

# STRONG LAWS OF LARGE NUMBERS FOR PRODUCTS OF RANDOM MATRICES

BY  
STEVE PINCUS

**ABSTRACT.** This work, on products of random matrices, is inspired by papers of Furstenberg and Kesten (Ann. Math. Statist. **31** (1960), 457–469) and Furstenberg (Trans. Amer. Math. Soc. **108** (1963), 377–428). In particular, a formula was known for almost sure limits for normalized products of random matrices in terms of a stationary measure. However, no explicit computational techniques were known for these limits, and little was known about the stationary measures.

We prove two main theorems. The first assumes that the random matrices are upper triangular and computes the almost sure limits in question. For the second, we assume the random matrices are  $2 \times 2$  and Bernoulli, i.e., random matrices whose support is two points. Then the second theorem gives an asymptotic result for the almost sure limits, with rates of convergence in some cases.

**1. Introduction and preliminaries.** In 1960, Furstenberg and Kesten proved a Strong Law of Large Numbers for random matrices, with some restrictive positivity conditions on the entries, and without computing the limits explicitly. In 1963, Furstenberg provided a formula for almost sure limits for normalized products of random matrices in terms of a stationary measure. Despite a considerable amount of further work, the invariant (stationary) measures, and hence the limits in the strong law, had not been found explicitly.

We will investigate computational techniques for suitably normalized products of random matrices. We first handle some preliminaries.

*Random matrices.* A random matrix is a random variable from a probability space  $(\Omega, B, P)$  into the space of real (or complex)  $n \times n$  matrices. During much of this work, we will assume that  $n = 2$ .  $A_i$  are always assumed to be i.i.d. random matrices, where  $i$  runs over  $Z_+$ , the positive integers. Let  $T_m := A_m \cdots A_1$ , and let  $\|\cdot\|$  be the usual operator norm ( $\|A\| := \sup_{|\nu|=1} |A\nu|$ ).

Furstenberg and Kesten proved the following useful theorem (1960). Let  $A_i$  be i.i.d. real random matrices; then  $\lim_{m \rightarrow \infty} m^{-1} E(\ln \|T_m\|) = E$  exists ( $E$  is not necessarily finite). Let  $\ln^+ t := \max(\ln t, 0)$ . Assume that  $E(\ln^+ \|A_1\|) < \infty$ . Then we have

**THEOREM 1.** *Almost surely (hereafter abbreviated a.s.)  $\lim_{m \rightarrow \infty} m^{-1} \ln \|T_m\| = E$ .*

---

Received by the editors October 6, 1980 and, in revised form, April 14, 1983.

1980 *Mathematics Subject Classification.* Primary 60F15.

*Key words and phrases.* Random matrices, Bernoulli matrices.

©1985 American Mathematical Society  
0002-9947/85 \$1.00 + \$.25 per page

For the second theorem, we handle some preliminaries. Let  $\mu$  be a measure on  $\text{Sl}_2(\mathbf{R})$ , and let  $G$  be the smallest closed subgroup of  $\text{Sl}_2(\mathbf{R})$  containing the support of  $\mu$ .

DEFINITION 1.  $G$  is *irreducible* if the only subspaces fixed by the matrices of  $G$  are  $\mathbf{R}^2$  and  $\{0\}$ .

DEFINITION 2.  $\mathbf{P}^1$  (projective one-space) = the quotient space of  $\mathbf{R}^2 \setminus \{0\}$  with the equivalence relation  $\nu \sim k\nu$  for any  $\nu \in \mathbf{R}^2 \setminus \{0\}$  and  $k \in \mathbf{R} \setminus \{0\}$ .

For  $u \in \mathbf{R}^2 \setminus \{0\}$  and  $g \in \text{Sl}_2(\mathbf{R})$ , let  $\sigma(g, u) := \|gu\|/\|u\|$ . Note that  $\sigma(g, u)$  depends only on  $g$  and on  $\bar{u} \in \mathbf{P}^1$ . Therefore, on  $\mathbf{P}^1$  define  $\sigma_1$  by  $\sigma_1(g, \bar{u}) = \sigma(g, u)$ . Set  $\rho = \ln \sigma$ ,  $\rho_1 = \ln \sigma_1$ .

Given a measure  $\mu$  on  $\text{Sl}_2(\mathbf{R})$ , we say that  $\nu$ , a probability measure on  $\mathbf{P}^1$ , is a *stationary measure* for  $\mu$  if for all continuous functions  $f$  from  $\mathbf{P}^1 \rightarrow \mathbf{R}$ ,  $\int f(x) d\nu(x) = \int \int f(gx) d\mu(g) d\nu(x)$ . Furstenberg (1963) proves that for  $\mu$  a probability measure on  $\text{Sl}_2(\mathbf{R})$ , there exists at least one stationary measure for  $\mu$  on  $\mathbf{P}^1$ . In the same paper he proves

THEOREM 2. *If  $G$  is irreducible, and*

$$\int \ln \|g\| d\mu(g) < \infty,$$

*$\int \int \rho_1(g, t) d\mu(g) d\nu(t)$  for a stationary measure  $\nu$  for  $\mu$  is independent of the stationary measure  $\nu$ . Denoting the common value by  $a_\mu(\rho_1)$ , we have (a.s.)*

$$\lim_{m \rightarrow \infty} m^{-1} \ln \|T_m u\| = a_\mu(\rho_1)$$

*for all nonzero vectors  $u \in \mathbf{R}^2$ , where  $A_i$  have the distribution  $\mu$ .*

In practice, this theorem is hard to use for exact computations, as stationary measures are hard to compute. We will use this theorem later to get some lower bounds.

We will need Theorem 2 in a slightly different form. Again we refer to the common value of  $\int \int \rho_1(g, t) d\mu(g) d\nu(t)$  by  $a_\mu(\rho_1)$ . It is easy to see that Theorem 2 implies the following, which we shall henceforth refer to as Theorem 2: *Given  $T_m := A_m \cdots A_1$ , if  $G$  is irreducible and  $\int \ln \|g\| d\mu(g) < \infty$ , then (a.s.)*

$$\lim_{m \rightarrow \infty} m^{-1} \ln \|T_m\| = a_\mu(\rho_1).$$

As we noted earlier, given  $A_i$  i.i.d. computations of (a.s.) limits of  $m^{-1} \ln \|T_m\|$  have been elusive. Nonetheless, there has been a good deal of work done in this area during the past twenty years. Representative are four papers of Tutubalin (1965, 1966, 1967, 1969), where he proves general results for normalized products of matrices, assuming that the distribution of  $A_i$  is absolutely continuous with respect to Haar measure on  $\text{Sl}_n(\mathbf{R})$ . Unfortunately, these papers do not shed light on techniques for explicit computations.

The first theorem we prove gives the (a.s.)  $\lim_{m \rightarrow \infty} m^{-1} \ln \|T_m\|$  whenever  $A_i$  are  $n \times n$ , (a.s.) upper triangular matrices. It was difficult, however, to come up with any other explicit computations, and so we turned our attention to an apparently simple case, the Bernoulli random  $2 \times 2$  matrix case; namely, we assumed that with probability (w.p.)  $p > 0$ ,  $A_i = A$ , and w.p.  $q := 1 - p > 0$ ,  $A_i = B$ , where  $A$  and  $B$  were arbitrary  $2 \times 2$  matrices. Exact calculations were quite difficult in this case

too. We were able to get a general asymptotic theorem in this setting, which we prove in §3. This proof is quite long and is broken up into more easily readable subsections.

We end this section with three definitions and a general theorem on matrices.

DEFINITION 3. Let  $[a_{jk}]$  be a complex  $n \times n$  matrix  $A$ . Define  $|A| := [|a_{jk}|]$ .

DEFINITION 4. Let  $(A_i)_{jk}$  denote the  $jk$ th entry of  $A_i$  (and similarly for any other matrix indexed by the positive integers).

DEFINITION 5. Let  $a, b \in \mathbb{R}$ . Define  $a \vee b := \max(a, b)$ , and  $a \wedge b := \min(a, b)$ .

The following general theorem can be found in Shilov (1971, p. 160).

THEOREM 3 (REAL CANONICAL FORM THEOREM FOR  $2 \times 2$  MATRICES). *Let  $B$  be a real  $2 \times 2$  matrix. Then there exists a real valued invertible matrix  $Q$ , and a matrix  $A$ , with  $B = QAQ^{-1}$ , and with  $A$  one of the following forms:*

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix},$$

where  $\lambda$  and  $\mu$  are arbitrary reals satisfying  $|\lambda| > |\mu|$ ;

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

for  $\lambda$  arbitrary real; or

$$\begin{pmatrix} \lambda \cos \phi & \lambda \sin \phi \\ -\lambda \sin \phi & \lambda \cos \phi \end{pmatrix},$$

for  $\lambda$  arbitrary nonzero, and  $\phi$  arbitrary in  $[0, \pi)$ . Furthermore, given  $B$ ,  $A$  is uniquely determined.

**2. An upper triangular theorem.** In this section we prove a theorem about  $n \times n$  upper triangular matrices.

THEOREM 4. *Let  $A_i$  be i.i.d.  $\mathbb{C}$ -valued  $n \times n$  random matrices, with  $A_i = [(A_i)_{jk}]$  and with  $(A_i)_{jk} = 0$  for  $j > k$ . Also, assume  $E(0 \vee \ln \|A_1\|) < \infty$ . Then (a.s.)*

$$\lim_{m \rightarrow \infty} m^{-1} \ln \|T_m\| = \alpha, \quad \text{where } \alpha := \max_{1 \leq i \leq n} E(\ln |(A_1)_{ii}|).$$

We will prove Theorem 4 in two cases. Case I is for  $\alpha \neq -\infty$ , while Case II is for  $\alpha = -\infty$ .

PROOF OF CASE I. Write  $A_i = D_i + S_i$ , where  $D_i$  is the diagonal part of  $A_i$ . Since  $\|A_m A_{m-1} \cdots A_1\| = \|D_m D_{m-1} \cdots D_1 + \tilde{S}\|$ , where  $\tilde{S}$  is a matrix whose elements on and below the diagonal are zero, it follows that  $\|A_m \cdots A_1\| \geq \|D_m \cdots D_1\|$ . Let the (a.s.)  $\lim_{m \rightarrow \infty} m^{-1} \ln \|T_m\| = \beta$ .

To see that  $\beta \geq \alpha$ , we need only see that a.s.  $\lim_{m \rightarrow \infty} m^{-1} \ln \|Q_m\| \geq \alpha$ , where  $Q_m = D_m \cdots D_1$ . But  $(Q_m)_{ii} = \prod_{j=1}^m (D_j)_{ii}$ , and hence

$$m^{-1} \ln \|Q_m\| \geq m^{-1} \sum_{j=1}^m (\ln |D_j|)_{ii} \quad \text{for any } i.$$

By the Strong Law of Large Numbers,

$$\lim_{m \rightarrow \infty} m^{-1} \sum_{j=1}^m (\ln |D_j|)_{ii} = E(\ln |D_{ii}|),$$

and hence  $\beta \geq \alpha$ .

So it remains to see that  $\beta \leq \alpha$ . We proceed as follows. Define  $X_i = \|S_i\|/\|D_i\|$ , and  $Y_i = 1 + X_i$ . Since  $E(\ln \|D_i\|) > -\infty$ , and  $E \ln \|A_i\| < \infty$ , it follows that  $EY_i < \infty$ . Now let

$$R_1 = \sum_{i=1}^m X_i, \quad R_2 = \sum_{1 \leq i < j \leq m} X_i X_j, \dots, \quad R_n = \sum_{1 \leq i_1 < i_2 < \dots < i_n \leq m} X_{i_n}.$$

Then a straightforward estimate shows that

$$\|A_m \cdots A_1\| \leq \|D_m \cdots D_1\| (1 + R_1 + R_2 + \dots + R_n).$$

Next, define  $I_m = \{(i_1, i_2, \dots, i_n): 1 \leq i_1 < i_2 < \dots < i_n \leq m\}$ ; recalling that  $Y_i = 1 + X_i$ , it follows easily that

$$1 + R_1 + R_2 + \dots + R_n \leq \sum_{I_m} Y_{i_1} Y_{i_2} \cdots Y_{i_n}.$$

So to see that  $\beta \leq \alpha$ , it remains to prove that

$$(*) \quad \overline{\lim}_{m \rightarrow \infty} m^{-1} \ln \left( \sum_{I_m} Y_{i_1} Y_{i_2} \cdots Y_{i_n} \right) = 0 \quad \text{a.s.}$$

Pick  $\varepsilon > 0$  arbitrary. Since  $E(\ln Y_i) < \infty$ , Borel-Cantelli allows us to deduce that  $\Pr(Y_i > e^{\varepsilon i} \text{ i.o.}) = 0$ . Define  $N = \max\{i: Y_i > e^{\varepsilon i}\}$ . We thus have that  $N$  is finite with probability 1. Let  $M = \max\{Y_i: 1 \leq i \leq n\}$ . Then

$$\begin{aligned} \sum_{I_m} Y_{i_1} Y_{i_2} \cdots Y_{i_n} &\leq M^N \sum_{I_m} \exp(\varepsilon \cdot (i_1 + i_2 + \dots + i_n)) \\ &\leq M^N \binom{m}{n} \exp(\varepsilon mn) \leq M^N m^n \exp(\varepsilon mn) \quad \text{a.s.} \end{aligned}$$

Hence

$$\lim_{m \rightarrow \infty} m^{-1} \ln \left( \sum_{I_m} Y_{i_1} Y_{i_2} \cdots Y_{i_n} \right) \leq \varepsilon n \quad \text{a.s.},$$

from which  $(*)$  follows.  $\therefore \beta = \alpha$ , and Case I is proved.

PROOF OF CASE II. Assume that  $E(\ln |(A_1)_{ii}|) = -\infty$  for all  $i = 1, \dots, n$ . To see that  $\lim_{m \rightarrow \infty} m^{-1} \ln \|T_m\| = -\infty$  (a.s.), it suffices to see that (a.s.)

$$\lim_{m \rightarrow \infty} m^{-1} \ln \|T_m\| < N, \quad \text{for } N \text{ arbitrary.}$$

Pick  $N \in \mathbb{R}$ . For each  $j \in \{1, \dots, n\}$ , there exists  $s_j$  such that  $E(s_j \vee \ln |(A_1)_{jj}|) \leq N - 1$ . Define

$$t_j := E(s_j \vee \ln |(A_1)_{jj}|);$$

$t_j$  is finite. We now define  $B_i(\omega)$ . For  $j > k$ ,

$$(B_i)_{jk}(\omega) := 0.$$

For  $j = k$ ,

$$(B_i)_{jj}(\omega) := \exp(s_j) \vee |(A_i)_{jj}(\omega)| \quad (\text{for } j = 1, \dots, n).$$

For  $j < k$ ,

$$(B_i)_{jk}(\omega) := |(A_i)_{jk}(\omega)|.$$

The  $B_i$  are i.i.d. upper triangular random matrices, with  $E(0 \vee \ln \|B_1\|) < \infty$ . Now define  $S_m := B_m \cdots B_1$ .

By Theorem 4, Case I, (a.s.)

$$\begin{aligned} \lim_{m \rightarrow \infty} m^{-1} \ln \|S_m\| &= \max_{1 \leq j \leq n} E(\ln |(B_1)_{jj}|) \\ &= \max_{1 \leq j \leq n} E(s_j \vee \ln |(A_1)_{jj}|) = \max_{1 \leq j \leq n} t_j \leq N - 1. \end{aligned}$$

It is easy to show that  $\|B_m(\omega) \cdots B_1(\omega)\| \geq \|A_m(\omega) \cdots A_1(\omega)\|$ . Hence (a.s.)

$$\lim_{m \rightarrow \infty} m^{-1} \ln \|T_m\| \leq \lim_{m \rightarrow \infty} m^{-1} \ln \|S_m\| \leq N - 1,$$

which completes the proof.

### 3. An asymptotic theorem for Bernoulli random matrices.

3.1. *Statement of the main theorem.* Let  $G(d)$  be the set of all  $d \times d$  matrices, let  $P(d)$  be the set of all probability measures on  $G(d)$ , and let  $L(d)$  be the subset of  $P(d)$  defined by

$$L(d) = \left\{ \mu \in P(d): \int_{G(d)} \ln^+ \|\omega\| \mu(d\omega) < \infty \right\}.$$

Now suppose  $\mu \in L(d)$ , let  $M_1, M_2, \dots$  be a sequence of i.i.d. random matrices all with distribution  $\mu$ , and define  $T_n$  by  $T_n = M_n M_{n-1} \cdots M_1$ . We know (by a theorem of Furstenberg and Kesten) that  $\lim_{n \rightarrow \infty} n^{-1} E(\ln \|T_n\|)$  exists. Denote the limit by  $R(\mu)$ . From another of Furstenberg's and Kesten's theorems (which is stated as Theorem 1 in this paper),  $\lim_{n \rightarrow \infty} n^{-1} \ln \|T_n\| = R(\mu)$  a.s. Two natural questions to raise in this context follow. Suppose  $\{\mu_n\}_1^\infty$  is a sequence of probability measures in  $L(d)$ ,  $\mu$  is another probability measure in  $L(d)$ , and  $\mu_n \rightarrow \mu$  (in some topology).

(1) Does it follow that  $R(\mu_n) \rightarrow R(\mu)$ ?

(2) If  $R(\mu_n) \rightarrow R(\mu)$ , what can be said about the rate of convergence?

Here we consider this problem under a restrictive set-up. Let  $B$  be an arbitrary, real valued, nonsingular  $2 \times 2$  matrix. Set

$$A_0 := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

and for  $\varepsilon > 0$ , define

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

Let  $0 < p < 1$ , and write  $q = 1 - p$ . We can now define a family  $\{\mu_\varepsilon\}_{\varepsilon \geq 0}$  of probability measures in  $L(2)$  simply by  $\mu_\varepsilon(B) = q$ ,  $\mu_\varepsilon(A_\varepsilon) = p$ . Since  $\mu_\varepsilon \in L(2)$ ,  $R(\mu_\varepsilon)$  exists for all  $\varepsilon \geq 0$ . To emphasize the dependence of  $B$  and  $\varepsilon$  let us

write  $R(\mu_\varepsilon) = R(\varepsilon, B)$ . We now state

*Assumption 1.*  $B$  has either real eigenvalues or complex eigenvalues with arguments which are rational numbers modulo  $\pi$ .

Then the main theorem we wish to prove is

THEOREM 5. (a)

$$R(0, B) = pq \sum_{n=1}^{\infty} pq^{n-1} \ln \|(B^n)_{11}\|.$$

( $(B^n)_{ij}$  means the  $ij$ th entry of  $B^n$ ).

(b)  $\lim_{\varepsilon \rightarrow 0} R(\varepsilon, B) = R(0, B)$ .

(c) Suppose  $R(0, B)$  is finite and Assumption 1 is satisfied. Then there exists a constant  $C$  such that  $|R(\varepsilon, B) - R(0, B)| \leq C\varepsilon$ , for all  $\varepsilon$  sufficiently small.

3.2. *Matrix lemmas.* In this section we prove two matrix lemmas that will be used in the proof of Theorem 5.

LEMMA 1. Let  $B$  be a  $2 \times 2$  real nonsingular matrix. Let  $\sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|$  be finite, where  $B$  is neither upper nor lower triangular, and let Assumption 1 hold. Then there exists  $C(B)$ , a positive constant just depending on  $B$ , such that for all  $n$ , the absolute value of all entries of  $B^n / (B^n)_{11}$  are  $\leq C(B)$ .

PROOF. The first case is where  $B = QCQ^{-1}$  for  $C = \text{diag}(\lambda, \mu)$ , with  $|\lambda| \geq |\mu|$ , and where we assume  $\det Q = 1$ . Write  $Q = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ .

Since  $B$  is not diagonal,  $\lambda \neq \mu$ . For  $\lambda = -\mu$ ,

$$\begin{aligned} |(B^n)_{21} / (B^n)_{11}| &\leq |2cd / (ad - bc(-1)^n)|, \\ |(B^n)_{12} / (B^n)_{11}| &\leq |2ab / (ad - bc(-1)^n)|, \quad \text{and} \\ |(B^n)_{22} / (B^n)_{11}| &= |(ad(-1)^n - bc) / (ad - bc(-1)^n)| = 1, \end{aligned}$$

provided that  $ad + bc \neq 0$ . Since  $ad + bc = (B)_{11} \neq 0$ , and  $|ad - bc| = 1$ , for all  $n$ , there exists the desired  $C(B)$ .

For  $|\lambda| > |\mu|$ , there exists  $N_0$  such that for all  $n > N_0$ ,  $|(B^n)_{11}| \geq |ad\lambda^n|/2$ . For all  $n$ ,

$$\begin{aligned} |(B^n)_{12}| &\leq |ab|(|\lambda^n| + |\mu^n|) \leq |2ab\lambda^n|, \\ |(B^n)_{21}| &\leq |cd|(|\lambda^n| + |\mu^n|) \leq |2cd\lambda^n|, \quad \text{and} \\ |(B^n)_{22}| &\leq |ad\mu^n| + |bc\lambda^n| \leq (|ad| + |bc|)|\lambda^n|. \end{aligned}$$

These two estimates allow us to find the desired  $C(B)$  in this case.

The cases where  $\det Q = -1$  and/or where  $C = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$  are proved similarly.

Define

$$\text{rot}(\lambda, \phi) := \begin{pmatrix} \lambda \cos \phi & \lambda \sin \phi \\ -\lambda \sin \phi & \lambda \cos \phi \end{pmatrix}.$$

The other case is that  $B = QCQ^{-1}$  for  $C = \text{rot}(\lambda, \phi)$ , where  $\lambda \neq 0$  and  $\phi/\pi$  is rational. Since  $(B^n)_{12}/(B^n)_{11}$ ,  $(B^n)_{21}/(B^n)_{11}$  and  $(B^n)_{22}/(B^n)_{11}$  are independent of  $\lambda$ , assume that  $\lambda = 1$ . Since  $(\text{rot}(1, \phi))^{2b} = I$ , we have that  $B^n = B^{n+2b}$  for all  $n$ . Lemma 1 follows at once, since we choose  $C(B)$  to maximize a finite set.

DEFINITION 6. For  $\theta \in (0, \pi)$ , let

$$e(\theta) := \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

LEMMA 2. *Suppose  $B$  has complex eigenvalues. Then there exists  $k(B)$ , a positive constant depending on  $B$ , such that for all  $n$  and all  $0 \leq \theta < \pi$ ,*

$$|(B^n/(B^n)_{11})e(\theta)| \geq k(B).$$

PROOF. By Theorem 3 we can assume that  $B = QAQ^{-1}$ , with  $A = \text{rot}(k, \phi)$ . Since  $|(B^n/(B^n)_{11})e(\theta)|$  is independent of  $k$ , we assume that  $k = 1$ .

Since  $B = Q \text{rot}(1, \phi) Q^{-1}$ , for all  $n$ ,  $|(B^n)_{11}| \leq \|Q\| \|Q^{-1}\|$ ; let  $C(B) := \|Q\| \|Q^{-1}\|$ .

Since  $Q$  and  $Q^{-1}$  both act on  $S^1$  (which is  $\{(\frac{\cos \theta}{\sin \theta}) : \theta \in (0, 2\pi)\}$ ), and since for any  $\nu \in S^1$ ,  $Q\nu \neq 0$  and  $Q^{-1}\nu \neq 0$ , there exist  $K(Q)$  and  $K(Q^{-1})$  (positive constants, both depending only on  $B$ ) such that for all  $\nu \in S^1$ ,  $|Q\nu| \geq K(Q)$  and  $|Q^{-1}\nu| \geq K(Q^{-1})$ .

Let  $D(B) := K(Q)K(Q^{-1})$ . Then

$$\begin{aligned} |B^n e(\theta)| &= |Q \text{rot}(1, n\phi) Q^{-1} e(\theta)| \\ &= (|Q \text{rot}(1, n\phi) Q^{-1} e(\theta)| / |\text{rot}(1, n\phi) Q^{-1} e(\theta)|) |\text{rot}(1, n\phi) Q^{-1} e(\theta)| \\ &\geq K(Q) |\text{rot}(1, n\phi) Q^{-1} e(\theta)| = K(Q) |Q^{-1} e(\theta)| \\ &\geq K(Q) K(Q^{-1}) = D(B). \end{aligned}$$

Define  $K(B) := D(B)/C(B)$ . Then for all  $n$  and  $0 \leq \theta < \pi$ ,  $|(B^n/(B^n)_{11})e(\theta)| \geq K(B)$ .

3.3. *A new process.* Let  $A_i: (\Omega, B, \mu) \rightarrow 2 \times 2$  matrices be i.i.d.,  $A_i := A$  w.p.  $p > 0 := B$  w.p.  $q := 1 - p > 0$  (assume  $A \neq B$ ). Consider  $V = \{\omega: A_1(\omega) = B\}$ , with  $\mu(V) = q$ .

Define  $\mu'$  by setting  $\mu'(B_0) = q^{-1}\mu(B_0)$  for  $B_0 \in B$ .

We now will define  $B_i(\omega)$  such that  $B_{2n}(\omega) = A^{f(2n, \omega)}$  and  $B_{2n+1}(\omega) = B^{f(2n+1, \omega)}$ . Let  $f(1, \omega) = n - 1$ , where  $n$  is the least integer such that  $A_n(\omega) = A$ . Let  $n_2$  be the least  $t > n$  such that  $A_t(\omega) = B$ . Then let  $f(2, \omega) = n_2 - 1 - f(1, \omega)$ . Recursively, we have  $B_{2n+1}(\omega) := B^{f(2n+1, \omega)}$  as follows: let  $m_{2n+1}$  be the least  $m > \sum_{i=1}^{2n} f(i, \omega)$  such that  $A_m(\omega) = A$ .

Then let

$$f(2n+1, \omega) := m_{2n+1} - 1 - \sum_{i=1}^{2n} f(i, \omega).$$

Define  $B_{2n}(\omega) := A^{f(2n, \omega)}$  as follows: let  $m_{2n}$  be the least  $m > \sum_{i=1}^{2n-1} f(i, \omega)$  such that  $A_m(\omega) = B$ . Then

$$f(2n, \omega) := m_{2n} - 1 - \sum_{i=1}^{2n-1} f(i, \omega).$$

Finally, let  $C_i := B_{2i} B_{2i-1}$ .

Since  $\mu(\{\omega: f(i, \omega) = \infty \text{ for some } i \in \mathbb{Z}_+\}) = 0$ , we can assume that  $f(i, \omega) < \infty$  for all  $i \in \mathbb{Z}_+$ .

Intuitively,  $B_{2i-1}(\omega) = B^{f(2i-1, \omega)}$ , where  $f(2i-1, \omega)$  is the length of the  $i$ th run of  $B$ 's, and  $B_{2i}(\omega) = A^{f(2i, \omega)}$ , where  $f(2i, \omega)$  is the length of the  $i$ th run of  $A$ 's.

Using stopping-time theory (Chung, 1974, p. 261), we see that the  $f(j, \cdot)$  are independent, and the  $\{f(2j, \cdot)\}$ ,  $j \in \mathbb{Z}_+$  are i.i.d., as are the  $\{f(2j-1, \cdot)\}$ ,  $j \in \mathbb{Z}_+$ .

Thus, the  $C_i$  are i.i.d. random matrices. Define  $T_n := A_n \cdots A_1$ ,  $S_n := C_n \cdots C_1$ ,  $m(i) := f(2i - 1)$ , and  $n(i) := f(2i)$ ,

$$\lim_{n \rightarrow \infty} n^{-1} \ln \|T_n\| := R(A, B, p) \quad (\text{a.s.}),$$

and

$$\lim_{n \rightarrow \infty} n^{-1} \ln \|S_n\| := S(A, B, p) \quad (\text{a.s.}).$$

It is straightforward to show (and left to the reader) that  $E(\ln^+ \|A_1\|) < \infty$  and  $E(\ln^+ \|C_1\|) < \infty$ . Thus by Theorem 1, these last two limits exist a.s.

Since  $E(m(1)) = \sum_{n=1}^{\infty} npq^{n-1} = 1/p$  and  $E(n(1)) = \sum_{n=1}^{\infty} nqp^{n-1} = 1/q$ ,

$$E(m(i) + n(i)) = 1/p + 1/q = 1/pq.$$

We now have

**THEOREM 6.**  $R(A, B, p) = pqS(A, B, p)$ .

In practice, computing  $S(A, B, p)$  is often more tractable than computing  $R(A, B, p)$ .

**PROOF.** Let

$$\begin{aligned} \Gamma = \left\{ \omega: \lim_{n \rightarrow \infty} n^{-1} \ln \|T_n(\omega)\| = R(A, B, p), \right. \\ \left. \lim_{n \rightarrow \infty} n^{-1} \ln \|S_n(\omega)\| = S(A, B, p), \right. \\ \left. \text{and } \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n m(i, \omega) + n(i, \omega) = \frac{1}{pq} \right\}. \end{aligned}$$

Pick  $\omega \in \Gamma$ . Define

$$\begin{aligned} g(n, \omega) &:= \sum_{i=1}^n m(i, \omega) + n(i, \omega), \\ R_n &:= g(n, \omega)/n, \quad \text{and} \\ Q_n &:= (g(n, \omega))^{-1} \ln \|A_{g(n, \omega)}(\omega) \cdots A_1(\omega)\|. \end{aligned}$$

Then

$$\begin{aligned} S(A, B, p) &= \lim_{n \rightarrow \infty} n^{-1} \ln \|S_n(\omega)\| \\ &= \lim_{n \rightarrow \infty} n^{-1} \ln \|A_{g(n, \omega)}(\omega) \cdots A_1(\omega)\| = \lim_{n \rightarrow \infty} R_n Q_n. \end{aligned}$$

But since  $\omega \in \Gamma$ ,  $\lim_{n \rightarrow \infty} Q_n$  exists, as does  $\lim_{n \rightarrow \infty} R_n$ . Hence

$$S(A, B, p) = \left( \lim_{n \rightarrow \infty} R_n \right) \left( \lim_{n \rightarrow \infty} Q_n \right) = R(A, B, p)/pq,$$

as desired.

Since

$$n^{-1} \ln \|T_n\| \leq n^{-1} \ln \prod_{i=1}^n \|A_i\| = n^{-1} \sum_{i=1}^n \ln \|A_i\|,$$

$$R(A, B, p) \leq E(\ln \|A_1\|) < \infty.$$

Theorem 6 is not so surprising if one compares the current set-up with the theory of runs (Feller, Vol. 1, pp. 322–328).



Let  $A_i$  be i.i.d.  $2 \times 2$  real random matrices,  $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$  w.p.  $p$ ,  $A_1 = B$  (arbitrary) w.p.  $q := 1 - p$ , where  $0 < p < 1$ . Then

$$R(0, B) = pq \sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|.$$

Let

$$\text{diag}(a, b) := \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

PROOF. Form the new process  $C_i$  as in this section. It suffices to see that

$$S(A, B, p) = \sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|,$$

where  $A = \text{diag}(1, 0)$ .  $C_i = A^{n(i)} B^{m(i)}$ ; since  $A^{n(i)} = \text{diag}(1, 0)$ ,  $C_i$  is upper triangular.

Thus, by Theorem 4

$$S(A, B, p) = E(\ln |(C_1)_{11}|) = \int \ln |(C_1)_{11}| d\mu'.$$

Let  $D_n := \{\omega: m(1, \omega) = n\}$ . Then

$$\int \ln |(C_1)_{11}| d\mu' = \sum_{n=1}^{\infty} \int_{D_n} \ln |(C_1)_{11}| d\mu'.$$

Hence

$$S(A, B, p) = \sum_{n=1}^{\infty} \mu'(D_n) \ln |(B^n)_{11}| = \sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|,$$

and so part (a) is proved.

3.4. *Theorem 5 (upper triangular case).* Theorem 5 is very easy for  $B$  upper triangular. Let  $B = [b_{jk}]$ , with  $b_{21} = 0$ . By Theorem 4,

$$R(\varepsilon, B) = (q \ln |b_{11}|) \vee (p \ln \varepsilon + q \ln |b_{22}|).$$

Thus, there exists  $\varepsilon_0 < 1$  such that for all  $\varepsilon < \varepsilon_0$ ,  $R(\varepsilon, B) = q \ln |b_{11}|$ . Since

$$\begin{aligned} pq \sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}| &= pq \sum_{n=1}^{\infty} pq^{n-1} \ln |b_{11}^n| \\ &= p^2 \sum_{n=1}^{\infty} nq^n \ln |b_{11}| = q \ln |b_{11}|, \end{aligned}$$

Theorem 5(b) is proved. Furthermore, since  $|R(\varepsilon, B) - R(0, B)| = 0$  for  $\varepsilon < \varepsilon_0$ . Theorem 5(c) follows immediately.

An analogous result can be deduced for lower triangular matrices.

3.5. *The finite sum case, with Assumption 1 (upper estimate).* In §§3.5 and 3.6 we consider  $B$  and  $p$  for which  $\sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|$  is finite,  $B$  is neither upper nor lower triangular, and for which Assumption 1 holds.

In this section we find  $C_1$  such that

$$R(\varepsilon, B) \leq pq \sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}| + \varepsilon C_1$$

for all  $\varepsilon$  sufficiently small. Start with the i.i.d. processes  $A_{\varepsilon,i} = A_\varepsilon$  w.p.  $p$ ,  $A_{\varepsilon,i} = B$  otherwise. Let  $T_{\varepsilon,n} = A_{\varepsilon,n} \cdots A_{\varepsilon,1}$ , and

$$R(\varepsilon, B) = (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|T_{\varepsilon,n}\|.$$

Given  $A_{\varepsilon,i}$ , form  $C_{\varepsilon,i}$  as described in §3.3. Let  $S_{\varepsilon,n} := C_{\varepsilon,n} \cdots C_{\varepsilon,1}$ , and let  $S(\varepsilon, B) := (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|S_{\varepsilon,n}\|$ . By Theorem 6,  $r(\varepsilon, B) = pqS(\varepsilon, B)$ . It thus suffices to find  $C_1 > 0$  such that

$$S(\varepsilon, B) \leq \sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}| + \varepsilon C_1$$

for all  $\varepsilon$  sufficiently small.

Define  $\chi_{\varepsilon,i} := (C_{\varepsilon,i})_{11}$ , and  $F_{\varepsilon,i} := C_{\varepsilon,i}/\chi_{\varepsilon,i}$ .  $(C_{\varepsilon,i})_{11} = B_{11}^{m(i)}$ , which is nonzero since  $\sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|$  is finite.

$$\begin{aligned} n^{-1} \ln \|S_{\varepsilon,n}\| &= n^{-1} \ln \left( \left\| \prod_{i=1}^n \chi_{\varepsilon,i} \right\| \right) \|F_{\varepsilon,n} \cdots F_{\varepsilon,1}\| \\ &= n^{-1} \left( \sum_{i=1}^n \ln |\chi_{\varepsilon,i}| + \ln \|F_{\varepsilon,n} \cdots F_{\varepsilon,1}\| \right). \end{aligned}$$

By the Strong Law of Large Numbers,

$$\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n |\chi_{\varepsilon,i}| = E(\ln |\chi_{\varepsilon,1}|) \text{ (a.s.)}.$$

LEMMA 3.

$$E(\ln |\chi_{\varepsilon,1}|) = \sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|.$$

The proof is left to the reader, as it is very similar to the proof of part (a) of Theorem 5.

Define  $H_{\varepsilon,n} := F_{\varepsilon,n} \cdots F_{\varepsilon,1}$ , and  $h(\varepsilon, B) := (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|H_{\varepsilon,n}\|$ . Since

$$n^{-1} \ln \|S_{\varepsilon,n}\| = n^{-1} \left( \sum_{i=1}^{\infty} \ln |\chi_{\varepsilon,i}| + \ln \|F_{\varepsilon,n} \cdots F_{\varepsilon,1}\| \right),$$

$$S(\varepsilon, B) = E(\ln |\chi_{\varepsilon,1}|) + h(\varepsilon, B).$$

So to complete our upper estimate, it suffices to find  $C_2 > 0$  with  $h(\varepsilon, B) \leq \varepsilon C_2$  for all  $\varepsilon$  sufficiently small.

Since  $C_{\varepsilon,i} = \text{diag}(1, \varepsilon^{n(i)})B^{m(i)}$ ,

$$F_{\varepsilon,i} = \text{diag}(1, \varepsilon^{n(i)})B^{m(i)}/(B^{m(i)})_{11}.$$

In this set-up, the conclusion of Lemma 1 holds.

We now have two lemmas whose proofs are straightforward and left to the reader.

LEMMA 4. *Given any  $m$  complex  $n \times n$  matrices  $A_1, \dots, A_m$ , then  $\|A_m \cdots A_1\| \leq \| |A_m| \cdots |A_1| \|$ .*

LEMMA 5. *Let  $A_i$  and  $D_i$  be complex  $n \times n$  matrices for all  $i$ , where  $A_i = |A_i|$  and  $D_i = |D_i|$  for all  $i$ , and where  $(D_i)_{jk} \geq (A_i)_{jk}$  for all  $i, j, k$ . Then  $\|A_n \cdots A_1\| \leq \|D_n \cdots D_1\|$ .*

By Lemmas 1, 4 and 5, we deduce that for all  $n$ ,

$$\|H_{\varepsilon, n}\| \leq \left\| \left( \text{diag}(1, \varepsilon) \begin{pmatrix} 1 & C(B) \\ C(B) & C(B) \end{pmatrix} \right)^n \right\|.$$

Let

$$E(B) := \begin{pmatrix} 1 & C(B) \\ C(B) & C(B) \end{pmatrix},$$

and define  $U(\varepsilon) := \text{diag}(1, \varepsilon)E(B)$ . We see that

$$h(\varepsilon, B) \leq \lim_{n \rightarrow \infty} n^{-1} \ln \|(U(\varepsilon))^n\|.$$

But since  $U(\varepsilon)$  is a constant matrix, by the spectral radius theorem (Loomis, 1953, p. 75),

$$\lim_{n \rightarrow \infty} n^{-1} \ln \|(U(\varepsilon))^n\| = \ln |t_{1, \varepsilon}| \vee \ln |t_{2, \varepsilon}|,$$

where  $t_{1, \varepsilon}$  and  $t_{2, \varepsilon}$  are the eigenvalues of  $U(\varepsilon)$ .

The eigenvalues of  $U(\varepsilon)$  satisfy

$$(1 - \lambda)(\varepsilon C(B) - \lambda) - \varepsilon C(B)^2 = 0;$$

or, equivalently,

$$\lambda^2 - \lambda(1 + \varepsilon C(B)) + \varepsilon C(B)(1 - C(B)) = 0.$$

Thus  $t_{1, \varepsilon}$  and  $t_{2, \varepsilon}$  are

$$((1 + \varepsilon C(B)) \pm ((1 + \varepsilon C(B))^2 - 4(\varepsilon C(B) - \varepsilon C(B)^2))^{1/2})/2.$$

Choose  $\varepsilon_0$  such that for all  $\varepsilon < \varepsilon_0$ ,  $\varepsilon C(B) \leq 0.01$  and  $\varepsilon C(B)^2 \leq 0.01$ . For  $\varepsilon < \varepsilon_0$ ,

$$|t_{1, \varepsilon}| \vee |t_{2, \varepsilon}| = |((1 + \varepsilon C(B)) + ((1 + \varepsilon C(B))^2 - 4(\varepsilon C(B) - \varepsilon C(B)^2))^{1/2})/2|.$$

Since

$$\begin{aligned} (1 + \varepsilon C(B))^2 - 4(\varepsilon C(B) - \varepsilon C(B)^2) &= (1 - \varepsilon C(B))^2 + 4\varepsilon C(B)^2 \\ &\leq 1 + 4\varepsilon C(B)^2 \leq (1 + 2\varepsilon C(B)^2)^2, \end{aligned}$$

$$|t_{1, \varepsilon}| \vee |t_{2, \varepsilon}| \leq |((1 + \varepsilon C(B)) + (1 + 2\varepsilon C(B)^2))/2|.$$

Thus for  $\varepsilon < \varepsilon_0$ ,

$$|t_{1, \varepsilon}| \vee |t_{2, \varepsilon}| \leq 1 + \varepsilon(C(B) + C(B)^2).$$

Since  $\ln(1 + x) \leq x$  for all  $x \geq 0$ ,

$$\ln |t_{1, \varepsilon}| \vee \ln |t_{2, \varepsilon}| \leq \varepsilon(C(B) + C(B)^2).$$

Thus for all  $\varepsilon < \varepsilon_0$ ,

$$h(\varepsilon, B) \leq \varepsilon C_2 \quad \text{for } C_2 := C(B) + C(B)^2,$$

which completes our upper estimate.

3.6. *The finite sum case, with Assumption 1 (lower estimate).* In this section we find a  $C_3 \geq 0$  such that

$$R(\varepsilon, B) \geq pq \sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}| - \varepsilon C_3$$

for all  $\varepsilon$  sufficiently small.

Using the same notation as in §3.5, we want  $C_4$  such that for all  $\varepsilon < \varepsilon_0$ ,  $h(\varepsilon, B) \geq -\varepsilon C_4$ . Choose  $C(B)$  to satisfy Lemma 1, and  $\varepsilon_0$  such that for all  $\varepsilon < \varepsilon_0$ ,  $\varepsilon < (100C(B)^2)^{-1}$ . Note that for any  $C(B)$  satisfying Lemma 1,  $C(B) \geq 1$ , since the  $(1, 1)$ th term of  $B^n/(B^n)_{11} = 1$ .

We now have a very important

CLAIM. For all  $n$ , for  $\varepsilon < \varepsilon_0$ ,  $|(H_{\varepsilon, n})_{11}| \geq (1 - 20C(B)^2\varepsilon)^n/2$ , and  $|(H_{\varepsilon, n})_{21}| \leq 3\varepsilon C(B)|(H_{\varepsilon, n})_{11}|$ .

PROOF. We see this by induction on  $n$ . First notice that by choice of  $\varepsilon_0$ ,  $1 - 20C(B)^2\varepsilon \geq 4/5$ .

For  $n = 1$ ,

$$|(H_{\varepsilon, 1})_{11}| = 1 \geq (1 - 20C(B)^2\varepsilon)/2,$$

and

$$|(H_{\varepsilon, 1})_{21}| = |\varepsilon^{n(1)}| |(B^{m(1)})_{21}/(B^{m(1)})_{11}| \leq C(B)\varepsilon < 3C(B)\varepsilon |(H_{\varepsilon, 1})_{11}|.$$

Now, assume the claim for  $n = 1, \dots, k$ . For  $n = k + 1$ ,  $H_{\varepsilon, k+1} = F_{\varepsilon, k+1}H_{\varepsilon, k}$ . Hence

$$H_{\varepsilon, k+1} = \text{diag}(1, \varepsilon^{n(k+1)})(B^{m(k+1)})_{11}/(B^{m(k+1)})_{11}H_{\varepsilon, k};$$

so

$$(H_{\varepsilon, k+1})_{11} = (H_{\varepsilon, k})_{11} + (B^{m(k+1)})_{12}/(B^{m(k+1)})_{11}(H_{\varepsilon, k})_{21}$$

and

$$(H_{\varepsilon, k+1})_{21} = \varepsilon^{n(k+1)}(B^{m(k+1)})_{21}(H_{\varepsilon, k})_{11} + (B^{m(k+1)})_{22}(H_{\varepsilon, k})_{21}/(B^{m(k+1)})_{11}.$$

Thus, using the induction hypothesis,

$$\begin{aligned} |(H_{\varepsilon, k+1})_{11}| &\geq |(H_{\varepsilon, k})_{11}| - C(B)|(H_{\varepsilon, k})_{21}| \\ &\geq |(H_{\varepsilon, k})_{11}| - C(B)3C(B)\varepsilon|(H_{\varepsilon, k})_{11}| \geq |(H_{\varepsilon, k})_{11}|(1 - 3C(B)^2\varepsilon) \\ &\geq ((1 - 20C(B)^2\varepsilon)^k/2)(1 - 20C(B)^2\varepsilon) \geq (1 - 20C(B)^2\varepsilon)^{k+1}/2. \end{aligned}$$

And since  $3C(B)^2\varepsilon < 3/100$ ,

$$\begin{aligned} |(H_{\varepsilon, k+1})_{11}| &\geq 9|(H_{\varepsilon, k})_{11}|/10. \\ |(H_{\varepsilon, k+1})_{21}| &\leq \varepsilon|(H_{\varepsilon, k})_{11}|C(B) + |(H_{\varepsilon, k})_{21}|C(B) \\ &\leq C(B)\varepsilon(|(H_{\varepsilon, k})_{11}|)(1 + 3C(B)\varepsilon) \\ &\leq 3C(B)\varepsilon|(H_{\varepsilon, k})_{11}|/2 \quad (\text{using that } C(B) \geq 1) \\ &\leq (3C(B)\varepsilon/2)(10/9)|(H_{\varepsilon, k+1})_{11}| \leq 3\varepsilon C(B)|(H_{\varepsilon, k+1})_{11}|, \end{aligned}$$

and hence the claim is proved.

Since  $\|H_{\varepsilon, n}\| \geq |(H_{\varepsilon, n})_{11}|$ ,

$$n^{-1} \ln \|H_{\varepsilon, n}\| \geq n^{-1}(\ln(1/2) + n \ln(1 - 20C(B)^2\varepsilon)).$$

Hence for all  $\varepsilon < \varepsilon_0$ ,  $h(\varepsilon, B) \geq \ln(1 - 20C(B)^2\varepsilon)$ .

For  $\varepsilon < \varepsilon_0$ ,  $0 < 20C(B)^2\varepsilon < 1/5$ . For  $0 < x < 1/5$ ,  $\ln(1 - x) \geq -2x$ . Hence for  $\varepsilon < \varepsilon_0$ ,  $h(\varepsilon, B) \geq -40C(B)^2\varepsilon$ . This completes our lower estimate.

Together with the upper estimate of §3.5, we have now completed the proof of Theorem 5(c), and, in the case where  $\sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|$  is finite and Assumption 1 holds, Theorem 6(b).

**3.7. The finite sum case, without Assumption 1 (upper estimate).** in §§3.7–3.9 we prove Theorem 5 where  $\sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|$  is finite, and where (by Theorem 3)  $B = QAQ^{-1}$ , for  $|\det Q| = 1$ ,  $A = \text{rot}(k, \phi)$ , with  $\phi/\pi$  irrational.

Using the same notation and set-up as in §§3.5 and 3.6, it suffices to see that  $\lim_{\varepsilon \rightarrow 0} h(\varepsilon, B) = 0$ . In this section we do an upper estimate. Given  $\varepsilon_1 > 0$ , we find  $\varepsilon_0$  such that for all  $0 < \varepsilon < \varepsilon_0$ ,  $h(\varepsilon, B) \leq \varepsilon_1$ .

Define  $L_{\varepsilon, i} := \text{diag}(1, \varepsilon^{n(i)}) |B^{m(i)}| / |(B^{m(i)})_{11}|$ ,  $M_{\varepsilon, n} := L_{\varepsilon, n} \cdots L_{\varepsilon, 1}$ , and  $m(\varepsilon, B) := (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|M_{\varepsilon, n}\|$ .

From Lemma 4, it follows that for all  $0 < \varepsilon \leq 1$ ,  $h(\varepsilon, B) \leq m(\varepsilon, B)$ . From Lemma 5, it follows that  $m(\varepsilon_2, B) \leq m(\varepsilon_3, B)$  if  $\varepsilon_2 < \varepsilon_3$ . We now see that there exists  $\varepsilon_0$  such that  $m(\varepsilon_0, B) \leq \varepsilon_1$ .

Let  $K_i := \text{diag}(1, 0) |B^{m(i)}| / |(B^{m(i)})_{11}|$ ,  $Q_n := K_n \cdots K_1$ , and  $q(B) := (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|Q_n\|$ .

Since  $K_i$  is upper triangular, with  $(K_i)_{11} = 1$  and  $(K_i)_{22} = 0$ ,  $q(B) = 0$ .

By Theorem 1,  $0 = \lim_{n \rightarrow \infty} n^{-1} (E(\ln \|Q_n\|))$ ; thus there exists  $N$  such that  $N^{-1} E(\ln \|Q_N\|) \leq \varepsilon_1/3$ . Define  $B := N^{-1} E(\ln \|Q_N\|)$ . Next, we claim that there exists  $\varepsilon_0$  such that  $N^{-1} E(\ln \|M_{\varepsilon_0, N}\|) \leq B + \varepsilon_1/3$ .

The basic idea here is the same as used by Furstenberg and Kesten in their 1960 paper (p. 458).

Define  $b_i: V \rightarrow R$  by  $b_i = \ln \|M_{1/i, N}\|$ , for all  $i \in Z_+$ . Then pointwise  $\ln \|M_{1/i, N}\| \rightarrow \ln \|Q_N\|$  as  $i \rightarrow \infty$ , since  $\lim_{i \rightarrow \infty} \|L_{1/i, n} - K_n\| = 0$  for all  $n$ . The  $b_i$  are monotone downwards. It is easy to check that  $E(\ln \|M_{\varepsilon, n}\|) < \infty$  for  $\varepsilon < 1$  and all  $n$ . By the Monotone Convergence Theorem, there exists  $\varepsilon_0$  such that  $N^{-1} E(\ln \|M_{\varepsilon_0, N}\|) \leq B + \varepsilon_1/3$ .

Thus, there exist  $N$  and  $\varepsilon_0$  such that  $N^{-1} E(\ln \|M_{\varepsilon_0, N}\|) \leq 2\varepsilon_1/3$ . We see that

$$m(\varepsilon_0, B) \leq N^{-1} E(\ln \|M_{\varepsilon_0, N}\|).$$

Since  $m(\varepsilon_0, B) = \lim_{n \rightarrow \infty} n^{-1} E(\ln \|M_{\varepsilon_0, n}\|)$ ,

$$m(\varepsilon_0, B) = \lim_{K \rightarrow \infty} (KN)^{-1} E(\ln \|M_{\varepsilon_0, KN}\|), \quad \text{a.s.}$$

Let  $V_{\varepsilon, m} := L_{\varepsilon, mN} \cdots L_{\varepsilon, (m-1)N+1}$ . Then  $M_{\varepsilon, KN} = V_{\varepsilon, K} \cdots V_{\varepsilon, l}$ . So

$$\begin{aligned} E(\ln \|M_{\varepsilon, KN}\|) &= E(\ln \|V_{\varepsilon, K} \cdots V_{\varepsilon, l}\|) \\ &\leq E(\ln \|V_{\varepsilon, K}\| \cdots \|V_{\varepsilon, 1}\|) = KE(\ln \|V_{\varepsilon, 1}\|). \end{aligned}$$

But  $V_{\varepsilon, 1} = M_{\varepsilon, N}$ , and hence

$$E(\ln \|M_{\varepsilon, KN}\|) \leq KE(\ln \|M_{\varepsilon, N}\|).$$

Thus for  $K \in Z_+$ ,

$$(KN)^{-1} E(\ln \|M_{\varepsilon_0, KN}\|) \leq N^{-1} E(\ln \|M_{\varepsilon_0, N}\|),$$

and so

$$m(\varepsilon_0, B) \leq N^{-1} E(\ln \|M_{\varepsilon_0, n}\|) \leq 2\varepsilon_1/3.$$

For any  $\varepsilon < \varepsilon_0$ ,  $h(\varepsilon, B) \leq m(\varepsilon, B) \leq m(\varepsilon_0, B) \leq 2\varepsilon_1/3 \leq \varepsilon_1$ , and hence we have our desired upper bound on  $h(\varepsilon, B)$ .

**3.8. Preliminaries for the lower estimate of the finite sum case, without Assumption 1.** In this section we reformulate the lower estimate question to one of estimating integrals against stationary measures. We start with two definitions.

**DEFINITION 7.** By Definition 2, we see that a point of  $\mathbf{P}^1$ , call it  $t_1$ , is the equivalence class of all vectors  $\nu$  such that

$$\nu = n \begin{pmatrix} \cos \theta_1 \\ \sin \theta_1 \end{pmatrix}$$

for some  $n \in \mathbf{R} \setminus \{0\}$  and some fixed  $\theta_1 \in [0, \pi)$ . Define  $z(t_1) := \theta_1$ .

**DEFINITION 8.** For  $g \in \text{Gl}_2(\mathbf{R})$ , define a bijection  $Z(g): [0, \pi) \rightarrow [0, \pi)$  by  $Z(g)(\theta) := zg z^{-1}(\theta)$ .

For a given  $g \in \text{Gl}_2(\mathbf{R})$ , we can be more explicit about  $Z(g)$ . Let  $ta^{-1}: \mathbf{R} \cup \{\infty\} \rightarrow [0, \pi)$  be the inverse of the map  $\tan: [0, \pi) \rightarrow \mathbf{R} \cup \{\infty\}$ . Let

$$\eta \begin{pmatrix} u \\ \nu \end{pmatrix} := ta^{-1} \left( \frac{\nu}{u} \right), \quad \text{for } \begin{pmatrix} u \\ \nu \end{pmatrix} \in \mathbf{R}^2 \setminus \{0\}.$$

Then it is easy to see that

$$Z(g)(\theta) = \eta \cdot g \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}.$$

Also recall Definition 6 ( $e(\theta)$ ).

We would like to use Theorem 2 to do the lower estimate; however, Theorem 2 assumes that the determinant of  $F_{\varepsilon, i}$  (introduced in §3.5) is one (a.s.), if we try to apply it directly. This motivates some preliminaries.

$F_{\varepsilon, 1}$  determines a measure on  $\text{Gl}_2(\mathbf{R})$ .  $\text{Det } F_{\varepsilon, 1}$  (a.s.)  $> 0$ , since  $C_{\varepsilon, 1} = \text{diag}(1, \varepsilon^{n(1)})B^{m(1)}$ , and  $\det B = k^2 > 0$ .

Since  $\text{Gl}_2(\mathbf{R})$  acts on  $\mathbf{P}^1$  (which is compact), we deduce from the Furstenberg (1963, p. 384) paper that there exists at least one stationary measure for the measure induced by  $F_{\varepsilon, 1}$  on  $\mathbf{P}^1$ . By a paper of Kaijser (1978, Theorem 7), the stationary measure (in this setting) is unique — call it  $\pi_\varepsilon$ .

We would like to claim

**FORMULA 1.**

$$h(\varepsilon, B) = \int_{\text{Gl}_2(\mathbf{R})} \int_{\mathbf{P}^1} \rho_1(g, t) d\mu_\varepsilon(g) d\pi_\varepsilon(t),$$

where the distribution of  $F_{\varepsilon, 1}$  is  $\mu_\varepsilon$ . (Recall the definition of  $\rho_1$  given in the preliminaries to Theorem 2.)

Given  $F_{\varepsilon, i}$ , define  $Y_{\varepsilon, i} := F_{\varepsilon, i}/(\det F_{\varepsilon, i})^{1/2}$ . Since  $F_{\varepsilon, 1}$  and  $Y_{\varepsilon, 1}$  induce the same transformation of  $\mathbf{P}^1$ ,  $\pi_\varepsilon$  is a stationary measure for the measure on  $\text{Gl}_2(\mathbf{R})$  induced by  $Y_{\varepsilon, 1}$ , which we will denote  $\mu'_\varepsilon$ .

Define  $Z_{\varepsilon, n} := Y_{\varepsilon, n} \cdots Y_{\varepsilon, 1}$ . Since  $Y_{\varepsilon, i}$  are i.i.d.,

$$z(\varepsilon, B) = (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|Z_{\varepsilon, n}\|,$$

which exists once we note that  $E(\ln^+ \|Y_{\varepsilon, 1}\|) < \infty$ .

It is easy to see that  $E(\ln^+ \|F_{\varepsilon,1}\|) < \infty$ , and with a little more work, the reader can verify that  $E(|\ln(\det F_{\varepsilon,1})^{1/2}|) < \infty$ . This implies that

$$E(\ln^+ \|F_{\varepsilon,1}/(\det F_{\varepsilon,1})^{1/2}\|) < \infty,$$

and hence that  $z(\varepsilon, B)$  is an (a.s.) limit.

By the construction of  $Y_{\varepsilon,i}, \mu'_\varepsilon$  is a measure on  $\text{Sl}_2(\mathbf{R})$ .

Since  $E(\ln \|Y_{\varepsilon,1}\|) \leq E(\ln^+ \|Y_{\varepsilon,1}\|) < \infty$ , we can apply Theorem 2 once we see that  $G_\varepsilon$  ( $:=$  the smallest closed subgroup of  $\text{Sl}_2(\mathbf{R})$  containing  $\text{supp } \mu'_\varepsilon$ ) is irreducible for all  $0 < \varepsilon \leq 1$ .

Since both  $m(1)$  and  $n(1)$  can equal 1 or 2 with positive probability,  $G_\varepsilon$  contains  $\text{diag}(\varepsilon^{-1/2}, \varepsilon^{1/2})B/|k|$  and  $\text{diag}(\varepsilon^{-1}, \varepsilon)B/|k|$ .

Suppose that for some  $0 < \varepsilon \leq 1$ ,  $G_\varepsilon$  is not irreducible. For this  $\varepsilon$ , pick  $\nu \in \mathbf{R}^2 \setminus \{0\}$  such that  $G_\varepsilon \nu \subset \{m\nu : m \in \mathbf{R} \setminus \{0\}\}$ . Then there exists  $m_1 \neq 0$  such that  $(\text{diag}(\varepsilon^{-1/2}, \varepsilon^{1/2})B/|k|)\nu = m_1\nu$  and  $m_2 \neq 0$  such that  $(\text{diag}(\varepsilon^{-1}, \varepsilon)B/|k|)\nu = m_2\nu$ . Define  $w := (B/|k|)\nu$ .

Since

$$\tan(\text{diag}(\varepsilon^{-1/2}, \varepsilon^{1/2})w) = \tan(\text{diag}(\varepsilon^{-1}, \varepsilon)w) = \tan \nu,$$

then

$$\varepsilon \tan w = \varepsilon^2 \tan w = \tan \nu.$$

If  $\varepsilon = 1$ ,  $\tan w = \tan \nu$ . For  $0 < \varepsilon < 1$ , either  $\tan w = 0$  or  $\infty$ . In either case,  $\tan w = \tan \nu$ .

Thus,  $\tan(B/|k|)\nu = \tan \nu$ , which implies that  $B$  has a real eigenvalue, contradicting assumptions. So  $G_\varepsilon$  is irreducible for all  $0 < \varepsilon \leq 1$ , and

$$z(\varepsilon, B) = \int_{\mathbf{P}^1} \int_{\text{Sl}_2(\mathbf{R})} \rho_1(g, t) d\mu'_\varepsilon(g) d\pi_\varepsilon(t),$$

since  $\pi_\varepsilon$  is a stationary measure for  $\mu'_\varepsilon$ . To see Formula 1 it suffices to see:

FORMULA 2.

$$h(\varepsilon, B) - E(\ln(\det F_{\varepsilon,1})^{1/2}) = z(\varepsilon, B).$$

FORMULA 3.

$$\begin{aligned} \int_{\mathbf{P}^1} \int_{\text{Gl}_2(\mathbf{R})} \rho_1(g, t) d\mu_\varepsilon(g) d\pi_\varepsilon(t) - E(\ln(\det F_{\varepsilon,1})^{1/2}) \\ = \int_{\mathbf{P}^1} \int_{\text{Gl}_2(\mathbf{R})} \rho_1(g, t) d\mu'_\varepsilon(g) d\pi_\varepsilon(t). \end{aligned}$$

We verify Formula 2:

$$\begin{aligned} h(\varepsilon, B) &= (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|H_{\varepsilon,n}\| = (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|F_{\varepsilon,n} \cdots F_{\varepsilon,1}\| \\ &= (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \left( \left| \prod_{i=1}^n (\det F_{\varepsilon,i})^{1/2} \right| \|Y_{\varepsilon,n} \cdots Y_{\varepsilon,1}\| \right) \\ &= (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \ln(\det F_{\varepsilon,i})^{1/2} + \lim_{n \rightarrow \infty} n^{-1} \ln \|Y_{\varepsilon,n} \cdots Y_{\varepsilon,1}\| \\ &= E(\ln(\det F_{\varepsilon,1})^{1/2}) + z(\varepsilon, B). \end{aligned}$$

To verify Formula 3 it suffices to see that for all  $t \in \mathbf{P}^1$  we have  
 FORMULA 4.

$$\int_{\mathrm{Gl}_2(\mathbf{R})} \rho_1(g, t) d\mu_\varepsilon(g) - E(\ln(\det F_{\varepsilon,1})^{1/2}) = \int_{\mathrm{Gl}_2(\mathbf{R})} \rho_1(g, t) d\mu'_\varepsilon(g).$$

PROOF.

$$\begin{aligned} \int_{\mathrm{Gl}_2(\mathbf{R})} \rho_1(g, t) d\mu'_\varepsilon(g) &= \int_{\mathrm{Gl}_2(\mathbf{R})} \ln \|ge(z(t))\| d\mu'_\varepsilon(g) \\ &= \int_V \ln \|Y_{\varepsilon,1}e(z(t))\| du' = \int_V \left( \ln \|F_{\varepsilon,1}e(z(t))\| - \ln(\det F_{\varepsilon,1})^{1/2} \right) d\mu' \\ &= \int_V \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' - E \left( \ln(\det F_{\varepsilon,1})^{1/2} \right) \\ &= \int_{\mathrm{Gl}_2(\mathbf{R})} \ln \|ge(z(t))\| d\mu_\varepsilon(g) - E \left( \ln(\det F_{\varepsilon,1})^{1/2} \right) \\ &= \int_{\mathrm{Gl}_2(\mathbf{R})} \rho_1(g, t) d\mu_\varepsilon(g) - E \left( \ln(\det F_{\varepsilon,1})^{1/2} \right). \end{aligned}$$

This verifies Formula 1. In the next section we do the lower estimate by showing that there exists  $\varepsilon_0$  such that for all  $0 < \varepsilon \leq \varepsilon_0$ ,

FORMULA 5.

$$\int_{\mathrm{Gl}_2(\mathbf{R})} \int_{\mathbf{P}^1} \rho_1(g, t) d\mu_\varepsilon(g) d\pi_\varepsilon(t) \geq -\varepsilon_1.$$

3.9. *Proof of lower estimate for finite sum case, without Assumption 1.* Before we start with the nontrivial estimates necessary to show Formula 5, we first prove the

CLAIM. There exists at least one matrix  $B$  with complex eigenvalues having irrational arguments modulo  $\pi$  for which the sum  $\sum_{n=1}^{\infty} pq^{n-1} \ln |(B^n)_{11}|$  is finite.

PROOF. It suffices to look among  $B = \mathrm{rot}(1, \phi)$ . We want  $\phi$  (irrational modulo  $\pi$ ) such that  $\sum_{n=1}^{\infty} pq^{n-1} \ln |\cos n\phi|$  is finite. First we see that there exists  $k < 0$  (not depending on  $n$ ) such that  $\int_0^{2\pi} \ln |\cos n\phi| d\phi > k$  for all  $n$ .

$$\begin{aligned} \int_0^{2\pi} \ln |\cos n\phi| d\phi &= 4n \int_0^{\pi/2n} \ln |\cos n\phi| d\phi \\ &= 4n \int_0^{\pi/2n} \ln |\sin n\phi| d\phi = \frac{4n}{n} \left[ \int_0^{\pi/2} \ln |\sin \theta| d\theta \right] = 4c > -\infty, \end{aligned}$$

since  $\int_0^{\pi/2} \ln |\sin \theta| d\theta > -\infty$ . Therefore,

$$\sum_{n=1}^{\infty} pq^{n-1} \left( \int_0^{2\pi} \ln |\cos n\phi| d\phi \right) \geq \sum_{n=1}^{\infty} pq^{n-1} k = k > -\infty.$$

However, by Fubini's Theorem,

$$\sum_{n=1}^{\infty} pq^{n-1} \left( \int_0^{2\pi} \ln |\cos n\phi| d\phi \right) = \int_0^{2\pi} \left( \sum_{n=1}^{\infty} pq^{n-1} \ln |\cos n\phi| \right) d\phi.$$

Thus  $\sum_{n=1}^{\infty} pq^{n-1} \ln |\cos n\phi|$  is a.s.  $(d\phi)$  finite, which verifies the claim.



Define  $r: V \rightarrow [0, \infty)$  by

$$r := \max(|(B^{m(1)})_{12}|, |(B^{m(1)})_{21}|, |(B^{m(1)})_{22}|) / |(B^{m(1)})_{11}|.$$

Define  $K_\varepsilon := \{t \in P^1: |\tan z(t)| < \varepsilon^{1/3}\}$ . Let  $R_\varepsilon := \{\omega: r(\omega) < \varepsilon^{-1/6}\}$ . Then

$$\begin{aligned} \int_{\mathbf{P}^1} \int_V \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_\varepsilon(t) &= \int_{K_\varepsilon} \int_{R_\varepsilon} \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_\varepsilon(t) \\ &\quad + \int_{\mathbf{P}^1 \setminus K_\varepsilon} \int_V \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_\varepsilon(t) \\ &\quad + \int_{K_\varepsilon} \int_{V/R_\varepsilon} \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_\varepsilon(t). \end{aligned}$$

Unraveling definitions,

$$\begin{aligned} \int_{\mathbf{P}^1} \int_{\text{Gl}_2(\mathbf{R})} \rho_1(g, t) d\mu_\varepsilon(g) d\pi_\varepsilon(t) &= \int_{\mathbf{P}^1} \int_{\text{Gl}_2(\mathbf{R})} \ln \|ge(z(t))\| d\mu_\varepsilon(g) d\pi_\varepsilon(t) \\ &= \int_{P^1} \int_V \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_\varepsilon(t). \end{aligned}$$

To complete the lower estimate, it suffices to find  $\varepsilon_{0,i} > 0$  such that for  $i = 1, 2, 3$ , for all  $0 < \varepsilon < \varepsilon_{0,i}$ ,

$$\int \int_{D_{\varepsilon,i}} \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_\varepsilon(t) > -\varepsilon_1/3,$$

where

$$D_{\varepsilon,1} := \{\omega, t: \omega \in R_\varepsilon, t \in K_\varepsilon\}, \quad D_{\varepsilon,2} := \{\omega, t: \omega \in V, t \in \mathbf{P}^1 \setminus K_\varepsilon\}$$

and

$$D_{\varepsilon,3} := \{\omega, t: \omega \in V \setminus R_\varepsilon, t \in K_\varepsilon\}.$$

On  $D_{\varepsilon,1}$

$$\begin{aligned} \ln \|F_{\varepsilon,1}e(z(t))\| &= \ln \|(\text{diag}(1, \varepsilon^{n(1)}) B^{m(1)} / (B^{m(1)})_{11}) e(z(t))\| \\ &\geq \ln |((B^{m(1)})_{11} \cos z(t) + (B^{m(1)})_{12} \sin z(t)) / (B^{m(1)})_{11}| \\ &= \ln |(\cos z(t))(1 + ((B^{m(1)})_{12} / (B^{m(1)})_{11}) \tan z(t))| \end{aligned}$$

(since for  $t \in K_\varepsilon$ ,  $\tan z(t) \neq \infty$ ). Here,

$$|\tan z(t)| |(B^{m(1)})_{12} / (B^{m(1)})_{11}| \leq \varepsilon^{1/3} \varepsilon^{-1/6} = \varepsilon^{1/6}.$$

Thus

$$\ln \|F_{\varepsilon,1}e(z(t))\| \geq \ln |(\cos z(t))(1 - \varepsilon^{1/6})|.$$

Since  $|\sin z(t) / \cos z(t)| < \varepsilon^{1/3}$ , we see that  $1 \leq \cos^2 z(t) + \varepsilon^{2/3} \cos^2 z(t)$ , and hence  $|\cos z(t)| \geq (1 + \varepsilon^{2/3})^{-1/2}$ . Thus on  $D_{\varepsilon,1}$

$$\ln \|F_{\varepsilon,1}e(z(t))\| \geq \ln \left( (1 + \varepsilon^{2/3})^{-1/2} (1 - \varepsilon^{1/6}) \right).$$

So

$$\int \int_{D_{\varepsilon,1}} \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_\varepsilon(t) \geq \ln \left( (1 + \varepsilon^{2/3})^{-1/2} (1 - \varepsilon^{1/6}) \right).$$

Since  $(1 + \varepsilon^{2/3})^{-1/2}(1 - \varepsilon^{1/6})$  converges to 1 as  $\varepsilon \rightarrow 0$ , there exists  $\varepsilon_{0,1}$  such that for all  $0 < \varepsilon < \varepsilon_{0,1}$ ,

$$\ln \left(1 + \varepsilon^{2/3}\right)^{-1/2} \left(1 - \varepsilon^{1/6}\right) > -\varepsilon_1/3,$$

and thus

$$\int \int_{D_{\varepsilon,1}} \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_\varepsilon(t) > -\frac{\varepsilon_1}{3}.$$

Next, consider  $\int_{D_{\varepsilon,3}} \int \ln \|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_\varepsilon(t)$ . First we will see that

$$\lim_{\varepsilon \rightarrow 0} (\ln \varepsilon) \mu'(V \setminus R_\varepsilon) = 0.$$

In order to do this, we now show that  $E(\ln^+ r) < \infty$ .

We have already seen that  $E(\ln^+ \|F_{\varepsilon,1}\|) < \infty$ ; it is easy to see that for all  $A$  ( $2 \times 2$  matrices),  $\| |A| \| \leq \sqrt{2} \|A\|$ .

$$\begin{aligned} \infty &> \int (0 \vee \ln \|F_{\varepsilon,1}\|) d\mu' \\ &= \int (0 \vee \ln(\varepsilon^{n(1)}) \|\text{diag}(1, \varepsilon^{n(1)}) B^{m(1)} / (B^{m(1)})_{11} \varepsilon^{n(1)}\|) d\mu' \\ &\geq \int (0 \vee (n(1) \ln \varepsilon + \ln(\sqrt{2})^{-1} \| |W| \|)) d\mu' \\ &\quad (\text{where } W := \text{diag}(1, \varepsilon^{n(1)}) B^{m(1)} / (B^{m(1)})_{11} \varepsilon^{n(1)}) \\ &\geq \int (0 \vee (n(1) \ln \varepsilon - \ln \sqrt{2} + \ln \| |B^{m(1)} / (B^{m(1)})_{11} \|)) d\mu' \\ &\geq \int (0 \vee (n(1) \ln \varepsilon - \ln \sqrt{2} + \ln r)) d\mu' \\ &\geq \int ((0 \vee \ln r) + n(1) \ln \varepsilon - \ln \sqrt{2}) d\mu' \\ &= \int (0 \vee \ln r) d\mu' - \ln \sqrt{2} + (\ln \varepsilon)/q. \end{aligned}$$

Hence  $E(\ln^+ r) < \infty$ .

So, given  $\varepsilon > 0$ , there exists  $M(\varepsilon)$  such that for any  $M \geq M(\varepsilon)$ ,  $\int_{r \geq M} (\ln r) d\mu' < \varepsilon$ , and thus for all  $M \geq M(\varepsilon)$ ,  $\int_{r \geq M} \ln M d\mu' < \varepsilon$ . Since  $\int_{r \geq M} \ln M d\mu' = \ln M \mu'(V \setminus R_{M^{-6}})$ , we deduce that  $\lim_{M \rightarrow \infty} (\ln M) \mu'(V \setminus R_{M^{-6}}) = 0$ , and hence that  $\lim_{\varepsilon \rightarrow 0} (\ln \varepsilon) \mu'(V \setminus R_\varepsilon) = 0$ .

By Lemma 2 there exists  $K(B)$  such that for all  $n$  and all  $0 \leq \theta < \pi$ ,

$$|(B^n / (B^n)_{11})e(\theta)| \geq K(B).$$

For any  $0 < B \leq 1$ , and for any  $u$  and  $\nu$  real,

$$\left\| \text{diag}(1, B) \begin{pmatrix} u \\ \nu \end{pmatrix} \right\| = (u^2 + B^2 \nu)^{1/2} \geq (B^2 u^2 + B^2 \nu^2)^{1/2} \geq B \left\| \begin{pmatrix} u \\ \nu \end{pmatrix} \right\|.$$

Thus,

$$\|F_{\varepsilon,1}e(\theta)\| = \|(\text{diag}(1, \varepsilon^{n(1)}) B^{m(1)} / (B^{m(1)})_{11})e(\theta)\| \geq \varepsilon^{n(1)} K(B),$$

since  $\|(B^{m(1)})/(B^{m(1)})_{11}e(\theta)\| \geq K(B)$ . Hence

$$\ln\|F_{\varepsilon,1}e(\theta)\| \geq \ln K(B) + n(1)\ln \varepsilon.$$

Thus

$$\begin{aligned} \int \int_{D_{\varepsilon,3}} \ln\|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_{\varepsilon}(t) &\geq \int \int_{D_{\varepsilon,3}} \ln K(B) + n(1)\ln \varepsilon d\mu' d\pi_{\varepsilon}(t) \\ &= \ln K(B)\pi_{\varepsilon}(K_{\varepsilon})\mu'(V \setminus R_{\varepsilon}) + \ln \varepsilon \int \int_{D_{\varepsilon,3}} n(1) d\mu' d\pi_{\varepsilon}(t) \\ &\geq -|\ln K(B)|\mu'(V \setminus R_{\varepsilon}) + \ln \varepsilon \pi_{\varepsilon}(K_{\varepsilon}) \int_{V \setminus R_{\varepsilon}} n(1) d\mu' \\ &= -|\ln K(B)|\mu'(V \setminus R_{\varepsilon}) + \ln \varepsilon \pi_{\varepsilon}(K_{\varepsilon}) \int n(1)1_{V \setminus R_{\varepsilon}} d\mu'. \end{aligned}$$

Since  $n(1)$  and  $m(1)$  are independent,  $n(1)$  and  $1_{V \setminus R_{\varepsilon}}$ , are independent. Since  $En(1) < \infty$ ,

$$\int n(1)1_{V \setminus R_{\varepsilon}} d\mu' = E(n(1))\mu'(V \setminus R_{\varepsilon}) = \mu'(V \setminus R_{\varepsilon})/q.$$

Thus

$$\begin{aligned} \int \int_{D_{\varepsilon,3}} \ln\|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_{\varepsilon}(t) &\geq -|\ln K(B)|\mu'(V \setminus R_{\varepsilon}) + \ln \varepsilon \pi_{\varepsilon}(K_{\varepsilon})\mu'(V \setminus R_{\varepsilon})/q \\ &\geq -|\ln K(B)|\mu'(V \setminus R_{\varepsilon}) + \ln \varepsilon \mu'(V \setminus R_{\varepsilon})/q. \end{aligned}$$

Since  $K(B)$  and  $q$  are fixed, and since  $\lim_{\varepsilon \rightarrow 0}(\ln \varepsilon)\mu'(V \setminus R_{\varepsilon}) = 0$ , there exists  $\varepsilon_{0,3}$  such that for all  $0 < \varepsilon < \varepsilon_{0,3}$ ,

$$\int \int_{D_{\varepsilon,3}} \ln\|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_{\varepsilon}(t) > -\frac{\varepsilon_1}{3}.$$

It remains to consider  $\int \int_{D_{\varepsilon,2}} \ln\|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_{\varepsilon}(t)$ .

$$\begin{aligned} \int \int_{D_{\varepsilon,2}} \ln\|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_{\varepsilon}(t) &\geq \int \int_{D_{\varepsilon,2}} \ln K(B) + n(1)\ln \varepsilon d\mu' d\pi_{\varepsilon}(t) \\ &= \ln K(B)\pi_{\varepsilon}(\mathbf{P}^1 \setminus K_{\varepsilon})\mu'(V) + \ln \varepsilon \int \int_{D_{\varepsilon,2}} n(1) d\mu' d\pi_{\varepsilon}(t) \\ &= \ln K(B)\pi_{\varepsilon}(\mathbf{P}^1 \setminus K_{\varepsilon}) + \ln \varepsilon \pi_{\varepsilon}(\mathbf{P}^1 \setminus K_{\varepsilon})/q. \end{aligned}$$

By the above calculation, it suffices to prove

**FORMULA 6.**  $\lim_{\varepsilon \rightarrow 0}(\ln \varepsilon)\pi_{\varepsilon}(\mathbf{P}^1 \setminus K_{\varepsilon}) = 0$ , to show that there exists  $\varepsilon_{0,2}$  such that for all  $0 < \varepsilon < \varepsilon_{0,2}$ ,

$$\int \int_{D_{\varepsilon,2}} \ln\|F_{\varepsilon,1}e(z(t))\| d\mu' d\pi_{\varepsilon}(t) > -\frac{\varepsilon_1}{3}.$$

To verify Formula 6, we first show that  $(B^2)_{12}/(B^2)_{11} \neq B_{12}/B_{11}$ .

Let

$$Q := \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{where } |\det Q| = 1.$$

If  $\det Q = 1$ , for any  $\theta$ ,

$$Q \operatorname{rot}(1, \theta) Q^{-1} = \begin{pmatrix} \cos \theta - (ac + bd) \sin \theta & (a^2 + b^2) \sin \theta \\ -(c^2 + d^2) \sin \theta & \cos \theta + (ac + bd) \sin \theta \end{pmatrix}.$$

Neither  $B_{11}$  nor  $(B^2)_{11}$  are zero, since  $\sum_{n=1}^{\infty} pq^{n-1} \ln|(B^n)_{11}|$  is finite.

$$(B^2)_{12}/(B^2)_{11} = (k^2(a^2 + b^2) \sin 2\phi)/(k^2(\cos 2\phi - (ac + bd) \sin 2\phi))$$

and

$$B_{12}/B_{11} = (k(a^2 + b^2) \sin \phi)/(k(\cos \phi - (ac + bd) \sin \phi)).$$

Since  $\phi/\pi$  is irrational,  $\sin \phi \neq 0$  and  $\sin 2\phi \neq 0$ . If  $(B^2)_{12}/(B^2)_{11} = B_{12}/B_{11}$ , then  $B_{11}/B_{12} = (B^2)_{11}/(B^2)_{12}$  (with the denominators nonzero). However,

$$B_{11}/B_{12} = (\cot \phi)/(a^2 + b^2) - (ac + bd)/(a^2 + b^2),$$

while

$$(B^2)_{11}/(B^2)_{12} = (\cos 2\phi)/(a^2 + b^2) - (ac + bd)/(a^2 + b^2).$$

Thus  $B_{11}/B_{12} = (B^2)_{11}/(B^2)_{12}$  would imply that  $\cot \phi = \cot 2\phi$ . However,  $\cot 2\phi = (\cot^2 \phi - 1)/2 \cot \phi$ , and hence we would deduce that  $2 \cot^2 \phi - \cot^2 \phi = -1$ , a contradiction. Hence if  $\det Q = 1$ ,  $(B^2)_{12}/(B^2)_{11} \neq B_{12}/B_{11}$ . The case  $\det Q = -1$  follows similarly.

Since  $\pi_\varepsilon$  is a stationary measure, we have

FORMULA 7.

$$\pi_\varepsilon(K_\varepsilon) = \int_{\operatorname{Gl}_2(\mathbf{R})} \pi_\varepsilon(g^{-1}(K_\varepsilon)) d\mu_\varepsilon(g) = \int_V \pi_\varepsilon(F_{\varepsilon,1}^{-1}(K_\varepsilon)) d\mu'.$$

For each  $\omega$ , define

$$B(\varepsilon, \omega) := \left\{ t \in \mathbf{P}^1: |\cot z(t) + (B^{m(1, \omega)})_{12}/(m(1, \omega))_{11}| \geq \varepsilon^{1/6} \right\},$$

where  $|\infty + m| > \varepsilon^{1/6}$  for any finite  $m$ .

We consider  $\varepsilon < t(B)$ . For  $\omega \in R_\varepsilon$ , we claim that

- (i) for  $\varepsilon < 2^{-6}$ ,  $F_{\varepsilon,1}(\omega)^{-1}(K_\varepsilon) \supset K_\varepsilon$ , and
- (ii) for  $\varepsilon < 2^{-24}$ ,  $F_{\varepsilon,1}(\omega)^{-1}(K_\varepsilon) \supset B(\varepsilon, \omega)$ .

PROOF OF (i). It suffices to see that  $F_{\varepsilon,1}(K_\varepsilon) \subset K_\varepsilon$ . Recall that  $Z(g)(\theta) = \eta g e(\theta)$ ; pick  $t \in K_\varepsilon$ .  $F_{\varepsilon,1}(t) \in K_\varepsilon$  if and only if  $|\tan z(F_{\varepsilon,1}(t))| < \varepsilon^{1/3}$  if and only if  $|\tan \eta F_{\varepsilon,1} e(z(t))| < \varepsilon^{1/3}$ .

Since

$$F_{\varepsilon,1} e(z(t)) = (\operatorname{diag}(1, \varepsilon^{n(1)}) B^{m(1)}) / (B^{m(1)})_{11} e(z(t)),$$

to see that  $F_{\varepsilon,1}(t) \in K_\varepsilon$ , it suffices to see that  $|J/L| < \varepsilon^{1/3}$ , where

$$J := \varepsilon^{n(1)}((B^{m(1)})_{21}/(B^{m(1)})_{11}) \cos z(t) + ((B^{m(1)})_{22}/(B^{m(1)})_{11}) \sin z(t)$$

and

$$L := \cos z(t) + ((B^{m(1)})_{12}/(B^{m(1)})_{11}) \sin z(t).$$

For  $z(t) = \pi/2$ ,  $t \notin K_\varepsilon$ . Otherwise, we have that  $|J/L| \leq |M/N|$ , where

$$M := \varepsilon((B^{m(1)})_{21}/(B^{m(1)})_{11}) + ((B^{m(1)})_{22}/(B^{m(1)})_{11}) \tan z(t)$$

and

$$N := 1 + ((B^{m(1)})_{12}/(B^{m(1)})_{11}) \tan z(t).$$

So (i) follows from showing that  $|M/N| < \varepsilon^{1/3}$ . For  $\omega \in R_\varepsilon$  and  $t \in K_\varepsilon$ ,

$$|\tan z(t)|r(\omega) \leq \varepsilon^{1/6}.$$

So

$$\begin{aligned} |M/N| &\leq \varepsilon \left( \varepsilon^{-1/6} + \varepsilon^{-1/6} \varepsilon^{1/3} \right) \left( 1 - \varepsilon^{1/3} \varepsilon^{-1/6} \right)^{-1} \leq \varepsilon \left( 2\varepsilon^{-1/6} \right) \left( 1 - \varepsilon^{1/6} \right)^{-1} \\ &= 2\varepsilon^{5/6} \left( 1 - \varepsilon^{1/6} \right)^{-1}. \end{aligned}$$

For

$$\varepsilon < 2^{-6}, \quad 2\varepsilon^{5/6} \left( 1 - \varepsilon^{1/6} \right)^{-1} < 4\varepsilon^{5/6} < \varepsilon^{1/3},$$

and hence (i) follows.

PROOF OF (ii). It suffices to see that  $F_{\varepsilon,1}B(\varepsilon, \omega) \subset K_\varepsilon$ . Using  $J$  and  $L$  as in the proof of (i), it suffices to see that  $|J/L| < \varepsilon^{1/3}$ , for  $t \in B(\varepsilon, \omega)$ . For  $z(t) = 0$ ,

$$|J/L| \leq \varepsilon |(B^{m(1)})_{21}/(B^{m(1)})_{11}| \leq \varepsilon \cdot \varepsilon^{-1/6} < \varepsilon^{1/3}.$$

Otherwise,  $|J/L| \leq |P/R|$ , where

$$P := \varepsilon((B^{m(1)})_{21}/(B^{m(1)})_{11})\cot z(t) + (B^{m(1)})_{22}/(B^{m(1)})_{11}$$

and

$$R := \cot z(t) + (B^{m(1)})_{12}/(B^{m(1)})_{11}.$$

Consider  $\varepsilon < 2^{-24}$ . If  $|\cot z(t)| < \varepsilon^{-1/4}$ ,

$$\begin{aligned} |P/R| &\leq \varepsilon \left( \varepsilon^{-1/6} \varepsilon^{-1/4} + \varepsilon^{-1/6} \right) / \varepsilon^{1/6} \\ &\leq \varepsilon^{5/6} \left( \varepsilon^{-5/12} + \varepsilon^{-1/6} \right) < \varepsilon^{5/6} \varepsilon^{-1/2} = \varepsilon^{1/3}, \end{aligned}$$

and hence  $|J/L| < \varepsilon^{1/3}$ , as desired.

If  $|\cot z(t)| \geq \varepsilon^{-1/4}$ ,

$$|\cot z(t) + (B^{m(1)})_{12}/(B^{m(1)})_{11}| \geq |\cot z(t)| - |(B^{m(1)})_{12}/(B^{m(1)})_{11}| \geq |\cot z(t)|/2,$$

since

$$|(B^{m(1)})_{12}/(B^{m(1)})_{11}| \leq |\cot z(t)|/2,$$

since

$$\left( \varepsilon^{-1/4} \right) / 2 \geq \varepsilon^{-1/6}.$$

Thus

$$\begin{aligned} |P/R| &\leq \varepsilon \left| \frac{((B^{m(1)})_{21}/(B^{m(1)})_{11})\cot z(t) + (B^{m(1)})_{22}/(B^{m(1)})_{11}}{(\cot z(t))/2} \right| \\ &\leq 2\varepsilon (|(B^{m(1)})_{21}/(B^{m(1)})_{11}| + |(B^{m(1)})_{22}/(B^{m(1)})_{11}|) \\ &\leq 2\varepsilon \left( 2\varepsilon^{-1/6} \right) \leq 4\varepsilon^{5/6} < \varepsilon^{1/3}, \end{aligned}$$

and hence (ii) follows.

We now assume that  $\varepsilon < 2^{-24}$ ,  $\varepsilon < t(B)$ ,

$$\varepsilon < ((B^2)_{11})^6 / \max(|(B^2)_{12}|^6, |(B^2)_{21}|^6, |(B^2)_{22}|^6),$$

and

$$\varepsilon < ((B)_{11})^6 / \max(|B_{12}|^6, |B_{21}|^6, |B_{22}|^6).$$

By Formula 7,

$$\int_{R_\varepsilon} \pi_\varepsilon(F_{\varepsilon,1}^{-1}(K_\varepsilon)) d\mu' \leq \pi_\varepsilon(K_\varepsilon).$$

Define  $U_{\varepsilon,1} := \{\omega: B^{m(1,\omega)} = B\}$ ,  $U_{\varepsilon,2} := \{\omega: B^{m(1,\omega)} = B^2\}$ .

By the assumptions on  $\varepsilon$ ,  $U_{\varepsilon,1} \cup U_{\varepsilon,2} \subset R_\varepsilon$ . Also, since  $B_{12}/B_{11} \neq (B^2)_{12}/(B^2)_{11}$ ,  $U_{\varepsilon,1} \cap U_{\varepsilon,2} = \emptyset$ . Thus, by (i) and (ii),

$$\begin{aligned} \pi_\varepsilon(K_\varepsilon) &\geq \int_{R_\varepsilon} \pi_\varepsilon(B(\varepsilon, \omega) \setminus K_\varepsilon) d\mu' + \int_{R_\varepsilon} \pi_\varepsilon(K_\varepsilon) d\mu' \\ &\geq \mu'(R_\varepsilon) \pi_\varepsilon(K_\varepsilon) + \int_{R_\varepsilon} \pi_\varepsilon(B(\varepsilon, \omega) \setminus K_\varepsilon) d\mu' \\ &\geq \mu'(R_\varepsilon) \pi_\varepsilon(K_\varepsilon) + \int_{U_{\varepsilon,1}} \pi_\varepsilon(B(\varepsilon, \omega) \setminus K_\varepsilon) d\mu' \\ &\quad + \int_{U_{\varepsilon,2}} \pi_\varepsilon(B(\varepsilon, \omega) \setminus K_\varepsilon) d\mu'. \end{aligned}$$

For all  $\omega \in U_{\varepsilon,1}$ ,

$$B(\varepsilon, \omega) = \left\{ t \in P^1: |\cot z(t) + B_{12}/B_{11}| \geq \varepsilon^{1/6} \right\},$$

independent of  $\omega$ . On  $U_{\varepsilon,1}$  let  $B(\varepsilon, \omega) \setminus K_\varepsilon := X_{\varepsilon,1}$ . Likewise, for all  $\omega \in U_{\varepsilon,2}$ ,

$$B(\varepsilon, \omega) = \left\{ t \in P^1: |\cot z(t) + (B^2)_{12}/(B^2)_{11}| \geq \varepsilon^{1/6} \right\},$$

so on  $U_{\varepsilon,2}$  let  $B(\varepsilon, \omega) \setminus K_\varepsilon := X_{\varepsilon,2}$ . Then

$$\pi_\varepsilon(K_\varepsilon) \geq \mu'(R_\varepsilon) \pi_\varepsilon(K_\varepsilon) + \mu'(U_{\varepsilon,1}) \pi_\varepsilon(X_{\varepsilon,1}) + \mu'(U_{\varepsilon,2}) \pi_\varepsilon(X_{\varepsilon,2}).$$

$$\mu'(U_{\varepsilon,1}) \geq \mu'(\{\omega: m(1, \omega) = 1\}) = p,$$

and

$$\mu'(U_{\varepsilon,2}) \geq \mu'(\{\omega: m(1, \omega) = 2\}) = pq \quad (< p).$$

So

$$\begin{aligned} \pi_\varepsilon(K_\varepsilon) &\geq \mu'(R_\varepsilon) \pi_\varepsilon(K_\varepsilon) + pq(\pi_\varepsilon(X_{\varepsilon,1}) + \pi_\varepsilon(X_{\varepsilon,2})) \\ &\geq \mu'(R_\varepsilon) \pi_\varepsilon(K_\varepsilon) + pq \pi_\varepsilon(X_{\varepsilon,1} \cup X_{\varepsilon,2}). \end{aligned}$$

CLAIM.  $X_{\varepsilon,1} \cup X_{\varepsilon,2} = \mathbf{P}^1 \setminus K_\varepsilon$ . To prove this, it is enough to show

$$\begin{aligned} &\left\{ t \in \mathbf{P}^1: |\cot z(t) + B_{12}/B_{11}| \geq \varepsilon^{1/6} \right\} \\ &\cup \left\{ t \in \mathbf{P}^1: |\cot z(t) + (B^2)_{12}/(B^2)_{11}| \geq \varepsilon^{1/6} \right\} = \mathbf{P}^1. \end{aligned}$$

If not, there exists  $t$  such that

$$|\cot z(t) + B_{12}/B_{11}| < \varepsilon^{1/6} \quad \text{and} \quad |\cot z(t) + (B^2)_{12}/(B^2)_{11}| < \varepsilon^{1/6}.$$

So

$$\begin{aligned} 2\varepsilon^{1/6} &\geq |\cot z(t) + B_{12}/B_{11}| + |-\cot z(t) - (B^2)_{12}/(B^2)_{11}| \\ &\geq |B_{12}/B_{11} - (B^2)_{12}/(B^2)_{11}|. \end{aligned}$$

But then

$$2\varepsilon^{1/6} \geq |B_{12}/B_{11} - (B^2)_{12}/(B^2)_{11}| < 2tB^{1/6} > 2\varepsilon^{1/6},$$

a contradiction.

Hence

$$\pi_\varepsilon(K_\varepsilon) \geq \mu'(R_\varepsilon)\pi_\varepsilon(K_\varepsilon) + pq\pi_\varepsilon(\mathbf{P}^1 \setminus K_\varepsilon),$$

which implies

$$\pi_\varepsilon(K_\varepsilon)(1 - \mu'(R_\varepsilon)) \geq pq\pi_\varepsilon(\mathbf{P}^1 \setminus K_\varepsilon),$$

yielding

$$(1 - \mu'(R_\varepsilon)) \geq pq\pi_\varepsilon(\mathbf{P}^1 \setminus K_\varepsilon) \quad \text{and} \quad \mu'(V \setminus R_\varepsilon) \geq pq\pi_\varepsilon(\mathbf{P}^1 \setminus K_\varepsilon).$$

Since we have already seen that  $\lim_{\varepsilon \rightarrow 0} (\ln \varepsilon) \mu'(V \setminus R_\varepsilon) = 0$ , we have

$$\lim_{\varepsilon \rightarrow 0} (\ln \varepsilon) pq\pi_\varepsilon(\mathbf{P}^1 \setminus K_\varepsilon) = 0, \quad \text{and} \quad \lim_{\varepsilon \rightarrow 0} (\ln \varepsilon) \pi_\varepsilon(\mathbf{P}^1 \setminus K_\varepsilon) = 0.$$

Thus Formula 7 has been verified, which completes the proof of Theorem 5 for  $\sum_{n=1}^\infty pq^{n-1} \ln |(B^n)_{11}|$  finite.

For a matrix with real eigenvalues, it is not generally true that there exists  $K(B)$  such that for all  $n$  and all  $0 \leq \theta < \pi$ ,  $|(B^n/(B^n)_{11})e(\theta)| \geq K(B)$ .

Hence it is not readily apparent how the above computation could be mimicked to prove Theorem 5 in the case where  $B$  has real eigenvalues.

3.10. *The  $-\infty$  sum case.* We finish the proof of Theorem 5 by handling the case where  $\sum_{n=1}^\infty pq^{n-1} \ln |(B^n)_{11}| = -\infty$ . By Theorem 5(a), it suffices to show that  $\lim_{\varepsilon \rightarrow 0} R(\varepsilon, B) = -\infty$ . As before, form  $A_{\varepsilon,i}, C_{\varepsilon,i}, S_{\varepsilon,n}$ , and define  $S(\varepsilon, B)$ . By Theorem 6,  $R(\varepsilon, B) = pqS(\varepsilon, B)$ . Thus, it suffices to show that for  $M$  arbitrary, there exists  $\varepsilon_0$  such that for all  $0 < \varepsilon < \varepsilon_0$ ,  $S(\varepsilon, B) \leq M$ . Defining  $f(i)$ ,  $m(i)$ , and  $n(i)$  as before, we have  $C_{\varepsilon,i} = \text{diag}(1, \varepsilon^{n(i)})B^{m(i)}$ .

It is easy to see that  $\sum_{n=1}^\infty pq^{n-1} (0 \vee \ln |(B^n)_{11}|) < \infty$ . Since

$$\sum_{n=1}^\infty pq^{n-1} \ln |B_{n,11}| = -\infty,$$

by the Monotone Convergence Theorem, there exists  $t > -\infty$  such that

$$\sum_{n=1}^\infty pq^{n-1} (t \vee \ln |(B^n)_{11}|) \leq M - 1.$$

Define  $G_i$  by

$$(G_i)_{jk} := |(B^{m(i)})_{jk}| \quad \text{for } (j, k) \neq (1, 1);$$

$$(G_i)_{11} = (e^t \vee |(B^{m(i)})_{11}|).$$

Define

$$D_{\varepsilon,i} := \text{diag}(1, \varepsilon n(i))G_i, \quad J_{\varepsilon,n} := D_{\varepsilon,n} \cdots D_{\varepsilon,1},$$

and

$$J(\varepsilon, B) = (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|J_{\varepsilon,n}\|.$$

By Lemmas 4 and 5,  $S(\varepsilon, B) \leq J(\varepsilon, B)$  for all  $0 < \varepsilon \leq 1$ , and for  $\varepsilon_1 < \varepsilon_2$ ,  $J(\varepsilon_1, B) \leq J(\varepsilon_2, B)$ .

We now show that there exists  $\varepsilon_0$  with  $J(\varepsilon_0, B) \leq M - 1/2$ .

Define  $K_i := \text{diag}(1, 0)G_i$ ,  $Q_n := K_n \cdots K_1$ , and

$$q(B) = (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|Q_n\|.$$

Since  $K_i$  is upper triangular, with  $(K_i)_{11} = e^t \vee |(B^{m(i)})_{11}|$  and  $(K_i)_{22} = 0$ , Theorem 4 implies that

$$g(B) = \int_V \ln(e^t \vee |(B^{m(i)})_{11}|) d\mu'.$$

Define  $D_n := \{\omega: m(1, \omega) = n\}$ . Since

$$\int_V \ln^+(e^t \vee |(B^{m(1)})_{11}|) d\mu' \leq E(\ln^+ \|J_{\varepsilon, 1}\|) < \infty,$$

we can rearrange, and hence

$$\int_V \ln(e^t \vee |(B^{m(1)})_{11}|) d\mu' = \sum_{n=1}^{\infty} \int_{D_n} \ln(e^t \vee |(B^{m(1)})_{11}|) d\mu'.$$

$(B^{m(1)})_{11} = (B^n)_{11}$  on  $D_n$ , and  $\mu'(D_n) = pq^{n-1}$ . Hence

$$\begin{aligned} q(B) &= \sum_{n=1}^{\infty} pq^{n-1} \ln(e^t \vee |(B^n)_{11}|) \\ &= \sum_{n=1}^{\infty} pq^{n-1} (t \vee \ln |(B^n)_{11}|) \leq M - 1. \end{aligned}$$

Furthermore,  $q(B) \geq \sum_{n=1}^{\infty} pq^{n-1} t > -\infty$ .

By Theorem 1,  $q(B) = \lim_{n \rightarrow \infty} n^{-1} E(\ln \|Q_n\|)$ , and since  $q(B) \leq M - 1$ , choose  $N$  such that  $N^{-1} E(\ln \|Q_N\|) \leq M - 3/4$ . Define  $b_i: V \rightarrow \mathbf{R}$  by  $b_i = \ln \|J_{1/i, N}\|$ , for all  $i \in \mathbf{Z}_+$ . Then pointwise  $\lim_{i \rightarrow \infty} \ln \|J_{1/i, N}\| = \ln \|Q_N\|$ .

By Lemma 5, the  $b_i$  are monotone downwards. Since  $E(\ln^+ \|J_{1/i, N}\|)$  is finite, the Monotone Convergence Theorem applies and there exists  $\varepsilon_0$  such that  $N^{-1} E(\ln \|J_{\varepsilon_0, N}\|) \leq M - 1/2$ . By a proof which is virtually identical to one already given in §3.7,  $J(\varepsilon_0, B) \leq M - 1/2$ .

For all  $0 < \varepsilon < \varepsilon_0$ , we have already seen that  $J(\varepsilon, B) \leq M - 1/2$ . Since  $S(\varepsilon, B) < J(\varepsilon, B)$ , for all  $\varepsilon < \varepsilon_0$ ,  $S(\varepsilon, B) < M$ . This completes the  $-\infty$  case.

**3.11. Remarks about Theorem 6.** The reason that we assumed  $B$  was nonsingular in Theorem 6 is that we can give an exact answer if  $\det B = 0$ .

Assume that  $A_i$  are i.i.d. random  $2 \times 2$  matrices,  $T_n := A_n \cdots A_1$ ,  $0 < p < 1$ , where  $A_1 = A$  w.p.  $p$ ,  $A_1 = B$  w.p.  $q := 1 - p$ , and  $\det B = 0$ . Let  $T(A, B) := (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|T_n\|$ .

By Theorem 3, there exists  $Q$  such that  $QBQ^{-1}$  is either 0,  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ , or  $\text{diag}(\alpha, 0)$  for some  $\alpha \neq 0$ .

Let  $B_i = QA_iQ^{-1}$ , and let  $S_n := B_n \cdots B_1$ . It is easy to see that (a.s.)  $\lim_{n \rightarrow \infty} n^{-1} \ln \|S_n\| = T(A, B)$ . Notice that  $B_i = QA_iQ^{-1}$  w.p.  $p$  and  $QBQ^{-1}$  w.p.  $q$ . Suppose that  $QBQ^{-1} = 0$  or  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ . Let  $L := \{\omega: A_i(\omega) = B \text{ for } i = 1 \text{ and } 2\}$ .  $\mu(L) = q^2 > 0$ . For  $\omega \in L$ ,  $B_2(\omega)B_1(\omega) = 0$ . Hence  $\lim_{n \rightarrow \infty} n^{-1} \ln \|S_n(\omega)\| = -\infty$  for any  $\omega \in L$ . Since (a.s.)  $\lim_{n \rightarrow \infty} n^{-1} \ln \|S_n\| = T(A, B)$ , and since with positive probability  $\lim_{n \rightarrow \infty} n^{-1} \ln \|S_n\| = -\infty$ , we deduce that  $T(A, B) = -\infty$ .



Finally, suppose that  $B_1 = QAQ^{-1}$  w.p.  $p$ , and  $QBQ^{-1} = \text{diag}(\alpha, 0)$  w.p.  $q$  ( $\alpha \neq 0$ ). Define  $g_i = \alpha$  if  $B_i = \text{diag}(\alpha, 0)$ ,  $g_i = 1$  otherwise, and  $E_i = \text{diag}(1, 0)$  if  $B_i = \text{diag}(\alpha, 0)$ ,  $E_i = QAQ^{-1}$  otherwise. Then for all  $i$ ,  $g_i E_i = B_i$ . Let  $F_n := E_n \cdots E_1$ , and  $F(A, B) := (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|F_n\|$ .

$$\begin{aligned} T(A, B) &= (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln \|S_n\| = (\text{a.s.}) \lim_{n \rightarrow \infty} n^{-1} \ln |g_n \cdots g_1| \|E_n \cdots E_1\| \\ &= \text{a.s. } E(\ln |g_1|) + F(A, B) = q \ln |\alpha| + F(A, B). \end{aligned}$$

$F(A, B)$  is the (a.s.) limit of  $n^{-1} \ln \|E_n \cdots E_1\|$  where the  $E_i$  are i.i.d., probability  $p$  of  $QAQ^{-1}$ , probability  $q$  of  $\text{diag}(1, 0)$ . By Theorem 5(a),

$$F(A, B) = pq \sum_{n=1}^{\infty} qp^{n-1} \ln |((QAQ^{-1})^n)_{11}|.$$

Hence

$$T(A, B) = q \ln |\alpha| + pq \sum_{n=1}^{\infty} qp^{n-1} \ln |((QAQ^{-1})^n)_{11}|.$$

Theorem 5 thus gives a first term in an asymptotic expansion for  $T(A, B)$ . It would seem very worthwhile to get sharper, higher order expansions for this elusive limit.

ACKNOWLEDGEMENT. A great debt of gratitude is owed to Dick Dudley. He introduced me to the work of Furstenberg and Kesten and offered both encouragement and sound advice to aid me during this work.

I would also like to thank the referee for a shorter proof of Theorem 3 than the original.

## REFERENCES

- K. L. Chung (1974), *A course in probability theory*, 2nd ed., Academic Press, New York.
- W. Feller (1966) *An introduction to probability theory and its applications*. Vol. 1, Wiley, New York.
- H. Furstenberg (1963), *Noncommuting random products*, Trans. Amer. Math. Soc. **108**, 377–428.
- H. Furstenberg and H. Kesten (1960), *Products of random matrices*, Ann. Math. Statist. **31**, 457–469.
- T. Kaijser (1978), *A limit theorem for Markov chains with applications to products of random matrices*, Duke Math. J. **45**, 311–349.
- L. Loomis (1953), *Abstract harmonic analysis*, Van Nostrand, New York.
- V. V. Sazonov and V. N. Tutubalin (1966), *Probability distributions on topological groups*, Theory Probab. Appl. **11**, 1–45.
- G. E. Shilov (1971), *Linear algebra*, Prentice-Hall, Englewood Cliffs, N.J.
- V. N. Tutubalin (1965), *On limit theorems for the product of random matrices*, Theory Probab. Appl. **10**, 15–27.
- , (1967), *On measures whose support is generated by a Lie algebra*, Theory Probab. Appl. **12**, 134–138.
- , (1969), *Some theorems of the type of the strong law of large numbers*, Theory Probab. Appl. **14**, 313–319.

BELL COMMUNICATIONS RESEARCH, IH-301, 185 MONMOUTH PARKWAY, WEST LONG BRANCH, NEW JERSEY 07764