

On the Extrapolated Jacobi or Simultaneous Displacements Method in the Solution of Matrix and Operator Equations

By W. V. Petryshyn

1. Introduction. Let $A = (a_{ij})$ be an $n \times n$ real or complex nonsingular matrix with $a_{ii} \neq 0$ and let b be a given column vector. If we put $A = D + Q$, where D is a diagonal matrix with $d_{ii} = a_{ii}$ and $Q \equiv A - D$, then one of the simplest and oldest iterative methods for the solution of equation

$$(i) \quad Ax = b$$

is the *Jacobi* or *simultaneous displacements method* according to which the iterants $x^{(m+1)}$ are determined by

$$(ii) \quad a_{ii}x_i^{(m+1)} = -\sum_{j \neq i} a_{ij}x_j^{(m)} + b_i, \quad 1 \leq i \leq n, \quad m \geq 0,$$

or, in matrix form, by

$$(iii) \quad Dx^{(m+1)} = -Qx^{(m)} + b,$$

where $x^{(0)}$ is an arbitrary initial approximation to the exact solution x^* of (i). If $e^{(m)}$ denotes the error vector $e^{(m)} \equiv x^{(m)} - x^*$, then from (iii) we derive

$$(iv) \quad e^{(m+1)} = Je^{(m)} = \dots = J^{m+1}e^{(0)},$$

where $J \equiv -D^{-1}Q$ is the *Jacobi iteration matrix* associated with A . It is known that the Jacobi method converges rather slowly. The standard technique of improving its convergence is that of extrapolation or relaxation which in our case leads to the extrapolated Jacobi or simultaneous displacement method

$$(v) \quad Dx^{(m+1)} = -\{(\omega - 1)D + \omega Q\}x^{(m)} + \omega b$$

with the extrapolated Jacobi matrix given by

$$(vi) \quad J_\omega \equiv -\{(\omega - 1) + \omega D^{-1}Q\},$$

where ω is the relaxation parameter which is a real number. It is known [17] that for an arbitrary $x^{(0)}$ the error vectors $e^{(m+1)}$ of these iterative methods tend to the zero vector if and only if the J_ω tend to the zero matrix, or equivalently, if and only if the spectral radius $r(J_\omega)$ ($\equiv \max_{1 \leq i \leq n} |\lambda_i(J_\omega)|$) of J_ω is less than unity.

However, since the evaluation of the eigenvalues of a general matrix is very complicated, it is difficult to guarantee in advance that the convergence conditions will be satisfied, and theoretical results for the Jacobi and extrapolated Jacobi methods, in spite of their being long in existence, are available only for very special classes of matrices. Furthermore, almost all of these results give only sufficient conditions for convergence. Thus, if the matrix A is strictly diagonally dominant,¹ von Mises

Received June 15, 1964.

¹ For the precise definitions of various concepts used in the Introduction, see the succeeding sections of this paper.

and Geiringer [8] proved that the Jacobi method is convergent (see also Wittmeyer [21] and Collatz [1]). Later, Geiringer [5] proved that both the Jacobi and Gauss-Seidel methods are convergent for irreducibly diagonally dominant¹ matrices A . Similar results were obtained by Collatz [2], Weissinger [19], and others.² If A is symmetric and possesses the property A ,¹ Young [22] showed that the Jacobi method converges if and only if A is positive definite while Stein and Rosenberg [14] proved its convergence in the case in which A is a real symmetric positive definite matrix with nonpositive off-diagonal entries. Only recently³ Householder [6] (see also Gavrilov [4], Wegner [18], and Newman [9]) showed that if $A = I + Q$ is symmetric then (iii) converges if and only if $(I + Q)$ and $(I - Q)$ are positive definite. Furthermore, Ostrowski [10], using a very complicated finite matrix theory, was able to derive both necessary and sufficient conditions for the convergence of (iii) for the class of the so-called real H -matrices A as well as other important results concerning free steering methods, variable parameter ω , etc. An analogue to the extrapolated method was investigated by Keller [7].

The purpose of this article is first to characterize and establish various properties of a class of K -symmetric and K -positive definite matrices \mathcal{K}_n and then by simple arguments to derive both the necessary and sufficient conditions for the convergence of the *generalized extrapolated Jacobi method* for A in \mathcal{K}_n of which the ordinary and extrapolated Jacobi methods become just special cases. Afterwards, we establish the connection between the conditions of the above authors and ours for the matrices of class \mathcal{K}_n and investigate the problem of determining the optimum parameter ω of the extrapolated process. In contradistinction to the procedures used by the above authors which essentially are restricted to matrices of finite order our procedures appear to be free of this restriction. Accordingly, the last section of the article deals with the extension of the results to operator equations in Hilbert space.

Let us point out that though the results derived in this article are mostly theoretical in nature, it is hoped that they will also prove to be of practical interest. We shall consider this aspect of the above results as well as other applications elsewhere.

2. K -symmetric and K -positive definite matrices. Let X_n be a real or complex n -dimensional linear vector space with inner product and norm defined by

$$(1) \quad x^*y \equiv (x, y), \quad \|x\| = (x, x)^{1/2},$$

where $x^* = (\bar{x}_1, \dots, \bar{x}_n)$ is the adjoint of the column vector x .

*Definition 1.*⁴ An $n \times n$ real or complex matrix $A = (a_{ij})$ will be called K -symmetric if there is at least one $n \times n$ Hermitian and positive definite matrix K such

² For the extensive treatment of various other contributions of these and other authors (especially the earlier ones) to the investigation of these and related iterative methods see the excellent monographs by Varga [17] and by Forsythe and Wasow [3].

³ The author is grateful to the referee for various suggestions concerning this paper and particularly for calling his attention to some of the papers on this subject.

⁴ For the properties of bounded and unbounded K -p.d. operators in Hilbert space H see Petryshyn [12], [13].

that for all x and y in X_n

$$(2) \quad (Ax, Ky) = (Kx, Ay)$$

and K -positive definite (K -p.d.) if for all $x \neq 0$ in X_n

$$(3) \quad (Ax, Kx) > 0.$$

It is easy to show that (2) is equivalent to the condition

$$(2_0) \quad KA = A^*K$$

while (3) is equivalent to the requirement that KA be positive definite, written

$$(3_0) \quad KA > 0.$$

The class of all real or complex $n \times n$ matrices A which have properties (2) and (3) will be denoted by \mathfrak{K}_n . The following theorem gives a convenient characterization of \mathfrak{K}_n .

THEOREM 1. *The following statements characterizing A in \mathfrak{K}_n are equivalent:*

- (a) A is K -symmetric and K -p.d.
- (b) A has positive real eigenvalues and a complete set of corresponding eigenvectors.
- (c) A is weakly positive in the sense of Wigner [20].
- (d) A can be written in the form $A = H_1H_2$, where H_1 and H_2 are two Hermitian and positive definite matrices.

Proof. We shall prove Theorem 1 by showing that

$$(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a).$$

(a) \Rightarrow (b): Let A be K -symmetric and K -p.d. Then, by Definition 1, there exists a Hermitian positive definite matrix K such that $T \equiv KA$ is Hermitian and positive definite. Hence $A = K^{-1}T$ and therefore, as is known, has the properties of (b).

(b) \Rightarrow (c): Let A have positive eigenvalues and a complete set of corresponding eigenvectors. Then A is similar to a diagonal matrix D , i.e., there exists a nonsingular matrix W and a diagonal matrix D whose diagonal elements are the positive eigenvalues of A which satisfy $A = WDW^{-1}$. Hence A is weakly positive in the sense of Wigner [20].

(c) \Rightarrow (d): Let A be weakly positive. Then there is a nonsingular matrix W and a positive diagonal matrix D such that $A = WDW^{-1}$. If we define H_1 and H_2 by $H_1 \equiv WW^*$ and $H_2 \equiv (W^*)^{-1}DW^{-1}$, then it is easy to verify that H_1 and H_2 are Hermitian and positive definite and that $A = H_1H_2$.

(d) \Rightarrow (a): Assuming A to be of the form $A = H_1H_2$ with H_1 and H_2 Hermitian and positive definite we see that when we let $K \equiv H_1^{-1}$, then $KA = A^*K$ and $KA = H_2 > 0$. Hence, by Definition 1 or its equivalent, A is K -symmetric and K -p.d. This completes the proof of Theorem 1.

Remark 1. Let us observe that when the matrices A and B belong to \mathfrak{K}_n then in general their sum $A + B$ does not necessarily belong to \mathfrak{K}_n . For example, it is easy to verify that the 2×2 matrices

$$(E_1) \quad A_1 = \begin{bmatrix} 11 & 1 \\ 7 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 11 & 8 \\ 1 & 1 \end{bmatrix}$$

belong to \mathcal{K}_2 while their sum

$$A_1 + A_2 = \begin{bmatrix} 22 & 9 \\ 8 & 2 \end{bmatrix}$$

does not. Of course, a sufficient condition for the matrix $aA + bB$ to belong to \mathcal{K}_n , whenever A and B belong to \mathcal{K}_n and $a > 0$ and $b > 0$, is that there is at least one common K with respect to which both A and B are K -symmetric and K -p.d. That this is not a necessary condition we see from the following example: The matrices

$$(E_2) \quad A_3 = \begin{bmatrix} 8 & 3 \\ 2 & 1 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 8 & 2 \\ 3 & 1 \end{bmatrix}$$

belong to \mathcal{K}_2 but there does not exist a common K with respect to which both A_3 and A_4 would be K -symmetric and K -p.d. For if there were such a K , say,

$$K = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$$

with $a_{11} > 0$, $a_{22} > 0$, and $a_{11}a_{22} - a_{12}^2 > 0$, then it would have to satisfy the conditions (2₀) and (3₀) which in our case reduce to the requirement that

$$3a_{11} - 2a_{22} = 2a_{11} - 3a_{22}.$$

The latter implies that $a_{11} = -a_{22}$. This is, however, impossible if K is to be a positive definite matrix. Nevertheless, the sum

$$A_3 + A_4 = \begin{bmatrix} 16 & 5 \\ 5 & 2 \end{bmatrix}$$

does belong to \mathcal{K}_2 with $K = I$, where I is the 2×2 identity matrix.

Remark 2. Let us also note that when A belongs to \mathcal{K}_n then A^{-1} exists and belongs to \mathcal{K}_n . This follows from Theorem 1(b).

For later use we shall need the following lemma.

LEMMA 1. *Let N and M be $n \times n$ real or complex matrices which are K -symmetric and let N be also K -p.d. Then the eigenvalues of*

$$(\alpha) \quad Mu = \lambda Nu$$

are real, the eigenvectors u_i and u_j corresponding to distinct eigenvalues λ_i and λ_j are orthogonal in the sense that

$$(\beta) \quad (Nu_i, Ku_j) = 0,$$

and the set of eigenvectors $\{u_i\}$ corresponding to the eigenvalues $\{\lambda_i\}$ of (α) is complete in X_n .

Proof. The proof of Lemma 1 can be carried out directly or by means of the following device. Evidently, the eigenvalue problem (α) is equivalent to the problem

$$(\alpha_0) \quad Tu = \lambda u,$$

where $T \equiv N^{-1}M$. If now in X_n we introduce the new inner product and the corresponding equivalent norm by

$$(\gamma) \quad [x, y] = (Nx, Ky), \quad |x| = [x, x]^{1/2}, \quad x, y \in X_n,$$

then it is easy to show that relative to the inner product (γ) the matrix T is Hermitian. Consequently, by the known results for Hermitian matrices, the eigenvalues of T are real, the eigenvectors u_i and u_j corresponding to two distinct eigenvalues λ_i and λ_j are orthogonal in the (γ) -metric, and the set of eigenvectors $\{u_i\}$ corresponding to the eigenvalues $\{\lambda_i\}$ is complete in the (γ) -norm and hence, in view of norm equivalence, in the X_n -norm. These are precisely the assertions of Lemma 1.

3. The Generalized Extrapolated Jacobi Method. Consider the problem of solving the matrix equation

$$(4) \quad Ax = b,$$

where b is a given column vector and A is a nonsingular $n \times n$ real or complex matrix of the form

$$(5) \quad A = D + Q$$

in which D is a K -symmetric and K -p.d. matrix and Q is K -symmetric. Let x_0 be an arbitrary initial approximation to the solution x^* of (4) and let ω be a real parameter. In analogy to (v) we determine the succeeding iterants $x_1, x_2, \dots, x_{n+1}, \dots$, by

$$(6) \quad Dx_{n+1} = -\{(\omega - 1)D + \omega Q\}x_n + \omega b$$

which we shall call the *generalized extrapolated Jacobi method*. Observe that D is not assumed to be a diagonal matrix. Evidently, the ordinary Jacobi and extrapolated Jacobi method, as well as other methods, are special cases of (6) for special choices of D, Q , and ω . The corresponding iteration matrix J_ω is given by

$$(7) \quad J_\omega \equiv -\{(\omega - 1) + \omega D^{-1}Q\}$$

and is called the *generalized extrapolated Jacobi matrix*.

THEOREM 2. *The necessary and sufficient conditions for the convergence of the generalized extrapolated Jacobi method (6) is that both matrices $A = D + Q$ and $G_\omega \equiv ((2 - \omega)/\omega)D - Q$ (or $-A$ and $-G_\omega$) be K -p.d. for some nonempty set Ω (or Ω^0) of ω .*

Proof. (Sufficiency): Let us define the matrix P_ω by $P_\omega \equiv -\{(\omega - 1)D + \omega Q\}$ and observe that, in view of the K -symmetry of D and Q , G_ω and P_ω are also K -symmetric. Furthermore, this and the definitions of G_ω and P_ω imply that for every vector z in X_n we have the identity

$$(8) \quad \omega^2(G_\omega z, Kz) (Az, Kz) = (Dz, Kz)^2 - (P_\omega z, Kz)^2.$$

To see this note that, in view of (5) and the definition of G_ω and P_ω , we have

$$\begin{aligned} \omega^2(G_\omega z, Kz) (Az, Kz) &= ([(2 - \omega)D - \omega Q]z, Kz) ([\omega D + \omega Q]z, Kz) \\ &= ([D - P_\omega]z, Kz) ([D + P_\omega]z, Kz) \end{aligned}$$

from which (8) follows. Assume now that A and G_ω (or $-A$ and $-G_\omega$) are both K -p.d. for $\omega \in \Omega$ (or Ω^0) and let $\lambda_\omega \equiv \lambda(\omega)$ be an arbitrary eigenvalue of J_ω and u its corresponding eigenvector. Then $J_\omega u = \lambda_\omega u$ or, in view of (7) and the definition of P_ω , $P_\omega u = \lambda_\omega Du$. Since D is k -symmetric and K -p.d. and P_ω is K -symmetric, Lemma 1 implies that λ_ω is real. Hence $(P_\omega u, Ku)^2 = \lambda_\omega^2 (Du, Ku)^2$ and, by (8),

$$(9) \quad \omega^2(G_\omega u, Ku) (Au, Ku) = (Du, Ku)^2 [1 - \lambda_\omega^2].$$

Since, by assumption, A and G_ω (or $-A$ and $-G_\omega$) are both K -p.d., (9) implies that $|\lambda_\omega| < 1$. Consequently, the method (6) converges.

(Necessity): Let us first note that in virtue of the properties of D and P_ω , Lemma 1 implies that all the eigenvalues $\lambda_i(\omega)$ of J_ω are real and the corresponding set of eigenvectors $\{u_i\}$ forms a complete set in X_n which we can assume to be orthogonal in the sense that whenever $i \neq j$

$$(10) \quad (Du_i, Ku_j) = 0.$$

Let z be an arbitrary vector in X_n . Then there are scalars α_i such that

$$z = \sum_{i=1}^n \alpha_i u_i$$

and, in view of (10),

$$(11) \quad (Dz, Kz) = \sum_{i=1}^n |\alpha_i|^2 (Du_i, Ku_i).$$

Similarly, since $P_\omega u_i = \lambda_i Du_i$,

$$(12) \quad (P_\omega z, Kz) = \sum_{i=1}^n \lambda_i |\alpha_i|^2 (Du_i, Ku_i).$$

If we define

$$(13) \quad \bar{\lambda}(\omega) \equiv \bar{\lambda}_\omega \equiv \max_{1 \leq i \leq n} |\lambda_i(\omega)|,$$

then from (12) and (11) we see that for every z in X_n

$$|(P_\omega z, Kz)| \leq \bar{\lambda}(\omega) \sum_{i=1}^n |\alpha_i|^2 (Du_i, Ku_i) = \bar{\lambda}(\omega) (Dz, Kz).$$

This inequality and the identity (8) show that for every z in X_n

$$\omega^2 (G_\omega z, Kz) (Az, Kz) \geq (Dz, Kz)^2 [1 - \bar{\lambda}(\omega)].$$

Assuming now that $\bar{\lambda}(\omega) < 1$, the last inequality shows that A and G_ω (or $-A$ and $-G_\omega$) are both K -p.d.; this completes the proof of Theorem 2.

Determination of the Sets Ω and Ω^0 . Let $\{\mu_i\}$, $i = 1, 2, \dots, n$, be the set of eigenvalues of $J_1 \equiv -D^{-1}Q$, and let $\{v_i\}$ be the corresponding set of eigenvectors. Then $Qv_i = -\mu_i Dv_i$ and for any vector z in X_n

$$(14) \quad (Az, Kz) = \sum_{i=1}^n (1 - \mu_i) |\alpha_i|^2 (Dv_i, Kv_i)$$

and

$$(15) \quad (G_\omega z, Kz) = \sum_{i=1}^n \left(\frac{2 - \omega}{\omega} + \mu_i \right) |\alpha_i|^2 (Dv_i, Kv_i).$$

Since D is K -p.d., (14) and (15) imply that A and G_ω are K -p.d. if and only if

$$(16) \quad 1 - \mu_i > 0, \quad \frac{2 - \omega}{\omega} + \mu_i > 0, \quad i = 1, 2, \dots, n,$$

respectively. Furthermore, in view of the fact that A , G_ω , and D are K -p.d. and $(Az, Kz) + (G_\omega z, Kz) = 2(Dz, Kz)/\omega$, ω must be positive and hence, by (16), must lie in the set Ω determined by

$$(17) \quad 0 < \omega < \frac{2}{1 - \mu_m},$$

where μ_m is the algebraically smallest eigenvalue of $J_1 = -D^{-1}Q$.

On the other hand, if D is K -p.d., then $-A$ and $-G_\omega$ are K -p.d. if and only if

$$(18) \quad 1 - \mu_i < 0, \quad \frac{2 - \omega}{\omega} + \mu_i < 0, \quad i = 1, 2, \dots, n,$$

respectively. By similar reasoning we find that in this case ω must be negative and, by (18), must lie in the set Ω^0 determined by

$$(19) \quad \frac{2}{1 - \mu_M} < \omega < 0,$$

where μ_M is the algebraically largest eigenvalue of J_1 .

Remark 3. Let us remark that Theorem 2 remains valid if instead of D we assume that $-D$ is K -p.d. As above, we again have to distinguish two cases. In the first case $-A$ and $-G_\omega$ are K -p.d. if and only if $\mu_i < 1$ and $0 < \omega < 2/(1 - \mu_m)$ while in the second case A and G_ω are K -p.d. if and only if

$$1 < \mu_i \quad \text{and} \quad \frac{2}{(1 - \mu_M)} < \omega < 0.$$

COROLLARY 1. *The generalized Jacobi method ($\omega = 1$):*

$$(20) \quad Dx_{n+1} = -Qx_n + b$$

converges if and only if $A = D + Q$ and $G_1 = D - Q$ are K -p.d.

Special Cases. Let us observe that by specializing D , Q , and ω we get from (6) and (20) different iterative procedures. Of course, in each case, D and Q have to satisfy the above conditions. In particular, if $a_{ii} > 0$ for all i and D is chosen to be the diagonal matrix with $d_{ii} = a_{ii}$ which is K -symmetric, then (20) and (6) reduce to the ordinary Jacobi and extrapolated Jacobi methods, respectively. We shall not dwell here on other procedures. However, we will mention one important special case.

COROLLARY 2. (a) *If A is Hermitian with $a_{ii} > 0$ and D is a diagonal matrix with $d_{ii} = a_{ii}$, then the extrapolated Jacobi method converges if and only if*

$$A = D + Q > 0 \quad \text{and} \quad G_\omega = \frac{2 - \omega}{\omega} D - Q > 0$$

or $A < 0$ and $G_\omega < 0$.

(b) *In particular, Jacobi method converges if and only if the matrices $A = D + Q$ and $G_1 = D - Q$ are both positive definite.*

Let us note³ that Corollary 2(b) was proved by Householder [6] for the case $D = I$. In [4] Gavrilov derived equivalent necessary and sufficient conditions for the validity of Corollary 2(b).

Example. To exhibit our simple necessary and sufficient conditions for the con-

vergence of ordinary Jacobi method consider the following example of a symmetric matrix A_0 discussed by Forsythe and Wasow [3]:

$$(E_3) \quad A_0 = \begin{bmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{bmatrix},$$

where a is a real number. Using the fact that $A_0 > 0$ if and only if its principal minors are positive, it is readily verified that $A_0 > 0$ when $-\frac{1}{2} < a < 1$; however, the Jacobi method converges only if $-\frac{1}{2} < a < \frac{1}{2}$ since, as was observed in [3], $\lambda_1 = -2a$ and $\lambda_2 = \lambda_3 = a$, being the eigenvalues of the Jacobi iteration matrix J , are in absolute value less than unity only for $-\frac{1}{2} < a < \frac{1}{2}$.

If in addition to $A_0 > 0$ we demand, in accordance with Corollary 2, that $G_1 = D - Q > 0$, then it is easy to verify that $G_1 > 0$ when $-1 < a < \frac{1}{2}$. Thus, $A_0 > 0$ and $G_1 > 0$ together if and only if $-\frac{1}{2} < a < \frac{1}{2}$. This verifies the condition on a obtained in [3] by explicitly computing the eigenvalues of J .

Remark 4. It should be remarked that Theorem 2 is essentially a special case of a more general Theorem A proved by the author [11] for operators in Hilbert space which for the space X_n can be formulated as follows:

THEOREM A. *Let D be a real or complex $n \times n$ K -symmetric matrix and let A be the matrix of the form $A = D + S + Q$. If S^* denotes the adjoint of S and Ω the set of real numbers $\omega > 0$ such that the matrices D , S , and Q have the properties that $(D + \omega S)$ is nonsingular and $G(\omega) \equiv ((2 - \omega)/\omega)D + S^* - Q$ is K -symmetric and K -p.d. for ω in Ω , then the sequence $\{x_{n+1}\}$ determined by the go-method,*

$$(j) \quad (D + \omega S)x_{n+1} = -\{(\omega - 1)D + \omega Qx_n\} + wb$$

converges to the solution x^* of $Ax = b$ if and only if A is K -p.d.

A closer examination reveals that the identities (9) and (10) derived in [11], on which Theorem A is based, remain valid when $S = 0$. Thus, in this case the go-method (j) reduces to the generalized extrapolated method (6) and Theorem A reduces to a theorem analogous to Theorem 2. However, it was to avoid the use of these identities, whose proof is rather involved and complicated, that for the matrices of class \mathcal{K}_n we presented here a simple proof of Theorem 2.

4. Connection Between Our and Other Conditions for $A \in \mathcal{K}_n$.

(A) *Strictly or Irreducibly Diagonally Dominant Matrices.* Let us recall [17] that the matrix A is called *diagonally dominant* if

$$(21) \quad |a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$$

for all i , *strictly diagonally dominant* if “ $>$ ” holds in (21) for all i , and *irreducibly diagonally dominant* if A is irreducible and “ $>$ ” holds in (21) for at least one i , where we say that A is *irreducible* if by no interchange of rows and of corresponding columns is it possible to reduce A to the form

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

in which A_{11} and A_{22} are square submatrices or, equivalently, if there does not exist a permutation matrix P such that the matrix PAP^{-1} has the above form. The example (E₃) for a in the interval $-\frac{1}{2} < a < \frac{1}{2}$ suggests that the strict diagonal dominance of A_0 implies our conditions: $A > 0$ and $G_1 > 0$. In fact, Theorem 3 below shows that this is the case for a general strictly or irreducibly diagonally dominant K -symmetric matrix A .

THEOREM 3. *If A is K -symmetric and strictly or irreducibly diagonally dominant with real positive diagonal elements and D is the diagonal matrix with $d_{ii} = a_{ii}$ which is also K -symmetric, then $A = D + Q$ and $G_1 = D - Q$ are both K -p.d.*

Proof. Let us first note that the K -symmetry of A and D implies the K -symmetry of G_1 . Furthermore, if $A = D + Q$ is strictly or irreducibly diagonally dominant, then

$$r_i \equiv \sum_{j \neq i} |a_{ij}| \leq a_{ii}$$

and, consequently, $G_1 = D - Q$ is also strictly or irreducibly diagonally dominant. Hence Gershgorin's principle and irreducibility imply that both A and G_1 are non-singular [15]. Since $a_{ii} > 0$, the union of the disks $|z - a_{ii}| \leq r_i$ contains only points in the complex plane z having their real parts positive. Hence, all the eigenvalues of A and of G_1 have positive real parts since, by Gershgorin's principle, they must lie in the union of these disks. But, being real, they are therefore positive. Finally, Theorem 1 implies that both A and G_1 are K -p.d.

Incidentally, as an immediate consequence of Theorem 3 we obtain the following generalization of the known result for Hermitian matrices [17].

COROLLARY 3. *If $A = (a_{ij})$ is a K -symmetric strictly or irreducibly diagonally dominant matrix with positive real diagonal elements a_{ii} , then A is K -p.d.*

(B) *Two-Cyclic Matrices.* Another sufficient condition for the convergence of the Jacobi method was derived by Young [22] for real symmetric and positive definite matrices A with the so-called "property A." In general we say that a given $n \times n$ matrix $A = (a_{ij})$ has *property A* or equivalently is *cyclic of order 2* [17] if there exists an $n \times n$ permutation matrix P such that

$$(22) \quad \tilde{A} \equiv PAP^T = \begin{bmatrix} D_{11} & -Q_{12} \\ -Q_{21} & D_{22} \end{bmatrix} = \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix} + \begin{bmatrix} 0 & -Q_{12} \\ -Q_{21} & 0 \end{bmatrix},$$

where D_{11} and D_{22} are square diagonal matrices of order p and $r = n - p$, respectively. Let us note that P permutes the diagonal elements of A among themselves. Observe that if A is K -symmetric and we define a new Hermitian and positive definite matrix \tilde{K} by

$$(23) \quad \tilde{K} \equiv PKP^T = \begin{bmatrix} K_{11} & K_{12}^* \\ K_{12} & K_{22} \end{bmatrix},$$

where the partition of \tilde{K} corresponds to that of \tilde{A} , then it is not hard to verify that \tilde{A} is \tilde{K} -symmetric if and only if A is K -symmetric and \tilde{A} is \tilde{K} -p.d. if and only if A is K -p.d. Hence, without loss of generality, we may assume that A is of the form (22), i.e., that A is a 2-cyclic matrix, and take K to be in the corresponding partitioned form (23). Furthermore, to simplify the discussion, we shall assume

that K is of the form

$$(24) \quad K = \begin{bmatrix} K_{11} & 0 \\ 0 & K_{22} \end{bmatrix}$$

and that the matrix

$$(25) \quad D = \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix}$$

is K -symmetric. Under these conditions the following theorem establishes the connection between our conditions and those of Young [22] for the matrices of class \mathfrak{K}_n .

THEOREM 4. *If A is 2-cyclic and K -symmetric, where K is of the form (24), then $A = D + Q$ is K -p.d. if and only if $G \equiv D - Q$ is K -p.d., where*

$$Q = \begin{bmatrix} 0 & -Q_{12} \\ -Q_{21} & 0 \end{bmatrix}.$$

Proof. Let us first observe that our conditions on A and D imply that G is K -symmetric. Let u be any vector in X_n and, corresponding to the partition of A , let u be partitioned into its p - and r -dimensional components u_1 and u_2 , respectively, so that

$$u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

A simple calculation shows that the quadratic form (Gu, Ku) is given by

$$(26) \quad (Gu, Ku) = (Du, Ku) + (Q_{12}u_2, K_{11}u_1) + (Q_{21}u_1, K_{22}u_2).$$

Consider the matrix

$$(27) \quad E = \begin{bmatrix} I_1 & 0 \\ 0 & -I_2 \end{bmatrix},$$

where I_1 and I_2 are the p - and r -dimensional identities, respectively, and then consider the transformation of variables

$$(28) \quad u' = \begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} I_1 & 0 \\ 0 & -I_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ -u_2 \end{bmatrix}.$$

It is not hard to see that the quadratic form (Au', Ku') is given by

$$(29) \quad (Au', Ku') = (Du, Ku) + (Q_{12}u_2, K_{11}u_1) + (Q_{21}u_1, K_{22}u_2).$$

Since E is a nonsingular matrix the identities (26) and (29) imply the validity of Theorem 4.

COROLLARY 4. *If A is 2-cyclic and symmetric, then $A = D + Q$ is positive definite if and only if $G = D - Q$ is positive definite.*

(C) *Real M -Matrices.* Following the notation of Varga [17] we shall say that a real $n \times n$ matrix $A = (a_{ij})$ with $a_{ij} \leq 0$ for all $i \neq j$ is an M -matrix if A is nonsingular and $A^{-1} \geq 0$, where " $>$ " means that all the entries r_{ij} of A^{-1} are non-negative real numbers. The following theorem establishes for A in \mathfrak{K}_n the

relationship between our conditions and those of an M -matrix and exhibits also a loose connection with some of the results of Stein and Rosenberg [14].

THEOREM 5. *Let $A = (a_{ij})$ be a real K -symmetric matrix with $a_{ij} \leq 0$ for all $i \neq j$ and such that the diagonal matrix D with $d_{ii} = a_{ii}$ is also K -symmetric. Then A is an M -matrix if and only if $A = D + Q$ and $G = D - Q$ are both K -p.d.*

Proof. The proof of Theorem 5 is essentially based on Corollary 1 above and Theorem 3.10 in [17] according to which A is an M -matrix if and only if the diagonal entries of A are positive real numbers and $J = -D^{-1}Q$ is a convergent matrix with non-negative entries (i.e., $J \geq 0$ and J^k converges to zero matrix when $k \rightarrow \infty$).

To prove Theorem 5 let us first observe that, by virtue of our conditions on A and D , G is also K -symmetric.

If we assume now that A is an M -matrix, then by Theorem 3.10 [17] the diagonal entries of A are real positive numbers and the Jacobi matrix $J = -D^{-1}Q$ is a convergent matrix with non-negative entries. Hence, by Corollary 1, both A and G are K -p.d.

To prove the converse, suppose that A and G are K -p.d. Then it follows that D is K -p.d. and, by assumption, also K -symmetric. Hence D is nonsingular and belongs to the class \mathcal{K}_n . Consequently, by Theorem 1, the eigenvalues λ_i of D are positive real numbers. Since D is a diagonal matrix, $\lambda_i = d_{ii} = a_{ii} > 0$. Thus, the entries of the Jacobi matrix $J = -D^{-1}Q$ are non-negative and, by Corollary 1, J is convergent. Consequently, by Theorem 3.10 [17], A is an M -matrix.

Let us note that if in Theorem 5 we have $D = I$, then $J = -Q$ is the Jacobi matrix with non-negative real entries considered by Stein and Rosenberg [14]. If G_s denotes the Gauss-Seidel matrix $(I - L)^{-1}U$, where L and U are, respectively, the strictly lower and strictly upper triangular matrices such that $L + U = J$, then by Stein-Rosenberg Theorem the matrices J and G_s are either both convergent or both divergent. Thus, combining this result with Theorem 5 we have the corollary:

COROLLARY 5. *Let $A = (a_{ij})$ be a real K -symmetric matrix with $a_{ij} \leq 0$ for all $i \neq j$ and $a_{ii} = 1$ for all i . Then the following three statements are equivalent:*

- (a) G_s is convergent,
- (b) $A = I + Q$ and $G = I - Q$ are K -p.d.,
- (c) A is an M -matrix.

Remark 5. Let us recall that if, in addition, A is symmetric (i.e., $K = I$), and positive definite, then it is called [17] a Stieltjes matrix and for such matrices Theorem 5 yields

COROLLARY 6. *If A is a Stieltjes matrix, then $G = D - Q$ is positive definite and, consequently, the Jacobi method converges.*

(D) *Example.* Consider the following simple example of 3×3 real symmetric matrix.

$$(E_4) \quad A_5 = \begin{bmatrix} 1 & \alpha & \beta \\ \alpha & 1 & \alpha \\ \beta & \alpha & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & \alpha & \beta \\ \alpha & 0 & \alpha \\ \beta & \alpha & 0 \end{bmatrix}$$

which, for a certain range of values of α and β , is neither strictly nor irreducibly

diagonally dominant, nor 2-cyclic nor an M -matrix, but for which the associated Jacobi matrix

$$(30) \quad J = -D^{-1}Q = \begin{bmatrix} 0 & -\alpha & -\beta \\ -\alpha & 0 & -\alpha \\ -\beta & -\alpha & 0 \end{bmatrix}$$

is convergent. Using again the fact that a symmetric matrix is positive definite if and only if its principal minors are positive, it is easy to verify that A_6 and

$$(31) \quad G_6 \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & \alpha & \beta \\ \alpha & 0 & \alpha \\ \beta & \alpha & 0 \end{bmatrix}$$

are simultaneously positive definite if and only if α and β are such that

$$(32) \quad -\frac{\sqrt{2}}{2} < \alpha < \frac{\sqrt{2}}{2}, \quad 2\alpha^2 - 1 < \beta < 1 - 2\alpha^2.$$

Consequently, by Corollary 1, for any α and β satisfying (32) the Jacobi method converges.

Observe that we would have come to the same conclusion had we computed or estimated the eigenvalues of J , i.e., the roots of the characteristic equation

$$p(\lambda) \equiv \lambda^3 - (2\alpha^2 + \beta^2) + 2\alpha^2\beta = 0.$$

In fact, it is not hard to show that the roots of $p(\lambda) = 0$ are in absolute value less than 1 if α and β are such that for all $\lambda \geq 1$

$$(33) \quad \lambda^3 - (2\alpha^2 + \beta^2) - 2\alpha^2|\beta| > 0.$$

Solving the inequality (33) we find that it holds precisely, for α and β satisfy conditions (32). Thus, Corollary 1 is numerically verified for this particular example.

5. The Best Iteration Process. In this section we consider the problem of determining the value of ω which gives the best convergence of the generalized extrapolated Jacobi method (6). We know that if e_n denotes the error vector $x_n - x^*$, then from (6) we obtain

$$(34) \quad e_{n+1} = J^{n+1} e_0,$$

where e_0 is the initial error, and

$$(35) \quad \|e_{n+1}\| \leq \|J_\omega\|^{n+1} \|e_0\|.$$

Therefore, the smaller the norm $\|J_\omega\|$ of the matrix J_ω the faster the process (6) converges. Hence, our problem is to determine the value of ω in Ω for which $\|J_\omega\|$ is smallest.

Observe that if, in analogy to (γ) , we define in X_n the new metric by

$$(36) \quad [x, y] = (Dx, Ky), \quad |x| = [x, x]^{1/2}, \quad x, y \in X_n,$$

then relative to (36), the matrices J_ω and J_1 are Hermitian and for any x in X_n

$$(37) \quad (1 - \omega + \omega\mu_m) |x|^2 \leq [J_\omega x, x] < (1 - \omega + \omega\mu_M) |x|^2.$$

Consequently,

$$(38) \quad E(\omega) \equiv \|J_\omega\| = \max \{ |1 - \omega + \omega\mu_m|, |1 - \omega + \omega\mu_M| \}.$$

Thus our problem is to find this value $\tilde{\omega} \in \Omega$ at which the function $E(\omega)$, which can be written in the form

$$(39) \quad E(\omega) = \frac{1}{2} \{ |1 - \omega + \omega\mu_m| + |1 - \omega + \omega\mu_M| + ||1 - \omega + \omega\mu_m| - |1 - \omega + \omega\mu_M|| \},$$

attains its minimum.

The representation (39) implies that for ω in $0 < \omega < 1/(1 - \mu_m)$ the function $E(\omega)$ is given by

$$(40) \quad \begin{aligned} E(\omega) &= \frac{1}{2} \{ 2 - 2\omega + (\mu_m + \mu_M)\omega + |(\mu_m - \mu_M)\omega| \} \\ &= (\mu_M - 1)\omega + 1 \equiv E_1(\omega). \end{aligned}$$

Similarly for $\omega > 1/(1 - \mu_M)$ ($> 1/(1 - \mu_m)$) we obtain

$$(41) \quad \begin{aligned} E(\omega) &= \frac{1}{2} \{ 2\omega - 2 - (\mu_m + \mu_M)\omega + |(\mu_M - \mu_m)\omega| \} \\ &= (1 - \mu_m)\omega - 1 \equiv E_2(\omega) \end{aligned}$$

and for ω such that $1/(1 - \mu_m) \leq \omega \leq 1/(1 - \mu_M)$ the function $E(\omega)$ becomes

$$(42) \quad \begin{aligned} E(\omega) &= \frac{1}{2} \{ 2(\mu_m - \mu_M)\omega + |2\omega - 2 - (\mu_m + \mu_M)\omega| \} \\ &= \begin{cases} (\mu_M - 1)\omega + 1 & \text{if } \frac{1}{1 - \mu_m} \leq \omega \leq \varpi, \\ (1 - \mu_m)\omega - 1 & \text{if } \varpi \leq \omega \leq \frac{1}{1 - \mu_M}, \end{cases} \end{aligned}$$

where

$$(43) \quad \varpi = \frac{2}{2 - (\mu_m + \mu_M)}.$$

Thus we see from (40)-(43) that in the interval $(0, \tilde{\omega}]$ the function $E(\omega)$ is given by the monotonically decreasing function $E_1(\omega)$ while in the interval (ϖ, ∞) it is given by the monotonically increasing function $E_2(\omega)$, i.e.,

$$(44) \quad E(\omega) = \begin{cases} E_1(\omega) \equiv (\mu_M - 1)\omega + 1 & \text{for } 0 < \omega \leq \varpi, \\ E_2(\omega) \equiv (1 - \mu_m)\omega - 1 & \text{for } \omega \geq \varpi. \end{cases}$$

The minimum of $E(\omega)$ is attained at $\omega = \varpi$ and is given by

$$(45) \quad E(\varpi) = E_1(\varpi) = E_2(\varpi) = \frac{\mu_M - \mu_m}{2 - (\mu_m + \mu_M)}.$$

For symmetric matrices relations (43) and (45) were derived in a slightly different way in [3], [7] and other papers.

Proceeding in a similar way we find that when $-A$ and $-G_\omega$ are K -p.d. then the value $\omega \in \Omega^0$ which minimizes the norm $E^0(\omega) \equiv \|J_\omega\|$ of the corresponding

operator J_ω is this point $\omega^0 \in \Omega^0$ at which the function

$$(46) \quad E^0(\omega) = \begin{cases} E_1^0(\omega) \equiv (\mu_m - 1)\omega + 1 & \text{for } \omega^0 \leq \omega < 0, \\ E_2^0(\omega) \equiv (1 - \mu_M)\omega - 1 & \text{for } \omega \leq \omega^0, \end{cases}$$

where

$$(47) \quad \omega^0 = \frac{2}{2 - (\mu_m + \mu_M)}.$$

Thus we see from (46) that $E^0(\omega) \equiv \|J_\omega\|$ attains its minimum at ω^0 given by

$$(48) \quad E(\omega) = E_1^0(\omega^0) = E_2^0(\omega^0) = \frac{\mu_m - \mu_M}{2 - (\mu_m + \mu_M)}.$$

The relations (44) and (46) show that $E(\omega)$ and $E^0(\omega)$ are represented graphically by Figures E and E⁰, respectively.

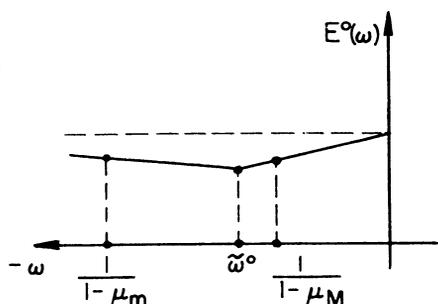


Fig. E⁰

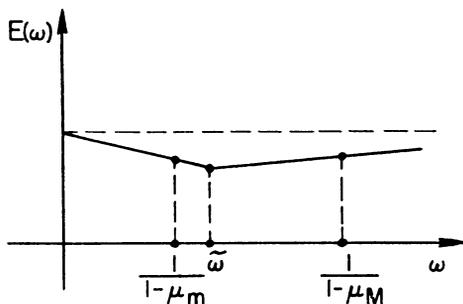


Fig. E

Remark 6. Let us remark that if, as is standard for matrices of finite order (see for example [7]), the rate of convergence R of the process (6) is defined by

$$(49) \quad R(\omega) \equiv -\log(\tilde{\lambda}(\omega)), \quad \tilde{\lambda}(\omega) \equiv \max_{1 \leq i \leq n} |\lambda_i(\omega)| < 1,$$

then it is known that the number of iterations necessary to reduce the initial error e_0 by a given factor is inversely proportional to R . Hence our problem in this case is to find an ω at which $R(\omega)$ assumes its maximum. Since $R = \log(1/\tilde{\lambda}(\omega))$ we see that the convergence is best for this value of ω^* for which $\lambda(\omega^*) = \min_\omega \tilde{\lambda}(\omega)$. It is interesting to observe that in our case we find, by simple manipulation, that $\omega^* = \omega$, where ω is given by (43), and

$$R(\omega^*) = \|J(\omega)\| = \min_{\omega \in \Omega} E(\omega) = \frac{\mu_M - \mu_m}{2 - (\mu_m + \mu_M)}.$$

We prefer, however, our approach to the problem of finding the best parameter since it is directly extendable to the iterative processes for solving linear operator equations in infinite-dimensional Hilbert spaces.

6. Skew- K -Symmetric Matrices. In analogy to (2₀) we say that a real or complex matrix $B = (b_{ij})$ is skew- K -symmetric if there exists at least one Hermitian and

positive definite matrix K such that

$$(50) \quad KB + B^*K = 0.$$

Let us remark that if we define $\tilde{B} \equiv iB$, where $i = \sqrt{-1}$, then $\tilde{B}^* = -iB^*$ and

$$(51) \quad K\tilde{B} - \tilde{B}^*K = i(KB + B^*K)$$

from which it follows that B is skew- K -symmetric if and only if \tilde{B} is K -symmetric. This and Theorem 1 yield the validity of the following

LEMMA 2. *B is skew- K -symmetric if and only if the eigenvalues of B are pure imaginary and the set of corresponding eigenvectors of B span the space X_n .*

If we now consider the problem of solving iteratively, by the generalized Jacobi method (20), the equation

$$(52) \quad Ax = b, \quad A \equiv D + Q,$$

where A is a nonsingular real or complex matrix for which D is K -symmetric and K -p.d. and Q is skew- K -symmetric, then the following theorem which is an analogue to Corollary 1 justifies the applicability of (20) to the solution of (52).

THEOREM 6. *If D is K -symmetric and K -p.d., Q is skew- K -symmetric, and $A = D + Q$ nonsingular, then the generalized Jacobi method (20) converges if and only if the matrices $\tilde{A} \equiv D + iQ$ and $\tilde{G} \equiv D - iQ$ are both K -p.d.*

Proof. Theorem 6 follows immediately from Corollary 1 and the above remark. In fact, $\tilde{A} = D + iQ$ and $\tilde{G} = D - iQ$ are both K -symmetric and consequently, by Corollary 1, the eigenvalues $\lambda(\tilde{J})$ of $\tilde{J} \equiv -iD^{-1}Q$ are in modulus less than 1 if and only if \tilde{A} and \tilde{G} are both K -p.d. Since the eigenvalues $\lambda(J)$ of $J = -D^{-1}Q$ are such that $|\lambda(J)| = |\lambda(\tilde{J})|$, the assertion of Theorem 6 follows.

Example. Let us apply Theorem 6 to the following simple example considered by Stein and Rosenberg [14]: $A \equiv I + Q$, where

$$(E_5) \quad Q = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & 0 & \frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \end{bmatrix}, \quad I = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Evidently, Q is skew- K -symmetric for $K = I$. It is not hard to verify that the matrices $\tilde{A} \equiv I + iQ$ and $\tilde{G} \equiv I - iQ$ are both Hermitian and positive definite. Hence, by Theorem 6, the ordinary Jacobi method converges, thus verifying the assertion by Stein and Rosenberg.

7. Extension to Linear Operator Equations in Hilbert Space. Let H be a real or complex Hilbert space with inner product $(,)$ and norm $\| \| = (,)^{1/2}$. A bounded linear operator A defined on all of H will be called K -symmetric if there exists a bounded linear Hermitian and positive definite operator K such that

$$(53) \quad (Ax, Ky) = (Kx, Ay), \quad x, y \in H,$$

and K -p.d. if for some constant $\alpha > 0$ and all $x \in H$

$$(54) \quad (Ax, Kx) \geq \alpha \| x \|^2.$$

It is known [13] that if A is K -p.d., then A^{-1} exists and is a bounded linear operator defined on its range $R(A) = H$.

Let us point out that most of the results obtained in the previous sections remain valid in a modified form for operators A in H . In fact, the following analogues to Theorems 1 and 2 are true:

THEOREM 7. *The following statements are equivalent.*

(a) A is K -symmetric and K -p.d.

(b) A can be written in the form $A = H_1H_2$, where H_1 and H_2 are two Hermitian and positive definite operators.

(c) A is weakly positive in the sense of Wigner, i.e., A can be expressed in the form $A = WDW^{-1}$, where D is positive definite and W and W^{-1} are bounded and defined on all of H .

Proof. (a) \Rightarrow (b). Since (53) holds for all x and y in H and the operators A and K are bounded, it follows that $KA = A^*K$, i.e., A^*K is Hermitian. Moreover, by (54), A^*K is also positive definite. Hence $A = H_1H_2$, where $H_1 = K^{-1}$ and $H_2 = A^*K$.

(b) \Rightarrow (c). This case was shown by Wigner [20]. In fact, if $A = H_1H_2$, where H_1 and H_2 are Hermitian and positive definite, then H_2 has the Hermitian and positive definite square root $H_2^{1/2}$. Hence, if we define $W \equiv H_2^{-1/2}$ and $D \equiv H_2^{1/2}H_1H_2^{1/2}$, then D is Hermitian and positive definite and $A = WDW^{-1}$.

(c) \Rightarrow (a). Let $A = WDW^{-1}$, where D is Hermitian and positive definite and W and W^{-1} are defined on all of H . Then, since $(W^{-1})^* = (W^*)^{-1}$, the operator $(W^*)^{-1}W^{-1}$ is Hermitian and positive definite and, as is easily verified, $KA = A^*K$ and KA is positive definite, i.e., A is K -symmetric and K -p.d.

Let us remark that, in analogy to Remark 2, it is easy to see that if A is K -symmetric and K -p.d., then so is A^{-1} .

Before stating a weaker version of Theorem 2 let us recall that a scalar λ is said to belong to the *resolvent set* $\rho(A)$ of A if λ is such that the range $R(\lambda - A) = H$ and $(\lambda - A)$ has a continuous inverse. The set $\rho(A)$ is an open set in the λ -plane and all λ not in $\rho(A)$ comprise the closed set $\sigma(A)$ called the *spectrum* of A . It is known [16] that when A is also Hermitian and the real numbers $m(A)$ and $M(A)$ are defined by

$$(56) \quad m(A) \equiv \inf_{\|x\|=1} (Ax, x), \quad M(A) \equiv \sup_{\|x\|=1} (Ax, x),$$

then the set $\sigma(A)$ is contained in the closed interval $[m(A), M(A)]$, the endpoints $m(A)$ and $M(A)$ actually belong to $\sigma(A)$, and

$$(57) \quad \|A\| = \max \{ |m(A)|, |M(A)| \}.$$

In solving the functional equation

$$(58) \quad Ax = b,$$

where b is a given element in H and A is a linear bounded operator in H of the form

$$(59) \quad A = D + Q,$$

where D is K -symmetric and K -p.d., and Q is K -symmetric, then in analogy to (6) we start with an arbitrary element x_0 in H and determine the succeeding iterants

$x_1, x_2, \dots, x_{n+1}, \dots$ by

$$(60) \quad Dx_{n+1} = -\{(\omega - 1)D + \omega Q\}x_n + \omega b,$$

or equivalently by

$$(61) \quad x_{n+1} = J_\omega x_n + g,$$

where $g \equiv \omega D^{-1}b$ and

$$(62) \quad J_\omega \equiv (1 - \omega)D - \omega D^{-1}Q.$$

Let us introduce a new metric in H by means of

$$(63) \quad [x, y] = (Dx, Ky), \quad |x| = [x, x]^{1/2}, \quad x, y \in H,$$

and denote H with this new metric by H_0 . In view of the fact that D is K -symmetric and K -p.d., the new metric (63) is well defined, the new norm $|x|$ is equivalent to the old $\|x\|$, and, considered in H_0 , the operators $J_1 \equiv -D^{-1}Q$ and $J_\omega \equiv (1 - \omega)I + \omega J_1$ are Hermitian. Let m and M be two real numbers such that for all x in H

$$(64) \quad m(Dx, Kx) \leq (-Qx, Kx) \leq M(Dx, Kx),$$

and let Ω_σ be the set of real numbers ω such that the operator $G_\omega \equiv ((2 - \omega)/\omega)D - Q$ is K -p.d. for each fixed ω in Ω_σ .

Observe that from (61) it follows by induction that

$$(65) \quad x_{n+1} = \sum_{i=0}^n J_\omega^i g + J_\omega^{n+1} x_0$$

from which we see that the sequence of approximations $\{x_{n+1}\}$ converges to the solution x^* of (58) for every b in H and any x_0 in H if the series $\sum_{i=0}^\infty J_\omega^i$ converges. As is well known [16], the latter converges if the spectrum $\sigma(J_\omega)$ of J_ω lies in the interior of the unit circle which, of course, in our present case reduces to the interior of the unit interval.

The following theorem, which as was already observed in Remark 4, is a special case of the Theorem in [11], is the closest analogue to Theorem 2 for general operators in H .

THEOREM 8. *If D, Q, K , and the set Ω_σ satisfy the above conditions, then the spectrum $\sigma(J_\omega)$ lies in the interior of the unit circle if and only if the operator A is K -p.d.*

To determine the optimum parameter for the process (60) let us assume that $1 - M > 0$. Then, by virtue of (64), it is easy to see that A is K -p.d., the set of Ω of ω determined by

$$(66) \quad 0 < \omega < \frac{2}{1 - m}$$

satisfies the conditions imposed on Ω_σ , and for such an ω and all x in H

$$(67) \quad (1 - \omega + \omega m) |x|^2 \leq [J_\omega x, x] \leq (1 - \omega + \omega M) |x|^2.$$

Hence, by Theorem 8, the method (60) converges and, in view of (57) and (67),

$$(68) \quad |J_\omega| \leq E(\omega) \equiv \max \{|(1 - \omega + \omega m)|, |(1 - \omega + \omega M)|\}.$$

Repeating the arguments of Section 5 we find that the smallest value of $E(\omega)$ is attained at

$$(69) \quad \omega = \frac{2}{2 - (M + m)}$$

and is given by

$$(70) \quad E(\omega) = \frac{M - m}{2 - (M + m)}.$$

Thus the best iteration process (60) is attained for ω given by (69) and the convergence is at least as fast as a geometric progression with ratio

$$\frac{M - m}{2 - (M + m)}.$$

Special Case. If $\omega = 1$, then (60) reduces to the generalized Jacobi method for operator equations in H .

Remark 7. As a final remark let us observe that when the extrapolated Jacobi method (60) is written in the form

$$(71) \quad Dx_{n+1} = Dx_n - \omega r_n$$

or in the equivalent form

$$(71_0) \quad x_{n+1} = x_n - \omega D^{-1}r_n,$$

where $r_n \equiv Ax_n - b$, then condition (66) indicates the range of ω for which convergence of (71₀) is assured and the generalized extrapolated Jacobi method written in the form (71₀) is at least formally identical with the *iterative method with relative minimal errors* (3.6) investigated by the author in [12] if in (3.6) all t_n are to take the same value ω and $B = D^{-1}$. Let us add that B so chosen satisfies all the conditions specified in [12].

Department of Mathematics and Institute for Computer Research
The University of Chicago

1. L. COLLATZ, "Fehlerabschätzung für das Iterationsverfahren zur Auflösung linearer Gleichungssysteme," *Z. Angew. Math. Mech.*, v. 22, 1942, p. 357-361. MR 5, 50.
2. L. COLLATZ, "Über die Konvergenzkriterien bei Iterationsverfahren für lineare Gleichungssysteme," *Math. Z.*, v. 53, 1950, p. 149-161. MR 12, 361.
3. G. W. FORSYTHE & W. R. WASOW, *Finite-Difference Methods for Partial Differential Equations*, Applied Mathematical Series, Wiley, New York, 1960. MR 23 #B3156.
4. YU. M. GAVRILOV, "On the convergence of iterative processes and criteria of sign-definiteness of quadratic forms," *Izv. Akad. Nauk SSSR Ser. Mat.*, v. 18, 1954, p. 87-94. (Russian) MR 16, 177.
5. H. GEIRINGER, "On the solution of systems of linear equations by certain iteration methods," *Reissner Anniversary Volume, Contributions to Applied Mechanics*, J. W. Edwards, Ann Arbor, Mich., 1948, p. 365-393. MR 10, 574.
6. A. S. HOUSEHOLDER, *On the Convergence of Matrix Iterations*, Rep. ORNL-1883, Oak Ridge National Laboratory, Oak Ridge, Tenn., 1955. MR 17, 790.
7. H. B. KELLER, *Numerical Methods*, Lecture Notes, New York University, 1961.
8. R. VON MISES & H. POLLACZEK-GEIRINGER, "Practische Verfahren der Gleichungsauflösung," *Z. Angew. Math. Mech.*, v. 9, 1929, p. 58-77.
9. M. NEWMAN, "Two Theorems on Matrices," *J. Res. Nat. Bur. Standards Sect. B*, v. 66B, 1962, p. 91-92.
10. A. M. OSTROWSKI, "Determinanten mit überwiegender Hauptdiagonale und die absolute Konvergenz von linearen Iterationsprozessen," *Comment. Math. Helv.*, v. 30, 1956, p. 175-210. MR 17, 898.

11. W. V. PETRYSHYN, "On the generalized over-relaxation method for operator equations," *Proc. Amer. Math. Soc.*, v. 14, 1963, p. 917-924.
12. W. V. PETRYSHYN, "Direct and iterative methods for the solution of linear operator equations in Hilbert space," *Trans. Amer. Math. Soc.*, v. 105, 1962, p. 136-175. MR 26 #3180.
13. W. V. PETRYSHYN, "On a class of K -p.d. and non- K -p.d. operators and operator equations," *J. Math. Anal. Appl.* (To be published.)
14. P. STEIN & R. L. ROSENBERG, "On the solution of linear simultaneous equations by iteration," *J. London Math. Soc.*, v. 23, 1948, p. 111-118. MR 10, 485.
15. O. TAUSSKY, "A recurring theorem on determinants," *Amer. Math. Monthly*, v. 56, 1949, p. 672-676. MR 11, 307.
16. A. E. TAYLOR, *Introduction to Functional Analysis*, Wiley, New York, 1958. MR 20 #5411.
17. R. S. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N. J., 1962. MR 28 #1725.
18. U. WEGNER, "Contributi alla teoria dei procedimenti iterativi per la risoluzione numerica dei sistemi di eguazioni lineari algebriche," *Atti Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez. I*, (8), v. 4, 1953, p. 1-48. MR 15, 66.
19. J. WEISSINGER, "Zur Theorie und Anwendung der Iterationsverfahrens," *Math. Nachr.*, v. 8, 1952, p. 193-212. MR 14, 478.
20. E. P. WIGNER, "On weakly positive matrices," *Canad. J. Math.*, v. 15, 1963, p. 313-317. MR 26 #4188.
21. H. WITTMAYER, "Über die Lösung von linearen Gleichungssysteme durch Iteration," *Z. Angew. Math. Mech.*, v. 16, 1936, p. 301-310.
22. D. YOUNG, "Iterative methods for solving partial difference equations of elliptic type," *Trans. Amer. Math. Soc.*, v. 76, 1954, p. 92-111. MR 15, 562.