Pacific Journal of Mathematics

p-CYCLIC MATRICES: A GENERALIZATION OF THE YOUNG-FRANKEL SUCCESSIVE OVERRELAXATION SCHEME

RICHARD STEVEN VARGA

Vol. 9, No. 2 June 1959

p-CYCLIC MATRICES: A GENERALIZATION OF THE YOUNG-FRANKEL SUCCESSIVE OVERRELAXATION SCHEME

RICHARD S. VARGA

1. Introduction. The Young-Frankel [11, 5] successive overrelaxation scheme, which has been shown [11, pp. 104–109] to be applicable to the numerical solution of partial equations of elliptic type, can be described as follows. If the system of linear equations to be solved is

$$M\vec{x} = \vec{k}\,, \tag{1}$$

where the $n \times n$ matrix $M = (m_{i,j})$ is such that $m_{i,i} \neq 0$ for $i = 1, 2, \cdots, n$, then the iterative sequence, defined by the successive overrelaxation scheme, is given by

$$x_i^{(l+1)} = \omega\left\{\sum_{j=1}^{i-1} b_{i,j}x_j^{(l+1)} + \sum_{j=i+1}^{n} b_{i,j}x_j^{(l)} + c_i\right\} + (1-\omega)x_i^{(l)} \tag{2}$$

where $x_i^{(0)}$ is arbitrary, $i = 1, 2, \cdots, n$, and where

$$b_{i,j} = \begin{cases} -m_{i,j}/m_{i,i}, & i \neq j \\ 0, & i = j \end{cases}, \tag{3}$$

and

$$c_i = k_i/m_{i,i}, \qquad i = 1, 2, \cdots, n\,. \tag{4}$$

With certain assumptions, Young [11] has shown that, for suitable choice of the *relaxation factor* ω, relatively rapid convergence of the iterative process (2) is assured. These hypotheses are satisfied by the usual five-point difference approximation to $-\nabla \cdot (k \nabla u) = S$, $k > 0$, in the plane [11, 9].

We show that successive overrelaxation can be considered as a special case of a more general iterative scheme applicable to the wider class of *p-cyclic* matrices, to be defined below. Indeed, ordinary successive (point) overrelaxation, as well its generalization [1] to successive block (line) overrelaxation, is just the special case $p = 2$ of the iterative scheme we shall now define.

2. p-cyclic matrices. We begin with the following

This paper was originally accepted by the Trans. Amer. Math. Soc. Presented to the American Mathematical Society, August 30, 1957, under the title "*The p-color problem: a generalization of the Young-Frankel successive overrelaxation scheme.*" Received by the editors of the Trans. Math. Soc. January 24, 1958.

DEFINITION 1. The $n \times n$ matrix $A = (a_{i,j})$ is *cyclic of index* p ($p \geqq 2$) if and only if, after some permutation of rows and columns of A, the matrix A assumes the cyclic block form

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & L_1 \\ L_2 & 0 & \cdots & 0 & 0 \\ 0 & L_3 & \cdots & 0 & 0 \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & \cdots & L_p & 0 \end{pmatrix} \tag{5}$$

where the zero diagonal submatrices are square.

This terminology was introduced by Romanovsky [8]. For A cyclic of index p, Romanovsky [8, pp. 162-166] proved that the characteristic polynomial $p_n(\lambda) \equiv |\lambda I - A|$ of A is of the form

$$p_n(\lambda) = \lambda^{\nu} \prod_{i=1}^{l} (\lambda^p - \sigma_i^p) , \tag{6}$$

where $\nu + \swarrow p = n$. (See Theorem 3.)

The following generalizations of the results of [11, 1] are easily established, and so we state them without proof.[1] The results will be stated only for generalizations of Young's (point) property (A); the extension to generalizations of block (line) property (A^{π}) is easily carried out.

DEFINITION 2. The $n \times n$ matrix $M = (m_{i,j})$ is *p-cyclic*, $p \geqq 2$, if and only if the diagonal entries of M are all non-zero, and there exists p disjoint non-empty subsets $S_{\swarrow}$, $\swarrow = 0, 1, \cdots, p-1$, of W, the set of the first n positive integers, such that $\bigcup_{\swarrow=0}^{p-1} S_{\swarrow} = W$, and if $m_{i,j} \neq 0$, then either $i = j$, or if $i \in S_{\swarrow}$, then $j \in S_{\swarrow-1}$, subscripts taken mod p.[2]

THEOREM 1. *The $n \times n$ matrix $M = (m_{i,j})$ is p-cyclic if and only if there exists a vector $\vec{\gamma} = (\gamma_1 \cdots, \gamma_n)$ with integral components such that if $m_{i,j} \neq 0$, $i \neq j$, then $\gamma_j - \gamma_i = -1$, or $\gamma_j - \gamma_i = p - 1$, and for each integer $\swarrow$, $0 \leqq \swarrow \leqq p-1$, there exists some γ_j such that $\gamma_j \equiv \swarrow (\mathrm{mod}\ p)$.*

Any vector $\vec{\gamma}$ with the properties above is called an *ordering vector* for the matrix M.

[1] Proofs and other details are given in Report WAPD-T-567 of the Bettis Plant of the Westinghouse Electric Corporation.

[2] For the case $p = 2$, this reduces to Young's (point) property (A). Young, however, does not assume that the sets S_0 and S_1 are non-empty. This distinction is trivial, since if S_0, say, is empty, then M is a diagonal matrix, and if all $m_{ii} \neq 0$, then the matrix B of (3) is the null matrix.

DEFINITION 3. The $n \times n$ p-cyclic matrix M is *consistently ordered* (relative to the ordering vector $\vec{\gamma}$) if and only if the following hold. If $m_{i,j} \neq 0$, $i \neq j$, then:

1. If $j > i$, then $\gamma_j - \gamma_i = p - 1$.
2. If $j < i$, then $\gamma_j - \gamma_i = -1$.

REMARK. If M is p-cyclic, then by the same permutation of its rows and columns, M can be consistently ordered.

THEOREM 2. *Let the $n \times n$ p-cyclic matrix M be consistently ordered. If $M' = (m'_{i,j})$, and $M'' = (m''_{i,j})$ are defined by*

$$(7) \qquad m'_{i,j} = \begin{cases} m_{i,j}, & i \leqq j \\ \lambda m_{i,j}, & i > j \end{cases}; \qquad m''_{i,j} = \begin{cases} m_{i,j}, & i = j \\ \lambda^{(p-1)/p} m_{i,j}, & i \neq j \end{cases},$$

then for all λ, $\det(M') = \det(M'')$.

The relationship between Definitions 1 and 2 is now brought out through the following

THEOREM 3. *The $n \times n$ matrix M is p-cyclic if and only if the matrix B of (3) is cyclic of index p. Thus, if M is p-cyclic, then the characteristic polynomial $p_n(\lambda) = |\lambda I - B|$ of B is of the form*

$$(8) \qquad p_n(\lambda) = \lambda^{\nu} \prod_{i=1}^{l} (\lambda^p - \mu_i^p) ,$$

where $\nu + lp = n$, and $\mu_i \neq 0$.

From this theorem we see that if M is p-cyclic then its non-diagonal entries define a matrix which is cyclic of index p. The second part of Theorem 3 follows from Romanovsky's result (6), and gives a new proof, for the case $p = 2$, of Young's lemma[3] [11, p. 98] which asserts that, for $p = 2$, the non-zero eigenvalues of the matrix B, of (3), occur in $\pm$ pairs.

Returning to equation (2), it may be written in the form

$$(9) \qquad \vec{x}^{(l+1)} = L_{\sigma,\omega} \vec{x}^{(l)} + \vec{f} ,$$

where $\vec{f}$ is a fixed vector and $L_{\sigma,\omega}$ is a linear operator, where σ denotes the dependence of the equations (2) on the ordering σ of the rows and columns of M.

We can now determine a relationship between the eigenvalues of the matrix B of (3) and the eigenvalues of the matrix $L_{\sigma,\omega}$ of (9).

THEOREM 4. *Let the matrix M be p-cyclic, and let σ denote a*

[3] This result has been presented elsewhere. See [2, pp. 368-369].

consistent ordering. If $\omega \neq 0$, *and if* λ *is a non-zero eigenvalue of* $L_{\sigma,\omega}$, *and if* μ *satisfies*

$$(\lambda + \omega - 1)^p = \lambda^{p-1}\omega^p\mu^p \,, \tag{10}$$

then μ *is an eigenvalue of* B. *If* μ *is an eigenvalue of* B, *and* λ *satisfies the equation above, then* λ *is an eigenvalue of* $L_{\sigma,\omega}$.

We define the *spectral radius*[4] of a matrix C, $\bar{\mu}(C)$, as the maximum of the moduli of its eigenvalues. If $\bar{\mu}(C) < 1$, then C is said to be *convergent*, and

$$R(C) \equiv -\log(\bar{\mu}(C)) \tag{11}$$

is defined as the *rate of convergence* of C.

The iterative *method of simultaneous displacements* [11, p. 100] is defined in general by the vector equation

$$\vec{x}^{(l+1)} = B\vec{x}^{(l)} + \vec{c} \,, \tag{12}$$

where the matrix B is defined in (3). The particular choice of $\omega = 1$ in (2), defining the operator $L_{\sigma,1}$, is called the *Gauss-Seidel* method. Theorem 4 leads immediately to the following generalization of Young's Corollary 2.1 [11, p. 100].

COROLLARY 1. *Let* M *to be a consistently ordered* p*-cyclic matrix. If* μ *is an eigenvalue of* B, *then* μ^p *is an eigenvalue of* $L_{\sigma,1}$; *if* λ *is a non-zero eigenvalue of* $L_{\sigma,1}$, *and if* $\mu^p = \lambda$, *then* μ *is an eigenvalue of* B. *The method of simultaneous displacements converges if and only if the Gauss-Seidel method converges, and if both converge, the latter converges exactly* p *times as fast.*

3. Determination of the optimum relaxation factor. With the $n \times n$ matrix $M = (m_{i,j})$ of (1) a consistently ordered p-cyclic matrix, whose associated matrix B of (3) is convergent ($\bar{\mu} \equiv \bar{\mu}(B) < 1$), we shall determine the optimum overrelaxation factor ω_b, producing fastest convergence in (2), assuming in particular that the eigenvalues of B^p are real and non-negative.[5] (See Corollary 2.) More precisely, ω_b is the unique positive real root (less than $p/(p-1)$) of the equation

[4] This is called the *spectral norm* of C by Young [**11**, p. 94] although it is not a norm in the usual sense. When the matrix B of (3) has non-negative entries, $\bar{\mu}(B)$ is called the *Jacobi constant for the matrix* M of (1) by Ostrowski, Comm. Math. Helv. (1956), 175-210.

[5] Similarly, in the case $p = 2$ of Young, Young [**11**, p. 102] assumes M to be symmetric and positive definite in finding a formula for ω_b. This implies that the eigenvalues of B^2 are real and non-negative.

$$\bar{\mu}^p\omega_b^p = p^p(p-1)^{1-p}(\omega_b - 1) . \tag{13}$$

Note that if $p = 2$, this reduces to Young's equation [11, p. 95] $\bar{\mu}^2\omega_b^2 = 4(\omega_b - 1)$. We shall also assume that $\bar{\mu} > 0$, since the case $\bar{\mu} = 0$ is trivial.

From (10), with $\mu = \bar{\mu}$ and $\lambda = z^p$, we obtain, after taking pth roots, the polynomial equation $z^p - \omega\bar{\mu}z^{p-1} + (\omega - 1) = 0$. We therefore define the polynomial $g_p(z; \omega)$ as

$$g_p(z; \omega) = z^p - \omega\bar{\mu}z^{p-1} + (\omega - 1) . \tag{14}$$

LEMMA 1. *If ω_b and $g_p(z; \omega)$ are defined respectively by* (13) *and* (14), *then:*

1. *For $1 < \omega < \omega_b$, $g_p(z; \omega)$ has precisely two positive real zeros.*
2. *For $\omega = \omega_b$, $g_p(z; \omega)$ has a unique positive real zero of multiplicity two.*
3. *For $\omega_b < \omega < p/(p-1)$, $g_p(z; \omega)$ has no positive real zeros.*

Proof. For $\omega > 1$, let $z = \zeta(\omega - 1)^{1/p}$, and we obtain

$$g_p(z; \omega) = (\omega - 1)[\zeta^p - \varepsilon(\omega)\zeta^{p-1} + 1] \equiv (\omega - 1)h_p(\zeta; \omega) , \tag{15}$$

which defines the polynomial $h_p(\zeta; \omega)$. For $\omega \neq 1$, we note that h_p and g_p have the same zeros. By definition,

$$\varepsilon(\omega) = \omega\bar{\mu}(\omega - 1)^{-1/p} . \tag{16}$$

From Descartes' rule of signs, $g_p(z; \omega)$ has at most two positive real zeros. Since

$$\frac{d\varepsilon^p(\omega)}{d\omega} = \frac{\omega^{p-1}\bar{\mu}^p}{(\omega-1)^2}\{(p-1)\omega - p\} ,$$

the function $\varepsilon(\omega)$ is strictly decreasing for $1 < \omega < \dfrac{p}{p-1}$. Now,

$$\varepsilon(\omega_b) = \frac{p}{(p-1)^{(p-1)/p}} .$$

If $\zeta_0 = (p-1)^{1/p}$, then $h_p(\zeta_0; \omega_b) = 0$, and it can be readily shown that $(p-1)^{1/p}$ is a zero of multiplicity two for $h_p(\zeta; \omega_b)$, which proves part 2. Using the monotonicity of $\varepsilon(\omega)$, we have that $h_p(\zeta_0; \omega) < 0$ for $1 < \omega < \omega_b$, from which we conclude part 1. For all $\zeta \geqq 0$ we have $h_p(\zeta; \omega_b) \geqq 0$, and again using the monotonicity of $\varepsilon(\omega)$, we have for all $\zeta \geqq 0$, that $h_p(\zeta; \omega) > 0$ for $\omega_b < \omega < p/(p-1)$, which proves part 3.

LEMMA 2. *For $1 < \omega < \omega_b$, the function $g_p(z; \omega)$ has a positive real zero greater in modulus than $[(\omega_b - 1)(p-1)]^{1/p}$.*

Proof. For $1 < \omega < \omega_b$, one shows directly that $h_p(\zeta_1\, \omega) < 0$, where

$$\zeta_1 = \left[\frac{(\omega_b - 1)(p-1)}{(\omega - 1)}\right]^{1/p},$$

and the conclusion follows from Lemma 1, part 1.

LEMMA 3. *For any $\omega \leqq 1$, the function $g_p(z; \omega)$ has a zero greater in modulus than $[(\omega_b - 1)(p-1)]^{1/p}$.*

Proof. By definition, $g_p(z; 1) = z^{p-1}(z - \bar{\mu})$. It is trivial to verify that for $\bar{\mu} < 1$ we have $\bar{\mu} > [(\omega_b - 1)(p-1)]^{1/p}$. For $\omega > 1$, it is easily shown that $g_p(\bar{\mu}; \omega) < 0$, and since $g_p(+\infty; \omega) = +\infty$, the result follows.

We now consider the case for $\omega \geqq \omega_b$. From Lemma 1, the real polynomial $g_p(z; \omega)$ has no positive real zeros for $\omega_b < \omega < p/(p-1)$, so we now consider the continuous image, for all $\omega > \omega_b$, of the double root in the upper half (complex) z plane $\mathcal{I}z \geqq 0$. This continuous image $z(\omega)$, $\omega > \omega_b$, of the double root in the upper half plane is shown by a somewhat tedious argument to lie in the annulus[6]

$$[(\omega_b - 1)(p-1)]^{1/p} < |z(\omega)| \leqq [(\omega - 1)(p-1)]^{1/p}, \tag{17}$$

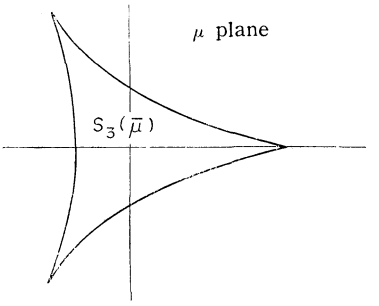
and we have

LEMMA 4. *For $\omega > \omega_b$, the function $g_p(z; \omega)$ has a zero z_0 satisfying $[(\omega_b - 1)(p-1)]^{1/p} < |z_0| \leqq [(\omega - 1)(p-1)]^{1/p}$.*

4. Spectral radius of $L_{\sigma,\omega}$. We consider, for $\omega = \omega_b$, the following equation derived from (10):

$$\mu(z) = \frac{1}{\omega_b}\left\{z + \frac{\omega_b - 1}{z^{p-1}}\right\}, \tag{18}$$

where again $z^p = \lambda$. If $|z| = r_0$, it can be directly verified from (18) that for $r_0^p = (\omega_b - 1)(p-1)$, the exterior of the circle $|z| = r_0$ is mapped conformally onto the exterior of the closed curve $\mu(r_0 e^{i\theta})$. We now denote by $S_p(\bar{\mu})$ the closed set of points which is the complement of the open set of image points $\mu(2)$, where $|z| > r_0$. The case $p = 3$ is illustrated in the adjacent figure.



THEOREM 5. *If all the eigenvalues of B lie in $S_p(\bar{\mu})$, where $0 < \bar{\mu} < 1$, and $\bar{\lambda}(\omega)$ denotes the spectral radius of the operator $L_{\sigma,\omega}$, then*

[6] For $p > 2$, it can be shown that strict inequality throughout is valid in (17).

1. $\bar{\lambda}(\omega_b) = (\omega_b - 1)(p - 1)$.
2. $\bar{\lambda}(\omega) > \bar{\lambda}(\omega_b) = (\omega_b - 1)(p - 1)$ *for all* $\omega \neq \omega_b$.
3. *For* $\omega > \omega_b$, $\begin{cases} \bar{\lambda}(\omega_b) < \bar{\lambda}(\omega) < (\omega - 1)(p - 1) & \text{for } p > 2. \\ \bar{\lambda}(\omega_b) < \bar{\lambda}(\omega) = (\omega - 1) & \text{for } p = 2. \end{cases}$

Proof. With the previous lemmas, the conclusions of Theorem 5 will follow if we show that $\bar{\lambda}(\omega_b) = (\omega_b - 1)(p - 1)$. Since, by the definition of $S_p(\bar{\mu})$, the image of the exterior of $|z| = r_0$ contains no eigenvalue of B, then $\bar{\lambda}(\omega_b) \leqq (\omega_b - 1)(p - 1)$. On the other hand, all eigenvalues μ_i of B satisfy $|\mu_i| \leqq \bar{\mu}$, with equality for some i. By definition, $r_0^p = (\omega_b - 1)(p - 1)$, and it follows from (18) that $\mu(r_0) = \bar{\mu}$. Since, by (18), $|\mu(z)| \leqq \bar{\mu}$ for all $|z| \leqq r_0$, and, by assumption, all eigenvalues of B lie in $S_p(\bar{\mu})$, it follows that the image of the closed exterior of $|z| = r_0$ contains at least one eigenvalue of B, so that $\bar{\lambda}(\omega_b) \geqq (\omega_b - 1)(p - 1)$, and thus $\bar{\lambda}(\omega_b) = (\omega_b - 1)(p - 1)$, which completes the proof.

The mapping $\mu(z)$ of (18) has p-fold symmetry in the sense that if we let $z = re^{i\theta}$, then

$$\mu(re^{i(\theta + 2\pi k/p)}) = e^{i2\pi k/p}\mu(re^{i\theta}) , \tag{19}$$

for $k = 0, 1, \cdots, p - 1$. We now assume that the eigenvalues of B^p are real and non-negative. The closed set $S_p(\bar{\mu})$ contains the segment $0 \leqq \tau \leqq \bar{\mu}$, as well as, by (19), the segments

$$\tau \exp\left(\frac{2\pi ik}{p}\right), \quad 0 \leqq \tau \leqq \bar{\mu}, \qquad k = 1, 2, \cdots, p - 1.$$

The assumption that the eigenvalues of B^p are real and non-negative implies, by Theorem 3, that the eigenvalues of B lie on the p segments $\tau \exp(2\pi ik/p)$, $0 \leqq \tau \leqq \bar{\mu}$, $k = 0, 1, \cdots, p - 1$. Thus, all the eigenvalues of B lie in $S_p(\bar{\mu})$, which gives us the following.

COROLLARY 2. *If all the eigenvalues of* B^p *are real and non-negative, and* $0 < \bar{\mu} < 1$, *then the conclusions of Theorem* 5 *are valid.*

From the proof of Theorem 5, we obtain the following useful result.

COROLLARY 3. *If all the eigenvalues of* B *lie in* $S_p(\bar{\mu}_1)$, $0 < \bar{\mu}_1 < 1$, *then* L_{σ,ω_1} *is convergent, where* ω_1 *is the solution of* (13) *with* $\bar{\mu} = \bar{\mu}_1$.

Using the definition of (11), we now compare the quantities $R(L_{\sigma,\omega_b})$ and $R(L_{\sigma,1})$.

THEOREM 6. *If ω_b satisfies* (13), *then as $\bar{\mu} \to 1-$, we have*

$$R(L_{\sigma,\omega_b}) \sim \left(\frac{2p}{p-1}\right)^{1/2} [R(L_{\sigma,1})]^{1/2} .$$

Proof. Since $R(L_{\sigma,\omega_b}) = -\log[(\omega_b - 1)(p-1)]$, and $R(L_{\sigma,1}) = -p/n\bar{\mu}$, the result follows upon applying L'Hospital's rule twice to

$$\frac{R(L_{\sigma,\omega_b})}{[R(_{\sigma,1})]^{1/2}} .$$

Combining the results of Theorem 6 and Corollary 1, we have the following

COROLLARY 4. *If $R(B)$ denotes the rate of convergence of the method of simultaneous displacements, then as $\bar{\mu} \to 1-$, we have*

$$R(L_{\sigma,\omega_b}) \sim \left[\frac{2p^2}{p-1}\right]^{1/2} [R(B)]^{1/2} .$$

We remark first that, by Theorem 6, generalized successive overrelaxation (with the optimum overrelaxation factor) always gives an order of magnitude improvement in the rate of convergence for the Gauss-Seidel method, even though the coefficient $[2p/(p-1)]^{1/2}$ strictly decreases with p. On the other hand, from the result of Corollary 4, generalized successive overrelaxation, while giving an order of magnitude improvement for the rate of convergence of the method of simultaneous displacements, is associated with the coefficient $[2p^2/(p-1)]^{1/2}$, which is strictly *increasing* with p.

Finally, we mention three facts, obtained by Young in the case $p = 2$, which extend to the general case. First, overestimating or underestimating $\bar{\mu}$, respectively, results in an overestimate or an underestimate of ω_b. Second, overestimating $\bar{\mu}$ by a small amount does not cause an appreciable decrease in $R(L_{\sigma,\omega_b})$, but, on the other hand, underestimating $\bar{\mu}$ causes a larger relative decrease in $R(_{\sigma,\omega_b})$. Finally, if the Jordan normal form of B^p is diagonal, then[7]

$$(20) \qquad \sup_{v \neq \vec{0}} \frac{\| L^m_{\sigma,\omega_b}[\vec{v}] \|}{\| \vec{v} \|} \sim m[(\omega_b - 1)(p-1)]^{m-1}, \qquad m \to \infty .$$

In other words, for any $p \geqq 2$, the largest degree of the elementary divisors [10, Chap. III] of L_{σ,ω_b} is two.

---

[7] The quantity $\| v \|$ is the *norm* of the vector $\vec{v}$. If the component of $\vec{v}$ are v_i, $i=1, 2, \cdots, n$, then $\| \vec{v} \| \equiv [\sum_{i=1}^{n} |v_i|^2]^{1/2}$.

5. Applications. As a first application, we limit ourselves to plane connected domains Ω, with boundary S, which are a union of a finite number of equilateral triangles of side h. Let S_h denote the nodes, or mesh points, which are points of S, and let R_h be all other nodes. If we consider the real function $u(x)$, where x is a node, $x \in R_h$, then interpreting x as a point in the complex plane, we have that $x + h\omega_j$ is also a node of the triangular mesh, where ω_j is any complex root of $z^3 + 1 = 0$.

Consider the following special numerical approximation to the Dirichlet problem for Ω:

$$u(x) - \frac{1}{3}\sum_{j=1}^{3} u(x + h\omega_j) = 0, \qquad x \in R_h , \tag{21}$$

where the values $u(x)$, $x \in S_h$, are prescribed. The matrix M determined by this system of linear equations is 3-*cyclic*, and generalized successive overrelaxation applies in this case.

As a second application, we consider the Peaceman-Rachford iteration scheme [7]. (See also [4] and [3].) We suppose that the matrix M of (1) can be expressed in form

$$M = H + V, \tag{22}$$

where H and V are symmetric and positive definite, and are similar to tridiagonal or Jacobi matrices. With equation (22), (1) is equivalent, for any scalar ρ, to each of the equations

$$(H + \rho I)\vec{x} = \vec{k} - (V - \rho I)\vec{x} , \tag{23}$$

$$(V + \rho I)\vec{x} = \vec{k} - (H - \rho I)\vec{x} . \tag{23'}$$

The Peaceman-Rachford implicit alternating direction method consists in using the implicit process defined by

$$(H + \rho_n I)\vec{x}_n^* = \vec{k} - (V - \rho_n I)\vec{x}_n , \tag{24}$$

$$(V + \rho_n I)\vec{x}_{n+1} = \vec{k} - (H - \rho_n I)\vec{x}_n^* , \tag{24'}$$

for a suitable sequence of non-negative scalars $\rho_1, \rho_2, \cdots$. Since the matrices H and V are tridiagonal, equations (24)-(24') can be rapidly solved by means of Gauss elimination. In typical computational applications, only a finite number of non-negative scalars $\rho_1, \rho_2, \cdots, \rho_l$ are used, these values being repeated cyclically in (24)-(24').

The concept of a cyclic matrix, as given in Definition 1, can now be profitably applied to Peaceman-Rachford iteration scheme. The basic equations (23)-(23') are clearly equivalent to

$$\begin{pmatrix} H+\rho I & V-\rho I \\ H-\rho I & V+\rho I \end{pmatrix}\begin{pmatrix} \vec{x} \\ \vec{x} \end{pmatrix} = \begin{pmatrix} \vec{k} \\ \vec{k} \end{pmatrix}, \tag{25}$$

where $\begin{pmatrix} \vec{k} \\ \vec{k} \end{pmatrix}$ is a column vector with $2n$ components.

For $\rho > 0$, the diagonal blocks of the matrix of (25) are non-singular, and (25) is equivalent to:

$$\vec{u} = B_\rho \vec{u} + \vec{g}\ , \tag{26}$$

where

$$B_\rho = \begin{pmatrix} 0 & (H+\rho I)^{-1}(\rho I - V) \\ (V+\rho I)^{-1}(\rho I - H) & 0 \end{pmatrix}, \tag{27}$$

$$\left[\vec{g} = \begin{pmatrix} \vec{g}_1 = (H+\rho I)^{-1} & \vec{k} \\ \vec{g}_2 = (V+\rho I)^{-1} & \vec{k} \end{pmatrix}. \right. \tag{27'}$$

By definition, B_ρ is cyclic of index 2. It can be show that B_ρ satisfies (block) property (A^π), and is consistently ordered in the sense of [1]. For fixed $\rho > 0$, the Gauss-Seidel (block) iteration scheme applied to (26) gives

$$\begin{aligned} \vec{x}_1^{(m+1)} &= (H+\rho I)^{-1}(\rho I - V)\vec{x}_2^{(m)} + (H+\rho I)^{-1}\vec{k}, \\ \vec{x}_2^{(m+1)} &= (V+\rho I)^{-1}(\rho I - H)\vec{x}_1^{(m+1)} + (V+\rho I)^{-1}\vec{k}\ , \end{aligned} \tag{28}$$

or equivalently,

$$\begin{aligned} (H+\rho I)\vec{x}_1^{(m+1)} &= (\rho I - V)\vec{x}_2^{(m)} + \vec{k}, \\ (V+\rho I)\vec{x}_2^{(m+1)} &= (\rho I - H)\vec{x}_1^{(m+1)} + \vec{k}\ . \end{aligned} \tag{29}$$

This, except for notation, is equivalent to (24)-(24') for a single fixed factor ρ. Generalizing, if the Peaceman-Rachford iteration scheme uses ℓ parameters $\rho_1, \rho_2, \cdots, \rho_\ell$ cyclically, then we consider

$$\vec{u} = B_{(\rho_\ell)}\vec{u} + \vec{g}\ , \tag{30}$$

where

$$B_{\{\rho_l\}} \equiv \begin{bmatrix} 0 & 0 & \cdots & 0 & T_1 \\ U_1 & 0 & & 0 & 0 \\ 0 & T_2 & & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & U_l & 0 \end{bmatrix}, \tag{31}$$

and

$$T_j \equiv (H + \rho_j I)^{-1}(\rho_j I - V);\ U_j \equiv (V + \rho_j I)^{-1}(\rho_j I - H)\ . \tag{31'}$$

By definition, $B_{\{\rho_l\}}$ is *cyclic of index* 2ℓ, and $\vec{u}$ is a column vector with $2n\ell$ components. Morever, if we use the extension of the results of §§2-3 to block overrelaxation, as was stated to be possible in the introduction, then $B_{\{\rho_l\}}$ is a consistently ordered 2ℓ-cyclic matrix. The application of the Gauss-Seidel block iteration scheme to (30), with $B_{\{\rho_l\}}$ defined in (31), is *exactly* the Peaceman-Rachford scheme with ℓ parameters used cyclically. This gives us the following

THEOREM 7. *The Peaceman-Rachford iteration scheme* (24)-(24′) *is the Gauss-Seidel block iteration scheme applied to the* 2ℓ*-cyclic matrix* $B_{\{\rho_l\}}$ *of* (31).

The special case $\rho_1 = \rho_2 = \cdots = \rho_l$ is admitted in the preceding result.

Since the Peaceman-Rachford scheme can be considered as the Gauss-Seidel block iteration scheme applied to a consistently order 2ℓ-cyclic matrix, the results of §4 strongly suggest the application of successive over-relaxation to the Peaceman-Rachford scheme. For the case $\ell = 1$, over-relaxation applied to (28) results in

$$\vec{x}_1^{(m+1)} = \omega\{(H + \rho I)^{-1}(\rho I - V)\vec{x}_2^{(m)} + (H + \rho I)^{-1}\vec{k} - \vec{x}_1^{(m)}\} + \vec{x}_1^{(m)}\ , \tag{32}$$

$$\vec{x}_2^{(m+1)} = \omega\{(V + \rho I)^{-1}(\rho I - H)\vec{x}_1^{(m+1)} + (V + \rho I)^{-1}\vec{k} - \vec{x}_2^{(m)}\} + \vec{x}_2^{(m)}\ . \tag{32'}$$

Unfortunately, it is hard to predict the rate of convergence of iterative schemes based on (32)-(32′), because their spectra are difficult to estimate.

REFERENCES

1. L. J. Arms, L. D. Gates, and B. Zondek, *A method of block iteration*, J. Soc. Ind. Appl. Math. **4** (1956), 220-229.

2. Garrett Birkhoff and Richard S. Varga, *Reactor criticality and non-negative matrices*, J. Soc. Ind. Appl. Math. **6** (1958), 354-377.

3. ———, *Implicit alternating direction methods*, to appear in the Trans. Amer. Math. Soc.
4. J. Douglas, Jr. and H. H. Rachford, Jr., *On the numerical solution of conduction problems in two and three space variables*, Trans. Amer. Math. Soc. **82** (1956), 421-439.
5. S. Frankel, *Convergence rates of iterative treatments of partial differential equations*, Math. Tables and other Aids to Cmputation **4** (1950), 65-75.
6. G. Frobenius, *Über Matrizen aus nicht negativen Elementen*, S.-B. Akad. Wiss. Berlin, (1912), 456-477.
7. D. W. Peaceman and H. H. Rachford, Jr., *The numerical solution of parabolic and elliptic differential equations*, J. Soc. Ind. Appl. Math. **3**, (1955), 28-41.
8. V. Romanovsky, *Recherches sur les chaines de Markoff*, Acta Math. **66** (1936), 147-251.
9. R. S. Varga, *Numerical solution of the two-group diffusion equations in $x - y$ geometry*, IRE Trans. Professional Group on Nuclear Science, NS-**4** (1957), 52-62.
10. J. Wedderburn, *Lectures on matrices*, Amer. Math. Soc. Colloquium Publications, **17**, New York, 1934.
11. David Young, *Iterative methods for solving partial difference equations of elliptic type*, Trans. Amer. Math. Soc. **76** (1954), 92-111.

BETTIS ATOMIC POWER DIVISION
OF
WESTINGHOUSE ELECTRIC CORPORATION

PACIFIC JOURNAL OF MATHEMATICS

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be typewritten (double spaced), and the author should keep a complete copy. Manuscripts may be sent to any one of the four editors. All other communications to the editors should be addressed to the managing editor, L. J. Paige at the University of California, Los Angeles 24, California.

50 reprints per author of each article are furnished free of charge; additional copies may be obtained at cost in multiples of 50.

The *Pacific Journal of Mathematics* is published quarterly, in March, June, September, and December. The price per volume (4 numbers) is $12.00; single issues, $3.50. Back numbers are available. Special price to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: $4.00 per volume; single issues, $1.25.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 2120 Oxford Street, Berkeley 4, California.

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

The Supporting Institutions listed above contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its content or policies.

Pacific Journal of Mathematics
Vol. 9, No. 2 June, 1959