

WORD-FUNCTIONS OF STOCHASTIC AND PSEUDO STOCHASTIC AUTOMATA

PAAVO TURAKAINEN

1. *Introduction.* It is well-known that any word-function $p: X^* \rightarrow R$ of finite rank n can be written in the form

$$(1) \quad p(P) = a b^{l(P)}(p_A(P) + c) \quad \forall P \in X^+$$

where a, b, c are constants and p_A is a word-function generated by a stochastic automaton. However, in the existing proofs (cf. [11], [12], [2], [3]) the number of states of A is much larger than n . The purpose of this paper is to consider the following problem: What is the minimal value of $k(n)$ such that any word-function p of rank n can be written in the form (1) for some $k(n)$ -state stochastic automaton A ?

It is proved that there exists an $(n + 2)$ -state actual doubly stochastic automaton A such that (1) holds for $c = -(n + 2)^{-1}$ and for some positive constants a, b . This is done by modifying the method we used in [11], [12] where we showed that every pseudo stochastic language is stochastic.

It is proved that there exists an $(n + 3)$ -state actual doubly stochastic automaton B such that

$$(2) \quad p(P) = b^{l(P)}(p_B(P) + c) \quad \forall P \in X^+$$

where b is a positive constant and $c = -(n + 3)^{-1}$. Moreover, B has a fixed initial state which is also the only final state. In fact this implies that every stochastic language L is quasidefinite in the sense of Paz [6] and that $L + \{\lambda\}$ is accepted by a stochastic automaton having a fixed initial state which is the same as the only final state. The last mentioned result has been known for pseudo stochastic automata (cf. [10]), but it has been open for stochastic automata.

If $|p(\lambda)| \leq 1$, then (2) holds even for an $(n + 2)$ -state stochastic automaton B (c depends on the sign of $p(\lambda)$). In the general case, the problem concerning the value $n + 2$ remains open. We show that n and $n + 1$ are not possible.

The problem concerning the value $n + 1$ for (1) remains open. The value n is not possible. As a partial result on $n + 1$ we show that

$$p(P) = a b^{l(P)}(p_{(A, \pi_0)}(P) - p_{(A, \rho_0)}(P)) \quad \forall P \in X^+$$

for an $(n + 1)$ -state stochastic automaton A and for some initial distributions π_0, ρ_0 . A corresponding result with $2n + 1$ states is established in [8] for word-functions of pseudo-Markov chains.

We also give a connection between the characteristic polynomials of any stochastic automaton A satisfying (1) (for every word $P \in X^*$) and of any minimal pseudo stochastic automaton generating p .

2. *Preliminaries and the main theorem.* By a *pseudo stochastic automaton* we mean any ordered quintuple $A = (X, S_n, \{A(x) \mid x \in X\}, \pi, f)$ where X is an alphabet, $S_n = \{s_1, \dots, s_n\}$ is the set of states, each $A(x)$ is an $n \times n$ matrix with real entries, π is an $1 \times n$ vector (*initial vector*), and f is an $n \times 1$ vector (*final vector*). In [11] and [12] this system was called a *generalized automaton*.

If π and the matrices $A(x)$ are stochastic, and if f is of the form $(\delta_1, \dots, \delta_n)^t$ where $\delta_i \in \{0, 1\}$, then A is called a *stochastic automaton* (f corresponds to the set of final states). If, in addition, also the column sums of the matrices $A(x)$ are equal to 1, then A is called a *doubly stochastic automaton*. Following Rabin [9], we say that A is an *actual stochastic automaton* if the entries of the matrices $A(x)$ are positive.

For any word $P = x_1 x_2 \dots x_k$ ($x_i \in X, k > 0$), denote $A(P) = A(x_1) A(x_2) \dots A(x_k)$ and define the word-function $p_A: X^* \rightarrow R$ generated by A as follows:

$$\begin{aligned} p_A(\lambda) &= \pi f, \\ p_A(P) &= \pi A(P) f \quad \forall P \in X^+. \end{aligned}$$

We write sometimes $p_{(A, \pi)}$ instead of p_A . By the notation $L(A, \eta)$, where $\eta \in R$, we mean the *pseudo stochastic language*

$$L(A, \eta) = \{P \in X^* \mid p_A(P) > \eta\}.$$

Theorem 1. *Let $A = (X, S_n, \{A(x) \mid x \in X\}, \pi, f)$ be any n -state pseudo stochastic automaton.*

(i) *There exists an $(n + 2)$ -state actual doubly stochastic automaton B with one final state such that*

$$(3) \quad p_A(P) = a b^{l(P)}(p_B(P) - 1/(n + 2)) \quad \forall P \in X^+$$

where a and b are positive constants which can be chosen arbitrarily large.

(ii) *There exist an $(n + 1)$ -state actual stochastic automaton B and two initial distributions π_0, ρ_0 such that, for some positive constant b ,*

$$(4) \quad p_A(P) = a b^{l(P)}(p_{(B, \pi_0)}(P) - p_{(B, \rho_0)}(P)) \quad \forall P \in X^+$$

where $a = 1$ if $|\pi f| \leq 1$, and $a = |\pi f|$ if $|\pi f| > 1$.

(iii) There exists an $(n + 3)$ -state actual doubly stochastic automaton C with a fixed initial state and one final state such that, for some positive constant b ,

$$(5) \quad p_A(P) = b^{l(P)}(p_C(P) - 1/(n + 3)) \quad \forall P \in X^+.$$

Moreover, the initial state is the same as the final state.

(iv) If $|\pi f| \leq 1$, then there exists an $(n + 2)$ -state actual doubly stochastic automaton C such that

$$(6) \quad p_A(P) = b^{l(P)}(p_C(P) + c) \quad \forall P \in X^+$$

where b is a positive constant, $c = -1/(n + 2)$ if $0 \leq \pi f \leq 1$, and $c = -(n + 1)/(n + 2)$ if $-1 \leq \pi f < 0$.

Proof of (i). B will be constructed in the following steps. We omit the trivial cases $\pi = 0$ or $f = 0$ and assume that $\pi \neq 0, f \neq 0$.

Step 1. There exists a nonsingular $n \times n$ matrix D such that $Df = (1, 0, \dots, 0)^t$. Define $A_1 = (X, S_n, \{A_1(x) \mid x \in X\}, \pi_1, f_1)$ where

$$\pi_1 = \pi D^{-1}, \quad f_1 = Df = (1, 0, \dots, 0)^t, \quad A_1(x) = DA(x)D^{-1} \quad \forall x \in X.$$

Then $p_{A_1} = p_A$.

Step 2. (Cf. Lemma 1 in [12].) Consider any matrix $A_1(x)$ in A_1 . There exist a column vector $u(x)$, a row vector $v(x)$ and a real number $a(x)$ such that in the matrix

$$(7) \quad A_2(x) = \begin{bmatrix} 0 & 0 & 0 \\ u(x) & A_1(x) & 0 \\ a(x) & v(x) & 0 \end{bmatrix}$$

all row sums and all column sums equal 0 ($a(x)$ = the sum of the entries of $A_1(x)$). Clearly, for any nonempty word P and for some vectors $u(P), v(P)$, the matrix $A_2(P)$ is obtained from (7) by replacing x by P . Hence, if we define $\pi_2 = (0, \pi_1, 0), f_2 = (0, f_1^t, 0)^t = (0, 1, 0, \dots, 0)$ and $A_2 = (X, S_{n+2}, \{A_2(x) \mid x \in X\}, \pi_2, f_2)$, then $p_{A_2} = p_{A_1} = p_A$.

Step 3. For any real number β , denote by $N(\beta)$ the $(n + 2) \times (n + 2)$ -matrix in which all entries are equal to β . Let α be any number such that the entries of every matrix $A_3(x) = A_2(x) + N(\alpha)$ are positive. Since the row sums and the column sums of each matrix $A_2(x)$ equal 0, it follows that the same holds for any matrix $A_2(P), P \in X^+$. Hence $N(\alpha)A_2(P) = A_2(P)N(\alpha) = 0$ whenever $P \in X^+$. Consequently,

$$(8) \quad A_3(P) = A_2(P) + N((n + 2)^{l(P)} - 1\alpha^{l(P)}) \quad \forall P \in X^+.$$

Let $B(x) = [(n+2)\alpha]^{-1}A_3(x)$ and $\pi_0 = a^{-1}(\pi_2 + (d, \dots, d))$ where a is the sum of the components of $\pi_2 + (d, \dots, d)$ and d is so large that these components are non-negative. Now define $B = (X, S_{n+2}, \{B(x) \mid x \in X\}, \pi_0, f_2)$. B is an actual doubly stochastic automaton, because the row and column sums of every matrix $A_3(x)$ equal $(n+2)\alpha$. Moreover, $f_2 = (0, 1, 0, \dots, 0)^t$, i.e. B has one final state, only. By (8) we now have

$$p_B(P) = a^{-1}[(n+2)\alpha]^{-l(P)}p_{A_3}(P) + 1/(n+2), \quad \forall P \in X^+.$$

Since $p_{A_3} = p_A$, we obtain the formula (3) where $b = (n+2)\alpha$. Since α and d can be chosen arbitrarily large, the same holds for a and b . The proof of (i) is now complete.

Proof of (iv). Assume first that $0 \leq \pi f \leq 1$. Thus $\pi_1 = (\pi f, 1 - \pi f, 0, \dots, 0)$ is a stochastic vector. We show that there is a nonsingular matrix D such that $\pi D^{-1} = \pi_1$ and $Df = (1, 0, \dots, 0)^t = f_1$. Denote the rows of D by R_1, \dots, R_n .

If $\pi f = 1$, we take any linearly independent vectors u_2, \dots, u_n orthogonal to f . Then choose $R_1 = \pi$, $R_2 = u_2, \dots, R_n = u_n$. Since $\pi f \neq 0$, the resulting matrix D is nonsingular. It satisfies the above conditions, because $\pi_1 = (1, 0, \dots, 0)$.

If $\pi f < 1$, choose $R_2 = (1 - \pi f)^{-1}(\pi - \pi f R_1)$ where R_1 is any vector such that $R_1 f = 1$ and $R_2 \neq 0$. Then $R_2 f = 0$ and there exist vectors R_3, \dots, R_n such that R_2, \dots, R_n are linearly independent and orthogonal to f . The corresponding matrix D is nonsingular and $Df = (1, 0, \dots, 0)^t = f_1$, $\pi_1 D = \pi f R_1 + (\pi - \pi f R_1) = \pi$, i.e. $\pi_1 = \pi D^{-1}$.

Now define $A_1 = (X, S_n, \{A_1(x) \mid x \in X\}, \pi_1, f_1)$ where $A_1(x) = DA(x)D^{-1}$ for every $x \in X$. Then $p_{A_1} = p_A$, and we continue by using the constructions of Steps 2 and 3. In Step 3, we can choose $d = 0$ and $a = 1$, because π_2 is a stochastic vector. In this way we have constructed an actual doubly stochastic automaton C satisfying the formula (6).

Note that if $\pi f = 1$, then the initial and final vectors of C are $\pi_0 = \pi_2 = (0, 1, 0, \dots, 0)$, $f_2 = (0, 1, 0, \dots, 0)^t$, i.e. the initial state is the same as the final state.

Finally, consider the case $-1 \leq \pi f < 0$. Instead of A , consider $A' = (X, S_n, \{A(x) \mid x \in X\}, \pi, -f)$ for which $0 < \pi(-f) \leq 1$. By the above proof,

$$p_{A'}(P) = b^{l(P)}(p_C(P) - 1/(n+2)) \quad \forall P \in X^+$$

for some actual doubly stochastic automaton C . But $p_{A'} = -p_A$ and $p_C(P) = 1 - p_{C'}(P)$ where the only difference between C and C' is the final vector. Consequently,

$$p_A(P) = b^{l(P)}(p_{C'}(P) - 1 + 1/(n+2)) \quad \forall P \in X^+$$

which completes the proof of (iv).

Proof of (iii). First replace A by $A' = (X, S_{n+1}, \{A'(x) \mid x \in X\}, \pi', f')$ where $\pi' = (1, 0, \dots, 0)$, $f' = (1, f')^t$ and, for each $x \in X$,

$$A'(x) = \begin{bmatrix} 0 & \pi A(x) \\ 0 & A(x) \end{bmatrix}.$$

Then $\pi'f' = 1$ and $p_{A'}(P) = p_A(P) \quad \forall P \in X^+$. Now continue as in the proof of (iv). Since $\pi'f' = 1$, the note we made in that proof implies that the constructed automaton satisfies the conditions of (iii).

Proof of (ii). By the symmetry of the condition (4), we may assume that $\pi f \geq 0$, because otherwise we can replace A by $\bar{A} = (X, S_n, \{A(x) \mid x \in X\}, \pi, -f)$, i.e. p_A is replaced by $-p_A$.

Let $A' = (X, S_n, \{A(x) \mid x \in X\}, \pi, f')$ where $f' = a^{-1}f$, and $a = \pi f$ if $\pi f > 1$, $a = 1$ if $\pi f \leq 1$. Now $0 \leq \pi f' \leq 1$, and we construct A_1 as in the proof of (iv). Thus $\pi_1 = (\pi f', 1 - \pi f', 0, \dots, 0)$ and $f_1 = (1, 0, \dots, 0)^t$. We continue as follows. Define $A_2 = (X, S_{n+1}, \{A_2(x) \mid x \in X\}, \pi_2, f_2)$ where $\pi_2 = (\pi_1, -1)$, $f_2 = (f_1^t, 0)^t = (1, 0, \dots, 0)^t$ and, for every $x \in X$,

$$A_2(x) = \begin{bmatrix} A_1(x) & u(x) \\ 0 & 0 \end{bmatrix}$$

where $u(x)$ is a column vector such that the row sums of $A_2(x)$ equal 0. Clearly $p_{A_2} = p_{A_1} = p_{A'}$. Now form the matrices $A_3(x) = A_2(x) + N(x)$ as in Step 3. Since the sum of the components of π_2 and the row sums in every matrix $A_2(P)$ ($P \in X^+$) equal 0, it follows that

$$(9) \quad \pi_2 A_3(P) = \pi_2 A_2(P) \quad \forall P \in X^+.$$

Define $\pi_0 = (\pi_1, 0)$ and $\rho_0 = (0, \dots, 0, 1)$. Then $\pi_2 = \pi_0 - \rho_0$. Consequently, by (9),

$$(10) \quad (\pi_0 - \rho_0) A_3(P) f_2 = p_{A'}(P) \quad \forall P \in X^+.$$

Let $B = (X, S_{n+1}, \{B(x) \mid x \in X\}, f_2)$ where $B(x) = [(n+1)\alpha]^{-1} A_3(x)$ for every $x \in X$. B is a stochastic automaton without a fixed initial vector. Since $p_{A_2} = p_{A'} = a^{-1}p_A$, it follows from (10) that (ii) holds.

The proof of Theorem 1 is now complete.

Corollary 1. For any word-function $p: X^* \rightarrow R$ of finite rank n , Theorem 1 holds when p_A and πf are replaced by p and $p(\lambda)$.

Corollary 2. For any pseudo stochastic language L accepted by an n -state pseudo stochastic automaton, there exists an actual $(n+4)$ -state doubly stochastic automaton C such that $L + \{\lambda\} = L(C, 1/(n+4))$. Moreover, C has a fixed initial state which is the same as the only final state. Consequently, any stochastic language L is quasidefinite in the sense of [6], and $L + \{\lambda\}$ is ε -approximable by a finite language if the above automaton C is used.

Proof. If A is an n -state pseudo stochastic automaton, then $L(A, \gamma)$ can be given in the form $L(A', 0)$ where the number of states in A' is $n + 1$. (If A is stochastic, then n is sufficient.) Now the part (iii) of Theorem 1 gives the desired automaton C .

Remark. Paz [7] has shown that there exist a stochastic automaton B and a cut-point γ such that $L(B, \gamma)$ is not ε -approximable by any regular language. Bertoni [1] and Kosaraju [5] proved that the same holds even for context-free languages.

Carlyle and Paz [2] have considered pseudo stochastic automata A having several final vectors f_1, \dots, f_k ($k > 0$). Let $p_A^{(i)}$ be the word-function corresponding to f_i . They showed that

$$p_A^{(i)}(P) = a b^{l(P)}(p_B^{(i)}(P) + c) \quad \forall P \in X^+$$

for a $2(k + 1)n$ -state stochastic automaton B . We show now that the same holds for B with less than $n + k + 2$ states.

Theorem 2. *Let $A = (X, S_n, \{A(x) \mid x \in X\}, \pi, \{f_1, \dots, f_k\})$ be any pseudo stochastic automaton with k final vectors. There exists an $(n + k - r + 2)$ -state doubly stochastic automaton $B = (X, S, \{B(x) \mid x \in X\}, \pi_0, \{g_1, \dots, g_k\})$ such that, for every i ,*

$$p_A^{(i)}(P) = a b^{l(P)}(p_B^{(i)}(P) - 1/(n + k - r + 2)) \quad \forall P \in X^+$$

where $a > 0$, $b > 0$, r is the maximum number of linearly independent vectors in $\{f_1, \dots, f_k\}$, and every g_i is of the form $(0, \dots, 0, 1, 0, \dots, 0)^t$.

Proof. We may assume that f_1, \dots, f_r are linearly independent. Clearly, there is a nonsingular $n \times n$ matrix D such that $Df_i = e_i$ ($i = 1, \dots, r$) where the i th component of e_i is 1 and other components equal 0. Let $\pi_1 = \pi D^{-1}$ and $A_1(x) = DA(x)D^{-1}$. Define $A_2 = (X, S_{n+k-r}, \{A_2(x) \mid x \in X\}, \pi_1, \{g_1, \dots, g_k\})$ where

$$A_2(x) = \begin{bmatrix} A_1(x) & A_1(x)Df_{r+1} & \dots & A_1(x)Df_k \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

and $g_1 = e_1, \dots, g_r = e_r, g_{r+1} = e_{n+1}, \dots, g_k = e_{n+k-r}$. Clearly $p_{A_2}^{(i)}(P) = p_A^{(i)}(P)$ for every $P \in X^+$. Hereafter, proceed as in Steps 2 and 3.

3. Impossibility of the bound $n + 1$ for formula (2). In this section we want to show that there exist n and an n -state pseudo stochastic automaton A such that the formula

$$(11) \quad p_A(P) = b^{l(P)}(p_B(P) + c) \quad \forall P \in X^+$$

is impossible for every $(n + 1)$ -state stochastic automaton even if b is allowed to be negative.

The formula

$$(12) \quad p_A(P) = a b^{l(P)}(p_B(P) + c) \quad \forall P \in X^+,$$

where A has n states, is not generally valid for any n -state stochastic automaton B . This follows, for example, from the well-known result that two-state stochastic automata over one letter accept regular languages, only, but this does not hold for all two-state pseudo stochastic automata and for the cut-point 0. Consequently, $n + 2$ is the only open value for (11), and so is $n + 1$ for (12) (see Theorem 1, (ii), (iv)).

Let us first consider a general two-state stochastic automaton $B = (\{x, y\}, \{s_1, s_2\}, \{B(x), B(y)\}, \pi_0, f_0)$ where $\pi_0 = (p_1, p_2)$, $f_0 = (0, 1)^t$ and

$$(13) \quad B(x) = \begin{bmatrix} 1 - a_1 & a_1 \\ b_1 & 1 - b_1 \end{bmatrix}, \quad B(y) = \begin{bmatrix} 1 - a_2 & a_2 \\ b_2 & 1 - b_2 \end{bmatrix}.$$

Let $\lambda = 1 - (a_1 + b_1)$ and $\mu = 1 - (a_2 + b_2)$. Then the eigenvalues of $B(x)$ and $B(y)$ are 1, λ and 1, μ , respectively. If $a_1 + b_1 > 0$, then

$$B(x^k) = \begin{bmatrix} \beta_1 & \alpha_1 \\ \beta_1 & \alpha_1 \end{bmatrix} + \begin{bmatrix} \alpha_1 - \alpha_1 \\ -\beta_1 & \beta_1 \end{bmatrix} \lambda^k \quad \forall k > 0$$

where $\alpha_1 = a_1/(a_1 + b_1)$ and $\beta_1 = b_1/(a_1 + b_1) = 1 - \alpha_1$. By denoting $\alpha_2 = a_2/(a_2 + b_2)$ and $\beta_2 = b_2/(a_2 + b_2)$ we can write the corresponding formula for $B(y^m)$, $m > 0$, assuming that $a_2 + b_2 > 0$. Consequently, if $a_1 + b_1 > 0$ and $a_2 + b_2 > 0$, we have

$$B(x^k y^m) = \begin{bmatrix} \beta_2 & \alpha_2 \\ \beta_2 & \alpha_2 \end{bmatrix} + \begin{bmatrix} \alpha_2 - \alpha_1 & \alpha_1 - \alpha_2 \\ \alpha_2 - \alpha_1 & \alpha_1 - \alpha_2 \end{bmatrix} \mu^m + \begin{bmatrix} \alpha_1 - \alpha_1 \\ -\beta_1 & \beta_1 \end{bmatrix} \lambda^k \mu^m$$

and

$$B(y^m x^k) = \begin{bmatrix} \beta_1 & \alpha_1 \\ \beta_1 & \alpha_1 \end{bmatrix} + \begin{bmatrix} \alpha_1 - \alpha_2 & \alpha_2 - \alpha_1 \\ \alpha_1 - \alpha_2 & \alpha_2 - \alpha_1 \end{bmatrix} \lambda^k + \begin{bmatrix} \alpha_2 - \alpha_2 \\ -\beta_2 & \beta_2 \end{bmatrix} \lambda^k \mu^m.$$

From the above formulas it follows that

$$(14) \quad p_B(x^k) = \alpha_1 + (\beta_1 - p_1) \lambda^k,$$

$$(15) \quad p_B(y^m) = \alpha_2 + (\beta_2 - p_1) \mu^m,$$

$$(16) \quad p_B(x^k y^m) = \alpha_2 + (\alpha_1 - \alpha_2) \mu^m + (\beta_1 - p_1) \lambda^k \mu^m,$$

$$(17) \quad p_B(y^m x^k) = \alpha_1 + (\alpha_2 - \alpha_1) \lambda^k + (\beta_2 - p_1) \lambda^k \mu^m$$

for any $k > 0$, $m > 0$ assuming that $a_1 + b_1 > 0$ and $a_2 + b_2 > 0$.

Theorem 3. *For an n -state pseudo stochastic automaton A there do not always exist an $(n + 1)$ -state stochastic automaton B and constants b, c such that*

$$p_A(P) = b^{l(P)}(p_B(P) + c) \quad \forall P \in X^+.$$

Proof. In our counterexample, we use the one-state pseudo stochastic automaton $A = (\{x, y\}, \{s_1\}, \{A(x), A(y)\}, \pi, f)$ where $\pi = (1)$, $f = (1)$, $A(x) = (-3)$, and $A(y) = (2)$. Now

$$p_A(P) = (-3)^{l_x(P)} 2^{l_y(P)} \quad \forall P \in \{x, y\}^+$$

where $l_x(P)$ and $l_y(P)$ denote the numbers of the occurrences of x and y in P .

In order to prove the theorem, we have to show that

$$(18) \quad (-3)^{l_x(P)} 2^{l_y(P)} = b^{l(P)}(p_B(P) + c) \quad \forall P \in \{x, y\}^+$$

is impossible for all two-state stochastic automata and for all values of b and c . Therefore, let $B = (\{x, y\}, \{s_1, s_2\}, \{B(x), B(y)\}, \pi_0, f_0)$ be an arbitrary two-state stochastic automaton. We use the notations (13). We may assume that $f_0 = (0, 1)^t$, because $(1, 1)^t$ is not possible. Consequently, formulas (14)–(17) hold if $a_1 + b_1 > 0$ and $a_2 + b_2 > 0$.

It is easy to see that, if $a_1 + b_1 = a_2 + b_2 = 0$, then (18) cannot hold. Hence, we assume that either $a_1 + b_1 > 0$ or $a_2 + b_2 > 0$.

Case 1. Assume that $a_1 + b_1 = 0$ and $a_2 + b_2 > 0$. Now $p_B(P)$ does not depend on the number of the letters x in P . Hence, by (15), the condition (18) gets the form

$$2^m(-3)^k = b^{k+m}(c + \alpha_2 + (\beta_2 - p_1)\mu^m) \quad \text{for all } k > 0, m > 0.$$

Thus

$$(-3/b)^k = (b/2)^m(c + \alpha_2 + (\beta_2 - p_1)\mu^m).$$

Since the right hand side does not depend on k , it follows that $b = -3$. Thus $c + \alpha_2 + (\beta_2 - p_1)\mu^m = (-2/3)^m$ for every $m > 0$. This is possible only if $c + \alpha_2 = 0$, $\mu = -2/3$ and $\beta_2 - p_1 = 1$. But if $\beta_2 - p_1 = 1$, then necessarily $\beta_2 = 1$, which means that $a_2 = 0$. Thus $\mu = 1 - b_2 \geq 0$, which is a contradiction, because $\mu = -2/3$.

Case 2. Assume that $a_1 + b_1 > 0$ and $a_2 + b_2 = 0$. This is analogous to Case 1.

Case 3. Assume that $a_1 + b_1 > 0$ and $a_2 + b_2 > 0$. By taking the difference of both sides of (17) and (16) it follows from (18) that

$$(\alpha_1 - \alpha_2)(1 - \lambda^k)(1 - \mu^m) = 0 \quad \text{for all } k > 0, m > 0.$$

Here $\lambda \neq 1$ and $\mu \neq 1$, because $a_1 + b_1 > 0$ and $a_2 + b_2 > 0$. Consequently, $\alpha_1 - \alpha_2$ must be equal to 0, i.e. $\alpha_1 = \alpha_2$, which implies that $\beta_1 = \beta_2$. The condition (18) implies now that

$$2^m(-3)^k = b^{k+m}(c + \alpha_1 + (\beta_1 - p_1)\lambda^k\mu^m),$$

which implies

$$(2/b)^m(-3/b)^k = c + \alpha_1 + (\beta_1 - p_1)\lambda^k\mu^m.$$

This is possible only if $c + \alpha_1 = 0$, $\beta_1 - p_1 = 1$, $\lambda = -3/b$ and $\mu = 2/b$. Thus $\beta_1 = 1$ and $\lambda\mu < 0$. On the other hand, $\beta_1 = \beta_2$, so that $\beta_1 = \beta_2 = 1$, which implies that $\lambda \geq 0$ and $\mu \geq 0$. This is a contradiction, because $\lambda\mu < 0$. The proof of Theorem 3 is now complete.

4. *Characteristic polynomials.* Let $A = (X, S_m, \{A(x) \mid x \in X\}, \pi, f)$ be any pseudo stochastic automaton. Let V_1 be the *reachable* part of R^m , i.e. V_1 is the linear space spanned by the vectors $\pi A(P)$, $P \in X^*$. Let \sim be an equivalence relation over V_1 defined as follows:

$$u \sim v \quad \text{if and only if} \quad u A(P)f = v A(P)f \quad \forall P \in X^*.$$

Then the number of states in any minimal pseudo stochastic automaton generating p_A is equal to $\dim(V_1/\sim)$ (cf. [13]). Assume now that A is in minimal form. This implies that $V_1 = R^m$ and $\dim(R^m/\sim) = m$. Consequently, it can be verified that every equivalence class in R^m/\sim contains one element only. We shall use this fact in the proof of the following theorem.

Theorem 4. *Let $A = (X, S_m, \{A(x) \mid x \in X\}, \pi, f)$ be a pseudo stochastic automaton in minimal form. Let $B = (X, S_n, \{B(x) \mid x \in X\}, \pi_0, f_0)$ be any stochastic automaton such that*

$$p_A(P) = a b^{l(P)}(p_B(P) + c) \quad \forall P \in X^*$$

for some constants a, b, c . Then, for any $P \in X^+$, the characteristic polynomials $\varphi_{A(P)}$, $\varphi_{B(P)}$ of $A(P)$ and $B(P)$ satisfy the formula

$$\varphi_{B(P)}(z) = \varphi_{A(P)}(b^{l(P)}z)\psi_P(z)$$

where $\psi_P(z)$ is a polynomial.

Proof. Let $C = (X, S_n, \{C(x) \mid x \in X\}, \pi_0, f_1)$ where, for each $x \in X$, $C(x) = bB(x)$, and $f_1 = a(f_0 + (c, \dots, c)^t)$. Then $p_A = p_C$. Let V_1 be the reachable part of R^n in C . Take any basis $\{\pi_0 C(P_1), \dots, \pi_0 C(P_r)\}$ for V_1 , and define a linear mapping $T: V_1 \rightarrow R^m$ as follows:

$$\pi_0 C(P_i)T = \pi A(P_i) \quad (i = 1, \dots, r).$$

(We write the mappings on the right hand side of the vectors.) We show that T is surjective and that the equality $C(x)T = TA(x)$ holds in V_1 for every $x \in X$.

Since $p_A = p_C$, we have for any words P, Q the equations

$$\begin{aligned} \pi A(P)A(Q)f &= \pi_0 C(P)C(Q)f_1 \\ &= \left(\sum_{i=1}^r a_i(P) \pi_0 C(P_i) \right) C(Q)f_1 \\ &= \left(\sum_{i=1}^r a_i(P) \pi A(P_i) \right) A(Q)f \end{aligned}$$

where $\sum a_i(P)\pi_0 C(P_i) = \pi_0 C(P)$. Consequently, the vectors $\pi A(P)$ and $\sum a_i(P)\pi A(P_i)$ are in the same equivalence class of the relation \sim . By the remark we made before Theorem 3, this implies that $\pi A(P) = \sum a_i(P)\pi A(P_i)$. Hence,

$$\pi A(P) = \sum a_i(P)\pi_0 C(P_i) T = \pi_0 C(P) T.$$

Consequently, T is a surjective mapping, because, by the minimality of A , the vectors $\pi A(P)$ span the space R^m . Furthermore,

$$\pi_0 C(P_i) C(x) T = \pi A(P_i) A(x) = \pi_0 C(P_i) T A(x)$$

for every $x \in X$. Thus the equality $C(x) T = T A(x)$ holds in V_1 . This implies that, for any $P \in X^+$, $C(P) T = T A(P)$ in V_1 . From this equality it follows that $\varphi'_{C(P)}(z) = \varphi_{A(P)}(z)\psi'_P(z)$ where $\psi'_P(z)$ is some polynomial and $\varphi'_{C(P)}(z)$ is the characteristic polynomial of $C(P)$ considered as a mapping of V_1 . Since V_1 is a subspace of R^n , it follows that $\varphi'_{C(P)}(z)$ is a factor of $\varphi_{C(P)}(z)$. Therefore, we have $\varphi_{C(P)}(z) = \varphi_{A(P)}(z)\psi''_P(z)$ for some polynomial $\psi''_P(z)$. By the definition of the matrices $C(x)$, this implies

$$\varphi_{A(P)}(b^{l(P)}z)\psi''_P(b^{l(P)}z) = b^{nl(P)}\varphi_{B(P)}(z),$$

which gives the required formula.

Acknowledgement. This work has been done during my visit in the Center for Mathematical System Theory, Department of Mathematics, University of Florida. I want to express my gratitude to Professor R. E. Kalman for offering me the opportunity to work in his Center.

References

- [1] BERTONI, A.: Complexity problems related to the approximation of probabilistic languages and events by deterministic machines. - Résumés des communications, Colloques IRIA, Théorie des automates des langages et de la programmation 3—7 juillet 1972.
- [2] CARLYLE, J. W., and A. PAZ: Realizations by stochastic finite automata. - J. Comput. System Sci. 5, 1971, 26—40.
- [3] CLAUS, V.: Stochastische Automaten. - Teubner Studienskripten, B. G. Teubner, Stuttgart, 1971.
- [4] INAGAKI, Y., and T. FUKUMURA: Some aspects of linear space automata. - Information and Control 20, 1972, 439—479.
- [5] KOSARAJU, S. R.: Probabilistic automata — A problem of Paz. - Information and Control 23, 1973, 97—104.
- [6] PAZ, A.: Some aspects of probabilistic automata. - Information and Control 9, 1966, 26—60.
- [7] — — Fuzzy star functions, probabilistic automata and their approximation by nonprobabilistic automata. - J. Comput. System Sci. 1, 1967, 371—390.
- [8] — — Word-functions of pseudo-Markov chains. - Manuscript 1972.

-
- [9] RABIN, M. O.: Probabilistic automata. - Information and Control 6, 1963, 230—245.
- [10] RICHARD, J.: Représentations matricielles des séries rationnelles en variables non commutatives. - C. R. Acad. Sci. Paris 270, Série A, 1970, 224—227.
- [11] TURAKAINEN, P.: On probabilistic automata and their generalizations. - Ann. Acad. Sci. Fenn. Ser. A I 429, 1968, 1—53.
- [12] —»— Generalized automata and stochastic languages. - Proc. Amer. Math. Soc. 21, 1969, 303—309.
- [13] —»— On the minimization of linear space automata. - Ann. Acad. Sci. Fenn. Ser. A I 506, 1972, 1—15.

University of Oulu
Department of Mathematics
SF-90100 Oulu 10
Finland

Received 5 September 1974