PRODUCTS OF RANDOM MATRICES

By H. Furstenberg and H. Kesten[1]

Princeton University

1. Introduction. Let $X^1, X^2, X^3, \cdots$ form a stationary stochastic process with values in the set of $k \times k$ matrices. The asymptotic behavior of the product ${}^nT^1 = X^nX^{n-1} \cdots X^1$ has been studied by Bellman [1] in certain cases. In particular, Bellman showed that, if the X^n are independent and have strictly positive entries, then, under certain conditions,

$$\lim_{n\to\infty} n^{-1}E\{\log({}^nY^1)_{i,j}\}$$

exists, where the subscripts ij refer to the ijth entry of the matrix. In addition, Bellman conjectured that

$$n^{-1/2}(\log\,({}^nY^1)_{i,j} - E\{\log\,({}^nY^1)_{i,j}\})$$

is asymptotically normally distributed. These assertions are motivated by considering the case where the matrices in the range of the X^n commute so that they may be simultaneously diagonalised.

We shall arrive at Bellman's result by considering first the behavior of the norms $\| {}^nY^1 \|$. We find that, under very general conditions, the limit of $n^{-1} \log \| {}^nY^1 \|$ exists almost everywhere, as well as $\lim_{n\to\infty} n^{-1}E\{\log \| {}^nY^1 \|\}$. Under certain positivity assumptions on the entries of the possible matrices, the asymptotic behavior of the $({}^nY^1)_{i,j}$ is deducible from that of $\| {}^nY^1 \|$, and this will enable us to strengthen Bellman's result to an almost everywhere statement. The conjecture regarding normality will be proven in certain cases and we shall give examples to show that the possibilities of further extension are limited.

2. Asymptotic behavior of the norm. For a $k \times k$ matrix A with real or complex entries we define the norm of A by $\| A \| = \max_i \sum_j | A_{i,j} |$. If B is another $k \times k$ matrix we have

$$\| AB \| \leqq \| A \| \, \| B \|. \tag{2.1}$$

This simple fact gives us:

Theorem 1: *If $X^1, X^2, X^3, \cdots$ form a stationary stochastic process with values in the set of $k \times k$ matrices, then*

$$\lim_{n\to\infty} n^{-1}E\{\log \| X^nX^{n-1} \cdots X^1 \|\} = E \tag{2.2}$$

Received October 12, 1959.

[1] Work supported by the U. S. Army Ordnance, Office of Ordnance Research, Contract DA 36-034-ORD-2001.

exists (E is not necessarily finite). If, in addition, the X-process is metrically transitive and $E\{\log^+ \| X^1 \|\} < \infty$,[2] *then*

$$\limsup_{n\to\infty} n^{-1} \log \| X^n X^{n-1} \cdots X^1 \| \leqq E \tag{2.3}$$

with probability 1.

PROOF: Set

$$ {}^nY^m = X^n X^{n-1} \cdots X^m. \tag{2.4}$$

Then, by (2.1),

$$\log \| X^{n+m} \cdots X^1 \| = \log \| {}^{n+m}Y^1 \| \leqq \log \| {}^{n+m}Y^{n+1} \| + \log \| {}^nY^1 \|,$$

and, since the process is stationary, $E\{\log \| {}^{n+m}Y^1 \|\} \leqq E\{\log \| {}^mY^1 \|\} + E\{\log \| {}^nY^1 \|\}$. This, however, is known ([4], p. 98) to imply (2.2).

Under the assumption that $E\{\log^+ \| X^1 \|\} < \infty$, it follows that

$$\sum_{n=1}^{\infty} P\{\log^+ \| X^n \| \geqq \epsilon n\} < \infty, \qquad \text{for all } \epsilon > 0$$

and therefore $\lim_{n\to\infty} n^{-1} \log^+ \| X^n \| = 0$, and so $\limsup_{n\to\infty} n^{-1} \log \| X^n \| \leqq 0$, with probability one. Hence in order to prove (2.3) it suffices to show that, for each $\epsilon > 0$,

$$\limsup_{n\to\infty} n^{-1}N^{-1} \log \| {}^{nN}Y^1 \| \leqq E + \epsilon \tag{2.5}$$

for some N.

By (2.2), given any $\epsilon > 0$, we can find an N such that $N^{-1}E\{\log \| {}^NY^1 \|\} \leqq E + \epsilon$ (if E is finite; $E = \infty$ is excluded by hypothesis, and if $E = -\infty$ only minor modifications need be made). If the process is metrically transitive, then, by the strong law of large numbers,

$$\lim_{r\to\infty} r^{-1}N^{-1} \sum_{k=0}^{r-1} \log \| {}^{k+N}Y^{k+1} \| \leqq E + \epsilon. \tag{2.6}$$

If r has the form nN, then, writing the sum in (2.6) as

$$\sum_{s=1}^{N} \sum_{t=0}^{n-1} \log \| {}^{(t+1)N+s-1}Y^{tN+s} \|$$

and using (2.1), we deduce

$$\limsup_{n\to\infty} r^{-1}N^{-1}(\log \| {}^{nN}Y^1 \| + \log \| {}^{nN+1}Y^2 \| + \cdots + \log \| {}^{(n+1)N-1}Y^N \|) \leqq E + \epsilon. \tag{2.7}$$

Since $\| {}^{nN+j}Y^{j+1} \| \leqq \| {}^{nN+j}Y^{nN+1} \| \, \| {}^{nN}Y^N \| \, \| {}^{N-1}Y^{j+1} \|$ and

$$\limsup_{n\to\infty} n^{-1} \log \| X^n \| \leqq 0,$$

[2] $\log^+ t = \max(\log t, 0)$.

we find that we may replace each $\| {}^{nN+j}Y^{j+1} \|$ in (2.7) by $\| {}^{nN}Y^{N} \|$. This gives

$$\limsup_{n\to\infty} (nN)^{-1} \log \| {}^{nN}Y^{N} \| \leqq E + \epsilon$$

from which (2.5) follows. This proves the theorem.

We would like to strengthen (2.3) to read

$$\lim_{n\to\infty} n^{-1} \log \| {}^{n}Y^{1} \| = E. \tag{2.8}$$

That this may be done will be seen in the next theorem. For the proof we shall require an auxiliary process which we proceed to define.[3]

Let Ω be the set of all sequences $\{(x_1, z_1), (x_2, z_2), \cdots\}$ where the x_n are matrices in the range of the X^n and the z_n are matrices satisfying

$$\| X_{n+1}Z_n \| Z_{n+1} = X_{n+1}Z_n, \quad \| Z_n \| = 0 \text{ or } 1. \tag{2.9}$$

The variables X^n, Z^n are now defined on Ω as the coordinate functions. The subset of Ω on which $z_1 = x_1/\| x_1 \|$ (with $z_1 = 0$ if $x_1 = 0$ and $z_{nn} = 0$ if $x_{n+1}z_n = 0$) may be taken as the sample space for the X process, since on this subset the Z^n are functions of the X^n. Consequently we may define a measure μ_1 on Ω by carrying over to this subset the given probability measure of the X process. Let T be the shift operator on $\Omega{:}T\{(x_n, z_n)\} = \{(x_{n+1}, z_{n+1})\}$. Note that since z_2 need not equal $x_2/\| x_2 \|$ when $z_1 = x_1/\| x_1 \|$, the subset considered above will generally not be invariant under T. Hence the measure μ_1 will generally not be an invariant measure. Now define the measures μ_k on Ω by

$$\mu_k(\Omega') = \mu_1(T^{-k+1}\Omega')$$

for $\Omega' \subset \Omega$. We then have

LEMMA 1: *Let $\nu_n = n^{-1} \sum_{k=1}^{n} \mu_k$. There exists a subsequence ν_{n_i} converging weakly to a probability measure μ on Ω in the sense that the finite dimensional joint distribution functions of the variables X^n, Z^n with respect to the ν_{n_i} converge to the corresponding distribution functions of the X^n, Z^n with respect to μ at each continuity point of the latter. The measure μ is stationary, i.e. invariant under T, and on subsets of Ω defined by the X^n alone, μ agrees with the given probability measure of the X process.*

PROOF: Any sequence of probability distributions on a finite dimensional Euclidean space has a convergent subsequence; however, the limiting distribution may not be a probability distribution, i.e. it may not assign probability 1 to the whole space. For the first part of the lemma it will suffice to show that a convergent subsequence of the distributions in question must converge to a probability distribution. Once this is shown the required subsequence n_i is obtained by a diagonal procedure. Since the X process is stationary we observe that μ_n all agree with the original X process measure on those subsets of Ω de-

[3] A similar idea occurs in a technical report entitled "Electron Levels in a One Dimensional Random Lattice," by H. L. Frisch and S. P. Lloyd of the Bell Telephone Laboratories.

fined by the X^n. Hence the same is true of the ν_n and of any limiting measure of the ν_n. Now, by definition, $\| Z^n \| = 0$ or 1 on the support of any of the μ_n, and so the range of $(Z^1, Z, \cdots, Z^m)$ is always compact. Let now $\{k_i\}$ be any sequence of integers such that the joint distribution functions of $Z_1, \cdots, Z_m$ and $X_1, \cdots, X_m$ with respect to ν_{k_i} converge. The convergent sequence of probability distributions is to be taken on the product of a compact space (the range of $(Z^1, Z^2, \cdots, Z^m)$) with a locally compact space (the range of $(X^1, \cdots, X^m)$) and they agree on sets defined by subsets of the locally compact space (i.e. sets defined in terms of $X^1, \cdots, X^m$). Moreover, by (2.9), the space Ω is a closed subset of the product space. Under these circumstances it is easily seen that the limiting distribution is a proper probability distribution. Thus the limiting measure μ exists; that it has the desired properties follows readily.

We may now prove (2.8).

THEOREM 2. *If $X^1, X^2, X^3, \cdots$ form a metrically transitive stationary stochastic process and*

$$E(\log^+ \| X^1 \|) < \infty,$$

then

$$\lim_{n\to\infty} n^{-1} \log \| {}^nY^1 \| = E$$

with probability 1, *where E is defined in Theorem* 1.

PROOF: Since $\log \| {}^nY^1 \| \leqq \sum_{k=1}^n \log \| X^k \|$, $E < \infty$. If $E = -\infty$, then the theorem is a consequence of (2.3). Hence we may assume that E is finite, and so

$$\inf_n n^{-1}E\{\log \| {}^nY^1 \|\} > -\infty. \tag{2.10}$$

This implies that $P\{\| {}^nY^1 \| = 0\} = 0$ and in particular $P\{\| X^1 \| = 0\} = 0$.

Consider now the space Ω introduced before and the probability measures ν_{n_i} converging to μ on Ω. The main step will be to prove the relations

$$E = \liminf_{i\to\infty} \int_\Omega \log \| X^2Z^1 \| \, d\nu_{n_i} \leqq \int_\Omega \log \| X^2Z^1 \| \, d\mu < \infty . \tag{2.11}$$

By definition of the ν_n, we have

$$\int_\Omega \log \| X^2Z^1 \| \, d\nu_{n_i} = n_i^{-1} \sum_{k=1}^{n_i} \int_\Omega \log \| X^{k+1}Z^k \| \, d\mu_1 .$$

On the support of μ_1 we have $Z^1 = X^1/\| X^1 \|$, and by induction, using (2.9) and the fact that $\| {}^nY^1 \|$ vanishes with probability 0, we find that $Z^n = {}^nY^1/\| {}^nY^1 \|$. Hence on the support of μ_1, $\| X^{k+1}Z^k \| = \| {}^{k+1}Y^1 \|/\| {}^kY^1 \|$ so that

$$\int_\Omega \log \| X^2Z^1 \| \, d\nu_{n_i} = n_i^{-1} \int_\Omega (\log \|{}^{n_i+1}Y^1 \| - \log \| X^1 \|) \, d\mu_1$$

$$= n_i^{-1} E\{\log \|{}^{n_i+1}Y^1\|\} - n_i^{-1} E\{\log \| X^1 \|\}.$$

The first equality of (2.11) now follows from the definition of E and from the fact that $E\{\log \| X^1 \|\}$ is finite by (2.10) and the hypothesis of the theorem.

The first inequality of (2.11) is derived by showing that $\log^+ \| X^2Z^1 \|$ is uniformly integrable with respect to the ν_{n_i}. In fact, since $\log \| X^2Z^1 \| \leqq \log \| X^2 \| + \log \| Z^1 \| = \log \| X^2 \|$ when $\| Z^1 \| = 1$, we have for $A \geqq 0$,

$$\begin{aligned}\int_{\log\|X^2Z^1\|\geqq A} \log \| X^2Z^1 \| \, d\nu_{n_i} &= n_i^{-1} \sum_{k=1}^{n_i} \int_{\log\|X^2Z^1\|\geqq A} \log \| X^2Z^1 \| \, d\mu_k \\ &\leqq n_i^{-1} \sum_{k=1}^{n_i} \int_{\log\|X^2Z^1\|\geqq A} \log \| X^2 \| \, d\mu_k \qquad (2.12)\\ &\leqq n_i^{-1} \sum_{k=1}^{n_i} \int_{\log\|X^2\|\geqq A} \log \| X^2 \| \, d\mu_k = \int_{\log\|X^1\|\geqq A} \log \| X^1 \| \, d\mu_1 .\end{aligned}$$

By hypothesis, $\int \log^+ \| X^1 \| \, d\mu_1 < \infty$, so that the last integral in (2.12) tends to zero as $A \to \infty$. But,

$$\int \log \| X^2Z^1 \| \, d\mu = \int_{\log\|X^2Z^1\|\geqq 0} \log \| X^2Z^1 \| \, d\mu + \int_{\log\|X^2Z^1\|\leqq 0} \log \| X^2Z^1 \| \, d\mu,$$

and similarly when μ is replaced by ν_{n_i}. By the uniform integrability of $\log^+ \| X^2Z^1 \|$,

$$\lim_{i\to\infty} \int_{\log\|X^2Z^1\|\geqq 0} \log \| X^2Z^1 \| \, d\nu_{n_i} = \int_{\log\|X^2Z^1\|\geqq 0} \log \| X^2Z^1 \| \, d\mu.$$

Since the following integrals have an integrand bounded from above and $\nu_{n_i} \to \mu$ one obtains

$$\liminf_{i\to\infty} \int_{\log\|X^2Z^1\|\leqq 0} \log \|X^2Z^1\| \, d\nu_{n_i} \leqq \int_{\log\|X^2Z^1\|\leqq 0} \log \|X^2Z^1\| \, d\mu.$$

This proves the first inequality of (2.11). The remaining inequality is a consequence of $E \log^+ \| X^1 \| < \infty$.

Our theorem now follows easily. For, since μ is a stationary measure on Ω and $\log \| X^2Z^1 \|$ is integrable with respect to this measure, we may apply the ergodic theorem ([3] p. 465) to find that

$$\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \log \| X^{k+1}Z^k \| = \psi$$

exists almost everywhere and $\int_\Omega \psi \, d\mu = \int_\Omega \log \| X^2Z^1 \| \, d\mu \geqq E$. Now by (2.11) and the assumption that $E > -\infty$, $\| {}^2Y^2Z^1 \| = \| X^2Z^1 \| \neq 0$ almost everywhere on Ω. Then $Z^2 = {}^2Y^2Z^1/\| {}^2Y^2Z^1 \|$ a.e. and $X^3Z^2 = {}^3Y^2Z^1/\| {}^2Y^2Z^1 \|$. Since, by (2.11), $\| X^3Z^2 \| \neq 0$ a.e., it follows that ${}^3Y^2Z^1 \neq 0$ a.e. and so $Z^3 = {}^3Y^2Z^1/\| {}^3Y^2Z^1 \|$ a.e. and $X^4Z^3 = {}^4Y^2Z^1/\| {}^3Y^2Z^1 \|$. Continuing in this manner we find for all k, $\| {}^kY^2Z^1 \| \neq 0$ a.e. and $X^{k+1}Z^k = {}^{k+1}Y^2Z^1/\| {}^kY^2Z^1 \|$ a.e. As a result, $\| X^{k+1}Z^k \| = \| {}^{k+1}Y^2Z^1 \| / \| {}^kY^2Z^1 \|$, and we have

$$\lim_{n\to\infty} n^{-1} \log \| {}^{n+1}Y^2Z^1 \| = \psi \text{ a.e.}$$

Hence

$$\liminf_{n\to\infty} n^{-1} \log \| {}^{n+1}Y^2 \| \geqq \psi \text{ a.e.} \tag{2.13}$$

with respect to μ (and hence with respect to μ_1). However, by Theorem 1, the left side of (2.13) is a.e. $\leqq E$, and since the integral of the right hand side of (2.13) is greater than or equal to E we find that the left hand side is equal to E a.e. This completes the proof.

3. Asymptotic behavior of the entries. In this section we always make the following assumption: A I: The possible matrix values M for X^1 all satisfy

$$M_{i,j} > 0, \tag{3.1}$$

and

$$1 \leqq (\max_{i,j} M_{i,j})/(\min_{i,j} M_{i,j}) \leqq C < \infty. \tag{3.2}$$

LEMMA 2. *If* A I *is satisfied, then*

$$({}^{n+m}Y^m)_{i_1,j_1} > 0, \tag{3.3}$$

and

$$C^{-2} \leqq (({}^{n+m}Y^m)_{i_1,j_1})/(({}^{n+m}Y^m)_{i_2,j_2}) \leqq C^2. \tag{3.4}$$

PROOF. (3.3) is obvious and so is (3.4) if $n = 0$. If $n \geqq 2$, we have

$$\frac{({}^{n+m}Y^m)_{i_1,j_1}}{({}^{n+m}Y^m)_{i_2,j_2}} = \frac{\sum_{r,s} (X^{n+m})_{i_1,r}({}^{n+m-1}Y^{m+1})_{r,s}(X^m)_{s,j_1}}{\sum_{r,s} (X^{n+m})_{i_2,r}({}^{n+m-1}Y^{m+1})_{r,s}(X^m)_{s,j_2}},$$

and this is $\leqq C^2$ by (3.2). A similar estimate holds for $n = 1$. This completes the proof of the lemma.

We remark that with A I, quotients of the form

$$\frac{({}^{n+m+1}Y^m)_{i_1,j_1}}{({}^{n+m}Y^m)_{i_2,j_2}}$$

are of the same order as $(X^{n+m+1})_{1,1}$ in the sense that the ratio of the two is bounded away from zero and infinity.

COROLLARY: *If X^1, X^2, $\cdots$ is a stationary stochastic process satisfying* A I, *then $\lim_{n\to\infty} n^{-1}E\{\log ({}^nY^1)_{i,j}\}$ exists and is the same for all i and j, say E. If, in addition, the X-process satisfies the conditions of Theorem 2, then for all i and j,*

$$\lim_{n\to\infty} n^{-1} \log ({}^nY^1)_{i,j} = E$$

with probability 1.

PROOF: (3.3) and (3.4) imply

$$\min_{i,j} ({}^nY^1)_{i,j} \leqq \|{}^nY^1\| \leqq k \max_{i,j} ({}^nY^1)_{i,j} \leqq kC^2 \min_{i,j} ({}^nY^1)_{i,j}. \tag{3.5}$$

Theorems 1 and 2 now give the required result.

The first part of this corollary is a generalization of the result in [1].

LEMMA 3. *If we define* $({}^{n+m}Y^m)_i = \sum_j ({}^{n+m}Y^m)_{i,j}$ *and if* A I *is satisfied, then*

$$\left| \frac{({}^{n+m}Y^m)_{i_1,j}}{({}^{n+m}Y^m)_{i_1}} - \frac{({}^{n+m}Y^m)_{i_2,j}}{({}^{n+m}Y^m)_{i_2}} \right| \leqq (1 - C^{-3})^n.$$

PROOF: A straightforward computation shows that

$$\begin{aligned} &\frac{({}^{m+r+1}Y^m)_{i_1,j}}{({}^{m+r+1}Y^m)_{i_1}} - \frac{({}^{m+r+1}Y^m)_{i_2,j}}{({}^{m+r+1}Y^m)_{i_2}} \\ &\quad = \sum_s \left\{ \frac{(X^{m+r+1})_{i_1,s}({}^{m+r}Y^m)_s}{({}^{m+r+1}Y^{m+r})_{i_1}} - \frac{(X^{m+r+1})_{i_2,s}({}^{m+r}Y^m)_s}{({}^{m+r+1}Y^{m+r})_{i_2}} \right\} \frac{({}^{m+r}Y^m)_{s,j}}{({}^{m+r}Y^m)_s}. \end{aligned} \tag{3.6}$$

But, for all i,

$$\sum_s \frac{(X^{m+r+1})_{i,s}({}^{m+r}Y^m)_s}{({}^{m+r+1}Y^m)_i} = 1; \tag{3.7}$$

all the summands in (3.7) are positive and by (3.4),

$$\frac{(X^{m+r+1})_{i_2,s}({}^{m+r}Y^m)_s}{({}^{m+r+1}Y^m)_{i_2}} \geqq C^{-3} \frac{(X^{m+r+1})_{i_1,s}({}^{m+r}Y^m)_s}{({}^{m+r+1}Y^m)_{i_1}}. \tag{3.8}$$

We proceed now as in [3], pp. 173–174. let S' be the set of indices s for which the summand in the right hand side of (3.6) is positive and S'' the set for which the summand is negative. By (3.3), (3.7) and (3.8)

$$\begin{aligned} 0 &\leq \sum_{s\varepsilon S'} \left\{ \frac{(X^{m+r+1})_{i_1,s}({}^{m+r}Y^m)_s}{({}^{m+r+1}Y^m)_{i_1}} - \frac{(X^{m+r+1})_{i_2,s}({}^{m+r}Y^m)_s}{({}^{m+r+1}Y^m)_{i_2}} \right\} \\ &= -\sum_{s\varepsilon S''} \left\{ \frac{(X^{m+r+1})_{i_1,s}({}^{m+r}Y^m)_s}{({}^{m+r+1}Y^m)_{i_1}} - \frac{(X^{m+r+1})_{i_2,s}({}^{m+r}Y^m)_s}{({}^{m+r+1}Y^m)_{i_2}} \right\} \\ &\leqq (1 - C^{-3}). \end{aligned}$$

Splitting the sum in (3.6) up into a sum over S' and one over S'' one obtains from (3.6)–(3.9),

$$\begin{aligned} \max_i \frac{({}^{m+r+1}Y^m)_{i,j}}{({}^{m+r+1}Y^m)_i} - \min_i \frac{({}^{m+r+1}Y^m)_{i,j}}{({}^{m+r+1}Y^m)_i} \leq (1 - C^{-3}) \Bigg\{ &\max_i \frac{({}^{m+r}Y^m)_{i,j}}{({}^{m+r}Y^m)_i} \\ &- \min_i \frac{({}^{m+r}Y^m)_{i,j}}{({}^{m+r}Y^m)_i} \Bigg\}. \end{aligned}$$

The proof is now completed by induction on r (compare [3], pp. 173–174).

Consider again the (X, Z) process defined by the stationary measure μ on Ω as in Lemma 1. Put

$$(3.10)\qquad a = \int \log \frac{(X^2Z^1)_{1,1}}{(Z^1)_{1,1}}\, d\mu = \int \log \frac{(X^{k+2}Z^{k+1})_{1,1}}{(Z^{k+1})_{1,1}}\, d\mu,$$

$$(3.11)\qquad \begin{aligned} c_r &= \int \left(\log \frac{(X^{r+1}Z^r)_{1,1}}{(Z^r)_{1,1}} - a\right)\left(\log \frac{(X^2Z^1)_{1,1}}{(Z^1)_{1,1}} - a\right) d\mu \\ &= \int \left(\log \frac{X^{r+k+1}Z^{r+k})_{1,1}}{(Z^{r+k})_{1,1}} - a\right)\left(\log \frac{(X^{k+2}Z^{k+1})_{1,1}}{(Z^{k+1})_{1,1}} - a\right) d\mu, \end{aligned}$$

and

$$(3.12)\qquad b = c_1 + 2\sum_{r=2}^{\infty} c_r .$$

(The convergence of the series in (3.12) will follow from the proof of Theorem 3.)

In order to prove the asymptotic normality of $\log ({}^nY^1)_{1,1}$ we introduce the following "independence assumption" A II: If Ω^1 is a measurable set in the sample space of the X-process, defined in terms of $X^{m+n+1}, X^{m+n+2}, \cdots$ only, then $| P\{\Omega^1 \mid X_1, \cdots, X_m\} - P\{\Omega^1\}| \leqq D_1\lambda_1^n P\{\Omega^1\}$, where D_1 and λ_1 are some fixed positive constants and $\lambda_1 < 1$.

Note that A II is satisfied if the X^i are mutually independent or if the X-process is an aperiodic Markov chain with finitely many states (i.e. X^1 can only take finitely many values) with one ergodic class ([3] Ch. V, §2).

In addition to A I and A II we shall need a condition regarding the moments of $\log (X^1)_{1,1}$. We then have

THEOREM 3. *If* A I *and* A II *are satisfied, and if*

$$(3.13)\qquad E \mid \log (X^1)_{1,1} \mid^{2+\delta} < \infty$$

for some $\delta > 0$, *and if* a *and* b *are given by* (3.10) *and* (3.12), *then*

$$(3.14)\qquad \lim_{n\to\infty} P\left\{\frac{\log({}^nY^1)_{i,j} - na}{(nb)^{1/2}} \leqq x\right\} = (2\pi)^{-1/2}\int_{-\infty}^{x} e^{-1/2t^2}\, dt$$

when $b \neq 0$. *If* $b = 0$, *then*

$$(3.15)\qquad (\log ({}^nY^1)_{i,j} - na)/(n)^{1/2} \to 0 \text{ in probability.}$$

REMARK. Since, by (3.4), $| \log ({}^nY^1)_{i_1,j_1} - \log ({}^nY^1)_{i_2,j_2} | \leqq 2 \log C$, it suffices to prove the result for $\log ({}^nY^1)_{1,1}$. This theorem will then give the joint limiting distribution of all $\log ({}^nY^1)_{i,j}$.

Moreover, (3.5) shows that (3.14) (or (3.15)) also holds for $\log \| {}^nY^1 \|$ instead of $\log ({}^nY^1)_{1,1}$.

PROOF. As remarked, we can take $i = j = 1$. Then $\log ({}^nY^1)_{1,1} = \sum_{k=1}^{n} \xi_k$ where

$$\xi_k = \log [({}^kY^1)_{1,1}]/[({}^{k-1}Y^1)_{1,1}] (k > 1) \quad\text{and}\quad \xi_1 = \log ({}^1Y^1)_{1,1} .$$

We shall show that Bernstein's central limit theorem [2] dealing with "almost independent" random variables is applicable. Strictly speaking, Bernstein's theorem would require $E \mid \log (X^1)_{1,1} \mid^3 < \infty$. We shall therefore follow the treatment given by Doob in [3], Ch. V, §7 for the special case of Markov chains. This will show that our conditions are sufficient. We may and will choose δ fixed and $0 < \delta < 1$ such that (3.13) is satisfied.

The heart of the proof is the expansion (cf [3] p. 38):

$$
\begin{aligned}
E\left\{ \exp it\gamma \sum_{s=m+n+1}^{m+n+k} (\xi_s - a) \mid X_1, \cdots X_m \right\} \\
&= 1 + it\gamma \sum_{s=m+n+1}^{m+n+k} E\{\xi_s - a \mid X_1, \cdots X_m\} \\
&\quad - (t\gamma)^2/2)E\left\{ \left(\sum_{s=m+n+1}^{m+n+k} \xi_s - a \right)^2 \middle| X_1, \cdots X_m \right\} \\
&\quad + O\left(\mid \gamma t \mid^{2+\delta} E\left\{ \left| \sum_{s=m+n+1}^{m+n+k} \xi_s - a \right|^{2+\delta} \middle| X_1, \cdots X_m \right\} \right).
\end{aligned}
\tag{3.16}
$$

We shall show that there exist positive constants D and λ with $0 < \lambda < 1$ such that

$$
\mid E\{\xi_s - a \mid X_1, \cdots, X_m\} \mid \leqq D\,\lambda\mid^{s-m}, \tag{3.17}
$$

$$
\left| E\left\{ \left(k^{-1/2} \sum_{s=m+n+1}^{m+n+k} \xi_s - a \right)^2 \mid X_1, \cdots, X_m \right\} - c_1 - 2k^{-1} \sum_{s=1}^{k} \sum_{r=s+1}^{k} c_{r-s+1} \right| \leqq D\lambda^n, \tag{3.18}
$$

and

$$
E\{\mid \sum_{s=m+1}^{m+k} (\xi_s - a)\mid^{2+\delta'} \mid X_1, \cdots, X_m\} \leqq Dk^{1+(\delta'/2)} \tag{3.19}
$$

for $\delta' = 0$ or δ. The relations (3.16)–(3.19) replace the Lemmas 7.1–7.4 in [3] pp. 222–228, and with their help one can then compute

$$
E\{\exp it\gamma \sum_{s=1}^{n} (\xi_s - a)\} \qquad \text{for } \gamma = (nb)^{-1/2}
$$

and complete the proof along the lines of pp. 228–230 in [3].

If A, B, C are $k \times k$ matrices, then

$$
\begin{aligned}
(ABC)_{1,1}/(BC)_{1,1} &= \sum_i A_{1,i}[(BC)_{i,1}/(BC)_{1,1}] \\
&= \sum_i A_{1,i}(B_i/B_1) + [\sum_i A_{1,i}B_i \sum_j (B_{i,j}/B_i - B_{1,j}/B_1)C_{j,1}]/[\sum_j B_{1,j}C_{j,1}],
\end{aligned}
\tag{3.20}
$$

where $B_i = \sum_j, B_{i,j}$. Now take[4]

[4] $[x]$ denotes the largest integer $\leqq x$.

$$A = X^{m+n+1},$$
$$B = {}^{m+n}Y^{m+[n/2]},$$
$$C = {}^{m+[n/2]}Y^{1}.$$

With these substitutions it follows from Lemma 2 applied to ratios such as $A_{1,i_1}/A_{1,i_2}$ and $B_{i,i_1}/B_{1,i_2}$, and from Lemma 3, that the second term in the right hand side of (3.20) is $O((X^{m+n+1})_{1,1}(1 - C^{-3})^{n/2})$ uniformly in the matrices $X_1, \cdots, X_{m+n}$. Hence, again using Lemma 2 and the identity

$$\log(\alpha + \beta) = \log \alpha + \log(1 + (\beta/\alpha)) = \log \alpha + O(\beta/\alpha)$$

as $\beta/\alpha \to 0$, one has, uniformly in $X_1, \cdots, X_{m+n}$

$$\begin{aligned}\xi_{m+n+1} &= \log\left\{\sum_i (X^{m+n+1})_{1,i} \frac{({}^{m+n}Y^{m+[n/2]})_i}{({}^{m+n}Y^{m+[n/2]})_i}\right\} \\ &+ O(1 - C^{-3})^{n/2} = \log \frac{({}^{m+n+1}Y^{m+[n/2]})_1}{({}^{m+n}Y^{m+[n/2]})_1} + O(1 - C^{-3})^{n/2}.\end{aligned} \tag{3.21}$$

This shows, in particular, that (again by Lemma 2) $|\xi_{m+n+1} - \log(X^{m+n+1})_{1,1}|$ is bounded and hence (using A II and (3.13)), $E\{|\xi_{m+n+1}|^{2+\delta} \mid X_1, \cdots, X_m\}$ is bounded.

We also obtain from (3.21)

$$\begin{aligned}E\left\{\xi_{m+n+1} \,\middle|\, X_1, \cdots, X_m\right\} &= E\left\{\log \frac{({}^{m+n+1}Y^{m+[n/2]})_1}{({}^{m+n}Y^{m+[n/2]})_1} \,\middle|\, X_1, \cdots, X_m\right\} \\ + O(1 - C^{-3})^{n/2} &= E\left\{\log \frac{({}^{m+n+1}Y^{n+[n/2]})_1}{({}^{m+n}Y^{m+[n/2]})_1}\right\} + O(1 - C^{-3})^{n/2} + O(\lambda_1^{n/2})\end{aligned} \tag{3.22}$$

uniformly in $X_1, \cdots, X_m$. The last equality is a consequence of A II because the expression between the braces depends only on $X_{m+[n/2]}, \cdots, X_{m+n+1}$, and the uniform boundedness of

$$E\left\{\left|\log \frac{({}^{m+n+1}Y^{m+[n/2]})_1}{({}^{m+n}Y^{m+[n/2]})_1}\right|\right\}$$

(cf. remark after Lemma 2). Setting $\lambda_2 = \{\max(1 - C^{-3}, \lambda_1)\}^{1/2} (< 1)$ and using the stationarity of the X-process, (3.22) can be rewritten as

$$E\{\xi_{m+n+1} \mid X_1, \cdots X_m\} = E\xi_{n-[n/2]+2} + O(\lambda_2^n) \tag{3.23}$$

uniformly in $X_1, \cdots, X_m$. But, by (3.21) and its analogue, derived by putting $A = X^{k+1}$, $B = {}^kY^2$, $C = Z^1$ in (3.20), one has

$$\begin{aligned}\xi_{k+1} &= \log \frac{(X^{k+1} \cdots X^1)_{1,1}}{(X^k \cdots X^1)_{1,1}} = \log \frac{(X^{k+1} \cdots X^2)_{1,1}}{(X^k \cdots X^2)_{1,1}} \\ &+ O(1 - C^{-3})^k = \log \frac{(X^{k+1} \cdots X^2 Z^1)_{1,1}}{(X^k \cdots X^2 Z^1)_{1,1}} + O(1 - C^{-3})^k,\end{aligned} \tag{3.24}$$

as long as $(Z^1)_{i,j} > 0$. However, by Lemma 2 and the construction of μ, $(Z^1)_{i,j} > 0$ a.e. with respect to μ. Therefore

$$
\begin{aligned}
E\xi_{k+1} &= \int \log \frac{(X^{k+1} \cdots X^1)_{1,1}}{(X^k \cdots X^1)_{1,1}}\, d\mu \\
&= \int \log \frac{(X^{k+1} \cdots X^2 Z^1)_{1,1}}{(X^k \cdots X^2 Z^1)_{1,1}}\, d\mu + O(1 - C^{-3})^k \qquad (3.25)\\
&= \int \log \frac{(X^{k+1} Z^k)_{1,1}}{(Z^k)_{1,1}}\, d\mu + O(1 - C^{-3})^k = a + Q(1 - C^{-3})^k.
\end{aligned}
$$

(3.22) and (3.25) prove (3.17) for a suitable D and $\lambda \geqq \lambda_2$.

As in (3.22) and (3.25) one shows

$$
\begin{aligned}
&E\{(\xi_{m+n+r} - a)(\xi_{m+n+1} - a) \mid X_1, \cdots, X_m\} \\
&= E\{(\xi_{n-[n/2]+r} - a)(\xi_{n-[n/2]+1} - a)\} + O(\lambda_2^n) = c_r + O(\lambda_2^n) \qquad (3.26)
\end{aligned}
$$

uniformly in $X_1, \cdots, X_m$. At the same time, using (3.17)

$$
\begin{aligned}
&E\{(\xi_{m+n+r} - a)(\xi_{m+n+1}) - a) \mid X_1, \cdots, X_m\} \\
&= E\{(\xi_{m+n+1} - a)E\{(\xi_{m+n+r} - a) \mid X_1 \cdots X_{m+n+1}\} X_1 \cdots X_m\} = O(\lambda_2^r) \qquad (3.27)
\end{aligned}
$$

uniformly in $X_1, \cdots, X_m$. Hence also $c_r = O(\lambda_2^r)$ so that the series in (3.12) converges. (3.18) can now be proved quite easily, as

$$
\begin{aligned}
E\left\{\left(k^{-1/2} \sum_{s=m+n+1}^{m+n+k} \xi_s - a\right)^2 \mid X_1 \cdots X_m\right\} &= k^{-1} \sum_{s=m+n+1}^{m+n+k} E\{(\xi_s - a)^2 \mid X_1 \cdots X_m\} \\
&\quad + 2k^{-1} \sum_{s=m+n+1}^{m+n+k} \sum_{r=s+1}^{m+n+k} E\{(\xi_s - a)(\xi_r - a) \mid X_1 \cdots X_m\},
\end{aligned}
$$

$$E\{(\xi_s - a)^2 \mid X_1 \cdots X_m\} = c_1 + 0(\lambda_2^{s-m}),$$

and

$$\sum_{r=s+1}^{m+n+k} E\{(\xi_s - a)(\xi_r - a) \mid X_1 \cdots X_m\} = \sum_{r=s+1}^{s+u} c_{r-s+1} + u0(\lambda_2^{s-m}) + O(\lambda_2^u).$$

Taking $u = \min(m + n + k - s, s - m)$ we obtain (3.18) for a $\lambda < 1$ satisfying $n\lambda_2^n \leqq D_2\lambda^n$ for some D_2 and all positive n. (3.18) of course implies (3.19) for $\delta' = 0$. The proof of (3.19) can be completed as in Lemma (7.4) p. 225 in [3]. In fact, if

$$u_n = \sum_{s=m+k+1}^{m+k+n} (\xi_s - a) \quad \text{and} \quad v_n = \sum_{s=m+k+n+1}^{m+k+2n} (\xi_s - a),$$

then

$$
\begin{aligned}
E\{|u_n + v_n|^{2+\delta} \mid X_1 \cdots X_m\} &\leqq E\{(u_n + v_n)^2(|u_n|^\delta + |v_n|^\delta) \mid X_1 \cdots X_m\} \\
&\leqq E\{|u_n|^{2+\delta} \mid X_1 \cdots X_m\} + E\{|v_n|^{2+\delta} \mid X_1 \cdots X_m\} \\
&\quad + 2E\{|u_n|^{1+\delta} |v_n| \mid X_1 \cdots X_m\} + 2E\{|u_n| \, |v_n|^{1+\delta} \mid X_1 \cdots, X_m\}.
\end{aligned}
$$

However, $E\{|u_n|^{1+\delta}|v_n| \mid X_1\cdots X_m\}$

$$\begin{aligned}
&\leqq E\{|u_n|^{1+\delta}E\{|v_n| \mid X_1\cdots X_{m+k+n}\} \mid X_1,\cdots,X_m\}\\
&\leqq E\{|u_n|^{1+\delta}(E\{|v_n|^2 \mid X_1\cdots X_{m+k+n}\})^{1/2} \mid X_1,\cdots,X_m\}\\
&\leqq Dn^{1/2}(E\{|u_n|^2 \mid X_1\cdots X_m\})^{(1+\delta)/2}\\
&\leqq Dn^{1/2}(Dn)^{(1+\delta)2} = D^{(3+\delta)/2}n^{(2+\delta)/2}
\end{aligned}$$

and similarly for $E\{|u_n||v_n|^{1+\delta} \mid X_1\cdots X_m\}$. Hence

$$\begin{aligned}
&E\{|u_n+v_n|^{2+\delta}|X_1,\cdots,X_m\}\\
(3.28)\quad &\leqq 2\sup_{k>0} E\left\{\left|\sum_{s=m+k+1}^{m+k+n}(\xi_s-a)\right|^{2+\delta}|X_1\cdots X_m\right) + 2D^{(3+\delta)/2}n^{(2+\delta)/2}.
\end{aligned}$$

The inequality (3.28) replaces (7.11), p. 226, in [3] if c_n in [3] is replaced by

$$\sup_{k\geqq 0} E\left\{\left|\sum_{s=m+k+1}^{m+k+n}(\xi_s-a)\right|^{2+\delta}|X_1\cdots X_m\right\}.$$

The remainder of the proof of (3.19) can be copied from pp. 226–227 of [3] and the theorem now follows from (3.16)–(3.19) as indicated before.

4. Two examples. The first example shows that one cannot prove the corollary to Theorem 2 without some positivity assumption as in (3.1), even though Theorems 1 and 2 show that the corresponding result for the norm is true without A I.

EXAMPLE 1. Take the X^i mutually independent with the distribution

$$P\left\{X^i=\begin{pmatrix}0&1\\1&0\end{pmatrix}\right\}=\tfrac{1}{2},$$

$$P\left\{X^i=\begin{pmatrix}2&0\\0&1\end{pmatrix}\right\}=\tfrac{1}{2}.$$

In this case ${}^nY^1$ is sometimes of the form $\begin{pmatrix}2^{k_1}&0\\0&2^{k_2}\end{pmatrix}$ and sometimes of the form $\begin{pmatrix}0&2^{k_1}\\2^{k_2}&0\end{pmatrix}$ where $0<\lim k_i/n<1$. Hence $n^{-1}\log({}^nY^1)_{1,1}$ has no limit, nor has $n^{-1}E\log({}^nY^1)_{1,1}$.

Since $\log\|{}^nY^1\|$ behaves slightly better than $\log({}^nY^1)_{1,1}$, as indicated by the last example, one might hope that the central limit theorem would hold at least for $\log\|{}^nY^1\|$ without AI. The following example shows that this is also false.

EXAMPLE 2. Take the X^i mutually independent with the same distribution. All X^i are of the form $\begin{pmatrix}\lambda_1&0\\0&\lambda_2\end{pmatrix}$ where λ_1 and λ_2 are independent strictly positive random variables with the same distribution, and $E\log\lambda_i=0$, $E(\log\lambda_i)^2=1$. Then $({}^nY^1)_{1,2}=({}^nY^1)_{2,1}=0$ while $n^{-1/2}\log({}^nY^1)_{1,1}$ and $n^{-1/2}\log({}^nY^1)_{2,2}$ are independent random variables, both asymptotically normal, with zero mean and unit variance. Since $n^{-1/2} \log \| {}^nY^1 \|$ is the maximum of these two variables it is not asymptotically normal, nor is $n^{-1/2}(\log \| {}^nY^1 \| - nd)$ for any d.

REFERENCES

[1] RICHARD BELLMAN, "Limit theorems for non-commutative operations I," *Duke Math. J.*, Vol. 21 (1954), pp. 491–500.

[2] SERGE BERNSTEIN, "Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes," *Math. Ann.* Vol. 97 (1927), pp. 1–59.

[3] J. L. DOOB, *Stochastic Processes*, John Wiley and Sons, New York, 1953.

[4] G. PÓLYA UND G. SZEGÖ, *Aufgaben und Lehrsätze ans der Analysis*, *Aufgabe I*, 2te Auflage, Springer-Verlag, Berlin, 1954.