ANNALES
POLONICI MATHEMATICI
XLVIII (1988)

A simple proof of a theorem on Markov operators

by Antoni Leon Dawidowicz (Kraków)
and Andrzej Turski (Katowice)

Abstract. We present a simple proof of one theorem of Koopman–von Neumann–Halmos characterizing weakly mixing Markov operators.

Introduction. In the ergodic theory we often encounter theorems which exhibit the interplay between properties of certain operators and properties of their Cartesian products. In the case of a deterministic operator the mixing property is equivalent to the ergodicity of its Cartesian square. This property does not extend to Markov operators. The corresponding results for the case of a Markov operator is the Koopman–von Neumann–Halmos theorem. A proof has been supplied by Aaronson, Lin and Weiss [1], using the idea of a Markov shift the theory of Hilbert spaces.

In this paper we present another proof of this theorem. This proof resorbs only to the properties of Markov operators in L^1.

1. Definitions, notation and basic theorem. Let (X, Σ, m) be a measure space with m being σ-finite.

Definition 1. A function P: $X \times \Sigma \to [0, 1]$ is called a *transition probability* if it satisfies the following conditions:

1. For every $x \in X$, $P(x, \cdot)$ is a measure on (X, Σ);
2. For every $A \in \Sigma$, $P(\cdot, A)$ is Σ-measurable;
3. For every $x \in X$, $P(x, \cdot)$ is absolutely continuous with respect to m.

Using this idea we can define two operators dual to each other called the *Markov operators* on X acting on $L^\infty(X)$ and $L^1(X)$ respectively. We shall denote these operators by the same character as the transition probability; however, in the case of an operator on $L^\infty(X)$ we shall write this symbol left to the argument and in the case of an operator on $L^1(X)$ we shall write it on the right.

Definition 2. $Pf(x) = \int_X f(y)P(x, dy)$ for $f \in L^\infty(X)$.

Definition 3. $\int_A uP(x)m(dx) = \int_X P(x, A)u(x)m(dx)$ for $u \in L^1(X)$, $A \in \Sigma$ (i.e., $uP = d\mu/dm$, where $\mu(A) = \int_X P(x, A)u(x)m(dx)$).

The correctness of these definitions (i.e., the independence on the choice of the representative function and the absolute continuity of μ) follows from property 3 of the definition of a transition probability.

The operator P acting on $L^1(X)$ will be also considered as an operator acting on absolute continuous measures,

$$\mu P(A) = \int_X P(x, A)\mu(dx).$$

We shall also write

$$\langle u, f\rangle = \int_X u(x) f(x) m(dx).$$

DEFINITION 4. A measure μ absolutely continuous with respect to m is called *invariant* if $\mu P = \mu$.

In subsequent considerations we shall need a characterization of the iterates of P. We define the operator of transition probability in n steps by induction.

DEFINITION 5. $P_1(x, A) = P(x, A)$, $P_{n+1}(x, A) = \int_X P(y, A) P_n(x, dy)$.

THEOREM 1. *P_n is a transition probability. Moreover, $P_n f = P^n f$ for every $f \in L^\infty(X)$ and $uP_n = uP^n$ for every $u \in L^1(X)$, where P^n denotes the n-fold iterate of P.*

This theorem shows that it is legitimate to denote by P^n also the transition probability in n steps.

DEFINITION 6. A Markov operator P is called *ergodic* if for every $u \in L^1(X)$ satisfying the equality

$$\int_X u(x) m(dx) = 0 \tag{1}$$

and for every $f \in L^\infty(X)$, the condition

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} \langle uP^k, f\rangle = 0$$

is satisfied.

DEFINITION 7. An operator P is called *weakly mixing* if for every $u \in L^1(X)$ satisfying equality (1) and for every $f \in L^\infty(X)$ the condition

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} |\langle uP^k, f\rangle| = 0$$

is satisfied.

We shall also use the following characterization of ergodicity.

THEOREM 2. *An operator P is ergodic if and only if every function $g \in L^\infty(X)$ satisfying the condition $Pg = g$ is constant almost everywhere.*

The proof of this theorem is generally known and presented for example in [2]. An analogous theorem characterizes weakly mixing operators.

THEOREM 3. *An operator P is weakly mixing if and only if for every function $g \in L^\infty(X)$ satisfying the condition $Pg = \lambda g$ for some complex λ with $|\lambda| = 1$ is constant a.e.*

The proof of this theorem is given in [3].

2. Formulation and proof of the Koopman–von Neumann–Halmos theorem. To formulate the theorem, we need the definition of the Cartesian product of Markov operators.

DEFINITION 8. Let P and Q be transition probabilities on (X, Σ, m) and $(Y, \mathfrak{M}, n)$, respectively. We denote by $P \times Q$ the transition probability on $(X \times Y, \Sigma \times \mathfrak{M}, m \times n)$ given by the formula

$$(P \times Q)\big((x, y), \cdot\big) = P(x, \cdot) \times Q(y, \cdot).$$

This transition probability induces the Markov operators $P \times Q$ on $L^\infty(X \times Y)$ and $L^1(X \times Y)$.

THEOREM 4. *An operator P on X is weakly mixing if and only if for every measure space $(Y, \mathfrak{M}, n)$ and for every ergodic Markov operator Q on Y admitting a finite invariant measure equivalent to n the operator $P \times Q$ is ergodic.*

Proof. *Necessity.* Let P be a weakly mixing operator on $L^\infty(X, \Sigma, m)$ and let Q be an ergodic operator on $L^\infty(Y, \mathfrak{M}, n)$. Let $\mu \sim n$ be a measure invariant with respect to Q, as in the condition of the theorem.

To prove the necessity of the condition it is enough to prove the ergodicity of $P \times Q$. We appeal to Theorem 2. Let a function $f \in L^\infty(X \times Y, \Sigma \times \mathfrak{M}, m \times n)$ satisfy the condition $(P \times Q) f = f$. We shall show that this function is constant almost everywhere. Let us consider an arbitrary function $u \in L^1(X)$ satisfying the condition

$$\int_X u(x)\, m(dx) = 0.$$

We have to show that

$$I_1 = \int_Y \Big| \int_X u(x) f(x, y)\, m(dx)\, \mu(dy) \Big| = 0.$$

Obviously,

$$\Big| \int_X u(x) f(x, y)\, m(dx) \Big| = \lambda(y) \int_X u(x) f(x, y)\, m(dx),$$

where $|\lambda(y)| = 1$ and $\lambda \in L^\infty(Y)$. Hence

$$I_1 = \int_Y \lambda(y) \int_X u(x) f(x, y)\, n(dx)\, \mu(dy)$$

$$= \int_Y \lambda(y) \int_X u(x) \frac{1}{N} \sum_{k=0}^{N-1} (P \times Q)^k f(x, y)\, m(dx)\, \mu(dy)$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} \int_X \lambda(y) \int_X u(x) \int_{X\times Y} f(s,t)(P\times Q)^k((x,y), d(s,t))\, m(dx)\,\mu(dy)$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} \int_Y \lambda(y) \int_X u(x) \int_X \left[\int_Y f(s,t)\, Q^k(y,dt) \right] P^k(x,ds)\, m(dx)\,\mu(dy)$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} \int_Y \lambda(y) \int_X uP^k(x) \left[\int_Y f(x,t)\, Q^k(y,dt) \right] m(dx)\,\mu(dy)$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} \int_X uP^k(x) \left[\int_Y \lambda(y) \int_Y f(x,t)\, Q^k(y,dt) \right] \mu(dy)\, m(dx).$$

We define a function $v \in L^1(Y)$ by the formula

$$v = d\mu/dn.$$

Since $\mu \sim n$, $v > 0$ almost everywhere. Moreover, $\lambda v \in L^1(Y)$. Hence

$$I_1 = \frac{1}{N}\sum_{k=0}^{N-1} \int_X uP^k(x) \left[\int_Y \lambda(y)\, v(y) \left(\int_Y f(x,t)\, Q^k(y,dt) \right) n(dy) \right] m(dx)$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} \int_X uP^k(x)\, \langle \lambda v, Q^k f(x,\cdot)\rangle\, m(dx)$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} \int_X uP^k(x)\, \langle (\lambda v) Q^k, f(x,\cdot)\rangle\, m(dx)$$

$$= \frac{1}{N}\sum_{k=0}^{N-1} \int_X uP^k(x) \int_Y f(x,y)(\lambda v) Q^k(y)\, n(dy)\, m(dx)$$

$$= \int_Y \frac{1}{N}\sum_{k=0}^{N-1} (\lambda v) Q^k(y) \int_X uP^k(x) f(x,y)\, m(dx)\, n(dy)$$

$$= \int_Y \frac{1}{N}\sum_{k=0}^{N-1} \frac{(\lambda v) Q^k(y)}{v(y)} \langle uP^k, f(\cdot, y)\rangle\, \mu(dy).$$

For every $A \in \mathfrak{M}$ we have

$$\left|\int_A (\lambda v) Q^k(y)\, n(dy)\right| = \left|\int_Y (\lambda v)(y)\, Q^k(y,A)\, n(dy)\right| \leqslant \int_Y |\lambda(y)\, v(y)\, Q^k(y,A)|\, n(dy)$$

$$= \int_Y v(y)\, Q^k(y,A)\, n(dy) = \mu Q^k(A) = \mu(A) = \int_A v(y)\, n(dy).$$

Hence $|(\lambda v)Q^k(y)| \leqslant v(y)$ for almost every $y \in Y$ and in consequence

$$|I_1| \leqslant \int_Y \frac{1}{N} \sum_{k=0}^{N-1} |\langle uP^k, f(\cdot, y)\rangle| \mu(dy)$$

for every positive integer N. Since

$$|\langle uP^k, f(\cdot, y)\rangle| \leqslant \|u\|_{L^1} \|f(\cdot, y)\|_{L^\infty} \leqslant \|u\|_{L^1} \cdot \|f\|_{L^\infty(X \times Y)},$$

μ is finite and

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} |\langle uP^k, f(\cdot, y)\rangle| = 0$$

almost everywhere it follows from the Lebesgue Dominated Convergence Theorem that $I_1 = 0$. Hence also $\int_X u(x) f(x, y) m(dx) = 0$ μ-a.e. Thus, for n-almost every y the function $f(\cdot, y)$ is constant m-almost everywhere. Write $g(y) = f(\cdot, y)$ m-a.e. Then $Qg(y) = (P \times Q) f(x, y) = f(x, y) = g(y)$ n-a.e. and it follows from the ergodicity of Q that the function g is constant almost everywhere. Consequently the operator $P \times Q$ is ergodic.

Sufficiency. To prove the sufficiency let us consider an operator P such that for every measure space $(Y, \mathfrak{M}, n)$ and for every ergodic Markov operator Q on Y admitting a finite invariant measure equivalent to n, the operator $P \times Q$ is ergodic. Let us consider a function $g \in L^\infty(X)$ such that $Pg = \lambda g$ for a complex number λ of modulus 1.

Case I. There exists $p \in N$ such that $\lambda^p = 1$.

Let us define

$$Y = \{0, 1, \ldots, p-1\}, \qquad \mathfrak{M} = \mathscr{P}(Y), \qquad n(A) = \operatorname{card} A,$$
$$\tau(y) = y - 1 \ (\operatorname{mod} p), \qquad Qf = f \circ \tau.$$

The Markov operator Q is ergodic and n is a finite measure invariant with respect to Q. Let $f(x, y) = \lambda^y g(x)$. Clearly, $f \in L^\infty(X \times Y)$ and $(P \times Q) f(x, y) = Pf(x, \tau(y)) = \lambda^{\tau(y)} \lambda g(x) = \lambda^y g(x) = f(x, y)$.

Since $P \times Q$ is ergodic, f is constant a.e.; consequently g is constant a.e. and P is weakly mixing.

Case II. $\lambda^p \neq 1$ for all $p \in N$.

Let $Y = \{y \in C\colon |y| = 1\}$. We denote by $\mathfrak{M}$ the σ-algebra of all Lebesgue measurable sets in Y and by n the one-dimensional Lebesgue measure on the unit circle. Let $\tau(y) = \lambda^{-1} y$, $Qf = f \circ \tau$. The operator Q is ergodic [4] and n is a finite invariant measure for Q. Let $f(x, y) = yg(x)$. Clearly, $f \in L^\infty(X \times Y)$ and $(P \times Q) f(x, y) = Pf(x, \lambda^{-1} y) = \lambda^{-1} y Pg(y) = \lambda^{-1} y \lambda g(x) = f(x; y)$ a.e. Since $P \times Q$ is ergodic, f is constant and so P is weakly mixing; this completes the proof.

References

[1] J. Aaronson, M. Lin, B. Weiss, *Mixing properties of Markov Operators and Ergodic Transformations and Ergodicity of Cartesian Product*, Israel J. Math. 33 (1979), 198–244.

[2] S. R. Foguel, *The Ergodic Theory of Markov Processes*, van Nostrand Reinhold, New York 1979.

[3] M. Lin, *On Weakly Mixing Markov Operators and Non-Singular Transformations*, Z. Warsch. und verw. Gebiete 55 (1981) 231–236.

[4] Ya. Sinai, *Introduction to Ergodic Theory*, University Press, Princeton 1976.

INSTITUTE OF MATHEMATICS OF JAGIELLONIAN UNIVERSITY, KRAKÓW
and
INSTITUTE OF MATHEMATICS OF SILESIAN UNIVERSITY, KATOWICE

Reçu par la Rédaction le 1985.10.15