Numerische Mathematik 3, 157—168 (1961)

Chebyshev semi-iterative methods, successive overrelaxation iterative methods, and second order Richardson iterative methods

Part II

By

GENE H. GOLUB and RICHARD S. VARGA

§ 4. Cyclic Matrices: The Cyclic Chebyshev Semi-Iterative Method

We now suppose that the $N \times N$ matrix B is cyclic, and in the form of (1.4). As we have already pointed out, the matrix B in this form satisfies YOUNG's property A, and is consistently ordered. Because B is real and symmetric, YOUNG's theory [26] can be applied to the solution of the matrix equation of (1.2). With B in the form (1.4), we partition the vectors $\vec{x}$ and $\vec{g}$ of (1.2) in a manner compatible with the partitioning in (1.4), and (1.2) is equivalent to

$$\begin{pmatrix} \vec{x}_1 \\ \vec{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & F \\ F^T & 0 \end{pmatrix} \begin{pmatrix} \vec{x}_1 \\ \vec{x}_2 \end{pmatrix} + \begin{pmatrix} \vec{g}_1 \\ \vec{g}_2 \end{pmatrix}. \tag{4.1}$$

Without using vectors with twice as many components, as was the case in §2, the successive overrelaxation iterative method can be rigorously applied directly to (4.1), giving

$$\begin{cases} \vec{x}_1^{(m+1)} = \omega\{F\vec{x}_2^{(m)} + \vec{g}_1 - \vec{x}_1^{(m)}\} + \vec{x}_1^{(m)} \\ \vec{x}_2^{(m+1)} = \omega\{F^T\vec{x}_1^{(m+1)} + \vec{g}_2 - \vec{x}_2^{(m)}\} + \vec{x}_2^{(m)}, \qquad m \geq 0, \end{cases} \tag{4.2}$$

where $\vec{x}_1^{(0)}$, $\vec{x}_2^{(0)}$ are arbitrary guesses. The best choice of ω is given by

$$\omega_b = \frac{2}{1+\sqrt{1-\varrho^2(B)}} = \frac{2}{1+\sqrt{1-\varrho(FF^T)}}. \tag{4.3}$$

We can also apply to (4.1) the Chebyshev semi-iterative method of (2.9), which gives, by vector components,

$$\begin{cases} \vec{x}_1^{(m+1)} = \omega_{m+1}\{F\vec{x}_2^{(m)} + \vec{g}_1 - \vec{x}_1^{(m-1)}\} + \vec{x}_1^{(m-1)}, \\ \vec{x}_2^{(m+1)} = \omega_{m+1}\{F^T\vec{x}_1^{(m)} + \vec{g}_2 - \vec{x}_2^{(m-1)}\} + \vec{x}_2^{(m-1)}, \qquad m \geq 1, \end{cases} \tag{4.4}$$

where $\vec{x}_1^{(1)} = F\vec{x}_2^{(0)} + \vec{g}_1$, and $\vec{x}_2^{(1)} = F^T\vec{x}_1^{(0)} + \vec{g}_2$, and these equations determine the vector sequences $\{\vec{x}_1^{(m)}\}_{m=0}^{\infty}$, and $\{\vec{x}_2^{(m)}\}_{m=0}^{\infty}$. It is interesting to observe that the proper subsequences $\{\vec{x}_1^{(2m+1)}\}_{m=0}^{\infty}$, and $\{\vec{x}_2^{(2m)}\}_{m=0}^{\infty}$ can be iteratively determined from

$$\begin{cases} \vec{x}_1^{(2m+1)} = \omega_{2m+1}\{F\vec{x}_2^{(2m)} + \vec{g}_1 - \vec{x}_1^{(2m-1)}\} + \vec{x}_1^{(2m-1)}, \qquad m \geq 1, \\ \vec{x}_2^{(2m+2)} = \omega_{2m+2}\{F^T\vec{x}_1^{(2m+1)} + \vec{g}_2 - \vec{x}_2^{(2m)}\} + \vec{x}_2^{(2m)}, \qquad m \geq 0, \end{cases} \tag{4.5}$$

where again $\vec{x}_1^{(1)} = F\vec{x}_2^{(0)} + \vec{g}_1$. Thus, this iterative method requires no add vector storage over the successive overrelaxation iterative method*, and re but the single vector guess $\vec{x}_2^{(0)}$.

We shall call this iterative method, obtained by selecting appropriat sequences of Chebyshev semi-iterative method, the *cyclic Chebyshev iterative method* for the matrix equation (4.1).

In the primitive case of §3, we considered the (primitive) successive relaxation iterative method, or equivalently the second order Richardson m $\alpha = \omega$ and $\beta = -1$, with the starting procedures

$$\vec{x}^{(1)} = B\vec{x}^{(0)} + \vec{g} \tag{4.6}$$

and

$$\begin{cases} \vec{x}^{(1)} = B\vec{x}^{(0)} + \vec{g}' \\ \vec{x}^{(2)} = B\vec{x}^{(1)} + \vec{g}. \end{cases} \tag{4.6'}$$

Here again, it is only necessary in the cyclic case to compute the proper sequences $\{\vec{x}_1^{(2m+1)}\}_{m=0}^{\infty}$ and $\{\vec{x}_2^{(2m)}\}_{m=0}^{\infty}$, and the starting procedures (4.6) (4.6′) become in this case

$$\vec{x}_1^{(1)} = F\vec{x}_2^{(0)} + \vec{g}_1 \tag{4.7}$$

and

$$\begin{cases} \vec{x}_1^{(1)} = F\vec{x}_2^{(0)} + \vec{g}_1 \\ \vec{x}_2^{(2)} = F^T\vec{x}_1^{(1)} + \vec{g}_2. \end{cases} \tag{4.7'}$$

If $\omega_m \equiv \omega$ then we see that (4.5) reduces to (4.2). Thus, for the Chebyshev semi-iterative method, a sequence of parameters ω_m is nece whereas for the successive overrelaxation method, only one parameter is neces The variant of the successive overrelaxation method with the starting proc (4.7′) has been studied by SHELDON [*15*] and the corresponding matrix ope for m iterative is denoted by $\mathfrak{L}_{\omega_b}^{m-1}\mathfrak{L}_1$. The relationship between the Chebyshev semi-iterative method and the successive overrelaxation m is quite close. Indeed, as given by (2.18), $\lim_{m\to\infty} \omega_m = \omega_b$, and it is in fact sl in [*7*], under simple assumptions, that the cyclic Chebyshev semi-iter method must degenerate *numerically* into the successive overrelaxation iter method.

As in §3, we will compare the successive overrelaxation iterative me of (4.2) for the starting procedures of (4.7) and (4.7′) with the cyclic Cheby semi-iterative method of (4.5), and as we shall see, using spectral norms basis for comparison, the cyclic Chebyshev semi-iterative method is superi to the successive overrelaxation iterative method.

* This idea has already been used by RILEY [*13*] to make the second Richardson iterative method competitive in storage with the successive overrelax iterative method.

** In relationship to [*18*], Theorem 1 of [*18*] shows with spectral radii as a for comparison, that the iterative method of (4.2) with $\omega = \omega_b$ is at least twi fast as the iterative method of (4.4). Using the cyclic Chebyshev semi-iter method of (4.5) eliminate this factor of 2 since, from (4.5), each *complete* iter of (4.5) increases the iteration indices of the vectors $\vec{x}_2$ and $\vec{x}_1$ by *two*.

§ 5. Cyclic matrices. Comparison of methods

The results in this section depend strongly upon the methods and results of §3, as well as the recent works of SHELDON [15]. For the Chebyshev semi-iterative method, the successive overrelaxation iterative method, and the second order Richardson iterative method of §2, we partition the error vector $\overset{*}{\vec{\varepsilon}}^{(m)}$ in a manner compatible with the form of the matrix B in (4.1), and we define

$$\overset{*}{\vec{\varepsilon}}^{(m)} = \begin{pmatrix} \overset{*}{\vec{\varepsilon}}_1^{(m)} \\ \overset{*}{\vec{\varepsilon}}_2^{(m)} \end{pmatrix}, \qquad m > 0, \tag{5.1}$$

where $\overset{*}{\vec{\varepsilon}}_1^{(0)} = \vec{\varepsilon}_1^{(0)}$ and $\overset{*}{\vec{\varepsilon}}_2^{(0)} = \vec{\varepsilon}_2^{(0)}$ are the vector components of the initial error vector. For these methods, we have that

$$\overset{*}{\vec{\varepsilon}}^{(m)} = p_m(B)\,\vec{\varepsilon}^{(0)}, \qquad m > 0, \tag{5.2}$$

where the matrix operator $p_m(B)$ corresponds respectively to the matrix operators $\tilde{p}_m(B)$, $r_m(B)$, $t_m(B)$ and $s_m(B)$ of §3. For the cyclic Chebyshev semi-iterative method, and the (cyclic) successive overrelaxation iterative method with the starting procedures of (4.7) and (4.7'), the corresponding error vector for the m-th complete iteration of these methods is defined by

$$\vec{\delta}^{(m)} = \begin{pmatrix} \overset{*}{\vec{\varepsilon}}_1^{(2m-1)} \\ \overset{*}{\vec{\varepsilon}}_2^{(2m)} \end{pmatrix}, \qquad m > 0. \tag{5.3}$$

From (2.8'), (3.21), and (3.28), it follows that the polynomials $p_m(x)$ of odd degree contain only odd powers of x, while the polynomials of even degree contain only even powers of x. Thus, we define polynomials U_m and V_m through

$$\begin{cases} p_{2m+1}(x) = x\,U_m(x^2), & m \geqq 0, \\ p_{2m}(x) = V_m(x^2), & m \geqq 0. \end{cases} \tag{5.4}$$

Since the matrix has the form (4.1), then

$$B^{2m} = \left(\begin{array}{c|c} (FF^T)^m & 0 \\ \hline 0 & (F^TF)^m \end{array}\right) \quad\text{and}\quad B^{2m+1} = \left(\begin{array}{c|c} 0 & (FF^T)^m F \\ \hline (F^TF)^m F^T & 0 \end{array}\right), \tag{5.5}$$

and the definitions of (5.4) and the properties of the powers of the matrix B allow us to express $\vec{\delta}^{(m)}$ in the simple form

$$\vec{\delta}^{(m)} = \left(\begin{array}{c|c} 0 & U_{m-1}(FF^T)\cdot F \\ \hline 0 & V_m(F^TF) \end{array}\right)\vec{\varepsilon}^{(0)}, \qquad m > 0. \tag{5.6}$$

Defining the matrix above as $P_m(B)$, this becomes

$$\vec{\delta}^{(m)} = P_m(B)\,\vec{\varepsilon}^{(0)}, \qquad m > 0. \tag{5.6'}$$

We analogously define the 2×2 matrix $Q_m(\mu)$ as

(5.7) $$Q_m(\mu) = \left(\begin{array}{c|c} 0 & U_{m-1}(\mu^2)\,\mu \\ \hline 0 & V_m(\mu^2) \end{array}\right), \qquad m>1,$$

whose spectral norm is easily seen to be

(5.8) $$\tau[Q_m(\mu)] = \{\mu^2 U_{m-1}^2(\mu^2) + V_m^2(\mu^2)\}^{\frac{1}{2}}, \qquad m \geqq 1.$$

From (5.4), this becomes

(5.8′) $$\tau[Q_m(\mu)] = \{p_{2m-1}^2(\mu) + p_{2m}^2(\mu)\}^{\frac{1}{2}}, \qquad m \geqq 1.$$

We now employ what is essentially a converse of Theorem 2 of the rec of SHELDON [15]*. Denoting the eigenvalues of the matrix B by μ_i, 1 then

(5.9) $$\tau[P_m(B)] = \max_{1\leqq i\leqq N} \{p_{2m-1}^2(\mu_i) + p_{2m}^2(\mu_i)\}^{\frac{1}{2}}, \qquad m \geqq 1.$$

Let us now denote the matrix operator of (5.6′) associated with the pol $\tilde{p}_m(B)$, $r_m(B)$, $l_m(B)$, and $s_m(B)$ of §3 as $\tilde{P}_m(B)$, $R_m(B)$, $T_m(B)$, and respectively. Then it follows immediately from the results of §3 that

(5.10) $$\begin{cases} \tau[\tilde{P}_m(B)] = \{\tau^2(\tilde{p}_{2m-1}(B)) + \tau^2(\tilde{p}_{2m}(B))\}^{\frac{1}{2}} \\ \tau[R_m(B)] = \{\tau^2(r_{2m-1}(B)) + \tau^2(r_{2m}(B))\}^{\frac{1}{2}} \\ \tau[T_m(B)] = \{\tau^2(l_{2m-1}(B)) + \tau^2(l_{2m}(B))\}^{\frac{1}{2}} \\ \tau[S_m(B)] = \{\tau^2(s_{2m-1}(B)) + \tau^2(s_{2m}(B))\}^{\frac{1}{2}}. \end{cases}$$

Since $\tau(\tilde{p}_m(B))$, $\tau(r_m(B))$, $\tau(l_m(B))$ and $\tau(s_m(B))$ decrease monotonical m, so do $\tau[\tilde{P}_m(B)]$, $\tau[R_m(B)]$, $\tau[T_m(B)]$**, and $\tau[S_m(B)[$. Furtherm Theorem 1, for $m>1$ and $0<\varrho<1$,

(5.11) $$\tau(\tilde{p}_m(B)) < \tau(r_m(B)) < \tau(l_m(B)) < \tau(s_m(B)),$$

so that

Lemma 2. For all $m>1$ and $0<\varrho<1$,

$$\tau[\tilde{P}_m(B)] < \tau[R_m(B)] < \tau[T_m(B)] < \tau[S_m(B)].$$

The spectral norm of the successive overrelaxation iterative method for the case when ω is fixed equal to ω_b has been recently calculated by S

* Specifically, in the notation of SHELDON [15], the result we are using in the following

Theorem. If λ is a non-zero eigenvalue of L, then λ is also an eigenvalue $T(\mu_i)$ where μ_i is an eigenvalue of the matrix B.

This result is tacitly assumed in [15], and we are indebted to Dr. SHELD supplying us with a proof of this result.

** The quantity $\tau[T_m(B)]$ in (5.10) is algebraically equivalent to the ex for $\tau[\mathfrak{L}_{\omega_b}^{m-1}\mathfrak{L}_1]$ in [15]. Thus, the monotonicity noted above strengthens SH Theorem 4 in [15].

[15], and if $\mathfrak{L}_{\omega_b}^m$ represents the corresponding matrix operator for m iterations, then*

$$\tau\left[\mathfrak{L}_{\omega_b}^m\right] = l_m^{\frac{1}{2}}(\omega_b - 1)^m, \qquad m \geqq 0, \tag{5.12}$$

where l_m is the larger root of

$$l^2 - \left[8m^2 + 4m^2\left(r^2 + \frac{1}{r^2}\right) + 2\right] l + 1 = 0, \tag{5.13}$$

and $r^2 = \omega_b - 1$, so that

$$\tau\left[\mathfrak{L}_{\omega_b}^m\right] = \left(\frac{2m}{\varrho} + \sqrt{\frac{4m^2}{\varrho^2} + 1}\right) \cdot (\omega_b - 1)^m, \qquad m \geqq 0. \tag{5.12'}$$

We observe that in obtaining the spectral norms for the four iterative methods just considered, no assumption has been made about a special form of the initial error $\vec{\varepsilon}^{(0)}$, and thus the four iterative methods can be directly compared.

Then we have

Theorem 2. In the cyclic case for all $m > 1$ and $0 < \varrho < 1$, with no special assumption on the form of the initial error vector $\vec{\varepsilon}^{(0)}$,

$$\begin{cases} \tau\left[\widetilde{P}_m(B)\right] < \tau\left[R_m(B)\right] < \tau\left[T_m(B)\right] < \tau\left[S_m(B)\right], \quad \text{and} \\ \tau\left[\widetilde{P}_m(B)\right] < \tau\left[\mathfrak{L}_{\omega_b}^m\right]. \end{cases} \tag{5.14}$$

Thus, the spectral norm of the matrix operator for the cyclic Chebyshev semi-iterative method is less than the spectral norm of the matrix operators for the successive overrelaxation iterative method and its modification by SHELDON.

Proof. From Lemma 2, it suffices to show that $\tau\left[\widetilde{P}_m(B)\right] < \tau\left[\mathfrak{L}_{\omega_b}^m\right]$ for all $m > 1$ and $0 < \varrho < 1$. By using the expressions of (3.8), (5.10), and (5.12'), this inequality reduces to

$$\left\{r^{-2}\left(\frac{2}{1+r^{4m-2}}\right)^2 + \left(\frac{2}{1+r^{4m}}\right)^2\right\}^{\frac{1}{2}} < \frac{2m}{\varrho} + \sqrt{\frac{4m^2}{\varrho^2} + 1}, \tag{5.15}$$

which is easily shown to be true for all $m > 1$, and $0 < \varrho < 1$. In fact, the proof of the above inequality shows that the ratio $\tau\left[\mathfrak{L}_{\omega_b}^m\right] / \tau\left[\widetilde{P}_m(B)\right]$ is a strictly increasing function of m, $m > 1$, for all $0 < \varrho < 1$. We strengthen the inequalities of (5.14) by including

Theorem 3. In the cyclic case with $0 < \varrho < 1$, and no special assumptions on the form of the initial error vector $\vec{\varepsilon}^{(0)}$, then the ratios

$$\frac{\tau\left[R_m(B)\right]}{\tau\left[\widetilde{P}_m(B)\right]} = \alpha_m; \qquad \frac{\tau\left[\mathfrak{L}_{\omega_b}^m\right]}{\tau\left[\widetilde{P}_m(B)\right]} = \beta_m \tag{5.16}$$

are strictly increasing for $m > 1$, and

$$\alpha_m = O(m), \qquad \beta_m = O(m), \qquad m \to \infty. \tag{5.17}$$

* Theorem 3 of [15] contains minor misprints, which we are now correcting.

Proof. It is an easy computation to show that $\tau[\widetilde{P}_m(B)] < 2r^{2m}(1+r^{-2})$, and that $2r^{2m}(1+r^{-2})$ is smaller than either $\tau[R_m(B)]$ or $\tau[\mathfrak{L}^m_{\omega_b}]$. The statements of (5.16) and (5.17) then follow immediately*.

§ 6. Applications

A great many physical and engineering problems lead to the numerical solution of matrix equations of the form

$$A\vec{x} = \vec{k}, \tag{6.1}$$

where A is an $N \times N$ real symmetric and positive definite matrix which can, after a suitable permutation of indices, be partitioned so that

$$A = \left[\begin{array}{cccc|cccc} A_{1,1} & 0 & \cdots & 0 & A_{1,p+1} & & \cdots & A_{1,s} \\ 0 & A_{2,2} & & 0 & \vdots & & & \vdots \\ \vdots & & \ddots & \vdots & & & & \\ 0 & 0 & \cdots & A_{p,p} & A_{p,p+1} & & \cdots & A_{p,s} \\ \hline A^T_{1,p+1} & \cdots & & A^T_{p,p+1} & A_{p+1,p+1} & 0 & \cdots & 0 \\ & & & & 0 & A_{p+2,p+2} & & 0 \\ \vdots & & & \vdots & \vdots & & \ddots & \vdots \\ A^T_{1,s} & \cdots & & A^T_{p,s} & 0 & 0 & \cdots & A_{s,s} \end{array}\right] \tag{6.2}$$

where the diagonal blocks $A_{j,j}$ are $n_j \times n_j$ matrices, $n_j \geq 1$ for $1 \leq j \leq s$, and $\sum_{j=1}^{s} n_j = N$. ARMS, GATES, and ZONDEK [*1*] extended the original analysis of YOUNG [*26*] and FRANKEL [*6*] to what is called the *successive block overrelaxation iterative method*, and it can be verified that the assumptions on the matrix A above are sufficient for the application of their theory. Let the vectors $\vec{x}$ and $\vec{k}$ of (6.1) be partitioned in a manner compatible with (6.2). Then, we can write (6.1) as

$$\begin{cases} A_{j,j}X_j + \sum_{k=1}^{s-p} A_{j,p+k}X_{p+k} = K_j, & 1 \leq j \leq p, \\ A_{p+j,p+j}X_{p+j} + \sum_{k=1}^{p} A^T_{k,p+j}X_k = K_{p+j}, & 1 \leq j \leq s-p. \end{cases} \tag{6.3}$$

The square submatrices $A_{j,j}$, $1 \leq j \leq s$, are evidently non-singular, so that the block diagonal matrix C is defined by

$$C = \begin{bmatrix} A_{1,1} & 0 & \cdots & 0 \\ 0 & A_{2,2} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & A_{s,s} \end{bmatrix}, \tag{6.4}$$

* Mr. DAVID FEINGOLD of Electricité de France (Paris) has recently proved (private communication) that the ratio $\{\tau[\mathfrak{L}^m_{\omega_b}]/\tau[R_m(B)]\}$ is strictly increasing for $m > 1$, $0 < \varrho < 1$, which strengthens Theorems 2 and 3.

then C is also non-singular. Now, $C^{-1}A$ has unit diagonal entries, and we define the matrix B as

(6.5) $$C^{-1}A = I - B,$$

so that the matrix B has zero diagonal entries. More precisely, B has the form

(6.6) $$B = \left[\begin{array}{ccc|ccc} 0 & \ldots & 0 & B_{1,p+1} & \ldots & B_{1,s} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & B_{p,p+1} & \ldots & B_{p,s} \\ \hline B_{p+1,1} & \ldots & B_{p+1,p} & 0 & \ldots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ B_{s,1} & \ldots & B_{s,p} & 0 & \ldots & 0 \end{array}\right].$$

With the definition of the matrix B in (6.5), (6.4) becomes

(6.7) $$\vec{x} = B\vec{x} + C^{-1}\vec{k},$$

The successive block overrelaxation iterative method applied to (6.7) is defined by

(6.8) $$\begin{cases} X_j^{(m+1)} = \omega\left[\sum_{k=1}^{s-p} B_{j,p+k} X_{p+k}^{(m)} + A_{j,j}^{-1} K_j - X_j^{(m)}\right] + X_j^{(m)}, & 1 \leqq j \leqq p, \\ X_{p+j}^{(m+1)} = \omega\left[\sum_{k=1}^{p} B_{p+j,k} X_k^{(m+1)} + A_{p+j,p+j}^{-1} K_{p+j} - X_{p+j}^{(m)}\right] + X_{p+j}^{(m)}, & \\ & 1 \leqq j \leqq s-p, \end{cases}$$

where the $X_j^{(0)}$, $1 \leqq j \leqq s$, are given vector components of the given initial vector guess $\vec{x}^{(0)}$. The optimum value of ω is computed from (4.3), where the $N\times N$ matrix B is defined in (6.5). Equivalently, the iterations of (6.8) can be defined also from

(6.9) $$X_j^{(m+1)} = \omega\left[\overset{*}{X}{}_j^{(m+1)} - X_j^{(m)}\right] + X_j^{(m)}, \qquad 1 \leqq j \leqq s,$$

where

(6.9′) $$\begin{cases} A_{j,j}\overset{*}{X}{}_j^{(m+1)} = -\sum_{k=1}^{s-p} A_{j,p+k} X_{p+k}^{(m)} + K_j, & 1 \leqq j \leqq p, \\ A_{p+j,p+j}\overset{*}{X}{}_{p+j}^{(m+1)} = -\sum_{k=1}^{p} A_{k,p+j}^T X_k^{(m+1)} + K_{p+j}, & 1 \leqq j \leqq s-p. \end{cases}$$

Equation (6.9′) shows that, in order to carry out the successive block overrelaxation iterative method, we have assumed that matrix equations of the form

(6.10) $$A_{j,j} X_j = G_j, \qquad 1 \leqq j \leqq s \cdot$$

can be solved directly for X_j, given G_j.

The matrix C defined in (6.4) is symmetric and positive definite, so that the matrices $C^{\frac{1}{2}}$ and $C^{-\frac{1}{2}}$ are uniquely defined. Forming the product $C^{-\frac{1}{2}}AC^{-\frac{1}{2}}$, we see that this product matrix also has unit diagonal entries, and in analogy with (6.5), we define the matrix $\hat{B}$ by

(6.11) $$C^{-\frac{1}{2}}AC^{-\frac{1}{2}} = I - \hat{B}.$$

The matrix $\widetilde{B}$ has the same cyclic form as does B of (6.7), and since $C^{-\frac{1}{2}} A C^{-\frac{1}{2}}$ is a definite and symmetric matrix, it follows from (6.11) that $\widetilde{B}$ is symmetric and convergent. Defining

$$C^{\frac{1}{2}} \vec{x} = \vec{y}, \qquad C^{-\frac{1}{2}} \vec{k} = \vec{l} \tag{6.12}$$

and using (6.11), (6.1) reduces to

$$\vec{y} = \widetilde{B} \vec{y} + \vec{l}. \tag{6.13}$$

The matrix $\widetilde{B}$ is similar to B, with

$$\widetilde{B} = C^{\frac{1}{2}} B C^{-\frac{1}{2}}. \tag{6.14}$$

Summarizing, we have reduced our original problem (6.1) by means of a change of variables to the form (6.13), where $\widetilde{B}$ is symmetric, cyclic, and convergent.

We now apply the cyclic Chebyshev semi-iterative method to the numerical solution of (6.13). If the vector components $Y_j^{(0)}$, $1 \leqq j \leqq p$, are given, then

$$(6.15)\quad \begin{cases} Y_{p+j}^{(2m+1)} = \omega_{2m+1} \left\{ \sum_{k=1}^{p} \widetilde{B}_{p+j,k} Y_k^{(2m)} + L_{p+j} - Y_{p+j}^{(2m-1)} \right\} + Y_{p+j}^{(2m-1)}, \\ \qquad\qquad 1 \leqq j \leqq s-p, \\ Y_j^{(2m+2)} = \omega_{2m+2} \left\{ \sum_{k=1}^{s-p} \widetilde{B}_{j,p+k} Y_{p+k}^{(2m+1)} + L_j - Y_j^{(2m)} \right\} + Y_j^{(2m)}, \\ \qquad\qquad 1 \leqq j \leqq p, \quad m \geqq 0, \end{cases}$$

defines the cyclic Chebyshev semi-iterative method. The ω's are calculated from (2.10), where $\varrho(B) = \varrho(\widetilde{B})$, since $\widetilde{B}$ is similar to B. To show now the relationship of this method to the successive block overrelaxation iterative method of (6.9)—(6.9′), we write (6.15) equivalently as

$$(6.16)\quad \begin{cases} Y_{p+j}^{(2m+1)} = \omega_{2m+1} \left(\overset{*}{Y}{}_{p+j}^{(2m+1)} - Y_{p+j}^{(2m-1)} \right) + Y_{p+j}^{(2m-1)}, & 1 \leqq j \leqq s-p, \quad m \geqq 0, \\ Y_j^{(2m+2)} = \omega_{2m+2} \left(\overset{*}{Y}{}_j^{(2m+2)} - Y_j^{(2m)} \right) + Y_j^{(2m)}, & 1 \leqq j \geqq p, \quad m \geqq 0, \end{cases}$$

where

$$(6.16')\quad \begin{cases} \overset{*}{Y}{}_{p+j}^{(2m+1)} = \sum_{k=1}^{p} \widetilde{B}_{p+j,k} Y_k^{(2m)} + L_{p+j}, & 1 \leqq j \leqq s-p, \quad m \geqq 0, \\ \overset{*}{Y}{}_j^{(2m+2)} = \sum_{k=1}^{s-p} \widetilde{B}_{j,p+k} Y_{p+k}^{(2m+1)} + L_j, & 1 \leqq j \leqq p, \quad m \geqq 0. \end{cases}$$

By using the definitions of (6.11) and (6.12), it follows that (6.15) is *equivalent* to (6.9)—(6.9′), provided the proper ω's are used in each step. In essence then, we can *indirectly* carry out the modified Chebyshev semi-iterative method of (6.15) by performing the iterations

$$(6.9'')\quad \begin{cases} X_{p+j}^{(m+1)} = \omega_{2m+1} \left(\overset{*}{X}{}_{p+j}^{(m+1)} - X_{p+j}^{(m)} \right) + X_{p+j}^{(m)}, & 1 \leqq j \leqq s-p, \quad m \geqq 0, \\ X_j^{(m+1)} = \omega_{2m+2} \left(\overset{*}{X}{}_j^{(m+1)} - X_j^{(m)} \right) + X_j^{(m)}, & 1 \leqq j \leqq p, \quad m \geqq 0, \end{cases}$$

where $\overset{*}{X}{}_j^{(m)}$, $1 \leqq j \leqq s$, is defined in (6.9′).

In terms of spectral norms, let $\vec{\delta}^{(m)} = \begin{pmatrix} \vec{\varepsilon}_1^{(2m)} \\ \vec{\varepsilon}_2^{(2m+1)} \end{pmatrix}$ denote the error vector for the m-th complete iterate of (6.15), relative to the matrix $\widetilde{B}$. From §5, we can state that

$$\|\vec{\delta}^{(m)}\| \leqq \tau[\widetilde{P}_m(\widetilde{B})] \cdot \|\delta^{(0)}\|, \qquad m \geqq 0. \tag{6.17}$$

If $\vec{\sigma}^{(m)} \equiv \begin{pmatrix} \vec{\alpha}_1^{(2m)} \\ \vec{\alpha}_2^{(2m+1)} \end{pmatrix}$ is the error for the m-th complete iteration of (6.9″), relative to the matrix B, then from $C^{\frac{1}{2}}\vec{x} = \vec{y}$, we have

$$\|C^{\frac{1}{2}}\vec{\sigma}^{(m)}\| \leqq \tau[\widetilde{P}_m(\widetilde{B})] \cdot \|C^{\frac{1}{2}}\vec{\sigma}^{(0)}\|, \qquad m \geqq 0. \tag{6.18}$$

Since both $C^{\frac{1}{2}}$ and $C^{-\frac{1}{2}}$ are symmetric and positive definite, their spectral radii coincide with their spectral norms, so that

$$\|C^{\frac{1}{2}} z\| \leqq \varrho(C^{\frac{1}{2}})\, \|z\|, \tag{6.19}$$

and

$$\|C^{\frac{1}{2}} z\| \geqq \frac{\|z\|}{\varrho(C^{-\frac{1}{2}})}, \tag{6.19'}$$

where equality is possible in both (6.19) and (6.19′). Combining these inequalities, we have*

$$\|\vec{\sigma}^{(m)}\| \leqq \tau[\widetilde{P}_m(\widetilde{B})]\,[\varrho(C^{-\frac{1}{2}}) \cdot \varrho(C^{\frac{1}{2}})]\, \|\vec{\sigma}^{(0)}\|, \qquad m \geqq 0. \tag{6.20}$$

From the results of §5, of the iterative methods studied, the cyclic Chebyshev semi-iterative method of (6.16)—(6.16′) gives the smallest spectral norm relative to the matrix equation of (6.13). Since actually iterating by means of (6.9′)—(6.9″) is equivalent to iterating by means of (6.16)—(6.16′), we arrive at the conclusion that the iterations of (6.9′)—(6.9″) are quite efficient.

We now list some well known problems which numerically give rise to matrix equations of the form (6.1), where the matrix A can be written as in (6.2). Clearly, such a list would include all problems which have been previously rigorously attacked by the successive overrelaxation iterative method, and its extensions.

A. Dirichlet problem in a plane bounded region, using a five point approximation to LAPLACE'S equation. Here, one can use successive point overrelaxation [*6, 19, 26*], successive line relaxation [*1, 3, 8*], or successive two line overrelaxation [*12, 21*], all these methods corresponding to different partitionings of the matrix A.

B. Dirichlet problem in a plane bounded region, using a ninepoint approximation to LAPLACE'S equation. Here, one can use successive line overrelaxation [*1, 21*], or successive two line overrelaxation [*8, 12, 21*].

C. Biharmonic problem in a plane bounded region, using a thirteen point approximation to the biharmonic equation. Here, one can use successive two line overrelaxation [*8, 12, 21*].

---

* The quantity $(\varrho(M^{-1})) \cdot \varrho(M)$ is also called the *P-condition number* [*17*] for a non-singular matrix M, and is denoted by $P(M)$.

In all these problems, the cyclic Chebyshev semi-iterative method can be used, and from the results of § 5, this iterative method gives the smaller spectral norm than the successive overrelaxation iterative methods.

Finally, matrix equations (6.1) do arise in which the matrix A cannot, after a permutation of indices, be put into the form of (6.2), even with proper partitioning. For example, in [21], a class of iterative methods called *primitive* iterative methods are studied, and for this class the results of § 2—3 are pertinent. It should also be said that even though the matrix A of (6.1) can be partitioned so that (6.2) holds, it can very well be the case that the diagonal blocks $A_{i,i}$, which must be directly inverted, as in (6.10), in order to apply the cyclic theory, are either too large in size or too complicated to permit such direct inversion. Thus, in solving the Dirichlet problem in a plane bounded region, if one chooses to use a nine point approximation to LAPLACE's equation, but is unwilling to directly invert more than one equation in one unknown, a primitive iterative method results. Here too the results of § 2—3 are pertinent.

§ 7. Numerical Results

We will now give results from both algebraic and numerical investigations, comparing the Chebyshev semi-iterative method with variants of the successive overrelaxation iterative method in the cyclic case. First, if $\vec{\varepsilon}^{(0)}$ is the vector error of our initial estimate $\vec{x}_0$ of the unique solution of $A\vec{x}=\vec{k}$, and $\vec{\delta}^{(m)}$ is the error vector for the m-th complete iteration, then from (5.6′),

$$\frac{\|\vec{\delta}^{(m)}\|}{\|\vec{\varepsilon}^{(0)}\|} \leqq \tau[P_m(B)], \qquad m>0. \tag{7.1}$$

Thus, if $m(\delta)$ is the least positive integer for which

$$\tau[P_m(B)] \leqq \delta, \qquad 0<\delta<1, \tag{7.2}$$

then $m(\delta)$ is an *upper bound* for the number of iterations necessary to reduce the Euclidean length of the initial error by the factor δ. Let $m_1(\delta)$, $m_2(\delta)$, $m_3(\delta)$, and $m_4(\delta)$ denote $m(\delta)$ when $P_m(B)$ is taken to be respectively $\widetilde{P}_m(B)$, $R_m(B)$, $T_m(B)$ and $\mathfrak{L}_{\omega_b}^m$. The tables 1—4 give $m_i(\delta)$ for various values of δ and $\varrho(B)$.

Table 1. $\omega_b=1.8195$; $\varrho=0.99507$

	$\delta=0.1$	$\delta=0.05$	$\delta=0.01$	$\delta=0.005$	$\delta=0.001$
$m_1(\delta)$	18	21	29	33	41
$m_2(\delta)$	22	27	36	40	49
$m_3(\delta)$	23	27	37	41	50
$m_4(\delta)$	37	41	50	54	63

Table 2. $\omega_b=1.93419$; $\varrho=0.999421$

	$\delta=0.1$	$\delta=0.05$	$\delta=0.01$	$\delta=0.005$	$\delta=0.001$
$m_1(\delta)$	50	60	84	94	117
$m_2(\delta)$	64	77	104	116	142
$m_3(\delta)$	65	77	105	116	143
$m_4(\delta)$	126	137	163	174	200

Table 3. $\omega_b=1.95218$; $\varrho(B)=0.9997$

	$\delta=0.1$	$\delta=0.05$	$\delta=0.01$	$\delta=0.005$	$\delta=0.001$
$m_1(\delta)$	69	93	116	130	163
$m_2(\delta)$	89	106	144	160	197
$m_3(\delta)$	89	107	145	161	198
$m_4(\delta)$	182	198	234	250	285

Table 4. $\omega_b=1.97211$; $\varrho(B)=0.9999$

	$\delta=0.1$	$\delta=0.05$	$\delta=0.01$	$\delta=0.005$	$\delta=0.001$
$m_1(\delta)$	119	143	200	225	282
$m_2(\delta)$	154	183	249	277	341
$m_3(\delta)$	154	184	250	278	341
$m_4(\delta)$	337	364	426	453	514

It is interesting to point out that the following

$$\lim_{\delta \to 0} \frac{m_i^{(\delta)}}{m_j^{(\delta)}} = 1 \tag{7.3}$$

can be proved* for all i, j. Thus, the cyclic Chebyshev semi-iterative method cannot require, for very small $\delta > 0$, percentagewise substantially different numbers of iterations than those required by the successive overrelaxation method. However, for slowly convergent problems, $\varrho(B)$ close to unity, there is a considerable advantage in using the cyclic Chebyshev in practical problems where δ is approximately 10^{-2}.

The above, while constituting an algebraic study of the various methods, does not give a complete picture of the comparison between these methods, because of the inequalities in (7.1) and (7.2). Although equality is attainable in (7.1) and (7.2), so that the numbers of iterations in Tables 1—4 are also attainable, we include results of numerical experiments in the cyclic case. In an effort to make the numerical experiments as up-to-date and practical as possible, we have compared the successive two line over-relaxation iterative method [*8*, *12*, *21*] with the cyclic Chebyshev semi-iterative method for the same partitioning of the matrix A of (6.2), in the numerical solution of self-adjoint partial differential equation

$$-\operatorname{div}\{D(x, y) \operatorname{grad} u(x, y)\} + \sigma(x, y)\, u(x, y) = S(x, y), \tag{7.4}$$

in a plane bounded region Ω, where D and σ are positive in Ω, with boundary conditions

$$\frac{\partial u(x, y)}{\partial n} = 0 \tag{7.5}$$

on the boundary Γ of Ω. These numerical problems involved non-constant mesh spacings. In part 1 of each problem, $S(x, y) \equiv 0$, so that the unique solution of the matrix problem of (6.1) is the null vector. With all the components of the initial vector $\vec{x}^{(0)}$ taken as 10^3, the iterations were continued until the maximum component of $\vec{x}^{(m)}$ was less than or equal to δ. In part 2 of each problem, $S(x, y) \equiv 1$ and with the same initial vector $\vec{x}^{(0)}$ as in part 1, the iterations were continued until

$$R^{(m+1)} = \sum_j \left| x_j^{(m+1)} - x_j^{(m)} \right| \tag{7.6}$$

satisfied $R^{(m+1)} \leq \delta R^{(0)}$.

Because the norms of both parts of the experiment are convenient in computation, but not the spectral norms of the comparison, the following comparisons are of interest in connection with the relationships exhibited in §6. The successive overrelaxation method is applied to two different orderings of the matrix A: the first, the σ_1 ordering, is the ordering of (6.2); the second is the "normal" ordering in which the double lines of mesh points are swept serially through the mesh.

* See [7] for details.

Table 5. *Problem A 121 interior mesh points*, $\omega_b = 1.8195$

Part 1

Method	$\delta=0.1$	$\delta=0.01$	$\delta=0.005$	$\delta=0.001$
Cyclic Chebyshev	17	28	31	39
SHELDON's Modified SOR	21	35	39	48
SOR with ω_b, σ_1 Ordering	20	34	37	46
SOR with ω_b, Normal Ordering	17	30	34	43

Part 2

Method	$\delta=0.1$	$\delta=0.01$	$\delta=0.005$
Cyclic Chebyshev	30	41	44
SHELDON's Modified SOR	39	52	55
SOR with ω_b, σ_1 Ordering	33	46	50
SOR with ω_b, Normal Ordering	32	45	49

Table 6. *Problem B 667 interior mesh points*, $\omega_b = 1.93419$

Part 1

Method	$\delta=0.1$	$\delta=0.01$	$\delta=0.005$	$\delta=0.001$
Cyclic Chebyshev	71	106	110	133
SHELDON's Modified SOR	88	123	134	157
SOR with ω_b, σ_1 Ordering	93	127	137	160
SOR with ω_b, Normal Ordering	81	121	133	155

Part 2

Method	$\delta=0.1$	$\delta=0.01$	$\delta=0.005$
Cyclic Chebyshev	83	113	119
SHELDON's Modified SOR	113	147	157
SOR with ω_b, σ_1 Ordering	97	133	143
SOR with ω_b, Normal Ordering	91	127	137

For references, see Part I **3**, 147 (1961).

Space Technology Laboratories, Inc.
Los Angeles 45, California
and
Case Institute of Technology
Cleveland 6, Ohio
10900 Euclid Avenue

(Received June 17, 1960)