11. An Ergodic Theorem

By Hikosaburo KOMATSU

University of Tokyo

(Comm. by Kunihiko KODAIRA, M. J. A., Feb. 12, 1968)

Let A be a bounded linear operator in a Banach space X with uniformly bounded powers:

(1) $\qquad \| A^n \| \leqq M < \infty, \qquad n=1, 2, \cdots.$

If X is reflexive, then the mean ergodic theorem of Yosida [4] and Kakutani [2] asserts that the mean of powers $\sum_{k=1}^{n} A^k/n$ converges strongly to a projection onto the eigen-space $N(1-A)=\{x \in X; (1-A)x=0\}$. Hille [1] showed also that the arithmetic mean may be replaced by any Cesaro mean of positive order. This seems to be the farthest one can expect under condition (1), for, obviously there are operators A with (1) such that A^n does not converge in any sense.

The purpose of this paper is to prove the following theorem which gives a sufficient condition in order that A^n converge strongly.

Assumptions. A is a bounded linear operator in a Banach space X which satisfies the following:

(a) The spectrum $\sigma(A)$ is contained in the unit disc $|\lambda| \leqq 1$;

(b) 1 is the only spectrum on the unit circle $|\lambda|=1$;

(c) There is an angle θ with $0 \leqq \theta < \pi/2$ such that the sector $S=\{\lambda \in \boldsymbol{C}; |\arg(\lambda-1)| < \pi-\theta\}$ is contained in the resolvent set $\rho(A)$ and

(2) $\qquad \|(\lambda-1)(\lambda-A)^{-1}\| \leqq K < \infty, \qquad \lambda \in S.$

Conclusions. (i) A satisfies (1);

(ii) $N(1-A)=N((1-A)^m)$, $m=1, 2, \cdots$;

(iii) The closure $\overline{R(1-A)}$ of the range $R(1-A)=\{(1-A)x; x \in X\}$ and $N(1-A)$ have only zero as the common elements;

(iv) $N(1-A)+\overline{R(1-A)}$ is a closed subspace of X, and coincides with X if X is reflexive;

(v) If $x=x_0+x_1$ with $x_0 \in N(1-A)$ and $x_1 \in \overline{R(1-A)}$, then $A^n x$ converges strongly to x_0 as $n \to \infty$;

(vi) Let $x \in X$. If there is a sequence $n_j \to \infty$ such that $A^{n_j}x$ converges weakly, then x belongs to $N(1-A)+\overline{R(1-A)}$.

Proof of (i). If Γ is a path which encircles $\sigma(A)$, we have

$$A^n = \frac{1}{2\pi i}\int_{\Gamma} \zeta^n (\zeta - A)^{-1} d\zeta.$$

Let Γ be the union of $\Gamma_1, \Gamma_2, \Gamma_3$, and Γ_4, where Γ_1 is the segment that connects $1+\cos\theta e^{i(\theta-\pi)}$ and $1+n^{-1}e^{i(\theta-\pi)}$, Γ_2 is the part of circle of radius n^{-1} with center at 1 that connects $1+n^{-1}e^{i(\theta-\pi)}$ and $1+n^{-1}e^{i(\pi-\theta)}$, Γ_3 is the segment that connects $1+n^{-1}e^{i(\pi-\theta)}$ and $1+\cos\theta e^{i(\pi-\theta)}$, and Γ_4 is a path in the interior of the unit disc outside $\sigma(A)$ that connects $1+\cos\theta e^{i(\pi-\theta)}$ and $1+\cos\theta e^{i(\theta-\pi)}$.

We have

$$\left\|\int_{\Gamma_1}\right\| \leqq \int_{1/n}^{\cos\theta}(1-2r\cos\theta+r^2)^{n/2}K dr/r$$
$$\leqq \int_1^{n\cos\theta}\left(1-\frac{s\cos\theta}{n}\right)^{n/2}K ds/s$$
$$\leqq K\int_1^{\infty} e^{-(s\cos\theta)/2} ds/s.$$

The integral over Γ_3 is estimated in the same way. Clearly we have

$$\left\|\int_{\Gamma_2}\right\| \leqq (1+n^{-1})^n 2K(\pi-\theta) \leqq 2\pi e K.$$

Lastly, let L be the length of Γ_4 and let R and N be the maxima on Γ_4 of $|\zeta|$ and $\|(\zeta-A)^{-1}\|$ respectively. Then

$$\left\|\int_{\Gamma_4}\right\| \leqq R^n N L \leqq L N.$$

Lemma. *For each $m=1,2,\cdots$, there is a constant M_m such that*

(3) $\|n^m(1-A)^m A^n\| \leqq M_m, \qquad n=1,2,\cdots.$

Proof. We have

$$n^m(1-A)^m A^n = \frac{1}{2\pi i}\int_\Gamma \zeta^n(n(1-\zeta))^m(\zeta-A)^{-1}d\zeta.$$

Since the integrand is bounded near 1 in this case, we can deform Γ_2 to the circle with arbitrarily small radius. Then,

$$\left\|\int_{\Gamma_1}\right\| \leqq \int_0^{\cos\theta}(1-2r\cos\theta+r^2)^{n/2}(nr)^m K dr/r$$
$$\leqq K\int_0^{\infty} e^{-(s\cos\theta)/2} s^m ds/s < \infty.$$

The integral over Γ_3 is treated in the same way. We have

$$\left\|\int_{\Gamma_4}\right\| \leqq R^n(2n)^m N L \leqq M_m/2$$

for some M_m, because $R<1$.

Proof of (v). If $x_0 \in N(1-A)$, $A^n x_0 = x_0$ tends to x_0. If $x_1 \in R(1-A)$, then it follows from Lemma that $A^n x_1 = O(n^{-1})$ and thus converges strongly to zero. Since A^n is uniformly bounded, the Banach-Steinhaus theorem shows that the same conclusion holds for $x_1 \in \overline{R(1-A)}$. Then,

Proof of (iii) is trivial.

Proof of (vi). Suppose that $A^{n_j}x$ converges weakly to x_0. Then, $(1-A)A^{n_j}x$ converges weakly to $(1-A)x_0$. On the other hand, it converges strongly to zero by Lemma. Hence we have $(1-A)x_0=0$.

Let $x_1 = x - x_0$. Clearly $A^{n_j}x_1 \to 0$ weakly. The equation

$$x_1=(A+(1-A))^{n_j}x_1$$
$$=A^{n_j}x_1+(1-A)\{n_jA^{n_j-1}x_1+\cdots\}$$

shows that x_1 is the weak limit of a sequence in $R(1-A)$.

Proof of (iv). By (v) and (vi) $N(1-A)+\overline{R(1-A)}$ is the set of all elements $x\in X$ such that $A^n x$ converges strongly. Since A^n is uniformly bounded, the closedness of $N(1-A)+\overline{R(1-A)}$ follows from the Banach-Steinhaus theorem.

If X is reflexive, $\{A^n x\}$ is sequencially weakly compact for any $x\in X$. Therefore, we can choose a subsequence $A^{n_j}x$ which converges weakly.

Proof of (ii). It is enough to prove that $(1-A)^2x=0$ implies $(1-A)x=0$ and this is immediate from (iii). However, this holds under a weaker condition than (2). In fact, suppose that there is a sequence $\lambda_j\to 1$ in $\rho(A)$ such that $\|(\lambda_j-A)^{-1}\|=o((\lambda_j-1)^{-2})$. Since

$$1=(\lambda-1)(\lambda-A)^{-1}+(1-A)(\lambda-A)^{-1}, \tag{4}$$

we have $(1-A)(\lambda_j-A)^{-1}=o((\lambda_j-1)^{-1})$. Multiplying $1-A$ on both sides, we get

$$(1-A)x=(\lambda_j-1)(1-A)(\lambda_j-A)^{-1}x+(\lambda_j-A)^{-1}(1-A)^2x.$$

Let $(1-A)^2x=0$. Then the second term vanishes. When λ_j tends to 1, the first term converges to zero. Thus we have $(1-A)x=0$.

Remarks. For normal operators A in a Hilbert space X estimate (2) in condition (c) holds automatically. In particular, self-adjoint operators A with $-1+\varepsilon\leqq A\leqq 1$ satisfy conditions (a), (b), and (c).

The whole theory depends on (i) and Lemma. Actually we need only (i) and the fact that $A^n(1-A)$ converges strongly to zero. The latter is much weaker than the statement of Lemma. Probably one can find weaker conditions than (a), (b), and (c) which ensure all conclusions of the theorem.

In [3] we will give a similar ergodic theorem in the case of continuous parameter with a detailed discussion on the order of convergence as the parameter tends to infinity.

References

[1] E. Hille: Remarks on ergodic theorems. Trans. Amer. Math. Soc., **57**, 246-269 (1945).

[2] S. Kakutani: Iteration of linear operators in complex Banach spaces. Proc. Imp. Acad. Tokyo, **14**, 295-300 (1938).

[3] H. Komatsu: Fractional powers of operators. III. Negative powers (to appear).

[4] K. Yosida: Mean ergodic theorem in Banach spaces. Proc. Imp. Acad. Tokyo, **14**, 292-294 (1938).