SIAM J. MATRIX ANAL. APPL.
Vol. 24, No. 1, pp. 68–77

COMPARISON OF CONVERGENCE OF GENERAL STATIONARY ITERATIVE METHODS FOR SINGULAR MATRICES*

IVO MAREK† AND DANIEL B. SZYLD‡

Abstract. New comparison theorems are presented comparing the asymptotic convergence factor of iterative methods for the solution of consistent (as well as inconsistent) singular systems of linear equations. The asymptotic convergence factor of the iteration matrix T is the quantity $\gamma(T) = \max\{|\lambda|, \lambda \in \sigma(T), \lambda \neq 1\}$, where $\sigma(T)$ is the spectrum of T. In the new theorems, no restrictions are imposed on the projections associated with the two iteration matrices being compared. The splittings of the well-known example of Kaufman [*SIAM J. Sci. Statist. Comput.*, 4 (1983), pp. 525–552] satisfy the hypotheses of the new theorems.

Key words. linear systems, iterative methods, comparison theorems, convergence factor, Markov processes, Markov chains, stochastic matrices

AMS subject classifications. 65F10, 15A48, 15A06, 15A51

PII. S0895479800375989

1. Introduction.

In this paper we study certain properties of iterative methods for the solution of $n \times n$ consistent (as well as inconsistent) singular linear systems of equations of the form

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

One case important in applications is when

$$A = I - B, \quad B^T e = e, \quad e^T = [1, 1, \ldots, 1], \tag{1.2}$$

B is the stochastic matrix representing a Markov chain, and the solution of (1.1), for $b = 0$, is the stationary probability distribution of the Markov chain (normalized so that $x^T e = 1$); see, e.g., [3], [25]. In this case, $\rho(B) = 1$, where $\rho(B)$ denotes the spectral radius of B.

Iterative methods for the solution of (1.1) based on splittings of the form $A = M - N$, where M is nonsingular, have been successfully used for this problem; see, e.g., [1], [2], [8], [10], [14], [21]. These methods include point and block versions of the classical Jacobi, Gauss–Seidel, and SOR methods [3], [25], [29] and can be written as the following iteration, starting from an initial vector $x_{(0)}$:

$$x_{(k+1)} = Tx_{(k)} + c, \quad c = M^{-1}b. \tag{1.3}$$

The matrix $T = M^{-1}N$ is called the iteration matrix, and it is generally assumed to be nonnegative (denoted $T \geq O$), e.g., when the splittings are weak regular [3], i.e., $M^{-1} \geq O$ and $M^{-1}N \geq O$. A regular splitting is such that $M^{-1} \geq O$ and $N \geq O$

*Received by the editors August 1, 2000; accepted for publication (in revised form) by M. Eiermann January 17, 2002; published electronically June 12, 2002.
http://www.siam.org/journals/simax/24-1/37598.html

†Czech Institute of Technology, School of Civil Engineering, Thakurova 7, 16000 Praha 1, Czech Republic (marek@ms.mff.cuni.cz). This work was supported by the Grant Agency of the Czech Republic, grant 201/02/595, and by grant CEZ J04:210000010.

‡Department of Mathematics, Temple University (038-16), 1805 N. Broad Street, Philadelphia, PA 19122-6094 (szyld@math.temple.edu). This work was supported by National Science Foundation grant DMS-9973219.

[29]. A weak splitting is such that $M^{-1}N \geq O$ [13] (some authors call these splittings nonnegative splittings; see, e.g., [6], [31]). Since $A = M(I - T)$ it follows that A singular implies that 1 is an eigenvalue of T, and $\rho(T) = 1$ is implied in the case of stochastic matrices such as in the case of Markov chains. It also follows that the null space of A, $\mathcal{N}(A)$, coincides with $\mathcal{N}(I - T)$, the null space of $I - T$.

The rate of convergence of these iterative methods is governed by the quantity $\gamma(T) = \max\{|\lambda|, \lambda \in \sigma(T), \lambda \neq 1\}$, where $\sigma(T)$ is the spectrum of T. When $\gamma(T) = 1$ convergence is not guaranteed. When $\gamma(T) < 1$ and $ind(I - T) = 1$, there is convergence; see, e.g., [3] and section 2. We call the quantity $\gamma(T)$ the *asymptotic convergence factor* of the iterative method (1.3).

In the case of nonsingular A, the quantity governing the rate of convergence of the iterative methods is $\rho(T)$. The Perron–Frobenius theory provides the first comparison theorem for two iteration matrices; see, e.g., [3], [29].

THEOREM 1.1. *Let* $0 \leq T_1 \leq T_2$; *then* $\rho(T_1) \leq \rho(T_2)$.

There exists a rich literature comparing two splittings of the same matrix; see, e.g., [6], [7], [9], [12], [13], [18], [30], [31]. The following result goes back forty years to Varga [29].

THEOREM 1.2. *Let A be a nonsingular matrix with $A^{-1} \geq O$ and let $A = M_1 - N_1 = M_2 - N_2$ be two regular splittings. If*

$$N_1 \leq N_2, \tag{1.4}$$

then $\rho(M_1^{-1}N_1) \leq \rho(M_2^{-1}N_2) < 1$.

The relation (1.4) means that $N_2 - N_1 \geq O$, i.e., that $(N_2 - N_1)x \geq 0$ whenever $x \geq 0$; in other words, if $\mathcal{K} = \mathbb{R}^n_+$, the nonnegative orthant, $(N_2 - N_1)\mathcal{K} \subset \mathcal{K}$. Woźnicki [30] was the first to prove that the hypothesis (1.4) can be replaced with

$$M_1^{-1} \geq M_2^{-1}; \tag{1.5}$$

see also [7], [31]. Condition (1.4) implies (1.5); see, e.g., [7], [15].

Comparison results such as Theorems 1.1 and 1.2 and their variants have been extended to nonnegative operators over Banach spaces, using partial orders defined by general cones $\mathcal{K}$ generating the appropriate Banach space; see, e.g., [6], [12], [22], [24], [27], [28]. See the appendix for the definition of a generating cone. The concept of nonnegativity carries over to any cone $\mathcal{K}$: $x \succeq O$ if $x \in \mathcal{K}$, and $T \succeq O$ if $T\mathcal{K} \subset \mathcal{K}$. The concepts of weak regular, regular splitting, etc., with respect to the cone $\mathcal{K}$ are based on this concept of $\mathcal{K}$-nonnegativity; see the mentioned references and [15].

When A is singular, several authors have provided examples where (1.4) holds, while $\gamma(M_1^{-1}N_1) \not\leq \gamma(M_2^{-1}N_2)$; see [4], [10]. The following example is due to Kaufman [10].

EXAMPLE 1.3. *Consider the matrix*

$$A = \begin{bmatrix} 1 & -1/2 & -1/2 & 0 \\ -1/2 & 1 & 0 & -1/2 \\ -1/2 & 0 & 1 & -1/2 \\ 0 & -1/2 & -1/2 & 1 \end{bmatrix}$$

and the two regular splittings $A = M_1 - N_1 = M_2 - N_2$ *defined by*

$$N_1 = \begin{bmatrix} 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad N_2 = \begin{bmatrix} 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then $N_1 \leq N_2$, *but* $\gamma(M_1^{-1}N_1) = 1/9 > \gamma(M_2^{-1}N_2) = 0$.

In [15] we showed that conditions of the form (1.4) or (1.5) would imply the relation $\gamma(M_1^{-1}N_1) \leq \gamma(M_2^{-1}N_2)$ if these conditions are interpreted using a specific partial order, which is different than the usual partial order defined by the nonnegative orthant $\mathcal{K} = \mathbb{R}^n_+$. The new partial order is derived from the projection matrix associated with the iteration matrix, as described in the next section. In [15], our results required that both iteration matrices $T_1 = M_1^{-1}N_1$ and $T_2 = M_2^{-1}N_2$ be associated with the *same* projection (onto $\mathcal{N}(A)$). The splittings of Example 1.3 do not have this property; see Example 2.3 below.

In section 3 we present new comparison results without the requirement that the two projections be the same. In particular, unlike the results in [15], no restriction is imposed on $\dim \mathcal{N}(A)$. In other words, the new theorems can be applied to a much more general collection of splittings of A. In particular, our theorems apply to Example 1.3.

In these theorems we implicitly assume that $\gamma(T) \in \sigma(T)$. In section 4 we extend our theory to some splittings where this assumption is not needed.

2. The partial order.

A matrix $T \in \mathbb{R}^{n\times n}$ is called *convergent* if $\lim_{k\to\infty} T^k$ exists. A splitting $A = M - N$ is called convergent if its iteration matrix $T = M^{-1}N$ is convergent. In this paper we consider the case where $\rho(T) = 1$. The following result indicates an equivalent definition of convergence; see [3, Lemma 7.6.9], [14] [15], [17]. For other equivalent conditions, see, e.g., [16], [19], [20], [26].

THEOREM 2.1. *Let* $T \in \mathbb{R}^{n\times n}$. *T is convergent if and only if*

$$T = P + Z, \text{ where } P^2 = P, \ PZ = ZP = O, \tag{2.1}$$

and $\rho(Z) < 1$. *Moreover, P is a projection onto* $\mathcal{N}(I - T)$.

It follows from Theorem 2.1 that $\lim_{k\to\infty} T^k = P$. In the case studied in this paper, i.e., when $A = M - N$ and $T = M^{-1}N$, the matrix P is a projection onto $\mathcal{N}(A)$. As is well known, an expression for this projection is $P = I - (I - T)^{\#}(I - T)$, where the notation $Q^{\#}$ stands for the (unique) group inverse of Q; see, e.g., [5], [16]. Thus, $I - P = (I - T)^{\#}(I - T)$.

REMARK 2.2. *If* $T \geq O$ *is irreducible, the Perron–Frobenius theorem implies that* $\dim \mathcal{N}(I - T) = \dim \mathcal{N}(A) = 1$. *In this case, any projection onto* $\mathcal{N}(A)$ *necessarily has the form*

$$P = \hat{x}\hat{z}^T, \quad \text{with} \quad \hat{z}^T\hat{x} = 1, \tag{2.2}$$

where $\hat{x} \in \mathcal{N}(A)$ *and* $\hat{z}$ *is some vector in* $\mathbb{R}^n$.

EXAMPLE 2.3. *Consider the matrix* $A = I - B$ *and the two splittings of Example* 1.3. *Let* $T_0 = B$, $T_i = M_i^{-1}N_i$, $i = 1, 2$. *We have* $\hat{x} = e \in \mathcal{N}(A)$, *e as in* (1.2). *Let* $T_i = P_i + Z_i$, *satisfying* (2.1), $i = 0, 1, 2$. *We obtain* $P_i = \hat{x}\hat{z}_i^T$, $i = 0, 1, 2$, *where* $\hat{z}_0^T = [1/4, 1/4, 1/4, 1/4]$, $\hat{z}_1^T = [0, 0, 1/2, 1/2]$, *and* $\hat{z}_2^T = [0, 1/4, 1/4, 1/2]$. *Note, however, that* $\rho(Z_0) = 1$ *and* T_0 *is not convergent.*

It follows from Example 2.3 that the iteration matrices obtained from different splittings of the same matrix A may have associated with them totally different projections P_i onto the same subspace $\mathcal{N}(A) = \mathcal{R}(P_i)$.

Given a convergent matrix $T_i = P_i + Z_i$ satisfying (2.1), the cone which we use for our comparison is a (pointed) cone generating the range of the projection $I - P_i = (I - T_i)^{\#}(I - T_i)$. In other words, we will use $\mathcal{K}_i$ such that for every element $u \in \mathcal{R}(I - P_i)$, there are v, $w \in \mathcal{K}_i$ (usually not unique) such that $u = v - w$, i.e.,

$\mathcal{K}_i - \mathcal{K}_i = \mathcal{R}(I - P_i)$. (We review the definition of a generating cone in the appendix.) Note that we always have $(I - P_i)\mathcal{K}_i = \mathcal{K}_i$, and furthermore $I - P_i$ is the identity operator on $\mathcal{K}_i$ and on $\mathcal{R}(I - P_i)$.

REMARK 2.4. *In the important and practical case of* $\dim \mathcal{N}(A) = 1$, *e.g., when* B *in* (1.2) *is irreducible, it was pointed out in* [15] *that we can compute* $\mathcal{R}(I - P_i)$ *even if we do not know* P_i. *This follows since from* (2.2), $P_i^T \hat{z} = \hat{z}$, *i.e., that* $(I - P_i)^T \hat{z} = 0$. *Then we can characterize* $\mathcal{R}(I - P_i)$ *as*

$$\mathcal{R}(I - P_i) = \{x \in \mathbb{R}^n \ : \ x^T \hat{z} = 0\}. \tag{2.3}$$

We can then choose

$$\mathcal{K}_i = \left\{ x \in \mathbb{R}^n \ : \ x = \sum_{k=1}^{n-1} \alpha_k v_k, \ \alpha_k \geq 0, k = 1, \ldots, n-1 \right\}, \tag{2.4}$$

where the $n - 1$ *vectors* $v_k \in \mathcal{R}(I - P_i)$ *(i.e.,* $v_k^T \hat{z} = 0$*) are linearly independent; cf.* (A.1).

Let $\mathcal{K}_i - \mathcal{K}_i = \mathcal{R}(I - P_i)$. By definition, the cone $\mathcal{K}_i$ generates a proper subspace, i.e., not the whole space. Therefore, to define a partial order on $\mathbb{R}^n$ using $\mathcal{K}_i$, these vectors and the matrices operating on them need to be restricted to the subspace $\mathcal{R}(I - P_i)$. Thus, we say that $x \overset{i}{\geq} y$, $x,\ y \in \mathbb{R}^n$, if $(I - P_i)(x - y) \in \mathcal{K}_i$. Similarly, a matrix $T \in \mathbb{R}^{n \times n}$ is said to be $\mathcal{K}_i$-nonnegative, denoted $T \overset{i}{\geq} O$ if $(I - P_i)Tx \in \mathcal{K}_i$ for all $x \in \mathcal{K}_i$. Similarly, a splitting $A = M - N$ is called $\mathcal{K}_i$-weak, $\mathcal{K}_i$-weak regular, or $\mathcal{K}_i$-regular if $M^{-1}N \overset{i}{\geq} O$, $M^{-1} \overset{i}{\geq} O$, and $M^{-1}N \overset{i}{\geq} O$, or $M^{-1} \overset{i}{\geq} O$ and $N \overset{i}{\geq} O$, respectively; see examples and further discussion in [15].

3. Comparison theorems. We begin with the observation that if one has two projections P_i and P_j onto the *same* subspace S, then

$$P_j P_i = P_i \ \text{ and consequently } \ (I - P_j)(I - P_i) = I - P_j \tag{3.1}$$

since for two projections P_i and P_j, there obviously holds $P_j P_i = P_i$ if and only if $\mathcal{R}(P_i) \subseteq \mathcal{R}(P_j)$.

In the particular case where S is one-dimensional and the two projections have the form (2.2), the identity (3.1) can be computed directly.

We are ready now to show an important tool for our comparisons.

LEMMA 3.1. *Let* A *be a singular matrix. Let* $A = M_1 - N_1 = M_2 - N_2$ *be two convergent splittings, and let* $T_i = M_i^{-1} N_i = P_i + Z_i$, $P_i^2 = P_i$, $P_i Z_i = Z_i P_i = O$, $\rho(Z_i) < 1$, $i = 1, 2$. *Then* $\sigma((I - P_i)Z_j) = \sigma(Z_j)$.

Proof. If $i = j$ there is nothing to prove. Thus, we assume $i \neq j$. Let $\lambda \in \sigma(Z_j)$ and x such that $Z_j x = \lambda x$. Since $Z_j(I - P_j) = Z_j$, we have, using (3.1), that

$$Z_j = Z_j(I - P_j)(I - P_i) = Z_j(I - P_i)$$

and therefore $Z_j(I - P_j)(I - P_i)x = \lambda x$. Consequently,

$$(I - P_i)Z_j x = (I - P_i)Z_j[(I - P_i)x] = \lambda(I - P_i)x,$$

and thus $\lambda \in \sigma((I - P_i)Z_j)$.

Conversely, let $\lambda \in \sigma((I - P_i)Z_j)$ and let v such that $(I - P_i)Z_j v = \lambda v$. Multiply the last equation by $(I - P_j)$ and, using (3.1), we have

$$(I - P_j)(I - P_i)Z_j v = (I - P_j)Z_j v = Z_j v = Z_j[(I - P_j)v] = \lambda(I - P_j)v$$

and thus $\lambda \in \sigma(Z_j)$. □

The following result was proved in [13], and the nonnegativity is with respect to any cone.

LEMMA 3.2. *Let $V \succeq O$, and let $x \succeq 0$, $x \neq 0$, be such that $Vx - \alpha x \succeq 0$. Then $\alpha \leq \rho(V)$.*

We can now proceed with the main result, which generalizes [15, Theorem 5.6] and is the general counterpart to Theorem 1.2 with the hypothesis (1.5).

THEOREM 3.3. *Let A be singular. Let $A = M_1 - N_1 = M_1(I - T_1) = M_2 - N_2 = M_2(I - T_2)$ be two (convergent) $\mathcal{K}_i$-regular splittings, where $\mathcal{K}_i$ is the cone generating $\mathcal{R}(I - P_i)$ for either $i = 1$ or $i = 2$, and $T_j = P_j + Z_j$, $P_j^2 = P_j$, $P_j Z_j = Z_j P_j = O$, $\rho(Z_j) < 1$, $j = 1, 2$. If*

$$M_1^{-1} \overset{i}{\succeq} M_2^{-1}, \tag{3.2}$$

then $\gamma(T_1) \leq \gamma(T_2)$.

Proof. We assume first that $i = 1$. If $\gamma(T_1) = 0$, there is nothing to prove, so we assume $\gamma(T_1) \neq 0$. Since $\mathcal{K}_1$ is the cone generating $\mathcal{R}(I - P_1)$, and by hypothesis $Z_1\mathcal{K}_1 = T_1\mathcal{K}_1 \subset \mathcal{K}_1$, there is a Perron eigenvector $x = (I - P_1)x \in \mathcal{K}_1$ for which $T_1 x = Z_1 x = \rho(Z_1)x = \gamma(T_1)x \succeq 0$. Here and in the rest of the proof we use the symbol $\succeq$ to indicate $\overset{1}{\succeq}$, since there is no possibility of confusion. Then

$$M_1 x = \frac{1}{\gamma(T_1)} N_1 x \succeq 0 \tag{3.3}$$

and

$$Ax = M_1(I - T_1)x = \frac{1 - \gamma(T_1)}{\gamma(T_1)} N_1 x \succeq 0.$$

Using (3.2), it follows that

$$(M_1^{-1} - M_2^{-1})Ax = (I - T_1)x - (I - T_2)x = T_2 x - \gamma(T_1)x \succeq 0. \tag{3.4}$$

Premultiply the last equation by $(I-P_1)$ which is not only $\mathcal{K}_1$-nonnegative but actually the identity on $\mathcal{K}_1$, and observe that because of (3.1), $(I - P_1)T_2 = (I - P_1)Z_2$. Thus, we have that

$$(I - P_1)Z_2 x \succeq \gamma(T_1)x,$$

which implies by Lemma 3.2 that $\rho((I - P_1)Z_2) \geq \gamma(T_1)$. Using Lemma 3.1, we can rewrite this as $\gamma(T_2) = \rho(Z_2) \geq \gamma(T_1)$, completing the proof for $i = 1$.

The proof for $i = 2$ is similar, using the eigenvector x of T_2, except that we need to require the additional hypothesis that x is in the interior of $\mathcal{K}_2$, so we can use [13, Lemma 3.3]. □

REMARK 3.4. *We point out that this theorem is valid with weaker hypotheses, using the same proof, namely, that the splittings be $\mathcal{K}_i$-weak splittings and convergent (or $\mathcal{K}_i$-weak regular splittings) and that if the Perron eigenvector x of Z_1 satisfies $N_1 x \overset{1}{\succeq} 0$. Alternatively the Perron eigenvector x of Z_2 (in the interior of $\mathcal{K}_2$) needs to satisfy $N_2 x \overset{2}{\succeq} 0$. We also remark that, as it can be seen from the hypotheses and the proof, no restriction on $\dim \mathcal{N}(A)$ is needed.*

The following result was shown in [15]; see also [7] or [31] for the nonsingular case.

Lemma 3.5. *Let $A = M_1 - N_1 = M_2 - N_2$ be two $\mathcal{K}_i$-weak regular splittings, where $\mathcal{K}_i$ is a cone generating $\mathcal{R}(I - P_i)$, for either $i = 1$ or $i = 2$, and $T_j = P_j + Z_j$, $P_j^2 = P_j$, $P_jZ_j = Z_jP_j = O$, $\rho(Z_j) < 1$, $j = 1, 2$. If $N_2 \overset{i}{\succeq} N_1$, then $M_1^{-1} \overset{i}{\succeq} M_2^{-1}$.*

We can write the counterpart to Theorem 1.2. The proof follows directly from Lemma 3.5 and Theorem 3.3.

Theorem 3.6. *Let A be singular. Let $A = M_1 - N_1 = M_1(I - T_1) = M_2 - N_2 = M_2(I - T_2)$ be two (convergent) $\mathcal{K}_i$-regular splittings, where $\mathcal{K}_i$ is the cone generating $\mathcal{R}(I - P_i)$, for either $i = 1$ or $i = 2$, and $T_j = P_j + Z_j$, $P_j^2 = P_j$, $P_jZ_j = Z_jP_j = O$, $\rho(Z_j) < 1$, $j = 1, 2$. If $N_2 \overset{i}{\succeq} N_1$, then $\gamma(T_1) \le \gamma(T_2)$.*

Again, this theorem is valid with weaker hypotheses; see Remark 3.4.

Example 3.7. *Consider the matrix A and the splittings of Example* 1.3. *The projections P_1 and P_2 are shown in Example* 2.3. *A simple computation gives the matrix*

$$(I - P_1)(N_1 - N_2) = \begin{bmatrix} 0 & -1/2 & 0 & 1/4 \\ 0 & 0 & 0 & 1/4 \\ 0 & 0 & 0 & -1/4 \\ 0 & 0 & 0 & 1/4 \end{bmatrix},$$

which is nonnegative with respect to the following cone generating $\mathcal{R}(I - P_1)$:

$$\mathcal{K}_1 = \left\{ \sum_{k=1}^{3} \alpha_k v_k, \alpha_k \ge 0, v_1^T = [-1, 0, 0, 0], v_2^T = [1, 1, -1, 1], v_3^T = [0, 1, 0, 0] \right\}.$$

Indeed, $(I - P_1)(N_1 - N_2)v_1 = 0$, $(I - P_1)(N_1 - N_2)v_2 = \frac{1}{2}v_1 + \frac{1}{4}v_2$, and $(I - P_1)(N_1 - N_2)v_3 = \frac{1}{2}v_1$. Furthermore, consider the matrix

$$(I - P_2)(N_1 - N_2) = \begin{bmatrix} 0 & -1/2 & 0 & 1/8 \\ 0 & 0 & 0 & 1/8 \\ 0 & 0 & 0 & -3/8 \\ 0 & 0 & 0 & 1/8 \end{bmatrix}.$$

This matrix is nonnegative with respect to the following cone generating $\mathcal{R}(I - P_2)$:

$$\mathcal{K}_2 = \left\{ \sum_{k=1}^{3} \alpha_k w_k, \alpha_k \ge 0, w_1^T = [-1, 0, 0, 0], w_2^T = [-2, 2, -2, 0], w_3^T = [1, 1, -3, 1] \right\}.$$

Indeed, $(I - P_2)(N_1 - N_2)w_1 = 0$, $(I - P_2)(N_1 - N_2)w_2 = w_1$, and $(I - P_2)(N_1 - N_2)w_3 = \frac{1}{2}w_1 + \frac{1}{8}w_3$.

We have shown in [15] examples when two matrices cannot be compared in the usual partial order but are comparable with the appropriate choice of generating cone. Example 3.7 indicates that even in the case when two matrices are comparable in the usual partial order, the direction of the comparison can be reversed with the appropriate cone, and thus the comparison of the asymptotic convergence factors can be obtained.

We note that in the special case when $P_1 = P_2$, Theorems 3.3 and 3.6 reduce to the comparison theorems in [15], but these do not apply to Example 1.3.

We now present the counterpart to Theorem 1.1 in the singular case.

Theorem 3.8. *Let A be singular. Let $A = M_1 - N_1 = M_1(I - T_1) = M_2 - N_2 = M_2(I - T_2)$ be two convergent $\mathcal{K}_i$-weak splittings, where $\mathcal{K}_i$ is the cone generating*

$\mathcal{R}(I-P_i)$ for either $i=1$ or $i=2$, and $T_j = P_j + Z_j$, $P_j^2 = P_j$, $P_jZ_j = Z_jP_j = O$, $\rho(Z_j) < 1$, $j = 1, 2$. If

$$T_2 \overset{i}{\geq} T_1, \tag{3.5}$$

then $\gamma(T_1) \leq \gamma(T_2)$.

Proof. We assume that $i = 1$. The proof for the case $i = 2$ is analogous. Premultiply (3.5) by $(I - P_1)$, the identity in $\mathcal{K}_1$, and using (3.1), we obtain

$$(I - P_1)Z_2 \overset{1}{\geq} Z_1 \overset{1}{\geq} O.$$

We now apply the Perron–Frobenius theorem in the subspace $\mathcal{R}(I - P_i)$ (see, e.g., [12], [22]) and obtain $\rho((I - P_1)Z_2) \geq \rho(Z_1)$. By Lemma 3.1 we then have

$$\gamma(T_2) = \rho(Z_2) = \rho((I - P_1)Z_2) \geq \rho(Z_1) = \gamma(T_1). \qquad \square$$

EXAMPLE 3.9. *Consider the matrix A and the splittings of Example 1.3. The projections P_1 and P_2 are shown in Example 2.3. One can directly compute the matrix*

$$(I - P_1)(T_1 - T_2) = \begin{bmatrix} 0 & -1/4 & -1/12 & 1/3 \\ 0 & 0 & 1/6 & -1/6 \\ 0 & 0 & 1/18 & -1/18 \\ 0 & 0 & -1/18 & 1/18 \end{bmatrix},$$

which is nonnegative with respect to the following cone generating $\mathcal{R}(I - P_1)$:

$$\mathcal{K}_1 = \left\{ \sum_{k=1}^{3} \alpha_k v_k, \alpha_k \geq 0, v_1^T = [0, -1, 0, 0], v_2^T = [1, 0, 0, 0], v_3^T = \left[-\frac{15}{4}, 0, 1, -1 \right] \right\}.$$

Indeed, $(I-P_1)(T_1-T_2)v_1 = \frac{1}{4}v_2$, $(I-P_1)(T_1-T_2)v_2 = 0$, and $(I-P_1)(T_1-T_2)v_3 = \frac{1}{3}v_2 + \frac{1}{9}v_3$. Furthermore, the matrix

$$(I - P_2)(T_1 - T_2) = \begin{bmatrix} 0 & -1/4 & -1/36 & 5/18 \\ 0 & 0 & -1/9 & 1/9 \\ 0 & 0 & 1/9 & -1/9 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is nonnegative with respect to the following cone generating $\mathcal{R}(I - P_2)$:

$$\mathcal{K}_2 = \left\{ \sum_{k=1}^{3} \alpha_k w_k, \alpha_k \geq 0, w_1^T = [-1, 0, 0, 0], w_2^T = [-1, 3, -1, -1], w_3^T = [-1, 1, -1, 0] \right\}.$$

Indeed, $(I-P_2)(T_1-T_2)w_1 = 0$, $(I-P_2)(T_1-T_2)w_2 = w_1$, and $(I-P_2)(T_1-T_2)w_3 = \frac{1}{9}w_1 + \frac{1}{3}w_3$.

4. Majorizing splittings. We conclude with some observations which enlarge the class of splittings for which we can compare the asymptotic convergence factors. In Theorems 3.3, 3.6, and 3.8, we assume that the splittings are convergent $\mathcal{K}_i$-weak, and thus, we are implicitly assuming that the asymptotic convergence factor belongs to the spectrum, i.e., that $\gamma(T_i) \in \sigma(T_i)$, $T_i = M_i^{-1}N_i$, $A = M_i - N_i$. We can capture some of the cases where the splittings are such that $\gamma(T_i) \notin \sigma(T_i)$ by the following construction.

DEFINITION 4.1. *Given a cone $\mathcal{K}$ and its induced partial order $\succeq$, one can define the absolute value of a matrix Z by $|Z| = Z^+ + Z^-$, where*

$$Z = Z^+ - Z^-, \quad \text{with } \ Z^+ \succeq O, \ Z^- \succeq O. \tag{4.1}$$

This definition of absolute value of an operator with respect to a partial order can be seen as a slight generalization of that defined in [23] in the case of a vector lattice space (Riesz space). Here, we do not need a vector lattice order but need only that the matrix Z be regular in the sense of [23, Definition 1.1]. The decomposition (4.1) is then possible (although not necessarily in a unique manner).

DEFINITION 4.2. *Let A be singular. Let $A = M_1 - N_1 = M_1(I - T_1) = M_2 - N_2 = M_2(I - T_2)$ be two splittings. Let $T_j = P_j + Z_j$, $P_j^2 = P_j$, $P_jZ_j = Z_jP_j = O$ for $j = 1, 2$. Let $A = M_2 - N_2$ be a $\mathcal{K}_2$-weak splitting, where $\mathcal{K}_2$ is a cone generating $\mathcal{R}(I - P_2)$. We say that the splitting $A = M_1 - N_1$ is majorized by the splitting $A = M_2 - N_2$ when $|Z_1| \overset{2}{\precsim} Z_2$. In a similar manner one defines a minorized splitting.*

REMARK 4.3. *Majorizing splittings were introduced in* [12, *section* 7] *for splittings of a nonsingular operator A and, in particular, were applied to SOR splittings. Many of the results from* [12] *using majorizing splittings can be easily extended to the singular case. Note that a basic hypothesis for deriving the results in* [12] *is the normality of the cones under consideration; see the appendix for definitions and comments.*

We are now ready to present a comparison result between a splitting for which the asymptotic convergence factor is not in the spectrum of the iteration matrix and another splitting for which it is.

THEOREM 4.4. *Let $A = M_1 - N_1 = M_1(I - T_1) = M_2 - N_2 = M_2(I - T_2)$ be two splittings. Let $T_j = P_j + Z_j$, $P_j^2 = P_j$, $P_jZ_j = Z_jP_j = O$ for $j = 1, 2$. Let $A = M_2 - N_2$ be a (convergent) $\mathcal{K}_2$-weak splitting, where $\mathcal{K}_2$ is a cone generating $\mathcal{R}(I - P_2)$, with $\rho(Z_2) < 1$. Assume that the splitting $A = M_2 - N_2$ majorizes the splitting $A = M_1 - N_1$. Then*

$$\gamma(T_1) = \rho(Z_1) \le \rho(|Z_1|) \le \rho(Z_2) = \gamma(T_2).$$

Proof. Relations

$$-|Z_1|y \ \overset{2}{\precsim} \ Z_1y \ \overset{2}{\precsim} \ |Z_1|y,$$

valid for any $y \in \mathcal{K}_2$, imply that $\rho(Z_1) \le \rho(|Z_1|)$. Further, by hypothesis, we have

$$0 \ \overset{2}{\precsim} \ |Z_1|y \ \overset{2}{\precsim} \ Z_2y \ \text{ for all } y \in \mathcal{K}_2,$$

and consequently $\rho(|Z_1|) \le \rho(Z_2)$. ☐

REMARK 4.5. *The fact that T_1 is convergent is a consequence of the hypothesis that Z_2 is convergent. Therefore, Z_1 need not be assumed to be convergent.*

5. Concluding remarks. We have demonstrated that the usual partial order ($\ge$) defined by the nonnegative orthant $\mathbb{R}^n_+$ is not the appropriate choice of order when comparing splittings of singular matrices.

We have provided two different partial orders with which the comparison of the splittings implies the comparison of the asymptotic convergence factors of the corresponding iteration matrices.

Example 1.3, due to Kaufman [10], was originally presented as a counterexample to possible theorems of the form of Theorem 1.2. It now becomes a good example to show that the alternative partial orders are the appropriate ones to use in the context of singular matrices.

Appendix.

DEFINITION A.1. *Let $\mathcal{E}$ be a real Banach space. A normal cone $\mathcal{K}$ is a subset of $\mathcal{E}$ with the following properties:*

(i) $\mathcal{K}+\mathcal{K}\subset\mathcal{K}$,

(ii) $\alpha\mathcal{K}\subset\mathcal{K}$ *for* $\alpha\geq 0$,

(iii) $\mathcal{K}\cap(-\mathcal{K})=\{0\}$, *i.e., it is pointed,*

(iv) $\bar{\mathcal{K}}=\mathcal{K}$, *where* $\bar{\mathcal{K}}$ *denotes the norm-closure of* $\mathcal{K}$, *and*

(v) $\exists\ \sigma>0$ *such that for* $x,\ y\ \in\mathcal{K}$ *one has* $\|x+y\|\geq\sigma\|x\|$.

We say that $\mathcal{K}$ is generating if $\mathcal{E}=\mathcal{K}-\mathcal{K}$. The typical example is $\mathcal{E}=\mathbb{R}^n$, and a generating cone is the standard cone

$$\begin{aligned}\mathcal{K}=\mathbb{R}^n_+&=\{x\in\mathbb{R}^n\ :\ x\geq 0\}\\ &=\left\{x\in\mathbb{R}^n\ :\ x=\sum_{k=1}^{n}\alpha_k e_k,\ \alpha_k\geq 0, k=1,\ldots,n\right\},\end{aligned}\tag{A.1}$$

where e_k is the standard kth canonical vector, i.e., the kth column of the identity.

We should remark that condition (v) is simply saying that the norm $\|.\|$ of the Banach space $\mathcal{E}$ is $\mathcal{K}$-semimonotone (and $\mathcal{K}$-monotone if it holds with $\sigma=1$). The following result, which can be found, e.g., in [11], indicates when it holds.

PROPOSITION A.2. *Assume $\mathcal{E}$ is a Banach space over the field of reals with the norm $\|.\|_{\mathcal{E}}$. A cone $\mathcal{K}\subset\mathcal{E}$ satisfying* (i)–(iv) *is normal, i.e., it fulfills* (v), *if and only if the norm on $\mathcal{E}$ $\|.\|_*$ defined by*

$$\|x\|_*=\mathrm{Max}\left(\inf\{\|u\|_{\mathcal{E}}:u\in\mathcal{E},(u-x)\in\mathcal{K}\},\sup\{\|v\|_{\mathcal{E}}:v\in\mathcal{E},(x-v)\in\mathcal{K}\}\right),\ x\in\mathcal{K},$$

is equivalent with $\|.\|_{\mathcal{E}}$.

As a consequence of Proposition A.2 we conclude that any closed cone in $\mathbb{R}^n$, i.e., any set satisfying (i)–(iv) of Definition A.1, is normal, since all the norms on a finite dimensional space are equivalent.

Acknowledgments. We thank Hans Schneider and Michael Eiermann, whose questions and observations led us to several improvements of the manuscript.

REFERENCES

[1] G. P. Barker and R. J. Plemmons, *Convergent iterations for computing stationary distributions of Markov chains,* SIAM J. Algebraic Discrete Methods, 7 (1986), pp. 390–398.

[2] V. A. Barker, *Numerical solution of sparse singular systems of equations arising from ergodic Markov chains,* Comm. Statist. Stochastic Models, 5 (1989), pp. 355–381.

[3] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, 3rd. ed., Academic Press, New York, 1979. Reprinted by SIAM, Philadelphia, 1994.

[4] J. J. Buoni, M. Neumann, and R. S. Varga, *Theorems of Stein–Rosenberg type.* III. *The singular case,* Linear Algebra Appl., 42 (1982), pp. 183–198.

[5] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*, Pitman, London, San Francisco, Melbourne, 1979. Reprinted by Dover, New York, 1991.

[6] J.-J. Climent and C. Perea, *Some comparison theorems for weak nonnegative splittings of bounded operators,* Linear Algebra Appl., 275/276 (1998), pp. 77–106.

[7] G. Csordas and R. S. Varga, *Comparisons of regular splittings of matrices,* Numer. Math., 44 (1984), pp. 23–35.
[8] T. Dayar and W. J. Stewart, *Comparison of partitioning techniques for two-level iterative methods on large, sparse Markov chains,* SIAM J. Sci. Comput., 21 (2000), pp. 1691–1705.
[9] L. Elsner, *Comparisons of weak regular splittings and multisplitting methods,* Numer. Math., 56 (1989), pp. 283–289.
[10] L. Kaufman, *Matrix methods for queuing problems,* SIAM J. Sci. Statist. Comput., 4 (1983), pp. 525–552.
[11] M. A. Krasnoselskii, E. A. Lifshits, and A. V. Sobolev, *Positive Linear Systems,* Nauka, Moscow, 1985 (in Russian); Heldermann-Verlag, Berlin, 1989 (in English).
[12] I. Marek, *Frobenius theory of positive operators: Comparison theorems and applications,* SIAM J. Appl. Math., 19 (1970), pp. 607–628.
[13] I. Marek and D. B. Szyld, *Comparison theorems for weak splittings of bounded operators,* Numer. Math., 58 (1990), pp. 387–397.
[14] I. Marek and D. B. Szyld, *Iterative and semi-iterative methods for computing stationary probability vectors of Markov operators,* Math. Comput., 61 (1993), pp. 719–731.
[15] I. Marek and D. B. Szyld, *Comparison theorems for the convergence factor of iterative methods for singular matrices,* Linear Algebra Appl., 316 (2000), pp. 67–87.
[16] C. D. Meyer and R. J. Plemmons, *Convergent powers of a matrix with applications to iterative methods for singular systems,* SIAM J. Numer. Anal., 14 (1977), pp. 699–705.
[17] V. Migallón, J. Penadés, and D. B. Szyld, *Block two-stage methods for singular systems and Markov chains,* Numer. Linear Algebra Appl., 3 (1996), pp. 413–426.
[18] V. A. Miller and M. Neumann, *A note on comparison theorems for nonnegative matrices,* Numer. Math., 47 (1985), pp. 427–434.
[19] M. Neumann and R. J. Plemmons, *Convergent nonnegative matrices and iterative methods for consistent linear systems,* Numer. Math., 31 (1978), pp. 265–279.
[20] R. Oldenburger, *Infinite powers of matrices and characteristic roots,* Duke Math. J., 6 (1940), pp. 357–361.
[21] D. P. O'Leary, *Iterative methods for finding the stationary vector for Markov chains,* in Linear Algebra, Markov Chains and Queuing Models, C. D. Meyer and R. J. Plemmons, eds., IMA Vol. Math. Appl. 48, Springer, New York, Berlin, 1993, pp. 125–136.
[22] W. C. Rheinboldt and J. S. Vandergraft, *A simple approach to the Perron–Frobenius theory for positive operators on general partially-ordered finite-dimensional linear spaces,* Math. Comput., 27 (1973), pp. 139–145.
[23] H. H. Schaefer, *Banach Lattices and Positive Operators*, Springer, Berlin, Heidelberg, New York, 1974.
[24] H. Schneider, *Positive operators and an inertia theorem,* Numer. Math., 7 (1965), pp. 11–17.
[25] W. J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, Princeton, NJ, 1994.
[26] D. B. Szyld, *Equivalence of convergence conditions for iterative methods for singular equations,* Numer. Linear Algebra Appl., 1 (1994), pp. 151–154.
[27] J. S. Vandergraft, *Spectral properties of matrices which have invariant cones,* SIAM J. Appl. Math., 16 (1968), pp. 1208–1222.
[28] J. S. Vandergraft, *Applications of partial orderings to the study of positive definiteness, monotonicity, and convergence of iterative methods for linear systems,* SIAM J. Numer. Anal., 9 (1972), pp. 97–104.
[29] R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962. 2nd ed., Springer, Berlin, 2000.
[30] Z. I. Woźnicki, *Two-Sweep Iterative Methods for Solving Large Linear Systems and Their Application to the Numerical Solution of Multi-Group Multi-Dimensional Neutron Diffusion Equations*, Ph.D. thesis, Institute of Nuclear Research, Otwock-Świerk, Poland, 1973.
[31] Z. I. Woźnicki, *Nonnegative splitting theory,* Japan J. Indust. Appl. Math., 11 (1994), pp. 289–342.