Czechoslovak Mathematical Journal

Petr Kratochvíl; Antonín Lešanovský

A contractive property in finite state Markov chains

Czechoslovak Mathematical Journal, Vol. 35 (1985), No. 3, 491–509

Persistent URL: http://dml.cz/dmlcz/102037

Czechoslovak Mathematical Journal, 35 (110) 1985, Praha

A CONTRACTIVE PROPERTY IN FINITE STATE MARKOV CHAINS

PETR KRATOCHVÍL, ANTONÍN LEŠANOVSKÝ, Praha

(Received August 3, 1984)

This paper deals with homogeneous Markov chains with finite number of states. Let $\mathbf{P} = (p_{ij})$ be the matrix of transition probabilities of a Markov chain and let $\mathbf{p}(0)$ denote its probability distribution (row) vector at the initial instant $t = 0$. Conditions on the matrix $\mathbf{P}$ guaranteeing the existence of a limit (say π) of the sequence $\{\mathbf{p}(t)\}_{t=0}^{\infty}$, where $\mathbf{p}(t) = \mathbf{p}(0)\,\mathbf{P}^t$ for $t = 0, 1, 2, \ldots$, are well known. If $\mathbf{P}$ fulfils these conditions then the probability distribution $\mathbf{p}(t)$ at time t for large t is frequently replaced by the limit π for the simplicity of calculations. It would be desirable to have an upper estimate of the differences of $\mathbf{p}(t)$ and π common for all t large enough, say $t \geqq t_0$. Perron's formula provides only a local result for individual t's and the asymptotical rate of convergence of $\mathbf{p}(t)$ to π.

We shall follow the idea used in the theory of non-homogeneous Markov chains. In that case the central role is played by ergodicity coefficients which have been studied by many authors. The most famous of them are the Birkhoff's contraction coefficient $\tau_B(\mathbf{P})$ – [2], and the ergodicity coefficient $\tau_1(\mathbf{P}) = \frac{1}{2} \max_{i,j} \sum_m |p_{im} - p_{jm}|$ first introduced by Dobrushin [4] for Markov chains with countable state spaces and exploited by Hajnal [5] and Sarymsakov [13], cf. Seneta [14] and [15]. The value of $\tau_1(\mathbf{P})$ was also evaluated by Paz [11]. Relations between $\tau_B(\mathbf{P})$, $\tau_1(\mathbf{P})$ and the eigenvalues of $\mathbf{P}$ were investigated by Bauer, Deutsch and Stoer [1].

The ergodicity coefficients express various effects of one-step transition matrices. We shall deal with the contractive effect of the matrix $\mathbf{P}$ on the L_1-norm of the differences $\mathbf{p}(t+1) - \mathbf{p}(t)$, $t = 0, 1, 2, \ldots$, i.e. with the relation between the L_1-norms of the vectors $\mathbf{p}(t+1) - \mathbf{p}(t) = \mathbf{p}(t)(\mathbf{P} - \mathbf{I})$ and $\mathbf{p}(t+2) - \mathbf{p}(t+1) = \mathbf{p}(t) \,.\, (\mathbf{P} - \mathbf{I})\,\mathbf{P}$.

Let k denote the number of rows (and of columns as well) of the matrix $\mathbf{P}$, and let R^k be the set of all row vectors with k components. We put

$$\mathscr{P} = \{\mathbf{p};\ \mathbf{p} \in R^k,\ p_i \geqq 0, \text{ for } i = 1, \ldots, k,\ \sum_{i=1}^{k} p_i = 1\},$$
$$\mathscr{P}_0 = \{\mathbf{p};\ \mathbf{p} \in \mathscr{P},\ \mathbf{p}(\mathbf{P} - \mathbf{I}) \neq 0\},$$

and

$$(1) \qquad \alpha(\mathbf{P}) = \sup\left[\left\{\frac{\|\mathbf{p}(\mathbf{P} - \mathbf{I})\,\mathbf{P}\|}{\|\mathbf{p}(\mathbf{P} - \mathbf{I})\|};\ \ \mathbf{p} \in \mathscr{P}_0\right\} \cup \{0\}\right].$$

In Section 4 we shall present an algorithm for determining the value of $\alpha(\mathbf{P})$, for an arbitrary stochastic matrix $\mathbf{P}$, based on Theorem 2. The procedure is illustrated by a numerical example given in Section 5. It is easy to show that

$$0 \leqq \alpha(\mathbf{P}) \leqq 1 \tag{2}$$

and that the relations

$$\|\mathbf{p}(t+2) - \mathbf{p}(t+1)\| \leqq \alpha(\mathbf{P}) \, \|\mathbf{p}(t+1) - \mathbf{p}(t)\| \,, \quad \text{for} \quad t = 0, 1, 2, \ldots,$$
$$\|\mathbf{p}(1) - \mathbf{p}(0)\| \leqq 2$$

hold for every initial distribution vector $\mathbf{p}(0)$. Further, in the case of $\alpha(\mathbf{P}) < 1$, the limit π of the sequence $\{\mathbf{p}(t)\}_{t=0}^{\infty}$ does exist for every $\mathbf{p}(0)$, and the inequality

$$\|\mathbf{p}(t) - \pi\| \leqq \frac{2[\alpha(\mathbf{P})]^t}{1 - \alpha(\mathbf{P})}$$

is true for all natural t. Thus, the number

$$\frac{2[\alpha(\mathbf{P})]^{t_0}}{1 - \alpha(\mathbf{P})}$$

is a common upper estimate of $\|\mathbf{p}(t) - \pi\|$ for all $t \geqq t_0$, and of $p_i(t) - \pi_i$ as well.

Finally, in Theorem 3 we shall present a characterization of matrices $\mathbf{P}$ such that the inequality

$$\|\mathbf{p}(t+2) - \mathbf{p}(t+1)\| < \|\mathbf{p}(t+1) - \mathbf{p}(t)\| \tag{3}$$

holds for all probability distribution vectors $\mathbf{p}(t)$, $t = 0, 1, 2, \ldots,$ possessing the property $\mathbf{p}(t) \neq \mathbf{p}(t+1)$. In other words, we shall find a necessary and sufficient condition for the validity of the relation $\alpha(\mathbf{P}) < 1$. This result is a generalization of that given in the paper [10] where only irreducible aperiodic matrices $\mathbf{P}$ are considered. Theorem 4 introduces, moreover, an upper bound, the evaluation of which is quite simple. The value of this upper bound is less than the unity whenever $\alpha(\mathbf{P})$ is.

1. NOTATION

We shall consider the following characteristics of a homogeneous Markov chain (all vectors are supposed to be row vectors):

k – the number of its states;

$S = \{1; 2; \ldots; k\}$ – its state-space;

$\mathbf{P} = (p_{ij})_{i,j \in S}$ – a matrix of its transition probabilities,

$$\sum_{j=1}^{k} p_{ij} = 1 \,;$$

$\mathbf{r}_i$, $i \in S$ – i-th row of the matrix $\mathbf{P}$, $\mathbf{r}_i = (p_{i1}, \ldots, p_{ik})$;

n – the number of its recurrent classes;

$\mathcal{N} = \{1; 2; \ldots; n\}$;
C_a, $a \in \mathcal{N}$ – its a-th recurrent class, $C_a \subset S$;
T – the set of its transient states, $T \subset S$;
$\boldsymbol{p}(t)$, $t = 0, 1, 2, \ldots$ – the vector of its probability distribution at time t.

Further, if $\mathbf{x} = (x_i)_{i \in S}$ is a vector then we put

$$C_a^+(\mathbf{x}) = \{i;\ i \in C_a,\ x_i > 0\} \text{ for } a \in \mathcal{N};$$
$$C_a^-(\mathbf{x}) = \{i;\ i \in C_a,\ x_i < 0\} \text{ for } a \in \mathcal{N};$$
$$T^+(\mathbf{x}) = \{i;\ i \in T,\ x_i > 0\};$$
$$T^-(\mathbf{x}) = \{i;\ i \in T,\ x_i < 0\};$$
$$f_a^+(\mathbf{x}) = \sum_{i \in C_a^+(\mathbf{x})} x_i \text{ for } a \in \mathcal{N};$$
$$f_a^-(\mathbf{x}) = \sum_{i \in C_a^-(\mathbf{x})} x_i \text{ for } a \in \mathcal{N};$$
$$\mathbf{x}^{(a)} = (x_i)_{i \in C_a} \text{ for } a \in \mathcal{N};$$
$$\mathbf{x}^{(n+1)} = (x_i)_{i \in T};$$
$$\|\mathbf{x}\| = \sum_{i=1}^{k} |x_i|, \text{ i.e. the } L_1\text{-norm of the vector } \mathbf{x};$$
$$\|\mathbf{x}^{(a)}\| = \sum_{i \in C_a} |x_i| \text{ for } a \in \mathcal{N};$$
$$\|\mathbf{x}^{(n+1)}\| = \sum_{i \in T} |x_i|.$$

Finally, we denote by $\boldsymbol{I}$ the unit matrices (without indicating their types) and by $\mathbf{e} = (e_i)_{i \in S}$ the vector such that $e_i = 1$, for each $i \in S$.

2. SOME REMARKS ON THE IRREDUCIBLE CASE

In this section, we shall assume that $\boldsymbol{P}$ is a matrix of transition probabilities of a homogeneous irreducible Markov chain. We shall first show that the assumption of the aperiodicity is not substantial in Theorem presented in the paper [10]. Indeed, if the chain is periodic (say with a period $d \geqq 2$) then it is well-known that there is a re-ordering of the state space producing the block form of the matrix $\boldsymbol{P}$

$$\boldsymbol{P} = \begin{pmatrix} 0 & \boldsymbol{P}_1 & 0 & \ldots & 0 \\ 0 & 0 & \boldsymbol{P}_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \boldsymbol{P}_{d-1} \\ \boldsymbol{P}_d & 0 & 0 & \ldots & 0 \end{pmatrix}. \tag{4}$$

It is easily seen that (4) implies that the product $\boldsymbol{PP}'$ must contain a zero element. Further, the vector $\boldsymbol{p}(0) \in \mathcal{P}$ such that $p_1(0) = 1$ and $p_i(0) = 0$ for each $i \in S - \{1\}$ possesses the properties

$$\boldsymbol{p}(0) \neq \boldsymbol{p}(1),$$

and

$$\|\mathbf{p}(2) - \mathbf{p}(1)\| = \|\mathbf{p}(1) - \mathbf{p}(0)\| = 2 .$$

Thus, we have just verified

Lemma 1. *Let* $\mathbf{P}$ *be an irreducible periodic matrix. Then*

(a) *the product* $\mathbf{PP}'$ *contains a zero element;*

(b) *there exists a vector* $\mathbf{p}(0) \in \mathscr{P}_0$ *such that the strong inequality* (3) *is not valid for* $t = 0$.

On the basis of Lemma 1, we obtain the generalization mentioned above of Theorem from the paper [10].

Theorem 1. *Let* $\mathbf{P}$ *be an irreducible matrix. Then the strict inequality* (3) *holds for every non-stationary vector* $\mathbf{p}(t) \in \mathscr{P}_0$ *if and only if the product* $\mathbf{PP}'$ *is a positive matrix.*

3. SEVERAL AUXILIARY ASSERTIONS

Let us consider a general homogeneous Markov chain with the finite state-space S and with $n \geqq 1$ recurrent classes of states $C_1, \ldots, C_n$ and with a set T of transient states. This includes also the irreducible case when $n = 1$ and $T = \emptyset$. We suppose that the states are labeled in such a way that the corresponding matrix of transition probabilities $\mathbf{P}$ has the block form

(5)
$$\mathbf{P} = \begin{pmatrix} \mathbf{P}^{(1)} & 0 & \ldots & 0 & 0 \\ 0 & \mathbf{P}^{(2)} & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & & \mathbf{P}^{(n)} & 0 \\ \mathbf{Q}^{(1)} & \mathbf{Q}^{(2)} & \ldots & \mathbf{Q}^{(n)} & \mathbf{Q}^{(n+1)} \end{pmatrix}.$$

We shall often need to split vectors from R^k into blocs (or, conversely, to define a vector from R^k blockwise) corresponding to the division of the state space S. If $\mathbf{x} \in R^k$ then we write

$$\mathbf{x} = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)}, \mathbf{x}^{(n+1)}) ,$$

where the vectors $\mathbf{x}^{(a)} \in R^{\operatorname{card} C_a}$ for $a \in \mathscr{N}$ and $\mathbf{x}^{(n+1)} \in R^{\operatorname{card} T}$ contain the components x_i of the vector $\mathbf{x}$ such that $i \in C_a$ and $i \in T$, respectively.

Let us first deal with the case that at least one of the matrices $\mathbf{P}^{(a)}$, $a \in \mathscr{N}$ (say $\mathbf{P}^{(a_0)}$) is periodic. The matrix $\mathbf{P}^{(a)}$ is evidently stochastic and irreducible. As such it may be regarded as a matrix of transition probabilities of a Markov chain. Lemma 1(b) implies the existence of a vector $\boldsymbol{\eta} = (\eta_i)_{i \in C_{a_0}}$ such that $\eta_i \geqq 0$ for $i \in C_{a_0}$, $\sum_{i \in C_{a_0}} \eta_i = 1$ and

(6)
$$\|\boldsymbol{\eta}[\mathbf{P}^{(a_0)}]^2 - \boldsymbol{\eta}\mathbf{P}^{(a_0)}\| = \|\boldsymbol{\eta}\mathbf{P}^{(a_0)} - \boldsymbol{\eta}\| ,$$

(7)
$$\boldsymbol{\eta}\mathbf{P}^{(a_0)} \neq \boldsymbol{\eta} .$$

We put

$$\boldsymbol{p}^{(a_0)} = \boldsymbol{\eta}\,,$$

$$\boldsymbol{p}^{(a)} = 0 \quad \text{for} \quad a \in [\mathcal{N} \cup \{n+1\}] - \{a_0\}\,.$$

It is easily seen that

$$\|\boldsymbol{p}(\boldsymbol{P} - \boldsymbol{I})\,\boldsymbol{P}\| = \|\boldsymbol{\eta}[\boldsymbol{P}^{(a_0)}]^2 - \boldsymbol{\eta}\boldsymbol{P}^{(a_0)}\|\,,$$

$$\|\boldsymbol{p}(\boldsymbol{P} - \boldsymbol{I})\| = \|\boldsymbol{\eta}\boldsymbol{P}^{(a_0)} - \boldsymbol{\eta}\|\,,$$

so that by virtue of (6) and (7) we obtain that $\boldsymbol{p} \in \mathscr{P}_0$ and

$$\|\boldsymbol{p}(\boldsymbol{P} - \boldsymbol{I})\,\boldsymbol{P}\| = \|\boldsymbol{p}(\boldsymbol{P} - \boldsymbol{I})\|\,.$$

With respect to (2), we have thus proved

Lemma 2. *If at least one matrix $\boldsymbol{P}^{(a)}$, where $a \in \mathcal{N}$, in (5) is periodic then $\alpha(\boldsymbol{P}) = 1$.*

Throughout the rest of this section, we shall suppose that all the matrices $\boldsymbol{P}^{(a)}$, $a \in \mathcal{N}$, in (5) are aperiodic. The following lemma yields a simple upper bound of the value of $\alpha(\boldsymbol{P})$.

Lemma 3. *The value of $\alpha(\boldsymbol{P})$ fulfils the inequality*

(8) $$\alpha(\boldsymbol{P}) \leqq \tfrac{1}{2} \max\,\{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|;\ i, j \in S\}\,.$$

Proof. Let $\boldsymbol{p} \in \mathscr{P}_0$ and let us put

$$\boldsymbol{x} = \boldsymbol{p}(\boldsymbol{P} - \boldsymbol{I})\,.$$

Then

$$\sum_{i=1}^{k} x_i = 0\,,$$

so that

(9) $$\sum_{i=1}^{k} x_i^+ = \sum_{i=1}^{k} x_i^- = \frac{\|\boldsymbol{x}\|}{2} \neq 0\,,$$

where

$$x_i^+ = \max\,\{x_i;\ 0\} \qquad \text{for} \quad i \in S\,,$$

$$x_i^- = \max\,\{-x_i;\ 0\} \quad \text{for} \quad i \in S\,.$$

We have

$$\|\boldsymbol{p}(\boldsymbol{P} - \boldsymbol{I})\,\boldsymbol{P}\| = \|\sum_{i=1}^{k} x_i^+ \boldsymbol{r}_i - \sum_{j=1}^{k} x_j^- \boldsymbol{r}_j\| =$$

$$= \frac{2}{\|\boldsymbol{x}\|}\,\|\sum_{i=1}^{k}\sum_{j=1}^{k} x_i^+ x_j^- \boldsymbol{r}_i - \sum_{i=1}^{k}\sum_{j=1}^{k} x_i^+ x_j^- \boldsymbol{r}_j\| \leqq$$

$$\leqq \frac{2}{\|\boldsymbol{x}\|} \sum_{i=1}^{k}\sum_{j=1}^{k} x_i^+ x_j^- \,.\, \max\,\{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|;\ i, j \in S\} =$$

$$= \|\boldsymbol{p}(\boldsymbol{P} - \boldsymbol{I})\|\,.\,\tfrac{1}{2}\max\,\{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|;\ i, j \in S\}\,.$$

This completes the proof.

Let the matrix $\mathbf{W} = (w_{ij})_{i,j \in T}$ be defined by

$$(10) \qquad \mathbf{W} = \sum_{v=0}^{\infty} [\mathbf{Q}^{(n+1)}]^v ,$$

where $[\mathbf{Q}^{(n+1)}]^v$ is the v-th power of $\mathbf{Q}^{(n+1)}$ for v natural, and $[\mathbf{Q}^{(n+1)}]^0 = \mathbf{I}$. The matrix $\mathbf{Q}^{(n+1)}$ corresponds to state transitions within the set T of transient states. Thus, according to Theorem 3.2.1 of [7], the matrix $\mathbf{W}$ is well-defined and

$$(11) \qquad \mathbf{W}^{-1} = \mathbf{I} - \mathbf{Q}^{(n+1)} .$$

We shall now characterize the set of vectors from R^k which can be expressed as $\mathbf{p}(\mathbf{P} - \mathbf{I})$, where $\mathbf{p} \in \mathscr{P}_0$.

Lemma 4. *We put*

$$(12a) \qquad Z = \{x;\ x \in R^k,\ x \neq \mathbf{0},\ x^{(a)} e^{(a)\prime} = 0 \ \textit{for}\ a \in \mathscr{N}\}$$

if $T = \emptyset$, *and*

$$(12b) \qquad Z = \{\mathbf{x};\ \mathbf{x} \in R^k,\ \mathbf{x} \neq \mathbf{0},\ \mathbf{x}^{(n+1)} \mathbf{W} \leqq \mathbf{0} ,$$
$$\mathbf{x}^{(a)} \mathbf{e}^{(a)\prime} = -\mathbf{x}^{(n+1)} \mathbf{W} \mathbf{Q}^{(a)} \mathbf{e}^{(a)\prime} \ \textit{for}\ a \in \mathscr{N}\}$$

provided $T \neq \emptyset$. *The following assertions are valid*:

(a) *If* $\mathbf{p} \in \mathscr{P}_0$ *then* $\mathbf{p}(\mathbf{P} - \mathbf{I}) \in Z$.

(b) *If* $\mathbf{x} \in Z$ *then there exist a positive number* c *and a vector* $\mathbf{p} \in \mathscr{P}_0$ *such that* $\mathbf{x} = c\mathbf{p}(\mathbf{P} - \mathbf{I})$.

(c) *The equality*

$$(13) \qquad \alpha(\mathbf{P}) = \sup \left\{ \frac{\|\mathbf{x}\mathbf{P}\|}{\|\mathbf{x}\|} ;\quad \mathbf{x} \in Z \right\}$$

is true.

Proof. (a) – Let $\mathbf{p} \in \mathscr{P}_0$ and let $\mathbf{x} = \mathbf{p}(\mathbf{P} - \mathbf{I})$.

By the definitions we have

$$\mathbf{x} \in R^k , \quad \mathbf{x} \neq 0 .$$

Further, it is easily seen that

$$[\mathbf{P}^{(a)} - \mathbf{I}]\, \mathbf{e}^{(a)\prime} = 0 \quad \text{for each} \quad a \in \mathscr{N} ,$$

so that if $T = \emptyset$ then

$$\mathbf{x}^{(a)} \mathbf{e}^{(a)\prime} = \mathbf{p}^{(a)} [\mathbf{P}^{(a)} - \mathbf{I}]\, \mathbf{e}^{(a)\prime} = 0 \quad \text{for} \quad a \in \mathscr{N} .$$

Finally, if $T \neq 0$ we obtain that

$$(14) \qquad \mathbf{x}^{(n+1)} \mathbf{W} = \mathbf{p}^{(n+1)} [\mathbf{Q}^{(n+1)} - \mathbf{I}]\, \mathbf{W} = -\mathbf{p}^{(n+1)} \leqq \mathbf{0}$$

and

$$\mathbf{x}^{(a)} \mathbf{e}^{(a)\prime} = \mathbf{p}^{(a)} [\mathbf{P}^{(a)} - \mathbf{I}]\, \mathbf{e}^{(a)\prime} + \mathbf{p}^{(n+1)} \mathbf{Q}^{(a)} \mathbf{e}^{(a)\prime} =$$
$$= -\mathbf{x}^{(n+1)} \mathbf{W} \mathbf{Q}^{(a)} \mathbf{e}^{(a)\prime} \quad \text{for}\ a \in \mathscr{N} .$$

(b) We shall consider the case of $T \neq \emptyset$ since if $T = \emptyset$ then it is necessary only to omit all the terms containing the $(n + 1)$-st blocks of vectors. Let $\boldsymbol{x} \in Z$ be given, and let us solve the matrix equation

$$\boldsymbol{q}(\boldsymbol{P} - \boldsymbol{I}) = \boldsymbol{x} . \tag{15}$$

It is easily seen that (15) is equivalent to the system of matrix equations

$$\boldsymbol{q}^{(a)}[\boldsymbol{P}^{(a)} - \boldsymbol{I}] + \boldsymbol{q}^{(n+1)}\boldsymbol{Q}^{(a)} = \boldsymbol{x}^{(a)} \quad \text{for} \quad a \in \mathcal{N} , \tag{16}$$
$$\boldsymbol{q}^{(n+1)}[\boldsymbol{Q}^{(n+1)} - \boldsymbol{I}] = \boldsymbol{x}^{(n+1)} .$$

The relation (11) implies

$$\boldsymbol{q}^{(n+1)} = -\boldsymbol{x}^{(n+1)}\boldsymbol{W} , \tag{17}$$

so that (16) may be rewritten as follows:

$$\boldsymbol{q}^{(a)}[\boldsymbol{P}^{(a)} - \boldsymbol{I}] = \boldsymbol{x}^{(a)} + \boldsymbol{x}^{(n+1)}\boldsymbol{W}\boldsymbol{Q}^{(a)} \quad \text{for} \quad a \in \mathcal{N} . \tag{18}$$

The matrix $\boldsymbol{P}^{(a)}$ corresponds to the recurrent class C_a. The well-known fact, quoted e.g. in [6], that the number 1 is a simple characteristic root of $\boldsymbol{P}^{(a)}$ implies that the rank of the matrix of the system (18) of linear equations, i.e. of $\boldsymbol{P}^{(a)} - \boldsymbol{I}$, is equal to card $C_a - 1$. Further, we have

$$[\boldsymbol{P}^{(a)} - \boldsymbol{I}]\, \mathbf{e}^{(a)'} = 0 \quad \text{for} \quad a \in \mathcal{N} ,$$

and

$$[\boldsymbol{x}^{(a)} + \boldsymbol{x}^{(n+1)}\boldsymbol{W}\boldsymbol{Q}^{(a)}]\, \mathbf{e}^{(a)'} = 0 \quad \text{for} \quad a \in \mathcal{N} ,$$

because of the assumption that $\boldsymbol{x} \in Z$. Thus, we see that the vector $\mathbf{e}^{(a)}$ is orthogonal to all the rows of the extended matrix of the system (18) which has the form

$$\begin{pmatrix} \boldsymbol{P}^{(a)} - \boldsymbol{I} \\ \boldsymbol{x}^{(a)} + \boldsymbol{x}^{(n+1)}\boldsymbol{W}\boldsymbol{Q}^{(a)} \end{pmatrix} \quad \text{for} \quad a \in \mathcal{N} .$$

So, we find for each $a \in \mathcal{N}$ that

$$\text{card } C_a - 1 = r(\boldsymbol{P}^{(a)} - \boldsymbol{I}) \leqq r\begin{pmatrix} \boldsymbol{P}^{(a)} - \boldsymbol{I} \\ \boldsymbol{x}^{(a)} + \boldsymbol{x}^{(n+1)}\boldsymbol{W}\boldsymbol{Q}^{(a)} \end{pmatrix} \leqq \text{card } C_a - 1$$

($r(.)$ denotes the rank of a matrix), i.e., Frobenius' condition is fulfilled. Let us fix a solution $\boldsymbol{q}^{(a)}$ of (18) for each $a \in \mathcal{N}$. The vector $\boldsymbol{q} = (\boldsymbol{q}^{(1)}, \ldots, \boldsymbol{q}^{(n)}, \boldsymbol{q}^{(n+1)})$, where the first n blocks have been just determined, and the $(n + 1)$-st block is given by (17), is a solution to the matrix equation (15), and moreover,

$$\boldsymbol{q}^{(n+1)} \geqq \boldsymbol{0} \tag{19}$$

because of the assumption that $\boldsymbol{x} \in Z$.

All the recurrent classes C_a, $a \in \mathcal{N}$, are assumed to be aperiodic. Thus, there is a single vector $\pi^{(a)}$ of stationary distribution of probability corresponding to a homogeneous Markov chain with the state-space C_a and with the matrix of transition probabilities $\mathbf{P}^{(a)}$ for each $a \in \mathcal{N}$. We put $\pi^{(n+1)} = \mathbf{0}$. The vector

$$\pi = (\pi^{(1)}, \ldots, \pi^{(n)}, \pi^{(n+1)})$$

possesses the properties

$$\pi_i > 0 \quad \text{for all} \quad i \in S - T,$$

$$\pi_i = 0 \quad \text{for all} \quad i \in T,$$

$$\sum_{i=1}^{k} \pi_i = n .$$

By virtue of the inequality (19), there is a positive number F such that

$$\pi + F \, . \, \mathbf{q} \geqq \mathbf{0} \tag{20}$$

and such that there is a recurrent state $i_0 \in S - T$ fulfilling

$$\pi_{i_0} + F\mathbf{q}_{i_0} > 0 . \tag{21}$$

Finally, we put

$$\mathbf{p} = \frac{1}{n + FG} (\pi + F \, . \, \mathbf{q}) , \tag{22}$$

where

$$G = \sum_{i=1}^{k} q_i . \tag{23}$$

The inequalities (20) and (21) imply

$$n + FG = \sum_{i=1}^{k} (\pi_i + Fq_i) > 0 .$$

The vector $\mathbf{p}$ defined by (22) evidently fulfils

$$p_i \geqq 0 \quad \text{for each} \quad i \in S ,$$

$$\sum_{i=1}^{k} p_i = 1 ,$$

and

$$\mathbf{p}(\mathbf{P} - \mathbf{I}) = \frac{F}{n + FG} \mathbf{q}(\mathbf{P} - \mathbf{I}) = \frac{F}{n + FG} \mathbf{x} ,$$

so that $\mathbf{p}$ meets all the requirements of the assertion (b) of this lemma (with $c = (n + FG)/F$).

The assertion (c) is a straightforward consequence of the statements (a) and (b).

The set Z contains infinitely many elements. For this reason, it is impossible to determine the value of $\alpha(\mathbf{P})$ directly by evaluating $\|\mathbf{xP}\|/\|\mathbf{x}\|$ for all $\mathbf{x} \in Z$. The following lemmas state, however, that it is sufficient to take into consideration only vectors belonging to a finite set $U \cap V$.

Lemma 5. *Let*

$$\beta = \max_{a \in \mathcal{N}} \max \{\|\mathbf{r}_i - \mathbf{r}_j\|; \; i, j \in C_a\} \tag{24}$$

and let i_1 and i_2 be two different states belonging to the same recurrent class (say C_{a_0}) which fulfil

$$\|\mathbf{r}_{i_1} - \mathbf{r}_{i_2}\| = \beta\,. \tag{25}$$

Then the vector $\mathbf{b}$ with components

$$b_{i_1} = \quad 1\,, \tag{26}$$

$$b_{i_2} = -1\,, \tag{27}$$

$$b_i = 0 \quad \textit{for each} \quad i \in S - \{i_1; i_2\}\,, \tag{28}$$

is an element of the set Z. Further, if $\mathbf{x} \in Z$ is such that $\mathbf{x}^{(n+1)} = 0$ (in the case $T = \emptyset$ let $\mathbf{x}$ be an arbitrary element of Z) then

$$\frac{\|\mathbf{xP}\|}{\|\mathbf{x}\|} \leqq \frac{\|\mathbf{bP}\|}{\|\mathbf{b}\|} = \frac{\beta}{2}\,. \tag{29}$$

Proof. The fact that the vector $\mathbf{b}$ determined above belongs to the set Z is quite evident. Further, let $\mathbf{x} \in Z$ and let $\mathbf{x}^{(n+1)} = 0\; T \neq \emptyset$.

We put

$$\mathcal{N}_0 = \{a;\; a \in \mathcal{N},\; \mathbf{x}^{(a)} \neq \mathbf{0}\}\,.$$

According to Lemmas 3 and 4(c) we obtain that

$$\frac{\|\mathbf{xP}\|}{\|\mathbf{x}\|} = \frac{\sum_{a\in\mathcal{N}_0} \|\mathbf{x}^{(a)}\mathbf{P}^{(a)}\|}{\sum_{a\in\mathcal{N}_0} \|\mathbf{x}^{(a)}\|} = \frac{\sum_{a\in\mathcal{N}_0} \frac{\|\mathbf{x}^{(a)}\mathbf{P}^{(a)}\|}{\|\mathbf{x}^{(a)}\|} \|\mathbf{x}^{(a)}\|}{\sum_{a\in\mathcal{N}_0} \|\mathbf{x}^{(a)}\|} \leqq$$

$$\leqq \frac{\sum_{a\in\mathcal{N}_0} \frac{1}{2} \max\{\|\mathbf{r}_i - \mathbf{r}_j\|;\; i, j \in C_a\}\, \|\mathbf{x}^{(a)}\|}{\sum_{a\in\mathcal{N}_0} \|\mathbf{x}^{(a)}\|} = \frac{\beta}{2}\,.$$

On the other hand, we have

$$\frac{\|\mathbf{bP}\|}{\|\mathbf{b}\|} = \tfrac{1}{2}\|b_{i_1} r_{i_1} + b_{i_2} r_{i_2}\| = \frac{\beta}{2}\,.$$

Lemma 6. *Let $T \neq \emptyset$ and let $\mathbf{y} \in Z$ be such that $\mathbf{y}^{(n+1)} \neq \mathbf{0}$. For each $a \in \mathcal{N}$ we put*

$$\mathbf{z}^{(a)}(\mathbf{y}) = \quad \mathbf{y}^{(n+1)}\mathbf{Q}^{(a)}\,, \tag{30}$$

$$f_a(\mathbf{y}) \;\; = -\mathbf{y}^{(n+1)}\mathbf{W}\mathbf{Q}^{(a)}\mathbf{e}^{(a)'}\,, \tag{31}$$

$$i(a, \mathbf{y}) = \min\{i;\; i \in C_a,\; \|f_a(\mathbf{y})\, \mathbf{r}_i^{(a)} + \mathbf{z}^{(a)}(\mathbf{y})\| \geqq \tag{32}$$
$$\geqq \|f_a(\mathbf{y})\, \mathbf{r}_j^{(a)} + \mathbf{z}^{(a)}(\mathbf{y})\| \; \textit{for each}\; j \in C_a\}\,.$$

If $\mathbf{x} \in Z$ is such that $\mathbf{x}^{(n+1)} \neq \mathbf{0}$ then

$$\frac{\|\mathbf{xP}\|}{\|\mathbf{x}\|} \leqq \sup\left\{\frac{\|\mathbf{yP}\|}{\|\mathbf{y}\|};\quad \mathbf{y} \in U \cup \{\mathbf{b}\}\right\}, \tag{33}$$

where

$$(34)\qquad U = \{\mathbf{y};\ \mathbf{y} \in Z,\ \mathbf{y}^{(n+1)} \neq \mathbf{0},\ y_{i(a,\mathbf{y})} = f_a(\mathbf{y}),$$

$$y_i = 0 \textit{ for } i \in C_a - \{i(a, \mathbf{y})\},\ \textit{for each } a \in \mathcal{N}\},$$

and the vector **b** *is determined by* (24) *through* (28).

Proof. Let $\mathbf{x} \in Z - U$ be such that $\mathbf{x}^{(n+1)} \neq \mathbf{0}$. We shall introduce a vector $\tilde{\mathbf{x}} \in R^k$ fulfilling

$$\tilde{\mathbf{x}}^{(n+1)} = \mathbf{x}^{(n+1)},$$

$$\tilde{x}_{i(a,\mathbf{x})} = f_a(\mathbf{x}) \quad \text{for} \quad a \in \mathcal{N},$$

$$\tilde{x}_i = 0 \quad \text{for} \quad i \in C_a - \{i(a, \mathbf{x})\},\quad a \in \mathcal{N}.$$

It is easily seen that $\tilde{\mathbf{x}} \in U$ and that $\tilde{\mathbf{x}}^{(n+1)}\mathbf{Q}^{(n+1)} = \mathbf{x}^{(n+1)}\mathbf{Q}^{(n+1)}$. Further, we have

$$(35)\qquad \|\mathbf{xP}\| = \sum_{a=1}^{n} \|\mathbf{x}^{(a)}\mathbf{P}^{(a)} + \mathbf{z}^{(a)}(\mathbf{x})\| + \|\mathbf{x}^{(n+1)}\mathbf{Q}^{(n+1)}\|.$$

Following the arguments used in the proof of Lemma 3, we obtain for each $a \in \mathcal{N}$

$$\|\mathbf{x}^{(a)}\mathbf{P}^{(a)} + \mathbf{z}^{(a)}(\mathbf{x})\| = \Big\| \sum_{j \in C_a^+(\mathbf{x})} \frac{x_j}{f_a^+(\mathbf{x})} (f_a(\mathbf{x})\, \mathbf{r}_j^{(a)} + \mathbf{z}^{(a)}(\mathbf{x})) +$$

$$+ f_a^-(\mathbf{x}) \Big(\sum_{i \in C_a^+(\mathbf{x})} \frac{x_i}{f_a^+(\mathbf{x})} \mathbf{r}_i^{(a)} + \sum_{j \in C_a^-(\mathbf{x})} \frac{x_j}{f_a^-(\mathbf{x})} \mathbf{r}_j^{(a)} \Big) \Big\| \leqq$$

$$\leqq \|f_a(\mathbf{x})\, \mathbf{r}_{i(a,\mathbf{x})}^{(a)} + \mathbf{z}^{(a)}(\tilde{\mathbf{x}})\| + |f_a^-(\mathbf{x})| \,.\, \beta,$$

so that according to (29)

$$(36)\qquad \frac{\|\mathbf{xP}\|}{\|\mathbf{x}\|} \leqq \frac{\|\tilde{\mathbf{x}}\mathbf{P}\| + \beta \sum_{a=1}^{n} |f_a^-(\mathbf{x})|}{\|\tilde{\mathbf{x}}\| + 2\sum_{a=1}^{n} |f_a^-(\mathbf{x})|} \leqq \max\left\{\frac{\|\tilde{\mathbf{x}}\mathbf{P}\|}{\|\tilde{\mathbf{x}}\|}; \frac{\|\mathbf{bP}\|}{\|\mathbf{b}\|}\right\} \leqq \sup\left\{\frac{\|\mathbf{yP}\|}{\|\mathbf{y}\|};\ \mathbf{y} \in U \cup \{\mathbf{b}\}\right\}.$$

Lemma 7. *Let* $T \neq \emptyset$ *and let* $\mathcal{U}, \mathcal{V} \subset T$ *be such that* $\mathcal{U} \neq T$, $\mathcal{V} \neq T$. *Let* $\mu(\mathcal{U}, \mathcal{V})$ *denote the dimension of the vector space of solutions* $\mathbf{g}^{(n+1)} = (g_i)_{i\in T}$ *to the system of homogeneous linear equations*

$$(37)\qquad g_i = 0 \quad \textit{for} \quad i \in \mathcal{U},$$

$$(38)\qquad [\mathbf{g}^{(n+1)}\mathbf{W}]_j = 0 \quad \textit{for} \quad j \in \mathcal{V},$$

where

$$[\mathbf{g}^{(n+1)}\mathbf{W}]_j = \sum_{h\in T} g_h w_{hj} \quad \textit{for} \quad j \in T.$$

Let us put

$$(39)\qquad \Phi = \{(\mathcal{U}, \mathcal{V});\ \mathcal{U} \subset T,\ \mathcal{V} \subset T,\ \mathcal{U} \neq T,\ \mathcal{V} \neq T,\ \mu(\mathcal{U}, \mathcal{V}) = 1\}$$

and

(40) $\quad V = \{\mathbf{y};\ \mathbf{y} \in Z,\ \mathbf{y}^{(n+1)} \neq \mathbf{0},\ \mathbf{y}^{(n+1)}$ *is a solution to the system* (37) *and* (38) *generated by a couple* $(\mathcal{U}, \mathcal{V}) \in \Phi\}$.

If $\mathbf{x} \in Z$ *is such that* $\mathbf{x}^{(n+1)} \neq \mathbf{0}$ *then*

$$\frac{\|\mathbf{x}P\|}{\|\mathbf{x}\|} \leqq \sup\left\{\frac{\|\mathbf{y}P\|}{\|\mathbf{y}\|};\quad \mathbf{y} \in V\right\}. \tag{41}$$

Proof. Let $\mathbf{x} \in Z - V$ be such that

$$\mathbf{x}^{(n+1)} \neq \mathbf{0}\,. \tag{42}$$

Let us put

$$\mathcal{U}(\mathbf{x}) = \{i;\ i \in T,\ x_i = 0\}\,,$$

$$\mathcal{V}(\mathbf{x}) = \{j;\ j \in T,\ [\mathbf{x}^{(n+1)}\mathbf{W}]_j = 0\}\,.$$

It is easy to see that $\mathcal{U}(\mathbf{x}), \mathcal{V}(\mathbf{x}) \subset T$. Further, we obtain by virtue of (42) and of the regularity of the matrix $\mathbf{W}$ that $\mathcal{U}(\mathbf{x}) \neq T$ and $\mathcal{V}(\mathbf{x}) \neq T$. Finally, the vector $\mathbf{x}^{(n+1)}$ is evidently a solution to the system (37) and (38) generated by the couple of sets $\mathcal{U}(\mathbf{x})$ and $\mathcal{V}(\mathbf{x})$. The assumption $\mathbf{x} \notin V$ thus implies that

$$\mu(\mathcal{U}(\mathbf{x}), \mathcal{V}(\mathbf{x})) > 1\,,$$

so that we find every system involving the equations (37) for $i \in \mathcal{U}(\mathbf{x})$, the equation (38) for $j \in \mathcal{V}(\mathbf{x})$, and an arbitrary single homogeneous linear equation, possesses a non-trivial solution. Let us denote by $\psi^{(n+1)} = (\psi_i)_{i \in T} \neq \mathbf{0}$ the vector fulfilling (37) with $\mathcal{U} = \mathcal{U}(\mathbf{x})$, (38) with $\mathcal{V} = \mathcal{V}(\mathbf{x})$, and, moreover,

$$\sum_{i \in T^-(\mathbf{x})} \psi_i = 0\,. \tag{43}$$

Let the vector $\varepsilon \in R^k$ have the components

$$\varepsilon_i = \psi_i \quad \text{for} \quad i \in T\,,$$

$$\varepsilon_i = -\frac{x_i}{f_a^+(\mathbf{x})}\,\psi^{(n+1)}\mathbf{W}\mathbf{Q}^{(a)}\mathbf{e}^{(a)\prime} \quad \text{for} \quad i \in C_a^+(\mathbf{x})\,, \quad a \in \mathcal{N}\,,$$

$$\varepsilon_i = 0 \quad \text{for} \quad i \in C_a - C_a^+(\mathbf{x})\,, \quad a \in \mathcal{N}\,.$$

Let us put

$$\lambda_1 = \sup\{\lambda;\ \lambda > 0,\ \operatorname{sing}(x_i + \lambda\varepsilon_i) = \operatorname{sign} x_i \text{ for each } i \in T,\ (\mathbf{x}^{(n+1)} + \lambda\varepsilon^{(n+1)})\,\mathbf{W} \leqq \mathbf{0}\}\,,$$

$$\lambda_2 = \sup\{\lambda;\ \lambda > 0,\ \operatorname{sign}(x_i - \lambda\varepsilon_i) = \operatorname{sign} x_i \text{ for each } i \in T,\ (\mathbf{x}^{(n+1)} - \lambda\varepsilon^{(n+1)})\,\mathbf{W} \leqq \mathbf{0}\}\,,$$

and

$$^\sharp\mathbf{x} = \mathbf{x} + \lambda_1\varepsilon\,, \quad ^\flat\mathbf{x} = \mathbf{x} - \lambda_2\varepsilon\,.$$

The reader can verify that

$$0 < \lambda_1 < \infty\,, \quad 0 < \lambda_2 < \infty\,,$$

$$^{\#}\mathbf{x}_i = {}^{\flat}\mathbf{x}_i = 0 \quad \text{for} \quad i \in \mathscr{U}(\mathbf{x})\,,$$

$$[^{\#}\mathbf{x}^{(n+1)}\mathbf{W}]_j = [^{\flat}\mathbf{x}^{(n+1)}\mathbf{W}]_j = 0 \quad \text{for} \quad j \in \mathscr{V}(\mathbf{x})\,,$$

$$^{\#}\mathbf{x}^{(n+1)}\mathbf{W} \leqq \mathbf{0}\,, \quad {}^{\flat}\mathbf{x}^{(n+1)}\mathbf{W} \leqq \mathbf{0}\,, \tag{44}$$

$$\mathbf{x} = \frac{\lambda_2}{\lambda_1 + \lambda_2}\, {}^{\#}\mathbf{x} + \frac{\lambda_1}{\lambda_1 + \lambda_2}\, {}^{\flat}\mathbf{x}\,. \tag{45}$$

Further, we have

$$\|^{\#}\mathbf{x}\| = \sum_{a=1}^{n} \sum_{i \in C_a{}^+(\mathbf{x})} \left(x_i - \lambda_1 \frac{x_i}{f_a^+(\mathbf{x})}\, \psi^{(n+1)}\mathbf{W}\mathbf{Q}^{(a)}\mathbf{e}^{(a)'} \right) + \tag{46}$$

$$+ \sum_{i \in T^+(\mathbf{x})} (x_i + \lambda_1 \psi_i) - \sum_{a=1}^{n} \sum_{i \in C_a{}^-(\mathbf{x})} x_i - \sum_{i \in T^-(\mathbf{x})} (x_i + \lambda_1 \psi_i) =$$

$$= \|\mathbf{x}\| - 2\lambda_1 \sum_{i \in T^-(\mathbf{x})} \psi_i +$$

$$+ \lambda_1 [\psi^{(n+1)}\mathbf{W}(\mathbf{I} - \mathbf{Q}^{(n+1)})\, \mathbf{e}^{(a)'} - \sum_{a=1}^{n} \psi^{(n+1)}\mathbf{W}\mathbf{Q}^{(a)}\mathbf{e}^{(a)'}] = \|\mathbf{x}\|$$

in view of the fact that $\psi^{(n+1)}$ is a solution to the system (37), (38) and (43). A similar reasoning produces the equality

$$\|^{\flat}\mathbf{x}\| = \|\mathbf{x}\|\,. \tag{47}$$

In this case we must, however, take into account the following relations:

$$x_i - \lambda_2 \varepsilon_i = x_i \quad \text{for each} \quad i \in C_a - C_a^+(\mathbf{x})\,, \quad a \in \mathscr{N}$$

and

$$x_i - \lambda_2 \varepsilon_i = \frac{x_i}{f_a^+(\mathbf{x})} [f_a(\mathbf{x}) + \lambda_2 \psi^{(n+1)}\mathbf{W}\mathbf{Q}^{(a)}\mathbf{e}^{(a)'}] =$$

$$= \frac{x_i}{f_a^+(\mathbf{x})} \left(-\mathbf{x}^{(n+1)} + \lambda_2 \varepsilon^{(n+1)}\right) \mathbf{W}\mathbf{Q}^{(a)}\mathbf{e}^{(a)'} =$$

$$= -\frac{x_i}{f_a^+(\mathbf{x})}\, {}^{\flat}\mathbf{x}^{(n+1)}\mathbf{W}\mathbf{Q}^{(a)}\mathbf{e}^{(a)'} \geqq 0$$

for each $i \in C_a^+(\mathbf{x})$, $a \in \mathscr{N}$. We used the assumption $\mathbf{x} \in Z$, the relation (44) and the evident fact that $\mathbf{Q}^{(a)}\mathbf{e}^{(a)'} \geqq 0$ for each $a \in \mathscr{N}$.

The relations (45), (46) and (47) imply the inequality

$$\frac{\|\mathbf{x}\mathbf{P}\|}{\|\mathbf{x}\|} \leqq \frac{\max\{\|^{\#}\mathbf{x}\mathbf{P}\|; \|^{\flat}\mathbf{x}\mathbf{P}\|\}}{\|\mathbf{x}\|} = \max\left\{\frac{\|^{\#}\mathbf{x}\mathbf{P}\|}{\|^{\#}\mathbf{x}\|}; \frac{\|^{\flat}\mathbf{x}\mathbf{P}\|}{\|^{\flat}\mathbf{x}\|}\right\}. \tag{48}$$

Let us suppose that

$$(49) \qquad \frac{\|{}^{\#}\mathbf{x}\mathbf{P}\|}{\|{}^{\#}\mathbf{x}\|} \geqq \frac{\|{}^{b}\mathbf{x}\mathbf{P}\|}{\|{}^{b}\mathbf{x}\|}.$$

(If the reverse inequality to (49) is true then substitute below ${}^{b}\mathbf{x}$ for ${}^{\#}\mathbf{x}$, and λ_2 for λ_1.) Let us put

$$\mathscr{U}(\mathbf{x}, {}^{\#}\mathbf{x}) = \{i;\ i \in T - \mathscr{U}(\mathbf{x}),\ \bar{x}_i = 0\},$$

$$\mathscr{V}(\mathbf{x}, {}^{\#}\mathbf{x}) = \{j;\ j \in T - \mathscr{V}(\mathbf{x}),\ [{}^{\#}\mathbf{x}^{(n+1)}\mathbf{W}]_j = 0\}.$$

From the definition of λ_1 we find that

$$(50) \qquad \mathscr{U}(\mathbf{x}, {}^{\#}\mathbf{x}) \cup \mathscr{V}(\mathbf{x}, {}^{\#}\mathbf{x}) \neq \emptyset.$$

The vector ${}^{\#}\mathbf{x}^{(n+1)}$ is obviously a solution to the system (37) with $\mathscr{U} = \mathscr{U}({}^{\#}\mathbf{x}) \supset \mathscr{U}(\mathbf{x})$ and (38) with $\mathscr{V} = \mathscr{V}({}^{\#}\mathbf{x}) \supset \mathscr{V}(\mathbf{x})$. This system arises from that generated by the couple of sets $\mathscr{U}(\mathbf{x})$ and $\mathscr{V}(\mathbf{x})$ by adding the equations $g_i = 0$ for $i \in \mathscr{U}(\mathbf{x}, {}^{\#}\mathbf{x})$, and $[\mathbf{g}^{(n+1)}\mathbf{W}]_j = 0$ for $j \in \mathscr{V}(\mathbf{x}, {}^{\#}\mathbf{x})$. In virtue of (50), the vector $\mathbf{x}^{(n+1)}$ is not a solution to this system so that

$$\mu(\mathscr{U}({}^{\#}\mathbf{x}), \mathscr{V}({}^{\#}\mathbf{x})) < \mu(\mathscr{U}(\mathbf{x}), \mathscr{V}(\mathbf{x})).$$

If $\mu(\mathscr{U}({}^{\#}\mathbf{x}), \mathscr{V}({}^{\#}\mathbf{x})) = 1$ then obviously ${}^{\#}\mathbf{x} \in V$, and (48) and (49) imply that ${}^{\#}\mathbf{x}$ fulfils

$$\frac{\|\mathbf{x}\mathbf{P}\|}{\|\mathbf{x}\|} \leqq \frac{\|{}^{\#}\mathbf{x}\mathbf{P}\|}{\|{}^{\#}\mathbf{x}\|}.$$

On the other hand, if $\mu(\mathscr{U}({}^{\#}\mathbf{x}), \mathscr{V}({}^{\#}\mathbf{x})) > 1$ then a finite number (not exceeding card $T - 2$) of applications of the process described above yields a vector $\mathbf{y} \in V$ such that

$$\frac{\|\mathbf{x}\mathbf{P}\|}{\|\mathbf{x}\|} \leqq \frac{\|\mathbf{y}\mathbf{P}\|}{\|\mathbf{y}\|}.$$

This completes the proof.

4. MAIN RESULTS

We are now able to express the value of $\alpha(\mathbf{P})$ for any quadratic stochastic matrix $\mathbf{P}$ as the maximum of a finite set. We shall give a method how to calculate $\alpha(\mathbf{P})$ from the elements p_{ij}, $i, j \in S$, of the matrix $\mathbf{P}$. The process will be illustrated by a numerical example in the next section.

Theorem 2. *If $T = \emptyset$ then*

$$(51) \qquad \alpha(\mathbf{P}) = \frac{\beta}{2},$$

and if $T \neq \emptyset$ *then*

$$\alpha(\mathbf{P}) = \max\left[\left\{\frac{\beta}{2}\right\} \cup \{\|\mathbf{yP}\|;\ \mathbf{y} \in U \cap V,\ \|\mathbf{y}\| = 1\}\right], \tag{52}$$

where β, U, *and* V *are defined by* (24), (34) *and* (40), *respectively.*

Proof. Let $T \neq \emptyset$ and let $\mathbf{y} \in U \cap V$. The components of the vector $\mathbf{y}$ corresponding to the recurrent states are uniquely determined, according to Lemma 6, by $\mathbf{y}^{(n+1)}$ because $f_a(\mathbf{y})$, $\mathbf{z}^{(a)}(\mathbf{y})$ and $i(a, \mathbf{y})$ obviously depend only on the $(n+1)$-st block $\mathbf{y}^{(n+1)}$ of the vector $\mathbf{y}$ for each $a \in \mathcal{N}$. Further, we know from Lemma 7 that $\mathbf{y}^{(n+1)}$ must be a solution to the system (37) and (38) generated by a couple$(\mathcal{U}, \mathcal{V}) \in \Phi$, where the set Φ is defined by (39). We conclude that the number of elements of the set

$$\{\mathbf{y};\ \mathbf{y} \in U \cap V,\ \|\mathbf{y}\| = 1\} \tag{53}$$

is less than or equal to the number of the couples $(\mathcal{U}, \mathcal{V})$ possessing the properties

$$\mathcal{U} \subset T,\quad \mathcal{V} \subset T,\quad \mathcal{U} \neq T,\quad \mathcal{V} \neq T, \tag{54}$$

which is equal to

$$(2^{\operatorname{card} T} - 1)^2 .$$

Thus, the set (53) is finite and the maximum on the right-hand side of (52) exists for every stochastic matrix $\mathbf{P}$ provided $T \neq \emptyset$.

The rest of the proof is divided into two parts:

1) If there is $a \in \mathcal{N}$ such that the matrix $\mathbf{P}^{(a)}$ is periodic then $\mathbf{P}^{(a)}$ may be modified to the form (4) so that $\beta = 2$. If $T \neq \emptyset$ then obviously $\|\mathbf{yP}\| \leqq 1$ holds for each $\mathbf{y} \in U \cap V$ such that $\|\mathbf{y}\| = 1$. Lemma 2 implies the validity of the relations (51) and (52).

2) Let the matrices $\mathbf{P}^{(a)}$ be aperiodic for all $a \in \mathcal{N}$. Then (51) and (52) are direct consequences of Lemmas 4(c), 5, 6 and 7.

Corollary. *If* $\mathbf{P}$ *is an irreducible stochastic matrix then*

$$\alpha(\mathbf{P}) = \tfrac{1}{2} \max\{\|\mathbf{r}_i - \mathbf{r}_j\|;\ i, j \in S\} .$$

Let us briefly describe the process of determining the value of $\alpha(\mathbf{P})$ based on Theorem 2 in the case $T \neq \emptyset$. We introduce the systems (37) and (38) of homogeneous linear equations generated by couples $(\mathcal{U}, \mathcal{V})$ fulfilling (54). Further, we find all the couples $(\mathcal{U}, \mathcal{V})$ such that $\mu(\mathcal{U}, \mathcal{V}) = 1$. For each couple $(\mathcal{U}, \mathcal{V})$ with this property we proceed as follows:

We solve the corresponding system (37) and (38). The solutions have the form $\tau \,.\, \xi^{(n+1)}$, where τ is a free real parameter and $\xi^{(n+1)} = (\xi_i)_{i \in T}$ is a fixed vector. There are two possibilities:

1) If $\xi^{(n+1)}\mathbf{W} \leqq 0$ or $\xi^{(n+1)}\mathbf{W} \geqq 0$, then we put $\mathbf{y}^{(n+1)} = \xi^{(n+1)}$ and $\mathbf{y}^{(n+1)} = -\xi^{(n+1)}$, respectively. Further, we evaluate $\mathbf{z}^{(a)}(\mathbf{y})$, $f_a(\mathbf{y})$ and $i(a, \mathbf{y})$ by (30),

(31) and (32). Finally, we put

$$y_{i(a,\mathbf{y})} = f_a(\mathbf{y}) \quad \text{for} \quad a \in \mathcal{N}\,,$$

$$y_i = 0 \quad \text{for} \quad i \in \bigcup_{a=1}^{n} \left[C_a - \{i(a, \mathbf{y})\}\right],$$

and calculate the value of

$$\frac{\|\mathbf{y}\mathbf{P}\|}{\|\mathbf{y}\|}\,. \tag{55}$$

2) If there exist $j_1, j_2 \in T$ such that $[\xi^{(n+1)}\mathbf{W}]_{j_1} > 0$ and $[\xi^{(n+1)}\mathbf{W}]_{j_2} < 0$ then no vector $\mathbf{y} \in V$ has the property that its $(n+1)$-st block $\mathbf{y}^{(n+1)}$ is a solution to the system (37) and (38) generated by the couple $(\mathcal{U}, \mathcal{V})$ just considered. We do not pay any more attention to it.

The application of the process just described produces a set of numbers (55). The value of $\alpha(\mathbf{P})$ is equal to the maximum of these numbers and of the number $\frac{1}{2}\beta$, where β is defined by (24).

We shall now turn to the problem of a characterization of stochastic (reducible) matrices such that the strong inequality (3) holds for every vector $\mathbf{p}(t) \in \mathcal{P}_0$.

Theorem 3. *Let*

$$\varrho_{ij} = \sum_{m=1}^{k} p_{im} p_{jm} \quad \textit{for all} \quad i, j \in S\,. \tag{56}$$

Then the strong inequality (3) *holds for every vector* $\mathbf{p}(t) \in \mathcal{P}_0$, $t = 0, 1, 2, \ldots$, *if and only if the following conditions are fulfilled*:

Condition $\mathcal{A}$: *For each* $a \in \mathcal{N}$ *and for each* $i, j \in C_a$, $\varrho_{ij} > 0$.

Condition $\mathcal{B}$: *For each* $i \in T$ *there exists* $a \in \mathcal{N}$ *such that* $\varrho_{ij} > 0$ *is true for each* $j \in C_a$.

Proof. We first prove the necessity of the conditions $\mathcal{A}$ and $\mathcal{B}$. It is seen that the relations $\varrho_{ij} = 0$ and $\|\mathbf{r}_i - \mathbf{r}_j\| = 2$ are equivalent. Thus, if the matrix $\mathbf{P}$ does not fulfil the condition $\mathcal{A}$ then $\beta = 2$ and the vector $\mathbf{b}$ defined in Lemma 5 has the properties $\mathbf{b} \in Z$ and

$$\|\mathbf{b}\mathbf{P}\| = \|\mathbf{b}\|\,. \tag{57}$$

On the other hand, if the matrix $\mathbf{P}$ does not fulfil the condition $\mathcal{B}$ (then $T \neq \emptyset$, of course) then we denote by u and v_a such states that $u \in T$, $v_a \in C_a$ for $a \in \mathcal{N}$, and

$$\varrho_{u,v_a} = 0 \quad \text{for each} \quad a \in \mathcal{N}\,. \tag{58}$$

A vector $\mathbf{y}$ with components

$$y_u = -1\,, \quad y_i = 0 \quad \text{for} \quad i \in T - \{u\}\,,$$

and

$$y_{v_a} = -\mathbf{y}^{(n+1)}\mathbf{W}\mathbf{Q}^{(a)}\mathbf{e}^{(a)\prime} \quad \text{for} \quad a \in \mathcal{N}\,,$$

$$y_i = 0 \quad \text{for} \quad i \in \bigcup_{i=1}^{n} \left[C_a - \{v_a\}\right],$$

is evidently an element of Z and

(59) $$y_{v_a} \geqq 0 \quad \text{for each} \quad a \in \mathcal{N}\,.$$

We put

$$M_u = \{m;\ m \in S,\ p_{um} = 0\}\,.$$

The relations (58) and (59) imply

(60) $$\|\mathbf{yP}\| = \sum_{m=1}^{k} \left|y_u p_{um} + \sum_{a=1}^{n} y_{v_a} p_{v_a m}\right| =$$

$$= \sum_{m \in M_u} \sum_{a=1}^{n} |y_{v_a}|\, p_{v_a m} + \sum_{m \in S - M_u} |y_u|\, p_{um} = \|\mathbf{y}\|\,.$$

In both the situations considered above, we have obtained a vector $\mathbf{x} \in Z$ ($\mathbf{x} = \mathbf{b}$ and $\mathbf{x} = \mathbf{y}$, respectively) fulfilling $\|\mathbf{xP}\| = \|\mathbf{x}\|$. The fact stated by Lemma 4(b) that there is $\mathbf{p}(0) \in \mathcal{P}_0$ such that $\mathbf{p}(1) - \mathbf{p}(0)$ is equal (up to a positive constant) to the vector $\mathbf{x}$ completes the proof of necessity of the conditions $\mathcal{A}$ and $\mathcal{B}$.

Their sufficiency is, due to Lemma 4, a consequence of the following theorem presenting an upper estimate of the value of $\alpha(\mathbf{P})$ This estimate is equal to 1 if and only if $\alpha(\mathbf{P}) = 1$. The simplicity of its evaluation makes it possible to avoid the rather complicated and time-consuming calculations described above in case that it is not necessary to know the exact value of $\alpha(\mathbf{P})$.

Theorem 4. *Let*

(61) $$a_i = \min\left[\{a;\ a \in \mathcal{N},\ \varrho_{ij} > 0 \text{ for all } j \in C_a\} \cup \{n\}\right] \quad \text{for} \quad i \in T\,,$$

(62) $$a_i = a \quad \text{for} \quad i \in C_a\,, \quad a \in \mathcal{N}\,,$$

(63) $$H = \min\left[\{1\} \cup \{\sum_{m \in C_{a_i}} p_{im};\ i \in T\}\right],$$

and

$$\sigma_a = \min\{\varrho_{ij};\ i, j \in C_a\} \quad \text{for} \quad a \in \mathcal{N}\,,$$

$$\omega_i = \min\{\varrho_{ij};\ j \in C_{a_i}\} \quad \text{for} \quad i \in T\,,$$

(64) $$L = \min\left[\{\sigma_a;\ a \in \mathcal{N}\} \cup \{\omega_i;\ i \in T\}\right].$$

Then the inequality

(65) $$\alpha(\mathbf{P}) \leqq 1 - \frac{HL}{k^2}$$

is true. In particular, if the matrix $\mathbf{P}$ *fulfils the conditions* $\mathcal{A}$ *and* $\mathcal{B}$ *then*

(66) $$\alpha(\mathbf{P}) < 1\,.$$

Proof. Let $\mathbf{x} \in Z$. We know from (9) that there is a state $i \in S$ such that

(67) $$x_i \leqq -\frac{\|\mathbf{x}\|}{2k}\,.$$

We shall prove the existence of a state $j \in C^+_{a_i}(x)$ such that

$$x_j \geqq \frac{\|\mathbf{x}\| H}{2k^2}. \tag{68}$$

There are two possibilities:

1) If i is a current state then the relation (67), the assumption $\mathbf{x} \in Z$ and the evident fact that $H \leqq 1$ imply

$$\sum_{m \in C_{a_i}^+(\mathbf{x})} x_m \geqq \mathbf{x}^{(a_i)}\mathbf{e}^{(a_i)'} - x_i \geqq -x_i \geqq \frac{\|\mathbf{x}\| H}{2k}. \tag{69}$$

2) Let $i \in T$. We know from Lemma 4(b) that there exist a vector $\mathbf{p} \in \mathscr{P}_0$ and a positive number c such that $\mathbf{x} = c\mathbf{p}(\mathbf{P} - \mathbf{I})$. We put ${}^b\mathbf{p} = \mathbf{p}\mathbf{P}$. Then

$$cp_i = c\, {}^b p_i - x_i \geqq -x_i,$$

and

$$\sum_{m \in C_{a_i}^+(\mathbf{x})} x_m \geqq c\mathbf{p}^{(n+1)}\mathbf{Q}^{(a_i)}\mathbf{e}^{(a_i)'} \geqq cp_i \sum_{m \in C_{a_i}} p_{im} \geqq \frac{\|\mathbf{x}\| H}{2k}. \tag{70}$$

The inequalities (69) and (70) imply the existence of a state $j \in C^+_{a_i}(\mathbf{x})$ fulfilling (68) because the set $C^+_{a_i}(\mathbf{x})$ cannot contain more than k elements. Further, it is easily seen that

$$|\gamma - \delta| = \gamma + \delta - 2\min\{\gamma; \delta\} \tag{71}$$

is true for every couple of non-negative numbers γ, δ. Thus, we obtain

$$\|\mathbf{x}\mathbf{P}\| = \sum_{a=1}^{n} \sum_{m \in C_a} \Big| \sum_{h \in C_a^+(\mathbf{x}) \cup T^+(\mathbf{x})} x_h p_{hm} + \sum_{h \in C_a^-(\mathbf{x}) \cup T^-(\mathbf{x})} x_h p_{hm} \Big| + \sum_{m \in T} \Big| \sum_{h \in T} x_h p_{hm} \Big| =$$

$$\leqq \|\mathbf{x}\| - 2 \sum_{m \in C_{a_i}} \min\{x_j p_{jm}; -x_i p_{im}\} \leqq$$

$$\leqq \|\mathbf{x}\| - \frac{\|\mathbf{x}\| H}{k^2} \sum_{m \in C_{a_i}} p_{jm} p_{im} = \|\mathbf{x}\| \left(1 - \frac{H}{k^2} \varrho_{ij}\right) \leqq \|\mathbf{x}\| \left(1 - \frac{HL}{k^2}\right).$$

This implies the validity of the inequality (65) due to Lemma 4.

Finally, if the matrix $\mathbf{P}$ fulfils the conditions $\mathscr{A}$ and $\mathscr{B}$ then obviously

$$\sum_{m \in C_{a_i}} p_{im} \geqq \omega_i > 0 \quad \text{for} \quad i \in T,$$

(provided $T \neq \emptyset$, of course), and

$$\sigma_a > 0 \quad \text{for} \quad a \in \mathscr{N},$$

so that

$$H > 0 \tag{72}$$

and

$$L > 0. \tag{73}$$

The relation (66) is a consequence of (65), (72), and (73).

5. AN ILLUSTRATIVE EXAMPLE

Let $k = 4$ and let

$$\mathbf{P} = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} & 0 & 0 \\ \frac{3}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

It is easily seen that $n = 1$, $C_1 = \{1; 2\}$, $T = \{3; 4\}$, and

$$\mathbf{W} = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}.$$

We shall follow the instructions given in Section 4. Table 1 has 9 rows corresponding to the nine couples $(\mathcal{U}, \mathcal{V})$ fulfilling (55). Each of these rows presents the form of the system (37) and (38), the value of $\mu(\mathcal{U}, \mathcal{V})$, and a particular solution $\xi^{(2)}$ to the system (37) and (38) if $\mu(\mathcal{U}, \mathcal{V}) = 1$. It is easy to verify that $\xi^{(2)}\mathbf{W} = 0$ holds in all the rows in question so that $\mathbf{y}^{(2)} = \xi^{(2)}$. Further, the vector $\mathbf{z}^{(1)}(\mathbf{y})$, the numbers $f_1(\mathbf{y})$ and $i(1, \mathbf{y})$, and the resultant vector $y^{(1)}$ are given. We find that it suffices to calculate the value of $\|\mathbf{yP}\|/\|\mathbf{y}\|$ for the following three vectors $\mathbf{y}$:

$$(0,\ 1,\ \ 0,\ -1)\,,$$
$$(1,\ 0,\ -1,\ \ 0)\,,$$
$$(1,\ 0,\ -2,\ \ 1)\,.$$

Table 1

$\mathcal{U}$	$\mathcal{V}$	the system (37), (38)	$\mu(\mathcal{U}, \mathcal{V})$	a solution $\mathbf{y}^{(2)} = \xi^{(2)}$ to (37), (38)	$\mathbf{z}^{(1)}(\mathbf{y})$	$f_1(\mathbf{y})$	$i(1, \mathbf{y})$	$\mathbf{y}^{(1)}$
$\{3\}$	$\{3\}$	$g_3 = 0$ $g_3 = 0$	1	$(0, -1)$	$(0, -1)$	1	2	$(0, 1)$
$\{3\}$	$\{4\}$	$g_3 = 0$ $\frac{1}{2}g_3 + g_4 = 0$	0	—	—	—	—	—
$\{3\}$	$\emptyset$	$g_3 = 0$	1	$(0, -1)$	$(0, -1)$	1	2	$(0, 1)$
$\{4\}$	$\{3\}$	$g_4 = 0$ $g_3 = 0$	0	—	—	—	—	—
$\{4\}$	$\{4\}$	$g_4 = 0$ $\frac{1}{2}g_3 + g_4 = 0$	0	—	—	—	—	—
$\{4\}$	$\emptyset$	$g_4 = 0$	1	$(-1, 0)$	$(-\frac{1}{2}, 0)$	1	1	$(1, 0)$
$\emptyset$	$\{3\}$	$g_3 = 0$	1	$(0, -1)$	$(0, -1)$	1	2	$(0, 1)$
$\emptyset$	$\{4\}$	$\frac{1}{2}g_3 + g_4 = 0$	1	$(-2, 1)$	$(-1, 1)$	1	1	$(1, 0)$
$\emptyset$	$\emptyset$	—	2	—	—	—	—	—

These values are equal to $\frac{3}{4}$, $\frac{3}{4}$ and $\frac{7}{8}$, respectively. Finally, $\frac{1}{2}\beta = \frac{1}{2}$. Thus, we obtain that

$$\alpha(\mathbf{P}) = \max\left\{\tfrac{3}{4};\ \tfrac{3}{4};\ \tfrac{7}{8};\ \tfrac{1}{2}\right\} = \tfrac{7}{8}\,.$$

Let us remark that $H = \frac{1}{2}$ and $L = \frac{1}{8}$ so that the upper estimate (68) of the value $\alpha(\mathbf{P})$ has the form $\alpha(\mathbf{P}) = 255/256$.

References

[1] *F. L. Bauer, E. Deutsch, J. Stoer:* Abschätzungen für die Eigenwerte positiver linearen Operatoren. Linear Algebra and Applications, *2* (1969), 275—331.

[2] *G. Birkhoff:* Lattice Theory. 3rd ed., Amer. Math. Soc. Colloq. Publ. Vol. *XXV*, Providence, R. I. (1967).

[3] *K. L. Chung:* Markov chains with stationary transition probabilities. Springer-Verlag, Berlin—Göttingen—Heidelberg (1960).

[4] *R. L. Dobrushin:* Central limit theorem for non-stationary Markov chains I, II. Theory Prob. Apl. *1* (1956), 63—80, 329—383 (English translation).

[5] *J. Hajnal:* Weak ergodicity in non-homogeneous Markov chains. Proc. Camb. Phil. Soc. *54* (1958), 233—246.

[6] *S. Karlin:* A first course in stochastic processes. Academic Press, New York and London (1968).

[7] *J. G. Kemeny, J. L. Snell:* Finite Markov chains. D. van Nostrand Comp., New York (1960).

[8] *D. G. Kendall:* Unitary dilatations of Markov transition operators, and the corresponding integral representations for transition-probability matrices. In: Probability and Statistics, The Harald Cramér Volume, U. Grenander (ed.), Stockholm Almqvist and Wiksell (New York: John Wiley and Sons) (1959).

[9] *D. G. Kendall:* Geometric ergodicity and the theory of queues. In: Mathematical methods in the social sciences, K. J. Arrow, S. Karlin, P. Suppes (ed.), Stanford, California (1960).

[10] *P. Kratochvíl:* On convergence of homogeneous Markov chains. Apl. mat. *28* (1983), 2, 116—119.

[11] *A. Paz:* Introduction to Probabilistic Automata. Academic Press, New York (1971).

[12] *Т. А. Сарымсаков:* Основы теории процессов Наркова. Государственное издателство технико-теоретической литературы, Москва (1954).

[13] *T. A. Sarymsakov:* On the theory of inhomogeneous Markov chains (in Russian). Dokl. Akad. Nauk Uzbek. S.S.R. *8* (1956), 3—7.

[14] *E. Seneta:* On the historical development of the theory of finite inhomogeneous Markov chains. Proc. Camb. Phil. Soc. *74* (1973), 507—513.

[15] *E. Seneta:* Coefficients of ergodicity. Structure and applications. Adv. Appl. Prob. *11* (1979), 576—590.

[16] *D. Vere-Jones:* Geometric ergodicity in denumerable Markov chains. Quart. J. Math. Oxford (2), *13* (1962), 7—28.

Authors' address: 115 67 Praha 1, Žitná 25, Czechoslovakia (Matematický ústav ČSAV).