MEAN ERGODIC THEOREMS FOR NONLINEAR OPERATORS

RAINER WITTMANN

(Communicated by William J. Davis)

ABSTRACT. The mean ergodic theorem is shown for nonlinear operators T: $K \longrightarrow K$ with $\|Tx + Ty\| \leq \|x + y\|$ for any $x, y \in K$ where K may be an arbitrary subset of a Hilbert space H.

1. INTRODUCTION

Let K be a subset of Banach space H and $T: K \longrightarrow K$ be a mapping. We are concerned with the mean ergodic theorem for such nonlinear operators, i.e. the norm convergence of

$$A_n(x) := \frac{1}{n+1} \sum_{i=0}^{n} T^i x \qquad (x \in K).$$

Such results are very well known for linear contractions. In order to generalize such results to nonlinear mappings the most natural constraint on T is to assume that

$$(1) \qquad \|Tx - Ty\| \leq \|x - y\| \qquad (x, y \in K).$$

Such mappings are called *nonexpansive*. Unfortunately there exist examples of nonexpansive operators where $K = H$ is a Hilbert space and T is nonexpansive with nonempty fixed point set but the mean ergodic theorem does not hold (although $A_n x$ converges weakly). The speed limit operators of Krengel and Lin [6] are such examples if the speed function is chosen appropriately. Moreover, these operators satisfy several additional natural conditions. On the other hand Baillon [1] was able to prove the mean ergodic theorem for nonexpansive operators T on a convex subset K of a Hilbert space if $-x \in K$ and $T(-x) = -Tx$ for any $x \in K$. Clearly the last condition and (1) imply

$$(2) \qquad \|Tx + Ty\| \leq \|x + y\| \qquad (x, y \in K).$$

In this paper we show that (2) alone guarantees the validity of the mean ergodic theorem if H is a Hilbert space. Note that (2) does not even imply continuity. On the other hand, if $-x \in K$ for any $x \in K$, then (2) implies (1). In fact, $\|Tx + T(-x)\| \le \|x + (-x)\| = 0$, implies $T(-x) = -Tx$ and therefore $\|Tx - Ty\| = \|T(-x) + Ty\| \le \|x - y\|$. But also in this case our result is better than that of Baillon because we need not impose any convexity assumption on K.

Hirano and Takahashi [4] have replaced (1) by the weaker condition

$(1')$ $\qquad\qquad\qquad \|T^n x - T^n y\| \le \alpha_n \|x - y\| \qquad (x, y \in K),$

$\qquad\qquad\qquad\qquad \lim_{n \to \infty} \alpha_n = 1,$

and proved some weak convergence results for $A_n x$. If $-x \in K$ and $T(-x) = -Tx$ then $(1')$ implies

$(2')$ $\qquad\qquad\qquad \|T^n x + T^n y\| \le \alpha_n \|x + y\| \qquad (x, y \in K),$

$\qquad\qquad\qquad\qquad \lim_{n \to \infty} \alpha_n = 1.$

We will show that the mean ergodic theorem still holds if only $(2')$ is satisfied.

Even the results about weak convergence of $A_n x$ for nonexpansive mappings imply that the limit is a fixed point. As Example 3.2 shows this need no more hold under our assumptions even if T is continuous. Unlike the case of non-expansive operators the fixed point set need no more be convex, but it is always nonempty if K is a convex and closed set. In this case the unique point $x_0 \in K$ with $\|x_0\| = \inf\{\|x\|: x \in K\}$ is always a fixed point because $\|Tx_0\| \le \|x_0\|$ by (2).

For further information on nonlinear ergodic theorems the reader is referred to Bruck [2], §9.3 of Krengel [5] and the literature cited there.

After having finished a first draft of this paper Professors U. Krengel and M. Lin drew the author's attention to the unpublished preprint [3] of Djafari–Rouhani and Kakutani. There, the particular case $\delta_k = 0$ of our main result Theorem 2.3 is stated, with the additional assumption $\|X_{n+1} - X_{m+1}\| \le \|X_n - X_m\|$ for any $n, m \in \mathbf{N}$, but since the proof was incomplete, [3] was never published (although it contains another important result which was later generalized in [6]).

2. Main stream

In the sequel H will always be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$.

Theorem 2.1. *Let $K \subset H$ and $T \colon K \longrightarrow K$ be a mapping satisfying*

(i) $\qquad\qquad\qquad \|Tx + Ty\| \le \|x + y\| \qquad (x, y \in K).$

Then for any $x \in K$

$$A_n x := \frac{1}{n+1} \sum_{i=0}^{n} T^i x$$

is norm convergent.

Theorem 2.1 is a special case of

Theorem 2.2. *Let* $K \subset H$ *and* $T\colon K \longrightarrow K$ *be a mapping satisfying*

(i) $$\|T^n x + T^n y\| \leq \alpha_n \|x + y\| \qquad (x, y \in K),$$

$$\lim_{n \to \infty} \alpha_n = 1.$$

Then for any $x \in K$

$$A_n x := \frac{1}{n+1} \sum_{i=0}^{n} T^i x$$

is norm convergent.

Theorem 2.2 is an immediate consequence of

Theorem 2.3. *Let* (X_n) *be a sequence in* H *such that*

(i) $$\|X_{n+k} + X_{m+k}\|^2 \leq \|X_n + X_m\|^2 + \delta_k \qquad (m, n, k \in \mathbf{N})$$

with

(ii) $$\lim_{k \to \infty} \delta_k = 0.$$

Then the sequence A_n *defined by*

$$A_n := \frac{1}{n} \sum_{i=1}^{n} X_i$$

is norm convergent.

Remark. Clearly Theorem 2.3 contains also the mean ergodic theorem for stationary processes in the wide sense (cf. [5], p. 32).

Proof. From (i) and (ii) we see that for any $m, n \in N$

$$\limsup_{k \to \infty} \|X_{n+k} + X_{m+k}\| = \liminf_{k \to \infty} \|X_{n+k} + X_{m+k}\| = \inf_{k \in N} \|X_n + X_m\|$$

Taking $m = n = 1$ this implies

$$n_\varepsilon := \inf\{n \in \mathbf{N}\colon \|X_{m_1}\|^2 \leq \|X_{m_2}\|^2 + \varepsilon/2 \quad \forall m_2 \geq m_1 \geq n_\varepsilon\} < \infty.$$

For any $\varepsilon > 0$ we set

$$k_\varepsilon := \inf\{k \in \mathbf{N}\colon \delta_n \leq \varepsilon \quad \forall n \geq k\}.$$

Since

$$\langle X_n, X_m \rangle = \tfrac{1}{2}(\|X_n + X_m\|^2 - \|X_n\|^2 - \|X_m\|^2)$$

we obtain for any $m, n \geq n_\varepsilon$ and $k \geq k_\varepsilon$

(1) $$\langle X_{n+k}, X_{m+k} \rangle = \tfrac{1}{2}(\|X_{n+k} + X_{m+k}\|^2 - \|X_{n+k}\|^2 - \|X_{m+k}\|^2)$$

$$\leq \tfrac{1}{2}(\|X_n + X_m\|^2 + \varepsilon - \|X_n\|^2 + \tfrac{\varepsilon}{2} - \|X_m\|^2 + \tfrac{\varepsilon}{2})$$

$$= \langle X_n, X_m \rangle + \varepsilon.$$

For any $\varepsilon > 0$ we define

$$K_\varepsilon := \left\{ \sum_{i=0}^{p} t_i X_{n_\varepsilon + i} : p \in \mathbf{N}, \quad 0 \le t_i \le 1, \quad \sum_{i=0}^{p} t_i = 1 \right\},$$

$$I_\varepsilon = \inf_{z \in K_\varepsilon} \|z\|^2$$

and choose $p_\varepsilon \in \mathbf{N}$, $(t_{\varepsilon, i})_{0 \le i \le p_\varepsilon}$ in $[0, 1]$ such that $\sum_{i=0}^{p_\varepsilon} t_{\varepsilon, i} = 1$ and

$$(2) \qquad \left\| \sum_{i=0}^{p_\varepsilon} t_{\varepsilon, i} X_{n_\varepsilon + i} \right\|^2 \le I_\varepsilon + \varepsilon.$$

Using (1) we obtain for any $\varepsilon > 0$ and $k \ge k_\varepsilon$

$$\left\| \sum_{i=0}^{p_\varepsilon} t_{\varepsilon, i} X_{n_\varepsilon + k + i} \right\|^2 = \sum_{i, j=0}^{p_\varepsilon} t_{\varepsilon, i} t_{\varepsilon, j} \langle X_{n_\varepsilon + k + i}, X_{n_\varepsilon + k + j} \rangle$$

$$\le \sum_{i, j=0}^{p_\varepsilon} t_{\varepsilon, i} t_{\varepsilon, j} \langle X_{n_\varepsilon + i}, X_{n_\varepsilon + j} \rangle + \sum_{i, j=0}^{p_\varepsilon} t_{\varepsilon, i} t_{\varepsilon, j} \varepsilon$$

$$= \left\| \sum_{i=0}^{p_\varepsilon} t_{\varepsilon, i} X_{n_\varepsilon + i} \right\|^2 + \varepsilon.$$

Together with the convexity of $x \longrightarrow x^2$ and (2) this implies

$$(3) \qquad \|z_{\varepsilon, m}\|^2 \le I_\varepsilon + \varepsilon \qquad (m \in \mathbf{N}, \varepsilon > 0)$$

where

$$z_{\varepsilon, m} := \frac{1}{m + 1} \sum_{k=k_\varepsilon}^{k_\varepsilon + m} \sum_{i=0}^{p_\varepsilon} t_{\varepsilon, i} X_{n_\varepsilon + k + i}.$$

Note that $\varepsilon \longrightarrow I_\varepsilon$ is decreasing. Hence

$$I_0 := \lim_{\varepsilon \to 0} I_\varepsilon$$

exists and for any $\alpha > 0$ we can define

$$\varepsilon_\alpha := \sup\{\varepsilon > 0 : I_\varepsilon \ge I_0 - \alpha\}.$$

Now let $\alpha > 0$ and $0 < \delta < \varepsilon < \varepsilon_\alpha$ be given. Since $\frac{1}{2}(z_{\varepsilon, m} + z_{\delta, n}) \in K_\varepsilon$ we have

$$\|\tfrac{1}{2}(z_{\varepsilon, m} + z_{\delta, n})\|^2 \ge I_\varepsilon \ge \tfrac{1}{2}(I_\varepsilon + I_0 - \alpha) \ge \tfrac{1}{2}(I_\varepsilon + I_\delta - \alpha).$$

Combining with (3) we obtain

$$2\|z_{\varepsilon, m}\|^2 + 2\|z_{\delta, n}\|^2 - \|z_{\varepsilon, m} + z_{\delta, n}\|^2 \le 2I_\varepsilon + 2\varepsilon + 2I_\delta + 2\delta - 2I_\varepsilon - 2I_\delta + 2\alpha \le 4\varepsilon + 2\alpha.$$

By the parallelogram identity the left-hand side of the last inequality equals $\|z_{\varepsilon, m} - z_{\delta, n}\|^2$. Thus we have shown

$$\|z_{\varepsilon, m} - z_{\delta, n}\|^2 \le 4\varepsilon + 2\alpha \qquad (0 < \delta < \varepsilon < \varepsilon_\alpha, \; m, n \in \mathbf{N}).$$

Since H is a complete space there exists $z \in H$ with

$$(4) \qquad \|z_{\varepsilon,m} - z\|^2 \leq 4\varepsilon + 2\alpha \qquad (0 < \varepsilon < \varepsilon_\alpha, m \in \mathbf{N}).$$

From the definition of $z_{\varepsilon,m}$ it is easy to see that

$$(m+1)z_{\varepsilon,m} - \sum_{i=n_\varepsilon+p_\varepsilon+k_\varepsilon}^{n_\varepsilon+k_\varepsilon+m} X_i = \sum_{i=0}^{p_\varepsilon-1} r_i X_{n_\varepsilon+k_\varepsilon+i} + \sum_{i=1}^{p_\varepsilon} s_i X_{n_\varepsilon+k_\varepsilon+m+i}$$

with $0 \leq r_i, s_i \leq 1$ if $m \geq p_\varepsilon$. Hence

$$\|z_{\varepsilon,m} - A_{m+1}\| \leq \frac{1}{m+1} \left(2p_\varepsilon \sup_{n \in \mathbf{N}} \|X_n\| + (n_\varepsilon + k_\varepsilon) \sup_{n \in \mathbf{N}} \|X_n\| \right).$$

Letting m tend to infinity (4) now yields

$$\limsup_{n \to \infty} \|A_m - z\|^2 \leq 2 \sup_{m \in \mathbf{N}} \|z_{\varepsilon,m} - z\|^2 + 2 \limsup_{n \to \infty} \|A_{m+1} - z_{\varepsilon,m}\|^2 \leq 8\varepsilon + 4\alpha + 0.$$

Since $\alpha > 0$ and $\varepsilon > 0$ can be made arbitrarily small the assertion follows. $\square$

3. Examples

The first two examples below are quite pathological. Their main purpose is to show the difference to the case of nonexpansive operators.

Example 3.1. Let μ be a σ-finite measure and L_+^2 be the convex cone of all positive functions of the Hilbert space $L^2(\mu)$. Further let $\phi: L_+^2 \longrightarrow [0,1]$ be an arbitrary function. Then the mapping $T_\phi f := \phi(f)f$ clearly satisfies

$$\|T_\phi f + T_\phi g\| \leq \|f + g\| \qquad (f, g \in L_+^2).$$

On the other hand, even when L_+^2 is one dimensional, i.e. μ is a point mass, one can easily give examples of functions ϕ such that the fixed point set is not convex and T_ϕ is not continuous. Moreover, ϕ may be chosen such that T_ϕ is order preserving on L_+^2.

Even if T seems to be nice the ergodic averages may behave pathologically:

Example 3.2. We define

$$T(x_1, x_2) = \begin{cases} \frac{\max(x_1^2, x_2^2)}{x_1^2 + x_2^2}(x_2, x_1) & \text{if } (x_1, x_2) \in \mathbf{R}_+^2 \setminus \{(0,0)\}, \\ 0 & \text{if } (x_1, x_2) = (0,0). \end{cases}$$

Again we have

$$\|T(x_1, x_2) + T(y_1, y_2)\| \leq \|(x_1, x_2) + (y_1, y_2)\| \qquad ((x_1, x_2), (y_1, y_2) \in \mathbf{R}_+^2)$$

and T is continuous on $\mathbf{R}_+^2$ with $(0,0)$ as the only fixed point.

On the other hand, if $x > 0$ the averages

$$\frac{1}{n+1} \sum_{i=0}^{n} T^i(x, 0)$$

converge to $(\frac{1}{2}x, \frac{1}{2}x)$ which is not a fixed point.

We now turn over to more natural examples. In the sequel μ will be a σ-finite measure. The following lemma is basic:

Lemma 3.3. *Let* $X, Y, X', Y' \in L^4(\mu)$ *and* $\delta \geq 0$ *such that*

(i) $\|X\|_4^4 \leq \|X'\|_4^4 + \delta/2$,

(ii) $\|Y\|_4^4 \leq \|Y'\|_4^4 + \delta/2$,

(iii) $\|X + Y\|_4^4 \leq \|X' + Y'\|_4^4 + \delta$,

(iv) $\|X - Y\|_4^4 \leq \|X' - Y'\|_4^4 + \delta$.

Then we have

$$\|X^2 + Y^2\|_2^2 \leq \|X'^2 + Y'^2\|_2^2 + \delta.$$

Proof. For simplicity we use the probabilistic notation

$$E(Z) = \int Z \, d\mu.$$

From (iii) and (iv) we get

$$
\begin{aligned}
2E(X^4) + 2E(Y^4) + 12E(X^2Y^2) &= E((X + Y)^4) + E((X - Y)^4) \\
&\leq E((X' + Y')^4) + E((X' - Y')^4) + 2\delta \\
&= 2E(X'^4) + 2E(Y'^4) + 12E(X'^2Y'^2) + 2\delta
\end{aligned}
$$

and therefore

$$
\begin{aligned}
\|X^2 + Y^2\|_2^2 &= E(X^4) + E(Y^4) + 2E(X^2Y^2) \\
&= \tfrac{2}{3}(E(X^4) + E(Y^4)) + \tfrac{1}{6}(2E(X^4) + 2E(Y^4) + 12E(X^2Y^2)) \\
&\leq \tfrac{2}{3}(E(X'^4) + \delta/2 + E(Y'^4) + \delta/2) + \tfrac{1}{6}(2E(X'^4) + 2E(Y'^4) \\
&\quad + 12E(X'^2Y'^2) + 2\delta) \\
&= E(X'^4) + E(Y'^4) + 2E(X'^2Y'^2) + \delta = \|X'^2 + Y'^2\|_2^2 + \delta. \quad \square
\end{aligned}
$$

Together with Theorem 2.1 and Theorem 2.3 we obtain the following corollaries:

Corollary 3.4. *Let* (X_n) *be a sequence in* $L^4(\mu)$ *such that*

(i) $\|X_{n+k} + X_{m+k}\|_4^4 \leq \|X_n + X_m\|_4^4 + \delta_k$ $(m, n, k \in \mathbf{N})$,

(ii) $\|X_{n+k} - X_{m+k}\|_4^4 \leq \|X_n - X_m\|_4^4 + \delta_k$ $(m, n, k \in \mathbf{N})$,
with

(iii) $\lim_{k \to \infty} \delta_k = 0$.

Then the sequence

$$\frac{1}{n+1} \sum_{i=1}^{n} X_i^2$$

is norm convergent in $L^2(\mu)$.

Corollary 3.5. *Let* $K \subset L^4(\mu)$ *and* $T: K \longrightarrow K$ *be such that*

(i) $\|Tf - Tg\|_4 \leq \|f - g\|_4$ $(f, g \in K)$,

(ii) $\|Tf + Tg\|_4 \leq \|f + g\|_4$ $(f, g \in K)$.

Then for any $f \in K$

$$\frac{1}{n+1} \sum_{i=0}^{n} (T^i f)^2$$

is convergent in $L^2(\mu)$.

If moreover $K \subset L^4_+(\mu)$ *then*

$$f \longrightarrow (T(f^{1/2}))^2$$

defines a mapping $\widetilde{T}$ *of*

$$\widetilde{K} := \{f^2 : f \in K\}$$

into itself satisfying

$$\|\widetilde{T}f + \widetilde{T}g\|_2 \leq \|f + g\|_2 \qquad (f, g \in \widetilde{K}).$$

Remark. The mapping $\widetilde{T}$ in Corollary 3.5 need not be nonexpansive, even if T is a positive, linear $L^1 - L^\infty$ contraction. To see this, let $(E, \mathscr{B}, \mu)$ be a probability space and define $Tf = \int f \, d\mu \, I_E$. If there exists a measurable set A with $\frac{1}{2} \leq \mu(A) < 1$ then

$$\|\widetilde{T}I_E - \widetilde{T}I_A\|_2 = 1 - \mu(A)^2 \geq \tfrac{3}{2}(1 - \mu(A)) > \sqrt{1 - \mu(A)} = \|I_E - I_A\|_2.$$

Note added in proof. For odd mappings Corollary 3.4 can already be found in J. B. Baillon, R. E. Bruck, S. Reich, *On the asymptotic behavior of nonexpansive mappings and semigroups in Banach spaces*, Houston J. Math. **4** (1978), 1–9. Of course, in this case 3.5(ii) follows from 3.5(i).

REFERENCES

1. J.-B. Baillon, *Quelques propriétés de convergence asymptotique pour les contractions impaires*, C. R. Acad. Sci. Paris **283** (1976), 587–590.

2. R. E. Bruck, *Asymptotic behavior of nonexpansive mappings*, Contemp. Math. **18** (1983), 1–47.

3. B. Djafari–Rouhani and S. Kakutani, *Ergodic theorems for nonexpansive nonlinear operators in a Hilbert space* (unpublished preprint 1984).

4. N. Hirano and W. Takahashi, *Nonlinear ergodic theorems for nonexpansive mappings in Hilbert spaces*, Kodai Math. J. **2** (1975), 11–25.

5. U. Krengel, *Ergodic theorems*, deGruyter, Berlin, New York, 1985.

6. U. Krengel and M. Lin, *Order preserving nonexpansive operators in* L^1 , Israel J. Math. **58** (1987), 170–192.

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY, FINE HALL, WASHINGTON ROAD, PRINCETON, NEW JERSEY 08544-1000