throughout V. By hypothesis, however, $|F|^2=\bar{F}(F^*K)$, and so $F=0$ identically throughout V.

In particular, then, F vanishes identically in a neighborhood of x_0. Since x_0 was arbitrary F vanishes identically and the proof is complete.

YESHIVA UNIVERSITY

DEFINITE AND QUASIDEFINITE SETS OF STOCHASTIC MATRICES[1]

AZARIA PAZ

1. **Introduction.** This note is concerned with asymptotic behaviour of long products of stochastic matrices of a given form. Its objects are:

(a) To prove that a theorem stated (but not proved) by the author in a previously-published paper [2] is equivalent to one proved by Wolfowitz in [1].

(b) To formulate a decision procedure for the above problem, preferable to that given by Wolfowitz in [1].

(c) To solve a related problem.

Familiarity with the above two papers is desirable.

2. **Definitions.** (We adopt here some of the definitions used by Wolfowitz.) A finite square matrix $P=\|p_{ij}\|$ is called stochastic if $p_{ij}\geqq 0$ for all i, j and $\sum_j p_{ij}=1$ for all i.

A stochastic matrix P is called indecomposable and aperiodic (S.I.A.) if

$$Q = \lim_{n\to\infty} P^n$$

exists and all rows of Q are the same. $|P|$ and $\delta(P)$ are defined as

$$|P| = \max_{ij} |p_{ij}|,$$
$$\delta(P) = \max_{j} \max_{i_1 i_2} |p_{i_1 j} - p_{i_2 j}|.$$

With every stochastic matrix P we associate a finite graph having n states (vertices)—n being the order of P—such that transition is

Received by the editors April 26, 1964.

[1] The note is based on part of the author's D.Sc. thesis submitted to the Senate of the Technion, Israel Institute of Technology, December 1963.

possible from state i to state j iff $p_{ij}>0$; in that case we say that state j is a *consequent* of state i.

We also say that two states i_1 and i_2 have a *common consequent* if there is a state j such that $p_{i_1j}>0$ together with $p_{i_2j}>0$.

Scrambling condition (Hajnal). A stochastic matrix is called scrambling if every pair of states in the associated graph has a common consequent.

H_1 *condition*. A stochastic matrix P satisfies the H_1 condition if there is k such that P^k is scrambling.

Let $A_1, \cdots, A_k$ be any square matrices of the same order. A product of t A's (repetitions permitted) is called a *word* of length t (in the A's). If $B=A_i \cdots A_a$ is a word, then the word $A_{a-r+1} \cdots A_a$ is called the r-suffix of the given word B.

Wolfowitz's condition. A finite set of stochastic matrices $A_1, \cdots, A_k$ of the same order satisfies the W-condition if any word in the A's is S.I.A.

H_4 *condition*. A set of matrices as above satisfies the H_4 condition (of order r) if there is r such that any word in the A's of length r or more is a scrambling word.

3. Some preliminary lemmas.

LEMMA 1. *Given a stochastic matrix P; if P satisfies* H_1, *then P is S.I.A.*

This is Corollary 1 in [2].

LEMMA 2. *If $A_1, \cdots, A_k$ is a set of stochastic matrices of the same order satisfying the W-condition, then there is t such that all the words in the A's of length $\geqq t$ are scrambling.*

This is Lemma 4 in [1].

LEMMA 3. *The* H_4 *condition is equivalent to the W-condition.*

PROOF. Assume that the W-condition is satisfied. Then there is t such that all words in the A's of length $\geqq t$ are scrambling by Lemma 2; hence the H_4-condition is satisfied. Assume now that the H_4-condition is satisfied and let B be any word in the A's; then B^k is scrambling for some k, this implying that B satisfies H_1. By Lemma 1, B is S.I.A. and the W-condition is satisfied.

4. Definite sets of stochastic matrices. Consider a finite set of stochastic matrices of the same order $N=\langle A_1, \cdots, A_k\rangle$. It may be the case that there exists an integer ν such that any word B (in the A's) of length $n\geqq\nu$ satisfies:

$$\delta(B) = 0$$

i.e., all rows of B are the same (this was noted by Hajnal in [3]). In that case we say that N is a definite set of matrices of order ν.

It can be shown that if N is a definite set, then the transition probabilities relative to any word B in the A's of length $\geqq \nu$ depend only on the ν-suffix of B.

Systems having similar properties are encountered in the theory of finite automata (see Perles, Rabin, Shamir [4]).

The following theorem provides a decision procedure for finding out whether a given set of matrices is definite.

THEOREM 1. *Let N be a finite set of stochastic matrices of the same order n. If N is a definite set, then N is definite of order $n-1$.*

PROOF. The above theorem was proved in [4] for the case where all matrices in N are degenerate stochastic matrices (i.e. matrices with exactly one unity in every row). The above proof may be used for the present theorem, with the following modifications: A *constant matrix* is a stochastic matrix with equal rows. (This generalizes the definition in [4] but does not interfere with the proof.) Also, the matrices in the set H (defined in [4]) are stochastic matrices (instead of degenerate stochastic matrices) and the linear space V (defined there) is taken over the real numbers (instead of the rational numbers).

EXAMPLE. Consider the following set of two 3×3 matrices:

$$A_1 = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/8 & 1/2 & 3/8 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/6 & 1/2 & 1/3 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}.$$

Straightforward computation shows that

$$A_1\cdot A_1 = \begin{bmatrix} 5/24 & 1/2 & 7/24 \\ 5/24 & 1/2 & 7/24 \\ 5/24 & 1/2 & 7/24 \end{bmatrix}, \qquad A_2\cdot A_2 = \begin{bmatrix} 3/16 & 1/2 & 5/16 \\ 3/16 & 1/2 & 5/16 \\ 3/16 & 1/2 & 5/16 \end{bmatrix}$$

and $A_2\cdot A_1 = A_1\cdot A_1$; $A_1\cdot A_2 = A_2\cdot A_2$.

This set is therefore 2-definite.

5. Quasidefinite sets of stochastic matrices. Let $N=\langle A_1, \cdots, A_k\rangle$ be a finite set of stochastic matrices of the same order. It may be the case that for any ϵ there exists an integer $\nu(\epsilon)$ such that any word B (in the A's) of length $n \geqq \nu(\epsilon)$ satisfies:

$$\delta(B) < \epsilon$$

i.e. all rows of B are approximately the same. In this case we shall say that N is a quasidefinite set.

It can be shown that for a quasidefinite set N the transition probabilities relative to any word B are approximated by those relative to its suffices—the longer the suffix, the closer the approximation.

It follows that quasidefinite systems are a natural generalization of definite systems.

The following theorem was proved in Wolfowitz's paper [1].

THEOREM 2. *A finite set of stochastic matrices of the same order is quasidefinite iff it satisfies the W-condition.*

The following theorem was stated (but not proved) in the author's paper [2]:

THEOREM 3. *A finite set of stochastic matrices of the same order is quasidefinite iff it satisfies the* H_4 *condition.*

Theorem 3 can be proved independently of Theorem 2, but by Lemma 3 the two theorems are equivalent and the proof is omitted.

6. Decision procedures. A decision procedure was given by Wolfowitz in [1] for ascertaining whether a given set of stochastic matrices is quasidefinite. The procedure is based on the following two assertions (proved in the above paper).

First, it is proved that if a set of matrices $A_1, \cdots, A_k$ of the same order is such that all words in the A's of length smaller than a specified integer t are S.I.A., then all words in the A's are S.I.A. (i.e. the set satisfies the W-condition).

Secondly, it is shown that any word having a scrambling matrix as a factor is S.I.A.

Using these two facts, one has to check all words in the A's of length $\leqq t$, discarding all those with a scrambling matrix as a factor. If, and only if, all these words are S.I.A., then the set is quasidefinite.

It can now be shown that the number t (definite in [1]) in the above decision procedure exceeds 2^{n^2-n}.

We give here a better decision procedure, based on the H_4 condition and on the following

THEOREM 4. *If a finite set N of stochastic matrices of order n satisfies the* H_4 *condition, then it satisfies this condition of order* v *for some integer* v *with:*

$$v \leqq \frac{1}{2}\left(3^n - 2^{n+1} + 1\right).$$

PROOF. Suppose that there is a word B (in the A's) of length exceeding the bound of the theorem and *not* scrambling. Then, there are two states i_1 and i_2 which do not have a common consequent by B. Set

$$B = B_1 \cdots B_\mu$$

(the B's being matrices taken from the given set), and consider the following sequence of unordered pairs of sets of states:

$$(\overset{1}{\alpha}_0, \overset{2}{\alpha}_0),\ (\overset{1}{\alpha}_1, \overset{2}{\alpha}_1),\ (\overset{1}{\alpha}_2, \overset{2}{\alpha}_2),\ \cdots,\ (\overset{1}{\alpha}_\mu, \overset{2}{\alpha}_\mu)$$

where $\overset{1}{\alpha}_0 = \langle i_1 \rangle$, $\overset{2}{\alpha}_0 = \langle i_2 \rangle$ and $\overset{1}{\alpha}_{i+1}$, $\overset{2}{\alpha}_{i+1}$ are the consequents of the states in $\overset{1}{\alpha}_i$, $\overset{2}{\alpha}_i$ respectively by matrix B_{i+1}.

By the definition of the α's and of B, we have that all α's are nonvoid sets and every pair of α's is a disjoint pair of sets. It can easily be shown by a combinatorial argument that the number of different unordered pairs of sets which are subsets of the set of all states and have the above properties is $\frac{1}{2}(3^n - 2^{n+1} + 1)$ (n being the number of states). This implies that in the above sequence there are two equal pairs, say

$$(\overset{1}{\alpha}_j, \overset{2}{\alpha}_j) = (\overset{1}{\alpha}_k, \overset{2}{\alpha}_k), \qquad j < k \leqq n.$$

It follows that any word of the form:

$$B_1 \cdots B_{j-1}(B_j \cdots B_k)^r\ B_{k+1} \cdots B_n, \qquad r = 1, 2 \cdots,$$

is not scrambling and the H_4 condition is not satisfied. This proves our theorem.

On the basis of this theorem a decision procedure can be formulated for ascertaining whether a given set of matrices satisfies the H_4 condition.

One has to check all words in the A's of length smaller than the bound given in the theorem above (i.e., $\frac{1}{2}(3^n - 2^{n+1} + 1)$). As in Wolfowitz's procedure, one may discard words having scrambling matrix as a factor, as the H_4 condition is directly based on the scrambling property of the matrices and, as is easily shown, any matrix with a scrambling matrix as a factor is itself scrambling. (This follows directly from our definition.)

REMARK. Wolfowitz's decision procedure is in fact an improvement on a procedure described by Thomasian in [5]. By Thomasian's procedure one must check *all* words of length $\leqq 2^{n^2}$. From a practical point of view, our procedure, although preferable to Wolfowitz's (e.g. for $n = 6$ our bound is 301, while the number t as defined by

Wolfowitz exceeds a billion), is difficult and for large n even impracticable; still, our bound is best, as can be shown by the following example.

7. **Example.** Fix n, let K be a set of n states, and let the following sequence be any enumeration of all *different unordered* pairs, of *nonvoid disjoint* sets of states from K:

$$(1) \qquad (\overset{1}{\alpha}_0, \overset{2}{\alpha}_0), \ (\overset{1}{\alpha}_1, \overset{2}{\alpha}_1), \cdots, (\overset{1}{\alpha}_k, \overset{2}{\alpha}_k),$$

such that the number of states in any set of the form $\alpha_i = \alpha_i^1 \cup \alpha_i^2$ is not smaller than that in the set α_{i-1} for $i = 1, 2, \cdots, k$. As stated before $k+1 = \frac{1}{2}(3^n - 2^{n+1} + 1)$.

If ϕ is a set of states and A a stochastic matrix, Be denote by $A(\phi)$ the set of states which are consequents of those in ϕ by A.

Let $A_1, \cdots, A_k$ be a set of stochastic matrices of order n, such that:

$$(2) \qquad A_i(\phi) = \begin{cases} K & \text{if } \phi \cap [K - \alpha_{i-1}] \neq \varnothing, \\ \overset{1}{\alpha}_i & \text{if } \phi \subseteq \overset{1}{\alpha}_{i-1}, \\ \overset{2}{\alpha}_i & \text{if } \phi \subseteq \overset{2}{\alpha}_{i-1}, \\ \alpha_i & \text{otherwise.} \end{cases}$$

Note that the number of states in $A_i(\phi)$ can be smaller than that in ϕ for the second or third case in (2) only. This follows from the definition of sequence (1), and we shall refer to this property as the conditional monotone property.

Note also that if (2) is satisfied for one element sets, it is satisfied for any sets.

We claim that any set of matrices defined as above is quasi-definite, but there is a word in the A's of length k which is not scrambling.

PROOF. The proof of the second assertion is immediate for the word:

$$A_1 \cdot A_2 \cdots A_k$$

is not scrambling by the definition of sequence (1) and by (2).

Assume now that the following word in the A's is not scrambling:

$$B = B_1 \cdot B_2 \cdots B_t, \qquad t > k,$$

i.e. there are two states i_1 and i_2 not having a common consequent by B. Set:

$$\langle i_1 \rangle = \overset{1}{\beta}_0; \quad \langle i_2 \rangle = \overset{2}{\beta}_0; \quad \overset{1}{\beta}_i = B_i(\overset{1}{\beta}_{i-1}) \quad \text{and} \quad \overset{2}{\beta}_i = B_i(\overset{2}{\beta}_{i-1})$$

and consider the following sequence:

$$(\overset{1}{\beta}_0, \overset{2}{\beta}_0), \cdots, (\overset{1}{\beta}_t, \overset{2}{\beta}_t). \tag{3}$$

This is a sequence of unordered pairs of nonvoid disjoint (by assumption) sets of states, and as $t \geqq k+1$, the sequence contains at least two equal pairs, say:

$$(\overset{1}{\beta}_p, \overset{2}{\beta}_p) = (\overset{1}{\beta}_q, \overset{2}{\beta}_q), \qquad p < q. \tag{4}$$

Consider the following subsequence of (3):

$$(\overset{1}{\beta}_p, \overset{2}{\beta}_p), \cdots, (\overset{1}{\beta}_{r-1}, \overset{2}{\beta}_{r-1}), (\overset{1}{\beta}_r, \overset{2}{\beta}_r), \cdots, (\overset{1}{\beta}_q, \overset{2}{\beta}_q). \tag{5}$$

As before, we shall denote by β_i the set $\beta_i = \overset{1}{\beta}_i \cup \overset{2}{\beta}_i$.

The matrix B_r transforms the set $(\overset{1}{\beta}_{r-1}, \overset{2}{\beta}_{r-1})$ into the set $(\overset{1}{\beta}_r, \overset{2}{\beta}_r)$, but B_r is one of the A's, say $B_r = A_h$. The following cases must be considered:

(a) $\quad \overset{1}{\beta}_{r-1} \cap (K - \alpha_{h-1}) \neq \emptyset, \quad$ or $\quad \overset{2}{\beta}_{r-1} \cap (K - \alpha_{h-1}) \neq \emptyset.$

This is impossible, for this would imply that $\overset{1}{\beta}_r \cap \overset{2}{\beta}_r \neq \emptyset$ by (2), contrary to the assumption that these sets are disjoint.

(b) $\quad \beta_{r-1} \subset \overset{1}{\alpha}_{h-1} \quad$ or $\quad \beta_{r-1} \subset \overset{2}{\alpha}_{h-1},$

which is also impossible, as in this case we get that

$$\overset{1}{\beta}_r = B_r(\overset{1}{\beta}_{r-1}) = \overset{1}{\alpha}_h (\text{or } \overset{2}{\alpha}_h) = B_r(\overset{2}{\beta}_{r-1}) = \overset{2}{\beta}_r,$$

contrary by (2) to our assumption that $\overset{1}{\beta}_r \cap \overset{2}{\beta}_r = \emptyset$.

(c) $\quad \overset{1}{\beta}_{r-1} \cap \overset{1}{\alpha}_{h-1} \neq \emptyset,\quad$ together with $\quad \overset{1}{\beta}_{r-1} \cap \overset{2}{\alpha}_{h-1} \neq \emptyset$, or
$\quad \overset{2}{\beta}_{r-1} \cap \overset{1}{\alpha}_{h-1} \neq \emptyset,\quad$ together with $\quad \overset{2}{\beta}_{r-1} \cap \overset{2}{\alpha}_{h-1} \neq \emptyset$

which is also impossible, as in this case we get by (2) that $\overset{1}{\beta}_r \cap \overset{2}{\beta}_r \neq \emptyset$.

(d) $\quad \overset{1}{\beta}_{r-1} \subseteq \overset{1}{\alpha}_{h-1},\quad$ together with $\quad \overset{2}{\beta}_{r-1} \subseteq \overset{2}{\alpha}_{h-1}$, or
$\quad \overset{1}{\beta}_{r-1} \subseteq \overset{2}{\alpha}_{h-1},\quad$ together with $\quad \overset{2}{\beta}_{r-1} \subseteq \overset{1}{\alpha}_{h-1},$

and the inclusion is proper in at least one part of the conditions, which is also impossible, since by the conditional monotone property and by the impossibility of case (b) (applying the same argument to all pairs in sequence (5)), we get that the number of states in β_q is larger than that in β_p, contrary to (4).

$$\text{(e)}\qquad \begin{aligned} &\overset{1}{\beta}_{r-1} = \overset{1}{\alpha}_{h-1}, \quad \text{together with} \quad \overset{2}{\beta}_{r-1} = \overset{2}{\alpha}_{h-1}, \text{ or} \\ &\overset{1}{\beta}_{r-1} = \overset{2}{\alpha}_{h-1}, \quad \text{together with} \quad \overset{2}{\beta}_{r-1} = \overset{1}{\alpha}_{h-1}. \end{aligned}$$

In this case we get that sequence (5) is a middle part of sequence (1), which is impossible as all the sets in (1) are different, contrary to (4).

All possible cases are covered by (a)–(e), and the proof is complete.

Acknowledgments. The author is indebted to Prof. H. Hanani and Dr. M. Yoeli of the Technion-Israel Institute of Technology for valuable advice and aid; to Professor M. O. Rabin of the Hebrew University for encouragement; and to Mr. E. Goldberg for editorial assistance.

BIBLIOGRAPHY

1. J. Wolfowitz, *Products of indecomposable, aperiodic, stochastic matrices*, Proc. Amer. Math. Soc. **14** (1963), 733–737.

2. A. Paz, *Graph-theoretic and algebraic characterizations of some Markow processes*, Israel J. Math. **1** (1963), 169–180.

3. J. Hajnal, *Weak ergodicity in nonhomogeneous Markow chains*, Proc. Cambridge Philos. Soc. **54** (1958), 233–246.

4. M. Perles, M. O. Rabin and E. Shamir, *The theory of definite automata*, IEEE Trans. Electronic Computers, **EC-12** (1963), 233–243.
Electronic Computers, **EC-12** (1963), 233–243.
5. A. J. Thomasian, *A finite criterion for indecomposable channels*, Ann. Math. Statist. **34** (1963), 337–338.

TECHNION-ISRAEL INSTITUTE OF TECHNOLOGY, HAIFA, ISRAEL