ISRAEL JOURNAL OF MATHEMATICS, Vol. 46, Nos. 1-2, 1983

RANDOM MATRIX PRODUCTS AND MEASURES ON PROJECTIVE SPACES

BY
H. FURSTENBERG AND Y. KIFER

ABSTRACT
The asymptotic behavior of $\|X_n X_{n-1} \cdots X_1 v\|$ is studied for independent matrix-valued random variables X_n. The main tool is the use of auxiliary measures in projective space and the study of markov processes on projective space.

Let $X_1, X_2, X_3, \cdots$ be a stationary stochastic process with values in the space of $m \times m$ real matrices. Let $\| \quad \|$ denote a (Banach algebra) norm on the space of these matrices. It was proved in [3] that if the expectation $E(\log^+ \|X_i\|)$ is finite, then

$$\lim_{n \to \infty} \|X_n X_{n-1} \cdots X_1\|^{1/n} \tag{1}$$

exists with probability one. If the variables $\{X_n\}$ are independent with common distribution μ, the limit in (1) is a constant (by the zero-one law) depending only on μ, say $\beta(\mu)$. Unlike the classical law of large numbers, the parameter $\beta(\mu)$ is not directly computable as an integral with respect to the measure μ. As a result the behavior of $\beta(\mu)$ as a function of μ is much more recondite and there are still simple questions that are unanswered.

Here we return to a theme that was taken up in [2] expressing $\beta(\mu)$ in terms of μ and some auxiliary measures on the $(m-1)$-dimensional projective space. These auxiliary measures will also enable us to study the asymptotic behavior of the vector norms $\|X_n X_{n-1} \cdots X_1 v\|$ for $v \in \mathbf{R}^m$. As an application of our analysis we present a stability result for $\beta(\mu)$ with respect to perturbation of μ.

We remark that since this paper was first written similar results have been obtained by Hennion [4].

Received June 26, 1983

The easiest way to prove the existence of the limit in (1) is to invoke Kingman's subadditive ergodic theorem ([7]).

THEOREM (Kingman). *Let* $W_{mn}, m, n = 0, 1, 2, \cdots, m < n$ *be a family of real-valued random variables satisfying*

(i) *The distribution of* W_{mn} *depends only on* $n - m$.

(ii) *For* $m < n < p$, $W_{mp} \leqq W_{mn} + W_{np}$.

(iii) *The expectation* $E(W_{01}^{+})$ *is finite.*

Then with probability 1, *the expression* $(1/n)W_{0n}$ *converges to a limit which may equal* $-\infty$.

The law of large numbers for matrix norms now follows by taking $W_{mn} = \log \| X_n X_{n-1} \cdots X_{m+1} \|$.

A more precise result may be proved using Osoledec's "multiplicative ergodic theorem" ([8], [9]) which is a far reaching generalization of the results of [3]. According to Osoledec's theorem, if $X_1, X_2, \cdots, X_n, \cdots$ is a matrix valued stationary stochastic process with $E(\log^+ \| X_i \|) < \infty$, then for almost every choice of the sequence $X_1, X_2, \cdots, X_n, \cdots$ the limit of $\| X_n X_{n-1} \cdots X_1 v \|^{1/n}$ as $n \to \infty$ exists for every vector v. Here $\| \quad \|$ is any (Banach space) norm on $\mathbf{R}^m$. From this follows the existence of $\lim \| X_n X_{n-1} \cdots X_1 \|^{1/n}$.

If the process $\{X_n\}$ is ergodic it is easy to see that $\lim \| X_n X_{n-1} \cdots X_1 \|^{1/n}$ is, with probability 1, a constant. On the other hand, in the Osoledec result, if we write

$$\beta(\omega, v) = \lim \| X_n(\omega) X_{n-1}(\omega) \cdots X_1(\omega) v \|^{1/n}, \tag{2}$$

for almost all ω, $\beta(\omega, v)$ usually depends non-trivially on v. We may therefore expect that for fixed v, $\beta(\omega, v)$ depends non-trivially on ω. One result of our analysis is that if the $\{X_n\}$ are independent then $\beta(\omega, v)$ depends only on v.

THEOREM A. *If the variables* $X_1, X_2, \cdots, X_n, \cdots$ *are independent and identically distributed in* $\mathrm{GL}(m, \mathbf{R})$ *and if* $E(\log^+ \| X_i \|) < \infty$, *then*

$$\lim_{n \to \infty} \| X_n X_{n-1} \cdots X_1 v \|^{1/n}$$

exists with probability one for every $v \in \mathbf{R}^m$, *and is a constant depending on* v *and on the distribution of* X_i.

In [5] it is shown that $\beta(\mu)$ is not a continuous function of μ with respect to the weak topology on probability measures μ. Namely, one may have $\mu_n \to \mu$ weakly, with μ_n, μ probability measures on unimodular $m \times m$ matrices (with uniformly bounded support) and nonetheless

$$\beta(\mu) > \limsup_{n\to\infty} \beta(\mu_n). \tag{3}$$

(The opposite inequality cannot take place!) We shall present here a family of measures μ for which the instability (3) does not occur.

We shall assume that $\mu\{\text{non-invertible matrices}\} = 0$. Then we can associate to μ the smallest closed subgroup $G_\mu \subset GL(m, \mathbf{R})$ for which $\mu(G_m) = 1$. Given a sequence of probability measures $\{\mu_k\}$ on $GL(m, \mathbf{R})$ we shall say that $\mu_k \to \mu$ *weakly and boundedly* if $\int f d\mu_k \to \int f d\mu$ for all continuous functions $f(g)$ with compact support on $GL(m, \mathbf{R})$, and if

$$\int_{\|g\|>N} \log^+ \|g\| d\mu_k(g) + \int_{\|g^{-1}\|>N} \log^+ \|g^{-1}\| d\mu(g) \to 0$$

as $N \to \infty$, uniformly in k.

THEOREM B. *Let μ be a probability measure on $GL(m, \mathbf{R})$ for which G_μ has the property that there exists at most one non-trivial subspace $V \subset \mathbf{R}^m$ for which $gV \subset V$ for all $g \in G_\mu$. Then if $\mu_k \to \mu$ weakly and boundedly, we will have $\beta(\mu_k) \to \beta(\mu)$.*

1. Random walks and laws of large numbers

Let M be a compact metric space and let $\mathcal{P}(M)$ be the space of probability (borel) measure on M. $\mathcal{P}(M)$ is a compact metric space in the topology of weak convergence. A continuous map $M \to \mathcal{P}(M)$ assigning to each $x \in M$ a measure μ_x defines a *random walk* on M. We define the corresponding Markov operator by

$$Pf(x) = \int f(y) d\mu_x(y)$$

and the adjoint operator P^* is defined on $\mathcal{P}(M)$ by

$$P^*\nu(A) = \int \mu_x(A) d\nu(x)$$

for A a borel set in M.

A stochastic process $\{Z_n, n = 0, 1, 2, \cdots\}$ is a *Markov process with transition probabilities* $\{\mu_x\}$ if

$$P\{Z_{n+1} \in A \mid Z_0, Z_1, \cdots, Z_n\} = \mu_{Z_n}(A).$$

The transition probabilities $\{\mu_x\}$ are determined by the operator P and so we can speak of a *Markov process corresponding to P*. The principal tool that we will use in our discussion is the following general result:

THEOREM 1.1. *Let $\{Z_n, n = 0, 1, 2, \cdots\}$ be a Markov process on M corresponding to the operator P on $\mathscr{C}(M)$ and let $f \in \mathscr{C}(M)$. With probability one,*

$$(1.1) \quad \limsup_{n\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} f(Z_n) \leqq \sup \left\{ \int f d\nu \;\middle|\; \nu \in \mathscr{P}(M) \text{ satisfying } P^*\nu = \nu \right\}.$$

The proof of the theorem is based on a lemma which appears in [2] and which we reproduce here for the reader's convenience.

LEMMA 1.2. *If $f = g - Pg$ for some $g \in \mathscr{C}(M)$ then with probability* 1,

$$\frac{1}{N+1} \sum f(Z_n) \to 0.$$

PROOF. Let $W_{n+1} = \sum_{k=0}^{n} (Pg(Z_k) - g(Z_{k+1}))/k$. We have

$$\begin{aligned} E(W_{n+1} \mid Z_0, Z_1, \cdots, Z_n) &= W_n + \frac{1}{n+1} E(Pg(Z_n) - g(Z_{n+1}) \mid Z_0, Z_1, \cdots, Z_n) \\ &= W_n + \frac{1}{n+1} Pg(Z_n) - \frac{1}{n+1} E(g(Z_{n+1}) \mid Z_0, Z_1, \cdots, Z_n) \\ &= W_n, \end{aligned}$$

since by the definition of a Markov process

$$E(g(Z_{n+1}) \mid Z_0, Z_1, \cdots, Z_n) = Pg(Z_n).$$

Hence $\{W_n\}$ forms a martingale which, by the boundedness of $g(x)$, has bounded second moments. This implies that W_n converges with probability 1, and by Kronecker's lemma it follows that

$$\sum_{k=1}^{n} \{Pg(Z_k) - g(Z_{k+1})\} = o(n)$$

and rearranging terms we have the assertion of the lemma.

LEMMA 1.3. *Let $f \in \mathscr{C}(M)$ be non-negative. Then for any $\varepsilon > 0$ we can write*

$$(1.2) \quad f = Pg - g + h$$

where $g, h \in \mathscr{C}(M)$, and where h satisfies

$$(1.3) \quad \| h \|_\infty \leqq \sup \left\{ \int f d\nu \;\middle|\; \nu \in \mathscr{P}(M) \text{ satisfies } P^*\nu = \nu \right\} + \varepsilon.$$

PROOF. Let $\mathscr{L} \subset \mathscr{C}(M)$ denote the subspace of functions of the form $Pg - g$. Let δ denote the distance of f to $\mathscr{L}$. By the Hahn–Banach theorem, there exists a continuous linear functional on $\mathscr{C}(M)$ vanishing on $\mathscr{L}$, having norm 1 and taking on the value δ at f. Thus there is a (signed) measure λ with $\|\lambda\| = 1$, $P^*\lambda = \lambda$, and $\int f d\lambda = \delta$. Now if we decompose λ into its positive and negative parts $\lambda = \lambda^+ - \lambda^-$ we find $P^*\lambda^+ - P^*\lambda^- = \lambda^+ - \lambda^-$, whence $P^*\lambda^+ \geqq \lambda^+$, $P^*\lambda^- \geqq \lambda^-$ which implies $P^*\lambda^\pm = \lambda^\pm$. Let $\nu_0 = \lambda^+/\|\lambda^+\|$. Then $P^*\nu_0 = \nu_0$ and $\int f d\nu_0 \geqq \int f d\lambda = \delta$. By definition of δ we can find $Pg - g \in \mathscr{L}$ so that

$$\|f - (Pg - g)\|_\infty \leqq \delta + \varepsilon.$$

Letting $h = f - (Pg - g)$ we obtain the assertion of the lemma once we check that

$$\delta \leqq \sup \left\{ \int f d\nu \;\middle|\; \nu \in \mathscr{P}(M) \textit{ and satisfies } P^*\nu = \nu \right\}.$$

But since $\nu = \nu_0$ satisfies $\int f d\nu_0 \geqq \delta$ this concludes the proof.

To prove Theorem 1.1 we first note that adjusting f by a constant, we may assume that f is non-negative. We then have the decomposition (1.2). Now clearly

$$\limsup \frac{1}{N+1} \sum_{n=0}^{N} h(Z_n) \leqq \|h\|_\infty$$

and so by (1.3) we complete the proof of the theorem.

Note that in the course of the proof we have also established the existence of measures μ with $P^*\nu = \nu$. This was based on the Hahn–Banach theorem. One can also prove this using an appropriate fixed-point theorem. In any case, the set on the right of (1.1) is non-empty.

THEOREM 1.4. *Let $\{Z_n, n = 0, 1, 2, \cdots\}$ be a Markov process on M corresponding to the operator P and let $f \in \mathscr{C}(M)$ be such that for all $\nu \in \mathscr{P}(M)$ for which $P^*\nu = \nu$ the integral $\int f d\nu$ takes on the same value $\bar{f}$. Then with probability one*

$$\frac{1}{N+1} \sum_{n=0}^{N} f(Z_n) \to \bar{f}$$

as $N \to \infty$.

This follows from Theorem 1.1 applied to the functions f and $-f$.

2. The behavior of $\|X_n X_{n-1} \cdots X_1 v\|$

We shall apply the results of the foregoing section to the following situation. Set $G = \mathrm{GL}(m, \mathbf{R})$, the locally compact group of invertible $m \times m$ matrices, and let P^{m-1} denote the $(m-1)$-dimensional projective space. We suppose we are given a probability measure μ on G, and we suppose for the moment that μ has support on a compact subset $Q \subset G$. Let $M = Q \times P^{m-1}$.

We may identify P^{m-1} with lines through the origin in $\mathbf{R}^m$ and since the matrices of G send these lines to themselves, we have a natural action of G on P^{m-1}. For $g \in G$ and $u \in P^{m-1}$ we shall indicate this action by gu. Let $X_1, X_2, \cdots, X_n, \cdots$ be a sequence of independent identically distributed G-valued random variables with distributions μ. The sequence of M-valued variables

$$\begin{aligned} &Z_0 = (e, u), \quad Z_1 = (X_1, X_1 u), \quad Z_2 = (X_2, X_2 X_1 u), \cdots \\ &Z_n = (X_n, X_n X_{n-1} \cdots X_1 u) \end{aligned} \tag{2.1}$$

defines a Markov process on M corresponding to the Markov operator

$$Pf(g, x) = \int f(g', g'x) d\mu(g') \tag{2.2}$$

and

$$\begin{aligned} E(f(Z_n) \mid Z_0, Z_1, \cdots, Z_{n-1}) &= E(f(X_n, X_n X_{n-1} \cdots X_1 u) \mid X_1, \cdots, X_{n-1}) \\ &= \int f(g', g'X_{n-1} \cdots X_1 u) d\mu(g') = Pf(Z_{n-1}). \end{aligned}$$

Let λ be a measure on M satisfying $P^*\lambda = \lambda$. By (2.2)

$$\int f d\lambda = \int Pf d\lambda = \iint f(g', g'x) d\mu(g') d\nu(x) \tag{2.3}$$

where ν is the projection of λ on P^{m-1}. If $f(g, x) = \varphi(x)$ then

$$\int \varphi d\nu = \int f d\lambda = \iint \varphi(g'x) d\mu(g') d\nu(x).$$

For any $\nu \in \mathcal{P}(P^{m-1})$ define $\mu * \nu$ by

$$\int \varphi d\mu * \nu = \iint \varphi(gx) d\mu(g) d\nu(x).$$

We say that ν is *μ-stationary* if $\mu * \nu = \nu$. Note that if Y is a P^{m-1}-valued random variable, and X is a G-valued random variable independent of Y and if X and Y have distributions μ and ν respectively, then XY has distribution $\mu * \nu$.

If conversely ν is a μ-stationary measure and λ is defined by

$$\int f d\lambda = \iint f(g, gx) d\mu(g) d\nu(x), \tag{2.4}$$

then $P^*\lambda = \lambda$. For

$$\begin{aligned}\int f dP^*\lambda &= \int Pf d\lambda \\ &= \iint Pf(g, gx) d\mu(g) d\nu(x) \\ &= \iiint f(g', g'gx) d\mu(g') d\mu(g) d\nu(x) \\ &= \int f(g', g'y) d\mu(g') d\nu(g) \\ &= \int f d\lambda.\end{aligned}$$

We thus obtain a one-to-one correspondence between $\lambda \in \mathcal{P}(M)$ with $P^*\lambda = \lambda$ and $\nu \in \mathcal{P}(P^{m-1})$ with $\mu * \nu = \nu$.

For $u \in P^{m-1}$ let $\hat{u} \in \mathbf{R}^m$ denote any vector $\neq 0$ on the line through the origin corresponding to u. For any $g \in C$ the expression $\|\hat{u}\|/\|g^{-1}\hat{u}\|$ depends only on g and u. We can therefore define a continuous function on M:

$$\rho(g, u) = \log \|\hat{u}\|/\|g^{-1}\hat{u}\|.$$

Consider now the averages

$$\frac{1}{N+1}\sum_{n=0}^{N} \rho(Z_n).$$

We have

$$\begin{aligned}\rho(Z_0) + \rho(Z_1) + \cdots + \rho(Z_N) &= (\log \|X_1\hat{u}\| - \log \|\hat{u}\|) \\ &\quad + (\log \|X_2X_1\hat{u}\| - \log \|X_1\hat{u}\|) + \cdots \\ &\quad + (\log \|X_NX_{N-1}\cdots X_1\hat{u}\| \\ &\qquad - \log \|X_{N-1}X_{N-2}\cdots X_1\hat{u}\|) \\ &= \log \|X_NX_{N-1}\cdots X_1\hat{u}\| - \log \|\hat{u}\|.\end{aligned}$$

Consequently

$$\frac{1}{N+1}\sum_{n=0}^{N} \rho(Z_n) = \frac{1}{N+1} \log \|X_NX_{N-1}\cdots X_1\hat{u}\| + O(1).$$

We wish to apply Theorems 1.1 and 1.4. We use the correspondence between λ and $P^*\lambda = \lambda$ and μ-stationary ν and we find that if λ corresponds to ν, then

$$\int \rho(g,u)d\lambda(g,u) = \iint \rho(g,gu)d\mu(g)d\nu(u)$$
$$= \iint \log\|g\hat{u}\|d\mu(g)d\nu(u).$$

We can now formulate the result.

THEOREM 2.1. *If μ is a measure of compact support in $\mathrm{GL}(m,\mathbf{R})$ and if $\{X_n\}$ is a sequence of i.i.d. random variables with values in $\mathrm{GL}(m,\mathbf{R})$ and with distribution μ, then for any $v \in \mathbf{R}^m$ we have with probability one,*

$$\limsup_{N\to\infty} 1/N \log\|X_N X_{N-1}\cdots X_1 v\| \tag{2.5}$$
$$\leqq \sup\left\{\iint \log\frac{\|g\hat{u}\|}{\|\hat{u}\|}d\mu(g)d\nu(u) \,\Big|\, \nu\in\mathcal{P}(P^{m-1}) \text{ and } \mu*\nu=\nu\right\}.$$

Moreover, if for all μ-stationary measures ν the expression

$$\iint \log\frac{\|g\hat{u}\|}{\|\hat{u}\|}d\mu(g)d\nu(u)$$

takes on the same value β, then with probability one

$$\frac{1}{N}\log\|X_N X_{N-1}\cdots X_1 v\| \to \beta. \tag{2.6}$$

If $v_1,\cdots,v_m$ form a basis of $\mathbf{R}^m$ we can identify $\|g\|$ with $\max\|gv_i\|$. (Note that any two norms $\|\ \|_1$, $\|\ \|_2$ either on matrices or on vectors will satisfy $(\|\ \|_1/\|\ \|_2)^{1/N}\to 1$.) Theorem 2.1 implies a corresponding result for the behavior of

$$\frac{1}{N}\log\|X_N X_{N-1}\cdots X_1\|.$$

But we shall immediately see that a stronger statement can be made here.

Let ν be some μ-stationary measure. Let U_0 be a P^{m-1}-valued random variable of distribution ν defined simultaneously with the i.i.d. sequence $\{X_n\}$ and assume it is independent of all the latter. Now define

$$Z_n^\nu = (X_n, X_n X_{n-1}\cdots X_1 U_0).$$

Z_n^ν is again a Markov process corresponding to the Markov operator P. If we set

$U_n = X_n X_{n-1} \cdots X_1 U_0$ we see inductively that the distribution of U_n is again V since $U_n = X_n U_{n-1}$ and X_n and U_{n-1} are independent. It follows that $Z_n^\nu = (X_n, U_n)$ is a stationary process. (A Markov process with stationary transition probabilities is stationary if the individual distributions remain the same.) We now apply the ergodic theorem to conclude that with probability one,

$$\frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 \hat{U}_0 \| \to \bar{\rho} \tag{2.7}$$

where $\bar{\rho}$ is a random variable satisfying

$$E(\bar{\rho}) = E(\rho(X_1, X_1 U_0)) = E(\log \| X_1 \hat{U}_0 \|) = \iint \log \frac{\| g\hat{u} \|}{\| \hat{u} \|} d\mu(g) d\nu(u). \tag{2.8}$$

Now $\bar{\rho}$ must take on values $\geqq E(\bar{\rho})$ and so with positive probability

$$\limsup_{N \to \infty} \frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 \| \geqq E(\bar{\rho}) = \iint \log \frac{\| g\hat{u} \|}{\| \hat{u} \|} d\mu(g) d\nu(u). \tag{2.9}$$

The lim inf in question is measurable with respect to the tail field of $\{X_n\}$ and so by the zero-one law it is a.e. constant. Thus (2.9) holds with probability one. Moreover, the expression to the left in (2.9) does not depend on the choice of ν. If we now set

$$\beta(\mu) = \sup \left\{ \iint \log \frac{\| g\hat{u} \|}{\| \hat{u} \|} d\mu(g) d\nu(u) \,\middle|\, \nu \in \mathcal{P}(P^{m-1}), \mu * \nu = \nu \right\} \tag{2.10}$$

then

$$\liminf \frac{1}{N} \| X_N X_{N-1} \cdots X_1 \| \geqq \beta(\mu).$$

Comparing this with (2.5) we obtain

THEOREM 2.2. *If μ is a measure of compact support in* GL(m, **R**) *and if $\{X_n\}$ is a sequence of independent identically distributed* GL(m, **R**)*-valued random variables with distribution μ, then with probability one,*

$$\lim \frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 \| = \beta(\mu) \tag{2.11}$$

where $\beta(\mu)$ is defined by (2.10).

The foregoing results may be extended to the case of non-compactly supported measures μ provided some boundedness restriction is imposed. We shall assume μ satisfies

$$\int [\log^+ \|g\| + \log^+ \|g^{-1}\| d\mu(g) < \infty. \tag{2.12}$$

Let $\bar{G}$ denote the one-point compactification of the locally compact group $\mathrm{GL}(m, \mathbf{R})$ and set $M = \bar{G} \times P^{m-1}$. On M we define transition probabilities as before, setting

$$Pf(g, x) = \int f(g', g'x) d\mu(g')$$

for all $g \in \bar{G}$. For each positive $T < \infty$ we consider the function $\rho_T(g, u)$ defined as before by $\log(\|\hat{u}\|/\|g^{-1}\hat{u}\|)$ for $\|g\|, \|g^{-1}\| \leqq T$ and extended continuously to all of M with the same bounds. At the same time consider $\rho(g, u) = \log(\|\hat{u}\|/\|g^{-1}\hat{u}\|)$ defined on $G \times P^{m-1} \subset M$. Form $\{Z_n\}$ as before. Then

$$\begin{aligned} &\left| \frac{1}{N+1} \sum \rho(Z_n) - \frac{1}{N+1} \rho_T(Z_n) \right| \\ &\leqq \frac{1}{N+1} \sum (\log^+ \|X_n\| + \log^+ \|X_n^{-1}\| + \log T) 1_{B_T}(X_n) \end{aligned} \tag{2.13}$$

where $B_T = \{g \mid \max\{\|g\|, \|g^{-1}\|\} > T\}$. By the ergodic theorem for the process $\{X_n\}$ the right-hand side of (2.13) converges to the limit

$$\begin{aligned} &\int_{B_T} (\log^+ \|g\| + \log^+ \|g^{-1}\|) d\mu(g) + \log T \mu(B_T) \\ &\leqq 2 \int_{B_T} (\log^+ \|g\| + \log^+ \|g^{-1}\|) d\mu(g). \end{aligned}$$

By (2.12) this $\to 0$ as $T \to \infty$. It follows that for large T, the asymptotic behavior of $\|X_N X_{N-1} \cdots X_1 v\|$ can be estimated by integrals of ρ_T with respect to measures λ satisfying $P^*\lambda = \lambda$. Comparing ρ with ρ_T we find that

$$\left| \int \rho d\lambda - \int \rho_T d\lambda \right| \leqq \int_{B_T} (\log^+ \|g\| + \log^+ \|g^{-1}\| + \log T) d\mu(g)$$

which converges to zero as $T \to \infty$. The result is the following

THEOREM 2.3. *The conclusions of Theorem* 2.1 *and Theorem* 2.2 *are valid if the measure* μ *satisfies*

$$\int [\log^+ \|g\| + \log^+ \|g^{-1}\|] d\mu(g) < \infty.$$

3. The filtration of $\mathbf{R}^m$ for random matrix products

We consider a fixed measure μ on $\mathrm{GL}(m, \mathbf{R})$ satisfying (2.12). For any μ-stationary measure ν on P^{m-1}, let $\alpha(\nu)$ be given by

$$\alpha(\nu)=\iint \frac{\|g\hat{u}\|}{\|\hat{u}\|}\,d\mu(g)d\nu(u).$$

We have

$$\beta(\mu)=\sup\{\alpha(\nu)\,|\,\nu\in\mathcal{P}(P^{m-1}) \text{ and } \mu*\nu=\nu\}.$$

We shall speak of $\beta(\mu)$ as the *rate of growth* of the matrix products for μ.

Let N denote the compact convex subset of $\mathcal{P}(P^{m-1})$ consisting of μ-stationary measures. To each $\nu\in N$ we can attach a stationary Markov process

$$Z_n^\nu=(X_n, X_nX_{n-1}\cdots X_1U_0)$$

where U_0 has distribution ν. Assume that Z_n^ν is an ergodic process. Then with probability one

$$\lim\frac{1}{N+1}\sum_{n=0}^{N}\rho(Z_n^\nu)=\lim\frac{1}{N}\log\|X_NX_{N-1}\cdots X_1\hat{U}_0\|$$

$$(3.1)\qquad =\iint\log\frac{\|g\hat{u}\|}{\|\hat{u}\|}\,d\mu(g)d\nu(u)$$

$$=\alpha(\nu)$$

by the ergodic theorem. We formulate this as follows:

LEMMA 3.1. *If $\nu\in N$ is such that Z_n^ν is ergodic, then for ν-almost every direction u, if $v\in\mathbf{R}^{m-1}$ is a vector in the direction u, then*

$$\frac{1}{N}\log\|X_NX_{N-1}\cdots X_1v\|\to\alpha(\nu).$$

LEMMA 3.2. *The measure ν with Z_n^ν ergodic are just the extremal points of N.*

PROOF. It is readily checked that if ν is non-extremal then Z_n^ν is non-ergodic. For the converse direction we use the fact that the invariant functions for a stationary Markov process $\{Z_n^\nu\}$ have the form $\varphi(Z_i^\nu)$ for some function φ on the state space ([1]). Since φ can be assumed to take on the values 0, 1, there will exist sets $A\subset\bar{G}\times P^{m-1}$ such that $0<\lambda(A)<1$ where $\lambda\in\mathcal{P}(M)$ corresponds to $\nu\in\mathcal{P}(P^{m-1})$, and with $(g',g'x)\in A$ for all $(g,x)\in A$ for μ-almost every g'. Then clearly A depends only on the second component: $A=\bar{G}\times A'$, with $g'A'\subset A'$ for a.e. $g'\in G$. We have $0<\nu(A')<1$ and the restriction of ν to A' is a μ-stationary measure. This shows that ν is not extremal.

LEMMA 3.3. $\beta(\mu)=\sup\{\alpha(\nu)\,|\,\nu\in N,\ \nu$ *extremal*$\}$.

PROOF. This follows from the Krein–Milman theorem which asserts that all of N is spanned by extremals.

Suppose now that for all extremal measures $\nu \in N$ we had $\alpha(\nu) = \beta$ for a fixed β. Then $\beta = \beta(\mu)$ and by Theorem 2.1 we would have with probability one

$$\lim_{N\to\infty} \frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 v \| = \beta(\mu) \tag{3.2}$$

for all $v \in \mathbf{R}^m$. If on the other hand (3.2) is not valid for all v, there must exist an extremal measure ν with $\alpha(\nu) < \beta(\mu)$. Lemma 3.1 is valid for this measure ν and we obtain a set of vectors v for which with probability one

$$\frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 v \| \to \alpha(\nu).$$

For any subspace $L \subset \mathbf{R}^m$, let $\bar{L}$ denote the corresponding set of points in P^{m-1}, i.e., the set of all directions represented in L. Let L be the set of vectors v for which

$$\limsup_{N\to\infty} \frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 v \| \leqq \alpha(\nu) \tag{3.3}$$

holds with probability one. L is clearly a subspace of $\mathbf{R}^m$ and by Lemma 3.1, $\nu(\bar{L}) = 1$. On the other hand since $\alpha(\nu) < \beta(\mu)$, by Theorem 2.2 it is clear that L is a proper subspace of $\mathbf{R}^m$. So if we let L_ν denote the minimal subspace for which $\nu(\bar{L}_\nu) = 1$, then $L_\nu \subset L$, so that L_ν is a proper subspace of $\mathbf{R}^m$ and (3.3) is valid for all $v \in L_\nu$.

LEMMA 3.4. *If ν is a μ-stationary measure and L_ν is the minimal subspace of $\mathbf{R}^m$ for which $\nu(\bar{L}_\nu) = 1$, then $gL_\nu = L_\nu$ for μ-almost every $g \in \mathrm{GL}(m, \mathbf{R})$.*

PROOF. Since ν is μ-stationary

$$\nu(\bar{L}) = \mu * \nu(\bar{L}) = \int \nu(g^{-1}\bar{L}) d\mu(g).$$

If this is 1, then $\nu(g^{-1}\bar{L}) = 1$ for a.e. g. But then $g^{-1}\bar{L} \supset \bar{L}$ and so $gL = L$.

Let us say that a subspace $L \subset \mathbf{R}^m$ is *μ-invariant* if $gL = L$ for μ-almost every $g \in \mathrm{GL}(m, \mathbf{R})$. We now have

THEOREM 3.5. *Assume the measure μ satisfies the condition of Theorem 2.3. Then either for every $v \in \mathbf{R}^m$ and with probability one*

$$\lim \frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 v \| = \beta(\mu), \tag{3.2 bis}$$

or for some proper μ-invariant subspace $L \subset \mathbf{R}^m$, for every $v \in L$,

$$\limsup \frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 v \| \leqq \alpha \tag{3.4}$$

where $\alpha < \beta(\mu)$, again with probability one.

If L is a μ-invariant subspace of $\mathbf{R}^m$ then for μ-almost every g we obtain a transformation $g \mid L$ of $L \to L$. Thus μ determines a measure μ_L on $\mathrm{GL}(L)$. For each μ-invariant subspace $L \subset \mathbf{R}^m$ we then have a rate of growth $\beta(\mu_L)$ which can be characterized as the infimum of α for which (3.4) holds.

We note that $\beta(\mu_{L'+L''}) = \max\{\beta(\mu_{L'}), \beta(\mu_{L''})\}$ if L', L'' are both μ-invariant subspaces. It follows that if there exists any subspace with $\beta(\mu_L) < \beta(\mu)$ then there is a unique maximal such subspace. We denote this subspace by L_1.

The alternative presented in Theorem 3.5 is that either (3.2) holds with probability one for every vector $v \in \mathbf{R}^m$, or L_1 is a non-trivial subspace of $\mathbf{R}^m$.

Once again let L be a μ-invariant subspace of $\mathbf{R}^m$. For μ-almost every g, $gL \subset L$ and g induces a transformation of $\mathbf{R}^m/L \to \mathbf{R}^m/L$. Thus μ determines a measure $\mu_{\mathbf{R}^m/L}$ on $\mathrm{GL}(\mathbf{R}^m/L)$.

LEMMA 3.6. *If L is a μ-invariant subspace of $\mathbf{R}^m$*

$$\beta(\mu) = \max\{\beta(\mu_L), \beta(\mu_{\mathbf{R}^m/L})\}.$$

PROOF. Choose a basis of $\mathbf{R}^m$ whose initial vectors form a basis of L. The matrices g in the support of μ have the form

$$g = \begin{pmatrix} g_{11} & g_{12} \\ 0 & g_{22} \end{pmatrix}$$

where g_{ij} are submatrices, g_{11} corresponding to the restriction of g to L and g_{22} corresponding to the action of g on $\mathbf{R}^m/L$. Forming the product $X_N, X_{N-1} \cdots X_1$ of matrices in this form, and identifying $\beta(\mu)$ with the rate of growth of the random product, we see immediately that

$$\beta(\mu_L) \leqq \beta(\mu), \qquad \beta(\mu_{\mathbf{R}^m/L}) \leqq \beta(\mu). \tag{3.5}$$

Now suppose that both of these inequalities are strict. Consider the product

$$(X_{2N} X_{2N-1} \cdots X_{N+1})(X_N X_{N-1} \cdots X_1) = \begin{pmatrix} A'_N & B'_N \\ 0 & C'_N \end{pmatrix} \begin{pmatrix} A''_N & B''_B \\ 0 & C''_N \end{pmatrix}.$$

Let $\varepsilon > 0$. When N is large, then with probability close to one we will have

$$\|A'_N\|, \|A''_N\| \leqq e^{(1+\varepsilon)N\beta(\mu_L)}, \qquad \|B'_N\|, \|B''_N\| \leqq e^{(1+\varepsilon)N\beta(\mu)},$$

$$\|C'_N\|, \|C''_N\| \leqq e^{(1+\varepsilon)N\beta(\mu_{\mathbf{R}^m/L})}.$$

Hence with probability close to one,

$$\| X_{2N}X_{2N-1}\cdots X_1\| \leqq e^{(1+\varepsilon)2N\beta(\mu_L)} + e^{(1+\varepsilon)2N\beta(\mu_{\mathbf{R}^m/L})}$$
$$+ e^{(1+\varepsilon)N(\beta(\mu)+\beta(\mu_L))} + e^{(1+\varepsilon)N(\beta(\mu)+\beta(\mu_{\mathbf{R}^m/L}))}.$$

But this is also larger than $e^{(1-\varepsilon)2N\beta(\mu)}$ which shows that both inequalities in (3.5) cannot be strict. This proves the lemma.

LEMMA 3.7. *With L_1 defined as above (L_1 may be $\{0\}$), the measure $\mu_{\mathbf{R}^m/L}$ has the property that with probability one*

$$\frac{1}{N}\log\|\bar{X}_N\bar{X}_{N-1}\cdots\bar{X}_1 z\| \to \beta(\mu)$$

for each $z \in \mathbf{R}^m/L_1$, where for $g \in \mathrm{GL}(m,\mathbf{R})$ with $gL_1 = L_1$ we denote by $\bar{g}$ the induced transformation on $\mathbf{R}^m/L_1$.

PROOF. We can write, with a harmless abuse of notation,

$$X_n = \begin{pmatrix} Y_n & Z_n \\ 0 & \bar{X}_n \end{pmatrix},$$

where Y_n denotes the restriction of X_n to L_1. By Lemma 3.6, since $\beta(\mu_{L_1}) < \beta(\mu)$ we must have $\beta(\mu_{\mathbf{R}^m/L_1}) = \beta(\mu)$. Now apply Theorem 3.5 to $\mu_{\mathbf{R}^m/L_1}$. If the assertion of Lemma 3.7 were not true, $\mathbf{R}^m/L_1$ would contain a proper subspace with a rate of growth $< \beta(\mu_{\mathbf{R}^m/L_1}) = \beta(\mu)$. We could then write

$$X_n = \begin{pmatrix} Y_n & Z_n & Z''_n \\ 0 & \bar{X}'_n & \bar{X}''_n \\ 0 & 0 & \bar{X}'''_n \end{pmatrix}$$

where the random products of $\bar{X}'_n$ have smaller rate of growth than $\beta(\mu)$. But also Y_n leads to a smaller rate of growth since $\beta(\mu_{L_1}) < \beta(\mu)$. The same is therefore true of

$$\begin{pmatrix} Y_n & Z'_n \\ 0 & \bar{X}'_n \end{pmatrix}$$

according to Lemma 3.6. But this would contradict the maximality of L_1. This proves Lemma 3.7.

This leads to

PROPOSITION 3.8. *For any $v \in \mathbf{R}^m$ with $v \notin L_1$*

$$\lim \frac{1}{N}\log\|X_N X_{N-1}\cdots X_1 v\|$$

exists with probability one and equals $\beta(\mu)$.

Now L_1 is invariant with respect to the matrices in the support of μ. We can carry out the same analysis for L_1 as we did for $\mathbf{R}^m$ and we find a proper subspace $L_2 \subset L_1$ containing all the μ-invariant subspaces with $\beta(\mu_L) < \beta(\mu_{L_1})$. For any $v \in L_2 \backslash L_1$ we will have

$$\lim \frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 v \| = \beta(\mu_{L_1}).$$

We repeat this procedure till we have exhausted $\mathbf{R}^m$. The final result is the following

THEOREM 3.9. *Let μ be a probability measure on* $\mathrm{GL}(m, \mathbf{R})$ *satisfying* (2.12). *There is a sequence of subspaces*

$$0 \subset L_r \subset L_{r-1} \subset \cdots \subset L_2 \subset L_1 \subset L_0 = \mathbf{R}^m$$

and a sequence of values $\beta(\mu) = \beta^0(\mu) > \beta'(\mu) > \beta''(\mu) > \cdots > \beta^{(r)}(\mu)$ *such that if $v \in L_i \backslash L_{i+1}$, then with probability one*

$$\lim_{N \to \infty} \frac{1}{N} \log \| X_N X_{N-1} \cdots X_1 v \| = \beta^{(i)}(\mu).$$

The subspaces $\{L_i\}$ and the growth rates $\beta^{(i)}(\mu)$ are related to the extremal μ-stationary measures ν in the following way. By Lemmas 3.1 and 3.2 we know that if ν is an extremal μ-stationary measure then

$$\alpha(\nu) = \iint \log \frac{\| g\hat{u} \|}{\| \hat{u} \|} \, d\mu(g) d\nu(u)$$

is the growth rate for some vector v. Therefore by the theorem $\alpha(\nu) = \beta^{(i)}(\mu)$ for some i. Since this growth rate characterizes the vectors of $L_i \backslash L_{i+1}$ we must have $\nu(\bar{L}_i \backslash \bar{L}_{i+1}) = 1$, or $\nu(\bar{L}_i) = 1$, $\nu(\bar{L}_{i+1}) = 0$.

Thus for extremal μ-stationary ν, the functional $\alpha(\nu)$ can take on only the values $\beta^0(\mu), \beta^1(\mu), \cdots, \beta^{(r)}(\mu)$. Consider all the measures ν with $\alpha(\nu) \geqq \beta^{(i)}(\mu)$, and let $\mathcal{L}_i$ be the set of all subspaces L with $\nu(\bar{L}) = 0$ for these ν. On each such subspace the growth rate is $< \beta^{(i)}(\mu)$ and therefore the same is true of the sum. These considerations lead to the following

THEOREM 3.10. *The subspaces $\{L_i\}$ and the $\beta^{(i)}(\mu)$ of the foregoing theorem can be obtained as follows. As ν ranges over all extremal μ-stationary measures, there are finitely many values of $\alpha(\nu)$ that can occur. These values are $\beta(\mu) = \beta^0(\mu) > \beta^1(\mu) > \cdots > \beta^{(r)}(\mu)$. Let $\mathcal{L}_i$ be the set of all μ-invariant subspaces satisfying $\nu(\bar{L}) = 0$ for all ν with $\alpha(\nu) > \beta^{(i)}(\mu)$. L_i is thus the sum of all subspaces in $\mathcal{L}_i$ and it is the unique maximal element of $\mathcal{L}_i$.*

Note that since for extremal measures $\alpha(\nu)$ can take on only finitely many values, the supremum

$$\beta(\mu) = \sup\{\alpha(\nu) \mid \nu \in \mathcal{P}(P^{m-1}), \mu * \nu = \nu\}$$

is attained for some extremal ν. Moreover, any ν for which $\nu(\bar{L}_1) = 0$ will satisfy $\alpha(\nu) = \beta(\mu)$. We have

COROLLARY. *There exist μ-stationary measures on $\mathcal{P}(P^{m-1})$ satisfying $\nu(\bar{L}_1) = 0$. For any such measure*

$$\iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g) d\nu(u) = \beta(\mu).$$

4. Perturbation of random matrix product

We let μ be a probability measure on $\mathrm{GL}(m, \mathbf{R})$ and, as in the preceding section, we say that a subspace $L \subset \mathbf{R}^m$ is μ-invariant if $gL = L$ for almost all $g \in \mathrm{GL}(m, \mathbf{R})$ with respect to μ. We denote by μ_L the induced measure on the group of linear transformations of L. The growth rate $\beta(\mu)$ is defined by (2.11).

PROPOSITION 4.1. *Assume that for every μ-invariant subspace $L \neq \{0\}$ of $\mathbf{R}^m$ we have $\beta(\mu_L) = \beta(\mu)$. Let $\{\mu_k\}$ be a sequence of probability measures on $\mathrm{GL}(m, \mathbf{R})$ satisfying*

$$\int_{\|g\|>T} \log^+ \|g\| d\mu_k(g) + \int_{\|g^{-1}\|>T} \log^+ \|g^{-1}\| d\mu_k(g) \to 0 \tag{4.1}$$

as $T \to \infty$ uniformly in k. If $\mu_k \to \mu$ weakly, i.e., if for all continuous functions $f(g)$ of compact support $\int f d\mu_k \to \int f d\mu$, then $\beta(\mu_k) \to \beta(\mu)$.

PROOF. By the corollary to Theorem 3.10, for each measure μ_k there exists a measure ν_k on P^{m-1} with $\mu_k * \nu_k = \nu_k$ and such that

$$\beta(\mu_k) = \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu_k(g) d\nu_k(u).$$

Passing to a subsequence one can assume, without loss of generality, that the sequence ν_k converges weakly to a measure ν^* on P^{m-1}. It is easy to deduce from the equalities $\mu_k * \nu_k = \nu_k$ that $\mu * \nu^* = \nu^*$.

Now apply Theorem 3.10 to the measure μ. Since by hypothesis $\beta(\mu_L) = \beta(\mu)$ for each μ-invariant subspace $L \subset \mathbf{R}^m$ the filtration $\{L_i\}$ corresponding to μ must be trivial, having only the subspaces $\{0\}$, $\mathbf{R}^m$. Hence for every extremal

μ-stationary measure ν, $\alpha(\nu)=\beta(\mu)$ and so too for all μ-stationary measures. In particular

$$\iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g)d\nu^*(u)=\beta(\mu).$$

On the other hand, if $\mu_k \to \mu$, $\nu_k \to \nu^*$, and (4.1) holds uniformly, then

$$\iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu_k(g)d\nu_k(u) \to \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g)d\nu^*(u).$$

This proves that $\beta(\mu_k)\to\beta(\mu)$ and thus completes the proof of the proposition.

In particular, if the measure μ is supported on a subgroup $G_\mu \subset \mathrm{GL}(m,\mathbf{R})$ which is irreducible, then $\beta(\mu)$ is "stable" in the sense that small perturbations of μ subject to (4.1) lead to small changes in $\beta(\mu)$. We can deduce the same result under the milder restriction that G_μ possess at most one non-trivial invariant subspace. For suppose that L is the unique μ-invariant subspace, and decompose the matrices $g \in G_\mu$ into the form

$$g=\begin{pmatrix} g_{11} & g_{12} \\ 0 & g_{22}\end{pmatrix}$$

with g_{11} the restriction of g to L and g_{22} the induced transformation on $\mathbf{R}^m/L$. By Proposition 4.1, if $\beta(\mu_L)=\beta(\mu)$ then we have stability. On the other hand, if $\beta(\mu_L)<\beta(\mu)$, then by Lemma 3.6, $\beta(\mu_{\mathbf{R}^m/L})=\beta(\mu)$. Now consider the transposed matrices

$${}^t g=\begin{pmatrix} {}^t g_{11} & 0 \\ {}^t g_{12} & {}^t g_{22}\end{pmatrix}$$

and the corresponding measure ${}^t\mu$. There is a unique non-trivial ${}^t\mu$-invariant subspace, $L^\perp$, and the restriction of ${}^t g$ to this subspace is given by ${}^t g_{22}$. Now if we regard $\beta(\mu)$ as a rate of growth of matrix products it is clear that $\beta(\mu)=\beta({}^t\mu)$ and for the same reason $\beta(\mu_{\mathbf{R}^m/L})=\beta({}^t\mu_{L^\perp})$. Since $\beta(\mu_{\mathbf{R}^m/L})=\beta(\mu)$, it follows that for the non-trivial ${}^t\mu$-invariant subspace $L^\perp$,

$$\beta({}^t\mu_{L_\perp})=\beta({}^t\mu).$$

Again we can apply Proposition 4.1 to conclude that we have stability for ${}^t\mu$. But clearly this is equivalent to stability for μ. This proves Theorem B of the Introduction:

THEOREM B. *Let μ be a probability measure on* $\mathrm{GL}(m,\mathbf{R})$ *and let G_μ be the smallest closed subgroup of* $\mathrm{GL}(m,\mathbf{R})$ *which supports the measure μ. If G_μ has the* *property that there exists at most one non-trivial subspace $V \subset \mathbf{R}^m$ for which $gV \subset V$ for $g \in G_\mu$, then whenever $\mu_k \to \mu$ weakly and boundedly, $\beta(\mu_k) \to \beta(\mu)$.*

A variant of Theorem B is the following result.

THEOREM B′. *Let μ be a probability measure on $\mathrm{GL}(m,\mathbf{R})$ and let G_μ be the smallest closed subgroup of $\mathrm{GL}(m,\mathbf{R})$ which supports the measure μ. Let $V \subset \mathbf{R}^m$ be an invariant subspace for G_μ such that V is contained in any other invariant subspace. Assume that $\beta(\mu_V) = \beta(\mu)$. Then whenever $\mu_k \to \mu$ weakly and boundedly, $\beta(\mu_k) \to \beta(\mu)$.*

Clearly, the hypotheses of this theorem imply those of Proposition 4.1.

A relevant example is where G_μ consists of all the matrices

$$g = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ 0 & g_{22} & g_{23} \\ 0 & 0 & g_{33} \end{pmatrix}.$$

Let $\gamma_i = \int \log|g_{ii}|\,d\mu(g)$ for $i = 1,2,3$. We claim that if either γ_1 or γ_3 is the largest of the three numbers γ_1, γ_2, γ_3 then we will have stability in the sense of the foregoing theorems. For if γ_1 is the largest of the three, then the conditions of Theorem B′ are fulfilled for μ. On the other hand, if γ_3 is the largest, the conditions of Theorem B′ are fulfilled for the transposed measure ${}^t\mu$, and since $\beta(\mu) = \beta({}^t\mu)$, we obtain the same result.

Consider now the case $m = 2$. There are three possibilities as regards the invariant subspaces for G_μ. Either there are no non-trivial subspaces, or there is just one, or there are two invariant subspaces. (We omit the trivial case of scalar matrices.) In the first two cases we will have stability of $\beta(\mu)$ by Theorem B. In the third case there need not be stability, as the example of [5] shows. Namely, we can suppose that μ is concentrated in diagonal matrices

$$g = \begin{pmatrix} g_{11} & 0 \\ 0 & g_{22} \end{pmatrix}.$$

Suppose

$$\int \log|g_{11}|\,d\mu(g) \neq \int \log|g_{22}|\,d\mu(g)$$

and let $\mu_\varepsilon = (1-\varepsilon)\mu + \varepsilon\delta_k$, where δ_k is the point measure attached to the element k which we take to be

$$k = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

One verifies that μ_ε has a unique stationary measure on $\mathcal{P}$ and one is led to the formula

$$\beta(\mu_\varepsilon)=\tfrac{1}{2}\left(\int \log|g_{11}|\,d\mu(g)+\int \log|g_{22}|\,d\mu(g)\right).$$

On the other hand, it is clear that

$$\beta(\mu)=\max\left(\int \log|g_{11}|\,d\mu(g),\int \log|g_{22}|\,d\mu(g)\right),$$

so that although $\mu_\varepsilon \to \mu$ as $\varepsilon \to 0$, $\beta(\mu_\varepsilon) \not\to \beta(\mu)$. Finally if

$$\int \log|g_{11}|\,d\mu(g)=\int \log|g_{22}|\,d\mu(g)$$

then the conditions of Proposition 4.1 are satisfied, so that in this case we do, indeed, obtain stability.

For $m=3$ we have a more specialized result. Suppose the matrices of G_μ have the form

$$g=\begin{pmatrix} g_{11} & g_{12} & g_{13} \\ 0 & g_{22} & g_{23} \\ 0 & 0 & g_{33} \end{pmatrix}$$

and that for some $g \in G_\mu$, $g_{12}\neq 0$, $g_{23}\neq 0$. Then if

$$\int \log|g_{11}|\,d\mu(g) \geqq \max\left\{\int \log|g_{22}|\,d\mu(g),\int \log|g_{33}|\,d\mu(g)\right\}$$

we can verify the conditions of Theorem B′. Thus in this case we will have stability.

For further discussion of stability with more restrictive conditions of convergence $\mu_k \to \mu$ we refer the reader to [6].

5. Appendix. Comparison with the Oseledec decomposition

In [8] Oseledec studies the rate of growth of products $X_nX_{n-1}\cdots X_1v$ for a stationary process $\{X_n\}$. This includes the situation we have studied, but the results are not quite the same. Oseledec considers the sequence $\{X_n\}$ fixed and regards the growth rate of $\|X_nX_{n-1}\cdots X_1v\|$ as a function of v. This leads to a "random" decomposition of $\mathbf{R}^m$ exhibiting different growth rates.

To illustrate the situation let us take an example when μ is concentrated on upper-triangular 2×2 matrices

$$X_i=\begin{pmatrix} a_i & c_i \\ 0 & b_i \end{pmatrix}.$$

Let $a = E \log |a_i|$ and $b = E \log |b_i|$. According to [8] a and b will be two characteristic exponents and they have corresponding directions with approximately e^{na} and e^{nb} rates of expanding (contracting).

The matrices X_i have the invariant subspace

$$\Gamma = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \right\}.$$

Clearly,

$$Y_n = X_n \cdots X_1 = \begin{pmatrix} a_n \cdots a_1 & \sum_{k=1}^{n} a_n \cdots a_{k+1} c_k b_{k-1} \cdots b_1 \\ 0 & b_n \cdots b_1 \end{pmatrix}. \tag{5.1}$$

If $a < b$, then all vectors of Γ grow with the speed e^{na}, but any vector $y \notin \Gamma$ grows as e^{nb}, i.e.,

$$\lim \frac{1}{n} \log \| X_n \cdots X_1 y \| \overset{\text{a.s.}}{=} \begin{cases} a & \text{if } y \in \Gamma, \\ b & \text{if } y \notin \Gamma. \end{cases} \tag{5.2}$$

In this case the filtration given by Theorem 3.9 is the same as in [8].

If $a > b$ then one direction in Oseledec's theorem is Γ and it is non-random. This direction corresponds to the growth rate of e^{na}. From (5.1) it is clear that the direction corresponding to the growth rate e^{nb} is determined by the vector

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$$

with

$$\frac{z_1}{z_2} = -\sum_{k=1}^{\infty} \frac{c_k}{d_k} \left(\frac{b_{k-1}}{d_{k-1}} \right) \cdots \left(\frac{b_1}{a_1} \right) \tag{5.3}$$

and it is random. If Γ is the only invariant subspace with respect to all $g \in \operatorname{supp} \mu$ (which is the generic case) then by Theorem B for any $v \in \mathbf{R}^m$ with the probability 1,

$$\lim_{n \to \infty} \frac{1}{n} \log \| X_n \cdots X_1 v \| = a$$

and so the direction (5.3) coincides with a fixed direction with probability 0.

References

1. J. L. Doob, *Stochastic Processes*, Wiley, New York, 1953.

2. H. Furstenberg, *Noncommuting random products*, Trans. Am. Math. Soc. **108** (1963), 377–428.

3. H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Stat. **31** (1960), 457–469.

4. H. Hennion, *Loi des grands nombres et perturbations pour des produit reductibles de matrices aléatoires indépendantes*, preprint, 1983.

5. Y. Kifer, *Perturbations of random matrix products*, Z. Wahrscheinlichkeitstheor. Verw. Geb. **61** (1982), 83–95.

6. Y. Kifer and E. Slud, *Perturbations of random matrix products in a reducible case*, in *Ergodic Theory and Dynamical Systems*, to appear.

7. J. F. C. Kingman, *Subadditive ergodic theory*, Ann. Probab. **1** (198?), 883–909.

8. V. I. Oseledec. *A multiplicative ergodic theorem, Lyapunov characteristic numbers for dynamical systems*, Trans. Mosc. Math. Soc. **19** (1968), 197–221.

9. M. S. Raghunathan, *A proof of Oseledec's multiplicative ergodic theorem*, Isr. J. Math. **32** (1979), 356–362.

INSTITUTE OF MATHEMATICS
THE HEBREW UNIVERSITY OF JERUSALEM
JERUSALEM, ISRAEL