JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 53, 313–342 (1976)

k-Part Splittings and Operator Parameter Overrelaxation

JOHN DE PILLIS

Department of Mathematics, University of California, Riverside, Calif. 92502, and Istituto Matematico, Viale Morgagni 67/A, I-50134 Firenze, Italy

Submitted by Ky Fan

This paper proceeds in two directions of attack for finding (iteratively) solutions for linear systems on Hilbert space. First, we consider scalar-dependent overrelaxation as a special case of operator-dependent overrelaxations. Secondly, we study "finer" splittings than the conventional two-part splittings and show where, in some cases, these new splittings can either accelerate convergence of approximating sequences derived from two-part splittings or else turn divergent sequences into convergent ones.

1. INTRODUCTION

Given the linear system

$$Ax = y_0, \tag{1.1}$$

where A is an invertible operator on Hilbert space $\mathscr{H}$ and y_0 is fixed in $\mathscr{H}$. To solve for x, we may split A into the two-part sum $A = A_1 + A_2$, where A_1 is an invertible operator on $\mathscr{H}$, and define the sequence of vectors $\{x_n\}$ recursively by

$$A_1 x_{n+1} + A_2 x_n = y_0, \qquad n = 0, 1, 2,.... \tag{1.2}$$

Once we fix the initial vector x_0, the sequence $\{x_n\}$ is uniquely defined (owing to the invertibility of A_1). We observe that if $\{x_n\}$ converges at all in $\mathscr{H}$, its limit is necessarily the solution vector x, for the system (1.1). We note that the sum decomposition (1.2) embraces the classical Gauss–Seidel iterative scheme (A is an $m \times m$ matrix, A_1 is the upper triangular part of A), the successive overrelaxation (SOR) method (A is an $m \times m$ matrix, A_1 equals the lower triangular part of A plus or minus a certain fraction of the diagonal part of A), and the regular splittings of Varga [8, Section 3.6; 9] (A is an $m \times m$ matrix, A_1^{-1} and $-A_2$ are matrices with nonnegative entries). In all cases, convergence obtains for $\{x_n\}$ defined by (1.2) for all initial vectors x_0, if and only if the spectral radius of $B = -A_1^{-1}A_2$ is less than 1.

409/53/2-9

In this paper, we consider k-part splittings

$$A = A_1 + A_2 + \cdots + A_k \tag{1.3}$$

for A given in (1.1), where A_1 is required to be invertible. Accordingly, once we are given $k - 1$ initial vectors $x_0, x_1, \ldots, x_{k-2}$, the sequence $\{x_n\}$ is uniquely defined (owing to the invertibility of A_1) by

$$A_1 x_{n+k-1} + A_2 x_{n+k-2} + \cdots + A_i x_{n+k-i} + \cdots + A_k x_n = y_0 \tag{1.4}$$

for $n = 0, 1, 2, \ldots$.

We find necessary and sufficient conditions for the k-part splitting (1.3) to guarantee convergence of the sequence $\{x_n\}$ defined by (1.4), for all sets of initial vectors $\{x_0, x_1, \ldots, x_{k-2}\}$.

After some preliminary definitions and theorems, Section 2 identifies convergence of a sequence $\{x_n\}$ induced by a k-part splitting, with convergence of a related sequence $\{Z_n\}$ induced by a certain two-part splitting. This straightforward result appears as Proposition 2.4.

Section 3 deals exclusively with four-part splittings

$$A = A_1 + A_2 + A_3 + A_4$$

for hermitian operator A. We assume the coefficient matrices A_1, A_2, A_3, A_4, are constrained in such a way that a certain operator-entried matrix is positive definite. (This does not necessarily imply that A must be positive definite.) Then a test matrix exists whose positive definiteness is equivalent to convergence of $\{x_n\}$ of (1.4), regardless of initial vectors $\{x_0, x_1, x_2\}$ (Theorem 3.2). With further constraints on the coefficient matrices, positive definiteness of our test matrix (hence, convergence of $\{x_n\}$ to the solution vector x for $Ax = y_0$) is equivalent to positive definiteness of A itself (Theorem 3.3).

Section 4, concerning three-part splittings, reveals a necessary restriction on hermitian $A = A_1 + A_2 + A_3$. As in the case of four-part splittings, we assume the coefficient matrices A_1, A_2, A_3 constrained so that a certain operator-entried matrix is positive definite. For these three-part splittings, convergence is equivalent to saying that A is positive definite (Theorem 4.2).

Section 5 introduces the notion of an operator parameter successive overrelaxation (SOR) decomposition (OPSORD?) for three-part (hence, for two-part) splittings of hermitian operator A. We answer the question: Under what conditions will a family of these overrelaxation three-part decompositions yield convergent sequences $\{x_n\}$? Theorem 5.2 is our most general result in this direction. A specialization of Theorem 5.3 is a recent theorem of Donelly [3, Theorem 2.1], which appears here as Theorem 5.4. In Donnelly's paper, he studies positive definite A and certain "periodic schemes," embodied in three-part scalar-parameter overrelaxation decompositions. Donnelly's paper is, in turn, a generalization of certain results of Chazin and Miranker [1].

Section 6 deals with three-part splittings that improve convergence over corresponding two-part splittings, even when A is not hermitian. That is, if $\{x_0, x_1', \ldots, x_n'\}$ arises from the splitting $A = A_1 + A_2'$, and $\{x_0, x_0, x_2, \ldots, x_n, \ldots\}$ arises from the splitting $A = A_1 + A_2 + A_3$ (invertible A_1 is fixed in both splittings), can A_3 be chosen so that $\{x_n\}$ converges faster than $\{x_n'\}$? We answer the question affirmatively in Theorem 6.3, where we show that if the spectrum of $A_1^{-1}A_2'$ lies in the circle $\{z: |z - 1| < 2\}$ and $\{x_n'\}$ diverges, then A_3 may be found so that $\{x_n\}$ converges (Case A). Also, the average reduction factor $\sigma'(m)$ for $\{x_n'\}$, after m iterations, has (generally) $\|(A_1^{-1}A_2')_m\|^{1/m}$ as an upper bound. Theorem 6.3 also shows that if the real part of the spectrum of $A_1^{-1}A_2'$ is nonnegative, then the average reduction factor $\sigma(m)$, for the three-part splitting sequence $\{x_n\}$, has an upper bound, which is about half that for $\sigma(m)$.

Section 7 presents an example illustrating the techniques of Section 6.

2. Preliminaries and Definitions

Our linear system $Ax = y_0$ is defined for bounded linear operator A on Hilbert space $\mathscr{H}$. The algebra of all bounded linear A on $\mathscr{H}$ is denoted $\mathscr{B}(\mathscr{H})$. A^* denotes the adjoint of A as defined by the inner product $\langle\ ,\ \rangle$ on $\mathscr{H}$. In the matrix case, if $A = (a_{ij})$, then $A^* = (\bar{a}_{ji})$, the conjugate transpose of A. Those hermitian A (i.e., $A = A^*$) such that for some $\delta > 0$, $\langle Ax, x\rangle \geqslant \delta$ for all unit vectors $x \in \mathscr{H}$, are called positive definite. This is denoted by $A > 0$. Since $A > 0$ if and only if $A = B^*B$ for some invertible $B \in \mathscr{B}(\mathscr{H})$, $A > 0$ if and only if, for all X invertible in $\mathscr{B}(\mathscr{H})$, $X^*AX > 0$. The operator X^*AX is said to be hermitian conjugate to A as long as X is ivertible in $\mathscr{B}(\mathscr{H})$. $\frac{1}{2}(A + A^*)$ is called the real part of A and is denoted $\operatorname{Re}(A)$. For integer k, $\bigoplus^k \mathscr{H}$ denotes the direct sum of Hilbert space $\mathscr{H}$ with itself k times, with induced inner product defined by $\langle (x_1, \ldots, x_k), (y_1, \ldots, y_k)\rangle = \sum_{i=1}^k \langle x_i, y_i\rangle$ for all $(x_1, \ldots, x_k), (y_1, \ldots, y_k) \in \bigoplus^k \mathscr{H}$.

For convenience, we state those results that we will use later. The first of these was proven by Stein for matrices [7].

THEOREM 2.1 (Stein [7]; see also [2, Theorem 2.1; 5, Theorem 3.1]). *Let $A = A^*$ and B belong to $\mathscr{B}(\mathscr{H})$. Suppose that*

$$T(A) = A - B^*AB > 0.$$

Then $A > 0$ if and only if $\rho(B)$, the spectral radius of B, is less than one.

Since positive definiteness is preserved under hermitian conjugacy, a useful result will be the following.

THEOREM 2.2 (de Pillis [2, Proposition 4.1)]. *Let $A = A^* = A_1 + A_2$, $A, A_1, A_2 \in \mathscr{B}(\mathscr{H})$. Let $B = -A_1^{-1}A_2$. Then $T(A) = A - B^*AB$ is hermitian conjugate to $A_1^* - A_2 = 2\mathrm{Re}(A_1) - A$.*

We conclude this section with the identification between two-part splittings on the direct sum $\mathscr{H} \oplus \mathscr{H} \oplus \cdots \oplus \mathscr{H} = \oplus^{k-1} \mathscr{H}$, k-part splittings on $\mathscr{H}$.

For the k-part splitting $A = A_1 + A_2 + \cdots + A_k$, define the induced linear operator $\mathcal{A}$ on $\oplus^{k-1} \mathscr{H}$ by the matrix

$$\mathcal{A} = \begin{bmatrix} A_1 & B_1 \\ 0 & C \end{bmatrix} + \begin{bmatrix} B_2 & A_k \\ -C & 0 \end{bmatrix}, \tag{2.1}$$

where A_1, A_k are linear operators on $\mathscr{H}$, C is an invertible linear operator on $\oplus^{k-2} \mathscr{H}$, and B_1, B_2 are linear maps sending $\oplus^{k-2} \mathscr{H}$ to $\mathscr{H}$, whose sum is the matrix

$$B_1 + B_2 = [A_2 A_3 \cdots A_{k-1}], \tag{2.2}$$

where $B_1 + B_2$ is the transformation

$$[A_2 A_3 \cdots A_{k-1}]: (x_{n-1}, x_{n-2}, \ldots, x_{n-k+2}) \to \sum_{i=2}^{k-1} A_i x_{n-i+1}$$

for all $(x_{n-1}, x_{n-2}, \ldots, x_{n-k+2}) \in \oplus^{k-2} \mathscr{H}$.

Remark. It is important to note that (2.1) yields noncorresponding partitioning. By way of illustration, let A be an $n \times n$ matrix (so that the dimension of $\mathscr{H}$ is n). The induced $\mathcal{A}$ of (2.1) acts on the $(k-1)\cdot(n)$-dimensional vector space $\oplus^{k-1} \mathscr{H}$. But note that A_1 and A_k are each $n \times n$ matrices, while B_1 and B_2 are both $n \times (k-2)\cdot n$ matrices. Accordingly, (2.2) is that $n \times (k-2)\cdot n$ matrix constructed by a "side-by-side union" of the $k-2$ matrices $A_2, A_3, \ldots, A_{k-1}$, each of which is $n \times n$.

Remark. As the referee has observed, methods based on k-part splittings, or "linear stationary methods of kth degree," can be found in [4, p. 214; 10, Chap. 16; 8, p. 154]. In fact, in [8] a reduction from $k = 3$ to $k = 2$ is established. Our method differs from each of these in that we are not necessarily restricted to those tools peculiar to the finite-dimensional workshop, e.g., the Jordan normal form and the determinant. In fact, we may note, as the referee has pointed out, that convergence obtains for our k-part splittings in finite dimensions if and only if the solutions of

$$\det(\lambda^{k-1}A_1 + \lambda^{k-2}A_2 + \cdots + A_k) = 0$$

all lie inside the unit circle of the complex plane, but this fact does not serve us for infinite dimensions.

Remark. For typographical reasons, vectors of $\bigoplus^{k-2}\mathscr{H}$ are written horizontally, e.g., as $x = (x_{n-1}, x_{n-2}, ..., x_{n-k+2})$ following (2.2). To be consistent with more standardized notation of finite dimensions, we may think of these horizontal displays as vertical, or column vectors x; thus the notation Ax may be viewed as matrix multiplication of matrix A with column vector x.

With the terminology of (2.1) and (2.2) in hand, we immediately obtain the proposition that establishes the imbedding of k-part splittings into a two-part split system.

PROPOSITION 2.3. *Suppose invertible linear operator*

$$A = A_1 + A_2 + \cdots + A_k,$$

where A_1 is invertible. Given $k - 1$ initial vectors $x_0, x_1, ..., x_{k-2}$, $k > 2$, and its induced sequence $\{x_n\}$ defined by (1.4), *i.e.,* $\sum_{i=1}^{k} A_k x_{n-i+1} = y_0$, $n = k - 1, k, k + 1, ...$ *Then*

$$\begin{bmatrix} A_1 & B_1 \\ 0 & C \end{bmatrix} Z_n + \begin{bmatrix} B_2 & A_k \\ -C & 0 \end{bmatrix} Z_{n-1} = Y_0, \tag{2.3}$$

where

$$Z_n = (x_{n+k-2}, x_{n+k-3}, ..., x_n), \qquad n = 0, 1, 2, ...$$

$Y_0 = (y_0, 0, ..., 0)$ *are vectors in* $\bigoplus^{k-1}\mathscr{H}$. *Accordingly, given an arbitrary initial (column) vector* $Z_0 = (x_{k-2}, x_{k-3}, ..., x_0) \in \bigoplus^{k-1}\mathscr{H}$, $\{Z_n\}_0^\infty$ *is that unique sequence generated by Z_0 relative to the two-part splitting* (2.3) *of linear operator $\mathfrak{A}$ acting on* $\bigoplus^{k-1}\mathscr{H}$.

Proof. Verification.

An immediate consequence is

PROPOSITION 2.4. *The sequence $\{Z_n\}$ in $\bigoplus^{k-1}\mathscr{H}$ defined by the two-part splitting* (2.3) *of the operator $\mathfrak{A}$ converges to the solution X of the linear system*

$$\mathfrak{A}X = Y_0$$

for all initial (column) vectors $Z_0 = (x_{k-2}, x_{k-3}, \ldots, x_0)$ *if and only if the sequence* $\{x_n\}$ *in* $\mathscr{H}$ *defined by the k-part splitting* (1.4) *of operator* A *converges to the solution vector* x *of the linear system*

$$Ax = y_0$$

for all initial vectors $x_0, x_1, \ldots, x_{k-2}$ *in* $\mathscr{H}$.

Proof. Immediate from Proposition 2.1.

The following conjugacy result will prove useful.

PROPOSITION 2.5. *Given*

$$A = \begin{bmatrix} M & R^* \\ R & N \end{bmatrix},$$

representing an operator on Hilbert space $\mathscr{H}_1 \oplus \mathscr{H}_2$, *where* $M \in \mathscr{B}(\mathscr{H}_1)$, $N \in \mathscr{B}(\mathscr{H}_2)$, *and* R *is bounded linear sending* $\mathscr{H}_1$ *to* $\mathscr{H}_2$. *Suppose* M *is invertible and* M *and* N *are hermitian (so that* A *is hermitian). Then* A *is hermitian conjugate to the operator*

$$B = \begin{bmatrix} M & 0 \\ 0 & N - RM^{-1}R^* \end{bmatrix} = \operatorname{diag}[M, N - RM^{-1}R^*].$$

Proof. $B = X^*AX$ for

$$X = \begin{bmatrix} I_1 & -M^{-1}R^* \\ 0 & I_2 \end{bmatrix},$$

where I_1, I_2 are the identities on $\mathscr{H}_1$ and $\mathscr{H}_2$, respectively.

We shall have need of order-isomorphisms. That is, ϕ is an order isomorphism if $\phi: \mathscr{B}(\mathscr{H}) \to \mathscr{B}(\mathscr{H})$ is an invertible linear map on $\mathscr{B}(\mathscr{H})$ that sends positive semidefinite operators and only positive semidefinite operators to positive semidefinite operators. We remark that since the cone of positive semidefinite operators has nonempty interior, ϕ is automatically continuous in the uniform norm topology (see [6, p. 228]).

3. FOUR-PART SPLITTINGS

We consider the situation $A = A_1 + A_2 + A_3 + A_4 \in \mathscr{B}(\mathscr{H})$. In this case, the two-part splitting of $\mathfrak{A}$ in (2.1) assumes the form

$$\mathfrak{A} = \underset{\mathfrak{a}_1}{\begin{bmatrix} A_1 & A_2 - X & A_3 - Y \\ 0 & a & b \\ 0 & c & d \end{bmatrix}} + \underset{\mathfrak{a}_2}{\begin{bmatrix} X & Y & A_4 \\ -a & -b & 0 \\ -c & -d & 0 \end{bmatrix}}, \tag{3.1}$$

where X, Y, a, b, c, and d belong to $\mathscr{B}(\mathscr{H})$. These operator-parameters can be suitably chosen so that $\mathfrak{A}$ is hermitian on $\oplus^3 \mathscr{H}$, whenever our four-splitting allows that $A_1^* - A_2 - A_3 - A_4$ is hermitian on $\mathscr{H}$. In fact, write hermitian $A_1^* - A_2 - A_3 - A_4$ as a sum of hermitian operators H_1, H_2, and H_3. That is,

$$H_1 + H_2 + H_3 = A_1^* - A_2 - A_3 - A_4. \tag{3.2}$$

Replace X, Y, a, b, c, d of (3.1) by requiring

$$\begin{aligned} X &= A_1^* - H_1, \\ Y &= A_1^* - A_2 - H_1 - H_2 + b, \\ a &= H_2 - b^*, \\ b &= b, \\ c &= H_3 + b^*, \\ d &= H_3. \end{aligned} \tag{3.3}$$

With new hermitian parameters H_1, H_2, H_3 (constrained by (3.2)) and b (arbitrary), the two-part splitting of $\mathfrak{A} = a_1 + a_2$ given in (3.1) is written as follows:

$$\mathfrak{A} = \underbrace{\begin{bmatrix} A_1 & -A_1^* + A_2 + H_1 & -A_4 - H_3 - b \\ 0 & H_2 - b^* & b \\ 0 & H_3 + b^* & H_3 \end{bmatrix}}_{a_1} + \underbrace{\begin{bmatrix} A_1^* - H_1 & A_1^* - A_2 - H_1 - H_2 + b & A_4 \\ -H_2 + b^* & -b & 0 \\ -H_3 - b^* & -H_3 & 0 \end{bmatrix}}_{a_2}. \tag{3.4}$$

In adding the terms a_1 and a_2 of (3.4), we reveal $\mathfrak{A}$ in its hermitian form

$$\mathfrak{A} = \begin{bmatrix} 2\,\mathrm{Re}\,(A_1) - H_1 & -H_2 + b & -H_3 - b \\ -H_2 + b^* & H_2 - 2\,\mathrm{Re}\,(b) & b \\ -H_3 - b^* & b^* & H_3 \end{bmatrix}, \tag{3.5}$$

where $2\,\mathrm{Re}\,(B) = B + B^*$, twice the real part of operator B.

Since $\mathfrak{A}$ is hermitian and has a well-defined two-part splitting (3.4) (induced by the four-part splitting $A = A_1 + A_2 + A_3 + A_4$), we are in a position to apply Theorem 2.2, which in our case reduces to

PROPOSITION 3.1. *Given* $\mathfrak{A} \in \mathscr{B}(\oplus^3 \mathscr{H})$ *with the two-part splitting* $\mathfrak{A} = a_1 + a_2$ *of* (3.4), *then* $T(\mathfrak{A}) = \mathfrak{A} - (a_1^{-1}a_2)^* \mathfrak{A}(a_1^{-1}a_2)$ *is hermitian conjugate to* $a_1{}^* - a_2$. *More specifically,*

$$T(\mathfrak{A}) \sim \begin{bmatrix} H_1 & -A_1{}^* + A_2 + H_1 + H_2 - b & -A_4 \\ -A_1 + A_2{}^* + H_1 + H_2 - b^* & H_2 & H_3 + b \\ -A_4{}^* & H_3 + b^* & H_3 \end{bmatrix}. \tag{3.6}$$

We consider situations where $T(\mathfrak{A})$ is positive definite. A consequence of the Stein theorem, Theorem 2.1, is that convergence of $\{Z_n\}$ of (2.3) is equivalent to positive definiteness of $\mathfrak{A}$. But convergence of sequence $\{Z_n\}$ is, in turn, equivalent to convergence of the sequence $\{x_n\}$ defined by the four-part splitting

$$A_1x_{n+3} + A_2x_{n+2} + A_3x_{n+1} + A_4x_n = y_0 \qquad \text{(Proposition 2.4)}.$$

The result of these observations is the following theorem.

THEOREM 3.2. *Given the four-part splitting* $A = A_1 + A_2 + A_3 + A_4$ *and the sequence* $\{x_n\}$ *defined iteratively by*

$$A_1x_{n+3} + A_2x_{n+2} + A_3x_{n+1} + A_4x_n = y_0, \qquad n = 0, 1, 2,..., \tag{3.7}$$

suppose

$$A_1{}^* - A_2 - A_3 - A_4 = H_1 + H_2 + H_3 \tag{3.8}$$

for certain hermitian operators H_1, H_2, $H_3 \in \mathscr{B}(\mathscr{H})$. *Suppose an operator* $b \in \mathscr{B}(\mathscr{H})$ *exists such that*

$$\begin{bmatrix} H_1 & -A_1{}^* + A_2 + H_1 + H_2 - b & -A_4 \\ -A_1 + A_2{}^* + H_1 + H_2 - b^* & H_2 & H_3 + b \\ -A_4{}^* & H_3 + b^* & H_3 \end{bmatrix} > 0 \tag{3.9}$$

as an operator on $\mathscr{H} \oplus \mathscr{H} \oplus \mathscr{H}$. *Then for any initial triple* $\{x_0, x_1, x_2\}$, *the sequence* $\{x_n\}$ *defined by* (3.7) *converges to the solution vector x for the system*

$$Ax = y_0, \qquad A = A_1 + A_2 + A_3 + A_4,$$

if and only if

$$\begin{bmatrix} 2\,\mathrm{Re}\,(A_1) - H_1 & -H_2 + b & -H_3 - b \\ -H_2 + b^* & -H_2 - 2\,\mathrm{Re}\,(b) & b \\ -H_3 - b^* & b^* & H_3 \end{bmatrix} > 0 \tag{3.10}$$

as an operator on $\mathscr{H} \oplus \mathscr{H} \oplus \mathscr{H}$.

Proof. Hypothesis (3.8) (which agrees with (3.2)) tells us that the four-part splitting of $A = A_1 + A_2 + A_3 + A_4$ induces the two-part splitting of $\mathcal{A} = a_1 + a_2$ (cf. (3.4)) on $\mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H}$. As we have seen, $\mathcal{A}$ of (3.4) reduces (see (3.5)) to form (3.10). Proposition 3.1 tells us that

$$T(\mathcal{A}) = \mathcal{A} - (a_1^{-1}a_2)^* \mathcal{A}(a_1^{-1}a_2)$$

is hermetian conjugate to (3.9) and is positive definite. Given $Y_0 = (y_0, 0, 0)$ in $\mathcal{H} \oplus \mathcal{H} \oplus \mathcal{H}$, the sequence $\{Z_n\}$, $Z_n = (x_{n+2}, x_{n+1}, x_n)$, defined by

$$a_1 Z_{n+1} + a_2 Z_n = Y_0$$

converges to the solution vector X for the system

$$\mathcal{A}(X) = (a_1 + a_2) X = Y_0$$

if and only if $\rho(a_1^{-1}a_2)$, the spectral radius of $a_1^{-1}a_2$, is less than one. Since $T(\mathcal{A}) > 0$, Stein's result (Theorem 2.1) applies, so that $Z_n \to X$ if and only if $\mathcal{A} > 0$. $Z_n \to X$ if and only if $\{x_n\}$ of (3.7) is such that $x_n \to x$, where $Ax = y_0$ (Proposition 2.4). That is, $x_n \to x$ if and only if $\mathcal{A} > 0$. Since $\mathcal{A}$ appears in the statement (3.10), our theorem is proved.

Under more restrictive conditions, our testing matrices become much more tractable. As an example, we present

THEOREM 3.3. *Given the four-part splitting* $A = A_1 + A_2 + A_3 + A_4$ *for hermitian* $A \in \mathcal{B}(\mathcal{H})$, *and the sequence* $\{x_n\}$ *defined iteratively by*

$$A_1 x_{n+3} + A_2 x_{n+2} + A_3 x_{n+1} + A_4 x_n = y_0, \qquad n = 1, 2, 3, \ldots \qquad (3.11)$$

Suppose the operators A_1, A_2, A_3, A_4 *are constrained as follows:*

(i) $A_1^* - A_2 = H_1 + H_2$ *for certain positive definite* $H_1, H_2 \in \mathcal{B}(\mathcal{H})$.

(ii) $A_3 + A_4 = -H_3$ *for positive definite* $H_3 \in \mathcal{B}(\mathcal{H})$.

(iii) *Relative to the positive definite* H_1, H_2, H_3 *above,*

$$\begin{bmatrix} H_2 - H_3 H_1^{-1} H_3 & H_3 H_1^{-1} A_4 \\ A_4^* H_1^{-1} H_3 & H_3 - A_4^* H_1^{-1} A_4 \end{bmatrix}$$

is positive definite as an operator on $\mathcal{H} \oplus \mathcal{H}$. *Then for any initial triple* $\{x_0, x_1, x_2\} \subset \mathcal{H}$, *the sequence* $\{x_n\}$ *defined by* (3.11) *converges to the solution vector* x *for the system* $Ax = y_0$, *if and only if* A *is positive definite.*

Proof. Once we choose $b = -H_3$, (3.9) reduces to the form

$$\begin{bmatrix} H_1 & H_3 & -A_4 \\ H_3 & H_2 & 0 \\ -A_4^* & 0 & H_3 \end{bmatrix}. \qquad (3.12)$$

With $M = H_1$, $R^* = [H_3 - A_4]$, $N = \text{diag}[H_2, H_3]$ in Proposition 2.5, we see that (3.12) is hermitian conjugate to

$$\begin{bmatrix} H_1 & 0 & 0 \\ 0 & H_2 - H_3 H_1^{-1} H_3 & H_3 H_1^{-1} A_4 \\ 0 & A_4^* H_1^{-1} H_3 & H_3 - A_4^* H_1^{-1} A_4 \end{bmatrix},$$

which is positive definite on $\mathscr{H} \oplus \mathscr{H} \oplus \mathscr{H}$ due to hypothesis (i) and (iii). We are assured, then, that the sequence $\{x_n\}$ converges if and only if (3.10) is positive definite. But with $b = -H_3$, along with hypotheses (i) and (ii), (3.10) assumes the form

$$C = \begin{bmatrix} 2\,\text{Re}\,(A_1) - H_1 & -H_2 - H_3 & 0 \\ -H_2 - H_3 & H_2 + 2H_3 & -H_3 \\ 0 & -H_3 & H_3 \end{bmatrix}.$$

For the identity operator I on $\mathscr{H}$, define nonsingular

$$X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

as an operator on $\mathscr{H} \oplus \mathscr{H} \oplus \mathscr{H}$. Compute X^*CX to obtain

$$X^*CX = \begin{bmatrix} 2\,\text{Re}\,(A_1) - H_1 - H_2 - H_3 & 0 & 0 \\ 0 & H_2 + H_3 & 0 \\ 0 & 0 & H_3 \end{bmatrix}.$$

Now,

$$\begin{aligned} 2\,\text{Re}\,(A_1) - (H_1 + H_2 + H_3) &= A_1 + A_1^* - (A_1^* - A_2 - A_3 - A_4) \\ &= A_1 + A_2 + A_3 + A_4 \\ &= A, \end{aligned}$$

so that X^*CX is the direct sum of the operators A, $H_2 + H_3$, and H_3. Hence,

$$\begin{aligned} A > 0 &\rightleftarrows X^*CX > 0 \\ &\rightleftarrows C > 0 \\ &\rightleftarrows \{x_n\} \quad \text{converges to solution vector } x \text{ (Theorem 3.2).} \end{aligned}$$

The theorem is proved.

4. Three-Part Splittings

We assume A is hermitian on $\mathscr{H}$ and enjoys the splitting

$$A = A_1 + A_2 + A_3 .$$

The two-part splitting of $\mathfrak{A}$ in (2.1) is of the form

$$\mathfrak{A} = \underset{a_1}{\left|\begin{matrix} A_1 & A_2 - X \\ 0 & b \end{matrix}\right|} + \underset{a_2}{\left|\begin{matrix} X & A_3 \\ -b & 0 \end{matrix}\right|}, \tag{4.1}$$

where X and b are operators on $\mathscr{H}$. In order that $\mathfrak{A} = \mathfrak{A}^*$ on $\mathscr{H} \oplus \mathscr{H}$, it is necessary and sufficient that $b = b^*$ on $\mathscr{H}$, and $X = A_2 + A_3 + b$. In other words, we consider $\mathfrak{A}$ on $\mathscr{H} \oplus \mathscr{H}$ in (4.1) in the form

$$\begin{aligned} \mathfrak{A} &= \underset{a_1}{\begin{bmatrix} A_1 & -A_3 - b \\ 0 & b \end{bmatrix}} + \underset{a_2}{\begin{bmatrix} A_2 + A_3 + b & A_3 \\ -b & 0 \end{bmatrix}} \\ &= \begin{bmatrix} A + b & -b \\ -b & b \end{bmatrix}. \end{aligned} \tag{4.2}$$

Note that $a_1{}^* - a_2$, the hermitian conjugate to $T(\mathfrak{A})$ (cf. Theorem 2.2), is written

$$\begin{aligned} a_1{}^* - a_2 &= \begin{bmatrix} A_1 - A_2{}^* - A_3{}^* - b & -A_3 \\ -A_3{}^* & b \end{bmatrix} \\ &= \begin{bmatrix} 2\,\mathrm{Re}\,(A_1) - A - b & -A_3 \\ -A_3{}^* & b \end{bmatrix}. \end{aligned} \tag{4.3}$$

Since convergence of iteration schemes will depend on positive definiteness of $\mathfrak{A}$ in (4.2), we present

Lemma 4.1. *Given $b = b^* \in \mathscr{B}(\mathscr{H})$. Then*

$$\mathfrak{A} = \begin{bmatrix} A + b & -b \\ b & b \end{bmatrix} > 0$$

on $\mathscr{H} \oplus \mathscr{H}$ if and only if $A > 0$ and $b > 0$ on $\mathscr{H}$.

Proof. Let nonsingular

$$X = \begin{bmatrix} I & 0 \\ I & I \end{bmatrix},$$

where I is the identity on $\mathscr{H}$. Then

$$X^*\mathfrak{A}X = \begin{bmatrix} A & 0 \\ 0 & b \end{bmatrix},$$

which is positive definite if and only if A and b are.

We state a general theorem for three-part splittings of hermitian systems $Ax = y_0$.

THEOREM 4.2. *Given the three-part splitting $A = A_1 + A_2 + A_3$, A_1^{-1} exists in $\mathscr{B}(\mathscr{H})$, $A = A^*$ in $\mathscr{H}$. Let the sequence $\{x_n\}$ be defined inductively by*

$$A_1x_{n+2} + A_2x_{n+1} + A_3x_n = y_0, \qquad n = 0, 1, 2,.... \tag{4.4}$$

Suppose positive definite $b \in \mathscr{B}(\mathscr{H})$ exists such that

$$\begin{bmatrix} 2\operatorname{Re}(A_1) - A - b & -A_3 \\ -A_3^* & b \end{bmatrix} > 0 \tag{4.5}$$

as an operator on $\mathscr{H} \oplus \mathscr{H}$. Then for any initial couple $\{x_0, x_1\} \subset \mathscr{H}$, the sequence $\{x_n\}$ defined by (4.4) *converges to the solution vector x for the system*

$$Ax = y_0, \qquad A = A_1 + A_2 + A_3,$$

if and only if A is positive definite on $\mathscr{H}$.

The positive definite operator (4.5) is exactly $a_1^* - a_2$ of (4.3), which, in turn, is hermitian conjugate to $T(\mathfrak{A})$ in Theorem 2.2. That is, (4.5) tells us that $T(\mathfrak{A}) > 0$, so that convergence of Z_n to X, $\mathfrak{A}X = Y_0$, where $a_1Z_{n+1} + a_2Z_n = Y_0$, is equivalent to $\mathfrak{A} > 0$ (Theorem 2.1). Thus,

$\{x_n\}$ of (4.4), converges

$\rightleftarrows \{Z_n\}$ of $a_1Z_{n+1} + a_2Z_n = Y_0$ converges (Proposition 2.4),

$\rightleftarrows \mathfrak{A} = a_1 + a_2 > 0$ (Theorem 2.1),

$\rightleftarrows A > 0$ (Lemma 4.1).

This proves the theorem.

5. OPERATOR-PARAMETER PARTITIONS

Our operator $A \in \mathscr{B}(\mathscr{H})$ will be given a four-part partitioning, which will induce a family of three-part partitionings (definition follows). In this section, we give conditions for which each splitting in this family results in a convergent iterative sequence.

DEFINITION 5.1. Given $A = D + S_1 + S_2 + S_3, \in \mathscr{B}(\mathscr{H})$. Let $\phi_\omega: \mathscr{B}(\mathscr{H}) \to \mathscr{B}(\mathscr{H})$ be a family of order isomorphisms where ω belongs to some index set Ω. Then the operator-parameter-successive-overrelaxation-decomposition is the ω-dependent three-part decomposition

$$\begin{aligned} A &= [\phi_\omega^{-1}(D) + S_1] + [D - \phi_\omega^{-1}(D) + S_2] + [S_3] \\ &= [A_1(\omega)] + [A_2(\omega)] + [A_3]. \end{aligned}$$

Our next theorem shows that in the event that the order isomorphisms are "small enough," convergence of the sequence $\{x_n\}$ given by

$$A_1(\omega)\, x_{n+2} + A_2(\omega)\, x_{n+1} + A_3 x_n = y_0$$

to the solution vector x for $Ax = y_0$ is equivalent to $A > 0$.

THEOREM 5.2. *Let $A = A^*$ belong to $\mathscr{B}(\mathscr{H})$, and suppose*

$$A = D + S_1 + S_2 + S_3 .$$

Let $\{\phi_\omega\}$, $\omega \in \Omega$, $\phi_\omega: \mathscr{B}(\mathscr{H}) \to \mathscr{B}(\mathscr{H})$ be a family of order isomorphisms, each of which induces the generalized overrelaxation decomposition

$$\begin{aligned} A &= [\phi_\omega^{-1}(D) + S_1] + [D - \phi_\omega^{-1}(D) + S_2] + [S_3] \\ &= A_1(\omega) + A_2(\omega) + A_3 , \end{aligned}$$

where

$$A_1(\omega) = \phi_\omega^{-1}(D) + S_1 \qquad \textit{is invertible}$$

$$A_2(\omega) = D - \phi_\omega^{-1}(D) + S_2 , \qquad \textit{and} \qquad A_3 = S_3 .$$

Suppose $\Phi(D) \in \mathscr{B}(\mathscr{H})$, where $\Phi: \mathscr{B}(\mathscr{H}) \to \mathscr{B}(\mathscr{H})$ is continuous in the operator norm, such that

$$\Phi(D) + S_1{}^* - S_2 - S_3 > 0. \tag{5.1}$$

Suppose, too, we can find operator $P > 0$ in $\mathscr{B}(\mathscr{H})$ such that the matrix

$$\mathscr{P} = \begin{bmatrix} P & S_3 \\ S_3{}^* & \Phi(D) + S_1{}^* - S_2 - S_3 - P \end{bmatrix} > 0$$

as an operator on $\mathscr{H} \oplus \mathscr{H}$. Then for all $y_0 \in \mathscr{H}$, and all order isomorphisms sufficiently small, i.e., those order isomorphisms ϕ_ω, $\omega \in \Omega$, such that

$$\phi_\omega(\Phi(D)) + \phi_\omega(D) < D + D^*, \tag{5.2}$$

the sequence $\{x_n\}$ *defined by*

$$[\phi_\omega^{-1}(D) + S_1]\, x_{n+2} + [D - \phi_\omega^{-1}(D) + S_2]\, x_{n+1} + [S_3]\, x_n = y_0 \tag{5.3}$$

converges to the solution vector x *of the system*

$$Ax = y_0$$

for every initial couple $\{x_0, x_1\}$, *if and only if* $A > 0$.

Proof. Let us set

$$\begin{aligned} A_1(\omega) &= \phi_\omega^{-1}(D) + S_1, \\ A_2(\omega) &= D - \phi_\omega^{-1}(D) + S_2, \\ A_3 &= S_3. \end{aligned}$$

For our positive definite b in (4.5), choose (5.1), diminished by sufficiently small $P > 0$, i.e.,

$$b = \Phi(D) + S_1{}^* - S_2 - S_3 - P > 0.$$

Theorem 4.2 applies. With the quantities A_1, A_2, A_3, b thus defined, our test matrix (4.5) of Theorem 4.2 assumes the form

$$\begin{aligned} &\begin{bmatrix} A_1{}^*(\omega) - A_2(\omega) - A_3 - b & -A_3 \\ -A_3{}^* & b \end{bmatrix} \\ &= \begin{bmatrix} \phi_\omega^{-1}(D + D^*) - D - \Phi(D) + P & -S_3 \\ -S_3{}^* & \Phi(D) + S_1{}^* - S_2 - S_3 - P \end{bmatrix} \\ &= \begin{bmatrix} P & -S_3 \\ -S_3{}^* & \Phi(D) + S_1{}^* - S_2 - S_3 - P \end{bmatrix} \\ &\quad + \begin{bmatrix} \phi_\omega^{-1}(D + D^*) - D - \Phi(D) & 0 \\ 0 & 0 \end{bmatrix} \\ &= \mathscr{P} + \begin{bmatrix} \phi_\omega^{-1}(D + D^*) - D - \Phi(D) & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{5.4}$$

Now,

$$\phi_\omega^{-1}(D + D^*) - D - \Phi(D) > 0 \rightleftarrows D + D^* - \phi_\omega(\Phi(D) + D) > 0$$

since ϕ_ω is an order isomorphism $\rightleftarrows D + D^* > \phi_\omega(\Phi(D)) + \phi_\omega(D)$. This last assertion is assumed for the class ϕ_ω, $\omega \in \Omega$, as hypothesis (5.2). We are therefore assured that the operator

$$\begin{bmatrix} \phi_\omega^{-1}(D + D^*) - D - \Phi(D) & 0 \\ 0 & 0 \end{bmatrix}$$

is positive semidefinite on $\mathscr{H} \oplus \mathscr{H}$. Since $\mathscr{P} > 0$, it follows that our test matrix (4.5) is positive definite on $\mathscr{H} \oplus \mathscr{H}$. From Theorem (4.2), this positive definiteness (a consequence of our hypotheses) equates the equivalence of the convergence of the sequence $\{x_n\}$ (defined by (5.3) to the solution vector x, where $Ax = (D + S_1 + S_2 + S_3)\,x = y_0$) with positive definiteness of A. This ends the proof.

For the order isomorphisms, $A \to W^*AW$, the following obtains.

THEOREM 5.3. *Given $A = A^*$ and $D > 0$ in $\mathscr{B}(\mathscr{H})$, where*

$$A = D + S_1 + S_2 + S_3,$$

let A be decomposed as a three-part splitting in operator-parameter overrelaxation form

$$\begin{aligned} A &= [W^*DW + S_1] + [D - W^*DW + S_2] + [S_3] \\ &= A_1(W) + A_2(W) + A_3, \end{aligned} \tag{5.5}$$

where W is invertible. Given $X = X^* \in \mathscr{B}(\mathscr{H})$ such that

$$XD + S_1^* - S_2 - S_3 > 0. \tag{5.6}$$

We suppose that the operator-parameter W and the hermitian operator X commute. If $P > 0$ in $\mathscr{B}(\mathscr{H})$ can be found such that

$$\begin{bmatrix} P & S_3 \\ S_3^* & XD + S_1^* - S_2 - S_3 - P \end{bmatrix} > 0 \tag{5.7}$$

as an operator on $\mathscr{H} \oplus \mathscr{H}$, then for all splittings (5.5) *where W is constrained relative to hermitian X in the operator norm by the condition*

$$\| W^{-1} \|^2 \, \| I + X \| < 2, \tag{5.8}$$

the sequence $\{x_n\}$ defined by

$$[W^*DW + S_1]\, x_{n+2} + [D - W^*DW + S_2]\, x_{n+1} + [S_3]\, x_n = y_0 \tag{5.9}$$

converges to the solution vector x of the system

$$Ax = y_0$$

for every initial couple $\{x_0, x_1\}$ if and only if $A > 0$.

Proof. In the statement of Theorem 5.2, choose $\phi_\omega(B) = (W^{-1})^* BW^{-1}$, so that $\phi_\omega^{-1}(B) = W^*BW$ for all $B \in \mathscr{B}(\mathscr{H})$. Set $\Phi(D) = XD$. Theorem 5.2

reduces to Theorem 5.3 once we show that (5.8) implies hypothesis (5.2) of Theorem 5.2. To see this we observe that

$$\| W^{-1} \|^2 \| I + X \| < 2 \rightleftarrows \|(W^{-1})^* W^{-1} \| \cdot \| I + X \| < 2 \quad \text{property of operator norm}$$
$$\rightleftarrows \|(W^{-1})^* W^{-1}(I + X)\| < 2$$
$$\rightleftarrows 2 \cdot I - (W^{-1})^* W^{-1}(I + X) > 0$$
$$\text{since } (W^{-1})^* W^{-1}(I + X) \text{ is hermitian,}$$
$$\rightleftarrows 2D - (W^{-1})^* W^{-1}(I + X) D > 0$$
$$\text{since } D > 0 \text{ commutes with } W, X,$$
$$\rightleftarrows \phi_\omega(\Phi(D)) + \phi_\omega(D) < 2D = D + D^*,$$

since $\phi_\omega(\ ) = (W^{-1})^* (\ ) W^{-1}$, and $\Phi(\ ) = X(\ )$.

The reduction of Theorem 5.2 to Theorem 5.3 is established, thus completing the proof.

Donnelly's result follows directly. To reproduce his statement, we assume $A > 0$ at the outset. Accordingly, we have

THEOREM 5.4 (Donnelly [3, Theorem 2.1]). *Given the positive definite operators A and $D \in \mathscr{B}(\mathscr{H})$, with the splitting depending on the scalar ω,*

$$\omega A = [D - \omega F - \omega G] + [(\omega - 1) D - \omega(E + E^* + F^*)] + [-\omega G^*]$$
$$= \omega A_1^* + \omega A_2^* + \omega A_3^*. \tag{5.10}$$

Let the sequence $\{x_n\}$ be defined iteratively by

$$A_1 x_{n+2} + A_2 x_{n+1} + A_3 x_n = y_0, \qquad n = 0, 1, 2, \ldots \tag{5.11}$$

The following constraint is assumed. There exists positive definite operator P on $\mathscr{H}$ and a scalar $\alpha > -1$ such that

$$\begin{bmatrix} P & G \\ G^* & \alpha D + E + E^* - P \end{bmatrix} > 0 \tag{5.12}$$

as an operator on $\mathscr{H} \oplus \mathscr{H}$. Then for all ω,

$$0 < \omega < 2/(1 + \alpha), \tag{5.13}$$

and for any initial couple $\{x_0, x_1\}$, the sequence $\{x_n\}$ defined by (5.11) *converges to the solution vector for the system $Ax = y_0$.*

Proof. Condition (5.10) is equivalent to

$$A = [(1/\omega) D - (F + G)] + [D - (1/\omega) D - (E + E^* + F^*)] + [-G^*].$$

Comparison of this decomposition with (5.5) leads us to define

$$W = \omega^{-1/2} I, \quad \omega > 0, \qquad (\text{i.e.}, \phi_\omega^{-1}(D) = W^* D W = (1/\omega)\, D) \tag{5.14}$$

and

$$X = \alpha I, \qquad \alpha > 1, \qquad (\text{i.e.}, \Phi(D) = \alpha D).$$

We also define

$$\begin{aligned} S_1 &= -(F + G), \\ S_2 &= -(E + E^* + F^*), \\ S_3 &= -G^*. \end{aligned} \tag{5.15}$$

Thus, condition (5.12) for $A > 0$ is equivalent to (5.7) of Theorem 5.3. Observe that constraint (5.12) implies that the lower right-hand corner, $\alpha D + D + E^* - P$, is positive definite. That is,

$$\begin{aligned} 0 < \alpha D + E + E^* - P \to 0 &< \alpha D + E + E^* && \text{since } P > 0, \\ \rightleftarrows 0 &< XD + S_1{}^* - S_2 - S_3 && \text{from (5.14) and (5.15)}, \end{aligned}$$

so that condition (5.6) obtains. Condition (5.7) also obtains, since the matrices of (5.7) and (5.12) agree. The constraint given in (5.8) for $W^{-1} = \omega^{1/2} I$, $\omega > 0$ and $X = \alpha I$, $\alpha > 1$ easily reduces to Donnelly's hypothesis (5.13) that $0 < \omega < 2/(1 + \alpha)$. Since all the hypotheses of Theorem 5.3 obtain in the statement of Donnelly's theorem, the proof is done.

6. Three-Part Splittings that Improve Convergence

In the last two sections, we dealt with three-part splittings

$$A = A_1 + A_2 + A_3$$

and identified convergence of the sequence $\{x_n\}$, where

$$A_1 x_{n+2} + A_2 x_{n+1} + A_3 x_n = y_0 ,$$

with positive definiteness of the test operator $T(\mathcal{O})$. With this strategy, however, we were constrained to consider only those bounded linear operators that were self-adjoint. The present section abandons the coupling of convergence and positive definiteness and considers arbitrary bounded, linear operators A.

In measuring "improvement of convergence" offered by a three-part splitting $A = A_1 + A_2 + A_3$, over a conventional two-part splitting $A = A_1' + A_2'$, we shall mean that the *average reduction factor* per iteration (cf. [8, p. 62]) of the three-part sequence $\{x_n\}$, $A_1 x_{n+2} + A_2 x_{n+1} + A_3 x_n = y_0$, is, in a sense, "better than," or less than, that of the induced two-part sequence $\{x_n\}$, $A_1' x_{n+1}' + A_2' x_n' = y_0$. (Specific definitions follow.)

A major point in resorting to iteration via splittings is that the operator A, not easily invertible, at least yields an additive piece, A_1, that *is* easy to invert. We therefore assume that A_1 remains invariant in construction of the two-part splitting $A = A_1 + A_2'$ and in all the three-part splittings $A = A_1 + A_2 + A_3$; that is, $A_2' = A_2 + A_3$. Although A_1 is fixed in all our constructions, the degree of freedom we have in choosing A_3, with three-part splittings, proves sufficient to improve the convergence rate of a convergent sequence $\{x_n'\}$, i.e., certain three-part sequences $\{x_n\}$ formed by the splittings $A = A_1 + A_2 + A_3$ converge faster than the two-part sequence $\{x_n'\}$ formed from the splitting $A = A_1 + A_2'$. As a further bonus, however, by selecting A_3 judiciously (A_1 is fixed), we can *always* construct a convergent three-part sequence, $\{x_0, x_0, x_2, ..., x_n, ...\}$, when the corresponding two-part sequence, $\{x_0, x_1', x_2', ..., x_n', ...\}$ is nonconvergent and the spectrum of $A_1^{-1} A_2'$ lies in the disk $\{z: | z - 1 | < 2\}$ (Theorem 6.3).

We split $A \in \mathscr{B}(\mathscr{H})$ by $A = A_1 + A_2 + A_3$. In this case, the associated operator $\mathfrak{A}$ of (2.1), in $\mathscr{B}(\mathscr{H} \oplus \mathscr{H})$, has the following family of two-part splittings:

$$\begin{aligned} \mathfrak{A} &= \begin{bmatrix} A_1 & A_2 + D \\ 0 & C \end{bmatrix} + \begin{bmatrix} -D & A_3 \\ -C & 0 \end{bmatrix} \\ &= a_1 + a_2 , \end{aligned} \tag{6.1}$$

where $D, C \in \mathscr{B}(\mathscr{H})$, C^{-1} exists in $\mathscr{B}(\mathscr{H})$, are arbitrary. Recall that for the system $\mathfrak{A}X = Y_0 = (y_0, 0)$ and for initial vector $Z_0 = (x_1, x_0) \in \mathscr{H} \oplus \mathscr{H}$, the two-part splitting $\mathfrak{A} = a_1 + a_2$ induces the sequence

$$\{Z_n = (x_{n+1}, x_n)\} \subset \mathscr{H} \oplus \mathscr{H},$$

which converges to the solution vector $X = (x, x)$ if and only if $\{x_n\} \subset \mathscr{H}$ defined by the three-part splitting

$$A = A_1 + A_2 + A_3$$

(i.e., $A_1 x_{n+2} + A_2 x_{n+1} + A_3 x_n = y_0$) converges to the solution vector x of the linear system $Ax = y_0$ for initial vectors x_0, x_1 in $\mathscr{H}$ (Proposition 2.4). Now convergence of the sequence $\{Z_n\}$ is equivalent to the operator $\mathscr{B} = a_1^{-1} a_2$

having spectral radius less than 1. Computation reveals that for all C, D of (6.1),

$$\mathscr{B} = a_1^{-1}a_2 = \begin{bmatrix} A_1^{-1}(A - A_1 - A_3) & A_1^{-1}A_3 \\ -I & 0 \end{bmatrix}$$

$$= \begin{bmatrix} (A_1^{-1} - I) - (A_1^{-1}A_3) & A_1^{-1}A_3 \\ -I & 0 \end{bmatrix} \qquad (6.2)$$

$$= \begin{bmatrix} B - T & T \\ -I & 0 \end{bmatrix},$$

where

$$B = A_1^{-1}A - I, \qquad T = A_1^{-1}A_3 .$$

For completeness we offer the following definition.

DEFINITION 6.1. Given the linear system $Ax = y_0$, $A, A^{-1} \in \mathscr{B}(\mathscr{H})$, where A has k-part splitting $A = A_1 + A_2 + \cdots + A_k$, $A_k \neq 0$. Given the sequence $\{x_n\}$ defined by $A_1x_{n+k-1} + A_2x_{n+k-2} + \cdots + A_kx_n = y_0$ with initial vectors $x_0 = x_1 = \cdots = x_{k-1}$. If x is the solution vector for the system, then the *average reduction factor per iteration, after m iterations*, denoted by $\sigma(m)$, is the quantity

$$\sigma(m) = \left(\frac{\| x_m - x \|}{\| x_0 - x \|}\right)^{1/m},$$

where $\| \ \|$ is any norm on vector space $\mathscr{H}_1$ that is compatible with the (fixed) norm on the operators (or matrices) A on $\mathscr{H}$. That is. for all $x \in \mathscr{H}$ and all operators A on $\mathscr{H}$, we have $\| A(x)\| \leqslant \| A \| \cdot \| x \|$.

Remark. Compatibility implies that $\lim_{n\to\infty} \| A^n \|^{1/n} = \rho(A)$, the spectral radius of A. For two-part splittings, $x_m - x = (-A_1^{-1}A_2)^m (x_0 - x)$, so that $\sigma(m)$ is bounded by the operator norm of $-A_1^{-1}A_2$ as follows, for large enough m:

$$\sigma(m) \leqslant \|(A_1^{-1}A_2)^m \|^{1/m}.$$

The following proposition considers the sequence $\{x_n\}$ induced by the three-part splitting $A = A_1 + A_1 + A_3$ and shows how the average reduction factor $\sigma(m)$ is bounded in terms of $\| \mathscr{B}^m \|$, the norm of $\mathscr{B}^m$, where $\mathscr{B} = a_1^{-1}a_2$ is the operator of (6.2).

PROPOSITION 6.2. *Given* A, $A^{-1} \in \mathscr{B}(\mathscr{H})$ *where* $A = A_1 + A_2 + A_3$, A_1, A_1^{-1}, A_2, $A_3 \in \mathscr{B}(\mathscr{H})$. *Let* $\{x\}$ *be defined by*

$$A_1x_{n+2} + A_2x_{n+1} + A_3x_n = y_0, \qquad n = 0, 1, 2,...,$$

with equal initial vectors $x_0 = x_1$. *Then* $\sigma(m)$, **the average reduction factor per iteration, after *m* iterations,** *is such that*

$$\sigma(m) = \left(\frac{\| x_m - x \|}{\| x_0 - x \|}\right)^{1/m} \leqslant (\sqrt{2} \, \| \mathscr{B}^m \|)^{1/m},$$

where x *is the solution vector for the system* $Ax = y_0$, $\mathscr{B} = a_1^{-1} a_2$ *as defined in* (6.2).

Proof. Since we have chosen equal initial vectors $x_0 = x_1$, the initial vector of our associated two-part splitting (cf. Proposition 2.4) is $Z_0 = (x_0, x_0)$. Recall that if x is the solution to the system $Ax = y_0$, necessarily, $X = (x, x)$ is the solution (column) vector of the two-part system $\mathfrak{A} X = (y_0, 0)$. Since $\mathfrak{A} = a_1 - a_2$, we have that

$$\begin{aligned}(-\mathscr{B})^m (x_0 - x, x_0 - x) &= (-a_1^{-1} a_2)^m (x_0 - x, x_0 - x) \\ &= (x_{m+1} - x, x_m - x).\end{aligned} \tag{6.3}$$

Consider the projection operator $\mathscr{P}(u, y) = (0, y)$ on $\mathscr{H} \oplus \mathscr{H}$, which applied to both sides of (6.3), yields

$$\mathscr{P} \cdot (-\mathscr{B})^m (x_0 - x, x_0 - x) = (0, x_m - x).$$

Allowing that $\mathscr{H} \oplus \mathscr{H}$ is normed by setting $\|(u, y)\|^2 = \| u \|^2 + \| y \|^2$ for $u, y \in \mathscr{H}$, the last equality implies

$$(2)^{1/2} \| \mathscr{B}^m \| \cdot \| x_0 - x \| \geqslant \| x_m - x \| .$$

That is,

$$\sigma(m) = \left(\frac{\| x_m - x \|}{\| x_0 - x \|}\right)^{1/m} \leqslant (\sqrt{2} \, \| \mathscr{B}^m \|)^{1/m},$$

proving the proposition.

Remark. Suppose that relative to the system $Ax = y_0$, the three-part splitting $A = A_1 + A_2 + A_3$, with initial vectors $x_0 = x_1$, induces the sequence $\{x_n\}$. Suppose the two-part splitting $A = A_1 + A_2'$, $A_2' = A_2 + A_3$, with the same initial vector x_0, induces the sequence $\{x_n'\}$. We already know that for $\{x_n'\}$,

$$\sigma'(m) = \left(\frac{\| x_m' - x \|}{\| x_0 - x \|}\right)^{1/m} \leqslant \|(-A_1^{-1}, A_2^1)^m \|^{1/m},$$

while Proposition 6.2 tells us that for $\{x_n\}$,

$$\sigma(m) = \left(\frac{\| x_m - x \|}{\| x_0 - x \|}\right)^{1/m} \leqslant \sqrt{2}^{1/m} \|(-a_1^{-1} a_2)^m \|^{1/m}.$$

We are ready to state our comparison theorem, which states in what sense (in terms of the average reduction factors, $\sigma(m)$ and $\sigma'(m)$) three-part splittings can be made better than the two-part splitting.

THEOREM 6.3. *Consider A invertible in $\mathscr{B}(\mathscr{H})$ and the linear system $Ax = y_0$. Given invertible $A_1 \in \mathscr{B}(\mathscr{H})$ and $x_0 \in \mathscr{H}$, which are fixed throughout, so that we have the splittings*

$$A = A_1 + A_2 + A_3, \qquad \textit{inducing sequence } \{x_0, x_0, x_2, \ldots, x_n, \ldots\},$$

and

$$A = A_1 + A_2', \qquad \textit{inducing the sequence } \{x_0, x_1', x_2', \ldots, x_n', \ldots\}.$$

Then,

Case A. *If $\{x_n'\}$ is divergent, ($\rho(A_1^{-1}A_2') > 1$), at least for the case where $\sigma(A_1^{-1}A_2')$, the spectrum of $A_1^{-1}A_2'$, lies in the open disk $\{z: |z - 1| < 2\}$, then A_3 may be chosen so that $\{x_n\}$ is convergent, that is, so that $\rho(a_1^{-1}a_2) < 1$.*

Case B. *If $\{x_n'\}$ is convergent and if $\sigma(A_1^{-1}A_2')$ lies in the right half of the open unit disk, $\{z: |z| < 1$ and $z + \bar{z} \geqslant 0\}$, then A_3 may be chosen so that $\sigma(m)$, for $\{x_n\}$, has about one-half the upper bound that $\sigma'(m)$ has for $\{x_n'\}$. That is, if $\sigma'(m) < \| B^m \|^{1/m}$, for all m, then A_3 may be chosen so that given $\epsilon > 0$, then for all integers m, sufficiently large,*

$$\sigma(m) < \tfrac{1}{2}\rho(B) + \epsilon,$$

where $B = A_1^{-1}A_2'$.

In both cases A_3 may be chosen in the form

$$(*) \qquad A_3 = A_1\phi(B)\,(I + \phi(B))^{-1}\,(B - \phi(B)).$$

For Case A, *$\phi(B) = p_1B - p_2I$, for certain scalars p_1, p_2 which meet the conditions that $p_1 > p_2 > 0$ and $p_1 + p_2 = 1$. For Case* B, *$\phi(B) = \frac{1}{2}B$.*

Proof. Let us consider the matrix operator of (6.2),

$$\mathscr{B} = a_1^{-1}a_2 = \begin{bmatrix} B - A_1^{-1}A_3 & A_1^{-1}A_3 \\ -I & 0 \end{bmatrix}, \tag{6.4}$$

whose spectral radius, $\rho(\mathscr{B})$, depends on $B = A_1^{-1}A_2'$ (which derives from the two-part splitting $A = A_1 + A_2'$), and on our choice of A_3.

Let $\phi(\ )$ be any complex analytic function on $C(\rho(B)) = \{z: |z| \leqslant \rho(B)\}$, the smallest closed disk centered at the origin, containing the spectrum of B.

We further assume that for all $z \in C(\rho(B))$, $\phi(z) \neq -1$. This allows us to define the operators U, V, and A_3 as follows.

$$\begin{aligned} U &= \boldsymbol{\phi}(B), \\ V &= (I + \boldsymbol{\phi}(B))^{-1}(B - \boldsymbol{\phi}(B)), \\ A_3 &= A_1 U V, \end{aligned} \tag{6.5}$$

where $\boldsymbol{\phi}(\ )$ is the corresponding operator-valued analytic function, induced by $\phi(\ )$, whose domain therefore includes any operator whose spectrum is contained in $C(\rho(B))$.

The definitions (6.5) imply that

$$B = U + V + UV,$$

so that (6.4) rewrites itself as

$$\mathscr{B} = \begin{bmatrix} U + V & UV \\ -I & 0 \end{bmatrix}. \tag{6.6}$$

It will now follow that $\sigma(\mathscr{B})$, the spectrum of $\mathscr{B}$, is just $\sigma(U) \cup \sigma(V)$, the union of the spectrums of U and V. To see this, use the fact that $\sigma(\mathscr{B}) = \sigma(W^{-1}\mathscr{B}W)$ for any invertible $W \in \mathscr{B}(\mathscr{H} \oplus \mathscr{H})$; then choose W to be the matrix-operator

$$\mathscr{W} = \begin{bmatrix} I & -V \\ 0 & I \end{bmatrix},$$

so that

$$\mathscr{W}^{-1}\mathscr{B}\mathscr{W} = \begin{bmatrix} U & 0 \\ -I & V \end{bmatrix}. \tag{6.7}$$

Thus, the operator $\mathscr{W}^{-1}\mathscr{B}\mathscr{W}$ of (6.7) (hence, operator $\mathscr{B}$ of (6.6) has spectral radius

$$\rho(\mathscr{B}) = \max\{\rho(U), \rho(V)\}. \tag{6.8}$$

Given (6.8), the question before us now is whether we can choose analytic $\phi(\ )$, which defines U, V and, hence, $A_3 = A_1 UV$, so that for Case A,

$$\{x_n'\} \text{ diverges}, \qquad \rho(A_1^{-1}A_2') > 1, \qquad \sigma(A_1^{-1}A_2') \subset \{z: |z - 1| < 2\},$$

and

$$\max\{p(U), \rho(V)\} < 1. \tag{6.9}$$

By the spectral mapping theorem and from (6.5), condition (6.9) obtains if for some analytic function $\phi(\ )$ whose domain of definition contains $\sigma(A_1^{-1}A_2')$, if and only if (Case A) $\{x_n'\}$ diverges,

$$\rho(A_1^{-1}A_2') > 1, \qquad \sigma(A_1^{-1}A_2') \subset \{z: | z - 1 | < 2\},$$

and for all $z \in \sigma(A_1^{-1}A_2')$,

$$| \phi(z)| < 1 \tag{6.10a}$$

and

$$| \phi(z) - z | < | \phi(z) + 1 | . \tag{6.10b}$$

To establish Case A of our theorem, it suffices to establish the conditions (6.10a), (6.10b).

Let us consider (6.10a) and (6.10b) geometrically. Let z be complex, and let H_z be the closed half-plane of all complex w that are at least as near to z as they are to -1. That is,

$$H_z = \{w: | w - z | \leqslant | w + 1 |\}.$$

If we denote by H^0, the set of all closed half-planes H_z that do not intersect the closed unit disk, then the union of all $H_z \in H^0$ is the "outside" complement of the cardioid

$$C = \{2z(1 + \text{Re}(z)) - 1: z = e^{i\theta}\}. \tag{6.11}$$

This means that should any point z_0 of the spectrum of $B = A_1^{-1}A_2'$ lie outside the cardioid $\mathscr{C}$, then for *all* functions ϕ (analytic or not) taking the complex plane to itself, if $\phi(z_0)$ lies inside the unit circle ((6.10a) obtains), then, necessarily, $| \phi(z_0) - z_0 | > | \phi(a_0) + 1 |$ ((6.10b) does not obtain). Thus, we have shown:

If a complex valued (analytic) function $\phi(\ )$ defined on the set $\sigma(A_1^{-1}A_2')$, say, enjoys properties (6.10a) *and* (6.10b) *for all $z \in \sigma(A_1^{-1}A_2')$, then, necessarily, $\sigma(A_1^{-1}A_2')$ lies in the interior of the cardiod $\mathscr{C}$ of* (6.11). (6.12)

Note that, consistent with constraint (6.12) on the position of $\sigma(A_1^{-1}A_2')$, Case A in the statement of our theorem carries the assumption that

$$\sigma(A_1^{-1}A_2') \subset \{z: | z - 1 | < 2\} \subset \mathscr{C}.$$

Under this assumption, an analytic $\phi(\ )$ on $\sigma(A_1^{-1}A_2')$ satisfying (6.10a) and (6.10b) is easy to find. In fact, choose $\phi(\ )$ to be that function that pulls z back to (almost) the midpoint between z itself and the fixed point -1. That is, subject to (6.10a), set

$$\phi(z) = p_1 z = p_2 , \tag{6.13}$$

where $p_1 > p_2 > 0$ and $p_1 + p_2 = 1$. Intuitively, in (6.13), we pull the disk$\{z: | z - 1 | < 2\}$ back toward the point -1 so that it covers the closed unit disk: we make the fit close enough (choose p_1 close enough to $\frac{1}{2}$) so that $\sigma(A_1^{-1}A_2')$ is carried to the interior of the unit disk, thus satisfying (6.10a). Since $p_1 > p_2$, (6.10b) will also be satisfied, as is easily verified. We have thus settled Case A, having converted the divergent sequence $\{x_0, x_1', x_2',..., x_n',...\}$, resulting from the splitting $A = A_1 + A_2'$, to a convergent sequence $\{x_0, x_1, x_2,..., x_n,...\}$, resulting from the splitting

$$A = A_1 + A_2 + A_1\phi(B)\frac{(B - \phi(B))}{(I + \phi(B))}, \tag{6.14}$$

where $B = A_1^{-1}A_2'$ and $\phi(\ )$ is proscribed in (6.13). Of course, once

$$A_3 = A_1\phi(B)(I + \phi(B))^{-1}(B - \phi(B))$$

is fixed, A_2 of (6.14) is uniquely determined.

To establish Case B, we need only exhibit complex valued $\phi(\ )$, analytic on $\{z: | z | < 1 \text{ and } \mathrm{Re}(z) \geqslant 0\}$ such that

$$| \phi(z)| < \tfrac{1}{2}\rho(A_1^{-1}A_2') \tag{6.15a}$$

and

$$|(\phi(z) - z)/(\phi(z) + 1)| < \tfrac{1}{2}\rho(A_1^{-1}A_2'). \tag{6.15b}$$

Such a ϕ is

$$\phi(z) = \tfrac{1}{2}z.$$

Conditions (6.15a) and (6.15b) together allow that

$$\rho(a_1^{-1}a_2) = \tfrac{1}{2}\rho(A_1^{-1}A_2') \qquad \text{(cf. (6.8))}. \tag{6.16}$$

We use the fact that for all m sufficiently large, and for any $A \in \mathscr{B}(\mathscr{H})$, $\| A^m \|^{1/m}$ converges (eventually downward) to $\rho(A)$, the spectral radius of A. Thus,

$$\sigma(m) = \left(\frac{\| x_m - x \|}{\| x_0 - x \|}\right)^{1/m} \leqslant ((2)^{1/2} \|(a_1^{-1}a_2)^m \|)^{1/m} \qquad \text{Proposition 6.2,}$$

$$\to \rho(a_1^{-1}a_2), \qquad \text{as} \qquad m \to \infty \text{ since } \|\ \| \text{ is compatible,}$$

$$= \tfrac{1}{2}\rho(A_1^{-1}A_{2'}) \qquad \text{from (6.16).}$$

That is, for any $\epsilon > 0$, there is an $N_\epsilon > 0$, such that for all integers $m > N_\epsilon$, with $B = A_1^{-1}A_2'$,

$$\sigma(m) < \tfrac{1}{2}\rho(B) + \epsilon,$$

proving the theorem.

7. An Example

In Theorem 6.3, Case B, we specify a three-part splitting that reduces the spectral radius by a factor of almost two. But we may pay the price of increased efficiency in the computation of the operator $(I + \phi(B))^{-1}$ in the construction of A_3 given in (*) and (5.14). In the following class of operators, this presents no problem.

Let us consider the $2n \times 2n$ matrix

$$A = \begin{bmatrix} u_1 I & F \\ F^* & u_2 I \end{bmatrix}, \tag{7.1}$$

where u_1, u_2 are scalars. We assume that F^{-1} is easy to find, and

$$\sigma(F^*F) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\},$$

the spectrum of F^*F relates to scalars u_1 and u_2 as follows: for each $i = 1, 2, \ldots, n$, λ_i/u_1u_2 lies in the right-half interior of the unit circle of the complex plane centered at $z = 2$. More precisely, for all $i = 1, 2, \ldots, n$, if $\lambda_i \in \sigma(F^*F)$, then

$$\frac{\lambda_i}{u_1u_2} \in \{z : |z - 2| < 1, \text{ and } z + \bar{z} \geqslant 4\}. \tag{7.2}$$

We now choose to provide A with the α-splitting introduced in [2, Section 3], with $\alpha = -1$. That is, set

$$A = \underset{A_1}{\begin{bmatrix} u_1 I & 0 \\ F^* & -u_2 I \end{bmatrix}} + \underset{A_2'}{\begin{bmatrix} 0 & F \\ 0 & 2u_2 I \end{bmatrix}}. \tag{7.3}$$

From this representation, we compute $B = A_1^{-1}A_2'$ as

$$B = A_1^{-1}A_2' = \begin{bmatrix} 0 & u_1^{-1}F \\ 0 & -2I + (u_1u_2)^{-1}F^*F \end{bmatrix},$$

which implies that $\sigma(B)$, the spectrum of B, equals

$$\sigma(B) = \{0, (u_1u_2)^{-1}\,\sigma(F^*F) - 2\}. \tag{7.4}$$

From the placement of $\sigma(F^*F)$ in (7.2), condition (7.4) is equivalent to saying that

$$\sigma(B) \subset \{z\colon |\, z \,| < 1 \text{ and } z + \bar{z} \geqslant 0\}.$$

This last condition allows us to employ Case B of Theorem 6.3. Accordingly, for $\phi(B) = \frac{1}{2}B$, we have

$$I + \phi(B) = \begin{bmatrix} I & \frac{1}{2}u_1^{-1}F \\ 0 & \frac{1}{2}(u_1u_2)^{-1}F^*F \end{bmatrix},$$

which has the tractable inverse

$$(I + \phi(B))^{-1} = \begin{bmatrix} I & -u_2(F^*)^{-1} \\ 0 & 2u_1u_2(F^*F)^{-1} \end{bmatrix}$$

since F^{-1} is presumed easy to find. Following (*) (or (6.14)) we compute the three-part splitting

$$\begin{aligned} A &= \begin{vmatrix} u_1I & F \\ F^* & u_2I \end{vmatrix} \\ &= \begin{bmatrix} u_1I & 0 \\ F^* & -u_2I \end{bmatrix} + \begin{bmatrix} 0 & \frac{1}{2}F + u_1u_2(F^*)^{-1} \\ 0 & u_2 + 2u_1u_1^2(F^*F)^{-1} \end{bmatrix} + \begin{bmatrix} 0 & \frac{1}{2}F - u_1u_2(F^*)^{-1} \\ 0 & u_2 - 2u_1u_2^2(F^*F)^{-1} \end{bmatrix} \\ &= A_1 + A_2 + A_3\,. \end{aligned} \tag{7.5}$$

Now with the two-part α-splitting of (7.3) with the constraints on the scalars u_1, u_2 and $\sigma(F^*F)$ as given in (7.2), we have a convergence rate $\rho(B)$ equal to

$$\max\{|\, \lambda_i(u_1u_2)^{-1} - 2 \,|\} < 1. \tag{7.6}$$

But the three-part splitting of (7.5) gives us a convergence rate equal to one-half of (7.6).

Now let us consider the particular family of six-by-six matrices $A(a)$, depending on real parameter a, of (7.1), where scalars $u_1 = u_2 = 1$, and

$$F(a) = \frac{a}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}. \tag{7.7}$$

Note that $U = F(a)/a$ is an orthogonal idempotent matrix ($U = U^t = U^{-1}$) with eigenvalues equal to ± 1. Effect the α-splitting ($\alpha = -1$) for $A(a)$ according to (7.3) so that the (family of) $B(a) = A_1^{-1}A_2'$ has the form

$$B(a) = \begin{bmatrix} 0 & 0 \\ 0 & (|a|^2 - 2)I \end{bmatrix}, \tag{7.8}$$

so that

$$\sigma(B(a)) = \{0, |a|^2 - 2\}.$$

We now tabulate some comparative readings of sequences $\{x_n'\}$ and $\{x_n\}$, induced by the two-part splitting of (7.3) and the three-part splitting of (7.5), respectively, with various values of the parameter a. Let us choose the linear system $A(a)\,x = 0$ with initial vector of each of the sequences $\{x_n'\}$, $\{x_n\}$ equal to $x_0 = \mathrm{col}(8, 4, -5, 4, 2, 0)$, taking

$$A(a) = \begin{bmatrix} I & F(a) \\ F(a) & I \end{bmatrix},$$

where I is the 3×3 identity matrix and the (family of) 3×3 matrices $F(a)$ are given in (7.7). Along with the values of the Euclidean norms of x_n' and x_n, we list (Definition 6.1) their respective average reduction factors per iteration after n iterations, $\sigma'(n)$ and $\sigma(n)$. Now,

$$\begin{bmatrix} I & 0 \\ F(a) & -I \end{bmatrix} x'_{n+1} + \begin{bmatrix} 0 & F(a) \\ 0 & 2I \end{bmatrix} x_n' = 0,$$

$$\begin{bmatrix} I & 0 \\ F(a) & -I \end{bmatrix} x_{n+2} + \begin{bmatrix} 0 & \frac{1}{2} + \frac{1}{a^2}F(a) \\ 0 & \left(1 + \frac{2}{a^2}\right) I \end{bmatrix} x_{n+1} + \begin{bmatrix} 0 & \left(\frac{1}{2} - \frac{1}{a^2}\right) F(a) \\ 0 & \left(1 - \frac{2}{a^2}\right) I \end{bmatrix} x_n = 0.$$

Initial $x_0 = x_1 = x_0' = \mathrm{col}(8, 4, -5, 4, 2, 0)$.
In summary,

$$B(a) = \begin{bmatrix} I & 0 \\ F(a) & -I \end{bmatrix}^{-1} \begin{bmatrix} 0 & F(a) \\ 0 & 2I \end{bmatrix},$$

$$\rho(B) = |a^2 - 2|,$$

(cf. (7.8)), and $\phi(B) = \frac{1}{2}B$, implying (Theorem 6.3, Case B) that $\sigma(n)$, the average reduction factor after n iterations on the three-part sequence, is about half of that for the two-part sequence, when n, the number of iterations, is large enough.

APPENDIX

TABLE I

$a = 1.33\ldots \quad \rho(B(a)) = 0.22\ldots$

n	$\|x_n\|/\|x_0\|$	$\|x_n'\|/\|x_0\|$	$\sigma(n)$	$\sigma'(n)$
1	1.000 000	0.608 276	1.000	0.608
2	0.608 276	0.152 069	0.779	0.389
3	0.109 827	0.038 017	0.478	0.336
4	0.017 483	0.009 504	0.363	0.312
5	0.002 602	0.002 376	0.304	0.298
6	0.000 371	0.000 594	0.268	0.289
10	0.000 000	0.000 002	0.203	0.273
11	0.000 000	0.000 000	0.195	0.271
18	0.000 000	0.000 000	0.166	0.262

$\sigma(18)/\sigma'(18) = 0.633$

TABLE II

$a = 1.71 \quad \rho(B(a)) = 0.9241$

(x_n' barely converges, x_n converges twice as fast)

n	$\|x_n\|/\|x_0\|$	$\|x_n'\|/\|x_0\|$	$\sigma(n)$	$\sigma'(n)$
1	1.000 000	0.777 489	1.000	0.777
2	0.777 489	0.718 477	0.881	0.847
3	0.482 093	0.663 945	0.784	0.872
4	0.261 576	0.613 551	0.715	0.885
5	0.133 131	0.566 983	0.668	0.892
10	0.003 279	0.382 089	0.564	0.908
20	0.000 001	0.173 522	0.511	0.916
21	0.000 000	0.160 351	0.508	0.916
50	0.000 000	0.016 252	0.481	0.920
78	0.000 000	0.001 782	0.474	0.922

$\sigma(78)/\sigma'(78) = 0.514$

TABLE III

n	$\lVert x_n \rVert / \lVert x_0 \rVert$	$\lVert x_n' \rVert / \lVert x_0 \rVert$	$\sigma(n)$	$\sigma'(n)$
		$a = 1.733 \quad \rho(B(a)) = 1.003289$ (x_n converges, but x_n' diverges)		
1	1.000 000	0.800 987	1.000	0.801
2	0.800 987	0.803 621	0.895	0.896
3	0.535 601	0.806 264	0.812	0.930
4	0.313 375	0.808 916	0.748	0.948
5	0.172 134	0.811 577	0.703	0.959
10	0.006 289	0.825 011	0.602	0.981
22	0.000 001	0.858 168	0.546	0.993
23	0.000 001	0.860 991	0.544	0.994
43	0.000 001	0.919 432	0.524	0.998

$\sigma(43)/\sigma'(43) = 0.525$

TABLE IV

n	$\lVert x_n \rVert / \lVert x_0 \rVert$	$\lVert x_n' \rVert / \lVert x_0 \rVert$	$\sigma(n)$	$\sigma'(n)$
		$a = 1.95 \quad \rho(B(a)) = 1.8025$ (x_n converges, x_n' diverges strongly)		
1	1.000 000	1.062	1.000	1.062
2	1.062 186	1.9	1.031	1.384
3	1.209 050	3.4	1.065	1.511
4	1.208 995	6.2	1.049	1.579
5	1.146 177	11.2	1.028	1.621
6	1.059 808	20.2	1.010	1.650
7	0.967 864	36.4	0.995	1.671
10	0.717 339	213.3	0.967	1.710
20	0.254 146	$	0.933	1.755
40	0.031 768	$	0.917	1.778
63	0.002 907	$	0.912	1.787

$\sigma(63)/\sigma'(63) = 0.511$

Acknowledgment

The author acknowledges many fruitful conversations with Professor F. T. Metcalf regarding this paper.

References

1. D. Chazin and W. Miranker, Chaotic relaxation, *Linear Algebra and Appl.* **2** (1969), 199–222.
2. J. de Pillis, Gauss–Seidel convergence for operators on Hilbert space, *SIAM J. Numer. Anal.* **10** (1973).
3. J. D. P. Donnelly, Periodic chaotic relaxation, *Linear Algebra and Appl.* **4** (1971), 117–128.
4. G. E. Forsythe and N. R. Wasow, "Finite Difference Methods for Partial Differential Equations," Wiley, New York, 1960.
5. M. Hanna, Generalized overrelaxation and Gauss–Seidel convergence on Hilbert space, *Proc. Amer. Math. Soc.* **35** (1972), 524–530.
6. H. H. Schaefer, "Topological Vector Spaces," Macmillan, New York, 1966.
7. P. Stein, The convergence of Seidel iterants of nearly symmetric matrices, *Math. Tables and Aids to Computation* **5** (1951), 237–240.
8. R. S. Varga, "Matrix Iterative Analysis," Prentice–Hall, Englewood Cliffs, N.J., 1962.
9. R. S. Varga, Factorization and normalized iterative methods, *in* "Boundary Problems in Differential Equations" (R. E. Langer, Ed.), pp. 121–142, Univ. of Wisconsin Press, Madison.
10. D. M. Young, "Iterative Solution of Large Linear Systems," Academic Press, New York, 1971.