Random Matrix Products

Alex Rutar[*]
University of St Andrews

Spring 2021[†]

[*] *alex@rutar.org*
[†] Last updated: June 11, 2021

Contents

Preface

These lecture notes on random matrix products are prepared for the reading group on random matrix products for the analysis group in Spring 2021. Much of the content is based on Alex Gorodnik's lecture notes for his course "Random walks on matrix groups". Any errors or omissions can be sent to the author.

I. The Multiplicative Ergodic Theorem

1 Random matrix products and the subadditive ergodic theorem

1.1 The Birkhoff ergodic theorem

Let Ω be a separable, second-countable metric space equipped with its Borel σ-algebra $\mathcal{B}$, and let μ be a Borel probability measure on Ω. Suppose we are given a measurable function $\theta : \Omega \to \Omega$. We denote the *pushforward* of μ by θ to denote the Borel probability measure defined by the rule

$$\theta_* \mu(E) = \mu(\theta^{-1}(E))$$

for Borel sets $E \subset \Omega$. We say that the function θ is *measure preserving* if $\theta_* \mu = \mu$. In this situation, we call the information (Ω, μ, θ) a *measure-preserving dynamical system*.

Given a Borel set $E \subset \Omega$, we say that E is θ-invariant if $\theta^{-1}(E) = E$, and denote the set of θ-invariant sets by $\mathcal{B}_\theta$. More generally, we say that a measurable function $f : \Omega \to K$ where K is a topological space is θ-invariant if $f(\omega) = f \circ \theta(\omega)$ for μ-a.e. ω. One can verify that $\mathcal{B}_\theta$ is a Borel σ-subalgebra of $\mathcal{B}$. In particular, f is θ-invariant if and only if f is $\mathcal{B}_\theta$-measurable. We say that (Ω, μ, θ) is *ergodic* if each θ-invariant set $E \in \mathcal{B}_\theta$ either has $\mu(E) = 0$ or $\mu(E) = 1$.

We will denote by θ^n the n-fold composition $\theta \circ \cdots \circ \theta$. Given a function f, we write $f = f^+ + f^-$ where $f^+ \geq 0$ and $f^- \leq 0$. A standard result is the following.

1.1 Theorem (Birkhoff Pointwise Ergodic). *Let (Ω, μ, θ) be an ergodic measure-preserving dynamical system and let $f = f^+ + f^-$ satisfy $f_+ \in L^1(\Omega, \mu)$. Then for μ-a.e. $\omega \in \Omega$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(\theta^i(\omega)) = \int_\Omega f \, \mathrm{d}\mu$$

where the limit may be attained at $-\infty$.

We have written Theorem 1.1 in additive notation, but it can be easily rephrased in multiplicative notation. Denote by $\log^+(x) = \max(0, \log x)$. Write $g = \exp(f)$ and note that $f_+ = \log^+(g)$. Then for μ-a.e. $\omega \in \Omega$,

$$\lim_{n\to\infty} (g(T^{n-1}\omega) \cdots g(\omega))^{1/n} = \exp\left(\int_\Omega \log g \, \mathrm{d}\mu\right).$$

Of course, here, the group written in product notation is still commutative. In the following section, we consider a more general setting where this is no longer the case.

1.2 Random matrix products

The setting of Theorem 1.1 is nice, but in these notes we are interested in a somewhat more general situation. First consider the following example. Let Ω denote the compact product space $\mathrm{GL}_d(\mathbb{C})^{\mathbb{N}}$ equipped with the left-shift map $\sigma : \Omega \to \Omega$ given by

$$\sigma((M_n)_{n=1}^{\infty}) = (M_n)_{n=2}^{\infty}$$

for a sequence of matrices $(M_n)_{n=1}^{\infty} \subset \Omega$. Let ν be a probability measure on $\mathrm{GL}_d(\mathbb{C})$ and let $X_i : \Omega \to \mathrm{GL}_d(\mathbb{C})$ for $i \in \mathbb{N}$ be independent random matrices with distribution ν. Asymptotic behaviour of random products of the form $X_n \cdots X_1$ can be interpreted as a matrix-valued generalization of the law of large numbers.

More generally, we are interested in matrix-valued measurable functions, i.e. functions $X : \Omega \to \mathrm{GL}_d(\mathbb{C})$ on a measure-preserving space (Ω, μ, θ). This setting is a generalization of the setting in Theorem 1.1, where we considered a measurable function $f : \Omega \to \mathbb{R}$ satisfying an integrability criteria. Let $\|\cdot\| : \mathrm{GL}_d(\mathbb{C})$ be a matrix norm. We will assume that $\|\cdot\|$ is *submultiplicative* (i.e. $\|AB\| \leq \|A\| \|B\|$), but we do not lose any generality since all matrix norms are equivalent. We also assume that X satisfies the integrability condition

$$\int_{\Omega} \log^{+} \|X(\omega)\| \, \mathrm{d}\omega < \infty.$$

As in the prior section, we are interested in determining statistical information concerning the limit of the random matrix product

$$S_n(\omega) = X(T^{n-1}\omega) \cdots X(\omega).$$

We will investigate various statistical properties of the random products $S_n(\omega)$. Here are three such examples which we will focus on:

(i) the growth rate of $\|S_n(\omega)\| = \left\|X(\theta^{n-1}\omega) \cdots X(\omega)\right\|$ for large n and "typical" ω.

(ii) the growth rate from a fixed starting point $\left\|X(\theta^{n-1}\omega) \cdots X(\omega)v\right\|$ for some $v \in \mathbb{C}^d$

(iii) the behaviour of the directions $\left\|X(\theta^{n-1}\omega) \cdots X(\omega)v\right\| / \left\|X(\theta^{n-1}\omega) \cdots X(\omega)v\right\|$ for some $v \in \mathbb{C}^d$.

Here are some settings where this theory is applicable.

Example. 1. Given fixed matrices $M_1, \ldots, M_\ell \in \mathrm{GL}_d(\mathbb{C})$, generate a sequence $S_0 = I$ and $S_{n+1} = M_i \cdot S_n$ where we take matrix M_i with probability $1/\ell$. The products S_n can be interpreted as a random walk on $\mathrm{GL}_d(\mathbb{C})$ (or $\mathbb{C}^d$) where the "steps" are given by multiplication by a matrix M_i.

2. If $U \subset \mathbb{R}^d$ is an open set and $F : U \to U$ is smooth, by the chain rule, the Jacobian of F^n at a point u satisfies

$$D(F^n)_u = (DF)_{F^{n-1}u} \cdots (DF)_u.$$

Here, $DF : U \to \mathrm{GL}_d(\mathbb{R})$ is a matrix-valued measurable function. The growth rate of DF is related to the entropy of F and the dimension of invariant measures.

3. If $T_i(x) = A_i x + t_i$ where $A_1, \ldots, A_\ell \in \mathrm{GL}_d(\mathbb{R})$ have operator norms $\|A_i\| < 1$ for $i = 1, \ldots, \ell$ and $t_i \in \mathbb{R}^n$, then there is a unique *self-affine set* K satisfying

$$K = \bigcup_{i=1}^{\ell} T_i(K)$$

and, given probabilities $p_1, \dots, p_\ell$, a unique *self-affine measure*, which is a Borel probability measure ν satisfying

$$\nu = \sum_{i=1}^{\ell} p_i (T_i)_* \mu.$$

Here, dimensional properties of the measure ν are related to properties of random products of the matrices $\{A_1, \dots, A_\ell\}$.

1.3 LYAPUNOV EXPONENTS

A fundamental statistical property associated with the matrix-valued function X is the following.

Definition. With notation as above, we define the *top Lyapunov exponent* $\lambda : \Omega \to \mathbb{R}$ by

$$\lambda(\omega) = \lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)\| . \tag{1.1}$$

We now have the following fundamental result.

1.2 Theorem (Furstenburg-Kesten). *The function λ is θ-invariant and satisfies*

$$\int_\Omega \lambda(\omega)\, \mathrm{d}\omega = \lim_{n\to\infty} \frac{1}{n} \int_\Omega \log \|S_n(\omega)\| \,\mathrm{d}\omega.$$

This result can be thought of as interchanging the limit with the integral, i.e. averaging over space is the same as averaging over time.

In fact, we will prove Theorem 1.2 as a consequence of a more general result. We first make some observations about the average $a_n := \int_\Omega \log \|S_n(\omega)\|$. Observe by submultiplicativity of the matrix norm that

$$\begin{aligned}
a_{n+m} &:= \int_\Omega \log \|S_{n+m}(\omega)\| \,\mathrm{d}\omega \\
&= \int_\Omega \log \left\|X(\theta^{n+m-1}\omega) \cdots X(\omega)\right\| \mathrm{d}\omega \\
&\le \int_\Omega \log \left\|X(\theta^{n+m-1}\omega) \cdots X(\theta^m\omega)\right\| \mathrm{d}\omega + \int_\Omega \log \left\|X(\theta^{m-1}\omega) \cdots X(\omega)\right\| \mathrm{d}\omega \\
&= \int_\Omega \log \|S_n(\theta^m\omega)\| \,\mathrm{d}\omega + \int_\Omega \log \|S_m(\omega)\| \,\mathrm{d}\omega \\
&= a_n + a_m
\end{aligned} \tag{1.2}$$

where the last line follows by the integrability condition on X along with the fact that θ is measure preserving.

Definition. We say that the sequence $(a_n)_{n=1}^\infty \subset \mathbb{R}$ is *subadditive* if $a_{n+m} \le a_n + a_m$ for each $n, m \in \mathbb{N}$. More generally, we say that a sequence of functions $\varphi_n : \Omega \to \mathbb{R}$ is *subadditive* if

$$\varphi_{n+m}(\omega) \le \varphi_n(\theta^m\omega) + \varphi_m(\omega). \tag{1.3}$$

The following lemma is straightforward.

1.3 Lemma. *If $(a_n)_{n=1}^{\infty}$ is a subadditive sequence, then $\lim_{n\to\infty} \frac{a_n}{n} = \inf_{n\geq 1} \frac{a_n}{n}$.*

In particular, implies that the limit

$$\lim_{n\to\infty} \frac{1}{n} \int_{\Omega} \log \|S_n(\omega)\| \,\mathrm{d}\omega$$

always exists. Moreover, if we set $\varphi_n(\omega) = \log \|S_n(\omega)\|$, we observed in (1.2) that the sequence of functions φ_n is subadditive. Thus Theorem 1.2 is a consequence of the following more general result.

1.4 THE SUBADDITIVE ERGODIC THEOREM

Throughout the statement and the proof, note that many inequalities implicity hold for μ-a.e. $\omega \in \Omega$.

1.4 Theorem (Kingman's Subadditive Ergodic). *Let $\varphi_n : \Omega \to \mathbb{R}$ be a subadditive sequence with $\varphi_1^+ \in L^1(\Omega, \mu)$. Then the limit $\varphi(\omega) := \lim_{n\to\infty} \frac{\varphi_n(\omega)}{n}$ exists for almost every $\omega \in \Omega$. Moreover, φ is θ-invariant and*

$$\int_{\Omega} \varphi(\omega) \,\mathrm{d}\omega = \lim_{n\to\infty} \frac{1}{n} \int_{\Omega} \varphi_n(\omega) \,\mathrm{d}\omega =: L.$$

Set

$$\varphi_-(\omega) = \liminf_{n\to\infty} \frac{\varphi_n(\omega)}{n} \qquad\qquad \varphi_+(\omega) = \limsup_{n\to\infty} \frac{\varphi_n(\omega)}{n}.$$

We first observe that φ_- (and by an analgous argument φ_+) is θ-invariant. By the subadditivity assumption (1.3) with $m = 1$,

$$\varphi_-(\omega) \leq \liminf_{n\to\infty} \frac{\varphi_n(\theta\omega) + \varphi(\omega)}{n+1} = \varphi_-(\theta\omega)$$

so with $X_a = \{\omega \in \Omega : \varphi_-(\omega) \geq a\}$ for any $a \in \overline{\mathbb{R}}$, we have $\theta^{-1}(X_a) \supset X_a$. But θ is measure-preserving, so this can forces $\mu(\theta^{-1}(X_a) \setminus X_a) = 0$, i.e. φ_- is θ-invariant.

Our general idea in this proof is to first establish the result for the function φ_-, and then use subadditivity and a repeat application of this result to obtain the result for φ_+. To subdivide the proof more clearly, we will first prove two intermediate lemmas.

1.5 Lemma. *We have $\int_{\Omega} \varphi_-(\omega) \,\mathrm{d}\omega = L$.*

Proof. Let $\epsilon > 0$ be arbitrary. For $k \in \mathbb{N}$, define

$$E_k = \Big\{\omega \in \Omega : \frac{\varphi_j(\omega)}{j} \leq \varphi_-(\omega) + \epsilon \text{ for some } j = 1, \dots, k\Big\}.$$

Note that $E_k \subset E_{k+1}$ and $\bigcup_k E_k = \omega$. Now set

$$\psi_k(\omega) = \begin{cases} \varphi_-(\omega) + \epsilon & : \omega \in E_k \\ \varphi_1(\omega) & : \omega \in E_k^c \end{cases}$$

Observe that $\psi_k \geq \varphi_-(\omega) + \epsilon$ by definition of E_k.

First, we will prove that for all $n > k$ and almost every $\omega \in \Omega$,

$$\varphi_n(\omega) \le \sum_{i=0}^{n-k-1} \psi_k(\theta^i\omega) + \sum_{i=n-k}^{n-1} \max\{\psi_k, \varphi_1\}(\theta^i\omega). \tag{1.4}$$

Since φ_- is θ-invariant, we may assume that $\varphi_-(\theta^n\omega) = \varphi_-(\omega)$ for all n.

We will inductively define a sequence $m_0 \le n_1 < m_1 \le n_2 < \cdots$ as follows. Let $m_0 = 0$. Inductively, let $n_j \ge m_{j-1}$ be the minimal integer such that $\theta^{n_j}\omega \in E_k$ (if it exists). By definition of E_k, there exists m_j such that $1 \le m_j - n_j \le k$ and

$$\varphi_{m_j-n_j}(\theta^{n_j}\omega) \le (m_j - n_j)(\varphi_-(\theta^{n_j}\omega) + \epsilon). \tag{1.5}$$

Let ℓ be maximal such that $m_\ell \le n$. By subadditivity, inductively applying the inequality

$$\varphi_i(\omega) \le \varphi_1(\theta^i\omega) + \varphi_{i-1}(\omega)$$

if $i \ne m_j$ for some j and the inequality

$$\varphi_{m_j}(\omega) \le \varphi_{n_j}(\omega) + \varphi_{m_j-n_j}(\theta^{n_j}\omega),$$

we obtain

$$\varphi_n(\omega) \le \sum_{i\in I} \varphi_1(\theta^i\omega) + \sum_{j=1}^{\ell} \varphi_{m_j-n_j}(\theta^{n_j}\omega) \tag{1.6}$$

where $I = \bigcup_{j=0}^{\ell-1}[m_j, n_{j+1}) \cup [m_\ell, n)$. Now if $i \in I$ with $i < n_{\ell+1}$, we have

$$\varphi_1(\theta^i\omega) = \psi_k(\theta^i\omega)$$

since $\theta^i\omega \notin E_k^c$. Since $\varphi_-(\theta^n\omega) = \varphi_-(\omega)$ and $\psi_k \ge \varphi_- + \epsilon$ by definition, by (1.5),

$$\varphi_{m_j-n_j}(\theta^{n_j}\omega) \le \sum_{i=n_j}^{m_j-1} (\varphi_-(\theta^i\omega) + \epsilon) \le \sum_{i=n_j}^{m_j-1} \psi_k(\theta^i\omega).$$

Thus (1.4) follows by (1.6) and the fact that $n - n_\ell < k$.

Now, suppose $\varphi_n/n \ge -C$ for some fixed constant $C > 0$. The upper bound follows by Fatou's Lemma:

$$\int_\Omega \varphi_-(\omega)\,\mathrm{d}\omega \le \liminf_{n\to\infty} \frac{1}{n}\int_\Omega \varphi_n(\omega)\,\mathrm{d}\omega = L.$$

To get the lower bound, by (1.4),

$$\frac{1}{n}\int_\Omega \varphi_n(\omega)\,\mathrm{d}\omega \le \frac{n-k}{n}\int_\Omega \psi_k(\omega)\,\mathrm{d}\omega + \frac{k}{n}\int_\Omega \max\{\psi_k, \varphi_1\}(\omega)\,\mathrm{d}\omega.$$

Thus taking the limit as n goes to infinity, we have

$$L \le \int_\Omega \psi_k(\omega)\,\mathrm{d}\omega$$

which holds for any $k \in \mathbb{N}$. Moreover, $\lim_{k\to\infty} \psi_k = \varphi_- + \epsilon$, so that $L \le \int_\Omega \varphi_-(\omega)\,\mathrm{d}\omega + \epsilon$. But $\epsilon > 0$ was arbitrary, giving the desired equality.

More generally, let $\varphi_n^{(C)} = \max\{\varphi_n, -Cn\}$ and $\varphi_-^{(C)} = \max\{\varphi_-, -C\}$. Then by the Monotone Convergence Theorem,

$$\begin{aligned}\int_\Omega \varphi_-(\omega)\,\mathrm{d}\omega = \inf_C \int_\Omega \varphi_-^{(C)}(\omega)\,\mathrm{d}\omega = \inf_C \inf_n \int_\Omega \frac{\varphi_n^{(C)}(\omega)}{n}\,\mathrm{d}\omega \\ = \inf_n \int_\Omega \frac{\varphi_n(\omega)}{n}\,\mathrm{d}\omega = L\end{aligned}$$

as required. □

1.6 Lemma. *We have* $\limsup_{n\to\infty} \frac{\varphi_{nk}(\omega)}{nk} = \varphi_+(\omega)$ *pointwise a.e.*

Proof. The upper bound follows since by subadditivity and invariance of φ_+,

$$\begin{aligned}\limsup_{n\to\infty} \frac{\varphi_{nk}(\omega)}{n} &\le \sum_{j=0}^{k-1} \limsup_{n\to\infty} \frac{\varphi_n(\theta^{nj}\omega)}{n} \\ &= k\varphi_+(\omega).\end{aligned}$$

Conversely, given $n \in \mathbb{N}$, write $n = kq_n + r_n$ where $r_n \in \{1, \dots, k\}$. By subadditivity,

$$\varphi_n(\omega) \le \varphi_{kq_n}(\omega) + \varphi_{r_n}(\theta^{kq_n}\omega) \le \varphi_{kq_n}(\omega) + \psi(\theta^{kq_n}\omega)$$

where $\psi = \max\{\varphi_1^+, \dots, \varphi_k^+\}$. By assumption, $\psi \in L^1$. Below, we will show that

$$\lim_{n\to\infty} \frac{\psi \circ \theta^{kq_n}}{q_n} = 0 \tag{1.7}$$

pointwise a.e. Assuming this result, we have

$$\limsup_{n\to\infty} \frac{\varphi_n}{n} \le \limsup_{n\to\infty} \frac{1}{n}\varphi_{kq_n} = \frac{1}{k}\limsup_{n\to\infty} \frac{1}{q_n}\varphi_{kq_n} \le \frac{1}{k}\limsup_{n\to\infty} \frac{\varphi_{nk}}{n}.$$

Let's prove (1.7). Let $\epsilon > 0$ be arbitrary. We first observe that

$$\begin{aligned}\sum_{n=1}^\infty \mu\big(\{\omega\in\Omega : |\psi(\theta^n\omega)| \ge \epsilon n\}\big) &= \sum_{n=1}^\infty \mu\big(\{\omega\in\Omega : |\psi(\omega)| \ge \epsilon n\}\big) \\ &= \sum_{n=1}^\infty \sum_{k=n}^\infty \mu\big(\{\omega\in\Omega : k\epsilon \le |\psi(\omega)| < (k+1)\epsilon\}\big) \\ &= \sum_{k=1}^\infty k\mu\big(\{\omega\in\Omega : k\epsilon \le |\psi(\omega)| < (k+1)\omega\}\big) \\ &\le \int_\Omega \frac{|\psi(\omega)|}{\epsilon}\,\mathrm{d}\omega < \infty.\end{aligned}$$

Thus the result follows by the Borel-Cantelli Lemma. □

Proof (of Theorem 1.4). We are now in position to complete the proof. As before, we first assume that $\varphi_n/n \ge -C$ for some fixed $C > 0$. Set

$$\phi_k = -\sum_{j=0}^{n-1} \varphi_k \circ \theta^{kj}.$$

By definition, $\phi_{n+m} = \phi_m + \phi_n \circ \theta^{km}$ and $\phi_1 = -\varphi_k \le Ck$, so $\phi_1^+ \in L^1(\Omega, \mu)$. Let $\phi_- = \liminf_{n\to\infty} \frac{\phi_n}{n} \, \mathrm{d}\omega$. Then by Lemma 1.5 and the fact that μ is θ-invariant,

$$\int_\Omega \phi_-(\omega) \, \mathrm{d}\omega = \lim_{n\to\infty} \int_\Omega \frac{\phi_n(\omega)}{n} \, \mathrm{d}\omega = \int_\Omega \varphi_k(\omega) \, \mathrm{d}\omega.$$

Now by the subadditivity assumption and Lemma 1.6,

$$-\phi_- = \limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi_k \circ \theta^{kj} \ge \limsup_{n\to\infty} \frac{\varphi_{kn}}{n} = k\varphi_+.$$

Combining the last two equations, we obtain

$$\int_\Omega \varphi_+ \, \mathrm{d}\omega \le -\frac{1}{k} \int_\Omega \phi_- \, \mathrm{d}\omega \le \frac{1}{k} \int_\Omega \varphi_k(\omega) \, \mathrm{d}\omega.$$

But this holds for any $k \in \mathbb{N}$, so that $\int_\Omega \varphi_+ \, \mathrm{d}\omega \le L$.

In general, as in the proof of Lemma 1.5, set $\varphi_n^{(C)} = \max\{\varphi_n, -Cn\}$ and $\varphi_\pm^{(C)} = \max\{\varphi_\pm, -C\}$. We just showed that $\int_\Omega -\varphi_-^{(C)} \, \mathrm{d}\omega = \int_\Omega \varphi_+^{(C)}(\omega) \, \mathrm{d}\omega$. But $\varphi_-^{(C)} \le \varphi_+^{(C)}$, so that $\varphi_-^{(C)} = \varphi_+^{(C)}$. Thus the result follows by the Monotone Convergence Theorem. □

Remark. This result generalizes Theorem 1.1 since, using the notation from that theorem, the function $\varphi_n(\omega) = \sum_{i=0}^{n-1} f(T^i\omega)$ is subadditive (since it is additive) and by invariance of T,

$$\int_\Omega f(T^i\omega) \, \mathrm{d}\omega = f(T^i\omega).$$

In fact, Theorem 1.1 follows directly from Lemma 1.5 since both $(\varphi_n)_{n=1}^\infty$ and $(-\varphi_n)_{n=1}^\infty$ are subadditive sequences of functions.

The argument in Lemma 1.6 can be interpreted as a "stability result" for subadditive sequences, which we then use to get control over φ_+ in the general case.

2 POSITIVITY OF LYAPUNOV EXPONENTS

2.1 NON-EXISTENCE OF INVARIANT MEASURES

In this section, we specialize slightly to the following setting. Let ν be a probability measure on $\mathrm{GL}_d(\mathbb{C})$. Then we take $\Omega = \mathrm{GL}_d(\mathbb{C})^\mathbb{N}$ equipped with the left-shift map σ, and μ is the infinite product $\mu = \nu^{\otimes\mathbb{N}}$. In this setting, the measure-preserving dynamical system (Ω, μ, σ) is ergodic. Since the Lyapunov exponent λ is σ-invariant, λ is constant μ-a.e. Abusing notation, we denote this constant by λ.

What can we say about the almost-everywhere value of λ? Of course, $\lambda \ge 0$, so we naturally specialize to distinguishing the cases where $\lambda = 0$ or $\lambda > 0$. There are some simple natural settings where $\lambda = 0$. Denote by G_ν the closure of the subgroup generated by the matrices in $\operatorname{supp} \nu$.

1. If G_ν is compact, then the norms of any random product is uniformly bounded above by a constant, so in fact $\lambda = 0$ everywhere.

2. If G_ν is contained in an abelian subgroup, then

$$\lambda = \int_\Omega \|M\| \,\mathrm{d}\nu(M)$$

which may be zero depending on the choice of ν.

3. If μ is the atomic measure with support

$$\operatorname{supp}\mu = \left\{ \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\},$$

then $\lambda = 0$ almost everywhere. More generally, if μ consists of a uniformly chosen random rational rotation, along with a uniformly chosen contraction or dilation depending on the angle, then $\lambda = 0$ almost everywhere.

Our main theorem in this section is that the three examples above are essentially the only ways in which we can have $\lambda = 0$ almost everywhere. We first state the following definition.

Definition. We say that a subgroup G of $\mathrm{GL}_d(\mathbb{C})$ is *totally irreducible* if there is no finite union of proper subspaces of $\mathbb{C}^d$ which are G-invariant.

We first observe a basic consquence of total irreducibility and non-compactness. Here, $\mathbb{P}(\mathbb{C}^d)$ is $d-1$-dimensional projective space, equipped with the projection map $[\cdot] : \mathbb{C}^d \setminus \{0\} \to \mathbb{P}(\mathbb{C}^d)$ taking $x \in \mathbb{C}^d$ to the equivalence class

$$[x] := \{y \in \mathbb{C}^d : y = \lambda x, \lambda \in \mathbb{C} \setminus \{0\}\}.$$

Of course, $M_d(\mathbb{C})$ acts naturally on $\mathbb{P}(\mathbb{C}^d)$ as well by $M \cdot [x] = [Mx]$ for $M \in M_d(\mathbb{C})$.

2.1 Lemma. *Suppose G_ν is totally irreducible and non-compact. Then there is no G_ν-invariant probability measure on $\mathbb{P}(\mathbb{C}^d)$.*

Proof. Suppose for contradiction μ is a G_ν-invariant probability measure on $\mathbb{P}(\mathbb{C}^d)$. Since G_ν is unbounded, there exists a sequences of matrices $(g_n)_{n=1}^\infty \subset G_\nu$ such that $\lim_{n\to\infty} \|g_n\| = \infty$. Let $u_n = g_n / \|g_n\|$, so that $\lim_{n\to\infty} \det u_n = 0$. Since $\|u_n\| = 1$ for each n, passing to a subsequence if necessary, we may assume

$$\lim_{n\to\infty} u_n = u \in M_d(\mathbb{C})$$

entry-wise. Write

$$V = [\ker u] \subset \mathbb{P}(\mathbb{C}^d) \text{ and } W = [\operatorname{im} u] \subset \mathbb{P}(\mathbb{C}^d)$$

and since $\|u\| = 1$ so that $u \neq 0$ and $\det u = 0$, V and W are proper projective subspaces of $\mathbb{P}(\mathbb{C}^d)$.

Decompose $\mu = \mu_1 + \mu_2$ where $\mu_1 = \mu|_V$ and $\mu_2 = \mu|_{V^c}$. If $[x] \in V^c$, then $g_n \cdot [x] = u_n \cdot [x]$ so $\lim_{n\to\infty} g_n \cdot [x] = u \cdot [x]$. Thus

$$\lim_{n\to\infty} (g_n)_* \mu = \lim_{n\to\infty} (g_n)_* \mu_1 + u_* \mu_2$$

where we recall $(g_n)_*\mu_1$ denotes the pushforward of μ_1 by g_n (and similarly for $u_*\mu_2$). Now, passing to a subsequence and using compactness of $\mathbb{P}(\mathbb{C}^d)$, we may assume

$$\lim_{n\to\infty} (g_n)_* \mu_1 = \mu_1^\infty \text{ and } \lim_{n\to\infty} g_n V = V^\infty$$

for some probability measure μ_1^∞ on $\mathbb{P}(\mathbb{C}^d)$ and projective subspace V^∞.

Since $\operatorname{supp}(g_n)_*\mu_1 \subset g_n V$, we have $\operatorname{supp}\mu_1^\infty \subset V^\infty$, and $\operatorname{supp} u_*\mu_2 \subset W$. Since each $g_n V$ is a proper projective subspace of $\mathbb{P}(\mathbb{C}^d)$, so is V^∞. But now $\operatorname{supp}\mu \subset V^\infty \cup W$ so that $\mu(V^\infty \cup W) = 1$. Let $F \subset V^\infty \cup W$ be the smallest finite union of proper projective subspaces such that $\mu(F) = 1$. Thus by invariance of μ under G_ν, we have $gF = F$ for any $g \in G_\nu$, contradicting the assumption of total irreducibility. ☐

2.2 POSITIVITY OF LYAPUNOV EXPONENTS

We now prove our main result on positivity of Lyapunov exponents. For simplicity, we will assume that $G_\nu \subset \mathrm{SL}_d(\mathbb{C})$.

2.2 Theorem (Furstenberg). *Suppose G_ν is totally irreducible and non-compact. Then*

$$\lambda(\omega) > 0$$

for μ-a.e. $\omega \in \Omega$.

It is meaningful to obtain the following operator-theoretic formulation of Theorem 2.2; this perspective will also reappear in **TODO: cite Furstenberg measures section**. Consider the Hilbert space

$$\mathcal{H} = L^2(\mathbb{C}^d) = \big\{ f : \mathbb{C}^d \to \mathbb{C} : \int_{\mathbb{C}^d} |f(x)|^2 \,\mathrm{d}m(x) < \infty \big\}.$$

Then a matrix $g \in \mathrm{SL}_d(\mathbb{C})$ acting on $\mathbb{C}^d$ induces a natural action $\pi(g) : \mathcal{H} \to \mathcal{H}$ by $\pi(g)f(x) = f(g^{-1}x)$, so we may define the operator $P_\nu : \mathcal{H} \to \mathcal{H}$ given by the Gelfand-Pettis integral

$$P_\nu f = \int_{G_\nu} \pi(g) f \,\mathrm{d}\nu(g).$$

Of course, by definition of the Gelfand-Pettis integral, $P_\nu f(x) = \int_{G_\nu(\mathbb{C})} f(g^{-1}x)\,\mathrm{d}\nu(g)$. One can interpret the operator P_ν as applying a random transformation of f by a matrix g chosen according to the probability measure ν. We first list some basic properties of the action π and the operator P_ν.

2.3 Lemma. *(i) $\|\pi(g)f\|_2 = \|f\|_2$ for any $g \in \mathrm{SL}_d(\mathbb{C})$*
(ii) $\|P_\nu\| \leq 1$
(iii) $P_{\nu_1}P_{\nu_2} = P_{\nu_1 * \nu_2}$
(iv) $P_\nu^* = P_{\nu^*}$ where $\mathrm{d}\nu^*(g) = \mathrm{d}\nu(g^{-1})$

Proof. Part (i) follows by a change of variables since $|\det g| = 1$, and parts (iii) and (iv) follow directly from the definition of P_ν.

It remains to see (ii). By Jensen's inequality and an application of Fubini's Theorem,

$$\begin{aligned}
\|P_\nu f\|_2^2 &= \int_{\mathbb{C}^d} \left| \int_{G_\nu} \pi(g) f(x)\,\mathrm{d}\nu(g) \right|^2 \mathrm{d}m(x) \\
&\le \int_{\mathbb{C}^d} \int_{G_\nu} |\pi(g) f(x)|^2 \,\mathrm{d}\nu(g)\,\mathrm{d}m(x) \\
&= \int_{G_\nu} \int_{\mathbb{C}^d} |\pi(g) f(x)|^2 \,\mathrm{d}m\,\mathrm{d}\nu(g) \\
&= \int_{G_\nu} \|\pi(g) f\|_2 \,\mathrm{d}\nu(g) \\
&= \|f\|_2
\end{aligned}$$

where the last line follows by (i) and the fact that ν is a probability measure. □

Our proof approach is bound $\|P_\nu\|$ and then relate Theorem 2.2 to the operator P_ν. We first need a standard result from analysis in Hilbert spaces, which we include for completeness.

2.4 Lemma. *Let P be a self-adjoint operator on a Hilbert space $\mathcal{H}$. Then*

$$\|P\| = \sup_{\|f\|=1} |\langle Pf, f\rangle|.$$

Proof. Set

$$\sup_{\|f\|=1} |\langle Pf, f\rangle| =: \alpha$$

Of course, we always have $\alpha \le \|P\|$ by the Cauchy-Schwarz inequality. Conversely, it suffices to show that $|\langle Pf, g\rangle| \le \alpha$ for any f, g with $\|f\| = \|g\| = 1$ (since taking $g = Pf / \|Pf\|$, $|\langle Pf, g\rangle| = \|P\|$). It suffices to prove the case where $\langle Pf, g\rangle \in \mathbb{R}$. Then since P is self-adjoint,

$$\langle Pf, g\rangle = \frac{\langle P(f+g), f+g\rangle - \langle P(f-g), f-g\rangle}{4}$$

so that

$$|\langle Pf, g\rangle| \le \alpha \cdot \frac{\|f+g\|^2 + \|f-g\|^2}{4} = \alpha$$

by the parallelogram identity. □

2.5 Lemma. *If $\|P_\nu\| = 1$, then there is a G_ν-invariant probability measure $\overline{\mu}$ on $\mathbb{P}(\mathbb{C}^d)$.*

Proof. We have that $P_\nu P_{\nu^*} = P_{\nu * \nu^*}$ is self adjoint, and $\|P_\nu P_\nu^*\| = \|P_\nu\|^2$ (this is just the C^* identity). Thus $\|P_\nu\| = 1$ if and only if $\|P_{\nu*\nu^*}\| = 1$, so without loss of generality, we may assume that P_ν is self-adjoint.

Suppose for contradiction $\|P_\nu\| = 1$. By Lemma 2.4, get $(f_n)_{n=1}^\infty \subset \mathcal{H}$ with $\|f_n\|_2 = 1$ and $\lim_{n\to\infty} |\langle Pf_n, f_n\rangle| = 1$. Since $|\langle P_\nu f_n, f_n\rangle| \le \langle P_\nu |f_n|, |f_n|\rangle \le 1$, we may assume $f_n \ge 0$. Now, by continuity and linearity of the inner product along with properties of the Gelfand-Pettis integral,

$$\lim_{n\to\infty} \int_{G_\nu} \langle \pi(g) f_n, f_n\rangle \,\mathrm{d}\nu(g) = \lim_{n\to\infty} \langle P_\nu f_n, f_n\rangle.$$

Since $\langle \pi(g) f_n, f_n \rangle \le 1$, we have $\lim_{n\to\infty} \langle \pi(g) f_n, f_n \rangle = 1$ ν-a.e.

In particular, for ν-a.e. g, we have since $f_n \ge 0$

$$\begin{aligned}
\lim_{n\to\infty} \|\pi(g) f_n - f_n\|_2^2 &= \lim_{n\to\infty} \left(\|\pi(g) f_n\|_2^2 + \|f_n\|_2^2 - 2\langle \pi(g) f_n, f_n \rangle \right) \\
&= 2 - 2 \lim_{n\to\infty} \langle \pi(g) f_n, f_n \rangle = 0
\end{aligned}$$

so by Cauchy-Schwarz,

$$\begin{aligned}
\lim_{n\to\infty} \left\| \pi(g) f_n^2 - f_n^2 \right\|_2 &\le \lim_{n\to\infty} \|\pi(g) f_n - f_n\|_2 \cdot \|\pi(g) f_n + f_n\|_2 \\
&\le 2 \lim_{n\to\infty} \|\pi(g) f_n - f_n\|_2 = 0. \qquad (2.1)
\end{aligned}$$

Now, consider the probability measures $\mathrm{d}\mu_n = f_n^2 \,\mathrm{d}m$ on $\mathbb{C}^d$, and let $\overline{\mu}_n$ denote the pushforward onto the projective space $\mathbb{P}(\mathbb{C}^d)$. Since $\mathbb{P}(\mathbb{C}^d)$ is compact, $\{\overline{\mu}_n\}_{n=1}^\infty$ has a weak*-accumulation point $\overline{\mu}$, and by (2.1), $\overline{\mu}$ is G_ν-invariant. □

We now finish the proof by relating the operator P_ν with Lyapunov exponents.

Proof (of Theorem 2.2). By Theorem 1.2, it suffices to show that

$$\lambda(\omega) = \lim_{n\to\infty} \frac{1}{n} \int \log \|S_n(\omega)\| \,\mathrm{d}\mu(\omega) = \lim_{n\to\infty} \frac{1}{n} \int_{\mathrm{GL}_d(\mathbb{C})} \log \|g\| \,\mathrm{d}\nu^{*n}(g) > 0$$

for μ-a.e. $\omega \in \Omega$.

Combining Lemma 2.1 and Lemma 2.5, we observe that $\gamma := \|P_\nu\| < 1$. Let

$$\begin{aligned}
f(x) &= \min\{C, |x|^{-\alpha}\} \\
K &= \{x : 1 \le |x| \le 2\}
\end{aligned}$$

where α is chosen so that $f \in L^2(\mathbb{C}^d)$ and $C > 0$ is a constant to be determined below. We then have

$$\begin{aligned}
\limsup_{n\to\infty} |\langle P_{\nu^{*n}} f, \mathbf{1}_K \rangle|^{1/n} &= \limsup_{n\to\infty} |\langle P_\nu^n f, \mathbf{1}_k \rangle|^{1/n} \\
&\le \limsup_{n\to\infty} \|P_\nu^n\|^{1/n} \cdot \|f\|_2^{1/n} \cdot \|\mathbf{1}_K\|_2^{1/n} \le \gamma.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\langle P_{\nu^{*n}} f, \mathbf{1}_K \rangle &= \int_{1\le|x|\le 2} \int_{\mathrm{GL}_d(\mathbb{C})} \min\{C, \left\|g^{-1}x\right\|^{-\alpha} \,\mathrm{d}\nu^{*n}(g) \,\mathrm{d}m(x) \\
&\ge \int_{1\le|x|\le 2} \int_{\mathrm{GL}_d(\mathbb{C})} \min\{C, \left\|g^{-1}\right\|^{-\alpha} \cdot \|x\|^{-\alpha}\} \,\mathrm{d}\nu^{*n}(g) \,\mathrm{d}m(x) \\
&\ge C_0 \int_{\mathrm{GL}_d(\mathbb{C})} \min\{C, \left\|g^{-1}\right\|^{-\alpha}\} \,\mathrm{d}\nu^{*n}(g)
\end{aligned}$$

for some constant C_0 depending only on α. Since $\inf_{g\in \mathrm{SL}_d(\mathbb{C})} \|g\| > 0$, we can take C sufficiently large so that $\min\{C, \left\|g^{-1}\right\|^{-\alpha}\} = \left\|g^{-1}\right\|^{-\alpha}$ for any $g \in \mathrm{SL}_d(\mathbb{C})$. We also use the fact that $\left\|g^{-1}\right\| \le C_0' \|g\|^{d-1}$, which follows by the adjoint formula for the matrix (since the entries in the adjoint are degree $d-1$ polynomial functions of the entries of g, and $|\deg g| = 1$). Thus there is some constant $C_1 > 0$ such that

$$\langle P_{\nu^{*n}} f, \mathbf{1}_K \rangle \geq C_1 \int_{\mathrm{GL}_d(\mathbb{C})} \|g\|^{-\alpha(d-1)} \, \mathrm{d}\nu^{*n}(g).$$

Thus taking logarithms, applying Jensen's inequality, and rearranging, we have

$$\int_{\mathrm{GL}_d(\mathbb{C})} \log \|g\| \, \mathrm{d}\nu^{*n}(g) \geq \frac{\log C_1}{\alpha(d-1)} - \frac{1}{\alpha(d-1)} \log \langle P_{\nu^*} f, \mathbf{1}_K \rangle$$

and therefore

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \int_{\mathrm{GL}_d(\mathbb{C})} \log \|g\| \, \mathrm{d}\nu^{*n}(g) &= -\frac{1}{\alpha(d-1)} \limsup_{n\to\infty} \frac{1}{n} \log \langle P_{\nu^{*n}} f, \mathbf{1}_K \rangle \\ &\geq -\frac{1}{\alpha(d-1)} \log \gamma > 0 \end{aligned}$$

as required. □

3 OSELEDEČ MULTIPLICATIVE ERGODIC THEOREM

3.1 SINGULAR VALUE DECOMPOSITIONS AND THE EXTERIOR ALGEBRA

If $M \in M_d(\mathbb{C})$ is any matrix, we can write $M = U\Sigma V^*$ where

$$\Sigma = \begin{pmatrix} \rho_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \rho_d \end{pmatrix} \text{ with } \rho_1 \geq \cdots \geq \rho_d \geq 0,$$

and U, V are unitary matrices. We refer to the this as the *singular value decomposition* of M, and the values $\rho_1, \ldots, \rho_d$ are the *singular values* of M. Note that if $M \in M_d(\mathbb{R})$, the matrices U and V can be chosen to be real-valued (so that they are orthogonal). Here, the singular values are the eigenvalues of $\sqrt{M^*M}$ which, by the continuous functional calculus, are the square roots of the eigenvalues of M^*M (which is self-adjoint and therefore has a real and positive spectrum). A standard exercise shows that $\|M\|_{op} = \rho_1$.

Recall that $X : \Omega \to \mathrm{GL}_d(\mathbb{C})$ is a matrix-valued function on a measure-preserving dynamical system (Ω, μ, θ), and

$$S_n(\omega) = X(\theta^{n-1}\omega) \cdots X(\omega).$$

In this section, we generally want to answer the following two questions:

(i) What is the exponential growth rate of the singular values of the random products $S_n(\omega)$?

(ii) What is the exponential growth rate of $\|S_n(\omega)v\|$ for some fixed starting vector $v \in \mathbb{C}^d$?

Of course, since $\|M\|_{op} = \rho_1$, (i) is a generalization of the discussion in Section 1.

In order to approach these questions, we want to convert statements about singular values into statements about norms of linear operators on some larger vector space. A natural way to do this is through the exterior algebra.

Given a vector space W, the k^{th} *exterior power* $\bigwedge^k W$ is the unique vector space satisfying the following universal property. If W' is any other vector space and $T : W^k \to W'$ is an alternating multilinear map (i.e. T is multilinear and $T(v_1, \ldots, v_k) = 0$ whenever $\{v_1, \ldots, v_k\}$ is linearly dependent), then there exists a unique linear map ϕ such that the following diagram commutes:

$$\begin{tikzcd} W^k \arrow[r, "\wedge^k"] \arrow[dr, "T"'] & \bigwedge^k W \arrow[d, "\phi"] \\ & W' \end{tikzcd}$$

In practice, we may define $\bigwedge^k W$ as the quotient of the k^{th} tensor product $W^{\otimes k}$ by the subspace generated by tensors of the form $v_1 \otimes \cdots \otimes v_k$ where $\{v_1, \ldots, v_k\}$ is linearly dependent in W. We denote the equivalence class of $[v_1 \otimes \cdots \otimes v_k]$ by $v_1 \wedge \cdots \wedge v_k$, and we have a natural wedge map $\wedge^k : W^k \to \bigwedge^k W$ given by $(v_1, \ldots, v_k) \mapsto v_1 \wedge \cdots \wedge v_k$. The wedge map induces a map $\wedge^k : \mathrm{Hom}(W) \to \mathrm{Hom}(\bigwedge^k W)$ by

$$\wedge^k M(v_1 \wedge \cdots \wedge v_k) = M(v_1) \wedge \cdots \wedge M(v_k).$$

Note that if W is d-dimensional, then

$$\wedge^d M(v_1 \wedge \cdots \wedge v_d) = (\det M) v_1 \wedge \cdots \wedge v_d.$$

We define an inner product on $\bigwedge^k W$ by

$$\langle v_1 \wedge \cdots \wedge v_k, w_1 \wedge \cdots \wedge w_k \rangle = \det(M).$$

where $M_{i,j} = \langle v_i, w_j \rangle$ and extend it to the whole space. Let $\{e_1, \ldots, e_d\}$ be an orthonormal basis for W consisting of eigenvectors of M^*M. Then one can show that $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \le i_1 < \cdots < i_k \le d\}$ is an orthonormal basis for $\bigwedge^k W$. Moreover, directly by definition, $\wedge^k(M^*) = (\wedge^k M)^*$. Thus

$$(\wedge^k M)^*(\wedge^k M)(e_{i_1} \wedge \cdots \wedge e_{i_k}) = \rho_{i_1}^2 \cdots \rho_{i_k}^2 e_{i_1} \wedge \cdots e_{i_k}$$

so $\wedge^k M^*M$ has eigenvectors $e_{i_1} \wedge \cdots \wedge e_{i_k}$ with corresponding eigenvalues $\rho_{i_1}^2 \cdots \rho_{i_k}^2$. In particular, $\left\| \wedge^k M \right\|_{op} = \rho_1 \cdots \rho_k$.

With this in mind, we may define the Lyapunov exponents $\lambda_1(\omega), \ldots, \lambda_d(\omega)$ inductively by the rule

$$\lambda_1(\omega) + \cdots + \lambda_k(\omega) = \lim_{n\to\infty} \frac{\log \left\| \wedge^k S_n(\omega) \right\|}{n}$$

for each $1 \le k \le d$. Of course, $\lambda_1(\omega) = \lambda(\omega)$ where $\lambda(\omega)$ is the Lyapunov exponent defined in (1.1). Note that these limits exist μ-a.e. by Theorem 1.2. The following result now follows immediately from the discussion above.

3.1 Theorem (Oseledeč Multiplicative Ergodic I). *Let $\rho_1^{(n)}(\omega) \ge \cdots \ge \rho_d^{(n)}(\omega) \ge 0$ be the singular values of $S_n(\omega)$. Then for μ-a.e. ω and all $j \in \{1, \ldots, d\}$,*

$$\lim_{n\to\infty} \frac{\log \rho_j^{(n)}(\omega)}{n} = \lambda_j(\omega).$$

3.2 GROWTH RATES OF SINGULAR VALUES

Fix $0 = \tau_{s+1} < \tau_s < \cdots < \tau_1 = d$. A *flag of type* τ is a sequence of subspaces $\{0\} = V_{s+1} \supset V_s \supset \cdots \supset V_1 = \mathbb{C}^d$ such that $\dim V_i = \tau_i$. Let $\mathcal{F}(\tau)$ denote the space of flags of type τ.

We can define a metric on $\mathcal{F}(\tau)$ as follows. Fix $\sigma_1, \ldots, \sigma_s$ where $\sigma_i \neq \sigma_j$ for $i \neq j$ and some $h > 0$. Suppose we are given flags $V^{(j)} = \{V_{s+1}^{(j)} \supset \cdots \supset V_1^{(j)}\}$ for $j = 1, 2$. Then for each $1 \leq i \leq s$ there are spaces $U_i^{(j)}$ so that

$$V_i^{(j)} = U_i^{(j)} \perp V_{i+1}^{(j)}$$

Where $A \perp B$ denotes the direct sum of orthogonal subspaces A and B. In particular, $\mathbb{C}^d = U_1^{(j)} \perp \cdots \perp U_s^{(j)}$. We may now define

$$d(V^{(1)}, V^{(2)}) = \max_{\substack{i \neq j, \|x\| = \|y\| = 1 \\ x \in U_i^{(1)}, y \in U_j^{(2)}}} |\langle x, y \rangle|^{h \cdot |\sigma_i - \sigma_j|^{-1}}. \tag{3.1}$$

Intuitively, the function d measures the degree of orthogonality between the flags $V^{(1)}$ and $V^{(2)}$, along with an exponential scaling factor. If $U_i^{(1)}$ and $U_j^{(2)}$ are orthogonal, then $|\langle x, y \rangle| = 0$ for any $x \in U_i^{(1)}$ and $y \in U_j^{(2)}$.

3.2 Lemma. *Suppose $h^{-1}|\sigma_i - \sigma_j| \geq s - 1$ for all $i \neq j$. Then d defines a metric on $\mathcal{F}(\tau)$, and $\mathcal{F}(\tau)$ is complete with respect to this metric.*

Proof. **TODO: write** □

We can now state and prove our main result in this section.

3.3 Theorem (Oseledeč Multiplicative Ergodic II). *Suppose $\log^+ \|X\| \in L^1(\Omega, \mu)$. Then for a.e. $\omega \in \Omega$,*

$$\lim_{n \to \infty} \|S_n^*(\omega) S_n(\omega)\|^{1/2n} =: \Lambda(\omega)$$

exists, and the eigenvalues of $\Lambda(\omega)$ are $e^{\lambda_i(\omega)}$.

Fix ω for which Theorem 3.1 holds, and set $X_n(\omega) = X(\theta^{n-1}\omega)$. Arguing similarly to (1.7), we may assume that

$$\limsup_{n \to \infty} \frac{\log \left\| X_n^{\pm 1} \right\|}{n} \leq 0. \tag{3.2}$$

Let $\alpha_1 > \cdots > \alpha_s$ denote the sorted distinct values of the $\{\lambda_i(\omega) : 1 \leq i \leq d\}$.

Let $\epsilon > 0$ be small and for each $1 \leq i \leq s$, let $U_i^{(s)}$ denote the subspace generated by the eigenvectors corresponding to the eigenvalues ρ of $(S_n^* S_n)^{1/2}$ satisfying

$$\left| \frac{\log \rho}{n} - \alpha_i \right| < \epsilon. \tag{3.3}$$

Let $P_i^{(n)}$ denote the orthogonal projection onto $U_i^{(n)}$.

We will need the following lemma, which heuristically states that the projections maps are, in the limit, pairwise orthogonal.

3.4 Lemma. *For all $i \neq j$ and all n sufficiently large,*

$$\left\| P_i^{(n)} P_j^{(n+1)} \right\| = \left\| P_j^{(n+1)} P_i^{(n)} \right\| \leq e^{(-|\alpha_i - \alpha_j| + 4\epsilon)n}.$$

Proof. Let $x \in \mathbb{C}^d$, $y = P_i^{(n)} x \in U_i^{(n)}$, and $z = P_j^{(n+1)} y$. First, suppose $i > j$. Since $y \in U_i^{(n)}$, applying (3.3), we have

$$\|S_{n+1} y\| \leq \|X_{n+1}\| \cdot \|S_n y\| \leq \|X_{n+1}\| e^{\alpha_i + \epsilon} \|y\|. \tag{3.4}$$

Since the spaces $U_k^{(n+1)}$ are invariant under the matrix $S_{n+1}^* S_{n+1}$, and $U_{k_1}^{(n+1)}$ is orthogonal to $U_{k_2}^{(n+1)}$ for any $k_1 \neq k_2$, we have

$$\langle S_{n+1} z, S_{n+1}(y - z) \rangle = \langle S_{n+1} z, S_{n+1} P_j^{(n+1)} (y - z) \rangle = 0.$$

Thus by the Pythagoras rule and again applying (3.3),

$$\begin{aligned} \|S_{n+1} y\| &= \sqrt{\|S_{n+1} z\|^2 + \|S_{n+1}(y - z)\|^2} \geq \|S_{n+1} z\| \\ &\geq e^{(\alpha_j - \epsilon)(n+1)} \|z\| \geq e^{(\alpha_j - 2\epsilon)n} \|z\|. \end{aligned}$$

Rearranging and applying (3.4), we have $\|z\| \leq \|X_{n+1}\| e^{(\alpha_i - \alpha_j + 3\epsilon)n} \|y\|$. Moreover, by (3.3), $\|X_{n+1}\| \leq e^{\epsilon n}$ for n sufficiently large. Thus

$$\begin{aligned} \left\| P_j^{(n+1)} P_i^{(n)} x \right\| &\leq e^{(\alpha_i - \alpha_j + 4\epsilon)n} \left\| P_i^{(n)} x \right\| \\ &\leq e^{(\alpha_i - \alpha_j + 4\epsilon)n} \|x\|. \end{aligned}$$

Otherwise let $i < j$. Then for $x \in \mathbb{C}^d$, $y = P_j^{(n+1)} x$, and $z = P_i^{(n)} y$, we have by (3.2) and (3.3) that

$$\begin{aligned} \|S_n y\| &= \left\| X_{n+1}^{-1} S_{n+1} y \right\| \leq \left\| X_{n+1}^{-1} \right\| \|S_{n+1} y\| \\ &\leq \left\| X_{n+1}^{-1} \right\| e^{(\alpha_j + \epsilon)(n+1)} \|y\| \leq e^{(\alpha_j + 3\epsilon)n} \|y\|. \end{aligned}$$

and for the lower bound, as argued above,

$$\|S_n y\| = \sqrt{\|S_n z\|^2 + \|S_n(y - z)\|^2} \geq \|S_n z\| \geq e^{\alpha_i - \epsilon} \|z\|.$$

Thus $\|z\| \leq e^{(\alpha_j - \alpha_i + 4\epsilon)n} \|y\|$ and it follows that $\left\| P_i^{(n)} P_j^{(n+1)} \right\| \leq e^{(\alpha_j - \alpha_i + 4\epsilon)n}$. □

Proof (of Theorem 3.3). Consider the sequence of flags $V^{(n)} = \{V_{s_n+1}^{(n)} \subset \cdots \subset V_1^{(n)}\}$ where

$$V_i = \bigoplus_{k=i}^{s_n} U_k^{(n)}.$$

Note that for n sufficiently large, $V^{(n)} \in \mathcal{F}(\tau)$ by (3.3) and the definition of $U_i^{(n)}$. By properties of projections, the Cauchy-Schwarz inequality, and Lemma 3.4, we have

$$|\langle x, y \rangle| = |\langle P_i^{(n)} x, P_j^{(n+1)} y \rangle| = \langle x, P_i^{(n)} P_j^{(n+1)} y \rangle| \leq e^{(-|\alpha_i - \alpha_j| + 4\epsilon)n}. \tag{3.5}$$

Fix a metric d on $\mathcal{F}(\tau)$ as in (3.1) by taking $\sigma_i = \alpha_i$ and h sufficiently small from Lemma 3.2. Thus by (3.5), we have

$$\begin{aligned} d(V^{(n)}, V^{(n+1)}) &\leq \max_{i \neq h} e^{(-|\alpha_i - \alpha_j| + 4\epsilon)n \cdot h|\alpha_i - \alpha_j|^{-1}} \\ &\leq e^{-(1-\epsilon')hn} \end{aligned}$$

for some small $\epsilon' > 0$ depending on ϵ. Thus for ϵ sufficiently small, $\{V^{(n)}\}_{n=1}^{\infty}$ is Cauchy so $\lim_{n\to\infty} V^{(n)} = V^{\infty} \in \mathcal{F}(\tau)$, and

$$d(V^{(n)}, V^{\infty}) \leq Ce^{-(1-\epsilon')hn} \tag{3.6}$$

for some fixed constant $C > 0$. Let $\rho_1^{(n)}, \ldots, \rho_{k_n}^{(n)}$ denote the distinct eigenvalues of $(S_n^* S_n)^{1/2}$ and for each $1 \leq i \leq s$ let $I_i^{(n)} \subset \{1, \ldots, k_n\}$ denote the indices corresponding to α_i. Again, for ϵ sufficiently small and n sufficiently large, $\bigcup_{i=1}^{s} I_i^{(n)} = \{1, \ldots, k_n\}$ where the union is disjoint. Since $S_n^* S_n$ is self-adjoint, by the spectral theorem,

$$(S_n^* S_n)^{1/2n} = \sum_{j=1}^{k_n} (\rho_j^{(n)})^{1/n} \cdot P_j^{(n)}$$

where $\lim_{n\to\infty} (\rho_j^{(n)})^{1/n} = \alpha_i$ for any $j \in I_i^{(n)}$ by Theorem 3.1, and $\lim_{n\to\infty} \sum_{j \in I_i^{(n)}} Q_j^{(n)} = P_i$ by (3.6). Thus

$$\lim_{n\to\infty} (S_n^* S_n)^{1/2n} = \sum_{i=1}^{s} \alpha_i P_i$$

and the desired result follows directly. □

3.3 Random walks of vectors

Using similar arguments as above, we can also determine the asymptotic growth rate of norms of images $S_n(\omega)x$.

3.5 Theorem (Oseledeč Multiplicative Ergodic III). *For a.e. $\omega \in \Omega$, there exists a flag $V(\omega) = \{V_{s+1} \supset \cdots \supset V_1\}$ such that for all $x \in V_{i+1} \setminus V_i$,*

$$\lim_{n\to\infty} \frac{\log \|S_n(\omega)x\|}{n} = \alpha_i(\omega)$$

where $\alpha_1(\omega) > \cdots > \alpha_s(\omega)$ are the distinct values of the Lyapunov exponents of ω.

Proof. Let $V(\omega)$ be the flag V^{∞} from the proof of Theorem 3.3. □