Numerische Mathematik 6, 106—119 (1964)

Block Iterative Methods for Cyclically Reduced Matrix Equations*

By
LOUIS A. HAGEMAN and RICHARD S. VARGA

* This paper includes results from the first author's Ph.D. thesis [7] at the University of Pittsburgh.

§ 1. Introduction

The purpose of this paper is to investigate block iterative methods for cyclically reduced matrix equations. To first describe the basis of the cyclic reduction methods, consider the partitioned system of n linear equations in n unknowns

$$\begin{bmatrix} I_{1,1} & -B_{1,2} \\ -B_{2,1} & I_{2,2} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}, \tag{1.1}$$

where $I_{1,1}$ and $I_{2,2}$ are respectively $r \times r$ and $(n-r) \times (n-r)$ identity matrices, $1 \leqq r < n$. We assume that the $n \times n$ matrix of (1.1) is non-singular. Since the $n \times n$ matrix of (1.1) can be written as $I - B$ where

$$B = \begin{bmatrix} 0 & B_{1,2} \\ B_{2,1} & 0 \end{bmatrix}, \tag{1.2}$$

then B is *weakly cyclic of index 2* [*15*, p. 39], and thus if μ is a non-zero eigenvalue of B, then so is $-\mu$ [*15*, p. 40]. Hence, the non-singularity of $I - B$ implies the non-singularity of $I + B$. Multiplying (1.1) on the left by the non-singular matrix $I + B$ yields, in partitioned form, the matrix equation

$$\left[\begin{array}{c|c} I_{1,1} - B_{1,2} B_{2,1} & 0 \\ \hline 0 & I_{2,2} - B_{2,1} B_{1,2} \end{array}\right] \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} k_1 + B_{1,2} k_2 \\ k_2 + B_{2,1} k_1 \end{bmatrix}, \tag{1.3}$$

and we see that the matrix of (1.3) is *completely reducible* [*5*; *15*, p. 43], i.e., (1.3) can be written as the pair of *uncoupled* matrix equations

$$(I_{1,1} - B_{1,2} B_{2,1})\, x_1 = k_1 + B_{1,2} k_2, \tag{1.4}$$

$$(I_{2,2} - B_{2,1} B_{1,2})\, x_2 = k_2 + B_{2,1} k_1, \tag{1.5}$$

in the unknowns x_1 and x_2, where the matrices $(I_{1,1} - B_{1,2} B_{2,1})$ and $(I_{2,2} - B_{2,1} B_{1,2})$ are non-singular. Hence, if we can solve for (or approximate closely) the unique vector x_1, in the lower order matrix equation of (1.4), then x_2 can be *explicitly formed from* (1.1) by means of

$$x_2 = B_{2,1} x_1 + k_2, \tag{1.6}$$

thus eliminating the need to solve *two* lower order matrix equations. We have thus *reduced* the solution of the matrix equation (1.1) of order n to the solution of a single lower order matrix equation (1.4) of order r, $1 \leqq r < n$. For the numerical solution of second-order elliptic partial differential equations with large numbers of mesh points, typical of such problems arising from practical settings, the order of the reduced matrix equation of (1.4) is approximately half that of (1.1); consequently, such reductions can be important in practical applications.

There has been a great deal of effort expended in making efficient the iterative solution of the original matrix equation (1.1). The basic successive overrelaxation iterative method of YOUNG [*17*] and FRANKEL [*4*] and its newer variant, the cyclic Chebyshev semi-iterative method [*6*], have both been theoretically and practically coupled successfully with the use of block or multi-line techniques, considered in [*1, 3, 8, 11, 13, 16*]. The first purpose of this paper is to theoretically couple the use of these block or multi-line techniques to the iterative solution of the cyclically reduced matrix equation (1.4), and to determine conditions which insure that these block techniques, applied to the cyclically reduced matrix problem of (1.4), are asymptotically iteratively faster than the corresponding block techniques applied to the original matrix problem of (1.1). The second purpose of this paper is to present the results of numerical experiments which indicate the practicality and utility of these newer iterative methods in solving second order elliptic difference equations.

The basic idea for this reduction of the order of the matrix equation of (1.1) to that of (1.4) might be attributed to FROBENIUS [*5*], who in essence proved that any weakly cyclic matrix B of index $p \geqq 2$ is such that B^p is completely reducible. (See also [*15*, p. 43].) SCHRÖDER [*14*] was apparently the first to couple the weakly cyclic nature of B with the complete reducibility of B^2 for Laplace-type elliptic difference equations. While SCHRÖDER suggested that cyclic reduction might improve rates of convergence for the point JACOBI and point Gauß-Seidel iterative methods, this was actually proved in [*15*, p. 157] for the matrix B non-negative, symmetric, and convergent, by applying the regular splitting theory of matrices [*16*]. We shall first extend this regular splitting theory so as to be able to apply it in the comparison of block iterative techniques applied to (1.1) and (1.4).

§ 2. Induced Regular Splittings for Cyclically Reduced Matrices

We begin by recalling [*15*, p. 87] that $M - N$ is a *splitting* of the $n \times n$ complex matrix A into two $n \times n$ matrices M and N if $A = M - N$ and M is non-singular. This splitting of the matrix A defines a natural iterative method

$$M x^{(n+1)} = N x^{(n)} + k, \quad n \geqq 0,$$

or equivalently

$$x^{(n+1)} = M^{-1} N x^{(n)} + M^{-1} k, \quad n \geqq 0, \tag{2.1}$$

for the iterative solution of the matrix problem

$$A x = k. \tag{2.2}$$

The process (2.1) is convergent [9] for any initial vector $x^{(0)}$ if and only if the spectral radius* of the iteration matrix $M^{-1}N$ is less than unity: $\varrho(M^{-1}N)<1$. If the splitting $M-N$ is such that the matrices N and M^{-1} have only non-negative real entries (written $N\geqq 0$, $M^{-1}\geqq 0$), then $M-N$ is called a *regular splitting* of A [*16*, *15*, p. 88]. While an arbitrary splitting $M-N$ of A does not in general give rise to a convergent iterative method (2.1), it has been shown [*16*, *15*, p. 89] that regular splittings on the other hand do if $A^{-1}\geqq 0$. We formally state this result as

Theorem A. If $M-N$ is a regular splitting of the $n\times n$ matrix A, and $A^{-1}\geqq 0$, then

$$\varrho(M^{-1}N)=\frac{\varrho(A^{-1}N)}{1+\varrho(A^{-1}N)}<1. \tag{2.3}$$

The comparison of spectral radii of iteration matrices arising from different regular splittings can be made by means of the following known result [*16*; *15*, p. 90]:

Theorem B. Let M_1-N_1 and M_2-N_2 be two regular splittings of A, where A^{-1} has only positive real entries (written $A^{-1}>0$). If $0\leqq N_1\leqq N_2$, equality excluded**, then

$$0<\varrho(M_1^{-1}N_1)<\varrho(M_2^{-1}N_2)<1. \tag{2.4}$$

The above result has been used [*16*] to compare particular block (or multi-line) iterative methods applied to a *fixed* matrix equation. Now, we seek to compare a particular block iterative method applied to the matrix equation of order n

$$\begin{bmatrix} I_{1,1} & -B_{1,2} \\ -B_{2,1} & I_{2,2} \end{bmatrix}\cdot\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} k_1 \\ k_2 \end{bmatrix} \tag{2.5}$$

with the *equivalent* or *induced* block iterative technique applied to its cyclically reduced matrix equation (1.4) of order r:

$$(I_{1,1}-B_{1,2}B_{2,1})\,x_1=k_1+B_{1,2}k_2. \tag{2.6}$$

We shall make this comparison by means of extensions of Theorems A and B.

Let $M-N$ be a regular splitting of the $n\times n$ matrix $A=I-B$ of (2.5), where the $n\times n$ matrices B and N have the weakly cyclic of index 2 form:

$$B=\begin{bmatrix} 0 & B_{1,2} \\ B_{2,1} & 0 \end{bmatrix},\qquad N=\begin{bmatrix} 0 & N_{1,2} \\ N_{2,1} & 0 \end{bmatrix}. \tag{2.7}$$

We assume that $N\leqq B$ and that B is irreducible*** and convergent, i.e., $\varrho(B)<1$.

---

* The spectral radius $\varrho(B)$ of an arbitrary $n\times n$ matrix B with eigenvalues λ_i is defined as $\varrho(B)=\max_{1\leqq i\leqq n}|\lambda_i|$.

** By this we mean that neither N_1 nor N_2-N_1 is the null matrix.

*** An $n\times n$ matrix S is *irreducible* [*15*, p. 20] if there exists no $n\times n$ permuation matrix P such that $PSP^T=\begin{bmatrix} S_{1,1} & S_{1,2} \\ 0 & S_{2,2} \end{bmatrix}$, where the diagonal submatrices are square and non-void.

We now define the $r \times r$ matrices $\hat{A}$, $\hat{M}$, and $\hat{N}$ as

$$\begin{aligned} \hat{M} &\equiv I_{1,1} - (B_{1,2} - N_{1,2})(B_{2,1} - N_{2,1}), \\ \hat{N} &\equiv B_{1,2} B_{2,1} - (B_{1,2} - N_{1,2})(B_{2,1} - N_{2,1}) = B_{1,2} N_{2,1} + N_{1,2} B_{2,1} - N_{1,2} N_{2,1}, \\ \hat{A} &\equiv \hat{M} - \hat{N} = I_{1,1} - B_{1,2} B_{2,1}. \end{aligned} \tag{2.8}$$

Note that the matrix $M = A + N = I - (B - N)$ is necessarily non-singular by hypothesis, and since $B - N$ is of the weakly cyclic of index 2 form (2.7), it follows that $\hat{M}$ is also non-singular. Thus, $\hat{M} - \hat{N}$ is at least a splitting of the $r \times r$ matrix $\hat{A}$. A stronger result, analogous to Theorem A, is given in

Theorem 1. Let $M - N$ be a regular splitting of the $n \times n$ matrix $A = I - B$, where the matrices B and N have the form (2.7), and B is irreducible and convergent. If $N \leqq B$, then $\hat{M} - \hat{N}$ is a regular splitting of the $r \times r$ matrix $\hat{A} = I_{1,1} - B_{1,2} B_{2,1}$, where $\hat{M}$ and $\hat{N}$ are defined in (2.8), and

$$\varrho(\hat{M}^{-1} \hat{N}) \leqq \varrho(M^{-1} N) < 1. \tag{2.9}$$

If, moreover, $N_{1,2} N_{2,1}$ has at least one positive entry, then

$$\varrho(\hat{M}^{-1} \hat{N}) < \varrho(M^{-1} N) < 1. \tag{2.10}$$

Proof. Using a result of FROBENIUS (see [15, p. 43]), the non-negative, irreducible, and convergent character of B enables us to state that the $r \times r$ matrix $B_{1,2} B_{2,1}$ is also non-negative, irreducible and convergent, and thus $\varrho(B_{1,2} B_{2,1}) = \varrho^2(B) < 1$. Next, since $N \leqq B$ by hypothesis, then

$$0 \leqq (B_{1,2} - N_{1,2})(B_{2,1} - N_{2,1}) \leqq B_{1,2} B_{2,1},$$

which shows that the $r \times r$ matrix $\hat{N}$ of (2.8) is non-negative. From the inequalities above, it follows from the Perron-Frobenius theory of non-negative matrices that $\varrho\{(B_{1,2} - N_{1,2})(B_{2,1} - N_{2,1})\} \leqq \varrho(B_{1,2} B_{2,1}) < 1$, and from this, we conclude [15, p. 83] that $\hat{M}^{-1} \geqq 0$. Thus, $\hat{M} - \hat{N}$ is a regular splitting of the $r \times r$ matrix $\hat{A}$, which proves the first part of the theorem.

We now show that the inequality of (2.10) is valid if $N_{1,2} N_{2,1}$ has at least one positive entry; from this, the inequality of (2.9) follows from a continuity argument. We first point out that a strengthened form of the proof above showing that $\hat{M}^{-1} \geqq 0$ also shows that $A^{-1} > 0$ and $\hat{A}^{-1} > 0$ [15, p. 84]. Hence, from the strict monotonicity of $x/(1+x)$ for $x \geqq 0$ and the expression (2.3) of Theorem A, it follows that

$$\varrho(\hat{M}^{-1} \hat{N}) < \varrho(M^{-1} N) \Leftrightarrow \varrho(\hat{A}^{-1} \hat{N}) < \varrho(A^{-1} N). \tag{2.11}$$

Consider now the $n \times n$ matrix $\frac{N}{\lambda} + B$, and the $r \times r$ matrix $\frac{\hat{N}}{\lambda} + B_{1,2} B_{2,1}$ as a function of $\lambda > 0$. Since B and $B_{1,2} B_{2,1}$ are non-negative and irreducible, and $N \geqq 0$ with $N_{1,2} N_{2,1}$ not the null matrix, then the spectral radii $\varrho\left\{\frac{N}{\lambda} + B\right\}$ and $\varrho\left\{\frac{\hat{N}}{\lambda} + B_{1,2} B_{2,1}\right\}$ of these matrices are *strictly decreasing* functions [2] of $\lambda > 0$. As

is easily verified [7], there exists positive numbers μ and ν such that

$$\varrho\left\{\frac{N}{\mu}+B\right\}=1 \quad \text{and} \quad \mu=\varrho(A^{-1}N)>0, \tag{2.12}$$

$$\varrho\left\{\frac{\hat{N}}{\nu}+B_{1,2}\,B_{2,1}\right\}=1 \quad \text{and} \quad \nu=\varrho(\hat{A}_1^{-1}\hat{N}_1)>0. \tag{2.13}$$

Now, since the $n\times n$ matrices B and N both have the form (2.7), so does the matrix $B+\frac{N}{\lambda}$. For any $\lambda>0$, we verify directly from (2.8) that

$$\begin{aligned}\varrho^2\left\{\left(B+\frac{N}{\lambda}\right)\right\}&=\varrho\left\{B_{1,2}\,B_{2,1}+\frac{\hat{N}}{\lambda}+\left(\frac{1+\lambda}{\lambda^2}\right)N_{1,2}\,N_{2,1}\right\}\\&>\varrho\left\{B_{1,2}\,B_{2,1}+\frac{1}{\lambda}\,\hat{N}\right\}.\end{aligned} \tag{2.14}$$

The last inequality in (2.14) follows because $N_{1,2}\,N_{2,1}$ is a non-negative matrix with at least one positive entry. Hence, if we let $\lambda=\mu$ of (2.12), the above inequality yields

$$1>\varrho\left\{B_{1,2}\,B_{2,1}+\frac{\hat{N}}{\mu}\right\}.$$

The strictly decreasing nature of $\varrho\left\{B_{1,2}\,B_{2,1}+\frac{\hat{N}}{\lambda}\right\}$ as a function of $\lambda>0$, coupled with (2.13), finally gives us that $\nu=\varrho(\hat{A}^{-1}\hat{N})<\varrho(A^{-1}N)=\mu$, which from (2.11) is equivalent to (2.10), completing the proof.

With the definitions of (2.8), we shall call $\hat{M}-\hat{N}$ the *induced regular splitting* of $\hat{A}$ derived from the regular splitting $M-N$ of A. These definitions of induced splittings of $\hat{A}$ were chosen so that ℓ-line iterative methods arising from the 5-point finite difference approximations to Laplace-type differential equations induce ℓ-line iterative methods for the reduced matrix equation (2.6). This will be described in more detail in § 4.

In analogy to Theorem B, we now compare the spectral radii of *two* induced regular splittings of $\hat{A}$.

Theorem 2. Let M_1-N_1 and M_2-N_2 be two regular splittings of the $n\times n$ matrix $A=I-B$, where the matrices B, N_1, and N_2 have the form (2.7), and B is irreducible and convergent. If $0\leqq N_1\leqq N_2\leqq B$, and $\hat{M}_1-\hat{N}_1$ and $\hat{M}_2-\hat{N}_2$ are the induced regular splittings of the $r\times r$ matrix $\hat{A}$, then

$$0\leqq\varrho(\hat{M}_1^{-1}\hat{N}_1)\leqq\varrho(\hat{M}_2^{-1}\hat{N}_2)<1. \tag{2.15}$$

If, moreover, $0\leqq\hat{N}_1\leqq\hat{N}_2$, equality excluded, then

$$0<\varrho(\hat{M}_1^{-1}\hat{N}_1)<\varrho(\hat{M}_2^{-1}\hat{N}_2)<1. \tag{2.16}$$

Proof. From the proof of Theorem 1, we have that $\hat{A}^{-1}>0$. Thus, the inequalities of (2.16) follow immediately from Theorem B, while those of (2.15) follow from (2.16) by an obvious continuity argument.

In the spirit of SCHRÖDER'S original observation [*14*] and its subsequent development in [*15*, p. 157], we can now define *new* splittings of the reduced $r\times r$

matrix $\hat{A}$ which lead to improved iterative methods for the cyclically reduced matrix problem of (2.6). If we let $\hat{K}$ be the non-negative diagonal $r \times r$ matrix defined by

$$\hat{K} = \text{diag}\,(B_{1,2}\, B_{2,1}) - \text{diag}\,\{(B_{1,2} - N_{1,2})\,(B_{2,1} - N_{2,1})\}, \tag{2.17}$$

then, in analogy to the definitions of (2.8), define

$$\tilde{M} = \hat{M} - \hat{K}, \quad \tilde{N} = \hat{N} - \hat{K}. \tag{2.18}$$

It is not difficult to verify, following the proof of Theorem 1, that $\tilde{M} - \tilde{N}$ is a regular splitting of the reduced $r \times r$ matrix $\hat{A}$. But, since $\hat{M} - \hat{N}$ is also a regular splitting of $\hat{A}$, and $\tilde{N} \leqq \hat{N}$ *by construction*, we immediately have from Theorem B the following corollary of Theorem 1.

Corollary 1. Let $M - N$ be a regular splitting of the $n \times n$ matrix $A = I - B$, where the matrices B and N have the form (2.7), and B is irreducible and convergent. If $N \leqq B$, then

$$\varrho\,(\tilde{M}^{-1}\tilde{N}) \leqq \varrho\,(\hat{M}^{-1}\hat{N}) < 1\,. \tag{2.19}$$

If morever, the matrix $\hat{K}$ of (2.18) has at least one diagonal entry positive, then

$$\varrho\,(\tilde{M}^{-1}\tilde{N}) < \varrho\,(\hat{M}^{-1}\hat{N}) < 1\,. \tag{2.20}$$

For additional results which, in particular, compare the Gauß-Seidel method applied to the original matrix problem (1.1) with various iterative methods applied to the cyclically reduced matrix problem (2.6), the reader is referred to [7].

In the practical applications of § 4, *only* the improved induced splittings $\tilde{M} - \tilde{N}$ of $\hat{A}$ will be discussed, since they give rise to iterative methods which are computationally more efficient as well as more rapidly convergent. For brevity, these splittings will henceforth be called simply *induced regular splittings*.

§ 3. The Block Successive Overrelaxation Iterative Method

Let the matrix problem $A\,x = k$ be written in the partitioned form

$$\begin{bmatrix} A_{1,1} & A_{1,2} \cdots A_{1,m} \\ A_{2,1} & A_{2,2} \cdots A_{2,m} \\ \vdots & \ddots \quad \vdots \\ A_{m,1} & A_{m,2} \cdots A_{m,m} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} = \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_m \end{bmatrix}, \tag{3.1}$$

where the diagonal submatrices $A_{j,j}$ are square and non-singular. If we let

$$D \equiv \begin{bmatrix} A_{1,1} & 0 & \cdots & 0 \\ 0 & A_{2,2} \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \cdots A_{m,m} \end{bmatrix}, \quad \text{and} \quad L \equiv -\begin{bmatrix} 0 & 0 & \cdots 0 \\ A_{2,1} & 0 & 0 \\ \vdots & \ddots & \vdots \\ A_{m,1} & A_{m,2} \cdots & 0 \end{bmatrix},$$

then the matrix A can be expressed as

$$A = D - (L + U), \tag{3.2}$$

where $U \equiv D - L - A$. Because of the non-singularity of the matrices $A_{i,i}$, D is also non-singular, and hence, letting $M = D$ and $N = L + U$, we see that (3.2) is a splitting of A.

The *block successive overrelaxation method* [1] associated with the decomposition (3.2) of A is then defined to be

$$Dx^{(n+1)} = \omega\{Lx^{(n+1)} + Ux^{(n)} + k\} + (1-\omega)\, Dx^{(n)}, \tag{3.3}$$

where the parameter ω is called the *relaxation factor*. This equation can also be written as

$$x^{(n+1)} = \mathfrak{L}_\omega x^{(n)} + \omega (D - \omega L)^{-1} k, \tag{3.4}$$

where $\mathfrak{L}_\omega \equiv (D - \omega L)^{-1} \{\omega U + (1-\omega) D\}$ denotes the *block successive overrelaxation iteration matrix*.

In what follows, we shall employ [17] the *asymptotic rate of convergence* $R_\infty(\mathfrak{L}_\omega)$,

$$R_\infty(\mathfrak{L}_\omega) \equiv -\ln \varrho(\mathfrak{L}_\omega),$$

as a measure of the effectiveness* of the iterative process (3.4). Upon defining the *block Jacobi matrix* J associated with the decomposition (3.2) of A as

$$J = D^{-1}(L + U), \tag{3.5}$$

we obtain the following well-known relationship [1; 15, p. 111; 17] between $R_\infty(\mathfrak{L}_\omega)$ and $\varrho(J)$.

Theorem C. Let A be a 2-cyclic matrix which is consistently ordered [15, p. 101]. If the eigenvalues of the block Jacobi matrix J are real and less than unity in modulus and if

$$\omega_b = \frac{2}{1 + \sqrt{1 - \varrho^2(J)}}, \tag{3.6}$$

then

$$\begin{aligned} &R_\infty(\mathfrak{L}_{\omega_b}) > R_\infty(\mathfrak{L}_\omega) \quad \text{if} \quad \omega \neq \omega_b; \\ &R_\infty(\mathfrak{L}_{\omega_b}) = -\ln(\omega_b - 1). \end{aligned} \tag{3.7}$$

Moreover, as $\varrho(J) \to 1-$, then

$$R_\infty(\mathfrak{L}_{\omega_b}) \sim 2\sqrt{2}\,[R_\infty(J)]^{\frac{1}{2}}. \tag{3.8}$$

Thus, if the conditions of Theorem C are satisfied, the asymptotic rate of convergence of the block successive overrelaxation method can be determined directly from the spectral radius of the associated block Jacobi matrix. If several block Jacobi splittings $D - (L + U)$ in (3.2), arising from different partitionings of A in (3.1), are all *regular splittings* of A, then the comparison theorems of § 2 can be used to compare the asymptotic convergence rates of the associated block successive overrelaxation methods.

§ 4. Applications of Cyclically Reduced Iterative Methods

We now apply cyclically reduced iterative methods to the numerical solution of the following second-order self-adjoint elliptic partial differential equation in

* If C is a convergent matrix, then the number of iterations required to reduce the initial error of the iterative method $x^{(n+1)} = Cx^{(n)} + g$ by a certain factor is roughly inversely proportional to $R_\infty(C)$. See, for example, [10].

two space variables

$$(4.1) \qquad -\left(P(x,y)\,u_x\right)_x-\left(Q(x,y)\,u_y\right)_y+\sigma(x,y)\,u(x,y)=f(x,y),\ (x,y)\in R,$$

defined in an open, bounded, and connected set R in the plane. For simplicity, we assume Dirichlet boundary conditions on Γ, the boundary of R:

$$(4.2) \qquad u(x,y)=\gamma(x,y), \qquad (x,y)\in\Gamma,$$

and that P, Q, σ, and f are given continuous functions in $\overline{R}$, the closure of R, with

$$(4.3) \qquad P(x,y)>0, \qquad Q(x,y)>0, \qquad \sigma(x,y)\geqq 0, \qquad (x,y)\in\overline{R}.$$

The application of cyclically reduced iterative methods to the numerical solution of (4.1) with more general *mixed* boundary conditions and *discontinuous* coefficients P, Q, σ, and f can similarly be made.

We now impose a non-uniform (in each coordinate direction) spatial mesh on $\overline{R}$. Following the discretization of (4.1) — (4.2) as given in [*15*, p. 183], a five-point finite difference approximation of (4.1) based on integration leads to a system of n linear equations in n unknowns of the form

$$(4.4) \qquad Sw=g,$$

where the $n\times n$ matrix S is a real symmetric irreducible Stieltjes matrix [*15*, p. 187]. It is geometrically evident that the mesh points of the discrete problem can be divided into two sets, one set consisting of *square* (or *black*) mesh points, the other consisting of *circle* (or *red*) mesh points, as shown in Fig. 1 below. The nature of the five-point approximation of (4.1) is such that each unknown w_i of (4.4), corresponding to an approximation of $u(x,y)$ at a mesh point of one set, is coupled to at most four other w_j's, which are approximations to $u(x,y)$ at mesh points of the *other* set. Because of this, there exists orderings for the mesh points of the discrete problem (see, for example, Fig. 1) such that the matrix S of (4.4) can be written in partitioned form as

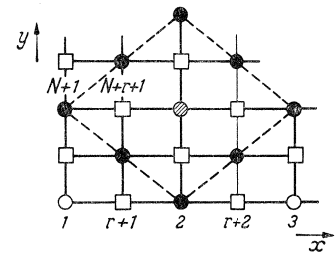


Fig. 1

$$(4.5) \qquad S=\begin{bmatrix} D_1 & -\widetilde{B}_1 \\ -\widetilde{B}_1^T & D_2 \end{bmatrix},$$

where D_1 and D_2 are positive $r\times r$ and $(n-r)\times(n-r)$ diagonal matrices, respectively. We now rewrite the matrix equation (4.4) as

$$(4.6) \qquad \begin{bmatrix} I_{1,1} & -B_1 \\ -B_1^T & I_{2,2} \end{bmatrix}\cdot\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}=\begin{bmatrix} k_1 \\ k_2 \end{bmatrix},$$

or equivalently,

$$(4.7) \qquad Ax\equiv(I-B)\,x=k,$$

where $I_{1,1}$ and $I_{2,2}$ are respectively $r\times r$ and $(n-r)\times(n-r)$ identity matrices, and

$$(4.8) \qquad D_1^{-\frac{1}{2}}\widetilde{B}_1 D_2^{-\frac{1}{2}}=B_1; \qquad x_i=D_i^{\frac{1}{2}}w_i, \qquad k_i=D_i^{-\frac{1}{2}}g_i; \qquad i=1,2.$$

It is clear that the $n \times n$ matrix B of (4.7) is *weakly cyclic of index 2* (cf. (1.2)); moreover, since the matrix S of (4.4) is an irreducible Stieltjes matrix, it follows that B is *convergent*, as well as irreducible and non-negative ($B \geqq 0$). Thus, as in § 1, we can cyclically reduce the $n \times n$ matrix equation (4.7) to the lower order matrix equation

$$(I_{1,1} - B_1 B_1^T)\, x_1 = k_1 + B_1\, x_2, \tag{4.9}$$

where $\hat{A} \equiv I_{1,1} - B_1 B_1^T$ is an $r \times r$ matrix. Because of our use of five point finite difference approximations, it follows that the $r \times r$ matrix $(I_{1,1} - B_1 B_1^T)$ corresponds to a *9-point approximation*. Fig. 1 indicates the typical coupling of a (shaded) circle mesh point to its eight (darkened) circle mesh points. The non-zero entries of the reduced matrix $\hat{A}$ are easily deduced from those of A by graphical considerations. For details, see [7]. Summarizing, we have shown that the discrete five-point finite difference approximations of (4.1) — (4.2) lead to matrix equations of the form (4.7), which can be cyclically reduced to matrix equations of the form (4.9).

We now apply the results of § 2 on induced regular splittings to the matrix equation of (4.9). First, consider the partitioning of the original $n \times n$ Stieltjes matrix S of (4.5) which couples into blocks the mesh points of ℓ successive horizontal mesh lines, for any $\ell \geqq 1$. With (4.8), this partitioning defines in a straightforward way (cf. § 3) an associated block Jacobi splitting $M_\ell - N_\ell$ for the related matrix A of (4.7); moreover, using the Stieltjes property of S, this splitting is necessarily a regular splitting [*15*, p. 90] of A, and the associated iterative method, which we call the *ℓ-line block Jacobi method*

$$M_\ell\, x^{(n+1)} = N_\ell\, x^{(n)} + k \tag{4.10}$$

is necessarily convergent.

From the results of § 2, there is associated with the regular splitting $M_\ell - N_\ell$ of A the *induced regular splitting* $\tilde{M}_\ell - \tilde{N}_\ell$ of the reduced $r \times r$ matrix $\hat{A}$. Moreover, from Theorem 1 and Corollary 1, we have that

$$\varrho\{\tilde{M}_\ell^{-1} \tilde{N}_\ell\} \leqq \varrho\{M_\ell^{-1} N_\ell\} < 1. \tag{4.11}$$

Stemming from the five-point finite difference approximations used to derive the matrix problem of (4.4), it is not difficult to verify (cf. [7]) that the induced regular splitting $\tilde{M}_\ell - \tilde{N}_\ell$ of the $r \times r$ matrix $\hat{A}$ corresponds to a partitioning of the matrix $\hat{A}$ which analogously couples in blocks the *circle* mesh points of ℓ successive horizontal mesh lines. Moreover, the iteration matrix deduced from the regular splitting $\tilde{M}_\ell - \tilde{N}_\ell$ of $\hat{A}$ is again just the ℓ-line block Jacobi matrix of the reduced matrix equation of (4.9). Because of this, we now use the shorter notation ${}^R J^{(\ell)} \equiv (\tilde{M}_\ell)^{-1} \tilde{N}_\ell$ and $J^{(\ell)} \equiv M_\ell^{-1} N_\ell$ to denote the block Jacobi iteration matrices for these related iterative methods. Thus, (4.11) becomes

$$\varrho({}^R J^{(\ell)}) \leqq \varrho(J^{(\ell)}) < 1.$$

Using the stronger form (2.10) of Theorem 1, which is always applicable when the total number of horizontal mesh lines interior to $\overline{R}$ is greater than ℓ, gives us

$$\varrho({}^R J^{(\ell)}) < \varrho(J^{(\ell)}) < 1. \tag{4.12}$$

To obtain a related inequality between the asymptotic rates of convergence of the associated block successive overrelaxation methods, we simply apply Theorem C. However, it is necessary to assume now that $\ell \geqq 2$, since it is easily shown that the matrix $\hat{A}$ is not a 2-cyclic matrix for 1-line horizontal mesh line partitionings. The result then is

$$R_\infty(\mathfrak{L}_{\omega_b}^{(\ell)}) < R_\infty({}^R\mathfrak{L}_{\omega_b}^{(\ell)}), \quad \ell \geqq 2, \tag{4.13}$$

which compares the asymptotic convergence rates of the block successive overrelaxation iterative methods (with optimum relaxation factors) applied both to the matrix equation (4.7) and to its associated reduced matrix equation (4.9). The main conclusions to be drawn here is that the ℓ-line block successive overrelaxation iterative method (with optimum relaxation factor) applied to the cyclically reduced matrix equation (4.9) is asymptotically a *faster* iterative method than the ℓ-line block successive overrelaxation iterative (with optimum relaxation factor) applied to the matrix equation (4.7), *independent* of the particular non-uniform mesh used on $\overline{R}$.

Unfortunately, the above application of the results of § 2, while showing the superiority of accelerated iteration methods for the reduced matrix problem, does not give any estimate of the precise value of the ratio of asymptotic rates of convergence in (4.13) in the general case. If we confine our attention now to the *model problem*, i.e., the solution of the Dirichlet problem in a square domain of side π, with a uniform mesh $h = \Delta x = \Delta y = \frac{\pi}{N}$ in each coordinate direction, the methods of [*16, 7*] based on regular splittings of matrices could be used to estimate their ratio for small h (i.e., $h \to 0$). However, PARTER [*11, 12*], by means of a different analysis, has shown the following *exact* asymptotic behavior:

$$\begin{aligned} &\varrho(J^{(\ell)}) \sim 1 - \ell h^2, \\ &\varrho({}^R J^{(\ell)}) \sim 1 - \tfrac{4}{3}\ell h^2, \quad h \to 0, \quad \ell \geqq 2. \end{aligned} \tag{4.14}$$

Consequently, from (3.8) we obtain

$$R_\infty(\mathfrak{L}_{\omega_b}^{(\ell)}) \sim 2\sqrt{2\ell}\, h, \quad \text{and} \quad R_\infty({}^R\mathfrak{L}_{\omega_b}^{(\ell)}) \sim 2\sqrt{\tfrac{8}{3}\ell}\, h, \quad h \to 0, \tag{4.15}$$

so that (4.13) becomes

$$R_\infty({}^R\mathfrak{L}_{w_b}^{(\ell)}) \sim \sqrt{\tfrac{4}{3}}\, R_\infty(\mathfrak{L}_{\omega_b}^{(\ell)}), \quad h \to 0, \quad \ell \geqq 2. \tag{4.16}$$

Actually, Parter's analysis yields asymptotic results like those of (4.14) for somewhat more general problems than the model problem, but in any event, his results compliment what might be called the discrete results such as (4.13) which hold for still more general problems.

It is also interesting to mention that PARTER [*12*] considers the partitioning of the reduced matrix $\hat{A}$ which couples into blocks the circle mesh points of ℓ successive *diagonal* mesh lines (cf. Fig. 1). This partitioning of $\hat{A}$ is a regular splitting of $\hat{A}$, but it is not an induced regular splitting of $\hat{A}$ in the sense of § 2. For the associated block Jacobi iterative method, PARTER gives the asymptotic result

$$\varrho({}^R J^{(\ell\text{-diag})}) \sim 1 - 2\ell h^2, \quad h \to 0, \tag{4.17}$$

which compares favorably with (4.14).

As in [16], the asymptotic rate of convergence of an iterative method must, for practical reasons, be considered *in conjunction* with the arithmetic requirements of the method. To this end, we define $R_E(C)$, the *effective rate of convergence* of a convergent matrix C, as

$$R_E(C) = \frac{R_\infty(C)}{M(C)}, \tag{4.18}$$

where $M(C)$ is the average number of multiplications required per mesh point to carry out one complete iteration of the associated iterative method*. Since actual time spent on a computer performing iterations is roughly inversely proportional to the effective rate of convergence, we shall use the effective rate of convergence as a *basis for comparison* of the various iteration methods introduced.

Table 1

ℓ	$R_\infty(\mathfrak{L}_\omega^{(\ell)})$	$R_E(\mathfrak{L}_\omega^{(\ell)})$	$R_\infty({}^R\mathfrak{L}_\omega^{(\ell)})$	$R_E({}^R\mathfrak{L}_w^{(\ell)})$	$R_\infty({}^R\mathfrak{L}_\omega^{(\ell\text{-diag})})$	$R_E({}^R\mathfrak{L}_\omega^{(\ell\text{-diag})})$
2	$\sqrt{16}\,h$	$.667\,h$	$\sqrt{\frac{64}{3}}\,h$	$1.026\,h$	$\sqrt{32}\,h$	$1.131\,h$
3	$\sqrt{24}\,h$	$.639\,h$	$\sqrt{32}\,h$	$1.170\,h$	$\sqrt{48}\,h$	$1.260\,h$
4	$\sqrt{32}\,h$	$.595\,h$	$\sqrt{\frac{128}{3}}\,h$	$1.188\,h$	$\sqrt{64}\,h$	$1.280\,h$
5	$\sqrt{40}\,h$	$.554\,h$	$\sqrt{\frac{160}{3}}\,h$	$1.159\,h$	$\sqrt{80}\,h$	$1.260\,h$

The following table, Table 1, gives the asymptotic and effective rates of convergence for the various methods considered in the numerical solution of five-point difference approximations of the model Dirichlet problem in a square domain of side π with a small uniform mesh spacing h; the asymptotic rates of convergence are computed using (4.14) and (4.17). For comparison with the line methods given in Table 1, the asymptotic and effective rates of convergence for the point successive overrelaxation method are, respectively, $2h$ and $0.4h$.

The arithmetic requirements of $\mathfrak{L}_{\omega_b}^{(\ell)}$, ${}^R\mathfrak{L}_{\omega_b}^{(\ell)}$ and ${}^R\mathfrak{L}_{\omega_b}^{(\ell\text{-diag})}$ are easily seen to increase linearly with ℓ. On the other hand, from (4.15), the asymptotic rates of convergence of these methods for the model problem are proportional to $\sqrt{\ell}$ for small h. Thus, we see from Table 1 that $R_E(\mathfrak{L}_{\omega_b}^{(\ell)})$ is maximized for $\ell=2$, $R_E({}^R\mathfrak{L}_{\omega_b}^{(\ell)})$ is maximized for $\ell=4$, and $R_E({}^R\mathfrak{L}_{\omega_b}^{(\ell\text{-diag})}$ is maximized for $\ell=4$. It should be emphasized that the results of Table 1 were obtained under the assumption of *small* uniform mesh spacings h for the model problem. For problems with discontinuous coefficients P, Q, and σ and general non-uniform mesh spacings, it is not possible to directly compare *all* the different iterative methods given in Table 1. Nevertheless, even for these more general problems, we have shown that

$$R_E({}^R\mathfrak{L}_\omega^{(\ell)}) > R_E(\mathfrak{L}_\omega^{(\ell)}), \qquad \ell \geqq 2,$$

if the number of horizontal mesh lines are greater than ℓ.

* The additional work required in passing from the original matrix equation of (4.4) to the reduced matrix equation of (4.9) and back again through (4.8) and (1.6) is neglected in the results given here. Generally, this additional work is about equivalent to the work required to perform two iterations on the original system. For more details, see [7].

§ 5. Numerical Results

We now give some results of numerical experiments which directly compare the 2-line cyclic Chebyshev semi-iterative method [6] applied to (4.4) with 2- and 3-line cyclic Chebyshev semi-iterative method applied to the cyclically reduced equation of (4.9). We have chosen to use the newer variant, viz. the cyclic Chebyshev semi-iterative method, of the successive overrelaxation iterative method because of its superior error norm behavior for smaller number of iterations. Its asymptotic rate of convergence is the same as that of the successive overrelaxation method [6], so that the results of (4.13) and Table 1 apply equally well to this newer variant.

The programs, which differed only in the method of solution, were written for the Philco-2000 digital computer. We shall refer to them as:

Program 1 Uses the 2-line cyclic Chebyshev method applied to the original system.

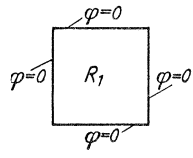
Program 2 Uses the 2-line cyclic Chebyshev method applied to the cyclically reduced system.

Program 3 Uses the 3-line cyclic Chebyshev method applied to the cyclically reduced system.

These programs were applied to the numerical solution of (4.1), where the fuctions $P \equiv Q$, and σ were chosen to be constant in each subregion R_i (cf. Fig. 2), and the region R and its boundary Γ take the various forms of Fig. 2, along with

Problem A

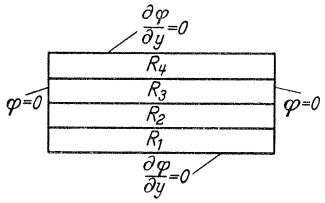
Region	1
$P=Q$	1.0
σ	0.0



Uniform Mesh: 2304 mesh points

Problem B

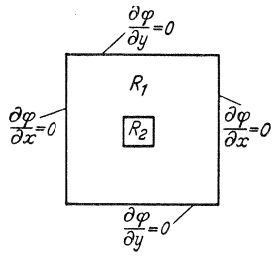
Region	1	2	3	4
$P=Q$	5.0	4.0	9.0	1.0
σ	0.0	0.0	0.0	0.0



Non-uniform Mesh: 2304 mesh points

Problem C

Region	1	2
$P=Q$	2.0	500.0
σ	.05	.05



Non-uniform Mesh: 576 mesh points

Fig. 2

the various prescribed boundary conditions. In each problem, we assumed that $f(x, y) \equiv 0$ in (4.1), so that the unique solution of each discrete problem is the null vector. All components of the initial vector were taken to be 10^3, and all iterations were continued until the maximum vector component was less than 10^{-3}. The results are given in Tables 2 and 3. For a more complete description of the numerical results, see [7].

Table 2

Program No.	Iteration routine		
	Time	Iterations	
1	.0346	61	Problem A
2	.0194	52	
3	.0161	43	
1	.0957	169	Problem B
2	.0537	145	
3	.0429	116	
1	.1016	792	Problem C
2	.0575	681	
3	.0467	554	

Table 3

Problem	No. of iterations ratio		Iteration time ratio	
	Program 1 / Program 2	Program 1 / Program 3	Program 1 / Program 2	Program 1 / Program 3
A	1.173	1.419	1.784	2.149
B	1.166	1.457	1.782	2.231
C	1.163	1.430	1.767	2.176

We remark that the ratios given in Table 3 agree quite well with the corresponding ratios determined from Table 1.

References

[1] ARMS, R. J., L. D. GATES, and B. ZONKEK: A Method of Block Iteration. J. of the Soc. for Indust. and Appl. Math. **4**, 220—229 (1956).

[2] BIRKHOFF, G., and R. S. VARGA: Reactor Criticality and Non-negative Matrices. J. of the Soc. for Indust. and Appl. Math. **6**, 354—377 (1958).

[3] CUTHILL, E. H., and R. S. VARGA: A Method of Normalized Block Iteration. J. of the Association for Computing Machinery. **6**, 236—244 (1959).

[4] FRANKEL, S. P.: Convergence Rates of Iterative Treatments of Partial Differential Equations. Mathematical Tables and Other Aids to Computation **4**, 67—75 (1950).

[5] FROBENIUS, G.: Über Matrizen aus nicht negativen Elementen. S.-B. Preuss. Akad. Wiss., Berlin, pp. 456—477 (1912).

[6] GOLUB, G. H., and R. S. VARGA: Chebyshev Semi-iterative Methods, Successive Overrelaxation Iterative Methods, and Second Order Richardson Iterative Methods. Numerische Mathematik **3**, 147—156 (1961); (Part I) and pp. 157—168 (Part II).

[7] HAGEMAN, L. A.: Block Iterative Methods for Two-cyclic Matrix Equations with Special Application to the Numerical Solution of the Second-Order Self-Adjoint Elliptic Partial Differential Equations in Two Dimensions. Report WAPD-TM-327, Bettis Atomic Power Laboratory, Pittsburgh 1962.

[8] HELLER, J.: Simultaneous, Successive, and Alternating Direction Iteration Schemes. J. of the Soc. for Indust. and Appl. Math. **8**, 150—173 (1960).
[9] HOUSEHOLDER, A. S.: The Approximate Solution of Matrix Problems. J. of the Association for Computing Machinery **5**, 205—243 (1958).
[10] KELLER, H. B.: On Some Iterative Methods for Solving Elliptic Difference Equations. Quart. Appl. Math. **16**, 209—226 (1958).
[11] PARTER, S. V.: Multi-line Iterative Methods for Elliptic Difference Equations and Fundamental Frequencies. Numerische Mathematik **3**, 305—319 (1961).
[12] — On Estimating the Rates of Convergence of Iterative Methods for Elliptic Difference Equations. To appear in: Trans. Amer. Math. Soc.
[13] PRICE, H. S., and R. S. VARGA: Recent Numerical Experiments Comparing Successive Overrelaxation Iterative Methods with Alternating Direction Implicit Methods. Report No. 91, Gulf Research and Development Company, Pittsburgh 1962.
[14] SCHRÖDER, J.: Zur Lösung von Potentialaufgaben mit Hilfe des Differenzenverfahrens. Z. Angewandte Mathematik und Mechanik. **34**, 241—253 (1954).
[15] VARGA, R. S.: Matrix Iterative Analysis. New York: Prentice-Hall Inc. 1962.
[16] — Factorization and Normalized Iterative Methods. Boundary Problems in Differential Equations, pp. 121—142. Madison: The University of Wisconsin Press 1960.
[17] YOUNG, D.: Iterative Methods for Solving Partial Difference Equations of Elliptic Type. Trans. Amer. Math. Soc. **76**, 92—111 (1954).

Bettis Atomic Power Laboratory
Box 1468
Pittsburgh, Pennsylvania
and
Case Institute of Technology
University Circle
Cleveland, Ohio

(Received December 16, 1963)

Druck der Universitätsdruckerei H. Stürtz AG., Würzburg