Zeitschrift für Analysis und ihre Anwendungen
Journal for Analysis and its Applications
Volume 19 (2000), No. 3, 873–887

On Ergodicity Coefficients of Infinite Stochastic Matrices

A. Rhodius

To Prof. T. Riedrich with appreciation on the occasion of his 65. birthday

Abstract. A class of ergodicity coefficients for infinite stochastic matrices is introduced and investigated with respect to connections to the well-known δ-coefficient. The theory yields results on the behaviour of infinite products of stochastic matrices, in particular on inhomogeneous Markov chains and Markov systems.

Keywords: *Extreme points, spectrum, product of matrices, weak ergodicity*

AMS subject classification: Primary 47 A 30, secondary 60 J 10

1. Introduction

Let $P = (p_{ij})_{i,j\in\mathbb{N}}$ be a stochastic matrix, S the set of all stochastic matrices P and l_1 the usual vector space of real sequences $x = (x_i)_{i\in\mathbb{N}}$ for which $\|x\|_1 = \sum_{i=1}^{\infty} |x_i| < \infty$. The set

$$H = \left\{ x \in l_1 : \sum_{i=1}^{\infty} x_i = 0 \right\}$$

is a closed subspace of $(l_1, \| \ \|_1)$. We denote the unit sphere of $(H, \| \ \|_1)$ by $S^{(1)}$. Every $P \in S$ defines a linear operator $T : l_1 \to l_1$ by $Tx = xP \ \ (x \in l_1)$ with

$$\|T\|_1 = \sup \left\{ \|xP\|_1 : x \in l_1 \text{ with } \|x\|_1 = 1 \right\} = 1$$

and H is an invariant subspace of T. An important tool for the investigation of inhomogeneous Markov chains is the so-called *δ-coefficient* defined by

$$\delta(P) = \frac{1}{2} \sup_{i,j} \sum_{l=1}^{\infty} |p_{il} - p_{jl}|.$$

It is well known that

$$\delta(P) = \sup \left\{ \|xP\|_1 : x \in H \text{ with } \|x\|_1 = 1 \right\} = \sup \left\{ \|xP\|_1 : x \in S^{(1)} \right\}.$$

The stochastic matrix $P \in S$ is said to be

A. Rhodius: FB Math. der TU, D-01062 Dresden; rhodius@math.tu-dresden.de

ISSN 0232-2064 / $ 2.50

scrambling if $\delta(P) < 1$

stable if $\tau(P) = 0$.

Let $\sigma(P)$ denote the spectrum of the operator T corresponding to P. In the cases of finite spaces one uses other norms, and there is a lot of papers about ergodicity coefficients with respect to arbitrary norms on $\mathbb{R}^n$ (see [5, 7 - 13, 16 - 19]).

Our purpose is to construct a theory of ergodicity coefficients for denumerable state spaces, this means for infinite stochastic matrices $P \in S$ and the set $\mathfrak{A}$ of all norms on l_1 which are equivalent to $\| \|_1$.

Definition 1. For any norm $\| \| \in \mathfrak{A}$ we call $\tau_{\| \|}(P) = \sup\{\|xP\| : x \in H \text{ with } \|x\| = 1\}$ the *ergodicity coefficient* for $P \in S$ with respect to the norm $\| \|$. Moreover, we put $C_{\| \|} = \sup\{\tau_{\| \|}(P) : P \in S\}$.

The ergodicity coefficient $\tau_{\| \|}(P)$ is the operator norm of the restricted operator $T|H$ and fulfils the following conditions:

(a) $\tau_{\| \|}(PQ) \le \tau_{\| \|}(P)\,\tau_{\| \|}(Q)$ for $P, Q \in S$.

(b) $\tau_{\| \|}(P) = 0$ if and only if P is stable (see Theorem 1).

(c) $|\lambda| \le \tau_{\| \|}(P)$ for $\lambda \in \sigma(P) \setminus \{1\}$.

One aim of the paper is to compare different ergodicity coefficients, in particular arbitrary ergodicity coefficients and the coefficient δ. Answers to these problems are derived using the set $\mathrm{Extr}B^{(1)}$ of all extreme points of the ball $B^{(1)} = \{x \in H : \|x\|_1 \le 1\}$, a theorem of H. Schneider and W. G. Strang [15] on the quotient of norms of a linear operator and properties of $C_{\| \|} = \sup\{\tau_{\| \|}(P) : P \in S\}$. Moreover, for a special class of norms $\| \| \in \mathfrak{A}$ the set $\mathrm{Extr}\,\{x \in H : \|x\| \le 1\}$ and explicit functional forms of $\tau_{\| \|}$ will be determined.

The theory yields new results on the weak ergodicity of inhomogeneous Markov chains, the behaviour of general infinite products built by sequences of stochastic matrices as well as Markov systems.

2. Ergodicity coefficients and bounds for spectral values

In the sequel $\sigma(P)$ means the spectrum $\sigma(T_{\mathbb{C}})$ of the complexification $T_{\mathbb{C}}$ if $T : l_1 \to l_1$ where $Tx = xP$ $(x \in l_1)$. We denote

$$l_{1,\mathbb{C}} = \big\{x + iy : x, y \in l_1\big\}$$

and put

$$\|x + iy\|_{\mathbb{C}} = \sup_{0 \le \Theta \le 2\pi} \|x \cos\Theta + y \sin\Theta\|$$

for $x + iy \in l_{1,\mathbb{C}}$. The set

$$H_{\mathbb{C}} = \big\{x + iy : x, y \in H\big\}$$

is a closed subspace of $(l_{1,\mathbb{C}}, \| \|_{\mathbb{C}})$. It is well known that for the complexification $T_{\mathbb{C}}$ of T where $T_{\mathbb{C}}(x + iy) = Tx + iTy$ $(x, y \in l_1)$ the relations

$$\|T_{\mathbb{C}}\| = \sup\Big\{\|T_{\mathbb{C}} z\|_{\mathbb{C}} : z \in l_{1,\mathbb{C}} \text{ with } \|z\|_{\mathbb{C}} = 1\Big\} = \|T\| = 1$$

$$\|T_{\mathbb{C}}|H_{\mathbb{C}}\| = \sup\left\{\|T_{\mathbb{C}}z\|_{\mathbb{C}} : z \in H_{\mathbb{C}} \text{ with } \|z\|_{\mathbb{C}} = 1\right\} = \|T|H\| = \tau_{\|\ \|}(P)$$

hold.

Lemma 1. *Let $\sigma_p(T_{\mathbb{C}})$ be the point spectrum of $T_{\mathbb{C}}$. Then:*

(1a) $\sigma_p(T_{\mathbb{C}}) \setminus \{1\} \subset \sigma_p(T_{\mathbb{C}}|H_{\mathbb{C}})$.

(1b) $\sigma(T_{\mathbb{C}}) \setminus (\sigma_p(T_{\mathbb{C}}) \cup \{1\}) \subset \sigma(T_{\mathbb{C}}|H_{\mathbb{C}}) \setminus \sigma_p(T_{\mathbb{C}}|H_{\mathbb{C}})$.

(2) $\sigma(T_{\mathbb{C}}|H_{\mathbb{C}}) \cup \{1\} = \sigma(T_{\mathbb{C}})$.

Proof.

(1a) If $\alpha \in \sigma_p(T_{\mathbb{C}}) \setminus \{1\}$, then there is a $z \in l_{1,\mathbb{C}}$ with $z \neq 0$ and $\alpha z = T_{\mathbb{C}} z$, that means $\alpha z = zP$. Therefore $\alpha z 1^T = zP1^T = \sum z_j$ and $(1-\alpha)\sum z_l = 0$ follow. Thus $z \in H_{\mathbb{C}}$.

(1b) In the case $\alpha \in \sigma(T_{\mathbb{C}})$ but $\alpha \notin \sigma_p(T_{\mathbb{C}})$ and $\alpha \neq 1$ we have $(T_C - \alpha I_{\mathbb{C}})l_{1,\mathbb{C}} \neq l_{1,\mathbb{C}_i}$. Let $\alpha \notin \sigma(T_{\mathbb{C}}|H_{\mathbb{C}}) \setminus \sigma_p(T_{\mathbb{C}}|H_{\mathbb{C}})$. Then $(T_{\mathbb{C}} - \alpha I_{\mathbb{C}})H_{\mathbb{C}} = H_{\mathbb{C}}$ follows, and

$$(1,0,0,\ldots)(P - \alpha I) = (p_{11} - \alpha, p_{12}, \ldots) \notin H_{\mathbb{C}}$$

would yield the contradiction $(T_{\mathbb{C}} - \alpha T_{\mathbb{C}})l_{1,\mathbb{C}} = l_{1,\mathbb{C}}$

(2) Because of (1a), (1b) we have $\sigma(T_{\mathbb{C}}) \subset \sigma(T_{\mathbb{C}}/H_{\mathbb{C}}) \cup \{1\}$.

Let be $z \in l_{1,\mathbb{C}}$. Then

$$(T_{\mathbb{C}} - I_{\mathbb{C}})z = zP - z = \left(\sum_l z_l p_{lj} - z_j\right)_{j\in\mathbb{N}}$$

and, because of $\sum_j(\sum_l z_l p_{lj} - z_j) = 0$, $(T_{\mathbb{C}} - I_{\mathbb{C}})z \in H_{\mathbb{C}}$, that means $(T_{\mathbb{C}} - I_{\mathbb{C}})l_{1,\mathbb{C}} \neq l_{1,\mathbb{C}}$ and $1 \in \sigma(T_{\mathbb{C}})$.

Let us assume $1 \neq \alpha \in \sigma(T_{\mathbb{C}}|H_{\mathbb{C}})$ but $\alpha \notin \sigma(T_{\mathbb{C}})$. As α can not be an eigenvalue of $T_{\mathbb{C}}|H_{\mathbb{C}}$ it follows $(T_{\mathbb{C}} - \alpha I_{\mathbb{C}})H_{\mathbb{C}} \neq H_{\mathbb{C}}$. Since $(T_{\mathbb{C}} - \alpha I_{\mathbb{C}})l_{1,\mathbb{C}} = l_{1,\mathbb{C}}$, there are $z_1 \in H_{\mathbb{C}}$ and $z_2 \in l_{1,\mathbb{C}}$ with $z_2 \notin H_{\mathbb{C}}$ and $(T_{\mathbb{C}} - \alpha I_{\mathbb{C}})z_2 = z_1$. Therefore

$$\sum_j \left(\sum_l z_l^{(2)} p_{lj} - \alpha z_j^{(2)}\right) = \sum_j z_j^{(1)} = 0$$

and $(1-\alpha)\sum_j z_j^{(2)} = 0$ follows, contradicting $\alpha \neq 1$ ∎

Relation (2) of Lemma 1 and the formula for the spectral radius applied to $T_{\mathbb{C}}/H_{\mathbb{C}}$ yield the following spectral estimation.

Theorem 1. *Let be $\|\ \| \in \mathfrak{A}$ and $P \in S$. Then the estimation*

$$|\lambda| \leq [\tau_{\|\ \|}(P^n)]^{\frac{1}{n}} \qquad (\lambda \in \sigma(P) \setminus \{1\})$$

holds for $n \in \mathbb{N}$.

3. Extreme points and explicit functional forms of ergodicity coefficients

In this section we use l_1-vectors

$$e_i = (\delta_{ij})_{j\in\mathbb{N}} \qquad \text{and} \qquad e_{ij} = \frac{1}{2}(e_i - e_j) \quad (i,j \in \mathbb{N}).$$

Let $a = (a_j)_{j\in\mathbb{N}}$ be a real bounded sequence with $0 < \inf_{j\in\mathbb{N}} a_j$. Further, let be

$$f_{ij}^{(a)} = \frac{2}{a_i + a_j} e_{ij} \quad (i,j \in \mathbb{N}) \qquad \text{and} \qquad \|x\|_a = \sum_{i=1}^{\infty} a_i |x_i| \quad (x \in l_1).$$

Moreover, we need the ball and sphere

$$B^{(a)} = \{x \in H : \|x\|_a \le 1\}$$
$$S^{(a)} = \{x \in H : \|x\|_a = 1\},$$

respectively.

Lemma 2. $\mathrm{Extr} B^{(a)} = \{ f_{ij}^{(a)} : i,j \in \mathbb{N} \text{ with } i \neq j \}$.

Proof. 1. Let $f_{ij}^{(a)} = \frac{1}{2}x + \frac{1}{2}y$ for some $x, y \in B^{(a)}$. Clearly, $\|f_{ij}^{(a)}\|_a = \|x\|_a = \|y\|_a = 1$. Because of $\|f_{ij}^{(a)}\|_a = 1$ and

$$\frac{1}{a_i + a_j} = \frac{1}{2}(x_i + y_i) = -\frac{1}{2}(x_j + y_j) \tag{1}$$

it follows $\frac{1}{2}(x_k + y_k) = 0$ and therefore $x_k = -y_k$ for $k \neq i,j$. Defining

$$\left.\begin{aligned} \tilde{x} &= x - x_i e_i - x_j e_j \\ \tilde{y} &= x - y_i e_i - y_i e_j \end{aligned}\right\}$$

we have

$$\|\tilde{x}\|_a = \sum_{k\neq i,j} a_k |x_k| = \|\tilde{y}\|_a < 1.$$

Then

$$1 = \|f_{ij}^{(a)}\|_a = \frac{1}{2} a_i |x_i + y_i| + \frac{1}{2} a_j |x_j + y_j| \le \frac{1}{2}(1 - \|\tilde{x}\|_a) + \frac{1}{2}(1 - \|\tilde{y}\|_a)$$

implies $\|\tilde{x}\|_a = \|\tilde{y}\|_a = 0$. Therefore $x_k = y_k = 0 \quad (k \neq i,j)$, $x_j = -x_i$ and $y_j = -y_i$ so that $|x_i| = |x_j| = \frac{1}{a_i + a_j}$ and $|y_i| = |y_j| = \frac{1}{a_i + a_j}$ follow. Thus, with (1) we obtain $x = y = f_{ij}^{(a)}$, that means $f_{ij}^{(a)}$ is an extreme point of $B^{(a)}$.

2. If we suppose $x \in S^{(a)}$ and $x \neq f_{ij}^{(a)}$ for all $i,j \in \mathbb{N}$ with $i \neq j$, then there are $i,j,k \in \mathbb{N}$ with $x_i, x_j, x_k \neq 0$. We may assume $i = 1, j = 2, k = 3$ and $x_{1,2} > 0$. There

is a $\delta_1 \neq 0$ such that $x_1 \pm \delta_1 > 0$, $x_2 \pm \delta_2 > 0$ and $x_2 \pm \delta_3 < 0$ hold for $\delta_2 = -\frac{a_1+a_3}{a_2+a_3}\delta_1$ and $\delta_3 = \frac{a_1-a_2}{a_2+a_3}\delta_1$. Let be

$$\left.\begin{array}{l} y = x - (\delta_1, \delta_2, \delta_3, 0, \ldots) \\ z = x + (\delta_1, \delta_2, \delta_3, 0, \ldots) \end{array}\right\}.$$

Then

$$\begin{aligned} \|y\|_a &= a_1x_1 + a_2x_2 - a_3x_3 + \cdots - (a_1\delta_1 + a_2\delta_2 - a_3\delta_3) = \|x\|_a = 1 \\ \|z\|_a &= a_1x_1 + a_2x_2 - a_3x_3 + \cdots + (a_1\delta_1 + a_2\delta_2 - a_3\delta_3) = \|x\|_a = 1 \end{aligned}$$

since $a_1\delta_1 + a_2\delta_2 - a_3\delta_3 = 0$. Because of $\delta_1 + \delta_2 + \delta_3 = 0$ it follows $y, z \in H$. Thus $x = \frac{1}{2}y + \frac{1}{2}z$ with $y, z \in S^{(a)}$, that means x is no extreme point of $B^{(a)}$ ∎

Lemma 3. *Let $x \in H$. Then:*

(a) *The series $\sum_{i,j} x_i^+ x_j^- e_{ij}$ is absolutely convergent in l_1.*

(b) $\|x\|_1 x = 4\sum_{i,j} x_i^+ x_j^- e_{ij}$.

(c) $x = (\sum_i x_i^+)^{-2} \cdot \sum_{i,j} x_i^+ x_j^- e_{ij}$ *for $x \in S^{(1)}$.*

Proof. (a) Clearly, $\sum x_i^+ = \sum x_j^- = \frac{1}{2}\|x\|_1$ and so

$$\sum_{i,j} \|x_i^+ x_j^- e_{ij}\|_1 = \sum_{ij} x_i^+ x_j^- = \sum_i x_i^+ \sum_j x_j^- = \frac{1}{4}\|x\|_1^2.$$

(b) We put $x = (x_1, x_2, \ldots)$ and $y = 4\sum_{i,j} x_i^+ x_j^- e_{ij}$. Statement (a) yields $y \in H$. Then

$$y_k = 4\Bigg(\sum_i x_i^+ x_k^- e_{ik} + \sum_j x_k^+ x_j^- e_{kj}\Bigg)_k.$$

For every k with $x_k \geq 0$ we obtain

$$y_k = 4\Bigg(\sum_j x_k^+ x_j^- e_{kj}\Bigg)_k = 4 \cdot \frac{1}{2} x_k^+ \sum_j x_j^- = x_k \|x\|_1,$$

and if $x_k < 0$, then

$$y_k = 4\Bigg(\sum_i x_i^+ x_k^- e_{ik}\Bigg)_k = 4 \cdot \Big(-\frac{1}{2}\Big) x_k^- \sum x_i^+ = -x_k^- \|x\|_1 = x_k \|x\|_1.$$

(c) Statement (c) immediately follows from (b) ∎

Theorem 2. *Let $x \in H$. Then:*

(a$_1$) $(\sum_i x_i^+) x = \sum_{i,j} x_i^+ x_j^- (a_i + a_j) f_{ij}^{(a)}$.

(a$_2$) $(\sum_i x_i^+) \|x\|_a = \sum_{i,j} x_i^+ x_j^- (a_i + a_j)$.

(b) $B^{(a)} = \overline{\text{conv}}\{f_{ij}^{(a)} : i, j \in \mathbb{N}$ *with* $i \neq j\}$.

Proof. (a) Since Lemma 3/(b) directly yields (a_1) it is only to show (a_2):

$$\begin{aligned}\sum_{i,j} x_i^+ x_j (a_i + a_j) &= \sum_i x_i^+ \left(a_i \sum_j x_j^- + \sum_j a_j x_j^- \right) \\ &= \sum_j x_j^- \cdot \sum_i a_i x_i^+ + \sum_i x_i^+ \sum_j a_j x_j^- \\ &= \sum_j x_j^+ \|x\|_a.\end{aligned}$$

(b) As $f_{ij}^{(a)} \in B^{(a)}$ we have only to show that

$$S^{(a)} \subset \overline{\text{conv}}\{ f_{ij}^{(a)} : i,j \in \mathbb{N} \text{ with } i \neq j \}.$$

Because of (a_1) we have

$$x = \frac{1}{\sum_{k=1}^{\infty} x_k^+} \sum_{i,j=1}^{\infty} x_i^+ x_j^- (a_i + a_j) f_{ij}^{(a)}$$

for $x \in S^{(a)}$. Defining

$$\begin{aligned}b^{(n)} &= \sum_{k,l=1}^{n} x_k^+ x_l^- (a_k + a_l) \\ y^{(n)} &= \sum_{i,j=1}^{n} x_i^+ x_j^- (a_i + a_j) f_{ij}^{(a)} \\ x^{(n)} &= \frac{1}{b^{(n)}} y^{(n)}\end{aligned}$$

for large n ($b^{(n)} \neq 0$) then $x^{(n)}$ are convex combinations of elements $f_{ij}^{(a)}$ and belong to the ball $B^{(a)}$ of H. Relations (a_2) and (a_1) imply $\lim_{n\to\infty} b^{(n)} = \sum_{i=1}^{\infty} x_i^+$ and $y^{(n)} \to (\sum_{i=1}^{\infty} x_i^+)x$ in l^1 such that $x^{(n)} \to x$ follows. Thus, we have $S^{(a)} \subset \overline{\text{conv}}\{f_{ij}^{(a)} : i,j \in \mathbb{N} \text{ with } i \neq j\}$ ∎

Theorem 3. *Let $P \in S$. Then*

$$\tau_{\|\ \|_a}(P) = \sup_{i,j} \|f_{ij}^{(a)} P\|_a = \sup_{i,j} \frac{1}{a_i + a_j} \sum_k a_k |p_{ik} - p_{jk}|.$$

Proof. This theorem holds since the functional $x \mapsto \|xP\|_a$ is convex, $B^{(a)} = \overline{\text{conv}}\{f_{ij}^{(a)} : i,j \in \mathbb{N} \text{ with } i \neq j\}$ and $f_{ij}^{(a)} = \frac{1}{a_i + a_j}(e_i - e_j)$ ∎

In investigations of Markov chains it is important to know whether for a stochastic matrix P there exists an ergodicity coefficient τ with $\tau(P) < 1$.

Example 1. We consider the stochastic matrix

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots & & & & & \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \cdots & & & & \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & \frac{1}{2} & 0 & \cdots & & & \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & \frac{1}{2} & \cdots & & & \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & \frac{1}{2} & \cdots & & \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{2} & \cdots & \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & \frac{1}{2} & \cdots \end{bmatrix}.$$

It is easy to see that $\delta(P) = 1$. On the other hand we obtain $\tau_{\| \|_a}(P) = \frac{35}{36} < 1$ for the sequence $a = (a_j)$ with $a_1 = 2, a_2 = 6$ and $a_j = 7$ for $j = 3, 4, 5 \dots$

Example 2. Let K be a set of infinite stochastic matrices $Q = (q_{ij})$ only consisting of elements zero and one with the following properties:

(1) $q_{11} = 1$.

(2) $q_{ij} = 0$ for alle $(i, j) \in N \times \mathbb{N}$, with $i \neq 1$ and $j \geq i$.

(3) There exists a natural number $n_Q \geq 2$ with $q_{ij} = 0$ for $i \in \mathbb{N}$ and $j \geq n_Q$.

It is easy to see that for every strictly increasing sequence $a = (a_j)$ we have $\tau_{\| \|}(Q) < 1$ for all $Q \in K$. On the other hand, there is a lot of $P \in K$ with $\delta(P) = 1$, which demonstrates the practical significance of other ergodicity coefficients than the δ-coefficient.

4. Sup$\{\tau_{\| \|}(P) : P \in S\}$ and the comparison of ergodicity coefficients

We start with some properties of linear operators acting on H and their operator norms.

Lemma 4. *Let $(H, \| \|)$ be a Banach space. Then the following assertions are equivalent:*

(a) $\sup_{P \in S} \sup \{\|xP\| : x \in H$ *with* $\|x\| = 1\} < \infty$.

(b) *The norm $\| \|$ is equivalent to $\| \|_1$ on H.*

Proof. We have only to prove that (a) implies (b). Let $x = (x_i)_1^\infty$ be a point of $H \setminus \{0\}$. Then there are $i, j \in \mathbb{N}$ with $x_i > 0$ and $x_j < 0$. The stochastic matrix $P^{(x)} = (p_{ij}^x)$ defined by $p_{i1}^x = 1$ for $x_i \geq 0$ and $p_{i2}^x = 1$ for $x_i < 0$ fulfils

$$xP^{(x)} = \Big(\sum x_i^+, -\sum x_i^-, 0, 0, 0, \dots\Big).$$

Because of $\|x\|_1 = 2 \sum x_i^+ = 2 \sum x_i^-$ one has

$$\|xP^{(x)}\| = \frac{1}{2}\|x\|_1 \cdot \|(1, -1, 0, 0, 0, \dots)\| = \frac{1}{2}\|x\|_1 \alpha$$

with $\alpha > 0$ such that

$$\frac{\alpha}{2}\|x\|_1 = \|xP^{(x)}\| \leq \sup_{P \in S} \tau_{\| \|}(P)\|x\|$$

holds for all $x \in H$. Assumption (a) and the completeness of $(H, \| \|)$ and $(H, \| \|_1)$ yield assertion (b) ∎

Lemma 5. *Let $L(H)$ be the space of bounded linear operators acting on H. Let be $A \in L(H)$. If A is an operator of rank 1, then there are a stochastic matrix $P \in S$ and $k > 0$ such that $(Ax) = k(xP)$ for all $x \in H$.*

Proof. Let $A \in L(H)$. For every $x \in l_1$ we put

$$\alpha_x = \sum x_i^+ - \sum x_i^- \qquad \text{and} \qquad h_x = x - \alpha_x(1,0,0,\ldots).$$

Then $x = h_x + \alpha_x(1,0,0,\ldots)$ with $h_x \in H$. The operator $\tilde{A}$ defined by

$$\tilde{A}x = Ah_x + \alpha_x(1,0,\ldots) \qquad (x \in l_1)$$

is a linear mapping from l_1 into l_1. Therefore there exists $(a_{ij})_{i,j\in\mathbb{N}}$ with $\sup_i \sum_{j=1}^{\infty} |a_{ij}| < \infty$ such that $\tilde{A}x = x(a_{ij})_{i,j\in\mathbb{N}}$ for $x \in l_1$. In particular, $Ah = h(a_{ij})$ for $h \in H$, and because of $e_l - e_m \in H \ \ (l, m \in \mathbb{N})$ and $A(H) \subset H$ the sums of the rows of (a_{ij}) are constant. Moreover, two matrices $(a_{ij}^{(1)})$ and $(a_{ij}^{(2)})$ represent the same operator A if and only if $(a_{ij}^{(1)} - a_{ij}^{(2)})$ is a stable matrix. Now let (a_{ij}) be a matrix representing an operator A of rank 1. We may assume that there are real sequences $(d_i)_{i\in\mathbb{N}}, (\lambda_j)_{j\in\mathbb{N}}$ with $\sum_{i=1}^{\infty} d_i = 0, (d_i)_{i\in\mathbb{N}} \neq (0), (\lambda_j)_{j\in\mathbb{N}} \neq (0)$ such that

$$(a_{ij}) = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ \lambda_1 d_1 & \lambda_1 d_2 & \lambda_1 d_3 & \cdots \\ \lambda_2 d_1 & \lambda_2 d_2 & \lambda_2 d_3 & \cdots \\ \lambda_3 d_1 & \lambda_3 d_2 & \lambda_3 d_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Because of $\sup_{j\in\mathbb{N}} |\lambda_j| \cdot \sum_{i=1}^{\infty} |d_i| < \infty$ the sequences $(d_i), (\lambda_j)$ are bounded. We put

$$c_i = \begin{cases} -\inf_j \lambda_j \cdot d_i & \text{if} \quad \lambda_j d_i < 0 \quad \text{for some} \quad j \in \mathbb{N} \\ 0 & \text{if} \quad \lambda_j d_i \geq 0 \quad \text{for all} \quad j \in \mathbb{N} \end{cases}$$

and have

$$\sum_{i=1}^{\infty} |c_i| \leq \sup_{j\in\mathbb{N}} |\lambda_j| \cdot \sum_{i=1}^{\infty} |d_i| < \infty.$$

The matrix $(a_{ij} + c_i)_{i,j\in\mathbb{N}}$ represents the operator A, $P = \frac{1}{k}(a_{ij} + c_i)$, where $k = \sum_{i=1}^{\infty} c_i$, is a stochastic matrix and $Ax = k(xP)$ for $x \in H$ ∎

Theorem 4 (compare [15]). *Let v_1 and v_2 be two equivalent norms on a vector space E and let $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the corresponding operator norms on $L(E)$, respectively. Moreover, let T be a subset of $L(E)$ which contains all bounded linear operators of rank 1. If one puts*

$$R_{21} = \sup_{0\neq x\in E} \frac{v_2(x)}{v_1(x)} \qquad \text{and} \qquad R_{12} = \sup_{0\neq x\in E} \frac{v_1(x)}{v_2(x)},$$

then one has

$$\sup_{0\neq A\in T} \frac{\|A\|_2}{\|A\|_1} = \sup_{0\neq A\in T} \frac{\|A\|_1}{\|A\|_2} = R_{21}R_{12}.$$

If we put $E = H$ and $v_1 = \|\ \|_1, v_2 = \|\ \| \in \mathfrak{A}$, then Theorem 4 and Lemma 5 immediately yield the following lemma.

Lemma 6. *Let* $\| \ \| \in \mathfrak{A}$. *Then*

$$\sup_{P\in S} \frac{\delta(P)}{\tau_{\| \ \|}} = \sup_{P\in S} \frac{\tau_{\| \ \|}(P)}{\delta(P)} = \sup_{0\neq x\in H} \frac{\|x\|}{\|x\|_1} \sup_{0\neq x\in H} \frac{\|x\|_1}{\|x\|}.$$

Lemma 7. *Let* $\| \ \|$ *be a norm on* H, $\| \ \| \sim \| \ \|_1$, *and* $S^{(1)} = \{x \in H : \|x\|_1 = 1\}$. *Then*

$$C_{\| \ \|} = \sup\left\{\tau_{\| \ \|}(P) : P \in S\right\} \geq \frac{\sup\{\|x\| : x \in S^{(1)}\}}{\inf\{\|x\| : x \in S^{(1)}\}}.$$

Proof. We consider $\mathrm{Extr} S$ of the convex set S as subset of the linear space $\mathbb{R}^{\mathbb{N}\times\mathbb{N}}$. $Q \in S$ is an extreme point of S if and only if Q has the number 1 as an element in each of its rows. Therefore $Q = (q_{ij}) \in S$ belongs to $\mathrm{Extr} S$ if and only if there is a denumerable system of sets $\{\wedge_n\}_{n\in\mathbb{N}}$ such that:

(i) $\mathbb{N} = \{1, 2, \ldots\} = \cup_{n\in\mathbb{N}} \wedge_n$.

(ii) $\wedge_i \cap \wedge_j = \emptyset$ for $i, j \in \mathbb{N}$ with $i \neq j$.

(iii) $q_{ij} = 1$ if and only if $i \in \wedge_j$ for $i, j \in \mathbb{N}$.

It follows

$$xQ = \left(\sum_{i\in\wedge_1} x_i, \sum_{i\in\wedge_2} x_i, \ldots\right) \qquad (x \in l_1)$$

and we obtain

$$\begin{aligned} C_{\| \ \|} &\geq \sup_{Q\in \mathrm{Extr} S} \ \sup_{x\in H, \|x\|=1} \|xQ\| \\ &\geq \sup_{x\in H, \|x\|=1} \ \sup_{(\wedge_1,\wedge_2,\ldots)\in\wedge} \left\| \left(\sum_{i\in\wedge_1} x_i, \sum_{i\in\wedge_2} x_i, \ldots\right)\right\| \end{aligned}$$

where $\wedge$ denotes the set of all decompositions $(\wedge_1, \wedge_2, \ldots)$ satisfying (i) and (ii). For every $\varepsilon > 0$ there exists an $x_\varepsilon \in H$, $\|x_\varepsilon\| = 1$ with $\|x_\varepsilon\|_1 \geq \sup_{x\in H, \|x\|=1} \|x\|_1 - \varepsilon$. Since $z_\varepsilon = \frac{x_\varepsilon}{\|x_\varepsilon\|_1} \in S^{(1)}$, we have

$$\sum (z_\varepsilon)^+ = \sum (z_\varepsilon)_i^- = \frac{1}{2}. \tag{2}$$

It follows

$$C_{\| \ \|} \geq \|x_\varepsilon\|_1 \left\| \left(\sum_{i\in\wedge_1} (z_\varepsilon)_i, \sum_{i\in\wedge_2} (z_\varepsilon)_i, \ldots\right)\right\|$$

for all decompositions $(\wedge_1, \wedge_2, \ldots) \in \wedge$.

Because of (2) we obtain $C_{\| \ \|} \geq \|x_\varepsilon\|_1 \|e_{ij}\|$ and with $\varepsilon \to 0$

$$C_{\| \ \|} \geq \sup_{x\in H, \|x\|=1} \|x\|_1 \sup_{i,j\in\mathbb{N}} \|e_{ij}\|.$$

Since $B^{(1)} = \{x \in H : \|x\|_1 \leq 1\} = \overline{\mathrm{conv}} \mathrm{Extr} S^{(1)}$ (see Theorem 2) and $\| \ \|$ is a continuous convex functional on $B^{(1)}$, it follows

$$C_{\| \ \|} \geq \sup_{x\in H, \|x\|=1} \|x\|_1 \sup_{x\in S^{(1)}} \|x\|.$$

The observation

$$\sup_{x\in H, \|x\|=1} \|x\|_1 = \left(\inf\{\|x\| : x \in S^{(1)}\}\right)^{-1}$$

completes the proof ∎

Theorem 5. *Let $\| \; \|$ be a norm on H, $\| \; \| \sim \| \; \|_1$ and $S^{(1)} = \{x \in H : \|x\|_1 = 1\}$. Then:*

(a) $C_{\| \, \|} = \sup\left\{\tau_{\| \, \|}(P) : P \in S\right\} = \dfrac{\sup\{\|x\| : x \in S^{(1)}\}}{\inf\{\|x\| : x \in S^{(1)}\}}$.

(b) $\inf\left\{\frac{\tau_{\| \, \|}(P)}{\delta(P)} : P \in S \text{ with } \delta(P) \neq 0\right\} = C_{\| \, \|}^{-1}$

$\sup\left\{\frac{\tau_{\| \, \|}(P)}{\delta(P)} : P \in S \text{ with } \delta(P) \neq 0\right\} = C_{\| \, \|}$.

Proof. Putting

$$r = \inf\{\|x\| : x \in S^{(1)}\}$$
$$R = \sup\{\|x\| : x \in S^{(1)}\}$$

Lemma 6 yields

$$\sup_{P \in S} \frac{\delta(P)}{\tau_{\| \, \|}(P)} = \sup_{P \in S} \frac{\tau_{\| \, \|}(P)}{\delta(P)} = R\frac{1}{r}$$

which together with $\delta(P) \leq 1$ implies the inequality $C_{\| \, \|} \leq R\frac{1}{r}$. The converse inequality $C_{\| \, \|} \geq R\frac{1}{r}$ follows from Lemma 7 ∎

Corollary 1. *For all norms $\| \; \| \in \mathfrak{A}$ one has:*

(1) *If $Q \in S$ is scrambling, then $\tau_{\| \, \|}(Q) < C_{\| \, \|}$.*

(2) *Any $Q \in S$ with $\tau_{\| \, \|}(Q) < \frac{1}{C_{\| \, \|}}$ is scrambling.*

Proof. This corollary follows directly from Theorem 5/(b) ∎

Corollary 2. *If $\| \; \|$ is a norm of $\mathfrak{A}$, then the following statements are equivalent:*

(a) $C_{\| \, \|} = 1$.

(b) $\tau_{\| \, \|} = \delta$.

(c) *There is a $K > 0$ with $\|x\| = K\|x\|_1 \quad (x \in H)$.*

Remark 1. Let $\tau_{\| \, \|}$ and $\tau_{\| \, \|_*}$ be two different ergodicity coefficients corresponding to the norms $\| \; \|$ and $\| \; \|_* \in \mathfrak{A}$, respectively. Then there are $P, Q \in S$ with $\tau_{\| \, \|}(P) < \tau_{\| \, \|_*}(P)$ and $\tau_{\| \, \|}(Q) > \tau_{\| \, \|_*}(Q)$.

Indeed, we may assume there is a $P \in S$ with $\tau_{\| \, \|}(P) < \tau_{\| \, \|_*}(P)$. Applying Theorem 5 we obtain

$$\sup\left\{\tau_{\| \, \|}(P)/\tau_{\| \, \|_*}(P) : P \in S \text{ non-stable}\right\} > 1$$

which means there exists a $Q \in S$ with $\tau_{\| \, \|}(Q) > \tau_{\| \, \|_*}(Q)$.

5. Applications

5.1 Inhomogeneous Markov chains. Let (X_n) denote an inhomogeneous Markov chain with denumerable infinite state space and 1-step transition matrices $P_n \in S$ $(n \in \mathbb{N})$. Further, let

$$^{m}P^{k} = P_{m+1}P_{m+2}\cdots P_{m+k} = (^{m}p_{ij}^{k})_{i,j\in\mathbb{N}} \qquad (m,k \in \mathbb{N})$$

be the k-step transition matrix. The Markov chain (X_n) is said to be

weakly ergodic if $\delta(^{m}P^{k}) = \frac{1}{2}\sup_{i,j}\sum_{l}|^{m}p_{il}^{k} -^{m} p_{jl}^{k}| \to 0$

strongly ergodic if there is a stable matrix $Q \in S$ with $\|^{m}P^{k} - Q\|_1 \to 0$

for $k \to \infty$ and every $m \in \mathbb{N}$.

Theorem 6. *Let $\| \ \|$ be equivalent to $\| \ \|_1$. A Markov chain (X_n) is weakly ergodic if and only if there exists a subdivision of the chain $(^{j_k}P^{j_{k+1}-j_k})$ such that*

$$\tau_{\| \|}(^{j_k}P^{j_{k+1}-j_k}) \leq 1 \quad (k \in \mathbb{N}) \qquad \textit{and} \qquad \sum_{k=1}^{\infty}\left(1-\tau_{\| \|}(^{j_k}P^{j_{k+1}-j_k})\right) = \infty.$$

Proof. Let $\delta(^{m}P^{k}) \to 0$ for $k \to \infty$ and every $m \in \mathbb{N}$. As $\tau_{\| \|}(^{m}P^{k}) \leq C_{\| \|}\delta(^{m}P^{k})$ there is a strongly increasing sequence (j_k) with

$$1-\tau_{\| \|}(^{j_k}P^{j_{k+1}-j_k}) > \frac{1}{2} \qquad (k \in \mathbb{N})$$

such that

$$\sum\left(1-\tau_{\| \|}(^{j_k}P^{j_{k+1}-j_k})\right) = \infty \qquad \text{and} \qquad \tau_{\| \|}(^{j_k}P^{j_{k+1}-j_k}) \leq 1 \quad (k \in \mathbb{N}).$$

Inversely, if these conditions are satisfied, then

$$\lim_{n\to\infty}\prod_{k=k_0}^{n}\tau_{\| \|}(^{j_k}P^{j_{k+1}-j_k}) = 0 \qquad (k_0 \in \mathbb{N}). \tag{3}$$

For each $n \in \mathbb{N}$ we choose a k_n with $j_{k_n} \geq n > j_{k_n} - 1$. Then

$$\tau_{\| \|}(^{m}P^{l}) = \tau_{\| \|}(P_{m+1}P_{m+2}\ldots P_{m+l}) \leq C_{\| \|}\cdot\tau_{\| \|}(P_{j_{k_{m+1}}}P_{j_{k_m}+1}\ldots P_{m+l}).$$

Because of (3) for each $\varepsilon > 0$ there is a $k_{m,\varepsilon} \in \mathbb{N}$ such that

$$\prod_{k=k_{m+1}}^{k_{m,\varepsilon}}\tau_{\| \|}(^{j_k-1}P^{j_{k+1}-j_k-1})\,C_{\| \|}^2 < \varepsilon.$$

Therefore we obtain for $m + l \geq j_{k_{m,\varepsilon}}$

$$\tau_{\| \|}(^{m}P^{l}) \leq C_{\| \|}^2\prod_{k=k_{m+1}}^{k_{m,\varepsilon}}\tau_{\| \|}\left(^{j_k-1}P^{j_{k+1}-j_k-1}\right) < \varepsilon$$

and (X_n) is weakly ergodic ∎

Corollary 3 (compare [10: Theorem 2.1). *A Markoc chain (X_n) is weakly ergodic if and only if there is a subdivision $({}^{j_k}P^{j_{k+1}-j_k})$ such that $\sum\big(1-\delta({}^{j_k}P^{j_{k+1}-j_k})\big)=\infty$.*

Example 3. Let (X_n) be an inhomogeneous Markov chain with 1-step transition matrices P_n taken from a finite subset of the set K (see Example 2). Using Theorem 6 with the trivial subdivision (P_n) one obtains directly the weak ergodicity of the chain (X_n).

Theorem 7. *A Markov chain (X_n) is weakly ergodic if and only if there exists a subdivision $({}^{j_k}P^{j_{k+1}-j_k})$ such that*

$$\sum_k\big(1-C_{\|\ \|\tau_{\|\ \|}}({}^{j_k}P^{j_{k+1}-j_k})\big)=\infty.$$

Proof. Let $\delta({}^mP^k)\to 0$ for $k\to\infty$ and every $m\in\mathbb{N}$. As $1-C_{\|\ \|\tau_{\|\ \|}}({}^mP_k)\to 1$ for $k\to\infty$ there exists a strongly increasing sequence (j_k) with

$$1-C_{\|\ \|\tau_{\|\ \|}}({}^{j_k}P^{j_{k+1}-j_k})>\frac{1}{2}$$

such that

$$\sum_k\big(1-C_{\|\ \|\tau_{\|\ \|}}({}^{j_k}P^{j_{k+1}-j_k})\big)=\infty.$$

Conversely, because of Theorem 5 one has

$$\infty=\sum_k\big(1-C_{\|\ \|\tau_{\|\ \|}}({}^{j_k}P^{j_{k+1}-j_k})\big)<\sum_k\big(1-\delta({}^{j_k}P^{j_{k+1}-j_k})\big)$$

and therefore (X_n) is weakly ergodic ∎

5.2 General products for sequences of stochastic matrices. Let $(Q_k)_{k=1}^{\infty}$ be a sequence of stochastic matrices $Q_k\in S$. For each permutation π of $\mathbb{N}$ we define inductively sequences (H_m) of products H_m by

$$\begin{aligned}\text{either}\quad H_{m+1}&=H_m\tilde{Q}_{m+1}\\ \text{or}\quad H_{m+1}&=\tilde{Q}_{m+1}H_m\end{aligned}\qquad(m\in\mathbb{N})$$

with $H_1=\tilde{Q}_1$ where $\tilde{Q}_k=Q_{\pi(k)(k\in\mathbb{N})}$. Such products are, e.g., the forward and backward products

$$\begin{aligned}P_m&=\tilde{Q}_1\tilde{Q}_2\cdots\tilde{Q}_{m-1}\tilde{Q}_m\\ M_m&=\tilde{Q}_m\tilde{Q}_{m-1}\cdots\tilde{Q}_2\tilde{Q}_1,\end{aligned}$$

respectively.

Definition 2. An infinite product (H_m) is said to be *weakly ergodic* [*stongly ergodic*] if $\delta(H_m)\to 0$ for $m\to\infty$ [if there exists a stable matrix $Q\in S$ with $\|H_m-Q\|_1\to 0$ for $m\to\infty$]. We say the weak [strong] ergodicity *obtains* for the infinite products of $(Q_k)_{k=1}^{\infty}$ if all products (H_m) constructed in the above described way are weakly [strongly] ergodic.

Theorem 8. *Let $\hat{Q} = \{P \in S : P = Q_k$ for some $Q_k\}$ be relatively compact in $(S, \| \|_1)$ and W be the set of all accumulation points of (Q_k). If there are a natural number $l \geq 1$ and $K < 1$ such that*

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

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

Proof. We consider the function f on S^l defined by

$$f(P_1, P_2, \dots, P_l) = \delta(P_1 P_2 \cdots P_l)$$

for $(P_1, P_2, \dots, P_l) \in S^l$. Because of

$$\left|\delta(P_1 P_2 \cdots P_l) - \delta(\tilde{P}_1 \tilde{P}_2 \cdots \tilde{P}_l)\right| \leq \left\|P_1 P_2 \cdots P_l - \tilde{P}_1 \tilde{P}_2 \cdots \tilde{P}_l\right\|_1,$$

using the compactness of the topological product $(\hat{Q} \cup W)^l$ it follows that $f|(\hat{Q} \cup W)^l$ is equicontinuous.

Let $\varepsilon > 0$ and $K + \varepsilon < 1$. As $f|(\hat{Q} \cup w)^l$ is equicontinuous there is a $\lambda > 0$ such that for all $(P_1, P_2, \dots, P_l) \in W^l$ and $(R_1, R_2, \dots, R_l) \in (W \cup Q)^l$ with $\max_{i=1,\dots,l} \|P_i - R_i\|_1 < \lambda$ the inequality $\delta(R_1 R_2 \cdots R_l) < K + \varepsilon$ holds.

Put

$$U_\lambda(T) = \left\{P \in S : \|P - T\|_1 < \lambda\right\} \qquad (T \in S).$$

Since W is a compact set of $(S, \| \|_1)$, there exists a finite set $\{T_1, T_2, \dots, T_k\} \subset W$ with $W \subset \cup_{i=1}^k U_\lambda(T_i)$. Now, let us consider a sequence $(Q_{\pi(k)})$. We may assume $(Q_{\pi(k)}) = (Q_k)$. There exists a k_0 with $Q_k \in \cup_{i=1}^k U_\lambda(T_i)$ for $k > k_0$ because otherwise a convergent subsequence (Q_{k_j}) of (Q_k), $\lim_{j \to \infty} Q_{k_j} = Q^*$ would exist such that $Q_{k_j} \notin \cup_{i=1}^k U_\lambda(T_i)$ 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 $Q_1, Q_2, \dots Q_{k_0}$ are factors of the products H_{m_0}. Then

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

for all $r \geq 2$ and $s = 0, 1, 2, \dots, l$ which implies $\lim_{m \to \infty} \delta(H_m) = 0$ ∎

Corollary 4 (compare [7, 12]). *Let $\| \| \in \mathfrak{A}$. If $\tau_{\| \|}(Q) < 1$ for all $Q \in W$, then the weak ergodicity obtains for all infinite products of $(Q_k)_{k=0}^{\infty}$.*

Proof. If $\tau_{\| \|}(P) < 1$ for all $P \in W$, then $\tau_{\| \|}(P) \leq \beta < 1$ for all $P \in W$ with some β. Theorem 5 yields $\delta(P_1 P_2 \cdots P_k) \leq C_{\| \|} \beta^k$ for every k-tuple $(P_1, P_2, \dots, P_k) \in W^k$, and $\delta(P_1 P_2 \cdots P_l) \leq C_{\| \|} \beta^l < 1$ in the case $l > -\frac{\log C_{\| \|}}{\log \beta}$ ∎

Remark 2. If there exists at least one accumulation point P of (Q_k) with $\delta(P) < 1$, then weak ergodicity obtains for the infinite products of (Q_k).

Indeed, let $\delta(P) < \beta < 1$ 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 $\delta(Q_{k_j}) < \beta$ for all $j > j_0$. For any product sequences (H_m) built by (Q_k) there is an m_0 such that the matrices $Q_{k_{j_0}+1}, \dots, Q_{k_{j_0}+r}$ are factors of the product H_{m_0}. Therefore $\delta(H_m) < \beta^r$ for all $m \geq m_0$ which implies $\lim_{m\to\infty} \delta(H_m) = 0$.

5.3 Markov systems. A *Markov system* over the finite alphabet $\sum$ is a pair $(\mathbb{N}, \{P(\sigma) : \sigma \in \sum\})$ where $\mathbb{N}$ is the set of states $1, 2, 3, \dots$ and $P(\sigma) \in S$ represents the transition probabilities between the states. Let $\sum^*$ be the set of all words $x = \sigma_1\sigma_2\cdots\sigma_k$ over $\sum$ and $P(x) = P(\sigma_1)P(\sigma_2)\cdots P(\sigma_k)$ the transition matrix associated with the word $x, l(x) = k$ denotes the length of the word x.

The Markov system is called *weakly ergodic* if there is an integer k such that the $P(x)$ are scrambling for all words x with $l(x) = k$ (see [10]). Obviously, $\{\mathbb{N}, \{P(\sigma) : \sigma \in \sum\}\}$ is weakly ergodic if and only if for every $\| \ \| \in \mathfrak{A}$ there exists an integer k with $\tau_{\| \ \|}(P(x)) < 1$ for all words $x \in \sum^*$ of length $l(x) = k$.

The weak ergodicity of Markov systems is important in context with perturbations of Markov systems. If $(\mathbb{N}, \{P(\sigma) : \sigma \in \sum\})$ is weakly ergodic and $(\mathbb{N}, \{\tilde{P}(\sigma) : \sigma \in \sum\})$ is an other Markov system over the same alphabet $\sum$, then for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\sup_{\sigma\in\sum} \|P(\sigma) - \tilde{P}(\sigma)\|_1 < \delta \quad \Longrightarrow \quad \sup_{x\in\sum^*} \|P(x) - \tilde{P}(x)\|_1 < \varepsilon.$$

Theorem 9. *Let $(\mathbb{N}, \{P(\sigma) : \sigma \in \sum\})$ be a Markov system over the finite alphabet $\sum$. Then the following assertions are equivalent:*

(1) $(\mathbb{N}, \{P(\sigma)\} : \sigma \in \sum)$ *is weakly ergodic.*

(2) *The backward products built by matrices from $\{P(\sigma)\}_{\sigma\in\Sigma}$ are strongly ergodic.*

(3) *The products built by matrices from $\{P(\sigma)\}_{\sigma\in\Sigma}$ are weakly ergodic.*

Proof. (1) $\Longrightarrow$ (3): Let (Q_k) be a sequence in $\{P(\sigma) : \sigma \in \sum\}$. All accumulation points of (Q_k) belong to $\{P(\sigma) : \sigma \in \sum\}$ such that (1) and Theorem 7 yield the weak ergodicity for all infinite products of Q_k.

(3) $\Longrightarrow$ (2) is true, since strong and weak ergodicity are equivalent for backward products of stochastic matrices.

(2) $\Longrightarrow$ (1): Let (2) be true and (1) false. For each $l \in \mathbb{N}$ there are $P_1^{(l)}, P_2^{(l)}, \dots, P_l^{(l)} \in \{P(\sigma)\}$ with $\delta(P_l^{(l)} \cdots P_1^{(l)}) = 1$. We show that there exists an infinite backward product built by matrices of $\{P(\sigma)\}$ which is not strongly ergodic. To prove this we construct a directed graph, whose vertices are certain finite products of the $P(\sigma)$. The root is the empty set, and there is a directed edge from $\emptyset$ to $P(\sigma)$ if $\delta(P_\sigma) = 1$. There is a directed edge from $P_lP_{l-1}\dots P_1$ to $QP_l\dots P_1$ if and only if $\delta(QP_l \cdots P_1) = 1$, such $Q \in \{P(\sigma)\}$ exists because of our assumption and as $\delta(QP_l \cdots P_1) = 1$ implies $\delta(P_l \cdots P_1) = 1$. Thus, we obtain an infinite tree, which therefore has an infinite path. This mean, there exists a not weakly ergodic backward product, which contradicts (2) ∎

Acknowledgement. The author is grateful to one of the referees for many suggestions which led to an improvement of this paper. Critical comments on the first version of Lemma 2 and Theorem 2/(b) are also appreciated. The author would also like to thank a referee for a critical question concerning the original version of Lemma 5.

References

[1] Cohn, H.: *Products of stochastic matrices and applications.* Int. J. Nath. and Math. Sci. 12 (1989), 209 – 233.

[2] Dobrushin, R. L.: *Central limit theorem for non-stationary Markov chains.* Parts I and II. Theor. Probab. Appl. I (1956), 63 – 80 and 329 – 383.

[3] Fleischer, I. and A. Joffe: *Ratio ergodicity for non-homogeneous Markov chains in general state spaces.* J. Theor. Probab. 8 (1995), 31 – 37.

[4] Gibert, S. and A. Mukherjea: *Products of infinite-dimensional nonnegative matrices; extension of Hajnal's results.* Math. Z. 196 (1987), 485 – 490.

[5] Hartfiel, D. J. and U. A. Rothblum: *Convergence of inhomogeneous products of matrices and coefficients of ergodicity.* Lin. Alg. Appl. 277 (1989), 1 – 9.

[6] Isaacson, D. L. and R. W. Madsen: *Markov Chains.* New York: Wiley Intersci. 1976.

[7] Leizarowitz, A.: *On infinite products of stochastic matrices.* Lin. Alg. Appl. 168 (1992), 189 – 219.

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

[9] Neumann, M. and H. Schneider: *The convergence of general products of matrices and the weak ergodicity of Markov chains.* Lin. Alg. Appl. 287 (1999), 307 – 314.

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

[11] Rhodius, A.: *On explicit forms for ergodicity coefficients.* Lin. Alg. Appl. 194 (1993), 71 – 83.

[12] Rhodius, A.: *On the maximum of ergodicity coefficients, the Dobrushin ergodicity coefficient, and products of stochastic matrices.* Lin. Alg. Appl. 253 (1997), 141 – 154.

[13] Rothblum, U. G. and C. P. Tan: *Upper bounds on the maximum modulus of subdominant eigenvalues of nonnegative matrices.* Lin. Alg. Appl. 66 (1985), 45 – 86.

[14] Sarymsakow, T. A.: *Inhomogeneous Markov chains* (in Russian). Teor. Veroyatnost. i Primenen. 6 (1961), 194 – 201.

[15] Schneider, H. and W. G. Strang: *Comparison theorems for supremum norms.* Num. Math. 4 (1962), 15 – 20.

[16] Seneta, E.: *Coefficients of ergodicity: Structure and applications.* Adv. Appl. Probab. 11 (1979), 576 – 590.

[17] Seneta, E.: *Nonnegative Matrices and Markov Chains.* New York: Springer-Verlag 1981.

[18] Seneta, E.: *Explicit forms for ergodicity coefficients of stochastic matrices.* Lin. Alg. Appl. 191 (1993), 245 – 252.

[19] Tan, C. P.: *A functional form for a particular coefficient of ergodicity.* J. Appl. Probab. 19 (1982), 858 – 863.

Received 04.05.1999; in revised form 30.05.2000