BULLETIN OF THE POLISH
ACADEMY OF SCIENCES
MATHEMATICS
Vol. 55, No. 1, 2007

An Application of Skew Product Maps to Markov Chains

by

Zbigniew S. KOWALSKI

Presented by Stanisław KWAPIEŃ

Summary. By using the skew product definition of a Markov chain we obtain the following results:

(a) Every k-step Markov chain is a quasi-Markovian process.
(b) Every piecewise linear map with a Markovian partition defines a Markov chain for every absolutely continuous invariant measure.
(c) Satisfying the Chapman–Kolmogorov equation is not sufficient for a process to be quasi-Markovian.

0. Introduction. Consider a stationary Markov chain with finite state space, i.e., a probability space (Ω, P), a sequence of $(X_n)_{n=0}^{\infty}$ random variables $X_n : \Omega \to \{1, \ldots, s\}$ a probability vector $\vec{p} = (p_1, \ldots, p_s)$, a stochastic matrix $\Pi = (p_{ij})_{s\times s}$ such that $\vec{p}\Pi = \vec{p}$ and

$$P(X_{n+1} = j \mid X_0 = i_0, \ldots, X_n = i_n) = P(X_{n+1} = j \mid X_n = i_n) = p_{i_n j}.$$

To apply ergodic theory here, we need a sequence representation of the process. Therefore we assume that $\Omega = \{1, \ldots, s\}^N$, $N = \{0, 1, 2, \ldots\}$, P is the measure given by

$$P(\omega : \omega_0 = i_0, \ldots, \omega_n = i_n) = p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}$$

and σ is the 1-sided shift on Ω, $(\sigma\omega)_i = \omega_{i+1}$. Finally, we get the measure preserving dynamical system $(\Omega, \mathcal{A}, P, \sigma, \alpha)$, where $\mathcal{A}$ is the σ-algebra generated by the cylinder sets $\{\omega : \omega_{i_1} = j_1, \ldots, \omega_{i_n} = j_n\}$ and $\alpha = \{A_i\}_{i=1}^{s}$, where $A_i = \{\omega : \omega_0 = i\}$, $i = 1, \ldots, s$. The process $(X'_n)_{n=0}^{\infty}$ given by $X'_n(\omega) = i \Leftrightarrow \sigma^n(\omega) \in A_i$ is a Markov chain with the same distribution

2000 *Mathematics Subject Classification*: Primary 37A05; Secondary 60J10.
Key words and phrases: skew product, Markov chain, quasi-Markovian process.
Research supported by grant MENII 1 P03A 021 29, Poland.

as $(X_n)_{n=0}^{\infty}$. The above process is denoted by (σ,α) and called a *Markov process.*

For our aims we will use more general notations. Let $(X,\mathcal{A},m,f,\alpha)$ be a measure preserving dynamical system where $(X,\mathcal{A},m)$ is a probability Lebesgue space and additionally f is a *positively nonsingular* map (i.e. $m(A)=0\Rightarrow m(f(A))=0$), and $\alpha=\{A_i\}_{i=1}^s$ is a generating partition, i.e. $\bigvee_{i=0}^{\infty} f^{-i}\alpha=\mathcal{A}$. We denote by (f,α) the process $(X_n)_{n=0}^{\infty}$ such that $X_n(x)=i\Leftrightarrow f^n(x)\in A_i$. For our aims it is convenient to use an equivalent definition of Bernoulli and Markov processes (see [4]). To this end let us consider another Lebesgue probability space $(Y,\mathcal{B},q)$ and let $\{T_i\}_{i=1}^s$ be a family of positively and negatively nonsingular maps of Y into Y (T_i is *negatively nonsingular* if $q(B)=0\Rightarrow q(T_i^{-1}B)=0$). The process (f,α) and the family $\{T_i\}_{i=1}^s$ define the skew product map

$$T(x,y)=(f(x),T_{a(x)}(y)), \tag{1}$$

where $a:X\to\{1,\dots,s\}$ is determined by $a(x)=i\Leftrightarrow x\in A_i$. Let $E(f,\alpha)$ be the class of functions g such that there exists a Lebesgue probability space $(Y,\mathcal{B},q)$ and a family of positively and negatively nonsingular maps $\{T_i\}_{i=1}^s$ such that g is the density of an absolutely continuous invariant measure (a.c.i.m.) under the skew product as in (1).

We also denote by $\widehat{\alpha}$ the field of unions of elements of α.

DEFINITION 1. The process (f,α) is

(i) a *Bernoulli process* if for every $g\in E(f,\alpha)$, g is measurable with respect to $\mathcal{B}$,

(ii) a *Markov process* if for every $g\in E(f,\alpha)$, g is measurable with respect to $\widehat{\alpha}\times\mathcal{B}$.

Reduction of the condition in (ii) allows us to introduce quasi-Markovian processes.

DEFINITION 2. We say that (f,α) is a *quasi-Markovian process* (q.m.p.) if for every $g\in E(f,\alpha)$ the set $\{g>0\}$ belongs to $\widehat{\alpha}\times\mathcal{B}$.

Some generalizations of Definition 2 and its applications can be found in [1, 2]. By using Definitions 1 and 2 we obtain the following facts.

THEOREM 1. *If α has the Markov property (i.e. $f(\alpha)\subset\widehat{\alpha}$) and if $(f^n,\bigvee_{i=0}^{n-1} f^{-i}\alpha)$ is a q.m.p. for some n with respect to m then (f,α) is also a q.m.p.*

THEOREM 2. *If $X=[0,1]$, f is piecewise linear, α has the Markov property and $f|A_i$ is linear for $i=1,\dots,s$ then (f,α) is a Markov process for every f-invariant probability measure which is absolutely continuous with respect to the Lebesgue measure.*

Let $(X_n)_{n=0}^{\infty}$ be a stationary process with s states. Moreover, let m be the measure on $\{1,\ldots,s\}^N$ determined by the finite-dimensional distributions of $(X_n)_{n=0}^{\infty}$.

DEFINITION 3. We say that $(X_n)_{n=0}^{\infty}$ is a *k-step Markov chain* if

$$m(\omega_{n+1}=j\,|\,\omega_0=i_0,\ldots,\omega_n=i_n) = m(\omega_{n+1}=j\,|\,\omega_{n-k+1}=i_{n-k+1},\ldots,\omega_n=i_n).$$

A 1-step Markov chain is a Markov chain. As a corollary to Theorem 1 we get

THEOREM 1′. *If a stationary process $(X_n)_{n=0}^{\infty}$ is a k-step Markov chain for some $k\ge 1$ then it is a q.m.p.*

Concerning Theorem 2 let us remark that if f is not piecewise linear but piecewise monotonic then (f,α) is not a Markov process in general. However, in many cases it turns out to be a q.m.p., for example if f is a Lasota–Yorke or Misiurewicz map ([4]). Let us turn to Theorem 1′. If we replace the Markov chain conditions by the Chapman–Kolmogorov equation (for the definition see below) then we may fall outside the class of quasi-Markovian processes (see Section 3).

DEFINITION 4. We say that a process (f,α) satisfies the *C-K-equation* if

$$m(f^{-n}A_j\,|\,A_i)=(\Pi^n)_{ij}\quad\text{ for } i,j=1,\ldots,s \text{ and } n\in\mathbb{N}.$$

Here $\Pi_{ij}=m(f^{-1}A_j\,|\,A_i)$ for $i,j=1,\ldots,s$.

1. Proof of Theorem 1. We will use the skew product description of processes. Let us consider the skew product map

$$T(x,y)=(f(x),T_{a(x)}(y))$$

where $a:X\to\alpha$ is determined by $a(x)=A\Leftrightarrow x\in A$. Here $\{T_A\}_{A\in\alpha}$ is a family of positively and negatively nonsingular maps of Y into Y. Let ν be a T-invariant measure absolutely continuous with respect to $m\times q$. The measure ν is also T^n-invariant. Here

$$T^n(x,y)=(f^n(x),T_{b(x)}(y))$$

where $b:X\to\beta=\bigvee_{i=0}^{n-1}f^{-i}\alpha$ is defined similarly to a. Here

$$T_B(y)=T_{A_n}\circ\cdots\circ T_{A_1}(y)\quad\text{ for } B=A_1\cap f^{-1}(A_2)\cap\cdots\cap f^{-(n-1)}(A_n).$$

Therefore,

$$\left\{\frac{d\nu}{d(m\times q)}>0\right\}=\bigcup_{B\in\beta}B\times D_B$$

if (f^n, m, β) is a q.m.p. by Definition 2. From
$$T^n\Big(\bigcup_{B\in\beta} B\times D_B\Big) = \bigcup_{B\in\beta} B\times D_B$$
we get
$$\bigcup_{B\in\beta} B\times D_B = \bigcup_{B\in\beta} f^n(B)\times T_B(D_B) = \bigcup_{A\in\alpha} A\times C_A.$$
Here we use $f^n(B)\in\widehat{\alpha}$ for $B\in\beta$. Therefore (f,m,α) is a q.m.p.

2. Proof of Theorem 2. Consider the skew product map
$$T(x,y) = (f(x), T_{a(x)}(y))$$
as in (1). Let P be the *Frobenius–Perron* (F-P) *operator* for T, i.e.
$$\int\limits_{T^{-n}E} G\,d(\lambda\times q) = \int\limits_{E} P(G)\,d(\lambda\times q)$$
for $E\in\mathcal{A}\times\mathcal{B}$ and $G\in L_1(\lambda\times q)$. Here λ is the Lebesgue measure. By the definition of T we have
$$P(wh)(x,y) = \sum_{i=1}^{s} a_i w(f_i^{-1}(x))1_{f(A_i)}(x)P_ih(y) \quad \text{for } w\in L_1(\lambda), h\in L_1(q),$$
where $f_i = f|A_i$, $a_i = 1/f_i'(x)$ for $x\in A_i$ and P_i denotes the F-P operator for T_i, $i=1,\ldots,s$. Let μ be f-invariant, $\mu\ll\lambda$. Moreover, let ν be T-invariant and $\nu\ll\mu\times q$. Then also $\nu\ll\lambda\times q$ and therefore for $G = d\nu/d(\lambda\times q)$ we obtain $P^nG = G$ for $n=1,2,\ldots$. Let $A_{i_1,\ldots,i_n} = A_{i_1}\cap f^{-1}(A_{i_2})\cap\cdots\cap f^{-(n-1)}(A_{i_n})$. Thus, as $f(\alpha)\subset\widehat{\alpha}$,
$$P(1_{A_{i_1\cdots i_n}}h) = a_{i_1}1_{A_{i_2\cdots i_n}}P_{i_1}h.$$
Consequently,
$$P^n(1_{A_{i_1\cdots i_n}}h) = \sum_{i=1}^{s} 1_{A_i}h_i.$$
By using approximation arguments as in [5] we conclude that
$$G = \sum_{i=1}^{s} 1_{A_i}g_i.$$
By repeating a similar reasoning for μ and the F-P operator for f we get
$$\frac{d\mu}{d\lambda} = \sum_{i=1}^{s} c_i 1_{A_i}.$$
Hence,
$$\frac{d\nu}{d(\mu\times\lambda)} = \frac{d\nu}{d(\lambda\times q)}\bigg(\frac{d(\mu\times q)}{d(\lambda\times q)}\bigg)^{-1} = \sum_{i=1}^{s} 1_{A_i}d_i.$$
Therefore (f,α) is a Markov process with respect to μ. ∎

Let us remark that Theorem 3 in [7] and Theorem 6.3 in [8] are special cases of Theorem 2. Here the Markov property of (f,α) has been obtained for some classes of piecewise linear transformations by using an explicit definition of invariant measure.

3. C-K-process which is not a q.m.p. We recall the construction of Courbage and Hamdan from [3], using their notation. Let $K=\{0,\dots,k-1\}$ and $\Omega=K^N=\{\omega:\omega(i)\in K,\ i=0,1,2,\dots\}$. We say that a probability measure μ on K^{n+1} is *invariant* if for any subsets $A_1,\dots,A_n$ of K,

$$\mu(K\times A_1\times\dots\times A_n)=\mu(A_1\times\dots\times A_n\times K).$$

Denote by $\mathrm{Inv}(K^{n+1})$ the set of invariant measures. For $\mu\in\mathrm{Inv}(K^{n+1})$ we define the measure $\nu_0=\Phi(\mu)$ on Ω by

$$\begin{aligned}&\nu_0(\{\omega\in\Omega:\omega_0=x_0,\dots,\omega_{pn}=x_{pn}\})\\&\quad=\mu(x_0,\dots,x_n)\mu(x_{n+1},\dots,x_{2n}\mid x_n)\cdots\mu(x_{(p-1)n+1},\dots,x_{pn}\mid x_{(p-1)n})\end{aligned}$$

for all $p\ge1$ and $x=(x_i)_{i=0}^{pn}\in K^{n+1}$. The measure ν_0 is σ^n-invariant where σ is the 1-sided shift on Ω. Now, we proceed to determine a suitable measure $\mu\in\mathrm{Inv}(K^{n+1})$. Let Π be a $k\times k$ stochastic strictly positive matrix and $\vec{p}$ be a row probability vector which is invariant under Π. By analyzing the proof of [4, Theorem 3.1] we conclude that there exists a measure $\mu\in\mathrm{Inv}(K^{n+1})$, for some prime number $n\ge3$, such that

$$\mu(A_0\times\dots\times A_n)>0 \tag{2}$$

for any $A_0,\dots,A_n\subseteq K$. The above holds because Π is strictly positive and $\mu|_{K^n}=\mu_\Pi|_{K^n}$ (by conditions (3.4), (3.5) from [4]). Here μ_Π denotes the $(\Pi,\vec{p})$ Markovian measure. The measure

$$\nu=\frac1n\sum_{i=0}^{n-1}\sigma^i\nu_0,$$

where $\nu_0=\Phi(\mu)$, is not $(\Pi,\vec{p})$ Markovian since $\mu\ne\mu_\Pi|_{K^{n+1}}$ (by [4, (3.18)]). Moreover the process $(\sigma,\nu,\mathcal{P})$ satisfies the C-K-equation with Π, i.e. $\nu(\sigma^{-n}P_j\mid P_i)=(\Pi^n)_{i,j}$ for $n=1,2,\dots$, by [4, (3.6)] and by the definition of ν. Here $P_i=\{\omega\in\Omega:\omega(0)=i\}$. We will prove that $(\sigma.\nu,\mathcal{P})$ is not a q.m.p. To this end we show that there exists a Borel invariant set $E\subset\Omega$ of measure one such that (E,ν,σ) is positively nonsingular and $\nu(\sigma(P_i\cap E))=1$ for $i=0,\dots,k-1$. Combining this with the ergodic properties of (σ,ν) we get the desired conclusion. Let $\mathcal{P}_0^n=\bigvee_{i=0}^{n-1}\sigma^{-i}\mathcal{P}$. By the definition of ν_0 and by (2) we get

LEMMA 1. *The process $(\sigma^n,\sigma^i\nu_0)$ is a Markov chain with vector $\vec{q}$:*

$$\vec{q}_{(x_0,\dots,x_{n-1})}=\mu(x_0,\dots,x_{n-i})\mu(x_{n-i+1},\dots,x_{n-1}\mid x_{n-i})$$

and the strictly positive matrix W:

$$W_{(x_0,\ldots,x_{n-1}),(y_0,\ldots,y_{n-1})} = \mu(y_0,\ldots,y_{n-i}\,|\,x_{n-i},\ldots,x_{n-1})\mu(y_{n-i+1},\ldots,y_{n-1}\,|\,y_{n-i}),$$

for $i=2,\ldots,n$.

The case $i=1$ is considered in [3]. Here $\vec{q}_{(x_0,\ldots,x_{n-1})} = \mu(x_0,\ldots,x_{n-1})$ and

$$W_{(x_0,\ldots,x_{n-1}),(y_0,\ldots,y_{n-1})} = \mu(y_0,\ldots,y_{n-1}\,|\,x_{n-1}).$$

By [6, Chapter 10] there exists an invariant set $\Omega' \subset \Omega$ with $\nu(\Omega')=1$ such that $\sigma|\Omega'$ is positive nonsingular. The Markov chains $(\sigma^n, \sigma^i\nu_0)$ are ergodic, being aperiodic, for $i=1,\ldots,n$. Therefore the supports of $\sigma^i\nu_0$ are pairwise disjoint sets with respect to ν. Hence there exists a Borel set $E_0 \subset \Omega'$ such that $\sigma^n(E_0) = E_0 = \sigma^{-n}(E_0)$, $\nu_0(E_0)=1$, the sets $\sigma^{-i}E_0$ are pairwise disjoint and $\sigma^i\nu_0(\sigma^i(E_0))=1$ for $i=1,\ldots,n$. Set $\mu_i=\sigma^i\nu_0$ and $E_i=\sigma^iE_0$, $i=0,\ldots,n-1$. Let $E=\bigcup_{i=0}^{n-1}E_i$. Then $\sigma^{-1}E=E$ and $\nu(E)=1$.

Remark. (σ,ν) is not totally ergodic as $\sigma^n(E_0)=E_0$ and $\nu(E_0)=1/n$.

Obviously (σ,ν) is positively nonsingular on E.

Lemma 2. *The partition* $\mathcal{P}_E = \{P_i\cap E\}_{i=0}^{k-1}$ *of* E *is a Markovian generator for* (σ,ν). *Moreover*, $\nu(\sigma(P_i\cap E))=1$ *for* $i=0,\ldots,k-1$.

Proof. It suffices to show that $\nu(\sigma(P_j\cap E))=1$ for $j=0,\ldots,k-1$. By the definition of ν,

$$\nu(\sigma(P_j\cap E)) = \frac{1}{n}\sum_{i=1}^{n}\mu_i(\sigma(P_j\cap E_{i-1})).$$

By (2) and Lemma 1, $\mu_i(\sigma^n([x_1,\ldots,x_{n-1},j]\cap E_i))=1$ for every block $[x_1,\ldots,x_{n-1}]$. Hence the inclusion

$$\sigma^{n-1}([x_1,\ldots,x_{n-1},j]\cap E_i) \subset P_j\cap E_{(i+n-1)\bmod n} = P_j\cap E_{i-1}$$

implies $\mu_i(\sigma(P_j\cap E_{i-1}))=1$. Hence $\mu_i(\sigma(P_j\cap E_{i-1}))=1$ for $i=1,\ldots,n$. This finishes the proof. ∎

Theorem 3. The process $(\sigma,\nu,\mathcal{P}_E)$ is not a q.m.p.

Proof. By the construction (σ,ν) is ergodic but not weakly mixing. The previous observations imply $\sigma|E$ is positively nonsingular and $\mathcal{P}$ is a Markovian generator for (σ,ν). Let $\mathcal{P}_\sigma$ denote the smallest field which contains $\{\sigma(P_i\cap E)\}_{i=0}^{k-1}$. By Lemma 2, $\mathcal{P}_\sigma=\{\emptyset,\Omega\}$. Assume that $(\sigma,\nu,\mathcal{P}_E)$ is a q.m.p. Then, by [4, Lemma 2], all eigenfunctions of σ are $\mathcal{P}_\sigma$-measurable, which implies that (σ,ν) is weakly mixing. This contradicts our assumption. ∎

References

[1] J. Aaronson and M. Denker, *Local limit theorems for Gibbs–Markov maps*, Stoch. Dyn. 1 (2001), 193–237.

[2] J. Aaronson, M. Denker, O. Sarig and R. Zweimüller, *Aperiodicity of cocycles and conditional local limit theorems*, ibid. 4 (2004), 31–62.

[3] M. Courbage and D. Hamdan, *An ergodic Markov chain is not determined by its two-dimensional marginal laws*, Statist. Probab. Lett. 37 (1998), 35–40.

[4] Z. S. Kowalski, *Quasi-Markovian transformations*, Ergodic Theory Dynam. Systems 17 (1997), 885–897.

[5] T. Morita, *Deterministic version lemmas in ergodic theory of random dynamical systems*, Hiroshima Math. J. 18 (1988), 15–29.

[6] W. Parry, *Entropy and Generators in Ergodic Theory*, Benjamin, New York, 1969.

[7] I. Shiokawa, *Ergodic properties of piecewise linear transformations*, Proc. Japan Acad. 46 (1970), 1122–1125.

[8] K. M. Wilkinson, *Ergodic properties of a class of piecewise linear transformations*, Z. Wahrsch. Verw. Gebiete 31 (1975), 303–328.

Zbigniew S. Kowalski
Institute of Mathematics and Informatics
Wrocław University of Technology
Wybrzeże St. Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: kowalski@pwr.wroc.pl

Received July 20, 2006;
received in final form February 23, 2007 (7541)