

GENERALIZED OVERRELAXATION AND GAUSS-SEIDEL CONVERGENCE ON HILBERT SPACE

MICHAEL P. HANNA¹

ABSTRACT. We present a generalized result on Gauss-Seidel convergence for bounded linear operators on Hilbert space \mathcal{H} . From this result, we obtain overrelaxation theorems of Petryshyn [4], [5], de Pillis [1], Ostrowski [3] and Reich [6].

1. Introduction. The general problem is to solve the system $Ax=y$ for x in Hilbert space \mathcal{H} , where A is some invertible linear operator on \mathcal{H} . A splitting of A is decided upon, so that $A=A_1+A_2$, where A_1 and A_2 are bounded linear operators on \mathcal{H} (possibly depending on some parameter) and where A_1 is easy to invert. Defining the sequence $\{x_n\}$ inductively by $A_1x_n+A_2x_{n-1}=y$, where x_0 is an arbitrary but fixed initial vector in \mathcal{H} , we see that $\{x_n\}$ converges to the solution vector x (for any x_0) if and only if $(A_1^{-1}A_2)^n \rightarrow 0$ in the uniform norm topology, i.e., if and only if the spectral radius of $A_1^{-1}A_2$ is strictly less than one. In our main theorem, Theorem 4.1, we consider "apparently" selfadjoint A . Let ϕ be any norm-continuous linear transformation sending positive definite and only positive definite operators to positive definite operators on \mathcal{H} . We then split $\phi(A)=A_1+A_2$ and state conditions under which $(A_1^{-1}A_2)^n \rightarrow 0$ is equivalent to selfadjoint A being "apparently" positive definite.

2. Definitions and notation. Let $\mathcal{B}(\mathcal{H})$ represent the full algebra of bounded linear operators on Hilbert space \mathcal{H} , which has inner product $\langle \cdot, \cdot \rangle$. An operator P is **positive definite (p.d.)** if and only if $\langle Px, x \rangle \geq \delta > 0$ for all unit vectors $x \in \mathcal{H}$. Any other inner product on \mathcal{H} , $[\cdot, \cdot]$, is necessarily induced by a unique p.d. H in the following sense:

(2.1) For all $x, y \in \mathcal{H}$, $[x, y] = \langle x, Hy \rangle = \langle Hx, y \rangle$.

For $A \in \mathcal{B}(\mathcal{H})$, we denote by A^* , and $A^{[*]}$, the Hilbert space adjoints of A relative to the original inner product $\langle \cdot, \cdot \rangle$, and the H -induced inner

Received by the editors August 30, 1971.

AMS 1969 subject classifications. Primary 6510, 6540, 4730; Secondary 1525, 1535.

Key words and phrases. Gauss-Seidel convergence for operators, overrelaxation method, iterative method, splittings, renormed Hilbert space.

¹ Deceased July 1971 at the age of 28. This work derives from Mike's notes, and represents that portion of his doctoral dissertation completed at the time at the University of California, Riverside, under the direction of Professor John de Pillis. Kindly address all communication to Professor de Pillis.

product $[\cdot, \cdot]$, respectively. We say A is **H -positive definite (H -p.d.)**, or **H -selfadjoint (H -s.a.)** if A is respectively p.d., or s.a., relative to the H -induced inner product $[\cdot, \cdot]$. We sometimes denote A p.d. by writing $A > 0$.

For convenience, we list some basic properties of these two induced inner products.

(2.2 a) For all $A \in \mathcal{B}(\mathcal{H})$, $A^{[*]} = H^{-1}A^*H$.

(2.2 b) $A^{[*]} = A^*$ if and only if HA and A^*H are s.a. (relative to $\langle \cdot, \cdot \rangle$).

(2.2 c) If A commutes with p.d. H , then $A^{[*]} = A$ if and only if $A^* = A$. Let $\phi: \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ be an invertible linear map on $\mathcal{B}(\mathcal{H})$ which sends the cone of positive semidefinite operators onto itself. (Note, ϕ is automatically continuous in the uniform norm topology, since the cone has interior; Schaefer, cf. [7, p. 228].) If, moreover, ϕ sends *only* positive semidefinite operators to positive semidefinite operators, the ϕ will be called an **order isomorphism**. By $\sigma(A)$, we mean the spectrum of A , and by $\rho(A)$, we mean the spectral radius of A , i.e., $\rho(A) = \sup\{|\lambda| : \lambda \in \rho(A)\}$.

3. Stein's theorem. A central result we use is an operator version of a theorem of P. Stein [8]. In [1, Theorem 2.1] there appears a proof of Stein's theorem for operators on a (possible infinite dimensional) Hilbert space which considers the point spectrum, continuous spectrum and residual spectrum of B separately. By renorming the Hilbert space, however, we are able to present a more cohesive proof, which we do now.

THEOREM 3.1. *Suppose $H = H^*$ and B are bounded operators on Hilbert space \mathcal{H} , and that $T(H) = H - B^*HB > 0$. Then $H > 0$ if and only if $\rho(B) < 1$.*

PROOF. $[\rho(B) < 1 \Rightarrow H > 0]$. We assume $B^k \rightarrow 0$ in the uniform operator topology as $k \rightarrow \infty$. Observe that since $T(H) = H - B^*HB$ is positive definite,

$$B^*(H - B^*HB)B = B^*HB - B^{*2}HB^2 \text{ is positive semidefinite.}$$

In fact, for all k ,

$$T(H) + B^*T(H)B + \dots + B^{*k}T(H)B^k = H - B^{*k}HB^k \geq 0$$

(is positive semidefinite).

Since $B^k \rightarrow 0$ as $k \rightarrow \infty$, it is immediate that $H \geq 0$. In fact, $H > 0$ (is invertible), since for all unit vectors x , $0 < \delta \leq \langle T(H)x, x \rangle = \langle Hx, x \rangle - \langle HBx, Bx \rangle$ which implies $\langle Hx, x \rangle \geq \delta + \langle HBx, Bx \rangle \geq \delta$, that is, $H > 0$.

$[H > 0 \Rightarrow \rho(B) < 1]$. Since $H > 0$, consider $[\cdot, \cdot]$, the inner product on \mathcal{H} induced by H . (Recall $[x, y] = \langle x, Hy \rangle = \langle Hx, y \rangle$.) Now hypothesis has it

that

$$\begin{aligned} \forall x, \quad \langle x, x \rangle = 1, \quad \langle Hx, x \rangle - \langle HBx, Bx \rangle &\geq \delta > 0 \\ \Leftrightarrow \forall x, \quad \langle x, x \rangle = 1, \quad [x, x] - [Bx, Bx] &\geq \delta > 0 \\ \Leftrightarrow \forall x, \quad \langle x, x \rangle = 1, \quad [Bx, Bx]/[x, x] &\leq 1 - \delta/[x, x] \\ \Leftrightarrow \forall x, \quad \langle x, x \rangle = 1, \quad [Bx, Bx]/[x, x] &\leq 1 - \delta/\langle Hx, x \rangle \\ &\leq 1 - \delta/\|H\| \\ &= 1 - \delta', \quad \delta' > 0. \end{aligned}$$

Multiplication of x by nonzero scalars allows us to remove the restriction, now that $\langle x, x \rangle = 1$, while leaving all inequalities intact. That is,

$$\forall x \neq 0, \quad [Bx, Bx]/[x, x] < 1 - \delta', \quad \delta' > 0,$$

which is to say, $\|B\|_H < 1$, where $\|B\|_H$ is the operator norm of B relative to the inner product $[\cdot, \cdot]$. Since the norm dominates the spectral radius, we have $\rho(B) < 1$, which completes the proof.

We shall need a conjugacy result of de Pillis [1, Theorem 2.2], which we set down for convenience.

THEOREM 3.2 (DE PILLIS). *Suppose A is selfadjoint in $\mathcal{B}(\mathcal{H})$, where $A = A_1 + A_2$, and $B = -A_1^{-1}A_2$ is defined. Then there exists an invertible $X \in \mathcal{B}(\mathcal{H})$ such that*

$$(3.1) \quad X^*(A - B^*AB)X = A_1 + A_1^* - A.$$

In other words, $T(A)$ is conjugate to $A_1 + A_1^ - A$, and hence $T(A)$ is p.d. if and only if $A_1 + A_1^* - A$ is.*

4. The main result. Throughout, we allow that ϕ is an order isomorphism sending $\mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$.

THEOREM 4.1. *Let $A = A^* = D + S + Q$. We assume:*

- (1) D is s.a. in $\mathcal{B}(\mathcal{H})$.
- (2) $2D - \phi(D - S^* + Q)$ is p.d. in $\mathcal{B}(\mathcal{H})$.
- (3) $D + \phi(S)$ has an inverse in $\mathcal{B}(\mathcal{H})$.

If $\mathcal{L}(\phi) = [D + \phi(S)]^{-1}[\phi(D + Q) - D]$, then $\rho(\mathcal{L}(\phi)) < 1$ if and only if $A > 0$.

REMARK. Note that $\mathcal{L}(\phi) = -A_1^{-1}A_2$ when $\phi(A)$ is split by writing

$$(4.1) \quad \Phi(A) = [\phi(S) + D] + [\phi(D + Q) - D] = [A_1] + [A_2].$$

In the context of Gauss-Seidel convergence, we now have a generalized overrelaxation decomposition for the system $\phi(A)x = y$.

PROOF OF THEOREM 4.1. Since ϕ preserves positive semidefiniteness, ϕ also sends s.a. operators to s.a. operators. That is, $\phi(A) = \phi(A)^*$ since

$A = A^*$. Theorem 3.2 now applies to s.a.

$$\phi(A) = [\phi(S) + D] + [\phi(D + Q) - D]$$

with $A_1 = \phi(S) + D$, $A_2 = \phi(D + Q) - D$. As usual $B = -A_1^{-1}A_2$, which, by hypothesis (3), exists. We therefore conclude that invertible $X \in \mathcal{B}(\mathcal{H})$ exists so that

$$\begin{aligned} X^*(\phi(A) - B^*\phi(A)B)X &= (A_1 + A_1^*) - \phi(A) \\ &= (\phi(S) + \phi(S^*) + 2D) - (\phi(D) + \phi(S) + \phi(Q)) \\ &= 2D - \phi(D - S^* + Q) \\ &> 0 \text{ by hypothesis (2).} \end{aligned}$$

Hence, $\phi(A) > 0$. From Theorem 3.1, this condition (which is equivalent to hypothesis (2)) is equivalent to $\rho(\mathcal{L}(\phi)) < 1$. Since ϕ is an order isomorphism $A > 0 \Leftrightarrow \phi(A) > 0 \Leftrightarrow \rho(\mathcal{L}(\phi)) < 1$, which completes the proof.

5. Special cases. The first instance of our Theorem 4.1 yields a result of Petryshyn where we substitute H -p.d. for p.d., and H -s.a. for s.a. In fact, the hypotheses of [5, Theorem 1] actually agree with our own for the special scalar order isomorphisms $\phi_\omega(A) = \omega A$ for $\omega > 0$, ω in some set Ω .

THEOREM 5.1 (PETRYSHYN [5, Theorem 1]). *Given $A = D + S + Q$ where S^* denotes the adjoint of S relative to the inner product $\langle \cdot, \cdot \rangle$ on \mathcal{H} . Suppose $H > 0$, and induces the inner product $[x, y] = \langle x, Hy \rangle = \langle Hx, y \rangle$, $\forall x, y \in \mathcal{H}$. Assume S commutes with H . The following properties are further assumed:*

- (1) $D = D^{[*]}$, i.e., D is H -s.a.
 - (2) $G(\omega) = [(2 - \omega)/\omega]$. $D + S^* - Q$ is H -p.d. for all positive ω in some set Ω .
 - (3) $(D + \omega S)^{-1}$ exists in $\mathcal{B}(\mathcal{H})$ for all $\omega \in \Omega$, $\omega > 0$.
- If $\mathcal{L}(\omega) = (D + \omega S)^{-1}[(\omega - 1)D + Q]$, then $\rho(\mathcal{L}(\omega)) < 1$ if and only if A is H -p.d.

PROOF. A consequence of these hypotheses is that $A = A^{[*]}$ is H -s.a. To see this, note that

$$\begin{aligned} A &= D + S + Q \\ &= D + S + \left(\frac{2 - \omega}{\omega}\right) D + S^* - G(\omega) \text{ from hypothesis (2).} \end{aligned}$$

Now use the hypothesis that S commutes with H , so that $S^* = S^{[*]}$ (cf. (2.2 c)), along with the H -s.a. property of D (hypothesis (1)) and the H -p.d. property of $G(\omega)$ (hypothesis (2)) to see that A is H -s.a.

We contend that hypotheses (1), (2), (3) above coincide with corresponding hypotheses (1), (2), (3) of Theorem 4.1, if our $\phi(A)$ is specialized to $\phi(A) = \omega A$, and if we read H -p.d. for p.d. This completes the proof.

If we specialize the order isomorphism ϕ to be the mapping $\phi_\omega(A) = W^*AW$, for W invertible in $\mathcal{B}(\mathcal{H})$, we obtain the following operator-parameter overrelaxation result of de Pillis [1, Theorem 4.2].

THEOREM 5.2. *Let $A = D - E - E^*$, where $D > 0$. Suppose W is invertible in $\mathcal{B}(\mathcal{H})$ where $\|W\|^2 < 2$, and W commutes with D . Suppose $D - X^*EX$ has an inverse in $\mathcal{B}(\mathcal{H})$. Then for*

$$\mathcal{L}(W) = (D - W^*EW)^{-1}(X^*DX - D - X^*EX),$$

we conclude that $\rho(\mathcal{L}(W)) < 1 \Leftrightarrow A > 0$.

PROOF. In Theorem 4.1, set $S = -E$, $Q = -E^*$ and $\phi(X) = W^*XW$ for all $X \in \mathcal{B}(\mathcal{H})$. Since hypotheses (1) and (3) of Theorem 4.1 are clearly fulfilled, we only pause to verify that hypothesis (2) is satisfied. Now

$$\begin{aligned} 2D - \phi(D - S^* + Q) &= 2D - W^*(D + E^* - E^*)W \\ &= 2D - W^*DW \\ &= D(2 \cdot I - W^*W) \quad (\text{since } WD = DW). \end{aligned}$$

From our assumption $\|W\|^2 = \|W^*W\| < 2$, we see that $2 \cdot I - W^*W$ is a p.d. operator commuting with $D > 0$. Hence the product is again p.d., which says that hypothesis (2) of Theorem 4.1 obtains under the assumptions above. The theorem is proved. For completeness, we state the corollary theorems of Ostrowski (which was shown in the matrix case) and Reich.

THEOREM 5.3 (OSTROWSKI [3], [9, p. 77]). *Let $A = D - E - E^*$ where $D > 0$. Suppose $0 < \omega < 2$ and $(D - \omega E)^{-1}$ exists. If*

$$\mathcal{L}(\omega) = (D - \omega E)^{-1}((\omega - 1)D - \omega E^*),$$

we conclude that $\rho(\mathcal{L}(\omega)) < 1 \Leftrightarrow A > 0$.

PROOF. Let the order isomorphism ϕ_ω be defined by $\phi_\omega(X) = \omega^{1/2}X$ for all $X \in \mathcal{B}(\mathcal{H})$.

REMARK. It has been pointed out by the referee, that Theorem 5.3 also follows from Corollary 2 of Petryshyn [5]. That is, for $D > 0$, let $Q = S^*$. Then $G(\omega) = ((2 - \omega)/\omega)D > 0$ for $\omega \in (0, 2)$.

THEOREM 5.4 (REICH [6]). *Let $A = (a_{ij})$, $i, j = 1, 2, \dots, n$, be an $n \times n$ hermitian matrix with positive main diagonal. Let*

$$D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}) > 0,$$

and let A_1 be the $n \times n$ matrix obtained from A by setting all entries of A above the main diagonal equal to zero.

Let $A_2 = A - A_1$. Then all eigenvalues of $A_1^{-1}A_2$ lie within the unit circle if and only if $A > 0$.

PROOF. In Theorem 4.1, set ϕ equal to the identity map,

$$D = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}),$$

$S = A_1 - D$, $Q = A_2$. $\mathcal{L}(\phi)$ now becomes $A_1^{-1}A_2$ and the proof is done.

We conclude this section with a new and direct proof of a theorem which appears in F. John's book [2, p. 21], where the following is shown:

THEOREM 5.5'. Given $n \times n$ matrices A and B , where $A + B > 0$, then $A - B^* > 0 \Rightarrow \rho(A^{-1}B) < 1$.

We extend Theorem 5.5' to $A, B \in \mathcal{B}(\mathcal{H})$ for arbitrary Hilbert space \mathcal{H} and obtain the converse statement as well.

THEOREM 5.5. Given $A, B \in \mathcal{B}(\mathcal{H})$, where $A + B > 0$, then $A - B^* > 0 \Leftrightarrow \rho(A^{-1}B) < 1$.

PROOF. Let $A_0 = A + B$, $A_1 = A$, $A_2 = B$. Then from Theorem 3.2 of de Pillis,

$$\begin{aligned} T(A_0) > 0 &\Leftrightarrow (A + A^*) - (A + B) > 0 \\ &\Leftrightarrow A - B^* > 0 \\ &\Leftrightarrow \rho(A^{-1}B) < 1 \end{aligned} \quad (\text{Theorem 3.1}).$$

The last two lines establish the result.

REMARK. The referee has indicated that a stronger result than Theorem 5.5 can be obtained if it is further assumed that there exist continuously invertible operators, M and L , such that MAL is selfadjoint, and $\omega = 1$. In this event, Theorem 5.5 finds its proof in [5, Theorem 5].

REFERENCES

1. J. de Pillis, *Gauss-Seidel convergence for operators on Hilbert space*, SIAM J. Numer. Anal. **10** (1973) (to appear).
2. F. John. *Lectures on advanced numerical analysis*, Gordon and Breach, New York, 1967. MR **36** #4773.
3. A. M. Ostrowski, *On the linear iteration procedures for symmetric matrices*, Rend. Mat. e Appl. (5) **14** (1954), 140-163. MR **16**, 1155.
4. W. V. Petryshyn, *On the generalized overrelaxation method for operation equations*, Proc. Amer. Math. Soc. **14** (1963), 917-924. MR **29** #6652.
5. ———, *Remarks on the generalized overrelaxation and the extrapolated Jacobi methods for operator equations in Hilbert space*, J. Math. Anal. Appl. **29** (1970), 558-568. MR **41** #7458.

6. E. Reich, *On the convergence of the classical iterative method of solving linear simultaneous equations*, Ann. Math. Statist. **20** (1949), 448–451. MR **11**, 136.

7. H. H. Schaefer, *Topological vector spaces*, Macmillan, New York, 1966. MR **33** #1689.

8. P. Stein, *The convergence of Seidel iterants of nearly symmetric matrices*, Math. Tables and Aids to Computation **5** (1951), 237–240.

9. R. S. Varga, *Matrix iterative analysis*, Prentice-Hall, Englewood Cliffs, N. J., 1962. MR **28** #1725.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, RIVERSIDE, CALIFORNIA
92502