Convergent and Stable Operators and Their Generalization

J. J. Koliha

Department of Mathematics, University of Melbourne, Parkville, Victoria 3052, Australia

Submitted by Gian-Carlo Rota

1. Introduction

An operator will always mean a bounded linear operator on a complex Hilbert space \( H \). In the second part of the paper we show that proper contractions \( T (\| T \| < 1) \) are characterized by the form \( T = B(I + B^*B)^{-1} \) for some operator \( B \), and convergent operators \( T (\| T^n \| \to 0) \) by the form \( T = (I + B^*B)^{-1/2} B \) for some \( B \). The result has a number of consequences that include a new proof of Rota's theorem [6], and the Stein Taussky theorem [7, 8]. Strictly stable and stable operators are shown to be connected with proper contractions and convergent operators respectively by means of the real Cayley transform \( \phi(A) = (I - A)(I + A)^{-1} \). We then obtain as a consequence the theorem of Lyapunov-Taussky [8] and a theorem of Williams [10, Corollary 3]. The theorem of Lyapunov-Taussky is further generalized (Theorem 3). In Section 3 we study asymptotically convergent and semistable operators. We derive some necessary and sufficient conditions for asymptotic convergence (valid also in Banach spaces), and prove certain analogs of the Stein-Taussky theorem for asymptotically convergent operators (Theorems 5, 6, and 7). The corresponding results for semistable operators are obtained with the aid of the real Cayley transform. Most of the results of this section appear to be new. Section 4 gives applications to the iterative solution of the linear equation \( u = Tu + f \) with an asymptotically convergent \( T \) (related to the results of [1]) and of the equation \( Au = h \) with \( A \) semistable. \( N(A) \) and \( R(A) \) will mean the null space and the range of an operator \( A \), respectively. An operator \( A \) will be called positive definite on a closed subspace \( H_1 \) of \( H \) if \( (Au, u) \geq \alpha \| u \|^2 \) for some \( \alpha > 0 \) and all \( u \in H_1 \).

2. Convergent and Stable Operators

Our first result characterizes proper contractions on a Hilbert space, i.e., operators \( T \) with \( \| T \| < 1 \).
Theorem 1. An operator $T$ is a proper contraction if and only if one of the following equivalent conditions is fulfilled:

(a) For each positive definite operator $P$ there is a unique operator $B$ such that $T = B(P + B^*B)^{-1/2}$.

(b) $T = B(I + B^*B)^{-1/2}$ for some operator $B$.

Proof. It is easily seen that $\| T \| < 1$ if and only if $I - T^*T$ is positive definite. Suppose $T$ is a proper contraction and $P$ positive definite. The equation

$$I - T^*T = XPX$$

has a unique positive definite solution $Q$, namely,

$$Q = P^{-1/2}(P^{1/2}(I - T^*T)P^{1/2})^{-1}P^{-1/2}.$$  

Put $B = TQ^{-1}$. Then

$$P + B^*B = P + Q^{-1}T^*TQ^{-1} = P + Q^{-1}(I - QPQ)Q^{-1} = Q^{-2},$$

and

$$B(P + B^*B)^{-1/2} = TQ^{-1}Q = T.$$  

If $T = C(P + C^*C)^{-1/2}$ for some $C$,

$$I - T^*T = (P + C^*C)^{-1/2} P(P + C^*C)^{-1/2},$$

i.e., $(P + C^*C)^{-1/2}$ is a positive definite solution of the equation $I - T^*T = XPX$, and $(P + C^*C)^{-1/2} = Q$. Then

$$C - T(P + C^*C)^{1/2} = TQ^{-1} = B.$$  

(a) obviously implies (b). Suppose $T = B(I + B^*B)^{-1/2}$ for some $B$. We observe that $I - T^*T = (I + B^*B)^{-1}$; hence, $I - T^*T$ is positive definite and $\| T \| < 1$.

We say that an operator $T$ is convergent if its spectrum $\sigma(T)$ is contained in the open unit disc $D$. Before deriving an analog of Theorem 1 for convergent operators, we mention some well known necessary and sufficient conditions that $T$ be convergent.

I. The series $\sum_{n=0}^{\infty} T^n$ converges in norm.

II. The series $\sum_{n=0}^{\infty} \| T^n \|$ converges.

III. The series $\sum_{n=0}^{\infty} T^{*n}WT^n$ converges in norm for each $W$.

IV. The series $\sum_{n=0}^{\infty} \| T^{*n}WT^n \|$ converges for each $W$.

V. $T^n \to 0$ in norm.
VI. \( T^n WT^n \to 0 \) in norm for each \( W \).

VII. \( \| T^p \| < 1 \) for some \( p > 1 \).

VIII. \( r(T) = \lim_n \| T^n \|^{1/n} < 1 \).

We can first establish the implications I \( \to \) V \( \to \) VII \( \to \) VIII \( \to \) I. V and VI are clearly equivalent. II, III, and IV follow from VIII when the Cauchy root test is applied, and each of II, III, and IV implies V.

**Theorem 2.** An operator \( T \) is convergent if and only if one of the following equivalent conditions is satisfied:

(a) For each positive definite operator \( P \) there is a unique operator \( B \) such that \( T = (P + B^*B)^{-1/2} B \).

(b) \( T = (I + B^*B)^{1/2} B \) for some \( B \).

**Proof.** Suppose that \( T \) is convergent and \( P \) positive definite. Define a sequence of operators by \( B_{n+1} = (P + B_n^*B_n)^{1/2} T \) starting with an arbitrary \( B_0 \). From the definition of \( B_n \) it follows that \( \{B_n\} \) converges in norm whenever \( \{B_n^*B_n\} \) converges in norm. Since \( B_{n+1}^*B_{n+1} - B_n^*B_n = T^n W T^n \) with \( W = T^* PT + T^* B_n^*B_n T - B_n^*B_n \), the series \( \sum_{n=0}^{\infty} (B_{n+1}^*B_{n+1} - B_n^*B_n) \) converges in norm according to III. Hence, \( \{B_n^*B_n\} \) converges, and \( B_n \to B \) in norm, where

\[
B = \left( \sum_{n=0}^{\infty} T^n PT^n \right)^{1/2} T, \tag{1}
\]

\( B = (P + B^*B)^{1/2} T \). If \( C \) is another operator satisfying \( C = (P + C^*C)^{1/2} T \), then the sequence \( C_{n+1} = (P + C_n^*C_n)^{1/2} T \) starting with \( C_0 = C \) converges to \( R \), while \( C_n = C \) for all \( n \). Hence, \( C = R \). \( R = (P + R^*R)^{1/2} T \) is equivalent to \( T = (P + B^*B)^{-1/2} B \), since \( P + B^*B \) is positive definite. (a) obviously implies (b). Let \( T = (I + B^*B)^{-1/2} B \) for some \( B \). Put \( S = (I + B^*B)^{-1/2} \). Then \( T = S(BS) S^{-1} \), where \( BS \) is a proper contraction by Theorem 1.

\[
\| T^n \| = \| S(BS)^n S^{-1} \| \leq \| S \| \| S^{-1} \| \| BS \|^n \to 0
\]
as \( n \to \infty \), and \( T \) is convergent by V.

A result related to our Theorem 1 and 2 has been obtained by Taussky who showed that every convergent matrix is unitarily similar to a matrix \( FA^{-1/2} \), where \( A \) is a diagonal matrix \( \geq I \), and \( FF^* = A - I \) [8, Theorem 7].

Note that \( FA^{-1/2} \) is a proper contraction.

Let us remark that Theorems 1 and 2 can be interpreted as theorems about the transformations

\[
F_I(X) = T(I + X^*X)^{1/2}, \quad G_I(X) = (I + X^*X)^{1/2} T.
\]
CONVERGENT AND STABLE OPERATORS

$F_T$ (resp. $G_T$) has a fixed point if and only if $T$ is a proper contraction (resp. convergent); the fixed point is then unique.

If $T$ is convergent, $T = (I + B^*B)^{-1/2} B$, where $B$ is defined by (1) with $P = I$. $T$ is similar to the proper contraction $T_1 = B(I + B^*B)^{-1/2}$ as

$$T = (I + B^*B)^{-1/2} T_1 (I + B^*B)^{1/2}. \quad (2)$$

This result sharpens a theorem due to Rota [6, Theorem 2], as (2) gives an explicit formula for the similarity.

**Corollary 1.** An operator $T$ is convergent if and only if $T$ is similar to a proper contraction.

Also the following result is a consequence of Theorem 2.

**Corollary 2.** An operator $T$ is convergent if and only if one of the following equivalent conditions is satisfied:

(a) The equation $X - T^*XT = P$ has a unique nonnegative solution for each positive definite $P$.

(b) The equation $X - T^*XT = I$ has at least one nonnegative solution.

**Proof.** If $T$ is convergent and $P$ positive definite, $T = (P + B^*B)^{-1/2} B$ with a uniquely determined $B$. The operator $X = P + B^*B$ is the unique positive definite solution of $X - T^*XT = P$. The implication (a) $\rightarrow$ (b) is obvious. Suppose that $X$ is a nonnegative operator satisfying $X - T^*XT = I$. $X$ is in fact positive definite as $(Xu, u) \geq ||u||^2$ for all $u \in H$. Put $B = X^{1/2}T$. Then $(I + B^*B)^{-1/2} B = (I + T^*XT)^{-1/2} B = X^{-1/2}X^{1/2}T = T$, and $T$ is convergent by Theorem 2.

Stein [7] proved that a matrix $T$ satisfies $T^n \rightarrow 0$ if and only if there is a positive definite matrix $Q$ such that $Q - T^*QT$ is also positive definite; Taussky [8] showed that the equation $X - T^*XT = P$ is uniquely solvable with a positive definite $X$ for each positive definite $P$ whenever $T^n \rightarrow 0$. A considerable generalization of Taussky's result was given by Redheffer [5]. In the sequel, Corollary 2 will be referred to as the theorem of Stein-Taussky.

An operator $A$ will be called strictly stable if there exists $\alpha > 0$ such that

$$\text{Re}(Ax, x) \geq \alpha ||x||^2 \quad (3)$$

for all $x \in H$. From this definition it follows that the closed numerical range of a strictly stable operator $A$ is contained in the open right half plane $H^+$, and, consequently, $\sigma(A) \subseteq H^+$. We show that strictly stable operators are related to proper contractions via the real Cayley transform

$$\phi(A) = (I - A)(I + A)^{-1}.$$
Lemma 1. If $A$ is an operator with $(-1) \notin \sigma(A)$, then $A$ is strictly stable if and only if its Cayley transform $T = (I - A)(I + A)^{-1}$ is a proper contraction.

Proof. If $(-1) \notin \sigma(A)$, $I + A$ is invertible on $H$ and $T = \phi(A)$ exists. Moreover, $I + T = 2(I + A)^{-1}$ is invertible on $H$, and $\| (I + T)y \| \geq c \| y \|$ for some $c > 0$ and all $y \in H$. For each $x \neq 0$ define $y = (I + T)^{-1}x$. Then $y \neq 0$. Using the inverse Cayley transform $A = (I - T)(I + T)^{-1}$ ($\phi$ is an involution), we obtain

$$\text{Re} \left( \frac{(Ax, x)}{\| x \|^2} \right) = \text{Re} \left( \frac{(I - T)y, (I + T)y}{\| (I + T)y \|^2} \right) = \frac{\| y \|^2 - \| Ty \|^2}{\| y \|^2 + \| Ty \|^2}.$$ 

If $\| T \| < 1$, $\text{Re}(Ax, x) \geq (1 - \| T \|)(1 + \| T \|)^{-1}\| x \|^2$, and $A$ is strictly stable. If, conversely, $\| T \| \geq 1$, there is a sequence $\{ y_n \}$ of unit vectors with $\| Ty_n \| \to -\| T \|$. For $x_n - (I + T)y_n$ we have

$$\text{Re} \left( \frac{(Ax_n, x_n)}{\| x_n \|^2} \right) = \frac{\| y_n \|^2 - \| Ty_n \|^2}{\| (I + T)y_n \|^2} \leq \frac{\| y_n \|^2 - \| Ty_n \|^2}{c^2 \| y_n \|^2} = \frac{1 - \| Ty_n \|^2}{c^2};$$

hence, the sequence $\text{Re}(Ax_n, x_n) \| x_n \|^{-2}$ either contains negative terms or comes arbitrarily close to 0, which shows that $A$ is not strictly stable.

If $\sigma(A)$ is contained in $H^+$, the operator $A$ is called (positive) stable. The usual definition of a (negative) stable matrix $A$ requires that the eigenvalues of $A$ be contained in the open left half plane, which condition plays an important role in the stability theory of differential equations (cf. S. Barnett and C. Storey, "Matrix Methods in Stability Theory," Nelson, Edinburgh, 1970). Clearly, $A$ is negative stable if and only if $-A$ is positive stable. Redheffer [5] defines a stable element $A$ in a unital ring with a valuation $| \cdot |$ by the requirement $\| (I - A)^{-1}(I + A) \| < 1$. In the case when the ring is the $B^*$-algebra of all linear bounded operators on $H$, this definition expresses the fact that $-A$ is strictly stable according to our terminology.

Also the transition from convergent to stable operators can be made with the aid of the Cayley transform. The following lemma follows from the fact that the transformation $\phi(\lambda) = (1 - \lambda)/(1 + \lambda)$ of the complex plane is an involution ($\phi \circ \phi = \text{identity}$) mapping the open unit disc $D$ onto $H^+$, and from the spectral mapping theorem.

Lemma 2. If $A$ is an operator with $(-1) \notin \sigma(A)$, then $A$ is stable if and only if its Cayley transform $T = (I - A)(I + A)^{-1}$ is convergent.

The foregoing results enable us to give an elementary proof a theorem obtained by Williams [10, Theorem 4, (1) and (2)].

Corollary 3. An operator $A$ is stable if and only if $A$ is similar to a strictly stable operator.
CONVERGENT AND STABLE OPERATORS

Proof. Let \( \phi(\lambda) = (1 - \lambda)/(1 + \lambda) \). If \( A \) is stable, \( T = \phi(A) \) is convergent, and from Corollary 1 it follows that for some invertible \( S \) the operator \( S^{-1}TS \) is a proper contraction. Then \( S^{-1}AS = S^{-1}\phi(T)S = \phi(S^{-1}TS) \), and \( S^{-1}AS \) is strictly stable by Lemma 1. Conversely, a strictly stable operator is also stable, and stability is preserved under similarity.

When \( (-1) \notin \sigma(T) \), the equation \( X - T^*XT = P \) is transformed into the equivalent equation \( XA + A^*X = 2(I + A^*)^{-1}P(I + A)^{-1} \) with \( A = \phi(T) \) (cf. [8]). This observation enables us to prove that the theorem of Stein-Taussky is equivalent to the following theorem of Lyapunov-Taussky proved originally for matrices [8]. The operator analog of the theorem is due to Williams [10, Theorem 5].

**Corollary 4.** An operator \( A \) is stable if and only if one of the following equivalent conditions is fulfilled:

(a) The equation \( XA + A^*X = P \) has a unique nonnegative solution for each positive definite \( P \).

(b) The equation \( XA + A^*X = I \) has at least one nonnegative solution.

Proof. If \( A \) is stable, (a) follows from Corollary 2 and Lemma 2. (a) obviously implies (b). We show that a nonnegative \( X \) satisfying \( XA + A^*X = P \) with a positive definite \( P \) is in fact positive definite. It is enough to prove that \( X \) is invertible on \( H \). \( XA \) is invertible as

\[
2\text{Re}(XAu, u) = (Pu, u)
\]

for all \( u \in H \).

Then \( XAL = I = L^*A^*X \) for some \( L \), which shows that \( X \) is invertible on \( H \) [10]. In view of Corollary 2 and Lemma 2, the proof that (b) implies the stability of \( A \) reduces to establishing \( (-1) \notin \sigma(A) \). Put \( M = X(I + A) \). Then

\[
\text{Re}(Mu, u) = (Xu, u) \quad \text{Re}(XAu, u) = (Xu, u) + \frac{1}{2} ||u||^2 \geq (\alpha + \frac{1}{2}) ||u||^2
\]

for all \( u \in H \), where

\[
\alpha = \inf_{||u||=1} (Xu, u) > 0.
\]

Hence, \( M \) is invertible on \( H \) and so is \( I + A \).

A generalization of Lyapunov's theorem was given by Ostrowski and Schneider [3] for matrices and by Williams [10] for operators. In Theorem 7 of [10] Williams shows that if \( 0 \notin \sigma(A) + \sigma(A^*) \), then the equation \( XA + A^*X = Y \) is uniquely solvable for each \( Y \), and that \( X \) is self-adjoint (resp. invertible) if \( Y \) is self-adjoint (resp. positive definite). This leads to the following generalization of the theorem of Stein-Taussky.
Theorem 3. Suppose $1 \notin \sigma(T) \sigma(T^*)$. Then:

(a) The equation $X - T^*XT = Z$ is uniquely solvable for each $Z$.

(b) $X$ is self-adjoint if and only if $Z$ is self-adjoint.

(c) If $Z$ is positive definite, $X$ is invertible.

Proof. By definition, $\sigma(T) \sigma(T^*) = \{\lambda \kappa \mid \lambda \in \sigma(T), \kappa \in \sigma(T^*)\}$. If $1 \notin \sigma(T) \sigma(T^*)$, $(-1) \notin \sigma(T)$. The equation $X - T^*XT = Z$ transforms into the equivalent equation $XA + A^*X = Y$ with $A = \phi(T)$ and $Y = 2(I + A^*)^{-1}Z(I + A)^{-1}$. Furthermore,

$$\sigma(A) + \sigma(A^*) = \left\{ \frac{1 - \lambda}{1 + \lambda} + \frac{1 - \kappa}{1 + \kappa} \right\} \lambda \in \sigma(T), \kappa \in \sigma(T^*)$$

Since $(1 - \lambda)/(1 + \lambda) + (1 - \kappa)/(1 + \kappa) = 2(1 - \lambda\kappa)/(1 + \lambda)(1 + \kappa))$, $0 \notin \sigma(A) + \sigma(A^*)$ as $1 \notin \sigma(T) \sigma(T^*)$. Hence, the result.

In closing the section we remark that $1 \notin \sigma(T) \sigma(T^*)$ is in fact equivalent to the assumption that the equation $X - T^*XT = Z$ is uniquely solvable for each (nonnegative) $Z$. For this purpose consider the operator $K$ defined on $B(H)$ by $K(X) = T^*XT$. Then it follows from the generalized Kleinecke’s theorem (cf. G. Lumer and M. Rosenblum, Proc. Amer. Math. Soc. 10 (1959), 32–41) that $\sigma(K) = \sigma(T) \sigma(T^*)$. Let us also remark that if $\sigma(T)$ does not meet the unit circle, then it can be deduced from Theorem 6 in [10] that there exists a self-adjoint invertible operator $X$ such that $X - T^*XT$ is positive definite.

3. Asymptotically Convergent and Semistable Operators

An operator $T$ is called asymptotically convergent if its spectrum $\sigma(T)$ lies in $D \cup \{1\}$ (where $D$ is the open unit disc), and $1$ is a simple pole of $(\lambda I - T)^{-1}$ whenever $1 \in \sigma(T)$. Clearly, convergent operators are asymptotically convergent. Let us consider an asymptotically convergent operator $T$ with $1 \in \sigma(T)$. If $C_1$ is a counterclockwise circle with center $1$ whose exterior contains $\sigma(T) - \{1\}$, the operator

$$P = \frac{1}{2\pi i} \int_{C_1} (\lambda I - T)^{-1} d\lambda$$

is the projection of $H$ onto $N(I - T)$ in the direction of $R(I - T)$, $H = N(I - T) \oplus R(I - T)$ with $R(I - T)$ closed, and $T$ is completely reduced by the two subspaces occurring in this direct sum [9, p. 306]. Writing $T_N$ and $T_R$ for the restrictions of $T$ to $N(I - T)$ and $R(I - T)$ respectively,
we have \( T = T_N \oplus T_R \); \( T_N \) is the identity operator \( I_N \) on \( N(I - T) \), and \( \sigma(T_R) \subset D \), so that \( T_R \) is convergent. Then \( T^n = I_N \oplus T_R^n \to I_N \oplus 0 = P \) in norm as \( n \to \infty \), that is
\[
\lim_{n \to \infty} T^n = P \quad \text{in norm.} \tag{5}
\]

**Theorem 4.** An operator \( T \) is asymptotically convergent if and only if \( T^n \to P \) in norm for some operator \( P \).

**Proof.** \( T^n \to 0 \) in norm is necessary and sufficient that \( T \) be convergent. Necessity has been also established for an asymptotically convergent operator \( T \) with \( 1 \in \sigma(T) \). To prove sufficiency, we suppose that \( T^n \to P \) in norm, \( P \neq 0 \). We show that
\[
PT = TP = P, \quad P^2 = P. \tag{6}
\]
Indeed \( PT = TP = \lim_n T^{n+1} = P \). Consequently, \( PT^n = P \) for all \( n \geq 0 \), and \( P^2 = \lim_n PT^n = P \). Since \( P \) is a projection, \( H = R(P) \oplus N(P) \). We prove that
\[
R(P) = N(I - T), \quad N(P) = R(I - T). \tag{7}
\]

The identity \( (I - T)P = 0 \) implies \( R(P) \subset N(I - T) \). If \( x \in N(I - T) \), \( x = T^nx \to Px \), and \( N(I - T) \subset R(P) \). Similarly, \( P(I - T) = 0 \) proves \( R(I - T) \subset N(P) \), and, consequently, \( R(I - T)^- \subset N(P) \) as \( N(P) \) is closed. (The bar indicates closure.) If \( x \in N(P) \), then
\[
x = x - Px = \lim_n (I - T^n)x \Rightarrow (I - T)(x + Tx + \cdots + T^{n-1}x),
\]
which shows that \( x \in R(I - T) \). So far we have proved that
\[
H = N(I - T) \oplus R(I - T)^-. \tag{\*}
\]

To establish that \( R(I - T) \) is closed, we observe that
\[
\|(T - P)^n\| = \|T^n - P\| \to 0
\]
as \( n \to \infty \), so that \( T - P \) is convergent. Then \( I - T + P \) is invertible on \( H \), \( R(I - T + P) = H \), and each element \( x \in H \) can be written in the form \( x = Pu + (I - T)u \), where \( Pu \in N(I - T) \) and \( (I - T)u \in R(I - T) \). When \( x \in R(I - T)^- \), \( Pu = 0 \) in view of the decomposition (\*), and
\[
x = (I - T)u \in R(I - T).
\]

According to [9, p. 310], 1 is a simple pole of \( (\lambda I - T)^{-1} \) whenever \( H = N(I - T) \oplus R(I - T) \) with \( R(I - T) \) closed. Finally, if \( T_N \) and \( T_R \)
The restrictions of \( T \) to \( N(I - T) \) and \( R(I - T) \), respectively, then \( T = T_N \oplus T_R \), and from \( T^n \to I_N \oplus 0 \) it follows that \( T_R^n \to 0 \) in norm, i.e., \( \sigma(T_R) \subset D \). The proof is completed by observation that

\[
\sigma(T) = \sigma(T_R) \cup \sigma(T_N) \subset D \cup \{1\}.
\]

**Remark 1.** If \( \lim_{n\to\infty} T^n = P \) for some \( P \neq 0 \) in the sense of the strong or even weak operator topology, an analysis of the foregoing proof discloses that in this case \( H = N(I - T) \oplus R(I - T) \) and that \( P \) is the projection of \( H \) onto \( N(I - T) \) in the direction of \( R(I - T) \).

**Remark 2.** Petryshyn [4] studied the convergence of \( \{T^n\} \) in norm for an operator \( T = I - \beta K \) with \( K \) self-adjoint. Theorem 2 of [4] states that \( \{T^n\} \) converges in norm to the orthogonal projector \( P \) of \( H \) onto \( N(K) \) if and only if \( K \) is positive definite on \( N(K)^\perp \) and \( 0 < \beta < 2 \|K\|^{-1} \). In Corollary 2 of the same paper it is proved that if \( \{T^n\} \) converges in norm, the range \( R(K) \) is closed.

We note that the equation (5) is equivalent to

\[
\sum_{n=0}^{\infty} T^n(I - T) = I - P
\]

in view of the identity

\[
I - T^n = (I - T)(I + T + \cdots + T^{n-1}). \tag{8}
\]

Hence, an operator \( T \) is asymptotically convergent if and only if the series \( \sum_{n=0}^{\infty} T^n(I - T) \) converges in norm. The following corollary gives yet another characterization of asymptotic convergence.

**Corollary 5.** An operator \( T \) is asymptotically convergent if and only if the restriction \( T_1 \) of \( T \) to the subspace \( R(I - T) \) is convergent.

**Proof.** We have already proved necessity. For the proof of sufficiency assume that \( T_1 \) is convergent on \( H_1 = R(I - T) \). (A correct assumption as \( H_1 \) is invariant under \( T \).) For each \( x \in H \) the series \( \sum_{n=0}^{\infty} T^n(I - T)x \) converges in norm, and (8) implies that also \( \{T^n x\} \) converges in norm for each \( x \in H \). According to Remark 1, \( H = N(I - T) \oplus R(I - T) \), \( P \) projects \( H \) onto \( N(I - T) \) (\( P = \lim_{n\to\infty} T^n \) in the strong operator topology), and \( I - P \) projects \( H \) onto \( R(I - T) \). Then

\[
T^n - P = T^n - T^n P = T^n (I - P),
\]

and

\[
\|T^n - P\| \leq \|T^n\| \|I - P\|.
\]
The last inequality shows that $T^n \to P$ in norm. Hence, the conclusion by Theorem 4.

All the results derived so far in this section are also valid in a Banach space. In the sequel we show that the theorem of Stein-Taussky can be extended to asymptotically convergent operators on Hilbert spaces in at least two ways. A somewhat more general version of the following theorem was derived in [2]. However, the proof given in [2] contains a slight mistake which will be rectified here.

**Theorem 5.** An operator $T$ is asymptotically convergent if and only if there exists an operator $X$ such that both $X$ and $X - T^*XT$ are positive definite on the subspace $R(I - T)$.\(^{-1}\).

**Proof.** Suppose that $T$ is asymptotically convergent. Then $T = I_N \oplus T_R$, where $I_N$ resp. $T_R$ is the restriction of $T$ to $N(I - T)$ resp. $R(I - T)$, while $R(I - T)$ is closed. $H_R = R(I - T)$ is itself a Hilbert space on which $T_R$ is convergent. According to Corollary 2, there is a positive definite operator $X_R$ on $H_R$ for which $X_R - (T_R)^* X_R T_R$ is positive definite. Take $X = S_N \oplus X_R$, where $S_N$ is an arbitrary operator on $N(I - T)$. Note that in general $(T_R)^* \neq (T^*)_R$, however

$$((T_R)^* u, v) = (u, T_R v) = (u, T v) = (T^* u, v) = ((T^*)_R u, v)$$

for all $u, v \in H_R$. Hence, for all $u \in H_R$

$$((X - T^*XT) u, u) = ((X_R - (T_R)^* X_R T_R) u, u),$$

and $X - T^*XT$ is positive definite on $H_R$. (In [2] an incorrect conclusion is drawn that the restriction of $X - T^*XT$ to $H_R$ agrees with $X_R - (T_R)^* X_R T_R$.) To prove the converse, suppose that $Y = X - T^*XT$ is positive definite on $H_1 = R(I - T)$ for some operator $X$ positive definite on $H_1$, i.e., that

$$(X u, u) \geq a \| u \|^2, \quad (Y u, u) \geq b \| u \|^2,$$

for all $u \in H_1$ and some positive real numbers $a$ and $b$. This time $H_1$ is not necessarily invariant under $X$, so that Corollary 2 cannot be applied. We use the fact that for some $\kappa$, $0 < \kappa < 1$,

$$\kappa (X u, u) \leq (Y u, u) \quad \text{for all } u \in H_1.$$

(Take for instance $\kappa = \min(\frac{1}{\kappa}, b \| X \|^{-1})$.) Furthermore,

$$(T^*XT u, u) = (X u, u) - (Y u, u) \leq (1 - \kappa) (X u, u) \quad \text{for all } u \in H_1,$$
and
\[(T^*nXT^n u, u) \leq (1 - \kappa)^n (X u, u),\]
for all \(u \in H_1\) and all \(n \geq 1\). Writing \(T_1\) for the restriction of \(T\) to \(H_1\), we obtain
\[\| T_1^n u \|^2 \leq a^{-1}(XT^n u, T^n u) \leq a^{-1}(1 - \kappa)^n \| X \| \| u \|^2, \quad u \in H_1,\]
which implies \(\| T_1^n \| \leq a^{-1}(1 - \kappa)^n \| X \|\), where the norm of \(T_1^n\) is taken on \(H_1\). Hence, \(\| T_1^n \| \to 0\) as \(n \to \infty\), \(T_1\) is convergent on \(H_1\) according to \(V\), and \(T\) is asymptotically convergent by virtue of Corollary 5.

Here is another version of the Stein-Taussky theorem for asymptotically convergent operators.

**Theorem 6.** Let \(T\) be asymptotically convergent. The equation
\[X - T^*XT - (I - T^*) W(I - T)\]
has a solution for each \(W\). If \(W\) is positive definite on \(R(I - T)\), so is \(X\). Any two solutions of (9) coincide on \(R(I - T)\).

**Proof.** For a given operator \(W\) define
\[Q = \sum_{n=0}^{\infty} (I - T^*) T^*nWT^n(I - T).\]

To establish the convergence in norm of the series we observe that for each \(x \in H\) not in \(N(I - T)\) we have
\[
\frac{\| T^n(I - T)x \|}{\| x \|} = \frac{\| T^n(I - T)x \|}{\| (I - T)x \|} \cdot \frac{\| (I - T)x \|}{\| x \|} \leq \| T_R^n \| \| I - T \|,
\]
so that \(\| T^n(I - T)\| \leq \| T_R^n \| \| I - T \|\), where \(T_R\) is the restriction of \(T\) to \(R(I - T)\). Since, by assumption, \(T_R\) is convergent on the Hilbert space \(H_R = R(I - T)\), the series \(\sum_{n=0}^{\infty} \| T_R^n \|\) converges in accordance with 11. Then
\[
\| (I - T^*) T^*nWT^n(I - T)\| \leq M \| I - T \|^2 \| W \| \| T_R^n \|
\]
for \(n \geq 1\), where \(M\) is a positive constant with \(\| T^*n \| = \| T^n \| \leq M\) whose existence is guaranteed by the principle of uniform boundedness. This proves that the series in (10) converges in norm. From (10) it follows that \(T^*QT = Q - (I - T^*) W(I - T)\). Suppose that \(W\) is positive definite on \(H_R\), i.e., that \((W u, u) \geq \alpha \| u \|^2\) for some \(\alpha > 0\) and all \(u \in H_R\). Since \(T_R\)
is convergent, \( I_R - T_R \) is invertible on \( H_R \), and \( \| (I - T) u \| \geq c \| u \|^2 \) for some \( c > 0 \) and all \( u \in H_R \). From (10) we obtain that

\[
(Q u, u) \geq (W(I - T) u, (I - T) u) \geq \alpha c^2 \| u \|^2 \quad \text{for all } u \in H_R,
\]

and \( Q \) is positive definite on \( H_R \). If \( X \) is another solution of \( X - T^*XT = C \) with \( C = (I - T^*) W(I - T) \), then

\[
X = C + T^*CT + \cdots + T^nCT^n + T^{n+1}XT^{n+1}.
\]

The right side converges in norm to \( Q + P^*XP \), where \( P \) is the uniform limit of \( \{T^n\} \). Since \( Pu = 0 \) for each \( u \in H_R \), \( Xu = Qu \) for all \( u \in H_R \).

Theorem 6 admits a converse in the case when \( T \) is compact.

**Theorem 7.** A compact operator \( T \) is asymptotically convergent if and only if there exists an operator \( Q \) positive definite on \( R(I - T) \) such that

\[
Q - T^*QT = (I - T^*) (I - T).
\]

**Proof.** The "only if" part follows from the preceding theorem. Let \( Q \) be an operator satisfying (11) and positive definite on \( R(I - T) \). First we show that \( N((I - T)^2) = N(I - T) \). Suppose that \( (I - T)^2 x = 0 \), and put \( y = (I - T)x \). Then

\[
(Qy, y) = ((Q - T^*QT)y, x) = ((I - T^*) (I - T)y, x) = 0,
\]

and \( y = (I - T)x = 0 \). By a general theorem on compact operators [9, p. 279], \( N((I - T)^2) = N(I - T) \) implies \( R((I - T)^2) = R(I - T) \), and \( H = N(I - T) \oplus R(I - T) \) with \( R(I - T) \) closed. Hence, \( 1 \) is a simple pole of \( (\lambda I - T)^{-1} \) provided that \( N(I - T) \neq \{0\} \). The operator \( (I - T^*) (I - T) \) is positive definite on \( R(I - T) \) as \( I_R - T_R \) is invertible on \( R(I - T) \). As in the proof of Theorem 5 we find that \( \| T_R^n \| \to 0 \) as \( n \to \infty \). Hence, \( \sigma(T) = \sigma(T_N) \cup \sigma(T_R) \subset \{1\} \cup D \).

The concept of stability for operators can be generalized in the following way. An operator \( A \) is called *semistable* if its spectrum \( \sigma(A) \) lies in \( H^+ \cup \{0\} \) (\( H^+ \) is the open right half plane), and \( 0 \) is a simple pole of \( (\lambda I - A)^{-1} \) whenever \( 0 \in \sigma(A) \). As we may expect, the real Cayley transform provides a link between asymptotically convergent and semistable operators.

**Lemma 3.** Let \( A \) be an operator with \((-1) \notin \sigma(A) \). \( A \) is semistable if and only if \( T = (I - A)(I + A)^{-1} \) is asymptotically convergent.
Proof. If \((-1) \notin \sigma(A)\), the Cayley transform

\[ T = \phi(A) = (I - A)(I + A)^{-1} \]

of \(A\) is defined, and

\[ I - T = 2A(I + A)^{-1} = 2(I + A)^{-1}A. \]

Hence,

\[ N(I - T) = N(A), \quad R(I - T) = R(A). \] (12)

The result then follows from the fact that the transformation

\[ \phi(\lambda) = (1 - \lambda)/(1 + \lambda) \]

maps \(D \cup \{1\}\) onto \(H^+ \cup \{0\}\) (and vice versa), the spectral mapping theorem, and Theorem 5.8-D in [10, p. 310].

If \(A\) is a semistable operator with \(0 \in \sigma(A)\), \(H = N(A) \oplus R(A)\) with \(R(A)\) closed. The projection \(P\) of \(H\) onto \(N(A)\) in the direction of \(R(A)\) is given by the formula

\[ P = \frac{1}{2\pi i} \int_{C_0} (\lambda I - A)^{-1} d\lambda, \] (13)

where \(C_0\) is a counterclockwise circle with center in the origin whose exterior contains \(\sigma(A) - \{0\}\), or alternatively by the formula (4) with \(T = \phi(A)\).

Hence, also \(T^n \to P\) in norm. If \(A_N\) and \(A_R\) are the restrictions of \(A\) to \(N(A)\) and \(R(A)\), respectively, \(A = A_N \oplus A_R = 0_N \oplus A_R\), where \(A_R\) is stable.

In the following theorem we may assume that \((-1) \notin \sigma(A)\). Indeed if \((-1)\) happens to be in \(\sigma(A)\), we consider an operator \(sA\) for some \(s > 0\) with \(-1/s \notin \sigma(A)\). (Such \(s\) always exists as \(A\) is bounded.) Then \((-1) \notin \sigma(sA)\), and \(A\) is semistable if and only if \(sA\) is semistable. With this convention Theorem 8 is a straightforward application of Lemma 3 to Corollary 5 and Theorem 5.

**Theorem 8.** An operator \(A\) is semistable if and only if one of the following equivalent conditions is satisfied:

(a) The restriction \(A_N\) of \(A\) to \(R(A)^{-}\) is stable.

(b) There exists an operator \(X\) such that both \(X\) and \(XA + A*X\) are positive definite on \(R(A)^{-}\).

The following result is obtained by applying Lemma 3 to Theorem 6.

**Theorem 9.** If \(A\) is semistable, the equation

\[ XA + A*X = 2A*WA \] (14)

has a solution for each \(W\). Any two solutions coincide on \(R(A)\).
In order to derive an analog of Theorem 7 for semistable operators, we suppose that I - A is compact and that Q is an operator positive definite on \( R(A) \) such that \( QA + A^*Q = 2A^*A \). First, we show that \( N(A^2) = N(A) \).

Put \( y = Ax \), where \( A^2x = 0 \). Then

\[
(Qy, y) = ((QA + A^*Q)y, x) = 2(A^*Ay, x) = 0,
\]

so that \( y = 0 \). Consequently, \( H = N(A) \oplus R(A) \) with \( R(A) \) closed as \( A = I - (I - A) \) with \( I - A \) compact. Furthermore, \((-1) \in \sigma(A)\) if and only if 2 is an eigenvalue of \( I - A \), i.e., if \( Ax = -x \) for some \( x \neq 0 \). Suppose that \( Ax = -x \) for some \( x \in H \). Then \( x = u + v \) with \( u \in N(A) \) and \( v \in R(A) \), and from the equality \( Av = -u - v \) it follows that \( u = 0 \). Hence, \( x \in R(A) \).

Finally, \( (QAx, x) + (A^*Qx, x) = -2(Qx, x) = 2 \| x \|^2 \), and \( x = 0 \) as \( Q \) is positive definite on \( R(A) \). This proves that \((-1) \notin \sigma(A)\). The operator \( T = (I - A)(I + A)^{-1} \) exists and is compact. Thus, we can apply Lemma 3 to Theorem 7 to obtain the following.

**Theorem 10.** If \( I - A \) is compact, then \( A \) is asymptotically stable if and only if there exists an operator \( Q \) positive definite on \( R(A) \) such that

\[
QA + A^*Q = 2A^*A.
\]

**4. Iterative Solution of Linear Equations**

We first consider an approximate solution of the equation

\[
u = Tu + f
\]

by means of the Picard iterations

\[
x_{n+1} = Tx_n + f, \quad n = 0, 1, 2, ..., \quad (16)
\]

assuming that \( T \) is asymptotically convergent. (16) can be also written in the form

\[
x_n = T^n x_0 + \sum_{k=0}^{n-1} T^k f.
\]

The sequence \( \{T^n x_0\} \) converges to \( Px_0 \) for each \( x_0 \in H \), where \( P \) is the projector of \( H \) onto \( N(I - T) \) in the direction of \( R(I - T) \). The series \( \sum_{k=0}^{\infty} T^k f \) converges to an element \( x^* \) if and only if \( f \in R(I - T) \). We observe that \( x^* \) is the unique solution of (15) on \( R(I - T) \). (A much more general result was
obtained by Browder and Petryshyn [1], where the sequence \( \{T^n\} \) was assumed to converge in the sense of the strong operator topology.) This proves the following theorem.

**THEOREM 11.** (a) If \( T \) is asymptotically convergent with \( 1 \in \sigma(T) \), the sequence \( \{x_n\} \) defined by (16) converges in norm for any \( x_0 \in H \) if and only if \( f \in R(I - T) \). The limit \( x \) of \( \{x_n\} \) in this case is a solution of (15) of the form

\[
x = Px_0 + x^*,
\]

where \( x^* \) is the unique solution of (15) on \( R(I - T) \).

(b) If \( T \) is convergent, (15) has a unique solution \( x \) for each \( f \in H \), and \( x_n \to x \) in norm for each \( x_0 \in H \). The rate of convergence is characterized by

\[
\| x_n - x \| = O(r^n),
\]

where \( r \) is any real number satisfying \( r(T) < r < 1 \). If, in addition, \( \| T \| < 1 \),

\[
\| x_n - x \| = O(\| T \|^n).
\]

The estimates (18) and (19) are standard results. Let us remark that \( f \in R(I - T) \) if and only if the sequence \( \{\sum_{k=0}^{N} T^k f\} \) is bounded in norm; this follows from the fact that the Hilbert space is reflexive and from Theorem 1(c) of [1].

The main objective of this section is to derive an analog of Theorem 11 for the approximate solution of the equation

\[
Au = h
\]

by means of the iterations

\[
(I + A) x_{n+1} = (I - A) x_n + 2h, \quad n = 0, 1, 2, \ldots,
\]

where \( A \) is a semistable operator. Let \( T = (I - A) (I + A)^{-1} \) be the Cayley transform of \( A \), and \( P \) the projector of \( H \) onto \( N(A) \) in the direction of \( R(A) \). As before, \( P = \lim_{n} T^n \) in norm. The equation (20) is equivalent to (15) with \( T = \phi(A) \) and \( f = (I + T) h \). Hence, the sequence \( \{x_n\} \) defined in (21) coincides with (16) for the described choice of \( T \) and \( f \). Recalling that \( N(A) = N(I - T) \) and \( R(A) = R(I - T) \), we have the following result.

**THEOREM 12.** (a) If \( A \) is semistable with \( 0 \in \sigma(A) \), the sequence \( \{x_n\} \) defined by (21) converges in norm for any \( x_0 \in H \) if and only if \( h \in R(A) \). If this is the case, the limit \( x \) of \( \{x_n\} \) is a solution of (20) of the form

\[
x = Px_0 + x^*,
\]

where \( x^* \) is the unique solution of (20) on \( R(A) \).
(b) If $A$ is stable, (20) has a unique solution $x$ for each $h \in H$, and $x_n \to x$ in norm for each $x_n \in H$. If for some $s$, $0 < s < 1$, $\sigma(A)$ is contained in the interior of the circle with diameter $[s, s^{-1}]$, then

$$\| x_n - x \| = O(r^n)$$

with $r = (1 - s)/(1 + s)$. If, in addition, $A$ is strictly stable with $\Re(x, x) \geq \alpha \| x \|^2$ for some $\alpha > 0$ and all $x \in H$,

$$\| x_n - x \| = O(M^n),$$

where

$$M = \left( \frac{1 + \| A \|^2 - 2\alpha}{1 + \| A \|^2 + 2\alpha} \right)^{1/2}.$$ 

Proof. (a) follows from the arguments preceding the statement of the theorem and from Theorem 11. To prove (b), assume first that $A$ is stable and that $\sigma(A)$ is contained in the interior of the circle $C$ with diameter $[s, s^{-1}]$, where $0 < s < 1$. It can be easily verified that the interior of $C$ is mapped under $(A) = (1 - A)/(1 + A)$ onto the interior of the circle $| \lambda | = r$ with $r = (1 - s)/(1 + s)$. Then (22) follows from the spectral mapping theorem combined with (18). Secondly, suppose that $\Re(Ax, x) \geq \alpha \| x \|^2$ for all $x \in H$, where $\alpha > 0$. We will estimate the norm of the operator $T = (I - A)(I + A)^{-1}$. For each $x \neq 0$ define $y = (I + A)x$. Then also $y \neq 0$, and

$$\frac{\| Tx \|^2}{\| x \|^2} = \frac{\| (I - A)y \|^2}{\| (I + A)y \|^2} = \frac{\| y \|^2 + \| Ay \|^2 - 2\Re(Ay, y)}{\| y \|^2 + \| Ay \|^2 + 2\Re(Ay, y)} \leq \frac{1 + \| A \|^2 - 2\alpha}{1 + \| A \|^2 + 2\alpha}.$$ 

Hence, $\| T \| \leq M < 1$, where $M$ is given by (24), and the estimate (23) follows from (19).

Acknowledgments

The author is grateful to Professor J. P. Williams for pointing out to him an error in the original proof of Theorem 1. The author also expresses his thanks to Professor B. E. Cain for suggesting a vital link in the present proof of Theorem 1, namely the solution of the equation $I - T^*T = XPX$. As the author learned from a private communication, Corollaries 2 and 4 appear also in the paper B. E. Cain (An inertia theory for operators in a Hilbert space, J. Math. Anal. Appl., to be published) based on a Ph.D. thesis.
REFERENCES