COLLOQUIUM MATHEMATICUM

VOL. 91 2002 NO. 2

A NOTE ON MARKOV OPERATORS AND TRANSITION SYSTEMS

BY

BARTOSZ FREJ (Wrocław)

Abstract. On a compact metric space X one defines a transition system to be a lower semicontinuous map $X \to 2^X$. It is known that every Markov operator on $C(X)$ induces a transition system on X and that commuting of Markov operators implies commuting of the induced transition systems. We show that even in finite spaces a pair of commuting transition systems may not be induced by commuting Markov operators. The existence of trajectories for a pair of transition systems or Markov operators is also investigated.

0. Preliminaries. Let X be a compact metrizable space with the Borel σ-algebra $\mathcal{B}$. Denote by $C(X)$ the space of all real-valued continuous functions on X, and by 2^X the collection of all nonempty closed subsets of X. Let $\mathbb{N}$ be the set of all natural numbers, and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. A *transition probability* on X is a function $P : X \times \mathcal{B} \to [0,1]$ such that $P(x,\cdot)$ is a probability measure for every $x \in X$, and $P(\cdot, A)$ is a Borel measurable function for every $A \in \mathcal{B}$. By the Ionescu-Tulcea theorem, every transition probability induces a random process on X in such a way that for every $x_0 \in X$ there exists a unique probability measure μ_{x_0} on the product space $X^{\mathbb{N}_0}$ satisfying for measurable rectangles $F = F_0 \times \ldots \times F_n \times X^{\mathbb{N}}$ the following condition:

$$\mu_{x_0}(F) = \chi_{F_0}(x_0) \int\limits_{F_1} P(x_0, dx_1) \int\limits_{F_2} P(x_1, dx_2) \ldots \int\limits_{F_n} P(x_{n-1}, dx_n).$$

Given a probability measure P_0 on X, we obtain a measure μ on $X^{\mathbb{N}_0}$, independent of the starting point of the process, by putting

$$\mu(E) = \int \mu_x(E)\, P_0(dx)$$

for all measurable sets $E \subset X^{\mathbb{N}_0}$ (see e.g. [N]). Then P_0 is a stationary measure for the process if and only if the resulting measure μ is invariant under the shift transformation on $X^{\mathbb{N}_0}$.

With a transition probability P we associate an operator on bounded Borel functions defined by $T_P f(x) = \int f(y)\, P(x, dy)$. We say that P is a *Feller transition probability* if the map sending a point $x \in X$ to the measure

2000 *Mathematics Subject Classification*: 37A30, 37B50.

Key words and phrases: Markov operator, transition probability, transition system.

$P(x, \cdot)$ is a continuous transformation from X into the set $\mathcal{M}(X)$ of all probability measures on X endowed with the weak* topology. Equivalently, P is Feller if and only if T_P sends $C(X)$ into $C(X)$.

From now on, by a transition probability we mean a Feller transition probability. It is known that T_P is a *Markov operator* on X, i.e. a bounded linear operator such that $T_P f$ is nonnegative for every nonnegative function f, and that $T_P f \equiv 1$ if $f \equiv 1$. Conversely, every Markov operator on $C(X)$ is of this form for some transition probability P. For simplicity, we use the same letter (usually T) to denote both a transition probability and the corresponding Markov operator. Let T_1, T_2 be transition probabilities on a space X. We define $T_1 T_2(x, A) = \int T_2(y, A)\, T_1(x, dy)$. It turns out that $T_1 T_2$ is a transition probability which corresponds to the composition of the Markov operators T_1 and T_2 on $C(X)$. We say that T_1 and T_2 *commute* if $T_1 T_2 = T_2 T_1$.

For a transition probability T the measure $T(x, \cdot)$ expresses our expectations concerning the next state of the random process associated with T with initial point x. In the topological setup we may be interested only in locating the smallest closed subset of the phase space X containing possible future states. In this way, one is led to investigate the map assigning to any $x \in X$ the topological support of $T(x, \cdot)$, namely

$$\phi_T : X \to 2^X, \qquad \phi_T(x) = \operatorname{supp} T(x, \cdot)$$

(see [I] and [D]). This map turns out to be lower semicontinuous, i.e. for every open set U the set $\{x : \phi_T(x) \cap U \neq \emptyset\}$ is open. In general, we define a *transition system* on X to be any lower semicontinuous map $\phi : X \to 2^X$. For instance, if $f : X \to X$ is a continuous map then $\phi(x) = \{f(x)\}$ defines a transition system. We will say that a transition system ϕ is *inscribed* in another transition system ψ (notation $\phi \prec \psi$) if $\phi(x) \subset \psi(x)$ for every $x \in X$. For a given transition system ϕ, the existence of a Markov operator whose transition system ϕ_T is inscribed in ϕ follows immediately from the classical Michael selection theorem. Moreover, from a more subtle Michael selection theorem (Th. 3.1‴ in [M]) we see that there exists a transition probability T for which $\phi_T = \phi$.

Notice that it is possible to extend a system $\phi : X \to 2^X$ to a map from 2^X into 2^X. Simply, put

$$\phi(F) = \overline{\bigcup_{x \in F} \phi(x)} \tag{1}$$

for $F \in 2^X$. In this manner, composition (and commuting) of transition systems is well defined; $(\phi \circ \psi)\{x\}$ is denoted by $\phi \circ \psi(x)$. It is not hard to verify that if S, T are Markov operators, then $\phi_{ST} = \phi_S \circ \phi_T$.

Consider a family $T_1, \ldots, T_m$ of commuting Markov operators. Then, by the last statement, the corresponding transition systems also commute. The converse is not true in general; commuting transition systems need not be associated with commuting Markov operators. We will provide an appropriate example in Section 2. Consequently, theorems stated for commuting transition systems are more general than those formulated for Markov operators. Conversely, counterexamples constructed in terms of Markov operators are automatically valid for transition systems.

An interesting class of transition systems is obtained by taking continuous maps $\phi : X \to 2^X$. Then the extended map $\phi : 2^X \to 2^X$ is continuous and $(2^X, \phi)$ can be regarded as a dynamical system. Such an approach was exploited in [D]. Continuity also implies that the closure in formula (1) is superfluous, since the union $\bigcup_{x \in F} \phi(x)$ is already closed. As we shall see, this makes the notion of commuting more meaningful. Obviously, transition systems defined on finite (discrete) spaces are continuous.

For any transition systems $\phi, \psi : X \to 2^X$, we will denote by $\phi \cup \psi$ the transition system on X given by the formula

$$(\phi \cup \psi)(x) = \phi(x) \cup \psi(x).$$

Clearly, $\phi \prec \phi \cup \psi$ and $\psi \prec \phi \cup \psi$. For instance, for convex combinations of transition probabilities S and T we have

$$\phi_{\alpha S + (1-\alpha) T} = \phi_S \cup \phi_T$$

if $0 < \alpha < 1$. If ϕ, ψ are induced by functions $f, g : X \to X$, respectively, then we simply write $f \cup g$.

In the case of finite spaces the investigation of transition systems and probabilities reduces to studying matrices. Namely, with every transition system ϕ on a finite k-element space we associate a 0-1 square matrix $[\phi_{ij}]_{k \times k}$ such that $\phi_{ij} = 1$ if and only if the jth point belongs to the image of the ith point. Composition of transition systems corresponds to the following rule of “multiplying” 0-1 matrices:

$$(A \odot B)(i,j) = \operatorname{sgn}(AB(i,j)).$$

One can also write $(A \cup B)(i,j) = A(i,j) \vee B(i,j)$ for the matrix representing the sum of transition systems. For example, let C_n be a cyclic group of integers modulo n. Denote by $I : C_n \to C_n$ the identity map, and by $R : C_n \to C_n$ the addition of the unit, i.e. $R(p) = p + 1 \pmod n$. If $n = 3$ then $I \cup R$, $I \cup R^2$, $R \cup R^2$ on C_3 correspond to the following 0-1 matrices:

$$I \cup R \sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}, \quad I \cup R^2 \sim \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad R \cup R^2 \sim \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix},$$

and the $\odot$-product of any two of them equals

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Likewise, transition probabilities on finite spaces are usually written in the form of stochastic matrices, and every such matrix defines a Feller transition probability. Here, composition of probabilities coincides with the usual product of matrices. For instance, if $\alpha, \beta, \gamma \in [0,1]$, $\alpha+\beta+\gamma = 1$, then $\alpha I + \beta R + \gamma R^2$ stands for the transition probability on C_3 defined by the stochastic matrix

$$\begin{bmatrix} \alpha & \beta & \gamma \\ \gamma & \alpha & \beta \\ \beta & \gamma & \alpha \end{bmatrix}.$$

If T is a stochastic matrix, then the 0-1 matrix of ϕ_T is obtained from T by replacing all nonzero entries by 1.

1. Trajectories of transition systems. For a transition system ϕ on a compact space X we define a *trajectory* as a sequence $(x^n)_{n\in\mathbb{N}_0} \subset X$ such that $x^{n+1} \in \phi(x^n)$ for every $n \in \mathbb{N}_0$. For a finite collection $\phi_1, \ldots, \phi_m$ of transition systems a *trajectory* is an m-dimensional array $(x^{n_1,\ldots,n_m}) \in X^{\mathbb{N}_0\times\ldots\times\mathbb{N}_0}$ such that for every $n_1, \ldots, n_m$ the following transition law holds:

$$\begin{aligned} x^{n_1+1,n_2,\ldots,n_m} &\in \phi_1(x^{n_1,\ldots,n_m}), \\ x^{n_1,n_2+1,\ldots,n_m} &\in \phi_2(x^{n_1,\ldots,n_m}), \\ &\vdots \\ x^{n_1,n_2,\ldots,n_m+1} &\in \phi_m(x^{n_1,\ldots,n_m}). \end{aligned}$$

A construction of trajectories for an arbitrary pair of commuting continuous transition systems can be found in [I] (the set Ω_∞ in the proof of Theorem 3).

For a finite family of Markov operators, by a trajectory we will mean a trajectory of the induced family of transition systems. It is shown in [DI] that if T and S are commuting Markov operators, then for every probability measure P_0 on X there exists a probability measure μ on $X^{N_0\times N_0}$ compatible with these operators, which is stationary if P_0 is invariant. Using this fact one can show the following theorem.

Theorem 1. *The measure μ of the set of all trajectories of the pair S, T is equal to* 1.

Proof. For fixed $k_0, l_0 \in \mathbb{N}_0$ consider the set $C_h(k_0,l_0)$ of all $(x^{k,l}) \in X^{\mathbb{N}_0\times\mathbb{N}_0}$ satisfying the condition $x^{k_0+1,l_0} \in \phi_T(x^{k_0,l_0})$. By Theorem 1 of [DI] the measure of this set is equal to

$$\int_X \int_X \ldots \int_X \int_{\phi_T(x^{k_0,l_0})} T(x^{k_0,l_0}, dx^{k_0+1,l_0}) \ldots T(x^{0,l_0}, dx^{1,l_0}) \\ \times S(x^{0,l_0-1}, dx^{0,l_0}) \ldots S(x^{0,0}, dx^{0,1}) P_0(dx^{0,0}) = 1,$$

since $\phi_T(x^{k_0,l_0})$ is the support of the measure $T(x^{k_0,l_0},\cdot)$. Similarly, the measure of the set $C_v(k_0,l_0) = \{(x^{k,l}) : x^{k_0,l_0+1} \in \phi_S(x^{k_0,l_0})\}$ is equal to 1. Clearly, the set of all trajectories of the pair S,T coincides with the intersection

$$\bigcap_{k\in\mathbb{N}_0}\bigcap_{l\in\mathbb{N}_0}[C_h(k,l)\cap C_v(k,l)],$$

hence has measure 1. ∎

As the following example shows, it is not true that an arbitrary pair of commuting transition systems has a nonempty set of trajectories.

Example 1. Consider the space

$$Y = \{p\}\cup\{a,a_1,a_2,\ldots\}\cup[0,\pi],$$

where $p,a,a_1,a_2,\ldots$ are different elements not belonging to $[0,\pi]$, p is an isolated point, and the sequence $a_1,a_2,\ldots$ converges to a. The space Y is compact. Define transition systems ϕ and ψ on Y in the following way:

$$\phi(y) = \begin{cases} \{a,a_1,a_2,\ldots\} & \text{if } y=p,\\ \{k/n : k\in\mathbb{N}\}\cap[0,\pi] & \text{if } y=a_n,\\ \{0\} & \text{if } y=a,\\ \{p\} & \text{if } y\in[0,\pi],\end{cases}$$

$$\psi(y) = \begin{cases} \{a,a_1,a_2,\ldots\} & \text{if } y=p,\\ \{\pi-k/n : k\in\mathbb{N}\}\cap[0,\pi] & \text{if } y=a_n,\\ \{\pi\} & \text{if } y=a,\\ \{p\} & \text{if } y\in[0,\pi].\end{cases}$$

Both ϕ and ψ are lower semicontinuous (the only discontinuity point is a). Moreover, ϕ and ψ commute, because

$$\phi\circ\psi(y) = \psi\circ\phi(y) = \begin{cases} [0,\pi] & \text{if } y=p,\\ \{p\} & \text{if } y\in\{a,a_1,a_2,\ldots\},\\ \{a,a_1,a_2,\ldots\} & \text{if } y\in[0,\pi].\end{cases}$$

We will show that the pair ψ,ϕ has no trajectories. Since for any $y\in Y$ one of the images $\phi(y),\phi^2(y),\phi^3(y)$ must be equal to $\{p\}$, it suffices to prove that there are no trajectories starting from p. Notice that constructing a trajectory with initial point $x^{0,0}=p$ we obtain $x^{1,0}\in\phi(x^{0,0})=\{a,a_1,a_2,\ldots\}$ and $x^{0,1}\in\psi(x^{0,0})=\{a,a_1,a_2,\ldots\}$. Then we have simultaneously $x^{1,1}\in\phi(x^{0,1})$, which consists of rational numbers, and $x^{1,1}\in\psi(x^{1,0})$ contained in the set of irrational numbers. A contradiction. ∎

In view of Theorem 1 it is clear that the pair ϕ,ψ in the preceding example is not associated with any pair of commuting Markov operators. In Section 2 we give another example indicating that even continuous (thus having trajectories) commuting transition systems need not be associated with commuting Markov operators. The existence of trajectories for commuting transition systems breaks completely at dimension three; in [DI] we find an example of three commuting continuous transition systems induced by three commuting Markov operators on a finite space with empty set of trajectories.

Let us mention some relations between transition systems and subshifts of finite type—a subject of extensive study in symbolic dynamics. It is known that the trajectories of a single transition system acting on a finite space can be identified with a subshift of finite type. On the other hand, every subshift of finite type may be recoded (using higher block codes) to obtain a shift space induced by a transition system (for information about subshifts of finite type we refer to [LM]). The subshift induced by a transition system ϕ is, in fact, the topological support of a certain stationary measure μ (as at the beginning of this paper) for the process corresponding to a transition probability T such that $\phi_T = \phi$. The set of trajectories of a finite collection of commuting transition systems on a finite space (if nonempty) can be identified with a multidimensional subshift of finite type. It seems to be an open question whether every multidimensional subshift of finite type can be represented in this way (it certainly can, if we drop the commutation requirement). As our next example shows, even when nonempty, a multidimensional subshift of finite type need not admit an invariant measure representing a multidimensional Markov process in the sense of [DI].

2. Commuting incompatibility. In the following, we construct an example of commuting 0-1 matrices not associated with commuting stochastic matrices.

Example 2. Let $X = C_3 \times C_3$. Define transition systems on X by

$$\phi = \begin{cases} I \times (I \cup R) & \text{on } \{0\} \times C_3, \\ I \times (I \cup R^2) & \text{on } \{1\} \times C_3, \\ I \times (R \cup R^2) & \text{on } \{2\} \times C_3, \end{cases}$$

$$\psi = \begin{cases} R \times (I \cup R) & \text{on } \{0\} \times C_3, \\ R \times (I \cup R^2) & \text{on } \{1\} \times C_3, \\ R \times (R \cup R^2) & \text{on } \{2\} \times C_3. \end{cases}$$

Since I and R commute, and superpositions of any two systems from among $I \cup R$, $I \cup R^2$, $R \cup R^2$ give $I \cup R \cup R^2$, the systems ϕ and ψ commute. We will show that there exist no commuting stochastic matrices T_ϕ, T_ψ inducing the transition systems ϕ, ψ.

Suppose such matrices exist. Enumerating the elements of X so that the point (p, q) has number $3p + q$, we can write

$$T_\phi = \left[\begin{array}{ccc|ccc|ccc} a & a' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b & b' & 0 & 0 & 0 & 0 & 0 & 0 \\ c' & 0 & c & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & d & 0 & d' & 0 & 0 & 0 \\ 0 & 0 & 0 & e' & e & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & f' & f & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & g' \\ 0 & 0 & 0 & 0 & 0 & 0 & h' & 0 & h \\ 0 & 0 & 0 & 0 & 0 & 0 & i & i' & 0 \end{array}\right],$$

$$T_\psi = \left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & r & r' & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & s & s' & 0 & 0 & 0 \\ 0 & 0 & 0 & t' & 0 & t & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & u & 0 & u' \\ 0 & 0 & 0 & 0 & 0 & 0 & v' & v & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & w' & w \\ \hline 0 & x & x' & 0 & 0 & 0 & 0 & 0 & 0 \\ y' & 0 & y & 0 & 0 & 0 & 0 & 0 & 0 \\ z & z' & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right],$$

where $a + a' = 1, b + b' = 1, \ldots, z + z' = 1$ and $a, b, c, \ldots, z \in (0,1)$; none of the numbers $a, b, \ldots, z$ is 0 or 1, since ϕ and ψ are supposed to give the precise supports of the transition measures. In order that the matrices T_ϕ and T_ψ commute, it is necessary that

$$\begin{bmatrix} a & a' & 0 \\ 0 & b & b' \\ c' & 0 & c \end{bmatrix} \cdot \begin{bmatrix} r & r' & 0 \\ 0 & s & s' \\ t' & 0 & t \end{bmatrix} = \begin{bmatrix} r & r' & 0 \\ 0 & s & s' \\ t' & 0 & t \end{bmatrix} \cdot \begin{bmatrix} d & 0 & d' \\ e' & e & 0 \\ 0 & f' & f \end{bmatrix},$$

i.e.

$$\begin{bmatrix} ar & ar' + a's & a's' \\ b't' & bs & bs' + b't \\ c'r + ct' & c'r' & ct \end{bmatrix} = \begin{bmatrix} rd + r'e' & r'e & rd' \\ se' & se + s'f' & s'f \\ t'd & tf' & tf + t'd' \end{bmatrix}. \tag{2}$$

Therefore

$$\begin{aligned} ar &= rd + r'e', & ar' + a's &= r'e, \\ bs &= se + s'f', & bs' + b't &= s'f, \\ ct &= tf + t'd', & ct' + c'r &= t'd, \end{aligned}$$

and further

$$\begin{aligned} r(a-d) &= r'e', & a's &= r'(e-a), \\ s(b-e) &= s'f', & b't &= s'(f-b), \\ t(c-f) &= t'd', & c'r &= t'(d-c). \end{aligned} \tag{3}$$

Since all the coefficients belong to $(0, 1)$, the right-hand sides are positive and we obtain $a > d > c > f > b > e > a$. A contradiction. ∎

In fact, there are no commuting matrices T_ϕ and T_ψ inducing transition systems inscribed in ϕ, ψ. To prove this one has to allow the possibility of some (or all) of $a, b, \ldots, z$ being 0 or 1. Notice that in this case (3) together with the additional assumption $r, s, t \in (0, 1)$ implies only weak inequalities between the coefficients. Hence, we obtain $a = b = \ldots = f = 1$, which, considering the whole matrices, leads to a contradiction. The proof is completed by showing that none of r, s, t can belong to $\{0, 1\}$.

REFERENCES

[D] T. Downarowicz, *Strictly nonpointwise Markov operators and weak mixing*, in: Ergodic Theory and its Connections with Harmonic Analysis (Alexandria, 1993), London Math. Soc. Lecture Note Ser. 205, Cambridge Univ. Press, 1995, 259–272.

[DI] T. Downarowicz and A. Iwanik, *Multiple recurrence for discrete time Markov processes II*, Colloq. Math. 55 (1988), 311–316.

[I] A. Iwanik, *Multiple recurrence for discrete time Markov processes*, ibid. 51 (1987), 165–174.

[LM] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge Univ. Press, 1995.

[M] E. Michael, *Continuous selections*, Ann. of Math. 63 (1956), 361–382.

[N] J. Neveu, *Mathematical Foundations of the Calculus of Probability*, Holden-Day, 1965.

Institute of Mathematics
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: frej@im.pwr.wroc.pl

Received 22 August 2000;
revised 5 March 2001 (3963)