq \beta} \prod_1^m \left| \frac{1 - k_\nu\lambda}{1 + k_\nu\lambda} \right|.$$

Proof. Cf. the proof of Lemma 3.1.

4. The Choice of k_ν. It is obvious from Lemma 3.2 why we are interested in making

$$\max_{\alpha \leq \lambda \leq \beta} \prod_1^m \left| \frac{1 - k_\nu\lambda}{1 + k_\nu\lambda} \right|$$

small. This problem has been extensively treated in the literature and hence we will give it only a brief treatment. We refer the reader to [2], [5] or [6] for details.

The optimal as well as several other good choices of k_ν are known. The perhaps simplest choice is

$$k_\nu = \frac{1}{\beta}\left(\frac{\beta}{\alpha}\right)^{(2\nu-1)/2m}, \qquad\qquad \nu = 1, 2, \cdots, m,$$

which corresponds to the parameters suggested in the original paper by Peaceman and Rachford [3].

With this choice we get

$$\left\| \prod_1^m T_{\nu i} \right\| \leq \max_{\alpha \leq \lambda \leq \beta} \prod_1^m \left| \frac{1 - k_\nu\lambda}{1 + k_\nu\lambda} \right| \leq \frac{1 - (\alpha/\beta)^{1/2m}}{1 + (\alpha/\beta)^{1/2m}}$$

(4.1)

$$= \frac{1 - q(h)}{1 + q(h)}, \quad \text{where} \quad 0 < q(h) = O(h^{1/m}).$$

The last equality follows from the fact that $\alpha/\beta = O(h^2)$. Using the results of Section 3 we see that

(4.2) $$\qquad\qquad k_\nu = O(h^{(2(m-\nu)+1)/m}), \qquad\qquad \nu = 1, 2, \cdots m.$$

It is easy to show that (4.1) and (4.2) are true also for the optimal parameters due to Jordan. Cf. Wachspress [6]. In the proof of our theorems we will suppose that the k_ν are chosen so that (4.1) and (4.2) hold. But we could by our technique prove convergence when $h \to 0$ for all sets of k_ν for which

$$\max_{\alpha \leq \lambda \leq \beta} \prod_1^m \left| \frac{1 - k_\nu \lambda}{1 + k_\nu \lambda} \right|^2 \leqq \left(\frac{1 - c(\alpha, \beta)}{1 + c(\alpha, \beta)} \right)^2$$

where $0 < c(\alpha, \beta) < 1$,

$$h(k_m)^{1/2}/c(\alpha, \beta) \to 0 \quad \text{when} \quad h \to 0,$$

and for $\nu = 1, 2, \cdots, m - 1$

$$(k_\nu)^{1/2}/c(\alpha, \beta) \to 0 \quad \text{when} \quad h \to 0.$$

5. A Fundamental Inequality. In this section we will derive an inequality for the operators

$$K_{\nu\mu} = k_\nu k_\mu (1 + k_\nu Q_1')^{-1}(Q_1' Q_2' - Q_2' Q_1')(1 + k_\mu Q_2')^{-1}$$

which will play an important part in the proofs of our theorems.

LEMMA 5.1.

$$\| K_{\nu\mu} \| \leqq 2(k_\mu)^{1/2} \max_{0 \leqq x_1, x_2 \leqq 1} \left| b(x) \frac{a_2(x)}{a_1(x)} \frac{\partial}{\partial x_2} (a_1(x)/b^2(x)) \right|$$

$$\times \left(4 \min_{0 \leqq x_1, x_2 \leqq 1} a_2 \left(1 + k_\mu \min_{0 \leqq x_1, x_2 \leqq 1} (b^{-2} c_2) \right) \right)^{1/2}$$

$$+ 2(k_\nu)^{1/2} \max_{0 \leqq x_1, x_2 \leqq 1} \left| b(x) \frac{a_1(x)}{a_2(x)} \frac{\partial}{\partial x_1} (a_2(x)/b^2(x)) \right|$$

$$\times \left(4 \min_{0 \leqq x_1, x_2 \leqq 1} a_1 \left(1 + k_\nu \min_{0 \leqq x_1, x_2 \leqq 1} (b^{-2} c_1) \right) \right)^{-1/2}$$

$$+ k_\mu \max_{0 \leqq x_1, x_2 \leqq 1} \left| \frac{1}{a_1(x)} \cdot \frac{\partial}{\partial x_2} \left(a_2(x) \frac{\partial}{\partial x_2} (a_1(x)/b^2(x)) \right) \right|$$

$$+ k_\nu \max_{0 \leqq x_1, x_2 \leqq 1} \left| \frac{1}{a_2(x)} \cdot \frac{\partial}{\partial x_1} \left(a_1(x) \frac{\partial}{\partial x_1} (a_2(x)/b^2(x)) \right) \right|$$

$$+ O(h(k_\nu)^{1/2} + h(k_\mu)^{1/2} + (k_\nu)^{1/2} k_\mu + (k_\mu)^{1/2} k_\nu).$$

The following corollary follows immediately from the lemma.

COROLLARY. *If $b^2 = a_1$ we have*

$$\| K_{\nu\mu} \| = O((k_\nu)^{1/2} + h(k_\mu)^{1/2}.$$

If furthermore $a_1 = a_2$ we have

$$\| K_{\nu\mu} \| = O(h(k_\nu)^{1/2} + h(k_\mu)^{1/2} + k_\nu(k_\mu)^{1/2} + k_\mu(k_\nu)^{1/2}).$$

Proof of the lemma. We first consider the operator $Q_1' Q_2' - Q_2' Q_1'$ which maps U_0 into U_0. For $x \in R(h)$ this operator is defined as the commutator of two second order difference operators and it can therefore be expressed as a third order differ-

ence operator. Using discrete counterparts of Leibnitz's rule we may write

$$(Q_1'Q_2' - Q_2'Q_1')u =$$

$$Q_1'(A_{21}^+(x, h)D_{+x_2}b^{-1} + A_{21}^-(x, h)D_{-x_2}b^{-1} + A_{20}(x, h))u$$

$$+ (b^{-1}D_{+x_1}A_{12}^+(x, h) + b^{-1}D_{-x_1}A_{12}^-(x, h) + A_{02}(x, h))Q_2'u$$

(5.1)
$$+ b^{-1}D_{+x_1}A_{11}^{++}(x, h)D_{+x_2}b^{-1}u + b^{-1}D_{-x_1}A_{11}^{-+}(x, h)D_{+x_2}b^{-1}u$$

$$+ b^{-1}D_{+x_1}A_{11}^{+-}(x, h)D_{-x_2}b^{-1}u + b^{-1}D_{-x_1}A_{11}^{--}(x, h)D_{-x_2}b^{-1}u$$

$$+ b^{-1}D_{+x_1}A_{10}^+(x, h)u + b^{-1}D_{-x_1}A_{10}^-(x, h)u + A_{01}^+(x, h)D_{+x_2}b^{-1}u$$

$$+ A_{01}^-(x, h)D_{-x_2}b^{-1}u + A_{00}(x, h)u,$$

for $x \in R(h)$ and $u \in U_0$. Formulas for the coefficients $A_{21}^+(x, h)$ etc. can be given in terms of a_1, a_2, b, c_1 and c_2 and certain of their divided differences. Because of our assumptions on the smoothness of a_1, etc. we see that all the coefficients in (5.1) are uniformly bounded functions of x and h. It can be shown that

$$A_{21}^+(x, h) = -b(x) \frac{a_2(x)}{a_1(x)} \cdot \frac{\partial}{\partial x_2}(a_1(x)/b^2(x)) + O(h),$$

$$A_{21}^-(x, h) = A_{21}^+(x, h) + O(h),$$

$$A_{12}^+(x, h) = b(x) \frac{a_1(x)}{a_2(x)} \frac{\partial}{\partial x_1}(a_2(x)/b^2(x)) + O(h),$$

$$A_{12}^-(x, h) = A_{12}^+(x, h) + O(h),$$

$$A_{20}(x, h) = -\frac{1}{a_1(x)} \frac{\partial}{\partial x_2}\left(a_2(x) \frac{\partial}{\partial x_2}(a_1(x)/b^2(x))\right) + O(h),$$

$$A_{02}(x, h) = -\frac{1}{a_2(x)} \frac{\partial}{\partial x_1}\left(a_1(x) \frac{\partial}{\partial x_1}(a_2(x)/b^2(x))\right) + O(h).$$

We can extend the formula (5.1) to be valid even for $x \in \partial R(h)$ if we replace $D_{\pm x_i}$ by $\tilde{D}_{\pm x_i}$ in (5.1), where

$$\tilde{D}_{\pm x_i}u = D_{\pm x_i}u, \qquad\qquad x \in R(h),$$

$$= 0, \qquad\qquad x \in \partial R(h).$$

The $\tilde{D}_{\pm x_i}$ are thus mappings of U_0 into U_0. We extend our coefficients, when needed, by defining them to be zero for $x \in \partial R(h)$. Our formula (5.1) will then not only give a correct representation for $(Q_1'Q_2' - Q_2'Q_1')u$ in the interior of our mesh but will also give a representation of $Q_1'Q_2' - Q_2'Q_1'$ as the sum of operators which have the form

$$T_1 A(x, h) T_2$$

where the T_i, $i = 1, 2$, map U_0 into U_0, T_i only operates in the x_i-direction and $A(x, h)$ is a uniformly bounded function. Now

$$\| k_\nu k_\mu (1 + k_\nu Q_1')^{-1} T_1 A(x, h) T_2 (1 + k_\mu Q_2')^{-1} \|$$

$$\leq \max_{0 \leq x_1, x_2 \leq 1; h \leq h_0} | A(x, h) | \cdot \| (1 + k_\nu Q_1')^{-1} k_\nu T_1 \| \cdot \| k_\mu T_2 (1 + k_\mu Q_2')^{-1} \|.$$

To finish the proof of Lemma 5.1 we only have to use Lemma 3.1 and the following lemma on the different terms of $K_{\nu\mu}$.

LEMMA 5.2. *For $k > 0$*

$$(5.2) \quad \| (1 + kQ_i')^{-1}kb^{-1}\tilde{D}_{\pm x_i} \| = \| k\tilde{D}_{\mp x_i}b^{-1}(1 + kQ_i')^{-1} \|$$

$$\leq k^{1/2}\Big(4 \min_{0 \leq x_1, x_2 \leq 1} a_i \cdot \Big(1 + k \min_{0 \leq x_1, x_2 \leq 1} b^{-2}c_i\Big)\Big)^{-1/2}.$$

Proof of Lemma 5.2. It is easy to see that $-\tilde{D}_{-x_i}$ is the adjoint operator of $\tilde{D}_{+x_i}$. Using Proposition 3.1 and the fact that adjoint operators have the same norm we are able to prove the equality in the lemma. Next define

$$v \equiv (1 + kQ_i')^{-1}k^{1/2}b^{-1}\tilde{D}_{-x_i}u, \qquad\qquad u \in U_0$$

so that

$$(1 + kQ_i')v = k^{1/2}b^{-1}\tilde{D}_{-x_i}u.$$

Form the scalar products of both sides with v

$$(v, (1 + kQ_i')v)_{i:1,N-1} = (v, k^{1/2}b^{-1}\tilde{D}_{-x_i}u)_{i:1,N-1}.$$

Observe that $v \in U_0$. Hence from considerations very similar to those in Section 3

$$(v, (1 + kQ_i')v)_{i:1,N-1} = \| v \|_{i:1,N-1}^2$$

$$- k(v, b^{-1}D_{-x_i}a_i(x + (h/2)e_i)D_{+x_i}b^{-1}v)_{i:1,N-1}$$

$$+ k(v, b^{-2}c_iv)_{i:1,N-1}$$

$$= \| v \|_{i:1,N-1}^2 + k(D_{+x_i}b^{-1}v, a_i(x + (h/2)e_i)D_{+x_i}b^{-1}v)_{i:0,N-1}$$

$$+ k(v, b^{-2}c_iv)_{i:1,N-1}$$

$$\geq \Big(1 + k \min_{0 \leq x_1, x_2 \leq 1} (b^{-2}c_i)\Big)\| v \|_{i:1,N-1}^2$$

$$+ \min_{0 \leq x_1, x_2 \leq 1} a_i k \| D_{+x_i}b^{-1}v \|_{i:0,N-1}^2.$$

Next we show that

$$(v, k^{1/2}b^{-1}\tilde{D}_{-x_i}u)_{i:1,N-1} = (v, k^{1/2}b^{-1}D_{-x_i}u)_{i:1,N} = -k^{1/2}(D_{+x_i}b^{-1}v, u)_{i:0,N-1}$$

$$\leq k^{1/2} \| D_{+x_i}b^{-1}v \|_{i:0,N-1} \| u \|_{i:0,N-1}$$

$$\leq k(C/2)\| D_{+x_i}b^{-1}v \|_{i:0,N-1}^2 + (1/2C)\| u \|_{i:1,N-1}^2$$

for all $C > 0$. If we choose $C = 2\min_{0 \leq x_1, x_2 \leq 1} a_i$ we get

$$\| (1 + kQ_i')^{-1}k^{1/2}b^{-1}\tilde{D}_{-x_i}u \|_{i:1,N-1}^2$$

$$\leq \Big\{4 \min_{0 \leq x_1, x_2 \leq 1} a_i \cdot \Big[1 + k \min_{0 \leq x_1, x_2 \leq 1} (b^{-2}c_i)\Big]\Big\}^{-1} \| u \|_{i:1,N-1}^2.$$

If we sum this inequality over R_j, $j \neq i$, we get one of the inequalities in (5.2). The case with $\tilde{D}_{+x_i}$ instead of $\tilde{D}_{-x_i}$ can be treated in a quite similar way.

6. Proofs of the Theorems. We will now estimate the norm of the remainder R which was defined by (2.9).

LEMMA 6.1.

(6.1)
$$\| R \| \leqq \prod_{\nu=1}^{m-1} (1 + \| R_\nu \|) - 1 + \sum_{\nu=1}^{m-1} \| R_\nu' \|$$

where

(6.2)
$$\| R_\nu \| \leqq 2\{\| K_{\nu\nu} \|/(1 - \| K_{\nu\nu} \|) + \| K_{\nu m} \|/(1 - \| K_{\nu m} \|)\},$$

and

(6.3)
$$\| R_\nu' \| \leqq 4 \sum_{\gamma=1}^{\nu} \| K_{\gamma,\nu+1} \|/(1 - \| K_{\gamma,\nu+1} \|).$$

Combining this lemma with (4.1) and Lemma 5.1 it is easy to finish the proof of Theorem 1.

Proof of the lemma. The remainder R was defined by

$$R = \prod_1^m S_\nu - \prod_1^m T_{\nu 1} \prod_1^m T_{\nu 2}$$

where

$$S_\nu = (1 + k_m Q_2')(1 + k_\nu Q_2')^{-1}(1 - k_\nu Q_1')(1 + k_\nu Q_1')^{-1}(1 - k_\nu Q_2')(1 + k_m Q_2')^{-1}$$

and

$$T_{\nu i} = (1 + k_\nu Q_i')^{-1}(1 - k_\nu Q_i').$$

R can be written as

(6.4)
$$R = \left(\prod_1^m S_\nu - \prod_1^m T_{\nu 1} T_{\nu 2} \right) + \left(\prod_1^m T_{\nu 1} T_{\nu 2} - \prod_1^m T_{\nu 1} \prod_1^m T_{\nu 2} \right).$$

Introduce

$$R_\nu \equiv S_\nu - T_{\nu 1} T_{\nu 2}.$$

Observe that $R_m = 0$. Obviously for $m = 2$

$$\prod_1^2 S_\nu - \prod_1^2 T_{\nu 1} T_{\nu 2} = T_{21} T_{22} R_1,$$

whence by Lemma 3.1

$$\left\| \prod_1^2 S_\nu - \prod_1^2 T_{\nu 1} T_{\nu 2} \right\| \leqq \| R_1 \|.$$

For $m = 3$,

$$\prod_1^3 S_\nu - \prod_1^3 T_{\nu 1} T_{\nu 2} = T_{31} T_{32}(T_{21} T_{22} R_1 + R_2 T_{11} T_{12} + R_2 \cdot R_1)$$

whence

$$\left\| \prod_1^3 S_\nu - \prod_1^3 T_{\nu 1} T_{\nu 2} \right\| \leqq \| R_1 \| + \| R_2 \| + \| R_1 \| \cdot \| R_2 \|$$

$$= \prod_1^2 (1 + \| R_\nu \|) - 1.$$

In exactly the same way we can show that

$$(6.5) \qquad \left\| \prod_1^m S_\nu - \prod_1^m T_{\nu 1} T_{\nu 2} \right\| \leqq \prod_1^{m-1} (1 + \| R_\nu \|) - 1.$$

Next consider the second term in (6.4). We easily get

$$(6.6) \qquad \prod_1^2 T_{\nu 1} T_{\nu 2} - \prod_1^2 T_{\nu 1} \prod_1^2 T_{\nu 2} = T_{21}(T_{22}T_{11} - T_{11}T_{22})T_{12}.$$

In the same way

$$\prod_1^3 T_{\nu 1} T_{\nu 2} - \prod_1^3 T_{\nu 1} \prod_1^3 T_{\nu 2}$$

$$= T_{31}T_{32} \left(\prod_1^2 T_{\nu 1} T_{\nu 2} - \prod_1^2 T_{\nu 1} \prod_1^2 T_{\nu 2} \right) + T_{31} \left(T_{32} \prod_1^2 T_{\nu 1} - \prod_1^2 T_{\nu 1}T_{32} \right) \prod_1^2 T_{\nu 2} .$$

By (6.6)

$$\prod_1^3 T_{\nu 1} T_{\nu 2} - \prod_1^3 T_{\nu 1} \prod_1^3 T_{\nu 2} = T_{31}T_{32}T_{21}(T_{22}T_{11} - T_{11}T_{22})T_{12}$$

$$+ T_{31}(T_{32}T_{21} - T_{21}T_{32})T_{11} \prod_1^2 T_{\nu 2} + T_{31}T_{21}(T_{32}T_{11} - T_{11}T_{32}) \prod_1^2 T_{\nu 2} .$$

Proceeding in this fashion we can show that

$$(6.7) \qquad \prod_1^m T_{\nu 1} T_{\nu 2} - \prod_1^m T_{\nu 1} \prod_1^m T_{\nu 2} = \sum_{N=1}^{m-1} R_\mu'$$

where

$$R_\mu' \equiv \prod_{\nu=\mu+2}^m T_{\nu 1} T_{\nu 2} \cdot T_{\mu+1,1} \sum_{\gamma=1}^\mu \left\{ \prod_{\delta=\gamma+1}^\mu T_{\delta 1} (T_{\mu+1,2}T_{\gamma 1} - T_{\gamma 1}T_{\mu+1,2}) \prod_{\rho=1}^{\gamma-1} T_{\rho 1} \right\} \prod_{\nu=1}^\mu T_{\nu 2} .$$

To prove (6.1), we use (6.4), (6.5) and (6.7).
We now prove (6.2). It can be shown by elementary calculations that

$$R_\nu = -2(1 - (k_\nu/k_m))$$

$$(6.8) \qquad \times \{(1 + k_\nu Q_1')^{-1}(1 + k_\nu Q_2')^{-1} - (1 + k_\nu Q_2')^{-1}(1 + k_\nu Q_1')^{-1}\}$$

$$+ 2(k_m/k_\nu)(1 - (k_\nu/k_m)^2)\{(1 + k_\nu Q_1')^{-1}(1 + k_\nu Q_2')^{-1}$$

$$- (1 + k_\nu Q_2')^{-1}(1 + k_\nu Q_1')^{-1}\}(1 + k_m Q_2')^{-1}.$$

We now show how to estimate the norm of the first term in terms of the norm of $K_{\nu\nu}$. Define P_ν and $P_{\nu\nu}$ by

$$P_\nu \equiv (1 + k_\nu Q_1')^{-1}(1 + k_\nu Q_2')^{-1} - (1 + k_\nu Q_2')^{-1}(1 + k_\nu Q_1')^{-1}$$

$$(6.9) \qquad = (1 - k_\nu Q_2')^{-1}\{(1 + k_\nu Q_1')^{-1}k_\nu^2(Q_1'Q_2' - Q_2'Q_1')$$

$$\cdot (1 + k_\nu Q_1')^{-1}(1 + k_\nu Q_2')^{-1}\} \equiv (1 + k_\nu Q_2')^{-1}P_{\nu\nu} .$$

But

$$P_{\nu\nu} = (1 + k_\nu Q_1')^{-1} k_\nu^2 (Q_1'Q_2' - Q_2'Q_1')(1 + k_\nu Q_1')^{-1}(1 + k_\nu Q_2')^{-1}$$

$$= (1 + k_\nu Q_1')^{-1} k_\nu^2 (Q_1'Q_2' - Q_2'Q_1')(1 + k_\nu Q_2')^{-1}(1 + k_\nu Q_1')^{-1}$$

$$+ (1 + k_\nu Q_1')^{-1} k_\nu^2 (Q_1'Q_2' - Q_2'Q_1')\{(1 + k_\nu Q_1')^{-1}(1 + k_\nu Q_2')^{-1}$$

$$- (1 + k_\nu Q_2')^{-1}(1 + k_\nu Q_1')^{-1}\}$$

$$= K_{\nu\nu} \cdot (1 + k_\nu Q_1')^{-1} + K_{\nu\nu} \cdot P_{\nu\nu}$$

where we used (6.9) to obtain the last identity. Thus by using Lemma 3.1

$$\| P_\nu \| \leqq \| P_{\nu\nu} \| \leqq \| K_{\nu\nu} \| + \| K_{\nu\nu} \| \cdot \| P_{\nu\nu} \|.$$

Hence

$$\| P_\nu \| \leqq \| P_{\nu\nu} \| \leqq \| K_{\nu\nu} \|/(1 - \| K_{\nu\nu} \|).$$

The estimation of the norm of the second term in (6.8) requires some special care. We use P_ν' and $P_{\nu m}$ defined by

$$P_\nu' \equiv (k_m/k_\nu)\{(1 + k_\nu Q_1')^{-1}(1 + k_\nu Q_2')^{-1}$$

$$- (1 + k_\nu Q_2')^{-1}(1 + k_\nu Q_1')^{-1}\}(1 + k_m Q_2')^{-1}$$

$$= (1 + k_\nu Q_2')^{-1}\{(1 + k_\nu Q_1')^{-1} k_\nu k_m (Q_1'Q_2' - Q_2'Q_1')(1 + k_\nu Q_1')^{-1}$$

$$\times (1 + k_m Q_2')^{-1}\}(1 + k_\nu Q_2')^{-1}$$

$$\equiv (1 + k_\nu Q_2')^{-1} \cdot P_{\nu m}(1 + k_\nu Q_2')^{-1}.$$

$P_{\nu m}$ can now be treated as $P_{\nu\nu}$ above. We get

$$\| P_\nu' \| \leqq \| P_{\nu m} \| \leqq \| K_{\nu m} \|/(1 - \| K_{\nu m} \|).$$

To finish our proof of (6.2) we use the definition of R_ν and remember that

$$k_\nu < k_m, \qquad\qquad \nu = 1, 2, \cdots, m - 1.$$

The proof of (6.3) is quite similar. Using the definition of R_μ' and Lemma 3.1 we get

$$\| R_\mu' \| \leqq \sum_{\gamma=1}^{\mu} \| T_{\gamma 1} T_{\mu+1,2} - T_{\mu+1,2} T_{\gamma 1} \|.$$

We can treat these terms almost exactly as we did the terms in R_ν after having observed that

$$T_{\gamma 1} T_{\mu+1,2} - T_{\mu+1,2} T_{\gamma 1}$$

$$= 4\{(1 + k_\gamma Q_1')^{-1}(1 + k_{\mu+1} Q_2')^{-1} - (1 + k_{\mu+1} Q_2')^{-1}(1 + k_\gamma Q_1')^{-1}\}.$$

This finishes the proof of Lemma 6.1 and thus also the proof of Theorem 1.

We now indicate how to modify the proof above to get a proof of Theorem 2. Our first modification consists in not carrying out the last similarity transformation in Section 2. We instead define our remainder as

$$R' = \prod_1^m (1 + k_\nu Q_2')^{-1}(1 - k_\nu Q_1')(1 + k_\nu Q_1')^{-1}(1 - k_\nu Q_2') - \prod_1^m T_{\nu 1} \prod_1^m T_{\nu 2}.$$

By carrying out an analysis of this remainder analogous to that of R in the proof of Theorem 1, we can prove a lemma corresponding to Lemma 6.1. The main difference in these two lemmas is that the estimate of $\| R' \|$ will contain terms proportional to $\| K_{mm} \|$. However as the corollary to Lemma 5.1 shows us this is not dangerous if $a_1 = a_2$.

To finish our proof we only have to note that

$$b\epsilon_{m+1} = b \prod_1^m (\rho_\nu + h^2 Q_2)^{-1}(\rho_\nu + h^2 Q_1)(\rho_\nu + h^2 Q_1)^{-1}(\rho_\nu - h^2 Q_2)\epsilon_1$$

$$= \prod_1^m (1 + k_\nu Q_2')^{-1}(1 - k_\nu Q_1')(1 + k_\nu Q_1')^{-1}(1 - k_\nu Q_2')b\epsilon_1$$

$$= \prod_1^m T_{\nu 1} \prod_1^m T_{\nu 2} b\epsilon_1 + R' b\epsilon_1 .$$

Courant Institute of Mathematical Sciences
New York University
New York, New York 10012

1. G. BIRKHOFF & R. S. VARGA, "Implicit alternating direction methods," *Trans. Amer· Math. Soc.*, v. 92, 1959, pp. 13–24. MR **21** #4549.
2. G. BIRKHOFF, R. S. VARGA & D. M. YOUNG, JR., *Alternating-Direction-Implicit Methods*, Advances in Computers, Vol. 3, Academic Press, New York, 1962.
3. D. W. PEACEMAN & H. H. RACHFORD, JR., "The numerical solution of parabolic and elliptic differential equations," *J. Soc. Indust. Appl. Math.*, v. 3, 1955, pp. 28–41. MR **17**, 196.
4. C. PEARCY, "On the convergence of alternating direction procedures," *Numer. Math.*, v. 4, 1962, pp. 172–176. MR **26** #3206.
5. R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N. J., 1962. MR **28** #1725.
6. E. L. WACHSPRESS, *Iterative Solution of Elliptic Systems and Applications to the Neutron Diffusion Equations of Reactor Physics*, Prentice-Hall, Englewood Cliffs, N. J., 1966.