

# Extensions of the Ostrowski-Reich Theorem for SOR Iterations

James M. Ortega\*

*Department of Mathematics  
North Carolina State University  
Raleigh, North Carolina 27650*

and

Robert J. Plemmons†

*Departments of Computer Science and Mathematics,  
University of Tennessee,  
Knoxville, Tennessee 37916*

Dedicated to Alston S. Householder  
on the occasion of his seventy-fifth birthday.

Submitted by Emeric Deutsch

---

## ABSTRACT

The Ostrowski-Reich theorem states that for a system  $Ax = b$  of linear equations with  $A$  nonsingular, if  $A$  is hermitian and if the diagonal of  $A$  is positive, then the SOR method converges for each relaxation parameter in  $(0, 2)$  if and only if  $A$  is positive definite. This is actually a special case of the Householder-John theorem, which states that for  $A = M - N$  with  $A, M$  nonsingular, if  $A$  is hermitian and  $M^* + N$  is positive definite, then  $M^{-1}N$  is a convergent matrix if and only if  $A$  is positive definite. Our purposes here are to generalize the Householder-John theorem and to provide an insight into how and why the SOR method can converge. As a result the Ostrowski-Reich theorem is extended in two directions; one is when  $A$  is hermitian but the diagonal of  $A$  is not necessarily positive, so that  $A$  is not necessarily positive definite, and the other is when  $A + A^*$  is positive definite but  $A$  is not necessarily hermitian. In the process, several other convergence results are obtained for general splittings of  $A$ . However, no claims are made concerning the ease in which the convergence results obtained here can be applied to practical situations.

---

---

\*Part of the work of the first author was supported under NASA Contract No. NAS1-14101 while in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665.

†The work of the second author was supported by the U.S. Army Research Office under Grant No. DAAG29-77-G-0166, and part of the work was accomplished while he was visiting the Department of Mathematical Sciences, Clemson University, Clemson, SC 29631.

## I. INTRODUCTION

Given a system  $Ax=b$  of linear equations with nonsingular coefficient matrix  $A$ , we will be concerned with the convergence of the iterative process

$$x^{k+1} = Hx^k + b, \quad k = 0, 1, \dots, \quad (1.1)$$

to the solution  $x=A^{-1}b$ . Here  $H$  is the matrix obtained from some splitting of  $A$ ; that is,

$$H = M^{-1}N, \quad A = M - N, \quad (1.2)$$

where it is assumed that  $M^{-1}$  exists.

Our primary interest will be in the SOR iterations. For that purpose we assume that the diagonals  $a_{ii}$  of  $A$  are all nonzero and split  $A$  into

$$A = D - L - U, \quad (1.3)$$

where, as usual,  $D$ ,  $-L$ , and  $-U$  denote the diagonal, strictly lower, and strictly upper triangular parts of  $A$ , respectively. Then for the SOR iterations we have, for the relaxation parameter  $\omega$ ,

$$M = \frac{1}{\omega}(D - \omega L), \quad N = \frac{1}{\omega}[(1 - \omega)D + \omega U] \quad (1.4)$$

and call

$$H_\omega = (D - \omega L)^{-1}[(1 - \omega)D + \omega U] \quad (1.5)$$

the SOR iteration matrix associated with  $\omega$ .

Now the Ostrowski-Reich theorem states that *if  $A$  is hermitian and if the diagonal of  $A$  is positive, then the SOR method converges* [i.e.,  $\rho(H_\omega) < 1$ , where  $\rho(H_\omega)$  denotes the spectral radius] *for all  $\omega$  in  $(0, 2)$ , if and only if  $A$  is positive definite*. The result was established for the Gauss-Seidel iteration (i.e.,  $\omega=1$ ) by Reich [16] and for the general case (i.e.  $0 < \omega < 2$ ) by Ostrowski [15].

But the Ostrowski-Reich theorem is often misquoted to conclude that the SOR method can converge only if  $A$  is definite. However, this is clearly not true, as will be illustrated in Sec. 2 by an analysis of the  $2 \times 2$  case.

It turns out that the Ostrowski-Reich theorem is a corollary to a more general result, which we will call the *Householder-John theorem*. The theorem states that *for the splitting (1.2), if  $A$  is hermitian and  $M^* + N$  is positive definite, then  $\rho(M^{-1}N) < 1$  if and only if  $A$  is positive definite*. The Ostrowski-Reich theorem follows immediately by noting that if  $M$  and  $N$  are specified by (1.4), then

$$M^* + N = \frac{1}{\omega}(2 - \omega)D,$$

which is then positive definite with  $D$  for  $0 < \omega < 2$ .

The sufficiency part of the above result was established by John [9], as reported by Keller [10]. A similar result was established earlier by Weissinger [19], with the matrix  $M^* + N$  replaced by  $M + N$  and under the stronger hypothesis that  $M + N$  is hermitian. The necessity part of the result was given by Householder [8], although he generously attributes it to Reich [16].

In this paper we generalize somewhat the Householder-John theorem and attempt to provide some insights into how and why the SOR method can converge. In this regard the Ostrowski-Reich theorem is extended in two directions; one is when  $A$  is hermitian but the diagonal  $A$  is not necessarily positive, so that  $A$  may be indefinite, and other is when  $A + A^*$  is positive definite but  $A$  is not necessarily hermitian. In the process, several other convergence results are obtained for general splittings of  $A$ .

## II. MAIN RESULTS

As indicated earlier, the Ostrowski-Reich theorem is sometimes misquoted to conclude that the SOR method can converge only if the coefficient matrix  $A$  is definite. That this is not the case will be illustrated by a series of  $2 \times 2$  examples. These simple examples are given in order to provide some insight into how and why the SOR method can converge.

As a first example we note that if

$$A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

then since  $A_1$  is symmetric and positive definite, the Ostrowski-Reich theorem applies and the SOR method converges for all  $0 < \omega < 2$ .

But the SOR method also converges for all  $0 < \omega < 2$  for each of

$$A_2 = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 2 & 1 \\ -1 & -2 \end{pmatrix},$$

although the Ostrowski-Reich theorem does not apply to any of these matrices. Here  $A_2$  is not hermitian,  $A_3$  is not definite, and  $A_4$  is neither hermitian nor definite.

Moreover, the Ostrowski-Reich theorem does not apply to either of

$$A_5 = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}, \quad A_6 = \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix},$$

since  $A_4$  is not hermitian and  $A_5$  is not definite. However, from the analysis given below, the SOR method converges in both of these cases for all

$$0 < \omega < \frac{4}{3},$$

and the upper bound on  $\omega$  is tight.

Notice that in our examples, although  $A_2$  and  $A_5$  are not hermitian,  $A_2 + A_2^*$  and  $A_5 + A_5^*$  are hermitian and positive definite. In addition, the diagonals of  $A_3$ ,  $A_4$ , and  $A_6$  are not positive, so that each matrix is necessarily indefinite. These situations will play a significant role in the convergence results given later in this section.

It turns out that the complete analysis of the  $2 \times 2$  real case is not difficult. If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad ad \neq 0, \quad ad - bc \neq 0,$$

then from (1.3), (1.4), and (1.5), the SOR iteration matrix  $H_\omega$  is given by

$$\begin{aligned} H_\omega &= \begin{pmatrix} a & 0 \\ \omega c & d \end{pmatrix}^{-1} \left[ \begin{pmatrix} (1-\omega)a & 0 \\ 0 & (1-\omega)d \end{pmatrix} + \begin{pmatrix} 0 & -\omega b \\ 0 & 0 \end{pmatrix} \right] \\ &= \begin{bmatrix} 1-\omega & \frac{-\omega b}{a} \\ \frac{-\omega(1-\omega)c}{d} & (1-\omega) + \frac{\omega^2 bc}{ad} \end{bmatrix}. \end{aligned}$$

Then setting

$$\beta = \frac{bc}{ad},$$

the characteristic equation for  $H_\omega$  is in this case

$$\lambda^2 - [2(1-\omega) + \omega^2\beta]\lambda + (1-\omega)^2 = 0.$$

It follows that  $\rho(H_\omega) < 1$  if and only if

$$0 \leq \beta < 1 \quad \text{and} \quad 0 < \omega < 2$$

or

$$\beta < 0 \quad \text{and} \quad \omega < \frac{2}{1 + \sqrt{|\beta|}}.$$

Note that, in particular, if  $\omega = 1$ , then the Gauss-Seidel iteration converges if and only if  $|\beta| < 1$ —that is,  $\det|A| > 0$ , where  $|A|$  is the matrix of absolute values.

In the examples given above,  $0 \leq \beta < 1$  for each of  $A_1$  through  $A_4$ , so that for these matrices the SOR method converges for all  $0 < \omega < 2$ . But  $\beta = -\frac{1}{4} < 0$  for  $A_5$  and  $A_6$ , so that the SOR method converges here only for  $0 < \omega < 2/(1 + \sqrt{|\beta|}) = \frac{4}{3}$ .

Before extending the Ostrowski-Reich theorem, it will be convenient to consider the more general Householder-John theorem and to provide some results concerning the convergence of the general method (1.1) and (1.2). In the process we will relax somewhat the assumptions made concerning the coefficient matrix  $A$ .

Throughout this section we will assume, without further mention, that  $A = M - N$  with  $A$  and  $M$  nonsingular  $n \times n$  complex matrices, and that  $H = M^{-1}N$ . We denote by  $A^*$  and  $A^{-*}$  the conjugate transpose and its inverse, and by  $\rho(H)$  the spectral radius of  $H$ . Let  $C^n$  denote the complex  $n$ -space. Then the matrix  $A$  will be called *positive definite*<sup>1</sup> if

$$\operatorname{Re}(x^*Ax) > 0 \quad \text{for all } x \in C^n, \quad x \neq 0,$$

or equivalently if the hermitian matrix  $A + A^*$  is positive definite in the usual sense. We say that a subset  $E$  of  $C^n$  is an *eigenset* of  $H$  if  $E$  consists of only (nonzero) eigenvectors of  $H$  with at least one eigenvector corresponding to each distinct eigenvalue of  $H$ .

We begin by establishing an identity which generalizes corresponding identities of Weissinger [19] and Householder [8].

LEMMA 2.1. *The matrices  $A$ ,  $M$ ,  $N$ , and  $H$  from (1.1) and (1.2) satisfy the identity*

$$A - H^*AH = (I - H)^*(M^*A^{-*}A + N)(I - H). \quad (2.1)$$

<sup>1</sup>We note that although nonhermitian matrices with this property appear in the literature in a variety of contexts, there is no generally accepted terminology for them.

*Proof.* Since  $AH = N(I - H)$ , it follows that

$$\begin{aligned} H^*AH &= H^*N(I - H) - N(I - H) + AH \\ &= -(I - H^*)N(I - H) + AH \\ &= -S + AH + (I - H^*)M^*(I - H), \end{aligned}$$

where we have set  $S = (I - H)^*(M^* + N)(I - H)$ . Then since  $(I - H)^*M^* = A^*$ , it follows that

$$\begin{aligned} A - H^*AH &= A + S - AH - A^*(I - H) \\ &= S + (A - A^*)(I - H) \\ &= (I - H)^*[M^* + N + M^*A^{-*}(A - A^*)](I - H) \\ &= (I - H)^*(M^*A^{-*}A + N)(I - H). \quad \blacksquare \end{aligned}$$

This identity now enables us to establish our most general convergence result. In the following theorem there is no assumption that  $A$  is hermitian or definite.

**THEOREM 2.2.** *If  $A$  and  $M^*A^{-*}A + N$  satisfy the condition*

$$x^*Ax \neq 0, \quad \frac{x^*(M^*A^{-*}A + N)x}{x^*Ax} > 0 \quad (2.2)$$

*for every  $x$  in some eigenset  $E$  of  $H$ , then  $\rho(H) < 1$ . Conversely, if  $\rho(H) < 1$ , then for each eigenvector  $x$  of  $H$  either (2.2) holds or else*

$$x^*Ax = x^*(M^*A^{-*}A + N)x = 0. \quad (2.3)$$

*Proof.* Suppose that (2.2) holds for some eigenset  $E$  of  $H$ , and let  $\lambda$  and  $x$  be any eigenvalue of  $H$  and corresponding eigenvector in  $E$ . Note that  $\lambda \neq 1$ ; otherwise  $Mx = Nx$ , which contradicts the nonsingularity of  $A$ . Then by (2.1),

$$x^*Ax - (Hx)^*A(Hx) = [(I - H)x]^*(M^*A^{-*}A + N)[(I - H)x],$$

or

$$(1 - |\lambda|^2)x^*Ax = |1 - \lambda|^2x^*(M^*A^{-*}A + N)x. \quad (2.4)$$

Thus by the assumption (2.2),

$$1 - |\lambda|^2 = |1 - \lambda|^2 \frac{x^*(M^*A^{-*}A + N)x}{x^*Ax} > 0,$$

so that  $|\lambda| < 1$ .

Conversely, if  $\rho(H) < 1$  and  $\lambda, x$  are any eigenpair of  $H$ , then (2.4) holds; and since  $1 - |\lambda| > 0$  and  $|1 - \lambda|^2 > 0$ , either (2.2) or (2.3) must be true. ■

The requirements that  $x^*Ax \neq 0$  for all  $x \in E$  are a drawback to Theorem 2.2, but one that we have been unable to overcome. Note that  $x^*Ax \neq 0$  for an eigenvector  $x$  of  $H$  implies that  $x^*Mx \neq 0$ ; conversely,  $x^*Mx \neq 0$  implies that either  $x^*Ax \neq 0$  or  $\lambda = 1$ . Note also that if  $Nx = \lambda Mx$ , then  $(M - N)x = (1 - \lambda)Mx$ , so that  $x$  is an eigenvector of the generalized eigenvalue problem  $Ax = SMx$ .

In general, the quantities  $x^*Ax$  and  $x^*(M^*A^{-*}A + N)x$ , as well as their product, are complex, although, as shown in the proof of Theorem 2.2, the quotient in (2.2) is always real for any eigenvector  $x$  of  $H$ . However, if  $A$  is hermitian, then  $M^*A^{-*}A + N = M^* + N$ , which is also hermitian. Thus the condition (2.2) is that the quantities  $x^*Ax$  and  $x^*(M^*A^{-*}A + N)x$ , which are necessarily real, have the same sign. It is useful to isolate this as the following definition.

**DEFINITION 2.3.** Let  $S$  be an arbitrary nonempty subset of  $C^n$ . Then two  $n \times n$  hermitian matrices  $P$  and  $Q$  are *quadratic form sign equivalent* (QFSE) on  $S$  if either

$$(x^*Px)(x^*Qx) > 0 \quad (2.5)$$

or

$$x^*Px = x^*Qx = 0 \quad (2.6)$$

for all  $x \in S$ , and *strongly* QFSE (SQFSE) on  $S$  if (2.5) holds for all  $x \neq 0$  in  $S$ .

In terms of this definition, Theorem 2.2 for hermitian  $A$  reduces to the following potentially useful result.

**COROLLARY 2.4.** If  $A$  is hermitian and  $A$  and  $M^* + N$  are SQFSE on some eigenset of  $H$ , then  $\rho(H) < 1$ . Conversely, if  $\rho(H) < 1$ , then  $A$  and  $M^* + N$  are QFSE on every eigenset of  $H$ .

We say that a hermitian matrix  $P$  is *positive definite* on a subset  $S$  of  $C^n$  if  $x^*Px > 0$  for all  $x \neq 0$  in  $S$ . In these terms, a special case of Corollary 2.4 is:

**COROLLARY 2.5.** *If  $A$  is hermitian and  $A$  and  $M^* + N$  are positive definite on some eigenset of  $H$ , then  $\rho(H) < 1$ . Conversely, if  $x^* Mx > 0$  for all  $x$  in some eigenset  $E$  of  $H$  and  $\rho(H) < 1$ , then  $A$  and  $M^* + N$  are positive definite on  $E$ .*

The proof of the converse of Corollary 2.5 follows immediately from the relation

$$x^* Ax + x^* (M^* + N)x = 2x^* Mx > 0 \quad (2.7)$$

which holds for every eigenvector  $x \neq 0$  of  $H$ . This shows that at least one of the terms of the left hand side of (2.7) is positive and hence, by Corollary 2.4 and the definition of QFSE, both terms must be positive.

We next turn briefly to the application of Theorem 2.2 to the SOR iterations. For that purpose we now let (1.3) be a block splitting of  $A$  where  $D$  is a block diagonal matrix with square diagonal blocks and where  $-L$  and  $-U$  are block strictly lower and upper triangular parts of  $A$ , respectively. As before, we assume that  $A$  and  $D$  are nonsingular. Then if  $A$  is hermitian,  $U = L^*$  and Theorem 2.2 leads to the following:

**THEOREM 2.6.** *Let  $A$  be hermitian and have a block SOR splitting of the form specified by (1.3), (1.4), and (1.5), and let  $\omega \in (0, 2)$ . If  $x^* Dx \neq 0$  for all  $x$  in some eigenset  $E$  of  $H_\omega$ , then  $\rho(H_\omega) < 1$  if and only if*

$$(x^* Dx)(x^* Ax) > 0 \quad \text{for all } x \in E. \quad (2.8)$$

*Proof.* Since  $A$  is hermitian,

$$(M^* A^{-*} A + N) = M^* + N = \frac{1}{\omega} (2 - \omega) D.$$

Also, if  $x$  belongs to some eigenset of  $H_\omega$ , then  $x^* Dx \neq 0$  implies  $x^* Ax \neq 0$ . The proof now follows immediately from Corollary 2.4. ■

We note that the case of interest to us is when  $A$  has elements of mixed signs on the diagonal (or, in general, when the block diagonal matrix  $D$  is not definite). Notice that this was the case in the  $2 \times 2$  examples  $A_3$  and  $A_6$  given earlier. In this case, the Ostrowski-Reich theorem of course does not apply. Our condition (2.8) states that there must be a certain "sign matching" between  $D$  and  $A$  in some eigenset  $E$  of  $A$  in order for the SOR iterations to converge. Note also that we could have phrased (2.8) in the alternate form

$$\begin{aligned} x^* Dx > 0 & \text{ implies } x^* Dx > x^* (L + L^*)x, \\ x^* Dx < 0 & \text{ implies } x^* Dx < x^* (L + L^*)x, \end{aligned} \quad (2.9)$$

for all  $x \in E$ .



It turns out that the Ostrowski-Reich theorem is not an immediate consequence of Theorem 2.6, and we shall return to this topic later. We now turn to another general approach that was used by Collatz [4] for the Gauss-Seidel iteration and by Albrecht [1] for the SOR method. The corresponding result for general splittings seems to be new. Here  $\text{Im}$  denotes the imaginary part of a complex number, as usual.

**THEOREM 2.7.** *Assume that  $x^*Mx \neq 0$  for all  $x$  in some eigenset  $E$  of  $H$ . Then  $\rho(H) < 1$  if and only if*

$$[\text{Re } x^*(M+N)x][\text{Re } x^*Ax] > -[\text{Im } x^*(M+N)x][\text{Im } x^*Ax] \quad (2.10)$$

for all  $x \in E$ .

*Proof.* Let  $x \in E$  have corresponding eigenvalue  $\lambda$ . Then  $Nx = \lambda Mx$  and, since  $x^*Mx \neq 0$  by assumption, we have

$$\lambda = \frac{x^*Nx}{x^*Mx}. \quad (2.11)$$

But  $2N = N - M - A$  and  $2M = M + N - A$ , so that (2.11) becomes

$$\lambda = \frac{x^*(M+N)x - x^*Ax}{x^*(M+N)x + x^*Ax}.$$

It now follows by some algebraic manipulation that  $|\lambda| < 1$  if and only if (2.10) holds. ■

Now if  $A$  is hermitian, then  $\text{Im } x^*Ax = 0$  and  $\text{Re } x^*(M+N)x = x^*(M^* + N)x$ ; hence (2.10) reduces to

$$[x^*(M^* + N)x][x^*Ax] > 0. \quad (2.12)$$

Moreover,  $x^*(M^* + N)x \neq 0$  implies that  $x^*Mx \neq 0$  for any eigenvector  $x$  of  $H$  by the relation  $Nx = \lambda Mx$ . Thus the first part of Corollary 2.4 also follows from Theorem 2.7.

If  $A$  is positive definite (but not necessarily hermitian), then the condition  $x^*Mx \neq 0$  is satisfied for every nonzero eigenvector  $x$  of  $H$ . For otherwise, as above,  $x^*Ax = 0$ . This yields a potentially useful corollary to Theorem 2.7.

COROLLARY 2.8. *If  $A$  is positive definite, then  $\rho(H) < 1$  if and only if (2.10) holds for all  $x$  in some eigenset of  $H$ .*

We now give a companion result to Theorem 2.2 under the condition that  $M^*A^{-*}A + N$  is positive definite. In this case we prove a somewhat different form of the converse which is more along the lines of the Ostrowski-Reich theorem.

THEOREM 2.9. *Assume that  $M^*A^{-*}A + N$  is positive definite. Then  $\rho(H) < 1$  if and only if  $A$  is positive definite.*

*Proof.* Assume that  $A$  is positive definite, and for any  $x \neq 0$ , let

$$x^*Ax = a + ib, \quad x^*(M^*A^{-*}A + N)x = c + id.$$

By assumption,  $a$  and  $c$  are positive. If  $x$  is an eigenvector of  $H$ , then (2.4) shows that the quotient

$$\frac{c + id}{a + ib}$$

is real. Then  $bc = ad$ , so that

$$ac + bd = ac + \frac{b^2c}{a} > 0,$$

and thus the quotient is positive. Then Theorem 2.2 applies, and so  $\rho(H) < 1$ .

Conversely, if  $\rho(H) < 1$ , then the sequence of vectors

$$x^{k+1} = Hx^k, \quad k = 0, 1, \dots,$$

converges to zero for any  $x^0$ . Then by (2.1) of Lemma 2.1,

$$(x^k)^*Ax^k - (x^{k+1})^*Ax^{k+1} = (x^k - x^{k+1})^*(M^*A^{-*}A + N)(x^k - x^{k+1}),$$

and, by the assumption that  $M^*A^{-*}A + N$  is positive definite,

$$\operatorname{Re}[(x^k)^*Ax^k - (x^{k+1})^*Ax^{k+1}] \geq 0.$$

Thus

$$\operatorname{Re}(x^{k+1})^*Ax^{k+1} \leq \operatorname{Re}(x^k)^*Ax^k, \quad k = 0, 1, \dots \quad (2.13)$$

Now if  $A$  were not positive definite, we could find an  $x$  such that  $\operatorname{Re}(x^0)^*Ax^0 \leq 0$ , and since  $x^1 \neq x^0$  [otherwise  $\rho(H)=1$ ], the positive definiteness of  $M^*A^{-1}A+N$  implies that

$$\operatorname{Re}(x^1)^*Ax^1 < \operatorname{Re}(x^0)^*Ax^0 \leq 0.$$

Thus (2.13) would preclude the convergence of the sequence  $\{x^k\}$  to 0. ■

For completeness we now recover the Householder-John and Ostrowski-Reich theorems from Theorem 2.9. If  $A$  is hermitian, then since  $M^*A^{-1}A+N=M^*+N$ , we have immediately

**COROLLARY 2.10** (Householder [8], John [9]). *If  $A$  is hermitian and if  $M^*+N$  is positive definite, then  $\rho(H)<1$  if and only if  $A$  is positive definite.*

As indicated in the Introduction, a weaker version of this result was established by Weissinger [19]. It is easy to see that if  $A$  is hermitian, then so is  $M^*+N$ , although this need not be true of  $M+N$ . Moreover,  $M^*+N$  is positive definite if and only if  $M+N$  is positive definite in our weak sense (see Footnote 1). Weissinger showed that if  $A$  and  $M+N$  are hermitian and  $M+N$  is positive definite, then the conclusion of Corollary 2.10 holds. The result was strengthened by Feingold and Spohn [7] by showing that the requirement that  $M+N$  be hermitian is unnecessary, and this was presented by Ortega [14] as the "*P*-regular splitting theorem." Here  $A=M-N$  is a *P*-regular splitting if  $M+N$  is positive definite. Thus an alternate statement to Corollary 2.10 is

**COROLLARY 2.10'.** *Assume that  $A$  is hermitian and that  $A=M-N$  is a *P*-regular splitting. Then  $\rho(H)<1$  if and only if  $A$  is positive definite.*

We note that Corollary 2.10 contains the following interesting result of Varga [17]: If  $A$  and  $D$  are hermitian,  $D$  is positive definite,  $S$  is skew-hermitian, and  $Q \equiv \frac{1}{2}(D-A+S)$ , then the spectral radius of

$$(D-\omega Q)^{-1}[(1-\omega)D+\omega Q^*]$$

is less than one for any  $\omega \in (0,2)$ , if and only if  $A$  is positive definite. Here

$$M = \frac{1}{\omega}(D-\omega Q), \quad N = \frac{1}{\omega}[(1-\omega)D+\omega Q^*]$$

and it is easy to see that, again,

$$M^* + N = \frac{1}{\omega}(2 - \omega)D,$$

and hence  $M^* + N$  is positive definite for  $\omega \in (0, 2)$ . Thus Varga's result is immediate from Corollary 2.10.

Finally, since  $M^* + N = (2 - \omega D)/\omega$  for the SOR splitting (1.3), (1.4), and (1.5), we have

**COROLLARY 2.11** (Ostrowski [15], Reich [16]). *If  $A$  is hermitian and if the block diagonal matrix  $D$  for  $A$  is positive definite, then the block SOR iteration matrix  $H_\omega$  satisfies  $\rho(H_\omega) < 1$  for all  $\omega \in (0, 2)$  if and only if  $A$  is positive definite.*

### III. SOME OBSERVATIONS

1. The main results given in this paper provide necessary and sufficient conditions for the convergence of general iterative methods in terms of the eigensystem of the iteration matrix. Although these results may give a deeper insight into how and why the SOR method can converge, they are deficient in that the conditions are not given in terms of the coefficient matrix  $A$  itself. In particular, this is true of Theorem 2.6, where we require that  $x^*Dx \neq 0$  and  $(x^*Dx)(x^*Ax) > 0$  in some eigenset  $E$  of the SOR iteration matrix  $H_\omega$ . Thus our results may not be totally satisfactory in terms of applications to practical problems, and we consider it an open problem to determine if necessary and sufficient conditions can be given in terms of  $A$  and  $\omega$  alone for the SOR method to converge, even for hermitian matrices.

We turn now to the case where  $A$  is not necessarily hermitian. It is tempting to conjecture that the SOR method will converge for any  $\omega \in (0, 2)$  if  $A$  is positive definite—but this is not the case. In fact, even the Gauss-Seidel method ( $\omega = 1$ ) fails to converge for the simple example

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & 1 \end{bmatrix},$$

which is clearly positive definite. We note that this matrix is of the form  $A = I - B$ , where  $B$  is skew symmetric, and that Niethammer and Schade [12,

13] have given bounds on the maximum size that  $\omega$  can be for this class of matrices. We note also that Broyden [3] and Niethammer [12] have shown that SOR will converge for any positive definite matrix provided that  $\omega$  is sufficiently small.

The application of the general results in Theorems 2.2 and 2.9 is difficult in the nonhermitian case because of the presence of the matrix  $A^{-*}A$ . For example, with  $A = D - L - U$ , as usual, the application of Theorem 2.9 to the SOR method requires that we ascertain whether the matrix

$$M^*A^{-*}A + N = \frac{1}{\omega}D^*A^{-*}A - L^*A^{-*}A + \frac{1}{\omega}[(1-\omega)D + \omega U] \quad (3.1)$$

is positive definite. One set of conditions for this is that  $D$  and  $DA^{-*}$  are positive definite, that  $L^*A^{-*} = UA^{-*}$ , and that  $\omega \in (0, 1]$ . But some of these conditions seem unlikely to be satisfied in problems of interest. Another condition is that  $A$  is skew-hermitian, so that  $A^{-*}A = -I$ . Then (3.1) reduces to

$$\frac{1}{\omega}D + L^* + \frac{1}{\omega}[(1-\omega)D - \omega L^*] = \frac{1}{\omega}(2-\omega)D.$$

Thus the application of Theorem 2.2 requires that

$$x^*Ax \neq 0 \quad \text{and} \quad \frac{x^*Dx}{x^*Ax} > 0$$

for all  $x$  in some eigenset of  $H_\omega$ .

Next, we note that the matrix  $A^{-*}A$ , which appears in our convergence criteria, has been studied by Deprima and Johnson [5] and by Fan [6], who shows, among other things, that  $A^{-*}A$  is similar to a unitary matrix if  $A$  is positive definite. We have not been able to use this or other such results to suitably simplify the expression  $M^*A^{-*}A + N$ .

2. Results analogous to those of this paper can be given for various other concrete iterations. For example, for the Jacobi iteration, let  $A$  have a splitting  $A = D - B$ , where  $D$  is a block diagonal part of  $A$ , and assume that  $x^*Dx \neq 0$  for all  $x$  in some eigenset  $E$  of  $H = D^{-1}B$ . Then  $\rho(H) < 1$  if and only if

$$|x^*Dx| > |x^*Bx| \quad (3.2)$$

for all  $x \in E$ .

If  $A$  is hermitian, then

$$D + B = D^* + B = 2D - A,$$

and (3.2) is equivalent to  $A$  and  $D^*$  being SQFSE on  $E$ . Hence, in this case, the result also follows from Corollary 2.4. If  $D$  is positive definite, then (3.2) is equivalent to  $x^*(D \pm B)x > 0$  for all  $x \in E$ . In this case the result becomes:  $\rho(H) < 1$  if and only if both  $A$  and  $2D - A$  are positive definite on some eigenset of  $H$ .

3. An alternate approach to the study of SOR convergence is by the use of the Perron-Frobenius theory of nonnegative matrices and, in particular, the theory of  $M$ -matrices. This approach does not directly depend upon whether the coefficient matrix  $A$  is hermitian or positive definite, but depends upon the determination of whether the comparison matrix  $\mathcal{M}(A) = 2|D| - |A|$  for  $A$  is a nonsingular  $M$ -matrix. For example, Kulisch [11] has shown that if  $\mathcal{M}(A)$  is a nonsingular  $M$ -matrix, then the SOR method converges for all

$$0 < \omega < \frac{2}{1 + \rho(|J|)},$$

where  $J$  is the Jacobi iteration matrix  $D^{-1}(L + U)$ . We note that this criterion can also be used to formulate the analysis of the convergence of the  $2 \times 2$  SOR iterations discussed earlier.

For a more detailed discussion of this approach the reader is referred to Ortega [14], Varga [17] or Young [20]. Also, Chapter 6 of Berman and Plemmons [2] provides fifty useful characterizations of nonsingular  $M$ -matrices, and Chapter 7 of that book contains the more recent developments in the study of convergence criteria for iterative methods, based upon both the positive definite approach and the  $M$ -matrix approach.

*We are grateful to Professor David Young for reading a preliminary version of this manuscript and making several useful comments.*

## REFERENCES

- 1 J. Albrecht, Fehlerabschätzungen bei Relaxations Verfahren zur numerische Auflösung linear Gleichungssysteme, *Numer. Math.* 3:188–201 (1961).
- 2 A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, New York, 1979.
- 3 C. Broyden, On convergence criteria for the method of successive over-relaxation. *Math. Comp.* 18:136–144 (1964).

- 4 L. Collatz, Über die Konvergenzkriterien bei Iterationsverfahren für Linearsysteme, *Math. Z.* 59:149–161 (1950).
- 5 C. Deprima and C. Johnson, The range of  $A^{-1}A^*$  in  $GL(n, C)$ , *Linear Algebra and Appl.* 9:209–222 (1974).
- 6 K. Fan, On strictly dissipative matrices, *Linear Algebra and Appl.* 9:223–241 (1974).
- 7 D. Feingold and D. Spohn, Un théorème simple sur les normes de matrices et sur conséquences, *C. R. Acad. Sci.* 256:2758–2760 (1963).
- 8 A. S. Householder, On the convergence of matrix iterations, *Oak Ridge National Laboratory Technical Report No. 1883*, 1955.
- 9 F. John, *Advanced Numerical Analysis*, Lecture Notes, Department of Mathematics, New York University, 1956.
- 10 H. B. Keller, On the solution of singular and semidefinite linear systems by iteration, *SIAM J. Numer. Anal.* 2:281–290 (1965).
- 11 U. Kulisch, Über reguläre zerlegungen von Matrizen und einige Anwendungen, *Numer. Math.* 11:444–449 (1968).
- 12 W. Niethammer, Relaxation bei nicht-symmetrischen Matrizen, *Math. Z.* 85:319–327 (1964).
- 13 W. Niethammer and J. Schade, On a relaxed SOR method applied to non-symmetric linear systems, *J. Computational and Applied Math.* 1:133–136 (1975).
- 14 J. Ortega, *Numerical Analysis: A Second Course*, Academic, New York, 1972.
- 15 A. M. Ostrowski, On the linear iteration procedures for symmetric matrices, *Rend. Mat. Appl.* 14:140–163. (1954).
- 16 E. Reich, On the convergence of the classical iterative procedures for symmetric matrices, *Ann. Math. Statistics* 20:448–451 (1949).
- 17 R. S. Varga, Extensions of the successive overrelaxation theory with applications to finite element approximations, in *Topics in Numerical Analysis* (J. J. H. Miller, Ed.) Academic, New York, 1973, pp. 329–343.
- 18 R. S. Varga, On recurring theorems on diagonal dominance, *Linear Algebra and Appl.* 13:1–10 (1976).
- 19 J. Weissinger, Verallgemeinerungen des Seidelschen Iterationsverfahrens, *Z. Angew. Math. Mech.* 33:155–163 (1953).
- 20 D. M. Young, *Iterative Solution of Large Linear Systems*, Academic, New York, 1971.

Received 20 June 1978; revised 16 July 1978