IMPLICIT ALTERNATING DIRECTION METHODS

BY

GARRETT BIRKHOFF AND RICHARD S. VARGA

1. **Introduction.** The arithmetic solution of self-adjoint elliptic difference equations, associated with differential equations of the form

$$\text{(A)} \qquad \Sigma u - \nabla\cdot(D\nabla u) = S, \qquad \Sigma(x) \geqq 0, \qquad D(x) > 0, \qquad S = S(\mathbf{x}),$$

in general plane regions and with respect to linear boundary conditions, is a classical problem of numerical analysis. Many such boundary value problems have been solved successfully on high-speed computing machines, using the (iterative) Young-Frankel "successive overrelaxation" (SOR) method as defined in [1] and [2], and variants thereof ("line" and "block" overrelaxation). For this method, estimates of the rate of convergence and the "optimum relaxation factor" can both be rigorously extended from the special case of $-\nabla^2 u = S$ in a rectangle, and Dirichlet-type boundary conditions, to the general case.

Recently, two variants [3; 4] of a new "implicit alternating direction" (IAD) method have been proved to converge much more rapidly, in the special case just mentioned, than the successive overrelaxation method and its variants. This fact has led to much speculation regarding the relative rates of convergence of SOR and IAD methods for more general elliptic boundary value problems. In [3, p. 41], success was reported in solving $-\nabla^2 u = S$ for "several examples involving more complex regions and less simple boundary conditions," but no theoretical analysis was given of the convergence rate of the "Peaceman-Rachford" process used in these examples. In [4, p. 421], the "Douglas-Rachford" process was asserted to be stable (i.e., convergent) for $-\nabla^2 u = S$ and Dirichlet-type boundary conditions in general plane regions, but again the cases of variable D, $\Sigma \neq 0$, and mixed boundary conditions were not covered.

Our main result below (Theorem 3) is that the convergence estimates given in [3] and [4] are applicable to the modified Helmholtz equation $c^2 u - \nabla^2 u = S$ in a rectangle with sides parallel to the axes, with the boundary condition $\partial u/\partial n + ku = 0$, $k \geqq 0$, and essentially to *no other case*. In dramatic contrast to the situation as regards SOR methods, the analysis of the rectangular case, via discrete Fourier analysis of eigenfunctions, gives no clue as to the general self-adjoint case.

Nevertheless, IAD methods work well for many other cases of great interest. We present in §§8–9 also a few positive results, giving partial theoretical justification for this success.

Presented to the Society, November 16, 1957; received by the editors January 15, 1958.

2. **The matrices** H, V, Σ. Though our analysis generalizes [3] and [4], by allowing D to be variable and Σ nonzero, it will also be limited to 5-point difference approximations to the expression $-\nabla\cdot D(\nabla u)$ in (A), on a square(1) mesh of side h. Thus, such approximations can be resolved into two 3-point difference operators H and V, representing

$$-h^2\partial(D\partial u/\partial x)/\partial x \text{ and } -h^2\partial(D\partial u/\partial y)/\partial y,$$

respectively, and given, at interior points, by

$$[uH](x,y) = -a(x,y)u(x+h,y) + 2b(x,y)u(x,y) - c(x,y)u(x-h,y), \tag{1}$$

$$[uV](x,y) = -\alpha(x,y)u(x,y+h) + 2\beta(x,y)u(x,y) - \gamma(x,y)u(x,y-h). \tag{1'}$$

The most natural choice for a, b, c is probably

$$a = D(x+h/2,y), \qquad c = D(x-h/2,y), \qquad 2b = a+c, \tag{2}$$

and similarly for α, β, γ

$$\alpha = D(x,y+h/2), \qquad \gamma = D(x,y-h/2), \qquad 2\beta = \alpha+\gamma. \tag{2'}$$

This choice makes H and V *symmetric* (obviously).

However, other choices for a, b, c, α, β, γ, are reasonable, such as

$$a = D(x,y) + D(x+h,y)/4 - D(x-h,y)/4, \tag{3a}$$

$$b = D(x,y) = \beta = (a+c)/2 = (\alpha+\gamma)/2, \tag{3b}$$

$$c = D(x,y) - D(x+h,y)/4 + D(x-h,y)/4, \tag{3c}$$

etc. We mention this, because of its relevance to the question of when $HV=VH$, discussed below. In the case $D=1$ of the Laplace operator, (3a)–(3c) and (2)–(2′) are clearly both equivalent to $a=c=\alpha=\gamma=1$, $b=\beta=2$.

Our analysis will be confined to the usual approximations (2)–(2′). In the case of *Dirichlet-type* problems (i.e., when $u(x_j, y_k)$ is given at the boundary mesh-points), the values of u at interior mesh-points will be interpreted as usual, as the components of an unknown "vector" $\boldsymbol{u}$. Hence Σ in (A) will be interpreted as multiplication of $\boldsymbol{u}$ by the *diagonal* matrix Σ, with ith diagonal entry $\Sigma_{ii}=\Sigma(x_j, y_k)$. Inspection of (2)–(2′) reveals H and V as symmetric matrices with positive diagonal entries, and nonpositive real, diagonally dominated(2) off-diagonal entries. Mesh-points adjacent to the boundary yield missing off-diagonal entries (see §7), from which the strict diagonal dominance follows. The given boundary values of u are combined as usual with the source terms $h^2S(x_j, y_k)$, to define $\boldsymbol{k}$ in the following, essentially known result.

(1) As in [3; 4] the case of a rectangular mesh offers no additional difficulties.

(2) This means that $a_{i,i} \geqq \sum_{j\neq i} |a_{i,j}| = \sum_{j\neq i}(-a_{i,j})$, strict inequality holding in at least one case.

LEMMA 1. *Dirichlet-type problems for* (1) *are approximately equivalent, for small h, to vector equations of the form*

$$u(H + V + \Sigma) = k. \tag{4}$$

Here Σ *is non-negative and diagonal*; H, V, $H+\Sigma$ *and* $V+\Sigma$ *are diagonally dominated Stieltjes matrices*; $H+V+\Sigma$ *has a positive inverse if the domain* R_h *of* (*interior*) *mesh-points is connected.*

EXPLANATION. A *Stieltjes*(3) *matrix* is a positive definite symmetric matrix, with nonpositive off-diagonal entries. The properties of Stieltjes matrices have been described in various forms in the literature. Every Stieltjes matrix H has positive diagonal entries, and is the direct sum of irreducible Stieltjes matrices H_i. Each H_i is then positive definite, and has a positive inverse; see [5, p. 602].

3. **Alternating direction methods.** Equation (4) is clearly equivalent, for any scalar ρ, to each of the two vector equations

$$u(H + \Sigma + \rho I) = k - u(V - \rho I), \tag{5}$$

$$u(V + \Sigma + \rho I) = k - u(H - \rho I). \tag{5'}$$

Following [3, pp. 33–34], where only the case $\Sigma = 0$ was considered, we define, with Sheldon and Wachspress [6], the Peaceman-Rachford method as the semi-iterative process defined by

$$u_n^*(H + \Sigma + \rho_n I) = k - u_n(V - \rho_n I), \tag{6}$$

$$u_{n+1}(V + \Sigma + \rho_n I) = k - u_n^*(H - \rho_n I), \tag{6'}$$

for some sequence of positive numbers ρ_n. Since the matrices which have to be inverted are similar to tridiagonal(4) or Jacobi matrices under permutation matrices, each of equations (6) and (6′) can be rapidly solved by Gauss elimination. The sequence of constants $\rho_1, \rho_2, \rho_3, \cdots$ in (6)–(6′) is intended to be chosen so as to make convergence of the u_n rapid.

If U is the unique solution of (4), then the error vector $E_n = u_n - U$ is multiplied through the performance of (6)–(6′) by the matrix

$$T_\rho = (V - \rho I)(H + \Sigma + \rho I)^{-1}(H - \rho I)(V + \Sigma + \rho I)^{-1}, \qquad \rho = \rho_n. \tag{6*}$$

Sheldon and Wachspress [6], also tried a variant of the preceding process, defined by

$$u_n^*(H + \Sigma/2 + \rho_n I) = k - u_n(V + \Sigma/2 - \rho_n I), \tag{7}$$

$$u_{n+1}(V + \Sigma/2 + \rho_n I) = k - u_n^*(H + \Sigma/2 - \rho_n I). \tag{7'}$$

(3) Such matrices were originally considered by T. J. Stieltjes, Acta Math. vol. 9 (1887) pp. 385–400.

(4) This observation was first applied by Crank and Nicolson, Proc. Cambridge Philos. Soc. vol. 43 (1947) pp. 50–67, to solve the heat equation in one space dimension. It is used also in line and block overrelaxation.

The error reduction is then expressed by

$$(7^*)\qquad \begin{aligned} S_\rho &= (V_1 - \rho I)(H_1 + \rho I)^{-1}(H_1 - \rho I)(V_1 + \rho I)^{-1}, \\ V_1 &= V + \Sigma/2, \qquad H_1 = H + \Sigma/2. \end{aligned}$$

Finally, Douglas and Rachford have proposed a third variant, defined by (6) and

$$(8)\qquad u_{n+1}(V + \Sigma + \rho_n I) = u_n(V + \Sigma) + \rho_n u_n^*;$$

see [4, p. 422, (2.3)]. The error reduction at the nth iteration then corresponds (for $\rho = \rho_n$) to multiplication by the matrix

$$U_\rho = \{(V + \Sigma)(H + \Sigma) + \rho\Sigma + \rho^2 I\}(H + \Sigma + \rho I)^{-1}(V + \Sigma + \rho I)^{-1}.$$

If $\rho = \rho_1 = \rho_2 = \rho_3 = \cdots$ is *fixed*, the rate of convergence of each of the three processes just defined is determined by the *spectral radius*(5) of the appropriate matrix S_ρ, T_ρ or U_ρ.

4. **Permutable case.** The analysis of [3] and [4] covered only the case $HV = VH$, $\Sigma = 0$ of the Dirichlet problem in a $(Mh \times Nh)$-rectangle. We shall now see how far it can be extended.

It is almost obvious that the analysis can be extended to any case of permutable operators—i.e., to

$$(9)\qquad HV = VH, \quad H\Sigma = \Sigma H, \quad \text{and} \quad V\Sigma = \Sigma V.$$

By a classic theorem of Frobenius, (9) is equivalent to a common basis of (real, orthogonal) eigenvectors for the operators H, V, and Σ; it is also equivalent to the symmetry of HV, $H\Sigma$ and $V\Sigma$.

In the case of the rectangle $0 \leqq x \leqq a$, $0 \leqq y \leqq b$, with Dirichlet-type boundary conditions, the functions $\sin(m\pi x/a)\sin(n\pi y/b)$ obviously form a basis of eigenfunctions for the modified Helmholtz equation $\nabla^2 u = (\Sigma/D)u$, with Σ and D constant. Hence (9) holds in this case, with symmetric HV, $H\Sigma$, and $V\Sigma$.

If commutativity (9) is assumed, estimates of the convergence rates of the schemes described above can be made easily. If $\|E\| = (\Sigma E_i^2)^{1/2}$ is the Euclidean length of a vector E, we can define the *Euclidean norm* of a matrix T as $\sup(\|ET\|/\|E\|)$ for nonzero E. If (9) holds, all rational functions of H, V and Σ are symmetric, and so the Euclidean norms of S_ρ, T_ρ and U_ρ equal their respective spectral radii. Further, let $\epsilon_1, \cdots, \epsilon_r$ be a basis of common (orthogonal) eigenvectors for H, V and Σ; let $\epsilon_k H = \sigma_k \epsilon_k$, $\epsilon_k V = \tau_k \epsilon_k$, and $\epsilon_k \Sigma = \nu_k \epsilon_k$. Then, under (6*), ϵ_k is multiplied by λ_{kn}, where

$$(10)\qquad |\lambda_{kn}| = \left|\frac{\sigma_k - \rho_n}{\sigma_k + \rho_n + \nu_1}\right| \cdot \left|\frac{\tau_k - \rho_n}{\tau_k + \rho_n + \nu_k}\right|.$$

(5) The spectral radius (or norm) of a matrix is the maximum of the moduli of its eigenvalues. For the relation to convergence, see [1, p. 94].

Similar formulas hold for S_ρ and U_ρ. Note that, in (10), the ν_k decrease the λ_{kn}; a similar result holds for U_ρ—though not for S_ρ.

Since the σ_k and τ_k are all positive, it follows that the spectral radii of all processes are less than unity for all $\rho_n > 0$. Further, since all matrices involved are symmetric, the error eigenvectors are orthogonal, and the spectral radius is identical with the ordinary Euclidean norm. Considering the cumulative effect on the ϵ_k, we thus conclude

THEOREM 1. *If* (9) *holds, then the error vector* $\mathbf{E}_N$ *after* N *application of the Peaceman-Rachford process* (6)–(6′) *satisfies*

$$\|\mathbf{E}_N\|/\|\mathbf{E}_0\| \leqq \sup_k \prod_{n=1}^{N} \left| \frac{\sigma_k - \rho_n}{\sigma_k + \rho_n} \cdot \frac{\tau_k - \rho_n}{\tau_k + \rho_n} \right|. \tag{11}$$

Similarly, after N *applications of* (6) *and* (8)

$$\|\mathbf{E}_N\|/\|\mathbf{E}_0\| \leqq \sup_k \prod_{n=1}^{N} \left| \frac{\sigma_k \tau_k + \rho_n^2}{(\sigma_k + \rho_n)(\tau_k + \rho_n)} \right|. \tag{12}$$

REMARK. The scheme (7)–(7′) while appearing to be less efficient than (6)–(6′), is equally efficient in the permutable case.

5. **Nonpermutable case.** Young [1] has shown that the convergence estimates obtained, as in [2], for the simple case of the Dirichlet problem for the rectangle, are applicable to the general problem described in §§1–2, if systematic overrelaxation (SOR) is used. We shall now show that the situation with IAD methods is entirely different. Essentially, the modified Helmholtz equation in a rectangle is the *only* case to which the estimates of §4 can be applied—at least with the difference approximation (2)–(2′).

LEMMA 1. *The relation* $HV = VH$ *implies*

$$\begin{aligned} D(x + h/2, y) &= D(x + h/2, y + h) \quad \text{and} \\ D(x, y + h/2) &= D(x + h, y + h/2), \end{aligned} \tag{13}$$

whenever (x, y), $(x+h, y)$, $(x+h, y+h)$, *and* $(x, y+h)$ *are all interior points of* R_h.

Proof. By direct comparison of the entries of HV and VH, (in the notation of Fig. 1), the equality $HV = VH$ implies

$$D_1 D_2 = D_3 D_4, \quad \text{and} \quad D_1 D_4 = D_2 D_3. \tag{14}$$

By hypothesis, all D_i's are positive; hence, we can divide the two equations (12) to get $D_2/D_4 = D_4/D_2$, or $D_2^2 = D_4^2$. Since D_2 and D_4 are positive, it follows that $D_2 = D_4$.

Similarly $D_1 = D_3$, which proves (11).

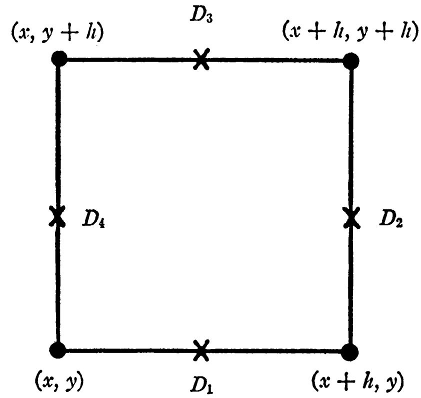

FIG. 1

LEMMA 2. *If $HV = VH$, and 3 mesh-points of a square of side h are interior mesh-points of R_h, then so is the fourth.*

Proof. For interior mesh points, H and V are defined by (1)–(1′) and (2)–(2′), with a, b, c, α, β, γ all positive. Now suppose that $(x, y+h)$, $(x+h, y+h)$, and $(x+h, y)$ are interior mesh-points of R_h, but that (x, y) is a boundary mesh-point. Then

$$[uHV](x, y + h) = a\gamma u(x + h, y) + \cdots,$$

where the dotted terms involve no contribution from $u(x+h, y)$. On the other hand, $[uVH](x, y+h)$ has no contribution from $u(x+h, y)$ whatsoever. Since $\alpha\gamma > 0$, this contradicts the hypothesis $HV = VH$. Similar arguments, involving $a\alpha$, $c\alpha$, and $c\gamma$, take care of the remaining three cases.

COROLLARY. *The relation $HV = VH$ implies that R_h is (if connected) a rectangle, with sides parallel to the coordinate axes*(6).

Proof. By induction, it is easy to show that any connected domain R_h with the "fourth mesh-point closure" property of Lemma 2, is a rectangle of the type specified.

THEOREM 2. *The relation $HV = VH$ holds in R_h if and only if R_h is a rectangle and*

$$D(x + h/2, y) = f(x + h/2) \quad \text{and} \quad D(x, y + h/2) = g(y + h/2).$$

Proof. If $HV = VH$, then R_h is a rectangle by the preceding corollary.

(6) This fact seems to have been overlooked in the discussion of nonrectangular domains in [4, p. 431, Equation (7.5)]. The statement "As A and B commute . . . " of [8, p. 410] is also incorrect.

Hence $D(x+h/2, y)$ is independent of y by repeated application of the first identity of (13). Similarly, $D(x, y+h/2)$ is independent of x.

The converse now follows by direct calculation (see [7]).

Consideration of the relations $\Sigma H = H\Sigma$ and $\Sigma V = V\Sigma$ leads to similar conclusions. We can prove

LEMMA 3. *If R_h is connected, then $\Sigma H = H\Sigma$ and $\Sigma V = V\Sigma$ if and only if Σ has the same value at all mesh-points.*

Proof. Letting Σ_i denote the ith diagonal entry of Σ, and h_{ij} the (i, j)-entry of H, clearly $\Sigma H = H\Sigma$ if and only if $\Sigma_i h_{ij} = h_{ij}\Sigma_j$ for all i, j, whence $\Sigma_i = \Sigma_j$ if $h_{ij} \neq 0$. By the transitivity of equality, this is the condition that $\Sigma(x, y)$ be constant on all interior mesh-points of a connected horizontal *row*. —Similarly, $\Sigma V = V\Sigma$ if and only if $\Sigma(x, y)$ is constant on the mesh-points of a connected *column*. The conclusion is now obvious.

6. **Applications.** The application of the preceding results to the implicit alternating direction methods for solving Dirichlet-type elliptic difference equations is quite straightforward. We first consider limiting conditions for the permutability of H, V, and Σ as $h \downarrow 0$. To deduce these, we remark the following.

LEMMA 4. *Let $D(x, y)$ be any piecewise continuous function, defined on a connected domain R. Then either $D(x, y)$ is constant or, for all sufficiently small $h > 0$, either the relation $D(x+h/2, y) = f(x+h/2)$ or $D(x, y+h/2) = g(y+h/2)$ fails on the interior mesh-points of R_h.*

For any two points in the same domain of continuity, we could otherwise find a sequence of $h_n \downarrow 0$, and corresponding horizontal-vertical paths having a fixed number of corner jumps of length h_n, such that $D(x, y)$ was constant on the straight segments—and hence changed by any arbitrarily small total amount. Therefore $D(x, y)$ would have to be piecewise constant. Finally, it is easy to show that a nonzero jump in $D(x, y)$ between adjacent regions of constancy, must violate one of the two relations (14).

Combining the preceding lemma with Theorem 2, we obtain

THEOREM 3. *If $D(x, y)$ is piecewise continuous, then $HV \neq VH$ for all sufficiently small h, in any connected domain R, except in the case $D(x, y) =$ const. of $D\nabla^2 u - \Sigma u + S = 0$ in a rectangle.*

Combining this with Lemma 4, we get the

COROLLARY. *Except in the case of the Helmholtz equation*

$$\nabla^2 u - \frac{\Sigma}{D} u + S = 0, \qquad \frac{\Sigma}{D} = const., \tag{15}$$

in a rectangle, for all sufficiently small h we must have either $HV \neq VH$, $\Sigma H \neq H\Sigma$, or $\Sigma V \neq V\Sigma$, in all Dirichlet-type problems.

Not only does permutability fail in the general case just described, but complex eigenvalues commonly occur for (15) in nonrectangular regions—as has been shown by Dr. Young(7) at the Ramo-Wooldridge Corp. Hence the convergence estimates of [3] are essentially limited to the case (15).

Similar discussions apply to the Sheldon-Wachspress scheme (7)–(7′) and the Douglas-Rachford process (6)–(8′). For sufficiently small h, the convergence estimates of [4] are limited to the case (15), contrary to the assertion made in [4].

7. **Mixed boundary conditions.** In §§4–6, we have treated only the case of given boundary values (i.e., Dirichlet-type problems). However, mixed boundary conditions, of the type

$$\partial u/\partial n + d(x, y)u = \delta(x, y), \qquad d > 0 \tag{16}$$

are also important. Thus, if $\delta(x, y) = 0$, (14) expresses the usual "extrapolated boundary" condition for the "extrapolation length" d.

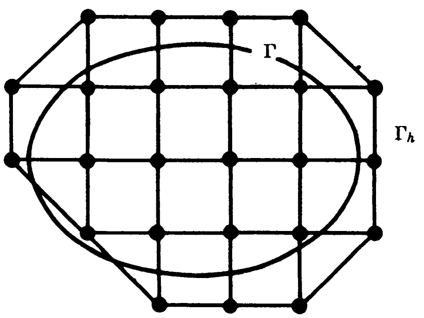

FIG. 2

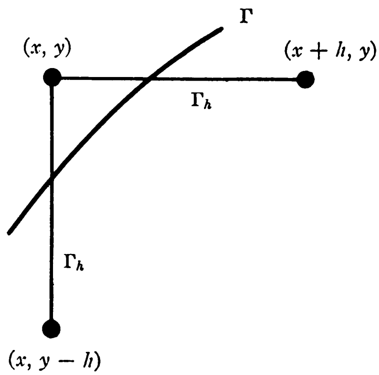

FIG. 3

To treat such mixed boundary conditions, we approximate the region R with boundary Γ, over which (1) and (16) are defined in the continuous problem, by sets R_h and Γ_h of "interior" and "boundary" mesh-points, as in Fig. 2. The values of u at all mesh-points of R_h *and* Γ_h must be taken as unknowns with mixed boundary conditions—whereas only the values in R_h are unknown in Dirichlet-type problems. Relative to this discretization, and the approximation (2)–(2′), we adopt a slight modification of the usual(8) approximation to (16) on Γ_h. Thus we take for the case depicted(9) in Fig. 3,

(7) The results of Dr. Young will be reported in a later joint paper. Mr. G. Bilodeau of the Westinghouse Corp. had already verified the possibility of $HV \neq VH$ in December, 1956.

(8) L. Fox, Quarterly of Applied Math. vol. 2 (1944) pp. 251–257; E. Batschelet, Z. Angew. Math. Phys. vol. 3 (1952) pp. 165–193; L. Collatz, *Numerische Behandlung von Differentialgleichungen*, Springer, Berlin, 2d ed., 1955.

(9) The other cases can also be treated in a similar manner.

$$
(17)\qquad
\begin{aligned}
[uH](x, y) &= u(x, y)[D(x + h/2, y) + hd(x, y)D(x, y)] \\
&\qquad - u(x + h, y)D(x + h/2, y), \\
[uV](x, y) &= u(x, y)[D(x, y - h/2) + hd(x, y)D(x, y)] \\
&\qquad - u(x, y - h)D(x, y - h/2).
\end{aligned}
$$

With reference to (2)–(2′), the matrices H and V so defined are symmetric, with positive diagonal elements, and nonpositive off-diagonal elements. Because h and $d(x, y)$ are positive, the off-diagonal elements are diagonally dominated, and thus H, V, $H+\Sigma$, and $V+\Sigma$ are tridiagonal Stieltjes matrices. We further observe that the matrices H and V, defined by the mixed boundary conditions (17)–(17′), have their off-diagonal entries determined in the same way as the off-diagonal entries of the matrices H and V of (2)–(2′). Because of this, Lemma 1 applies to the case of mixed boundary conditions. The analog of Lemma 2 is

LEMMA 5. *If $HV = VH$, and 3 mesh-points of a square of side h are mesh-points for unknowns, then so is the fourth.*

It is easily seen that the proof of Lemma 2 applies without change to Lemma 5. It thus results that the corollary to Lemma 2, as well as Theorem 2, and Lemma 3, remain valid for the boundary conditions of (16).

LEMMA 6. *If $HV = VH$, and R_h is connected, then $d(x, y)D(x, y) = \text{const.}$ at all mesh-points.*

Proof. Using Fig. 3, and the definitions of (17)–(17′), we have:

$$
\begin{aligned}
[uHV](x, y + h) &= - D_4(D_3 + hd(x, y + h)D(x, y + h))u(x, y) + \cdots, \\
[uVH](x, y + h) &= - D_4(D_1 + hd(x, y)D(x, y))u(x, y) + \cdots.
\end{aligned}
$$

Since the coefficients of $u(x, y)$ must be equal, assuming $HV = VH$, then we conclude that

$$
(18)\qquad d(x, y + h)D(x, y + h) = d(x, y)D(x, y),
$$

since $D_1 = D_3$ by Lemma 1. Since this is true for all adjacent pairs of mesh-points, and since R_h is connected, then $d(x, y)D(x, y)$ is constant at all mesh-points.

Combining the above lemmas, we obtain

THEOREM 4. *Subject to the boundary conditions* (16), *if $D(x, y)$ is piecewise continuous, then, except in the case $D(x, y) = \text{const.}$, $d(x, y) = \text{const.}$ of $D\nabla^2 u - \Sigma u + S = 0$ in a rectangle, $HV \neq VH$ for all sufficiently small h, in any connected domain R.*

8. **Spectra of S_ρ, T_ρ, U_ρ.** The preceding negative results show that available comparisons of the rates of convergence for SOR and IAD methods

([3, p. 30] and [4, p. 436]), which depend strongly on the reality of the eigenvalues of the matrices involved, must be essentially modified for nonrectangular domains. Nevertheless, indications are that IAD can be made to converge more rapidly for most problems. Even if just one ρ is used, the comparison is surprisingly favorable, as we now show(10).

THEOREM 5. *If $\rho>0$, and Σ is constant, then the spectral radius of T_ρ is less than one. In any case, the spectral radius of S_ρ is less than one.*

Proof. We use the fact that H, V, etc. are symmetric and positive definite. Defining

$$\begin{aligned} \tilde{T}_\rho &= (V + \rho I + \Sigma)^{-1} T_\rho (V + \rho I + \Sigma), \\ &= [(V + \rho I + \Sigma)^{-1}(\rho I - V)][(H + \rho I + \Sigma)^{-1}(\rho I - H)], \end{aligned} \tag{20}$$

we see that $\tilde{T}_\rho$, which has the same spectral radius as T_ρ, is the product of two transformations $\tilde{V}_\rho$ and $\tilde{H}_\rho$ defined by symmetric matrices, each of which has Euclidean norm less than unity if Σ is a constant, by the formulas of §4. Hence, so does their product $\tilde{T}_\rho$, whence $\tilde{T}_\rho$ (and thus T_ρ) have spectral radii less than unity; $HV=VH$ need not be assumed.

As regards S_ρ, we need only repeat the above argument, replacing H by $H+\Sigma/2$, V by $V+\Sigma/2$, and Σ by 0.

THEOREM 6. *If $\rho>0$, and $\Sigma=0$, then the spectrum of U_ρ lies in the circle $|\lambda-1/2|<1/2$.*

Proof. Clearly, $U_\rho=(I+\hat{U}_\rho)/2$, where

$$\hat{U}_\rho = (\rho I - V)(\rho I - H)(\rho I + H)^{-1}(\rho I + V)^{-1}. \tag{21}$$

Also, $\hat{U}_\rho$ is similar to

$$\tilde{U}_\rho = [(\rho I + V)^{-1}(\rho I - V)][(\rho I - H)(\rho I + H)^{-1}] = \tilde{T}_\rho, \tag{21'}$$

since the two terms in the second square bracket are permutable. The conclusion follows, and shows the sharper

COROLLARY(11). *If $\Sigma=0$, the eigenvalues λ_k of U_ρ are obtained from those μ_k of T_ρ, by the transformation $\lambda_k=(1+\mu_k)/2$.*

When $HV\neq VH$ (e.g., for $-\nabla^2 u=S$ in nonrectangular domains), we cannot even prove that all products $S_\rho S_{\rho'}$ (ρ, $\rho'>0$) have spectral radii less than one. However, under the hypotheses of Theorem 5, if $\rho_1\geqq\rho_2\geqq\cdots\geqq\rho_n$ and $(\rho_1-\rho_n)$ does not exceed any eigenvalue of V in magnitude, then we can prove that $S_{\rho_1}\cdots S_{\rho_n}$ has spectral radius less than one.

(10) The proof of Theorem 5 is essentially due to Sheldon and Wachspress [6].

(11) This has been observed, when $HV=VH$, by D. R. Peaceman (oral communication).

9. **Other positive results.** For the Stieltjes matrices H and V, we now define

$$\Lambda_0 = \max_j \{h_{jj}, v_{jj}\}. \tag{22}$$

Thus, for the Poisson equation $-\nabla^2 u = S$ in the plane, using (2)–(2′), evidently $\Lambda_0 = 2$.

THEOREM 7. *If $\rho > \Lambda_0$, then T_ρ is a non-negative matrix.*

Proof. Since H and V are Stieltjes matrices, then for any $\rho > 0$, $(H+\rho I+\Sigma)^{-1}$ and $(V+\rho I+\Sigma)^{-1}$ are non-negative matrices. By definition, for $\rho > \Lambda_0$, both $(-H+\rho I)$ and $(-V+\rho I)$ are non-negative matrices, with positive diagonal entries. It is therefore clear from (6*) that T_ρ is non-negative.

A non-negative matrix T is *primitive* [5, p. 606], if and only if some power of T is positive. It is also known(12) that any non-negative irreducible matrix with positive diagonal entries is primitive. The factors of (6*) all have positive diagonal entries if $\rho > \Lambda_0$. If the domain R_h consisting of interior mesh-points is *connected*, so that $A = H + V$ is irreducible ("indecomposable" [5, p. 598]), it follows that T_ρ is primitive. As a special case, there follows the

COROLLARY. *For $-\nabla^2 u = S$ in a convex plane domain, if $\rho > 2$, T_ρ is a positive matrix.*

In the preceding theorem, $\rho > \Lambda_0$ corresponds to *underrelaxation*—i.e., to $\omega = <1$ in the sense of [1]. For, if A is *any* irreducible Stieltjes matrix, and L_ω is the corresponding overrelaxation matrix, as derived in [1], then for $0 < \omega < 1$, L_ω is non-negative and primitive, as is T_ρ for $\rho > 2$.

Again, let us define Λ_1 by the formula

$$\Lambda_1 = \max_{k,l} \{\sigma_k, \tau_l\} > 0, \tag{23}$$

where σ_k and τ_l are the eigenvalues of H and V. Thus, in the case of the Poisson equation $-\nabla^2 u = S$ in any plane domain, $\Lambda_1 \leqq 4$ (as may be shown using Gerschgorin's lemma).

THEOREM 8. *If $\Sigma =$ const. and $\rho > \Lambda_1$, then all eigenvalues of T_ρ are real and positive.*

Proof. Let $\tilde{T}_\rho$ be as in (20). Since Σ is const., each factor in square brackets is symmetric and, for $\rho > \Lambda_1$, positive definite. Thus $\tilde{T}_\rho = \tilde{V}_\rho \tilde{H}_\rho$, being similar to the positive definite symmetric matrix $\tilde{V}_\rho^{1/2} \tilde{H}_\rho \tilde{V}_\rho^{1/2}$, has all positive eigenvalues.

(12) H. Wielandt, Math. Z. vol. 52 (1950) p. 544.

Bibliography

1. David Young, *Iterative methods for solving partial difference equations of elliptic type,* Trans. Amer. Math. Soc. vol. 76 (1954) pp. 92–111.

2. S. Frankel, *Convergence rates of iterative treatments of partial differential equations,* Math. Tables and Other Aids to Computation vol. 4 (1950) pp. 65–75.

3. D. W. Peaceman and H. H. Rachford, Jr., *The numerical solution of parabolic and elliptic differential equations,* J. Soc. Indust. Appl. Math. vol. 3 (1955) pp. 28–41.

4. J. Douglas, Jr. and H. H. Rachford, Jr., *On the numerical solution of heat conduction problems in two and three space variables,* Trans. Amer. Math. Soc. vol. 82 (1956) pp. 421–439.

5. Gerard Debreu and I. N. Herstein, *Nonnegative square matrices,* Econometrica vol. 21 (1953) pp. 597–607.

6. E. L. Wachspress, *CURE: A generalized multigroup coding for the IBM*-704, Report KAPL-1724 of the Knolls Atomic Power Laboratory, August, 1957.

7. G. Birkhoff and R. S. Varga, *Implicit alternating direction methods,* Report WAPD-T-650 of Westinghouse Electric Corp., October, 1957.

8. Jim Douglas, Jr., *A note on the alternating direction implicit method,* Proc. Amer. Math. Soc. vol. 8 (1957) pp. 409–411.

Harvard University,
Cambridge, Mass.
Westinghouse, Bettis Plant,
Pittsburgh, Pa.