ON ESTIMATING THE "RATES OF CONVERGENCE" OF ITERATIVE METHODS FOR ELLIPTIC DIFFERENCE EQUATIONS (1)

BY
SEYMOUR V. PARTER

1. **Introduction.** The numerical solution of elliptic difference equations is still a very important problem. Indeed, the last several years have seen the development of many new iterative methods for these problems (see [3], [4], [5], [6]). In order to compare these methods it is necessary to estimate the rates of convergence.

In [9] we obtained such estimates for certain multi-line iterative methods. However, second thoughts on that work have led us to the following conclusions:

(a) While the results of [9] are all correct, there are some flaws in some of the proofs given there. These are easily repaired by more careful arguments along the same general lines as the arguments given there.

(b) The results of [9] are but a special case of some very general theorems which apply to many more important problems.

In this work we give a very general (and correct) analysis of certain aspects of this general estimation problem. This general approach enables us to discuss Neumann problems as well as Dirichlet problems.

Most of the concepts and arguments used in this work have appeared elsewhere in the vast literature on elliptic differential operators, finite-difference approximations to such operators, and iterative procedures for solving elliptic difference equations. However, there is one very important new concept. Namely, given a sequence of matrices $\{A_n\}$ which arise in these problems we introduce the concept of a sequence of "splittings" of A_n which satisfy property B. This concept is developed in §4. As we will see in §§5, 6 and 7, many of the natural splittings of A_n do satisfy property B.

In §2 we collect some basic facts from the theory of Hilbert space and the Hilbert space approach to elliptic differential operators. §3 is devoted to the formulation of a very general approach to discrete boundary-value problems.

In §4 we turn to iterative methods and the fundamental theorems on the asymptotic behavior of certain eigenvalues associated with these methods. These theorems give formulae of the form

---

Received by the editors February 27, 1963 and, in revised form, August 15, 1963.

(1) This research was partially supported by the National Science Foundation under NSF Contract NSF G-18837 and partially supported by the Office of Naval Research under Contract Nonr 225(37).

$$\rho \approx 1 - \Lambda\,\Delta x\,\Delta y + o(\Delta x\,\Delta y), \tag{1}$$

where Λ is the minimal eigenvalue of a variational problem associated with the underlying elliptic operator.

§5 is devoted to a brief discussion of the earlier results of [9] and some related problems.

In §6 we turn to a Neumann problem and give estimates for the rates of convergence of certain iterative methods.

In §7 we consider the cyclically reduced equations studied by Hageman [5]. We obtain a general asymptotic formula for the rates of convergence of the multi-line iterative methods applied to these equations. As a special case we verify the conjecture of Hageman that

$$\rho({}^{R}J^{lL}) \approx 1 - \frac{4}{3}\, l h^2, \qquad l \geqq 2 \tag{1.1}$$

for the case of the Laplace operator in a square of side π when Dirichlet data is prescribed.

One interesting aspect of the analysis in the case of these cyclically reduced equations is that we are not given the underlying differential operator in advance. Rather, one must show that these equations correspond to the equations which would arise in a particular finite-difference approach to a problem concerning an elliptic differential operator.

While we have aimed at a clear, essentially independent, exposition, we do assume that the reader is somewhat familiar with the general theory of iterative methods for elliptic difference equations. Some of this background material is contained in [9] and the references found there. A rather complete discussion of these topics is available in Forsythe and Wasow [2], Varga [11] and Young [12].

2. **Preliminaries.** Let G be a bounded domain in the (x,y)-plane bounded by a finite number of "smooth" Jordan curves. Let L_2 denote the space of real-valued measurable functions f defined on G which satisfy

$$\|f\|^2 = \iint_G |f|^2\, dxdy < \infty. \tag{2.1}$$

The "inner product" of two functions $f, g \in L_2$ is defined by

$$\langle f,g\rangle = \iint_G fg\, dxdy. \tag{2.1a}$$

A sequence of functions $\{f_n\} \in L_2$ is said to converge "strongly" or converge in norm if there is a function $f \in L_2$ such that

(2.2) $$\text{Lim } \|f_n - f\| = 0.$$

A sequence of functions $\{f_n\} \in L_2$ is said to converge "weakly" to a function $f \in L_2$ if

(2.2a) $$\text{Lim } \langle f_n, g\rangle = \langle f, g\rangle$$

for every $g \in L_2$.

The following lemma collects some of the basic facts about convergence in L_2 which we require.

LEMMA 1. *The space L_2 is complete, i.e., if $\{f_n\} \in L_2$ and $\|f_n - f_m\| \to 0$ as n and $m \to \infty$, then there is a function $f \in L_2$ such that $\|f_n - f\| \to 0$ as $n \to \infty$.*

If a sequence $\{f_n\} \in L_2$ converges strongly, then it converges weakly.

If the sequence $\{f_n\} \in L_2$ satisfies $\|f_n\| \leq k$, for some constant k, there is a subsequence $f_{n'}$ which converges weakly to a function $f \in L_2$. Moreover, if $\|f_n\| \to \|f\|$ as $n \to \infty$, then the sequence $\{f_{n'}\}$ also converges strongly to f.

Proof. See [10].

A function $f \in L_2$ is said to have a strong derivative g in the ξ-direction $(\partial f/\partial \xi)$ if there is a sequence of functions $f_m \in C_1(G)$ such that

(2.4) $$\|f_n - f\| + \left\| \frac{\partial f_n}{\partial \xi} - g \right\| \to 0.$$

LEMMA 2. *Let $\{f_n\} \in L_2$ be a sequence of functions having strong first derivatives and satisfying*

(2.5) $$D[f_n] = \int\int_G \left[\left(\frac{\partial f_n}{\partial x}\right)^2 + \left(\frac{\partial f_n}{\partial y}\right)^2 \right] dxdy \leq k, \qquad \|f_n\| \leq k$$

for some constant k. Then, there is a subsequence $\{f_{n'}\}$ and a function $f \in L_2$ such that

(2.6) $$\|f_{n'} - f\| \to 0.$$

Moreover, the function f has strong first derivatives and

(2.6a) $$D[f] = \int\int_G \left[\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 \right] dxdy \leq k.$$

Proof. This is the Rellich selection theorem. See [1].

Consider the self-adjoint, uniformly elliptic operator

(2.7) $$M \equiv \frac{\partial}{\partial x} a \frac{\partial}{\partial x} + \frac{\partial}{\partial x} b \frac{\partial}{\partial y} + \frac{\partial}{\partial y} b \frac{\partial}{\partial x} + \frac{\partial}{\partial y} c \frac{\partial}{\partial y},$$

where $a(x,y)$, $b(x,y)$, and $c(x,y) \in C_2(\bar{G})$, i.e., are twice(2) continuously differentiable in $\bar{G}$, the closure of G, and satisfy

$$a \geqq a_0 > 0, \quad ac - b^2 \geqq a_0 > 0, \tag{2.7a}$$

for some positive constant a_0.

A function $u \in L_2$ is called a "strong" solution of $Mu = f$ $(f \in L_2)$ if u has strong second derivatives and $Mu = f$ a.e.

DEFINITION. The space $\mathscr{T}$ of "test" functions is the set of all functions $\phi \in C_\infty(G)$ having compact support in G.

DEFINITION. A function $u \in L_2$ is called a "weak" solution of $Mu = f$, $(f \in L_2)$ if

$$\langle M\phi, u\rangle = \langle \phi, f\rangle \tag{2.8}$$

for every test function $\phi \in \mathscr{T}$.

DEFINITION. The space $H^0 \subset L_2$ of functions which "vanish at the boundary of G" is the set of all functions $u \in L_2$ which are the strong limits of functions $u_n \in L_2$ which are zero a.e. in some neighborhood of the boundary ∂G of G. The space H_k^0 is the subspace of H^0 consisting of functions having strong derivatives of order k.

REMARK. It is clear that H^0 is a closed subspace of L_2.

LEMMA 3. *If $u \in H^0$ is a weak solution of $Mu = f (f \in L_2)$, then u is a strong solution of $Mu = f$ and u is a continuous function on $\bar{G}$.*

Proof. While this result is well known to those who work in the field of partial differential equations, there is no ready reference in which these facts can be found in this precise form. Hence we sketch a proof.

If $f(x,y) \in C_1(\bar{G})$ then there is a solution $u \in C_2(G)$ and $Mu = f$, $u = 0$ on ∂G. Moreover, one may easily establish the estimate

$$\left\|\frac{\partial u}{\partial x}\right\| + \left\|\frac{\partial u}{\partial y}\right\| \leqq k\|f\|, \tag{*}$$

where the constant k is independent of f. Thus, we have

$$a\frac{\partial^2 u}{\partial x^2} + 2b\frac{\partial^2 u}{\partial x \partial y} + c\frac{\partial^2 u}{\partial y^2} = g,$$

where $\|g\| \leqq k'\|f\|$, for some constant k' depending only on the functions a, b, c. On applying the argument of [8, p. 295] we obtain the estimate

(2) It is not essential that we have such smooth coefficients. However, it is convenient to assume sufficient continuity to avoid the necessity of detailed arguments about fine points.

$$(**) \qquad \left\| \frac{\partial^2 u}{\partial x^2} \right\| + \left\| \frac{\partial^2 u}{\partial x \partial y} \right\| + \left\| \frac{\partial^2 u}{\partial y^2} \right\| \leqq k'' \|f\| .$$

Now, let $f^{(n)}$ be a sequence of functions in $C_1(\overline{G})$ converging strongly to f. Let $u^{(n)}$ be the solution of $Mu^{(n)} = f^{(n)}$, $u^{(n)} \in H^0$. Using (*) and (**) it is an easy matter to obtain the strong convergence of a subsequence $u^{(n')}$ to a function $u \in H_2^0$ which is a strong solution of our problem. Moreover, from the Sobolev lemma (see [7, p. 655]) u is continuous, and in fact is a uniform limit of the subsequence $u^{(n')}$. Moreover, a strong solution is certainly a weak solution. However, in this case, the weak solution is unique (see [7, Theorem 1′, p. 659]).

COROLLARY. *If $u \in L_2$ is a weak solution of $Mu = f$, $f \in L_2$, then u is a strong solution of $Mu = f$ and u is a continuous function on G.*

Proof. Consider the function $u_1 = u \cdot \zeta$, where $\zeta \in \mathscr{T}$ and $\zeta \equiv 1$ on a compact subdomain K. Then $u_1 \in H^0$ and u_1 is a weak solution of $Mu_1 = f_1 \in L_2$. Thus $u_1 \in H_2^0$ and is continuous on $\overline{G}$. However, $u_1 = u$ on K. Therefore u is continuous on every compact subset K'.

We now consider eigenvalue problems associated with the operator M. Let

$$B[u] \equiv \iint_G \left[a\left(\frac{\partial u}{\partial x}\right)^2 + 2b \frac{\partial u}{\partial x}\frac{\partial u}{\partial y} + c\left(\frac{\partial u}{\partial y}\right)^2 \right] dxdy. \tag{2.9}$$

Let $Q = Q(x,y) \in C_2(\overline{G})$ be a strictly positive function in $\overline{G}$, i.e., there is a $q_0 > 0$ such that

$$Q \geqq q_0 > 0 \quad \text{in } \overline{G}. \tag{2.10}$$

We consider two eigenvalue problems.

Dirichlet boundary conditions.

$$\begin{aligned} Mu + \Lambda Q u &= 0 \quad \text{in } G, \\ u &= 0 \quad \text{on } \partial G, \\ u &\not\equiv 0. \end{aligned} \tag{2.11}$$

Neumann boundary conditions.

$$\begin{aligned} Mu + \Lambda Q u &= 0 \quad \text{in } G, \\ L^{(1)} u &= 0 \quad \text{on } \partial G, \\ u &\not\equiv 0, \end{aligned} \tag{2.12}$$

where

$$L^{(1)} u = \xi_1(a u_x + b u_y) + \xi_2(c u_y + b u_x) \tag{2.12a}$$

and (ξ_1, ξ_2) is the normal vector on ∂G.

LEMMA 4. *There are solutions of* (2.11) *and* (2.12). *These solutions u are in* $C_4(G)$. *If Λ is an eigenvalue of* (2.11), *then* $\Lambda > 0$. *If Λ is an eigenvalue of* (2.12), *then* $\Lambda \geqq 0$. *Moreover, if* $\{\Lambda_k\}$ *are the eigenvalues of* (2.11) *or* (2.12) *with associated eigenfunctions* $\{\phi_k\}$, *then the set* $\{\phi_k\}$ *is complete in* L_2.

The minimal eigenvalue of (2.11) *is characterized by*

$$\Lambda = \operatorname{Min} B[u]/\langle u, Qu\rangle, \tag{2.13}$$

where the minimum is taken over all functions $u \in H_1^0$.

The function $u_0 \equiv 1$ *is the only eigenfunction associated with the value* $\Lambda = 0$. *The next eigenvalue is characterized by*

$$\Lambda = \operatorname{Min} B[u]/\langle u, Qu\rangle, \tag{2.14}$$

where now the minimum is taken over all functions $u \in L_2$ *having strong derivatives and satisfying*

$$\langle u_0, Qu\rangle = 0. \tag{2.14a}$$

Proof. See [1, Volume II, Chapter 7]. The results for the Dirichlet boundary conditions are also discussed in [13, p. 217].

REMARK. For functions $u \in H_1^0$ and the eigenfunction ϕ_k of (2.12) one has

$$B[u] = -\langle Mu, u\rangle.$$

REMARK. The eigenfunctions $\phi_k(x,y)$ of (2.11) or (2.12) may be extended as continuous functions, in fact continuously differentiable functions, defined on an open set $\Omega \supset \bar{G}$.

REMARK. Let ϕ_k be an eigenfunction of (2.11). Then there is a sequence $\phi^{(m)} \in C_2(\Omega)$ so that $\phi^{(m)}$ vanishes outside a compact set $\Sigma_m \subset G$ and

$$\|\phi^{(m)} - \phi_k\| + |B[\phi_k] - B[\phi^{(m)}]| \to 0 \quad \text{as} \quad m \to \infty.$$

This follows from the fact that these functions are dense in H_1^0.

3. **Discrete boundary-value problems.** Let the (x,y)-plane be divided into rectangles by the lines $x = m\Delta x, y = j\Delta y$. The points $(x_m, y_j) = (m\Delta x, j\Delta y)$ *together with the points in which ∂G intersects these lines* are called general mesh points. A mesh region $\bar{G}(h)$ is obtained by selecting certain of these general mesh points subject to the conditions

(C.1) *There is a constant* $R > 0$ *such that*

$$\operatorname*{Min}_{Q \in \bar{G}} |P - Q| < R(\Delta x + \Delta y)$$

for all $P \in \bar{G}(h)$.

(C.2) *All of the general mesh points* $P \in G$ *which satisfy*

$$\text{Min}\,|P - Q| > R(\Delta x + \Delta y), \qquad Q \in \partial G,$$

belong to $\bar{G}(h)$.

Certain of the points of $\bar{G}(h)$ are called "interior" points while the remaining points are called "boundary" points. The set of interior points is denoted by $G(h)$ and the set of boundary points is denoted by $\partial G(h)$. For consistency, we require

(C.3) $P \in \partial G(h)$ *implies that*

$$\text{Min}\,|P - Q| < R(\Delta x + \Delta y), \qquad Q \in \partial G.$$

The mesh region $G(h)$ may be "triangulated" by drawing certain diagonals. After this has been done (in some fashion) one may sensibly speak about the space $\mathscr{A}$, depending on the triangulation, of piecewise linear, real-valued functions defined on $G(h)$. Or, more precisely, defined on $\mathscr{G}$, the union of all the triangles of $\bar{G}(h)$. The functions of $\mathscr{A}$ are defined by their values at the mesh points of $\bar{G}(h)$ *and the requirement that they be linear over each triangle*.

For any two functions g, $h \in \mathscr{A}$ we define the "inner product"

$$[g, h] \equiv \Delta x\,\Delta y \sum_{P \in \bar{G}(h)} g(P)\,h(P). \tag{3.1}$$

DEFINITION. A sequence $\{\bar{G}_n(h)\} = \{\bar{G}_n(h;\ \Delta x_n, \Delta y_n)\}$ of mesh regions is said to converge to G if

(a) $\Delta x_n / \Delta y_n = r$, a constant for all n.

(b) $\Delta x_n \to 0$ as $n \to \infty$.

(c) The points of $\bigcup G_n(h)$ are dense in G and nowhere dense in the complement of $\bar{G}$.

(d) The points of $\bigcup \partial G_n(h)$ are dense in ∂G and nowhere dense in its complement.

(e) There are two positive constants C_1, C_2, which are independent of n, such that

$$C_1\langle f, f\rangle \leqq [f, f]_n \leqq C_2\langle f, f\rangle, \qquad f \in \mathscr{A}_n.$$

For any function u defined and continuous on an open set $\Omega \supset \bar{G}$ we obtain a function $\hat{u} \in \mathscr{A}_n$ by setting

$$\hat{u}(P) = u(P), \qquad P \in \bar{G}_n(h) \tag{3.2}$$

provided n is large enough so that $\bar{G}_n(h) \subset \Omega$. And, for any function u defined and continuous on G we obtain a function $\hat{u}_1 \in \mathscr{A}_n$ by setting

$$\begin{cases} \hat{u}_1(P) = u(P), & \text{if } P \in \bar{G}_n(h) \cap G, \\ \hat{u}_1(P) = 0, & \text{otherwise.} \end{cases} \tag{3.2a}$$

DEFINITION. A linear operator M_n mapping $\mathscr{A}_n$ into $\mathscr{A}_n$ is called a *con-*

sistent approximation to M over $G_n(h)$ if

$$M_n[\hat{u}](P) = M[u](P) + o(1) \quad \text{as } n \to \infty \tag{3.3}$$

for all $P \in G_n(h) \cap G$ and all $u \in C_2(\Omega)$.

DEFINITION. A linear operator L_n^0 mapping $\mathscr{A}_n$ into $\mathscr{A}_n$ is called a *consistent approximation to the identity on* $\partial G_n(h)$ if

$$L_n^0[\hat{u}](P) = u(P) + o(1) \quad \text{as } n \to \infty \tag{3.4}$$

for all $P \in \partial G_n(h)$ and all $u \in C(\Omega)$.

DEFINITION. A linear operator L_n^1 mapping $\mathscr{A}_n$ into $\mathscr{A}_n$ is called a *consistent approximation to* $L^{(1)}$ (see (2.12a) of §2) *on* $\partial G_n(h)$ if

$$L_n^{(1)}[\hat{u}](P) = L^{(1)}u(P) + o(1) \quad \text{as } n \to \infty \tag{3.5}$$

for all $P \in \partial G_n(h)$ and all $u \in C_1(\Omega)$.

Using these consistent approximations one may consider the discrete boundary-value problems

$$\begin{aligned} M_n[v] &= \hat{f} \quad \text{in } G_n(h) \\ L_n^{(k)}[v] &= \hat{\phi} \quad \text{on } \partial G_n(h), \qquad k = 0, 1. \end{aligned} \tag{3.6}$$

DEFINITION. The subspaces $\mathscr{A}_n^0$, $\mathscr{A}_n^{(1)}$ of $\mathscr{A}_n$ are defined by

$$\mathscr{A}_n^0 \equiv [v \in \mathscr{A}_n;\ L_n^0 v = 0 \text{ on } \partial G_n(h)],$$

$$\mathscr{A}_n^{(1)} \equiv [v \in \mathscr{A}_n;\ L_n^1 v = 0 \text{ on } \partial G_n(h)].$$

We let $P_n^{(k)}$, $k = 0, 1$ denote the projection into $\mathscr{A}_n^{(k)}$, and define

$$M_n^{(k)} \equiv P_n^{(k)} M_n P_n^{(k)}, \qquad k = 0, 1. \tag{*}$$

DEFINITION. The operator M_n^0 mapping $\mathscr{A}_n^0$ into $\mathscr{A}_n^0$ is a consistent approximation to M^0, the operator M restricted to H^0, if for every $u \in C_2(\Omega)$ which vanishes outside a compact set $\Sigma \subset G$ whose distance to ∂G is greater than $2R(\Delta x + \Delta y)$, we have

$$\hat{u} \in \mathscr{A}_n^0 \tag{3.7}$$

and

$$M_n^0 \hat{u} = (Mu)\hat{} + o(1). \tag{3.7a}$$

DEFINITION. The operator M_n^1 is a consistent approximation to M^1, the operator M restricted to those functions u satisfying $L^1 u = 0$ on ∂G, if for every function $u \in C_2(\Omega)$ satisfying $L^1 u = 0$ on ∂G one may construct a function $\hat{u}_2 \in \mathscr{A}_n^1$ which satisfies

$$M_n^1 \hat{u}_2 = (M^1 u)\hat{} + o(1) \tag{3.8}$$

for all points $P \in G_n(h) \cap G$ whose distance to ∂G is greater than

$$2R(\Delta x + \Delta y),$$

and

$$M_n^1 \hat{u}_2 = (M^1 u)\hat{} + o[(\Delta x + \Delta y)^{-1}] \tag{3.8a}$$

for those points $P \in G_n(h)$ whose distance from ∂G is less than $3R(\Delta x + \Delta y)$. Moreover, if u vanishes outside a compact set $\Sigma \subset G$, then we require that $\hat{u}_2 = \hat{u}$ and (3.8) holds everywhere when n is large enough.

In most cases the operators $L_n^{(k)}$ have an important property, namely: given a function $v \in \mathscr{A}_n$ one may construct a function $\tilde{v} \in \mathscr{A}_n$ which satisfies (for an arbitrary ϕ)

$$L_n^{(k)} \tilde{v}(P) = \phi(P), \qquad P \in \partial G_n(h)$$

$$\tilde{v}(P) = v(P), \qquad P \in G_n(h).$$

In such a case, the boundary-value problems (3.6) may be reduced to problems of the form

$$M_n^{(k)}[w] = g \quad \text{in } G_n(h). \tag{3.9}$$

4. **Iterative methods and general estimates.** After ordering the points of $G_n(h)$ the equation (3.9) becomes a system of linear equations of the form

$$A_0 X = Y_0, \tag{4.1}$$

where A_0 is the matrix representation of the operator $M_n^{(k)}$. Since the elements of A_0 are generally unbounded as $\Delta x_n \to 0$ one usually sets

$$A_n = -\Delta x_n \Delta y_n A_0$$

$$Y = -\Delta x_n \Delta y_n Y_0$$

to obtain a system

$$A_n X = Y, \tag{4.1a}$$

where A_n is the matrix representation of $-\Delta x_n \Delta y_n M_n^{(k)}$.

One obtains an iteration procedure by "splitting" A_n in the form

$$A_n = T_n - N_n, \tag{4.2}$$

where T_n is nonsingular and T_n^{-1} is relatively easy to obtain. After choosing an initial guess X^0 we obtain $X^{\nu+1}$ from relation

$$T_n X^{\nu+1} = N_n X^\nu + Y. \tag{4.2a}$$

LEMMA 5. *If A_n is nonsingular, then the vectors X^ν converge to the unique solution X of* (4.1a) *for all initial vectors X^0 if and only if*

$$\rho_n \equiv \operatorname{Max} |\lambda| < 1, \tag{4.3}$$

where λ is an eigenvalue of

(4.3a) $$\det \{\lambda T_n - N_n\} = 0.$$

If A_n is singular and there is a solution X of (4.1a), *then this solution X is not unique. Moreover, the vectors X^ν converge to a solution X of* (4.1a) *if and only if*

(4.4) $$\text{Max} \, |\lambda| \leqq 1.$$

And for all eigenvalues λ of absolute value one, we have

(4.4a) $$\lambda = 1$$

and the eigenspace of $T_n^{-1}N_n$ associated with the eigenvalue $\lambda = 1$ is spanned by the associated eigenvectors. That is, the elementary divisors of $T_n^{-1}N_n$ associated with the eigenvalue $\lambda = 1$ are simple.

Proof. See [2].

In the case when A_n is nonsingular we define the rate of convergence of the iterative method as

$$-\log \rho_n.$$

In the case when A_n is singular, and the iteration procedure is convergent, we define

(4.5) $$\mu_n = \text{Max} \, |\lambda|, \qquad |\lambda| < 1$$

and the rate of convergence is

$$-\log \mu_n.$$

Let $\mathscr{P}_n$ and $\mathscr{N}_n$ be the operators mapping $\mathscr{A}_n^k$ into $\mathscr{A}_n^k$ associated with the matrices T_n and N_n, respectively.

Definition. Let $\{\overline{G}_n(h)\}$ be a sequence of mesh regions converging to G. Let $M_n^{(k)}$ be a consistent approximation to $M^{(k)}$ over $G_n(h)$. The sequence of splittings $A_n = T_n - N_n$ is said to satisfy *property* B if

(a) The operators $\mathscr{N}_n$ are self-adjoint and uniformly bounded, i.e., there is a constant C_2 such that

(4.6) $$[\mathscr{N}_n w, \mathscr{N}_n w]_n \leqq C_2 [w, w]_n, \qquad w \in \mathscr{A}_n^{(k)}.$$

(b) There is a $Q(x,y) \in C_2(\Omega)$ which is strictly positive, i.e., $Q(x,y) \geqq q_0 > 0$. When $k = 0$

(4.7.0) $$\text{Lim} \, [\mathscr{N}_n \hat{\phi}, \hat{w}_1]_n = \langle \phi, Qw \rangle$$

whenever $\phi \in C(\Omega)$ and vanishes outside a compact set $\Sigma \subset G$. When $k = 1$

(4.7.1) $$\text{Lim} \, [\mathscr{N}_n \hat{\phi}_2, \hat{w}_1]_n = \langle \phi, Qw \rangle$$

whenever $\phi \in C_2(\Omega)$ and satisfies $L^1 \phi = 0$ on ∂G, and $w \in C(G) \cap L_2$.

LEMMA 6. *Let the splitting $A_n = T_n - N_n$ satisfy property* B. *Let $v^{(n)} \in \mathscr{A}_n^{(k)}$ converge strongly to a function $v \in C(G) \cap L_2$. If $k = 0$, let $\phi \in C(\Omega)$ and vanish outside a compact set $\Sigma \subset G$. If $k = 1$, let $\phi \in C_2(\Omega)$ and satisfy $L^1 \phi = 0$ on ∂G. Then, if $k = 0$*

$$\text{Lim } [\hat{\phi}, \mathscr{N}_n v^{(n)}]_n = \langle \phi, Qv \rangle, \tag{4.8a}$$

and if $k = 1$

$$\text{Lim } [\hat{\phi}_2, \mathscr{N}_n v^{(n)}]_n = \langle \phi, Qv \rangle. \tag{4.8b}$$

Proof. Let $\hat{\psi}$ denote $\hat{\phi}$ or $\hat{\phi}_2$. Then, for n large enough $\hat{\psi} \in \mathscr{A}_n^{(k)}$. Since $\mathscr{N}_n$ is self-adjoint

$$[\hat{\psi}, \mathscr{N}_n v^{(n)}]_n = [\mathscr{N}_n \hat{\psi}, v^{(n)}]_n.$$

But

$$[\mathscr{N}_n \hat{\psi}, v^{(n)}]_n = [\mathscr{N}_n \hat{\psi}, \hat{v}_1]_n + [\mathscr{N}_n \hat{\psi}, v^{(n)} - \hat{v}_1]_n,$$

and the lemma follows from the fact that the mesh regions $\overline{G}_n(h)$ converge to G (see (e) of §3).

LEMMA 6b. *Let the splittings $A_n = T_n - N_n$ satisfy property* B. *Let $v^{(n)} \in \mathscr{A}_n^{(k)}$ converge strongly to a function $v \in L_2$. Let $\phi \in C(\Omega)$ vanish outside a compact set $\Sigma \subset G$. Then*

$$\text{Lim} [\hat{\phi}, \mathscr{N}_n v^{(n)}] = \langle \phi, Qv \rangle.$$

Proof. If $v \notin C(G)$ we "smooth" v to obtain $V \in C(G)$. Then can be done via the Friedrichs mollifiers, see [7]. Then

$$[\hat{\phi}, \mathscr{N}_n v^{(n)}] = [\hat{\phi}, \mathscr{N}_n \hat{V}] + [\hat{\phi}, \mathscr{N}_n (v^{(n)} - \hat{V})].$$

By hypothesis,

$$[\hat{\phi}, \mathscr{N}_n \hat{V}] \to \langle \hat{\phi}, QV \rangle.$$

And

$$|[\hat{\phi}, \mathscr{N}_n (v^{(n)} - \hat{V})]| \leqq C_2 \|\phi\| \cdot \|\mathscr{N}_n\| \cdot \|v^{(n)} - \hat{V}\|.$$

Letting $V \to v$, we obtain the desired results.

LEMMA 7. *Let the mesh regions $\{\overline{G}_n(h)\}$ converge to G. Let $M_n^{(k)}$ be a consistent approximation to $M^{(k)}$ over $G_n(h)$. Let $M_n^{(k)}$ be self-adjoint. Let $v^{(n)} \in \mathscr{A}_n^{(k)}$ converge strongly to a function $v \in C(G) \cap L_2$. If $k = 0$, let $\phi \in C_2(\Omega)$ and vanish outside a compact set $\Sigma \subset G$. If $k = 1$, let $\phi \in C_2(\Omega)$ and satisfy $L^{(1)} \phi = 0$ on ∂G. Let $\hat{\psi} = \hat{\phi}$ or $\hat{\phi}_2$, depending on whether $k = 0$ or 1. Then*

$$\text{Lim } [\hat{\psi}, M_n^{(k)} v^{(n)}]_n = \langle M\phi, v \rangle. \tag{4.9}$$

Proof. In either case $\hat{\psi} \in \mathscr{A}_n^{(k)}$ for n sufficiently large and

$$M_n^{(k)}\hat{\psi} = (M\phi)\hat{\ } + o(1)$$

for $P \in G_n(h) \cap G$ and the distance from P to ∂G is greater than $2R(\Delta x + \Delta y)$, while

$$M_n^{(k)}\hat{\psi} = (M\phi)\hat{\ } + o[(\Delta x + \Delta y)^{-1}]$$

for $P \in G_n(h)$ and the distance from P to ∂G is less than $3R(\Delta x + \Delta y)$.

The number of points P satisfying this latter condition is $O[(\Delta x + \Delta y)^{-1}]$. Therefore

$$\operatorname{Lim}\ [M_n^{(k)}\hat{\psi}, \hat{v}_1]_n = \langle M\phi, v\rangle.$$

But

$$[\hat{\psi}, M_n^{(k)} v^{(n)}]_n = [M_n^{(k)}\hat{\psi}, v^{(n)}]_n$$

and

$$[M_n^{(k)}\hat{\psi}, v^{(n)}]_n = [M_n^{(k)}\hat{\psi}, \hat{v}_1]_n + [M_n^{(k)}\hat{\psi}, (v^{(n)} - \hat{v}_1)]_n.$$

The lemma now follows from the fact that

$$[M_n^{(k)}\hat{\psi}, M_n^{(k)}\hat{\psi}]_n \leqq K$$

for some constant K and the fact that the mesh regions converge to G.

In many cases, particularly when the difference equations have been derived from a variational principle, the operators $M_n^{(k)}$ satisfy a "coerciveness" condition. Namely, there is a constant C_3 such that

$$|[M_n^{(k)} v, v]_n| \leqq K \tag{4.10a}$$

implies that

$$D[v] \leqq C_3 K, \qquad v \in \mathscr{A}_n^{(k)}, \tag{4.10b}$$

where $D[v]$ is the Dirichlet integral of v defined in §2.

LEMMA 7b. *Let the mesh regions $\{\bar{G}_n(h)\}$ converge to G. Let $M_n^{(k)}$ be a consistent approximation to $M^{(k)}$ over $G_n(h)$. Let $M_n^{(k)}$ be self-adjoint and satisfy a coerciveness condition.*

Let $v^{(n)} \in \mathscr{A}_n^{(k)}$ converge strongly to a function $v \in L_2$ while $\|M_n^{(k)} v^{(n)}\| \leqq B$ for some constant B. Let $\phi \in C_4(G)$ and vanish outside a compact set $\Sigma \subset G$. Then

$$\operatorname{Lim}\ [\hat{\phi}, M_n^{(k)} v^{(n)}] = \langle M\phi, v\rangle.$$

Proof. Let us observe that $w \in C_1(\bar{G})$ implies that

$$[M_n^{(k)}\hat{\phi}, w] = \langle M\phi, w\rangle + \epsilon_n,$$

where

$$\|\epsilon_n\|^2 \leqq D[(M^{(k)}\phi)w] \cdot o(1).$$

Hence this estimate can easily be extended to the case where $D[w] < \infty$. However, the coerciveness condition guarantees that there is a constant B_1 so that

$$D[v^{(n)}] \leqq B_1.$$

Since $M_n^{(k)}$ is self-adjoint we have

$$[\hat{\phi}, M_n^{(k)} v^{(n)}] = [M_n^{(k)} \phi, v^{(n)}]$$

and the lemma follows from the remarks above.

In the case $k = 0$, the matrix A_n is usually nonsingular. We consider this case first.

THEOREM 1. *Let the sequence $\{\overline{G}_n(h)\}$ converge to G. Let M_n^0 be a consistent approximation to M^0 over $G_n(h)$. Let A_n be the matrix representation of $-\Delta x_n \Delta y_n M_n^0$ and let $A_n = T_n - N_n$ be a splitting of A_n. We assume*

(i) *The matrices A_n and T_n are positive definite, hence M_n^0, $\mathscr{P}_n$ and $\mathscr{N}_n$ are self-adjoint.*

(ii) *The "splitting" $A_n = T_n - N_n$ satisfies property* B. *Then*

$$\rho_n \geqq 1 - \Lambda \Delta x_n \Delta y_n + o(\Delta x_n \Delta y_n), \tag{4.11}$$

where Λ is the smallest eigenvalue of

$$\begin{aligned} Mu + \Lambda Q u &= 0 \quad \text{in } G, \\ u &= 0 \quad \text{on } \partial G, \\ u &\not\equiv 0. \end{aligned} \tag{4.11a}$$

Proof. Let ϕ_1 be the eigenfunction of (4.11a) associated with the minimal eigenvalue Λ. Given any $\epsilon > 0$ one may find a function $\phi \in C_2(\Omega)$ having compact support $\Sigma \subset G$ so that

$$\|\phi - \phi_1\| + |B[\phi] - B[\phi_1]| < \epsilon. \tag{4.12}$$

Using Lemmas 6 and 7 we find

$$\begin{aligned} [M_n^0 \hat{\phi}, \hat{\phi}]_n &= \langle M\phi, \phi\rangle + o(1), \\ [\mathscr{N}_n \hat{\phi}, \hat{\phi}]_n &= \langle \phi, Q\phi\rangle + o(1). \end{aligned}$$

Using (4.12), and the fact that ϕ_1 is an eigenfunction, we have

$$\begin{aligned} [M_n^0 \hat{\phi}, \hat{\phi}]_n &= -\Lambda\langle \phi_1, Q\phi_1\rangle + o(1), \\ [\mathscr{N}_n \hat{\phi}, \hat{\phi}]_n &= \langle \phi_1, Q\phi_1\rangle + o(1). \end{aligned} \tag{4.13}$$

In this case,

$$\rho_n = \operatorname{Max} \frac{|[\mathscr{N}_n w, w]_n|}{[\mathscr{P}_n w, w]_n}, \qquad w \in \mathscr{A}_n^0. \tag{4.14}$$

Since $\mathscr{P}_n = -\Delta x_n \Delta y_n M_n^0 + \mathscr{N}_n$, the lemma follows from inserting (4.13) into (4.14).

Whenever the splitting $A_n = T_n - N_n$ satisfies property **A** (see [12]) and A_n and T_n are positive definite one has

$$\rho_n = \text{Max} \frac{[\mathscr{N}_n w, w]_n}{[\mathscr{P}_n w, w]_n}, \qquad w \in \mathscr{A}_n^0. \tag{4.15}$$

Theorem 2. *Under the same hypothesis as in Theorem* 1, *let*

(iii) *The operators* M_n^0 *satisfy a "coerciveness" condition.*

(iv) *Equation* (4.15) *applies instead of* (4.14). *Then*

$$\rho_n \approx 1 - \Lambda \Delta x_n \Delta y_n + o(\Delta x_n \Delta y_n) \tag{4.16}$$

where, once more, Λ *is the minimal eigenvalue of* (4.11a).

Proof. Under these hypotheses one may easily prove that the iterative method is convergent, i.e., $\rho_n < 1$. Let $v^{(n)} \in \mathscr{A}_n^0$ be the eigenfunction associated with ρ_n. Let

$$\lambda_n = [(1 - \rho_n)/\rho_n \Delta x_n \Delta y_n] > 0. \tag{4.17}$$

The eigenvalue equation may now be rewritten as

$$M_n^0 v^{(n)} = -\lambda_n \mathscr{N}_n v^{(n)}. \tag{4.17a}$$

Using Theorem 1 we find that

$$\text{Lim Sup } \lambda_n \leqq \Lambda, \tag{4.17b}$$

while this theorem asserts that $\text{Lim } \lambda_n$ exists and

$$\text{Lim } \lambda_n = \Lambda. \tag{4.17c}$$

Assume that (4.17c) is not true. Then there is a sequence (n^1) with $n^1 \to \infty$ so that

$$\text{Lim } \lambda_{n^1} = \Lambda_0 < \Lambda. \tag{4.18}$$

We may normalize the eigenfunctions $v^{(n)}$ so that $\|v^{(n)}\| = 1$. Since the mesh regions $\overline{G}_n(h)$ converge to G, there is a constant K such that

$$[v^{(n)}, v^{(n)}]_n \leqq K.$$

Since the $\{\mathscr{N}_n\}$ are uniformly bounded, there is a constant K_1 such that

$$\begin{aligned} |[M_n^0 v^{(n)}, v^{(n)}]_n| &= \lambda_n |[\mathscr{N}_n v^{(n)}, v^{(n)}]| \\ &\leqq K_1. \end{aligned} \tag{4.19}$$

Thus, using the Rellich selection principle and condition (iii) of the hypothesis, we may select a subsequence n'' so that

(4.20a) $$\lambda_{n''} \longrightarrow \Lambda_0$$

and the subsequence $v^{(n'')}$ converges strongly to a function $v \in L_2$. Moreover, the uniformly bounded sequence $\mathcal{N}_{n''} v^{(n'')}$ converges weakly to a function $q \in L_2$. Using Lemmas 6b and 7b we find that $q = Qv$ and v is a weak solution of

$$Mv = \Lambda_0 q.$$

Thus $v \in C(G) \cap L_2$. Let ϕ_k be the kth eigenfunction of (4.11a). Let $\phi^{(m)} \in C_2(\Omega)$ be a sequence of functions such that $\phi^{(m)}$ has compact support $\Sigma_m \subset G$ and

$$\|\phi^{(m)} - \phi_k\| + |B[\phi^{(m)}] - B[\phi_k]| \longrightarrow 0 \quad \text{as } m \longrightarrow 0.$$

From (4.17a) and Lemmas 6 and 7 we have

(4.21) $$\begin{aligned} \operatorname{Lim} [\hat{\phi}^m, M^0_{n''} v^{(n'')}]_{n''} &= -\Lambda_0 \operatorname{Lim} [\hat{\phi}^{(m)}, \mathcal{N}_{n''} v^{(n'')}]_{n''} \\ \langle M\phi^{(m)}, v\rangle &= -\Lambda_0 \langle \phi^{(m)}, Qv\rangle. \end{aligned}$$

On letting $m \longrightarrow \infty$ we find that

$$\Lambda_k \langle \phi_k, Qv\rangle = \Lambda_0 \langle \phi_k, Qv\rangle.$$

Thus, $\Lambda_0 = \Lambda_k$ for some k. Therefore

$$\Lambda_0 \geqq \Lambda = \min_{(k)} \Lambda_k,$$

which proves the theorem.

REMARK. We observe that the proofs of Theorems 1 and 2 do not require that M^0_n arise in the manner described in §3 and culminating in equation (*). Indeed the proofs use *only* the properties (3.7) and (3.7a) of the definition of a consistent approximation to M^0.

Most of the literature concerning iterative methods for elliptic difference equations is restricted to the case $k = 0$. Therefore, before developing the estimates for the case $k = 1$ we digress to discuss some relevant facts.

Suppose that A_n is positive semi-definite and of the form

(4.22) $$A_n = T_n - [L_n + L^*_n] = T_n - N_n,$$

where T_n is block diagonal and positive definite while L_n is lower (block) triangular and L^*_n is its transpose which is upper triangular. This representation leads to three natural iterative procedures for (4.1a).

Block Jacobi method.

(4.23a) $$T_n X^{\nu+1} = [L_n + L^*_n] X^\nu + Y.$$

Block Gauss-Seidel method.

(4.23b) $$(T_n - L_n) X^{\nu+1} = L^*_n X^\nu + Y.$$

Block over-relaxation method. A parameter $\omega > 0$ is chosen, and

$$(T_n - \omega L_n)X^{\nu+1} = [\omega L_n^* + (1-\omega)T_n]X^{\nu} + \omega Y. \tag{4.23c}$$

LEMMA 8. *If the splitting* (4.22) *satisfies block property* A, *then*
(a) *If* λ *is an eigenvalue of*

$$(\lambda T_n - N_n)X = 0 \tag{4.24a}$$

then $-\lambda$ *is also.*
(b) *The eigenvalues* μ *of*

$$[\mu(T_n - L_n) - L_n^*]X = 0 \tag{4.24b}$$

are of the form

$$\mu = \lambda^2,$$

where λ *is an eigenvalue of* (4.24a).
(c) *The eigenvalues* σ *of*

$$\{\sigma(T_n - \omega L_n) - [\omega L_m^* + (1-\omega)T_n]\}X = 0$$

satisfy the equation

$$(\sigma + \omega - 1)^2 = \omega^2\lambda^2\sigma,$$

where λ *is an eigenvalue of* (4.24a).

Proof. See [12].

In particular, if A_n is actually singular and the vector e is the only eigenvector associated with the eigenvalue zero, then ± 1 are both eigenvalues of (4.24a). Moreover, e is the only eigenvector associated with $+1$ and there is a unique eigenvector associated with -1. Thus, the iterative method (4.23a) does not converge while the iterative methods given by (4.23b) and (4.23c) do converge, provided

$$1 \leqq \omega < 2.$$

More important, the quantity μ_n is a function of the quantity $\tilde{t}_n$ defined by

$$\tilde{t}_n = \operatorname{Max} \frac{[\mathcal{N}_n w, w]_n}{[\mathcal{P}_n w, w]_n},$$

where $w \in \mathcal{A}_n^1$ and satisfies

$$[\tilde{e}, \mathcal{P}_n \omega]_n = [\tilde{e}, \mathcal{N}_n w]_n = 0,$$

where $\tilde{e}$ is the function associated with the vector e.

THEOREM 3. *Let the sequence of mesh regions* $\{\overline{G}_n(h)\}$ *converge to* G. *Let* $M_n^{(1)}$ *be a consistent approximation to* M^1 *over* $G_n(h)$. *Let* A_n *be the matrix* *representation of* $-\Delta x_n \Delta y_n M_n^{(1)}$ *and let* $A_n = T_n - N_n$ *be a splitting of* A_n. *We assume*

(i) *The matrix* A_n *is positive semi-definite and the only eigenvector associated with the eigenvalue zero is the vector* $e = \tilde{1}$, *all of whose components are* 1.

(ii) *The matrix* T_n *is positive definite.*

(iii) *The splitting* $A_n = T_n - N_n$ *satisfies property* B. *Let*

$$\tilde{\rho}_n \equiv \operatorname{Max} \frac{|[\mathcal{N}_n w, w]_n|}{[\mathcal{P}_n w, w]_n}, \tag{4.25}$$

where the maximum is taken over all functions $w \in \mathscr{A}_n^1$ *which also satisfy*

$$[\mathcal{P}_n w, \hat{1}]_n = 0. \tag{4.25a}$$

Then

$$\tilde{\rho}_n \geqq 1 - \Lambda \Delta x_n \Delta y_n + o(\Delta x_n \Delta y_n), \tag{4.26}$$

where Λ *is the smallest nonvanishing eigenvalue of*

$$\begin{aligned} Mu + \Lambda Qu &= 0 \quad \text{in } G, \\ L^1 u &= 0 \quad \text{on } \partial G, \\ u &\not\equiv 0. \end{aligned} \tag{4.26a}$$

Proof. Let ϕ be the eigenfunction of (4.26) associated with Λ. In general, $\hat{\phi}_2$ need not satisfy (4.25a). However, since the splittings satisfy property B we have

$$[\hat{1}, \mathcal{N}_n \hat{\phi}_2]_n = o(1).$$

Therefore

$$[\hat{1}, \mathcal{P}_n \hat{\phi}_2]_n / [\hat{1}, \mathcal{P}_n \hat{1}]_n = o(1). \tag{4.27}$$

Let $\psi \in \mathscr{A}_n^1$ be defined by

$$\omega = \hat{\phi}_2 - ([\hat{1}, \mathcal{P}_n \hat{\phi}_2]_n / [\hat{1}, \mathcal{P}_n \hat{1}]_n) \cdot \hat{1}. \tag{4.28}$$

The function ψ does satisfy (4.25a) and we may now repeat the argument of Theorem 1, using ψ as a "test" function.

In a similar way, using such modified test functions, we may modify the proof of Theorem 2 and obtain

THEOREM 4. *Under the same hypotheses as in Theorem* 3, *let*

(iv) *The operators* $M_n^{(1)}$ *satisfy a coerciveness condition.*

(v) $\tilde{\rho}_n = \tilde{t}_n$ *defined above.*

Then

$$\tilde{\rho}_n \approx 1 - \Lambda \Delta x_n \Delta y_n + o(\Delta x_n \Delta y_n),$$

where, as before, Λ is the smallest nonvanishing eigenvalue of (4.26a).

REMARK. A careful study of the proofs of Theorems 1, 2, 3, 4 reveals that it is not really essential that the boundary curves of G be smooth (see §2). Indeed, these theorems are all true if the eigenfunctions ϕ_k of (2.11) and (2.12) $\{$(4.11a) and (4.26a)$\}$ are in $C_2(\overline{G})$. In that case they may easily be extended as functions in $C_2(\Omega)$, where $\Omega \supset \overline{G}$. In particular, Theorems 1, 2, 3, 4 are all valid when G is a rectangle.

We conclude this section with one more remark. For computational purposes it is sometimes convenient to work with the matrix

$$B_n = DA_nD,$$

where D is a positive definite matrix, usually a diagonal matrix. However, this has no effect on the value of ρ. For, let $B_n = T_0 - N_0$ and

$$\rho^1 = \operatorname*{Max}_X \frac{(N_0X, X)}{(T_0X, X)},$$

where $N_0 = DND$ and $T_0 = DTD$. Then

$$\rho^1 = \operatorname*{Max}_Y \frac{(NY, Y)}{(TY, Y)} = \rho,$$

where $Y = DX$.

5. **Multi-line iterative methods and related results.** In this section we return to the problem considered in [9]. The mesh region $\overline{G}(h)$ is obtained by taking all points $(x_m, y_j) = (m\Delta_x, j\Delta_y)$ inside $\overline{G}$. A point $(x_m, y_j) \in \overline{G}(h)$ is called a boundary point if not all of its eight neighbors are in G. All other points are called interior points.

The mesh region $\overline{G}(h)$ is triangulated by drawing all the diagonals with positive (or negative, it doesn't matter) slope.

The operator M is approximated by the finite-difference operator l_h defined by

$$(5.1)\quad l_h[u] = (Au_x)_{\bar{x}} + \frac{1}{2}\{(bu_x)_y + (bu_{\bar{x}})_y + (bu_{\bar{y}})_x + (bu_{\bar{y}})_{\bar{x}}\} + (C^0u_y)_{\bar{y}},$$

where

$$(5.1\text{a})\quad \begin{aligned} A(x,y) &\equiv a\left(x + \frac{1}{2}\Delta x, y\right), \\ C^0(x,y) &\equiv c\left(x, y + \frac{1}{2}\Delta y\right) \end{aligned}$$

while u_x and $u_{\bar{x}}$ denote forward and backward difference quotients respectively, i.e.,

$$V_x = \frac{1}{\Delta x}\{V(x+\Delta x, y) - V(x,y)\},\quad V_{\bar{x}} = \frac{1}{\Delta x}\{V(x,y) - V(x-\Delta x, y)\},$$

(5.1b)

$$V_y = \frac{1}{\Delta y}\{V(x, y+\Delta y) - V(x,y)\},\quad V_{\bar{y}} = \frac{1}{\Delta y}\{V(x,y) - V(x, y-\Delta y)\},$$

for any function $V(x,y)$ defined on $\bar{G}(h)$.

In this case summation by parts gives

$$[l_h[u], u] = -B_0[u], \qquad u \in \mathscr{A}^0,$$

where

$$B_0[u] = \Delta x \Delta y \sum \{A(u_x)^2 + \frac{1}{2} b[u_{\bar{x}} u_{\bar{y}} + u_x u_{\bar{x}} + u_{\bar{y}} u_{\bar{x}} + u_x u_{\bar{y}}] + C^0(u_y)^2\}.$$

Thus, the uniform ellipticity of M shows that (for Δx, Δy small enough) l_h satisfies the "coerciveness" condition over $\mathscr{A}^0$.

If one takes $\Delta x_n / \Delta y_n = r$ and considers a sequence of such mesh regions $\bar{G}_n(h)$ for which $\Delta x_n \to 0$ as $n \to \infty$, one immediately finds that this sequence converges to G.

The operators L_n^0 are taken as the identity operators over $\mathscr{A}_n$.

One may easily verify that the operators M_n^0 defined on $\mathscr{A}_n^0$ by

$$M_n^0[u] = l_h[u] \quad \text{in } G_n(h)$$
$$= 0 \qquad \text{on } \partial G_n(h)$$

are negative definite and are consistent approximations to M^0 over $G_n(h)$.

In [9] we developed the multi-line (k) iterative method associated with this difference approximation to M. The interested reader is referred to the details there. However, we note that in that work we worked with the matrix representation of

(5.2)
$$-\left[\frac{\Delta x_n \Delta y_n}{\Delta x_n^2 + \Delta y_n^2}\right] \Delta x_n \Delta y_n M_n^0.$$

Thus, there is a factor of $r/(r^2+1)$ multiplying all equations in [9].

Since these k-line iterative methods satisfy property A, the dominant eigenvalue of the Jacobi method (which we denote by $\lambda_R(k)$) satisfies condition (iv) of Theorem 2.

Following the development given in [9], we find that

(5.3a)
$$[u, \mathscr{N}_n v]_n = 2\Delta y \Delta x \sum_{(\nu)} I_\nu,$$

where

$$I_\nu = -\frac{1}{2}\sum_{(m)} (b_{m,\nu k+1} + b_{m-1,\nu k})\, v_{m-1,\nu k+1}\, u_{m,\nu k}$$

(5.3b)
$$+\frac{1}{r}\sum_{(m)} C^0_{m,\nu k}\, v_{m,\nu k+1}\, u_{m,\nu k}$$

$$+\frac{1}{2}\sum_{(m)} (b_{m,\nu k+1} + b_{m+1,\nu k})\, v_{m+1,\nu k+1}\, u_{m,\nu k}.$$

In this formula we have used the familiar notation

(5.3c)
$$F(m\Delta x, j\Delta y) = F_{m,j}.$$

If $u \in C(\Omega)$ and vanishes outside a compact set $\Sigma \subset G$ and $v \in C(G)$ we find that

(5.4)
$$\Delta x I_\nu = \frac{1}{r}\int C(\xi, \nu k\,\Delta y)\, u(\xi, \nu k\,\Delta y)\, v(\xi, \nu k\,\Delta y)\, d\xi + o(1),$$

where the integration is carried out over the intersection of the line $y = \nu k\Delta y$ and the set Σ. Since only $1/k$ of the lines enter into the sum in (5.3a) we see that

(5.5)
$$[u, \mathcal{N}_n v]_n = \frac{2}{kr}\int\int C(x,y)\, u(x,y)\, v(x,y)\, dxdy + o(1),$$

i.e., the splittings in the k-line iterative method satisfy property B with

(5.5a)
$$Q(x,y) = \frac{2}{kr} C(x,y).$$

Thus, using Theorem 2, we find, as we found in [9],

(5.6)
$$\lambda_R(k) \approx 1 - \frac{k}{2}\Lambda(\Delta y)^2,$$

where Λ is the minimal eigenvalue of

(5.6a)
$$\begin{aligned} Mu + \Lambda c u &= 0 \quad \text{in } G,\\ u &= 0 \quad \text{on } \partial G,\\ u &\not\equiv 0. \end{aligned}$$

If $b(x,y) \equiv 0$, then the point iterative methods also satisfy property A. If we consider the point Jacobi iterative method we find that condition (iv) of Theorem 2 is satisfied. Moreover, a direct calculation shows that

(5.7a)
$$[u, \mathcal{N}_n v]_n = \Delta x_n \Delta y_n (I_1 + I_2),$$

where

$$
(5.7b) \qquad I_1 = \frac{2}{r} \sum a_{m+1/2,j} u_{m,j} v_{m+1,j}, \qquad I_2 = 2r \sum c_{m,j+1/2} u_{m,j} v_{m,j+1}.
$$

Thus, the splittings associated with the point iterative method satisfy property B with

$$
(5.7c) \qquad Q = 2\left(rc + \frac{1}{r} a\right).
$$

On applying Theorem 2, we find that the dominant eigenvalue for this method is given by

$$
(5.7d) \qquad \rho \approx 1 - \frac{1}{2} \Lambda \Delta x \Delta y,
$$

where Λ is the minimal eigenvalue of

$$
(5.7e) \qquad \begin{aligned} Mu + \Lambda \left(rc + \frac{1}{r} a\right) u &= 0 \text{ in } G, \\ u &= 0 \text{ on } \partial G, \\ u &\not\equiv 0. \end{aligned}
$$

6. **A Neumann problem.** Let G be the rectangle

$$
(6.1) \qquad G \equiv \{(x,y); X_0 < x < X_1, Y_0 < y < Y_1\},
$$

and let $b(x,y) \equiv 0$, i.e.,

$$
(6.1a) \qquad M = \frac{\partial}{\partial x} a \frac{\partial}{\partial x} + \frac{\partial}{\partial y} c \frac{\partial}{\partial y}.
$$

The mesh region $\bar{G}(h)$ is chosen as in §5. We notice that there are four constants m_0, m_1, J_0, and J_1 which describe the points of $\partial G(h)$. That is

$$
(6.2) \qquad \partial G(h) = S_1 + S_2 + S_3 + S_4,
$$

where

$$
\begin{aligned}
S_1 &= \{(x_m, y_j); m = m_0, J_0 < j < J_1\}, \\
S_2 &= \{(x_m, y_j); m_0 < m < m_1, j = J_0\}, \\
S_3 &= \{(x_m, y_j); m = m_1, J_0 < j < J_1\}, \\
S_4 &= \{(x_m, y_j); m_0 < m < m_1, j = J_1\}.
\end{aligned}
$$

Once more the operator M is approximated by

$$
(6.3) \qquad l_h[u] = (Au_x)_{\bar{x}} + (C^0 u_y)_{\bar{y}}.
$$

The operators $L_n^{(1)}$ are defined by

(6.3a) $$L_n^{(1)}[u] = u \quad \text{in } G(h),$$

(6.3b) $$(L_n^{(1)}[u])_{m_0,j} = \frac{a}{\Delta x}(u_{m_0+1,j} - u_{m_0,j}) \text{ on } S_1,$$

(6.3c) $$(L_n^{(1)}[u])_{J_0,m} = \frac{c}{\Delta y}(u_{m,J_0+1} - u_{m,J_0}) \text{ on } S_2,$$

(6.3d) $$(L_n^{(1)}[u])_{m_1,j} = \frac{a}{\Delta x}\{u_{m_1-1,j} - u_{m_1,j}\} \text{ on } S_3,$$

(6.3e) $$(L_n^{(1)}[u])_{J_1,m} = \frac{c}{\Delta y}\{u_{m,J_1-1} - u_{m,J_1}\} \text{ on } S_4.$$

If $u \in C_2(\Omega)$ and $L^{(1)}[u] = 0$ on ∂G, then we define

$$\hat{u}_2(P) = u(P) \quad \text{if } P \in G(h)$$

and for $P \in \partial G$ we define $\hat{u}_2(P)$ by the condition that $L_n^{(1)}[\hat{u}_2](P) = 0$. For example, if P is the point (x_{m_0}, y_j) then

(6.4) $$\hat{u}_2(x_{m_0}, y_j) = u(x_{m_0+1}, y_j).$$

Using this definition of $\hat{u}_2$, we see that $M_n^{(1)}$ mapping $\mathscr{A}_n^1$ into $\mathscr{A}_n^1$ is a consistent approximation to M^1 on $\partial G(h)$. At all points P in $G(h)$ *which do not have one of their four neighbors on* $\partial G(h)$

(6.4a) $$M_n^{(1)}[\hat{u}_2](P) = M[u](P) + o(1).$$

At those points P in $G(h)$ which do have a neighbor on $\partial G(h)$ we have

(6.4b) $$M_n^{(1)}[\hat{u}_2](P) = M[u](P) + O(1).$$

We carry out this computation for a point (x_{m_0+1}, y_j), and for simplicity, take $a = c = 1$,

$$(M_n^1[\hat{u}])_{m_0+1,j} = \frac{1}{\Delta x^2}\{u_{m_0+1,j} - 2u_{m_0+1,j} + u_{m_0+2,j}\}$$

$$+ \frac{1}{(\Delta y)^2}\{u_{m_0+1,j-1} - 2u_{m_0+1,j} + u_{m_0+1,j+1}\}.$$

The second term in this expression equals

$$\left(\frac{\partial^2}{\partial y^2}[u]\right)_{m_0+1,j} + o(1).$$

Turning to the first term, we observe that

$$u_{m_0+1,j} = u_{m_0,j} + \Delta x\left(\left(\frac{\partial u}{\partial x}\right)^{\sim}\right).$$

Since $(\partial u/\partial x) = 0$ on $x = X_0$ and this formula involves $(\partial u/\partial x)$ at a point whose distance from $x = X_0$ is at most $2\Delta x$, we see that

$$u_{m_0+1,j} = u_{m_0,j} + O(\Delta x^2).$$

An immediate computation verifies (6.4b).

Note. Although one could have chosen the mesh region so that $\partial G(h) = \partial G$ *in this case,* this is not necessary for our analysis. Hence we deliberately consider the more general situation in order to indicate how this analysis could be extended to a more general region G.

Moreover, $-M_n^1$ is positive semi-definite and the function $\tilde{e} \equiv 1$ is the only function in $\mathscr{U}_n^1$ for which $M_n^1[u] \equiv 0$.

Once more, summation by parts shows that $M_n^{(1)}$ satisfies a "coerciveness" condition over $\mathscr{U}_n^1$.

Let us now consider the point Jacobi iteration method and the multi-line (horizontal) Jacobi iteration methods. In both cases the splittings of A_n (the matrix representation of $-\Delta x_n \Delta y_n M_n^1$) satisfy property A (assuming the equations are consistently ordered). Hence, the analysis of §4 shows that these methods *do not converge*. However, the corresponding Gauss-Seidel methods and extrapolated relaxation methods do converge. Moreover, the rates of convergence of these convergent methods can be determined from the eigenvalues of the Jacobi procedures.

In order to apply Theorem 4, we must show that these splittings also satisfy property B. This is indeed the case. The proof is essentially the same proof as was given for these methods in §5 for the more general region. The only possible source of difficulty arises because we are no longer restricted to a compact subdomain on which both functions (u and v) are uniformly continuous. However, we need only observe that the contribution to $[u, \mathscr{N}_n v]_n$ from a strip near the boundary ∂G goes to zero.

Thus we obtain

THEOREM 5. *Let the points of $G_n(h)$ be consistently ordered. Consider the point Gauss-Seidel method applied to the problem*

$$M_n^1 v = f. \tag{6.5}$$

The method is convergent, and the quantity μ_n defined in §4 *is given by*

$$\mu_n = [\rho_n]^2, \tag{6.5a}$$

where

$$\rho_n \approx 1 - \Lambda\, \Delta x_n \Delta y_n + o(\Delta x_n \Delta y_n). \tag{6.5b}$$

In (6.5b) *the quantity Λ is the smallest nonvanishing eigenvalue of*

$$Mu + \Lambda\left(rc + \frac{1}{r}\,a\right)u = 0 \quad in\ G,$$

$$L^1[u] = 0 \quad on\ \partial G, \tag{6.6}$$

$$u \not\equiv 0.$$

Theorem 6. *Consider the multi-line (horizontal) Gauss-Seidel method applied to* (6.5a). *The method is convergent. Furthermore,*

$$\mu_n = [\rho_n]^2, \tag{6.7}$$

where

$$\rho_n \approx 1 - \frac{k}{2}\,\Lambda(\Delta y_n)^2 + o(\Delta x_n \Delta y_n). \tag{6.7a}$$

In (6.7a) *the quantity* Λ *is the smallest nonvanishing eigenvalue of*

$$Mu + \Lambda cu = 0 \quad in\ G,$$

$$L^{(1)}[u] = 0 \quad on\ \partial G, \tag{6.8}$$

$$u \not\equiv 0.$$

7. **The cyclically reduced equations.** Let M have the special form

$$M \equiv \frac{\partial}{\partial x}\,a\,\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\,c\,\frac{\partial}{\partial y}\,. \tag{7.1}$$

Let the mesh regions $\overline{G}_n(h)$ and the operators l_h and L_n^0 be chosen as in §5. At an interior point (x_m, y_j) the operators M_n^0 has the form

$$\begin{aligned}(M_n^0[v])_{m,j} = {} & \left(\frac{1}{\Delta x}\right)^2 \{a_{m+1/2,j}v_{m+1,j} - (a_{m+1/2,j} + a_{m-1/2,j})v_{m,j} - a_{m-1/2,j}v_{m-1,j}\} \\ & + \left(\frac{1}{\Delta y}\right)^2 \{c_{m,j+1/2}v_{m,j+1} - (c_{m,j+1/2} + c_{m,j-1/2})v_{m,j} + c_{m,j-1/2}v_{m,j-1}\}.\end{aligned} \tag{7.2}$$

Defintion . The lattice points $(x_m, y_j) \in \overline{G}_n(h)$ are divided into two classes, the red points $\mathscr{R}$ and the black points $\mathscr{B}$ (see [12]) defined by

$$\begin{aligned}\mathscr{R} &\equiv \{(x_m, y_j);\ m + j \equiv 1 \pmod 2\}, \\ \mathscr{B} &\equiv \{(x_m, y_j);\ m + j \equiv 0 \pmod 2\}.\end{aligned} \tag{7.3}$$

Let the points in $\overline{G}_n(h)$ be ordered as follows: the black points are enumerated first in "typewriter" ordering along the horizontal lines and then the red points are enumerated in "typewriter" ordering along the horizontal lines. We consider the discrete boundary-value problem (3.6). Let $X = \{v_{m,j}\}$ be a vector of unknowns *ordered in this way.* Then

(7.4a)
$$X = \begin{pmatrix} X_B \\ X_R \end{pmatrix},$$

where X_B denotes the vector of unknowns at the black points and X_R denotes the vector of unknowns at the red points. If A_n is the matrix representation of $-\Delta x_n \Delta y_n M_n^0$ (based on this ordering), then A_n has the form

(7.4b)
$$A_n = \begin{bmatrix} D_B & -B_0 \\ -B_0^T & D_R \end{bmatrix}.$$

The matrices D_B and D_R are diagonal matrices. The boundary-value problem is thus reduced to a system of linear equations of the form

(7.4c)
$$A_n \begin{pmatrix} X_B \\ X_R \end{pmatrix} = \begin{pmatrix} Y_B \\ Y_R \end{pmatrix} = Y,$$

where the vector Y is determined by the functions $\hat{f}$ and $\hat{\phi}$.

As we mentioned in §4, for computational purposes it is convenient to consider the equivalent system

(7.5)
$$\tilde{A}_n W = \tilde{Y},$$

where, having set

(7.5a)
$$D_n = \begin{bmatrix} D_R & 0 \\ 0 & D_B \end{bmatrix},$$

we define

(7.5b)
$$\begin{aligned} \tilde{A}_n &\equiv D_n^{-1/2} A_n D_n^{-1/2}, \\ W &\equiv A_n^{1/2} X, \\ \tilde{Y} &\equiv D_n^{-1/2} Y. \end{aligned}$$

Now we find that

(7.6)
$$\tilde{A}_n = \begin{bmatrix} I_B & -B_1 \\ -B_1^T & I_R \end{bmatrix},$$

where I_B and I_R are the identity matrices over the black points and the red points respectively.

The significant fact which is emphasized by the form of A_n and $\tilde{A}_n$ is: *the values of $v_{m,j}$ at the black (red) points depend only on the vector Y and the values of $v_{m,j}$ at the red (black) points.* Thus, one may easily eliminate one set of unknowns. To be explicit, we find that

(7.7a)
$$(2I - \tilde{A}_n)\tilde{A}_n W = (2I - \tilde{A}_n)\tilde{Y}$$

takes the form

(7.7b) $$(I_B - B_1 B_1^T) W_B = \widetilde{Y}_B + B_1 \widetilde{Y}_R,$$

(7.7c) $$(I_R - B_1^T B_1) W_R = \widetilde{Y}_R + B_1^T \widetilde{Y}_B.$$

However, one may solve (7.7b) for W_B and obtain W_R from

(7.7d) $$W_R = B_1^T W_B + \widetilde{Y}_R.$$

This reduction of (7.4c) to (7.7b) is called a cyclical reduction by Hageman [5] and Varga [11] (see [6] also). The equations (7.7b) are called the cyclically reduced equations. Let us consider these equations.

Let

(7.8a) $$\mathscr{A}_n = D_B^{1/2}(I_B - B_1 B_1^T) D_B^{1/2}, \qquad Z = D_B^{1/2}(\widetilde{Y}_B + B_1 \widetilde{Y}_R).$$

Then (7.7b) takes the form

(7.8b) $$\mathscr{A}_n X_B = Z,$$

where, of course, X_B is the original vector of unknowns $\{v_{m,j}\}$ *at the black points.*

Unfortunately, the set of black points in $\overline{G}_n(h)$ do not constitute a mesh region as this concept is defined in §3. The difficulty arises from condition (C.2). However, consider the coordinate system

(7.9) $$\xi = x + ry, \qquad \eta = x - ry,$$

where, as usual, $r = \Delta x_n / \Delta y_n$. Set

(7.10) $$\Delta\xi_n = \Delta\eta_n = [(\Delta x_n)^2 + (\Delta y_n)^2]^{1/2}.$$

The points $P_{m,j}$ defined by $\xi = m\Delta\xi_n \eta = j\Delta\eta_n$ which lie inside $\overline{G}$ are precisely the black points in $\overline{G}_n(h)$ which we denote by $\overline{G}_n(h; B)$. Thus in the $(\xi; \eta)$ coordinate system, $\overline{G}_n(h; B)$ is a mesh region.

However, if $\Delta x_n \neq \Delta y_n$, this is not an orthogonal coordinate system. But, this need not disturb us. We may interpret the original problem in the coordinate system

$$x^1 = x, \qquad y^1 = ry.$$

In this orthogonal coordinate system $\Delta x^1 = \Delta y^1$. Moreover, there is an obvious isomorphism between the domain G and the domain G^1, the elliptic operator M and the elliptic operator M^1, etc. On the other hand, it is important that we observe that the area of a basic rhombus of side $\Delta\xi$ is (in general) not $\Delta\xi_n \Delta\mathscr{N}_n$ but $2\Delta x_n \Delta y_n$. Hence, after triangulating $\overline{G}_n(h; B)$ by drawing the horizontal (or vertical) diagonals, we must take the inner product $[f, g]_n$ on $\mathscr{A}_n(B)$ (the space of piecewise linear functions) as

(7.11) $$[f, g]_n = 2\Delta x_n \Delta y_n \sum f(P) g(P), \qquad P \in \overline{G}_n(h; B).$$

The subspace $\mathscr{A}_n^0(B)$ is defined as the set of all functions $f \in \mathscr{A}_n(B)$ which vanish on $\partial G_n(h; B)$.

Let $\mathscr{S}_n^0$ be the linear operator which maps $\mathscr{A}_n^0(B)$ into $\mathscr{A}_n^0(B)$ which corresponds to the matrix $-(\Delta x_n \Delta y_n)^{-1} \mathscr{A}_n$.

Lemma 8. *The operator $\mathscr{S}_n^0$ is a consistent approximation to $2M^0$ over $G_n(h; B)$ in the sense that the conditions* (3.7) *and* (3.7a) *are met. We take $R = 2$.*

Proof. Condition (3.7) is obvious. We turn to condition (3.7a). Let

$$u \in C_2(\Omega)$$

and vanish outside a compact set $\Sigma \subset G$ whose distance from ∂G is greater than $4(\Delta x + \Delta y)$. Let $(x_m, y_j) \in G_n(h; B)$. By retracing the steps which led to (7.8b) we find that

$$(\mathscr{S}_n^0[\hat{u}])_{m,j} = 2M_n^0[\hat{u}]_{m,j} + \Delta x_n \Delta y_n (M_n^0 \tilde{D}^{-1} M_n^0[\hat{u}])_{m,j}, \tag{7.12}$$

where $\tilde{D}$ is the diagonal operator mapping $\mathscr{A}_n^0$ in $\mathscr{A}_n^0$ corresponding to the matrix D_n defined in (7.5). Hence we must only show that

$$\Delta x_n \Delta y_n M_n^0\{\tilde{D}^{-1} M_n^0[\hat{u}]\} = o(1). \tag{7.12a}$$

It is an easy matter to verify that

$$(\tilde{D}^{-1} v)_{m,j} = \left[\frac{2}{r} a_{m,j} + 2rc_{m,j}\right]^{-1} v_{m,j} + o(1) \tag{7.12b}$$

for all bounded $v \in \mathscr{A}_n^0$. Thus

$$\Delta x_n \Delta y_n M_n^0\{\tilde{D}^{-1} M_n^0[\hat{u}]\} = \Delta x_n \Delta y_n M_n^0[\phi + \psi], \tag{7.13}$$

where

$$\phi = \left[\frac{2}{r} a + 2rc\right]^{-1} \cdot M[u] \in C(G) \tag{7.13a}$$

and vanishes outside Σ, and

$$\psi = o(1) \quad \text{as } n \to \infty. \tag{7.13b}$$

However, a detailed look at M_n^0 shows that for n large enough

$$\Delta x_n \Delta y_n M_n^0[\phi] = o(1) \quad \text{as } n \to \infty$$

and

$$|\Delta x \Delta y M_n^0[\psi]| \leq K \operatorname{Max} |\psi|,$$

where

$$K = 4 \operatorname{Max}\left\{\frac{1}{r} a + rc\right\}.$$

LEMMA 9. *The operator $\mathcal{S}_n^0$ is self-adjoint and the matrix $\mathcal{A}_n$ is positive definite.*

Proof. See Hageman [5, Theorem 3.1].

LEMMA 10. *Let $\mathcal{A}_n = T_n - N_n$ be the splitting of $\mathcal{A}_n$ corresponding to the l-line (horizontal) Jacobi iterative method, $l \geqq 2$. This sequence of splittings satisfies property* B *with*

$$Q = \frac{2}{l}\left[\frac{2ac + r^2c^2}{rc + \frac{1}{r}a}\right]. \tag{7.14}$$

Moreover, this sequence of splittings also satisfies block property A.

Proof. See Hageman [5] for the proof that these splittings satisfy block property A. Let $u \in C_2(\Omega)$ vanish outside a compact set $\Sigma \subset G$ and let $v \in C(G)$. Then, after a tedious computation using the results of [5, pp. 78-79], to determine the elements of $\mathcal{A}_n$, we find that

$$[\hat{u}, \mathcal{N}_n \hat{v}]_n = 4\Delta x_n \Delta y_n \sum_{(\nu)} I_\nu, \tag{7.15}$$

where

$$\begin{aligned} I_\nu = {} & \sum_{(m)}{}' [\beta_T(m, \nu l - 1) v_{m,\nu l+1} u_{m,\nu l-1}] \\ & + \sum_{(m)}{}' [\beta_T(m, \nu l) v_{m,\nu l+2} u_{m,\nu l}] \\ & + \sum_{(m)}{}' [\alpha_{TL}(m, \nu l) v_{m-1,\nu l+1} u_{m,\nu l}] \\ & + \sum_{(m)}{}' [\alpha_{TR}(m, \nu l) v_{m+1,\nu l+1} u_{m,\nu l}] \end{aligned} \tag{7.15a}$$

and the α's and β's satisfy, for $(x_m, y_j) \in \Sigma$,

$$\beta_T(m,j) = \left(\frac{r^2c^2}{2\left[\frac{1}{r}a + rc\right]}\right)_{m,j} + o(1) \quad \text{as } \Delta x_n \to 0,$$

$$\alpha_{TR}(m,j) = \frac{2ac}{2\left[\frac{1}{r}a + rc\right]} + o(1) \quad \text{as } \Delta x_n \to 0, \tag{7.15b}$$

$$\alpha_{TR}(m,j) = \alpha_{TL}(m,j) + o(1) \quad \text{as } \Delta x_n \to 0$$

and the $'$ on the summation sign indicates that this summation is over the black points. Thus

$$2\Delta x_n I_\nu = \frac{l}{2}\int Q(\xi, y_{\nu l})\, u(\xi, y_{\nu l})\, v(\xi, y_{\nu l})\; d\xi + o(1). \tag{7.16}$$

We observe that only $1/l$ of the lines $y = j\Delta y$ enter into the sum (7.15). Thus the lemma is proven.

Lemma 11. *Let $\mathscr{A}_n = T_n - N_n$ be the splitting of $\mathscr{A}_n$ corresponding to the point Jacobi iterative method. In this case the splitting does not satisfy property* A. *However, from the theory of positive matrices (see* [11]*) we may deduce that* (4.15) *applies. Moreover, the splitting satisfies property* B *with*

$$Q(x,y) = 2\left\{\frac{2ac + r^2c^2}{\frac{1}{r}a + rc}\right\}.$$

Proof. Let u and v be the functions in Lemma 10. Then

$$[\hat{u}, \mathscr{N}_n \hat{v}]_n = 2\Delta x_n \Delta y_n\left[\sum_1 + \sum_2\right], \tag{7.17}$$

where

$$\begin{aligned}\textstyle\sum_1 = {} & \sum_{(m,j)}{}' \left[\beta_T(m,j)v_{m,j+2}u_{m,j} + \beta_B(m,j)v_{m,j-2}u_{m,j}\right] \\ & + \sum_{(m,j)}{}' \left[\beta_R(m,j)v_{m+2,j}u_{m,j} + \beta_L(m,j)v_{m-2,j}u_{m,j}\right], \\ \textstyle\sum_2 = {} & \sum_{(m,j)}{}' \left[\alpha_{TR}(m,j)v_{m+1,j+1}u_{m,j} + \alpha_{TL}(m,j)v_{m-1,j+1}u_{m,j}\right] \\ & + \sum_{(m,j)}{}' \left[\alpha_{BR}(m,j)v_{m+1,j-1}u_{m,j} + \alpha_{BL}(m,j)v_{m-1,j-1}u_{m,j}\right]\end{aligned} \tag{7.17a}$$

and all the α's approach $\alpha_{TR}(m,j)$ as $\Delta x_n \to 0$ and all the β's approach $\beta_T(m,j)$ as $\Delta x_n \to 0$.

Thus using (7.15b), we have

$$[\hat{u}, \mathscr{N}_n \hat{v}]_n = 2\Delta x_n \Delta y_n \sum{}' \left\{\frac{2r^2c^2 + 4ac}{\frac{1}{r}a + rc}\right\} u_{m,j}v_{m,j} + o(1),$$

which proves the lemma.

Lemma 12. *The operator $\mathscr{S}_n^0$ satisfies a coerciveness condition.*

Proof. We use the notation which has been used implicitly in the two preceding lemmas. That is, we write

$$
\begin{aligned}
-\Delta x_n \Delta y_n(\mathscr{S}_n{}^0[u])_{m,j} = {} & (\alpha_0(m,j) + \beta_0(m,j))u_{m,j} \\
& - \{\alpha_{TR}(m,j)u_{m+1,j+1} + \alpha_{TL}(m,j)u_{m-1,j+1} \\
& + \alpha_{BR}(m,j)u_{m+1,j-1} + \alpha_{BL}(m,j)u_{m-1,j-1}\} \\
& - \{\beta_T(m,j)u_{m,j+2} + \beta_B(m,j)u_{m,j-2} \\
& + \beta_R(m,j)u_{m+2,j} + \beta_L(m,j)u_{m-2,j}\},
\end{aligned}
\tag{7.18}
$$

where $\alpha_0(m,j)$ and $\beta_0(m,j)$ are defined by setting

$$\alpha_0(m,j) = \alpha_{TR}(m,j) + \alpha_{TL}(m,j) + \alpha_{BR}(m,j) + \alpha_{BL}(m,j). \tag{7.18a}$$

Using the formulas given by Hageman [5, pp. 78, 79] or by making a direct calculation, one may verify that all the α's and all the β's, with the possible exception of β_0, are positive. We claim that, in fact,

$$\beta_0(m,j) \geqq \beta_T(m,j) + \beta_B(m,j) + \beta_R(m,j) + \beta_L(m,j). \tag{7.19}$$

Suppose this has been verified. Then we may write $\mathscr{S}_n{}^0$ as

$$\mathscr{S}_n{}^0 = \mathscr{R}_n^0 + C_n^0, \tag{7.20}$$

where C_n^0 is the operator determined by the α's and $\mathscr{R}_n^0$ is the operator determined by the β's. By construction, C_n^0 satisfies a coerciveness condition in the (ξ,η) coordinates. And, this is sufficient for our purposes. Moreover, (7.19) implies that $\mathscr{R}_n^0$ and C_n^0 are both negative semi-definite. Thus

$$|[u, \mathscr{S}_n{}^0 u]_n| = |[u, \mathscr{R}_n^0 u]_n| + |[u, C_n^0 u]_n|.$$

If

$$|[u, \mathscr{S}_0{}^n u]_n| \leqq K$$

then, of course

$$|[u, C_n^0 u]_n| \leqq K$$

and the coerciveness condition on C_n^0 implies that $\mathscr{S}_n{}^0$ satisfies a coerciveness condition.

We return to the proof of (7.19). Let $u_0 \equiv 1$ in the interior, $G_n(h;B)$. Then, if $(x_m, y_j) \in G_n(h;B)$ and all of the neighbors of (x_m, y_j) are also in $G_n(h;B)$, the representation (7.12) shows that

$$(\mathscr{S}_0{}^n[v])_{m,j} = 0. \tag{7.21}$$

By construction,

$$(C_n^0[v])_{m,j} = 0.$$

Hence, from the definition of $\mathscr{R}_n^0$, we have

(7.22) $$\beta_0(m,j) = \beta_T(m,j) + \beta_B(m,j) + \beta_R(m,j) + \beta_L(m,j).$$

For those points $(x_m, y_j) \in G_n(h; B)$ who have at least one neighbor on $\partial G_n(h; B)$ one must look again at the formulas of Hageman. One easily concludes that in this case

$$\beta_0(m,j) > \beta_T(m,j) + \beta_B(m,j) + \beta_R(m,j) + \beta_L(m,j).$$

THEOREM 7. *Let $\rho_n({}^RJ)$ denote the dominant eigenvalue of the point Jacobi iterative method applied to the cyclically reduced equations* (7.7b). *Let $\rho_n({}^RJ^{lL})$ denote the dominant eigenvalue of the l-line $(l \geqq 2)$ Jacobi iterative method applied to the cyclically reduced equation* (7.7b). *Then*

(7.23a) $$\rho_n({}^RJ) \approx 1 - \Lambda \Delta x_n \Delta y_n + o(\Delta x_n \Delta y_n),$$

and

(7.23b) $$\rho_n({}^RJ^{lL}) \approx 1 - l\Lambda \Delta x_n \Delta y_n + o(\Delta x_n \Delta y_n),$$

where Λ is the smallest eigenvalue of

(7.24) $$Mu + \Lambda \left[\frac{2ac + r^2c^2}{\frac{1}{r}a + rc} \right] u = 0 \quad \text{in } G, \qquad u = 0 \quad \text{on } \partial G, \qquad u \not\equiv 0.$$

Proof. From the remarks following Theorem 2, we see that Lemmas 8, 9, 11, 12 and Theorem 2 imply (7.23a). Similarly, Lemmas 8, 9, 10, 12 and Theorem 2 imply (7.23b).

COROLLARY. *Let G be a square of side π. Let M be the Laplace operator. Let $\Delta x = \Delta y = h$. Then*

(7.25) $$\rho({}^RJ^{lL}) \approx 1 - \frac{4}{3} lh^2.$$

Proof. In this case Λ is the smallest eigenvalue of

$$\Delta u + \left[\frac{3}{2} \Lambda \right] u = 0.$$

And, as is well known, if $[3\Lambda/2]$ is denoted by Λ_1, then

$$\Lambda_1 = 2.$$

COROLLARY. *Let G be a square of side π. Let M be the Laplace operator. Then*

$$\rho({}^RJ) \approx 1 - \frac{4}{3} h^2.$$

In addition to the horizontal multi-line methods for these reduced equations, one may also consider the multi-line iterative methods which arise when we take blocks of l-lines in the oblique directions $\xi =$ constant or $\eta =$ constant.

Using arguments similar to those used in Lemmas 10 and 11, we obtain the following result.

THEOREM 8. *Consider the l-line ($\xi =$ constant) Jacobi iterative method, $l \geqq 1$. Let $\rho({}^RJ^{lL};\ \xi)$ designate the dominant eigenvalue of the method. Then*

$$\rho({}^RJ^{lL};\ \xi) \approx 1 - l\Lambda_0(\Delta x\, \Delta y), \qquad l \geqq 1,$$

where Λ_0 is the smallest eigenvalue of

$$Mu + \frac{1}{2}\left(\frac{1}{r}\,a + rc\right)\Lambda u = 0 \quad \textit{in } G,$$

$$u = 0 \quad \textit{on } \partial G.$$

REMARK. Observe that in this case $\rho({}^RJ^{lL};\ \xi) \approx \rho({}^RJ^{lL};\ \eta)$.

COROLLARY. *In the special case when M is the Laplace operator, G is a square of side π and $\Delta x = \Delta y = h$ we have*

$$\rho({}^RJ^{lL};\ \xi) \approx 1 - 2lh^2, \qquad l \geqq 1.$$

These results have been obtained without interpreting the operator $\mathscr{S}_n{}^0$ as an operator which arises in a natural way from an operator M_n and a "boundary" operator L_n^0. However, if G is a convex domain, this can be done also. We will illustrate this remark with the case described above.

Let G be the square of side π, $0 \leqq x, y \leqq \pi$. Let M be the Laplace operator. Let

(7.26) $$\Delta x_n = \Delta y_n = h_n = \pi/(n+1).$$

Let

(7.27) $$\overline{G}_n(h;\ E, B) \equiv \{(mh_n, jh_n);\ m + j \equiv 0\ (\mathrm{mod}\, 2),\ -1 \leqq m, j \leqq n+2\}.$$

Set

(7.28)
$$\begin{aligned} G_n(h;\ E, B) &\equiv \{(x_m, y_j) \in \overline{G}_n(h;\ E, B);\ 1 \leqq m, j \leqq n\},\\ \partial G_n(h;\ E, B, 0) &\equiv \{(x_m, y_j) \in \overline{G}_n(h;\ E, B);\ m \cdot j = 0\},\\ \partial G_n(h;\ E, B, 1) &\equiv \overline{G}_n(h;\ E, B) - \{G_n(h; E, B) + \partial G_n(h; E, B, 0)\},\\ \partial G_n(h;\ E, B) &\equiv \partial G_n(h;\ E, B, 0) + \partial G_n(h;\ E, B, 1). \end{aligned}$$

We let l_k be the usual 5-point approximation to M (see (7.2)) and define

(7.29) $$\mathscr{M}_n \equiv 2l_h + \Delta x_n\, \Delta y_n\, l_h^2.$$

The boundary operator $\mathscr{L}_n{}^0$ is defined as the identity on

$$G_n(h;\ E,B) + \partial G_n(h;\ E,B,0).$$

The definition of $\mathscr{L}_n{}^0$ on $\partial_n(h;\ E,B,1)$ is a bit complicated to describe, but simple enough. Let

$$\partial_n(h;\ E,B,1) = S_1 + S_2 + S_3 + S_4 + K, \tag{7.30}$$

where

$$\begin{aligned} S_1 &\equiv \{(x_m,y_j);\ 1 \leqq j \leqq n;\ m = -1\}, \\ S_2 &\equiv \{(x_m,y_j);\ 1 \leqq m \leqq n; j = -1\}, \\ S_3 &\equiv \{(x_m,y_j);\ 1 \leqq j \leqq n;\ m = n+2\}, \\ S_4 &\equiv \{(x_m,y_j);\ 1 \leqq m \leqq n;\ j = n+2\}. \end{aligned} \tag{7.30a}$$

The operator $\mathscr{L}_n{}^0$ is taken as the identity over K. We will describe $\mathscr{L}_n{}^0$ over S_1 and assume that the reader can make the similar definition over S_2, S_3 and S_4. Let $1 \leqq j \leqq n$; then

$$(\mathscr{L}_n{}^0[u])_{-1,j} \equiv u_{0,j+1} + u_{0,j-1} - u_{1,j}. \tag{7.30b}$$

As usual, we set

$$\mathscr{A}_n^0(E,B) \equiv \{f \in \mathscr{A}_n(E,B);\ \mathscr{L}_n{}^0 f = 0 \text{ on } \partial G_n(h;\ E,B)\}.$$

A simple computation verifies that $\mathscr{S}_n{}^0$ is the operator mapping $\mathscr{A}_n^0(E,B)$ into $\mathscr{A}_n^0(E,B)$ and

$$(\mathscr{S}_n{}^0[v])_{m,j} = (\mathscr{M}_n[v])_{m,j} \quad \text{if } (x_m,y_j) \in G_n(h;\ E,B). \tag{7.31}$$

It will be instructive to consider another operator related to $\mathscr{S}_n{}^0$. Let $\bar{G}_n(h;\ E,B)$ be the set of black points in the plane which have neighbors in $\bar{G}_n(h;B)$. Clearly, $\bar{G}_n(h;E,B)$ is a mesh region in the (ξ,η) coordinate sys- em. We define the interior of $G_n(h;E,B)$ by

$$G_n(h;\ E,B) = G_n(h;B). \tag{7.32}$$

Let the coefficients $a(x,y)$, $c(x,y)$ be defined over an open set $\Omega \supset \bar{G}$ so that the operator l_h defined in §5 is well defined over $\bar{G}_n(h;E,B)$. Let

$$\mathscr{M}_n \equiv 2l_h + \Delta x_n \Delta y_n l_h \tilde{D}^{-1} l_h, \tag{7.33}$$

where $\tilde{D}$ is the diagonal operator defined on $\bar{G}_n(h;E,B)$ by

$$(\tilde{D}[v])_{m,j} = \left[\frac{1}{r}(a_{m-1/2,j} + a_{m+1/2,j}) + r(c_{m,j-1/2} + c_{m,j+1/2})\right] v_{m,j}. \tag{7.33a}$$

Let L_n^0 be the identity over $\bar{G}_n(h;\ E,B)$. Let $\mathscr{M}_m^0$ be the operator defined on $\mathscr{A}_n^0(E,B)$ by

$$\begin{aligned} \mathscr{M}_n^0[u] &= \mathscr{M}_n[u] \quad \text{in } G_n(h;E,B), \\ &= 0 \quad \text{on } \partial G_n(h;E,B). \end{aligned} \tag{7.33b}$$

One may easily show that $\mathscr{M}_n^0$ is a consistent approximation to $2M^0$ and Lemmas 8, 9, 10, 11, 12 all hold. Thus, if $\sigma_n({}^RJ)$ and $\sigma_n({}^RJ^{lL})$ denote the dominant eigenvalues of the point and l-line (horizontal) Jacobi iterative methods applied to the matrix associated with $-\Delta x_n \Delta y_n \mathscr{M}_n^0$, then

$$\begin{aligned} \sigma_n({}^RJ) &\approx \rho_n({}^RJ) + o(\Delta x_n \Delta y_n), \\ \sigma_n({}^RJ^{lL}) &\approx \rho_n({}^RJ^{lL}) + o(\Delta x_n \Delta y_n), \qquad l \geqq 2. \end{aligned} \tag{7.34}$$

8. **Concluding remarks.** In general, one cannot determine the exact value of the quantity Λ which appears in our estimates. Nevertheless, in the case of Dirichlet boundary data, one can make comparisons by considering the variational characterization of Λ given in Lemma 4. Moreover, one can estimate Λ from above and below by considering inscribed and circumscribed rectangles and the quantities

$$\begin{aligned} Q_{\max} &\equiv \operatorname*{Max}_{\bar{G}} Q(x,y), \\ Q_{\min} &\equiv \operatorname*{Min}_{\bar{G}} Q(x,y). \end{aligned} \tag{8.1}$$

In those cases where the splitting $A_n = T_n - N_n$ satisfies block property A, a good estimate for the ρ associated with the Jacobi iterative method enables one to obtain good estimates for the optimal value of ω. Using such a "good" value for ω can result in very great savings (see [12]).

Theorems 1 and 3 can be very valuable in more general situations. For example, Theorem 1 remains valid when the sequence $\bar{G}_n(h)$ converges to G in a somewhat weaker sense. To be exact, since the "test" function $\hat{u}$ used in the proof of Theorem 1 has compact support $\Sigma \subset G$, one need not require condition (e). Moreover, in some cases an estimate of the form (4.11) is sufficient for one's needs. For example, in the so-called "model" problem, the Laplace operator in a rectangle, one may obtain the exact formula for $\lambda_R(1)$ (see §5) by elementary means. And, as we have commented in [9], the formulas for $\lambda_R(k)$ lead one to conclude that $k = 1$ is probably the optimal choice. However, this conclusion follows from the estimates (4.11) *and the exact value for* $\lambda_R(1)$. In many cases, "upper bounds" for ρ_n may be obtained by other methods. In such cases, the lower bounds provided by Theorems 1 and 3 may complete the picture without further efforts. For example, in the model problem studied by Hageman [5] for the cyclically reduced equations he found

$$\rho({}^{R}J^{2L}) \lesssim 1 - \frac{8}{3}h^2 + o(h^2). \tag{8.2}$$

Thus, the estimate (4.11) applied to this case gives the correct asymptotic formula (see §7, (7.25)).

Finally, we comment on the significance of the formulas (7.34). This result, together with the characterization of Λ shows that the first term in the asymptotic expansion for ρ is a geometric term and essentially insensitive to the treatment of the boundary values. Indeed, one can prove that

$$\sigma_n({}^{R}J) > \rho_n({}^{R}J),$$

$$\sigma_n({}^{R}J^{lL}) > \rho_n({}^{R}J^{lL}), \qquad l \geqq 2,$$

for all n!. Nevertheless, the first terms in the asymptotic expansions are identical.

References

1. R. Courant and D. Hilbert, *Methoden der mathematischen Physik,* Vols. I, II, Springer, Berlin, 1931.
2. G. E. Forsythe and W. Wasow, *Finite-difference methods for partial differential equations,* Wiley, New York, 1960.
3. G. Golub and R. S. Varga, *Chebyshev semi-iterative methods, successive overrelaxation iterative methods, and second order Richardson iterative methods.* I, Numer. Math. **3** (1961), 147-156.
4. ———, *Chebyshev semi-iterative methods, successive overrelaxation iterative methods, and second order Richardson iterative methods.* II, Numer. Math. **3** (1961), 157-168.
5. L. A. Hageman, *Block iterative methods for two-cyclic matrix equations with special application to the numerical solution of the second-order self-adjoint elliptic partial differential equation in two dimensions,* WAPD-TM-327, Bettis Atomic Power Lab., Pittsburgh, Pa., 1962, 135 pp.
6. L. A. Hageman and R. S. Varga, *Block iterative methods for cyclically reduced matrix equations,* Numer. Math. **6** (1964), 106-119.
7. L. Nirenberg, *Remarks on strongly elliptic partial differential equations,* Comm. Pure Appl. Math. **8** (1955), 648-674.
8. J. Nitsche and J. C. C. Nitsche, *Error estimates for the numerical solution of elliptic differential equations,* Arch. Rational Mech. Anal. **5** (1960), 293-306.
9. S. V. Parter, *Multi-line iterative method for elliptic difference equations and fundamental frequencies,* Numer. Math. **3** (1961), 305-319.
10. F. Riesz and B. Sz. Nagy, *Functional analysis,* Ungar, New York, 1955.
11. R. S. Varga, *Matrix iterative analysis,* Prentice-Hall, Englewood Cliffs, N. J., 1962.
12. D. Young, *The numerical solution of elliptic and parabolic partial differential equations,* A survey of numerical analysis, McGraw-Hill, New York, 1962.
13. G. Hellwig, *Partielle Differential-gleichungen: Eine Einführung,* Teubner, Stuttgart, 1960.

STANFORD UNIVERSITY,
STANFORD, CALIFORNIA

CORNELL UNIVERSITY,
ITHACA, NEW YORK