

A VECTOR-VALUED RANDOM ERGODIC THEOREM

ANATOLE BECK AND J. T. SCHWARTZ

The first object of this paper is to give a proof of the following theorem, which generalizes a result (unpublished) of S. Kakutani proven by other means. This result is then applied to a problem in Banach spaces which generalizes the strong law of large numbers.

I. 1. DEFINITION. Let \mathfrak{X} be a B -space and (S, Σ, m) a measure space.¹ Let $L_p(S, \Sigma, m, \mathfrak{X})$, $1 \leq p < \infty$, hereinafter called simply $L_p(S, \mathfrak{X})$, denote the space of all strongly measurable \mathfrak{X} -valued functions X defined on S for which the norm $\|X\|_p = (\int_S \|X(s)\|_{\mathfrak{X}}^p m(ds))^{1/p} < \infty$.

2. THEOREM. Let \mathfrak{X} be a reflexive B -space and let (S, Σ, m) be a σ -finite measure space. Let there be defined on S a strongly measurable function T_s with values in the B -space $B(\mathfrak{X})$ of bounded linear operators on \mathfrak{X} . Suppose that $\|T_s\| \leq 1$ for all $s \in S$. Let h be a measure-preserving transformation (m.p.t.) in (S, Σ, m) . Then for each $X \in L_1(S, \mathfrak{X})$ there is an $\bar{X} \in L_1(S, \mathfrak{X})$ such that $\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n T_s T_{h(s)} \cdots T_{h^{i-1}(s)} \cdot (X(h^i(s))) = \bar{X}(s)$ strongly in \mathfrak{X} a.e. in S ,² and $\bar{X}(s) = T_s(\bar{X}(h(s)))$ a.e. in S . Moreover, if $m(S) < \infty$, then \bar{X} is also the limit in the mean of order 1.

It should be noted that this statement includes the case in which $X(s)$ is constant and $\{T_{h^i(s)}\}$, $i=1, 2, \dots$, is an independent sequence, in the sense of probability.³

PROOF. $L_2(S, \mathfrak{X})$ is a reflexive B -space.⁴ Define the operator U in $L_2(S, \mathfrak{X})$ by $UX(s) = T_s(X(h(s)))$, $s \in S$, $X \in L_2(S, \mathfrak{X})$. Then U is clearly linear. Moreover,

Presented to the Society February 25, 1956; received by the editors August 28, 1955 and, in revised form, February 8, 1957.

¹ We shall make free use throughout this paper of the terminology introduced in *Measure theory* by Paul R. Halmos (Van Nostrand, 1950). As far as the terminology and theory of integration of vector-valued functions goes, we follow E. Hille, *Functional analysis and semi-groups*, Amer. Math. Soc. Colloquium, Publications, no. 31, especially pp. 40-50.

² In this paper, all convergence is as $n \rightarrow \infty$, unless otherwise noted.

³ Cf. S. Kakutani, *Ergodic theory*, Proceedings of the International Congress of Mathematicians, Cambridge, vol. 2, 1950, pp. 128-141.

⁴ J. Dieudonné, *Sur le théorème de Lebesgue-Nikodym (V)*, Canadian J. Math. vol. 3 (1951) pp. 129-139; also R. S. Phillips, *On weakly compact subsets of a Banach space*, Amer. J. Math. vol. 65 (1943) pp. 108-136.

$$\begin{aligned} \|UX\| &= \left(\int_S \|T_s(X(h(s)))\|^2 m(ds) \right)^{1/2} \\ &\leq \left(\int_S \|X(h(s))\|^2 m(ds) \right)^{1/2} \\ &= \left(\int_S \|X(s)\|^2 m(ds) \right)^{1/2} \\ &= \|X\|, \end{aligned}$$

so that $\|U\| \leq 1$. By the mean ergodic theorem of Yosida and Kakutani,⁵ we find that $L_2(S, \mathfrak{X}) = \mathfrak{M} \oplus \mathfrak{N}$, where $UX = X$ for all⁶ $X \in \mathfrak{M}$ and \mathfrak{N} is the closure of the range of $I - U$.

Let \mathfrak{A} be the set of elements of $L_2(S, \mathfrak{X})$ which are (essentially) bounded in norm for all $s \in S$. Then \mathfrak{A} is dense in $L_2(S, \mathfrak{X})$, so that $\mathfrak{M} \oplus (I - U)\mathfrak{A}$ is dense in $L_2(S, \mathfrak{X})$. Suppose $X = X_1 + X_2$, where $X_1 = UX_1$ and $X_2 = (I - U)Y$, $Y \in \mathfrak{A}$. Then

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n U^i(X) &= \frac{1}{n} \sum_{i=1}^n U^i(X_1) + \frac{1}{n} \sum_{i=1}^n U^i(X_2) \\ &= \frac{1}{n} \sum_{i=1}^n X_1 + \frac{1}{n} \sum_{i=1}^n U^i(I - U)(Y) \\ &= X_1 + \frac{1}{n} (U - U^{n+1})(Y). \end{aligned}$$

Therefore,

$$\begin{aligned} \left\| X_1(s) - \frac{1}{n} \sum_{i=1}^n U^i X(s) \right\| &= \left\| \frac{1}{n} (U - U^{n+1})(Y(s)) \right\| \\ &\leq \frac{2}{n} \left(\text{ess sup}_S \|Y(s)\| \right), \end{aligned}$$

for almost all $s \in S$, so that

$$\left\| X_1(s) - \frac{1}{n} \sum_{i=1}^n U^i X(s) \right\| \rightarrow 0 \text{ uniformly in } S.$$

Moreover, for all $X \in L_2(S, \mathfrak{X})$, we have

$$\begin{aligned} \|U^i X(s)\| &= \|T_s T_{h(s)} \cdots T_{h^{i-1}(s)}(X(h^i(s)))\| \\ &\leq \|X(h^i(s))\| \qquad \text{for all } s \in S. \end{aligned}$$

⁵ K. Yosida and S. Kakutani, *Operator theoretical treatment of Markoff's process and the mean ergodic theorem*, Ann. of Math. vol. 42 (1941) pp. 188-228.

⁶ $UX = X$ is equivalent to the assertion that $T_s(X(h(s))) = X(s)$ a.e. in S .

Therefore,

$$\begin{aligned} \sup_{0 < n < \infty} \left\| \frac{1}{n} \sum_{i=1}^n U^i X(s) \right\| &\leq \sup_{0 < n < \infty} \frac{1}{n} \sum_{i=1}^n \|X(h^i(s))\| \\ &\leq \infty \text{ a.e. in } S \end{aligned}$$

by the Birkhoff ergodic theorem, since

$$\int_S \|X(s)\|^2 m(ds) < \infty \text{ and } h \text{ is a m.t.p. in } (S, \Sigma, m).$$

At this point, we make use of the following

THEOREM (BANACH). *Let $\{T_n\}$ be a sequence of continuous linear maps from a B -space \mathfrak{Y} into the space $L_p(S, \mathfrak{X})$. Suppose that for each $y \in \mathfrak{Y}$, we have $\sup_n \|(T_n(y))(s)\| < \infty$ for almost all $s \in S$. Suppose that for each y in a dense set \mathfrak{Y}_0 in \mathfrak{Y} , the limit $\lim_n (T_n(y))(s)$ exists for almost all $s \in S$. Then for every $y \in \mathfrak{Y}$, the limit $\lim_n (T_n(y))(s)$ exists a.e. in S .⁷*

If in Banach's theorem, we take $\mathfrak{Y} = L_2(S, \mathfrak{X})$ and take for T_n the mapping from $L_2(S, \mathfrak{X})$ into $L_2(S, \mathfrak{X})$ which is defined by $T_n = n^{-1} \cdot \sum_{i=1}^n U^i$, then $n^{-1} \sum_{i=1}^n U^i X(s)$ has a limit a.e. in S for all $X \in \mathfrak{Y}_0 = \mathfrak{M} \oplus (I - U)\mathfrak{A}$, a dense set in $L_2(S, \mathfrak{X})$. Also,

$$\sup_n \|(T_n X)(s)\| = \sup_{0 < n < \infty} \left\| \frac{1}{n} \sum_{i=1}^n U^i X(s) \right\| < \infty \text{ a.e. in } S,$$

as proved above. Therefore, it follows immediately that $n^{-1} \cdot \sum_{i=1}^n U^i X(s)$ has a limit a.e. in S for all $X \in L_2(S, \mathfrak{X})$.

Let us now look at $L_1(S, \mathfrak{X})$. Since the limit $\lim n^{-1} \sum_{i=1}^n U^i X(s)$ exists a.e. in S for all $X \in L_2(S, \mathfrak{X})$, it exists in particular for every simple function X . However, if Y be any element of $L_1(S, \mathfrak{X})$, we have as before

$$\sup_{0 < n < \infty} \left\| \frac{1}{n} \sum_{i=1}^n U^i Y(s) \right\| \leq \sup_{0 < n < \infty} \frac{1}{n} \sum_{i=1}^n \|Y(h^i(s))\| < \infty \text{ a.e. in } S$$

by the Birkhoff ergodic theorem. Since the simple functions are dense

⁷ This theorem was stated by Banach for the special case of real-valued functions, but the result may easily be extended to the vector-valued case. The reader is referred to *Linear operators* by N. Dunford and J. Schwartz, New York, Interscience Press (forthcoming) Vol. I. where a proof, stated for the complex-valued functions but valid verbatim for the vector-valued case, is given. The original Banach theorem is proven in his *Sur la convergence presque partout des fonctionnelles lineaires*, Bull. Sci. Math. vol. 50 (1926) pp. 27-32, 36-43.

in $L_1(S, \mathfrak{X})$, another application of the Banach theorem yields the first part of this theorem.

If X is a simple function, $X \in L_2(S, \mathfrak{X})$, so that $U\bar{X} = \bar{X}$, i.e. $T_s(\bar{X}(h(s))) = \bar{X}(s)$ a.e. in S . Thus, the bounded linear mappings $X \rightarrow \bar{X}$ and $X \rightarrow T_s(\bar{X}(h(\cdot)))$ agree on a dense set in $L_1(S, \mathfrak{X})$. Thus, $\bar{X}(s) = T_s(\bar{X}(h(s)))$ a.e. in S for all $X \in L_1(S, \mathfrak{X})$.

If $m(S) < \infty$, then the convergence in the mean of order 2 implies convergence in the mean of order 1 for all X in the dense subset $L_2(S, \mathfrak{X})$ of $L_1(S, \mathfrak{X})$. It follows immediately from the uniform boundedness of the operators $n^{-1} \sum_{i=1}^n U^i$ that $\lim n^{-1} \sum_{i=1}^n U^i(X)$ exists in the mean of order 1 for all $X \in L_1(S, \mathfrak{X})$. Q.E.D.

Adapting the usual concepts of probability theory to the present context, we make the following definition.

3. DEFINITION. Let \mathfrak{X} be a B -space and let $(\Omega, \mathfrak{B}, \text{Pr})$ be a probability space, i.e. a measure space of total measure 1. If X is a strongly measurable function defined on Ω with values in \mathfrak{X} , X will be called a *strongly measurable \mathfrak{X} -valued random variable* (defined on Ω). If $X \in L_1(\Omega, \mathfrak{X})$, X will be called a *strongly integrable \mathfrak{X} -valued random variable*. If $\{X_i\}$, $-\infty < i < \infty$, is a two-sided sequence of strongly measurable \mathfrak{X} -valued random variables such that

$$\text{Pr} \{ \omega \mid X_i(\omega) \in A_i, |i| \leq n \} = \text{Pr} \{ \omega \mid X_{i+1}(\omega) \in A_i, |i| \leq n \}$$

for each finite collection $\{A_i\}$, $|i| \leq n$, of Borel subsets of \mathfrak{X} , the sequence $\{X_i\}$ is said to be *stationary*. If $\text{Pr} \{ \omega \mid X_i(\omega) \in A_i, |i| \leq n \} = \prod_{i=-n}^n \text{Pr} \{ \omega \mid X_i(\omega) \in A_i \}$ for each such finite collection $\{A_i\}$, the sequence $\{X_i\}$ is said to be *independent*.

A number of other concepts and terms of probability theory which carry over without change from the ordinary case of real or complex valued random variables will be used freely in what follows.

We are especially interested in the following corollary.

4. COROLLARY (TO THEOREM 2). Let \mathfrak{X} be a reflexive B -space. Let T be a linear operator in \mathfrak{X} with $\|T\| \leq 1$. Let $\{X_i\}$, $-\infty < i < \infty$, be a stationary sequence of strongly integrable \mathfrak{X} -valued random variables. Then there is a strongly integrable \mathfrak{X} -valued random variable Y such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n T^i X_i = Y \text{ almost surely.}$$

PROOF. From the definition of strong measurability,⁸ it is clear that there is a closed separable subspace \mathfrak{X}_0 of \mathfrak{X} such that each X_i , $-\infty < i < \infty$, lies in \mathfrak{X}_0 with probability 1. Thus we can assume with-

⁸ E. Hille, *loc. cit.*

out loss of generality that \mathfrak{X} is separable. Let $(\Omega, \mathfrak{B}, \text{Pr})$ be the measure space on which all the random variables X_i are defined. Let $\{O_j\}$ be a countable basis for the open sets of \mathfrak{X} , and let \mathfrak{B}_0 be the σ -field of subsets of Ω generated by the countable family of sets $\{\omega \mid X_i(\omega) \in O_j\}$, $-\infty < i < \infty, j = 1, 2, \dots$. Then $\{X_i\}$ may clearly be regarded as a stationary sequence of strongly measurable \mathfrak{X} -valued random variables on the measure space $(\Omega, \mathfrak{B}_0, \text{Pr})$. This measure space is separable in the sense of Halmos.

There is no loss of generality involved in the assumption that $\mathfrak{B}_0 = \mathfrak{B}$. Let $(\mathfrak{B}(\text{Pr}), \text{Pr})$ denote the measure algebra associated with $(\Omega, \mathfrak{B}, \text{Pr})$. Let $\{\Omega_i\}$ be an enumeration of the atoms of \mathfrak{B} and let Ω_∞ be the complement in Ω of $\bigcup_{i=1}^\infty \Omega_i$.

It is clear from the definition of \mathfrak{B}_0 (which equals \mathfrak{B}) that there is an automorphism of the measure algebra $(\mathfrak{B}(\text{Pr}), \text{Pr})$ which induces in $L_1(\Omega, \mathfrak{B}, \text{Pr}, \mathfrak{X})$ an isometry which takes X_i into X_{i+1} for every $-\infty < i < \infty$. Call this automorphism H and denote by U the isometry it generates in $L_1(\Omega, \mathfrak{X})$.

It is clear that H takes each atom into an atom of equal measure. Thus, H rotates the atoms, each one taking only a finite number of positions, and these in order. Thus, the theorem is completely trivial in the union of the atoms and it only remains to prove it in the atom-free part Ω_∞ .

In order to simplify our notation, we will assume that $\Omega = \Omega_\infty$, i.e. that $\mathfrak{B} = \mathfrak{B}_0$ is atom-free. Then, by a well-known theorem,⁹ $(\Omega, \mathfrak{B}, \text{Pr})$ is isomorphic with (I, Σ, μ) , the Lebesgue measure space of the real unit interval. This isomorphism induces

(i) an automorphism \widehat{H} of the measure algebra of (I, Σ, μ) corresponding to H .

(ii) an isometric isomorphism between $L_1(\Omega, \mathfrak{X})$ and $L_1(I, \mathfrak{X})$ which preserves convergence in measure and almost uniform convergence.¹⁰

Furthermore, if $X_i \rightarrow \widehat{X}_i$ under the isomorphism (ii) above, then the isometry induced in $L_1(I, \mathfrak{X})$ by \widehat{H} takes \widehat{X}_i into \widehat{X}_{i+1} .

At this point, we make use of a theorem of von Neumann,¹¹ which tells us that there is a measure-preserving (point) transformation h of I into itself so that for each set $A \in \Sigma$,

$$h(A) = \widehat{H}(A) \text{ modulo a set of 0 measure.}$$

⁹ Cf. P. R. Halmos, *op. cit.* pp. 173-174.

¹⁰ For the definition of almost uniform convergence and the fundamental theorems about it, used herein, see E. Hille, *loc. cit.*

¹¹ J. von Neumann, *Einige Sätze über messbare Abbildungen*, Ann. of Math. vol. 33 (1932) pp. 547-586, especially p. 582.

Then $\hat{X}_{i+1}(t) = \hat{X}_i(h(t))$ for almost all $t \in I$. By Theorem 2, there is an $\bar{X} \in L_1(I, \mathfrak{X})$ such that

$$\begin{aligned} \lim \frac{1}{n} \sum_{i=1}^n T^i \hat{X}_i(t) &= \lim \frac{1}{n} \sum_{i=1}^n T^i \hat{X}_0(h^i(t)) \\ &= \bar{X}(t) \text{ a.e. in } I. \end{aligned}$$

Then

$$\frac{1}{n} \sum_{i=1}^n T^i \hat{X}_i \rightarrow \bar{X} \text{ almost uniformly,}$$

since all the \hat{X}_i and \bar{X} are strongly measurable and $\mu(I) < \infty$.

Let $Y \in L_1(\Omega, \mathfrak{X})$ be the pre-image of \bar{X} under the isomorphism (ii). Then

$$\frac{1}{n} \sum_{i=1}^n T^i X_i \rightarrow Y \text{ almost uniformly.}$$

Therefore,

$$\frac{1}{n} \sum_{i=1}^n T^i X_i \rightarrow Y \text{ almost surely. Q.E.D.}$$

II. There are two important cases in which it is possible to evaluate the limit in Theorem 2 and Corollary 4. The first is the case in which the operator function of Theorem 2 is everywhere equal to the identity. The other is the case in which the random variables X_i of Corollary 4 are independent. We shall need

5. DEFINITION. Let (S, Σ, m) be a measure space of total measure 1, and let Σ' be a sub- σ -field of the σ -field Σ . Then (S, Σ', m) is a measure space and $L_1(S, \Sigma', m, \mathfrak{X}) \subset L_1(S, \Sigma, m, \mathfrak{X})$. If $X \in L_1(S, \Sigma, m, \mathfrak{X})$ and $X_0 \in L_1(S, \Sigma', m, \mathfrak{X})$ and

$$\int_A X_0(s)m(ds) = \int_A X(s)m(ds)$$

for every $A \in \Sigma'$, then we write $X_0 = E(X | \Sigma')$ and call X_0 the conditional expectation of X with respect to Σ' . If Σ' contains only sets of measure 0 and 1, so that $E(X | \Sigma') = \int_S X(s)m(ds)$ a.e. in S , we will use the symbol $E(X)$ ambiguously to mean either $E(X | \Sigma')$ or $\int_S X(s)m(ds)$.

6. THEOREM. Let \mathfrak{X} be a reflexive B -space and let (S, Σ, m) be a measure space with $m(S) = 1$. Let h be a *m.p.t.* in S and denote by Σ_0

the collection of all sets S_0 in Σ for which $h(S_0) = S_0$. Let $X \in L_1(S, \mathfrak{X})$. Then if we set $\bar{X} = E(X | \Sigma_0)$,

$$\lim \frac{1}{n} \sum_{i=1}^n X(h^i(s)) = \bar{X}(s) \text{ a.e. in } S.$$

PROOF. By Theorem 2, the limit above exists a.e. in S . If we can show it has the properties of $E(X | \Sigma_0)$, we will be through. We know that $\bar{X}(s) = \bar{X}(h(s))$ a.e. so that \bar{X} is Σ_0 -measurable. Let A be a set invariant under h . Then, for all finite n ,

$$\begin{aligned} \int_A \frac{1}{n} \sum_{i=1}^n X(h^i(s))m(ds) &= \frac{1}{n} \sum_{i=1}^n \int_A X(h^i(s))m(ds) \\ &= \frac{1}{n} \sum_{i=1}^n \int_A X(s)m(ds) \\ &= \int_A X(s)m(ds). \end{aligned}$$

Since, according to Theorem 2, $\lim n^{-1} \sum_{i=1}^n X(h^i(\cdot))$ exists in the mean of order 1, we may let $n \rightarrow \infty$ in the above equation and find that

$$\int_A \bar{X}(s)m(ds) = \int_A X(s)m(ds).$$

The desired result now follows directly from Definition 5. Q.E.D.

7. COROLLARY. *If, in addition to the hypothesis of Theorem 6, we assume that h is ergodic, i.e. leaves no set of measure other than 0 or 1 invariant, then*

$$\bar{X} = E(X) = \int_S X(s)m(ds) \text{ a.e. in } S.$$

PROOF. Clear.

Note the similarity of this corollary to the strong law of large numbers.

The other special case we will discuss deals with the situation when the sequence $\{X_i\}$ of Corollary 4 is not only stationary but independent as well. In the following lemmas, T is a bounded linear operator in the reflexive B -space \mathfrak{X} , and $\|T\| \leq 1$. L is the strong limit of $n^{-1} \sum_{i=1}^n T^i$. $\{a_i\}$ is an independent sequence of random variables and

$$\Pr \{ a_i = 1 \} = \alpha, \quad \Pr \{ a_i = 0 \} = 1 - \alpha.$$

The two-sided sequence $\{X_i\}$, $-\infty < i < \infty$, is an independent sequence of random variables with the same distribution (and therefore a stationary sequence). We intend to prove the following generalization of the mean ergodic theorem:

$$\frac{1}{n} \sum_{i=1}^n T^i X_i \rightarrow E(LX_0) \text{ almost surely.}$$

8. LEMMA. *If $\{c_i\}$ is a bounded sequence of complex numbers and $n^{-1} \sum_{i=1}^n c_i \rightarrow 0$, then $n^{-1} \sum_{i=1}^n a_i c_i \rightarrow 0$ almost surely.*

PROOF. Let $b_i = a_i c_i - \alpha c_i$, $i = 1, 2, \dots$. Then $\{b_i\}$ is a bounded independent sequence of random variables with 0 mean. By the strong law of large numbers,

$$\frac{1}{n} \sum_{i=1}^n a_i c_i - \frac{1}{n} \sum_{i=1}^n \alpha c_i = \frac{1}{n} \sum_{i=1}^n (a_i c_i - \alpha c_i) = \frac{1}{n} \sum_{i=1}^n b_i \rightarrow 0 \text{ almost surely.}$$

Since

$$\lim \frac{1}{n} \sum_{i=1}^n \alpha c_i = \alpha \lim \frac{1}{n} \sum_{i=1}^n c_i = 0,$$

$$\lim \frac{1}{n} \sum_{i=1}^n a_i c_i = 0 \text{ almost surely. Q.E.D.}$$

9. LEMMA. *If $\{x_i\}$ is a sequence of elements in \mathfrak{X} and $n^{-1} \sum_{i=1}^n x_i$ converges weakly to 0, then $n^{-1} \sum_{i=1}^n a_i x_i$ almost surely converges weakly to 0.*

PROOF. Let \mathfrak{X}' be the closed linear subspace generated by the elements x_i . We may clearly assume without loss of generality that $\mathfrak{X}' = \mathfrak{X}$. Then the dual \mathfrak{X}^* of \mathfrak{X} is separable, since \mathfrak{X} is reflexive and separable. For each $x^* \in \mathfrak{X}^*$, $x^*(n^{-1} \sum_{i=1}^n a_i x_i) = n^{-1} \sum_{i=1}^n a_i (x^* x_i) \rightarrow 0$ by Lemma 8, since $n^{-1} \sum_{i=1}^n x^* x_i = x^*(n^{-1} \sum_{i=1}^n x_i) \rightarrow 0$. Since \mathfrak{X}^* is separable, $n^{-1} \sum_{i=1}^n a_i x_i$ almost surely converges weakly to 0. Q.E.D.

10. LEMMA. *If $x \in \mathfrak{X}$, then $n^{-1} \sum_{i=1}^n a_i T^i x$ almost surely converges weakly to αLx .*

PROOF. By the mean ergodic theorem, $x = x_1 + x_2$, where $Tx_1 = x_1 = Lx$ and $n^{-1} \sum_{i=1}^n T^i x_2 \rightarrow 0$. Then, *a fortiori*, $n^{-1} \sum_{i=1}^n T^i x_2$ converges weakly to 0. Hence

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n a_i T^i x &= \frac{1}{n} \sum_{i=1}^n a_i T^i x_1 + \frac{1}{n} \sum_{i=1}^n a_i T^i x_2 \\ &= \frac{1}{n} \sum_{i=1}^n a_i x_1 + \frac{1}{n} \sum_{i=1}^n a_i T^i x_2 \end{aligned}$$

almost surely converges weakly to $\alpha x_1 = \alpha Lx$ by the strong law of large numbers and Lemma 9. Q.E.D.

11. LEMMA. $n^{-1} \sum_{i=1}^n T^i X_i$ converges strongly to $E(LX_0)$ almost surely if the X_i take on only a finite number of values.

PROOF. Let x_1, \dots, x_m be the values taken by the X_i , $-\infty < i < \infty$. Then we define random variables $a_{i,j}$, $-\infty < i < \infty$, $j=1, \dots, m$ by

$$\begin{aligned} a_{i,j} &= 1 && \text{if } X_i = x_j, \\ &= 0 && \text{if } X_i \neq x_j. \end{aligned}$$

Since $\{X_i\}$ is a stationary sequence of independent random variables, the same is true for the sequence $\{a_{i,j}\}$ for each j . We have

$$X_i = \sum_{j=1}^m a_{i,j} x_j.$$

Let $\alpha_j = \Pr \{a_{i,j} = 1\}$, $j=1, \dots, m$. Then $\sum_{j=1}^m \alpha_j = 1$. It follows by Lemma 10 that

$$\frac{1}{n} \sum_{i=1}^n T^i X_i = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^m a_{i,j} T^i x_j = \sum_{j=1}^m \frac{1}{n} \sum_{i=1}^n a_{i,j} T^i x_j$$

almost surely converges weakly to

$$\sum_{j=1}^m \alpha_j Lx_j = \sum_{j=1}^m (Lx_j)(\Pr\{X_0 = x_j\}) = E(LX_0).$$

Since, by Corollary 4, $n^{-1} \sum_{i=1}^n T^i X_i$ almost surely converges strongly, the present lemma is proved. Q.E.D.

12. THEOREM. Let \mathfrak{X} be a reflexive B -space and let $\{X_i\}$, $-\infty < i < \infty$ be an independent sequence of \mathfrak{X} -valued random variables having the same distribution. Let T denote a bounded linear operator in \mathfrak{X} , and $\|T\| \leq 1$. Let L denote the strong limit of $n^{-1} \sum_{i=1}^n T^i$. Then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n T^i X_i$$

exists and is equal to $E(LX_0)$ almost surely.

PROOF. As in the proof of Corollary 4, we can show that we may assume without loss of generality that \mathfrak{X} is separable. Choose any $\epsilon > 0$ and let $\{B_j\}$ be a partition of \mathfrak{X} into a countable family of disjoint Borel sets of diameter less than ϵ and let $\{x_j\}$ be a sequence of elements in \mathfrak{X} such that $x_j \in B_j, j = 1, 2, \dots$. Let $X^{(n)}$ be a random variable defined as follows:

$$X^{(n)} = X_0 \text{ if } X_0 \in B_n, \\ = 0 \text{ otherwise.}$$

Then $X_0 = \sum_{j=1}^\infty X^{(j)}$ and in fact $\sum_{j=1}^n X^{(j)}$ converges to X_0 almost uniformly and therefore in the norm of $L_1(\Omega, \mathfrak{X})$. Choose N so large that

$$\left\| \sum_{j=1}^N X^{(j)} - X_0 \right\|_1 = E \left(\left\| \sum_{j=N+1}^\infty X^{(j)} \right\| \right) < \epsilon.$$

Let \bar{X}_i be the random variable defined by the requirement that $\bar{X}_i = x_j$ if $X_i \in B_j, j = 1, \dots, N; \bar{X}_i = 0$ otherwise. Then clearly $\{\bar{X}_i\}$ is a stationary sequence to which Lemma 11 applies. Thus,

$$\limsup \left(\left\| \frac{1}{n} \sum_{i=1}^n T^i \bar{X}_i - E(L\bar{X}_0) \right\| \right) = 0.$$

Moreover $E(\|X_i - \bar{X}_i\|) < 2\epsilon$. Since $\|T\| \leq 1, \|L\| \leq 1$. Hence, it also follows that $\|E(L\bar{X}_0) - E(LX_0)\| \leq 2\epsilon$. Also, since $\{\|X_i - \bar{X}_i\|\}$ is a stationary sequence of independent real-valued random variables, we may apply the strong law of large numbers to conclude that

$$\lim \frac{1}{n} \sum_{i=1}^n \|X_i - \bar{X}_i\| = E(\|X_0 - \bar{X}_0\|) < 2\epsilon$$

almost surely. We know that

$$\frac{1}{n} \sum_{i=1}^n \|T^i X_i - T^i \bar{X}_i\| \leq \frac{1}{n} \sum_{i=1}^n \|X_i - \bar{X}_i\|.$$

Therefore, $\limsup \|n^{-1} \sum_{i=1}^n T^i X_i - E(LX_0)\| < 2\epsilon + 0 + 2\epsilon = 4\epsilon$ almost surely. Since ϵ is arbitrary, the present theorem follows immediately. Q.E.D.

It would be convenient if this same evaluation of the limit of $n^{-1} \sum_{i=1}^n T^i X_i$ could be made on the hypothesis that the X_i form a stationary sequence. Using the notation of Corollary 4 and Theorem 6, we should like to be able to say that $\bar{X} = E(LX_0 | \mathfrak{B}')$, where \mathfrak{B}' is the field of sets invariant under h . The following example will show

that this is not always true, even if \mathcal{B}' contains only sets of probability 0 and 1. We shall use an ergodic transformation in the unit interval to generate our stationary sequence.

13. EXAMPLE. Let h_1 be an ergodic m.p.t. in the Lebesgue measure space of the interval $(0, 1/2)$. Define the map h of $(0, 1)$ into itself by

$$\begin{aligned} h(s) &= s + 1/2, & 0 < s < 1/2, \\ &= 1/2, & s = 1/2, \\ &= h_1(s - 1/2), & 1/2 < s < 1. \end{aligned}$$

Then h is ergodic and the only sets invariant under h have measure 0 or 1. Let \mathfrak{X} be an arbitrary B -space and define T by $T(x) = -x$, all $x \in \mathfrak{X}$. Let y be any nonzero element of \mathfrak{X} and define the random variable X by putting

$$\begin{aligned} X(s) &= y, & 0 < s < 1/2, \\ &= -y, & 1/2 \leq s < 1. \end{aligned}$$

Then $\lim n^{-1} \sum_{i=1}^n T^i X(h^i(s)) = \lim n^{-1} \sum_{i=1}^n X(s) = X(s) \neq 0$ while $L=0$, so that $E(LX)=0$. Difficulties of this sort, arising from the spectral properties of h and T , are dealt with in a forthcoming paper by Beck.

YALE UNIVERSITY