C O L L O Q U I U M   M A T H E M A T I C U M

VOL. 90 2001 NO. 2

A FAMILY OF STATIONARY PROCESSES WITH INFINITE MEMORY HAVING THE SAME p-MARGINALS. ERGODIC AND SPECTRAL PROPERTIES

BY

M. COURBAGE and D. HAMDAN (Paris)

Abstract. We construct a large family of ergodic non-Markovian processes with infinite memory having the same p-dimensional marginal laws of an arbitrary ergodic Markov chain or projection of Markov chains. Some of their spectral and mixing properties are given. We show that the Chapman–Kolmogorov equation for the ergodic transition matrix is generically satisfied by infinite memory processes.

1. Introduction. According to the Kolmogorov theorem a stochastic process is defined by the family of all compatible finite-dimensional distributions. It is a natural question to ask about the existence or non-existence of some subfamily of finite-dimensional marginal laws which completely determine the process. In the case of a finite state space, there exist stationary ergodic processes which are completely determined by the family of 2-dimensional marginal laws [8].

In this paper, we characterize a class of Markovian and non-Markovian processes non-uniquely determined by the above family. This means that if $X = (X_n)$, $n \in \mathbb{Z}$, is an ergodic stationary process on a finite state space $K = \{0, 1, \ldots, k-1\}$, then there is an ergodic process $Y = (Y_n)$, $Y_n \in K$, distinct from X such that the law of $(Y_{n_1}, \ldots, Y_{n_p})$ is equal to the law of $(X_{n_1}, \ldots, X_{n_p})$, for any p-uple $(n_1, \ldots, n_p)$ of integers. For Bernoulli processes, well known examples of non-independent processes which are pairwise independent have been given by Lévy, Feller, Janson and Bradeley (see references in [1], [3], [4]). Ergodic examples of such processes have been given by Robertson and Womack [12], Flaminio [9], Bretagnolle and Kłopotowski [1] and the authors [3]–[6]. In [3], [5] we constructed, in the case of Markov chains with strictly positive transition matrix, a continuum of non-Markovian processes having the same 2-dimensional marginal laws, and we extended [4] that result to mixing Markov chains. This implies that

2000 *Mathematics Subject Classification*: 60Gxx, 28Dxx.

Key words and phrases: Chapman–Kolmogorov measure, Markov chain, infinite memory process, mixing subsets, spectral type, integral over an automorphism.

the Chapman–Kolmogorov equation is also satisfied by non-Markovian processes.

The construction given in [3]–[5] is extended here in order to provide new ergodic processes having the same p-dimensional marginals as any ergodic Markov chain or some family of non-Markovian functions of Markov chains. Let σ be the left-shift transformation on $K^{\mathbb{Z}}$. In Sections 2 and 4, we give a family of non-Markovian processes resulting from a mapping ϕ_n acting on the set M_n of shift-invariant measures on K^{n+1}, $n \geq 1$, into the set $M(K^{\mathbb{Z}}, \sigma)$ of all shift-invariant probability measures on $K^{\mathbb{Z}}$. We prove new properties of $\phi_n(\mu)$, $\mu \in M_n$ (Lemmas 2 and 3 and Theorem 4), namely, ϕ_n is one-to-one and marginals preserving. This allows us to reduce the construction of the above family of processes to M_n. In [7] and [10], the last problem is solved for 2-marginals and p-marginals respectively. In Section 5, we prove that these processes have infinite memory and that the set of measures with infinite memory is dense in the class of measures having the same 2-dimensional marginals of any ergodic Markov chain. These shift dynamical systems are also characterized as factors of an integral automorphism. This allows us to show in Section 3 that they have countable Lebesgue spectrum and finite point spectrum. We show that they have a continuum of mutually mixing subsets, finite Pinsker σ-algebra and consequently positive entropy. We characterize a family of functions with decaying correlations.

2. The class $\phi_n M_n$. Let $K = \{0, \ldots, k-1\}$, $k \geq 2$. Let $\Omega = K^{\mathbb{Z}}$ be the set of all doubly infinite sequences $\omega = (\omega_i)$, $\omega_i \in K$, $i \in \mathbb{Z}$, and $\mathcal{A}$ the σ-algebra generated by the cylindric sets $\{\omega : \omega_{i_1} = j_1, \ldots, \omega_{i_n} = j_n\}$, also denoted by $\{\omega_{i_1} = j_1, \ldots, \omega_{i_n} = j_n\}$. The shift transformation acts on Ω by $(\sigma(\omega))_i = \omega_{i+1}$. We denote by $M(\Omega, \sigma)$ the set of all σ-invariant probability measures on $(\Omega, \mathcal{A})$. We also denote by M_n, $n \geq 2$, the set of all probability measures on K^{n+1} that are invariant under the shift, that is,

$$\sum_y \mu(\{\omega_0 = y, \omega_1 = x_0, \ldots, \omega_{p+1} = x_p\}) = \mu(\{\omega_0 = x_0, \ldots, \omega_p = x_p\}), \qquad 0 \leq p \leq n-1.$$

We assume that all the probability measures μ we consider are such that $\mu(\{\omega_0 = y\}) > 0$ for all y.

We define a probability measure $\nu_0 = \nu_0(\mu)$ on $K^{\mathbb{N}}$ by the identity

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

for any positive integer p. Here and in what follows, $\mu(x_0, \ldots, x_t)$ means $\mu(\{\omega \in K^{n+1} : \omega_0 = x_0, \ldots, \omega_t = x_t\})$ and $\mu(x_s, \ldots, x_{s+t} \mid x_{s-1})$ means the conditional probability of the set $\{\omega \in K^{n+1} : \omega_1 = x_s, \ldots, \omega_{1+t} = x_{s+t}\}$ given the set $\{\omega \in K^{n+1} : \omega_0 = x_{s-1}\}$.

As ν_0 is σ^n-invariant: $\sigma^n \nu_0 := \nu_0 \circ \sigma^{-n} = \nu_0$, the measure

$$\nu := \frac{1}{n} \sum_{i=0}^{n-1} \sigma^i \nu_0$$

is σ-invariant. We denote by $\phi_n \mu$ the measure on $K^{\mathbb{Z}}$ defined by

$$(\phi_n \mu)(\{\omega \in \Omega : \omega_l = x_0, \ldots, \omega_{l+t} = x_t\}) \\ := \nu(\{\omega \in K^{\mathbb{N}} : \omega_0 = x_0, \ldots, \omega_t = x_t\})$$

for any $l \in \mathbb{Z}$.

The mapping ϕ_n can be identified with a mapping defined on the set $M(\Omega, \sigma)$ of all σ-invariant measures on Ω. In fact, any $\mu \in M_n$ can be extended to a stationary n-order Markov chain on K. Conversely, the restriction to the coordinates $(\omega_0, \ldots, \omega_n)$ of any measure from $M(\Omega, \sigma)$ belongs to M_n. Thus, $\{\phi_n : n \geq 2\}$ is a family of mappings from $M(\Omega, \sigma)$ into itself.

The measure $\phi_n \mu$ is a function of a stationary Markov chain on an extended state space. The state space of this Markov chain is $L = \bigcup_{m=2}^{n+1} K^m$ where the only transition permitted leads from $(x_0, \ldots, x_m)$ to $(x_0, \ldots, x_{m+1})$ with transition probability

$$P_{(x_0,\ldots,x_m),(x_0,\ldots,x_{m+1})} = \mu(x_{m+1} \,|\, x_0, \ldots, x_m) \tag{2}$$

if $1 \leq m \leq n-1$, and

$$P_{(x_0,\ldots,x_n),(x_n,x_{n+1})} = \mu(x_{n+1} \,|\, x_n) \tag{3}$$

otherwise. The stationary probability row vector $p_{(x_0,\ldots,x_m)}$ is equal to $n^{-1}\mu(x_0, \ldots, x_m)$, $1 \leq m \leq n$. Denoting this Markov chain by (z_i), $z_i \in L$, $i \in \mathbb{Z}$, we consider the function $y_i = f(z_i)$, where

$$f(x_0, \ldots, x_m) = x_m, \quad 1 \leq m \leq n.$$

It is straightforward to verify that the finite joint distributions of (y_i) are the same as those of $\phi_n \mu$.

There is another relation between $\phi_n \mu$ and Markov chains. In fact, it is straightforward to verify that ν_0 is a σ^n-invariant measure isomorphic to the Markov chain on the state space formed by the points $(x_0, \ldots, x_{n-1})$ of K^n such that $\mu(x_0, \ldots, x_{n-1}) > 0$ with transition matrix W defined by

$$W_{(x_0,\ldots,x_{n-1}),(y_0,\ldots,y_{n-1})} = \mu(y_1, \ldots, y_{n-1} \,|\, y_0) \cdot \mu(y_0 \,|\, x_0, \ldots, x_{n-1}) \tag{4}$$

and the invariant row probability vector $\mu(x_0, \ldots, x_{n-1})$. The mapping $\theta : K^{\mathbb{Z}} \to (K^n)^{\mathbb{Z}}$ defined by

$$(\theta(\omega))_i = (\omega_{in}, \ldots, \omega_{(i+1)n-1}) \tag{5}$$

realizes the isomorphism between the Markov chain $((K^n)^{\mathbb{Z}}, \sigma, \mu_W)$ and $(K^{\mathbb{Z}}, \sigma^n, \nu_0)$, since $\theta\sigma^n = \sigma\theta$, and $\mu_W = \nu_0\theta^{-1}$.

It can be shown as in [5] that $\phi_n\mu$ converges to μ in the w^*-topology so that $\bigcup_{n=1}^{\infty} \phi_n M(\Omega, \sigma)$ is dense in $M(\Omega, \sigma)$.

We now characterize the case in which $\phi_n\mu$ is a Markov chain.

We introduce some notation. If Π is a $k \times k$ stochastic matrix and p is an invariant row probability vector, we denote by μ_Π the stationary Markov chain with initial probability distribution p and transition matrix Π. If $\mu \in M(\Omega, \sigma)$, we define the stochastic transition matrices $A_r(\mu)$ by

$$(A_r(\mu))_{i,j} := \mu(\omega_r = j \mid \omega_0 = i) \tag{6}$$

for $i, j \in K$ and $r \in \mathbb{N}$; $A_1(\mu)$ will be denoted by $\Pi(\mu)$ or simply Π. We also denote by $R_n\mu$ the restriction of μ to K^{n+1}.

Lemma 1. *Let $\mu \in M_n$ and $\Pi := A_1(\mu)$. Then we have the implications* (i)⇒(ii)⇒(iii)⇔(iv)⇔(v)⇒(vi), *where*:

(i) *$\phi_n\mu$ is strongly mixing*,
(ii) *$\phi_n\mu$ is weakly mixing*,
(iii) *$\nu_0(\mu)$ is σ-invariant*,
(iv) $R_n\mu = R_n\mu_\Pi$,
(v) $\nu_0(\mu) = \mu_\Pi$,
(vi) $\phi_n\mu = \mu_\Pi$.

If moreover Π is irreducible and aperiodic these assertions are equivalent.

Therefore, if a measure μ from M_n is distinct from the Markov measure, then $\phi_n\mu$ is *non-weakly mixing and non-Markovian.* We also know from [5] that *the only fixed points of $\phi_n R_n$ are the Markov chains.*

The next result shows that the mapping $\phi_n : M_n \to M(\Omega, \sigma)$ is one-to-one.

Lemma 2. *Let $\mu_1, \mu_2 \in M_n$. If $\phi_n\mu_1 = \phi_n\mu_2$, then $\mu_1 = \mu_2$.*

Proof. The proof is by induction. Let μ_1, μ_2 be such that $\phi_n\mu_1 = \phi_n\mu_2$. Then $\mu_1(x_0, x_1) = \mu_2(x_0, x_1)$. Suppose that $\mu_1(x_0, \ldots, x_p) = \mu_2(x_0, \ldots, x_p)$ for $(x_0, \ldots, x_p)$-cylinders and $1 \le p \le m \le n-1$. We now compute

$$\begin{aligned}\phi_n\mu_1(x_0, \ldots, x_{m+1}) &= \frac{1}{n}\sum_{i=0}^{n-1} \sigma^i\nu_0(x_0, \ldots, x_{m+1}) \\ &= \frac{1}{n}\sum_{i=0}^{n-1} \nu_0(\mu_1)(\{\omega_i = x_0, \ldots, \omega_{i+m+1} = x_{m+1}\})\end{aligned}$$

$$= \frac{1}{n} \sum_{i=0}^{n-m-1} \nu_0(\mu_1)(\{\omega_0 = x_0, \dots, \omega_{m+1} = x_{m+1}\})$$

$$+ \frac{1}{n} \sum_{i=n-m}^{n-1} \nu_0(\mu_1)(\{\omega_i = x_0, \dots, \omega_{i+m+1} = x_{m+1}\})$$

$$= \frac{n-m}{n} \mu_1(x_0, \dots, x_{m+1})$$

$$+ \frac{1}{n} \sum_{i=n-m}^{n-1} \mu_1(x_0, \dots, x_{n-i}) \mu_1(x_{n-i+1}, \dots, x_{m+1} \mid x_{n-i}).$$

Comparing this with the same expression for μ_2, since in the above sum $1 \le n - i \le m$, we obtain by induction

$$\mu_1(x_0, \dots, x_{m+1}) = \mu_2(x_0, \dots, x_{m+1}). \ \blacksquare$$

Now, we characterize the measures $\phi_n \mu$ as factors of integrals of some automorphisms of Ω. This provides a new proof of the ergodicity of $\phi_n \mu$. We recall some useful definitions and results [2].

Definition 1. A *dynamical system* $(Y, \mathcal{G}, \nu, S)$ is a measure preserving automorphism S of the measure space $(Y, \mathcal{G}, \nu)$. It is said to be a *factor* of the dynamical system $(X, \mathcal{F}, m, T)$ if there exists a measurable map φ from X onto Y such that $m \circ \varphi^{-1} = \nu$ and $\varphi \circ T = S \circ \varphi$.

A factor of an ergodic system is ergodic.

Definition 2. Let $(X, \mathcal{F}, m, T)$ be a dynamical system and $f \in L^1(m)$. The *integral* of $(X, \mathcal{F}, m, T)$ corresponding to f is defined to be the dynamical system $(X^f, \mathcal{F}^f, m^f, T^f)$ where

$$X^f = \{(x, i) : x \in X,\ 1 \le i \le f(x)\}, \qquad m^f(A \times \{i\}) = \frac{1}{\int f\, dm} m(A)$$

and

$$T^f(x, i) = \begin{cases} (x, i+1) & \text{if } i + 1 \le f(x), \\ (Tx, 1) & \text{if } i + 1 > f(x). \end{cases}$$

(X^f, T^f, m^f) is ergodic if and only if (X, T, m) is ergodic.

We consider a particular case relevant to our systems: let (X, T, m) be the dynamical system with $X = \Omega$, $m = \nu_0(\mu)$, $T = \sigma^n$ and take f to be the constant function on Ω with value n; we shall use the following notations:

$$\widetilde{\Omega} = \Omega^f = \Omega \times \{1, \dots, n\},$$

$$\widetilde{T}(\omega, i) = \begin{cases} T^f(\omega, i) = (\omega, i+1) & \text{if } i + 1 \le n, \\ (\sigma^n \omega, 1) & \text{if } i = n, \end{cases}$$

and $\widetilde{m}(A \times \{i\}) = m^f(A \times \{i\}) = m(A)/n$. Let us define, as above, $\nu := n^{-1}\sum_{i=0}^{n-1} m \circ \sigma^{-i}$.

THEOREM 1. *With the notations above*, (Ω, ν, σ) *is a factor of* $(\widetilde{\Omega}, \widetilde{m}, \widetilde{T})$.

Proof. Define the map $\varphi : \widetilde{\Omega} \to \Omega$ by $\varphi(\omega, i) = \sigma^i\omega$. If A is a measurable subset of Ω, we have $\varphi^{-1}(A) = \bigcup_{i=1}^n \sigma^{-i}A \times \{i\}$; therefore

$$\widetilde{m}(\varphi^{-1}(A)) = \frac{1}{n}\sum_{i=1}^{n} m(\sigma^{-i}A) = \frac{1}{n}\sum_{i=1}^{n-1} m(\sigma^{-i}A) + \frac{1}{n}m(A) = \nu(A).$$

Here we have used the invariance of m under $T = \sigma^n$. On the other hand, $\varphi \circ \widetilde{T} = \sigma \circ \varphi$. In fact, if $i + 1 \leq n$ we have

$$\varphi \circ \widetilde{T}(\omega, i) = \varphi(\omega, i + 1) = \sigma^{i+1}\omega = \sigma(\sigma^i\omega) = \sigma \circ \varphi(\omega, i),$$

and if $i = n$, we have

$$\varphi \circ \widetilde{T}(\omega, n) = \varphi(\sigma^n\omega, 1) = \sigma^{n+1}\omega = \sigma \circ \varphi(\omega, n). \ ■$$

THEOREM 2. *Let* $\mu \in M(\Omega, \sigma)$ *and* $n \geq 2$. *If* $A_n(\mu)$ *is irreducible the system* $(\Omega, \phi_n\mu, \sigma)$ *is ergodic.*

Proof. By the above theorem it suffices to show that $(\Omega, \nu_0(\mu), \sigma^n)$ is ergodic. To do this, recall (see (4)) that this system is isomorphic to the Markov chain μ_W on the state space K^n with the transition matrix W defined by

$$W_{(x_0,\ldots,x_{n-1}),(y_0,\ldots,y_{n-1})} = \mu(y_1, \ldots, y_{n-1} \,|\, y_0)\mu(y_0 \,|\, x_0, \ldots, x_{n-1})$$

and the invariant row probability vector $\mu(x_0, \ldots, x_{n-1})$.

Note that it is straightforward to verify that the dynamical system $(\Omega, \sigma\nu_0(\mu), \sigma^n)$ (which is isomorphic to the first one by σ) is isomorphic to the Markov chain $\mu_{W^{(1)}}$ on the state space K^n given by the transition matrix

$$W^{(1)}_{(x_0,\ldots,x_{n-1}),(y_0,\ldots,y_{n-1})} = \mu(y_0, \ldots, y_{n-1} \,|\, x_{n-1})$$

and the invariant row probability vector $\mu(x_0, \ldots, x_{n-1})$. The theorem now results from the following lemma:

LEMMA 3. *With the above notations*:

(i) W *is irreducible* $\Leftrightarrow$ $W^{(1)}$ *is irreducible* $\Leftrightarrow$ $A_n(\mu)$ *is irreducible.*

(ii) W *is irreducible and aperiodic* $\Leftrightarrow$ $W^{(1)}$ *is irreducible and aperiodic* $\Leftrightarrow$ $A_n(\mu)$ *is irreducible and aperiodic.*

Proof. (i) The first equivalence follows from the isomorphism mentioned in the proof of the above theorem. Let us prove the second equivalence. By definition $W^{(1)}$ is irreducible if and only if for every $(x_0, \ldots, x_{n-1})$ and $(y_0, \ldots, y_{n-1})$ in K^n such that $\mu(x_0, \ldots, x_{n-1})\mu(y_0, \ldots, y_{n-1}) > 0$, there exists an integer l and l points $(x_n,\dots,x_{2n-1}),\dots,(x_{ln},\dots,x_{(l+1)n-1})$ in K^n such that

$$W^{(1)}_{(x_0,\dots,x_{n-1}),(x_n,\dots,x_{2n-1})}\dots W^{(1)}_{(x_{ln},\dots,x_{(l+1)n-1}),(y_0,\dots,y_{n-1})}>0.$$

This inequality means

$$\begin{aligned}(7)\qquad &\mu(x_n,\dots,x_{2n-1}\,|\,x_{n-1})\mu(x_{2n},\dots,x_{3n-1}\,|\,x_{2n-1})\dots\\ &\quad\dots\mu(x_{ln},\dots,x_{(l+1)n-1}\,|\,x_{ln-1})\mu(y_0,\dots,y_{n-1}\,|\,x_{(l+1)n-1})>0.\end{aligned}$$

Now for every $j=1,\dots,l$ we have

$$\begin{aligned}(8)\qquad \sum_{x_{jn},\dots,x_{(j+1)n-2}}&\mu(x_{jn},\dots,x_{(j+1)n-1}\,|\,x_{jn-1})\\ &=\mu(\omega_n=x_{(j+1)n-1}\,|\,\omega_0=x_{jn-1})\\ &=(A_n(\mu))_{x_{jn-1},x_{(j+1)n-1}}.\end{aligned}$$

Suppose that $A_n(\mu)$ is irreducible. Then there is an integer l such that $[(A_n(\mu))^{l+1}]_{x_{n-1},y_{n-1}}>0$, that is, there are $x_{2n-1},\dots,x_{(l+1)n-1}$ such that

$$(A_n(\mu))_{x_{n-1},x_{2n-1}}(A_n(\mu))_{x_{2n-1},x_{3n-1}}\dots(A_n(\mu))_{x_{(l+1)n-1},y_{n-1}}>0$$

and in view of (7) and (8) we see that $W^{(1)}$ is irreducible.

Conversely, suppose that $W^{(1)}$ is irreducible. Then, as $\mu(a)\mu(b)>0$ for any a,b in K, there exist $x_0,\dots,x_{n-2},y_0,\dots,y_{n-2}$ in K such that $\mu(x_0,\dots,x_{n-2},a)\mu(y_0,\dots,y_{n-2},b)>0$, so, by the irreducibility of $W^{(1)}$ there exists an integer l and l points in K^n, $(x_{jn},\dots,x_{(j+1)n-1})$, $j=1,\dots,l$, such that (7) is satisfied with $x_{n-1}=a$ and $y_{n-1}=b$. But in view of (8) this implies

$$\begin{aligned}&(A_n(\mu))_{a,x_{2n-1}}(A_n(\mu))_{x_{2n-1},x_{3n-1}}\dots\\ &\qquad\dots(A_n(\mu))_{x_{ln-1},x_{(l+1)n-1}}(A_n(\mu))_{x_{(l+1)n-1},b}>0.\end{aligned}$$

Therefore $[(A_n(\mu))^l]_{a,b}>0$ and the proof of (i) is complete.

The proof of (ii) is similar and will be omitted. ∎

Another proof. We have shown above that $\phi_n\mu$ is a function of a Markov chain on the extended state space L. It can be seen, by calculations which we omit, that this chain is irreducible if A_n is irreducible. But a function of an ergodic process is ergodic, therefore, this gives a third proof of the ergodicity of $\phi_n\mu$.

3. Decay of correlations, spectral and mixing properties. Let $(X,m,T=S^n)$ be a dynamical system and (X^f,m^f,T^f) its integral automorphism defined in the preceding section, where $f\equiv n$. We denote (X^f,m^f,T^f) by $(\widetilde{X},\widetilde{m},\widetilde{T})$. We shall relate the spectrum of $(\widetilde{X},\widetilde{m},\widetilde{T})$ to the spectrum of (X, m, T). Denote by U_T the unitary operator on L^2_m induced by $T : U_T f = f \circ T$. We define $U_{\widetilde{T}}$ similarly.

It can be checked that the characteristic functions $1_{\Omega\times\{i\}}$, which we denote by e_i, are translated by $U_{\widetilde{T}}$ according to the relations

$$\begin{aligned} &U_{\widetilde{T}} e_i = e_{i-1} \quad \text{for } 2 \le i \le n, \\ &U_{\widetilde{T}} e_1 = e_n. \end{aligned}$$

We denote by V the subspace of $L^2_{\widetilde{m}}$ generated by $\{e_i\}$.

THEOREM 3. *If U_T has countable Lebesgue spectrum in $\{1\}^\perp$, then $U_{\widetilde{T}}$ has countable Lebesgue spectrum in the orthocomplement of V and a discrete spectrum on V concentrated on the n roots of unity. The maximal spectral type of $\widetilde{T}$ is $\lambda + \sum_{k=0}^{n-1} \delta(2k\pi/n)$ where λ is the Lebesgue measure.*

Proof. For every $i = 1, \ldots, n$ denote by H_i the closed linear subspace of all $f \in L^2_{\widetilde{m}}$ such that $f(x, j) = 0$ for m-almost all x and all $j \neq i$. Clearly the H_i are mutually orthogonal and $L^2_{\widetilde{m}} = \bigoplus_i H_i$. Let $\{e_{p,q} : p \in \mathbb{Z},\ q \in \mathbb{N}\}$ be an orthonormal basis of $L^2_m \ominus 1$ such that $U_T e_{p,q} = e_{p+1,q}$ for all p and q. For every $i = 1, \ldots, n$ and all $(p, q) \in \mathbb{Z} \times \mathbb{N}$ define $e^i_{p,q} \in L^2_{\widetilde{m}}$ by $e^i_{p,q}(x, j) = e_i(x, j) e_{p,q}(x)$ a.e. Then $\{\sqrt{n}\, e^i_{p,q} : p \in \mathbb{Z},\ q \in \mathbb{N}\}$ is an orthonormal basis of $H_i \ominus e_i$. Furthermore, $U_{\widetilde{T}} e^i_{p,q} = e^{i-1}_{p,q}$ for $2 \le i \le n$ and $U_{\widetilde{T}} e^1_{p,q} = e^n_{p+1,q}$. In fact, let $(x, j) \in \widetilde{X}$; then for $j < n$ and $i \neq 1$,

$$\begin{aligned} U_{\widetilde{T}} e^i_{p,q}(x, j) &= e^i_{p,q}(\widetilde{T}(x, j)) = e^i_{p,q}(x, j+1) = e_i(x, j+1) e_{p,q}(x) \\ &= e_{i-1}(x, j) e_{p,q}(x) = e^{i-1}_{p,q}(x, j), \end{aligned}$$

and for $j = n$ we have

$$\begin{aligned} U_{\widetilde{T}} e^i_{p,q}(x, n) &= e^i_{p,q}(\widetilde{T}(x, n)) = e^i_{p,q}(Tx, 1) = e_i(Tx, 1) e_{p,q}(Tx) = 0 \\ &= e_{i-1}(x, n) e_{p,q}(x) = e^{i-1}_{p,q}(x, n). \end{aligned}$$

Now for $i = 1$ and $j + 1 \le n$,

$$\begin{aligned} U_{\widetilde{T}} e^1_{p,q}(x, j) &= e^1_{p,q}(\widetilde{T}(x, j)) = e^1_{p,q}(x, j+1) = e_1(x, j+1) e_{p,q}(x) = 0 \\ &= e_n(x, j) e_{p+1,q}(x) = e^n_{p+1,q}(x, j) \end{aligned}$$

and, finally,

$$\begin{aligned} U_{\widetilde{T}} e^1_{p,q}(x, n) &= e^1_{p,q}(\widetilde{T}(x, n)) = e^1_{p,q}(Tx, 1) \\ &= e_1(Tx, 1) e_{p,q}(Tx) = e_n(x, n) e_{p+1,q}(x) = e^n_{p+1,q}(x, n). \end{aligned}$$

Thus, we have shown that $U_{\widetilde{T}}$ has a countable Lebesgue spectrum in the orthocomplement of V.

Now, $U_{\widetilde{T}}$, acting on V, is represented by the matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \multicolumn{5}{c}{\dots\dots\dots\dots\dots} \\ 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \end{bmatrix}$$

which has the n roots of the unity as eigenvalues. Thus, the spectrum of $U_{\widetilde{T}}$ on V is simple and concentrated on the n roots of the unity. Then the maximal spectral type of $U_{\widetilde{T}}$ is $\lambda + \sum_{k=0}^{n-1} \delta(2k\pi/n)$, where $\delta(a)$ is the Dirac measure concentrated on $\{a\}$ and λ is the Lebesgue measure. ■

The spectrum of $U_{\widetilde{T}}$ in $V^{\perp}$ can be analyzed in terms of the spectrum of $U_T|\{1\}^{\perp}$, when T has an arbitrary spectrum.

THEOREM 4. *Let $h \in L^2_m$ and let $h_i \in L^2_{\widetilde{m}}$ be defined by $h_i(x,j) = h(x)e_i(x,j)$. Then*:

(i) σ_{h_i} *is singular* (*resp. absolutely continuous*, *with respect to the Lebesgue measure*) *if and only if σ_h is singular* (*resp. absolutely continuous*, *with respect to the Lebesgue measure*).

(ii) σ_{h_i} *is discrete if and only if σ_h is discrete.*

Proof. For $t \in \mathbb{Z}$ and $r \in \{0, 1, \dots, n-1\}$ we have

$$\widehat{\sigma}_{h_i}(nt + r) = \langle h_i, U_{\widetilde{T}}^{(nt+r)} h_i \rangle = \begin{cases} 0 & \text{if } r \neq 0, \\ n^{-1}\widehat{\sigma}_h(t) & \text{if } r = 0. \end{cases}$$

Now we use the following lemmas to complete the proof of the theorem:

LEMMA 4. *Let σ_1 and σ_2 be two measures on the circle S^1 whose Fourier coefficients satisfy $\widehat{\sigma}_1(tn) = \widehat{\sigma}_2(t)$ for any $t \in \mathbb{Z}$, where n is a fixed positive integer. Then $\sigma_2 = \sigma_1\tau^{-1}$ where τ is the mapping $x \mapsto \tau x = nx \pmod{2\pi}$ from S^1 onto itself.*

LEMMA 5. *Let σ_1 and σ_2 be two measures such that $\sigma_2 = \sigma_1\tau^{-1}$ where $\tau x = nx \pmod 1$. Then*

(i) σ_1 *is singular iff σ_2 is singular to the Lebesgue measure λ.*

(ii) σ_1 *is absolutely continuous* (*with respect to the Lebesgue measure λ*) *iff σ_2 is a.c. with respect to λ.*

(iii) σ_1 *is discrete iff σ_2 is discrete.*

Now, we turn to the dynamical system $(X, \mathcal{F}, \nu, S)$, where $\nu = n^{-1}\sum_{i=0}^{n-1} m \circ S^{-i}$. We shall investigate mixing subsets of this system.

DEFINITION 3. Two subsets A, B are said to be *mutually mixing* (or, simply, *mixing*) for $(X, \mathcal{F}, \nu, S)$ if $\nu(A \cap S^{-n}B) \to \nu(A)\nu(B)$ as $|n| \to \infty$. A set A is called a *mixing set* if it is mixing with itself.

THEOREM 5. *With the above notations, let* $(X,\mathcal{F},m,S^n)$ *be a dynamical system with absolutely continuous spectrum. Consider* $(X,\mathcal{F},\nu,S)$ *where* $\nu = n^{-1}\sum_{i=0}^{n-1} m\circ S^{-i}$.

(i) *The spectral measure of* $f\in L^2_\nu$ *is absolutely continuous with respect to the Lebesgue measure if and only if*

$$\int f(S^i x)\,dm(x)=0,\qquad \forall i.$$

(ii) *A subset* $A\in\mathcal{F}$ *is mixing for* $(X,\mathcal{F},\nu,S)$ *if and only if*

$$m(S^{-i}A)=m(A),\qquad \forall i=1,\ldots,n.$$

(iii) *If a subset* A *is mixing then* A *mixes with every set* B.

Proof. (i) Recall that (X,ν,S) is a factor of $(\widetilde{X},\widetilde{m},\widetilde{T})$ where $T=S^n$, the factor map being $(x,i)\mapsto\varphi(x,i)=S^ix$. Now the spectral measure of $f\in L^2_\nu$ (relative to the unitary operator U_S) is also the spectral measure of $f\circ\varphi$ relative to the unitary operator $U_{\widetilde{T}}$. By the theorem above the spectral measure of $f\circ\varphi$ is absolutely continuous with respect to the Lebesgue measure if and only if $f\circ\varphi$ is orthogonal to the e_i's, that is, $\langle f\circ\varphi,e_i\rangle=0$. But

$$\langle f\circ\varphi,e_i\rangle=\frac{1}{n}\int f\circ\varphi(x,i)\,dm(x)=\frac{1}{n}\int f(S^ix)\,dm(x).$$

(ii) The subset A is mixing if the Fourier coefficients of the spectral measure of $f:=1_A-\nu(A)$ go to zero at infinity.

By the above theorem the Fourier coefficients of σ_f tend to zero at infinity if and only if σ_f is absolutely continuous with respect to Lebesgue measure and by (i) this is equivalent to $\int f(S^ix)\,dm(x)=0$. That is to say, $\int 1_A(S^ix)\,dm(x)-\nu(A)=0$ for all i, which means $m(S^{-i}A)=\nu(A)$ for all i. This is equivalent to $m(S^{-i}A)=m(A)$ for all i.

(iii) This is a general result for any dynamical system. ∎

The next theorem shows the existence of mixing subsets in (X,ν,S).

THEOREM 6. *With the notations of the above theorem, for any* $t\in[0,1]$ *there exists a set* $A\in\mathcal{F}$ *such that* $m(S^{-i}A)=t$ *for all* i.

Proof. As m is weakly mixing for S^n it is non-atomic. The same holds for $m\circ S^{-i}$ for all $i=1,\ldots,n-1$. Consider the $\mathbb{R}^n$-valued measure F defined on $(\Omega,\mathcal{F})$ by $F(A)=(m(A),m(S^{-1}A),\ldots,m(S^{-(n-1)}A))$. Then F is non-atomic. Therefore, according to the Lyapunov convexity theorem [11], F has a convex compact range in $\mathbb{R}^n$. As $F(\Phi)=0$ and $F(X)=(1,1,\ldots,1)$, $tF(X)+(1-t)F(\Phi)=(t,t,\ldots,t)$ is in the range of F for every $t\in[0,1]$; the theorem is proved. ∎

THEOREM 7. *Let* $n\ge 2$ *be an integer. Let* S *be a measurable transformation on the probability measure space* $(X,\mathcal{F},m)$ *such that* S^n *preserves* m.

We suppose that $(X, \mathcal{F}, m, S^n)$ has an absolutely continuous spectrum and that there exist $C > 0$ and $\lambda \in [0,1[$ such that $|m(A \cap S^{-nt}B) - m(A)m(B)| \le C\lambda^t$ for all A, B in $\mathcal{F}$. Let $\nu = n^{-1}\sum_{i=0}^{n-1} m \circ S^{-i}$. Then if A and B are two mixing sets in $(X, \mathcal{F}, \nu, S)$ we have $|\nu(A \cap S^{-l}B) - \nu(A)\nu(B)| \le C\lambda^t$, where t is defined by $l = nt + r$ with $r = 0, 1, \ldots, n-1$.

Proof. Let A, B be two mixing sets in $(X, \mathcal{F}, \nu, S)$. Then $m(S^{-i}A) = \nu(A)$ for all i and the same holds for B. Now for $l = nt + r$ with $r = 0, 1, \ldots, n-1$ we have

$$\begin{aligned}
&|\nu(A \cap S^{-nt-r}B) - \nu(A)\nu(B)| \\
&\quad = \left| \frac{1}{n}\sum_{i=0}^{n-1} m(S^{-i}A \cap S^{-nt}S^{-i-r}B) - m(A)m(B) \right| \\
&\quad \le \frac{1}{n}\sum_{i=0}^{n-1} |m(S^{-i}A \cap S^{-nt}(S^{-r-i}B)) - m(A)m(B)| \\
&\quad = \frac{1}{n}\sum_{i=0}^{n-1} |m(S^{-i}A \cap S^{-nt}S^{-r-i}B) - m(S^{-i}A)m(S^{-r-i}B)| \\
&\quad \le \frac{1}{n}\sum_{i=0}^{n-1} C\lambda^t = C\lambda^t. \ \blacksquare
\end{aligned}$$

COROLLARY. *Let $n \ge 2$ and let $\mu \in M_n$ be such that $A_n(\mu)$ irreducible and aperiodic. If two sets A, B are mixing for the dynamical system $(\Omega, \mathcal{B}(\Omega), \phi_n\mu, \sigma)$ then they mix with an exponential rate.*

Proof. The hypothesis $A_n(\mu)$ irreducible and aperiodic implies that $(\Omega, \nu_0(\mu), \sigma^n)$ is a mixing Markov chain. It is well known that a mixing Markov chain has a countable Lebesgue spectrum and that it is exponentially mixing with a uniform rate. Now the result follows from the theorem above. $\blacksquare$

Theorems 5 and 6 give a description of the family of mixing subsets in $(\Omega, \mathcal{B}(\Omega), \phi_n\mu, \sigma)$. A subset $A \in \mathcal{B}(\Omega)$ is mixing if and only if

$$\nu_0(A) = \nu_0(\sigma^{-1}A) = \ldots = \nu_0(\sigma^{-n+1}A).$$

Therefore, mixing cylindric subsets can be easily characterized in terms of the measure μ. In particular, *any cylinder based on two successive coordinates is mixing.* Moreover, it follows by direct calculations from the above equalities that if any cylindric set of length $n+1$ (equal to the dimension of μ) is mixing then μ is Markovian. Therefore, for non-Markovian μ, there is at least one cylindric set of length $n+1$ which is not mixing.

Theorem 6 says that there is a continuum of mixing subsets for $\phi_n\mu$. Nevertheless, we shall show the following proposition:

PROPOSITION. *If a dynamical system* $(\Omega, \mathcal{A}, \mu, T)$ *is non-mixing then in the space of measurable subsets endowed with the Nikodym metric* $d(A, B) = \mu(A \triangle B)$, *the family of all non-mixing subsets is open and dense* (*i.e.*, *it is generic*).

Proof. Put $f = 1_A - \mu(A)$, and denote the Fourier coefficients of the spectral measure σ_f by $\widehat{\sigma}_f(n) = \langle U_T^n f, f\rangle$. Then

$$\widehat{\sigma}_f(n) = \mu(A \cap T^{-n}A) - \mu(A)^2.$$

Therefore, by the Wiener lemma, σ_f is continuous if A is mixing. Now, *the convergence of* A_n *to* A *with respect to the Nikodym metric is equivalent to the* L^2*-convergence of the corresponding* f_n *to* f. As the subspace H_c of continuous spectrum of U is closed in L^2_μ, this implies that the family of mixing subsets of T is closed. Now, as the system is non-mixing, for any $\varepsilon > 0$, there is a non-mixing measurable subset A_ε of measure $\mu(A_\varepsilon) < \varepsilon$. Let A be a mixing subset, and write $A^0_\varepsilon = A_\varepsilon \cap A$ and $A^1_\varepsilon = A^{\mathrm{c}} \cap A_\varepsilon$. Either A^0_ε or A^1_ε is non-mixing. In the first case, the subset $A \setminus A^0_\varepsilon$ is non-mixing and in the second case $A \cup A^1_\varepsilon$ is non-mixing. In both cases, we obtain non-mixing subsets as close to the mixing subset A as we want. This shows that the family of non-mixing subsets is dense. ■

Therefore, although according to Sinai's theorem, the above system having positive entropy has a Bernoullian factor, this factor corresponds to a rare family of subsets, since it is contained in the family of mixing subsets.

The Pinsker σ*-algebra of* $\phi_n\mu$. Recall that the *Pinsker* σ*-algebra* of a dynamical system $(X, \mathcal{F}, \mu, T)$ is the smallest σ-algebra $\pi(T)$ that contains all finite partitions $\mathcal{P}$ such that

$$h(T, \mathcal{P}) = 0.$$

The Pinsker σ-algebra is invariant under T.

It follows that a set $A \in \mathcal{F}$ is in $\pi(T)$ if and only if $h(T, \{A, A^{\mathrm{c}}\}) = 0$.

Now, we describe the Pinsker σ-algebra of $(\Omega, \phi_n\mu, \sigma)$ when $(\Omega, \sigma^n, \nu_0)$ is ergodic. In this case, the measures $\nu_0 \circ \sigma^{-i}$ are mutually singular. Let $B_0, B_1, \ldots, B_{n-1}$ be measurable subsets such that $\nu_0(\sigma^{-i}B_j) = \delta_{ij}$.

THEOREM 8. *Under the above hypothesis*, *set* $\nu = \phi_n\mu$. *Then*:

(i) *A set* B *is in the Pinsker* σ*-algebra of* (Ω, ν, σ) *if and only if it is periodic*, *that is*, *there exists* $m > 0$ *such that* $\sigma^{-m}B = B$ ν*-a.e.*

(ii) *If* $(\Omega, \nu_0, \sigma^n)$ *has absolutely continuous spectrum*, *then the Pinsker* σ*-algebra of* $(\Omega, \phi_n\mu, \sigma)$ *is generated by* $B_0, \ldots, B_{n-1}$.

Proof. (i) As explained in Section 2, $(K^{\mathbb{Z}}, \nu, \sigma)$ is a factor of the (p, P) Markov chain $(L^{\mathbb{Z}}, \nu_1, \sigma)$. Denote by φ the factor map and by $\mathcal{P}$ [resp. $\mathcal{P}_1$] the Pinsker σ-algebra of $(K^{\mathbb{Z}}, \nu, \sigma)$ [resp. $(L^{\mathbb{Z}}, \nu_1, \sigma)$]. It is known that the Pinsker σ-algebra of a Markov chain is the largest invariant σ-algebra contained in the time zero σ-algebra of the chain, so $\mathcal{P}_1$ is finite. It follows that for every A in $\mathcal{P}_1$, $\sigma^{-m}A = A$ for some integer $m > 0$, that is, the elements of $\mathcal{P}_1$ are periodic. On the other hand it is clear that every periodic set is in $\mathcal{P}_1$ so $\mathcal{P}_1$ is the σ-algebra of periodic sets with respect to σ (in fact since $\mathcal{P}_1$ is finite we can take the same period). It follows that a set B is in $\mathcal{P}$ if and only if there is $m > 0$ such that $\sigma^{-m}\varphi^{-1}B = \varphi^{-1}B$ ν_1-a.e.; but this equality is equivalent to $\sigma^{-m}B = B$ ν-a.e. so (i) is proved.

(ii) Let $B \in \mathcal{P}$. By (i) there is $m > 0$ such that $\sigma^{-m}B = B$ ν-a.e., and for all $t \in \mathbb{Z}$ and $r = 0, 1, \ldots, m-1$ we have

$$\langle 1_B, U_\sigma^{mt+r} 1_B \rangle = \nu(B \cap \sigma^{-mt-r}B) = \nu(B \cap \sigma^{-r}B).$$

This proves that 1_B has discrete spectral measure. Therefore, $1_B \circ \varphi$ is in V and there exist scalars $\alpha_1, \ldots, \alpha_n$ such that $1_B \circ \varphi = \sum_j \alpha_j e_j$; hence for every i and $x, 1_B(\sigma^i x) = \alpha_i$, that is, $1_{\sigma^{-i}B}(x) = \alpha_i$.

This implies each α_i is 0 or 1 and $\nu_0(\sigma^{-i}B)$ is 0 or 1. Let $I = I(B) := \{i : \nu_0(\sigma^{-i}B) = 1\}$. We shall show that $B = \bigcup_{i \in I} B_i$ ν-a.e. In fact, if $j \in I$ we have

$$\begin{aligned}
\nu_0 \circ \sigma^{-j}\Big(B \triangle \bigcup_{i\in I} B_i\Big) &\le \nu_0 \circ \sigma^{-j}(B \setminus B_j) + \nu_0 \circ \sigma^{-j}\Big(\bigcup_{i\in I} B_i \setminus B\Big) \\
&\le \nu_0(\sigma^{-j}B \setminus \sigma^{-j}B_j) + \sum_{i\in I} \nu_0(\sigma^{-j}B_i \setminus \sigma^{-j}B) \\
&= \nu_0(\sigma^{-j}B \setminus \sigma^{-j}B_j) + \nu_0(\sigma^{-j}B_j \setminus \sigma^{-j}B) \\
&= \nu_0(\sigma^{-j}B \triangle \sigma^{-j}B_j) = 0,
\end{aligned}$$

and if $j \notin I$ we have

$$\begin{aligned}
\nu_0 \circ \sigma^{-j}\Big(B \triangle \bigcup_{i\in I} B_i\Big) &\le \nu_0(\sigma^{-j}B) + \nu_0\Big(\sigma^{-j}\bigcup_{i\in I} B_i\Big) \\
&\le \nu_0(\sigma^{-j}B) + \sum_{i\in I} \nu_0(\sigma^{-j}B_i) = 0.
\end{aligned}$$

On the other hand B_i is in $\mathcal{P}$ for every $i = 0, \ldots, n-1$. In fact, in virtue of (i) we have to show that there exists $m > 0$ such that $\sigma^{-m}B_i = B_i$ ν-a.e. or equivalently $\sigma^{-m}B_i = B_i$ $\nu_0 \circ \sigma^{-j}$-a.e., for every $j = 0, 1, \ldots, n-1$. But since ν_0 is σ^n-invariant we have

$$\nu_0(\sigma^{-j}(\sigma^{-n}B_i)) = \nu_0(\sigma^{-j}B_i) = \delta_{ij},$$

which proves $\sigma^{-n}B_i = B_i$ $\nu_0 \circ \sigma^{-j}$-a.e. and so B_i is in $\mathcal{P}$. ∎

REMARKS. (i) If ν gives positive measure to all cylinders then the Pinsker σ-algebra of (Ω, ν, σ) contains no finite union B of cylinders, for otherwise we must have $\nu(\sigma^{-m} B \cap B^{\mathrm{c}}) = 0$ for some $m > 0$, but B^{c} contains cylinders and therefore $\sigma^{-m} B \cap B^{\mathrm{c}}$ contains cylinders.

(ii) If ν_0 gives positive measure to all cylinders then so does ν. It follows that $h(\sigma, \{A, A^{\mathrm{c}}\}) > 0$ for any cylinder A.

4. The class of p-dimensional marginals. Given a stochastic process on a finite state space, is there another one, having identical p-dimensional marginals? The family $\phi_n \mu$, $\mu \in M_n$, gives an answer to this problem. Lemma 2 states that ϕ_n is one-to-one and Theorem 9 (below) states that it is marginals preserving. This allows us to reduce the construction of a distinct process to the class M_n. In [7] and [10], the last problem is solved for 2-marginals and p-marginals respectively for the ergodic Markov chains, and more generally, for the $\phi_n \mu$ processes.

DEFINITION 4. (i) We define the *family of p-marginals* of $\mu \in M(\Omega, \sigma)$ to be the family of joint distributions $\mu(\omega_{n_1}, \ldots, \omega_{n_p})$ for any p-uple $(n_1, \ldots, n_p)$ of integers. We say a measure $\lambda \in M(\Omega, \sigma)$ *has the same p-marginals as* μ if $\mu(\omega_{n_1}, \ldots, \omega_{n_p}) = \lambda(\omega_{n_1}, \ldots, \omega_{n_p})$ for any p-uple as above.

(ii) We say that a measure μ in $M(\Omega, \sigma)$ has the *Chapman–Kolmogorov property* (for short we write "μ is C.K.") if the family $(A_r(\mu))_{r \in \mathbb{N}}$ forms a semigroup of matrices, that is, $A_r(\mu) = (A_1(\mu))^r$ for every $r \in \mathbb{N}$, where

$$(A_r(\mu))_{i,j} := \mu(\omega_r = j \mid \omega_0 = i).$$

We start by proving a lemma which gives an expression of any two-dimensional marginals of $\phi_n \mu$ in terms of finitely many two-dimensional marginals of μ in M_n.

LEMMA 6. (i) *We have* $A_1(\phi_n \mu) = A_1(\mu)$ *and for* $r = 2, \ldots, n$,

$$A_r(\phi_n \mu) = \frac{n - r + 1}{n} A_r(\mu) + \frac{1}{n} \sum_{i=1}^{r-1} A_i(\mu) A_{r-i}(\mu).$$

(ii) *For* $l \in \mathbb{N}^*$ *and* $r \in \{2, \ldots, n-1\}$ *we have*

$$\begin{aligned} A_{ln+r}(\phi_n \mu) = \frac{1}{n} \Big[& \sum_{i=0}^{n-r} A_{n-i}(\mu) A_n^{l-1}(\mu) A_{r+i}(\mu) \\ & + \sum_{i=n-r+1}^{n-1} A_{n-i}(\mu) A_n^{l}(\mu) A_{r+i-n}(\mu) \Big], \end{aligned}$$

and for $r = 0, 1$,

$$A_{ln+r}(\phi_n\mu) = \frac{1}{n}\sum_{i=0}^{n-1} A_{n-i}(\mu)A_n^{l-1}(\mu)A_{r+i}(\mu).$$

Proof. Let $x, y \in K$, and denote $A_r(\mu)$ by A_r.

(i) Let $i \in \{0, \ldots, n-1\}$. Then

$$\begin{aligned}\sigma^i\nu_0(\omega_0 = x,\ \omega_r = y) &= \nu_0(\omega_i = x,\ \omega_{r+i} = y)\\ &= \sum_z \nu_0(\omega_0 = z_0, \ldots, \omega_{i-1} = z_{i-1},\ \omega_i = x,\\ &\qquad \omega_{i+1} = z_{i+1}, \ldots, \omega_{r+i-1} = z_{r+i-1},\ \omega_{r+i} = y).\end{aligned}$$

We consider two cases: $r + i \le n$ and $r + i > n$. In the first case we obtain

$$\begin{aligned}\sigma^i\nu_0(\omega_0 = x,\ \omega_r = y) &= \sum_z \mu(z_0, \ldots, z_{i-1}, x, z_{i+1}, \ldots, z_n)\\ &= \mu(\omega_i = x,\ \omega_{r+i} = y) = \mu(\omega_0 = x,\ \omega_r = y)\\ &= \mu(x)(A_r)_{x,y}\end{aligned}$$

and in the second case we can write

$$\begin{aligned}&\sigma^i\nu_0(\omega_0 = x,\ \omega_r = y)\\ &\quad= \sum_z \mu(z_0, \ldots, z_{i-1}, x, z_{i+1}, \ldots, z_{r+i-1}, y)\mu(z_{n+1}, \ldots, z_{r+i-1}, y \,|\, z_n)\\ &\quad= \sum_z \mu(\omega_i = x,\ \omega_n = z_n)\mu(\omega_{r+i-n} = y \,|\, \omega_0 = z_n)\\ &\quad= \sum_z \mu(x)\mu(\omega_{n-i} = z_n \,|\, \omega_0 = x)(A_{r+i-n})_{z_n,y}\\ &\quad= \sum_z \mu(x)(A_{n-i})_{x,z_n}(A_{r+i-n})_{z_n,y} = \mu(x)(A_{n-i}A_{r+i-n})_{x,y}.\end{aligned}$$

Hence

$$\begin{aligned}\phi_n\mu(\omega_0 = x,\ \omega_r = y) &= \frac{1}{n}\sum_{i=0}^{n-r} \mu(x)(A_r)_{x,y}\\ &\qquad + \frac{1}{n}\sum_{i=n-r+1}^{n-1} \mu(x)(A_{n-i}A_{r+i-n})_{x,y}.\end{aligned}$$

So if we put $n - i = j$ we obtain $n - r + 1 \le i \le n - 1 \Leftrightarrow 1 \le j \le r - 1$; therefore

$$\phi_n\mu(\omega_0 = x,\ \omega_r = y) = \mu(x)\left[\frac{n-r+1}{n}(A_r)_{x,y} + \frac{1}{n}\sum_{j=1}^{r-1}(A_jA_{r-j})_{x,y}\right].$$

(ii) We have

$$\sigma^i\nu_0(\omega_0 = x,\ \omega_{ln+r} = y) = \nu_0(\omega_i = x,\ \omega_{ln+r+i} = y)$$
$$= \sum_z \nu_0(\omega_0 = z_0, \ldots, \omega_{i-1} = z_{i-1},\ \omega_i = x,\ \omega_{i+1} = z_{i+1}, \ldots,$$
$$\omega_{ln+r+i-1} = z_{ln+r+i-1},\ \omega_{ln+r+i} = y).$$

Then if $r + i \leq n$ we can write

$$\begin{aligned}
\sigma^i\nu_0(\omega_0 = x,\ \omega_{ln+r} = y) &= \sum_z \mu(z_0, \ldots, z_{i-1}, x, z_{i+1}, \ldots, z_n) \\
&\quad \times \mu(z_{n+1}, \ldots, z_{2n} \,|\, z_n) \ldots \mu(z_{(l-1)n+1}, \ldots, z_{ln} \,|\, z_{(l-1)n}) \\
&\quad \times \mu(z_{ln+1}, \ldots, z_{ln+r+i-1}, y \,|\, z_{ln}) \\
&= \sum_z \mu(\omega_i = x,\ \omega_n = z_n)\mu(\omega_n = z_{2n} \,|\, \omega_0 = z_n) \ldots \\
&\quad \ldots \mu(\omega_n = z_{ln} \,|\, \omega_0 = z_{(l-1)n})\mu(\omega_{r+i} = y \,|\, \omega_0 = z_{ln}) \\
&= \mu(x) \sum_z (A_{n-i})_{x,z_n}(A_n)_{z_n,z_{2n}} \ldots (A_n)_{z_{(l-1)n},z_{ln}}(A_{r+i})_{z_{ln},y} \\
&= \mu(x)(A_{n-i}A_n^{l-1}A_{r+i})_{x,y},
\end{aligned}$$

and if $r + i > n$ we have

$$\begin{aligned}
&\sigma^i\nu_0(\omega_0 = x,\ \omega_{ln+r} = y) \\
&= \sum_z \mu(z_0, \ldots, z_{i-1}, x, z_{i+1}, \ldots, z_n) \\
&\quad \times \mu(z_{n+1}, \ldots, z_{2n} \,|\, z_n) \ldots \mu(z_{ln+1}, \ldots, z_{(l+1)n} \,|\, z_{ln}) \\
&\quad \times \mu(z_{(l+1)n+1}, \ldots, z_{ln+r+i-1}, y \,|\, z_{(l+1)n}) \\
&= \sum_z \mu(\omega_i = x,\ \omega_n = z_n)\mu(\omega_n = z_{2n} \,|\, \omega_0 = z_n) \ldots \\
&\quad \ldots \mu(\omega_n = z_{(l+1)n} \,|\, \omega_0 = z_{ln})\mu(\omega_{ln+r+i-(l+1)n} = y \,|\, \omega_0 = z_{(l+1)n}) \\
&= \mu(x) \sum_z (A_{n-i})_{x,z_n}(A_n)_{z_n,z_{2n}} \ldots (A_n)_{z_{ln},z_{(l+1)n}}(A_{r+i-n})_{z_{(l+1)n},y} \\
&= \mu(x)(A_{n-i}(A_n)^l A_{r+i-n})_{x,y}
\end{aligned}$$

so

$$\begin{aligned}&\phi_n\mu(\omega_0 = x,\ \omega_{ln+r} = y)\\&\quad= \mu(x)\frac{1}{n}\Big[\sum_{i=0}^{n-r}(A_{n-i}A_n^{l-1}A_{r+i})_{x,y} + \sum_{i=n-r+1}^{n-1}(A_{n-i}A_n^{l}A_{r+i-n})_{x,y}\Big].\end{aligned}$$

It is evident that for $r = 0, 1$ we have $r + i \le n$ so the second term on the right side does not exist; so for $r = 0, 1$,

$$\phi_n\mu(\omega_0 = x,\ \omega_{ln+r} = y) = \frac{1}{n}\sum_{i=0}^{n-1}(A_{n-i}A_n^{l-1}A_{r+i})_{x,y}. \ \blacksquare$$

COROLLARY. *For $\mu \in M(\Omega, \sigma)$, the following are equivalent*:

(i) $A_r(\mu) = (A_1(\mu))^r$ *for* $r = 1, \ldots, n$.
(ii) $A_r(\phi_n\mu) = (A_1(\phi_n\mu))^r$ *for* $r = 1, \ldots, n$.
(iii) $A_r(\phi_n\mu) = (A_1(\phi_n\mu))^r$ *for* $r \in \mathbb{N}$.

Proof. The implication (iii)⇒(ii) is trivial. The implication (i)⇒(iii) follows from Lemma 6. We prove that (ii) implies (i) by induction on r. For $r = 1$, $A_r(\mu) = (A_1(\mu))^r$. Let $r \le n - 1$ and suppose $A_s(\mu) = (A_1(\mu))^s$ for $s \le r$. As $A_1(\phi_n\mu) = A_1(\mu)$, the property (ii) and Lemma 6 imply

$$(A_1(\mu))^{r+1} = \frac{n-r}{n}A_{r+1}(\mu) + \frac{1}{n}\sum_{i=1}^{r}A_i(\mu)A_{r+1-i}(\mu).$$

Now $1 \le i \le r \Leftrightarrow 1 \le r + 1 - i \le r$; therefore for such an i,

$$A_i(\mu)A_{r+1-i}(\mu) = (A_1(\mu))^i(A_1(\mu))^{r+1-i} = (A_1(\mu))^{r+1}$$

so we obtain the equality

$$(A_1(\mu))^{r+1} = \frac{n-r}{n}A_{r+1}(\mu) + \frac{r}{n}(A_1(\mu))^{r+1},$$

which is equivalent to $A_{r+1}(\mu) = (A_1(\mu))^{r+1}$. $\blacksquare$

THEOREM 9. (i) *For every n, two measures $\lambda, \mu \in M_n$ have the same p-marginals if and only if $\phi_n\lambda$ and $\phi_n\mu$ have the same p-marginals.*

(ii) *Let $\lambda, \mu \in M(\Omega, \sigma)$. If $\phi_n\lambda$ and $\phi_n\mu$ have the same p-marginals for infinitely many n, then λ and μ have the same p-marginals.*

Proof. (i) Suppose $\lambda, \mu \in M_n$ have the same p-marginals. It follows from the definitions (see Lemma 6 for the case $n = 2$) that $\nu_0(\lambda)$ and $\nu_0(\mu)$ have the same p-marginals. The same is true for $\sigma^j\nu_0(\mu)$ and $\sigma^j\nu_0(\lambda)$ for all j, and thus for their respective arithmetic means. For (ii) we give the proof for the case $p = 2$. Let $s \in \mathbb{N}$; so there is $n \ge s$ such that $\phi_n\lambda$ and $\phi_n\mu$ have the same 2-marginals; hence for every r, $A_r(\phi_n\lambda) = A_r(\phi_n\mu)$; therefore it follows by Lemma 6 that $A_1(\lambda) = A_1(\mu)$, and for every $r = 2, \ldots, n$,

$$\frac{n-r+1}{n} A_r(\lambda) + \frac{1}{n} \sum_{i=1}^{r-1} A_i(\lambda) A_{r-i}(\lambda) = \frac{n-r+1}{n} A_r(\mu) + \frac{1}{n} \sum_{i=1}^{r-1} A_i(\mu) A_{r-i}(\mu),$$

i.e.

$$(n-r+1)(A_r(\lambda) - A_r(\mu)) = \sum_{i=1}^{r-1} [A_i(\mu)A_{r-i}(\mu) - A_i(\lambda)A_{r-i}(\lambda)],$$

which by induction on r implies $A_r(\lambda) = A_r(\mu)$ for every $r \le n$ and so for $r = s$. This also proves the second part of (i). ■

Is it possible to obtain, by using ϕ_n, measures distinct from any ergodic non-Markovian measure μ and having the same 2-marginals? This would be possible if $\phi_n\mu$ had the same marginals as μ, for $\phi_n\mu$ is necessarily distinct from μ (recall that the Markov chains are the only fixed points of ϕ_n). But, surprisingly, there are no measures other than the Chapman–Kolmogorov measures for which $\phi_n\mu$ has the same 2-marginals as μ. This is shown in the following theorem:

THEOREM 10. *For a measure $\mu \in M(\Omega, \sigma)$, let C_μ be the equivalence class consisting of the measures in $M(\Omega, \sigma)$ having the same 2-marginals as μ. Then the following are equivalent*:

(i) *μ is a Chapman–Kolmogorov (C.K.) measure.*
(ii) *There is $n \ge 2$ such that $\phi_n\mu \in C_\mu$.*
(iii) *$\phi_n\mu \in C_\mu$ for all $n \ge 2$.*
(iv) *There is n_0 such that $\phi_n\mu$ is C.K. for all $n \ge n_0$.*
(v) *$\phi_n\mu$ is C.K. for all $n \ge 2$.*

Proof. The implications (iii)⇒(ii) and (v)⇒(iv) are trivial. By (i)⇒(iii) of the Corollary above we see that (i) implies (v). We complete the proof by showing that (ii) implies (i) and that (iv) implies (iii).

Suppose (ii) and let $n \ge 2$ be such that $\phi_n\mu \in C_\mu$ so $A_r(\phi_n\mu) = A_r(\mu)$ for $r \in \mathbb{N}$. By Lemma 6(i) we have $A_1(\phi_n\mu) = A_1(\mu)$ and

$$A_r(\phi_n\mu) = \frac{n-r+1}{n} A_r(\mu) + \frac{1}{n} \sum_{i=1}^{r-1} A_i(\mu) A_{r-i}(\mu) \quad \text{for } r = 2, \ldots, n.$$

Therefore

$$A_r(\mu) = \frac{n-r+1}{n} A_r(\mu) + \frac{1}{n} \sum_{i=1}^{r-1} A_i(\mu) A_{r-i}(\mu) \quad \text{for } r = 2, \ldots, n.$$

As $i \in \{1, \ldots, r-1\}$ if and only if $r-i \in \{1, \ldots, r-1\}$ it follows by induction on r that

$$A_r(\mu) = (A_1(\mu))^r \quad \text{for } r = 1, \ldots, n.$$

Now if $l \in \mathbb{N}^*$ and $r \in \{0, 1, \ldots, n-1\}$ we see directly by Lemma 6(ii) that

$$A_{ln+r}(\mu) = (A_1(\mu))^{ln+r};$$

thus (ii) implies (i).

Finally suppose (iv), so for every $n \geq n_0$ we have $A_r(\phi_n\mu) = (A_1(\phi_n\mu))^r$ for $r = 1, \ldots, n$; but in view of (ii)⇒(i) of the Corollary above this implies $A_r(\mu) = (A_1(\mu))^r$ for $r = 1, \ldots, n$ and so (iv)⇒(i); but (i) implies (iii) by the Corollary above and the equalities $A_1(\phi_m\mu) = A_1(\mu)$, $m \geq 2$, so (iv) implies (iii). ■

Remark. We can replace (iv) by

(iv)′ There are infinitely many n such that $\phi_n\mu$ is C.K.

5. Infinite memory chains. An invariant measure which is not a Markov chain of a finite order is called a *chain of infinite memory*. We show that the $\phi_n\mu$'s have infinite memory and that the set of measures with infinite memory is dense in the class of 2-marginals of ergodic Markov chains. We first need the following lemma proved essentially in [5].

Lemma 7. *Let $\mu_1, \ldots, \mu_r$ be a family of distinct irreducible Markov chains on the same finite state space E. Suppose that there exist two distinct indices i and j such that for every state $y \in E$ there is a state x with $\mu_i(x,y)\mu_j(x,y) > 0$. Then every non-trivial convex combination $\sum_i \alpha_i\mu_i$ of μ_i's has infinite memory.*

Theorem 11. *Let $\mu \in M_n$ and $n \geq 2$. If μ is not Markovian and $A_n(\mu)$ is irreducible then $\phi_n\mu$ has infinite memory.*

Proof. Let θ be the canonical isomorphism between n-order Markov chains and Markov chains of order 1 (described in §2 for the measure ν_0). If $\phi_n\mu$ is a p-order Markov chain it is a pn-order Markov chain and therefore $\theta(\phi_n\mu)$ is also a p-order Markov chain on K^n. Then $\theta(\phi_n\mu)$ is a barycenter of the Markov chains $\theta(\sigma^i\nu_0)$, $i = 1, \ldots, n$. Since the $\sigma^i\nu_0$ for distinct i's are isomorphic, so also are the $\theta(\sigma^i\nu_n)$. Then Lemma 3 asserts that they are irreducible. Also Lemma 1 shows that ν_0 and $\sigma\nu_0$ are distinct. Then all we need to show is that they satisfy the last hypothesis in Lemma 7. So, let $b = (y_0, \ldots, y_{n-1}) \in K^n$ and let us prove that there exists $a = (x_0, \ldots, x_{n-1}) \in K^n$ such that

$$\theta(\nu_0)(a,b)\theta(\sigma\nu_0)(a,b) > 0. \tag{9}$$

Now

$$\theta(\nu_0(a,b)) = \mu(x_0,\ldots,x_{n-1})\mu(y_0 \,|\, x_0,\ldots,x_{n-1})\mu(y_1,\ldots,y_{n-1} \,|\, y_0),$$
$$\theta(\sigma\nu_0(a,b)) = \mu(x_0,\ldots,x_{n-1})\mu(y_0,\ldots,y_{n-1} \,|\, x_{n-1}).$$

Thus (9) is true if and only if there is some $(x_0,\ldots,x_{n-1}) \in K^n$ such that

$$\mu(x_0,\ldots,x_{n-1})\mu(x_0,\ldots,x_{n-1},y_0)\mu(y_0,\ldots,y_{n-1}) \times \mu(x_{n-1},y_0,\ldots,y_{n-1}) > 0,$$

that is, if and only if for some $(x_0,\ldots,x_{n-1}) \in K^n$,

$$\mu(x_0,\ldots,x_{n-1},y_0)\mu(x_{n-1},y_0,\ldots,y_{n-1}) > 0. \tag{10}$$

But this is true: in fact, take $x_{n-1} \in K$ such that $\mu(x_{n-1},y_0,\ldots,y_{n-1}) > 0$. It follows that $\mu(x_{n-1},y_0) > 0$; thus we can find $x_0,\ldots,x_{n-2} \in K$ such that $\mu(x_0,\ldots,x_{n-1},y_0) > 0$ so for this $a = (x_0,\ldots,x_{n-1})$ we obtain the two inequalities

$$\mu(x_0,\ldots,x_{n-1},y_0) > 0, \quad \mu(x_0,\ldots,x_{n-1})\mu(x_{n-1},y_0,\ldots,y_{n-1}) > 0,$$

which are equivalent to (10). ■

COROLLARY. *If Π is a $k \times k$ stochastic matrix, denote C_{μ_Π} by C_Π.*

(i) *If Π^n is irreducible and μ in $C_{\Pi,n}$ is such that $\mu \neq \mu_\Pi$, then $\phi_n\mu$ has infinite memory.*

(ii) *If Π is irreducible and aperiodic and μ in C_Π is such that $R_n\mu \neq \mu_\Pi$, then $\phi_n\mu$ has infinite memory.*

Thus, the set $S_\Pi = \bigcup \phi_n C_{\Pi,n} \setminus \mu_\Pi$ consists of infinite memory chains if Π is aperiodic and irreducible.

THEOREM 12. *If Π is irreducible then the set of all measures in C_Π with infinite memory is a G_δ-dense set in C_Π.*

Proof. Let $\mu \in C_\Pi$ and $\mu \neq \mu_\Pi$. Then there exist at most a finite number of n such that $R_n\mu = \mu_\Pi$. We know that for all n sufficiently large, $\phi_n R_n \mu \neq \mu_\Pi$ and converges to μ. The result will follow from (i) of the above Corollary. In fact, one can show that the irreducibility of Π implies that there are an infinite number of irreducible powers, denoted by Π^{n_k}, of Π. Then $\phi_{n_k} R_{n_k}\mu$ has infinite memory and converges to μ.

If $\mu = \mu_\Pi$, let $\mathcal{V}$ be a neighborhood of μ_Π. We have seen in [7] that C_Π contains measures distinct from μ_Π. Then there is such a measure also in $\mathcal{V}$ because C_Π is convex. We call it ν. Again, all but a finite number of $R_n\nu$ are non-Markovian. As $\phi_n R_n \nu \to \nu$, it follows, as above, that the $\phi_{n_k} R_{n_k}\nu$ have infinite memory and belong to $\mathcal{V}$ for large k. ■

Acknowledgments. We are indebted to J. P. Thouvenot for bringing to our attention the relation between $\phi_n\mu$ and an integral over a transformation. We thank him for this and for several discussions.

REFERENCES

[1] J. Bretagnolle and A. Kłopotowski, *Sur l'existence des suites de variables aléatoires s à s indépendantes échangeables ou stationnaires*, Ann. Inst. H. Poincaré 31 (1995), 325–350.

[2] I. P. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic Theory*, Springer, New York, 1981.

[3] M. Courbage and D. Hamdan, *A class of nonmixing dynamical systems with monotonic semigroup property*, Lett. Math. Phys. 22 (1991), 101–106.

[4] —, —, *Non-Markovian reversible Chapman–Kolmogorov measures on subshifts of finite type*, J. Statist. Phys. 74 (1994), 1281–1292.

[5] —, —, *Chapman–Kolmogorov equation for non-Markovian shift-invariant measures*, Ann. Probab. 22 (1994), 1662–1677.

[6] —, —, *Decay of correlations and mixing properties in a dynamical systems with zero K-S entropy*, Ergodic Theory Dynam. Systems 17 (1997), 87–103.

[7] —, —, *An ergodic Markov chain is not determined by its two-dimensional marginal laws*, Statist. Probab. Lett. 37 (1998), 35–40.

[8] —, —, *Examples of ergodic process uniquely determined by their two-dimensional marginal laws*, ibid. 52 (2001), 341–345.

[9] L. Flaminio, *Mixing k-fold independent processes of zero entropy*, Proc. Amer. Math. Soc. 118 (1993), 1263–1269.

[10] D. Hamdan, *Markov chains with positive transitions are not determined by any p-marginals*, Monatsh. Math. 130 (2000), 189–199.

[11] J. Lindenstrauss, *A short proof of Liapounoff's convexity theorem*, J. Math. Mech. 15 (1966), 971–972.

[12] J. B. Robertson and J. M. Womack, *A pairwise independent stationary stochastic process*, Statist. Probab. Lett. 3 (1985), 195–199.

M. Courbage
L.P.T.M.C.
Université Paris 7
2, Place Jussieu
75251 Paris Cedex 05, France
E-mail: courbage@ccr.jussieu.fr

D. Hamdan
Laboratoire de Probabilités
4, Place Jussieu
75252 Paris Cedex 05, France
E-mail: hamdan@ccr.jussieu.fr

Received 26 September 2000 (3974)