On the convergence of superpositions of a sequence of operators

by

M. REICHAW (Halfa)

Introduction. Let \( T_i : X \to X, \ i = 1, 2, \ldots, \) be continuous linear mappings of a Banach space \( X \) into itself and write \( T^n = T_n \cdots T_1. \) The problem of finding conditions under which the sequences

\[
\{T^n\}_{n=1,2,\ldots} \quad \text{and} \quad \left\{ \frac{1}{n} \sum_{i=1}^{n} T^i \right\}_{n=1,2,\ldots}
\]

converge for \( n \to \infty \) has been investigated by several authors. For example if \( T_i = T \) are all equal, then the following well-known mean ergodic theorem (9) or (13), p. 692, corollaries 2 and 3) holds.

Let \( T : X \to X \) be a linear continuous mapping of a Banach space \( X \) into itself. If the sequence of averages

\[
\frac{1}{n} \sum_{i=1}^{n} T^i
\]

is bounded, \( T^n x/n \to 0 \) as \( n \to \infty, \) and the sequence

\[
z_n = \frac{1}{n} \sum_{i=1}^{n} T^i x
\]

contains a weakly convergent subsequence \( \{z_{n_k}\} \) for \( x \) in a fundamental set, then

\[
\frac{1}{n} \sum_{i=1}^{n} T^i x
\]

* This research was supported by the Air Office of Scientific Research.

I would like to thank Professor J. Wolffowitz for his kind help in writing this paper. I am also indebted to Dr. S. Rolewicz for several useful remarks.
converges strongly to a linear continuous operator \( P : X \to X \) mapping \( X \) onto the manifold \( \{ x \in X : T x = x \} \) of fixed points of \( T \). Strong results in the case of a semigroup of operators \( T_n : X \to X \) have been obtained in [2].

The general case of different operators \( T_i, i = 1, 2, \ldots \), without the assumptions of any algebraic structure has been considered in [3], [4], [7] and [8] in the terminology of stochastic processes. In these works \( T_1 \) is a finite stochastic matrix and can be regarded as a linear operator mapping a finite-dimensional space into itself.

In section 1 of the present paper simple and easily applicable sufficient conditions for the convergence of sequences \( (T^n)_{n=1,2,\ldots} \) and \( \{ \sum \frac{T^j}{j!} \}_{j=1,2,\ldots} \) are given. In the case where \( T_i \) commute \( (i.e., T_iT_j = T_jT_i \) for all \( i, j = 1, 2, \ldots \) these conditions turn out to be also necessary for the convergence of \( (T^n)_{n=1,2,\ldots} \). In section 2 the notion of a stable operator, generalizing the notion of a stable matrix, is introduced and the results of section 1 are applied to the theory of stochastic matrices.

**Notation.** We denote by \( \| x \| \) (or \( \| T \| \)) the norm of the point \( x \) (of the operator \( T \), \( x \to x \) denotes \( |x_n - x| \) → 0 and \( T \to x \) denotes either the strong convergence of \( T_n \) to \( T \) \( (i.e., \forall x, \| T_n x - T x \| \to 0) \) or the norm convergence \( \| T_n - T \| \to 0 \).

A matrix \( S = (s_{ij}) \) (finite or not) will be called stable if all the rows of \( S \) are identical, \( i.e., s_{ij} = s_{ik} \) for all \( i, k \) and \( j \). A stochastic matrix \( T = (t_{ij}) \) is such that \( t_{ij} \geq 0 \) and \( \sum t_{ij} = 1 \) for every \( i \). It is easily seen that

\[ (a) \] If \( S \) is a stable matrix and \( T \) any matrix such that \( ST \) exists, then \( ST \) is a stable matrix,

\[ (b) \] If \( T \) is stochastic and \( S \) stable, then \( TS = S \),

\[ (c) \] If \( T \) is stochastic and \( S \) is stable and stochastic, then \( TS = S = S^2 \).

We shall consider a linear operator \( \Delta : X \to X \) and sequences \( \{ T_n \}_{n=1,2,\ldots} \) and \( \{ S_n \}_{n=1,2,\ldots} \) of linear operators mapping a Banach space \( X \) into itself and we shall consider also the following properties:

\[ (a) \] \( T_n S = T_n S_j + A S_j \) for all \( i, j = 1, 2, \ldots \),

\[ (b) \] \( \prod (T_n - S_j) \to 0 \) as \( n \to \infty \), where convergence means either strong convergence or norm-convergence \((1)\).

In the sequel \( T^n \) always stands for \( T_1 T_2 \ldots T_n \). Thus, if \( T_i = T \) are all equal, \( T^n = T \cdot T \ldots T \) is simply the \( n \)-th iterate of \( T \).

---

(1) One may illustrate condition \((a)\) geometrically by regarding \( S_t \) as "projections" of \( X \) onto the space \( T = S_t(X) = S_t(x) \) (common to all operators \( S_t \)) such that \( T x = x \) and \( A x = x \) for all \( x \in Y \).

(2) \( \prod (T_i - S_j) \) denotes \( (T_n - S_0)(T_{n-1} - S_{n-1}) \ldots (T_1 - S_1) \).
By letting \( n \to \infty \), we infer from (c), (b) and (d) that 
\( (d) \quad P = T_k P \).

Thus \( T_k P = P, \) \( T_k T_k P = P \) and by induction \( T^n P = P \) for every \( n = 1, 2, \ldots \). Therefore, by (d), \( P^1 = P \) and by (d) we infer that (d) holds.

It remains to show that \( P \) maps \( X \) onto

\[
F = \bigcap_{k=1}^\infty F_k \quad \text{and} \quad F = \{ z; Px = x \}.
\]

But this is quite trivial. Indeed, if
\[
x \in \bigcap_{k=1}^\infty F_k \quad \text{then} \quad T_k x = x \quad \text{for every} \quad k = 1, 2, \ldots \]

Consequently, \( T_k x = x \) and, by (d), \( Px = x \). Conversely, if \( x = Px \) for some \( x \in X \), then, by (d), \( T_k x = T_k Px = Px = x \). Thus \( x \in F \). Theorem 2 is proved.

**Theorem 3.** Let \( T_i : X \to X \) be linear continuous operators mapping a Banach space \( X \) into itself satisfying assumption (c) of Theorem 2 and such that

(a) \( T_i \) and \( T_j \) commute for all \( i, j = 1, 2, \ldots \) (i.e., \( T_i T_j = T_j T_i \)).

A necessary and sufficient condition for the existence of a linear operator \( P: X \to X \) such that

\[
(d) \quad T^n \to P \quad \text{as} \quad n \to \infty
\]

and

\[
(d') \quad P = T_k P = P^1 = PT_k \quad \text{for every} \quad k = 1, 2, \ldots,
\]

is the existence of a sequence \( (S_i)_{i=1}^\infty \) of linear operators \( S_i : X \to X \) such that (a) and (b) hold.

**Proof of sufficiency.** By Theorem 2 there exists a linear operator \( P: X \to F \) mapping \( X \) onto the set

\[
F = \bigcap_{k=1}^\infty F_k = \{ z; Px = x \},
\]

where \( F_k = \{ z; T_k x = x \} \) such that (d) and (d') are satisfied. Thus (d) holds and it remains to show that \( P T_k = P \) for all \( k = 1, 2, \ldots \). This, however, is a simple consequence of (c), (d') and (d). Indeed, by (c) we have \( T^n T_k = T_k T^n \) for all \( k \) and \( n \). Letting \( n \to \infty \) we infer by (d') and (d) that \( P T_k = T_k P = P \).

**Proof of necessity.** To prove the necessity of (a) and (b) let us put \( S_i = P \) for all \( j = 1, 2, \ldots \). Then by (d') we infer that (a) is satisfied and it remains to show that (b) holds. Indeed, by (d) we have
\[(T_1 - P)(T_1 - P) = T_1 T_1 - PT_1 - T_P + P^2 = T_2 T_1 - P \text{ and by induction}
\]
\[\prod_{i=1}^{n} (T_i - P) = T^n - P.\]

Thus by (a'), we obtain
\[\prod_{i=1}^{n} (T_i - P) \to 0 \text{ as } n \to \infty.\]

Theorem 3 is proved.

Using the fact that in the proof of the necessity in Theorem 3 it is possible to put \(S_j = P\) for all \(j = 1, 2, \ldots\) and putting, in Theorem 3, \(T_i = T\) and \(S_j = S\), where \(T : X \to X\) and \(S : X \to X\) are linear operators of \(X\) into itself, we obtain as a consequence of Theorem 3 the following

**Theorem 4.** Let \(T : X \to X\) be a linear operator of a Banach space \(X\) into itself satisfying

(a') there exists a constant \(C\) such that \(\|T^n\| \leq C\) for all \(n = 1, 2, \ldots\), where \(T^n = T \cdot T \cdot \ldots \cdot T\).

A necessary and sufficient condition for the existence of a linear operator \(P : X \to X\) such that

(d') \(T^n P = P T^n\) as \(n \to \infty\)

and

(e') \(T^n P = P T^n\) as \(n \to \infty\).

Let us conclude this section by the following

**Theorem 5.** If \(T : X \to X\) is a linear operator of a Banach space \(X\) into itself such that there exists a linear operator \(S : X \to X\) satisfying (a') and if

\[\frac{1}{k} \|T^n\| \to 0 \text{ as } k \to \infty\]

and \(\sum_{i=1}^{k} (T - S)^i\) is bounded, then

\[\left\|\frac{1}{k} \sum_{i=1}^{k} T^i - \frac{1}{k} \sum_{i=1}^{k} S^i\right\| \to 0 \text{ as } k \to \infty.\]

**Proof.** Applying Theorem 1 to \(A = T^{k+1}, T_j = T\) and \(S_j = S\) we infer by (a') that

\[T^n - T^{n+k} = (T - T^{k+1}) (T - S)^{n-1}.\]

Summation over \(j\) and division by \(k\) yields

\[\frac{1}{k} \sum_{i=1}^{k} T^i - \frac{1}{k} \sum_{i=1}^{k} S^i = \frac{1}{k} (T - T^{k+1}) \sum_{i=1}^{k} (T - S)^i.\]

Since by the assumptions of the theorem, the right-hand side of this equality tends to zero as \(k \to \infty\), Theorem 5 is proved.

**Remark 2.** Theorem 5 may be interpreted in the following manner.

Let \(T\) denote the transition probability matrix from one state of a given system to another state during a time unit. Then \(T^n\) denotes the transition probability matrix after \(n\) units of time. Suppose that we know the average \(\frac{1}{k} \sum_{i=1}^{k} T^i\). Then, if the assumptions of Theorem 5 hold, we find that the average \(\frac{1}{k} \sum_{i=1}^{k} T^i\) for \(k\) sufficiently large will be close to \(\frac{1}{k} \sum_{i=1}^{k} S^i\).

A similar reasoning holds when we interchange the roles of \(\frac{1}{k} \sum_{i=1}^{k} T^i\) and \(\frac{1}{k} \sum_{i=1}^{k} S^i\).

Let us also note that knowing the function \(\|T^n\|/k\) and the upper bound of \(\sum_{i=1}^{k} (T - S)^i\), one can estimate

\[\left\|\frac{1}{k} \sum_{i=1}^{k} T^i - \frac{1}{k} \sum_{i=1}^{k} S^i\right\|.\]

2. To give some applications of the results obtained in section 1 we will introduce the notion of a stable operator \(S : X \to X\) mapping a Banach space \(X\) into itself. This notion is a generalization of the notion of a stable matrix. Let us denote by \(X^*\) the conjugate of \(X\) (i.e., \(X^*\) is the space of all linear continuous functionals on \(X\)).

**Definition.** Let \(e = (\alpha_i)_{i=1}^{\infty}\) and \(f = (\beta_j)_{j=1}^{\infty}\) be biorthogonal sequences (finite or not), where \(\alpha_i : X \to \mathbb{R}\) and \(f_j : X^* \to \mathbb{R}\). We say that \(S : X \to X\) is stable relatively to the pair \((e, f)\) if and only if for all \(i, j\), we have \((\alpha_i, \beta_j)\).
Thus by (b') and by $T^n \to P$ it follows that $ST^n \to P$ as $n \to \infty$. But, by (b''), $ST^n$ is stable relatively to $(e, f)$ and therefore, by (b), $P$ is stable relatively to $(e, f)$. The theorem is proved.

Before the next Remark let us recall the well-known fact that for a finite dimensional Banach space all norms are equivalent. In particular, for the space $[X]$ of linear operators $T : X \to X$ mapping a finite dimensional Banach space $X$ into itself all norms are equivalent. Each operator belonging to $[X]$ is represented by a finite matrix; thus if $X$ is $n$-dimensional, $[X]$ is $n^2$-dimensional. Therefore

(i) if $T^n \to P$ in one norm, then $T^n \to P$ in any other norm provided that $X$ maps a finite dimensional Banach space into itself.

Remark 3. A finite stochastic matrix $T$ is called indecomposable and aperiodic (SIA) if $P = \lim T^n$ exists and $P$ is stable (see for instance [8], p. 733). It is a trivial consequence of (f), (a), (a) and Theorem 4 that $T$ is SIA if and only if for some norm || || there exists a stable stochastic matrix $S$ such that $||T^n|| \leq C$ for all $n = 1, 2, \ldots$ and $(T - S) \to 0$ as $n \to \infty$. This gives a simple characterization of SIA matrices.

References