JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **204,** 631–645 (1996)
ARTICLE NO. 0458

Generalizations of Dobrushin's Inequalities and Applications

Marius Radulescu* and Sorin Radulescu

Centre of Mathematical Statistics, Casa Academiei, Calea 13 Septembrie nr. 13, Bucharest 5, RO-76100, Romania

Received January 8, 1996

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a seminorm and let $(e_i)_{1\le i\le n}$ be the canonical base of $\mathbb{R}^n$. Denote $M = \frac{1}{2}\max_{r,s} f(e_r - e_s)$, $K = \max_r f(e_r)$. We prove the inequality

$$f(x) \le M\left(\sum_{i=1}^{n} |x_i|\right) + (K-M)\left|\sum_{i=1}^{n} x_i\right|, \qquad x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

We use the above inequality to prove some generalizations of Dobrushin's inequalities and a generalization of an inequality due to J. E. Cohen *et al.* (*Linear Algebra Appl.* **179**, 1993, 211–235). Hilbert space generalizations of the above inequalities are proved using Levi's reduction theorem. As special cases of our results we obtain several inequalities given previously by Adamovici, Djokovic, Hlawka, and Hornich. © 1996 Academic Press, Inc.

1. INTRODUCTION

A nonhomogeneous Markov chain is described by a sequence $(P_k)_{k\ge 1}$, $P_k = (p_{ij}(k))$, $i, j \in S$, of stochastic matrices. Throughout this paper by a stochastic matrix we shall mean a column stochastic matrix, that is, a matrix whose entries are nonnegative and the sum of the entries in each column is equal to one.

The notion of weak ergodicity of the sequence $(P_k)_{k\ge 1}$ introduced by Kolmogorov [12] in 1931 requires that for every i, j, p, r,

$$\lim_{s\to\infty}\left(t_{pi}^{(r,s)} - t_{pj}^{(r,s)}\right) = 0, \tag{1.1}$$

where $T_{rs} = (t_{ij}^{(r,s)}) = P_{r+1}P_{r+2}\ldots P_{r+s}$ represents a matrix product.

* E-mail address: radulescu@roearn.ici.ro.

0022-247X/96 $18.00

Thus the columns of T_{rs} tend to be identical as $s \to \infty$. Hence a weakly ergodic Markov chain exhibits some kind of stable behavior after a long time.

Bernstein [2, 3] in his textbook of 1946 began the development of the theory of both weak and strong ergodicity.

A very important notion in the study of ergodic theory of Markov chains is that of the ergodicity coefficient. It was partly crystallized in a paper of W. Doeblin of 1937, but was put into its most powerful form by Dobrushin [8] and thoroughly exploited in 1958 by Hajnal [9], unaware of earlier Soviet work.

The notion of an ergodicity coefficient simplifies and makes more elegant and complete the theory of nonhomogeneous Markov chains.

There is a rich literature on the notion of a coefficient of ergodicity [11, 15, 16, 18–20, 22] and on a specialized coefficient (Dobrushin's) which is in a sense the optimal tool in the study of the ergodic theory of Markov chains.

If $A = (a_{ij})$ is a real $m \times n$ matrix, then let us denote

$$\alpha_{m,n}(A) = \min_{r,s}\left[\sum_{i=1}^{m} \min(a_{ir}, a_{is})\right]$$

$$\overline{\alpha}_{m,n}(A) = \frac{1}{2}\max_{r,s}\left(\sum_{i=1}^{m} |a_{ir} - a_{is}|\right).$$

In the case where A is a stochastic matrix, using the equality $\min(a, b) = \frac{1}{2}(a + b - |a - b|)$, one can easily see that $\alpha_{m,n}(A) = 1 - \overline{\alpha}_{m,n}(A)$.

$\alpha_{m,n}(A)$ is called Dobrushin's coefficient of ergodicity of the stochastic matrix A. Sometimes in this paper we shall refer to the functional $A \mapsto \overline{\alpha}_{m,n}(A)$ as Dobrushin's coefficient of ergodicity.

Condition (1.1) may be written by means of Dobrushin's coefficient as

$$\lim_{s \to \infty} \overline{\alpha}_{m,n}(T_{r,s}) = 0 \qquad \text{for every } r \geq 1. \tag{1.2}$$

This coefficient is important not only in the study of the asymptotic behavior of Markov systems but also in comparisons of stochastic matrices as communication channels [5, 6], consensus problems [4], or dynamic programming.

Let $A = (a_{ij})$ be a real $m \times n$ matrix. We shall associate with A a linear operator A: $\mathbb{R}^n \to \mathbb{R}^m$ defined as $Ax = ((Ax)_1, (Ax)_2, \ldots, (Ax)_m)$, where

$$(Ax)_i = \sum_{j=1}^{n} a_{ij} x_j, \qquad i = 1, 2, \ldots, m,\ x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n.$$

For $p \in [1, \infty]$, $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, we shall denote by $\|x\|_p$ the l^p-norm of x, that is

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{1/p} \qquad \text{if } p \in [1, \infty)$$

$$\|x\|_p = \max_{1 \le j \le n} (|x_j|) \qquad \text{if } p = +\infty.$$

The l^p-norm of A is defined as

$$\|A\|_p = \sup\{\|Ax\|_p : x \in \mathbb{R}^n, \|x\|_p = 1\}.$$

It is well known that

$$\|A\|_1 = \max_{1 \le j \le n} \left(\sum_{i=1}^{m} |a_{ij}| \right). \tag{1.3}$$

For every integer $k \ge 1$ put $H_k = \{x \in \mathbb{R}^k : x_1 + x_2 + \cdots + x_k = 0\}$. If A: $\mathbb{R}^n \to \mathbb{R}^m$, then we shall denote by $\tilde{A}$ the restriction of A to H_n. The l^p-norm of $\tilde{A}$ is defined as

$$\|\tilde{A}\|_p = \sup\{\|Ax\|_p : x \in H_n, \|x\|_p = 1\}.$$

The following theorem is due to Dobrushin [8].

THEOREM 1.1. *If $A = (a_{ij})$ is an $m \times n$ stochastic matrix, then $\|\tilde{A}\|_1 = \overline{\alpha}_{m,n}(A)$.*

The above equality may be written as an inequality

$$\|Ax\|_1 \le \overline{\alpha}_{m,n}(A) \cdot \|x\|_1 \qquad \text{for every } x \in H_n. \tag{1.4}$$

We shall refer to inequality (1.4) as Dobrushin's first inequality.

The following result is also due to Dobrushin [8].

THEOREM 1.2. *Let $A = (a_{ij})$ be an $m \times n$ stochastic matrix and $B = (b_{ij})$ be an $n \times p$ stochastic matrix. Then*

$$\overline{\alpha}_{m,n}(AB) \le \overline{\alpha}_{m,n}(A) \cdot \overline{\alpha}_{n,p}(B). \tag{1.5}$$

For a proof of the above result see Iosifescu [11, p. 58]. We shall refer to inequality (1.5) as Dobrushin's second inequality.

A generalization of Dobrushin's first inequality has been recently given as follows.

THEOREM 1.3 [5, Lemma 3.2]. *Let $A = (a_{ij})$ be an $m \times n$ stochastic matrix. Then*

$$\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \le \overline{\alpha}_{m,n}(A) \left(\sum_{j=1}^{n} |x_j| \right) + \alpha_{m,n}(A) \left| \sum_{j=1}^{n} x_j \right| \tag{1.6}$$

for every $x_1, x_2, \ldots, x_n \in \mathbb{R}$.

Another generalization of Dobrushin's first inequality asserts that (1.4) holds for arbitrary real matrices A. This follows at once from the two theorems below.

THEOREM 1.4. *Let C be a compact convex set in $\mathbb{R}^n$ and let $f\colon C \to \mathbb{R}$ be a convex map. Then*

$$\sup_{x \in C} f(x) = \sup_{x \in \operatorname{ext} C} f(x).$$

(*Here by* ext C *we denote the set of extreme points of C.*)

THEOREM 1.5 [17, Corollary 3.2]. *Let $Q = \{x \in H_n\colon \|x\|_1 \le 1\}$. Then* $\operatorname{ext} Q = \{\frac{1}{2}(e_r - e_s)\colon r, s \in \{1, 2, \ldots, n\},\ r \ne s\}$. (*Here by e_r we denote the vector $(\delta_{rj})_{1 \le j \le n}$ of $\mathbb{R}^n$ where δ_{rj} is the Kronecker symbol.*)

One of the main results of our paper uses Levi's reduction theorem to prove that in an inner product space H the inequality

$$\sum_{i=1}^{m} \left\| \sum_{j=1}^{n} a_{ij} x_j \right\| \le \overline{\alpha}_{m,n}(A) \left(\sum_{j=1}^{n} \|x_j\| \right) + \left(\|A\|_1 - \overline{\alpha}_{m,n}(A) \right) \left\| \sum_{j=1}^{n} x_j \right\| \tag{1.7}$$

holds for every real $m \times n$ matrix $A = (a_{ij})$ and every $x_j \in H$, $j = 1, 2, \ldots, n$.

One can easily see that inequality (1.7) supersedes the above mentioned generalizations of Dobrushin's first inequality.

It is very interesting to note that a series of inequalities belonging to Hlawka [14, p. 171], Hornich [10], Adamovic [1], Djokovic [7], and Mitrinovici [14, pp. 171–177] are also special cases of inequality (1.7).

Another result of our paper contains a generalization of Dobrushin's second inequality. More precisely, we prove that inequality (1.5) holds if A is an arbitrary real matrix and all the sums of the entries in each column of B are equal.

2. GENERALIZATIONS OF DOBRUSHIN'S INEQUALITIES: THE SCALAR CASE

Let $A = (a_{ij})$ be an $m \times n$ stochastic matrix; it is well known that $\overline{\alpha}_{m,n}(A) = 1$ if and only if A has two orthogonal columns.

A generalization of the above property is contained in the following proposition.

PROPOSITION 2.1. *Let $A = (a_{ij})$ be a real $m \times n$ matrix. Denote $P = \{r \in \{1, 2, \dots, n\}: \sum_{i=1}^{m} |a_{ir}| = \|A\|_1\}$. Then the following assertions are equivalent*:

(1) $\overline{\alpha}_{m,n}(A) = \|A\|_1$.

(2) *There exist distinct points $r, s \in P$ such that $a_{ir}a_{is} \le 0$ for every $i \in \{1, 2, \dots, m\}$.*

Proof. This is obvious.

In Lemma 2.2 and Theorem 2.3 we shall use the following notation. Let $n \ge 1$ be a natural number and let $I = \{1, 2, \dots, n\}$. Denote by $\mathscr{P}(I)$ the family of all subsets of I. If J is a subset of I we shall denote by J' the complement of J with respect to I and by $|J|$ the cardinal of J.

For every $a \ge 0$, $b \in \mathbb{R}$, and $J \in \mathscr{P}(I)$, put

$$Z(a, b, J) = \left\{ x \in \mathbb{R}^n: x_j \ge 0 (\forall) j \in J,\ x_j \le 0 (\forall) j \in J',\ \sum_{j=1}^{n} |x_j| = a, \sum_{j=1}^{n} x_j = b \right\}.$$

Consider the sets

$$W_1 = \{(a, b, J) \in \mathbb{R}_+ \times \mathbb{R} \times \mathscr{P}(I): 0 \le |b| \le a, 1 \le |J| \le n - 1\}$$
$$W_2 = \{(a, b, J) \in \mathbb{R}_+ \times \mathbb{R} \times \mathscr{P}(I): 0 \le a = b, |J| = n\}$$
$$W_3 = \{(a, b, J) \in \mathbb{R}_+ \times \mathbb{R} \times \mathscr{P}(I): 0 \le a = -b, J = \varnothing\}$$
$$W = W_1 \cup W_2 \cup W_3.$$

For every $i \in I$, put $e_i = (\delta_{ij})_{j \in J}$. (Here by δ_{ij} we denote the Kronecker symbol.)

LEMMA 2.2. *The following assertions hold*:

(i) $(a, b, J) \in W$ *if and only if* $Z(a, b, J) \ne \varnothing$.

(ii) *If* $(a, b, J) \in W$ *then the set* $Z(a, b, J)$ *is compact and convex*

$\operatorname{ext}[Z(a,b,J)] = \{u(r,s,a,b)\colon\ r \in J,\ \ s \in J'\}$, *where* $u(r,s,a,b) = ((a+b)/2)e_r + ((b-a)/2)e_s$.

(iii) *The family* $\{Z(a,b,J)\colon (a,b,J) \in W\}$ *is a covering of* $\mathbb{R}^n$.

Proof. To prove (i) consider $(a,b,J) \in W$. If $(a,b,J) \in W_1$, then put $|J| = m$, $x_1 = x_2 = \cdots = x_m = ((a+b)/2m) \geq 0$, $x_{m+1} = x_{m+2} = \cdots x_n = (b-a)/2(n-m) \leq 0$, $x = (x_1, x_2, \ldots, x_n)$ and note that $x \in Z(a,b,J)$. For $(a,b,J) \in W_2$, let $x_1 = x_2 = \cdots = x_n = a/n$ and note that $x = (x_1, x_2, \ldots, x_n) \in Z(a,b,J)$. For $(a,b,J) \in W_3$, let $x_1 = x_2 = \cdots = x_n = b/n$ and note that $x = (x_1, x_2, \ldots, x_n) \in Z(a,b,J)$.

To prove (ii) consider $(a,b,J) \in W$ and put $m = |J|$, $e'_j = (\delta_{ij})_{1 \leq j \leq m}$, $i = 1,2,\ldots,m$, $e''_i = (\delta_{ij})_{1 \leq j \leq n-m}$ $i = 1,2,\ldots,n-m$.

$$Y_1 = \left\{ x \in \mathbb{R}^m\colon x_j \geq 0,\ j = 1,2,\ldots,m,\ \sum_{j=1}^{m} x_j = \frac{a+b}{2} \right\}$$

$$Y_2 = \left\{ x \in \mathbb{R}^{n-m}\colon x_j \leq 0,\ j = 1,2,\ldots,n-m,\ \sum_{j=1}^{n-m} x_j = \frac{b-a}{2} \right\}.$$

Let $\sigma\colon I \to I$ be a bijection such that $\sigma(\{1,2,\ldots,m\}) = J$, $\sigma(\{m+1, m+2, \ldots, n\}) = J'$ and consider the linear isomorphism $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$, $\varphi(x_1, x_2, \ldots, x_n) = (x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})$, $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Note that $Z(a,b,J) = \varphi(Y_1 \times Y_2)$, hence

$$\begin{aligned}\operatorname{ext}[Z(a,b,J)] &= \operatorname{ext}(\varphi(Y_1 \times Y_2)) = \varphi(\operatorname{ext}(Y_1 \times Y_2)) \\ &= \varphi(\operatorname{ext} Y_1 \times \operatorname{ext} Y_2).\end{aligned}$$

Since $\operatorname{ext} Y_1 = \{((a+b)/2)e'_r\colon\ r \in \{1,2,\ldots,m\}\}$, $\operatorname{ext} Y_2 = \{((b-a)/2)e''_r\colon r \in \{1,2,\ldots,n-m\}\}$ we obtain

$$\begin{aligned}\operatorname{ext}[Z(a,b,J)] &= \varphi\left(\left\{ \frac{a+b}{2} e_r + \frac{b+a}{2} e_s \colon r \in \{1,2,\ldots,m\},\right.\right. \\ &\qquad\qquad \left.\left. s \in \{m+1, m+2, \ldots, n\} \right\}\right) \\ &= \{u(r,s,a,b)\colon r \in J,\ s \in J'\}.\end{aligned}$$

To prove (iii), let $x \in \mathbb{R}^n$ and put $J = \{j \in I\colon\ x_j \geq 0\}$, $a = \sum_{j=1}^{n} |x_j|$, $b = \sum_{j=1}^{n} x_j$. Then $x \in Z(a,b,J)$.

THEOREM 2.3. *Let* $f\colon \mathbb{R}^n \to \mathbb{R}$ *be a map with the properties*

$$f(x+y) \leq f(x) + f(y), \qquad x, y \in \mathbb{R}^n \tag{2.1}$$

$$f(\lambda x) = |\lambda| f(x), \qquad \lambda \in \mathbb{R},\ x \in \mathbb{R}^n. \tag{2.2}$$

Denote $M = \frac{1}{2}\max_{r,s}[f(e_r - e_s)]$, $K = \max_r[f(e_r)]$, *and* $L = \{\lambda e_i: f(e_i) = K$ *and* $\lambda \in \mathbb{R}\}$. *Then the inequality*

$$f(x) \le M\left(\sum_{j=1}^{n} |x_j|\right) + (K - M)\left|\sum_{j=1}^{n} x_j\right|, \qquad x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \tag{2.3}$$

holds. For every $x \in L$ *we have equality in* (2.3).

Proof. By (2.1) and (2.2) one can easily see that $f(x) \ge 0$, $x \in \mathbb{R}^n$. Let $(a, b, J) \in W$. By Theorem 1.4 we have that the inequality

$$f(x) \le \max\{f(v): v \in \operatorname{ext}[Z(a, b, J)]\} = \max_{\substack{r \in J \\ s \in J'}}\left[f\left(\frac{a+b}{2}e_r + \frac{b-a}{2}e_s\right)\right]$$

holds for every $x \in Z(a, b, J)$.

Put $w_{rs} = f(((a+b)/2)e_r + ((b-a)/2)e_s)$ and note that

$$\begin{aligned}
w_{rs} &= f\left(\frac{a+b}{2}(e_r - e_s) + be_s\right) \le f\left(\frac{a+b}{2}(e_r - e_s)\right) + f(be_s) \\
&= \frac{a+b}{2}f(e_r - e_s) + |b|f(e_s) \\
&\le (a+b)M + K|b|, \\
w_{rs} &= f\left(\frac{a+b}{2}e_r + \frac{b-a}{2}e_s\right) = f\left(\frac{b-a}{2}(e_s - e_r) + be_r\right) \\
&\le f\left(\frac{b-a}{2}(e_s - e_r)\right) + f(be_r) \\
&= \frac{a-b}{2}f(e_r - e_s) + |b|f(e_r) \le (a-b)M + K|b|.
\end{aligned}$$

Thus $w_{rs} \le \min\{(a+b)M + K|b|, (a-b)M + K|b|\} = aM + K|b| + \min(bM, -bM) = aM + K|b| - M|b| = aM + (K - M)|b|$ and hence

$$f(x) \le aM + (K - M)|b| \qquad \text{for every } x \in Z(a, b, J). \tag{2.4}$$

From (2.4) and the preceding lemma inequality (2.3) follows.

The following corollary of Theorem 2.3 is a generalization of Theorem 1.3.

COROLLARY 2.4. *Let $A = (a_{ij})$ be a real $m \times n$ matrix and $1 \le p \le +\infty$. Put*

$$\overline{\alpha}^{(p)}_{m,n}(A) = \frac{1}{2} \max_{r,s} \left(\sum_{i=1}^{m} |a_{ir} - a_{is}|^p \right)^{1/p} \qquad \text{if } 1 \le p \le +\infty$$

$$\overline{\alpha}^{(p)}_{m,n}(A) = \frac{1}{2} \max_{i,r,s} |a_{ir} - a_{is}| \qquad \text{if } p = +\infty$$

$$\beta^{(p)}_{m,n}(A) = \max_{r} \left(\sum_{i=1}^{m} |a_{ir}|^p \right)^{1/p} \qquad \text{if } 1 \le p < +\infty$$

$$\beta^{(p)}_{m,n}(A) = \max_{i,r} |a_{ir}| \qquad \text{if } p = +\infty$$

$$T(m,n,p) = \left\{ r \in \{1,2,\ldots,n\} \colon \beta^{(p)}_{m,n}(A) = \left(\sum_{i=1}^{m} |a_{ir}|^p \right)^{1/p} \right\} \qquad \text{if } 1 \le p < +\infty$$

$$T(m,n,p) = \left\{ r \in \{1,2,\ldots,n\} \colon \beta^{(p)}_{m,n}(A) = \max_{i,r} \left(|a_{ir}| \right) \right\} \qquad \text{if } p = +\infty$$

$$L(m,n,p) = \left\{ \lambda e_r \colon r \in T(m,n,p), \lambda \in \mathbb{R} \right\} \qquad \text{if } 1 \le p \le +\infty.$$

Then for every $x_1, x_2, \ldots, x_n \in \mathbb{R}$ the following inequalities hold:

$$\left(\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right|^p \right)^{1/p} \le \overline{\alpha}^{(p)}_{m,n}(A) \left(\sum_{j=1}^{n} |x_j| \right) + \left(\beta^{(p)}_{m,n}(A) - \overline{\alpha}^{(p)}_{m,n}(A) \right) \left| \sum_{j=1}^{n} x_j \right| \qquad \text{if } 1 \le p \le +\infty \tag{2.5}$$

$$\max_{1 \le i \le m} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \le \overline{\alpha}^{(p)}_{m,n}(A) \left(\sum_{j=1}^{n} |x_j| \right) + \left(\beta^{(p)}_{m,n}(A) - \overline{\alpha}^{(p)}_{m,n}(A) \right) \left| \sum_{j=1}^{n} x_j \right| \qquad \text{if } p = +\infty \tag{2.6}$$

$$\sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \le \overline{\alpha}_{m,n}(A) \left(\sum_{j=1}^{n} |x_j| \right) + \left(\|A\|_1 - \overline{\alpha}_{m,n}(A) \right) \cdot \left| \sum_{j=1}^{n} x_j \right|. \tag{2.7}$$

For every $p \in [1, \infty]$, $x = (x_1, x_2, \dots, x_n) \in L(m, n, p)$ *we have equality in* (2.5) *and* (2.6). *For every* $x = (x_1, x_2, \dots, x_n) \in L(m, n, 1)$ *we have equality in* (2.7).

Proof. For every $p \in [1, \infty]$ consider the map $f_p(x) = \|Ax\|_p$, $x \in \mathbb{R}^n$. Then using the notation from the preceding theorem we have that $M = \overline{\alpha}_{m,n}^{(p)}(A)$, $K = \beta_{m,n}^{(p)}(A)$. Applying Theorem 2.3 we obtain inequalities (2.5) and (2.6).

If in inequality (2.5) we put $p = 1$, then we obtain inequality (2.7).

Generalizations of inequality (2.7) to spaces of integrable functions or to spaces of measures may be found in Zaharopol and Zbăganu [24] and Zbăganu [25].

Remark 2.5. $\overline{\alpha}_{m,n}^{(p)}(A)$ and $\beta_{m,n}^{(p)}(A)$ have the following remarkable interpretation. Let E, F, H be three Banach spaces defined as follows: E is $\mathbb{R}^n$ endowed with the l^1-norm, F is endowed with the l^p-norm, and H is the subspace $\{x \in R\colon \sum_{j=1}^n x_j = 0\}$ of E. Then the norm of the linear map $A\colon E \to F$, $(Ax)_i = \sum_{j=1}^n a_{ij} x_j$, $i = 1, 2, \dots, m$ is $\beta_{m,n}^{(p)}(A)$, while the norm of the restriction of A to H is $\overline{\alpha}_{m,n}^{(p)}(A)$.

COROLLARY 2.6. *Let $B = (b_{ij})$ be an invertible real $n \times n$ matrix. Then the inequality*

$$\sum_{i=1}^n |y_i| \le \overline{\alpha}_{n,n}(B^{-1}) \sum_{i=1}^n \left| \sum_{j=1}^n b_{ij} y_j \right| + \left(\|B^{-1}\|_1 - \overline{\alpha}_{n,n}(B^{-1}) \right) \left| \sum_{i=1}^n \sum_{j=1}^n b_{ij} y_j \right| \tag{2.8}$$

holds for every $y_1, y_2, \dots, y_n \in \mathbb{R}$.

Proof. Let $A = (a_{ij}) = B^{-1}$. Inequality (2.8) follows at once from inequality (2.7) if we put $y_i = \sum_{j=1}^n a_{ij} x_j$, $i = 1, 2, \dots, n$. For every real $m \times n$ matrix $A = (a_{ij})$ put

$$\gamma_{m,n}(A) = \frac{1}{2} \max_{r,s} \left| \sum_{i=1}^m (a_{ir} - a_{is}) \right|$$

$$\delta_{m,n}(A) = \max_r \left| \sum_{i=1}^m a_{ir} \right|.$$

One can easily see that the inequalities

$$\gamma_{m,n}(A) \le \overline{\alpha}_{m,n}(A), \qquad \delta_{m,n}(A) \le \|A\|_1$$

hold for every real $m \times n$ matrix A.

Using the functionals $\gamma_{m,n}$ and $\delta_{m,n}$ we can state the following generalization of Dobrushin's second inequality:

THEOREM 2.7. *Let $p \in [1,\infty]$, $A = (a_{ij})$ be a real $m \times n$ matrix and let $B = (b_{ij})$ be a real $n \times k$ matrix. Then the following inequalities hold*:

$$\overline{\alpha}^{(p)}_{m,k}(AB) \le \overline{\alpha}^{(p)}_{m,n}(A)\,\overline{\alpha}_{n,k}(B) + \gamma_{n,k}(B)\big(\beta^{(p)}_{m,n}(A) - \overline{\alpha}^{(p)}_{m,n}(A)\big) \tag{2.9}$$

$$\beta^{(p)}_{m,k}(AB) \le \overline{\alpha}^{(p)}_{m,n}(A)\|B\|_1 + \big(\beta^{(p)}_{m,n}(A) - \overline{\alpha}^{(p)}_{m,n}(A)\big)\delta_{n,k}(B) \tag{2.10}$$

$$\overline{\alpha}_{m,k}(AB) \le \overline{\alpha}_{m,n}(A)\,\overline{\alpha}_{n,k}(B) + \gamma_{n,k}(B)\big(\|A\|_1 - \overline{\alpha}_{m,n}(A)\big) \tag{2.11}$$

$$\|AB\|_1 \le \overline{\alpha}_{m,n}(A)\|B\|_1 + \big(\|A\|_1 - \overline{\alpha}_{m,n}(A)\big)\delta_{n,k}(B). \tag{2.12}$$

Proof. Put

$$\varphi(A,B) = \overline{\alpha}^{(p)}_{m,n}(A)\,\overline{\alpha}_{n,k}(B) + \gamma_{n,k}(B)\big(\beta^{(p)}_{m,n}(A) - \overline{\alpha}^{(p)}_{m,n}(A)\big)$$

$$\psi(A,B) = \overline{\alpha}^{(p)}_{m,n}(A)\big(\|B\|_1 - \overline{\alpha}_{n,k}(B)\big) + \big(\beta^{(p)}_{m,n}(A) - \overline{\alpha}^{(p)}_{m,n}(A)\big)\big(\delta_{n,k}(B) - \gamma_{n,k}(B)\big).$$

Let $x = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^k$ and note that

$$\left|\sum_{j=1}^{n}\left(\sum_{i=1}^{k} b_{ji}x_i\right)\right| = \left|\sum_{i=1}^{k}\left(\sum_{j=1}^{n} b_{ji}\right)x_i\right| \le \gamma_{n,k}(B)\|x\|_1 + \big(\delta_{n,k}(B) - \gamma_{n,k}(B)\big)\left|\sum_{i=1}^{k} x_i\right|.$$

By the above inequality and by (2.5) and (2.6) we obtain

$$\begin{aligned}\|ABx\|_p &\le \overline{\alpha}^{(p)}_{m,n}(A)\|Bx\|_1 + \big(\beta^{(p)}_{m,n}(A) - \overline{\alpha}^{(p)}_{m,n}(A)\big)\left|\sum_{j=1}^{n}(Bx)_j\right| \\ &\le \overline{\alpha}^{(p)}_{m,n}(A)\left[\overline{\alpha}_{n,k}(B)\|x\|_1 + \big(\|B\|_1 - \overline{\alpha}_{n,k}(B)\big)\left|\sum_{i=1}^{k} x_i\right|\right. \\ &\quad + \big(\beta^{(p)}_{m,n}(A) - \overline{\alpha}^{(p)}_{m,n}(A)\big)\left|\sum_{j=1}^{n}\left(\sum_{i=1}^{k} b_{ji}x_i\right)\right| \\ &\le \varphi(A,B)\|x\|_1 + \psi(A,B)\left|\sum_{i=1}^{k} x_i\right|.\end{aligned}$$

If in the above inequality we take the supremum over $\{x \in \mathbb{R}^k; \|x\|_1 \leq 1, \sum_{i=1}^k x_i = 0\}$, then we obtain inequality (2.9). Since $\|ABx\|_p \leq \varphi(A, B)\|x\|_1 + \psi(A, B) \cdot |\sum_{i=1}^k x_i| \leq (\varphi(A, B) + \psi(A, B))\|x\|_1$ we obtain that

$$\beta_{m,k}^{(p)}(AB) \leq \varphi(A, B) + \psi(A, B)$$
$$= \overline{\alpha}_{m,n}^{(p)}(A)\|B\|_1 + \left(\beta_{m,n}^{(p)}(A) - \overline{\alpha}_{m,n}^{(p)}(A)\right) \cdot \delta_{n,k}(B).$$

If in (2.9) and (2.10) we put $p = 1$ we obtain (2.11) and (2.12).

COROLLARY 2.8. *Let $A = (a_{ij})$ be a real $m \times n$ matrix and let $B = (b_{ij})$ be a real $n \times k$ matrix. If for some $b \in \mathbb{R}$ we have $\sum_{i=1}^n b_{ij} = b$ for every $j \in \{1, 2, \ldots, k\}$, then*

$$\overline{\alpha}_{m,k}(AB) \leq \overline{\alpha}_{m,k}(A) \cdot \overline{\alpha}_{n,r}(B). \tag{2.13}$$

Proof. The condition that all the sums of the entries of B in each column are equal implies that $\gamma_{n,k}(B) = 0$. Now, inequality (2.13) follows at once from (2.11).

Remark 2.9. One can easily see that inequality (2.12) improves the classical inequality $\|AB\|_1 \leq \|A\|_1 \cdot \|B\|_1$.

3. A HILBERT SPACE GENERALIZATION OF DOBRUSHIN'S FIRST INEQUALITY

The following theorem is known as Levi's reduction theorem.

THEOREM 3.1 [13; 14, p. 175]. *Let $A = (a_{ij})$ be a real $m \times r$ matrix and let $B = (b_{ij})$ be a real $n \times r$ matrix. Consider on $\mathbb{R}^k$ the l^2-norm. Suppose that*

$$\sum_{i=1}^m \left| \sum_{j=1}^r a_{ij} t_j \right| \leq \sum_{i=1}^n \left| \sum_{j=1}^r b_{ij} t_j \right| \tag{3.1}$$

for every $t_1, t_2, \ldots, t_r \in \mathbb{R}$.

Then for every $x_1, x_2, \ldots, x_r \in \mathbb{R}^k$ we have

$$\sum_{i=1}^m \left\| \sum_{j=1}^r a_{ij} x_j \right\|_2 \leq \sum_{i=1}^n \left\| \sum_{j=1}^r b_{ij} x_j \right\|_2. \tag{3.2}$$

Remark 3.2. Since every two finite dimensional Hilbert spaces of the same dimension are isometric, it follows that inequality (3.1) implies that inequality (3.2) holds in any finite dimensional Hilbert space.

One can easily see that this implies that inequality (3.2) holds in pre-Hilbert spaces of arbitrary dimension.

Levi's reduction theorem allows us to obtain a Hilbert space generalization of Dobrushin's first inequality.

THEOREM 3.3. *Let H be a pre-Hilbert space and let $A = (a_{ij})$ be a real $m \times n$ matrix. Then*

$$\sum_{i=1}^{m} \left\| \sum_{j=1}^{n} a_{ij} x_j \right\| \leq \overline{\alpha}_{m,n}(A) \left(\sum_{j=1}^{n} \|x_j\| \right) + \big(\|A\|_1 - \overline{\alpha}_{m,n}(A) \big) \left\| \sum_{j=1}^{n} x_j \right\| \tag{3.3}$$

for every $x_1, x_2, \ldots, x_n \in H$.

Proof. The above inequality follows at once from (2.7) and Levi's reduction theorem.

COROLLARY 3.4 (Hlawka's Inequality [14, p. 171]. *Let H be a pre-Hilbert space. Then*

$$\|x + y\| + \|y + z\| + \|z + x\| \leq \|x\| + \|y\| + \|z\| + \|x + y + z\| \tag{3.4}$$

for every $x, y, z \in H$.

Proof. Inequality (3.4) follows at once from (3.3) if we take

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

and note that $\overline{\alpha}_{3,3}(A) = 1$, $\|A\|_1 = 2$.

COROLLARY 3.5. *Let H be a pre-Hilbert space, $n \geq 3$, and $k \in \{2, 3, \ldots, n-1\}$. Then*

$$\sum_{1 \leq i < j \leq n} \|x_i + x_j\| \leq (n-2) \left(\sum_{i=1}^{n} \|x_i\| \right) + \left\| \sum_{i=1}^{n} x_i \right\| \tag{3.5}$$

$$\sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \|x_{i_1} + x_{i_2} + \cdots + x_{i_k}\| \leq C_{n-2}^{k-2} \left(\frac{n-k}{k-1} \sum_{i=1}^{n} \|x_i\| + \left\| \sum_{i=1}^{n} x_i \right\| \right) \tag{3.6}$$

for every $x_1, x_2, \ldots, x_n \in H$.

Proof. One can easily see that inequality (3.5) is a special case of inequality (3.6). To prove (3.6) let $m = C_n^k$ and consider the set

$$L = \left\{(i_1, i_2, \ldots, i_k) \in \{1, 2, \ldots, n\}^k : 1 \le i_1 < i_2 < \cdots < i_k \le n\right\}.$$

Let $\sigma : \{1, 2, \ldots, m\} \to L$, $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_k)$, be a bijection and consider the $m \times n$ matrix $A = (a_{ij})$ whose entries are defined as

$$a_{ij} = \begin{cases} 1 & \text{if } j \in \{\sigma_1(i), \sigma_2(i), \ldots, \sigma_k(i)\} \\ 0 & \text{if } j \notin \{\sigma_1(i), \sigma_2(i), \ldots, \sigma_k(i)\}. \end{cases}$$

Note that $\overline{\alpha}_{m,n}(A) = C_{n-1}^{k-1} - C_{n-2}^{k-2}$ and $\|A\|_1 = C_{n-1}^{k-1}$. An application of inequality (3.3) to the matrix A yields inequality (3.6).

Inequality (3.5) was established by Adamovic [1]. This inequality contains Hlawka's inequality as a special case when $n = 3$. Adamovic's proof is based on an identity in an inner-product space. A straightforward proof by induction of inequality (3.5) was given by Vasic [23].

Inequality (3.6) was proved by Djokovic [7] and independently by Smiley and Smiley in [21]. Conditions for equality in (3.6) were given in [7, 21].

COROLLARY 3.6. *Let H be a pre-Hilbert space, $a \in H$, $x_1, x_2, \ldots, x_n \in H$. If*

$$\sum_{i=1}^{n} x_i = -ta \qquad (t \ge 1) \tag{3.7}$$

then the inequality

$$\sum_{i=1}^{n} \left(\|x_i + a\| - \|x_i\|\right) \le (n-2)\|a\| \tag{3.8}$$

holds. If $t < 1$ in (3.7), then (3.8) need not necessarily hold.

Proof. Let $A = (a_{ij})$ where $a_{ij} = \delta_{ij} - 1/t$, $1 \le i, j \le n$, and note that $\overline{\alpha}_{n,n}(A) = 1$, $\|A\|_1 = 1 + (n-2)/t$. By (3.3) we obtain that

$$\begin{aligned} \sum_{i=1}^{n} \|x_i + a\| &= \sum_{i=1}^{n} \left\| \sum_{j=1}^{n} a_{ij} x_j \right\| \\ &\le \overline{\alpha}_{n,n}(A) \sum_{i=1}^{n} \|x_i\| + \left(\|A\|_1 - \overline{\alpha}_{n,n}(A)\right) \left\| \sum_{j=1}^{n} x_j \right\| \\ &= \sum_{i=1}^{n} \|x_i\| + \frac{n-2}{t} \left\| \sum_{i=1}^{n} x_i \right\| = \sum_{i=1}^{n} \|x_i\| + (n-2)\|a\|. \end{aligned}$$

The result from the above corollary is known as Hornich's inequality. Hornich [10] gave his inequality in a different form, which is a special case of (3.8).

A comprehensive discussion on the inequalities of Hlawka, Adamovic, Djokovic, and Hornich may be found in Mitrinovici [14, pp. 171–173].

REFERENCES

1. D. D. Adamovic, Generalisation d'une identite de Hlawka et de l'inegalite correspondante, *Mat. Vesnik* **1**, No. 16 (1964), 39–43.
2. S. N. Bernstein, ''Teoriya Veroiatnostei,'' 4th ed., Gostehizdat, Moskow/Leningrad, 1946 (relevant portions are reprinted in [3]).
3. S. N. Bernstein, Classification of Markov chains and their matrices, *in* ''Sobranie Sochineniy, Tom IV, Teoriya Veroiatnostei i Matematicheskaia Statistika (1911–1946),'' pp. 455–483, Nauka, Moscow, 1964. [In Russian]
4. S. Chatterjee and E. Seneta, Towards consensus: Some convergence theorems on repeated averaging, *J. Appl. Probab.* **14** (1977), 89–97.
5. J. E. Cohen, Y. Iwasa, Gh. Răutu, M. B. Ruskai, E. Seneta, and Gh. Zbăganu, Relative entropy under mappings by stochastic matrices, *Linear Algebra Appl.* **179** (1993), 211–235.
6. J. E. Cohen, R. H. Kemperman, and Gh. Zbăganu, Comparisons of stochastic matrices as communication channels, submitted for publication.
7. D. Z. Djokovic, Generalizations of Hlawka's inequality, *Glas. Mat. Fiz. Astronom. Ser. II Drustvo Mat. Fiz. Hrvatske* **18** (1963), 169–175.
8. R. L. Dobrushin, Central limit theorem for nonstationary Markov chains, I, II, *Theory Probab. Appl.* **1** (1956), 65–80, 329–383.
9. J. Hajnal, Weak ergodicity in non-homogeneous Markov chains, *Math. Proc. Cambridge Philos. Soc.* **54** (1958), 233–246.
10. H. Hornich, Eine Unglachung für Vektorlangen, *Math. Z.* **48** (1942), 264–274.
11. M. Iosifescu, ''Finite Markov Processes and Their Applications,'' Wiley, New York, and Editura Tehnică, Bucharest, 1980.
12. A. N. Kolmogorov, Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung, *Math. Ann.* **104** (1931), 415–458.
13. F. W. Levi, Ein Reductionsverfahren für lineare Vektorungleichungen, *Arch. Math. (Basel)* **2** (1949), 24–26.
14. D. S. Mitrinovici, ''Analytic Inequalities,'' Springer-Verlag, Berlin, 1970.
15. A. Paz and M. Reichaw, Ergodic theorems for sequences of infinite stochastic matrices, *Math. Proc. Cambridge Philos. Soc.* **63** (1967), 777–786.
16. A. Paz, ''Introduction to Probabilistic Automata,'' Academic Press, New York, 1971.
17. U. G. Rothblum, Explicit solutions to optimisation problems on the intersections of the unit ball of the l_1- and l_∞-norms with a hyperplane, *SIAM J. Discrete Math.* **5** (1984), 619–632.
18. U. G. Rothblum and C. P. Tan, Upper bounds on the maximum modulus of subdominant eigenvalues of nonnegative matrices, *Linear Algebra Appl.* **66** (1985), 45–86.
19. E. Seneta, Coefficients of ergodicity: Structure and applications, *Adv. in Appl. Probab.* **11** (1979), 576–590.

20. E. Seneta, ‘‘Non-negative Matrices and Markov Chains,’’ Springer-Verlag, New York, 1981.
21. D. M. Smiley and M. F. Smiley, The polygonal inequalities, *Amer. Math. Monthly* **71** (1964), 755−760.
22. C. P. Tan, ‘‘Coefficients of Ergodicity and GI Stochastic Automata,’’ Ph.D. Thesis, Univ. of Malaya, Kuala Lumpur, 1992.
23. P. M. Vasic, Les inégalities pour les fonctions convexes d’ordre n, *Mat. Vesnik* **5**, No. 20 (1968), 327−331.
24. R. Zaharopol and Gh. Zbăganu, Dobrushin coefficients of ergodicity and asymptotic stability of $-L^1$ contractions, to appear.
25. Gh. Zbăganu, Divergence and contraction coefficients, to appear.