Theorems of Stein-Rosenberg Type. III. The Singular Case*

John J. Buoni[†]
Youngstown State University
Youngstown, Ohio 44555

Michael Neumann[‡]
University of South Carolina
Columbia, South Carolina 29208

and

Richard S. Varga[§]
Kent State University
Kent, Ohio 44242

Submitted by Hans Schneider

ABSTRACT

In the theory of iterative methods, the classical Stein-Rosenberg theorem can be viewed as giving the simultaneous convergence (or divergence) of the extrapolated Jacobi (JOR) matrix J_ω and the successive overrelaxation (SOR) matrix $\mathcal{L}_\omega$, in the case when the Jacobi matrix J_1 is nonnegative. As has been established recently by Buoni and Varga, necessary and sufficient conditions for the simultaneous convergence (or divergence) of J_ω and $\mathcal{L}_\omega$ have been established which do not depend on the assumption that J_1 is nonnegative. Our aim here is to extend these results to the singular case, using the notion of semiconvergence. In particular, for a real singular matrix A with nonpositive off-diagonal entries, we find conditions (Theorem 3.4) which imply that J_ω and $\mathcal{L}_\omega$ simultaneously semiconverge for all ω in the real interval $[0,1)$.

*Dedicated to Professor John Todd on his seventieth birthday, 16 May 1981.

[†] Research supported by the Air Force Office of Scientific Research at Kent State University during summer 1979 and summer 1980.

[‡] This author would like to thank Kent State University for making funds available for a visit to its Department of Mathematics.

[§] Research supported by the Air Force Office of Scientific Research and the Department of Energy.

1. INTRODUCTION

In Buoni and Varga [3], necessary and sufficient conditions have been given for the *simultaneous* convergence (and divergence) of the successive overrelaxation (SOR) iteration matrix $\mathcal{L}_\omega$ and the extrapolated Jacobi (JOR) iteration matrix J_ω. Such results, of course, are strongly similar in spirit to the classical Stein-Rosenberg theorem (cf. [2, 10, 11, 12]). The main purpose of this paper is to extend the results of [3] to the *singular* case. In particular, for a real singular matrix A with nonpositive off-diagonal entries, we find conditions (Theorem 3.4) which imply that the associated iteration matrices $\mathcal{L}_\omega$ and J_ω simultaneously semiconverge for all ω in the real interval $[0,1)$.

The remainder of this section is devoted to explaining notation and conventions. We denote by $\mathbb{C}^{n,n}$ the collection of all $n\times n$ complex matrices $A=[a_{i,j}]$. Similarly, $\mathbb{C}^n$ denotes the complex n-dimensional vector space of all column vectors $\mathrm{v}:=[v_1,\dots,v_n]^T$, where $v_i\in\mathbb{C}$ for all $1\leqslant i\leqslant n$. The restriction to real entries or components defines $\mathbb{R}^{n,n}$ and $\mathbb{R}^n$. Next, for any A in $\mathbb{C}^{n,n}$, its *spectrum* is denoted as usual by $\sigma(A):=\{\lambda\in\mathbb{C}:\det(\lambda I-A)=0\}$, and its *spectral radius* is denoted by $\rho(A):=\max\{|\lambda|:\lambda\in\sigma(A)\}$. We further set

$$\gamma(A):=\max\{|\lambda|:\lambda\in\sigma(A)\text{ and }\lambda\neq 1\}. \tag{1.1}$$

If $A=[a_{i,j}]$ in $\mathbb{R}^{n,n}$ has only nonnegative real entries, we write $A\geqslant\mathcal{O}$, where $\mathcal{O}$ denotes the null matrix in $\mathbb{C}^{n,n}$.

Next, if $N(A):=\{\mathrm{x}\in\mathbb{C}^n: A\mathrm{x}=\mathbf{0}\}$ denotes the null space of any $A\in\mathbb{C}^{n,n}$, then (cf. Ben-Israel and Greville [1, p. 170])

$$\operatorname{index}(A):=\min\{k: k=0,1,2,\dots,\text{ and } N(A^k)=N(A^{k+1})\},$$

where as usual $A^0:=I$. Note that $\operatorname{index}(A)=0$ iff A is nonsingular, while $\operatorname{index}(A)=k(\geqslant 1)$ iff the maximum of all orders of those Jordan blocks of A which correspond to zero eigenvalues of A is precisely k.

A matrix $A\in\mathbb{C}^{n,n}$ is said to be *convergent* if

$$\lim_{k\to\infty} A^k \tag{1.2}$$

exists and is the zero matrix. It is well known that A is convergent iff $\rho(A)<1$. More generally, if the limit in (1.2) exists, we say that A is *semiconvergent*. Hensel [5], and later Oldenberger [8], have shown that A is semiconvergent iff (i) $\rho(A)\leqslant 1$, (ii) if $\rho(A)=1$, then $\lambda\in\sigma(A)$ with $|\lambda|=1$

implies that $\lambda=1$, and (iii) if $\rho(A)=1$, then index$(I-A)=1$, i.e., all elementary divisors associated with the eigenvalue 1 of A are linear (cf. Berman and Plemmons [2, p. 152]). We note that (i) and (ii) can be equivalently replaced by (iv) $\gamma(A)<1$.

As in Buoni and Varga [3], we split a matrix $A\in\mathbb{C}^{n,n}$ into

$$A=D-L-U \qquad (D, L, U \text{ in } \mathbb{C}^{n,n}). \tag{1.3}$$

We assume throughout that D in (1.3) *is always nonsingular.* Note that we do *not* in general assume that D is diagonal, or that L and U are triangular. Associated with this splitting (1.3) for A is the (generalized) *extrapolated Jacobi* (JOR) *matrix* J_ω, defined for all $\omega\in\mathbb{C}$ by

$$J_\omega:=I-\omega D^{-1}A, \tag{1.4}$$

and the (generalized) *successive overrelaxation* (SOR) *matrix* $\mathcal{L}_\omega$, defined for all $\omega\in\tilde{\mathbb{C}}$ by

$$\mathcal{L}_\omega:=(D-\omega L)^{-1}\{(1-\omega)D+\omega U\}, \tag{1.5}$$

where, for convenience of notation, we set

$$\tilde{\mathbb{C}}:=\{\omega\in\mathbb{C}: D-\omega L \text{ is nonsingular}\}. \tag{1.6}$$

Because D is nonsingular, we note that $\tilde{\mathbb{C}}$ contains all sufficiently small ω. Of course, in the "usual" splitting of (1.3) where D is diagonal and L is strictly lower triangular, we have $\tilde{\mathbb{C}}=\mathbb{C}$.

Next, set $Z^{n,n}:=\{A=[a_{i,j}]\in\mathbb{R}^{n,n}: a_{i,j}\leqslant 0 \text{ for all } i\neq j\}$. Then A is said to be an *M-matrix* if $A\in Z^{n,n}$ and if A can be expressed as

$$A=sI-B, \qquad \text{with } B\geqslant\mathcal{O} \text{ and with } s\geqslant\rho(B). \tag{1.7}$$

It is well known that A is a nonsingular [singular] M-matrix if (1.7) holds with $s>\rho(B)$ [$s=\rho(B)$]. In addition (cf. [2, p. 152]), A is said to be an *M-matrix with property c* if, for some $s>0$, $A=sI-B$ with $B\geqslant\mathcal{O}$ where B/s is semiconvergent.

Next, as in [3], we set

$$\begin{aligned}\Omega_J&:=\{\omega\in\mathbb{C}:\rho(J_\omega)<1\},\\ \mathfrak{D}_J&:=\{\omega\in\mathbb{C}:\rho(J_\omega)>1\}\end{aligned} \tag{1.8}$$

and

$$\begin{aligned}\Omega_{\mathfrak{L}} &:= \{\omega \in \tilde{\mathbb{C}} : \rho(\mathfrak{L}_\omega) < 1\},\\ \mathfrak{D}_{\mathfrak{L}} &:= \{\omega \in \tilde{\mathbb{C}} : \rho(\mathfrak{L}_\omega) > 1\}.\end{aligned} \tag{1.9}$$

With this notation, we can state part of the classical Stein-Rosenberg theorem (cf. [2, 3, 10, 11, 12]) as the following

THEOREM 1.1. *Given $A \in \mathbb{C}^{n,n}$, assume that the splitting of A in* (1.3) *is such that $D^{-1}L$ and $D^{-1}U$ are respectively strictly lower and strictly upper triangular matrices, and assume $J_1 \geqslant \mathcal{O}$. Then*

$$\Omega_J \cap \Omega_{\mathfrak{L}} \supset (0,1] \qquad \text{if} \quad \rho(J_1) < 1, \tag{1.10}$$

and

$$\mathfrak{D}_J \cap \mathfrak{D}_{\mathfrak{L}} \supset (0,1] \qquad \text{if} \quad \rho(J_1) > 1. \tag{1.11}$$

To conclude our discussion of notation and conventions, we set

$$S_J := \{\omega \in \mathbb{C} : J_\omega \text{ is semiconvergent}\} \tag{1.12}$$

and

$$S_{\mathfrak{L}} := \{\omega \in \tilde{\mathbb{C}} : \mathfrak{L}_\omega \text{ is semiconvergent}\}, \tag{1.13}$$

which will be of later use to us.

2. RELATIONSHIPS BETWEEN J_ω AND $\mathfrak{L}_\omega$

In this section, we derive certain relationships between J_ω and $\mathfrak{L}_\omega$.

LEMMA 2.1. *Given any $A \in \mathbb{C}^{n,n}$, then*

$$N(D^{-1}A) = N(I - J_\omega) \quad \textit{and} \quad \text{index}(D^{-1}A) = \text{index}(I - J_\omega) \tag{2.1}$$

for any $0\neq\omega\in\mathbb{C}$. *Similarly,*

$$N(D^{-1}A)=N(I-\mathcal{L}_\omega) \tag{2.2}$$

for any $0\neq\omega\in\tilde{\mathbb{C}}$. *In addition, if* index$(D^{-1}A)=\nu$, *where* $\nu\leqslant 1$, *then also*

$$\text{index}(I-\mathcal{L}_\omega)=\nu \tag{2.3}$$

for all $0\neq\omega\in\mathbb{C}$ *sufficiently small.*

Proof. Since $I-J_\omega=\omega D^{-1}A$ from (1.4), then (2.1) immediately follows for any $0\neq\omega\in\mathbb{C}$. Next, since it can be verified from (1.5) that

$$Q(\omega):=\frac{1}{\omega}(I-\mathcal{L}_\omega)=(I-\omega D^{-1}L)^{-1}D^{-1}A \tag{2.4}$$

for any $0\neq\omega\in\tilde{\mathbb{C}}$, then (2.2) directly follows.

Next, assume that index$(D^{-1}A)=1$ and that $D^{-1}A$ has precisely m $(m>0)$ zero eigenvalues. Thus, in the Jordan normal form of $D^{-1}A$, there are exactly m Jordan blocks, corresponding to the eigenvalue zero, these blocks being all 1×1. Hence, there are m linearly independent eigenvectors of $D^{-1}A$ corresponding to these m zero eigenvalues, and, from (2.4), these m eigenvectors are evidently also eigenvectors of $Q(\omega)$, corresponding to m zero eigenvalues. Because these eigenvectors are linearly independent, the Jordan normal form of $Q(\omega)$ contains at least m Jordan blocks corresponding to the eigenvalue zero, for each $0\neq\omega\in\tilde{\mathbb{C}}$. If $m=n$, so that all eigenvalues of $D^{-1}A$ are zero, then the hypothesis index$(D^{-1}A)=1$ implies that $D^{-1}A\equiv\mathcal{O}$. Thus, from (2.4), $I-\mathcal{L}_\omega\equiv\mathcal{O}$ for all $\omega\in\tilde{\mathbb{C}}$, from which (2.3) follows for $\nu=1$. Hence, we may assume that $m<n$. Now, for small $\omega\neq 0$, we can also write (2.4) as

$$Q(\omega)=D^{-1}A+\omega D^{-1}L(I-\omega D^{-1}L)^{-1}D^{-1}A, \tag{2.5}$$

so that $Q(\omega)$ can be viewed as a perturbation of $D^{-1}A$ for small $\omega\neq 0$. As such, to the remaining $n-m$ nonzero eigenvalues $\{\xi_j\}_{j=1}^{n-m}$ of $D^{-1}A$, we can, by a classical result of Ostrowski [9, p. 334], associate $n-m$ eigenvalues $\{\tau_j(\omega)\}_{j=1}^{n-m}$ of $Q(\omega)$ such that

$$|\xi_j-\tau_j(\omega)|=O(|\omega|^{1/n}) \qquad \text{for all} \quad 1\leqslant j\leqslant n-m, \tag{2.6}$$

for all $\omega\neq 0$ sufficiently small. Because of (2.6), we see that the Jordan normal form of $Q(\omega)$, for $\omega \neq 0$ sufficiently small, then has precisely m Jordan blocks associated with the eigenvalue zero, these blocks being all 1×1. Thus index$(Q(\omega))=1$ for all $\omega\neq 0$ sufficiently small, which gives (2.3) for $\nu=1$.

Finally, if index$(D^{-1}A)=0$, a similar use of (2.6) gives (2.3) for $\nu=0$. ■

To conclude this section, we state without proof the following lemma, which is a slight modification of Buoni and Varga [3, Theorem 2.2].

LEMMA 2.2. *Given any* $A\in\mathbb{C}^{n,n}$ *with* index$(D^{-1}A)\leq 1$, *then, for each* $1\neq\lambda(\omega)\in\sigma(\mathcal{L}_\omega)$, *there exists a* $1\neq\mu(\omega)\in\sigma(J_\omega)$ *such that*

$$|\lambda(\omega)-\mu(\omega)|=O\left(|\omega|^{1+1/n}\right) \tag{2.7}$$

for all $\omega\neq 0$ *sufficiently small.*

3. MAIN RESULTS

We now extend Theorem 3.1 of Buoni and Varga [3] to the simultaneous semiconvergence of J_ω and $\mathcal{L}_\omega$.

THEOREM 3.1. *Given any* $A\in\mathbb{C}^{n,n}$ *with* index$(D^{-1}A)\leq 1$, *assume that if* $\sigma(D^{-1}A)\setminus\{0\}$ *is not empty, then there is a real* $\hat{\theta}$ *with* $0\leq\hat{\theta}<2\pi$ *for which*

$$\min\left[\operatorname{Re}\left(e^{i\hat{\theta}}\xi\right):\xi\in\sigma(D^{-1}A)\setminus\{0\}\right]=:\eta>0. \tag{3.1}$$

Then

$$[S_J\cap S_{\mathcal{L}}]\setminus\{0\}\neq\varnothing. \tag{3.2}$$

More precisely, if $\sigma(D^{-1}A)\setminus\{0\}$ *is empty, then*

$$S_J\cap S_{\mathcal{L}}=\tilde{\mathbb{C}}, \tag{3.3}$$

while if $\sigma(D^{-1}A)\setminus\{0\}$ *is not empty so that* (3.1) *applies, then there exists an* $r_0>0$ *for which*

$$S_J\cap S_{\mathcal{L}}\supset\left\{\omega=re^{i\hat{\theta}}:0\leq r<r_0\right\}. \tag{3.4}$$

Proof. First, consider the case when index $(D^{-1}A)=0$, i.e., $D^{-1}A$ is nonsingular. Obviously, $\sigma(D^{-1}A)\setminus\{0\}$ is nonempty and the hypothesis (3.1) applies. But then, from Theorem 3.1 of [3], $\Omega_J \cap \Omega_{\mathfrak{L}} \supset \{\omega = re^{i\hat{\theta}} : 0<r<r_0\}$ for some $r_0>0$. On the other hand, as $J_0=\mathfrak{L}_0=I$ from (1.4) and (1.5), we always trivially have that $0\in S_J \cap S_{\mathfrak{L}}$. Hence, these two facts imply more than (3.4) in this case. Thus, we may assume that index$(D^{-1}A)=1$. In this case, if $\sigma(D^{-1}A)=\{0\}$, then $D^{-1}A\equiv \mathcal{O}$, so that $J_\omega = I = \mathfrak{L}_\omega$ for all $\omega\in\tilde{\mathbf{C}}$ from (1.4) and (2.4). Thus, $S_J\cap S_{\mathfrak{L}}=\tilde{\mathbf{C}}$ in this case, which gives (3.3). Hence, we may assume in what follows that index$(D^{-1}A)=1$ and that $\sigma(D^{-1}A)\setminus\{0\}$ is not empty.

Consider any $\omega := re^{i\hat{\theta}}$ with $r>0$. From (1.4), we can express any eigenvalue $\mu(\omega)$ of J_ω as

$$\mu(\omega)=1-re^{i\hat{\theta}}\xi, \qquad \text{where} \quad \xi\in\sigma(D^{-1}A). \tag{3.5}$$

Direct computation with (3.5) and (3.1) yields

$$|\mu(\omega)|^2=1-2r\mathrm{Re}\left(e^{i\hat{\theta}}\xi\right)+r^2|\xi|^2\leqslant 1-2r\eta+r^2|\xi|^2, \tag{3.6}$$

for any $0\neq\xi\in\sigma(D^{-1}A)$, from which it follows that there is an $r_1>0$ such that

$$\gamma(J_{re^{i\hat{\theta}}})<1 \qquad \text{for all} \quad 0<r<r_1. \tag{3.7}$$

Next, for any eigenvalue $\xi=0$ of $D^{-1}A$, its associated eigenvalue $\mu(\omega)$ of $J(\omega)$ is necessarily unity from (3.5). Moreover, the hypothesis index$(D^{-1}A)=1$ implies from (2.1) of Lemma 2.1 that index$(I-J_\omega)=1$ for any $0\neq\omega\in\mathbf{C}$. Thus, from the conditions in Section 1 characterizing semiconvergence, J_ω is semiconvergent for all $\omega=re^{i\hat{\theta}}$ with $0\leqslant r<r_1$, i.e.,

$$S_J\supset\left\{\omega=re^{i\hat{\theta}}:0\leqslant r<r_1\right\}. \tag{3.8}$$

Continuing, consider now $\mathfrak{L}_\omega$ for $\omega:=re^{i\hat{\theta}}$ with $r>0$. From (2.3) of Lemma 2.1, it follows that, for all $\omega\neq 0$ sufficiently small, the Jordan blocks corresponding to any eigenvalue unity of $\mathfrak{L}_\omega$ are necessarily 1×1. Moreover, from (2.7) of Lemma 2.2, if $\lambda(\omega)$ is any eigenvalue of $\mathfrak{L}_\omega$ with $\lambda(\omega)\neq 1$, there is an associated eigenvalue $\mu(\omega)$ of J_ω with $\mu(\omega)\neq 1$ such that

$$|\lambda(\omega)-\mu(\omega)|=\mathcal{O}(r^{1+1/n}),$$

for all $\omega\neq 0$ sufficiently small, so that

$$|\lambda(\omega)|^2\leqslant|\mu(\omega)|^2+\mathcal{O}(r^{1+1/n}).$$

Because $\mu(\omega)\neq 1$, we can apply (3.6), which gives

$$|\lambda(\omega)|^2\leqslant 1-2r\eta+\mathcal{O}(r^{1+1/n}). \tag{3.9}$$

Thus, there is an $r_2>0$ such that

$$\gamma(\mathcal{L}_{re^{i\hat{\theta}}})<1 \qquad \text{for all} \quad 0<r<r_2. \tag{3.10}$$

Thus, by the conditions characterizing semiconvergence, it follows that

$$S_{\mathcal{L}}\supset\left\{\omega=re^{i\hat{\theta}}:0\leqslant r<r_2\right\}. \tag{3.11}$$

On choosing $r_0:=\min(r_1;\, r_2)$, then (3.8) and (3.11) give the desired result of (3.4). ■

It is convenient to make the following

DEFINITION 3.2. Given any $A\in\mathbb{C}^{n,n}$, then $D^{-1}A$ is said to be *strongly semistable* if (i) index$(D^{-1}A)\leqslant 1$ and if (ii)

$$\min\left[\operatorname{Re}(\xi):\xi\in\sigma(D^{-1}A)\setminus\{0\}\right]>0, \tag{3.12}$$

provided that $\sigma(D^{-1}A)\setminus\{0\}$ is not empty.

With the above definition, we have, from Theorem 3.1, the particular result of

COROLLARY 3.3. *Given any $A\in\mathbb{C}^{n,n}$ for which $D^{-1}A$ is strongly semistable, then there is an $r_0>0$ for which*

$$S_J\cap S_{\mathcal{L}}\supset[0,r_0). \tag{3.13}$$

In the above Corollary 3.3, it is natural to ask if there is a class of matrices for which, as in the original Stein-Rosenberg theorem [cf. (1.10)], the inclusion of (3.13) holds for the particular interval $[0,1)$. This is given in

THEOREM 3.4. *Given any $A=[a_{i,j}]\in Z^{n,n}$ with $a_{i,i}>0$ for all $1\leqslant i\leqslant n$, let the matrices D, $-L$, and $-U$ in the splitting* (1.3) *of A be respectively the diagonal, the strictly lower triangular, and the strictly upper triangular parts of A. If A is an M-matrix with* $\operatorname{index}(D^{-1}A)\leqslant 1$, *then*

$$S_J\cap S_{\mathfrak{L}}\supset[0,1). \tag{3.14}$$

Proof. Since D is a positive diagonal matrix, the hypotheses imply that $D^{-1}A$ is also an M-matrix with $\operatorname{index}(D^{-1}A)\leqslant 1$. As such, $D^{-1}A$ is strongly semistable, so that (3.13) of Corollary 3.3 gives the existence of an $r_0>0$ with $S_J\cap S_{\mathfrak{L}}\supset[0,r_0)$. Thus, the object of this proof is to show that we can, in fact, choose $r_0=1$. Of course, if $D^{-1}A$ is nonsingular, i.e., $\operatorname{index}(D^{-1}A)=0$, then the statement $\Omega_J\cap\Omega_{\mathfrak{L}}\supset(0,1]$ from (1.10), coupled with the fact that 0 is always in $S_J\cap S_{\mathfrak{L}}$, is stronger than that of (3.14). Hence, we may assume in what follows that $D^{-1}A$ is a singular M-matrix with $\operatorname{index}(D^{-1}A)=1$.

Since F an M-matrix with $\operatorname{index}(F)\leqslant 1$ is equivalent (cf. [2, p. 153]) to F an M-matrix with property c, then evidently $D^{-1}A$ is an M-matrix with property c. Thus every regular splitting of $D^{-1}A=M-N$ (i.e., $M^{-1}\geqslant\mathcal{O}$ and $N\geqslant\mathcal{O}$) satisfies $\rho(M^{-1}N)\leqslant 1$ and $\operatorname{index}(I-M^{-1}N)\leqslant 1$ (cf. [2, p. 200]). However, since $D^{-1}A$ is singular, there is an $\mathbf{x}\neq\mathbf{0}$ such that $D^{-1}A\mathbf{x}=M\mathbf{x}-N\mathbf{x}=\mathbf{0}$, whence $M^{-1}N\mathbf{x}=\mathbf{x}$. Thus, $1\in\sigma(M^{-1}N)$, so that in fact $\rho(M^{-1}N)=1$ and $\operatorname{index}(I-M^{-1}N)=1$. Writing

$$D^{-1}A=I-\tilde{L}-\tilde{U},\qquad \text{where}\quad \tilde{L}:=D^{-1}L \text{ and } \tilde{U}:=D^{-1}U,$$

consider the two splittings $D^{-1}A=M_1-N_1=M_2-N_2$ of $D^{-1}A$, such that

$$M_1:=\frac{1}{r}I \text{ and } N_1:=\frac{1}{r}\left[(1-r)I+r\tilde{L}+r\tilde{U}\right],\qquad \text{where}\quad 0<r\leqslant 1;$$

$$M_2:=\frac{1}{s}(I-s\tilde{L}) \text{ and } N_2:=\frac{1}{s}\left[(1-s)I+s\tilde{U}\right],\qquad \text{where}\quad 0<s\leqslant 1.$$

By definition, $\tilde{L}$ and $\tilde{U}$ are respectively strictly lower and strictly upper triangular nonnegative matrices, and it is readily verified that these two splittings of $D^{-1}A$ are regular splittings for the range of parameters considered. But, as it can be verified that $M_1^{-1}N_1=J_r$ and $M_2^{-1}N_2=\mathfrak{L}_s$, then

$$\begin{aligned}&\rho(J_r)=1 \text{ and } \operatorname{index}(I-J_r)=1 &&\text{for all}\quad 0<r\leqslant 1,\\ &\rho(\mathfrak{L}_s)=1 \text{ and } \operatorname{index}(I-\mathfrak{L}_s)=1 &&\text{for all}\quad 0<s\leqslant 1.\end{aligned} \tag{3.15}$$

We next claim that $\gamma(J_r)<1$ for all $0\leq r<1$. Obviously, as $\gamma(J_0)=0$ since $J_0=I$, we may assume that $0<r<1$. Now from (1.4), $J_r=rJ_1+(1-r)I\geq\mathcal{O}$ for any $0<r<1$. If J_r is irreducible for some (and hence every) r with $0<r<1$, the positivity of the diagonal entries for $0<r<1$ implies that J_r is primitive (cf. [11, Theorem 2.9, p. 49]). Therefore all eigenvalues $\mu(r)\neq 1$ of J_r satisfy $|\mu(r)|<1$ for all $0<r<1$, whence $\gamma(J_r)<1$ for $0\leq r<1$. Similarly, if J_r is reducible for some (and hence every) r with $0<r<1$, each irreducible diagonal block in its reduced normal form (cf. [11, p. 46]) will again have positive diagonal entries and hence be primitive, so that all eigenvalues $\mu(r)\neq 1$ of J_r again satisfy $|\mu(r)|<1$ for all $0<r<1$, whence $\gamma(J_r)<1$ for $0\leq r<1$. Combining this with the first statement of (3.15), we deduce that $S_J\supset[0,1)$.

Finally, $\mathcal{L}_s$ can be expressed

$$\mathcal{L}_s=\{I+s\tilde{L}+\cdots+s^{n-1}\tilde{L}^{n-1}\}\{(1-s)I+s\tilde{U}\},$$

so that $\mathcal{L}_s$, for $0\leq s<1$, is a nonnegative matrix, all of whose diagonal entries are at least $1-s$. As the above argument showing that $\gamma(J_r)<1$ for $0\leq r<1$ similarly shows that $\gamma(\mathcal{L}_s)<1$ for $0\leq s<1$, we deduce, with the second statement of (3.15), that $S_{\mathcal{L}}\supset[0,1)$. Thus, $S_J\cap S_{\mathcal{L}}\supset[0,1)$, giving the desired result of (3.14). ■

We first remark that Theorem 3.4 is equally valid with $\text{index}(D^{-1}A)\leq 1$ replaced by $\text{index}(A)\leq 1$. Also, the real inclusion of (3.14), is *sharp*, i.e., it is not in general possible to increase the real interval $[0,1)$ in (3.14), as examples in Section 4 will amply show.

We next derive a divergence-type result analogous to that of Theorem 3.4.

THEOREM 3.5. *Given any* $A\in\mathbb{C}^{n,n}$ *with* $\text{index}(D^{-1}A)>1$, *then*

$$S_J=\{0\}. \tag{3.16}$$

If $\sigma(D^{-1}A)\setminus\{0\}$ *is not empty, assume for each real* θ *with* $0\leq\theta\leq 2\pi$ *that*

$$\min\left[\operatorname{Re}(e^{i\theta}\xi):\xi\in\sigma(D^{-1}A)\setminus\{0\}\right]\leq 0. \tag{3.17}$$

Then [cf. (1.8)]

$$\mathfrak{D}_J=\mathbb{C}\setminus\{0\}. \tag{3.18}$$

Proof. If index$(D^{-1}A)>1$, it follows from (2.1) of Lemma 2.1 that index$(I-J_\omega)>1$ for any $0\neq\omega\in\mathbb{C}$. By definition, J_ω then fails to be semiconvergent for any $0\neq\omega\in\mathbb{C}$. On the other hand, since $J_0=I$ is trivially semiconvergent, then (3.16) follows.

Next, any eigenvalue $\mu(\omega)$ of J_ω can be expressed, using (1.4), as $\mu(\omega)=1-\omega\xi$ where $\xi\in\sigma(D^{-1}A)$. Writing $\omega=re^{i\theta}$, then

$$|\mu(\omega)|^2=1-2r\mathrm{Re}\left(e^{i\theta}\xi\right)+r^2|\xi|^2. \tag{3.19}$$

If $\sigma(D^{-1}A)\setminus\{0\}$ is not empty, then any $\xi\in\sigma(D^{-1}A)\setminus\{0\}$ is nonzero. With (3.19) and the hypothesis of (3.17), we have, for each $r>0$ and each real θ, that there is $\mu(re^{i\theta})\in\sigma(J_{re^{i\theta}})$ for which

$$|\mu(re^{i\theta})|^2\geqslant 1+r^2|\xi|^2>1.$$

Consequently, $\rho(J_\omega)>1$ for each $0\neq\omega\in\mathbb{C}$, which gives (3.18). ■

Finally, as in [3], we can interpret the conditions (3.1) and (3.17) *geometrically*. Assuming $\sigma(D^{-1}A)\setminus\{0\}$ is not empty, set

$$K\left[\sigma(D^{-1}A)\setminus\{0\}\right]:=\text{closed convex hull of } \sigma(D^{-1}A)\setminus\{0\}. \tag{3.20}$$

With this notation, we establish the following analog of [3, Theorem 3.4].

THEOREM 3.6. *Given $A\in\mathbb{C}^{n,n}$, assume that $\sigma(D^{-1}A)\setminus\{0\}$ is not empty. Then*

$$[S_J\cap S_{\mathcal{L}}]\setminus\{0\}\neq\varnothing \quad \textit{iff} \quad 0\notin K\left[\sigma(D^{-1}A)\setminus\{0\}\right] \textit{ and } \operatorname{index}(D^{-1}A)\leqslant 1. \tag{3.21}$$

Proof. With the hypothesis that $\sigma(D^{-1}A)\setminus\{0\}$ is not empty, assume that index$(D^{-1}A)\leqslant 1$ and that $0\notin K[\sigma(D^{-1}A)\setminus\{0\}]$. This latter assumption implies that there is a real $\hat{\theta}$ with $0\leqslant\hat{\theta}<2\pi$ for which (3.1) is valid. Consequently, from (3.2) of Theorem 3.1, then $[S_J\cap S_{\mathcal{L}}]\setminus\{0\}\neq\varnothing$.

Conversely, suppose that $[S_J\cap S_{\mathcal{L}}]\setminus\{0\}\neq\varnothing$. Then, from Theorem 3.5, it follows that index$(D^{-1}A)\leqslant 1$. Thus, it remains to show that $0\notin K[\sigma(D^{-1}A)\setminus\{0\}]$. Suppose, on the contrary, that $0\in K[\sigma(D^{-1}A)\setminus\{0\}]$, where $\sigma(D^{-1}A)\setminus\{0\}$ is nonempty. It easily follows that (3.17) is then valid, so that from (3.18) of Theorem 3.5, $\mathfrak{D}_J = \mathbb{C}\setminus\{0\}$, whence $S_J = \{0\}$. As this contradicts the assumption that $[S_J \cap S_{\mathfrak{L}}]\setminus\{0\} \neq \varnothing$, then $0 \notin K[\sigma(D^{-1}A)\setminus\{0\}]$. ∎

4. SOME EXAMPLES

In this section, we present five examples, the first three of which show that the real inclusion (3.14) of Theorem 3.4 is *sharp*, i.e., it cannot be in general increased.

Consider first the matrix

$$A_1 := \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{with} \quad D_1 := I. \tag{4.1}$$

Clearly, A_1 is an element of $Z^{5,5}$, and, on writing $A_1 =: I - B_1$, the directed graph of the nonnegative matrix B_1 shows that B_1 is irreducible and primitive with $\rho(B_1) = 1$. Thus, A_1 is a singular M-matrix, and it also follows that $\operatorname{index}(D_1^{-1}A_1) = 1$. Therefore the associated iteration matrices $J_r^{A_1}$ and $\mathfrak{L}_r^{A_1}$ [for the "usual" splitting of (1.3)] are both necessarily semiconvergent for any r in $[0,1)$, from (3.14) of Theorem 3.4. In addition, since $J_1^{A_1} = B_1$, the above properties for B_1 give that $J_1^{A_1}$ is semiconvergent. On the other hand, direct computations give us that

$$\mathfrak{L}_1^{A_1} = \left[\begin{array}{c|cccc} 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$=: \left[\begin{array}{c|cccc} 0 & 1 & 0 & 0 & 0 \\ \hline 0 & & & & \\ 0 & & & & \\ 0 & & E & & \\ 0 & & & & \end{array}\right], \tag{4.2}$$

so that $\sigma(\mathfrak{L}_1^{A_1}) = \{0\} \cup \sigma(E)$. Since the 4×4 matrix E above is nonnegative, irreducible, and cyclic of index 2 with $\rho(E) = 1$, then (cf. [11, p. 38]) ± 1 are eigenvalues of E and $\mathcal{L}_1^{A_1}$. Thus, $\mathcal{L}_1^{A_1}$ cannot be semiconvergent, and thus $1\notin S_J\cap S_{\mathcal{L}}$.

Next, a well-known result of Kahan (cf. [11, p. 75]) gives that

$$\rho\left(\mathcal{L}_r^{A_1}\right)\geqslant|r-1| \qquad \text{for any complex number } r. \tag{4.3}$$

Consequently, for any $\delta>0$, $\rho(\mathcal{L}_{-\delta}^{A_1})\geqslant 1+\delta>1$, whence $-\delta\notin S_J\cap S_{\mathcal{L}}$ for any $\delta>0$. This establishes the sharpness of the inclusion (3.14) in Theorem 3.4 for the matrix A_1 of (4.1).

We remark that Professor Hans Schneider (personal communication) originally used the matrix A_1 of (4.1) to negatively answer a question, posed by Neumann and Plemmons [7, p. 273], on whether $\gamma(\mathcal{L}_1)\leqslant\gamma(J_1)$ is valid for any $A\in Z^{n,n}$ having (i) all diagonal entries of A positive, and (ii) $\rho(J_1)=\rho(\mathcal{L}_1)=1$. The matrix A_1 of (4.1), which satisfies these hypotheses, provides a counterexample, since

$$\gamma\left(J_1^{A_1}\right)<\gamma\left(\mathcal{L}_1^{A_1}\right)=1. \tag{4.4}$$

More precisely, calculations we have performed for this matrix of (4.1) give that

$$\begin{aligned} &\gamma\left(J_r^{A_1}\right)<\gamma\left(\mathcal{L}_r^{A_1}\right)<1 \qquad \text{for all} \quad 0.925943<r<1,\\ &\gamma\left(\mathcal{L}_r^{A_1}\right)<\gamma\left(J_r^{A_1}\right)<1 \qquad \text{for all} \quad 0<r<0.925943. \end{aligned} \tag{4.5}$$

We remark that the inequality in (4.4) also provides a *counterexample* to a result in Berman and Plemmons [2, p. 200, Theorem 6.21, part 2].

Next, consider the matrix

$$A_2:=\begin{bmatrix} 1 & -1 & 0\\ -\frac{1}{2} & 1 & -\frac{1}{2}\\ 0 & -1 & 1 \end{bmatrix}, \qquad \text{with} \quad D_2:=1. \tag{4.6}$$

On writing $A_2=:I-B_2$, the directed graph of the nonnegative matrix B_2 shows that B_2 is cyclic of index 2 with $\sigma(B_2)=\{-1,0,1\}$. Thus, A_2 is a singular M-matrix, and it also follows that index$(D_2^{-1}A_2)=1$. Thus, (3.14) of Theorem 3.4 again gives that the iteration matrices $J_r^{A_2}$ and $\mathcal{L}_r^{A_2}$ both semiconverge for any r in $[0,1)$. Now, however, since $J_1^{A_2}=B_2$, the above cyclic property of B_2 shows that $J_1^{A_2}$ does *not* semiconverge (cf. [6]). On the other hand, since

$$\mathcal{L}_1^{A_2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \tag{4.7}$$

then $\mathcal{L}_1^{A_2}$ semiconverges. Since the inequality (4.3) also applies in this case, we see that the inclusion (3.14) of Theorem 3.4 is sharp for the matrix A_2 of (4.6). In this case, we remark that

$$\gamma\left(\mathcal{L}_r^{A_2}\right) < \gamma\left(J_r^{A_2}\right) < 1 \qquad \text{for all} \quad 0 < r < 1, \tag{4.8}$$

which resembles one of the inequalities of the classical Stein-Rosenberg theorem in the nonsingular case.

Next, consider the matrix

$$A_3 := \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}, \qquad \text{with} \quad D_3 := I. \tag{4.9}$$

It is easily seen that $\sigma(D_3^{-1}A_3) = \{0, 1-\eta, 1-\eta^2\}$, where η is any primitive root of $\eta^3 = 1$, and that A_3 is a singular M-matrix with index$(D_3^{-1}A_3) = 1$. Thus, from Theorem 3.4, $J_r^{A_3}$ and $\mathcal{L}_r^{A_3}$ both semiconverge for all $0 \leqslant r < 1$. However, since $\sigma(J_1^{A_3}) = \{1, \eta, \eta^2\}$ and since $\sigma(\mathcal{L}_1^{A_3}) = \{-1, 0, +1\}$, neither $J_1^{A_3}$ nor $\mathcal{L}_1^{A_3}$ semiconverges. Since the inequality (4.3) also applies in this example, we thus see that the inclusion (3.14) of Theorem 3.4 is sharp for the matrix A_3 of (4.9). In this case, as in the previous example, we have that

$$\gamma\left(\mathcal{L}_r^{A_3}\right) < \gamma\left(J_r^{A_3}\right) < 1 \qquad \text{for all} \quad 0 < r < 1. \tag{4.10}$$

We remark that the three examples given above were selected to illustrate all possible situations concerning the nonsemiconvergence of $\mathcal{L}_1$ and/or J_1 for singular M-matrices with index$(D^{-1}A) = 1$.

Finally, in each of the three examples given above, it is the case that there exists an $r_i > 0$ such that [cf. (4.5), (4.8), (4.10)]

$$\gamma\left(\mathcal{L}_r^{A_i}\right) < \gamma\left(J_r^{A_i}\right) < 1 \quad \text{for all } 0 < r < r_i, \qquad i = 1, 2, 3. \tag{4.11}$$

This, in spirit, resembles a consequence of the classical Stein-Rosenberg theorem in the convergent case, i.e., $\rho(\mathfrak{L}_r^A)<\rho(J_r^A)<1$ for all $0<r\leqslant 1$. That (4.11) *fails* to be true for every singular M-matrix with index$(D^{-1}A)=1$ is the point of our next example.

Consider the matrix

$$A_4:=\begin{bmatrix} 1 & -1 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}, \quad \text{with} \quad D_4:=I. \tag{4.12}$$

Reasoning as in the previous examples, A_4 is a singular M-matrix with index $(D_4^{-1}A)=1$, so that $S_J \cap S_{\mathfrak{L}} \supseteq [0,1)$ from (3.14) of Theorem 3.4. In this case, it can be vierfied that $\sigma(A_4)=\{0,\frac{3}{2},\frac{3}{2}\}$, and that index$(\frac{3}{2}I-A_4)=2$. Moreover, it can be further verified that

$$\begin{aligned} &\gamma(J_r^{A_4})<\gamma(\mathfrak{L}_r^{A_4})<1 \quad &&\text{for all} \quad 0<r<\tfrac{3}{4}, \\ &\gamma(\mathfrak{L}_r^{A_4})<\gamma(J_r^{A_4})<1 \quad &&\text{for all} \quad \tfrac{3}{4}<r<\tfrac{4}{3}. \end{aligned} \tag{4.13}$$

The first of the inequalities of (4.13) thus provides a counterexample to (4.11) holding for all singular M-matrices with index$(D^{-1}A)=1$. It is interesting to remark that there is further intertwining of $\gamma(J_r^{A_4})$ and $\gamma(\mathfrak{L}_r^{A_4})$ in this example, in that

$$\begin{aligned} &\gamma(J_r^{A_4})<\gamma(\mathfrak{L}_r^{A_4}) \quad &&\text{for all} \quad 0<r<\tfrac{3}{4}, \\ &\gamma(\mathfrak{L}_r^{A_4})<\gamma(J_r^{A_4}) \quad &&\text{for all} \quad \tfrac{3}{4}<r<3.154701, \\ &\gamma(J_r^{A_4})<\gamma(\mathfrak{L}_r^{A_4}) \quad &&\text{for all} \quad 3.154701<r<\infty. \end{aligned} \tag{4.14}$$

We remark that the inequalities of the first display of (4.13) are just the *reverse* of what one expects in the classical Stein-Rosenberg theorem, and this is due in this case to the fact that index$(\frac{3}{2}I-A_4)=2$. (For an explanation of this inequality reversal, see Buoni and Varga [4].)

We finally remark that it is *not necessary* that the index of some nonzero eigenvalue [of a singular M-matrix with index$(D^{-1}A)=1$] exceed unity to achieve both a counterexample to (4.11) and an analog of the intertwining of (4.14). Specifically, with

$$A_5:=\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad \text{with} \quad D_5:=I, \tag{4.15}$$

it can be verified that

$$\begin{aligned}
&\gamma(J_r^{A_5})<\gamma(\mathcal{L}_r^{A_5}) && \text{for all} \quad -\infty<r<2-\sqrt{2}, \\
&\gamma(\mathcal{L}_r^{A_5})<\gamma(J_r^{A_5}) && \text{for all} \quad 2-\sqrt{2}<r<2+\sqrt{2}, \\
&\gamma(J_r^{A_5})<\gamma(\mathcal{L}_r^{A_5}) && \text{for all} \quad 2+\sqrt{2}<r<\infty.
\end{aligned} \tag{4.16}$$

REFERENCES

1 A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Wiley, New York, 1974.
2 A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic, New York, 1979.
3 J. J. Buoni and R. S. Varga, Theorems of Stein-Rosenberg type, in *Numerical Mathematics* (R. Ansorge, K. Glashoff, and B. Werner, eds.), ISNM 49, Birkäuser, Basel, 1979, pp. 65–75.
4 J. J. Buoni and R. S. Varga, Theorems of Stein-Rosenberg type II. Optimal paths of relaxation in the complex domain, in *Elliptic Problem Solvers* (M. H. Schultz, ed.), Academic Press, New York, 1981, pp. 231–240.
5 K. Hensel, Über Potenzreihen von Matrizen, *J. Reine Angew. Math.* 155:107–110 (1926).
6 C. D. Meyer and R. J. Plemmons, Convergent powers of a matrix with applications to iterative methods for singular linear systems, *SIAM J. Numer. Anal.* 14:699–705 (1977).
7 M. Neumann and R. J. Plemmons, Convergent nonnegative matrices and iterative methods for consistent linear systems, *Numer. Math.* 31:265–279 (1978).
8 R. Oldenburger, Infinite powers of matrices and characteristic roots, *Duke Math. J.* 6:357–361 (1940).
9 A. M. Ostrowski, *Solution of Equations in Euclidean and Banach Spaces*, Academic, New York, 1973.
10 P. Stein and R. Rosenberg, On the solution of linear simultaneous equations by iteration, *J. London Math. Soc.* 23:111–118 (1948).
11 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J., 1962.
12 D. M. Young, *Iterative Solution of Large Linear Systems*, Academic, New York, 1971.

Received 18 February 1981; revised 26 May 1981