

# Markov Operators and the Nevo–Stein Theorem

Alexander Bufetov

Vienna, Preprint ESI 1077 (2001)

September 26, 2001

Supported by Federal Ministry of Science and Transport, Austria  
Available via <http://www.esi.ac.at>

# Markov operators and the Nevo – Stein theorem

Alexander Bufetov

## 1 Introduction

In this note the Nevo – Stein Ergodic Theorem for actions of free groups is derived from Rota's "Alternierende Verfahren" for Markov operators.

Pointwise ergodic theorems for actions of arbitrary countable groups were first obtained by V.I. Oseledec [1] in 1965, who took a probability distribution on the group, averaged the group elements with respect to convolutions of this distribution and under certain conditions proved almost everywhere convergence of these averages to a group-invariant function.

In 1969 Y. Guivarc'h [2] (motivated by Arnold and Krylov [3]) studied spherical averages in the free group. Let  $(X, \nu)$  be a probability space and suppose a free group  $F_m$  with  $m$  generators acts on  $(X, \nu)$  by measure-preserving transformations. Fix a set  $a_1, \dots, a_m$  of free generators in our group. For an element  $g \in F_m$  let  $|g|$  be the *length* of  $g$ , that is, the length of the irreducible word representing  $g$  over the alphabet  $a_1, \dots, a_m, a_1^{-1}, \dots, a_m^{-1}$ . Let  $C(n)$  be the sphere of radius  $n$  in our group, that is,  $C(n) = \{g \in F_m : |g| = n\}$ .  $C(n)$  has  $(2m)(2m-1)^{n-1}$  elements. For  $g \in F_m$ , let  $T_g$  be the transformation corresponding to  $g$ . For  $\varphi \in L_1(X, \nu)$  and  $g \in F_m$  write  $T_g\varphi = \varphi \circ T_g^{-1}$ .

Set

$$s_n\varphi = \frac{1}{2m(2m-1)^{n-1}} \sum_{g \in C(n)} T_g\varphi \quad (1)$$

Denote by  $F_m^2$  the subgroup of elements of even length in  $F_m$  with respect to generators  $\{a_1, \dots, a_m\}$  (in other words,  $F_m^2$  is generated by the elements  $a_i a_j$ ,  $i, j \in \{1, \dots, m\}$ ). The theorem of Y. Guivarc'h [2] says that for  $\varphi \in L_2(X, \nu)$  the sequence  $s_{2n}\varphi$  converges in  $L_2$  to a function invariant under the action of  $F_m^2$ .

In 1986 R.I. Grigorchuk [4] announced that for  $\varphi \in L_1$  the sequence  $c_N\varphi = \frac{1}{N} \sum_{n=0}^{N-1} s_n\varphi$  converges almost everywhere to an  $F_m$ -invariant function; in 1994 Nevo and Stein [5] published a proof of this theorem.

Let  $\mathcal{I}_2$  be the sigma-algebra of subsets of  $X$  invariant under the action of the subgroup  $F_m^2$ . In 1994 A. Nevo and E.M. Stein [5] proved

**Theorem 1** *Let  $p > 1$  and let  $\varphi \in L_p(X, \nu)$ . Then  $s_{2n}\varphi \rightarrow E(\varphi|\mathcal{I}_2)$  almost everywhere as  $n \rightarrow \infty$ .*

In this note, we give another proof of Theorem 1 and prove a slightly stronger

**Theorem 2** *Theorem 1 holds for  $\varphi \in L \log L(X, \nu)$ .*

(recall that  $\varphi \in L \log L(X, \nu)$  means  $|\varphi| \log |\varphi| \in L_1$ ).

To prove this theorem, we assign to the group action a special Markov operator  $P$  and then Theorem 2 follows from Rota's "Alternierende Verfahren" applied to  $P$ . This proof is given in Sec.2. In Sec.3, we prove convergence of even powers of  $P$  and give a criterion of the triviality of tail sigma-algebra of  $P$ .

## 2 Proof of Theorem 2

Recall that we have fixed a set  $a_1, \dots, a_m$  of free generators in our group; let  $\mathcal{A}$  be the set of the generators and their inverses:  $\mathcal{A} = \{a_1, \dots, a_1^{-1}, \dots, a_m^{-1}\}$ .

For  $g \in F_m$ ,  $|g| = n$ , let  $g_1, g_2, \dots, g_n$  be, respectively, the first, the second, ..., the  $n$ -th symbol of the shortest word representing  $g$  over the generators. We have then  $g = g_1 \dots g_n$ ,  $g_i g_{i+1} \neq 1$  for  $i = 1, \dots, n$ , and  $T_g = T_{g_1} \dots T_{g_n}$ .

Now we connect the spherical averages  $s_n$  with powers of a specially chosen Markov operator. Note that Oseledets's ergodic theorem [1] is proved in a similar way; the operator used here was used by Grigorchuk [7], Thouvenot (oral communication), and myself [8].

Let  $Y = X \times \mathcal{A}$ , and let the measure  $\eta$  on  $Y$  be the product of  $\nu$  and the uniform distribution on  $\mathcal{A}$ .

Define an operator  $P$  on  $L_1(Y, \eta)$  by the formula

$$P\varphi(x, a) = \frac{1}{2m-1} \sum_{b \in \mathcal{A}, b \neq a^{-1}} \varphi(T_a x, b) \quad (2)$$

$P$  is a measure-preserving Markov operator on  $(Y, \eta)$ .

To obtain Theorem 1, we shall use the following

**Theorem 3 (Alternierende Verfahren)** *Let  $(X, \nu)$  be a probability space and let  $P$  be a measure-preserving Markov operator on  $L_1(X, \nu)$ . Then for any  $\varphi \in L \log L(X, \nu)$  the sequence  $P^n(P^*)^n \varphi$  converges  $\nu$ -almost everywhere and in  $L_1$  as  $n \rightarrow \infty$ .*

This theorem was proved by Gian-Carlo Rota [6] for  $\varphi \in L_p$ ,  $p > 1$ , but the proof, based on convergence of conditional expectations, works without any changes for  $\varphi \in L \log L$ .

(Let us briefly recall Rota's proof [6]: set  $X^{\mathbb{Z}}$  to be the set of bi-infinite trajectories of  $P$ :  $X^{\mathbb{Z}} = \{x = \dots x_{-n} \dots x_0 \dots x_n \dots, x_n \in X\}$ , and let  $P^{\mathbb{Z}}$  be the stationary Markov measure on  $X^{\mathbb{Z}}$  corresponding to the operator  $P$  and the stationary distribution  $\nu$ . For any  $k, m \in \{-\infty\} \cup \mathbb{Z} \cup \{\infty\}$ ,  $k \leq m$ , denote by  $\mathcal{F}_k^m$  the sigma-algebra on  $X^{\mathbb{Z}}$  generated by the random variables  $x_l$ ,  $k \leq l \leq m$ . For any  $\varphi \in L_1(X, \nu)$  and any  $n > 0$  we have  $E(\varphi(x)|\mathcal{F}_n^\infty) = (P^*)^n \varphi(x_{-n})$ , and  $E(E(\varphi(x)|\mathcal{F}_n^\infty)|\mathcal{F}_0) = P^n(P^*)^n \varphi(x_0)$ . If  $\varphi \in L_1(X, \nu)$  then the sequence  $\varphi_n(x) = E(\varphi(x)|\mathcal{F}_n^\infty)$  converges almost everywhere in  $X^{\mathbb{Z}}$  as  $n \rightarrow \infty$ . If  $\varphi \in L \log L(X, \nu)$  then  $\sup_n |\varphi_n| \in L_1$ , therefore the sequence  $E(\varphi_n|\mathcal{F}_0)$

also converges almost everywhere in  $X^{\mathbb{Z}}$  as  $n \rightarrow \infty$ . Since  $E(\varphi_n(x)|\mathcal{F}_0) = P^n(P^*)^n\varphi(x_0)$ , the sequence  $P^n(P^*)^n\varphi(x_0)$  converges for  $\nu$ -almost all  $x_0$  as  $n \rightarrow \infty$ , and Theorem 3 is proved.)

Now we derive Theorems 1, 2 from Theorem 3 applied to the operator  $P$ .

The operator adjoint to  $P$  is given by

$$P^*\varphi(x, a) = \frac{1}{2m-1} \sum_{b \in \mathcal{A}, b \neq a} \varphi(T_b x, b^{-1}) \quad (3)$$

A straightforward calculation gives

$$P^n\varphi(x, a) = \frac{1}{(2m-1)^n} \sum_{g \in C(n), b \in \mathcal{A}, g_n = a, g_1 b \neq 1} \varphi(T_g x, b) \quad (4)$$

$$(P^*)^n\varphi(x, a) = \frac{1}{(2m-1)^n} \sum_{g \in C(n), b \in \mathcal{A}, g_n \neq a, g_1 b = 1} \varphi(T_g x, b) \quad (5)$$

$$(P^*)^n P^n \varphi(x, a) = \frac{1}{(2m-1)^{2n}} \sum_{g, h \in C(n), g_n = h_1^{-1}, h_n a \neq 1, g_1 b \neq 1} \varphi(T_g T_h x, b) \quad (6)$$

Set  $D(n) = \{(g, h), g, h \in C(n), g_n h_1 = 1\}$  and

$$S_{2n} = \frac{1}{2m(2m-1)^{2n-2}} \sum_{(g, h) \in D(n)} T_{gh} \quad (7)$$

**Lemma 1** *Let  $\varphi \in L_1(X, \nu)$ , set  $\Phi(x, a) = \varphi(x)$  for all  $a \in \mathcal{A}$ , and let  $\Psi = P^n(P^*)^n\Phi$ . Then*

$$\frac{1}{2m} \sum_{a \in \mathcal{A}} \Psi(x, a) = S_{2n}\varphi.$$

Indeed, (6) yields

$$\frac{1}{2m} \sum_{a \in \mathcal{A}} \Psi(x, a) = \frac{1}{2m(2m-1)^{2n-2}} \sum_{g, h \in C(n), g_n h_1 = 1} \varphi(T_g T_h x) = S_{2n}\varphi.$$

(the last equality holds because if  $(g, h) \in D(n)$ , then  $(h^{-1}, g^{-1}) \in D(n)$ .)

Theorem 3 and Lemma 1 yield

**Lemma 2** *For  $\varphi \in L \log L$  the sequence  $S_{2n}\varphi$  converges  $\nu$ -almost everywhere as  $n \rightarrow \infty$ .*

Now we connect  $S_n$  and  $s_n$ .

**Proposition 1**

$$S_{2n} = \frac{2m-2}{2m-1} s_{2n-2} + \frac{1}{2m-1} S_{2n-2}.$$

Set  $E(n) = \{g, h \in D(n) \mid g_{n-1}h_2 \neq 1\}$ ;  $F(n) = \{g, h \in D(n) \mid g_{n-1}h_2 = 1\}$ . Clearly,  $D(n) = E(n) \cup F(n)$ . The map  $(g, h) \rightarrow gh$  sends  $E(n)$  onto  $C(2n-2)$ , and every point of  $C(2n-2)$  has precisely  $2m-2$  preimages, which gives

$$\frac{1}{2m(2m-1)^{2n-2}} \sum_{(g,h) \in E(n)} T_{gh} = \frac{2m-2}{2m-1} s_{2n-2} \quad (8)$$

The map  $(g, h) \rightarrow (g_1 \dots g_{n-1}, h_2 \dots h_n)$  sends  $F(n)$  onto  $D(n-1)$  and every point of  $D(n-1)$  has exactly  $2m-1$  preimages, whence we have

$$\frac{1}{2m(2m-1)^{2n-2}} \sum_{(g,h) \in F(n)} T_{gh} = \frac{1}{2m-1} S_{2n-2} \quad (9)$$

Adding (8) and (9), we obtain the claim.

Lemma 2 and Proposition 1 imply convergence of  $s_{2n}\varphi$  for  $\varphi \in L \log L(X, \nu)$ , that is, convergence in Theorems 1, 2. The invariance of the limit under  $F_m^2$  follows from Guivarc'h's theorem [2] and the approximation of integrable functions by square-integrable ones. The proof of Theorems 1, 2 is complete.

### 3 Convergence of powers of $P$

The group  $F_m$  naturally acts on the space  $Y = X \times \mathcal{A}$  by  $T_g(x, a) = (T_g x, a)$  and on  $L_1(Y, \eta)$  by  $T_g \varphi = T_{g^{-1}} \varphi$ . This action is measure-preserving.

**Definition.** A function  $\psi \in L_1(Y, \eta)$  *does not depend on  $\mathcal{A}$*  if there exists  $\varphi \in L_1(X, \nu)$  such that  $\psi(x, a) = \varphi(x)$  for all  $a \in \mathcal{A}$ .

**Theorem 4** *For any  $\psi \in L \log L(Y, \eta)$ , the sequences  $P^{2n}\psi$ ,  $P^{2n+1}\psi$ ,  $(P^*)^{2n}\psi$ ,  $(P^*)^{2n+1}\psi$  converge  $\nu$ -almost everywhere and in  $L_1$  as  $n \rightarrow \infty$ . If  $\bar{\psi}$  is the limit function for any of these sequences, then  $\bar{\psi}$  is  $F_m^2$ -invariant and does not depend on  $\mathcal{A}$ .*

**Lemma 3** *For  $\varphi \in L \log L(X, \nu)$  the sequence  $s_{2n+1}\varphi$  converges almost everywhere and in  $L_1$  as  $n \rightarrow \infty$ .*

Indeed, the identity

$$s_{n+1} = \frac{2m}{2m-1} s_1 s_n - \frac{1}{2m-1} s_{n-1}$$

readily yields

$$s_{2n+1} = \frac{2m}{2m-1} s_1 (s_{2n} - \frac{1}{2m-1} s_{2n-2} + \frac{1}{(2m-1)^2} s_{2n-4} - \dots),$$

and convergence of  $s_{2n}\varphi$  implies convergence of  $s_{2n+1}\varphi$ .

Now for  $a \in \mathcal{A}$  denote

$$s_n^a = \frac{1}{(2m-1)^{n-1}} \sum_{g \in C(n), g_1=a} T_g$$

The obvious relation

$$s_n^a = \frac{2m}{2m-1} T_a s_{n-1} - \frac{2m}{(2m-1)^2} s_{n-2} + \frac{1}{(2m-1)^2} s_{n-2}$$

yields

$$s_n^a = T_a \left( \sum_{k=1}^{[(n-1)/2]} \frac{2m}{(2m-1)^{1+2k}} s_{n-1-2k} \right) - \sum_{k=1}^{[n/2]} \frac{1}{(2m-1)^{2k}} s_{n-2k}^a \quad (10)$$

(here  $[k]$  stands for the integer part of  $k$ ); (10) implies

**Lemma 4** *For any  $\varphi \in L \log L(X, \nu)$  and any  $a \in \mathcal{A}$ , the sequences  $s_{2n}^a \varphi$ ,  $s_{2n+1}^a \varphi$  converges both  $\nu$ -almost everywhere and in  $L_1(X, \nu)$  as  $n \rightarrow \infty$ .*

Now for  $a, b \in \mathcal{A}$  let  $C(n, a, b) = \{g \in C(n), g_1 = a, g_n = b\}$  and set

$$S_n^{a,b} = \sum_{g \in C(n, a, b)} T_g, \quad s_n^{a,b} = \frac{S_n^{a,b}}{(2m-1)^{n-1}}$$

**Lemma 5** *For any  $\varphi \in L \log L(X, \nu)$  and any  $a, b \in \mathcal{A}$ , the sequences  $s_{2n}^{a,b} \varphi$ ,  $s_{2n+1}^{a,b} \varphi$  converge  $\nu$ -almost everywhere and in  $L_1(X, \nu)$  as  $n \rightarrow \infty$ .*

The obvious relation  $S_n^{a,b} = S_{n-1}^a T_b - S_{n-2}^a + S_{n-2}^{a,b}$  yields (after normalization and iteration) the equation

$$s_n^{a,b} = \left( \sum_{k=1}^{[(n-1)/2]} \frac{1}{(2m-1)^{1+2k}} s_{n-1-2k}^a T_b - \sum_{k=1}^{[n/2]} \frac{1}{(2m-1)^{2k}} s_{n-2k}^a \right), \quad (11)$$

and Lemma 5 follows from Lemma 4.

**Lemma 6** *Fix  $r \in \mathcal{A}$ ,  $\varphi \in L \log L(X, \nu)$ , and set  $\psi(x, a) = 0$  if  $a \neq r$  and  $\psi(x, a) = \varphi(x)$  for  $a = r$ . Then*

$$P^n \psi(x, a) = \frac{1}{(2m-1)^{n-1}} \sum_{l \in \mathcal{A}, l \neq r} S_n^{a^{-1}, l},$$

$$(P^*)^n \psi(x, a) = \frac{1}{(2m-1)^{n-1}} \sum_{b \neq a^{-1}} S_n^{r, b}.$$

Proof: By (4) and the definition of  $\psi$ , we have

$$P^n \psi(x, a) = \frac{1}{(2m-1)^{n-1}} \sum_{g \in C(n), g_n = a, g_1 \neq 1} \psi(T_g x, b) = \frac{1}{(2m-1)^{n-1}} \sum_{g \in C(n), g_n = a, g_1 \neq 1} \varphi(T_g x) =$$

$$= \frac{1}{(2m-1)^{n-1}} \sum_{g_1=a^{-1}, g_n \neq r} T_g \varphi(x) = \frac{1}{(2m-1)^{n-1}} \sum_{l \in \mathcal{A}, l \neq r} S_n^{a^{-1}, l}.$$

The case of  $(P^*)^n$  is treated similarly.

Lemma 6 and Lemma 5 imply convergence of  $P^{2n}\psi$ ,  $(P^*)^{2n}\psi$ ,  $P^{2n+1}\psi$ ,  $(P^*)^{2n+1}\psi$  for  $\psi$  such as in the formulation of Lemma 6 and, consequently, for all  $\psi \in L \log L(Y, \eta)$ .

Convergence in Theorem 4 is proved; it remains to identify the limit.

Let  $\bar{\psi} = \lim_{n \rightarrow \infty} P^{2n}\psi$ . Then, clearly,  $P^2\bar{\psi} = \bar{\psi}$ . We have the following easy

**Proposition 2** *Suppose  $\bar{\psi} \in L_1(Y, \eta)$  satisfies  $P\bar{\psi} = \bar{\psi}$ . Then  $\bar{\psi}$  does not depend on  $\mathcal{A}$  and  $T_g\bar{\psi} = \bar{\psi}$  for all  $g \in F_m$ .*

*Suppose  $\bar{\psi} \in L_1(Y, \eta)$  satisfies  $P^2\bar{\psi} = \bar{\psi}$ . Then  $\bar{\psi}$  does not depend on  $\mathcal{A}$  and  $T_g\bar{\psi} = \bar{\psi}$  for all  $g \in F_m^2$ .*

For proof, see Lemma 10 in [8] or [2].

This completes the proof of Theorem 4.

**Corollary 1** *If  $F_m^2$  acts ergodically on  $(X, \nu)$ , then for any  $\psi \in L \log L(Y, \eta)$  we have  $P^n\psi \rightarrow \int_Y \psi d\eta$ ,  $(P^*)^n\psi \rightarrow \int_Y \psi d\eta$  almost everywhere and in  $L_1(Y, \eta)$  as  $n \rightarrow \infty$ .*

Since  $L \log L$  is dense in  $L_1$ , we have  $P^n\psi \rightarrow \int_Y \psi d\eta$ ,  $(P^*)^n\psi \rightarrow \int_Y \psi d\eta$  in  $L_1$  for any  $\psi \in L_1$ . By the 0 – 2 laws for Markov operators (see [9], [10]), Corollary 1 implies

**Proposition 3** *Suppose  $F_m^2$  acts ergodically on  $(X, \nu)$ . Then the tail sigma-algebra of  $P$  is trivial and the tail sigma-algebra of  $P^*$  is trivial.*

Say that the action of  $F_m$  on  $(X, \nu)$  has eigenvalue  $-1$  if there exists  $\varphi \in L_2(X, \nu)$  such that  $T_a\varphi = -\varphi$  for all  $a \in \mathcal{A}$ . Clearly,  $F_m^2$  acts ergodically iff  $F_m$  acts ergodically and does not have eigenvalue  $-1$ ; then, the action of  $F_m$  has eigenvalue  $-1$  iff the operator  $P$  (considered as an operator in  $L_2$ ) has eigenvalue  $-1$ ; so, the tail sigma-algebra of  $P$  is trivial iff  $P$  is ergodic and does not have eigenvalue  $-1$ .

Finally, if the action of  $F_m^2$  on  $(X, \nu)$  is not ergodic then nonconstant  $F_m^2$ -invariant functions in  $L_1(Y, \eta)$  that do not depend on  $\mathcal{A}$  are nontrivial tail functions; Proposition 3 implies that these functions span the tail sigma-algebra of  $P$ .

## 4 Generalizations

The method of Markov operators allows to generalize the Nevo–Stein theorem in the following way.

Let, again,  $(X, \nu)$  be a probability space and suppose a free group  $F_m$  with  $m$  generators acts on  $(X, \nu)$  by measure-preserving transformations; as above,

$\{a_1, \dots, a_m\}$  is a fixed set of free generators for  $F_m$  and  $T_1, \dots, T_m : X \rightarrow X$  are transformations corresponding to the generators, and  $T_{-i} = T_i^{-1}$ .

Consider the set  $W_{\mathcal{A}}$  of all finite words over the alphabet  $\mathcal{A} = \{-m, \dots$ :

$$W_m = \{w = w_1 w_2 \dots w_n \mid w_i \in \mathcal{A}\}$$

Denote by  $|w|$  the length of the word  $w$  and for any positive integer  $n$ , let  $W_{\mathcal{A}}(n) = \{w \in W_{\mathcal{A}}, |w| = n\}$ .

For each  $w \in W_{\mathcal{A}}$ ,  $w = w_1 \dots w_n$ , define a transformation

$$T_w = T_{w_n} T_{w_{n-1}} \dots T_{w_1}. \quad (12)$$

Let  $\Pi$  be a stochastic  $2m \times 2m$  matrix, whose rows and columns are indexed by elements of  $\mathcal{A}$ , that is,  $\Pi = (p_{ij}), i, j \in \mathcal{A}$ . Assume that  $\Pi$  has a stationary distribution  $(p_{-m}, \dots, p_{-1}, p_1, \dots, p_m)$  such that all  $p_i > 0$ .

For  $w \in W_{\mathcal{A}}$ ,  $w = w_1 \dots w_n$ , denote

$$p(w) = p_{w_1 w_2} p_{w_2 w_3} \dots p_{w_{n-1} w_n}, \quad \pi(w) = p_{w_1} p(w).$$

Consider the operators

$$s_n^\Pi = \sum_{|w|=n} \pi(w) T_w \quad (13)$$

The matrix  $\Pi$  is said to **generate the free group** if  $p_{ij} = 0$  is equivalent to  $i + j = 0$ .

Assume the self-adjointness condition

$$p_i = p_{-i}, \quad p_{-i, -j} = \frac{p_j p_{ji}}{p_i} \quad (14)$$

Relation (14) is equivalent to saying that all operators  $\bar{s}_{2n}$  are self-adjoint.

Let  $F_m^2$  be the subgroup of words of even length in  $F_m$ , that is, the subgroup generated by  $a_i a_j$ ,  $i, j \in \{1, \dots, m\}$ .

**Theorem 5** *Assume the matrix  $\Pi$  generates the free group and satisfies (14). Then for any  $\varphi \in L \log L(X, \nu)$ , the sequence  $s_{2n}^\Pi \varphi$  converges as  $n \rightarrow \infty$  both  $\nu$ -almost everywhere and in  $L_1(X, \nu)$  to an  $F_m^2$ -invariant function.*

The condition that  $\Pi$  generates the free group in Theorem 5 can be weakened in the following way.

**Definition 1** *A matrix  $\Pi$  with nonnegative entries will be called irreducible if for some  $n > 0$  all entries of the matrix  $\Pi + \Pi^2 + \dots + \Pi^n$  are positive (if  $\Pi$  is stochastic then this is equivalent to saying that in the corresponding Markov chain any state is attainable from any other state).*

**Definition 2** *A matrix  $\Pi$  with nonnegative entries will be called strictly irreducible if  $\Pi$  is irreducible and  $\Pi \Pi^T$  is irreducible (here  $\Pi^T$  stands for the transpose of  $\Pi$ ).*



**Definition 3** A Markov chain will be called strictly irreducible if its matrix of transition probabilities is strictly irreducible. A Markov measure will be called strictly irreducible if the corresponding chain is strictly irreducible.

It is clear that a matrix generating the free group is strictly irreducible. Theorem 5 is a particular case of

**Theorem 6** Assume the matrix  $\Pi$  is strictly irreducible and satisfies (14). Then for any  $\varphi \in L \log L(X, \nu)$ , the sequence  $s_{2n}^\Pi \varphi$  converges as  $n \rightarrow \infty$  both  $\nu$ -almost everywhere and in  $L_1(X, \nu)$  to an  $F_m^2$ -invariant function.

The proofs of Theorems 5, 6 are similar to that of Theorem 2; they will appear in a forthcoming paper.

*Acknowledgement.* I am grateful to Amos Nevo for the suggestion to use Rota's theorem. I am grateful to Victor A. Kleptsyn, Valeriy I. Oseledets, Elias M. Stein, and Yakov G. Sinai for useful discussions. I am grateful to Vadim A. Kaimanovich for his careful reading of the manuscript and very helpful comments.

## References

- [1] V.I. Oseledets, Markov chains, skew-products, and ergodic theorems for “general” dynamical systems, *Theory of Probability and its Applications*, 1965, 10, 3, 551–557.
- [2] Y. Guivarc’h, Généralisation d’un théorème de von Neumann, *C.R.A.S. Paris* 268 (1969), 1020–1023.
- [3] V.I. Arnold, A.L. Krylov, Equidistribution of points on a sphere and ergodic properties of solutions of ordinary differential equations in the complex plane. *Doklady AN SSSR* 1963, 148(1), 9–12.
- [4] R.I. Grigorchuk, Pointwise ergodic theorems for actions of free groups, *Proceedings of the Tambov Workshop in the Theory of Functions*, 1986.
- [5] A. Nevo, E.M. Stein, A generalization of Birkhoff’s pointwise ergodic theorem, *Acta mathematica*, 173(1994), no.1, 135 – 154.
- [6] G.-C. Rota, An “Alternierende Verfahren” for general positive operators, *Bulletin AMS*, 68, 1962, 95 – 102.
- [7] R.I. Grigorchuk, Ergodic theorems for actions of free semigroups and groups, *Mathematical Notes*, 1999, 65, no.5, 779–783.
- [8] A. I. Bufetov, Operator ergodic theorems for actions of free semigroups and groups, *Functional Analysis and its Applications*, 2000, vol. 34, no.4, pp. 239–251.

- [9] V.A. Kaimanovich, Measure-theoretic boundaries of Markov chains, 0 — —2 laws and entropy, in *Harmonic Analysis and Discrete Potential Theory*, ed. M.A. Picardello, Plenum, 1992, pp. 145-181.
- [10] Y. Deriennic, Lois “zéro-deux” pour les processus de Markov, applications aux marches aléatoires, *Annales Inst. H. Poincaré, Sect. B*, 12(1976), 111-129.
- [11] S. Kakutani, Random ergodic theorems and Markoff processes with a stable distribution, *Proc. 2nd Berkeley Symp. Math.Stat. and Prob.*, 1951, 247 - 261.