On the convergence of superpositions of a sequence of operators*

by

M. REICHAW (Haifa)

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

$$\{T^n\}_{n=1,2,\ldots} \quad \text{and} \quad \left\{\frac{1}{n}\sum_{j=1}^{n} T^j\right\}_{n=1,2,\ldots}$$

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 [1], p. 662, 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

$$\frac{1}{n}\sum_{j=1}^{n} T^j$$

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

$$x_n = \frac{1}{n}\sum_{j=1}^{n} T^j x$$

contains a weakly convergent subsequence $\{x_{n_\nu}\}$ for x in a fundamental set, then

$$\frac{1}{n}\sum_{j=1}^{n} T^j x$$

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

I would like to thank Professor J. Wolfowitz 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 $\{x; Tx = x\}$ of fixed points of T. Strong results in the case of a semigroup of operators $T_a: X \to X$ 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_i 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 $\{n^{-1}\sum_{j=1}^{n} T^j\}_{n=1,2\ldots}$ 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 $\|x\|$ (by $\|T\|$) the norm of the point x (of the operator T), $x_n \to x$ denotes $\|x_n - x\| \to 0$ and $T_n \to T$ denotes either the strong convergence of T_n to T (i. e. for every x, $\|T_nx - Tx\| \to 0$) 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., if $s_{i,j} = s_{k,j}$ for all i, k and j. A *stochastic matrix* $T = (t_{i,j})$ is such that $t_{ij} \geqslant 0$ and $\sum_{j=1}^{\infty} t_{ij} = 1$ for every i. It is easily seen that

(a_0) if S is a stable matrix and T any matrix such that ST exists, then ST is a stable matrix,

(a_1) if T is stochastic and S stable, then $TS = S$,

(a_2) if T is stochastic and S is stable and stochastic, then $TS = S = S^2$.

We shall consider a linear operator $A: X \to X$ and sequences $\{T_i\}_{i=1,2,\ldots}$ and $\{S_j\}_{j=1,2,\ldots}$ of linear operators mapping a Banach space X into itself and we shall consider also the following properties:

(a) $T_iS_j = S_j = S_iS_j$ and $AS_j = S_j$ for all $i, j = 1, 2, \ldots$ [1],

(b) $\prod_{i=1}^{n} (T_i - S_i) \to 0$ as $n \to \infty$, where convergence means either strong convergence or norm-convergence [2].

In the sequel T^n always stands for $T_n \cdot T_{n-1} \cdot \ldots \cdot T_1$. Thus, if $T_i = T$ are all equal, $T^n = \underbrace{T \cdot T \cdot \ldots \cdot T}_{n}$ is simply the n-th iterate of T.

---

[1] One may illustrate condition (a) geometrically by regarding S_i as "projections" of X onto the space $Y = S_i(X) = S_j(X)$ (common to all operators S_i) such that $T_ix = x$ and $Ax = x$ for all $x \in Y$.

[2] $\prod_{i=1}^{n} (T_i - S_i)$ denotes $(T_n - S_n)(T_{n-1} - S_{n-1}) \ldots (T_1 - S_1)$.

1. We begin with the following

THEOREM 1. *Let $A: X \to X$, $T_i: X \to X$ and $S_j: X \to X$ be linear mappings of a linear space X into itself satisfying property* (a).

Then

$$(\bar{a}) \qquad T^n - AT^{n-1} = (T_n - A)\prod_{i=1}^{n-1} (T_i - S_i).$$

Proof. We will first show by induction that

$$(\bar{\bar{a}}) \qquad T^n - S_nT^{n-1} = \prod_{i=1}^{n} (T_i - S_i).$$

Indeed, denoting by T^0 the identity operator, we see that $(\bar{\bar{a}})$ holds for $n = 1$. Now suppose that $(\bar{\bar{a}})$ holds for n. Then

$$\prod_{i=1}^{n+1} (T_i - S_i) = (T_{n+1} - S_{n+1})(T^n - S_nT^{n-1})$$
$$= T^{n+1} - S_{n+1}T^n - T_{n+1}S_nT^{n-1} + S_{n+1}S_nT^{n-1}.$$

By (a) we have $T_{n+1}S_n = S_n = S_{n+1}S_n$ and thus $-T_{n+1}S_nT^{n-1} + S_{n+1}S_nT^{n-1} = 0$. Therefore

$$\prod_{i=1}^{n+1} (T_i - S_i) = T^{n+1} - S_{n+1}T^n$$

and $(\bar{\bar{a}})$ is proved.

To prove $(\bar{a})$ let us denote by T^0 the identity operator and by $S_0 = 0$ the null operator.

Then $(\bar{a})$ is trivially satisfied for $n = 1$. Using $(\bar{\bar{a}})$ with n replaced by $n-1$ we obtain

$$(T_n - A)\prod_{i=1}^{n-1} (T_i - S_i) = (T_n - A)(T^{n-1} - S_{n-1}T^{n-2})$$
$$= T^n - AT^{n-1} - T_nS_{n-1}T^{n-2} + AS_{n-1}T^{n-2}.$$

But by (a) we have $-T_nS_{n-1}T^{n-2} + AS_{n-1}T^{n-2} = -S_{n-1}T^{n-2} + S_{n-1}T^{n-2} = 0$. Thus $(\bar{a})$ holds. Theorem 1 is proved.

Theorem 1 is the main theorem of this paper. The following theorems are simple consequences of Theorem 1 and are obtained by applying $(\bar{a})$ to some particular operators A and sequences $\{T_i\}_{i=1,2}$ and $\{S_j\}_{j=1,2,\ldots}$. In the sequel "linear operator" stands for "linear and continuous operator".

THEOREM 2. *Let $T_i: X \to X$ be linear operators of a Banach space X into itself satisfying*

(c) *there exists a constant C such that $\|T_{n+k}\cdot T_{n+k-1}\cdot\ldots\cdot T_n\| \leqslant C$ for all n and k.*

A sufficient condition for the existence of a linear operator $P\colon X\to X$ such that

(d$_1$) *$T^n\to P$ as $n\to\infty$,*

and

(d$_2$) *$P = T_kP = P^2$ for every $k = 1, 2, \ldots,$*

is the existence of a sequence $\{S_j\}_{j=1,2,\ldots}$ of linear operators $S_j\colon X\to X$ satisfying (b) *and*

(a*) *$T_iS_j = S_iS_j = S_j$ for all $i, j = 1, 2, \ldots$*

Moreover, the operator P maps X onto the set

$$F = \bigcap_{k=1}^{\infty} F_k$$

where $F_k = \{x; T_kx = x\}$ is the set of fixed points of T_k and $F = \{x; Px = x\}$.

Proof. By (a*) we have $T_{n+k}\cdot T_{n+k-1}\ldots T_nS_j = S_j$ and therefore property (a) is satisfied for $A = T_{n+k}\cdot T_{n+k-1}\ldots T_n$. By Theorem 1 it follows that

$$T^n - T^{n+k} = (T_n - T_{n+k}\cdot T_{n+k-1}\ldots T_n)\prod_{i=1}^{n-1}(T_i - S_i).$$

Hence by (c) we have

(d$_3$) $\|T^n - T^{n+k}\| \leqslant 2C\left\|\prod_{i=1}^{n-1}(T_i - S_i)\right\|$

and thus for every $x\in X$

(d$_3'$) $\|T^nx - T^{n+k}x\| \leqslant 2C\left\|\prod_{i=1}^{n-1}(T_i - S_i)x\right\|.$

The space of linear operators mapping X into itself being complete both in the norm-topology and in the strong topology (see [6], p. 140 and p. 142), it follows by (b), (d$_3$) and (d$_3'$) that

(d$_4$) there exists a linear operator $P\colon X\to X$ such that $T^n\to P$ as $n\to\infty$.

Thus (d$_1$) holds.

To show (d$_2$) let us apply Theorem 1 to the operator $A = T_k$ where k is fixed, and to the sequences $\{T_i\}_{i=1,2,\ldots}$ and $\{S_j\}_{j=1,2,\ldots}$. Then by (a*) we obtain

$$T^n - T_kT^{n-1} = (T_n - T_k)\prod_{i=1}^{n-1}(T_i - S_i).$$

By letting $n\to\infty$, we infer from (c), (b) and (d$_1$) that

(d$_5$) $P = T_kP$.

Thus $T_1P = P$, $T_2T_1P = P$ and by induction $T^nP = P$ for every $n = 1, 2, \ldots$ Therefore, by (d$_1$), $P^2 = P$ and by (d$_5$) we infer that (d$_2$) holds.

It remains to show that P maps X onto

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

But this is quite trivial. Indeed, if

$$x\in\bigcap_{k=1}^{\infty} F_k = F,$$

then $T_kx = x$ for every $k = 1, 2, \ldots$

Consequently, $T^nx = x$ and, by (d$_1$), $Px = x$. Conversely, if $x = Px$ for some $x\in X$, then, by (d$_2$), $T_kx = T_kPx = Px = x$. Thus $x\in F$. Theorem 2 is proved.

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

(e) *T_i and T_j commute for all $i, j = 1, 2, \ldots$ (i. e. $T_iT_j = T_jT_i$).*

A necessary and sufficient condition for the existence of a linear operator $P\colon X\to X$ such that

(d$_1'$) *$T^n\to P$ as $n\to\infty$*

and

(d$_2'$) *$P = T_kP = P^2 = PT_k$ for every $k = 1, 2, \ldots,$*

is the existence of a sequence $\{S_j\}_{j=1,2,\ldots}$ of linear operators $S_j\colon X\to X$ such that (a*) *and* (b) *hold.*

Proof of sufficiency. By Theorem 2 there exists a linear operator $P\colon X\to F$ mapping X onto the set

$$F = \bigcap_{k=1}^{\infty} F_k = \{x; Px = x\},$$

where $F_k = \{x; T_kx = x\}$ such that (d$_1$) and (d$_2$) are satisfied. Thus (d$_1'$) holds and it remains to show that $PT_k = P$ for all $k = 1, 2, \ldots$ This, however, is a simple consequence of (e), (d$_1'$) and (d$_2$). Indeed, by (e) we have $T^nT_k = T_kT^n$ for all k and n. Letting $n\to\infty$ we infer by (d$_1'$) and (d$_2$) that $PT_k = T_kP = P$.

Proof of necessity. To prove the necessity of (a*) and (b) let us put $S_j = P$ for all $j = 1, 2, \ldots$ Then by (d$_2'$) we infer that (a*) is satisfied and it remains to show that (b) holds. Indeed, by (d$_2'$) we have

$(T_2-P)(T_1-P) = T_2T_1-PT_1-T_2P+P^2 = T_2T_1-P$ and by induction

$$\prod_{i=1}^{n}(T_i-P) = T^n-P.$$

Thus by (d_1') we obtain

$$\prod_{i=1}^{n}(T_i-P)\to 0 \quad \text{as} \quad n\to\infty.$$

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*

(c′) *there exists a constant C such that $\|T^n\|\leqslant C$ for all $n = 1, 2, \ldots$, where $T^n = \underbrace{T\cdot T\ldots T}_{n}$ is the n-th iterate of T.*

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

(d′) $T^n\to P$ *as* $n\to\infty$

and

(d″) $P = TP = P^2 = PT$

is the existence of a linear operator $S: X\to X$ such that

(a′) $TS = S = S^2$

and

(b′) $(T-S)^n\to 0$ *as* $n\to\infty$.

Remark 1. Let us note that by Theorem 2 the limit P of T^n maps the space X onto the subspace $\{x; Tx = x\} = \{x; Px = x\}$ of all fixed points of T. By property (d_3) we obtain also the following inequality:

$$\|T^n-P\|\leqslant 2C\|(T-S)^{n-1}\|.$$

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*

$$\frac{1}{k}\|T^k\|\to 0 \quad \text{as} \quad k\to\infty$$

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

$$\Big\|\frac{1}{k}\sum_{j=1}^{k}T^j-\frac{1}{k}\sum_{j=k+1}^{2k}T^j\Big\|\to 0 \quad \text{as} \quad k\to\infty.$$

Proof. Applying Theorem 1 to $A = T^{k+1}$, $T_j = T$ and $S_j = S$ we infer by $(\bar{a})$ that

$$T^j-T^{j+k} = (T-T^{k+1})(T-S)^{j-1}.$$

Summation over j and division by k yields

$$\frac{1}{k}\sum_{j=1}^{k}T^j-\frac{1}{k}\sum_{j=k+1}^{2k}T^j = \frac{1}{k}(T-T^{k+1})\sum_{j=1}^{k}(T-S)^j.$$

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 $k^{-1}\sum_{j=1}^{k}T^j$. Then, if the assumptions of Theorem 5 hold, we find that the average $k^{-1}\sum_{k+1}^{2k}T^j$ for k sufficiently large will be close to $k^{-1}\sum_{j=1}^{k}T^j$. A similar reasoning holds when we interchange the roles of $k^{-1}\sum_{j=1}^{k}T_j$ and $k^{-1}\sum_{j=k+1}^{2k}T^j$. Let us also note that knowing the function $\|T^k\|/k$ and the upper bound of $\|\sum_{j=1}^{k}(T-S)^{j-1}\|$ one can estimate

$$\Big\|\frac{1}{k}\sum_{j=1}^{k}T^j-\frac{1}{k}\sum_{j=k+1}^{2k}T^j\Big\|.$$

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 $e = \{e_i\}_{i=1,2,\ldots}$ and $f = \{f_j\}_{j=1,2,\ldots}$ be biorthogonal sequences (finite or not), where $e_i\in X$ and $f_i\in X^*$. We say that $S: X\to X$ is *stable relatively to the pair* (e, f) if and only if for all i, j, k we have (f_i, Se_j)

$= (f_k, Se_j)$, where $(\varphi, x) = \varphi(x)$ is the value of the functional $\varphi \in X^*$ at the point $x \in X$.

For example, let $X = m$ be the Banach space of all bounded sequences $x = (\xi_1, \xi_2, \ldots)$ of real numbers ξ_i with norm $\|x\| = \sup_{1\leqslant i<\infty} |\xi_i|$. Let $e_i = (0, 0, \ldots, 0, \underset{i\text{-th place}}{1}, 0, \ldots)$ and let $f_i(x) = \xi_i$, $i = 1, 2, \ldots$ Then each stable and stochastic matrix $S = (s_{i,j})$ is a stable operator $S: X \to X$ relatively to the pair (e, f) and, as can easily be seen, we have $\|S\| = 1$.

Similarly, if X is a ν-dimensional Banach space and $e_i = (0, 0, \ldots, 0, \underset{i}{1}, 0, \ldots, 0) = \{\delta_{ij}\}_{j=1,2,\ldots,\nu}$ where

$$\delta_{ij} = \begin{cases} 0 & \text{for} \quad i \neq j \\ 1 & \text{for} \quad i = j \end{cases} \quad \text{and} \quad f_j(x) = \xi_j \quad \text{for} \quad x = \sum_{i=1}^{\nu} \xi_i e_i,$$

we infer that a stable matrix $S = (s_{ij})_{i,j=1,\ldots,\nu}$ is a stable operator mapping X into itself. A trivial consequence of the definition of a stable operator is:

($\bar{\text{e}}$) If $S_n \to S$ and $S_n: X \to X$ is a stable operator relatively to (e, f), $n = 1, 2, \ldots$, then $S: X \to X$ is a stable operator relatively to (e, f).

Let us prove the following

THEOREM 6. *If $T: X \to X$ is a linear operator of a Banach space X into itself such that*

(c′) $\|T^n\| \leqslant C$ *for all* $n = 1, 2, \ldots$

and if there exists a continuous linear operator $S: X \to X$ satisfying

(a′) $TS = S = S^2$,

(b′) $(T-S)^n \to 0$

and

(b″) ST^n *is a stable operator relatively to* (e, f),

then there exists a linear continuous operator $P: X \to X$ such that $T^n \to P$, $P = TP = P^2 = PT$ and P is a stable operator relatively to (e, f). Moreover, P maps X onto the set $\{x; Tx = x\} = \{x; Px = x\}$ of fixed points of T and in the case where X is a Hilbert space and $T = T^*$ is selfadjoint, the operator P is also selfadjoint.

Proof. By assumptions (c′), (a′) and (b′) we infer from Theorem 4 that $T^n \to P: X \to X$ and $P = TP = P^2 = PT$. Moreover, by Remark 1, P maps X onto the set $\{x; Tx = x\} = \{x; Px = x\}$. Further, in the case where X is a Hilbert space and T is selfadjoint, P is also selfadjoint as a limit of selfadjoint operators T^n. It remains to show that P is stable relatively to (e, f). For this purpose let us apply Theorem 1 to the operators $A = S$, $T_i = T$ and $S_j = S$. Then by ($\bar{\text{a}}$) we obtain

$$T^n - ST^{n-1} = (T-S)^n.$$

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 ($\bar{\text{e}}$), P is stable relatively to (e, f). The theorem is proved.

Before the next Remark 3 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 ν-dimensional, $[X]$ is ν^2-dimensional.) Therefore

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

Remark 3. A finite stochastic matrix T is called *indecomposable* and *aperiodic* (SIA) if $P = \lim_{n\to\infty} T^n$ exists and P is stable (see for instance [8], p. 733). It is a trivial consequence of (f), (a_1), (a_2) and Theorem 4 that T is SIA if and only if for some norm $\| \ \|$ there exists a stable stochastic matrix S such that $\|T^n\| \leqslant C$ for all $n = 1, 2, \ldots$ and $(T-S)^n \to 0$ as $n \to \infty$. This gives a simple characterization of SIA matrices.

References

[1] N. Dunford and J. T. Schwartz, *Linear operators*, Part I, I. N. C., New York 1958.

[2] W. F. Eberlein, *Abstract ergodic theorems and weak almost periodic functions*, Trans. Amer. Math. Soc. 67 (1949), p. 217-240.

[3] J. Hajnal, *The ergodic properties of non-homogeneous finite Markov chains*, Proc. Cambridge Phil. Soc. 52 (1956), p. 67-77.

[4] — *Weak ergodicity in non-homogeneous Markov chains*, ibidem 54 (1958), p. 233-246.

[5] Л. В. Канторович и Г. П. Акилов, *Функциональный анализ в нормированных пространствах*, Москва 1959.

[6] Л. А. Люстерник и В. И. Соболев, *Элементы функционального анализа*, Москва 1951.

[7] J. L. Mott and H. Scheider, *Matrix norms applied to weakly ergodic Markov chains*, Archiv der Mathematik 8 (1957), p. 331-333.

[8] J. Wolfowitz, *Products of indecomposable aperiodic stochastic matrices*, Proc. Amer. Math. Soc. 14 (1963), p. 733-737.

[9] K. Yosida and S. Kakutani, *Operator theoretical treatment of Markov process and mean ergodic theorem*, Ann. Math. 42 (1941), p. 188-228.

TECHNION — I. I. T. HAIFA
CORNELL UNIVERSITY, ITHACA, NEW YORK

Reçu par la Rédaction le 3. 8. 1964