Numer. Math. 35, 69–79 (1980)

Numerische Mathematik

On the Sharpness of Some Upper Bounds for the Spectral Radii of S.O.R. Iteration Matrices

Michael Neumann[1] and Richard S. Varga[2]*

[1] Department of Mathematics, The University of Nottingham, Nottingham NG 7 2 RD, England
[2] Department of Mathematics, Kent State University, Kent, OH 44242, USA

Summary. Sharpness is shown for three upper bounds for the spectral radii of point S.O.R. iteration matrices resulting from the splitting (i) of a nonsingular H-matrix A into the 'usual' $D-L-U$, and (ii) of an hermitian positive definite matrix A into $D-L-U$, where D is hermitian positive definite and $L=\frac{1}{2}(A-D+S)$ with S some skew-hermitian matrix. The first upper bound (which is related to the splitting in (i)) is due to Kahan [6], Apostolatos and Kulisch [1] and Kulisch [7], while the remaining upper bounds (which are related to the splitting in (ii)) are due to Varga [11]. The considerations regarding the first bound yield an answer to a question which, in essence, was recently posed by Professor Ridgway Scott: What is the largest interval in ω, $\omega \geqq 0$, for which the point S.O.R. iterative method is convergent for all strictly diagonally dominant matrices of arbitrary order? The answer is, precisely, the interval (0, 1].

Subject Classifications: AMS(MOS): 65F10, CR: 5.14.

1. Introduction and Notations

The importance of *relaxation methods* for the iterative solution of nonsingular systems

$$Ax=b, \tag{1.1}$$

arising in the (possibly approximate) solution to practical problems is well known, e.g., [4, 10, 13 and 14]. This is further reflected in the works of many other authors concerning particular aspects or applications of these methods.

As indicated in the Summary, we are concerned here with the sharpness of three known upper bounds for the spectral radii of S.O.R. iteration matrices arising from two classes of matrices, the class of all nonsingular complex H-matrices, and the class of all hermitian positive definite matrices. However, the

* Research supported in part by the Air Force Office of Scientific Research, and the Department of Energy

0029-599X/80/0035/0069/$02.20

S.O.R. iteration matrices relating to the latter class, will be slightly more general than the conventional S.O.R. iteration matrices.

Let $\mathbb{C}^{n,n}$ denote the set of all $n \times n$ complex matrices. A matrix $A=(a_{i,j}) \in \mathbb{C}^{n,n}$ is called a *nonsingular H-matrix* if it can be scaled, via multiplication by a nonsingular diagonal matrix, to be a strictly diagonally dominant matrix, that is, if a nonsingular diagonal matrix $E=(e_{i,j}) \in \mathbb{C}^{n,n}$ can be found such that

$$|a_{i,i} e_{i,i}| > \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{i,j} e_{j,j}|, \qquad i=1, 2, \ldots, n.$$

(Thus, every strictly diagonally dominant matrix is, in particular, a nonsingular H-matrix.) The theory of nonsingular H-matrices, as has been shown in [12], is intimately related to the theory of nonsingular M-matrices, and many equivalent conditions for a matrix $A \in \mathbb{C}^{n,n}$ to be a nonsingular H-matrix can be found there, and also in Berman and Plemmons [2].

If $A \in \mathbb{C}^{n,n}$ is a nonsingular H-matrix then, evidently, its diagonal entries are nonzero, so that A may be split into

$$A = D - L - U := D(I - B^A), \tag{1.2}$$

where D, $-L$ and $-U$ are diagonal, strictly lower triangular and strictly upper triangular matrices, respectively. Furthermore, for any scalar ω, we can form the S.O.R. iteration matrix

$$\mathscr{L}_\omega^A := (D - \omega L)^{-1} [(1-\omega) D + \omega U].$$

It is well known that the iterative scheme

$$x_i = \mathscr{L}_\omega^A x_{i-1} + \omega (D - \omega L)^{-1} b, \qquad i = 1, 2, \ldots,$$

will converge to the unique solution of the system (1.1) from any starting vector x_0 if and only if the spectral radius of $\mathscr{L}_\omega^A$, $\rho(\mathscr{L}_\omega^A)$, satisfies $\rho(\mathscr{L}_\omega^A) < 1$. We further remark that the matrix B^A in (1.2) is known in the literature as the point Jacobi iteration matrix associated with A, and as shown in [10] (see also [12]), if A is a nonsingular H-matrix, then $\rho(|B^A|) < 1$, where for $T = (t_{i,j}) \in \mathbb{C}^{n,n}$, $|T| := (|t_{i,j}|)$.

In Kahan [6], Apostolatos and Kulisch [1], and Kulisch [7], it was established that, for a nonsingular H-matrix A,

(I) $\quad \rho(\mathscr{L}_\omega^A) \leqq |\omega - 1| + \omega \rho(|B^A|) \qquad$ for all $0 \leqq \omega \leqq 2/(1 + \rho(|B^A|))$,

and furthermore, that for any $0 < \omega < 2/(1 + \rho(|B^A|))$,

$$\rho(\mathscr{L}_\omega^A) < 1.$$

The first objective of this paper will be to show that inequality (I) is *sharp* for the class of all nonsingular H-matrices of arbitrary orders. This and some resulting corollaries will be established in Sect. 2.

Generalizations of the 'conventional' splitting of (1.2) (where D, $-L$, and $-U$ are diagonal, strictly lower triangular and strictly upper triangular, resp.) have been suggested and studied by several authors, e.g., [3, 5, 10, 11, 13 and 14].

The particular generalization which will concern us here was developed in [11]. There, $A \in \mathbb{C}^{n,n}$ is assumed hermitian positive definite (that is, $A = A^*$, the conjugate transpose of A, and the eigenvalues of A are all positive) and

$$A = D - L - U := D(I - \tilde{B}), \tag{1.3}$$

where D is (also) hermitian positive definite, and

$$L = \tfrac{1}{2}(D - A + S) \tag{1.4}$$

with S some skew-hermitian matrix (i.e., $S = -S^*$). It follows from (1.3) and (1.4) that $U = L^*$.

For the above generalization and for $\omega \in [0, 2]$, Varga [11] considered the S.O.R. iteration matrix

$$\begin{aligned} \mathscr{L}_{\omega, A} &:= (D - \omega L)^{-1} [(1-\omega) D + \omega U] \\ &= (D - \omega L)^{-1} [(1-\omega) D + \omega L^*]. \end{aligned}$$

Let

$$E_\omega = \{v \in \mathbb{C}^n \colon (v, Dv) = 1,\ \mathscr{L}_{\omega, A} v = \xi v,\ |\xi| = \rho(\mathscr{L}_{\omega, A})\},$$

where $\mathbb{C}^n$ denotes the n-dimensional complex space and $(.,.)$ denotes the (usual) inner product in $\mathbb{C}^n$, and set

$$\tau_\omega := \inf \{|(v, Sv)| \colon v \in E_\omega\}. \tag{1.5}$$

Under the assumption that

$$\tau_\omega = 0 \tag{1.6}$$

for all $\omega \in [0, 2]$, it was shown in [11] that for $\rho(\tilde{B}) < 1$ and with

$$\omega_b := \frac{2}{1 + \sqrt{1 - \rho^2(\tilde{B})}},$$

then

(II) $$\rho(\mathscr{L}_{\omega, A}) \leqq \frac{2(1-\omega)\, \omega\, \rho(\tilde{B})}{2 - \omega\, \rho(\tilde{B})} \quad \text{for all } \omega \in [0, \omega_b],$$

and that

(III) $$\min \{\rho(\mathscr{L}_{\omega, A}) \colon 0 \leqq \omega \leqq 2\} \leqq \sqrt{\omega_b - 1}.$$

In Sect. 3, we shall prove that both (II) and (III) are *sharp* upper bounds (for the aforementioned type of S.O.R. splitting satisfying the above requirements). However, our approach here will not employ Rayleigh quotients which were the main vehicle used in [11] for obtaining upper bounds (II) and (III), but rather, we shall make a more direct use of the spectral mapping theorem.

Finally, for a matrix $A \in \mathbb{C}^{n,n}$, $\sigma(A)$ will denote the spectrum of A.

2. The Sharpness of Upper Bound (I)

Let ν be any number in $[0, 1)$, and let $\mathscr{H}_\nu$ be the subset of nonsingular H-matrices defined by

$\mathscr{H}_\nu := \{A \in \mathbb{C}^{n,n}$, n arbitrary: A is a nonsingular H-matrix with $\rho(|B^A|) = \nu\}$, where B^A is given in (1.2). It follows from (I) that

$$\sup\{\rho(\mathscr{L}_\omega^A) : A \in \mathscr{H}_\nu\} \leqq |\omega - 1| + \omega\nu, \quad \forall\, 0 \leqq \omega \leqq 2/(1+\nu). \tag{2.1}$$

The sharpness of this inequality is established in

Theorem 1. *For each* $0 \leqq \nu < 1$, *and for all* $0 \leqq \omega \leqq 2/(1+\nu)$,

$$\sup\{\rho(\mathscr{L}_\omega^A) : A \in \mathscr{H}_\nu\} = |\omega - 1| + \omega\nu. \tag{2.2}$$

Remark. Although Kahan [6] originally established the case $0 < \omega \leqq 1$ of Theorem 1, the remaining case $1 < \omega \leqq 2/(1+\nu)$ of Theorem 1 has remained open in the literature. Because [6] is generally not accessible, we give a proof of (2.2) for the complete range $0 \leqq \omega \leqq 2/(1+\nu)$.

Proof. For $\omega = 0$, (2.2) trivially holds for each $\nu \in [0, 1)$ since $1 = \rho(I) = \rho(\mathscr{L}_0^A)$. Next, fix ν in $[0, 1)$ and assume first that $0 < \omega \leqq 1$. For any positive integer $p \geqq 2$, consider the $p \times p$ matrices

$$A := I - B^A, \tag{2.3}$$

where

$$B^A := \nu \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & 0 & 1 & \\ & & \ddots & \ddots \\ 0 & & & 1 \\ 1 & 0 & \dots & 0 \end{bmatrix}.$$

Then, in (2.3), B^A is an irreducible nonnegative cyclic of index p matrix (e.g., [10]) with $\rho(B^A) = \rho(|B^A|) = \nu$, whence $A \in \mathscr{H}_\nu$. Because of the cyclic character of B^A, the eigenvalues, λ, of its associated S.O.R. matrix $\mathscr{L}_\omega^A$ and the eigenvalues, μ, of B^A are related, e.g., [9, 10 and 8] by

$$(\lambda + \omega - 1)^p = \lambda\,\omega^p \mu^p.$$

More precisely, since ν is an eigenvalue of B^A by the Perron-Frobenius Theorem, each solution λ, e.g., [9, 10], of

$$(\lambda + \omega - 1)^p = \lambda\,\omega^p \nu^p \tag{2.4}$$

is an eigenvalue of $\mathscr{L}_\omega^A$. Now, because $0 < \omega \leqq 1$, it follows that, as $p \to \infty$, there is, for p sufficiently large, a nonnegative solution λ' of (2.4) (which is an eigenvalue of $\mathscr{L}_\omega^A$ and) which is given by

$$\lambda' = \delta(\omega, \nu) - \frac{\omega\nu \ln(\delta(\omega, \nu))^{-1}}{p} + O\left(\frac{1}{p^2}\right),$$

where, for convenience, $\delta(\omega, \nu) := |\omega - 1| + \omega \nu$. Consequently,

$$\rho(\mathscr{L}_\omega^A) \geqq \lambda'.$$

But since $\rho(\mathscr{L}_\omega^A) \leqq \delta(\omega, \nu)$ by (I), then evidently

$$\sup\{\rho(\mathscr{L}_\omega^A) : A \in \mathscr{H}_\nu\} = \delta(\omega, \nu)$$

for all $0 \leqq \omega \leqq 1$, so that (2.2) is valid in this range.

For the remaining case $1 \leqq \omega \leqq 2/(1+\nu)$, for each positive integer p, consider the $(2p) \times (2p)$ matrices

$$\tilde{A} = I - B^{\tilde{A}}, \tag{2.5}$$

where

$$B^{\tilde{A}} := \nu \begin{bmatrix} 0 & 1 & & & 0 \\ & & 0 & \ddots & \\ \vdots & & & \ddots & \\ 0 & & & & 1 \\ -1 & 0 & \dots & & 0 \end{bmatrix},$$

so that, in this case, $B^{\tilde{A}}$ is a weakly cyclic matrix of index $2p$ (e.g., [10]), with $\rho(|B^{\tilde{A}}|) = \nu$. Thus, $\tilde{A} \in \mathscr{H}_\nu$ for every $p \geqq 1$. Here, $-\nu^{2p}$ is an eigenvalue of $(B^{\tilde{A}})^{2p}$, and, as before, any solution λ of the equation

$$(\lambda + \omega - 1)^{2p} = -\lambda \omega^{2p} \nu^{2p} \tag{2.6}$$

is an eigenvalue of $\mathscr{L}_\omega^{\tilde{A}}$. But then, we can again verify that for large enough p,

$$\lambda'' = -\left[\delta(\omega, \nu) - \frac{\omega \nu \ln(\delta(\omega, \nu))^{-1}}{2p} + O\left(\frac{1}{p^2}\right)\right]$$

is a negative real solution to (2.6), so that with (I),

$$-\lambda'' \leqq \rho(\mathscr{L}_\omega^{\tilde{A}}) \leqq \delta(\omega, \nu).$$

Thus, (2.2) is also valid for $\omega \in [1, 2/(1+\nu)]$. ■

Theorem 1 yields several corollaries.

Corollary 1. $\rho(\mathscr{L}_\omega^A) < 1$ *for all* $A \in \mathscr{H}_\nu$ *if and only if* $0 < \omega < 2/(1+\nu)$.

As an illustration that the conclusion of Corollary 1 is, indeed, not valid for the case when $\omega = 2/(1+\nu)$, consider the 2×2 matrix

$$\tilde{A} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & \nu \\ -\nu & 0 \end{pmatrix} := I - B^{\tilde{A}}$$

where $0 < \nu < 1$. Clearly, $\tilde{A}$ is in $\mathscr{H}_\nu$. Next, since $B^{\tilde{A}}$ is weakly cyclic of index 2 with $i\nu \in \sigma(B^{\tilde{A}})$, any solution λ of the equation

$$(\lambda + \omega - 1)^2 = -\lambda \omega^2 \nu^2 \tag{2.7}$$

is an eigenvalue of $\mathscr{L}_\omega^{\tilde{A}}$. Rewriting (2.7) as a quadratic equation in λ, one obtains

$$\lambda^2+\lambda[2(\omega-1)+\omega^2 v^2]+(\omega-1)^2=0,$$

which, for $\omega=2/(1+v)$, has two real roots $\lambda_1=-1$ and $\lambda_2=-[(1-v)/(1+v)]^2$. Thus, $\rho(\mathscr{L}_{2/(1+v)})=1$.

Next, since $2/(1+v)\to 1$ as $v\to 1$, we may further state

Corollary 2. $\rho(\mathscr{L}_\omega^A)<1$ *for all*

$$A\in\bigcup_{0\leqq v<1}\mathscr{H}_v$$

if and only if $0<\omega\leqq 1$.

Moreover, since in the proof of Theorem 1 only strictly diagonally dominant matrices were used, Corollary 2 may be strengthened, so as to provide an answer to Professor Scott's question, as follows.

Corollary 3. *Let SDD be the set of all complex strictly diagonally dominant matrices of all orders. Then* $\rho(\mathscr{L}_\omega^A)<1$ *for all* $A\in SDD$ *if and only if* $0<\omega\leqq 1$.

The authors wish to remark that R. Scott (in a personal communication) has now provided a different proof of the result of Corollary 3.

3. The Sharpness of Upper Bounds (II) and (III)

We shall approach the problem of showing that the remaining upper bounds (namely (II) and (III)) are sharp by considering a special type of splitting for nonsingular matrices in $\mathbb{C}^{n,n}$ which are not necessarily hermitian positive definite. It should be remarked, though, that the sharpness of these upper bounds could be inferred from the material developed in [11].

Let $A\in\mathbb{C}^{n,n}$ be a nonsingular matrix, and consider the splitting of A into

$$A=D-L-U, \tag{3.1}$$

where D is some nonsingular matrix and where

$$L=U:=(D-A)/2. \tag{3.2}$$

Set

$$A:=D(I-B'), \tag{3.3}$$

and define the following set

$$\Omega_D:=\{\omega\neq 0 \text{ real}: \det(D-\omega L)\neq 0\}.$$

It is easy to verify that with L and U chosen as in (3.2) and with B' given by (3.3), the point S.O.R. iteration matrix associated with any $\omega\in\Omega_D$ has the following representation

$$\mathscr{L}'_\omega=(I-\omega B'/2)^{-1}[(1-\omega)I+\omega B'/2]. \tag{3.4}$$

Before proceeding to establish the main objective of this section, we present several noteworthy facts concerning the S.O.R. iteration matrix of (3.4). First, since $1 \notin \sigma(\mathscr{L}'_\omega)$ for any $\omega \in \Omega_D$, $I - \mathscr{L}'_\omega$ is nonsingular, and we find that

$$(I - \mathscr{L}'_\omega)^{-1}(I + \mathscr{L}'_\omega) = \left(\frac{2-\omega}{\omega}\right) A^{-1} D. \tag{3.5}$$

Recall that a matrix in $\mathbb{C}^{n,n}$ is called *positive stable* if all its eigenvalues have a positive real part. Thus, we have

Proposition 1. *Let A and D be nonsingular matrices in $\mathbb{C}^{n,n}$ and consider the set Ω_D. Then, for each $\omega \in \Omega_D$ for which the r.h.s. of* (3.5) *is positive stable, we have*

$$\rho(\mathscr{L}'_\omega) < 1. \tag{3.6}$$

In particular, if $A^{-1}D$ is positive stable and $(0, 2) \subset \Omega_D$, then (3.6) *holds for each $\omega \in (0, 2)$.*

Proof. The proof follows at once from Theorem 6.3 in [14, p. 82]. ∎

Corollary 4. *If A is an hermitian positive definite matrix and D is a matrix in $\mathbb{C}^{n,n}$ such that $D + D^*$ is (hermitian) positive definite and such that $(0, 2) \subset \Omega_D$, then* (3.6) *holds for each $\omega \in (0, 2)$.*

Corollary 5. *If A is a nonsingular M-matrix and D is a diagonal matrix whose diagonal entries are all positive and $(0, 2) \subset \Omega_D$, then* (3.6) *holds for each $\omega \in (0, 2)$.*

Remark. For the definition and the many pertinent properties of a nonsingular M-matrix, see Berman and Plemmons [2].

The next proposition is a Stein-Rosenberg type result. Its proof requires simple use of the spectral mapping theorem and is therefore omitted.

Proposition 2. *Let A and D be nonsingular matrices in $\mathbb{C}^{n,n}$, and define $B'_\omega := (1 - \omega) I + \omega B'$ to be the Jacobi overrelaxation matrix, where B' is given by* (3.3). *If $\rho(B') < 1$, then $(0, 2) \in \Omega_D$, and*

$$\rho(\mathscr{L}'_\omega) \leqq \rho(B'_\omega) \quad \textit{for all } \omega \in (0, 1],$$

with strict inequality holding if $\rho(B') \neq 0$.

We now return to the main theme of this section. Suppose then that $A \in \mathbb{C}^{n,n}$ is a nonsingular matrix and that $D \in \mathbb{C}^{n,n}$ is a nonsingular matrix such that $\rho(B')$ (see (3.3)) is less than unity. Then, as indicated in Proposition 2, $(0, 2) \in \Omega_D$ and it follows that for each $\omega \in (0, 2)$, $\lambda_\omega \in \sigma(\mathscr{L}'_\omega)$ implies that $\lambda_\omega \neq -1$. Thus, by (3.4) for each $\omega \in (0, 2)$, $\mu \in \sigma(B')$ if and only if

$$\mu = \frac{2(\lambda_\omega + \omega - 1)}{\omega(\lambda_\omega + 1)} \tag{3.7}$$

for some $\lambda_\omega \in \sigma(\mathscr{L}'_\omega)$, in which case each $\lambda_\omega \in \sigma(\mathscr{L}'_\omega)$, $\omega \in (0, 2)$, has the representation

$$\lambda_\omega = \frac{\omega\mu - 2\omega + 2}{2 - \omega\mu} \tag{3.8}$$

for some $\mu \in \sigma(B')$. Furthermore, if we assume that

$$\nu := \rho(B') \neq 0,$$

then λ_ω given by (3.8) satisfies

$$\lambda_\omega = \frac{\omega(\mu/\nu) + (2/\nu)(1-\omega)}{(2/\nu) - \omega(\mu/\nu)}.$$

Hence, for $\omega \in (0, 2)$,

$$\rho(\mathscr{L}'_\omega) = \max_{\mu \in \sigma(B)} \left| \frac{\omega(\mu/\nu) + (2/\nu)(1-\omega)}{\omega(\mu/\nu) - (2/\nu)} \right|. \tag{3.9}$$

The next lemma will allow us to obtain a more explicit expression for the spectral radius of $\mathscr{L}'_\omega$ than that of (3.9). Its proof, being straightforward, is omitted.

Lemma 1. *Let α, β and γ be real numbers with $|\gamma| > |\alpha| > 0$. Then,*

$$\max_{|z| \leqq 1} \left| \frac{\alpha z + \beta}{\alpha z + \gamma} \right| = \max \left\{ \left| \frac{\alpha + \beta}{\alpha + \gamma} \right|; \left| \frac{\alpha - \beta}{\alpha - \gamma} \right| \right\}. \tag{3.10}$$

In addition to our previous assumptions (that $0 < \nu = \rho(B') < 1$), suppose now that $\pm \nu \in \sigma(B')$. Then since $2/\nu > \omega$ for all $\omega \in (0, 2)$, Lemma 1 has the implication that the l.h.s. of (3.9) can be (further) determined from the expression

$$\rho(\mathscr{L}'_\omega) = \max \left\{ \left| \frac{\omega + (2/\nu)(1-\omega)}{(2/\nu) - \omega} \right|; \left| \frac{\omega - (2/\nu)(1-\omega)}{\omega + (2/\nu)} \right| \right\}, \tag{3.11}$$

for each $\omega \in (0, 2)$. To establish the precise behaviour of $\rho(\mathscr{L}'_\omega)$ as a function of ω in the interval $(0, 2)$, we make use of the following (technical) lemma.

Lemma 2. *For $\omega \in (0, 2)$, define*

$$M_1(\omega) := \frac{\omega + (2/\nu)(1-\omega)}{(2/\nu) - \omega}, \tag{3.12}$$

and

$$M_2(\omega) := \frac{\omega - (2/\nu)(1-\omega)}{(2/\nu) + \omega}.$$

Then:

(i) $M_1(\omega) = |M_1(\omega)| > |M_2(\omega)|$, $\omega \in (0, 1]$,
(ii) $M_2(\omega) = |M_2(\omega)| > |M_1(\omega)|$, $\omega \in (\omega', 2)$,

where

$$\omega' := \frac{(2/\nu)}{(2/\nu) - 1}.$$

(iii) *In the interval* $(0, \omega']$, *the function* $M_1(\omega)$ *is nonnegative and strictly decreasing.*

(iv) *In the interval* $[1, 2)$, *the function* $M_2(\omega)$ *is positive and strictly increasing.*

Proof. The proof is obtained by verification. ■

The findings in Lemma 2 together with (3.11) imply that there exists a point $\omega_b \in (1, \omega')$ such that as ω varies in $(0, 2)$, $\rho(\mathscr{L}'_\omega) = M_1(\omega)$ in $(0, \omega_b)$ and $\rho(\mathscr{L}'_\omega) = M_2(\omega)$ in $(\omega_b, 2)$, while $M_1(\omega_b) = M_2(\omega_b)$ and

$$\rho(\mathscr{L}'_{\omega_b}) = \min_{\omega \in (0, 2)} \rho(\mathscr{L}'_\omega).$$

To find ω_b, consider the set of $\omega \in (0, 2)$ for which

$$M_1(\omega) - M_2(\omega) = \frac{2[\omega^2 - (4/v^2)\,\omega + (4/v^2)]}{[(2/v) - \omega]\,[(2/v) + \omega]} = 0. \tag{3.13}$$

Since the denominator of (3.13) is nonzero for all $\omega \in (0, 2)$, the set of points in $(0, 2)$ which satisfy (3.13) are the solution in that interval of the quadratic equation

$$\omega^2 - (4/v^2)\,\omega + (4/v^2) = 0. \tag{3.14}$$

Of the two solutions to (3.14), the only one which lies in $(0, 2)$ (and, indeed, as can be verified, belongs to $(1, \omega')$) is given by

$$\omega_b = \frac{2}{v^2} - \frac{2}{v}\sqrt{\frac{1}{v^2} - 1} = \frac{2}{1 + \sqrt{1 - v^2}}, \tag{3.15}$$

(which coincides with the optimal relaxation parameter in the classical S.O.R. theory, e.g., [10] and [14]). We may now prove

Lemma 3. *For* ω_b *given by* (3.15),

$$\rho(\mathscr{L}'_{\omega_b}) = \sqrt{\omega_b - 1}. \tag{3.16}$$

Proof. Since $\rho(\mathscr{L}'_{\omega_b}) = M_1(\omega_b)$ from our comments following Lemma 2 and (3.13), we have that

$$\rho(\mathscr{L}'_{\omega_b}) = \frac{\omega_b + (2/v)(1 - \omega_b)}{(2/v) - \omega_b}. \tag{3.17}$$

Substituting the r.h.s. of (3.15) in (3.17), we obtain that

$$\rho(\mathscr{L}'_{\omega_b}) = \frac{\sqrt{(1 - v^2)} - (1 - v)}{\sqrt{(1 - v^2)} + (1 - v)}.$$

But then

$$\rho^2(\mathscr{L}'_{\omega_b}) + 1 = \frac{2}{1 + \sqrt{1 - v^2}} = \omega_b$$

by (3.15), establishing (3.16). ■

Theorem 2. *The upper bounds* (II) *and* (III) *are sharp.*

Proof. Let $\hat{B}\in\mathbb{C}^{n,n}$ be an hermitian matrix with $\sigma(\hat{B})=\{\pm\hat{v}\}$, where $0<\hat{v}<1$, and consider the matrix

$$\hat{A}:=D-\hat{B}:=I-\hat{B}. \tag{3.18}$$

Let $L=\frac{1}{2}(I-\hat{A})=\frac{1}{2}\hat{B}$, so that on setting

$$\hat{A}=I-L-U, \tag{3.19}$$

we have that $U=L$. From the definition of $\hat{B}$, it follows that $0\neq\pm\hat{v}=\pm\rho(\hat{B})\in\sigma(\hat{B})$, while $\rho(\hat{B})<1$. Thus, L, U and $\hat{B}$ satisfy all the various requirements leading up to and including those of Lemmas 2 and 3. Hence, by (3.13) and the interpretation of the results of Lemma 2 (see immediately below the proof of this lemma), the splitting (3.19) satisfies

$$\rho(\mathcal{L}'_\omega)=\frac{\omega+(2/\hat{v})(1-\omega)}{(2/\hat{v})-\omega}=\frac{\omega\,\hat{v}+2(1-\omega)}{2-\omega\,\hat{v}}$$

in the interval $(0,\omega_b)$, where ω_b is given by (3.15). Moreover,

$$\rho(\mathcal{L}'_{\omega_b})=\sqrt{\omega_b-1}$$

by Lemma 3.

Next, from the construction of $\hat{B}$ and the definition of $\hat{A}$, $\hat{A}$ is hermitian positive definite. Furthermore, because of the definition of L, $L=\frac{1}{2}(I-A+S)$ where $S\equiv 0$ is (obviously) skew-hermitian and $U=L=L^*$. Thus, the splitting (3.19) (also) satisfies the requirements of (1.3) through (1.6). Hence, by observations made in the preceding paragraph, the upper bounds (II) and (III) are sharp. ■

References

1. Apostolatos, N., Kulisch, U.: Über die Konvergenz des Relaxationsverfahrens bei nicht-negativen und diagonaldominanten Matrizen. Computing **2**, 17–24 (1967)
2. Berman, A., Plemmons, R.J.: Nonnegative matrices in the mathematical sciences. New York: Academic Press 1979
3. de Pillis, J.: Gauss-Seidel convergence for operators on Hilbert space. SIAM J. Numer. Anal. **10**, 112–122 (1973)
4. Forthyse, G.E., Wasow, W.R.: Finite difference methods for partial differential equations. New York: Wiley 1960
5. Hanna, M.: Generalized overrelaxation and Gauss-Seidel convergence on Hilbert space. Proc. Amer. Math. Soc., **35**, pp. 524–530 (1972)
6. Kahan, W.: Gauss-Seidel methods for solving large systems of linear equations. Doctoral Thesis, University of Toronto, 1958
7. Kulisch, U.: Über reguläre Zerlegungen von Matrizen und einige Anwendungen. Numer. Math. **11**, 444–449 (1968)
8. Nichols, N.K., Fox, L.: Generalized consistent ordering and optimum successive overrelaxation factor. Numer. Math. **13**, 425–433 (1969)

9. Varga, R.S.: *p*-cyclic matrices, a generalization of the Young-Frankel successive overrelaxation scheme. Pacific J. Math. **9**, 617–628 (1959)
10. Varga, R.S.: Matrix iterative analysis, Englewood Cliffs, NJ: Prentice Hall 1962
11. Varga, R.S.: Extensions of the successive overrelaxation theory with applications to finite element approximation. In: Topics in numerical analysis (J.J.H. Miller, ed.), pp. 329–343. New York: Academic Press 1973
12. Varga, R.S.: On recurring theorems on diagonal dominance, Linear Algebra and Appl. **13**, 1–9 (1976)
13. Wachspress, E.L.: Iterative solution of elliptic systems, Englewood Cliffs, NJ: Prentice Hall, 1966
14. Young, D.M.: Iterative solution of large linear system, New York: Academic Press, 1971

Received September 11, 1979