The Annals of Mathematical Statistics
1968, Vol. 39, No. 5, 1448–1462

ON THE ERGODICITY FOR NON-STATIONARY MULTIPLE MARKOV PROCESSES

By Hélène Deberghes and Bùi-Trọng-Liễu

Université de Lille

0. Introduction. In recent years, among many problems concerning non-stationary Markov processes with discrete time, much attention has been drawn to the problem of ergodicity. The problem is treated by Hajnal [3], Mott [6], Kozniewska [5], in the case of simple (i.e. first order) Markov processes with finite state space, the state space being the same for every instant t. Recently, one of the authors, in collaboration with Dorel [1], has extended the results of [5] to the case where the state space, for each t, is any measurable space. Recently also, Iosifescu [4], studied the uniform ergodicity of non-stationary multiple Markov processes. Using a definition of Iosifescu, we extend the results of [1] to kth order processes, $k \geqq 1$.

In the course of the present investigation, we introduce the notion of ergodicity of power h, h being a positive integer, and show that under certain conditions of uniformity, every non-stationary Markov process of order k which is ergodic of power h, with $h \geqq k$, is ergodic in the sense defined by Iosifescu. This simplifies considerably the eventual verification of ergodicity of a process (in the sense of Iosifescu). It is also shown, by means of examples, that there exist kth order Markov processes, ergodic of power h, with $h < k$, which are not ergodic in the sense of Iosifescu, thus justifying the notion of ergodicity of power h. The use of the "associated" simple process corresponding to the kth order process simplifies many proofs and, moreover, in the finite case, allows the passage from certain rectangular matrices to square matrices, thus facilitating the computations and the practical verification of the results. This article also shows that there are problems concerning kth order Markov processes which cannot be reduced to problems concerning first order Markov processes (for example, ergodicity of power $0 < h < k$).

1. Definitions, notation and preliminary remarks. In the following, we shall denote by $\mathbf{N}$ the set of non-negative integers, $\mathbf{N}^*$ the set of positive integers, and $\mathbf{1}_B$ the indicator of the set B.

Let $(\mathfrak{X}', \mathfrak{B}')$ and $(\mathfrak{X}'', \mathfrak{B}'')$ be two measurable spaces. We recall that a transition probability P from $(\mathfrak{X}', \mathfrak{B}')$ to $(\mathfrak{X}'', \mathfrak{B}'')$ is a mapping $P : \mathfrak{X}' \times \mathfrak{B}'' \rightarrow [0, 1]$ such that $\forall\, x \in \mathfrak{X}'$, $P(x, \cdot)$ is a probability measure on $\mathfrak{B}''$, and $\forall\, B \in \mathfrak{B}''$, $P(\cdot, B)$ is a real random variable defined on $(\mathfrak{X}', \mathfrak{B}')$.

Let $((\mathfrak{X}_t, \mathfrak{B}_t))_{t \in \mathbf{N}^*}$ be a sequence of measurable spaces. We shall denote the product σ-algebra of $(\mathfrak{B}_t)$, $1 \leqq t \leqq n$, (resp. $t \in \mathbf{N}^*$), by $\times_{t=1}^{n} \mathfrak{B}_t$ (resp. by $\times_{t \in \mathbf{N}^*} \mathfrak{B}_t$). In the particular case when $\mathfrak{X}_t = \mathfrak{X}$, we let $\times_{t=1}^{n} \mathfrak{B}_t = \times^{n} \mathfrak{B}$.

Received 4 August 1967.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

1.1. Let $k \varepsilon \mathbf{N}^*$ be a fixed number. Suppose that, $\forall t \varepsilon \mathbf{N}^*$, $t \geqq k$, there exists a transition probability $P_{t,t+1}^{(k)}$ from $(\prod_{j=t-k+1}^{t} \mathfrak{X}_j, \times_{j=t-k+1}^{t} \mathfrak{B}_j)$ to $(\mathfrak{X}_{t+1}, \mathfrak{B}_{t+1})$. *The sequence* $((\mathfrak{X}_t, \mathfrak{B}_t), P_{t+k-1,t+k}^{(k)})_{t\varepsilon\mathbf{N}^*}$ *is called a non-stationary kth order Markov process with discrete time and with state spaces* $(\mathfrak{X}_t, \mathfrak{B}_t)$, $t \varepsilon \mathbf{N}^*$.

From the $P_{t+k-1,t+k}^{(k)}$, we define the $P_{s,t}^{k,h}$ for $s \varepsilon \mathbf{N}^*$, $s \geqq k$; $h \varepsilon \mathbf{N}^*$ and $t \geqq s - k + 1$ as follows: $P_{s,t}^{k,h}$ is a mapping from $\prod_{j=s-k+1}^{s} \mathfrak{X}_j \times \times_{j=t}^{t+h-1} \mathfrak{B}_j$ to $[0, 1]$ such that $\forall x = (x_{s-k+1}, \cdots, x_s) \varepsilon \prod_{j=s-k+1}^{s} \mathfrak{X}_j$ and $\forall B \varepsilon \times_{j=t}^{t+h-1} \mathfrak{B}_j$

$$P_{s,t}^{k,h}(x, B) = \mathbf{1}_{\prod_{j=s-k+1}^{t-1} \mathfrak{X}_j \times B \times \prod_{i=t+h}^{s} \mathfrak{X}_i}(x) = \mathbf{1}_B(x_t, \cdots, x_{t+h-1}), \qquad \text{if} \quad t + h - 1 \leqq s$$

$$= \int_{\mathfrak{X}_{s+1}} P_{s,s+1}^{(k)}(x, dx_{s+1}) \int_{\mathfrak{X}_{s+2}} P_{s+1,s+2}^{(k)}((x_{s-k+2}, \cdots, x_{s+1}), dx_{s+2}) \cdots \int_{\mathfrak{X}_{t+h-1}} P_{t+h-2,t+h-1}^{(k)}((x_{t+h-k-1}, \cdots, x_{t+h-2}), dx_{t+h-1}) \cdot \mathbf{1}_B(x_t, \cdots, x_{t+h-1}), \qquad \text{if} \quad t + h - 1 > s.$$

We easily verify that $P_{s,t}^{k,h}$ is a transition probability from $(\prod_{j=s-k+1}^{s} \mathfrak{X}_j, \times_{j=s-k+1}^{s} \mathfrak{B}_j)$ to $(\prod_{j=t}^{t+h-1} \mathfrak{X}_j, \times_{j=t}^{t+h-1} \mathfrak{B}_j)$, and that $\forall s, t, u \varepsilon \mathbf{N}^*$, $k \leqq s$, $s - k + 1 \leqq u \leqq t$, $\forall x \varepsilon \prod_{j=s-k+1}^{s} \mathfrak{X}_j$, $\forall B \varepsilon \times_{j=t}^{t+h-1} \mathfrak{B}_j$,

$$\text{(1.1.a)} \qquad P_{s,t}^{k,h}(x, B) = \int_{\prod_{j=u}^{u+k-1} \mathfrak{X}_j} P_{s,u}^{k,k}(x, dy) P_{u+k-1,t}^{k,h}(y, B).$$

1.2. Let us consider the non-stationary first order Markov process $((\mathcal{Y}_t, \mathfrak{Z}_t), Q_{t,t+1})_{t\varepsilon\mathbf{N}^*, t \geqq k}$, where $\mathcal{Y}_t = \prod_{j=t-k+1}^{t} \mathfrak{X}_j$, $\mathfrak{Z}_t = \times_{j=t-k+1}^{t} \mathfrak{B}_j$, and $Q_{t,t+1}(x, B) = P_{t,t-k+2}^{k,k}(x, B)$, $\forall (x, B) \varepsilon \mathcal{Y}_t \times \mathfrak{Z}_{t+1}$. This process is called the *associated process of the* kth *order Markov process* $((\mathfrak{X}_t, \mathfrak{B}_t), P_{t+k-1,t+k}^{(k)})_{t\varepsilon\mathbf{N}^*}$. From the $Q_{t,t+1}(x, B)$, we define as usual,

$$Q_{s,s}(x, B) = \mathbf{1}_B(x), \qquad \forall (x, B) \varepsilon \mathcal{Y}_s \times \mathfrak{Z}_s,$$

and

$$Q_{s,t}(x, B) = \int_{\mathcal{Y}_{s+1}} Q_{s,s+1}(x, dx_{s+1}) \cdots \int_{\mathcal{Y}_{t-1}} Q_{t-2,t-1}(x_{t-2}, dx_{t-1}) Q_{t-1,t}(x_{t-1}, B)$$
$$\forall s, t \varepsilon \mathbf{N}^*, \quad k \leqq s < t \quad \text{and} \quad \forall (x, B) \varepsilon \mathcal{Y}_s \times \mathfrak{Z}_t.$$

Let us also recall that a kth order Markov process $((\mathfrak{X}_t, \mathfrak{B}_t), P_{t+k-1,t+k}^{(k)})_{t\varepsilon\mathbf{N}^*}$ is said to be stationary if

$$\mathfrak{X}_t = \mathfrak{X}, \qquad \mathfrak{B}_t = \mathfrak{B} \quad \text{and} \quad P_{t+k-1,t+k}^{(k)} = P^{(k)}, \qquad \forall t \varepsilon \mathbf{N}^*.$$

We then denote it by $((\mathfrak{X}, \mathfrak{B}), P_n^{(k)})_{n\varepsilon\mathbf{N}^*}$.

1.3. In the particular case of a finite state space (i.e. that of $\mathfrak{X} = \{1, 2, \cdots, r\}$ with $\mathfrak{B} = \mathfrak{P}(\mathfrak{X}) =$ class of all subset of $\mathfrak{X}$) we denote the process by $(\mathfrak{X}, P_{t,t+1}^{(k)})_{t\varepsilon\mathbf{N}^*, t \geqq k}$. The transition probability $P_{t,t+1}^{(k)}$ from $(\mathfrak{X}^k, \times^k \mathfrak{B})$ to $(\mathfrak{X}, \mathfrak{B})$ is entirely defined by the r^{k+1} numbers $P_{t,t+1}^{(k)}((i_1, \cdots, i_k), \{j\})$ where $i_1, \cdots, i_k$, $j \varepsilon \mathfrak{X}$. In the same way, the transition probability $P_{s,t}^{(k,h)}$, for $s, t, h \varepsilon \mathbf{N}^*$, $s \geqq k$, $t \geqq s - k + 1$, is entirely defined by the r^{k+h} numbers $P_{s,t}^{(k,h)}((i_1, \cdots, i_k), \{(j_1, \cdots, j_h)\})$ where $i_1, \cdots, i_k, j_1, \cdots, j_h \varepsilon \mathfrak{X}$.

It is convenient to introduce the mapping φ_h from $\mathfrak{X}^h$ to $\{1, 2, \cdots, r^h\}$, defined by

$$\varphi_h(i_1, \cdots, i_h) = \textstyle\sum_{j=1}^{h-1} (i_j - 1)r^{h-j} + i_h, \qquad \forall (i_1, \cdots, i_h) \in \mathfrak{X}^h.$$

Then, we can express the transition probability $P_{t,t+1}^{(k)}$ as a rectangular matrix, called transition matrix. Now, $\forall t \in \mathbf{N}^*$, $t \geqq k$, it is easy to construct the matrix $Q_{t,t+1}$ of the associated process from the $(P_{t,t+1}^{(k)})_{t\in\mathbf{N}^*, t\geqq k}$. Indeed, $\forall i, j \in \{1, \cdots, r^k\}$, where $i = \varphi_k(i_1, \cdots, i_k)$ and $j = \varphi_k(j_1, \cdots, j_k)$, $(i_1, \cdots, i_k)$ and $(j_1, \cdots, j_k) \in \mathfrak{X}^k$,

$$\begin{aligned} Q_{t,t+1}(i, \{j\}) &= P_{t,t-k+2}^{k,k}((i_1, \cdots, i_k), \{(j_1, \cdots, j_k)\}) \\ &= \textstyle\prod_{\alpha=2}^{k} \delta_{i_\alpha, j_{\alpha-1}} P_{t,t+1}^{(k)}(i, \{j_k\}), \end{aligned}$$

where $\delta_{\alpha,\beta} = 1$ if $\alpha = \beta$ and $\delta_{\alpha,\beta} = 0$ if $\alpha \neq \beta$.

2. Weak ergodicity. In the following, we assume that $\mathfrak{X}_t = \mathfrak{X}$, $\forall t \in \mathbf{N}^*$. Then $\prod_{j=1}^{h} \mathfrak{X}_j$ becomes $\mathfrak{X}^h$, and $\mathcal{Y} = \mathfrak{X}^k$. We also assume that the σ-algebra $\bigcap_{t\in\mathbf{N}^*} \mathcal{B}_t$ is not reduced to the σ-algebra $\{\emptyset, \mathfrak{X}\}$. Indeed, in this case, all the following definitions are verified. This assumption implies that, $\forall h \in \mathbf{N}^*$, the σ-algebra $\mathcal{A}_h = \bigcap_{t\in\mathbf{N}^*} \times_{j=t}^{t+h-1} \mathcal{B}_j$ is not reduced to the σ-algebra $\{\emptyset, \mathfrak{X}^k\}$. If $\mathcal{B}_t = \mathcal{B}$, $\forall t \in \mathbf{N}^*$, then $\mathcal{A}_k = \times^k \mathcal{B} = \mathfrak{I}$.

2.1. Definitions. A non-stationary kth order Markov process $((\mathfrak{X}, \mathcal{B}_t), P_{t+k-1,t+k}^{(k)})_{t\in\mathbf{N}^*}$ is said to be

(i) weakly ergodic of power h, $(h \in \mathbf{N}^*)$, if

$$\lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - P_{s,t}^{k,h}(y, B)] = 0,$$

$$\forall s \in \mathbf{N}^*, \qquad s \geqq k, \qquad \forall x, y \in \mathfrak{X}^k \text{ and } \forall B \in \mathcal{A}_h;$$

(ii) weakly ergodic, if it is weakly ergodic of power h, $\forall h \in \mathbf{N}^*$;

(iii) weakly and uniformly ergodic of power h, $(h \in \mathbf{N}^*)$, if $\mathcal{B}_t = \mathcal{B}$, $\forall t \in \mathbf{N}^*$ and if the process is weakly ergodic of power h, and if moreover the limit

$$\lim_{n\to\infty} [P_{s,s+n}^{k,h}(x, B) - P_{s,s+n}^{k,h}(y, B)] = 0$$

holds uniformly with respect to s, x, B;

(iv) weakly and uniformly ergodic, if it is weakly and uniformly ergodic of power h, $\forall h \in \mathbf{N}^*$ and if moreover, the limit

$$\lim_{n\to\infty} [P_{s,s+n}^{k,h}(x, B) - P_{s,s+n}^{k,h}(y, B)] = 0$$

holds uniformly with respect to s, x, B and h.

2.2. Proposition. *If a kth order Markov process $((\mathfrak{X}, \mathcal{B}_t), P_{t+k-1,t+k}^{(k)})_{t\in\mathbf{N}^*}$ (respectively $((\mathfrak{X}, \mathcal{B}), P_{t+k-1,t+k}^{(k)})_{t\in\mathbf{N}^*}$) is weakly (respectively weakly and uniformly) ergodic of power h, where $h \in \mathbf{N}^*$, it is weakly (respectively weakly and uniformly) ergodic of power h', $\forall h' \in \mathbf{N}^*$, $h' \leqq h$.*

We give the proof only for the non-uniform case (that of the other case uses analogous arguments). Let $h' \leqq h$. $\forall A_{h'} \in \mathcal{A}_{h'}$, let us consider the set $A_h = A_{h'} \times \mathfrak{X}^{h-h'}$. Then we have, $\forall s \in \mathbf{N}^*$, $s \geqq k$, and $\forall x, y \in \mathfrak{X}^k$,

$\lim_{t\to\infty} [P^{k,h}_{s,t}(x, A_h) - P^{k,h}_{s,t}(y, A_h)] = 0$. But, by the definition of $P^{k,h}_{s,t}$, $P^{k,h}_{s,t}(z, A_h) = P^{k,h'}_{s,t}(z, A_{h'})$, $\forall z \varepsilon \mathfrak{X}^k$, so that $\lim_{t\to\infty} [P^{k,h'}_{s,t}(x, A_{h'}) - P^{k,h'}_{s,t}(y, A_{h'})] = 0$.

2.3. PROPOSITION. *A kth order Markov process* $((\mathfrak{X}, \mathfrak{B}_t), P^{(k)}_{t+k-1,t+k})_{t\varepsilon\mathbf{N}^*}$ *(respectively* $((\mathfrak{X}, \mathfrak{B}), P^{(k)}_{t+k-1,t+k})_{t\varepsilon\mathbf{N}^*}$*) is weakly (respectively weakly and uniformly) ergodic of power k if and only if its associated process* $((\mathfrak{Y}, \mathfrak{I}_t), Q_{t,t+1})_{t\varepsilon\mathbf{N}^*, t\geq k}$ *(respectively* $((\mathfrak{Y}, \mathfrak{I}), Q_{t,t+1})_{t\varepsilon\mathbf{N}^*, t\geq k}$*) is weakly (respectively weakly and uniformly) ergodic of power* 1.

The proof, for the nonuniform case, for example, follows directly from the fact that $\forall s \varepsilon \mathbf{N}^*$, $s \geqq k$, $\forall (x, B) \varepsilon \mathfrak{X}^k \times \mathfrak{A}_k$, i.e. $\forall (x, B) \varepsilon \mathfrak{Y} \times \bigcap_{j\varepsilon\mathbf{N}^*, j\geq k} \mathfrak{I}_j$, $P^{k,k}_{s,t}(x, B) = Q_{s,t+k-1}(x, B)$ for $t \geqq s - k + 1$.

The following results generalize those concerning the first order Markov processes (cf. [1]).

2.4. PROPOSITION. *A necessary and sufficient condition for a k*th *order Markov process* $((\mathfrak{X}, \mathfrak{B}_t), P^{(k)}_{t+k-1,t+k})_{t\varepsilon\mathbf{N}^*}$ *(respectively* $((\mathfrak{X}, \mathfrak{B}), P^{(k)}_{t+k-1,t+k})_{t\varepsilon\mathbf{N}^*}$*) to be weakly (respectively weakly and uniformly) ergodic of power h,* $h \varepsilon \mathbf{N}^*$*, is the existence, for every* $s \varepsilon \mathbf{N}^*$, $s \geqq k$, *of a sequence of probability measures* $(\pi^{(h)}_{s,t})_{t\varepsilon\mathbf{N}, t\geq s-k+1}$, *each* $\pi^{(h)}_{s,t}$ *being defined on* $\times_{j=t}^{t+h-1} \mathfrak{B}_j$, *such that* $\forall x \varepsilon \mathfrak{X}^k$, $\forall B \varepsilon \mathfrak{A}_h$, $\forall s \varepsilon \mathbf{N}^*$, $s \geqq k$,

$$\text{(A)} \qquad \lim_{t\to\infty} [P^{k,h}_{s,t}(x, B) - \pi^{(h)}_{s,t}(B)] = 0,$$

(respectively that limit being uniform in s, x and B).

PROOF. We give the proof only for the nonuniform case. If there exists, for every $s \varepsilon \mathbf{N}$, $s \geqq k$, a sequence of probability measures $(\pi^{(h)}_{s,t})_{t\varepsilon\mathbf{N}, t\geq s-k+1}$ such that (A) is verified. Then, $\forall s \varepsilon \mathbf{N}^*$, $s \geqq k$; $\forall x, y \varepsilon \mathfrak{X}^k$ and $\forall B \varepsilon \mathfrak{A}_h$,

$$\lim_{t\to\infty} [P^{k,h}_{s,t}(x, B) - P^{k,h}_{s,t}(y, B)]$$
$$= \lim_{t\to\infty} [P^{k,h}_{s,t}(x, B) - \pi^{(h)}_{s,t}(B)] + \lim_{t\to\infty} [\pi^{(h)}_{s,t}(B) - P^{k,h}_{s,t}(y, B)] = 0,$$

and the sufficiency of the condition is proved. Conversely, suppose that the process is weakly ergodic of power h. $\forall s \varepsilon \mathbf{N}^*$, $s \geqq k$, let $\mu_s^{(k)}$ be a probability measure on $\times_{j=s-k+1}^{s} \mathfrak{B}_j$, and let us define, for every $t \varepsilon \mathbf{N}^*$, $t \geqq s - k + 1$, $\pi^{(h)}_{s,t}$ on $\times_{j=t}^{t+h-1} \mathfrak{B}_j$ as follows

$$\pi^{(h)}_{s,t}(B) = \int_{\mathfrak{X}^k} d\mu_s^{(k)}(x) P^{k,h}_{s,t}(x, B), \qquad \forall B \varepsilon \times_{j=t}^{t+h-1} \mathfrak{B}_j.$$

Then, $\forall x \varepsilon \mathfrak{X}^k$ and $\forall A \varepsilon \mathfrak{A}_h$,

$$\lim_{t\to\infty} [P^{k,h}_{s,t}(x, A) - \pi^{(h)}_{s,t}(A)]$$
$$= \lim_{t\to\infty} \int_{\mathfrak{X}^k} d\mu_s^{(k)}(y) (P^{k,h}_{s,t}(x, A) - P^{k,h}_{s,t}(y, A))$$
$$= \int_{\mathfrak{X}^k} d\mu_s^{(k)}(y) \lim_{t\to\infty} [P^{k,h}_{s,t}(x, A) - P^{k,h}_{s,t}(y, A)] = 0.$$

The necessity of the condition is thus proved.

2.5. PROPOSITION. *A kth order Markov process* $((\mathfrak{X}, \mathfrak{B}_t), P^{(k)}_{t+k-1,t+k})_{t\varepsilon\mathbf{N}^*}$ *is weakly ergodic of power h,* $h \varepsilon \mathbf{N}^*$, *if and only if,* $\forall s \varepsilon \mathbf{N}^*$, $s \geqq k$, *there exists* $v(s) \varepsilon \mathbf{N}^*$, $v(s) \geqq k$, *and a probability measure* $\mu^{(k)}_{v(s)}$ *on* $\times_{j=v(s)-k+1}^{v(s)} \mathfrak{B}_j$, *such that*

$$\lim_{t\to\infty} [P^{k,h}_{s,t}(x, B) - \int_{\mathfrak{X}^k} d\mu^{(k)}_{v(s)}(y) P^{k,h}_{v(s),t}(y, B)] = 0, \qquad \forall x \varepsilon \mathfrak{X}^k, \qquad \forall B \varepsilon \mathfrak{A}_h.$$

PROOF. *Necessity.* First, let us prove that if the Markov process is weakly ergodic of power h, $h \varepsilon \mathbf{N}^*$, then for any probability measure $\mu^{(k)}$ on $\times_{j=1}^{k} \mathcal{B}_j$,

$$\lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - \int_{\mathfrak{X}^k} d\mu^{(k)}(y) P_{k,t}^{k,h}(y, B)] = 0,$$
$$\forall s \varepsilon \mathbf{N}^*, \quad s \geqq k, \quad \forall x \varepsilon \mathfrak{X}^k, \quad \forall B \varepsilon \mathfrak{A}_h.$$

Indeed, $\forall s \varepsilon \mathbf{N}^*$, $s \geqq k$, let us define a probability measure $\pi_{k,s-k+1}^{(k)}$ on $\times_{j=s-k+1}^{s} \mathcal{B}_j$ by

$$\pi_{h,s-k+1}^{(k)}(A) = \int_{\mathfrak{X}^k} d\mu^{(k)}(y) P_{k,s-k+1}^{k,k}(y, A), \qquad \forall A \varepsilon \times_{j=s-k+1}^{s} \mathcal{B}_j,$$

and a probability measure $\pi_{s,t}^{(h)}$ on $\times_{j=t}^{t+h-1} \mathcal{B}_j$ by

$$\pi_{s,t}^{(h)}(A) = \int_{\mathfrak{X}^k} d\pi_{k,s-k+1}^{(k)}(y) P_{s,t}^{k,h}(y, A), \qquad \forall A \varepsilon \times_{j=t}^{t+h-1} \mathcal{B}_j.$$

2.4 shows that $\forall x \varepsilon \mathfrak{X}^k$ and $\forall B \varepsilon \mathfrak{A}_h$,

$$\lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - \pi_{s,t}^{(h)}(B)]$$
$$= \lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - \int_{\mathfrak{X}^k} d\pi_{k,s-k+1}^{(k)}(y) P_{s,t}^{k,h}(y, B)] = 0.$$

Applying the Lebesgue convergence theorem to $P_{s,t}^{k,h}(\cdot, B)$, we have

$$\lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - \int_{\mathfrak{X}^k} d\mu^{(k)}(z) \int_{\mathfrak{X}^k} P_{s,t}^{k,h}(y, B) P_{k,s-k+1}^{k,k}(z, dy)] = 0,$$

viz., by (1.1.a),

$$\lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - \int_{\mathfrak{X}^k} d\mu^{(k)}(z) P_{k,t}^{k,h}(z, B)] = 0.$$

Hence, $\forall s \varepsilon \mathbf{N}^*, s \geqq k; \forall v \varepsilon \mathbf{N}^*, v \geqq k$ (then *a fortiori*, we can say that $\exists\, v \varepsilon \mathbf{N}^*$, $v \geqq k$), $\exists$ a probability measure $\mu_v^{(k)}$ on $\times_{j=v-k+1}^{v} \mathcal{B}_j$ such that

$$\lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - \int_{\mathfrak{X}^k} d\mu_v^{(k)}(y) P_{v,t}^{k,h}(y, B)] = 0, \qquad \forall x \varepsilon \mathfrak{X}^k, \qquad \forall B \varepsilon \mathfrak{A}_h.$$

Indeed, we take: $\mu_v^{(k)} = \mu^{(k)}$, if $v = k$, where $\mu^{(k)}$ is any probability measure on $\times_{j=1}^{k} \mathcal{B}_j$,

$$\mu_v^{(k)}(B) = \int_{\mathfrak{X}^k} d\mu^{(k)}(y) P_{k,v-k+1}^{k,k}(y, B), \qquad \forall B \varepsilon \times_{j=v-k+1}^{k} \mathcal{B}_j \quad , \quad \text{if} \quad v > k.$$

Sufficiency. $\forall s \varepsilon \mathbf{N}^*$, $s \geqq k$, let us define from the $\mu_{v(s)}^{k}$, a sequence of probability measures $(\pi_{s,t}^{(h)})_{t\varepsilon\mathbf{N}^*, t\geqq s-k+1}$ as follows:

If $t > v(s)$, $\pi_{s,t}^{(h)}(A) = \int_{\mathfrak{X}^k} d\mu_{v(s)}^{(k)}(y) \cdot P_{v(s),t}^{k,h}(y, A)$, $\forall A \varepsilon \times_{j=t}^{t+h-1} \mathcal{B}_j$. If $t \leqq v(s)$, $\pi_{s,t}^{(h)} = \mu_t^{(h)}$, where $\mu_t^{(h)}$ is an arbitrary probability measure on $\times_{j=t}^{t+h-1} \mathcal{B}_j$. This sequence $(\pi_{s,t}^{(h)})_{t\varepsilon\mathbf{N}^*, t\geqq s-k+1}$ verifies the sufficiency of the condition of 2.4, so that the Markov process is weakly ergodic of power h. And this completes the proof.

2.6. PROPOSITION. *A necessary and sufficient condition for a* kth *order Markov process* $((\mathfrak{X}, \mathcal{B}_t), P_{t+k-1,t+k}^{(k)})_{t\varepsilon\mathbf{N}^*}$ *to be weakly ergodic of power* h, $h \varepsilon \mathbf{N}^*$, *is that whenever there exist* $s \varepsilon \mathbf{N}^*$, $s \geqq k$, $B \varepsilon \mathfrak{A}_h$, *and an increasing sequence* $(t_j)_{j\varepsilon\mathbf{N}}$ *of indices such that* $P_{s,t_j}^{k,h}(x, B)$ *converges to a limit when* $j \to \infty$, *for some* $x \varepsilon \mathfrak{X}^k$, *then* $P_{s,t_j}^{k,h}(y, B)$ *converges* $\forall y \varepsilon \mathfrak{X}^k$ *to the same limit. Moreover, this common limit is independent of* s.

PROOF. *Necessity.* Let us suppose that the Markov process is weakly ergodic of power h, and let $\mu^{(k)}$ be a probability measure on $\times_{j=1}^{k} \mathcal{B}_j$. We have seen, in the proof of 2.5, that $\forall\, s \in \mathbf{N}^*$, $s \geqq k$; $\forall\, x \in \mathfrak{X}^k$; $\forall\, B \in \mathcal{A}_h$,

$$\lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - \int_{\mathfrak{X}^k} d\mu^{(k)}(z) P_{k,t}^{k,h}(z, B)] = 0.$$

Let $\sigma = (t_i)_{i\in\mathbf{N}}$ be an increasing sequence of indices such that the sequence $(P_{s,t_i}^{k,h}(x, B))_{i\in\mathbf{N}, t_i \geqq s-k+1}$ converges. Such an increasing sequence σ always exists, for $(P_{s,t}^{k,h}(x, B))_{t\in\mathbf{N}, t\geqq s-k+1}$ is a sequence of numbers of the compact interval $[0, 1]$. Then,

$$\lim_{i\to\infty} [P_{s,t_i}^{k,h}(x, B) - \int_{\mathfrak{X}^k} d\mu^{(k)}(z) P_{k,t_i}^{k,h}(z, B)] = 0.$$

The sequence $(\int_{\mathfrak{X}^k} d\mu^{(k)}(z) P_{k,t_i}^{k,h}(z, B))_{i\in\mathbf{N}}$ then also converges, and we have

$$\lim_{i\to\infty} P_{s,t_i}^{k,h}(x, B) = \lim_{i\to\infty} \int_{\mathfrak{X}^k} d\mu^{(k)}(z) P_{k,t_i}^{k,h}(z, B).$$

This equality shows that if y is another state of $\mathfrak{X}^k$, we also have

$$\lim_{i\to\infty} P_{s,t_i}^{k,h}(y, B) = \lim_{i\to\infty} \int_{\mathfrak{X}^k} d\mu^{(k)}(z) P_{k,t_i}^{k,h}(z, B),$$

so that we can write $\lim_{i\to\infty} P_{s,t_i}^{k,h}(x, B) = \pi^{(h)}(\sigma, B)$, independent of s and x. The necessity of the condition is thus proved.

Sufficiency. $\forall\, s \in \mathbf{N}^*$, $s \geqq k$, $\forall\, x, y \in \mathfrak{X}^k$ and $\forall\, B \in \mathcal{A}_h$, $(P_{s,t}^{k,h}(x, B) - P_{s,t}^{k,h}(y, B))_{t\in\mathbf{N}^*, t\geqq s-k+1}$ is a sequence of numbers of the compact interval $[-1, +1]$. Then, we can find a subsequence $(P_{s,t_j}^{k,h}(x, B) - P_{s,t_j}^{k,h}(y, B))_{j\in\mathbf{N}^*}$ which converges to a limit. Let us denote this limit by L. For the same reason, we can find a subsequence $(P_{s,t_{j_l}}^{k,h}(x, B))_{l\in\mathbf{N}^*}$ of the subsequence $(P_{s,t_j}^{k,h}(x, B))_{j\in\mathbf{N}^*}$ which converges. But, by hypothesis, for another $y \in \mathfrak{X}^k$, the subsequence $(P_{s,t_{j_l}}^{k,h}(y, B))_{l\in\mathbf{N}^*}$ converges to the same limit. Hence,

$$\lim_{l\to\infty} [P_{s,t_{j_l}}^{k,h}(x, B) - P_{s,t_{j_l}}^{k,h}(y, B)] = 0.$$

This relation, and the existence of a limit for the sequence $(P_{s,t_j}^{k,h}(x, B) - P_{s,t_j}^{k,h}(y, B))_{j\in\mathbf{N}^*}$ implies that $L = 0$. Every subsequence of $(P_{s,t}^{k,h}(x, B) - P_{s,t}^{k,h}(y, B))_{t\in\mathbf{N}^*, t\geqq s-k+1}$ which converges has then limit 0. This implies that

$$\lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - P_{s,t}^{k,h}(y, B)] = 0,$$
$$\forall\, s \in N^*, \qquad s \geqq k, \qquad \forall\, x, y \in \mathfrak{X}^k, \qquad \forall\, B \in \mathcal{A}_h,$$

viz. the weak ergodicity of power h of the Markov process.

We state now the following theorem:

2.7. PROPOSITION. *If a* kth *order Markov process* $((\mathfrak{X}, \mathcal{B}), P_{t+k-1,t+k}^{(k)})_{t\in\mathbf{N}^*}$ *is weakly ergodic of power* k *and if moreover, the limit* $\lim_{t\to\infty} [P_{s,t}^{k,k}(x, B) - P_{s,t}^{k,k}(y, B)]$ *holds uniformly in* B, $B \in \times^k \mathcal{B}$, $\forall\, s \in \mathbf{N}^*$, $\forall\, x$ *and* $y \in \mathfrak{X}^k$, *then this process is weakly ergodic and moreover, the limit* $\lim_{t\to\infty} [P_{s,t}^{k,h}(x, B) - P_{s,t}^{k,h}(y, B)]$ *holds uniformly in* B, $\forall\, h \in \mathbf{N}^*$, $\forall\, s \in \mathbf{N}^*$, $\forall\, x$ *and* $y \in \mathfrak{X}^k$.

PROOF. For every $x, y \in \mathfrak{X}^k$, let $\mu_{s,n}(x, y; \cdot)$ be the signed measure, defined on $\times^k \mathcal{B}$ by

$$\mu_{s,n}(x, y; \cdot) = P_{s,s+n}^{k,k}(x, \cdot) - P_{s,s+n}^{k,k}(y, \cdot).$$

By the Jordan decomposition,

$$\mu_{s,n} = \overset{+}{\mu}_{s,n} - \overset{-}{\mu}_{s,n} .$$

Then, $\forall\, h \varepsilon \mathbf{N}^*$, $\forall\, x, y \varepsilon \mathfrak{X}^k$, $\forall\, B \varepsilon \mathsf{X}^h \mathcal{B}$, $\forall\, s, n \varepsilon \mathbf{N}^*$ with $s \geqq k$,

$$\begin{aligned} &P^{k,h}_{s,s+n}(x, B) - P^{k,h}_{s,s+n}(y, B) \\ &\quad = \int_{\mathfrak{X}^k} P^{k,k}_{s,s+n}(x, dz) P^{k,h}_{s+n+k-1,s+n}(z, B) - \int_{\mathfrak{X}^k} P^{k,k}_{s,s+n}(y, dz) P^{k,h}_{s+n+k-1,s+n}(z, B) \quad \text{by (1.1.a)} \\ &\quad = \int_{\mathfrak{X}^k} \mu_{s,n}(x, y; dz) P^{k,h}_{s+n+k-1,s+n}(z, B) \\ &\quad = \int_{\mathfrak{X}^k} \overset{+}{\mu}_{s,n}(x, y; dz) P^{k,h}_{s+n+k-1,s+n}(z, B) - \int_{\mathfrak{X}^k} \overset{-}{\mu}_{s,n}(x, y; dz) P^{k,h}_{s+n+k-1,s+n}(z, B). \end{aligned}$$

Since $0 \leqq P^{k,h}_{s+n+k-1,s+n}(z, B) \leqq 1$, we have

$$\begin{aligned} &-\overset{-}{\mu}_{s,n}(x, y; \mathfrak{X}^k) \\ &\quad \leqq \int_{\mathfrak{X}^k} \overset{+}{\mu}_{s,n}(x, y; dz) P^{k,h}_{s+n+k-1,s+n}(z, B) - \int_{\mathfrak{X}^k} \overset{-}{\mu}_{s,n}(x, y; dz) P^{k,h}_{s+n+k-1,s+n}(z, B) \\ &\quad \leqq \overset{+}{\mu}_{s,n}(x, y; \mathfrak{X}^k). \end{aligned}$$

But $\mu_{s,n}(x, y; \mathfrak{X}^k) = 0$, so that $\overset{+}{\mu}_{s,n}(x, y; \mathfrak{X}^k) = \overset{-}{\mu}_{s,n}(x, y; \mathfrak{X}^k)$. Consequently

$$|P^{k,h}_{s,s+n}(x, B) - P^{k,h}_{s,s+n}(y, B)| \leqq \overset{+}{\mu}_{s,n}(x, y; \mathfrak{X}^k).$$

By hypothesis, $\forall\, \epsilon > 0$, $\forall\, x, y \varepsilon \mathfrak{X}^k$, $\forall\, s \varepsilon \mathbf{N}^*$, $s \geqq k$, there exists $n_0(\epsilon, x, y, s) \varepsilon \mathbf{N}^*$ such that $n \geqq n_0(\epsilon, x, y, s) \Rightarrow |\mu_{s,n}(x, y; A)| < \epsilon$, $\forall\, A \varepsilon \mathsf{X}^k \mathcal{B}$. Now, if $\{H_{s,n}, \complement H_{s,n}\}$ is any Hahn decomposition for $\mu_{s,n}$, then

$$\begin{aligned} \overset{+}{\mu}_{s,n}(x, y; \mathfrak{X}^k) &= \mu_{s,n}(x, y; \mathfrak{X}^k \cap H_{s,n}) \\ &= \mu_{s,n}(x, y; H_{s,n}). \end{aligned}$$

$H_{s,n}$ being one of the A's cited above, we have consequently: $\forall\, \epsilon > 0$, $\forall\, x, y \varepsilon \mathfrak{X}^k$, $\forall\, s \varepsilon \mathbf{N}^*$, $s \geqq k$, $\exists\, n_0(\epsilon, x, y, s) \varepsilon \mathbf{N}^*$ such that $n \geqq n_0(\epsilon, x, y, s) \Rightarrow |P^{k,h}_{s,s+n}(x, B) - P^{k,h}_{s,s+n}(y, B)| < \epsilon$, $\forall\, h \varepsilon \mathbf{N}^*$ and $\forall\, B \varepsilon \mathsf{X}^h \mathcal{B}$, and this completes the proof.

One would be led to believe it possible to prove this result for a first order process (although this does not seem to have been proved elsewhere) and then to apply it to kth order processes by means of the associated process. However, the complicated form of the relations between $P^{k,h}_{s,t}$ and $Q^{1,l}_{s,t}$ can easily convince the reader how well-founded is our way of tackling the proof of this result.

This theorem has two interesting corollaries. The first of them is concerned with the uniform case, as follows:

2.8. PROPOSITION. *A necessary and sufficient condition for a* kth *order Markov process* $((\mathfrak{X}, \mathcal{B}), P^{(k)}_{t+k-1,t+k})_{t\varepsilon\mathbf{N}^*}$ *to be weakly and uniformly ergodic is the existence of some* $h \varepsilon \mathbf{N}^*$, $h \geqq k$, *such that the process is weakly and uniformly ergodic of power* h.

The necessary condition follows obviously from the definition 2.1. The sufficient condition uses the Proposition 2.2 and arguments analogous to those of the proof of 2.7.

The second corollary of 2.7 concerns the finite case. Indeed, if $\mathfrak{X}$ contains only a finite number of states, the condition of uniformity imposed in 2.7 is always satisfied, so that we have

2.9. PROPOSITION. *Let* $(\mathfrak{X}, P^{(k)}_{t+k-1,t+k})_{t\varepsilon\mathbf{N}^*}$ *be a* kth *order Markov process with a finite state space. Then it is weakly ergodic if and only if it is weakly ergodic of power* h, *for some integer* $h \geqq k$.

2.10. As mentioned in the introduction, we give now an example of multiple Markov processes which are *weakly ergodic of power* h, $h < k$, *without being weakly ergodic.*

In the following, we denote by $\{i \bmod j\}$ the set of positive integers which are equal to i modulo j.

Consider the Markov process $(\mathfrak{X}, P^{(4)}_{t+3,t+4})_{t\varepsilon\mathbf{N}^*}$, of order 4, with two states, $\mathfrak{X} = \{1, 2\}$, and with transition matrices

$$P^{(4)}_{t,t+1} = 1_{\{1 \bmod 3\}}(t)\cdot\begin{pmatrix} M' \\ M' \\ M' \\ M' \end{pmatrix} + 1_{\{2 \bmod 3\}}(t)\cdot\begin{pmatrix} M_1 \\ M_2 \\ M_2 \\ M_1 \end{pmatrix} + 1_{\{0 \bmod 3\}}(t)\cdot\begin{pmatrix} M'' \\ M'' \\ M'' \\ M'' \end{pmatrix}$$

where

$$M' = \begin{pmatrix} \frac12 & \frac12 \\ \frac12 & \frac12 \\ \frac12 & \frac12 \\ \frac12 & \frac12 \end{pmatrix}, \quad M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad M'' = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

These are rectangular matrices with 16 rows and two columns. The associated process $(\mathcal{Y}, Q_{t,t+1})_{t\varepsilon\mathbf{N}^*, t\geqq 4}$ has 16 states, $\mathcal{Y} = \{1, \cdots, 16\}$. It is easy to see that $Q_{t,t+1}$ is, depending on the value taken by t, equal to one of the three following square matrices (with 16 rows and 16 columns):

$$S_1 = Q_{3p+1,3p+2,}, \quad S_2 = Q_{3p+2,3p+3}, \quad S_3 = Q_{3p+3,3p+4}, \quad \forall\, p\ \varepsilon\ \mathbf{N}^*, \quad p \geqq 2.$$

Let now $S = S_3 S_1 S_2$. A computation shows that

$$S^n = U_1 = \begin{pmatrix} U_1' & 0 & U_1' & 0 \\ U_1' & 0 & U_1' & 0 \\ U_1' & 0 & U_1' & 0 \\ U_1' & 0 & U_1' & 0 \end{pmatrix}$$

where

$$U_1' = \begin{pmatrix} \frac14 & 0 & 0 & \frac14 \\ 0 & \frac14 & \frac14 & 0 \\ 0 & \frac14 & \frac14 & 0 \\ \frac14 & 0 & 0 & \frac14 \end{pmatrix}.$$

Consequently, $p\ \varepsilon\ N^*$, $p \geqq 2$,

(1°) $\lim_{m\to\infty} Q_{3p,3p+3m} = \lim_{m\to\infty} S^m = U_1$;

(2°) $\lim_{m\to\infty} Q_{3p,3p+3m+1} = U_1 S_3 = U_2$, where

$$U_2 = \begin{pmatrix} U_2' & 0 \\ U_2' & 0 \\ U_2' & 0 \\ U_2' & 0 \end{pmatrix} \quad \text{and} \quad U_2' = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix};$$

(3°) $\lim_{m\to\infty} Q_{3p,3p+3m+2} = U_2 S_1 = U_3$, where

$$U_3 = \begin{pmatrix} U_3' \\ U_3' \\ U_3' \\ U_3' \end{pmatrix} \quad \text{and} \quad U_3' = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 \end{pmatrix}.$$

U_1, U_2 and U_3 not being matrices with identical rows, 2.4 shows that the associated process $(\mathcal{Y}, Q_{t,t+1})_{t\in\mathbf{N}^*, t\geq 4}$ is not weakly ergodic of power 1. Consequently, by 2.3, the Markov process $(\mathcal{X}, P^{(4)}_{t+3,t+4})_{t\in\mathbf{N}^*}$ is *not weakly ergodic*. It is then not weakly ergodic of power h, for $h \geqq 4$. We now show that this process is *weakly and uniformly ergodic of power* 1, and that it is *not weakly ergodic of power* 2. Indeed, we have

(a) $\lim_{m\to\infty} P^{4,1}_{3p,3p+3m} = \lim_{m\to\infty} Q_{3p,3p+3m-1} \cdot P^{4,1}_{3p+3m-1,3p+3m} = U_3 P^{(4)}_{3p+2,3p+3} = \pi_1$,

where π_1 is a matrix with 16 rows and 2 columns, the elements of which are identical and equal to $\frac{1}{2}$.

(b) $$\lim_{m\to\infty} P^{4,1}_{3p,3p+3m+1} = U_1 P^{(4)}_{3p+3,3p+4} = \pi_2,$$

where π_2 is a matrix with 16 identical rows and 2 columns,

$$\pi_2 = \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \end{pmatrix}.$$

(c) $$\lim_{m\to\infty} P^{4,1}_{3p,3p+3m+2} = U_2 P^{(4)}_{3p+1,3p+2} = \pi_1.$$

So that, by 2.4, we have the weak and uniform ergodicity of power 1 for the Markov process $(\mathcal{X}, P^{(4)}_{t+3,t+4})_{t\in\mathbf{N}^*}$. We also have

$$\lim_{m\to\infty} P^{4,2}_{3p,3p+3m+2} = U_3 \cdot P^{4,2}_{3p+3m+5,3p+3m+2} = U_3 U_4,$$

where

$$U_4 = \begin{pmatrix} \mu_1 & \cdots & & 0 \\ \vdots & \mu_1 & & \vdots \\ & & \mu_1 & \\ 0 & \cdots & & \mu_1 \end{pmatrix}, \qquad \mu_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

The rows of the matrix $U_3 U_4$ not being identical, the process $(\mathcal{X}, P^{(4)}_{t+3,t+4})_{t\in\mathbf{N}^*}$ is not weakly ergodic of power 2.

3. Strong ergodicity. As in Section 2, we assume that $\mathcal{X}_t = \mathcal{X}$, $\forall\, t \in \mathbf{N}^*$, and that the σ-algebra $\bigcap_{t\in\mathbf{N}^*} \mathcal{B}_t$ is not reduced to the trivial σ-algebra $\{\varnothing, \mathcal{X}\}$.

3.1. DEFINITIONS. A non-stationary kth order Markov process $((\mathfrak{X}, \mathfrak{B}), P^{(k)}_{t+k-1,t+k})_{t\varepsilon \mathbf{N}^*}$ is said to be

(i) strongly ergodic of power h, $h \varepsilon \mathbf{N}^*$, if

$$\lim_{t\to\infty} P^{k,h}_{s,t}(x, B) = \pi_s^{(h)}(B), \qquad \forall\, s \varepsilon N^*,\ s \geqq k,\ \forall\ (x, B) \varepsilon \mathfrak{X}^k \times \mathfrak{A}_h$$

where $\pi_s^{(h)}$ is a probability measure on $\mathfrak{A}_h$;

(ii) strongly ergodic, if it is strongly ergodic of power h, $\forall\, h \varepsilon \mathbf{N}^*$.

(iii) strongly and uniformly ergodic of power h, $h \varepsilon \mathbf{N}^*$, if $\mathfrak{B}_t = \mathfrak{B}$, $\forall\, t \varepsilon \mathbf{N}^*$ and if the process is strongly ergodic of power h and if moreover, the limit $\lim_{n\to\infty} P^{k,h}_{s,s+n}(x, B)$ holds uniformly with respect to s, x, B;

(iv) strongly and uniformly ergodic, if it is strongly and uniformly ergodic of power h, $\forall\, h \varepsilon \mathbf{N}^*$, and if moreover the limit $\lim_{n\to\infty} P^{k,h}_{s,s+n}(x, B)$ holds uniformly with respect to h.

It is clear that

3.2. *The strong (respectively strong and uniform) ergodicity (respectively ergodicity of power h, $h \varepsilon \mathbf{N}^*$) implies the weak (respectively weak and uniform) ergodicity (respectively ergodicity of power h).*

3.3. PROPOSITION. *If a non-stationary* kth *order Markov process* $((\mathfrak{X}, \mathfrak{B}_t), P^{(k)}_{t+k-1,t+k})_{t\varepsilon \mathbf{N}^*}$ *is strongly ergodic of power h, $h \varepsilon \mathbf{N}^*$, then, $\forall\, s \varepsilon \mathbf{N}^*$, $s \geqq k$, $\forall\ (x, B) \varepsilon \mathfrak{X}^k \times \mathfrak{A}_h$,*

$$\lim_{t\to\infty} P^{k,h}_{s,t}(x, B) = \pi^{(h)}(B),$$

where $\pi^{(h)}$ is a probability measure on $\mathfrak{A}_h$, independent of s.

In other words, in Definition 3.1, we have $\pi_s^{(h)} = \pi^{(h)}$, $\forall s \varepsilon \mathbf{N}^*$, $s \geqq k$.

The proof of 3.3 does not differ from that of the simple case: it makes use of the Fatou-Lebesgue theorem. Indeed, $\forall\ s, u \varepsilon \mathbf{N}^*$, $s, u \geqq k$, let us assume that $s < u$, then by using (1.1.a), we have

$$\pi_s^{(h)}(B) = \lim_{t\to\infty} P^{k,h}_{s,t}(x, B) = \lim_{t\to\infty} \int_{\mathfrak{X}^k} P^{k,k}_{s,u-k+1}(x, dy) P^{k,h}_{u,t}(y, B)$$
$$= \int_{\mathfrak{X}^k} P^{k,k}_{s,u-k+1}(x, dy) \lim_{t\to\infty} P^{k,h}_{u,t}(y, B) = \pi_u^{(h)}(B), \qquad \forall\, B \varepsilon \mathfrak{A}_h.$$

3.4. PROPOSITION. *A necessary and sufficient condition for a non-stationary* kth *Markov process $((\mathfrak{X}, \mathfrak{B}_t), P^{(k)}_{t+k-1,t+k})_{t\varepsilon \mathbf{N}^*}$ to be strongly ergodic of power k is that its associated process is strongly ergodic of power* 1.

The proof follows directly from the fact that $\forall\ s \varepsilon \mathbf{N}^*$, $s \geqq k$, $\forall\ (x, B) \varepsilon \mathfrak{X}^k \times \mathfrak{A}_k$, i.e., $\forall\ (x, B) \varepsilon \mathfrak{Y} \times \bigcap_{j\varepsilon \mathbf{N}^*, j \geqq k} \mathfrak{J}_j$,

$$P^{k,k}_{s,t}(x, B) = Q_{s,t+k-1}(x, B) \quad \text{for } t \geqq s - k + 1.$$

3.5. PROPOSITION. *If a* kth *order Markov process* $((\mathfrak{X}, \mathfrak{B}), P^{(k)}_{t+k-1,t+k})_{t\varepsilon \mathbf{N}^*}$ *is strongly (respectively strongly and uniformly) ergodic of power h, with $h \varepsilon \mathbf{N}^*$, then it is strongly (respectively strongly and uniformly) ergodic of power h', $\forall\ h' \leqq h$.*

We omit the proof which is analoguous to the proof in the case of weak ergodicity.

We give now a sufficient condition for strong and uniform ergodicity.

3.6. PROPOSITION. *Let* $((\mathfrak{X}, \mathfrak{B}), P^{(k)}_{t+k-1,t+k})_{t\varepsilon \mathbf{N}^*}$ *be a* kth *order Markov process.*

If it is strongly and uniformly ergodic of power h for some integer $h \geqq k$, *and if moreover*, $\sum_{n \varepsilon \mathbf{N}} a_n < +\infty$, *where*

$$a_n = \sup |P^{k,1}_{m'+n,m'+n+1}(x, B) - P^{k,1}_{m''+n,m''+n+1}(x, B)|,$$

(*where the sup is taken over all* $x \varepsilon \mathfrak{X}^k$, m', $m'' \varepsilon \mathbf{N}^*$ *and* $\geqq k - n$, $B \varepsilon \mathcal{B}$).

Then, it is strongly and uniformly ergodic.

PROOF. Let us recall the condition $(C_k)_{m_0}$ introduced by Iosifescu in his paper [4]. This condition $(C_k)_{m_0}$, where $m_0 \varepsilon \mathbf{N}^*$, $m_0 \geqq k$, means that: $\exists\, n_0 \varepsilon \mathbf{N}^*$ and $\delta \varepsilon]0, 1[$ such that $|P^{k,k}_{m',m'+n_0}(x, B) - P^{k,k}_{m'',m''+n_0}(y, B)| < 1 - \delta$, $\forall\, m', m'' \varepsilon \mathbf{N}^*$, $m', m'' \geqq m_0$; $\forall\, x, y \varepsilon \mathfrak{X}^k$; $\forall\, B \varepsilon \mathbf{X}^k\mathcal{B}$.

Now, by hypothesis and by 3.5, the process is strongly and uniformly ergodic of power k, so that $\forall\, \epsilon > 0$, $\exists\, n_0(\epsilon) \varepsilon \mathbf{N}^*$ such that $n \geqq n_0(\epsilon) \Rightarrow |P^{k,k}_{s,s+n_0}(x, B) - \pi^{(k)}(B)| < \epsilon/2$, $\forall\, s \varepsilon \mathbf{N}^*$, $s \geqq k$, $\forall\, x \varepsilon \mathfrak{X}^k$ and $\forall\, B \varepsilon \mathbf{X}^k\mathcal{B}$, $\pi^{(k)}$ being a probability measure on $\mathbf{X}^k\mathcal{B}$, Then, $\forall\, m', m'' \varepsilon \mathbf{N}^*$, $m', m'' \geqq k$, $\forall\, x, y \varepsilon \mathfrak{X}^k$ and $\forall\, B \varepsilon \mathbf{X}^k\mathcal{B}$,

$$|P^{k,k}_{m',m'+n}(x, B) - P^{k,k}_{m'',m''+n}(y, B)|$$
$$\leqq |P^{k,k}_{m',m'+n}(x, B) - \pi^{(k)}(B)| + |P^{k,k}_{m'',m''+n}(y, B) - \pi^{(k)}(B)| < \epsilon,$$

when $n \geqq n_0(\epsilon)$. This is the condition $(C_k)_k$ cited above, with $\delta = 1 - \epsilon$. The Theorem 1 of [4], which states the strong and uniform ergodicity of the process under the conditions $\sum_{n \varepsilon \mathbf{N}} a_n < \infty$ and $(C_k)_{m_0}$, achieves the proof.

Naturally, there exist other kinds of sufficient condition, such as the following concerning finite Markov processes:

3.7. PROPOSITION. *Let* $(\mathfrak{X}, P^{(k)}_{t+k-1,t+k})_{t \varepsilon \mathbf{N}^*}$ *be a* kth *order Markov process with finite state space. If it is strongly ergodic of power h, for some* $h \varepsilon \mathbf{N}^*$, $h \geqq k$, *and if moreover* $\lim_{t\to\infty} P^{(k)}_{t,t+1}$ *exists, then it is strongly ergodic.*

PROOF. As in the preceding proposition, it is sufficient to show that, under the stated conditions, if the process is strongly ergodic of power k, then it is strongly ergodic. By hypothesis, $\lim_{t\to\infty} P^{(k)}_{t,t+1} = P^{(k)}$. Therefore, $\forall\, h \varepsilon \mathbf{N}^*$, $\lim_{t\to\infty} P^{k,h}_{t,t+1} = P^{k,h}$ also exists. Since the process is strongly ergodic of power k, we have $\lim_{t\to\infty} P^{k,k}_{s,t} = \pi^{(k)}$, where $\pi^{(k)}$ is a stochastic matrix indexed by i and $j \varepsilon \mathfrak{X}^k$, with identical rows. Now, $\forall\, s, t, h \varepsilon \mathbf{N}^*$, $s \geqq k$, $t \geqq s - k + 1$, we have $P^{k,h}_{s,t} = P^{k,k}_{s,t-k} \cdot P^{k,h}_{t-1,t}$. Then,

$$\lim_{t\to\infty} P^{k,h}_{s,t} = \lim_{t\to\infty} P^{k,k}_{s,t-k} \cdot \lim_{t\to\infty} P^{k,h}_{t-1,t} = \pi^{(k)} \cdot P^{k,h} = \pi^{(h)},$$

where $\pi^{(h)}$ is a matrix with identical rows. So, the process is strongly ergodic.

This proposition has the following corollary: under the same conditions as in 3.7, *the* kth *order Markov process is strongly ergodic if and only if the associated process is strongly ergodic of power* 1.

3.8. As mentioned in the introduction, we give now an example of multiple Markov processes which are *strongly ergodic of power h, h < k, without being strongly ergodic.*

Consider a non-stationary Markov process $(\mathfrak{X}, P^{(3)}_{t+2,t+3})_{t \varepsilon \mathbf{N}^*}$ of order 3, with 2 states, its transition matrices being of the form

$$P^{(3)}_{t,\,t+1} = \tfrac{1}{2}\cdot[1 - (-1)^t]P_1 + \tfrac{1}{2}\cdot[1 + (-1)^t]P_2$$

for $t \varepsilon \mathbf{N}^*$, $t \geqq 3$, where P_1 and P_2 are matrices with 8 rows and 2 columns; whose transposes are respectively

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

The passage to the associated process gives us

$$Q_{t,t+1} = \tfrac{1}{2}\cdot[1 - (-1)^t]Q_1 + \tfrac{1}{2}\cdot[1 + (-1)^t]Q_2$$

and a computation shows that

(1°) $\lim_{m\to\infty} Q_{2p,2p+2m} = U_1$, where

$$U_1 = \begin{pmatrix} U_1' & U_1' \\ U_1' & U_1' \end{pmatrix} \quad \text{and} \quad U_1' = \begin{Bmatrix} \frac{1}{4} & 0 & 0 & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} \end{Bmatrix}.$$

(2°) $\lim_{m\to\infty} Q_{2p;2p+2m+1} = U_2$, where

$$U_2 = \begin{pmatrix} U_2' \\ U_2' \end{pmatrix}, \quad \text{with} \quad U_2' = \begin{Bmatrix} \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} \end{Bmatrix}.$$

U_1 and U_2 not being matrices with identical rows, the associated process is not weakly ergodic of power 1, and *a fortiori* not strongly ergodic of power 1. The Markov process $(\mathfrak{X}, P^{(3)}_{t+2\,,t+3})_{t\varepsilon\mathbf{N}^*}$ is then *not strongly ergodic*. It is also easy to see that it is *not strongly ergodic of power* 2. But it is *strongly and uniformly ergodic of power* 1. Indeed,

$$\lim_{t\to\infty} P^{3,1}_{s,t} = P_1, \qquad \forall\ s \varepsilon \mathbf{N}^*, \quad s \geqq 3.$$

(In detail, one examines separately $\lim_{m\to\infty} P^{3,1}_{2p,2p+2m}$, $\lim_{m\to\infty} P^{3,1}_{2p+1,2p+2m}$, $\lim_{m\to\infty} P^{3,1}_{2p,2p+2m+1}$ and $\lim_{m\to\infty} P^{3,1}_{2p+1,2p+2m+1}$). In fact, the limits are reached uniformly, for:

(a) $\quad P^{3,1}_{2p,2p+2m} = Q_{2p,2p+2m-2}\cdot U_2\cdot P_1 = (U_2U_1)^{m-1}\cdot U_2\cdot P_1$

$$= P_1, \quad \text{for} \quad 2m \geqq 6,$$

$$\text{(b)} \quad P^{3,1}_{2p+1,2p+2m} = U_1(U_2U_1)^{m-2} \cdot U_2 \cdot P_1,$$

$$= P_1, \quad \text{for} \quad 2m - 1 \geqq 7,$$

$$\text{(c)} \quad P^{3,1}_{2p,2p+2m+1} = (U_2U_1)^m P_2 = P_1, \quad \text{for} \quad 2m + 1 \geqq 5,$$

$$\text{(d)} \quad P^{3,1}_{2p+1,2p+2m+1} = U_1(U_2U_1)^{m-1} P_2 = P_1, \quad \text{for} \quad 2m \geqq 6.$$

Consequently, the process is strongly and uniformly ergodic of power 1.

4. Relations between weak and strong ergodicities. In this section, we shall suppose that $\forall\, t \in \mathbf{N}^*$, $\mathcal{B}_t = \mathcal{B}$. The σ-algebra $\mathcal{A}_h$ will then be equal to $\times^h \mathcal{B}$, $\forall\, h \in \mathbf{N}^*$. The relations between weak ergodicity and strong ergodicity have been studied in detail in [1], for first order Markov processes. Different notions of "K-stationarity", (not to be confused with "stationary" Markov processes, which is synonymous with "homogeneous") are used to this effect.

4.1. Definitions. A kth order Markov process $((\mathfrak{X}, \mathcal{B}), P^{(k)}_{t+k-1,t+k})_{t \in \mathbf{N}^*}$ is said to be

(i) K-stationary, if $\exists$ a probability measure $\mu^{(k)}$ on $\times^k \mathcal{B}$, such that

$$\int_{\mathfrak{X}^k} P^{k,k}_{t,t-k+2}(x, B)\, d\mu^{(k)}(x) = \mu^{(k)}(B), \qquad \forall\, t \in \mathbf{N}^*,\ t \geqq k \quad \text{and} \quad \forall\, B \in \times^k \mathcal{B}.$$

(ii) K-asymptotically stationary if $\exists$ a probability measure on $\times^k \mathcal{B}$, such that

$$\lim_{t \to \infty} \int_{\mathfrak{X}^k} P^{k,k}_{s,t}(x, B)\, d\mu^{(k)}(x) = \mu^{(k)}(B), \qquad \forall\, s \in \mathbf{N}^*, s \geqq k \quad \text{and} \quad \forall\, B \in \times^k \mathcal{B}.$$

The results concerning kth order Markov processes which generalize those obtained for first order processes, are easily proved by passage to the associated processes. We shall only quote the following essential ones:

4.2. Proposition. *A necessary and sufficient condition for a nonstationary* kth *order Markov process* $((\mathfrak{X}, \mathcal{B}), P^{(k)}_{t+k-1,t+k})_{t \in \mathbf{N}^*}$ *to be strongly ergodic of power k is that it is weakly ergodic of power k and K-asymptotically stationary.*

Consider now the stationary case. As indicated in 1.2, the transition probabilities $P^{(k)}_{s,t}$ and $P^{k,h}_{s,t}$ depend only on the difference $t - s$. They are respectively denoted by $P^{(k)}_{t-s}$ *and* $P^{k,h}_{t-s}$.

4.3. Proposition. *Let* $((\mathfrak{X}, \mathcal{B}), P_n^{(k)})_{n \in \mathbf{N}^*}$ *be a stationary* kth *order Markov process such that its associated process verifies the condition* (*D*) *of Doeblin* (cf. [2]). *Then, it is strongly ergodic of power k if and only if it is weakly ergodic of power k.*

In fact, the condition (D) implies the K-stationarity of the process (cf. [1]) and the equivalence between the weak ergodicity and the strong ergodicity follows from 4.2.

We finally state the following result, concerning the stationary case, which improves a similar result obtained for first order processes in [1].

4.4. Proposition. *For a stationary* kth *order Markov process, the strong and uniform ergodicity of power h, $h \in \mathbf{N}^*$, and the weak and uniform ergodicity of power h, are equivalent.*

PROOF. It is sufficient to show that the weak and uniform ergodicity of power h implies the strong and uniform ergodicity of power h. For this purpose, we first prove that $\forall\, x \in \mathfrak{X}^k$ and $\forall\, B \in \times^h\mathcal{B}$, $P_n^{k,h}(x, B)$ converges uniformly in x and B to a limit independent of x, when $n \to \infty$, and then, that this limit is a probability measure on $\times^h\mathcal{B}$.

By the weak and uniform ergodicity of power h, $\forall\, \epsilon > 0$, $\exists\, n_0(\epsilon) \in \mathbf{N}^*$ such that $n \geqq n_0(\epsilon) \Rightarrow \sup_{x,y\in\mathfrak{X}^k} |P_n^{k,h}(x, B) - P_n^{k,h}(y, B)| < \epsilon/2$. Then $n \geqq n_0(\epsilon) \Rightarrow$

$$\sup_{x\in\mathfrak{X}^k} P_n^{k,h}(x, B) - \inf_{x\in\mathfrak{X}^k} P_n^{k,h}(x, B) < \epsilon/2. \tag{4.2.a}$$

Now, for $B \in \times^h\mathcal{B}$, $(\sup_{x\in\mathfrak{X}^k} P_n^{k,h}(x, B))_{n\in\mathbf{N}^*}$ is a decreasing sequence while $(\inf_{x\in\mathfrak{X}^k} P_n^{k,h}(x, B))_{n\in\mathbf{N}^*}$ is an increasing one, for if n_1 and n_2 are two positive integers such that $n_1 > n_2$, then

$$P_{n_1}^{k,h}(x, B) = \int_{\mathfrak{X}^k} P_{n_1-n_2-k+1}^{k,k}(x, dz)P_{n_2}^{k,h}(z, B).$$

Consequently,

$$\inf_{x\in\mathfrak{X}^k} P_{n_1}^{k,h}(x, B) \leqq \inf_{x\in\mathfrak{X}^k} P_{n_2}^{k,h}(x, B) \leqq \sup_{x\in\mathfrak{X}^k} P_{n_2}^{k,h}(x, B) \leqq \sup_{x\in\mathfrak{X}^k} P_{n_1}^{k,h}(x, B)$$

and by (4.2.a), the two sequences converge to a same limit when $n \to \infty$.

Let $\pi^{(h)}(B)$ be this common limit. We have $n \geqq n_0(\epsilon) \Rightarrow$

$$|P_n^{k,h}(x, B) - \pi^{(h)}(B)| < \epsilon/2. \tag{4.2.b}$$

It remains to show that $\pi^{(h)}$ is a probability measure on $\times^h\mathcal{B}$. It is clear that $\pi^{(h)}$ is a positive set function, taking its values in [0, 1], and that $\pi^{(h)}(\mathfrak{X}^k) = 1$. $\pi^{(h)}$ is also finitely additive, since for every finite family of pairwise disjoint sets of $\times^h\mathcal{B}$, say $(B_i)_{i\in I}$,

$$\begin{aligned}\pi^{(h)}(\textstyle\bigcup_{i\in I} B_i) &= \lim_{n\to\infty} P_n^{k,h}(x, \textstyle\bigcup_{i\in I} B_i)\\ &= \lim_{n\to\infty} \textstyle\sum_{i\in I} P_n^{k,h}(x, B_i)\\ &= \textstyle\sum_{i\in I} \lim_{n\to\infty} P_n^{k,h}(x, B_i) = \textstyle\sum_{i\in I} \pi^{(h)}(B_i).\end{aligned}$$

Finally, if $(A_m)_{m\in\mathbf{N}}$ is a decreasing sequence of sets of $\times^h\mathcal{B}$, such that $A_m \downarrow \varnothing$, then $\lim_{m\to\infty} \pi^{(h)}(A_m) = 0$. Indeed, $\forall\, n \in \mathbf{N}^*$ and $\forall\, x \in \mathfrak{X}^k$, $P_n^{k,h}(x, \cdot)$ being a probability measure, verifies this monotone sequential condition; therefore, $\forall\, \epsilon > 0$, $\exists\, m_0(\epsilon, n, x)$ such that $m \geqq m_0(\epsilon, n, x) \Rightarrow P_n^{k,h}(x, A_m) < \epsilon/2$. Now, let $m_0(\epsilon, x) = m_0(\epsilon, n_0(\epsilon), x)$; $n_0(\epsilon)$ being the same in (4.2.b). There exists a $z \in \mathfrak{X}^k$ such that $m_0(\epsilon, z) = \inf_{x\in\mathfrak{X}^k} m_0(\epsilon, x)$. Then, $m \geqq m_0(\epsilon, z) \Rightarrow \pi^{(h)}(A_m) \leqq |P_{n_0}^{k,h}(z, A_m) - \pi^{(h)}(A_m)| + P_{n_0}^{k,h}(z, A_m) < \epsilon$, that is to say $\lim_{m\to\infty} \pi^{(h)}(A_m) = 0$ and this completes the proof.

4.5. Consider now a stationary kth order Markov process with a finite state space, say $(\mathfrak{X}, P_n^{(k)})_{n\in\mathbf{N}^*}$. It is clear that the weak ergodicity of power h, $h \in \mathbf{N}^*$, and the strong one are always uniform, so that by 4.2, in this case, *the weak ergodicity of power h and the strong ergodicity of power h are equivalent* and are always uniform. Consequently, we can henceforth suppress the adjectives "weak" and "strong", and mention only whether $(\mathfrak{X}, P_n^{(k)})_{n\in\mathbf{N}^*}$ is ergodic of power h. We also remark that, by the stationarity, $\lim_{t\to\infty} P_{t,t+1}^{(k)}$ always exists. Using 3.7, 2.9, 2.7 and the preceding remarks, we conclude that:

For a stationary kth *order Markov process with a finite state space, weak and strong ergodicity are equivalent.*

4.6. We should like to end this article by pointing out that relations (for first order and stationary Markov processes) between various notions of ergodicity (of power 1) and the existence of a single ergodicity set with or without cyclic subsets have been examined in detail in [1].

REFERENCES

[1] BÙI-TRONG-LIỄU and DOREL, M. (1967). Sur le comportement asymptotique de processus de Markov non homogènes. *Studia Math.* **28** 253–274.

[2] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[3] HAJNAL, J. (1958). Weak ergodicity in nonhomogeneous Markov chains. *Proc. Cambridge Phil. Soc.* **54** 233–246.

[4] IOSIFESCU, M. (1966). Conditions nécessaires et suffisantes pour l'ergodicité uniforme des chaînes de Markov variables et multiples. *Rev. Roum. Math. Pures et Appl.* **11** No. 3, 325–330.

[5] KOZNIEWSKA, I. (1962). Ergodicité et stationnarité des chaînes de Markov variables à un nombre fini d'états possibles. *Col. Math.* **9** 333–345.

[6] MOTT, J. L. (1957). Conditions for the ergodicity of nonhomogeneous finite Markov chains. *Proc. Royal Edinburgh Soc. Ser. A* **64** 369–380.