NORTH-HOLLAND

On the Maximum of Ergodicity Coefficients, the Dobrushin Ergodicity Coefficient, and Products of Stochastic Matrices

Adolf Rhodius
Technische Universität Dresden
Fachrichtung Mathematik
Mommsenstraße 13
01062 Dresden, Germany

Submitted by Richard A. Brualdi

ABSTRACT

For a sequence of stochastic matrices we consider conditions for weak ergodicity of infinite products taken in an arbitrary order. The main tools of the investigations are ergodicity coefficients corresponding to a norm. Comparisons between different ergodicity coefficients show the important role played by the Dobrushin ergodicity coefficient. In particular, the application of the Dobrushin ergodicity coefficient yields generalizations of Leizarowitz's ergodicity conditions. © *Elsevier Science Inc., 1997*

1. INTRODUCTION

Let $(Q_k)_{k=0}^{\infty}$ be a sequence of $n \times n$ stochastic matrices Q_k. For each permutation π of the set $N = \{0, 1, 2, \dots\}$ we define inductively sequences (H_m) of products H_m by either $H_{m+1} = H_m \tilde{Q}_{m+1}$ or $H_{m+1} = \tilde{Q}_{m+1} H_m$ $(m \in N)$, $H_0 = \tilde{Q}_0$, where $\tilde{Q}_k = Q_{\pi(k)}$ $(k \in N)$. Such products are, e.g., the forward products

$$P_m = \tilde{Q}_0 \tilde{Q}_1 \cdots \tilde{Q}_{m-1} \tilde{Q}_m$$

and the backward products

$$M_m = \tilde{Q}_m \tilde{Q}_{m-1} \cdots \tilde{Q}_1 \tilde{Q}_0,$$

655 Avenue of the Americas, New York, NY 10010 0024-3795/97/$17.00
SSDI 0024-3795(95)00706-7

whose asymptotic behavior has been investigated in many works (see e.g. [1–10, 15–17]). For example, it is well known that weak ergodicity obtains for the backward products M_m if and only if strong ergodicity obtains for M_m (see [17, p. 154]). An infinite product (H_m) is said to be weakly ergodic [strongly ergodic] if $(H_m)_{ik} - (H_m)_{jk} \to 0$ for all $i, j, k = 1, 2, \dots, n$ [if (H_m) is weakly ergodic and $(H_m)_{ij}$ tends to a limit as $m \to \infty$ for every i and j]. We say the weak [strong] ergodicity obtains for the infinite products of $(Q_k)_{k=0}^{\infty}$ if all products (H_m) constructed in the above-described way are weakly [strongly] ergodic (see [7]).

One aim of this paper is to prove sufficient conditions for the weak ergodicity of the infinite products of $(Q_k)_{k=0}^{\infty}$ on the basis of the set W of all accumulation points of $(Q_k)_{k=0}^{\infty}$, using ergodicity coefficients; a second purpose concerns relations between different ergodicity coefficients; and a third relates to subsets V of S_n for which every finite product formed by members of V is ergodic. S_n denotes the set of all $n \times n$ stochastic matrices.

We call a stochastic matrix $P \in S_n$ ergodic if and only if P^m tends to a stable matrix Q as $m \to \infty$. Q is said to be stable if Q has rank one. Ergodic matrices are often called regular [17] or SIA (stochastic, indecomposable, aperiodic); see [8]. P is ergodic if and only if P^m is scrambling for some $m \in N$.

Let $n \geqslant 2$, $\|\cdot\|$ be a norm on $\mathbb{R}^n$, $H = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n : \sum_{i=1}^n x_i = 0\}$, and S_n be the set of all $n \times n$ stochastic matrices. Then

$$\tau_{\|\cdot\|}(P) = \max\{\|xP\| : x \in H, \|x\| = 1\} \qquad (P \in S_n)$$

is the ergodicity coefficient corresponding to the norm $\|\cdot\|$ (see [16]). An important ergodicity coefficient is the Dobrushin ergodicity coefficient $\delta(P)$; it corresponds to the l_1-norm $\|x\|_1 = \sum_{i=1}^n |x_i|$ $(x \in \mathbb{R}^n)$ and is often preferable to other ergodicity coefficients because of its simple explicit form

$$\delta(P) = \tfrac{1}{2} \max_{i,j} \sum_{k=1}^{n} |p_{ik} - p_{jk}|,$$

where $P = (p_{ij})_{i,j=1,2,\dots,n}$. Moreover, the Dobrushin coefficient δ is the only ergodicity coefficient corresponding to a norm $\|\cdot\|$ with

$$\max\{\tau_{\|\cdot\|}(P) : P \in S_n\} = 1;$$

in all other cases one has $\max\{\tau_{\|\cdot\|}(P) : P \in S_n\} > 1$ (see [6, 8]). A stochastic matrix $P \in S_n$ with $\delta(P) < 1$ is called scrambling. P is scrambling if and only if no two rows of P are orthogonal. Probably Leizarowitz [7] was the first to use the set W of accumulation points of a sequence (Q_k) for the investigation of the corresponding Markov chains; the case of singleton sets W was treated by Mott [9]. The basic Theorem A in [7] gives two sufficient conditions for the weak ergodicity of all infinite products $(H_m)_{m=0}^{\infty}$:

I. There exists at least one matrix $Q \in W$ which satisfies

$$\tau_{\|\cdot\|}(Q) < 1/C_{\|\cdot\|}, \qquad \text{where} \quad C_{\|\cdot\|} = \max\{\tau_{\|\cdot\|}(P) : P \in S_n\}.$$

II. $\tau_{\|\cdot\|}(Q) < 1$ for every $Q \in W$.

If a $Q \in W$ is scrambling, then I holds for the l_1-norm. Our Theorem 1 will show that any $Q \in S_n$ fulfilling I for some norm is scrambling. Therefore condition I should be replaced by the condition that at least one $Q \in W$ is scrambling. This new form of condition I is more applicable than the original one, since—in contrast to an arbitrary norm—one has no problems in computing $\tau_{\|\cdot\|_1}$ $(= \delta)$ and $C_{\|\cdot\|_1}$ $(= 1)$. Theorem 1 gives the following relation between an arbitrary ergodicity coefficient and Dobrushin's coefficient: $m = M = C_{\|\cdot\|}$ are the smallest numbers $m, M > 0$ such that $m^{-1}\delta(P) \leqslant \tau_{\|\cdot\|}(P) \leqslant M\delta(P)$ holds for all $P \in S_n$. Since for every ergodic matrix $P \in S_n$ of interest there are ergodicity coefficients τ_1, τ_2 such that $\tau_1(P) < 1$, $\tau_2(P) > 1$ (Corollary 3), it is useful to take another ergodicity coefficient that δ in case no $P \in W$ is scrambling. On the other hand, our Theorem 3, generalizing Leizarowitz's theorem, only needs the Dobrushin ergodicity coefficient δ. This emphasizes that δ is rather prominent in the class of all ergodicity coefficients corresponding to a norm.

We say $V \subset S_n$ has the finite product ergodicity (FPE) property if all finite products of members of V are ergodic matrices. It is well known that every inhomogeneous Markov chain whose transition matrices are chosen from a FPE property set is weakly ergodic (see [1]). Our results are a generalization of this fact.

2. THE MAXIMUM OF ERGODICITY COEFFICIENTS AND THE COMPARISON OF DIFFERENT ERGODICITY COEFFICIENTS

LEMMA 1. *Let $L(H)$ be the set of all linear operators $B : H \to H$. Then $A \in L(H)$ if and only if there are a stochastic matrix $P \in S_n$ and $k_A > 0$ such that $Ax = (xP)k_A$ $(x \in H)$.*

Proof. For every $A \in L(H)$ there exists an $n \times n$ matrix $Q_A = (q_{ij})$ and a real number c_A with

$$Ax = xQ_A \quad (x \in H), \qquad \sum_{j=1}^{n} q_{ij} = c_A \quad (i = 1, 2, \dots, n).$$

Q_A and $\tilde{Q}_A$ represent the same operator A if and only if $Q_A - \tilde{Q}_A$ is a stable matrix. Therefore there exists a matrix Q_A representing A with $q_{ij} \geqslant 0$, $c_A > 0$. Then $P = (1/c_A)Q_A$ is a stochastic matrix and $k_A = c_A$. ■

LEMMA 2. *Let $\|\cdot\|$ be norm on $\mathbb{R}^n$, and $S^{(1)} = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i^+ = \sum_{i=1}^n x_i^- = \frac{1}{2}\}$. Then*

$$C_{\|\cdot\|} = \max\{\tau_{\|\cdot\|}(P) : P \in S_n\}$$

$$= \max\{\tau_{\|\cdot\|}(P) : P \in S_n,\ \delta(P) = 1\}, \tag{a}$$

$$C_{\|\cdot\|} \geqslant \frac{\max_{x \in S^{(1)}} \|x\|}{\min_{x \in S^{(1)}} \|x\|}. \tag{b}$$

Proof. (a): Since $\tau_{\|\cdot\|}$ is a convex functional on the convex subset S_n of $\mathbb{R}^{n^2}$, it assumes its maximum at an extremal point of S_n. If Q_0 is an extremal point of S_n with

$$\tau_{\|\cdot\|}(Q_0) = \max\{\tau_{\|\cdot\|}(P) : P \in S_n\},$$

then Q_0 is nonscrambling, because a stochastic matrix Q is an extremal point of S_n if and only if Q has the number 1 as an element in each of its rows. Therefore only $\delta(Q_0) = 1$ is possible; the case $\delta(Q_0) = 0$ would imply $\tau_{\|\cdot\|}(Q_0) = 0$.

(b): Let Extr S_n be the set of all extremal points of S_n. We have

$$C_{\|\cdot\|} = \max\{\tau_{\|\cdot\|}(P) : P \in S_n\}$$

$$= \max_{x \in H,\ \|x\| = 1} \left(\max_{P \in S_n} \|xP\| \right).$$

Since the functionals $f_x(P) = \|xP\|$ are convex on $S_n = \operatorname{conv}(\operatorname{Extr} S_n)$, we obtain

$$C_{\|\cdot\|} = \max_{x\in H,\, \|x\|=1} \Big(\max_{Q \in \operatorname{Extr} S_n} \|xQ\| \Big).$$

The stochastic matrix $Q = (q_{ij}) \in S_n$ belongs to $\operatorname{Extr} S_n$ if and only if there are n sets $\Lambda_1, \Lambda_2, \dots, \Lambda_n$ such that the following conditions hold:

(i) $\{1, 2, \dots, n\} = \Lambda_1 \cup \Lambda_2 \cup \cdots \cup \Lambda_n$.
(ii) $\Lambda_i \cap \Lambda_j = \varnothing$ $(i \neq j,\ i, j = 1, 2, \dots, n)$.
(iii) For each $j = 1, 2, \dots, n$ we have $q_{ij} = 1$ if and only if $i \in \Lambda_j$.

Therefore it follows that

$$xQ = \Big(\sum_{i\in\Lambda_1} x_i, \sum_{i\in\Lambda_2} x_i, \dots, \sum_{i\in\Lambda_n} x_i \Big) \qquad (x \in \mathbb{R}^n),$$

and we obtain

$$C_{\|\cdot\|} = \max_{x\in H,\, \|x\|=1} \left(\max_{(\Lambda_1, \Lambda_2, \dots, \Lambda_n)\in\Lambda} \left\| \Big(\sum_{i\in\Lambda_1} x_i, \sum_{i\in\Lambda_2} x_i, \dots, \sum_{i\in\Lambda_n} x_i \Big) \right\| \right),$$

where Λ denotes the set of all decompositions $(\Lambda_1, \Lambda_2, \dots, \Lambda_n)$ of $\{1, 2, \dots, n\}$ satisfying the conditions $\{1, 2, \dots, n\} = \Lambda_1 \cup \Lambda_1 \cup \cdots \cup \Lambda_n$, $\Lambda_i \cap \Lambda_j = \varnothing$ for $i \neq j$, $i, j = 1, 2, \dots, n$. We choose an $x_0 \in H$ with

$$\|x_0\| = 1 \quad \text{and} \quad \|x_0\|_1 = \max_{x\in H,\, \|x\|=1} \|x\|_1.$$

Since $z = x_0/\|x_0\|_1 \in S^{(1)}$, we have $\sum_{i=1}^n z_i^+ = \sum_{i=1}^n z_i^- = \frac{1}{2}$. It follows that

$$C_{\|\cdot\|} \geqslant \|x_0\|_1 \cdot \left\| \Big(\sum_{i\in\Lambda_1} z_i, \sum_{i\in\Lambda_2} z_i, \dots, \sum_{i\in\Lambda_n} z_i \Big) \right\|$$

for all decomposition $(\Lambda_1, \Lambda_2, \dots, \Lambda_n)$. On the other hand (see [14, 12]) we have

$$\operatorname{Extr} S^{(1)} = \left\{ \tfrac{1}{2}(e_i - e_j) : i \neq j,\ i, j = 1, 2, \dots, n \right\},$$

where

$$e_i = (\delta_{i1}, \delta_{i2}, \dots, \delta_{in}), \qquad i = 1, 2, \dots, n,$$

the set of the extremal points of $S^{(1)}$. The equation $\sum_{i=1}^n z_i^+ = \sum_{i=1}^n z_i^-$ implies the inclusion

$$\operatorname{Extr} S^{(1)} \subset \left\{ \left(\sum_{i \in \Lambda_1} z_i, \dots, \sum_{i \in \Lambda_n} z_i \right) : (\Lambda_1, \Lambda_2, \dots, \Lambda_n) \in \Lambda \right\};$$

it therefore follows that

$$C_{\|\cdot\|} \geqslant \max_{x \in H,\, \|x\| = 1} \|x\|_1 \max_{x \in H,\, x \in S^{(1)}} \|x\|.$$

The observation $\max_{x \in H,\, \|x\|=1} \|x\|_1 = \dfrac{1}{\min_{x \in S^{(1)}} \|x\|}$ completes the proof. ∎

THEOREM 1. *Let* $\|\cdot\|$ *be a norm on* $\mathbb{R}^n$ *and*

$$S^{(1)} = \left\{ x \in \mathbb{R}^n : \sum_{i=1}^n x_i^+ = \sum_{i=1}^n x_i^- = \frac{1}{2} \right\}.$$

Then:

(a) *One has*

$$\begin{aligned} C_{\|\cdot\|} &= \max\{\tau_{\|\cdot\|}(P) : P \in S_n\} \\ &= \max\{\tau_{\|\cdot\|}(P) : P \in S_n,\ \delta(P) = 1\} \\ &= \frac{\max_{x \in S^{(1)}} \|x\|}{\min_{x \in S^{(1)}} \|x\|}. \end{aligned}$$

(b) *The smallest numbers* m, $M > 0$ *fulfilling*

$$\frac{1}{m} \delta(P) \leqslant \tau_{\|\cdot\|}(P) \leqslant M \delta(P) \qquad (P \in S_n)$$

are $m = M = C_{\|\cdot\|}$

Proof. Let

$$R = \max_{x \in M \setminus \{0\}} \frac{\|x\|}{\|x\|_1}, \qquad R_1 = \max_{x \in H \setminus \{0\}} \frac{\|x\|_1}{\|x\|}.$$

Obviously

$$R = \max_{x \in S^{(1)}} \|x\|, \qquad R_1 = \frac{1}{\min_{x \in S^{(1)}} \|x\|}. \tag{$*$}$$

We denote $|||\cdot|||_1$, $|||\cdot|||$ the operator norms induced by the vector norms $\|\cdot\|_1$, $\|\cdot\|$, respectively, that is,

$$|||A|||_1 = \max\{\|Ax\|_1 : x \in S^{(1)}\},$$

$$|||A||| = \max\{\|Ax\| : x \in H, \|x\| = 1\}$$

for all $A \in L(H)$. The vector norms $\|\cdot\|$, $\|\cdot\|_1$ on H and the corresponding operator norms $|||A|||$, $|||A|||_1$ are related like the vector norms and the induced matrix norms considered in Theorem 5.6.18 of [5]. Thus this theorem may be applied so that we obtain

$$\max_{A \in L(H),\, A \neq 0} \frac{|||A|||_1}{|||A|||} = \max_{A \in L(H),\, A \neq 0} \frac{|||A|||}{|||A|||_1} = R_1 R.$$

Lemma 1 yields

$$\max_{P \in S^n} \frac{\delta(P)}{\tau_{\|\cdot\|}(P)} = \max_{P \in S_n} \frac{\tau_{\|\cdot\|}(P)}{\delta(P)} = R_1 R.$$

Thus, we have

$$\frac{1}{m}\delta(P) \leqslant \tau_{\|\cdot\|}(P) \leqslant M \cdot \delta(P) \qquad (P \in S_n) \tag{$**$}$$

for $m = M = R_1 R$, where $R_1 R$ is the smallest number for m and M such that the inequality holds.

It remains to verify $R_1 R = C_{\|\cdot\|}$. Lemma 2 and $(*)$ imply $C_{\|\cdot\|} \geqslant R_1 R$. The inequality $(**)$ and the property $\delta(P) \leqslant 1$ $(P \in S_n)$ yield

$$\pi_{\|\cdot\|}(P) \leqslant R_1 R \delta(P) \leqslant R_1 R,$$

and therefore $C_{\|\cdot\|} \leqslant R_1 R$. ■

COROLLARY 1. *For all norms* $\|\cdot\|$ *on* $\mathbb{R}^n$ *one has*:

(1) *If* $Q \in S_n$ *is scrambling, then* $\tau_{\|\cdot\|}(Q) < C_{\|\cdot\|}$.
(2) *A matrix* $Q \in S_n$ *with* $\tau_{\|\cdot\|}(Q) < 1/C_{\|\cdot\|}$ *is scrambling.*

Proof. The corollary follows directly from Theorem 1(b). The right inequality of (b) implies (1); the left inequality yields (2). ■

COROLLARY 2. *If* $\|\cdot\|$ *is a norm on* $\mathbb{R}^n$, *the following statements are equivalent*:

(a) $C_{\|\cdot\|} = 1$.
(b) $\tau_{\|\cdot\|} = \delta$.
(c) *There is a* $K > 0$ *with* $\|x\| = K\|x\|_1$ $(x \in H)$.

Proof. The equation $C_{\|\cdot\|} = \max_{x \in S^{(1)}}\|x\| / \min_{x \in S^{(1)}}\|x\|$ yields the equivalence of (a), (b), and (c). ■

REMARK 1. *Let* τ, σ *be two different ergodicity coefficients corresponding to the norms* $\|\cdot\|_\tau, \|\cdot\|_\sigma$, *respectively. Then*:

(a) *There are stochastic matrices* P, Q *with* $\tau(P) < \sigma(P)$ *and* $\tau(Q) > \sigma(Q)$.
(b) *If* $C_{\|\cdot\|_\sigma} > C_{\|\cdot\|_\tau}$, *then there exists a stochastic matrix* P *with* $\tau(P) < 1$, $\sigma(P) \geqslant 1$.

Proof. (a): See [13, Theorem 5.2]. (b): The condition $C_{\|\cdot\|_\sigma} > C_{\|\cdot\|_\tau}$ implies the existence of a stochastic matrix $\tilde{P}$ with $C_{\|\cdot\|_\sigma} = \sigma(\tilde{P}) > 1$ and $0 < \tau(\tilde{P}) < \sigma(\tilde{P})$. We put $\sigma(\tilde{P}) = \alpha^{-1}$. For any stable matrix $\tilde{Q}_0 \in S_n$ we have $P = \alpha\tilde{P} + (1-\alpha)Q_0 \in S_n$ and $\tau(P) = \alpha\tau(\tilde{P}) < \alpha\sigma(\tilde{P}) = \sigma(P) = 1$. ■

THEOREM 2. *Let* $P \in S_n$ *be nonstable and ergodic*, $n \geqslant 3$. *The following alternative holds*:

(A) *either* $\tau(P) < 1$ *for all ergodicity coefficients* τ, *or*
(B) *for each real number* $K \geqslant 1$ *there exists an ergodicity coefficient* τ *with* $\tau(P) \geqslant K$.

Each of the following properties is equivalent to (*A*):

(A$_1$) $zP \in \mathrm{Lin}(\{z\})$ $(z \in H)$.

(A$_2$) *There is a real number* λ $(0 < |\lambda| < 1)$ *such that* $zP = \lambda z$ $(z \in H)$.

(A$_3$) *There is a real number* λ $(0 < |\lambda| < 1)$ *such that* $\tau(P) = |\lambda|$ *for all ergodicity coefficients* τ.

(A$_4$) *There exists a* $\lambda \in \mathbb{R}$ *such that*

$$P = Q + \lambda I,$$

where Q *is a nonnegative row-constant matrix.*

The following property is equivalent to (B):

(B$_1$) *There exists a* $z_0 \in H$ *with* $z_0 P \notin \mathrm{Lin}(\{z_0\})$.

Proof. Clearly either (A$_1$) or (B$_1$) is true.

I. Let (B$_1$) be true. For any $K \geqslant 1$ there is a norm $\|\cdot\|$ on H fulfilling $\|z_0\| = 1$, $\|z_0 P\| = K$ such that

$$\tau_{\|\cdot\|}(P) = \max\{\|xP\| : \|x\| = 1,\ x \in H\} \geqslant K.$$

II. Assuming (A$_1$), we have for each $z \in H$ a real number λ_z with $zP = \lambda_z z$. Let $0 \neq z_0 \in H$, and let $z \in H$ be linearly independent of z_0. Then $zP = \lambda_z z$, $z_0 P = \lambda_{z_0} z_0$, and $(z + z_0)P = \lambda_{z+z_0}(z + z_0)$ yield $\lambda_z = \lambda_{z_0}$. The ergodicity of P implies $|\lambda_{z_0}| < 1$. Thus, we have (A$_2$), (A$_3$).

The implication (A$_3$) $\to$ (A$_1$) holds because otherwise (B) would be true. Obviously (A$_4$) $\to$ (A$_2$). Let (A$_2$) be true; then we have $z(P - \lambda I) = 0$ $(z \in H)$, and since the vectors $(\delta_{ki} - \delta_{kj})_{k=1,\dots,n}$ belong to H $(i \neq j,\ i, j = 1, 2, \dots, n)$, the matrix $P - \lambda I$ is row-constant. $P - \lambda I$ is nonnegative because $P \in S_n$. ∎

COROLLARY 3. *Let* $n \geqslant 3$. *If* $P \in S_n$ *is a nonstable ergodic matrix with property* (B$_1$), *then there are ergodicity coefficients* τ_1, τ_2 *with* $\tau_1(P) < 1$ *and* $\tau_2(P) > 1$.

3. THE WEAK ERGODICITY OF THE INFINITE PRODUCTS CORRESPONDING TO $(Q_k)_{l=0}^{\infty}$

The following theorem is a generalization of Theorem A in [7] and needs only the Dobrushin ergodicity coefficient; part I is known and simple, but necessary to obtain a complete generalization of Theorem A.

THEOREM 3.

I. *If there exists at least one* $P \in W$ *with* $\delta(P) < 1$, *then weak ergodicity obtains for the infinite products of* $(Q_k)_{k=0}^{\infty}$.

II. *If there are a natural number* $l \geqslant 1$ *and a* $K < 1$ *such that*

$$\delta(P_1 P_2 \cdots P_l) \leqslant K$$

for all l-tuples $(P_1, P_2, \ldots, P_l) \in W^l$, *then weak ergodicity obtains for all infinite products of* $(Q_k)_{k=0}^{\infty}$.

Proof. I: Let $\tau(P) < \beta < 1$ and $P = \lim_{j\to\infty} Q_{k_j}$ for the subsequence (Q_{k_j}) of (Q_k). Then there exists a j_0 with

$$\delta(Q_{k_j}) < \beta \qquad \text{for all} \quad j > j_0.$$

For any product sequences (H_m) built by (Q_k) there is an m_0 such that the matrices $Q_{k_{j_0+1}}, \ldots, Q_{k_{j_0+r}}$ are factors of the product H_{m_0}. Therefore

$$\delta(H_m) < \beta^r \qquad \text{for all} \quad m \geqslant m_0,$$

which implies $\lim_{m\to\infty} \delta(H_m) = 0$.

II: The functional f on S_n^l defined by

$$f(P_1, P_2, \ldots, P_l) = \delta(P_1 P_2 \cdots P_l) \text{ for } (P_1, P_2, \ldots, P_l) \in S_n^l$$

is equicontinuous on the compact set S_n^l. Let $\varepsilon > 0$ and $K + \varepsilon < 1$. Then there exists a $\lambda > 0$ such that for all $(P_1, P_2, \ldots, P_l) \in W^l$ the following assertion holds: For any $(R_1, R_2, \ldots, R_l) \in S_n^l$ fulfilling

$$R_i \in U_\lambda(P_i), \qquad i = 1, 2, \ldots, l,$$

where $U_\lambda(P_i) = \{Q \in S_n : d(Q, P_i) < \lambda\}$ (d is the euclidean metric on $\mathbb{R}^{n^2}$), we have

$$\delta(R_1 R_2 \cdots R_l) \leqslant K + \varepsilon.$$

Since W is a compact subset of S_n, there exists a finite set $\{T_1, T_2, \ldots, T_h\} \subseteq W$ with

$$W \subseteq \bigcup_{i=1}^{h} U_\lambda(T_i).$$

Now, let us consider the sequence $(Q_k)_{k=0}^{\infty}$. There exists a k_0 with

$$Q_k \in \bigcup_{i=1}^{h} U_\lambda(T_i) \qquad \text{for} \quad k > k_0,$$

because otherwise a convergent subsequence (Q_{k_j}) of (Q_k), $\lim_{j\to\infty} Q_{k_j} = Q^*$, would exist such that $Q_{k_j} \notin \bigcup_{i=1}^{h} U_\lambda(T_i)$ for all j and therefore $Q^* \notin W$.

Finally, let (H_m) be an arbitrary sequence of products constructed from $(Q_k)_{k=1}^{\infty}$. There exists an m_0 such that the matrices $Q_0, Q_1, \dots, Q_{k_0}$ are factors of the products H_{m_0}. Then

$$\delta\big(H_{m_0+rl+s}\big) \leqslant \delta\big(H_{m_0}\big)\cdot (K+\varepsilon)^{r-1}$$

for all $r \geqslant 2$, $s = 0, 1, 2, \dots, l-1$, which implies $\lim_{m\to\infty} \delta(H_m) = 0$. ■

REMARK 2. Part I of Leizarowitz's Theorem A follows directly from our Theorem 3 (part I) and Corollary 1. Much more, there exists a norm $\|\cdot\|$ on $\mathbb{R}^n$ with $\tau_{\|\cdot\|}(Q) < 1/C_{\|\cdot\|}$ if and only if Q is scrambling.

REMARK 3. Part II of Leizarowitz's Theorem A follows from our Theorem 3, part II. If we have $\tau_{\|\cdot\|}(P) < 1$ for all $P \in W$, then $\tau_{\|\cdot\|}(P) \leqslant \beta < 1$ for all $P \in W$ with some β. Theorem 1 yields

$$\delta(P_1 P_2 \cdots P_k) \leqslant C_{\|\cdot\|}\,\beta^k$$

for every k-tuple $(P_1, P_2, \dots, P_k) \in W^k$, and

$$\delta(P_1 P_2 \cdots P_l) \leqslant C_{\|\cdot\|}\,\beta^l < 1$$

in the case $l > -(\log C_{\|\cdot\|})/\log\beta$.

A set $V \subset S_n$ is said to have the left convergence property (LCP) if all backward products built by members of V converge (see [2, p. 230]). We say that set $V \subset S_n$ has the finite product ergodicity (FPE) property if all finite products of members of V are ergodic matrices.

Special classes of sets $V \subset S_n$ having the FPE property are the Sarymsakow class [16] and a class described by Anthonisse and Tijms [1]. In both cases all products of at least $n-1$ members of V are scrambling matrices.

REMARK 4 (Compare [2]). *If $V \subset S_n$ is a finite set, then the following properties are equivalent*:

(1) *V has the FPE property.*

(2) *There are a natural number l and a $K < 1$ such that $\delta(P_1 P_2 \cdots P_l) \leqslant K$ for all $(P_1, P_2, \dots, P_l) \in V^l$.*

(3) *The products $P_1 P_2 \cdots P_l$ of arbitrary matrices of V are scrambling for $l \geqslant \frac{1}{2}(3^n - 2^{n+1} - 1)$.*

(4) *V has the LCP property, and 1 is a simple eigenvalue of each $P \in V$.*

Proof. The equivalence of the first three assertions is well known (see e.g. [1]). (2) yields the weak ergodicity of all infinite products built by V, which implies the strong ergodicity of the backward products. We obtain therefore the LCP property of V and (4), because each $P \in V$ is ergodic. On the other hand, (4) implies that each sequence $(P_k P_{k-1} \cdots P_1)^m$ tends to a stable matrix as $m \to \infty$, so that (1) follows. ∎

COROLLARY 4. *Let the set W of all accumulation points of Q_k have the* FPE *property and*

$$\min_{i,j}{}^{+} (P)_{ij} \geqslant \lambda > 0 \qquad (P \in W),$$

where min$^+$ *means the minimum over all positive elements. Then weak ergodicity obtains for all infinite products of (Q_k).*

Proof. The question whether or not a $P \in S_n$ is ergodic only depends of the distribution of the zero in P. The set of different patterns corresponding to ergodic matrices is therefore finite. By Remark 4 and the FPE property of W there exists a natural number l such that the products $P_1 P_2 \cdots P_l$ are scrambling for all $(P_1, P_2, \dots, P_l) \in W^l$. This property, the second assumption, and the explicit form of $\delta(P)$ yield easy that

$$\delta(P_1 \cdot P_2 \cdot \cdots \cdot P_l) \leqslant 1 - \lambda^l \qquad \left[(P_1, P_2, \dots, P_l) \in W^l\right].$$

Theorem 3 implies the assertion. ∎

REMARK 5. Corollary 4 is a slight generalization of the well-known theorem (see [17, p. 142]) which shows that the FPE property of $\{Q_k : k = 1, \dots\}$ and $\min^+_{i,j}(Q_k)_{ij} \geqslant \lambda > 0$ $(k = 1, \dots)$ imply the weak ergodicity of all infinite products of (Q_k).

For the simple sequence (Q_k) defined by

$$Q_{2h-2} = \begin{pmatrix} 1 & 0 \\ 1 - \frac{1}{h} & \frac{1}{h} \end{pmatrix}, \quad Q_{2h-3} = \begin{pmatrix} 0 & 1 \\ \frac{1}{h} & 1 - \frac{1}{h} \end{pmatrix}, \qquad h = 2, 3, \dots,$$

we have

$$W = \left\{ \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \right\}$$

and $\min^+_{i,j} (P)_{ij} = 1 > 0$ $(P \in W)$, but $\inf_{Q_k} \min^+_{i,j} (Q_k)_{ij} = 0$.

The author is grateful to the referee, whose comments led to an improvement of the first version of this paper.

REFERENCES

1 J. M. Anthonisse and H. Tijms, Exponential convergence of products of stochastic matrices, *J. Math. Anal. Appl.* 598:360–364 (1977).

2 I. Daubechies and J. C. Lagarias, Sets of matrices all infinite products of which converge, *Linear Algebra Appl.* 161:227–263 (1992).

3 R. L. Dobrushin, Central limit theorem for non-stationary Markov chains I, II, *Theorem Probab. Appl.* I:63–80, 329–383 (1956).

4 J. Hajnal, Weak ergodicity in nonhomogeneous Markov-chains, *Proc. Cambridge Philos. Soc.* 54:233–246 (1958).

5 R. A. Horn and Ch. A. Johnson, *Matrix Analysis*, Cambridge U.P., New York, 1985.

6 R. Kühne and A. Rhodius, A characterization of Dobrushin's coefficient of ergodicity, *Z. Anal. Anwendungen* 9:187–188 (1990).

7 A. Leizarowitz, On infinite products of stochastic matrices, *Linear Algebra Appl.* 168:189–219 (1992).

8 A. Lesanovsky, Coefficients of ergodicity generated by non-symmetrical vector norms, *Czechoslovak Math. J.* 40:284–294 (1990).

9 J. L. Mott, Conditions for the ergodicity of non-homogeneous finite Markov chains, *Proc. Roy. Soc. Edinburgh Sect.* A 64:369–380 (1957).

10 A. Paz, *Introduction to Probabilistic Automata*, Academic, New York, 1971.

11 A. Rhodius, The maximal value for coefficients or ergodicity, *Stochastic Process. Appl.* 29:141–145 (1988).

12 A. Rhodius, Über explizite Formeln für Ergodizitätskoeffizienten, *Wiss. Z. Pädagog. Hochsch. "Karl Friedrich Wilhelm Wander" Dresden* 23:69–74 (1989).

13 A. Rhodius, On explicit forms for ergodicity coefficients, *Linear Algebra Appl.* 194:71–83 (1993).

14 U. G. Rothblum and C. P. Tan, Upper bounds on the maximum modulus of subdominant eigenvalues of nonnegative matrices, *Linear Algebra Appl.* 66:45–86 (1985).
15 T. A. Sarymsakov, Inhomogeneous Markov chains (in Russian), *Teor. Veroyatnost. i Primenen.* 6:194–201 (1961).
16 E. Seneta, Coefficients of ergodicity: Structure and applications, *Adv. Appl. Probab.* 11:576–590 (1979).
17 E. Seneta, *Nonnegative Matrices and Markov Chains*, Springer-Verlag, New York, 1981.
18 E. Seneta, Explicit forms for ergodicity coefficients of stochastic matrices, *Linear Algebra Appl.* 191:245–252 (1993).
19 C. P. Tan, A functional form for a particular coefficient of ergodicity, *J. Appl. Probab.* 19:858–863 (1982).

Received 21 February 1995; final manuscript accepted 25 September 1995