Exact Convergence and Divergence Domains for the Symmetric Successive Overrelaxation Iterative (SSOR) Method Applied to *H*-Matrices

A. Neumaier
Institut für Angewandte Mathematik
Universität Freiburg
D-7800 Freiburg, West Germany

and

R. S. Varga*
Mathematisches Institut
Universität Würzburg
D-8700 Würzburg am Hubland, West Germany

Submitted by W. Niethammer

ABSTRACT

In this paper, exact convergence and divergence domains for the SSOR iterative method, as applied to the class of *H*-matrices, are obtained. The theory of regular splittings and the recent results of Varga, Niethammer, and Cai are used as tools in establishing these convergence and divergence domains.

1. INTRODUCTION

Today, a popular preconditioning method, used in conjunction with the conjugate gradient method, is one or more sweeps of the symmetric successive overrelaxation (SSOR) iterative method (cf. [2]). This new use of SSOR has, interestingly enough, sparked recent interest into the general theory of this method. The purpose of this paper is to obtain exact domains for the convergence and divergence of the SSOR iterative method, as it pertains to *H*-matrices. As is well known, the classes of *M*-matrices and *H*-matrices were introduced by A. M. Ostrowski in his fundamental work [8].

*Research supported in part by the Alexander von Humboldt Stiftung, by the Air Force Office of Scientific Research, and by the Department of Energy.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

for any A in $\mathcal{H}_v$, unless $v=0$, indicating that the interval $(0,2/[1+v])$ cannot be the largest such interval, unless $v=0$. On the other hand, Varga, Niethammer, and Cai [12] have shown that for each v with $\frac{1}{2}<v<1$ and for each ω satisfying

$$\frac{2}{1+\sqrt{2v-1}}<\omega<2, \tag{1.12}$$

there is a matrix B in $\mathcal{H}_v$ for which $(S_\omega(B))>1$.

The main result of this paper in fact precisely determines the largest interval in ω for which $S_\omega(A)$ is convergent for any A in $\mathcal{H}_v$, answering the above question. We remark that the proof of our Theorem of Section 2 makes use of the theory of *regular splittings* of matrices (cf. [9], [10]) and the recent results of Varga, Niethammer, and Cai [12].

2. STATEMENT OF THE MAIN RESULT

Our main result (to be proved in §3) is the

THEOREM. *For each v with $0\leqslant v<1$, set*

$$\hat{\omega}(v):=\begin{cases} 2 & \text{if } 0\leqslant v\leqslant \frac{1}{2}, \\ \dfrac{2}{1+\sqrt{2v-1}} & \text{if } \frac{1}{2}<v<1. \end{cases} \tag{2.1}$$

Then, for each matrix A in $\mathcal{H}_v$ and for each ω with $0<\omega<\hat{\omega}(v)$,

$$\rho(S_\omega(A))<1, \tag{2.2}$$

i.e., $S_\omega(A)$ is convergent. On the other hand, for each ω with $\omega\leqslant 0$ or with $\omega>\hat{\omega}(v)$, there is a matrix B in $\mathcal{H}_v$ for which

$$\rho(S_\omega(B))\geqslant 1, \tag{2.3}$$

i.e., $S_\omega(B)$ is divergent.

From our Theorem, we see that the curve $\hat{\omega}(v)$, for $0\leqslant v<1$, as defined in (2.1), separates the convergence and divergence domains for matrices in

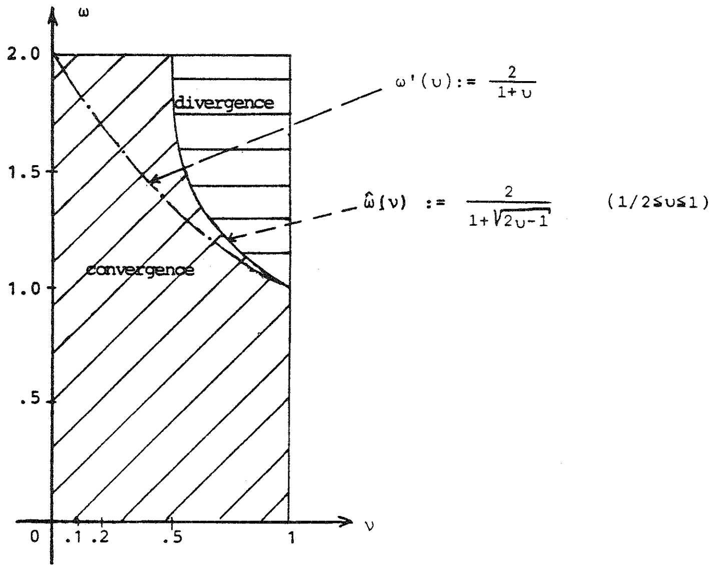

FIG. 1.

$\mathscr{H}_v$, as shown in Figure 1. Only *on* this curve is the convergence or divergence of matrices in $\mathscr{H}_v$ unsettled. For comparison purposes, we have also drawn the curve $\omega'(v)$ of (1.11) in Figure 1.

For any A in $\mathscr{H}_v$, the upper bound (of Neumann [6]) in (1.9) gives that

$$\sup\{\rho(S_\omega(A)) : A \in \mathscr{H}_v\} \leqslant \omega v + |1-\omega| \qquad (\text{for all} \quad 0<\omega<2/[1+v]). \tag{2.4}$$

In light of our Theorem, it can be verified that equality *cannot* hold in (2.4) for any ω satisfying $2/[1+v]<\omega<\hat{\omega}(v)$ [cf. (2.1)]. In a later paper, we propose to find sharper upper bounds, as a function of ω and v, for $\sup\{\rho(S_\omega(A) : A \in \mathscr{H}_v\}$, much in the spirit of sharp upper bounds which have been found for the related successive overrelaxation (SOR) iterative method for matrices in $\mathscr{H}_v$ (cf. Kahan [4], Kulisch [5], and Neumann and Varga [7]).

3. PROOF OF THE THEOREM

In this section, we first establish some needed preliminary results for the proof of our Theorem. In what follows, let $J = L + U$ be an $n \times n$ Jacobi matrix [cf. (1.4)], and set

$$v = \rho(|J|), \qquad \text{where we assume that} \quad 0 \leqslant v < 1. \tag{3.1}$$

Next, for arbitrary nonnegative real numbers s and t, set

$$Q_s := (I - sL)(I - sU), \qquad \overline{Q}_s := (I - s|L|)(I - s|U|), \tag{3.2}$$

$$P_s := s^2 Q_s^{-1} \cdot L \cdot U, \qquad \overline{P}_s := s^2 \overline{Q}_s^{-1} \cdot |L| \cdot |U|, \tag{3.3}$$

$$M_{s,t} := I - sJ + tLU, \qquad \overline{M}_{s,t} := I - s|J| + t|L| \cdot |U|, \tag{3.4}$$

$$N_{s,t} := (1-s)J + tLU, \qquad \overline{N}_{s,t} := |1-s|\,|J| + t|L| \cdot |U|. \tag{3.5}$$

Because of the strictly triangular character of the matrices L and U, it is evident that Q_s and $\overline{Q}_s$ are nonsingular for any $s \geqslant 0$. Moreover, as

$$\begin{aligned} Q_s^{-1} &= (I - sU)^{-1}(I - sL)^{-1} \\ &= \left[I + sU + \cdots + (sU)^{n-1}\right]\left[I + sL + \cdots + (sL)^{n-1}\right], \end{aligned}$$

it follows that

$$|Q_s^{-1}| \leqslant (I - s|U|)^{-1}(I - s|L|)^{-1} = \left(\overline{Q}_s\right)^{-1}, \qquad \text{whence} \quad \left(\overline{Q}_s\right)^{-1} \geqslant \mathcal{O}. \tag{3.6}$$

Hence, with (3.6) and (3.3),

$$|P_s| \leqslant s^2 |Q_s^{-1}| \cdot |L| \cdot |U| \leqslant s^2 \left(\overline{Q}_s\right)^{-1} \cdot |L| \cdot |U| =: \overline{P}_s. \tag{3.7}$$

LEMMA 1. *For any s satisfying $0 \leqslant s < 1/v$ (where $1/v := \infty$ if $v = 0$), then*

$$\rho(P_s) \leqslant \rho(|P_s|) \leqslant \rho\left(\overline{P}_s\right) < 1. \tag{3.8}$$

Proof. The first inequality of (3.8) is a well-known consequence of the Perron-Frobenius theory of nonnegative matrices (cf. [10, p. 47]), while the second inequality of (3.8) follows from (3.7). Now, the assumption $0 \leqslant s < 1/v$

implies that $\rho(s|J|)<1$, so that $I-s|J|$ is, by definition [cf. (1.6)], a nonsingular M-matrix, whence $(I-s|J|)^{-1}\geqslant \mathcal{O}$ (cf. [3, p. 137]). Next, from (3.2) and (3.3), it can be verified that

$$I-\bar{P}_s=\bar{Q}_s^{-1}\{I-s|J|\}. \tag{3.9}$$

Thus, $I-\bar{P}_s$ is nonsingular, and it follows from (3.9) and (3.2) that

$$\left(I-\bar{P}_s\right)^{-1}=I+s^2(I-s|J|)^{-1}|L|\cdot|U|\geqslant \mathcal{O}. \tag{3.10}$$

But as $I-\bar{P}_s$ is a matrix in $Z^{n,n}$ satisfying (3.10), then $I-\bar{P}_s$ is a nonsingular M-matrix (cf. [3, p. 137]), so that [cf. (1.6)] $\rho(\bar{P}_s)<1$. ∎

LEMMA 2. *For any* s *and* t *satisfying* $0<s<1/v$ *and* $0\leqslant t\leqslant s^2$,

$$|M_{s,t}^{-1}|\leqslant\left(\bar{M}_{s,t}\right)^{-1}, \qquad \textit{whence} \quad \left(\bar{M}_{s,t}\right)^{-1}\geqslant \mathcal{O}. \tag{3.11}$$

Proof. Let $\varepsilon:=(s^2-t)/s^2$, so that $0\leqslant\varepsilon\leqslant 1$. From the definitions (3.2)–(3.3), it can be verified that

$$M_{s,t}=Q_s(I-\varepsilon P_s) \quad\text{and}\quad \bar{M}_{s,t}=\bar{Q}_s\left(I-\varepsilon\bar{P}_s\right). \tag{3.12}$$

As $0\leqslant\varepsilon\leqslant 1$, Lemma 1 gives that both $I-\varepsilon P_s$ and $I-\varepsilon\bar{P}_s$ are nonsingular. Thus, from (3.12) and from (3.6) and (3.7),

$$|M_{s,t}^{-1}|\leqslant\left|(I-\varepsilon P_s)^{-1}\right|\cdot|Q_s^{-1}|\leqslant(I-\varepsilon|P_s|)^{-1}|Q_s^{-1}|\leqslant\left(I-\varepsilon\bar{P}_s\right)\left(\bar{Q}_s\right)^{-1},$$

so that with (3.12)

$$|M_{s,t}^{-1}|\leqslant\left(\bar{M}_{s,t}\right)^{-1}.$$

∎

LEMMA 3. *For any* s *and* t *satisfying* $0<s<(v+1)/2v$ *and* $0\leqslant t\leqslant s^2$,

$$\rho\left(M_{s,t}^{-1}N_{s,t}\right)\leqslant\rho\left(\left(\bar{M}_{s,t}\right)^{-1}\bar{N}_{s,t}\right)<1. \tag{3.13}$$

Proof. Setting $\bar{A}_{s,t}:=\bar{M}_{s,t}-\bar{N}_{s,t}$, it follows from the definitions (3.4) and (3.5) that

$$\bar{A}_{s,t}=I-(s+|1-s|)|J|. \tag{3.14}$$

Now, as $\bar{A}_{s,t}$ is evidently in $Z^{n,n}$ with $\rho\{(s+|1-s|)|J|\}<1$ since $0<s<(v+1)/2v$, we have [cf. (1.6)] $\bar{A}_{s,t}$ is a nonsingular M-matrix with $(\bar{A}_{s,t})^{-1}\geqslant \mathcal{O}$. Next, clearly $\bar{N}_{s,t}\geqslant \mathcal{O}$ from (3.5), and as $0<s<(v+1)/(2v)$ implies $0<s<1/v$, then (3.11) of Lemma 2 gives that $(\bar{M}_{s,t})^{-1}\geqslant \mathcal{O}$. Thus, $\bar{A}_{s,t}=\bar{M}_{s,t}-\bar{N}_{s,t}$ is a regular splitting of $\bar{A}_{s,t}$ with $(\bar{A}_{s,t})^{-1}\geqslant \mathcal{O}$. Consequently (cf. [10, p. 89]), $\rho((\bar{M}_{s,t})^{-1}\bar{N}_{s,t})<1$. But with (3.11), then

$$\rho\left(M_{s,t}^{-1}N_{s,t}\right)\leqslant\rho\left(|M_{s,t}^{-1}|\cdot|N_{s,t}|\right)\leqslant\rho\left(\left(\bar{M}_{s,t}\right)^{-1}\bar{N}_{s,t}\right)<1, \tag{3.15}$$

establishing (3.13). ∎

For the SSOR iterative method, it is well known (cf. [13, p. 462]) that the associated (point) SSOR iteration matrix S_ω can also be represented as

$$S_\omega=I-\omega(2-\omega)\{(I-\omega L)(I-\omega U)\}^{-1}(I-J) \tag{3.16}$$

for any ω. On setting

$$C_\omega:=Q_\omega^{-1}(I-J), \tag{3.17}$$

it is evident from (3.2) that S_ω in (3.16) can also be expressed as

$$S_\omega=I-\omega(2-\omega)C_\omega. \tag{3.18}$$

For any s with $1<s<(v+1)/2v$ and for any ω with $1<\omega<2$, set

$$\hat{t}:=\frac{\omega^2(s-1)}{\omega-1}. \tag{3.19}$$

Then it can be verified from the various definitions above that

$$\begin{aligned} M_{s,\hat{t}}&=\left(\frac{s-1}{\omega-1}\right)Q_\omega\left[I+\left(\frac{\omega-s}{s-1}\right)C_\omega\right],\\ N_{s,\hat{t}}&=\left(\frac{s-1}{\omega-1}\right)Q_\omega\left[I-C_\omega\right], \end{aligned} \tag{3.20}$$

so that

$$(M_{s,\hat{t}})^{-1}\cdot N_{s,\hat{t}}=\left[I+\left(\frac{\omega-s}{s-1}\right)C_\omega\right]^{-1}\cdot(I-C_\omega). \tag{3.21}$$

Now, if $1 < s \leqslant \omega$, then $\hat{t}$ of (3.19), regarded as a function of ω, is positive and decreasing on $[s, 2)$. Hence, the maximum value of $\hat{t}$ on $[s, 2]$ occurs when $\omega = s$, so that $0 < \hat{t} \leqslant s^2$ for all ω in $[s, 2)$. But then, Lemma 3 applies, so that $\rho(M_{s,\hat{t}}^{-1} N_{s,\hat{t}}) < 1$ for all $1 < s \leqslant \omega$, where $s < (v+1)/2v$ and where $\omega < 2$. This, however, implies that if λ is any eigenvalue of the matrix C_ω, then from (3.21)

$$\left| \frac{1-\lambda}{1+\left(\frac{\omega - s}{s-1}\right)\lambda} \right| < 1. \tag{3.22}$$

As λ cannot be zero in (3.22), the above reduces to

$$\frac{2 \operatorname{Re} \lambda}{|\lambda|^2} > 1 - \frac{\omega - s}{s-1} =: \gamma, \tag{3.23}$$

and if γ is positive, the above is equivalent to

$$|1 - \gamma\lambda| < 1 \tag{3.24}$$

for any eigenvalue λ of C_ω.

Now, the eigenvalues μ of S_ω from (3.18) are evidently connected to the eigenvalues λ of C_ω through

$$\mu = 1 - \omega(2-\omega)\lambda, \tag{3.25}$$

and on rewriting (3.24) in terms of μ, we obtain, assuming $\gamma > 0$, that

$$|\mu_0 - \mu| < 1 - \mu_0, \qquad \text{where} \quad \mu_0 := 1 - \frac{\omega(2-\omega)}{\gamma}. \tag{3.26}$$

Geometrically, $|\mu_0 - \mu| < 1 - \mu_0$ is an open disk, with center μ_0 and radius $1 - \mu_0$, which lies completely in $|z| < 1$ if $\mu_0 \geqslant 0$. But then, if $\mu_0 \geqslant 0$ and if $\gamma > 0$, all eigenvalues μ of S_ω lie in $|z| < 1$, so that S_ω is convergent.

Now, the condition that μ_0 is nonnegative is equivalent, from (3.26) and (3.23), to the condition that

$$s \geqslant \frac{\omega^2 - \omega + 1}{\omega^2 - 2\omega + 2}. \tag{3.27}$$

On the other hand, as we have assumed that $s<(v+1)/2v$, we must also have from (3.27) that

$$\frac{\omega^2-\omega+1}{\omega^2-2\omega+2}<\frac{v+1}{2v},$$

or, equivalently, that

$$(1-v)\omega^2-2\omega+2>0, \tag{3.28}$$

where $1<s\leqslant\omega<2$.

Two cases arise from (3.28). If $0\leqslant v<\frac{1}{2}$, the quadratic equation in ω in (3.28) is positive for all real ω, while if $\frac{1}{2}\leqslant v<1$, then (3.28) is satisfied for

$$0<\omega<\frac{2}{1+\sqrt{2v-1}} \qquad \left(\tfrac{1}{2}\leqslant v<1\right). \tag{3.29}$$

This brings us to the

Proof of our Theorem. Choose any ω satisfying

$$1<\omega<\begin{cases}2 & \text{if } 0\leqslant v\leqslant\frac{1}{2},\\[4pt] \dfrac{2}{1+\sqrt{2v-1}} & \text{if } \frac{1}{2}<v<1,\end{cases} \tag{3.30}$$

and define [cf. (3.27)]

$$\hat{s}(\omega):=\frac{\omega^2-\omega+1}{\omega^2-2\omega+2}, \tag{3.31}$$

where $\hat{s}(\omega)$ places the role of s in our previous discussion [cf. (3.19)–(3.27)]. As $\hat{s}(\omega)>1$ is from (3.31) equivalent to $\omega>1$, then $\hat{s}(\omega)>1$ for all ω satisfying (3.30). Next, from (3.19), we find that

$$\hat{t}(\omega):=\frac{\omega^2(\hat{s}(\omega)-1)}{\omega-1}=\frac{\omega^2}{\omega^2-2\omega+2}, \tag{3.32}$$

so that $\hat{t}(\omega)<[\hat{s}(\omega)]^2$ for all ω satisfying (3.30). Note that as $\hat{s}(\omega)\leqslant\omega$ is equivalent, by (3.31), with $(\omega-1)^3 \geqslant 0$, then $\hat{s}(\omega) \leqslant \omega$ for all ω satisfying (3.30). Next, we compute from (3.23) and (3.31) that

$$\hat{\gamma} := 1 - \frac{\omega - \hat{s}(\omega)}{\hat{s}(\omega) - 1} = \omega(2-\omega) > 0 \tag{3.33}$$

for all ω satisfying (3.30). Finally, we note that (3.27) is trivially satisfied from (3.31), and our choice of ω in (3.30) was selected so that (3.28) is also satisfied. We therefore conclude that, for any $n \times n$ Jacobi matrix J with $\upsilon := \rho(|J|)$ satisfying $0 \leqslant \upsilon < 1$, the associated SSOR matrix S_ω is convergent for all ω satisfying (3.30).

From the equivalence (iii) and (i) of § 1, it is clear that [cf. (1.8)] for any A in $\mathscr{H}_\upsilon$, $\rho(S_\omega(A)) < 1$ for all $0 < \omega \leqslant 1$. Thus, with the above development, the first part of Theorem 1, concerning convergence, has been established.

To complete the proof of Theorem 1, suppose $\omega > \hat{\omega}(\upsilon)$, where $\hat{\omega}(\upsilon)$ is defined in (2.1). It has been shown in [12] that, for n sufficiently large, a particular matrix E can be found in $\mathscr{H}_\upsilon$ (whose associated Jacobi matrix is weakly cyclic of index n) for which

$$\rho(S_\omega(E)) > 1, \tag{3.34}$$

i.e., $S_\omega(E)$ is divergent. Finally, for each υ with $0 \leqslant \upsilon < 1$, it is easy to verify that there is a real positive definite matrix E in $\mathscr{H}_\upsilon$. But it is known (cf. [13, p. 463]) that $\rho(S_\omega(E)) < 1$ implies $0 < \omega < 2$, so that for any $\omega \leqslant 0$ we have $\rho(S(E)) \geqslant 1$, completing the proof. ∎

REFERENCES

1 G. Alefeld and R. S. Varga, Zur Konvergenz des symmetrischen Relaxationsverfahren, *Numer. Math.* 25:291–295 (1976).

2 O. Axelsson, Conjugate gradient type methods for unsymmetric and inconsistent systems of linear equations, *Linear Algebra Appl.* 29:1–16 (1980).

3 A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, New York, 1979.

4 W. Kahan, Gauss-Seidel methods for solving large systems of linear equations, Doctoral Thesis, Univ. of Toronto, 1958.

5 U. Kulisch, Über reguläre Zerlegungen von Matrizen und einige Anwendungen, *Numer. Math.* 11:444–449 (1968).

6 M. Neumann, On bounds for the convergence of the SSOR method for H-matrices, *Linear and Multilinear Algebra* (to appear).

7 M. Neumann and R. S. Varga, On the sharpness of some upper bounds for the spectral radii of S.O.R. iteration matrices, *Numer. Math.* 35:69–79 (1980).

8 A. M. Ostrowski, Über die Determinanten mit überwiegender Hauptdiagonale, *Comment. Math. Helv.* 10:69–96 (1937).
9 R. S. Varga, Factorizations and normalized iterative methods, in *Boundary Problems in Differential Equations* (R. E. Langer, Ed.), Univ. of Wisconsin Press, Madison, 1960, pp. 121–142.
10 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.
11 R. S. Varga, On recurring theorems on diagonal dominance, *Linear Algebra Appl.* 13:1–9 (1976)
12 R. S. Varga, W. Niethammer, and D.-Y. Cai, *p*-cyclic matrices and the symmetric successive overrelaxation method, *Linear Algebra Appl.* 58:425–439 (1984).
13 D. M. Young, *Iterative Solution of Large Linear Systems*, Academic, New York, 1971.

Received 12 August 1983