Ergodic and Algebraic Properties of Transfer Operators for Products of Random Matrices



Fan Wang

St Anne's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2018

Acknowledgements

I would like to thank my supervisor Professor David Steinsaltz for introducing this intriguing maths problem to me and for providing his constant support. My DPhil studies were not very smooth, in which I experienced culture shock, self denial and reconstruction, "failures" in the transfer and the confirmation of status, and worried enormously about my ignorance. However, in addition to providing academic guidance, my supervisor was always there encouraging me, telling me that I had done a good job, and giving me an appropriate amount of freedom to decide my own research plan. I cannot be thankful enough for his help.

I would like to thank the examiners, Professor Christina Goldschmidt and Dr Ian Morris, who helped me considerably improve this dissertation. I am quite talented in making various kinds of mistakes in proofs, so I really appreciate their patience and tolerance in pointing them out.

I would like to thank Dr Aditi Kar and Professor Nikolay Nikolov from the Mathematics Institute, who were willing to spare their precious time discussing maths problems with me. And I would like to thank those mathematicians (especially Professor Fuhai Zhu) who had inspired me by their passions in maths.

I would like to thank my landlady Dr Rekha Wadekar, who generously invited me to her family and treated me as a member when I was almost homeless at the final stage of my DPhil life.

I also would like to thank the colour of the trees outside Radcliffe Science Library during the sunset, as well as Georgios.

Thanks are also due to National University of Defense Technology, which provided me financial support for three and half years.

Contents

1 Introduction **9**

2 Products of Random Matrices **19**
2.1 Lyapunov Exponents 20
2.1.1 Existence of the Lyapunov Exponent 20
2.1.2 Matrix Actions and Invariant Measures 21
2.1.3 Existence of an Invariant Measure 23
2.1.4 Furstenberg-Kesten Theorem 27
2.2 Lyapunov Spectrum 30
2.3 Invariant Measures on Projective Space 33
2.3.1 Irreducibility 34
2.3.2 Contracting Property 37
2.3.3 Integral Formula for the Lyapunov Exponent 43
2.3.4 Contraction on the Projective Space 46

3 Other Perspectives **51**
3.1 The Algebraic Perspective 51
3.1.1 Decomposition of Semigroups 52
3.1.2 Decomposition of Invariant Measures 54
3.2 The Geometric Perspective 58

4 Markovian Products of Matrices **65**
4.1 Shift Spaces 65
4.2 Random Dynamical Systems 67
4.3 Markovian Products of Matrices 70

5 Pollicott's Algorithm **79**
5.1 Projective Actions 81
5.2 Transfer Operators 90
5.3 Parametrised Transfer Operators 99

5.4 Nuclear Operators . . . 110
5.5 Nuclearity of Transfer Operators . . . 116
5.6 Fredholm Theory . . . 120
5.7 Algorithm for the Lyapunov Exponent . . . 125
5.8 Modified Lemma . . . 128
5.9 Numerical Results . . . 131

6 Generalizations with Transfer Operators **133**
6.1 Markovian Products of Positive Matrices . . . 133
6.1.1 Markovian Transfer Operators . . . 134
6.1.2 Parametrised Markovian Transfer Operators . . . 141
6.1.3 The Trace, the Determinant and the Algorithm . . . 148
6.2 Markovian Products of General Matrices . . . 150
6.2.1 Function Space $G^+_{\theta,\alpha}$ . . . 151
6.2.2 Markovian Transfer Operators on $G^+_{\theta,\alpha}$ . . . 153
6.2.3 Spectral Properties . . . 161

7 Algebras of Transfer Operators **167**
7.1 Overview . . . 167
7.2 C*-Algebras . . . 170
7.2.1 Definition of C*-Algebras . . . 170
7.2.2 Gelfand Representation Theorem . . . 172
7.2.3 Positive Linear Functionals and *-Representations . . . 174
7.3 Von Neumann Algebras . . . 175
7.3.1 Definition of von Neumann Algebras . . . 175
7.3.2 Projections . . . 177
7.3.3 Standard Traces . . . 178
7.4 Type I Factors . . . 179
7.5 Abstract Traces and Type II_1 Factors . . . 181
7.6 Crossed Products . . . 182
7.7 Transfer Operator Algebras . . . 184
7.8 Determinant Function . . . 189
7.9 Discussion . . . 193

A Code **195**
A.1 Pollicott's Algorithm in R . . . 195

Bibliography **201**

Index **207**

Chapter 1

Introduction

Noncommutative Random Variables

In elementary probability theory, the random variables we study are mostly commutative (with respect to summation). For example, when we use the experiments of "flipping coins" to verify the law of large numbers, a random variable ξ taking values 0 or 1 is usually assumed to represent the result (head or tail) of one single experiment. After n independent trials we will get n samples $\xi_1, \ldots, \xi_n$, whereas the sum of the these samples equals the overall number of tails, denoted by $f(n) = \sum_{i=1}^{n} \xi_i$. The law of large numbers studies the asymptotic behaviour of $f(n)/n$ when $n \to \infty$.

The value of $f(n)$ here does not depend on the order of those experiments. Expressed formally, if σ is an element in the permutation group S_n of order n, we have

$$\sum_{i=1}^{n} \xi_i = \sum_{i=1}^{n} \xi_{\sigma(i)}.$$

Essentially, it is due to the fact that these random variables are commutative.

In many high-dimensional problems, on the other hand, we have to deal with non-commutative random variables. Let us start with an example from demography.

Assume there are d geographically isolated populations with migration. At time $t \in \mathbb{N}$, the individual numbers at these sites can be written as a d-dimension vector

$$N^{(t)} = (N_1^{(t)}, \cdots, N_d^{(t)}).$$

Assume in each specific type of environment, the populations would behave according to the a certain rule which only depends on the type of the environment. Say during time t to $t+1$, the environment variable is ω_t, the individual number grows with a rate of $a_{ii}(\omega_t)$ at site i, and the

individuals move from site j to site i with a rate of $a_{ij}(\omega_t)$ $(1 \leq i, j \leq d)$. Now we have the following equation for the evolution of the system

$$N_i^{(t+1)} = a_{ii}(\omega_t)N_i^{(t)} + \sum_{j \neq i} a_{ij}(\omega_t)N_j.$$

Denote $A(\omega) = (a_{ij}(\omega))_{1 \leq i,j \leq d}$, then we can get the recursive formula

$$N^{(t+1)} = A(\omega_t)N^{(t)}$$

as well as the general formula

$$N^{(t+1)} = A(\omega_t)A(\omega_{t-1}) \cdots A(\omega_0)N^{(0)}.$$

Similarly to the problem of "flipping coins", we are still interested in the asymptotic behaviour of $N^{(t)}$ under normalisation. For example, $(\|N^{(t)}\|_1/\|N^{(0)}\|_1)^{1/t}$ represents the average growth rate of individual numbers in all d sites up to time t, where $\|\cdot\|_1$ is the ℓ_1-norm of vectors. But what is different in this model is, as matrices are non-commutative with respect to multiplication in general, the value of $N^{(t)}$ depends on the order of the sequence $\omega_1, \omega_2, \ldots$.

Definition of the Lyapunov Exponent

Now let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which we can assume $A_1(\omega), A_2(\omega), \ldots$ is a stationary process taking values in $M_d(\mathbb{R})$, the set of $d \times d$ real matrices, and denote $S_n(\omega) = A_n(\omega)A_{n-1}(\omega) \cdots A_1(\omega)$.

Note here by Kolmogorov's extension theorem we can without loss of generality identify Ω with the sequences of environment variables, i.e. $\omega = (\omega_0, \omega_1, \ldots)$. Let $\sigma : \Omega \to \Omega$ be the shift map, i.e. $(\sigma\omega)_i = \omega_{i+1}$ $(i \geq 0)$. Assume A only depends on the current environment, and write $A(\omega) = A(\omega_0)$ when there is no ambiguity. Then we can set $A_i(\omega) = A(\sigma^i\omega)$ $(i \geq 1)$.

Then under the condition $\mathbb{E}[\max\{\log \|A(\omega)\|, 0\}] < \infty$, we can define

$$\gamma := \lim_{n \to \infty} \frac{1}{n}\mathbb{E}[\log \|S_n(\omega)\|]. \tag{1.1}$$

The existence of this limit (taking values in $\mathbb{R} \cup \{-\infty\}$) can be obtained by the fact that $a_n := \mathbb{E}[\log \|S_n\|]$ is sub-additive, i.e. $a_{m+n} \leq a_m + a_n$. Moreover, it can be shown that the value of the limit does not depend on the choice of the norm $\|\cdot\|$. This limit is usually called the Lyapunov exponent.

Similar to $(\|N^{(t)}\|_1/\|N^{(0)}\|_1)^{1/t}$ representing the average growth rate of the overall individual numbers up to time t, the Lyapunov exponent characterises the logarithm of it as time goes to infinity. This statement will be further confirmed in eq. (1.2).

For the rest of this chapter, we assume $A(\omega)$ to be a random invertible matrix defined on $(\Omega, \mathcal{F}, \mathbb{P})$, let μ be the measure of $A(\omega)$ on $GL(d, \mathbb{R})$, and $A_1(\omega), A_2(\omega), \dots$ be i.i.d. samples of $A(\omega)$. Denote $S_n(\omega) := A_n(\omega)A_{n-1}(\omega)\cdots A_1(\omega)$, $f^+ := \sup\{f, 0\}$ i.e., the positive part of the function f. $\mathbb{E}[\log^+ \|A(\omega)\|] < \infty$ is also called the finite expectation condition, and γ is used to denote the corresponding Lyapunov exponent defined in eq. (1.1). A matrix is said to be positive if all of its entries are positive.

Classical Theory

The asymptotic behaviour of non-commutative products was firstly studied in [2], in which Bellman focused on 2×2 positive matrices and described the asymptotic behaviour of the entries of random products.

Soon after that in 1960, Furstenberg and Kesten [19] proved that, under the finite expectation condition $\mathbb{E}[\log^+ \|A(\omega)\|] < \infty$, the limit

$$\lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)\|, \quad \mathbb{P}\text{-a.s. exists and equals } \gamma. \tag{1.2}$$

Moreover, Furstenberg also proved in [18] that, under the same finite expectation condition and, further more, some irreducibility condition,

$$\lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)x\|, \quad \mathbb{P}\text{-a.s. exists and equals } \gamma. \tag{1.3}$$

for all non-zero $x \in \mathbb{R}^d$.

Both of the two conclusions above can be regarded as the non-commutative version of the law of large numbers. In fact, if we set

$$x_0(\omega) = x, \quad x_n(\omega) := S_n(\omega)x/\|S_n(\omega)x\| \ (n \geq 1),$$

we can see

$$\begin{aligned}
\frac{1}{n}\log\|S_n(\omega)x\| &= \frac{1}{n}\log\left\|A_n(\omega)\frac{S_{n-1}(\omega)x}{\|S_{n-1}(\omega)x\|}\right\| + \frac{1}{n}\log\|S_{n-1}(\omega)x\| \\
&= \frac{1}{n}\left(\log\|A_n(\omega)x_{n-1}(\omega)\| + \cdots + \log\|A_2(\omega)x_1(\omega)\| + \log\|A_1(\omega)x_0(\omega)\|\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\log\|A_i(\omega)x_{i-1}(\omega)\|
\end{aligned}$$

If we denote by $\mathbb{R}P^{d-1}$ the real projective space induced by $\mathbb{R}^d$, and define

$$\Sigma : \Omega \times \mathbb{R}P^{d-1} \to \Omega \times \mathbb{R}P^{d-1}$$
$$(\omega, x) \mapsto \left(\sigma\omega, \frac{A(\omega)x}{\|A(\omega)x\|}\right)$$

and

$$F : \Omega \times \mathbb{R}P^{d-1} \to \mathbb{R}$$
$$(\omega, x) \mapsto \log \|A(\omega)x\|,$$

then

$$\frac{1}{n}\sum_{i=1}^{n} \log \|A_i(\omega)x_{i-1}(\omega)\| = \frac{1}{n}\sum_{i=1}^{n} F \circ \Sigma^i(\omega, x)$$

might remind us of Birkhoff's ergodic theorem, which in turn is a generalisation of the law of large numbers.

In 1983, Furstenberg and Kifer [20] gave the integral formula for the Lyapunov exponent. They proved that under the finite expectation condition and the irreducibility condition,

$$\gamma = \iint \log \|A(\omega)x\| \mathbb{P}(\mathrm{d}\omega)\nu(\mathrm{d}x), \tag{1.4}$$

where ν is a $\mathbb{P}$-invariant measure on the projective space $\mathbb{R}P^{d-1}$, i.e., for each bounded Borel function f on $\mathbb{R}P^{d-1}$, we have

$$\iint f\left(\frac{A(\omega)x}{\|A(\omega)x\|}\right) \mathbb{P}(\mathrm{d}\omega)\nu(\mathrm{d}x) = \int f(x)\nu(\mathrm{d}x).$$

When we deal with scalar random variables, the integral formula of expectations helps us compute the values of them. However, the "closed form" of the Lyapunov exponent here does not contribute to the evaluation, since the invariant measure ν cannot be computed explicitly in general [38]. Therefore, we have to find ways to computing the Lyapunov exponent numerically.

Numerical Methods

Common methods of computing the Lyapunov exponent include Monte-Carlo method, weak-disorder expansion [8], cycle-expansion method [37] and a recent approach proposed by Mark Pollicott [48] based on the properties of transfer operators.

Monte-Carlo method is the most general one among them all. In fact, it can be applied to any i.i.d. random matrix products (or even stationary random matrix products) because the essence of it is just sampling and applying eq. (1.1) or eq. (1.2). However, Monte-Carlo method has a slow rate of convergence in the sense that the approximation using n samples has only the accuracy of $O(n^{-1/2})$ [4, 35].

The weak-disorder expansion [8] considers possible values of $A(\omega)$ to be small perturbations of the average $\bar{A} = \mathbb{E}[A(\omega)]$, and write $A(\omega) = \bar{A} + \epsilon R(\omega)$ and $R_i(\omega) := R(\omega_i)$ for each $i \geq 1$. Then the product can be written as

$$\begin{aligned} S_n(\omega) &= A_n(\omega)A_{n-1}(\omega)\cdots A_1(\omega) \\ &= (\bar{A} + \epsilon R_n(\omega))(\bar{A} + \epsilon R_{n-1}(\omega))\cdots(\bar{A} + \epsilon R_1(\omega)) \end{aligned}$$

which can be seen as a polynomial with respect to ϵ with matrix coefficients. [8] proves that when $\bar{A}$ can be diagonalised and possesses a simple top eigenvalue, we can get an expansion of the Lyapunov exponent γ with respect to ϵ. Therefore the estimates can be obtained by truncating the expansion. For example, the first and the second estimates are

$$\begin{aligned} \gamma^{(1)} &= \log \lambda_1 + O(\epsilon), \\ \gamma^{(2)} &= \log \lambda_1 - \frac{\epsilon^2}{2\lambda_1^2}\mathbb{E}[R(\omega)_{1,1}^2] + O(\epsilon^3), \end{aligned}$$

where λ_1 is the top eigenvalue of $\bar{A}$, $R(\omega)_{1,1}$ is the $(1,1)$-entry of the matrix $R(\omega)$. We can easily see, the problems of this method are that, firstly it assumes specific diagonalisation property on the average matrix; secondly the accuracy of this method strongly depends on the value of ϵ, which cannot be small if the matrices are not close to their average; and finally the estimates depend on the choice of the basis with respect to which the matrices are represented, however the top Lyapunov exponent is invariant under change of basis.

The cycle-expansion method [37] focuses on the special case of random products of positive matrices. It begins from the observation that, by the Perron-Frobenius theorem, any positive matrix M possesses a top eigenvalue $\rho(A) > 0$, and for any positive matrices M_1, M_2, ρ satisfies $\rho(M_1M_2) = \rho(M_2M_1)$ as well as $\rho(M_1^n) = \rho(M_1)^n$ $(n \geq 1)$. More importantly,

$$\gamma = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\log \rho(S_n(\omega))].$$

Now if μ is a discrete measure, we can explicitly give the p-th approximation by

$$\gamma^{(p)} = \frac{1}{p}\mathbb{E}[\log \rho(S_p(\omega))]. \tag{1.5}$$

By the nice properties of ρ mentioned above, it is easier to compute this value as non-commutativity disappears. Moreover, it is proved in [37] that $|\gamma^{(p)} - \gamma| \sim O(e^{-\epsilon p})$ for some $\epsilon > 0$.

This rate of convergence is faster than the Monte-Carlo method. However, cycle-expansion can only be applied to the case when μ is a discrete measure on positive matrices.

Pollicott's Approach

Mark Pollicott proposed an algorithm in 2010 [48], applying Ruelle's theory of transfer operators and Grothendieck's classical work on nuclear operators to compute the Lyapunov exponent associated to the random products of positive matrices. It is pointed out that if μ is a discrete measure with the support $\{M_1, M_2, \ldots, M_m\}$, write $\mu = \sum_{i=1}^m p_i \delta_{M_i}$ with $\sum_{i=1}^m p_i = 1, p_i \geq 0, 1 \leq i \leq m$, then we can define a corresponding transfer operator $\mathcal{L}_t$, acting on some function space $\mathcal{G}$ on some subset of $\mathbb{R}P^{d-1}$ (we will specify how to choose $\mathcal{G}$ in later chapters), satisfying[1]

$$(\mathcal{L}_t f)(x) = \sum_{i=1}^m p_i e^{t\mathbb{E}[\log \|A(\omega)x\|]} f\left(\frac{M_i x}{\|M_i x\|}\right), \quad f \in \mathcal{G},$$

where t is a real perturbation parameter and $\mathbb{E}$ is the expectation with respect to ω. After choosing a proper function space $\mathcal{G}$, it can be proved that $\mathcal{L}_t$ possesses an isolated and simple top eigenvalue $\lambda(t)$, and that $\lambda'(0) = \gamma$. Therefore the problem of computing the Lyapunov exponent has been transformed into the problem of calculating the top eigenvalue of the transfer operator $\mathcal{L}_t$.

Next, since it can be proved that $\mathcal{L}_t$ is a trace-class operator, Pollicott defined the determinant function

$$d(z, t) := \det(\mathcal{I} - z\mathcal{L}_t),$$

and obtained the top eigenvalue of $\mathcal{L}_t$ by calculating the largest zero of $d(z, t)$. More specifically, by the determinant-trace formula, $d(z, t)$ can be written as

$$d(z, t) = \exp\left(-\sum_{n=1}^{\infty} \frac{\operatorname{tr}(\mathcal{L}_t^n)}{n} z^n\right), \tag{1.6}$$

and can be proved to be an analytic function with respect to z when $|t|$ is sufficiently small. We

[1]The following definition of transfer operators is given by Pollicott [48], and we will see in Chapter 5 that it is not an optimal choice.

now consider the Laurent expansion of $d(z,t)$,

$$d(z,t) = 1 + \sum_{i=1}^{\infty} a_i(t) z^i. \tag{1.7}$$

By $\lambda'(0) = \gamma$ and using the partial sum of the first p terms in eq. (1.7), we can get the the p-th estimate of the Lyapunov exponent

$$\gamma^{(p)} = \frac{\sum_{i=1}^{p} a_i'(0)}{\sum_{i=1}^{p} i a_i(0)}, \tag{1.8}$$

and it is shown that $|\gamma^{(p)} - \gamma| \sim O(e^{-\theta p^{1+1/d}})$ for some $\theta > 0$ [48][29].

A New Outline

Although Pollicott's algorithm can only be applied to i.i.d. products of positive random matrices, it provides a new outline for computing the Lyapunov exponent, which can be summarised as the following three steps:

(S-a) Convert the Lyapunov exponent problem for products of random matrices to the top eigenvalue problem of the corresponding transfer operator;

(S-b) Transform the top eigenvalue problem of the transfer operator to the problem of computing the largest zero of the determinant function;

(S-c) Prove that the determinant function is analytic and that we can get the estimates from the truncated polynomials.

As far as the author knows, finding an algorithm to efficiently and accurately compute the Lyapunov exponent for random products of general matrices is still an open problem. However, Pollicott's work, which raises awareness of the relationship of the Lyapunov exponent and the largest zero of the determinant function, suggests a possible generalisation for this approach. Correspondingly, there are three main problems which we would encounter:

(Q-a) Given a measure μ on $GL(d,\mathbb{R})$, can the corresponding transfer operator be defined? Can the problem of computing the Lyapunov exponent for μ be transformed into the eigenvalue problem of the transfer operator?

(Q-b) Can the determinant function of the transfer operator be defined? If so, is the largest zero of the determinant function still the top eigenvalue of the transfer operator?

(Q-c) Is the determinant function analytic? Can we use the truncated polynomials to obtain the estimates?

Main Results

There are three main results in this dissertation, namely Theorem 6.1.6, Theorem 6.2.17 and Theorem 7.8.4.

Theorem 6.1.6 is a direct generalisation of Pollicott's algorithm to Markovian products of positive matrices. In order to give an explicit algorithm, we construct a "matrix" of simple transfer operators, prove its spectral properties in Theorem 6.1.1 and Theorem 6.1.2, and give the trace formula in Theorem 6.1.5.

Theorem 6.2.17 aims at solving (Q-a) for products of general matrices. It is proved that, after choosing a proper function space $\mathcal{G}$, the positive matrix condition can be weakened to strongly irreducibility and contracting condition, and the i.i.d. products can be replaced by Markovian products, whereas we can still define a corresponding transfer operator given a measure μ on $GL(d,\mathbb{R})$, and this transfer operator has the same spectral properties. That is, the transfer operator $\mathcal{L}_t$ has an isolated and simple top eigenvalue $\lambda(t)$ and $\lambda'(0) = \gamma$.

However (Q-b) cannot be solved directly. This is due to the fact that, when μ is supported by positive matrices, $\mathcal{L}_t$ acting on some analytic function space is a trace-class operator, while for transfer operators associated to products of general matrices, the traces usually do not exist. We will see in Section 7.1 that, even if when $\mathcal{L}_t$ is induced by a singleton measure $\mu = \delta_M$ supported on a single non-negative matrix M, $\mathrm{tr}(\mathcal{L}_t)$ sometimes can not be defined. This means that we have to find other ways to replace the calculation of normal traces in the original problem.

We turn then to a more general definition for traces of operators [10][41][42]. Given a closed operator algebra $\mathcal{A}$ (under some specific topology), write $\mathcal{A}^+ := \{a^*a : a \in \mathcal{A}\}$ for the positive portion of $\mathcal{A}$. A linear functional φ is said to be tracial if $\varphi(a^*a) = \varphi(aa^*)$ for any $a \in \mathcal{A}$. An abstract trace is defined to be a positive linear functional which is tracial.

Then we can define the corresponding abstract determinant function similar to eq. (1.6),

$$\Delta(z,t) := \exp\left(-\sum_{n=1}^{\infty} \frac{\phi(\mathcal{L}_t^n)}{n} z^n\right).$$

We will prove that, if ϕ satisfies some specific conditions, $\Delta(z,t)$ is analytic on $U = \{w \in \mathbb{C} : |w| > \lambda(t)\}$, where $\lambda(t)$ is the top eigenvalue of $\mathcal{L}_t$, and moreover,

$$\lim_{U \ni z \to \lambda(t)} \Delta(z,t) = 0,$$

i.e., the top eigenvalue of $\mathcal{L}_t$ can be regarded as a zero of the abstract determinant function and we can still have the Laurent expansion of $\Delta(z,t)$ in order to obtain the estimates as in eq. (1.8).

Before stating these the main results, we will state the classical results on the Lyapunov exponents developed by Furstenberg, Kesten, Kifer, Guivarc'h, Raugi, Le Page etc. in Chapter 2.

In Chapter 3, some algebraic and geometric perspectives on the problem of random matrix products are provided. The algebraic perspective is based on Mukherjea's work [40], which draws a connection between our problem and the completely simple semigroup theory. The geometric perspective section aims at providing a concrete visualisation for readers who might be new to the theory of random matrix products. No prior knowledge of the classical theory is required for this section.

In Chapter 4, we generalise most of the results in Chapter 2 to the Markovian setting, using the language of random dynamical systems.

Pollicott's efficient algorithm will be introduced in Chapter 5 together with brief introductions to nuclear operators and Fredholm theory. We also provide a complete proof of Perron-Frobenius-Ruelle theorem for transfer operators acting on the space of analytic functions with continuous extension on the boundary in Theorem 5.2.4 and Theorem 5.3.2.

Transfer operators for Makrovian products of positive matrices are investigated in Section 6.1. The spectral properties for standard and parametrised transfer operators are given by Theorem 6.1.1 and Theorem 6.1.2 respectively. The trace formula can be found in Theorem 6.1.5 as well as the algorithm in Theorem 6.1.6.

We will redefine the transfer operators and the function spaces on which they are acting in a more general setting in Section 6.2. The main results — spectral properties, are given by Theorem 6.2.17.

An introduction to the theory of operator algebras can be found in the beginning of Chapter 7. Then we prove that we can identify the operator algebra generated by transfer operators with a specific operator algebra called the crossed product, and provide an analogous definition for a determinant function based on the trace theory in von Neumann algebras. In the end, we give the third main result in Theorem 7.8.4.

Chapter 2

Products of Random Matrices

In this chapter, we summarise some classical results on products of i.i.d. random matrices. The problem is usually stated as follows:

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which we can define a sequence of i.i.d. samples $A_1(\omega), A_2(\omega), \ldots$ chosen according to some probability measure μ on $d \times d$ (invertible) matrices. We want to understand the limiting behaviour of S_n as $n \to \infty$. More specifically, we want to know the existence and values (if any exist) of the limits

$$\lim_{n\to\infty} \frac{1}{n} \log \|A_n(\omega)A_{n-1}(\omega)\cdots A_1(\omega)\|, \tag{2.1}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \log \frac{\|A_n(\omega)A_{n-1}(\omega)\cdots A_1(\omega)x\|}{\|x\|}, \quad 0 \neq x \in \mathbb{R}^d. \tag{2.2}$$

We remark here that the matrix norm and the vector norm $\|\cdot\|$ in problem (2.1) (2.2) can be chosen arbitrarily since all norms on finite-dimensional spaces are equivalent. Most of the results in this chapter are originally due to Furstenberg, Kesten and Kifer in [19] [18] [20], and then to later developments of Guivarc'h and Raugi [24], and Le Page [35]. We mainly focus on those results which would be useful for later sections. Readers who might be interested in more detail on classical results can also refer to [4].

Denote by $M_d(\mathbb{R})$ the set of $d \times d$ real matrices, and $GL(d, \mathbb{R})$ the subset of invertible matrices in $M_d(\mathbb{R})$. We write $S_n(\omega) := A_n(\omega)A_{n-1}(\omega)\cdots A_1(\omega)$ as the product of the first n elements by multiplying on the left.

2.1 Lyapunov Exponents

2.1.1 Existence of the Lyapunov Exponent

We start with a lemma which is commonly seen as an exercise in a calculus textbook. This lemma is attributed to Michael Fekete [12]. We provide a proof here because the same basic idea of this proof will appear again in the matrix setting in Theorem 2.1.13.

Lemma 2.1.1. *Let $(a_n)_{n\geq 1}$ be a sequence in $\mathbb{R}\cup\{-\infty\}$ satisfying $a_{m+n}\leq a_m+a_n$ for each $m,n\geq 1$, then the limit $\lim_{n\to\infty} a_n/n$ exists.*

Proof. If there exists an m such that $a_m=-\infty$, then the sequence $(a_n)_{n\geq 1}$ will be $-\infty$ for all $n\geq m$, therefore the limit exists and is equal to $-\infty$. Otherwise, we want to prove that

$$\lim_{n\to\infty}\frac{a_n}{n}=\inf_n\frac{a_n}{n}.$$

Fix an integer $k\geq 1$. For any $n\geq 1$, we can write $n=kq(n)+r(n)$ for some integer $q(n)\geq 0$ and $0\leq r(n)\leq k-1$. By subadditivity, we have

$$\frac{a_n}{n}=\frac{a_{kq(n)+r(n)}}{n}\leq\frac{q(n)a_k}{n}+\frac{a_{r(n)}}{n}.$$

Note that

$$0\leq\frac{a_{r(n)}}{n}\leq\frac{\max_{0\leq i\leq k-1}a_i}{n}\to 0\quad(n\to\infty)$$

and that $\lim_{n\to\infty} q(n)a_k/n=a_k/k$ since

$$\left|\frac{q(n)a_k}{n}-\frac{a_k}{k}\right|=\frac{r(n)|a_k|}{nk}\leq\frac{|a_k|}{n}\to 0\quad(n\to\infty),$$

we have for any integer $k\geq 1$

$$\overline{\lim_{n\to\infty}}\frac{a_n}{n}\leq\frac{a_k}{k},$$

which gives

$$\overline{\lim_{n\to\infty}}\frac{a_n}{n}\leq\inf_k\frac{a_k}{k}\leq\underline{\lim_{k\to\infty}}\frac{a_k}{k}.$$

Therefore the result follows. □

The inequality condition described in Lemma 2.1.1 is usually called the **subadditive inequality**. This is useful in our problem since if we assume we have a stationary process $\{A_n(\omega) : n \geq 1\}$ with values in $d \times d$ matrices, define $S_j^i(\omega) := A_j(\omega) \cdots A_i(\omega)$, for integers $i < j$, and let $S_n(\omega) := S_n^1(\omega)$ for $n \geq 1$. Then for $m, n \geq 1$

$$\log \|S_{m+n}(\omega)\| \leq \log \|S_{m+n}^{n+1}(\omega)\| + \log \|S_n(\omega)\|.$$

As $S_{m+n}^{n+1}(\omega)$ has the same distribution as $S_m(\omega)$, thus if we denote $a_n = \mathbb{E}[\log \|S_n(\omega)\|]$, we have $a_{m+n} \leq a_m + a_n$. Now as long as $a_1 \neq \infty$, we can define $\lim_n a_n/n$ and have the following Definition 2.1.2 of the Lyapunov exponent. Here for a real function f, we denote by $f^+(x) = \sup\{0, f(x)\}$ the positive part of $f(x)$, and $f^-(x) = \sup\{0, -f(x)\}$.

Definition 2.1.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which we can define $\{A_n(\omega) : n \geq 1\}$ to be a stationary process taking values from $d \times d$ matrices and satisfying $\mathbb{E}[\log^+ \|A_1(\omega)\|] < \infty$, then the limit

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}[\log \|A_n(\omega) \cdots A_1(\omega)\|] \tag{2.3}$$

exists and takes value in $\mathbb{R} \cup \{-\infty\}$ by the discussion above. We call it the **Lyapunov exponent**.

Remark 2.1.3. The condition $\mathbb{E}[\log^+ \|A_1(\omega)\|] < \infty$ is usually called the **finite expectation condition**. It is equivalent to $a_1 \neq \infty$.

The limit defined in (2.3) sometimes is also called the top Lyapunov exponent, as we will see definitions for other Lyapunov exponents in Section 2.2.

2.1.2 Matrix Actions and Invariant Measures

In 1960, Furstenberg and Kesten [19] studied the asymptotic behaviour of $S_n(\omega)$ and proved the following Theorem 2.1.13. This can be a corollary of Kingman's subadditive ergodic theorem (which was discovered 13 years later after Furstenberg's work, see e.g. [33][11]), but here we follow the proof given in [4], which is a simplification of the original proof of Furstenberg and Kesten, and provides an intuition for how to characterise an invariant measure. We start with the definitions of matrix actions and invariant measures.

Definition 2.1.4. We shall denote by $\mathbb{R}P^{d-1}$ the **real projective space** defined by $(\mathbb{R}^d \backslash \{0\})/\sim$, where for any two non-zero vectors u and v, we define the equivalence relation $u \sim v$ if and only if there exists a real number k such that $u = kv$. For a non-zero vector $u \in \mathbb{R}^d$, we denote by $\bar{u}$ the equivalent class containing u.

We say a semigroup S endowed with some topology is a **topological semigroup** if the multiplication map $S \times S \to S, (s_1, s_2) \mapsto s_1 s_2$ is continuous.

Given a topological semigroup S and a topological space X, a **semigroup action** φ of S on X is a continuous function $\varphi : S \times X \to X$ satisfying: $\varphi(ts, x) = \varphi(t, \varphi(s, x))$ for any $s, t \in S$ and $\varphi(e, x) = x$ if S contains an identity e. When there is no ambiguity, we sometimes write $s.x$ instead of $\varphi(s, x)$ for short.

There are two types of action commonly defined for a matrix semigroup S. The first one is its action on $M_d(\mathbb{R})$ by $(A, B) \mapsto AB$, where $A \in S, B \in M_d(\mathbb{R})$, and AB denotes the usual multiplication of matrices A and B. The second is its action on $\mathbb{R}^d$ by $(A, x) \mapsto Ax$ for $A \in S, x \in \mathbb{R}^d$, and Ax denotes the usual multiplication of the matrix A and the vector x. We will call these two **natural actions** of matrices.

If $S \subseteq GL(d, \mathbb{R})$, we can define its **action on matrices of norm one** by

$$(A, B) \mapsto A \cdot B := \frac{AB}{\|AB\|}, \quad A \in S,\ \|B\| = 1,$$

as well as its **action on the real projective space** $\mathbb{R}P^{d-1}$ by

$$(A, x) \mapsto A \cdot x := \frac{Ax}{\|Ax\|}, \quad A \in S,\ x \in \mathbb{R}P^{d-1}.$$

We also call these two **normalised actions** of matrices. Note, as A is invertible, neither $\|AB\|$ nor $\|Ax\|$ can be 0.

Definition 2.1.5. Let S be a topological semigroup acting on a topological space X by $\varphi : S \times X \to X$. Let μ be a Borel probability measure on S and ν on X. Then we can define a probability measure $\mu *_\varphi \nu$ on X satisfying

$$\int_X f(x)\, \mathrm{d}(\mu *_\varphi \nu)(x) = \int_X \int_G f(\varphi(g, x))\, \mathrm{d}\mu(g)\, \mathrm{d}\nu(x),$$

for any bounded Borel function f on X. Here $\mu *_\varphi \nu$ is called the **convolution** of μ and ν with respect to the action φ.

In particular, when φ is a natural action of matrices, we denote by $\mu * \nu$ the convolution of μ and ν; and when φ is a normalised action, we write $\mu \dot{*} \nu$.

If $\mu *_\varphi \nu = \nu$, we say ν is a **μ-invariant measure** with respect to the action φ. When the action φ is specified without ambiguity, we will just say ν is a μ-invariant measure for short.

When we consider the action of a topological semigroup S on itself and assume μ is a probability measure on S, we can inductively define $\mu_\varphi^i := \mu_\varphi^{i-1} *_\varphi \mu$ for each $i \geq 2$. And

particularly, when S is the matrix semigroup, we use μ^i and $\mu^{\cdot i}$ to denote the powers of μ when φ is the natural action and the normalised action respectively.

2.1.3 Existence of an Invariant Measure

We do not yet know when a μ-invariant measure exists for a given μ, but the procedure introduced by Proposition 2.1.8 will produce one, provided we can find a limit point of a certain sequence in $\mathcal{P}(X)$, the set of Borel probability measures on X. In order to do so, we must impose some topological properties on X.

If X is compact and Hausdorff, and if we denote by $\mathcal{C}(X)$ the set of continuous functions on X, then by the Riesz representation theorem, the continuous dual of $\mathcal{C}(X)$ (under the topology of uniform convergence) can be identified with the space of regular Borel measures $\mathcal{M}(X)$ on X.

Moreover, by the Banach-Alaoglu theorem, the closed unit ball $\mathcal{B}(X)$ of $\mathcal{M}(X)$ given by

$$\mathcal{B}(X) = \left\{\mu \in \mathcal{M}(X) : \sup_{\substack{f\in\mathcal{C}(X);\\ \|f\|_\infty\leq 1}} \left|\int f \,\mathrm{d}\mu\right| \leq 1\right\}$$

is compact in the weak*-topology, where the weak*-topology is defined as the weakest topology that makes all the linear maps

$$\begin{aligned}\mathcal{L}_f : \mathcal{M}(X) &\to \mathbb{R}\\ \mu &\mapsto \int f \,\mathrm{d}\mu,\end{aligned}$$

induced by $f \in \mathcal{C}(X)$, continuous.

It is proved in [45] that,

Theorem 2.1.6. *X is compact and metrisable if and only if $\mathcal{P}(X)$, the space of probability measures defined on X, is compact and metrisable with respect to the weak*-topology.*

Since for a metric space, compactness is equivalent to sequential compactness, we have

Corollary 2.1.7. *Let X be a compact and metrisable topological space, then $\mathcal{P}(X)$ is sequentially compact with respect to the weak*-topology.*

Now with the sequential compactness of $\mathcal{P}(X)$, we are ready to introduce the following procedure. This is also called the Krylov-Bogolyubov procedure. (See e.g. [1] Section 1.5 for a more detailed introduction.)

Proposition 2.1.8. *Assume S is a topological semigroup acting on a compact and metrisable space X by $\varphi : S \times X \to X$. Let m be a probability measure on X and μ on S, then a μ-invariant measure on X exists, and each limit point of $\{n^{-1}\sum_{i=1}^{n}\mu^i *_\varphi m : n \geq 1\}$ in the weak*-topology is a μ-invariant measure on X.*

Proof. By Corollary 2.1.7, as X is compact and metrisable, the set $\mathcal{P}(X)$ of probability measures on X is sequentially compact in the weak*-topology.

Let $\nu_n = n^{-1}\sum_{i=1}^{n}\mu^i *_\varphi m$ be a sequence in $\mathcal{P}(X)$, there exists a subsequence ν_{n_i} converging to a probability measure, say ν, then

$$\mu *_\varphi \nu_n = \frac{1}{n}\sum_{i=2}^{n+1}\mu^i *_\varphi m = \nu_n + \frac{1}{n}(\mu^{n+1} *_\varphi m - \mu *_\varphi m). \tag{2.4}$$

Notice that, if for any $f \in \mathcal{C}(X), m \in \mathcal{P}(X)$ write $m(f)$ for $\int f \,\mathrm{d}m$,

$$\frac{1}{n}\left|(\mu^{n+1} *_\varphi m - \mu *_\varphi m)(f)\right| \leq \frac{2}{n}\|f\|_\infty \to 0 \ (n \to \infty),$$

$$\left|(\mu *_\varphi \nu_n - \mu *_\varphi \nu)(f)\right| = \left|(\nu_n - \nu)\left(\int f(\varphi(s,\cdot))\,\mathrm{d}\mu(s)\right)\right| \to 0 \ (n \to \infty).$$

Let $n \to \infty$ along the subsequence in (2.4), we get $\mu *_\varphi \nu = \nu$. □

Remark 2.1.9. When X is only locally compact, what we can do is consider the set of continuous functions with compact supports, denoting by $\mathcal{C}_c(X)$. We still have $\mathcal{C}_c(X)^* = \mathcal{M}(X)$, but $\mathcal{P}(X)$ is not necessarily closed in $\mathcal{M}(X)$ in the weak*-topology. In this case, the concept of tightness is usually introduced. Namely, for $\Gamma \subset \mathcal{P}(X)$, we say Γ is **tight** if for any $\epsilon > 0$ there exists a compact subset $K_\epsilon \subset X$ such that whenever $\mu \in \Gamma, \mu(K_\epsilon) > 1 - \epsilon$.

Intuitively, tightness is intended to prevent the limit points of sequences in Γ from lying outside of $\mathcal{P}(X)$. Whenever $\Gamma \subset \mathcal{P}(X)$ is tight, the weak*-closure of Γ is weak*-compact in $\mathcal{P}(X)$.

However, in this thesis, we mainly consider matrix actions on the projective spaces or other compact spaces induced by normalization. Readers interested in the existence of invariant measures for actions on locally compact spaces can find more information in [45][27].

With the existence of an invariant measure with respect to a probability measure on invertible matrices, we give the following Proposition 2.1.10 for later use.

Proposition 2.1.10. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and $\{A_n(\omega)\}_{n\geq 1}$ be a sequence of i.i.d random variables with a common distribution μ on $GL(d,\mathbb{R})$. Let ν be a μ-invariant measure on $\mathbb{R}P^{d-1}$. If*

$$\iint |\log \|Ax\|| \, \mathrm{d}\mu(A) \, \mathrm{d}\nu(x) < \infty,$$

then

$$\lim_{n\to\infty} \frac{1}{n} \log \|A_n(\omega) \cdots A_1(\omega)x\|$$

exists for $\mathbb{P}\otimes\nu$-almost all (ω, x) and in $L^1(\mathbb{P}\otimes\nu)$. Moreover, if we denote the limit by $\gamma(\omega, x)$, then

$$\iint \gamma(\omega, x) \, \mathrm{d}\mathbb{P}(\omega)\nu(x) = \iint \log \|Ax\| \, \mathrm{d}\mu(A) \, \mathrm{d}\nu(x).$$

Proof. Denote the collection of Borel sets in $GL(d,\mathbb{R})$ by $\mathcal{B}$ and the support of measure μ by Ω_0. As $\mathbb{P}(A_n(\omega) \in B) = \mu(B)$ for any $B \in \mathcal{B}$, for each n, A_n is a measurable map from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\Omega_0, \mathcal{B}, \mu)$. Furthermore, the sequence $(A_n)_{n\geq 1}$ is a measure preserving transformation from $(\Omega, \mathcal{F}, \mathbb{P})$ to $\prod_{n\geq 1}(\Omega_0, \mathcal{B}, \mu)$.

Without loss of generality, we identify Ω with $\prod_{n\geq 1} \Omega_0 = \{\omega = (\omega_0, \omega_1, \ldots)\}$, i.e., the collection of all admissible sample sequences, and define σ as the shift map of Ω such that $\sigma(\omega_0, \omega_1, \ldots) = (\omega_1, \omega_2, \ldots)$. In this case, for each $n \geq 1$, A_n is the projection map sending an $\omega \in \Omega$ to its n-th coordinate ω_n. To emphasize that ω_n, the n-th experiment outcome following the same law as all the A's, is a matrix, we write $A(\omega_n) := \omega_n$. Thus $A_n(\omega) = A(\omega_n)$. Now define a map

$$\Sigma : \Omega \times \mathbb{R}P^{d-1} \to \Omega \times \mathbb{R}P^{d-1}$$

$$(\omega, x) \mapsto (\sigma\omega, A(\omega_0) \cdot x).$$

Then we can see that Σ preserves $\mathbb{P}\otimes\nu$. In fact, for any $f \in \mathcal{F}$ and any Borel set C in $\mathbb{R}P^{d-1}$

$$(\mathbb{P}\otimes\nu)(\sigma\omega \in F, A_1(\omega)\cdot x \in C) = \mathbb{P}(\sigma\omega \in F) \cdot \int 1_C(A \cdot x) \, \mathrm{d}\mu(A) \, \mathrm{d}\nu(x)$$

$$= (\mathbb{P}\otimes\nu)(\omega \in F, x \in C),$$

where the first equality is due to the fact that σ preserves $\mathbb{P}$ and the second follows from the μ-invariance of ν. Let $f(\omega, x) = \log \|A(\omega_0)x\|$ and notice

$$f(\Sigma^p(\omega, x)) = \log \left\| A(\omega_p) \frac{S_p(\omega)x}{\|S_p(\omega)x\|} \right\|,$$

then $n^{-1}\log\|S_n(\omega)x\|$ can be written as

$$\frac{1}{n}(f\circ\Sigma^{n-1}+\cdots+f\circ\Sigma+f)(\omega,x),$$

which, by Birkhoff's ergodic theorem (See e.g. [11] Section 7.2), converges $\mathbb{P}\otimes\nu$-a.s. and in $L^1(\mathbb{P}\otimes\nu)$, thus the last equation follows. □

If we replace the real projective space by the set of matrices of norm one, and consider the corresponding action, we can have a similar result with almost the same proof. In fact, we can even generalise this result to additive cocycles.

Definition 2.1.11. Given a semigroup action of S on a topological space X, then a function $\delta : S\times X\to\mathbb{R}$ is called an **additive cocycle** on $S\times X$ if it satisfies

$$\delta(ts,x)=\delta(t,s.x)+\delta(s,x).$$

Obviously $f(A,x)=\log\|Ax\|$ defined in the last proposition is an additive cocycle. When we consider the normalised action on matrices of norm one, we can also define an additive cocycle by $g(A,M)=\log\|AM\|$.

For a general additive cocyle, the following corollary can be proved in a similar way to Proposition 2.1.10.

Corollary 2.1.12. *Let S be a topological semigroup acting on a compact metrisable space X, $\{Y_n(\omega)\}_{n\geq1}$ be a sequence of i.i.d random variables with a common distribution μ on S, let ν be a μ-invariant measure on X (the existence of ν is guaranteed by Proposition 2.1.8), and let δ be an additive cocycle on $S\times X$. If*

$$\int\left|\delta(Y,x)\right|\mathrm{d}\mu(Y)\,\mathrm{d}\nu(x)<\infty,$$

then

$$\lim_{n\to\infty}\frac{1}{n}\delta(Y_n(\omega)\cdots Y_1(\omega),x)$$

exists for $\mathbb{P}\otimes\nu$-almost all (ω,M) and in $L^1(\mathbb{P}\otimes\nu)$. Moreover, if we denote the limit by $\gamma(\omega,x)$, then

$$\iint\gamma(\omega,x)\,\mathrm{d}\mathbb{P}(\omega)\,\mathrm{d}\nu(x)=\iint\delta(Y,x)\,\mathrm{d}\mu(Y)\,\mathrm{d}\nu(x).$$

2.1.4 Furstenberg-Kesten Theorem

In this section, we will give a proof for the first important theorem on products of random matrices due to Furstenberg and Kesten [19]. The theorem states that under the finite expectation condition, we can drop the expectation symbol in the definition of the Lyapunov exponent and still obtain a.s. convergence. More precisely, let $S_n = A_n \cdots A_1$ be an i.i.d. product of random matrices, then the limit $\lim_{n\to\infty} n^{-1} \log \|S_n\|$ almost surely exists and equals the Lyapunov exponent defined by $\lim_{n\to\infty} n^{-1}\mathbb{E}[\log \|S_n\|]$. Since the main ideas in the original proof will be repeatedly used in later sections, we decide to follow them here instead of directly applying the famous subadditive ergodic theorem [33] due to Kingman.

The essence of this proof is an application of Corollary 2.1.12. We consider the matrix action on matrices of norm one with the additive cocycle $\delta(A, M) = \log \|AM\|$. Moreover, we still need the existence of an invariant measure ν on matrices of norm one and the condition that $|\delta(A, M)| \in L^1(\mu \otimes \nu)$. In the following proof, we will construct such an invariant measure in step 1, and prove the integrability condition in step 2. In step 3, Corollary 2.1.12 is applied to obtain the limit $\gamma(\omega, M)$. Then we prove this limit almost surely equals to the Lyapunov exponent in step 4, using the same technique as in Lemma 2.1.1.

Theorem 2.1.13. *(Furstenberg, Kesten)*
Let $\{A_n(\omega)\}_{n\geq 1}$ be a sequence of i.i.d. random variables with a common distribution μ on $GL(d, \mathbb{R})$. If $\mathbb{E}[\log^+ \|A_1(\omega)\|] < \infty$, then the limit

$$\lim_{n\to\infty} \frac{1}{n} \log \|A_n(\omega) \cdots A_1(\omega)\|$$

exists for $\mathbb{P}$-almost all ω. If we denote by $\gamma(\omega)$ the limit when it exists, then $\gamma(\omega)$ equals the top Lyapunov exponent $\mathbb{P}$-a.s..

Proof. Denote by γ the Lyapunov exponent. We first assume $\gamma \neq -\infty$.

Step 1. Consider the matrix action on the set of matrices of norm one. If we let $m = \delta_{I_d}$ be the Dirac measure on the identity matrix I_d, then by Proposition 2.1.8, $\nu_n = n^{-1} \sum_{k=1}^{n-1} \mu^k \dot{*} m$ weakly converges to a μ-invariant measure ν through a subsequence (ν_{n_i}). Notice

$$\begin{aligned}\iint \log \|AM\| \, \mathrm{d}\mu(A) \, \mathrm{d}\nu_n(M) &= \frac{1}{n} \sum_{k=1}^{n-1} \iint \log \|AM\| \, \mathrm{d}\mu(A) \, \mathrm{d}(\mu^k \dot{*} m)(M) \\ &= \frac{1}{n} \sum_{k=1}^{n-1} \int \cdots \int \log \left\| A_{k+1} \frac{A_k \cdots A_1}{\|A_k \cdots A_1\|} \right\| \mathrm{d}\mu(A_{k+1}) \cdots \mathrm{d}\mu(A_1)\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{n}\int\cdots\int \log\|A_nA_{n-1}\cdots A_1\|\,\mathrm{d}\mu(A_n)\cdots\mathrm{d}\mu(A_1)\\
=&\frac{1}{n}\mathbb{E}[\log\|A_n(\omega)\cdots A_1(\omega)\|].
\end{aligned}
$$

Through the subsequence $n_i\to\infty$, we would have

$$
\lim_{i\to\infty}\iint \log\|AM\|\,\mathrm{d}\mu(A)\,\mathrm{d}\nu_{n_i}(M)=\gamma. \tag{2.5}
$$

Step 2. As $\mathbb{E}[\log^+\|A_1(\omega)\|]<\infty$ and $\|M\|=1$, we have

$$
\int \log^+\|AM\|\,\mathrm{d}\mu(A)\le\mathbb{E}[\log^+\|A_1(\omega)\|]<\infty \tag{2.6}
$$

is a bounded continuous function with respect to M. Thus, since $\nu_{n_i}\to\nu$ weakly,

$$
\lim_{i\to\infty}\iint \log^+\|AM\|\,\mathrm{d}\mu(A)\,\mathrm{d}\nu_{n_i}(M)=\iint \log^+\|AM\|\,\mathrm{d}\mu(A)\,\mathrm{d}\nu(M). \tag{2.7}
$$

As for the negative part $\log^-\|AM\|=\log^+\|AM\|-\log\|AM\|$, we can set $f_k=\min\{\log^-\|AM\|,k\}$ for $k\ge1$ and obtain

$$
\lim_{i\to\infty}\iint f_k\,\mathrm{d}\mu(A)\,\mathrm{d}\nu_{n_i}(M)=\iint f_k\,\mathrm{d}\mu(A)\,\mathrm{d}\nu(M). \tag{2.8}
$$

Notice $f_k\ge0$ and $f_k\to\log^-\|AM\|$ a.s. as $k\to\infty$, by the monotone convergence theorem and Fatou's lemma, we have

$$
\begin{aligned}
\iint \log^-\|AM\|\,\mathrm{d}\mu(A)\,\mathrm{d}\nu(M)&=\lim_{k\to\infty}\iint f_k\,\mathrm{d}\mu(A)\,\mathrm{d}\nu(M) && \text{(by the MCT)}\\
&=\lim_{k\to\infty}\lim_{i\to\infty}\iint f_k\,\mathrm{d}\mu(A)\,\mathrm{d}\nu_{n_i}(M) && \text{(by (2.8))}\\
&\le\varliminf_{i\to\infty}\iint \log^-\|AM\|\,\mathrm{d}\mu(A)\,\mathrm{d}\nu_{n_i}(M). && \text{(by Fatou's Lemma)}
\end{aligned}
$$

Together with (2.7), we have

$$
\varlimsup_{i\to\infty}\iint \log\|AM\|\,\mathrm{d}\mu(A)\,\mathrm{d}\nu_{n_i}(M)\le\iint \log\|AM\|\,\mathrm{d}\mu(A)\,\mathrm{d}\nu(M).
$$

By (2.5),

$$\iint \log \|AM\| \, \mathrm{d}\mu(A) \, \mathrm{d}\nu(M) \geq \gamma > -\infty. \tag{2.9}$$

Therefore (2.9) and (2.6) show that $\log \|AM\| \in L^1(\mu \otimes \nu)$.

Step 3. By Corollary 2.1.12, as $n \to \infty$,

$$\frac{1}{n} \log \|A_n(\omega) \cdots A_1(\omega) M\| \to \gamma(\omega, M) \quad \mathbb{P} \otimes \nu\text{-a.s.},$$

such that

$$\iint \gamma(\omega, M) \, \mathrm{d}\mathbb{P}(\omega) \, \mathrm{d}\nu(M) = \iint \log \|AM\| \, \mathrm{d}\mu(A) \, \mathrm{d}\nu(M).$$

Step 4. Notice (2.9) shows that

$$\gamma \leq \iint \log \|AM\| \, \mathrm{d}\mu(A) \, \mathrm{d}\nu(M) = \iint \gamma(\omega, M) \, \mathrm{d}\mathbb{P}(\omega) \, \mathrm{d}\nu(M).$$

Thus it is sufficient to prove $\gamma(\omega, M) \leq \gamma, \ \mathbb{P}$-a.s. for each M.

For a fixed integer k, write $n = kq_n + r_n$ for some integers q_n and $0 \leq r_n \leq k-1$, write $S_n^m(\omega) = A_n(\omega) \cdots A_m(\omega)$ and $S_n(\omega) = S_n^1(\omega)$, then

$$\begin{aligned}
\frac{1}{n} \log \|S_n(\omega)\| &\leq \frac{1}{n} \log \|A_n(\omega) \cdots A_{kq_n+1}(\omega)\| + \frac{1}{n} \log \|S_{kq_n}(\omega)\| \\
&\leq \frac{1}{n} \sum_{i=kq_n+1}^{n} \log \|A_i(\omega)\| + \frac{1}{n} \left(\sum_{j=1}^{q_n-1} \log \|S_{(j+1)k}^{jk+1}(\omega)\| + \log \|S_k(\omega)\| \right).
\end{aligned}$$

For the first term, note for any fixed $0 \leq r \leq k-1$ and any $\epsilon > 0$,

$$\sum_{q=1}^{\infty} \mathbb{P} \left(\frac{1}{kq+r} \sum_{i=kq+1}^{kq+r} \log \|A_i(\omega)\| > \epsilon \right) = \sum_{q=1}^{\infty} \mathbb{P} \left(\sum_{i=1}^{r} \log \|A_i(\omega)\| > (kq+r)\epsilon \right) < \infty,$$

since $\mathbb{E}\left[\sum_{i=1}^{r} \log \|A_i(\omega)\|\right] \leq r\mathbb{E}[\log^+ \|A_1(\omega)\|] < \infty$. Therefore,

$$\frac{1}{n} \sum_{i=kq+1}^{kq+r} \log \|A_i(\omega)\| \to 0 \quad \text{a.s.}, \quad \text{as } n \to \infty.$$

By the law of large numbers, the second term converges to $\frac{1}{k}\mathbb{E}[\log \|S_k(\omega)\|]$ $\mathbb{P}$-a.s. as $q \to \infty$. Now we get

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \|S_n(\omega)\| \leq \lim_{k\to\infty} \frac{1}{k}\mathbb{E}[\log \|S_k(\omega)\|] = \gamma. \tag{2.10}$$

Therefore, by step 3 and (2.10)

$$\gamma(\omega, M) = \lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)M\| \leq \underline{\lim_{n\to\infty}} \frac{1}{n} \log \|S_n(\omega)\| \leq \gamma.$$

Finally, when $\gamma = -\infty$, by the last inequality, $\gamma(\omega, M) = -\infty$. □

2.2 Lyapunov Spectrum

In this section, we introduce singular value decompositions, exterior powers and the Lyapunov spectrum. The Lyapunov exponent introduced in Section 2.1 is sometimes also called the top Lyapunov exponent. In later sections, we will see that, analogous to the top eigenvalue of a matrix, the (top) Lyapunov exponent of a system solely characterises the asymptotic behaviour in the single dominant direction. And we have to introduce the Lyapunov spectrum, analogous to the spectrum of a matrix, in order to understand more details on the asymptotic behaviour of a system.

Singular Value Decomposition

Given a real matrix A, we denote by A^* the transpose of A.

Definition 2.2.1. Let $A \in M_d(\mathbb{R})$, then we say that

$$A = UDV$$

is a **singular value decomposition** (or an **SVD**) of A if U, V are orthogonal (i.e., $U^*U = V^*V = I_d$) and $D = \operatorname{diag}(c_1, c_2, \ldots, c_d)$ with $c_1 \geq c_2 \geq \cdots \geq c_d \geq 0$. We call $c_1, \ldots, c_d$ the **singular values** of the matrix A.

Remark 2.2.2. Such a decomposition exists for any matrix. In fact, for $A \in GL(d, \mathbb{R})$, as A^*A is a positive-definite matrix, there exists an orthogonal matrix V such that $A^*A = V^*\Delta V$, where $\Delta = \operatorname{diag}(c_1^2, \ldots, c_d^2)$ with $c_1^2 \geq \cdots \geq c_d^2 > 0$. Now let $D = \operatorname{diag}(|c_1|, \ldots, |c_d|)$, and define $U = AV^*D^{-1}$, we can see $U^*U = D^{-1}VA^*AV^*D^{-1} = I_d$, thus U is orthogonal. Therefore $A = UDV$ is a singular value decomposition.

And in general for $A \in M_d(\mathbb{R})$, similarly write $A^*A = V^*\Delta V$. There will be zeros on the diagonal of Δ thus D^{-1} is not well-defined, but we can write $\Delta = \mathrm{diag}(\Delta_0, O)$, where $\Delta_0 \in GL(r, \mathbb{R})$. Write $V^* = \begin{pmatrix} V_1^* & V_2^* \end{pmatrix}$ such that $V_1 \in \mathbb{R}^{r\times d}, V_2 \in \mathbb{R}^{(d-r)\times d}$. We can deduce $AV_2^* = 0$ and define $U_1 = AV_1^*\Delta_0^{-1/2} \in \mathbb{R}^{d\times r}$. Now U_1 consists of r orthogonal vectors, we can add other $d-r$ vectors with which they together form an orthogonal basis. Denote the matrix formed by the other $d-r$ vectors by U_2, then $U = \begin{pmatrix} U_1 & U_2 \end{pmatrix}$ is the orthogonal matrix we are looking for.

We can also notice, for the ℓ_2-norm $\|\cdot\|_2$,

$$\|A\|_2 = \sup_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sup_{x\neq 0} \frac{\|UDVx\|_2}{\|x\|_2} = \sup_{x\neq 0} \frac{\|DVx\|_2}{\|Vx\|_2} = \|D\|_2 = c_1.$$

This implies that c_1 is invariant under different SVDs of A. Similarly, $c_2, \ldots, c_d$ are also invariant, which will follow from (2.11) below.

Exterior Power

Definition 2.2.3. Let W be a d-dimensional linear space, for $2 \leq k \leq d$, we define the **k-th exterior power** $\bigwedge^k(W)$ of W to be the linear space spanned by elements of the form $x_1 \wedge x_2 \wedge \cdots \wedge x_k$, for $x_i \in W, 1 \leq i \leq k$, satisfying

(a) $x_1 \wedge \cdots \wedge (x_j + x_j') \wedge \cdots \wedge x_k = (x_1 \wedge \cdots \wedge x_j \wedge \cdots \wedge x_k) + (x_1 \wedge \cdots \wedge x_j' \wedge \cdots \wedge x_k)$;

(b) $x_1 \wedge \cdots \wedge cx_j \wedge \cdots \wedge x_k = c(x_1 \wedge \cdots \wedge x_j \wedge \cdots \wedge x_k)$ for $c \in \mathbb{R}$;

(c) for any permutation σ of $\{1, 2, \ldots, k\}$

$$x_{\sigma(1)} \wedge \cdots \wedge x_{\sigma(k)} = \mathrm{sgn}(\sigma) x_1 \wedge \cdots \wedge x_k,$$

where $\mathrm{sgn}(\sigma)$ is the sign of the permutation σ.

Remark 2.2.4. The k-th exterior power space $\bigwedge^k(W)$ can be considered as the collection of all k-dimensional subspaces in W. In fact, if we assume $\{e_1, \ldots, e_d\}$ is a basis of W, then $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : i_1 > i_2 > \cdots > i_k\}$ is a basis for $\bigwedge^k(W)$. An inner product on $\bigwedge^k(W)$ can be introduced by

$$\langle x_1 \wedge \cdots \wedge x_k, y_1 \wedge \cdots \wedge y_k \rangle = \det\left[\left(\langle u_i, v_j \rangle\right)_{i,j}\right].$$

We still use $\|\cdot\|_2$ to denote the norm on $\bigwedge^k(W)$ induced by this inner product. (See e.g. [36] Chap. 9 for more details on exterior powers.)

Definition 2.2.5. Given a matrix $A \in GL(d, \mathbb{R})$, for $2 \leq k \leq d$ we can define an automorphism $\wedge^k A$ on $\bigwedge^k(\mathbb{R}^d)$ by

$$(\wedge^k A)(x_1 \wedge \cdots \wedge x_k) = Ax_1 \wedge \cdots \wedge Ax_k.$$

Remark 2.2.6. We will denote $\| \wedge^k A\|_2 = \sup\{\| \wedge^k Aw\|_2 : w \in \bigwedge^k(\mathbb{R}^d), \|w\|_2 = 1\}$. We observe that $\wedge^k(A_1A_2) = (\wedge^k A_1)(\wedge^k A_2)$ for $A_1, A_2 \in GL(d, \mathbb{R})$. And therefore $\| \wedge^k (A_1A_2)\|_2 \leq \| \wedge^k A_1\|_2 \| \wedge^k A_2\|_2$.

Usually given a matrix A, we can simply compute $\wedge^k A$ by computing the $(d-k)$-th minor determinants of A under a certain basis. Here is a simple example. Let e_1, e_2, e_3 be the natural basis for $\mathbb{R}^3$ and

$$A\begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Then consider the basis $e_1 \wedge e_2, e_1 \wedge e_3, e_2 \wedge e_3$ of $\bigwedge^2(\mathbb{R}^3)$, and notice by Definition 2.2.3 property (c), $e_i \wedge e_j = -e_j \wedge e_i$ for $i \neq j$ and $e_i \wedge e_i = 0$, we have

$$\begin{aligned} A(e_1 \wedge e_2) &= (a_{11}e_1 + a_{21}e_2 + a_{31}e_3) \wedge (a_{12}e_1 + a_{22}e_2 + a_{32}e_3) \\ &= (a_{11}a_{22} - a_{21}a_{12})e_1 \wedge e_2 + (a_{11}a_{32} - a_{31}a_{12})e_1 \wedge e_3 + (a_{21}a_{32} - a_{31}a_{22})e_2 \wedge e_3 \\ &= M_{3,3}e_1 \wedge e_2 + M_{2,3}e_1 \wedge e_3 + M_{1,3}e_2 \wedge e_3, \end{aligned}$$

where $M_{i,j}$ denotes the determinant of the submatrix of A, constructed by deleting the i-th row and j-th column of A. We can do the same for $A(e_1 \wedge e_3)$ and $A(e_2 \wedge e_3)$.

SVD for Random Products

Now we apply the discussion above to a singular value decomposition $A = UDV$ of $A \in GL(d, \mathbb{R})$, we can get $\wedge^k A = (\wedge^k U)(\wedge^k D)(\wedge^k V)$. We can notice $(\wedge^k U)^*(\wedge^k U) = \wedge^k I_d = \mathrm{id}_{\wedge^k(\mathbb{R}^d)}$, therefore $\wedge^k U, \wedge^k V$ are both orthogonal matrices. By the computation introduced above, $\wedge^k D$ is still a diagonal matrix with $c_1c_2 \cdots c_k$ as its largest diagonal entry. Therefore $\wedge^k A = (\wedge^k U)(\wedge^k D)(\wedge^k V)$ is a singular value decomposition of $\wedge^k A$, and by the same arguments as in Remark 2.2.2 we have

$$\| \wedge^k A\|_2 = \| \wedge^k D\|_2 = c_1c_2 \cdots c_k. \tag{2.11}$$

Now for the matrix $S_n(\omega)$, i.e. the product of an i.i.d. sequence $\{A_n(\omega)\}$ of random matrices, we can write $S_n(\omega) = U_n(\omega)D_n(\omega)V_n(\omega)$ as an SVD for $S_n(\omega)$ where $D_n(\omega) =$

$\mathrm{diag}(c_1(n,\omega),\ldots,c_d(n,\omega))$. (For simplicity, in the following discussion we omit ω when there is no ambiguity.)

We recall the fact that the definition of the Lyapunov exponent and Theorem 2.1.13 do not depend on the choice of the norm $\|\cdot\|$. An obvious consequence is that, when $\mathbb{E}[\log^+ \|A_1(\omega)\|] < \infty$,

$$\gamma = \lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)\|_2 = \lim_{n\to\infty} \frac{1}{n} \log c_1(n,\omega) \quad \mathbb{P}\text{-a.s.},$$

by Theorem 2.1.13. Notice

$$\log^+ \| \wedge^k A_1(\omega)\|_2 = \log^+ c_1 \cdots c_k \le k \log^+ c_1,$$

therefore $\mathbb{E}[\log^+ \|A_1(\omega)\|] < \infty$ implies that $\mathbb{E}[\log^+ \| \wedge^k A_1(\omega)\|] < \infty$. Thus we can also apply Theorem 2.1.13 to $\wedge^k S_n(\omega)$ and get that

$$\lim_{n\to\infty} \frac{1}{n} \log \| \wedge^k S_n(\omega)\|_2 = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{k} \log c_i(n,\omega)$$

exists $\mathbb{P}$-a.s. for each $2 \le k \le d$. We denote this limit by ζ_k. Inductively, set $\zeta_1 = \gamma$ and for each $2 \le k \le d$, we can define

$$\gamma_k = \zeta_k - \zeta_{k-1} = \lim_{n\to\infty} \frac{1}{n} \log c_k(n,\omega). \tag{2.12}$$

Clearly, $\gamma \ge \gamma_2 \ge \cdots \ge \gamma_d$ since $c_1(n,\omega) \ge c_2(n,\omega) \ge \cdots \ge c_d(n,\omega)$. Set $\gamma_1 = \gamma$, then the set $\{\gamma_k\}_{1\le k\le d}$ is called the **Lyapunov spectrum**. γ_k is called the k**-th Lyapunov exponent** and γ_1 is sometimes called the top Lyapunov exponent.

2.3 Invariant Measures on Projective Space

In this section, we want to study the limit

$$\lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)x\|$$

for some non-zero vector $x \in \mathbb{R}^d$, where $S_n(\omega)$ is defined as in previous sections. We expect this limit to exist ($\mathbb{P}$-a.s.) and to be related to the Lyapunov exponent, but we can easily find examples in which the limit does not equal the Lyapunov exponent for some choices of x in $\mathbb{R}^d$.

We consider the most trivial case, in which μ is a Dirac measure on the real matrix A. Then all the possible random matrix products reduce to the deterministic product A^n. Suppose

$A = \mathrm{diag}(c_1, \ldots, c_d)$, where $c_1 > c_2 > \cdots > c_d > 0$, then the Lyapunov exponent equals $\log c_1$. However, let $\{e_i : 1 \leq i \leq d\}$ be the standard basis of $\mathbb{R}^d$, then we can easily see that

$$\lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)e_i\| = \log c_i,$$

for each $1 \leq i \leq d$.

The phenomenon in the example is caused by the fact that A is a diagonal matrix, where the asymptotic behaviour of A^n can be reduced to d smaller problems, namely the asymptotic behaviour of each c_i^n. We will later introduce the notion of irreducibility and strong irreducibility to exclude this kind of example in which matrix action can decompose into actions that stabilise subspaces.

Definition 2.3.1. Given a probability measure μ on $M_d(\mathbb{R})$, the **support** of the measure μ is defined to be

$$\mathrm{supp}(\mu) = \{A \in M_d(\mathbb{R}) : \text{for every open neighbourhood } N_A \ni A, \mu(N_A) > 0\}.$$

We denote by $\mathcal{S}_\mu$ the **semigroup generated by** $\mathrm{supp}(\mu)$, i.e., $\mathcal{S}_\mu = \bigcup_{n\geq 1} \mathrm{supp}(\mu^n)$ is the smallest semigroup which contains $\mathrm{supp}(\mu)$.

We can notice that all the elements of the form $S_n(\omega)$ for $\omega \in \Omega, n \geq 1$ lie in $\mathcal{S}_\mu$. Thus the limiting behaviour of $S_n(\omega)$ relates to the "boundary" of $\mathcal{S}_\mu$.

2.3.1 Irreducibility

Definition 2.3.2. Let $\mathcal{S}$ be a subset of $GL(d, \mathbb{R})$, then we say $\mathcal{S}$ is **irreducible** if there is no proper linear subspace V of $\mathbb{R}^d$ such that $A(V) = V$ for any $A \in \mathcal{S}$. We say $\mathcal{S}$ is **strongly irreducible** if there does not exist a finite family of proper linear subspaces $V_1, \ldots, V_k$ of $\mathbb{R}^d$ such that

$$A(V_1 \cup \cdots \cup V_k) = V_1 \cup \cdots \cup V_k$$

for any $A \in \mathcal{S}$.

Remark 2.3.3. It is not hard to see that the irreducibility (or strong irreducibility) of a set $\mathcal{S}$ is equivalent to the irreducibility (strong irreducibility *resp.*) of the semigroup generated by $\mathcal{S}$. Thus without loss of generality we assume $\mathcal{S}$ is a semigroup.

When the semigroup $\mathcal{S}$ is not irreducible, let V be the common proper invariant subspace of all matrices in $\mathcal{S}$. Let $e_1, \ldots, e_k$ be a basis of V, and complete it with $e'_{k+1}, \ldots, e'_d$ so that they

together form a basis of $\mathbb{R}^d$. Then in this basis all the matrices in $\mathcal{S}$ can be written in the form

$$\begin{pmatrix} C & U \\ 0 & D \end{pmatrix},$$

where $C \in GL(k, \mathbb{R}), D \in GL(d-k, \mathbb{R}), U \in \mathbb{R}^{k\times(d-k)}$ and 0 is a zero matrix.

When the semigroup $\mathcal{S}$ is not strongly irreducible, by definition there exists a family of proper linear subspaces $V_1, \ldots, V_k$ such that $A(V_1 \cup \cdots \cup V_k) = V_1 \cup \cdots \cup V_k$ for each $A \in \mathcal{S}$. Let $\mathcal{V}$ be the collection of all the families which have this property for $\mathcal{S}$, as $\mathcal{V}$ is non-empty. Thus we can define

$$s = \min\{r \in \mathbb{N} : \mathcal{L} = \{W_1, \ldots, W_l\} \in \mathcal{V},\ \dim W_i \leq r, 1 \leq i \leq l\}.$$

Let $\mathcal{L} = \{W_1, \ldots, W_l\}$ be a family for which the minimum is reached. Assume $\dim W_1 = s$, for some integer $1 \leq t \leq l$ and

$$\bigcup_{A\in\mathcal{S}} A(W_1) = W_1 \cup \cdots \cup W_t.$$

Furthermore, we can assume $W_i \cap W_j = \{0\}$ for each $i \neq j$ without loss of generality. In fact, for each $A \in \mathcal{S}$, $\dim A(W_1) = s$ since A is invertible, and for each $1 \leq i \leq t$, the dimension of W_i is at most s, henceforth $A(W_1) = W_j$ for some j. Similarly, for each $A \in \mathcal{S}, 1 \leq i \leq t$, there exists $1 \leq j(i) \leq t$ such that $AW_i = W_{j(i)}$. Now we have

$$A\left(\bigcup_{1\leq i<j\leq t} W_i \cap W_j\right) = \bigcup_{1\leq i<j\leq t} W_i \cap W_j$$

for each $A \in \mathcal{S}$. Therefore, unless $W_i \cap W_j = \{0\}$ for each pair $i \neq j$, we have $0 < \dim W_i \cap W_j < s$ which is contradictory to the minimality of s.

In short, when the semigroup $\mathcal{S}$ is not strongly irreducible, there exists a family of proper linear subspaces $W_1, \ldots, W_t$ with the same dimension, such that $W_i \cap W_j = \{0\}$ for $i \neq j$, and each matrix in $\mathcal{S}$ permutes $W_1, \ldots, W_t$.

For example, we assume we have a semigroup $\mathcal{S}$ which is not strongly irreducible, and there exist proper linear subspaces W_1, W_2 which satisfy the properties described above. Assume $A_1 \in \mathcal{S}$ sends W_1 to W_2 and W_2 to W_1, while $A_2 \in \mathcal{S}$ fixes both subspaces, then after choosing a basis of $\mathbb{R}^d$ (by choosing bases of W_1, W_2 separately and then complete them), we can write

down their matrix representations

$$A_1 = \begin{pmatrix} 0 & C_1 & * \\ D_1 & 0 & * \\ 0 & 0 & * \end{pmatrix}, \quad A_2 = \begin{pmatrix} C_2 & 0 & * \\ 0 & D_2 & * \\ 0 & 0 & * \end{pmatrix}.$$

In a lot of cases, strong irreducibility condition can be weakened to irreducibility. As we have seen above, "not irreducible" implies that the problem can be decomposed into several sub-problems, i.e., we can consider submatrix action on subspaces; meanwhile, "not strongly irreducible" implies that we can "almost" decompose the problem, but permutations of subspaces are also involved. The"irreducible but not strongly irreducible" case can be dealt with if we define an equivalence relation among those subspaces which are permuted by the semigroup. In [4], stopping times are also defined to mark when the permutations happen. Interested readers can refer to Section III.7 of the book for more details.

In this dissertation, we will only assume strong irreducibility for convenience.

The rest of the section discusses the relationship between (strong) irreducibility and properties of the corresponding invariant measures.

Intuitively, if the problem can be decomposed into sub-problems (i.e., being reducible), we can find proper subspaces on which those submatrices are acting. And then we can find invariant measures supported on those subspaces. For example, let μ be supported by a single diagonal 2×2 matrix $D = \operatorname{diag}(a, b)$ and $\{e_1, e_2\}$ be a standard basis for $\mathbb{R}^2$. Then both of the Dirac measures $\delta_{e_1}, \delta_{e_2}$ are μ-invariant, but meanwhile they vanish on e_2 and e_1 respectively, i.e., they are not proper if we introduce the following definition.

Definition 2.3.4. A measure ν on $\mathbb{R}P^d$ is said to be **proper** if for any proper subspace V of $\mathbb{R}^d$, $\nu(\overline{V}) = 0$, where we denote by $\overline{V}$ the image of $V \setminus \{0\}$ under the natural quotient map $\pi : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}P^{d-1}$.

The following proposition states that the concept of strong irreducibility allows us to obtain proper invariant measures.

Proposition 2.3.5. *Let μ be a probability measure on $GL(d, \mathbb{R})$ such that $\mathcal{S}_\mu$ is strongly irreducible, and ν be a μ-invariant measure on $\mathbb{R}P^{d-1}$, then ν is proper.*

Proof. Let $\mathcal{W}$ be the collection of all the proper linear subspaces W of $\mathbb{R}^d$ such that $\nu(\overline{W}) \neq 0$, where $\overline{W} = \pi(W)$. If $\mathcal{W}$ is non-empty, define

$$l_0 = \inf\{\dim W : W \in \mathcal{W}\}.$$

Denote $\mathcal{W}(l_0) = \{W \in \mathcal{W} : \dim W = l_0\}, m = \sup\{\nu(\overline{W}) : W \in \mathcal{W}(l_0)\} > 0$, and $\Lambda = \{W \in \mathcal{W}(l_0) : \nu(\overline{W}) = m\}$.

We claim that $\Lambda \neq \emptyset$. Otherwise for each $n \geq 1$, there exists $W_n \in \mathcal{W}(l_0)$ such that $\nu(\overline{W}_n) > m - 1/n^2$. Note for any $W_i, W_j, 1 \leq i, j \leq n$, if $W_i \neq W_j$ then $\dim(W_i \cap W_j) < l_0$, thus $\nu(\overline{W}_i \cap \overline{W}_j) = 0$. Thus if $\{W_n\}_{n\in I}$ is an infinite set (consisting of different elements), then

$$\nu\left(\bigcup_{n\in I} \overline{W}_n\right) = \sum_{n\in I} \nu(\overline{W}_n) > \sum_{n\in I} m - \frac{1}{n^2} > \infty,$$

which contradicts the fact that ν is a probability measure. Thus the supremum m can be reached by some $W \in \mathcal{W}(l_0)$. Moreover, with the same arguments, we can deduce that Λ is a finite set.

For $W \in \Lambda$, we have

$$m = \nu(\overline{W}) = \iint 1_{\overline{W}}(A \cdot x)\, \mathrm{d}\mu(A)\, \mathrm{d}\nu(x) = \int \nu(A^{-1} \cdot \overline{W})\, \mathrm{d}\mu(A). \tag{2.13}$$

By the maximality of m, $\nu(A^{-1} \cdot \overline{W}) = m$ for μ-almost all $A \in \mathrm{supp}(\mu)$. If we write $L = \bigcup\{W : W \in \Lambda\}$, then $A(L) = L$ for all $A \in \mathcal{S}_\mu$. This contradicts the strong irreducibility of $\mathcal{S}_\mu$. Therefore $\mathcal{W} = \varnothing$. □

Corollary 2.3.6. *Let μ be a probability measure on $GL(d, \mathbb{R})$ such that $\mathcal{S}_\mu$ is strongly irreducible, and ν be a μ-invariant measure on $\mathbb{R}P^{d-1}$. Let A be a non-zero $d \times d$ matrix, then the probability measure $A\nu$ on $\mathbb{R}P^{d-1}$ satisfying*

$$\int f(x)\, \mathrm{d}(A\nu)(x) = \int f(A \cdot x)\, \mathrm{d}\nu(x)$$

for each bounded Borel function f on $\mathbb{R}P^{d-1}$, is well-defined.

Proof. Note $A \cdot x$ is well-defined if and only if $Ax \neq 0$. By Proposition 2.3.5, we have $\nu(\pi(\ker A)) = 0$, where $\pi : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}P^{d-1}$ is the natural quotient map. Therefore the conclusion follows. □

2.3.2 Contracting Property

In order to prove that under the same finite expectation condition, for a nonzero vector $x \in \mathbb{R}^d$,

$$\lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)x\| = \gamma \quad \mathbb{P}\text{-a.s..}$$

We follow the idea provided by Guivarc'h and Raugi in [24], which is to prove that

$$C(n,\omega) \leq \frac{\|S_n(\omega)x\|_2}{\|S_n(\omega)\|_2} \leq \|x\|_2,$$

for some $C(n,\omega)$ satisfying $\lim_{n\to\infty} n^{-1}\log C(n,\omega) = 0$.

We start with a singular value decomposition of $S_n(\omega) = U_n(\omega)D_n(\omega)V_n(\omega)$, where $D_n(\omega) = \operatorname{diag}(c_1(n,\omega),\ldots,c_d(n,\omega))$. Then we can get (we omit ω when there is no ambiguity)

$$\frac{\|S_n(\omega)x\|_2^2}{\|S_n(\omega)\|_2^2} = \frac{\|D_nV_nx\|_2^2}{\|D_n\|_2^2} = \frac{\sum_{i=1}^d c_i(n)^2\langle V_nx, e_i\rangle^2}{c_1(n)^2} = \sum_{i=1}^d \frac{c_i(n)^2}{c_1(n)^2}\langle x, V_n^*e_i\rangle^2, \tag{2.14}$$

where $\{e_i : 1 \leq i \leq d\}$ is the standard basis for $\mathbb{R}^d$. Now, there are two objects whose limiting behaviour we are interested in: $c_i(n)/c_1(n)$ for each $1 \leq i \leq k$ and $V_n^*e_i$.

Observe that

$$\frac{S_n^*}{\|S_n^*\|_2} = V_n^* \operatorname{diag}\left(1, \frac{c_2(n)}{c_1(n)}, \ldots, \frac{c_d(n)}{c_1(n)}\right) U_n^*, \tag{2.15}$$

in which both of the objects appear. Thus it might be a good idea to first study $S_n^*/\|S_n^*\|_2$. We start with the following definition which characterises the limiting behaviour of the rank of $S_n^*/\|S_n^*\|_2$.

Definition 2.3.7. Let $\mathcal{S}$ be a subset of $M_d(\mathbb{R})$, we can define the **index** of $\mathcal{S}$ to be

$$\min\{r : \text{there exists } (M_n) \subset \mathcal{S} \text{ s.t. } M_n/\|M_n\|_2 \text{ converges to a rank } r \text{ matrix}\}.$$

If the index of $\mathcal{S}$ is one, then we say $\mathcal{S}$ is **contracting**.

Remark 2.3.8. Now with the same notation as before, we assume that for some fixed ω, $S_n^*(\omega)/\|S_n^*(\omega)\|_2$ has a limit point of a rank $r(\omega)$ matrix $M(\omega)$. Now observe in equation (2.15)

$$\frac{S_n^*}{\|S_n^*\|_2} = V_n^* \operatorname{diag}\left(1, \frac{c_2(n)}{c_1(n)}, \ldots, \frac{c_d(n)}{c_1(n)}\right) U_n^*.$$

Since the set of $d\times d$ orthogonal matrices is compact and $0 \leq c_i(n)/c_1(n) \leq 1$ for $2 \leq i \leq d$, we can assume that along a subsequence (n_j), $U_{n_j}(\omega) \to U_\infty(\omega)$, $V_{n_j}(\omega) \to V_\infty(\omega)$, $c_i(n_j,\omega)/c_1(n_j,\omega) \to \alpha_i(\omega)$ and we can write

$$M(\omega) = V_\infty^*(\omega) \operatorname{diag}\big(1, \alpha_1(\omega), \ldots, \alpha_d(\omega)\big) U_\infty^*(\omega), \tag{2.16}$$

with $1 \geq \alpha_1(\omega) \geq \cdots \geq \alpha_d(\omega) \geq 0$ and U_∞, V_∞ both orthogonal. Now as $M(\omega)$ is of rank $r < d$, we can see that $\alpha_{r+1} = \cdots = \alpha_d = 0$ and $\alpha_r > 0$, $\mathbb{P}$-a.s..

As for the limiting behaviour of $V_n^* e_i, 1 \leq i \leq r$, note equation (2.16) actually shows that the linear space spanned by $V_\infty^* e_i, 1 \leq i \leq r$ is exactly the range of $M(\omega)$.

Now our problem is reduced to understanding the range of each limit point of $S_n^*/\|S_n^*\|_2$. Note here $S_n^*(\omega) = A_1^*(\omega) \cdots A_n^*(\omega)$, which has a different order of product from $S_n(\omega)$. Therefore in the following lemma and proposition, we define $T_n(\omega) = A_1(\omega) \cdots A_n(\omega)$. We should bear in mind the difference between $S_n(\omega)$ and $T_n(\omega)$ and that we will apply the lemma and proposition below to the transposes of matrices. Also, these results are not true for $S_n(\omega)$, which will be further explained in Section 3.2.

Lemma 2.3.9. *(Guivarc'h, Raugi [24]) Let $\{A_n(\omega)\}$ be an i.i.d sequence of random variables with a common distribution μ on $GL(d, \mathbb{R})$ and ν be a μ-invariant measure on $\mathbb{R}P^{d-1}$. Then for $\mathbb{P}$-almost all ω, there exists a probability measure ν_ω on $\mathbb{R}P^{d-1}$ such that*

$$A_1(\omega) \cdots A_n(\omega) Q\nu \to \nu_\omega, \quad (n \to \infty)$$

in the weak topology for η-almost all $Q \in GL(d, \mathbb{R})$, where $\eta = \sum_{m=0}^{\infty} 2^{-(m+1)} \mu^m$, and μ^0 denotes a Dirac measure on the identity matrix I_d. Moreover, for any bounded Borel function f on $\mathbb{R}P^{d-1}$,*

$$\mathbb{E} \int f \, \mathrm{d}\nu_\omega = \int f \, \mathrm{d}\nu.$$

Proof. Let f be a continuous real function on $\mathbb{R}P^{d-1}$ and define a map $F : GL(d, \mathbb{R}) \to \mathbb{R}$ by

$$F(Q) = \int f(Q \cdot x) \, \mathrm{d}\nu(x).$$

Let $T_n(\omega) = A_1(\omega) \cdots A_n(\omega)$, $\mathcal{F}_n$ be the σ-algebra generated by $A_1(\omega), \ldots, A_n(\omega)$, then

$$\begin{aligned}
\mathbb{E}[F(T_{n+1})|\mathcal{F}_n] &= \int F(T_n Q) \, \mathrm{d}\mu(Q) \\
&= \iint f(T_n Q \cdot x) \, \mathrm{d}\mu(Q) \, \mathrm{d}\nu(x) \\
&= \int f(T_n \cdot x) \, \mathrm{d}\nu(x) \\
&= F(T_n).
\end{aligned}$$

Therefore $\{F(T_n) : n \geq 1\}$ is a martingale. As $F(T_n)$ is bounded, it converges $\mathbb{P}$-a.s. to a limit which we write as $\Gamma f(\omega)$, such that

$$\mathbb{E}[\Gamma f(\omega)] = \mathbb{E}[F(T_1)] = \int f \, \mathrm{d}\nu. \tag{2.17}$$

Now for any positive integers k and r, since $\{F(T_n)\}_{n\geq 1}$ is a bounded martingale,

$$\mathbb{E}\left[\left|F(T_{k+r}) - F(T_k)\right|^2\right] = \mathbb{E}[F(T_{k+r})^2] - \mathbb{E}[F(T_k)^2].$$

Thus when N is sufficiently large,

$$\sum_{k=1}^{N} \mathbb{E}\left[\left|F(T_{k+r}) - F(T_k)\right|^2\right] = \sum_{k=1}^{r} \mathbb{E}[F(T_{k+N})^2] - \sum_{k=1}^{r} \mathbb{E}[F(T_k)^2] \leq 2r\|f\|_\infty^2. \tag{2.18}$$

Consider now $\sum_{k=1}^{\infty} |F(T_kQ) - F(T_k)|^2$, when Q is sampled from η, this quantity has expectation given by

$$\begin{aligned}
&\sum_{k=1}^{\infty} \mathbb{E}\left[\int \left|F(T_kQ) - F(T_k)\right|^2 \mathrm{d}\eta(Q)\right] \\
=&\sum_{k=1}^{\infty}\sum_{m=1}^{\infty} 2^{-(m+1)} \mathbb{E}\left[\int \left|F(T_kQ) - F(T_k)\right|^2 \mathrm{d}\mu^m(Q)\right] \\
=&\sum_{k=1}^{\infty}\sum_{m=1}^{\infty} 2^{-(m+1)} \mathbb{E}\left[\left|F(T_{k+m}) - F(T_k)\right|^2\right] \\
=&\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} 2^{-(m+1)} \left\{\mathbb{E}\left[F(T_{k+m})\right]^2 - \mathbb{E}\left[F(T_k)\right]^2\right\} \\
\leq&\sum_{m=1}^{\infty} m2^{-m}\|f\|_\infty^2 < \infty,
\end{aligned} \tag{2.19}$$

where the last inequality follows from (2.18). Therefore $\sum_{k=1}^{\infty} |F(T_kQ) - F(T_k)|^2 < \infty$ and $F(T_kQ)$ converges to Γf, $\mathbb{P}\otimes\eta$-almost surely.

Choose a dense sequence $\{f_q : q = 1, 2, \dots\}$ in the space $\mathcal{C}(\mathbb{R}P^{d-1})$ of continuous functions on $\mathbb{R}P^{d-1}$. There exists a measurable subset $Y \subset \Omega \times \mathbb{R}P^{d-1}$ such that $\mathbb{P}\otimes\nu(Y) = 1$, and for each q and for each $(\omega, Q) \in Y$ we have

$$\int f_q(T_n(\omega)Q \cdot x) \, \mathrm{d}\nu(x) \to \Gamma f_q(\omega).$$

Choose now a fixed $Q \in GL(d,\mathbb{R})$ and $\omega \in \Omega$. If $\nu_{Q,\omega}$ is a limit point of $\{T_n(\omega)Q\nu : n \geq 1\}$ in the weak* topology, then

$$\int f_q \, \mathrm{d}\nu_{Q,\omega} = \Gamma f_q(\omega),$$

for each q, which implies that $\nu_{Q,\omega}$ does not depend on Q and is the limit point of $\{T_n(\omega)Q\nu : n \geq 1\}$. Thus we denote it by ν_ω. By (2.17), $\mathbb{E}[\Gamma f_q] = \int f_q \, \mathrm{d}\nu$. Therefore

$$\mathbb{E}[\int f_q \, \mathrm{d}\nu_\omega] = \int f_q \, \mathrm{d}\nu.$$

Now the conclusion follows by the fact that continuous functions are dense in the set of bounded Borel functions. □

We remark here the support $\mathrm{supp}(\eta)$ of the measure $\eta = \sum_{m=0}^{\infty} 2^{-(m+1)}\mu^m$ is exactly the union of $\{I_d\}$ and the semigroup $\mathcal{S}_\mu$ generated by the support of the measure μ.

Notice, in Remark 2.3.8, we fixed the ω at the start of our discussion and assumed the limit point is of rank $r(\omega)$. The next proposition will help us pass from one ω to $\mathbb{P}$-a.s.. Intuitively, let us say there exists some order of products which would converges to the limit point with the minimal rank, then when the product is sufficiently long, we can expect that this order will appear almost surely.

Since in Lemma 2.3.9 we have the freedom to choose (η-almost) any $Q \in \mathrm{supp}(\eta) = \mathcal{S}_\mu \cup \{I_d\}$, we might want to try the sequence $\{M_n\} \subset \mathcal{S}_\mu$ which, after normalisation, converges to the limit point which has the rank of the index.

Proposition 2.3.10. *Let $\{A_n(\omega)\}$ be an i.i.d sequence of random variables with a common distribution μ on $GL(d,\mathbb{R})$. Assume the semigroup $\mathcal{S}_\mu$ generated by $\mathrm{supp}(\mu)$ is strongly irreducible and has index r. Let $T_n(\omega) = A_1(\omega)A_2(\omega)\cdots A_n(\omega)$, then for $\mathbb{P}$-almost all ω, there exists an r-dimensional subspace $W(\omega)$ of $\mathbb{R}^d$ such that it is the range of any limit point of $T_n(\omega)/\|T_n(\omega)\|$, and for any non-zero vector $x \in \mathbb{R}^d$, $\mathbb{P}\{\omega : x \perp W(\omega)\} = 0$.*

Proof. Let ν be a μ-invariant measure on $\mathbb{R}P^{d-1}$. By Lemma 2.3.9, for $\mathbb{P}$-almost all ω there exists a probability measure ν_ω such that,

$$T_n(\omega)Q\nu \to \nu_\omega, \quad (n \to \infty)$$

in the weak* topology, for η-almost all $Q \in GL(d,\mathbb{R})$, where $\eta = \sum_{m=0}^{\infty} 2^{-(m+1)}\mu^m$.

Fix ω and let $M(\omega)$ (of rank $r(\omega)$) be a limit point of $T_n(\omega)/\|T_n(\omega)\|$. Then through a subsequence, $T_n(\omega)Q\nu$ converges to $M(\omega)Q\nu$ (note $M(\omega)Q \neq 0$, thus $M(\omega)Q\nu$ is well-defined by Corollary 2.3.6). Therefore

$$M(\omega)Q\nu = \nu_\omega, \tag{2.20}$$

for each $Q \in \operatorname{supp}(\eta) = \mathcal{S}_\mu \cup \{I_d\}$. Since $\mathcal{S}_\mu$ has index r, there exists a sequence $Q_n/\|Q_n\|$ which converges to a rank r matrix P. As $QQ_n \in \mathcal{S}_\mu$, we have

$$M(\omega)QQ_n\nu = \nu_\omega.$$

By Corollary 2.3.6, either $M(\omega)QP = 0$ or $M(\omega)QP\nu$ is well-defined and we can take $n \to \infty$ and get $M(\omega)QP\nu = \nu_\omega$.

If the first case is true for all $Q \in \mathcal{S}_\mu$, then define $L = \{QPx : Q \in \mathcal{S}_\mu, x \in \mathbb{R}^d\}$, then $\{0\} \neq L \subset \ker M(\omega)$ is a proper linear subspace of $\mathbb{R}^d$ and $Q(L) = L$ for each $Q \in \mathcal{S}_\mu$. This contradicts the strong irreducibility of $\mathcal{S}_\mu$.

Now $M(\omega)QP\nu = \nu_\omega$ for some $Q \in \mathcal{S}_\mu$. We will prove $\operatorname{supp}\nu_\omega \subset \overline{\operatorname{im}(M(\omega)QP)}$. Say if $\bar{z} \in \operatorname{supp}\nu_\omega$ but $z \notin \operatorname{im}(M(\omega)QP)$, since $\operatorname{im}(M(\omega)QP)$ is closed, there exists a neighbourhood N_z of z such that $N_z \cap \operatorname{im}(M(\omega)QP) = \emptyset$. However, $\bar{z} \in \operatorname{supp}\nu_\omega$ implies that $\nu_\omega(N_z) > 0$, which is contradictory to the fact that

$$\nu_\omega(N_z) = \nu\left(\{y : M(\omega)QPy \in N_z\}\right) = 0.$$

Define $W(\omega) = \operatorname{span}\{x \in \mathbb{R}^d : \bar{x} \in \operatorname{supp}\nu_\omega\}$, then $W(\omega)$ is a subspace of $\operatorname{im}(M(\omega)QP)$. Since P is of rank r, $\dim W(\omega) \leq r$.

On the other hand, let $Q = I_d$ in equation (2.20), then $M(\omega)\nu = \nu_\omega$. Thus,

$$\nu\left(\bar{x} : M(\omega)\cdot x \in \overline{W(\omega)}\right) = \nu_\omega\left(\overline{W(\omega)}\right) = 1.$$

Notice that $H := \{x : M(\omega)x \in W(\omega)\}$ is a linear space. Since ν is proper and $\nu(\overline{H}) \neq 0$, we have $H = \mathbb{R}^d$, i.e., $W(\omega) = \operatorname{im} M(\omega)$. Therefore $\dim W(\omega) = \operatorname{rank} M(\omega) = r(\omega) \geq r$. So $\dim W(\omega) = r$.

Therefore, we have proved that for any limit point $M(\omega)$ of $T_n(\omega)/\|T_n(\omega)\|$, the range $W(\omega)$ of $M(\omega)$ is of dimension r.

Now for each non-zero vector $x \in \mathbb{R}^d$, define $x^\perp = \{\bar{y} \in \mathbb{R}P^{d-1} : \langle x, y\rangle = 0\}$. Then

$$\mathbb{P}\{x \perp W(\omega)\} = \mathbb{P}\{\operatorname{supp}\nu_\omega \subset x^\perp\} = \mathbb{P}\{\nu_\omega(x^\perp) = 1\} = \nu(x^\perp) = 0,$$

by Lemma 2.3.9 and Proposition 2.3.5. □

2.3.3 Integral Formula for the Lyapunov Exponent

The main theorem in this section consists of three results. The first one, which is firstly proved by Furstenberg [18] in 1963, claims that the random matrix action on the real projective space would generate the same Lyapunov exponent. Having worked to establish necessary tools in previous sections, this result is immediate. The rest are firstly proved in [20] by Furstenber and Kifer. The second one gives an integral formula of the Lyapunov exponent using the invariant measure, and the third one would be useful for the next section in proving "uniformly contracting" results.

Theorem 2.3.11. *(Furstenberg, Kifer)*
Let $\{A_n(\omega)\}$ be a sequence of i.i.d invertible random matrices with a common distribution μ such that $\mathbb{E}[\log^+ \|A_1(\omega)\|] < \infty$ and $\mathcal{S}_\mu$ is strongly irreducible, then

(a) For any sequence $\{x_n : n \geq 1\}$ of vectors which converges to a non-zero vector $x \in \mathbb{R}^d$,

$$\lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)x_n\| = \gamma \quad \mathbb{P}\text{-a.s.}.$$

(b) Let ν be a μ-invariant measure on the real projective space, then

$$\gamma = \iint \log \|Ax\| \, \mathrm{d}\mu(A) \, \mathrm{d}\nu(x).$$

(c) If moreover $\mathbb{E}[\log^+ \|A_1^{-1}(\omega)\|] < \infty$, then

$$\frac{1}{n}\mathbb{E}\left[\log \|S_n(\omega)x\|\right] \to \gamma \quad (n \to \infty)$$

uniformly on $\{x \in \mathbb{R}^d : \|x\| = 1\}$.

Proof. Without loss of generality, we prove (a) for the ℓ_2-norm. By the discussion in the last section, we have

$$\frac{\|S_n(\omega)x_n\|_2^2}{\|S_n(\omega)\|_2^2} = \sum_{i=1}^{d} \frac{c_i(n)^2}{c_1(n)^2} \langle x_n, V_n^* e_i \rangle^2.$$

Assume $\mathcal{S}_\mu$ has rank r. Then by Remark 2.3.8, $\lim_{n\to\infty} c_i(n)/c_1(n) = 0$ for $r+1 \leq i \leq d$ and $\inf_n c_r(n)/c_1(n) > 0$, $\mathbb{P}$-a.s.. Let $M(\omega)$ be a limit point of $S_n^*(\omega)/\|S_n^*(\omega)\|$, then the range of $M(\omega)$ equals the linear space spanned by $\{V_\infty^* e_i : 1 \leq i \leq r\}$.

By Proposition 2.3.10, the range of $M(\omega)$ is a r-dimensional linear space $W(\omega)$ such that for any non-zero $x \in \mathbb{R}^d$, $\mathbb{P}\{x \perp W(\omega)\} = 0$. Thus given the sequence $x_n \to x$, let $y_n(\omega)$ be the projection of x_n on $W(\omega)$,

$$\frac{\|S_n(\omega)x_n\|_2}{\|S_n(\omega)\|_2} \geq \inf_n \frac{c_r(n)}{c_1(n)} \left(\sum_{i=1}^{r} \langle x_n, V_n^* e_i \rangle^2 \right)^{1/2} = \inf_n \frac{c_r(n)}{c_1(n)} \|y_n(\omega)\|_2, \quad \mathbb{P}\text{-a.s..}$$

By the continuity of the projection operator, we have $y_n(\omega) \to y(\omega)$, where $y(\omega)$ denotes the projection of x onto $W(\omega)$, then by Proposition 2.3.5,

$$\mathbb{P}\left\{ \lim_{n\to\infty} \frac{1}{n} \log \|y_n(\omega)\|_2 \neq 0 \right\} = \mathbb{P}\left\{ \|y(\omega)\|_2 = 0 \right\} = \mathbb{P}\left\{ x \perp W(\omega) \right\} = 0.$$

Thus

$$\inf_n \frac{c_r(n)}{c_1(n)} \|y_n(\omega)\|_2 \leq \frac{\|S_n(\omega)x_n\|_2}{\|S_n(\omega)\|_2} \leq \|x_n\|_2, \quad \mathbb{P}\text{-a.s.} \tag{2.21}$$

implies that

$$\lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)x_n\|_2 = \lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)\|_2 = \gamma, \quad \mathbb{P}\text{-a.s..}$$

For (b), by Proposition 2.1.10, as $n \to \infty$, $n^{-1} \log \|S_n(\omega)x\|$ converges $\mathbb{P} \otimes \nu$-a.s. to a random variable $\gamma(\omega, x)$ such that

$$\iint \gamma(\omega, x) \, \mathrm{d}\mathbb{P}(\omega) \, \mathrm{d}\nu(x) = \iint \log \|Ax\| \, \mathrm{d}\mu(A) \, \mathrm{d}\nu(x).$$

By part (a), $\gamma(\omega, x) = \gamma$, thus

$$\gamma = \iint \log \|Ax\| \, \mathrm{d}\mu(A) \, \mathrm{d}\nu(x).$$

For (c), let $\{x_n : n \geq 1\}$ be a converging sequence on the unit circle, then by (a)

$$\lim_{n\to\infty} \frac{1}{n} \log \|S_n(\omega)x_n\| = \gamma, \quad \mathbb{P}\text{-a.s..}$$

Note for any invertible matrix M and $\|z\| = 1$,

$$0 = \log \|z\| \leq \log \|M^{-1}\| + \log \|Mz\|,$$

therefore $-\log\|Mz\| \le \log\|M^{-1}\| \le \log^+\|M^{-1}\|$, and

$$\log^-\|Mz\| \le \log^+\|M^{-1}\|.$$

Together with $\log^+\|Mz\| \le \log^+\|M\|$, we have

$$\big|\log\|Mz\|\big| = \log^+\|Mz\| + \log^-\|Mz\| \le \log^+\|M\| + \log^+\|M^{-1}\|, \tag{2.22}$$

Now

$$\begin{aligned}\left|\frac{1}{n}\log\|S_n(\omega)x_n\|\right| &\le \frac{1}{n}\sum_{i=1}^{n}\Big[\log^+\|A_i(\omega)\| + \log^+\|A_i^{-1}(\omega)\|\Big] \\ &\to \mathbb{E}[\log^+\|A_1(\omega)\|] + \mathbb{E}[\log^+\|A_1^{-1}(\omega)\|] \quad (n\to\infty),\end{aligned}$$

by the law of large numbers. Thus by the dominated convergence theorem,

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\log\|S_n(\omega)x_n\|] = \gamma. \tag{2.23}$$

Now if the conclusion is not true, then there exists $\epsilon > 0$ and subsequence n_i such that

$$\sup_{\|x\|=1}\left|\frac{1}{n_i}\mathbb{E}\left[\log\left\|S_{n_i}(\omega)x\right\|\right] - \gamma\right| \ge \epsilon.$$

As $\{x : \|x\| = 1\}$ is compact, the supremum on the left-hand side can be reached, say at the point y_i, therefore

$$\left|\frac{1}{n_i}\mathbb{E}\left[\log\left\|S_{n_i}(\omega)y_i\right\|\right] - \gamma\right| \ge \epsilon, \quad \text{for all } i \ge 1.$$

Then there exists a convergent subsequence $\{y_{i_k}\}$ of $\{y_i\}$ which converges to a non-zero vector. The contradicts (2.23), which tells us that

$$\lim_{k\to\infty}\frac{1}{n_{i_k}}\mathbb{E}\left[\log\left\|S_{n_{i_k}}(\omega)y_{i_k}\right\|\right] = \gamma.$$

□

2.3.4 Contraction on the Projective Space

Now we introduce a metric on the real projective space, namely for $x, y \in \mathbb{R}P^{d-1}$

$$d(x, y) = (1 - \langle x, y\rangle^2)^{1/2} = |\sin \alpha(x, y)|,$$

since $\langle x, y\rangle = \cos \alpha(x, y)$, where $\alpha(x, y)$ is the angle starting from direction x to $y \pmod{2\pi}$ such that $\alpha(x, y) = -\alpha(y, x)$.

It is a metric since for any $-2\pi \leq \alpha, \beta \leq 2\pi$, we have

$$|\sin(\alpha + \beta)| = |\sin \alpha \cos \beta + \sin \beta \cos \alpha| \leq |\sin \alpha| + |\sin \beta|.$$

Now for any $x, y, z \in \mathbb{R}P^{d-1}$, we have $\alpha(x, z) = \alpha(x, y) + \alpha(y, z)$, therefore $d(x, z) \leq d(x, y) + d(y, z)$.

Note we can also write, for $x, y \in \mathbb{R}^d$

$$d(\bar{x}, \bar{y}) = \frac{\|x \wedge y\|}{\|x\|\|y\|},$$

as $\|x \wedge y\|$ is the area of the parallelogram spanned by vectors x and y. Note here and for the rest of this section, $\|\cdot\|$ denotes the ℓ_2-norm.

Given a matrix $A \in M_d(\mathbb{R})$ with its singular value decomposition $A = UDV$ such that $D = \operatorname{diag}(c_1, \ldots, c_d)$. Then we can easily deduce

$$d(A \cdot x, A \cdot y) = \frac{\|Ax \wedge Ay\|}{\|Ax\|\|Ay\|} \leq \frac{\| \wedge^2 A\|\|x \wedge y\|}{\|Ax\|\|Ay\|} = \frac{c_1 c_2 \|x\|\|y\|}{\|Ax\|\|Ay\|} d(x, y). \tag{2.24}$$

If we replace A by $S_n(\omega)$, then this inequality might help us understand the asymptotic behaviour of the distance between two vectors under the action of $S_n(\omega)$. More precisely, if we denote the first two Lyapunov exponent by γ_1 and γ_2, then we have

$$\frac{d(S_n(\omega) \cdot x, S_n(\omega) \cdot y)}{d(x, y)} \leq \frac{c_1(n)c_2(n)}{\|S_n(\omega)\bar{x}\|\|S_n(\omega)\bar{y}\|},$$

therefore

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \frac{d(S_n(\omega) \cdot x, S_n(\omega) \cdot y)}{d(x, y)} \leq \gamma_1 + \gamma_2 - 2\gamma_1 = \gamma_2 - \gamma_1 \quad \mathbb{P}\text{-a.s.}, \tag{2.25}$$

by equation (2.12) and Theorem 2.3.11 when $\mathbb{E}[\log^+ \|A_1(\omega)\|] < \infty$ and $\mathcal{S}_\mu$ is strongly irreducible. Hence, when $\gamma_2 < \gamma_1$, we can expect $d(S_n(\omega) \cdot x, S_n(\omega) \cdot y)$ would converge to

0 exponentially fast. Intuitively, this could be related to the contracting property by Proposition 2.3.10, since when the matrix set is contracting (i.e. has index one), for any non-zero x, $S_n(\omega) \cdot x$ converges to a one-dimensional direction $Z(\omega)$. Therefore, we can expect any two directions $S_n(\omega) \cdot x$ and $S_n(\omega) \cdot y$ would get closer as $n \to \infty$.

The following proposition validates our intuition. It was first proved by Furstenberg and Kifer [18]. Here we provide the simplified proof due to Guivarc'h and Raugi [24].

Proposition 2.3.12. *(Guivarc'h, Raugi) Let μ be a probability measure on $GL(d, \mathbb{R})$ such that $\mathbb{E}[\log^+ \|A\|] < \infty$. If $\mathcal{S}_\mu$ is strongly irreducible and contracting, then $\gamma_1 > \gamma_2$.*

Proof. Consider the normalised action of matrices on matrices (instead of normalised action of matrices on vectors) in Theorem 2.3.11. By (2.21), the assumptions imply that for any matrix $M \neq 0$,

$$\inf_n \frac{\|S_n(\omega)M\|}{\|S_n(\omega)\|} \neq 0, \quad \mathbb{P}\text{-a.s..}$$

And furthermore, when $\mathcal{S}_\mu$ is contracting

$$\lim_{n\to\infty} \frac{c_2(n,\omega)}{c_1(n,\omega)} = \frac{\| \wedge^2 S_n(\omega)\|}{\|S_n(\omega)\|^2} = 0, \quad \mathbb{P}\text{-a.s.} \qquad (\text{since the index is } 1).$$

Therefore for any nonzero matrix M, we have

$$\lim_{n\to\infty} \frac{\|S_n(\omega)M\|^2}{\| \wedge^2 S_n(\omega)\|} = +\infty, \quad \mathbb{P}\text{-a.s..} \tag{2.26}$$

Now consider a normalised matrix action on the following space

$$X = \{(M, N) : \|M\| = \|N\| = 1, M \in M_d(\mathbb{R}), N \in \mathrm{End}(\textstyle\bigwedge^2 \mathbb{R}^d)\}$$

(where $\mathrm{End}(V)$ denotes the set of endomorphisms on vector space V) defined by

$$A(M, N) = (A \cdot M, (\wedge^2 A) \cdot N), \text{ for } A \in GL(d, \mathbb{R}),$$

as well as the additive cocycle defined by

$$\delta(A, (M, N)) = \log \frac{\|AM\|^2}{\|(\wedge^2 A)N\|}, \quad \text{for } A \in GL(d, \mathbb{R}).$$

Let ν be a μ-invariant measure under this action. Note the projections of ν on each component of space X, namely ν_1, ν_2, are also μ-invariant measures under the matrix action on each component. Therefore,

$$\lim_{n\to\infty} \frac{1}{n}\delta(S_n(\omega), (M, N)) = \lim_{n\to\infty} \frac{2}{n} \log \|S_n(\omega)M\| - \lim_{n\to\infty} \frac{1}{n} \log \|(\wedge^2 S_n(\omega))N\|$$

$$= 2\gamma_1 - (\gamma_1 + \gamma_2) = \gamma_1 - \gamma_2.$$

Note $\lim_{n\to\infty} \delta(S_n(\omega), (M, N)) = +\infty$ (by equation (2.26)) implies

$$\gamma_1 - \gamma_2 = \iint \delta(A, (M, N))\, \mathrm{d}\mu(A)\, \mathrm{d}\nu(M, N) > 0$$

by Theorem 2.3.11 (b) and Lemma 2.2 of Chapter II in [4]. Here we state the lemma without a proof. The intuitive explanation is that if an i.i.d. random walk can almost surely reach infinity, then the expectation of each step is positive.

Lemma 2.3.13. *Let S be a topological semigroup acting on a topological space X, $\{Y_n(\omega)\}$ be a sequence of i.i.d. elements of S with distribution μ and δ an additive cocycle on $S \times X$. Suppose that ν is a μ-invariant measure on X satisfying*

(a) $\iint \delta^+(s, x)\, \mathrm{d}\mu(s)\, \mathrm{d}\nu(x) < \infty$;

(b) For $\mathbb{P} \otimes \nu$-almost all (ω, x), $\lim_{n\to\infty} \delta(Y_n(\omega) \cdots Y_1(\omega), x) = +\infty$.

Then $\delta \in L^1(\mathbb{P} \otimes \nu)$ and $\iint \delta(s, x)\, \mathrm{d}\mu(s)\, \mathrm{d}\nu(x) > 0$.

□

Proposition 2.3.14. *Let $\{A_n(\omega) : n \geq 1\}$ be a sequence of i.i.d random matrices with a common distribution μ such that $\mathbb{E}[\log^+ \|A_1(\omega)\|] < \infty$. If $\mathcal{S}_\mu$ is strongly irreducible and contracting then*

(a) For any $x, y \in \mathbb{R}P^{d-1}$,

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \frac{d(S_n(\omega) \cdot x, S_n(\omega) \cdot y)}{d(x, y)} < 0, \quad \mathbb{P}\text{-a.s.},$$

(b) If, moreover, $\mathbb{E}[\log^+ \|A_1^{-1}(\omega)\|] < \infty$, then

$$\overline{\lim_{n\to\infty}} \sup_{x,y\in\mathbb{R}P^{d-1}} \frac{1}{n} \mathbb{E}\left[\log \frac{d(S_n(\omega) \cdot x, S_n(\omega) \cdot y)}{d(x, y)}\right] < 0.$$

Proof. We only need to prove the second statement as the first one is a direct result of (2.25) and Proposition 2.3.12. Now

$$\overline{\lim_{n\to\infty}} \sup_{x,y\in\mathbb{R}P^{d-1}} \frac{1}{n} \mathbb{E}\left[\log \frac{d(S_n(\omega) \cdot x, S_n(\omega) \cdot y)}{d(x, y)}\right]$$

$$\leq \lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\log c_1(n,\omega)c_2(n,\omega)\right] + 2\lim_{n\to\infty}\left\{\sup_{x\neq 0}\mathbb{E}\left[\frac{1}{n}\log\frac{\|x\|}{\|S_n x\|}\right]\right\}$$

$$=\gamma_1+\gamma_2-2\gamma_1<0,$$

by equation (2.12) and statement (c) of Theorem 2.3.11. □

In the rest of this thesis, for any matrix $A \in GL(d,\mathbb{R})$, we denote

$$\ell(A) = \sup\left\{\log^+\|A\|, \log^+\|A^{-1}\|\right\}.$$

It is not hard to prove that for any $M_1, M_2 \in GL(d,\mathbb{R})$

$$\ell(M_1M_2) \leq \ell(M_1) + \ell(M_2);$$

and for any integer $1 \leq p \leq d$, $M \in GL(d,\mathbb{R})$,

$$\left|\log\|\wedge^p M\|\right| \leq p\ell(M). \tag{2.27}$$

In the end of this section, we state a result taken from Section V.2 of [4]. A generalized version of this proposition will be used in Chapter 5.

Proposition 2.3.15. *Let μ be a probability measure on $GL(d,\mathbb{R})$ such that $\mathcal{S}_\mu$ is strongly irreducible and contracting. If for some $\beta > 0$, $\int e^{\beta\ell(M)}\,\mathrm{d}\mu(M) < \infty$ then there exists some α_0 such that when $0 < \alpha \leq \alpha_0$,*

$$\lim_{n\to\infty}\left\{\sup_{\substack{x,y\in\mathbb{R}P^{d-1}\\ x\neq y}}\mathbb{E}\left[\frac{d^\alpha(S_n(\omega)\cdot x, S_n(\omega)\cdot y)}{d^\alpha(x,y)}\right]\right\}^{1/n} < 1.$$

Proof. Consider the normalised action of matrices on the space

$$B = \left\{(x,y) : x \neq y, x, y \in \mathbb{R}P^{d-1}\right\}$$

by

$$A(x,y) = (A\cdot x, A\cdot y), \quad \text{for } A \in GL(d,\mathbb{R}), (x,y) \in B.$$

Then

$$\delta(A,(x,y)) = \log\frac{d(A\cdot x, A\cdot y)}{d(x,y)}$$

defines an additive cocycle on $GL(d,\mathbb{R}) \times B$. Note

$$\sup_{(x,y)\in B} \delta(A,(x,y)) = \sup_{(x,y)\in B} \log \frac{\|Ax \wedge Ay\|}{\|Ax\|\|Ay\|} \frac{\|x\|\|y\|}{\|x \wedge y\|} \leq 4\ell(A),$$

by equation (2.27). Now for $\alpha \in (0, \beta/4)$ define $a_n = \log \left\{ \sup_{(x,y)} \int e^{\alpha\delta(M,(x,y))} \, \mathrm{d}\mu^n(M) \right\}$, then

$$\int e^{\alpha\delta(M,(x,y))} \, \mathrm{d}\mu^{n+m}(M) = \int e^{\alpha\delta(M_1,M_2(x,y))} e^{\alpha\delta(M_2,(x,y))} \, \mathrm{d}\mu^m(M_1) \, \mathrm{d}\mu^n(M_2) \leq e^{a_m} e^{a_n}.$$

Therefore $\{a_n : n \geq 1\}$ satisfies the subadditive inequality. By Lemma 2.1.1, we know a_n/n converges to $\inf_n a_n/n$ as $n \to \infty$. Now we want to prove $\inf_n a_n/n < 0$, or for some n,

$$\sup_{(x,y)\in B} \int e^{\alpha\delta(M,(x,y))} \, \mathrm{d}\mu^n(M) < 1.$$

In fact, by the inequality $e^x \leq 1 + x + x^2 e^{|x|}/2$, we have

$$\begin{aligned} &\sup_{(x,y)\in B} \int e^{\alpha\delta(M,(x,y))} \, \mathrm{d}\mu^n(M) \\ &\leq 1 + \alpha \sup_{(x,y)\in B} \int \delta(M,(x,y)) \, \mathrm{d}\mu^n(M) + \frac{\alpha^2}{2} \int 16[\ell(M)]^2 e^{4\alpha\ell(M)} \, \mathrm{d}\mu^n(M) \end{aligned}$$

Proposition 2.3.14 (b) shows that the second term is negative for some n. Finally the third term can be arbitrarily small if α is sufficiently small. □

Chapter 3

Other Perspectives

In this chapter, we provide two different perspectives on the classical problem of random matrix products, namely the algebraic perspective and the geometric perspective.

Most of the results in Section 3.1 can be found in [27],. We omit the proofs and refer the readers to the original book as the proofs of these results are mostly algebraic and not necessary for this thesis, but we add some intuitive explanations (in the form of remarks) and examples to hopefully help readers understand.

Section 3.2, which aims at providing an intuitive explanation for the products of positive matrices, is original and can be read without any prior knowledge of the classical theory of random matrix products.

3.1 The Algebraic Perspective

In the introduction of Section 2.3, we claimed that the limiting behaviour of

$$S_n(\omega) = A_n(\omega) \cdots A_1(\omega)$$

relates to the "boundary" of $\mathcal{S}_\mu$, the semigroup generated by $\operatorname{supp}(\mu)$. And we have somewhat confirmed this statement by deducing that the range of a limit point of $T_n(\omega)/\|T_n(\omega)\|$ (i.e., the "boundary" of $\mathcal{S}_\mu$ under normalisation), where $T_n(\omega) = A_1(\omega) \cdots A_n(\omega)$, is an r-dimensional linear space $V(\omega)$ for $\mathbb{P}$-almost all ω (and that we could apply this result to $S_n^*(\omega)$), provided that the semigroup $\mathcal{S}_\mu$ has index r. In this section, we will give a different perspective on how to understand $\mathcal{S}_\mu$ and its "boundary".

We denote by $\mathcal{P}(X)$ the set of Borel probability measures on topological space X.

3.1.1 Decomposition of Semigroups

We begin with the definitions of the algebraic terms necessary for the decomposition theorem.

Definition 3.1.1. A **two-sided ideal** of a semigroup $(S, \cdot)$ is a subset I of S such that $IS \cup SI \cup SIS \subset I$. A semigroup $(S, \cdot)$ is called **simple** if it has no proper two-sided ideals. An element $e \in S$ is **idempotent** if $e^2 = e$. Let $E(S)$ denote the set of idempotents in S. An idempotent in S is said to be **primitive** if it is minimal with respect to the partial order $\leq$ on the set $E(S)$ defined by

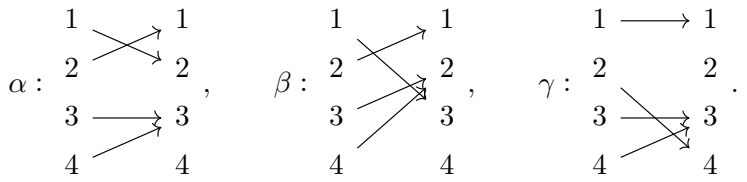
$$e \leq f \quad \text{iff} \quad ef = fe = e \quad e, f \in E(S).$$

A semigroup is called **completely simple** if it is simple and contains a primitive idempotent.

Theorem 3.1.2. *(The Rees-Suschkewitsch Representation Theorem)*
Let $(S, \cdot)$ be a completely simple semigroup. Let e be a primitive idempotent in S. Let $X = E(Se), Y = E(eS)$ and $G = eSe$, where $E(A)$ denotes the set of idempotent elements in A. Then G is a group and S is isomorphic to $X \times G \times Y (\triangleq S_1$, say), where the multiplication in S_1 is given by $(x_1, g_1, y_1)(x_2, g_2, y_2) = (x_1, g_1 y_1 x_2 g_2, y_2)$, and the isomorphism is given by $(x, g, y) \mapsto xgy$.

Remark 3.1.3. The proof can be found in [7] or Chapter 1 of [27]. This theorem might seem abstract at the first glance to readers who are not familiar with semigroup theory. Here we give an explicit example showing how to do the decomposition.

Let $X = \{1, 2, 3, 4\}$ and $(\mathcal{J}_X, \circ)$ be the semigroup of maps from X to X where $\circ$ denotes composition of maps. We denote by id_X the map which sends every element to itself. Obviously id_X is an idempotent. Let $\alpha, \beta, \gamma \in \mathcal{J}_X$ as below,



For each $f \in \mathcal{J}_X$, we can define an equivalence relation $\pi(f)$ by $a\pi(f)b$ iff $f(a) = f(b)$. In the example we have $3\pi(\alpha)4$, $3\pi(\beta)4$ and $3\pi(\gamma)4$.

For $f \in \mathcal{J}_X$, we can also define $R(f)$ to be the range of f and call the cardinality of $R(f)$ the rank of f. Therefore, $R(\alpha) = R(\beta) = \{1, 2, 3\}, R(\gamma) = \{1, 3, 4\}$. And α, β, γ are all of rank 3.

Now we consider the semigroup $S\langle\alpha,\beta\rangle$ generated by α,β. Note first $e := \alpha^2$ is an idempotent as

$$\alpha^2 = \begin{array}{ccccc} 1 & & 1 & & 1 \\ 2 & \searrow\nearrow & 2 & \searrow\nearrow & 2 \\ 3 & \longrightarrow & 3 & \longrightarrow & 3 \\ 4 & \nearrow & 4 & \nearrow & 4 \end{array} = \begin{array}{ccc} 1 & \longrightarrow & 1 \\ 2 & \longrightarrow & 2 \\ 3 & \longrightarrow & 3 \\ 4 & \nearrow & 4 \end{array} =: e.$$

Here e is of rank 3, thus $e\,\mathrm{id}_X = \mathrm{id}_X\,e = e$, i.e., $e \leq \mathrm{id}_X$ by Definition 3.1.1.

Notice $3\pi(\alpha)4, 3\pi(\beta)4$, i.e., both α and β cannot distinguish 3 from 4 in the domain. Now we quotient out this equivalence of 3 and 4 in the domain (and consequently in the range as well), and then the maps can be written as

$$\tilde{\alpha}: \begin{array}{ccc} 1 & & 1 \\ 2 & \searrow\nearrow & 2 \\ 3,4 & \longrightarrow & 3,4 \end{array} \qquad \tilde{\beta}: \begin{array}{ccc} 1 & & 1 \\ 2 & \nearrow\searrow & 2 \\ 3,4 & & 3,4 \end{array},$$

Now $\tilde{\alpha}, \tilde{\beta}$ are of full rank and together can generate a group, namely the symmetric group on 3 symbols. In fact, it can be proved that $S\langle\alpha,\beta\rangle$ is completely simple and this is the group G as described in the theorem, and e works as the quotient action which makes 3 and 4 equivalent. Finally we can write

$$\alpha = \begin{array}{ccccc} 1 & \longrightarrow & 1 & & 1 \\ 2 & \longrightarrow & 2 & \searrow\nearrow & 2 \\ 3 & \longrightarrow & 3 & \longrightarrow & 3 \\ 4 & \nearrow & 4 & & \end{array} = \tilde{\alpha}e, \qquad \beta = \begin{array}{ccccc} 1 & \longrightarrow & 1 & & 1 \\ 2 & \longrightarrow & 2 & \nearrow\searrow & 2 \\ 3 & \longrightarrow & 3 & & 3 \\ 4 & \nearrow & 4 & & \end{array} = \tilde{\beta}e,$$

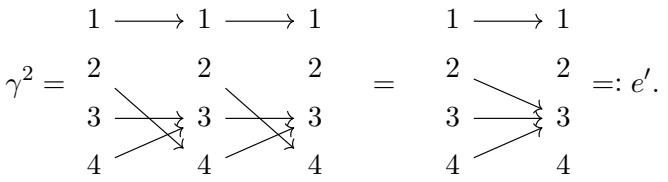
which is the decomposition we desire.

However, if we follow the same procedure for γ, we can notice, although $3\pi(\gamma)4$ in the domain, $\{3,4\}$ in the range has three different pre-images $2,3,4$. Therefore, if we see 3 and 4 as the same element, we will obtain

$$\tilde{\gamma}: \begin{array}{ccc} 1 & \longrightarrow & 1 \\ 2 & \searrow & 2 \\ 3,4 & \longrightarrow & 3,4 \end{array},$$

which is not of full rank. Therefore, $\tilde{\gamma}$ is not invertible like $\tilde{\alpha}$ and $\tilde{\beta}$, i.e., $\tilde{\gamma}$ is not suitable as an element in a group.

As a matter of fact, in the semigroup $S\langle\alpha,\beta,\gamma\rangle$ we can find an idempotent smaller than e, namely



Obviously e' is an idempotent of rank 2 identifying $\{2,3,4\}$ with $e' \leq e$. We can also notice that after identifying $\{2,3,4\}$ both in the domain and in the range, α and β become ill-defined, since $\{2,3,4\}$ is mapped to both $\{1\}$ and $\{2,3,4\}$ at the same time, which implies that we might be able to find an even smaller idempotent. In fact, $e'' := e'\alpha e'$, for example, maps every element to $\{3\}$, which is an idempotent of rank 1.

The procedure cannot be applied to γ (and consequently not α,β,γ simultaneously) simply because γ generates a smaller idempotent $e' \leq e$. This means that $S\langle\gamma\rangle$ (and therefore $S\langle\alpha,\beta,\gamma\rangle$ as well) is not a completely simple semigroup. In fact, if we denote by Q_2 the set of elements of rank not greater than 2 in $S\langle\gamma\rangle$, then Q_2 is a proper two-sided ideal of $S\langle\gamma\rangle$.

In general, let X be a finite set and $(\mathcal{J}_X,\circ)$ be the semigroup of maps from X to X. Let S be a sub-semigroup of $\mathcal{J}_X$. If S is completely simple, then we can see all the elements in S are of the same rank. If S is not completely simple, we can instead consider its proper two-sided ideal Q consisting of elements in S of the lowest rank.

The result is similar in the case of matrices. Let S be a multiplicative matrix semigroup. Then if it is completely simple, it only consists of matrices of the same rank. One might think the assumption of being completely simple is rather strong. In fact, we will see soon that the case we care about the most, matrices in the support of an invariant measure form a completely simple semigroup. The intuition is, if a matrix is on the "boundary" of a semigroup, it has to be of the lowest rank, or we can always multiply it by a matrix of the lowest rank and obtain an element "closer" to the "boundary", which is contradictory to the concept of a "boundary".

3.1.2 Decomposition of Invariant Measures

Now for an invariant probability measure λ on the semigroup S, we have a similar decomposition Theorem 3.1.4 corresponding to Theorem 3.1.2. In the following theorem, we use the notation $S(\lambda)$ instead of our usual notation $\operatorname{supp}\lambda$ to denote the support of λ, in order to make it obvious that $S(\lambda)$ is semigroup. The proof can be found in Theorem 2.4 of [27].

Theorem 3.1.4. *Let $\lambda \in \mathcal{P}(S)$ be such that $\lambda = \lambda * \lambda$. Then $S(\lambda) = \operatorname{supp}\lambda$, the support of λ, is a closed completely simple sub-semigroup of S. Also, if $S(\lambda) = X \times G \times Y$ is the*

product representation for $\operatorname{supp}\lambda$*, then* G *is necessarily a compact subgroup of* S *and there exist* $\lambda_1 \in \mathcal{P}(X)$ *and* $\lambda_3 \in \mathcal{P}(Y)$ *such that* $\lambda = \lambda_1 * \lambda_2 * \lambda_3$*, where* λ_2 *denotes the normed Haar measure on* G*.*

Recall the definition and the significance of tightness in Remark 2.1.9. The proof of the following theorem can be found in Theorem 2.7 of [27].

Theorem 3.1.5. *(Convergence in Distribution for Products of Semigroups)*
Denote $S = \operatorname{cl}(\mathcal{S}_\mu)$*, the closure of* $\mathcal{S}_\mu$*. Let* $\mu \in \mathcal{P}(S)$ *and suppose that the sequence* (μ^n) *is tight. Let* $\mathcal{K} = \{\lambda \in \mathcal{P}(S) : \lambda \text{ is a weak limit point of the sequence } (\mu^n)\}$*. Also define:* $S_0 = \bigcup\{S(\lambda) : \lambda \in \mathcal{K}\}$ *and* $S_1 = \operatorname{cl}(S_0)$*. Then the following assertions hold:*

(a) The sequence $\frac{1}{n}\sum_{k=1}^{n}\mu^k$ *converges weakly to a probability measure* ν *such that* $\nu = \nu * \nu = \nu * \mu = \mu * \nu$*, and* $S(\nu)$ *is the closed completely simple kernel of* S *with a compact group factor.*

(b) The set $\mathcal{K}$ *is a group with respect to convolution as multiplication. The support* $S(\nu)$ *of* ν *coincides with the set* S_1*. If* η *is the identity of* $\mathcal{K}$*, then* $S(\eta) \subset S_1$*, and they have the following product representations:*

$$S_1 = X \times H \times Y, \quad S(\eta) = X \times H_1 \times Y$$

(the usual product representation for completely simple semigroups), where H_1 *is a closed normal subgroup of the compact group* H*, and* X, Y *have their usual meanings. Also if* $\lambda \in \mathcal{K}$ *and* $\lambda \neq \eta$*, then*

$$S(\lambda) = X \times gH_1 \times Y, \quad g \in H \setminus H_1,$$

if the product representation of the idempotent probability measure η *(see Theorem 3.1.4) is* $\eta_1 * \eta_2 * \eta_3$*, then the product representation of* λ *is* $\eta_1 * (\eta_2 * \delta_g) * \eta_3$*, where* $\eta_1 \in \mathcal{P}(X), \eta_3 \in \mathcal{P}(Y)$*, and* η_2 *is the Haar probability measure on* H_1*. (Here, equality for sets is used in the sense of a topological isomorphism.) Furthermore, the set* S_0 *is also a completely simple sub-semigroup of* S*, and* $\mathcal{K} = \{\eta\}$ *if and only if* $\mu * \eta = \eta$*.*

Remark 3.1.6. It is not hard to see that when S is the multiplicative semigroup of $d \times d$ real matrices with the usual topology, $S(\nu)$ consists of matrices of the lowest rank from the semigroup S. Since the matrices of lowest rank always form an ideal, it must be $S(\nu)$ itself since otherwise it would contradict the fact that $S(\nu)$ is simple.

Example 1. Let $X_1 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, X_2 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}$ and $\mu = p\delta_{X_1} + q\delta_{X_2}$ where $p, q > 0, p+q = 1$. Since $X_2X_1 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ therefore $(X_2X_1)^n \to 0$, thus by Theorem 3.1.5 the support $S(\nu)$ of the equilibrium measure ν consists of all the matrices of the minimal rank, which will be a singleton at the zero matrix. This example shows that understanding the asymptotic behaviour of S_n is not enough for us to understand the limit (2.1). However we can try to use the semigroup theory on $S_n/\|S_n\|$.

Now we define a new matrix multiplication: for any two $d \times d$ real matrices A, B, let

$$A \cdot B = \begin{cases} AB/\|AB\|, & \text{if } AB \neq 0; \\ 0, & \text{if } AB = 0. \end{cases} \tag{3.1}$$

Note this is just the normalised action of matrices on themselves given in Definition 2.1.4. Define $T_0 = \{x/\|x\| : x \in S\}$, then $(T_0 \cup \{0\}, \cdot)$ is a semigroup.

Example 2. Let $X_1 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}, X_2 = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}$ and $\mu = p\delta_{X_1} + q\delta_{X_2}$ where $p, q > 0, p+q = 1$. Since $X_1 \cdot X_2 = I_2$, thus under the new matrix multiplication, the set of the matrices of minimal rank is [1]

$$\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\},$$

which we denote by K. Since K is an ideal of $S = \langle X_1, X_2 \rangle$, we have $S \cdot K \subset K$ under the new matrix multiplication (3.1), therefore $S_n \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is a random walk on the set K, where $S_n = A_n \cdots A_1, A_i \overset{i.i.d}{\sim} \mu$. Specifically,

$$X_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

[1] The set of the matrices of minimal rank should be a sub-semigroup of S. However, readers may notice the product of these two elements is zero according to definition (3.1) of the new matrix multiplication. We shall note here the zero matrix actually is not in the support of the invariant measure ν. This fact is due to our new definition of zero in T_0. We should see this 0 as a new element instead of the zero matrix. This semigroup $K \cup \{0\}$ is sometime called a **semigroup with zero**, we can define respective 0-simple, completely 0-simple and have similar Rees-Suschkewitsch Representation Theorem for completely 0-simple semigroup. More detail can be found in Chapter 1 of [27].

and

$$X_2 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

That is to say, every time when X_2 is chosen, it provides a normalisation constant $1/2$. Therefore

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| S_n \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\| = -q \log 2.$$

Similarly,

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| S_n \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\| = -p \log 2.$$

These are equivalent to

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| S_n \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| = -q \log 2,$$

$$\lim_{n\to\infty} \frac{1}{n} \log \left\| S_n \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\| = -p \log 2.$$

From this trivial example we can see the limit (2.2) comes from the normalisation constant when we consider the direction process and the random walk on the completely 0-simple kernel.

We conclude with a simple application of semigroup decomposition theory from [40].

Proposition 3.1.7. *Let $A_n(\omega)$ be an i.i.d. sequence of $d \times d$ non-negative matrices with a common distribution μ such that*

(a) $\mathbb{P}(A_1(\omega)$ has a zero row or a zero column$) = 0$;

(b) $\mathbb{P}(S_n(\omega)$ is strictly positive$) > 0$, for some n.

Then the direction process $Q_n = S_n(\omega)/\|S_n(\omega)\|$ has a limit distribution which is concentrated on the strictly positive matrices of rank one.

Proof. Let J be the set of all $d \times d$ non-negative matrices which do not possess any zero rows or zero columns, and P the set of $d \times d$ strictly positive matrices. Note P is an ideal of J and $\mu^k(P) > 0$ for some k. Now it follows that

$$\mu^{n+1}(P) = \mu^{n+1}(PJ) \geq \mu^n(P)\mu(J) = \mu^n(P) > 0, \quad \text{for } n \geq k.$$

Moreover, if we let $P^{\complement}$ denote the complement of P,

$$\begin{aligned}\mu^{(m+1)k}(P^{\complement}) &= \int_{xy\in P^{\complement}} \mu^{mk}(\mathrm{d}y)\mu^{k}(\mathrm{d}x)\\ &\leq \mu^{mk}(P^{\complement})\mu^{k}(P^{\complement})\\ &\leq [\mu^{k}(P^{\complement})]^{m+1} \to 0 \ (k\to\infty).\end{aligned}$$

Therefore $\mu^{n}(P)$ increases monotonically to 1. Then

$$\lim_{n\to\infty} \mu^{\cdot n}(P) = \lim_{n\to\infty} \mathbb{P}(S_n/\|S_n\| \in P) = \lim_{n\to\infty} \mathbb{P}(S_n \in P) = 1.$$

Hence $\{\mu^{\cdot n} : n \geq 1\}$ is tight. By Theorem 3.1.5, we have

$$\cdot\mu_n := \frac{1}{n}\sum_{k=2}^{n} \mu^{\cdot k} \to \lambda, \quad \text{weakly as } n\to\infty,$$

and $S(\lambda)$ is a completely simple semigroup in J with respect to the normalised matrix product (3.1). Thus, if we consider the factorisation guaranteed by Theorem 3.1.4, the group factor of $S(\lambda)$ is a compact group in J. Since a compact non-negative matrix group with at least one strictly positive matrix must be a singleton (See e.g. [27] Corollary 1.8), the conclusion follows. □

Remark 3.1.8. This theorem is also true for the direction process $Q'_n = T_n(\omega)/\|T_n(\omega)\|$, where $T_n(\omega) = A_1(\omega)\cdots A_n(\omega)$. We can compare this result with Proposition 2.3.10, which seems to be more general but are only applicable to invertible matrices.

This theorem also shows that for any two non-zero vectors $x, y \in \mathbb{R}^d$, the angle between $S_nx/\|S_nx\|$ and $S_ny/\|S_ny\|$ converges to zero (mod π) as $n \to \infty$, which can be seen as a consequence of (2.25). This contracting property of positive matrices product can also be explained in a geometric way, and this will be introduced in the next section.

3.2 The Geometric Perspective

In this section, we will specifically consider the product of an infinite matrix sequence randomly chosen from a set of two 2×2 positive matrices. Let

$$\mathbb{R}^{+}_{2} := \{(x,y) : x>0, y>0\},$$

and let $\sim$ denote the same equivalence relation defining $\mathbb{R}P^{1}$.

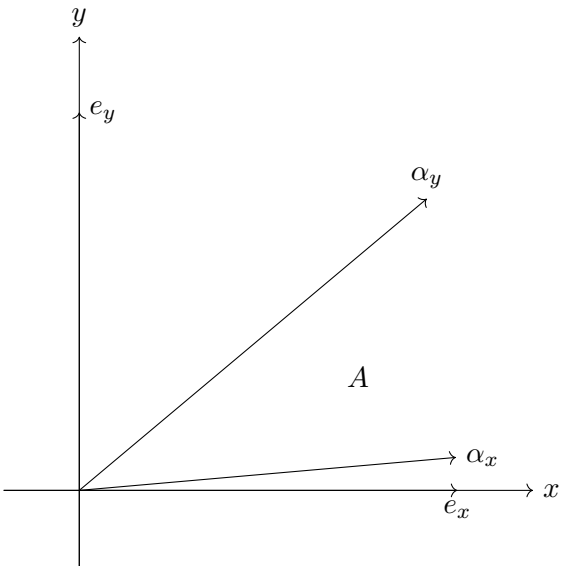


Figure 3.1: A positive 2×2 matrix $A = \begin{pmatrix} \alpha_x & \alpha_y \end{pmatrix}$ with $\det A > 0$.

Given a 2×2 positive matrix A, we can write it as $A = \begin{pmatrix} \alpha_x & \alpha_y \end{pmatrix}$, where α_x and α_y are two 2×1 column vectors, and we can draw these two vectors in a coordinate system as in Figure 3.1. Here we assume α_x is closer to the x-axis than α_y (this assumption is equivalent to $\det(A) > 0$, in this section we assume all the matrices have determinants strictly greater than zero). Also we mark the standard orthonormal basis $\{e_x, e_y\}$. It is easy to see that matrix A takes e_x to α_x, e_y to α_y and all the vectors between e_x and e_y to vectors between α_x and α_y. In other words, A maps $\mathbb{R}^2_+$ to the area encompassed by α_x and α_y, and these two vectors can be seen as the new coordinate basis.

Now if we consider the projective action: $\tilde{A} : x \to Ax/\|Ax\|$ from $\mathbb{R}^2_+/\sim$ to itself, and we take the quarter-circle shown in the Figure 3.2 as a section of $\mathbb{R}^2_+/\sim$. Then $\tilde{A}$ actually maps the whole quarter-circle (without endpoints) to the thick arc curve (without endpoints) shown in the figure.

If we have another 2×2 matrix B which can be written as $B = \begin{pmatrix} \beta_x & \beta_y \end{pmatrix}$ like A, we can draw both of them in the same coordinate system as in Figure 3.3. Here we assume there is no intersection between the area encompassed by α_x and α_y and that encompassed by β_x and β_y except at the origin. It is not hard to prove that the results in this section are still valid after some modification if without this assumption.

Now since

$$AB = A \begin{pmatrix} \beta_x & \beta_y \end{pmatrix} = \begin{pmatrix} A\beta_x & A\beta_y \end{pmatrix},$$

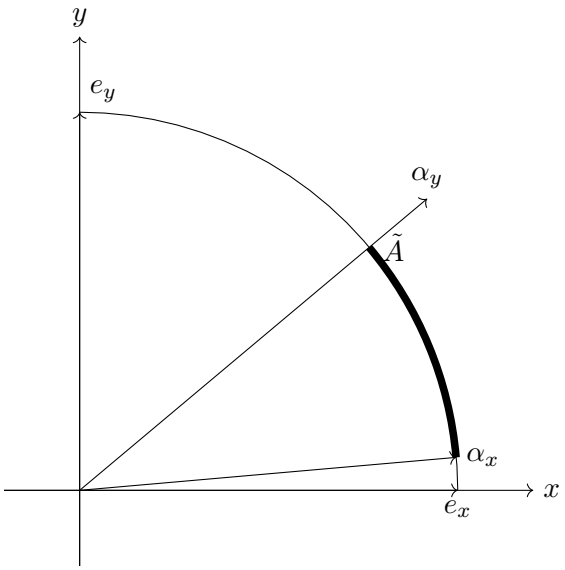


Figure 3.2: $\tilde{A}$ maps $\mathbb{R}^d_+ / \sim$ to the thick arc curve.

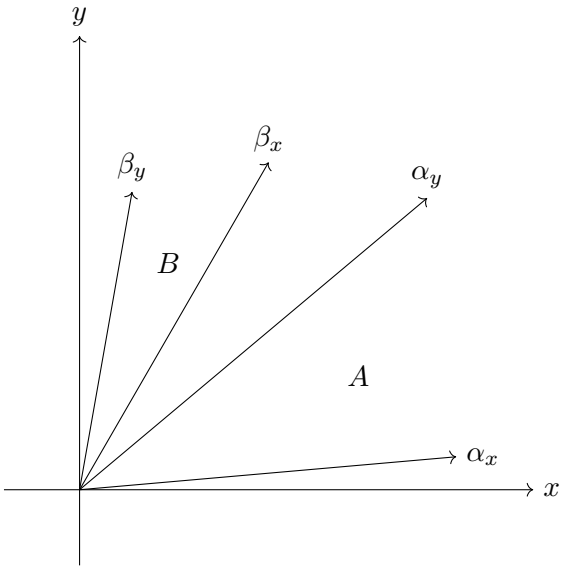


Figure 3.3: Positive matrices A and B at the same coordinate system with no intersection assumed.

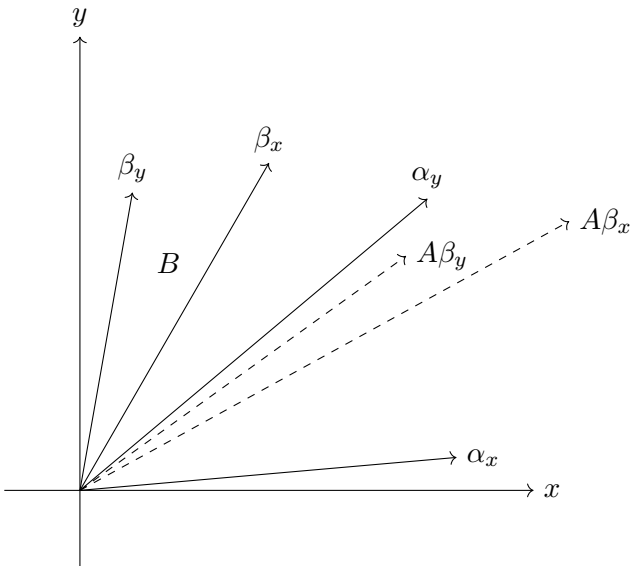


Figure 3.4: The product $AB = \begin{pmatrix} A\beta_x & A\beta_y \end{pmatrix}$ can be seen as the matrix B under the new coordinate system spanned by α_x and α_y.

we can see the product of A and B as a new matrix which maps $\mathbb{R}^2_+$ to the area encompassed by two vectors $A\beta_x$ and $A\beta_y$, where these are vectors β_x and β_y under the new coordinate system $\{\alpha_x, \alpha_y\}$. (See Figure 3.4.)

Similarly, $A^2 = \begin{pmatrix} A\alpha_x & A\alpha_y \end{pmatrix}$ can be found in Figure 3.5. Both of the areas representing AB and A^2 are subsets of the area representing A.

For the same reason, both areas representing B^2 and BA are shown in Figure 3.6, and lie in the area representing B.

To make it clearer, we normalise these vectors and view these matrices as arc curves in the quarter-circle which is a section of the projective space $\mathbb{R}^2_+/\sim$. For example, as shown in Figure 3.7, the thick arc curve (without endpoints) beside the notation $\widetilde{AA}$ is the image of the quarter-circle (without endpoints) under the map of $\widetilde{AA}$. In general, each arc curve(without endpoints) is a duplicate of the quarter-circle(without endpoints) under a contracting map. Note that $\tilde{A} \cdot \tilde{B} = \widetilde{AB}$, so there is no ambiguity of this notation.

With the same technique, we can compute any product of a finite sequence of A and B. And we can see that this procedure is actually analogous to the construction of a Cantor set (See Figure 3.8).

There is a common technique in dynamical systems which can be applied to understand the equilibrium state of this procedure, called symbolic dynamics. The idea here is, each point in

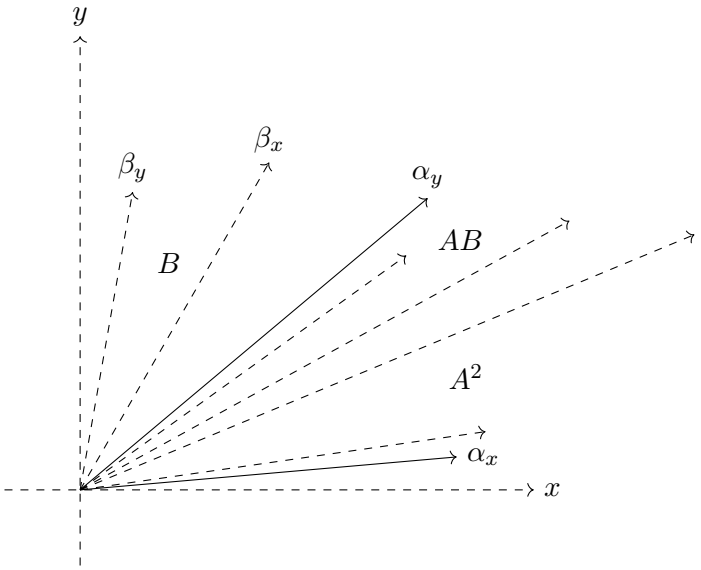


Figure 3.5: The products AB and A^2.

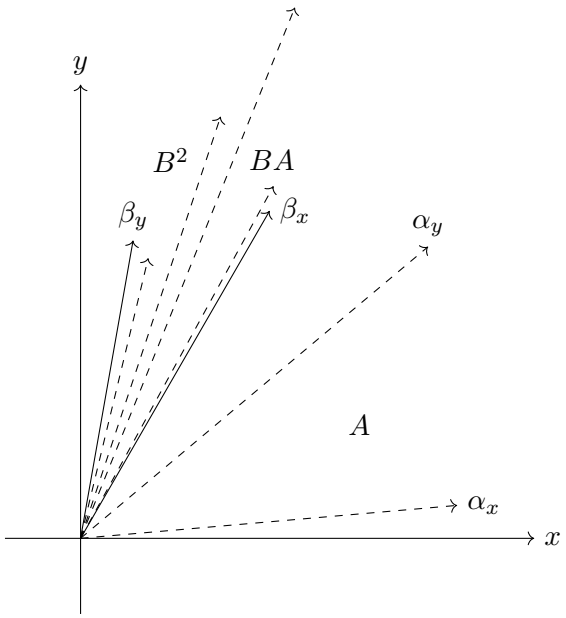


Figure 3.6: The products BA and B^2.

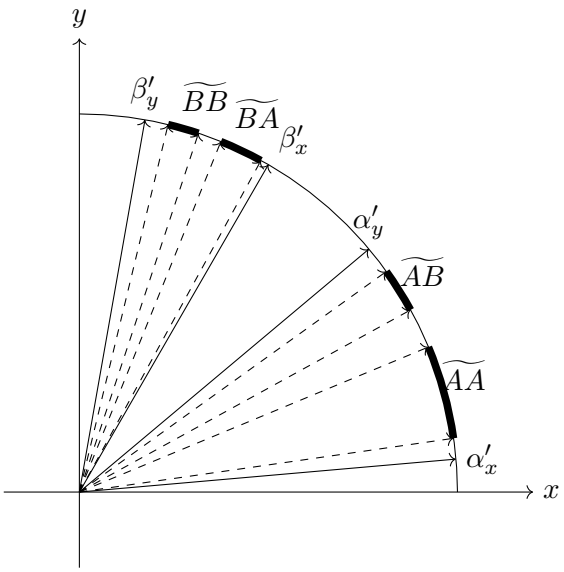


Figure 3.7: The images of $\mathbb{R}^2_+$ under the transformations $\widetilde{AA}, \widetilde{AB}, \widetilde{BA}, \widetilde{BB}$.

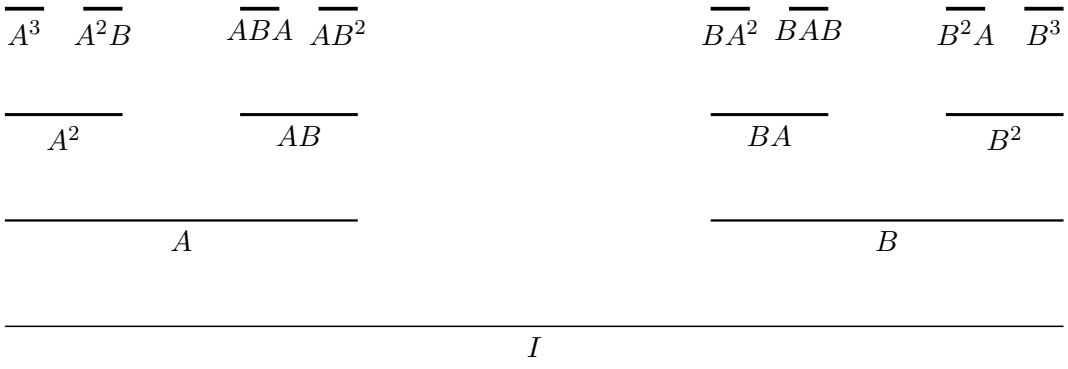


Figure 3.8: The creation of a Cantor set.

the support of the invariant measure corresponds to a unique symbolic sequence. For example, a point p in the Cantor set corresponds to a unique sequence of A and B such that the product of the sequence maps $\mathbb{R}^2_+/\sim$ to that point p. Readers interested in symbolic dynamics can refer to e.g. [5], [44] and [55] .

In a product $M_1 \cdots M_n$ the matrices on the left characterise the general locations of the limit points, while the matrices on the right characterise the detail (locations in higher resolution). For example, in Figure 3.8, we can see that all the limit points of products starting with matrix A are on the left-hand side (obviously, they have to be in the range of A), and then the second "digit" (whether it is A or B) tells whether it should be on the left or right inside the range of A. Therefore reading a product from left to right is like reading a postal address in the Chinese format, where you get increasingly detailed information about what the exact destination is; while reading a product from right to left is similar to reading (from left to right) a postal address in the UK format, it starts with where the receiver stands in his room but might finally end in another universe. Now we can see why Lemma 2.3.9 and Proposition 2.3.10 are true for $T_n(\omega) = A_1(\omega) \cdots A_n(\omega)$ instead of $S_n(\omega) = A_n(\omega) \cdots A_1(\omega)$.

We note here this beautiful construction of the limiting set is straightforward only for positive matrices (or, more generally, a matrix set which is uniformly contracting). In the general setting of $\mathcal{S}_\mu$ merely contracting we can only deduce the average contraction (Proposition 2.3.15) or contraction after sufficiently long time (Proposition 2.3.14).

Chapter 4

Markovian Products of Matrices

In this chapter, we address generalisations of the classical results stated in Chapter 2 to the Markovian setting. Let $A_n(\omega)$ be a Markovian sequence of matrices defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We study the limits (if they exist) (2.1) and (2.2). Figure 4.1 is the outline for the relation between main results in Chapter 2 and those in this chapter. A horizontal arrow signifies a generalisation and a vertical arrow an implication.

Some of these generalised results have already been proved in earlier literatures. For example, the existence of the Lyapunov exponent and Theorem 4.3.3 were given by Furstenberg and Kesten [19], and they follow immediately from the subadditive ergodic theorem. A version of Theorem 4.3.7 and Proposition 4.3.8 when the matrix set is contracting, due to Guivarc'h and Raugi, can be found in [24] [25].

However, we decide to choose an alternative route to state these results using the language of random dynamical systems. It turns out to be more friendly for a discussion of possible further generalisations in this and later chapters.

In Section 4.1, we introduce the definition of a shift space. The essential background on random dynamical systems will be given in Section 4.2. We will assume our probability spaces to be full shift spaces (whose definition is given by Definition 4.1.2). The generalised results will be stated in Section 4.3. In this chapter, $\mathbb{N}$ denotes the set of non-negative integers.

4.1 Shift Spaces

Definition 4.1.1. Let E be a measure space and $\mathbb{T} = \mathbb{N}$ or $\mathbb{Z}$. Then $(\Omega, \mathcal{F}, \sigma)$ is said to be a **shift space** over E if

$$\Omega = E^{\mathbb{T}} = \{\omega = (\omega_i)_{i\in\mathbb{T}} : \omega_i \in E\}$$

endowed with the product topology, the Borel sigma-algebra $\mathcal{F}$ generated by cylinder sets, as well as the **shift map** $\sigma : \Omega \to \Omega$ satisfying $(\sigma\omega)_n = \omega_{n+1}$ for each $n \in \mathbb{T}, \omega \in \Omega$.

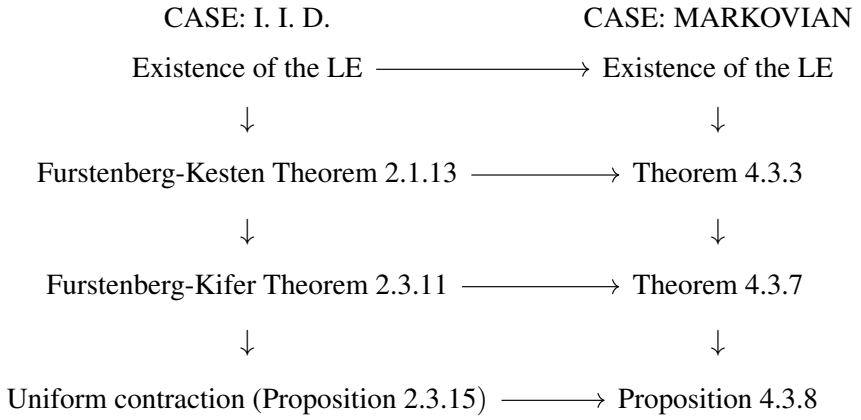


Figure 4.1: Outline of Chapter 4.

The shift space $(\Omega, \mathcal{F}, \sigma)$ is said to be a **shift of finite type** if E is a finite set with the natural discrete topology. In particular, if the cardinality of E is k, we say $(\Omega, \mathcal{F}, \sigma)$ is a **shift over** k **symbols**. When $\mathbb{T} = \mathbb{Z}$, a shift space is said to be **two-sided**; and when $\mathbb{T} = \mathbb{N}$, we say it is **one-sided**. A one-sided shift is sometimes denoted by $(\Omega^+, \mathcal{F}^+, \sigma)$.

Definition 4.1.2. Assume Ω is a one-sided shift over k symbols. Given a $k \times k$ non-negative matrix Q, i.e. the i, j-th entry $Q(i, j) \geq 0$ for any $1 \leq i, j \leq k$, define

$$\Omega_Q = \{\omega \in \Omega : Q(\omega_i, \omega_{i+1}) > 0 \text{ for each } i \in \mathbb{T}\}.$$

An element in Ω_Q is called an **admissible sequence**. Since Ω_Q is a subspace of Ω, let $\mathcal{F}_Q$ be the sigma-algebra $\mathcal{F}$ restricted to the subspace Ω_Q. Then $(\Omega_Q, \mathcal{F}_Q, \sigma)$ is called a **shift induced by matrix** Q. Here Q is usually taken to be a 0-1 matrix without loss of generality.

If entries of Q are all positive, we also say Ω_Q is a **full shift** and we can obviously identify it with Ω.

Definition 4.1.3. Let $\Omega = E^{\mathbb{T}}$ be a shift space. Endow $(\Omega, \mathcal{F}, \sigma)$ with a probability measure $\mathbb{P}$. If such a $\mathbb{P}$ is invariant under σ, i.e., for any finite $m \geq 1$ and measurable cylinder set $C_0 \times C_1 \times \cdots \times C_{m-1} \subset E^m$

$$\mathbb{P}(\omega_0 \in C_0, \ldots, \omega_{m-1} \in C_{m-1}) = \mathbb{P}(\omega_1 \in C_0, \ldots, \omega_m \in C_{m-1}),$$

we say $(\Omega, \mathcal{F}, \mathbb{P})$ is **stationary**. For the rest of this thesis, whenever we write $(\Omega, \mathcal{F}, \mathbb{P})$, it is assumed that this space is stationary.

We say the shift space $(\Omega, \mathcal{F}, \mathbb{P})$ is **Markovian** (resp. **i.i.d.**, or **ergodic**) if $\mathbb{P}$ is Markovian (resp. i.i.d., or ergodic).

4.2 Random Dynamical Systems

Definition 4.2.1. Let $(X, \mathcal{B}, \mu)$ be a measure space, $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ be a shift space. Let

$$\varphi : \mathbb{T} \times \Omega \times X \to X, \quad (n, \omega, x) \mapsto \varphi(n, \omega, x)$$

be a $(\mathcal{B}(\mathbb{T}) \otimes \mathcal{F} \otimes \mathcal{B}, \mathcal{B})$-measurable function. We may also write $\varphi(n, \omega, x)$ as $\varphi(n, \omega)x$ when we consider $\varphi(n, \omega)$ as an endomorphism of X. We call φ a **cocycle**, or a (measurable) **discrete-time random dynamical system** (RDS) acting on X if it satisfies the following conditions:

(C1) $\varphi(0, \omega) = \mathrm{id}_X$ for any $\omega \in \Omega$;

(C2) $\varphi(m + n, \omega) = \varphi(m, \sigma^n \omega) \circ \varphi(n, \omega)$ for any $m, n \in \mathbb{T}, \omega \in \Omega$.

Moreover, if X is a finite-dimensional vector space and $\varphi(n, \omega)$ is a linear transformation of X for each $n \in \mathbb{T}, \omega \in \Omega$, then we call φ a **linear cocycle**, or a **linear random dynamical system** (LRDS).

Given an RDS φ on the measurable space $(X, \mathcal{B})$ over a one-sided shift $(\Omega^+, \mathcal{F}^+, \mathbb{P}, \sigma)$, if we denote $\varphi(1, \omega)$ by $M(\omega)$, then by the cocycle property (C2) of φ, the random dynamical system is fully characterised by $M(\omega)$ in the sense that

$$\varphi(n, \omega) = M(\sigma^{n-1}\omega) \circ M(\sigma^{n-2}\omega) \circ \cdots \circ M(\omega). \tag{4.1}$$

If φ is an RDS over two-sided time $\mathbb{T} = \mathbb{Z}$, by (C1) and (C2), we have $\mathrm{id}_X = \varphi(0, \omega) = \varphi(-1, \omega) \circ \varphi(1, \sigma^{-1}\omega)$, i.e., if again we denote $\varphi(1, \omega)$ by $M(\omega)$, $\varphi(-1, \omega) = M(\sigma^{-1}\omega)^{-1}$. Note this determines the behaviour of φ in negative time, therefore φ is also fully characterised by $M(\omega)$ in the case of two-sided time.

If φ is an LRDS, after fixing a basis, $\varphi(1, \omega) = M(\omega)$ can be regarded as a matrix with a variable ω, whereas the operation "$\circ$" is the matrix multiplication. When $M(\omega)$ depends only on the first coordinate ω_0 and $\mathbb{P}$ is i.i.d., $\varphi(n, \omega)$ reduces to an i.i.d. product of matrices which has been discussed in Chapter 2. Similarly as in Chapter 2, if we consider the concept of an LRDS as being induced by a natural matrix action on some vector space, we can also consider the normalised action and give the following definition.

Definition 4.2.2. Let $\mathrm{Proj}(X)$ be the associated projective space of a finite-dimensional vector space X and φ is a linear cocyle. Then we can define another cocyle $\bar{\varphi}$ on $\mathbb{T} \times \Omega \times \mathrm{Proj}(X)$ by

$$\bar{\varphi}(n, \omega)\pi(x) = \frac{\varphi(n, \omega)x}{\|\varphi(n, \omega)x\|}, \quad 0 \neq x \in X,$$

where $\pi : X \to \mathrm{Proj}(X)$ in the natural quotient map. Such a $\bar{\varphi}$ is called a **normalised cocyle**, or a **normalised random dynamical system** (NRDS) induced by the LRDS φ. It is obvious that this definition does not depend on the choice of the representative x of $\pi(x)$. Sometimes $\bar{\varphi}(n,\omega)\pi(x)$ is also written as $\varphi(n,\omega)\cdot x$.

Remark 4.2.3. Recall that in Chapter 2 and Section 3.1, we study the i.i.d. products of random matrices by investigating the semigroup generated by them. This is plausible due to the fact that the set of possible finite products forms a semigroup. More precisely, if we use the new language here, since the corresponding one-sided LRDS is fully characterised by $\varphi(1,\omega) = M(\omega)$ which depends only on the first coordinate ω_0, for any $m, n \geq 0$, $\omega, \omega' \in \Omega^+$, we have

$$\varphi(n,\omega') \circ \varphi(m,\omega) = \varphi(m+n,\omega''), \tag{4.2}$$

where $\omega'' \in \Omega^+$ can be given by $(\omega'')_k = \omega_k$ when $0 \leq k \leq m-1$ and $(\omega'')_k = (\omega')_{k-m}$ for $k \geq m$. Thus $\{\varphi(n,\omega) : n \geq 0, \omega \in \Omega^+\}$ forms a semigroup.

We might expect to do the same for a general product, but we will soon realise that when Ω^+ is a general one-sided shift space, $\{\varphi(n,\omega) : n \geq 0, \omega \in \Omega^+\}$ does not necessarily form a semigroup. For example, in the construction of ω'' above, if the jump from ω_{k-1} to ω'_k has probability 0, this ω'' is forbidden. In general, we cannot find any ω'' satisfying (4.2) if the concatenation of $(\omega_0, \ldots, \omega_{k-1})$ and ω' is not allowed.

Remark 4.2.4. In this dissertation, unless otherwise indicated, we will mainly discuss when the shift space is a one-sided full shift of finite type, and therefore will not distinguish Ω^+ from Ω_M^+. Obviously under this assumption $\{\varphi(n,\omega) : n \geq 0, \omega \in \Omega^+\}$ still forms a semigroup.

Note a simple technique we can try when the shift space is not full is that, we can fix a state $j \in E$ and find all the admissible strings $\{s_\iota\}_{\iota \in I}$ in each of which j returns to j for the first time. By grouping each string s_ι as a new state, we can regard the shift space as a full shift space over $|I|$ symbols (it is not of finite type when $|I| = \infty$) without loss of generality.

For instance, consider a shift space Ω^+ over 2 symbols consisting of transitions $1 \to 1, 1 \to 2$ and $2 \to 1$ but with $2 \to 2$ forbidden, together with an LRDS φ given by

$$\varphi(1,\omega) = M_1, \text{ when } \omega_0 = 1; \quad \varphi(1,\omega) = M_2, \text{ when } \omega_0 = 2.$$

Choose state $j = 1$, then we can see all the possible paths from 1 to 1 (for the first time) are

$$\{s_\iota\}_{\iota \in I} = \{1(\to 1),\ 1 \to 2(\to 1)\},$$

then we can group $1 \to 2$ as a new state, consider a shift space $\tilde{\Omega}^+$ over $|I| = 2$ symbols $\{1, 1 \to 2\}$, and assume the new LRDS $\tilde{\varphi}$ to be given by

$$\tilde{\varphi}(1,\tilde{\omega}) = M_1, \text{ when } \tilde{\omega}_0 = 1; \quad \tilde{\varphi}(1,\tilde{\omega}) = M_1 M_2, \text{ when } \tilde{\omega}_0 = (1 \to 2).$$

It is not hard to see that $\tilde{\Omega}^+$ is a full shift, and the Lyapunov exponents (whose definition will be given by (4.3)) associated with φ and $\tilde{\varphi}$ are the same after assigning an appropriate probability measure on $\tilde{\Omega}^+$.

Naturally, when $\{\varphi(n,\omega) : n \geq 0, \omega \in \Omega^+\}$ forms a semigroup, we can define irreducibility and the contracting property as follows.

Definition 4.2.5. Given an LRDS (or an NRDS) φ, if $\varphi(1,\omega) = M(\omega)$ is invertible for each ω, we say φ is **(spatially) invertible**.

Definition 4.2.6. Assume Ω is a full shift and φ is a spatially invertible LRDS fully charcterised by $\varphi(1,\omega) = M(\omega)$, denote by $\mathcal{S}_\varphi$ the semigroup

$$\mathcal{S}_\varphi := \{\varphi(n,\omega) : n \in \mathbb{T}, \omega \in \Omega\}$$

under the usual matrix multiplication, then for the sake of convenience when $\mathcal{S}_\varphi$ is irreducible (resp. strongly irreducible, contracting, of index r), we say φ is **irreducible** (resp. **strongly irreducible**, **contracting**, **of index r**).

Definition 4.2.7. Given an RDS φ acting on X, then the mapping

$$\Sigma : (\omega, x) \mapsto (\sigma\omega, \varphi(1,\omega)x)$$

is a measurable dynamical system on $(\Omega \times X, \mathcal{F} \times \mathcal{B})$, which is called the **skew product** of the shift space Ω and the cocycle φ, or the **skew product induced by the RDS** φ.

Remark 4.2.8. Let φ be an LRDS over one-sided time. If $\varphi(1,\omega) = M(\omega)$ depends only on the first coordinate ω_0 then for any $n \geq 1, \omega = (\omega_0, \omega_1, \dots)$, we have

$$\Sigma^n : (\omega, x) \mapsto \big(\sigma^n\omega, M(\omega_{n-1}) \cdots M(\omega_0)x\big) = (\sigma^n\omega, \varphi(n,\omega)x).$$

Note an LRDS also induces an NRDS $\bar{\varphi}$ on the projective space $\mathrm{Proj}(X)$ by Definition 4.2.2. Therefore we have the corresponding skew product $\Theta : (\omega, x) \mapsto (\sigma\omega, M(\omega) \cdot x)$.

Definition 4.2.9. Given an RDS φ on $(X, \mathcal{B})$ over a shift space Ω, denote by Σ the skew product of Ω and φ on X, then a probability ν on $(\Omega \times X, \mathcal{F} \times \mathcal{B})$ is called an **invariant measure** for φ (or for Σ) if it satisfies:

$$(1)\ \ \Sigma\nu = \nu, \qquad (2)\ \ \pi_\Omega\nu = \mathbb{P},$$

where $\pi_\Omega : \Omega \times X \to \Omega$ is the natural projection $\pi_\Omega(\omega, x) = \omega$.

Remark 4.2.10. When φ is an NRDS on $\mathrm{Proj}(X)$, since $\mathrm{Proj}(X)$ is compact and φ is continuous, by Theorem 1.5.10 of [1] (or the first result Proposition 4.3.1 in next section) an invariant measure exists; and since X is Polish, by Proposition 1.4.3 of [1], there exists a unique factorisation of the invariant measure, i.e., symbolically we can write $\nu(\mathrm{d}\omega, \mathrm{d}\bar{x}) = \nu_\omega(\mathrm{d}\bar{x})\mathbb{P}(\mathrm{d}\omega)$.

Remark 4.2.11. With the decomposition above, we can obtain ν_ω from a φ-invariant measure ν. And we can notice that in Lemma 2.3.9, we also introduced a measure named ν_ω satisfying

$$\nu = \int \nu_\omega \mathbb{P}(\mathrm{d}\omega).$$

We remark there these two symbols do not represent the same thing, but are closely related.

Remark 4.2.12. Note one can easily restrict an RDS φ on $\mathbb{Z} \times \Omega$ with two-sided time to $\varphi^+ = \varphi|_{\mathbb{N}\times\Omega^+}$ with one-sided time.

If φ is a spatially invertible RDS with one-sided time, then we can extend it to an RDS $\tilde{\varphi}$ with two-sided time by defining $\tilde{\varphi}(1, \omega) = \varphi(1, r^+(\omega))$ and $\tilde{\varphi}(-1, \omega) = \varphi(1, r^+(\sigma^{-1}\omega))^{-1}$ for $\omega \in \Omega$, where $r^+ : \Omega \to \Omega^+$ is the natural restriction map. By Theorem 1.7.2 of [1], there exists a one-to-one correspondence between the invariant measures for φ and those for $\tilde{\varphi}$.

4.3 Markovian Products of Matrices

In this section, we will generalise the results in Chapter 2 to the Markovian products of matrices with the language of random dynamical systems. We remark here the finite expectation condition $\mathbb{E}[\log^+ \|M(\omega)\|] < \infty$ is not necessary when Ω^+ is a shift over finite symbols.

We shall also note that the Lyapunov exponent is well-defined for any stationary process of random matrices according to Definition 2.1.2, thus we can write

$$\gamma = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\log \|\varphi(n, \omega)\|] \tag{4.3}$$

as the Lyapunov exponent for φ over a one-sided shift.

We start with the existence of an invariant measure for an RDS.

Proposition 4.3.1. *(Generalisation of Proposition 2.1.8) Let $(\Omega^+, \mathcal{F}^+, \mathbb{P}, \sigma)$ be a one-sided shift, φ be an RDS on a compact, metrisable space X, then an invariant measure for φ exists. Furthermore, if m is a probability measure on $\Omega^+ \times X$ such that $\pi_{\Omega^+} m = \mathbb{P}$, then each limit point of*

$$\nu_n = \frac{m + \Sigma m + \cdots + \Sigma^{n-1} m}{n}$$

in the weak*-topology is an invariant measure for φ, where Σ is the skew product of Ω^+ and φ.

Proof. The proof is similar, and we only need to notice $\Sigma\nu_n = \nu_n + (\Sigma^n m - m)/n$. □

Now we prove a generalisation of the first theorem of Furstenberg and Kesten to the ergodic case. Note an irreducible Markovian shift is always ergodic (See e.g. [46] pp.51).

Lemma 4.3.2. *(Generalisation of Corollary 2.1.12) Let S be a topological semigroup acting on a topological space X, $(\Omega^+, \mathcal{F}^+, \mathbb{P}, \sigma)$ be a one-sided shift, φ be an RDS on X fully characterised by $\varphi(1, \omega) = Y(\omega) \in S$, and ν be an invariant measure for φ. Let δ be an additive cocycle on $S \times X$. If*

$$\int |\delta(Y(\omega), x)| \nu(\mathrm{d}\omega, \mathrm{d}x) < \infty, \tag{4.4}$$

then

$$\lim_{n\to\infty} \frac{1}{n} \delta(Y(\omega_{n-1}) \cdots Y(\omega_0), x)$$

exists for ν-almost all (ω, x) and in $L^1(\nu)$. Moreover, if we denote the limit by $\gamma(\omega, x)$, then

$$\iint \gamma(\omega, x)\nu(\mathrm{d}\omega, \mathrm{d}x) = \iint \delta(Y(\omega), x)\nu(\mathrm{d}\omega, \mathrm{d}x).$$

Proof. By Definition 4.2.9, the skew product Σ preserves the invariant measure ν on X. Thus we can still apply Birkhoff's ergodic theorem and obtain that

$$\frac{1}{n}(\delta + \delta \circ \Sigma + \cdots + \delta \circ \Sigma^{n-1})(\omega, x) \to \gamma(\omega, x) \in L^1(\nu), \quad \nu\text{-a.s.}$$

□

Theorem 4.3.3. *(Generalisation of Theorem 2.1.13) Let $(\Omega^+, \mathcal{F}^+, \mathbb{P}, \sigma)$ be a one-sided ergodic shift, φ be an invertible LRDS determined by $\varphi(1, \omega) = M(\omega)$. If $\mathbb{E}[\log^+ \|M(\omega)\|] < \infty$, Then the limit*

$$\lim_{n\to\infty} \frac{1}{n} \log \|\varphi(n, \omega)\|$$

exists for $\mathbb{P}$-almost all ω. If we denote the limit by $\gamma(\omega)$, then $\gamma(\omega)$ equals the Lyapunov exponent $\mathbb{P}$-a.s..

Proof. Let $\Theta : (\omega, x) \mapsto (\sigma\omega, M(\omega) \cdot x)$ be the skew product for the NRDS induced by the LRDS φ. Assume ν is the invariant measure with respect to Θ with factorisation $\nu(\mathrm{d}\omega, \mathrm{d}M) = \nu_\omega(\mathrm{d}M)\mathbb{P}(\mathrm{d}\omega)$. Note the existence of such a factorisation is guaranteed by Remark 4.2.10.

As before, we first assume $\gamma \neq -\infty$ and go through the similar four steps of the proof of Theorem 2.1.13.

Step one, let $m(\mathrm{d}M, \mathrm{d}\omega) = \rho_{I_d}(\mathrm{d}M)\mathbb{P}(\mathrm{d}\omega)$ where ρ_{I_d} is the singleton measure on I_d and let $\nu_n = (m + \Theta m + \cdots + \Theta^{n-1}m)/n$ as in Proposition 4.3.1. Then by Proposition 4.3.1, ν_n weakly converges to an invariant measure for Θ through a subsequence.

Note for $i \geq 1$,

$$\begin{aligned}\int \log \|M\|(\Theta^i m)(\mathrm{d}\omega, \mathrm{d}M) &= \int \log \frac{\|\varphi(i,\omega)M\|}{\|\varphi(i-1,\omega)M\|} m(\mathrm{d}\omega, \mathrm{d}M) \\ &= \int \log \frac{\|\varphi(i,\omega)\|}{\|\varphi(i-1,\omega)\|} \mathbb{P}(\mathrm{d}\omega).\end{aligned}$$

Thus

$$\int \log \|M\| \nu_n(\mathrm{d}\omega, \mathrm{d}M) = \frac{1}{n}\mathbb{E}[\log \|\varphi(n,\omega)\|] \to \gamma, \quad (n \to \infty)$$

by the definition of the Lyapunov exponent.

In step 2 and 3 we shall apply Lemma 4.3.2 to the additive cocycle $\delta(Y, x) = \log \|Yx\|$, and verify the integrability condition (4.4). The proof is the same as the step 2 and 3 of Theorem 2.1.13.

Finally, we can see step 4 for this theorem only requires a generalisation of equation (2.10). In fact, it is still valid if we notice $\varphi((j+1)k, \omega)\varphi^{-1}(jk, \omega) = \varphi(k, \sigma^{jk}\omega)$ and that

$$\frac{1}{kq}\sum_{j=1}^{q-1} \log \|\varphi((j+1)k,\omega)\varphi^{-1}(jk,\omega)\| \to \frac{1}{k}\mathbb{E}[\log \|\varphi(k,\omega)\|], \quad \mathbb{P}\text{-a.s.},$$

by Birkhoff's ergodic theorem. The case when $\gamma = -\infty$ is also proved consequently. □

Theorem 2.1.13 can also be a direct consequence of the subadditve ergodic theorem. In order to generalise other main results such as Theorem 2.3.11 and Proposition 2.3.15, we need the following lemma to understand the invariant measure ν on $\Omega^+ \times \mathcal{X}$ in the case of Markovian shifts. This can be considered as a generalisation for Lemma 2.3.9.

Lemma 4.3.4. *Let $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ be a two-sided Markovian full shift over k symbols, ψ be an RDS acting on the measure space X, determined by $\psi(1, \omega) = M(\omega)$, and assume $M(\omega)$ depends only on the first coordinate ω_0. Let ν be an invariant measure for $\psi|_{\mathbb{N}\times\Omega^+}$, and $\nu(\mathrm{d}\omega, \mathrm{d}x) = \nu_\omega(\mathrm{d}x)\mathbb{P}(\mathrm{d}\omega)$ be its unique factorisation by Remark 4.2.10. Then*

(a) ν_ω only depends on the coordinate ω_0. In other words, there exist k probability measures $\nu_1, \ldots, \nu_k$ on X such that $\nu = \sum_{i=1}^k q_i \nu_i$, where $q_i = \mathbb{P}(\omega_0 = i)$;

(b) $\psi(n,\sigma^{-n}\omega)\nu_{\sigma^{-n}\omega}$ *weakly converges to a probability measure* $\tilde{\nu}_\omega$ *as* $n\to\infty$*, where* $\tilde{\nu}$ *is an invariant measure for* ψ *with two-sided time;*

(c) For $1\le j\le k$*, define* $\Lambda_j := \{SM_j : S\in\mathcal{S}_\psi\}$*, and* $\eta_j = \sum_{m=0}^{\infty} 2^{-(m+1)}P^m * e_j$*, where* e_j *denotes the singleton measure on the* j*-th matrix, then for* η_j*-almost all* $Q_j\in\Lambda_j$

$$\lim_{n\to\infty}\psi(n,\sigma^{-n}\omega)Q_j\nu_j = \tilde{\nu}_\omega.$$

Proof. For (a), since ν_ω is $\mathcal{F}^+$-measurable, $\nu_\omega = \mathbb{E}[\nu_\omega|\mathcal{F}^+] = \mathbb{E}[\nu_\omega|\omega_0]$, is a function depending only on ω_0. Let ν_i be ν_ω when $\omega_0 = i$, then

$$\nu = \int \nu_\omega \mathbb{P}(\mathrm{d}\omega) = \sum_{i=1}^{k} q_i\nu_i,$$

where $q_i = \mathbb{P}(\omega_0 = i)$.

The proof of (b) can be found in [1] Theorem 1.7.2. Define

$$\mu_{n,\omega} = \psi(n,\sigma^{-n}\omega)\nu_{\sigma^{-n}\omega},\quad \mathcal{F}^+_{-m} = \sigma^m\mathcal{F}^+,$$

then we can see $\mu_{n,\omega}$ is a martingale with respect to the filtration $\{\mathcal{F}^+_{-m} : m\ge 0\}$.

For (c), for any bounded Borel function f on X,

$$\begin{aligned}
&\sum_{n=1}^{\infty}\mathbb{E}\left[\int\left|\psi(n,\sigma^{-n}\omega)Q_j\nu_j(f) - \psi(n,\sigma^{-n}\omega)\nu_{\sigma^{-n}\omega}(f)\right|^2 \mathrm{d}\eta_j(Q_j)\right]\\
\le&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}2^{-(m+1)}\mathbb{E}\left[\int\left|\psi(n,\sigma^{-n}\omega)\psi(m,\sigma^{-m-n}\omega)M_j\nu_j(f) - \psi(n,\sigma^{-n}\omega)\nu_{\sigma^{-n}\omega}(f)\right|^2\right]\\
=&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}2^{-(m+1)}\mathbb{E}\left[\int\left|\psi(m+n,\sigma^{-m-n}\omega)M_j\nu_j(f) - \psi(n,\sigma^{-n}\omega)\nu_{\sigma^{-n}\omega}(f)\right|^2\right]\\
\le&\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}2^{-(m+1)}\mathbb{E}\left[\int\left|\psi(m+n,\sigma^{-m-n}\omega)\nu_{\sigma^{-(m+n)}\omega}(f) - \psi(n,\sigma^{-n}\omega)\nu_{\sigma^{-n}\omega}(f)\right|^2\right]\\
\le&\sum_{m=1}^{\infty}m2^{-m}\|f\|_\infty^2 < \infty,
\end{aligned}$$

where the last inequality follows by the fact that $\mu_{n,\omega}$ is a martingale with respect to the filtration $\{\mathcal{F}^+_{-m} : m\ge 0\}$ and (2.19). The rest of the proof follows in the same way as in Lemma 2.3.9. □

Similar to Proposition 2.3.5, we can study the properness of an invariant measure on $\Omega^+ \times \mathrm{Proj}(X)$ under an LRDS.

Proposition 4.3.5. *If $(\Omega^+, \mathcal{F}^+, \mathbb{P}, \sigma)$ is a one-sided Markovian full shift over k symbols, φ is a strongly irreducible and spatially invertible LRDS determined by $\varphi(1, \omega) = M(\omega)$, and $M(\omega)$ depends only on the first coordinate ω_0. Let $\bar{\varphi}$ be the NRDS acting on $\mathrm{Proj}(X)$ induced by φ, and ν be the invariant measure for $\bar{\varphi}$, then for any hyperplane H of X, we have*

$$0 = \int \nu_\omega(\overline{H})\mathbb{P}(\mathrm{d}\omega),$$

where $\overline{H}$ is the image of H under the natural quotient map $\pi : X \to \mathrm{Proj}(X)$.

Proof. Denote by $\Gamma(l)$ the set of l-dimensional linear subspaces of X. By Lemma 4.3.4, there exist k probability measures $\nu_1, \ldots, \nu_k$ such that $\nu = \sum_{i=1}^k q_i \nu_i$. For each $1 \le i \le k$, define

$$l_i := \inf\{l \le d : \nu_i(\overline{W}) \neq 0, \text{ for some } W \in \Gamma(l)\},$$

$$r_i := \sup\left\{\nu_i(\overline{W}) : W \in \Gamma(l_i)\right\}.$$

Here we want to prove that $l_1 = \cdots = l_k = d$.

We first prove that for each $1 \le i \le k$, then there exists some $W_i \in \Gamma(l_i)$ such that $\nu_i(\overline{W}_i) = r_i$. If the supremum $r_i > 0$ is not attainable, assume there exists an infinite sequence of l_i-dimensional linear subspaces $\{V_\alpha\}_{\alpha \ge 1}$ such that

$$\nu_i(V_\alpha) > r_i - \frac{1}{\alpha^2}.$$

Note if $V_\alpha \neq V_\beta$, then $\dim(V_\alpha \cap V_\beta) < l_i$, then $\nu_i(V_\alpha \cap V_\beta) = 0$ by the minimality of l_i, and thus

$$\nu_i(V_\alpha \cup V_\beta) = \nu_i(V_\alpha) + \nu_i(V_\beta) > 2r_i - \frac{1}{\alpha^2} - \frac{1}{\beta^2}.$$

Since ν_i is a finite measure, $\{V_\alpha\}_{\alpha \ge 1}$ cannot be mutually different and must consist of only finitely many different elements, which is a contradiction. Therefore such a $W_i \in \Gamma(l_i)$ exists.

Define $l_{\min} = \min\{l_i : 1 \le i \le k\}$ and assume $l_{\min} < d$. Let

$$K = \{1 \le i \le k : \dim W_i = l_{\min}\},$$

and let $f(\omega, x) = 1(\omega_0 \in K, x \in W_{\omega_0})$ for $\{W_i\}_{1 \le i \le k}$ satisfying $r_i = \nu_i(\overline{W}_i)$, then

$$\int f(\omega, x)\nu(\mathrm{d}\omega, \mathrm{d}x) = \int_{\omega_0 \in K} \nu_\omega(\overline{W}_{\omega_0})\mathbb{P}(\mathrm{d}\omega) = \sum_{i \in K} q_i r_i.$$

On the other hand,

$$\int f(\omega,x)(\Sigma\nu)(\mathrm{d}\omega,\mathrm{d}x)=\int f(\sigma\omega,\hat{M}(\omega)x)\nu(\mathrm{d}\omega,\mathrm{d}x)$$
$$=\int_{\omega_1\in K}\nu_{\omega_0}(\hat{M}(\omega)^{-1}\overline{W}_{\omega_1})\mathbb{P}(\mathrm{d}\omega).$$

Since when $\omega_1\in K$, $\dim[M(\omega)^{-1}W_{\omega_1}]=\dim W_{\omega_1}=l_{\min}\le l_{\omega_0}$, we have

$$\nu_{\omega_0}(\hat{M}(\omega)^{-1}\overline{W}_{\omega_1})=0,\quad \text{if } \omega_0\notin K;$$

and

$$\nu_{\omega_0}(\hat{M}(\omega)^{-1}\overline{W}_{\omega_1})\le r_{\omega_0},\quad \text{if } \omega_0\in K.$$

Thus

$$\sum_{i\in K}q_ir_i=\int f\nu=\int(f\circ\Sigma)\nu\le\sum_{i,j\in K}q_ip_{ij}r_i\le\sum_{i\in K}q_ir_i.$$

In order for the above equality to hold, we must have $K=\{1,2,\dots,k\}$ and $r_i=\nu_i(\hat{M}_i^{-1}\overline{W}_j)$ for any $i,j\in K$. This shows that

$$\Lambda:=\bigcup\Big\{W\in\Gamma(l_{\min}):\nu_i(\overline{W})=r_i \text{ for some } 1\le i\le k\Big\}$$

is invariant under $\mathcal{S}_\varphi$. Note Λ is a finite union, since for any $W_1,W_2\in\Lambda$, we have $\dim W_1\cap W_2<l_{\min}$ thus

$$\nu(\overline{W}_1\cup\overline{W}_2)=\nu(\overline{W}_1)+\nu(\overline{W}_2)>2\int_K r(\omega)\mathbb{P}(\mathrm{d}\omega).$$

Now Λ being a finite union of proper linear subspaces invariant under $\mathcal{S}_\varphi$ contradicts the strong irreducibility of $\mathcal{S}_\varphi$. □

Now with the properness of the invariant measure, we are ready to give the asymptotic behaviour of Markovian products in direction, which is analogous to Proposition 2.3.10.

Proposition 4.3.6. *Let $(\Omega^+,\mathcal{F}^+,\mathbb{P},\sigma)$ be a one-sided Markovian full shift over k symbols, φ be a strongly irreducible and spatially invertible LRDS, which is determined by $\varphi(1,\omega)=M(\omega)$ and has index r. We assume $M(\omega)$ depends only on the first coordinate ω_0.*

Then for $\mathbb{P}$-almost all ω, there exists an r-dimensional subspace $V(\omega)$ of $\mathbb{R}^d$ such that it is the range of any limit point of $\varphi(n,\sigma^{-n}\omega)/\|\varphi(n,\sigma^{-n}\omega)\|$, and for any non-zero $x\in\mathbb{R}^d$, $\mathbb{P}\{x:x\perp V(\omega)\}=0$.

Proof. By Lemma 4.3.4, we have for η_j-almost all $Q_j \in \Lambda_j = \{SM_j : S \in \mathcal{S}_\varphi\}$

$$\lim_{n\to\infty} \varphi(n, \sigma^{-n}\omega)Q_j\nu_j = \tilde{\nu}_\omega, \quad 1 \le j \le k.$$

Fix ω and let $R(\omega)$ be a limit point of $\varphi(n, \sigma^{-n}\omega)/\|\varphi(n, \sigma^{-n}\omega)\|$ of rank $r(\omega)$. Then through a subsequence, $\varphi(n, \sigma^{-n}\omega)Q_j\nu_j$ converges to $R(\omega)Q_j\nu_j$. Note by Proposition 4.3.5, ν is proper, so is ν_j for each $1 \le j \le k$, and thus $R(\omega)Q_j\nu_j$ is well-defined. Therefore for each $1 \le j \le k$, and any $Q_j \in \Lambda_j := \{SM_j : S \in S_\varphi\}$,

$$R(\omega)Q_j\nu_j = \tilde{\nu}_\omega. \tag{4.5}$$

Since φ has index r, assume there exists a sequence $S_j^{(n)}/\|S_j^{(n)}\|$ which converges to a rank r matrix P as $n \to \infty$, where $S_j^{(n)} \in \Lambda_j$. Note $Q_j S_j^{(n)} \in \Lambda_j$, we have

$$R(\omega)Q_j S_j^{(n)}\nu_j = \tilde{\nu}_\omega.$$

Now either $R(\omega)Q_jP = 0$ or $R(\omega)Q_jP\nu_j = \tilde{\nu}_\omega$.

If $R(\omega)Q_jP = 0$ for some $Q_j \in \Lambda_j$, then define $L := \{Q_jPx : Q_j \in \Lambda_j, x \in \mathbb{R}^d\}$, then $\{0\} \neq L \subset \ker R(\omega)$ is a proper linear subspace of $\mathbb{R}^d$ and $Q(L) = L$ for each $Q \in \mathcal{S}_\varphi$. This contradicts the strong irreducibility.

Therefore $R(\omega)Q_jP\nu_j = \tilde{\nu}_\omega$ for each $Q_j \in \Lambda_j$. Thus $\operatorname{supp}\tilde{\nu}_\omega \subset \overline{\operatorname{im}(R(\omega)Q_jP)}$. Let

$$V(\omega) = \operatorname{span}\{x \in \mathbb{R}^d : \bar{x} \in \operatorname{supp}\tilde{\nu}_\omega\},$$

then $\dim V(\omega) \le r$, since the rank of $R(\omega)Q_jP$ is not greater than $r = \operatorname{rank}(P)$.

On the other hand, let $Q_j = M_j$ in (4.5), we have $R(\omega)M_j\nu_j = \tilde{\nu}_\omega$. Thus

$$1 = \tilde{\nu}_\omega(V(\omega)) = \nu_j\{x : R(\omega)M_jx \in V(\omega)\}.$$

Since ν_j is proper, $\{x : R(\omega)M_jx \in V(\omega)\}$ is a linear space, thus has to be $\mathbb{R}^d$. Therefore $V(\omega) \supset \operatorname{im}(R(\omega)M_j) = \operatorname{im}(R(\omega))$ and $\dim V(\omega) \ge r(\omega) \ge r$.

These show that $V(\omega) = \operatorname{im}(R(\omega))$ and that $\dim V(\omega) = r$. In the end,

$$\mathbb{P}\{x : x \perp V(\omega)\} = \mathbb{P}\{\operatorname{supp}\tilde{\nu}_\omega \subset x^\perp\} = \mathbb{P}\{\tilde{\nu}_\omega(x^\perp) = 1\} = \nu(x^\perp) = 0.$$

□

Theorem 4.3.7. *(Generalisation of Theorem 2.3.11) If $(\Omega^+, \mathcal{F}^+, \mathbb{P}, \sigma)$ is a one-sided Markovian full shift of finite type, φ is a strongly irreducible and spatially invertible LRDS determined by $\varphi(1, \omega) = M(\omega)$, and $M(\omega)$ depends only on the first coordinate ω_0. If $\mathbb{E}[\log^+ \|M(\omega)\|] < \infty$, then for a sequence (x_n) of vectors converging to a non-zero $x \in \mathbb{R}^d$ we have*

$$\gamma = \lim_{n\to\infty} \frac{1}{n} \log \|\varphi(n, \omega)x_n\|, \quad \mathbb{P}\text{-a.s.}.$$

If ν is the invariant measure on $\Omega^+ \times \mathbb{R}P^{d-1}$ for the NRDS $\bar{\varphi}$ induced by φ, then

$$\gamma = \int \log \|M(\omega)\bar{x}\| \nu(\mathrm{d}\omega, \mathrm{d}\bar{x}).$$

Moreover, if $\mathbb{E}[\log^+ \|M^{-1}(\omega)\|] < \infty$, then

$$\frac{1}{n}\mathbb{E}\left[\log \|\varphi(n, \omega)x\|\right] \to \gamma \quad (n \to \infty)$$

uniformly on $\{x \in \mathbb{R}^d : \|x\| = 1\}$.

Proof. Just as in Theorem 2.3.11, we have

$$\frac{\|\varphi(n, \omega)x_n\|^2}{\|\varphi(n, \omega)\|^2} = \sum_{i=1}^{d} \frac{c_i(n)^2}{c_1(n)^2} \langle x_n, V_n^* e_i\rangle.$$

Let $W(\omega)$ be the $r(\omega)$-dimensional range of $Y(\omega)$, then by Proposition 4.3.5, for any non-zero vector $x \in \mathbb{R}^d$, $\mathbb{P}\{x \perp W(\omega)\} = 0$.

For the second statement,

$$\int_{\Omega^+\times X} \log^+ \|M(\omega)x\| \nu(\mathrm{d}\omega, \mathrm{d}\bar{x}) \le \int_{\Omega^+} \int_X \log^+ \|M(\omega)\| \nu_\omega(\mathrm{d}\bar{x}) \mathbb{P}(\mathrm{d}\omega) < \infty.$$

By Lemma 4.3.2, $n^{-1} \log \|\varphi(n, \omega)x\|$ converges to a random variable $\Lambda(\omega, x)$, ν-almost surely. The rest of the proof is the same. □

Now with the tools above, we can generalise Proposition 2.3.12 and Proposition 2.3.14 without much work. Note here in [4] the proof of Lemma 2.3.13 uses Birkhoff's ergodic theorem without assuming ergodicity, which is still valid in the Markovian setting. Finally for a generalised version of Proposition 2.3.15, we give a complete description as it will be used in Section 6.

Proposition 4.3.8. *(Generalisation of Proposition 2.3.15) If $(\Omega^+, \mathcal{F}^+, \mathbb{P}, \sigma)$ is a one-sided Markovian full shift of finite type, φ is a strongly irreducible, contracting and spatially invertible LRDS determined by $\varphi(1, \omega) = M(\omega)$, where $M(\omega)$ depends only on the first coordinate ω_0. Write $\ell(M) := \sup\{\log^+ \|M\|, \log^+ \|M^{-1}\|\}$. If for some $\beta > 0$,*

$$\int_{\Omega^+} \exp\Big(\beta \ell\left(M(\omega)\right)\Big)\, \mathbb{P}(\mathrm{d}\omega) < \infty,$$

then there exists some α_0 such that when $0 < \alpha \leq \alpha_0$,

$$\lim_{n\to\infty} \left\{ \sup_{\substack{x,y\in\mathbb{R}P^{d-1} \\ x\neq y}} \mathbb{E}\left[\frac{d^\alpha(\varphi(n,\omega)\cdot x, \varphi(n,\omega)\cdot y)}{d^\alpha(x,y)}\right] \right\}^{1/n} < 1.$$

Proof. According to the original proof, we only need to redefine a_n by

$$a_n = \log \sup_{x,y} \int e^{\alpha\delta(\varphi(n,\omega),(x,y))} \mathbb{P}(\mathrm{d}\omega).$$

And by the cocycle property of φ together with a generalised version of Proposition 2.3.14, the result follows. □

Chapter 5

Pollicott's Algorithm

The purpose of this chapter is to give an explicit description of Pollicott's algorithm [48] in computing the Lyapunov exponent for i.i.d. products of positive matrices (i.e., matrices whose entries are all positive) on a discrete measure.

More precisely, let $\{M_i\}_{1\leq i\leq k}$ be a finite set of invertible positive matrices and $(p_1,\ldots,p_k)$ be a k-dimensional vector such that $p_1+\cdots+p_k=1$, which will be called a **probability vector** in the following. Define a random matrix $A(\omega)$ on a probability space $(\Omega,\mathcal{F},\mathbb{P})$, with its law given by $\mu=\sum_{i=1}^k p_i\delta_{M_i}$, where δ_M denotes a singleton measure on the matrix M. Pollicott's approach aims at computing the Lyapunov exponent associated with this type of problem.

The idea of the algorithm can be briefly described as follows. Define a family of parametrised operators $\mathcal{L}_t$ on some function space $\mathcal{G}$ on the projective space $\mathbb{R}P^{d-1}$ by

$$(\mathcal{L}_t w)(x)=\sum_{i=1}^k p_i e^{t\log\|M_i x\|} w\left(\frac{M_i x}{\|M_i x\|}\right).$$

Then it can be proved that, when $|t|$ is sufficiently small, $\mathcal{L}_t$ possesses a simple maximal top eigenvalue $\beta(t)$ (Theorem 5.3.2) and $\beta'(0)$ equals the corresponding Lyapunov exponent (Theorem 5.3.3). Now our problem is transformed into the calculation of $\beta(t)$. Ruelle [49] applied Grothendieck's theory [22] on nuclear operators on Banach spaces, and first proved that this type of operators are nuclear of order zero, which means that they share many characteristics with matrices and operators of finite rank. Naturally, just as we would expect from matrices, the top eigenvalue $\beta(t)$ of $\mathcal{L}_t$ is related to the zeros of the determinant function $\det(I-z\mathcal{L}_t)$. The trace-determinant formula for matrices can also be generalised for this type of operators as

below

$$d(z,t) := \det(I - z\mathcal{L}_t) = \exp\left(-\sum_{n=1}^{\infty} \frac{\mathrm{tr}(\mathcal{L}_t^n)}{n} z^n\right).$$

It can be proved that the determinant function $d(z,t)$ is an entire function with respect to z according to Fredholm theory. If we expand the determinant function as a series

$$d(z,t) = 1 + \sum_{n=1}^{\infty} a_n(t) z^n,$$

then the Lyapunov exponent $\gamma = \beta'(0)$ can be given by (Theorem 5.7.1)

$$\beta'(0) = \frac{\sum_{n=1}^{\infty} a_n'(0)}{\sum_{n=1}^{\infty} n a_n(0)}.$$

The p-th estimate of γ is

$$\gamma^{(p)} = \frac{\sum_{n=1}^{p} a_n'(0)}{\sum_{n=1}^{p} n a_n(0)}.$$

In Section 5.1, we will discuss the identification of the real (resp. complex) projective space [on the positive orthant] with subspaces of real (resp. complex) spaces, and introduce prerequisites for defining the function space $\mathcal{G}$ on which transfer operators are acting. Transfer operators are defined in Section 5.2, in which the spectral properties of transfer operators are given by Theorem 5.2.4. Parametrised transfer operators and their spectral properties are introduced in Section 5.3. At the end of this section, we will prove that $\beta'(0) = \gamma$ in Theorem 5.3.3. Section 5.4 aims at providing a brief introduction to the concept of nuclear operators, and we will see the nuclearity of transfer operators in Theorem 5.5.1 of Section 5.5.The eigenvalues of simple transfer operators are given in Theorem 5.5.4, therefore the trace formula for transfer operators can be naturally given by Corollary 5.5.5. Fredholm theory is briefly introduced in Section 5.6. And finally we can give the algorithm in Section 5.7. Some numerical results are given in Section 5.9.

The results in this chapter are mostly based on the ideas from [48] [44] [22] [49] [28]. It is worth mentioning that there are a number of flaws in Pollicott's original paper [48], for example, the wrong choice of the domain U (which will be corrected in Section 5.1); the non-optimal choice of the weight functions in parametrised transfer operators (which will be modified in Section 5.3); the incomplete proof of the Perron-Frobenius-Ruelle theorem (which will be given by Theorem 5.2.4 and Theorem 5.3.2); and the lemma (Lemma 4.2) in [48] which can be applied to only 2×2 matrices (a general form with its proof will be given in Section 5.8).

5.1 Projective Actions

In Section 2.1.2, we introduced the normalised matrix actions on vector spaces. Namely, given an invertible matrix $A \in GL(d, \mathbb{R})$, we can define a map

$$\tilde{A} : \mathbb{R}P^{d-1} \to \mathbb{R}P^{d-1}$$

$$\pi(x) \mapsto \pi(Ax),$$

where π is the natural quotient map from $\mathbb{R}^d$ to $\mathbb{R}P^{d-1}$. In this section, we always assume the given matrix A is positive, and we plan to investigate some analytic properties of this type of induced maps.

First of all, we note that, if A is positive and we denote

$$\mathbb{R}^d_+ = \{(x_1, \dots, x_d) : x_i > 0, 1 \le i \le d\}, \quad \Delta = \pi(\mathbb{R}^d_+),$$

since for any $x \in \mathbb{R}^d_+$, $Ax \in \mathbb{R}^d_+$, $\tilde{A}$ restricted to Δ is a well-defined map.

It is also worth noting that $\mathbb{R}P^{d-1}$ as a manifold is not convenient for computation. Therefore we will usually choose representatives in $\mathbb{R}^d$ in one of the two following ways.

(1) Denote $\{x \in \mathbb{R}^d : \|x\|_2^2 := \sum_{i=1}^d x_i^2 = 1\}$ by S^{d-1}. We can naturally define a surjective map

$$\bar{\pi} : S^{d-1} \to \mathbb{R}P^{d-1}, \qquad \bar{x} \mapsto \bar{\pi}(\bar{x}) := \pi(\bar{x}).$$

It is surjective because for any $\pi(x) \in \mathbb{R}P^{d-1}$, we can find a unit vector $\bar{x} \in S^{d-1}$ of the same direction to represent $\pi(x)$. Since there are two directions we can choose from, an equivalence relation can be defined on S^{d-1} by $u \sim (-u)$ for any $u \in S^{d-1}$. Then we obtain a diffeomorphism between $S^{d-1}/\sim$ and $\mathbb{R}P^{d-1}$. Moreover, we denote by $\bar{A}$ the induced map

$$\bar{A} : S^{d-1} \to S^{d-1}, \qquad \bar{x} \mapsto A\bar{x}/\|A\bar{x}\|_2. \tag{5.1}$$

When restricted on the positive part, let $S^{d-1}_+ := S^{d-1} \cap \mathbb{R}^d_+$, then S^{d-1}_+ is diffeomorphic to Δ, and $\bar{A}|_{S^{d-1}_+}$ can be identified with $\tilde{A}|_\Delta$.

(2) Given $x = (x_1, \dots, x_d) \in \mathbb{R}^d_+$, then we can choose $x_1 \neq 0$ as the normalisation constant and define

$$R^{d-1}_+ := \{x = (x_1, \dots, x_d) \in \mathbb{R}^d_+ : x_1 = 1\}, \tag{5.2}$$

$$\hat{A} : R^{d-1}_+ \to R^{d-1}_+, \qquad x = (1, x_2 \dots, x_d) \mapsto Ax/(Ax)_1, \tag{5.3}$$

where we denote by $(Ax)_i$ $(1 \leq i \leq d)$ the i-th entry of the vector (Ax). This map is well-defined as $(Ax)_1 \neq 0$. $\hat{A}$ can be identified with $\tilde{A}|_\Delta$ with the same reason.

Example 3. Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. Although it is hard to write down $\tilde{A}|_\Delta$, we can explicitly express $\bar{A}|_{S_+^{d-1}}$ and $\hat{A}$. In fact,

$$\bar{A}(x_1, x_2) = \left(\frac{2x_1 + x_2}{\sqrt{5x_1^2 + 6x_1x_2 + 2x_2^2}}, \frac{x_1 + x_2}{\sqrt{5x_1^2 + 6x_1x_2 + 2x_2^2}} \right), \text{ subject to } x_1^2 + x_2^2 = 1.$$

$$\hat{A}(1, x_2) = \left(1, \frac{x_2 + 1}{x_2 + 2}\right), \quad \text{or} \quad \hat{A}(x_2) = \frac{x_2 + 1}{x_2 + 2}, \text{ when } \mathbb{R}_+^1 \text{ is identified with } R_+^1.$$

It is not hard to see that the two different ways to write $\hat{A}$ are equivalent. We can observe that with the second representation, $\hat{A}$ is a real Möbius transformation. In general $\hat{A}$ can be regarded as a projective transformation.

Remark 5.1.1. (1) In Section 4.2, we stated that a random dynamical system acting on a compact metrisable space can generate an invariant measure. It is worth mentioning that, it can be proved invariant measures on different choice of sections induced by the same RDS are equivalent. More detail on the equivalence of these choices can be found in e.g. [1] Chapter 6.2.

(2) R_+^{d-1} can be naturally identified with $\mathbb{R}_+^{d-1}$ by restricting the first entry to be 1.

The first result for positive matrices appears in Pollicott's paper [48] as a lemma.

Lemma 5.1.2. *Let A be positive matrix, then the induced map $\tilde{A} : \mathbb{R}P^{d-1} \to \mathbb{R}P^{d-1}$ restricted to $\Delta = \pi(\mathbb{R}_+^d)$ is a real analytic function $\tilde{A}_j|_\Delta : \Delta \to \Delta$.*

Proof. With the same notation as above, we consider $\bar{A}|_{S_+^{d-1}}$. Note $x \mapsto Ax$ and $x \mapsto \|Ax\|_2$ are both real analytic on S_+^{d-1} and $\|Ax\|_2 \neq 0$ for any $x \in S_+^{d-1}$, we know $\hat{A}|_{S_+^{d-1}}$ is real analytic.

As $\tilde{A}|_\Delta$ can be identified with $\bar{A}|_{S_+^{d-1}}$, the conclusion follows. □

With the intuition of Section 3.2 in mind (where we used the identification of $\tilde{A}|_\Delta$ with $\bar{A}|_{S_+^{d-1}}$ and showed a contraction property for products of positive matrices), we can obtain a stronger result than the lemma above. This result will be essential in defining the function space

on which transfer operators are acting. In the following, we say an analytic function ψ maps an open set U **strictly inside** U if the closure $\overline{\psi(U)}$ of the image of ψ satisfies $\overline{\psi(U)} \subset U$.

Notice the real map $\hat{M}$ induced by the positive matrix M can be extended to a complex map by replacing the real variables by complex ones. More precisely, let

$$C^{d-1} := \{z = (1, z_2, \ldots, z_d) : z_i \in \mathbb{C},\ 2 \leq i \leq d\},$$

then we can define

$$\hat{M} : C^{d-1} \to C^{d-1}$$

$$z = (1, z_2, \ldots, z_d) \mapsto \frac{Mz}{(Mz)_1}.$$

It is well-defined as $\mathrm{Re}(Mz)_1 = (M \,\mathrm{Re}\, z)_1 \neq 0$, so $(Mz)_1 \neq 0$. Here we still use $\hat{M}$ to denote the extended map. To express $\hat{M}$ explicitly,

$$\frac{Mz}{(Mz)_1} = \left(1, \frac{(Mz)_2}{(Mz)_1}, \ldots, \frac{(Mz)_d}{(Mz)_1}\right),$$

and, if $M = (a_{ij})_{1 \leq i,j \leq d}$, we have, for each $i \geq 1$,

$$\frac{(Mz)_i}{(Mz)_1} = \frac{a_{i1} + \sum_{j=2}^{d} a_{ij} z_j}{a_{11} + \sum_{j=2}^{d} a_{1j} z_j}.$$

Thus $\hat{M}$ can be regarded as a projective transformation of the form

$$(z_2, \ldots, z_d) \mapsto \left(\frac{a_{21} + \sum_{j=2}^{d} a_{2j} z_j}{a_{11} + \sum_{j=2}^{d} a_{1j} z_j}, \ldots, \frac{a_{d1} + \sum_{j=2}^{d} a_{dj} z_j}{a_{11} + \sum_{j=2}^{d} a_{1j} z_j}\right).$$

Denote

$$C_+^{d-1} := \{z = (1, z_2, \ldots, z_d) : \mathrm{Re}\, z \in \mathbb{R}_+^d\},$$

then we have the following lemma.

Lemma 5.1.3. *Let M be a positive matrix, and $\hat{M}$ be the induced map of M acting on C^{d-1}. Then there exists a bounded and connected open set U in C_+^{d-1} such that $\hat{M}|_U$ maps U strictly inside U and $\hat{M}|_U$ is a complex analytic function. Moreover, U can be chosen such that $\hat{M}|_V$ maps $V \subset R_+^{d-1}$ strictly inside V as a real analytic map, where $V = \mathrm{Re}\, U$.*

Proof. Denote

$$U_0 := \{(z_2, \dots, z_d) \in C_+^{d-1} : \mathrm{Re}(z_j/z_i) > 0,\ 2 \le i, j \le d\}. \tag{5.4}$$

If we write $z_j = x_j + \mathrm{i}y_j$, where $x_j \in \mathbb{R}_+, y_j \in \mathbb{R}$ for each $2 \le j \le d$, then we can see that $\mathrm{Re}(z_j/z_i) > 0$ is equivalent with

$$x_i x_j + y_i y_j > 0, \quad 2 \le i, j \le d. \tag{5.5}$$

We first notice $\hat{M}|_{U_0}$ maps U_0 into U_0 as a complex map. In fact, if we write $\hat{M}$ in the form of

$$(z_2, \dots, z_d) \mapsto \left(\frac{a_{21} + \sum_{j=2}^d a_{2j} z_j}{a_{11} + \sum_{j=2}^d a_{1j} z_j}, \dots, \frac{a_{d1} + \sum_{j=2}^d a_{dj} z_j}{a_{11} + \sum_{j=2}^d a_{1j} z_j} \right) =: (w_2, \cdots, w_d),$$

where $M = (a_{ij})_{1 \le i,j \le d}$. Then,

$$\frac{w_\gamma}{w_\delta} = \frac{(a_{\gamma 1} + \sum_j a_{\gamma j} x_j) + \mathrm{i} \sum_j a_{\gamma j} y_j}{(a_{\delta 1} + \sum_j a_{\delta j} x_j) + \mathrm{i} \sum_j a_{\delta j} y_j}, \quad 2 \le \gamma, \delta \le d.$$

Therefore,

$$\begin{aligned} \mathrm{Re}\left(\frac{w_\gamma}{w_\delta}\right) &= \frac{(a_{\gamma 1} + \sum_j a_{\gamma j} x_j)(a_{\delta 1} + \sum_j a_{\delta j} x_j) + \sum_j a_{\gamma j} y_j \sum_j a_{\delta j} y_j}{(a_{\delta 1} + \sum_j a_{\delta j} x_j)^2 + (\sum_j a_{\delta j} y_j)^2} \\ &= \frac{a_{\gamma 1} a_{\delta 1} + \sum_{i,j} a_{\gamma i} a_{\delta j} (x_i x_j + y_i y_j) + \sum_j a_{\delta 1} a_{\gamma j} x_j + \sum_j a_{\gamma 1} a_{\delta j} x_j}{(a_{\delta 1} + \sum_j a_{\delta j} x_j)^2 + (\sum_j a_{\delta j} y_j)^2} > 0, \end{aligned}$$

by the positivity of M, $x_j > 0$ $(2 \le j \le d)$, and (5.5).

Moreover, for $a, b, c, d \in \mathbb{R}_+$, we have the following inequality

$$\frac{a+b}{c+d} \le \max\left\{\frac{a}{c}, \frac{b}{d}\right\}. \tag{5.6}$$

Therefore for each $2 \le \gamma \le d$, by (5.5) and (5.6),

$$\begin{aligned} \left| \frac{a_{\gamma 1} + \sum_{j=2}^d a_{\gamma j} z_j}{a_{11} + \sum_{j=2}^d a_{1j} z_j} \right|^2 &= \frac{(a_{\gamma 1} + \sum_j a_{\gamma j} x_j)^2 + (\sum_j a_{\gamma j} y_j)^2}{(a_{11} + \sum_j a_{1j} x_j)^2 + (\sum_j a_{1j} y_j)^2} \\ = \frac{a_{\gamma 1}^2 + \sum_{j,l} a_{\gamma j} a_{\gamma l} (x_j x_l + y_j y_l) + 2 \sum_j a_{\gamma 1} a_{\gamma j} x_j}{a_{11}^2 + \sum_{j,l} a_{1j} a_{1l} (x_j x_l + y_j y_l) + 2 \sum_j a_{11} a_{1j} x_j} \end{aligned}$$

$$\leq \max_{1\leq j,l\leq d}\left\{\frac{a_{\gamma j}a_{\gamma l}}{a_{1j}a_{1l}}\right\} =: C_\gamma.$$

Hence $\hat{M}$ maps U_0 into a bounded open set

$$U = U_0 \cap \left\{(w_2,\ldots,w_d) \in C_+^{d-1} : |w_\gamma| < C_\gamma + 1,\ 2 \leq \gamma \leq d\right\}.$$

Thus, when restricted to U, $\hat{M}$ maps U strictly inside U.

Finally, note $\operatorname{Re} U_0 = R_+^{d-1}$ and $\operatorname{im} \hat{M}|_{R_+^{d-1}} \subset U$. Therefore $\hat{M}|_{\operatorname{Re} U}$ maps $\operatorname{Re} U$ strictly inside $\operatorname{Re} U$. □

Corollary 5.1.4. *Let $\{M_i\}_{i\in I}$ be a finite set of positive matrices, and $\hat{M}_i$ be the induced map of M_i on C^{d-1}. Then there exists a bounded and connected open set $U \subset C_+^{d-1}$ such that for each $i \in I$, $\hat{M}_i|_U$ maps U strictly inside U and each $\hat{M}_i|_U$ is a complex analytic function. Moreover, U can be chosen such that $\hat{M}_i|_V$ maps $V \subset R_+^{d-1}$ strictly inside V as a real analytic map, where $V = \operatorname{Re} U$.*

Proof. By the proof of Lemma 5.1.3, for each $i \in I$, we can find $U_i \subset C_+^{d-1}$ such that $\hat{M}_i$ maps U_0 to $U_0 \cap U_i$, where U_0 is given by (5.4) and U_i is bounded and connected with the form

$$U_i = \{(w_2,\ldots,w_d) \in C_+^{d-1} : |w_\gamma| < C_\gamma^{(i)}, 2 \leq \gamma \leq d\},$$

for some positive constants $C_\gamma^{(i)}, 1 \leq i \leq k,\ 2 \leq \gamma \leq d$.

Now take the bounded and connected open set $U = \left(\bigcup_{i\in I} U_i\right) \cap U_0$, then obviously $\hat{M}_i$ maps U_0 strictly inside U, and thus maps U strictly inside U. □

For the rest of the article, we will call an open connected set in $\mathbb{C}^d$ (or in $\mathbb{R}^d$) a *domain*. Some common choices of norms on $\mathbb{C}^d$ are, for $z = (z_1,\ldots,z_d)$,

$$\|z\|_0 := \max_i |z_i|, \quad \|z\|_1 := \sum_{i=1}^d |z_i|, \quad \|z\|_2 := \left(\sum_{i=1}^d |z_i|^2\right)^{1/2}.$$

It is not hard to see that they are equivalent. When we identify a domain $U \subset C^{d-1}$ as a subset of $\mathbb{C}^{d-1}$ (or $V \subset R_+^{d-1}$ as a subset of $\mathbb{R}^{d-1}$), the norms can be naturally inherited.

However, we shall also notice that in the beginning of Section 2.3.4, another distance on the real projective space is defined by

$$d(x,y) = (1 - \langle x, y\rangle^2)^{1/2} = |\sin\alpha(x,y)| = \frac{\|x \wedge y\|}{\|x\|\|y\|}, \quad x, y \in \mathbb{R}P^{d-1}, \tag{5.7}$$

where $\alpha(x,y)$ is the angle starting from direction x to $y \pmod{2\pi}$ such that $\alpha(x,y) = -\alpha(y,x)$. Now we prove that when a domain $V \subset R_+^{d-1}$ is bounded, this distance is equivalent with the distances induced by common norms.

Lemma 5.1.5. *Let W be a bounded domain in R_+^{d-1}. Then there exist constants $\kappa_1, \kappa_2 > 0$ such that for any $x = (x_1, \ldots, x_d), y = (y_1, \ldots, y_d) \in \mathbb{R}^d$ with $x_1, y_1 \neq 0$, if we have $\hat{x} := x/x_1, \hat{y} := y/y_1$ are both in W, then*

$$\kappa_1\|\hat{x} - \hat{y}\| \leq d(x,y) \leq \kappa_2\|\hat{x} - \hat{y}\|.$$

Proof. By the equivalence of norms on $\mathbb{R}^d$, we can without loss of generality choose $\|\cdot\|$ to be the ℓ_2-norm in this proof. Apply the law of cosines to the triangle formed by $\hat{x}, \hat{y}$ and the origin in $\mathbb{R}^d$, then we have

$$\begin{aligned}\|\hat{x} - \hat{y}\|^2 &= \|\hat{x}\|^2 + \|\hat{y}\|^2 - 2\|\hat{x}\|\|\hat{y}\|\cos\alpha(x,y) \\ &= (\|\hat{x}\| - \|\hat{y}\|)^2 + 2\|\hat{x}\|\|\hat{y}\|(1 - \cos\alpha(x,y)) \\ &\geq 2\|\hat{x}\|\|\hat{y}\|(1 - \cos\alpha(x,y)).\end{aligned}$$

Note $\|\hat{x}\| \geq 1, \|\hat{y}\| \geq 1$ and

$$1 - \cos\alpha(x,y) = \frac{d(x,y)^2}{1 + \cos\alpha(x,y)} \geq \frac{d(x,y)^2}{2},$$

we have $\|\hat{x} - \hat{y}\| \geq d(x,y)$.

On the other hand, applying the law of sines to the same triangle in $\mathbb{R}^d$ shows that $\|\hat{x} - \hat{y}\|/d(x,y)$ equals the diameter of this triangle's circumcircle. As W is bounded, the diameter must have an upper bound as well. □

Corollary 5.1.6. *Proposition 2.3.14 and Proposition 2.3.15 are still true when the distance d is replaced by a distance induced by a norm on a bounded domain W in R_+^{d-1} (when it is regarded as a subset of $\mathbb{R}^{d-1}$).*

Proof. This follows easily by Lemma 5.1.5. □

We notice here Proposition 2.3.14 and Proposition 2.3.15 can only be applied to a matrix set which is strongly irreducible and contracting. In order to obtain analogous results for products of positive matrices, we provide two different approaches here.

The first approach investigates when a set of positive matrices is strongly irreducible and contracting. However, as we have restricted the action of positive matrices to the positive orthant, we shall slightly generalise the definition of strong irreducibility given by Definition 2.3.2.

We say a subset $C \subset \mathbb{R}^d$ is a **cone** if C can be written in the form of

$$C = \{r_1 v_1 + \cdots + r_s v_s : r_1, \ldots, r_s \geq 0\},$$

for some vectors $v_1, \ldots, v_s \in \mathbb{R}^d$. The **dimension** of a cone is defined to be the dimension of the linear space generated by the cone. We can observe that the positive orthant $\mathbb{R}^d_+$ is a d-dimensional cone spanned by the natural basis of $\mathbb{R}^d$.

If both C and $C_0 \subset C$ are cones of $\mathbb{R}^d$, we say C_0 is a **subcone** of C. C_0 is said to be **proper** if $\dim C_0 < \dim C$.

Now let $\mathcal{S}$ be a subset of $d \times d$ invertible matrices acting on a cone C of $\mathbb{R}^d$, i.e., for each $M \in \mathcal{S}$, $MC \subset C$. We say $\mathcal{S}$ is **strongly irreducible on the cone** C if there does not exist a finite family of proper subcones $C_1, \ldots, C_s$ of C, such that

$$M(C_1 \cup \cdots \cup C_s) \subset C_1 \cup \cdots \cup C_s.$$

for each $M \in \mathcal{S}$. $\mathcal{S}$ is said to be **irreducible on the cone** C if there does not exist a proper subcone C_0 of C, such that $MC_0 \subset C_0$.

Proposition 5.1.7. *Let $\{M_i\}_{i\in I}$ be a set of invertible and positive $d \times d$ matrices. Then*

(a) $\{\hat{M}_i\}_{i\in I}$ is contracting.

(b) $\{\hat{M}_i\}_{i\in I}$ is strongly irreducible on $\mathbb{R}^d_+$ if and only if it is irreducible.

Proof. The first conclusion follows from Proposition 3.1.7. Now we prove the second.

With the same reason given by Remark 2.3.3, if $\{M_i\}_{i\in I}$ is not strongly irreducible on $\mathbb{R}^d_+$, then we can without loss of generality assume that there exist proper subcones $C_1, \ldots, C_s$ with the same dimension $l \geq 1$, such that $C_i \cap C_j = \{0\}$ for each $1 \leq i, j \leq s$, and that each $M \in \mathcal{S}$ corresponds to a permutation $\rho(M)$ in the symmetric group of order s. By reordering the index, we can assume

$$MC_1 \subset C_2, \ \ldots, \ MC_{\alpha-1} \subset C_\alpha, \ MC_\alpha \subset C_1,$$

for some integer $\alpha|s$. Therefore $M^\alpha C_i \subset C_i$ for $1 \le i \le \alpha$. By the Perron-Frobenius theorem [3], we know there exists a unique non-zero fixed point $v_i \in C_i$ for the action of M^α on C_i. However, applying the Perron-Frobenius theorem to the action of M^α on $\mathbb{R}_+^d$ shows that $v_1 = \cdots = v_\alpha \in C_1 \cap \cdots \cap C_\alpha$. Since $C_i \cap C_j = \{0\}$ and v_i is non-zero, we must have $\alpha = 1$ and therefore $s = 1$. That is to say, there exists a proper cone C of $\mathbb{R}_+^d$ such that $MC \subset C$ for each $M \in \mathcal{S}$. □

It is not hard to see that the theory built in Section 2.3 still works for the actions of matrices on the cones if we consider the corresponding invariant measures to be supported on a subset of the cones. When the matrix set fails to be (strongly) irreducible, we can try to decompose the problem into subproblems as is described in Remark 2.3.3. Or we could try the second approach.

This approach is inspired by Birkhoff's geometric proof for the Perron-Frobenius theorem [3], in which Hilbert's projective metric was applied to obtain a contraction property for positive linear maps.

For $x, y \in \mathbb{R}_+^d$, if we define

$$d_H(x,y) := \log \frac{\max_{1\le i\le d}\{x_i/y_i\}}{\min_{1\le i\le d}\{x_i/y_i\}},$$

then d_H induces a distance on R_+^{d-1}. Moreover, for any positive matrix A, we have

$$d_H(Ax, Ay) \le k(A) d_H(x,y), \quad x, y \in \mathbb{R}_+^d, \tag{5.8}$$

where $k(A) = \tanh(\Delta(A)/4) < 1$ and $\Delta(A) := \sup\{d_H(Ax, Ay) : x, y \in \mathbb{R}_+^d\}$.

Proposition 5.1.8. *Let $\{M_i\}_{i\in I}$ be a finite set of positive matrices and (A_n) be a a sequence of matrices taking values in $\{M_i\}_{i\in I}$. Define $S_n = A_n A_{n-1} \cdots A_1$. Then*

(a) for $x, y \in \mathbb{R}_+^d$, there exists a constant $0 < \delta < 1$ such that

$$d_H(S_n x, S_n y) \le \delta^n d_H(x,y);$$

(b) for $x, y \in \mathbb{R}_+^d$, there exist constants $0 < \delta < 1, \kappa > 0$ such that when n is sufficiently large, we have

$$\|\bar{S}_n x - \bar{S}_n y\| \le \kappa \delta^n d_H(x,y),$$

where $\bar{S}_n$ is the induced map on S^{d-1} defined by (5.1)*;*

(c) for $x, y \in \mathbb{R}^d_+$, there exist constants $0 < \delta < 1, \kappa > 0$ such that when n is sufficiently large, we have

$$d(\hat{S}_n x, \hat{S}_n y) \leq \kappa \delta^n d_H(x, y),$$

where $\hat{S}_n$ is the induced map on R^{d-1}_+ and d is the angle distance given by (5.7)*;*

(d) for a bounded domain $W \subset R^{d-1}_+$ and $x, y \in W$, there exist constants $0 < \delta < 1, \kappa > 0$ such that when n is sufficiently large, we have

$$\|\hat{S}_n x - \hat{S}_n y\| \leq \kappa \delta^n d_H(x, y),$$

where $\hat{S}_n$ is the induced map on R^{d-1}_+.

All the δ's above can be chosen to be independent of the choice of the sequence $(A_n)_n$.

Proof. (a) Choose $\delta = \max_{i \in I} k(M_i) < 1$, then the conclusion follows by (5.8).

(b) Here we choose $\|\cdot\|$ to be the ℓ_2-norm. Observe first that for any $\bar{x} = (x_1, \ldots, x_d), \bar{y} = (y_1, \ldots, y_d) \in S^{d-1}_+$, we have

$$\min_{1 \leq i \leq d} \frac{x_i}{y_i} \leq 1 \leq \max_{1 \leq i \leq d} \frac{x_i}{y_i}.$$

Therefore

$$\begin{aligned}
\|\bar{x} - \bar{y}\| = \left(\sum_{i=1}^d |x_i - y_i|^2 \right)^{1/2} &= \left(\sum_{i=1}^d \left| \frac{x_i}{y_i} - 1 \right|^2 y_i^2 \right)^{1/2} \\
&\leq \max_{1 \leq i \leq d} \left| \frac{x_i}{y_i} - 1 \right| \left(\sum_{1 \leq i \leq d} y_i^2 \right)^{1/2} \leq \left| \max_{1 \leq i \leq d} \frac{x_i}{y_i} - \min_{1 \leq i \leq d} \frac{x_i}{y_i} \right| \\
&\leq \exp(d_H(x, y)) - 1. \qquad (5.9)
\end{aligned}$$

Now we have,

$$\|\bar{S}_n x - \bar{S}_n y\| \leq \exp(d_H(S_n x, S_n y)) - 1 \leq d_H(S_n x, S_n y) \exp(d_H(S_n x, S_n y)).$$

When n is sufficiently large, by (a), there exists $\delta < 1$ such that

$$d_H(S_n x, S_n y) < \delta^n d_H(x, y) \quad \text{and} \quad \exp(d_H(S_n x, S_n, y)) < 2.$$

Thus $\|\bar{S}_n x - \bar{S}_n y\| \leq 2\delta^n d_H(x, y)$.

(c) We only need to prove that for $x, y \in \mathbb{R}^d$, there exists some constant $\kappa_0 > 0$ such that

$$d(x, y) \leq \kappa_0 \|\bar{x} - \bar{y}\|.$$

We choose $\|\cdot\|$ to be the ℓ_2-norm. Then applying the law of sines to the triangle formed by $\bar{x}, \bar{y}$ and the origin O in $\mathbb{R}^d$, then we can see $\|\bar{x} - \bar{y}\|/d(x, y)$ equals the diameter of this triangle's circumcircle. As this circle has a chord of length $1 = \|\bar{x} - O\|$, the diameter must be no smaller than 1. That is to say, $d(x, y) \leq \|\bar{x} - \bar{y}\|$. Thus by applying (b),

$$d(\hat{S}_n x, \hat{S}_n y) = d(\bar{S}_n x, \bar{S}_n y) \leq \|\bar{S}_n x - \bar{S}_n y\| \leq \kappa \delta^n d_H(x, y).$$

(d) This conclusion follows easily from Lemma 5.1.5.

□

Here the inequality (5.9) is due to [6]. More properties on Hilbert's projective metric can also be found in this article.

5.2 Transfer Operators

In Chapter 2, we proved that if μ is a probability measure on $GL(d, \mathbb{R})$, then there exists a μ-invariant measure ν on $\mathbb{R}P^{d-1}$, in the sense that $\mu * \nu = \nu$, i.e., for any bounded Borel function w on $\mathbb{R}P^{d-1}$, we have

$$\int w(x) \, \mathrm{d}\nu(x) = \int \mathrm{d}\nu(x) \int w(A \cdot x) \, \mathrm{d}\mu(A). \tag{5.10}$$

Now if we define a linear map $\mathcal{L}$ on bounded Borel functions by

$$(\mathcal{L}w)(x) = \int w(A \cdot x) \, \mathrm{d}\mu(A),$$

(5.10) can be written as

$$\int w(x) \, \mathrm{d}\nu(x) = \int (\mathcal{L}w)(x) \, \mathrm{d}\nu(x), \tag{5.11}$$

for each bounded Borel function w on $\mathbb{R}P^{d-1}$. More generally, we can give the following definition.

Definition 5.2.1. Let $\mathcal{G}$ be a function space on some subspace of $\mathbb{R}P^{d-1}$ (or of the complex projective space $\mathbb{C}P^{d-1}$), μ be a probability measure on $GL(d,\mathbb{R})$. If for any $w \in \mathcal{G}$, $w(A \cdot x)$ is integrable with respect to μ and

$$\left(x \mapsto \int w(A \cdot x)\, \mathrm{d}\mu(A)\right) \in \mathcal{G},$$

then we call define $\mathcal{L}$ on $\mathcal{G}$ by

$$(\mathcal{L}w)(x) = \int w(A \cdot x)\, \mathrm{d}\mu(A),$$

and say $\mathcal{L}$ is a **transfer operator** with respect to μ on $\mathcal{G}$.

When μ is a discrete measure supported by a finite set of invertible and positive matrices $\{M_1, \ldots, M_k\}$ with the corresponding probability vector $(p_1, \ldots, p_k)$ such that $p_1 + \cdots + p_k = 1$. We can write the transfer operator $\mathcal{L}$ as [1]

$$(\mathcal{L}w)(x) = \sum_{i=1}^{k} p_i w(\hat{M}_i x), \tag{5.12}$$

where $\hat{M}_i$ is the induced map of M_i on R_+^{d-1} given by (5.3).

It would be convenient to restrict $\mathcal{L}$ to subspaces of the bounded Borel functions. In this chapter, we will focus mainly on the case when $\mathcal{G}$ is chosen to be the space $\mathcal{C}^\omega(V)$ of real analytic functions on $V \subset R_+^{d-1}$ with a continuous extension on $\overline{V}$, or $\mathcal{A}_\infty(U)$ of complex analytic functions on $U \subset C^{d-1}$ with continuous extensions on $\overline{U}$, where V and U are chosen according to Corollary 5.1.4. For any $w \in \mathcal{A}_\infty(U)$, we define $\|w\| := \sup_{z \in U} |w(z)| < \infty$, then $(\mathcal{A}_\infty(U), \|\cdot\|)$ is a Banach space.

Later, we will also study the case when $\mathcal{G} = \mathscr{H}(U)$, the space of analytic functions on the domain $U \subset C^{d-1}$. We can easily see that $\mathcal{A}_\infty(U)$ is a subspace of $\mathscr{H}(U)$.

Denote by $\mathcal{C}(\overline{V}, \mathbb{R})$ the space of continuous real functions on $\overline{V}$ equipped with the uniform norm $|\cdot|_\infty$. Then by the properties of V and U given in Corollary 5.1.4, $\mathcal{L}$ is a well-defined bounded linear operator on $\mathcal{C}^\omega(V)$, $\mathcal{A}_\infty(U)$, $\mathscr{H}(U)$, and $\mathcal{C}(\overline{V}, \mathbb{R})$.

The main goal of this section is to introduce Theorem 5.2.4, which gives the spectral properties of transfer operators. The following lemma in complex analysis will be frequently used in the proof of Theorem 5.2.4 and in later sections. It is also known as Montel's Theorem.

[1] Unless otherwise indicated, we will always identify $\Delta \subset \mathbb{R}P^{d-1}$ with R_+^{d-1} in the rest of the dissertation.

Lemma 5.2.2. *Let W be a domain of $\mathbb{C}^d$ and Λ be a subset of $\mathcal{A}_\infty(W)$ which is uniformly bounded. Then for any compact subset $K \subset W$, Λ is equicontinuous on K. In other words, for any $\epsilon > 0$, there exists $\delta > 0$ (not depending on w) such that whenever $z, z' \in K$ satisfies $\|z - z'\| < \delta$, we have $|w(z) - w(z')| < \epsilon$, for any $w \in \Lambda$. In fact, there exist $\delta' > 0$ and $C_K > 0$ (both not depending on w) such that whenever $\|z - z'\| < \delta'$, we have*

$$|w(z) - w(z')| \leq C_K \|z - z'\|.$$

Consequently, there exists some $C'_K > 0$ (not depending on w) such that

$$\left|\frac{\partial w}{\partial z_i}(z)\right| \leq C'_K, \quad 1 \leq i \leq d,\ z \in K.$$

Proof. By the equivalence of norms, we have the freedom to choose $\|z\| := \max_i |z_i|$ in this proof. Define

$$\Gamma_i(z_i, r) := \{\zeta_i \in \mathbb{C} : |\zeta_i - z_i| = r\},$$

$$\Omega_i(z_i, r) = \{\zeta_i \in \mathbb{C} : |\zeta_i - z_i| < r\},$$

for each $1 \leq i \leq d$. Denote $\Gamma(z, r) := \prod_{i=1}^d \Gamma_i(z_i, r)$ and $\Omega(z, r) := \prod_{i=1}^d \Omega_i(z_i, r)$.

Fix a compact subset $K \subset W$, then there exists $\delta_K > 0$ such that $\Omega(z, 3\delta_K) \subset W$ for each $z \in K$. Now if $z = (z_1, \ldots, z_d), z' = (z'_1, \ldots, z'_d) \in K$ satisfies

$$\|z - z'\| = \max_i |z_i - z'_i| < \delta_K,$$

Then for $\zeta_i \in \Gamma_i(z_i, 2\delta_K)$ $(1 \leq i \leq d)$ we have

$$|\zeta_i - z'_i| \geq |\zeta_i - z_i| - |z_i - z'_i| > \delta_K.$$

Since $\overline{\Omega(z, 2\delta_K)} \subset W$, by Cauchy's integral formula,

$$\begin{aligned}
&|w(z_1, z_2 \ldots, z_d) - w(z'_1, z_2, \ldots, z_d)| \\
\leq &\frac{1}{2\pi} \int_{|\zeta_1 - z_1| = 2\delta_K} \left| \frac{w(\zeta_1, z_2, \ldots, z_d)}{\zeta_1 - z_1} - \frac{w(\zeta_1, z_2, \ldots, z_d)}{\zeta_1 - z'_1} \right| |\,\mathrm{d}\zeta_1| \\
\leq &\frac{\|w\|}{2\pi} \int_{|\zeta_1 - z_1| = 2\delta_K} \frac{|z_1 - z'_1|}{|(\zeta_1 - z_1)(\zeta_1 - z'_1)|} |\,\mathrm{d}\zeta_1| \\
\leq &\frac{\|w\|}{2\pi} \frac{|z_1 - z'_1|}{(2\delta_K) \cdot \delta_K} \cdot (4\pi\delta_K) = \frac{\|w\|}{\delta_K} |z_1 - z'_1|
\end{aligned}$$

$$\leq \delta_K^{-1}\|w\|\|z-z'\|.$$

Therefore

$$\begin{aligned}|w(z)-w(z')| \leq & \sum_{i=1}^{d} |w(z_1',\ldots,z_{i-1}',z_i\ldots,z_d)-w(z_1',\ldots,z_i',z_{i+1},\ldots,z_d)| \\ \leq & \delta_K^{-1} d\|w\|\|z-z'\| \leq \delta_K^{-1} dB\|z-z'\|, \end{aligned} \tag{5.13}$$

where B denotes the uniform bound for functions in Λ.

Thus for $\epsilon > 0$, choose $\delta = \min\{\delta_K, \delta_K\epsilon/dB\} > 0$, then whenever $z, z' \in K$ satisfies $\|z-z'\| < \delta$, we have $|w(z)-w(z')| < \epsilon$.

Moreover, (5.13) implies that, for any $z \in K$, and each $1 \leq i \leq d$,

$$\left|\frac{\partial w}{\partial z_i}(z)\right| = \lim_{\Delta z_i \to 0} \frac{|w(z_1,\ldots,z_i+\Delta z_i,\ldots,z_d)-w(z_1,\ldots,z_i,\ldots,z_d)|}{|\Delta z_i|} \leq \frac{dB}{\delta_K}.$$

Therefore, the conclusion follows by defining $C_K = C_K' = dB/\delta_K$.

□

Lemma 5.2.3. *Let M be a positive $d \times d$ matrix, and $\hat{M}$ be the induced map on a bounded domain $W \subset C_+^{d-1}$. Then there exist constants $0 < D_M < \infty$ and $\eta > 0$ such that whenever $z, z' \in W$ satisfies $\|z-z'\| < \eta$, we have $\|\hat{M}z - \hat{M}z'\| \leq D_M\|z-z'\|$.*

Proof. If $M = (a_{ij})_{1\leq i,j\leq d}$, and write $\hat{M}$ as

$$(z_2,\ldots,z_d) \mapsto \left(\frac{a_{21}+\sum_{j=2}^d a_{2j}z_j}{a_{11}+\sum_{j=2}^d a_{1j}z_j}, \ldots, \frac{a_{d1}+\sum_{j=2}^d a_{dj}z_j}{a_{11}+\sum_{j=2}^d a_{1j}z_j}\right).$$

For some $2 \leq k \leq d$, write $\alpha_l = a_{k1}+\sum_{j\neq l} a_{kj}z_j$, $\beta_l = a_{kl}$, $\gamma_l = a_{11}+\sum_{j\neq l} a_{1j}z_j$, $\delta_l = a_{1l}$, for $2 \leq l \leq d$, then

$$\frac{\partial(\hat{M}z)_k}{\partial z_l} = \frac{\partial}{\partial z_l}\left(\frac{a_{k1}+\sum_{j=2}^d a_{kj}z_j}{a_{11}+\sum_{j=2}^d a_{1j}z_j}\right) = \frac{\partial}{\partial z_l}\left(\frac{\alpha_l+\beta_l z_l}{\gamma_l+\delta_l z_l}\right) = \frac{\beta_l\gamma_l-\alpha_l\delta_l}{(\gamma_l+\delta_l z_l)^2}$$

As $W \subset C_+^{d-1}$ is bounded, $|\alpha_l|$ and $|\gamma_l|$ are both bounded. Since for $z \in W$ we have $\operatorname{Re} z_j > 0$ $(2 \leq j \leq d)$ and M is positive, we have

$$|\gamma_l+\delta_l z_l|^2 \geq |\operatorname{Re}(\gamma_l+\delta_l z_l)|^2 \geq a_{11}^2,$$

thus

$$\left|\frac{\partial(\hat{M}z)_k}{\partial z_l}\right| = \frac{|\beta_l\gamma_l - \alpha_l\delta_l|}{|\gamma_l + \delta_l z_l|^2} \leq \frac{|\beta_l||\gamma_l| + |\delta_l||\alpha_l|}{a_{11}^2} < b_k < \infty, \tag{5.14}$$

for some constant $b_k > 0$ (not depending on $z \in W$ or $2 \leq l \leq d$). Thus by (5.14) and the definition of partial derivatives, there exists $\eta_l > 0$ such that whenever $|\Delta z_l| < \eta_l$,

$$\left|\left(\hat{M}(z_2, \ldots, z_l + \Delta z_l, \ldots, z_d)\right)_k - \left(\hat{M}(z_2, \ldots, z_l, \ldots, z_d)\right)_k\right| \leq b_k|\Delta z_l|.$$

If we choose $\|z\| := \max_{2\leq i\leq d}|z_i|$ and let $\eta = \min_{2\leq l\leq d}\eta_l, b_{\max} = \max_{2\leq k\leq d} b_k$, then when $\|z - z'\| < \eta$, we have for each $2 \leq k \leq d$

$$|(\hat{M}z - \hat{M}z')_k| \leq db_{\max}\|z - z'\|.$$

That is to say, $\|\hat{M}z - \hat{M}z'\| \leq D_M\|z - z'\|$ if we define $D_M := db_{\max} < \infty$.

□

Now we are ready for the following properties of transfer operators [48]. The proof is given by the author, inspired by Chapter 2 of [44].

Theorem 5.2.4. *(Perron-Frobenius-Ruelle Theorem) Let $\mathcal{L}$ be the linear operator on $\mathcal{A}_\infty(U)$ defined as* (5.12)*, then*

(a) The constant function $\mathbf{1}$ is the eigenvector of $\mathcal{L}$ associated to the simple maximal positive eigenvalue 1, and the rest of the spectrum of $\mathcal{L}$ lies in the unit open ball $\{z \in \mathbb{C} : |z| < 1\}$;

(b) $\mathcal{L}$ is a compact operator on $\mathcal{A}_\infty(U)$;

(c) There is a unique bounded positive linear functional ν on $\mathcal{A}_\infty(U)$ satisfying $\|\nu\| = 1$ and that for any $w \in \mathcal{A}_\infty(U)$,

$$\nu(\mathcal{L}w) = \nu(w);$$

(d) In the uniform topology,

$$\lim_{n\to\infty}(\mathcal{L}^n w)(z) = \nu(w), \quad w \in \mathcal{A}_\infty(U).$$

Moreover, when $\mathcal{L}$ is regarded as acting on $\mathcal{C}(\overline{V}, \mathbb{R})$, then

(e) There is a unique probability measure ν^* on $\overline{V}$ satisfying that for any continuous real function $f \in \mathcal{C}(\overline{V}, \mathbb{R})$,

$$\int \mathcal{L}f d\nu^* = \int f d\nu^*;$$

(f) In the uniform topology,

$$\lim_{n\to\infty}(\mathcal{L}^n f)(x) = \int f \,\mathrm{d}\nu^*, \quad f \in \mathcal{C}(\overline{V}, \mathbb{R}).$$

Proof. Part 1. It is obvious that the constant function $\mathbf{1}$ on $\overline{U}$ is an eigenfunction corresponding to the eigenvalue 1. Now we prove 1 is a simple eigenvalue by discussing the following two cases.

Case (i). There is another eigenfunction $g \in \mathcal{A}_\infty(U)$ corresponding to the eigenvalue 1. Since $\hat{M}_i$ maps U strictly inside U for each $1 \le i \le k$, we can assume $K \subset C^{d-1}$ is a compact set satisfying $\bigcup_{i=1}^k \hat{M}_i U \subset K \subset U$. Assume $\sup_{z\in U}|g(z)|$ is attained at some $\zeta \in \partial U$, then $\hat{M}_i\zeta \in K$ for each $1 \le i \le k$, and

$$|g(\zeta)| = |(\mathcal{L}g)(\zeta)| \le \sum_{i=1}^k p_i |g(\hat{M}_i\zeta)| \le \sup_{z\in U}|g(z)|.$$

Therefore g has to be a constant function on $\overline{U}$ by the maximum modulus principle.

Case (ii). The eigenvalue 1 admits a Jordan block. Then there exists some $g \in \mathcal{A}_\infty(U)$ such that $\mathcal{L}g = \mathbf{1} + g$. This implies that $\mathcal{L}^n g = n + g$ for each $n \ge 1$, and that $\|\mathcal{L}^n g\| > \|g\|$ when n is sufficiently large. However, if $\sup_{z\in U}|(\mathcal{L}g)(z)|$ is attained at $\zeta \in \partial U$, then

$$\|\mathcal{L}g\| = |(\mathcal{L}g)(\zeta)| \le \sum_{i=1}^k p_i |g(\hat{M}_i\zeta)| \le \|g\|. \tag{5.15}$$

Therefore $\|\mathcal{L}^n g\| \le \|g\|$ for any $n \ge 1$, which leads to a contradiction. This proves that 1 is a simple eigenvalue of $\mathcal{L}$.

Part 2. Now let (w_n) be a uniformly bounded sequence in $\mathcal{A}_\infty(U)$. Let K be a compact set satisfying $\bigcup_{i=1}^k \hat{M}_i U \subset K \subset U$. Note $\|\mathcal{L}w_n\| \le \|w_n\|$ by (5.15), i.e., $(\mathcal{L}w_n)$ is uniformly bounded. And by Lemma 5.2.2 and Lemma 5.2.3, whenever $z, z' \in U$ satisfies that $\|z - z'\|$ is sufficiently small, we have $\hat{M}z, \hat{M}z' \in K$ can be sufficiently close and

$$|(\mathcal{L}w_n)(z) - (\mathcal{L}w_n)(z')| \le \sum_{i=1}^k p_i |w_n(\hat{M}_i z) - w_n(\hat{M}_i z')|$$

$$\le \sum_{i=1}^k p_i C_K \|\hat{M}_i z - \hat{M}_i z'\|$$

$$\leq C_K \left(\sup_{1\leq i\leq k} D_{M_i} \right) \|z - z'\|,$$

for constants $C_K, D_{M_i} > 0$ $(1 \leq i \leq k)$. Therefore $(\mathcal{L}w_n)$ is also equicontinuous. By the Arzelà-Ascoli theorem, $(\mathcal{L}w_n)$ contains a convergent subsequence. This proves that $\mathcal{L}$ is a compact operator on $\mathcal{A}_\infty(U)$.

Part 3. By Schauder's theorem, the adjoint operator $\mathcal{L}^* : (\mathcal{A}_\infty(U))' \to (\mathcal{A}_\infty(U))'$ is compact with a fixed point $\nu \in (\mathcal{A}_\infty(U))'$ corresponding to the eigenvalue 1 with $\nu(\mathbf{1}) = 1$. That is to say, for $w \in \mathcal{A}_\infty(U)$

$$\nu(\mathcal{L}w) = (\mathcal{L}^*\nu)(w) = \nu(w). \tag{5.16}$$

Part 4. Given any $w \in \mathcal{A}_\infty(U)$, by (5.15) we know the sequence $(\mathcal{L}^n w)_{n\geq 1}$ of functions is uniformly bounded by $\|w\|$. By letting $w_n = \mathcal{L}^{n-1}w$ $(n \geq 1)$ in part 2 of the proof, we know $(\mathcal{L}^n w)_{n\geq 1}$ is also equicontinuous. Therefore by the Arzelà-Ascoli theorem, there exists a convergent subsequence $(\mathcal{L}^{n_\alpha} w)_\alpha$. Write w^* as the limit point of $(\mathcal{L}^{n_\alpha} w)_\alpha$ in the uniform topology.

On the one hand, by (5.15), we have $\|\mathcal{L}^n w^*\| \leq \|w^*\|$ for each $n \geq 0$. On the other hand, (5.15) also implies $\|w^*\| \leq \|\mathcal{L}^{n+m} w\|$ for each $m, n \geq 0$. And letting $m = n_\alpha$ and $\alpha \to \infty$ gives $\|w^*\| \leq \|\mathcal{L}^n w^*\|$ for each $n \geq 0$. Consequently, $\|w^*\| = \|\mathcal{L}^n w^*\|$ for each $n \geq 0$.

In particular, when $n = 1$, assume $\sup_{z\in U} |(\mathcal{L}w^*)(z)|$ is attained at $\zeta' \in \partial U$ and $\sup_{z\in U} |w^*(z)|$ is attained at $\zeta \in \partial U$, then

$$\|w^*\| = |w^*(\zeta)| = |(\mathcal{L}w^*)(\zeta')| \leq \sum_{i=1}^{k} p_i |w^*(\hat{M}_i\zeta')| \leq \|w^*\|.$$

This shows the supremum $\|w^*\| = |w^*(\hat{M}_i\zeta')|$ can be reached at $\hat{M}_i\zeta' \in U$, which does not lie on the boundary. Thus by the maximum modulus principle w^* is a constant function on $\overline{U}$. Now by part 3, we have

$$\nu(w) = \lim_{\alpha\to\infty} \nu(\mathcal{L}^{n_\alpha} w) = \nu(w^*) = w^*.$$

We can apply the same proof to every subsequence of $(\mathcal{L}^n w)_n$ and conclude that every subsequence of $(\mathcal{L}^n w)_n$ has a further subsequence uniformly converging to the same limit, a constant function $w^* = \nu(w)$. Therefore for $w \in \mathcal{A}_\infty(U)$, in the uniform topology we have

$$\nu(w) = \lim_{n\to\infty} \mathcal{L}^n w. \tag{5.17}$$

As the evaluation of ν is completely determined by (5.17), it is unique. Moreover, since $|\nu(w)| = \lim_{n\to\infty} \|\mathcal{L}^n w\| \le \|w\|$ and $\nu(\mathbf{1}) = 1$, we have $\|\nu\| = 1$.

Part 5. In this part, we prove the existence and uniqueness of a probability measure on $\overline{V}$ satisfying (e).

Starting with the uniqueness, let ν^* be a probability measure on $\overline{V}$ satisfying (e). Then by applying (5.17) to $f \in \mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$, we have

$$\nu^*(f) = \lim_{n\to\infty} \nu^*(\mathcal{L}^n f) = \nu^*(\nu(f)) = \nu(f).$$

Since $\mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$ is dense in $\mathcal{C}(\overline{V}, \mathbb{R})$ in the uniform topology by the Stone-Weierstrass theorem, we can see the evaluation $\nu^*(f)$ for each $f \in \mathcal{C}(\overline{V}, \mathbb{R})$ is completely determined. Thus the uniqueness is proved.

Now we construct a probability measure ν^* satisfying (e). Define a linear functional ν_0 to be ν restricted to $\mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$. Since $\mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$ is dense in $\mathcal{C}(\overline{V}, \mathbb{R})$ in the uniform topology, we can extend ν_0 on $\mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$ to ν^* on $\mathcal{C}(\overline{V}, \mathbb{R})$ by defining

$$\nu^*(f) = \lim_{\alpha\to\infty} \nu_0(f_\alpha), \tag{5.18}$$

for $f_\alpha \in \mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$ satisfying $|f_\alpha - f|_\infty = \sup_{x\in V} |f_\alpha(x) - f(x)| \to 0$ $(\alpha \to \infty)$.

This extension is well-defined. In fact, let $f_\alpha, f'_\alpha \in \mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$ be two sequences of real functions converging to the same limit f uniformly. Then

$$|f_\alpha - f'_\alpha|_\infty \le |f_\alpha - f|_\infty + |f'_\alpha - f|_\infty \to 0 \ (\alpha \to \infty),$$

and

$$|\mathcal{L}^n f_\alpha - \mathcal{L}^n f'_\alpha| \le \mathcal{L}^n |f_\alpha - f'_\alpha| \le |f_\alpha - f'_\alpha|_\infty \to 0 \ (\alpha \to \infty).$$

Therefore, by applying (5.17) restricted on $\overline{V}$,

$$\lim_{\alpha\to\infty} |\nu_0(f_\alpha) - \nu_0(f'_\alpha)| = \lim_{\alpha\to\infty} \lim_{n\to\infty} |\mathcal{L}^n f_\alpha - \mathcal{L}^n f'_\alpha| = \lim_{n\to\infty} \lim_{\alpha\to\infty} |\mathcal{L}^n f_\alpha - \mathcal{L}^n f'_\alpha| = 0.$$

This implies that the definition (5.18) does not depend on the choice of (f_α) if the limit exists. And if we replace f'_α by $f_{\alpha+\alpha'}$ for $\alpha' \ge 0$, then with the exactly same proof we can see $\nu_0(f_\alpha)$ is a Cauchy sequence, therefore the limit (5.18) always exists.

Note when $f \in \mathcal{C}(\overline{V}, \mathbb{R})$ is non-negative, we obviously have $\nu^*(f) \ge 0$ by choosing a sequence of non-negative f_α uniformly converging to f and noticing $\nu_0(f_\alpha) = \lim_{n\to\infty} \mathcal{L}^n f_\alpha \ge$

0. That is to say, ν is a positive linear functional on $\mathcal{C}(\overline{V},\mathbb{R})$. Then by the Riesz representation theorem, there exists a unique positive measure (which we still denote by ν^*) on $\overline{V}$ satisfying

$$\nu^*(f) = \int f \,\mathrm{d}\nu^*, \quad f \in \mathcal{C}(\overline{V},\mathbb{R}).$$

As $\nu^*(\mathbf{1}) = \nu(\mathbf{1}) = 1$, ν^* is in fact a probability measure.

Now we verify that ν^* satisfies the property given in (e). Notice, for $f \in \mathcal{C}(\overline{V},\mathbb{R})$ with $|f_\alpha - f|_\infty \to 0$ in $\mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$,

$$\begin{aligned}|\nu^*(\mathcal{L}f) - \nu^*(f)| &\le |\nu^*(\mathcal{L}f) - \nu^*(\mathcal{L}f_\alpha)| + |\nu^*(\mathcal{L}f_\alpha) - \nu^*(f_\alpha)| + |\nu^*(f_\alpha) - \nu^*(f)| \\ &\to 0 \quad (\alpha \to \infty),\end{aligned}$$

by (5.18), (5.16), and that $|\mathcal{L}f_\alpha - \mathcal{L}f|_\infty \le |f_\alpha - f|_\infty \to 0$. Thus $\nu^*(\mathcal{L}f) = \nu^*(f)$ for $f \in \mathcal{C}(\overline{V},\mathbb{R})$.

Moreover, for $f \in \mathcal{C}(\overline{V},\mathbb{R})$ and $f_\alpha \in \mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$ with $|f_\alpha - f|_\infty \to 0$,

$$\begin{aligned}|\mathcal{L}^n f - \nu^*(f)|_\infty &\le |\mathcal{L}^n f - \mathcal{L}^n f_\alpha|_\infty + |\mathcal{L}^n f_\alpha - \nu^*(f_\alpha)|_\infty + |\nu^*(f_\alpha) - \nu^*(f)| \\ &\le |f - f_\alpha|_\infty + \|\mathcal{L}^n f_\alpha - \nu(f_\alpha)\| + |\nu_0(f_\alpha) - \nu^*(f)|\end{aligned}$$

Letting $n \to \infty$ and applying (5.17) shows

$$\limsup_{n\to\infty} |\mathcal{L}^n f - \nu^*(f)|_\infty \le |f - f_\alpha|_\infty + |\nu_0(f_\alpha) - \nu^*(f)|.$$

Then letting $\alpha \to \infty$ shows $|\mathcal{L}^n f - \nu^*(f)|_\infty \to 0$ by (5.18). This proves (f).

Part 6. In the end, we prove that the rest of the spectrum has modulus strictly smaller than 1. Define

$$\Gamma := \{w \in \mathcal{A}_\infty(U) : \|w\| \le 1, \nu(w) = 0\}, \quad \mathcal{L}\Gamma := \{\mathcal{L}w : w \in \Gamma\} \subset \Gamma.$$

By (5.17), we know $\mathcal{L}^n w$ converges to 0 for $w \in \Gamma$ as $n \to \infty$. And by part 2 and the Arzelà-Ascoli Theorem, we know $\mathcal{L}\Gamma$ is relatively compact in the uniform topology . Thus for some $0 < \delta < 1$, there exists $N > 0$ such that $\|\mathcal{L}^N u\| \le \delta < 1$ for any $u \in \mathcal{L}\Gamma$. That is to say $\|\mathcal{L}^{N+1} w\| \le \delta < 1$ for any $w \in \Gamma$. Thus $\|(\mathcal{L}|_\Gamma)^{N+1}\|^{1/(N+1)} \le \delta^{1/(N+1)}$, i.e., the spectral radius of $\mathcal{L}|_\Gamma$ is strictly smaller than 1. This completes the proof.

□

Remark 5.2.5. (1) We will see again that transfer operators are compact with a different proof in Section 5.5.

(2) Note that statement (e) is similar to (5.11), although in (5.11), the μ-invariant measure is supported on $\mathbb{R}P^{d-1}$; while in (e), ν is supported on $V \subset R_+^{d-1}$. We mentioned in Remark 5.1.1 that the μ-invariant measure on $\mathbb{R}P^{d-1}$ and the one supported by $V \subset R_+^{d-1}$ are equivalent, thus these two invariant measures can be identified.

(3) By Theorem 2.3.11, we can see that the Lyapunov exponent induced by the discrete measure $\mu = \sum_{i=1}^{k} p_i \delta_{M_i}$ can be given by

$$\gamma = \sum_{i=1}^{k} p_i \nu^*(f_i),$$

where

$$f_i(x) := \log \|M_i x\| \in \mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U).$$

5.3 Parametrised Transfer Operators

In this section, we study the family of parametrised transfer operators. They can be regarded as a perturbed version of the previously defined transfer operators. We will prove that they also possess a simple maximal top eigenvalue, and this perturbed top eigenvalue can be used to compute the Lyapunov exponent. In this section, V and U are chosen according to Corollary 5.1.4 as usual.

Definition 5.3.1. Given a finite set $\{M_i\}_{1\le i\le k}$ of invertible and positive matrices, and a probability vector $(p_i)_{1\le i\le k}$. Choose a proper domain $U \subset C^{d-1}$ and define the Banach space $\mathcal{A}_\infty(U)$ as before. Let $\varphi_i(z) \in \mathcal{A}_\infty(U)$ for $1 \le i \le k$, and define an operator $\mathcal{L} : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$ by

$$(\mathcal{L}w)(z) = \sum_{i=1}^{k} e^{t\varphi_i(z)} p_i w(\hat{M}_i z). \tag{5.19}$$

They are called the **weighted transfer operators** (or **parametrised transfer operators**) with weights $(e^{t\varphi_i})_{1\le i\le k}$, the probability vector $(p_i)_{1\le i\le k}$ and the parameter t.

It is easy to see that

$$\|\mathcal{L}w\| \le \sum_{i=1}^{k} p_i e^{|t|\|\varphi_i\|} \|w\|,$$

and therefore $\mathcal{L}$ is a bounded linear operator. Now we introduce the following theorem analogous to Theorem 5.2.4.

Theorem 5.3.2. *Let $\mathcal{L}$ be a weighted transfer operator on $\mathcal{A}_\infty(U)$ defined in Definition 5.3.1 with weights $(e^{t\varphi_i})_{1\le i\le k}$ $(t\in\mathbb{R})$ satisfying*

(i) $\varphi_i \in \mathcal{A}_\infty(U)$ *are real analytic when restricted on* V, $(1\le i\le k)$;

(ii) $\dfrac{\partial\varphi_i}{\partial z_j}\in\mathcal{A}_\infty(U)$, *i.e.,* $\sup\limits_{z\in U}\left|\dfrac{\partial\varphi_i}{\partial z_j}(z)\right|<\infty$ *for* $1\le i\le k, 2\le j\le d$;

then when $|t|$ is sufficiently small,

(a) $\mathcal{L}$ *possesses a simple maximal positive eigenvalue β and the corresponding eigenfunction h can be chosen to be strictly positive on $\overline{V}$;*

(b) The rest of the spectrum lies in the open ball $\{z\in\mathbb{C}:|z|<\beta\}$;

(c) There exists a unique bounded positive linear functional ν on $\mathcal{A}_\infty(U)$ satisfying $\|\nu\|=1$ and that

$$\nu(\mathcal{L}w)=\beta\nu(w),\quad w\in\mathcal{A}_\infty(U);$$

(d) For any $w\in\mathcal{A}_\infty(U)$, in the uniform topology,

$$\lim_{n\to\infty}\beta^{-n}\mathcal{L}^n w=h(z)\nu(w),$$

where β and h are given in (a) with $\nu(h)=1$.

Moreover, when $\mathcal{L}$ is regarded as acting on $\mathcal{C}(\overline{V},\mathbb{R})$, then

(e) There exists a unique probability measure ν^* on $\overline{V}$ satisfying

$$\int\mathcal{L}f\,\mathrm{d}\nu^*=\int f\,\mathrm{d}\nu^*,\quad f\in\mathcal{C}(\overline{V},\mathbb{R});$$

(f) For any $f\in\mathcal{C}(\overline{V},\mathbb{R})$, in the uniform topology,

$$\lim_{n\to\infty}\beta^{-n}\mathcal{L}^n f=h(x)\int f\,\mathrm{d}\nu^*\quad(n\to\infty),$$

where β is given in (a), and $h(x)$ is the restriction of h to $\overline{V}$ with $\int h\,\mathrm{d}\nu^*=1$.

Proof. Part 1. In this part, we construct an eigenvalue β with a corresponding eigenfunction h; we show that $\beta > 0$ and h is non-negative on $\overline{V}$.

Let $w \in \mathcal{A}_\infty(U)$ with $\|w\| \leq 1$, and K be the compact set such that $\bigcup_{i=1}^k \hat{M}_i U \subset K \subset U$, then by Lemma 5.2.2 and Lemma 5.2.3, for $z, z' \in U$ sufficiently close, since $\hat{M}_i z, \hat{M}_i z' \in K$ can be sufficiently close,

$$|w(\hat{M}_i z) - w(\hat{M}_i z')| \leq C_K \|\hat{M}_i z - \hat{M}_i z'\| \leq C_K \left(\sup_{1 \leq i \leq k} D_{M_i} \right) \|z - z'\|,$$

for constants $C_K, D_{M_i} > 0$. Denote $D = \sup_{1 \leq i \leq k} D_{M_i} C_K < \infty$. Now for $z, z' \in U$ sufficiently close,

$$\begin{aligned}
&\left| \mathcal{L}\left(w + \frac{1}{n}\right)(z) - \mathcal{L}\left(w + \frac{1}{n}\right)(z') \right| \\
\leq &\sum_{i=1}^k p_i \left| e^{t\varphi_i(z)} \left(w(\hat{M}_i z) + \frac{1}{n} \right) - e^{t\varphi_i(z')} \left(w(\hat{M}_i z') + \frac{1}{n} \right) \right| \\
\leq &\sum_{i=1}^k p_i \left| e^{t\varphi_i(z)} - e^{t\varphi_i(z')} \right| \left(\|w\| + \frac{1}{n} \right) + \sum_{i=1}^k p_i \|e^{t\varphi_i}\| |w(\hat{M}_i z) - w(\hat{M}_i z')| \\
\leq &\sup_{1 \leq i \leq k} \| \mathrm{d} e^{t\varphi_i} \| \left(\|w\| + \frac{1}{n} \right) \|z - z'\| + D \sup_{1 \leq i \leq k} \|e^{t\varphi_i}\| \|z - z'\|.
\end{aligned}$$

Note $\|w\| \leq 1$. Define

$$D' := 2 \sup_{1 \leq i \leq k} \| \mathrm{d} e^{t\varphi_i} \| + D \sup_{1 \leq i \leq k} \|e^{t\varphi_i}\| < \infty,$$

by condition (ii). We have

$$\left| \mathcal{L}\left(w + \frac{1}{n}\right)(z) - \mathcal{L}\left(w + \frac{1}{n}\right)(z') \right| \leq D' \|z - z'\|. \tag{5.20}$$

Moreover, if w is non-negative (therefore real) on $\overline{V}$, then for some $x_0 \in \overline{V}$

$$\begin{aligned}
\sup_{z \in U} \left| \mathcal{L}\left(w + \frac{1}{n}\right) \right| \geq \sup_{x \in V} \left| \mathcal{L}\left(w + \frac{1}{n}\right) \right| &= \sum_{i=1}^k p_i e^{t\varphi_i(x_0)} \left(w(\hat{M}_i x_0) + \frac{1}{n} \right) \\
&\geq \frac{1}{n} \sum_{i=1}^k p_i e^{-|t\varphi_i(x)|_\infty} =: D_n > 0.
\end{aligned} \tag{5.21}$$

Define

$$\Lambda_n := \{w \in \mathcal{A}_\infty(U) : \|w\| \le 1, \|\partial w/\partial z_i\| \le D'/D_n,\ 2 \le i \le d,$$
$$\text{and } w \text{ is (real and) non-negative on } \overline{V}\}.$$

And define a map L_n on Λ_n by

$$L_n(w) = \frac{\mathcal{L}(w + 1/n)}{\|\mathcal{L}(w + 1/n)\|}.$$

Note functions in

$$\Gamma_n := \{w \in \mathcal{A}_\infty(U) : \|w\| \le 1, \|\partial w/\partial z_i\| \le D'/D_n,\ 2 \le i \le d\}$$

are uniformly bounded and equicontinuous, thus by the Arzelà-Ascoli theorem, Γ_n is compact in the uniform topology. Λ_n is closed in Γ_n, therefore is compact as well. Moreover, it is not hard to see that Λ_n is convex; and by (5.20) and (5.21), when $z, z' \in U$ are sufficiently close

$$\begin{aligned}
&|L_n(w)(z) - L_n(w)(z')| \\
=&\|\mathcal{L}(w + 1/n)\|^{-1} \left| \mathcal{L}\left(w + \frac{1}{n}\right)(z) - \mathcal{L}\left(w + \frac{1}{n}\right)(z') \right| \qquad (5.22) \\
\le& D_n^{-1} D' \|z - z'\|.
\end{aligned}$$

Thus L_n maps Λ_n into Λ_n. As Λ_n is compact and convex, by the Schauder-Tychonoff fixed point theorem, there exists a fixed point $h_n \in \Lambda_n$ such that $L_n(h_n) = h_n$, i.e.,

$$\mathcal{L}\left(h_n + \frac{1}{n}\right) = \beta_n h_n,$$

where

$$\beta_n = \|\mathcal{L}(h_n + 1/n)\| \ge 0. \qquad (5.23)$$

As $h_n \in \Lambda_n$ is non-negative on $\overline{V}$, we have

$$\beta_n h_n = \sum_{i=1}^{k} p_i e^{t\varphi_i(x)} \left(h_n(\hat{M}_i x) + \frac{1}{n} \right) \ge \left(\inf_{x \in V} h_n + \frac{1}{n} \right) \sum_{i=1}^{k} p_i e^{-|t\varphi_i|_\infty} > 0.$$

Therefore $\inf_{x \in V} h_n(x) > 0$ and

$$\beta_n \inf_{x \in V} h_n \ge \left(\inf_{x \in V} h_n + \frac{1}{n} \right) \sum_{i=1}^{k} p_i e^{-|t\varphi_i|_\infty},$$

which implies

$$\beta_n \geq \sum_{i=1}^{k} p_i e^{-|t\varphi_i|_\infty} =: D''. \tag{5.24}$$

Note $D'' > 0$ does not depend on n. Define

$$\Lambda'' := \{w \in \mathcal{A}_\infty(U) : \|w\| \leq 1, \|\partial w/\partial z_i\| \leq D'/D'',\ 2 \leq i \leq d,$$
$$\text{and } w \text{ is (real and) non-negative on } \overline{V}\}.$$

Then by (5.22) and (5.23), we have $h_n \in \Lambda''$ for all $n \geq 1$. Note Λ'' is compact with respect to the uniform topology with the same reasons for Λ_n being compact. Now we can choose a convergent subsequence $(h_{n_\alpha})_\alpha$ of $(h_n)_n$ such that $h_{n_\alpha} \to h$ uniformly as $\alpha \to \infty$, then

$$\mathcal{L}h = \beta h,$$

where

$$\beta = \|\mathcal{L}h\| \geq D'' = \sum_{i=1}^{k} p_i e^{-|t\varphi_i|_\infty} > 0, \tag{5.25}$$

by (5.24) and continuity. Moreover, $h \in \Lambda''$ is non-negative on $\overline{V}$.

Part 2. Now we prove the eigenfunction h is strictly positive on $\overline{V}$. Denote by $\mathcal{Z}$ the set of zeros of h and assume $\mathcal{Z} \neq \emptyset$ with $x_0 \in \mathcal{Z}$. Then by

$$0 = \beta h(x_0) = \sum_{i=1}^{k} p_i e^{t\varphi_i(x_0)} h(\hat{M}_i x_0),$$

we have $\hat{M}_i x_0 \in \mathcal{Z}$ for each $1 \leq i \leq k$. By induction, if $x_0 \in \mathcal{Z}$,

$$\{\hat{M}_{i_n} \cdots \hat{M}_{i_1} x_0 : n \geq 1, 1 \leq i_1, \ldots, i_n \leq k\} \subset \mathcal{Z}.$$

Note since $h \in \Lambda''$, we have for each $2 \leq i \leq d$,

$$\|\partial h/\partial z_i\| \leq D''' := D'/D'' < \infty.$$

And by Proposition 5.1.8 (d), we can see $\|\hat{S}_n x - \hat{S}_n y\|$ can be arbitrarily small by choosing n sufficiently large, where S_n is an n-length product of matrices taking values in $\{M_i\}_{1 \leq \leq k}$. Therefore $|h(\hat{S}_n x) - h(\hat{S}_n y)| \leq D'''\|\hat{S}_n x - \hat{S}_n y\|$.

Again by Proposition 5.1.8 (d), we can choose $N_1 > 0$ and $0 < \delta < 1, \kappa > 0$ such that when $n > N_1$,

$$\mathbb{E}[\|\hat{S}_n(\omega)x - \hat{S}_n(\omega)y\|] \leq \kappa\delta^n d_H(x, y).$$

Define $\rho := \sup_{1\leq i\leq k, x\in V} |\varphi_i(x)|$. For any $x \in V$, $\epsilon > 0$, choose $y \in \mathcal{Z}$ and $|t| < -\log\delta/\rho$, then there exists $N = \max\{N_1, \log(\epsilon/\kappa D''')/(\log\delta + |t|\rho)\}$ such that when $n > N$ we have,

$$\begin{aligned}
|h(x)| &\leq \sum_{i_1,\ldots,i_n} p_{i_1}\cdots p_{i_n} e^{|t|n\rho} |h(\hat{M}_{i_n}\cdots\hat{M}_{i_1}x)| \\
&= \sum_{i_1,\ldots,i_n} p_{i_1}\cdots p_{i_n} e^{|t|n\rho} |h(\hat{M}_{i_n}\cdots\hat{M}_{i_1}x) - h(\hat{M}_{i_n}\cdots\hat{M}_{i_1}y)| \\
&\leq D''' e^{|t|n\rho}\mathbb{E}\left[\|\hat{M}_{i_n}\cdots\hat{M}_{i_1}x - \hat{M}_{i_n}\cdots\hat{M}_{i_1}y\|\right] \\
&\leq \kappa D''' e^{|t|n\rho}\delta^n d_H(x, y) < \epsilon d_H(x, y).
\end{aligned}$$

Since $\epsilon > 0, x \in V$ are arbitrary, h has to be a constant zero function on $\overline{V}$, and therefore can be uniquely analytically extended to the constant zero function on U, which contradicts $\beta = \|\mathcal{L}h\| > 0$.

Part 3. Now we turn to the uniqueness of the eigenfunction h (up to complex scalars) corresponding to the eigenvalue β. Assume $g(z) \in \mathcal{A}_\infty(U)$ is another eigenfunction corresponding to β, then $(\operatorname{Re} g)(x), (\operatorname{Im} g)(x) \in \mathcal{C}^\omega(V)$ are both eigenfunctions of $\mathcal{L}$ acting on $\mathcal{C}^\omega(V)$ associated with the eigenvalue β. Since $h(x)$ is strictly positive on $\overline{V}$, we know $(\operatorname{Re} g)(x)/h(x) \in \mathcal{C}^\omega(V)$. Let

$$u = \inf_{x\in V} \frac{(\operatorname{Re} g)(x)}{h(x)}.$$

Assume this infimum is attained at $x_0 \in \overline{V}$, then we have $(\operatorname{Re} g)(x) - uh(x) \geq 0$ with $(\operatorname{Re} g)(x_0) - uh(x_0) = 0$. Then by the same proof as in part 2, we know $(\operatorname{Re} g)(x) - uh(x)$ is identically zero on V and similarly we can prove that $(\operatorname{Im} g)(x) = vh(x)$ on V for some constant $v \in \mathbb{R}$. So $g(x) = (u + \mathrm{i}v)h(x)$ on V and $g(z) = (u + \mathrm{i}v)h(z)$ on U as the extension is unique.

Part 4. Note $h(z)$, as the top eigenfunction of $\mathcal{L}$ given by (5.19), depends on t. To clarify this dependence, in this part, for each given t (we suppose that $|t|$ is sufficiently small, i.e., there exists $\gamma > 0$ and $|t| < \gamma$), we will denote the top eigenfunction by $h_t(z)$; and similarly the

parametrised transfer operator $\mathcal{L}$ by $\mathcal{L}_t$, the set Λ'' by Λ''_t and top eigenvalue β by β_t. We will prove that β_t and $h_t(z)$ are continuous at $t = 0$.

Notice first by (5.24),

$$\beta_t \geq \sum_{i=1}^{k} p_i e^{-|t||\varphi_i|_\infty} \geq \sum_{i=1}^{k} p_i e^{-\gamma|\varphi_i|_\infty} =: B, \tag{5.26}$$

and $B > 0$ does not depend on t.

Note as $h_t(z) \in \Lambda''_t$ for each $|t| < \gamma$, we have $\|h_t\| \leq 1$, i.e., $\{h_t : |t| < \gamma\}$ is uniformly bounded. By Lemma 5.2.2, for the compact set K satisfying $\bigcup_{i=1}^{k} \hat{M}_i U \subset K \subset U$, we have that there exists $C'_K > 0$ (not depending on t) such that whenever $z, z' \in K$ are sufficiently close,

$$|h_t(z) - h_t(z')| \leq C'_K \|z - z'\|.$$

Therefore for $z, z' \in U$ sufficiently close,

$$\begin{aligned}
|h_t(z) - h_t(z')| &= \frac{1}{\beta_t} \sum_{i=1}^{k} p_i e^{t\varphi_i(z)} |h_t(\hat{M}_i z) - h(\hat{M}_i z')| \\
&\leq \frac{1}{\beta_t} \sum_{i=1}^{k} p_i \|e^{t\varphi_i}\| \cdot C'_K \|\hat{M}_i z - \hat{M}_i z'\| \\
&\leq \frac{C'_K}{B} \left(\sup_{1 \leq i \leq k} e^{\gamma \|\varphi_i\|} \right) \left(\sup_{1 \leq i \leq k} D_{M_i} \right) \|z - z'\|.
\end{aligned}$$

This shows that $\{h_t : |t| < \gamma\}$ is equicontinuous. Thus by the Arzelà-Ascoli Theorem, it is relatively compact.

Now let $\{t_m\}$ be a sequence in $\{|t| < \gamma\}$ converging to 0, then the sequence $\{h_{t_m}\}_m$ has a convergent subsequence $\{h_{t_{m_\alpha}}\}_\alpha$ such that $h_{t_{m_\alpha}}$ uniformly converges to some function $\tilde{h} \in \mathcal{A}_\infty(U)$ as $\alpha \to \infty$.

By (5.25) and (5.26), we have $\beta_t = \|\mathcal{L}_t h_t\| \geq B$ for each $|t| < \gamma$. Then letting $t = t_{m_\alpha}$ and $\alpha \to \infty$ shows that

$$\tilde{\beta} := \lim_{\alpha \to \infty} \beta_{t_{m_\alpha}} = \|\mathcal{L}_0 \tilde{h}\| \geq B > 0.$$

Therefore $\tilde{h}$ is not identically zero.

Moreover, letting $\alpha \to \infty$ in

$$\beta_{t_{m_\alpha}} h_{t_{m_\alpha}}(z) = \sum_{i=1}^{k} p_i e^{t_{m_\alpha}\varphi_i(z)} h_{t_{m_\alpha}}(\hat{M}_i z)$$

gives that

$$\tilde{\beta}\tilde{h}(z) = \sum_{i=1}^{k} p_i \tilde{h}(\hat{M}_i z). \tag{5.27}$$

On the one hand, (5.27) shows that $\tilde{h}$ is an eigenfunction of $\mathcal{L}_0$, thus by Theorem 5.2.4, we have $\tilde{\beta} \leq 1$; on the other hand, letting $t = t_{m_\alpha} \to 0$ in (5.25), we have $\tilde{\beta} \geq 1$. Thus $\tilde{\beta} = 1$ and by the uniqueness of the top eigenfunction for $\mathcal{L}_0$, we have $\tilde{h} = h_0$. This proves the continuity of β_t and h_t at $t = 0$.

Part 5. Note when $t = 0$, $h_0(z)$ equals the constant function $\mathbf{1}$ on U by Theorem 5.2.4. Thus by the continuity proved in part 4, we have, for any $z \in \overline{U}$

$$\lim_{|t|\to 0} h_t(z) = \mathbf{1}(z).$$

As $\overline{U}$ is compact, we furthermore have, for any $\epsilon > 0$, there exists $\gamma' > 0$ (not depending on $z \in U$) such that when $|t| < \gamma'$,

$$|h_t(z) - \mathbf{1}(z)| < \epsilon,\ z \in \overline{U}.$$

This shows that when $|t|$ is sufficiently small, $h_t(z)$ takes values in $\{\zeta \in \mathbb{C} : |\zeta - 1| < 1\}$, thus $\log h_t(z)$ is analytic on U.

Now by the analyticity of $\log h_t$ on U and the positivity of β_t, we can define

$$\phi_i(t, z) = t\varphi_i(z) + \log h_t(\hat{M}_i z) - \log h_t(z) - \log \beta_t \in \mathcal{A}_\infty(U),\ 1 \leq i \leq k.$$

Then we have

$$\sum_{i=1}^{k} p_i e^{\phi_i(t,z)} = \frac{1}{\beta_t} \sum_{i=1}^{k} p_i e^{t\varphi_i(z)} \frac{h_t(\hat{M}_i z)}{h_t(z)} = 1.$$

Moreover, since $h_t(z)$ is strictly positive when restricted on $\overline{V}$ by part 2, the restriction of $\phi_i(t, z)$ on $\overline{V}$ is also well-defined and lies in $\mathcal{C}^\omega(V)$.

Now, consider the weighted transfer operator $\tilde{\mathcal{L}}$ with the normalised weights $(e^{\phi_i})_{1\leq i\leq k}$. We have (here we omit t again for convenience)

$$\tilde{\mathcal{L}} = \beta^{-1} M_h^{-1} \mathcal{L} M_h,$$

where M_h is the multiplication operator by h. Therefore, the spectrum of $\tilde{\mathcal{L}}$ is the spectrum of $\mathcal{L}$ scaled by β^{-1}. We can use the same proof as for Theorem 5.2.4 to prove the rest of the statements of this theorem for the normalised transfer operator $\tilde{\mathcal{L}}$. Therefore they are also true for $\mathcal{L}$.

□

Now we are ready to study a special type of parametrised transfer operators by choosing

$$\varphi_i(z) = \log \|M_i z\|,$$

where $\|\cdot\|$ here is chosen to be the extension of the ℓ_2-norm on $\mathbb{R}^d$ to $\mathbb{C}^d$ by replacing real variables to complex ones. More specifically, for $z = (z_2, \ldots, z_d) \in U$ and a matrix $A = (a_{ij})_{1\leq i,j\leq d}$,

$$\log \|Az\| := \frac{1}{2} \log \left[\sum_{i=1}^{d} \left(a_{i1} + \sum_{j=2}^{d} a_{ij} z_j \right)^2 \right]. \tag{5.28}$$

It is not hard to see that since $\{M_i\}_{1\leq i\leq k}$ are all positive matrices and $z \in U$ (where U is chosen according to Corollary 5.1.4), $\|M_i z\|$ is complex analytic and does not take values in the negative real line or zero, thus $\varphi_i(z) \in \mathcal{A}_\infty(U)$.

In order to apply Theorem 5.3.2, we need to verify the two conditions. When $x \in V$, we can see $\|M_i x\|$ is never 0, thus condition (i) is satisfied. For condition (ii), if we write $A = (a_{ij})$, we have for $2 \leq l \leq d$,

$$\frac{\partial}{\partial z_l} \log \|Az\| = \frac{1}{\|Az\|^2} \sum_{i=1}^{d} \left(a_{i1} + \sum_{j=2}^{d} a_{ij} z_j \right) a_{il}.$$

Since each z_j is bounded by the choice of U, and $\left|\|Az\|^2\right| \geq \left(\sum_{i=1}^{d} a_{i1}^2\right)^{1/2} > 0$ by $\operatorname{Re} z_j \geq 0$, we can see condition (ii) is also satisfied.

Therefore, according to Theorem 5.3.2, for each $t \in \mathbb{R}$, the parametrised transfer operators, denoted by

$$\mathcal{L}_t : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$$

$$w(z) \mapsto \sum_{i=1}^{k} p_i e^{t \log \|M_i z\|} w(\hat{M}_i z), \tag{5.29}$$

has a simple maximal eigenvalue $\beta(t) > 0$. And by Theorem 5.2.4, we can easily see that $\beta(0) = 1$.

Since $\beta(t)$ is simple and the rest of the spectrum lies in the open ball $\{z : |z| < \beta(t)\}$, we can decompose $\mathcal{L}_t$ as

$$\mathcal{L}_t = \beta(t)Q(t) + R(t).$$

Here $Q(t)$ is the rank-one projection given by $w \mapsto h\nu(w)$, where h is the simple maximal eigenfunction and ν is the positive linear functional on $\mathcal{A}_\infty(U)$ given by Theorem 5.3.2, and $R(t) := \mathcal{L}_t - \beta(t)Q(t)$ has spectral radius strictly smaller than $\beta(t)$.

Moreover, for any $w \in \mathcal{A}_\infty(U)$, by Theorem 5.3.2,

$$Q(t)R(t)w = Q(t)\mathcal{L}_t w - \beta(t)Q(t)w = h\nu(\mathcal{L}_t w) - \beta(t)h\nu(w) = 0,$$

$$R(t)Q(t)w = \mathcal{L}_t Q(t)w - \beta(t)Q(t)w = \mathcal{L}_t(\nu(w)h) - \beta(t)h\nu(w) = 0.$$

Therefore, we have $Q(t)R(t) = R(t)Q(t) = 0$. This shows that

$$\mathcal{L}_t^n = \beta^n(t)Q(t) + R^n(t). \tag{5.30}$$

On the other hand, we would have

$$(\mathcal{L}_t^n w)(z) = \sum_{|\underline{i}|=n} p_{\underline{i}} e^{t \log \|S_{\underline{i}} z\|} w(\hat{S}_{\underline{i}} z), \tag{5.31}$$

where for each sequence $\underline{i}$ of elements in $\{1, \ldots, k\}$, $|\underline{i}|$ denotes the length of the sequence. If $\underline{i} = (i_1, \ldots, i_n)$, $p_{\underline{i}}$ denotes the product $p_{i_1} \cdots p_{i_n}$, $S_{\underline{i}}$ denotes $M_{i_n} \cdots M_{i_1}$ and $\hat{S}_{\underline{i}}$ denotes $\hat{M}_{i_n} \cdots \hat{M}_{i_1}$.

Now we are ready to state a theorem that connects the top eigenvalue of parametrised transfer operators and the corresponding Lyapunov exponent.

Theorem 5.3.3. *Let $\mathcal{L}_t$ be a family of parametrised transfer operators defined as in eq.* (5.29)*, then when $|t|$ is sufficiently small, there exists a decomposition of $\mathcal{L}_t$ with the form of*

$$\mathcal{L}_t = \beta(t)Q(t) + R(t),$$

where $\beta(t)$ is the simple maximal eigenvalue of $\mathcal{L}_t$, $Q(t)$ is a rank-one projection, $R(t)$ has spectral radius strictly smaller than $\beta(t)$, and $Q(t)R(t) = R(t)Q(t) = 0$. Then there exists

$\epsilon > 0$, *such that* $\beta(t)$ *is analytic on* $(-\epsilon, \epsilon)$ *and* $\beta'(0)$ *equals the Lyapunov exponent associated with the matrix set* $\{M_i\}_{1\le i\le k}$ *and the probability vector* $(p_i)_{1\le i\le k}$.

Proof. The existence of the decomposition follows from Theorem 5.3.2 and the discussion above. The fact that $\beta(t)$ is analytic follows from the perturbation theory of operators. (See e.g. [43] Section VII.6.)

Now write $\mathbf{1}(x)$ as the constant function on V. Then by (5.30) and (5.31), we have

$$\beta^n(t)Q(t)\mathbf{1} + R^n(t)\mathbf{1} = \sum_{|\underline{i}|=n} p_{\underline{i}} e^{t\log\|S_{\underline{i}}x\|} = \mathbb{E}\left[e^{t\log\|S_n(\omega)x\|}\right]. \tag{5.32}$$

Differentiating both sides of (5.32) with respect to t gives

$$\mathbb{E}[\log\|S_nx\| \cdot e^{t\log\|S_nx\|}] = n\beta^{n-1}(t)\beta'(t)Q(t)\mathbf{1} + \beta^n(t)\frac{\mathrm{d}}{\mathrm{d}t}(Q(t)\mathbf{1}) + \frac{\mathrm{d}}{\mathrm{d}t}(R^n(t)\mathbf{1}).$$

Letting $t = 0$ and using $Q(0)\mathbf{1} = \mathbf{1}, \beta(0) = 1$ gives

$$\frac{1}{n}\mathbb{E}[\log\|S_nx\|] = \beta'(0) + \frac{1}{n}\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(Q(t)\mathbf{1}) + \frac{1}{n}\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}(R^n(t)\mathbf{1}). \tag{5.33}$$

By perturbation theory (See e.g. [4] Chapter V. Lemma 3.2), both

$$\frac{\mathrm{d}}{\mathrm{d}t}(Q(t)\mathbf{1}) \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}t}(R^n(t)\mathbf{1})$$

are bounded when $|t|$ is sufficiently small. Letting $n \to \infty$ in (5.33) and applying the definition of the Lyapunov exponent, we get

$$\gamma = \lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\log\|S_nx\|] = \beta'(0).$$

□

Here Theorem 5.3.3 suggests an approach of computing the Lyapunov exponent by defining the corresponding parametrised transfer operator and differentiating its maximal top eigenvalue at 0. But how to compute the top eigenvalue for parametrised transfer operators becomes the new problem. In 1976, Ruelle [49] applied Grothendieck's theory on nuclear operators and Fredholm theory on determinant functions to dynamical systems, and proved that, analogous to matrices, we can define the determinant functions for transfer operators and study the eigenvalues of transfer operators by studying the zeros of the determinant function.

5.4 Nuclear Operators

In Section 2.2, we introduced the singular value decompositions (or SVD) for a real matrix. Namely, let A be a $d \times d$ real matrix, then it can be written as a product

$$A = UDV,$$

where U, V are orthogonal matrices, $D = \mathrm{diag}(a_1, a_2, \ldots, a_d)$ is a diagonal matrix. Here $a_1 \geq \cdots \geq a_d \geq 0$ are also called the singular values of A. In this section, we plan to discuss the possibility of decomposing an operator in a similar style.

Now we write U as d column vectors and V as d row vectors, i.e.,

$$U = \begin{pmatrix} u_1 & u_2 & \ldots & u_d \end{pmatrix}, \quad V = \begin{pmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_d^* \end{pmatrix}$$

where $*$ denotes the transpose of matrices. Then we have

$$A = \sum_{i=1}^{d} a_i u_i v_i^*.$$

Note $u_i v_i^*$ can be identified with the tensor product $u_i \otimes v_i$, we also write

$$A = \sum_{i=1}^{d} a_i u_i \otimes v_i,$$

where $\|u_i\|_2 = \|v_i\|_2 = 1$ for each $1 \leq i \leq d$.

For any $w \in \mathbb{R}^d$, we can write

$$Aw = \sum_{i=1}^{d} a_i (u_i \otimes v_i)(w) = \sum_{i=1}^{d} a_i \langle v_i, w \rangle u_i.$$

And

$$\mathrm{tr}(A) = \sum_{i=1}^{d} \langle u_i, A u_i \rangle = \sum_{i=1}^{d} a_i \langle v_i, u_i \rangle \leq \sum_{i=1}^{d} a_i. \tag{5.34}$$

Moreover, since $A^*A = V^* D U^* U D V = V^* D^2 V$, we have

$$\mathrm{tr}(A^*A) = \mathrm{tr}(D^2) = \sum_{i=1}^{d} a_i^2. \tag{5.35}$$

If we informally regard operators on Hilbert spaces as infinite matrices, from the results above, let $\mathcal{H}$ be a separable Hilbert space over $\mathbb{C}$, naturally we want to define the decomposition of an operator $\mathcal{L} : \mathcal{H} \to \mathcal{H}$ by

$$\mathcal{L} = \sum_{i=1}^{\infty} \rho_i u_i \otimes v_i, \tag{5.36}$$

in the sense that for any $w \in \mathcal{H}$,

$$\mathcal{L}w = \sum_{i=1}^{\infty} \rho_i \langle v_i, w \rangle u_i, \tag{5.37}$$

where a norm $\| \cdot \|$ is defined on $\mathcal{H}$ induced by its inner product, the series (5.36) and (5.37) converge in the respective norm topologies, $\rho_i \in \mathbb{C}$, $\{u_i\}$ and $\{v_i\}$ are two orthonormal bases in $\mathcal{H}$, $\|u_i\| = \|v_i\| = 1$ for $i \geq 1$.

We can easily show that, if the decomposition (5.36) exists, with the convention

$$\langle \lambda u, v \rangle = \bar{\lambda} \langle u, v \rangle = \langle u, \bar{\lambda} v \rangle,$$

the adjoint operator $\mathcal{L}^*$ can be given by

$$\mathcal{L}^* = \sum_{i=1}^{\infty} \bar{\rho_i} v_i \otimes u_i.$$

And $\mathcal{L}^* \mathcal{L} v_i = |\rho_i|^2 v_i$ for each $i \geq 1$. That is to say, $\{|\rho_i|^2 : i \geq 1\}$ is a subset of point spectrum of $\mathcal{L}^* \mathcal{L}$.

However, regarding operators as infinite matrices is wrong in general and such a decomposition might not exist for all operators. For example let $\mathcal{H} = L^2([0, 1])$ and $(\mathcal{M}_x f)(x) = x f(x)$. Since the point spectrum of $\mathcal{M}_{x^2}$ is empty, if the decomposition exists for $\mathcal{M}_x$ then $\mathcal{M}_x$ has to be 0, which is absurd. Therefore we give the following definitions for some subclasses of operators.

Definition 5.4.1. Let $\mathcal{L}$ be an operator action on a separable Hilbert space over $\mathbb{C}$. If $\mathcal{L}$ can be represented in the form

$$\mathcal{L} = \sum_{i=1}^{\infty} \rho_i u_i \otimes v_i, \tag{5.38}$$

where $\rho_i \in \mathbb{C}$, $\{u_i\}$ and $\{v_i\}$ are orthonormal basis of $\mathcal{H}$, then the elements of $\{|\rho_i| : i \geq 1\}$ are called **singular values** of $\mathcal{L}$.

If $\sum_{i=1}^{\infty} |\rho_i| < \infty$, then $\mathcal{L}$ is said to be a **trace-class** operator on $\mathcal{H}$.

If $\mathcal{L}^*\mathcal{L}$ is a trace-class operator, or equivalently $\sum_{i=1}^{\infty} |\rho_i|^2 < \infty$, then $\mathcal{L}$ is said to be a **Hilbert-Schmidt operator** on $\mathcal{H}$.

Remark 5.4.2. (1) The equation (5.38) is also called the Schmidt representation of an operator $\mathcal{L}$. In fact, all compact operators can be represented in this way. Thus the class of trace-class operators is a subclass of the compact operators. Moreover, similar to the case of matrices, the singular values $\{|\rho_i| : i \geq 1\}$ of a given compact operator $\mathcal{L}$ in its Schmidt representation do not depend on the choice of the orthonormal basis ([28] Section 4.2).

(2) When $\mathcal{L}$ is a trace-class operator, the trace $\mathrm{tr}(\mathcal{L})$ can also be given by

$$\mathrm{tr}(\mathcal{L}) = \sum_{i=1}^{\infty} \langle v_i, \mathcal{L}v_i \rangle. \tag{5.39}$$

The definition does not depend on the choice of the orthonormal basis.

(3) The condition $\sum_{i=1}^{\infty} \langle v_i, \mathcal{L}v_i \rangle < \infty$ cannot be used to replace $\sum_{i=1}^{\infty} |\rho_i| < \infty$ to define a trace-class operator. For example, let $\mathcal{T}$ be an operator acting on a separable Hilbert space $\mathcal{H}$ equipped with an orthonormal basis $\{v_i\}_{i=1}^{\infty}$ by

$$\mathcal{T}v_{2k} = \frac{1}{\sqrt{2k}} v_{2k}, \quad \mathcal{T}v_{2k+1} = -\frac{1}{\sqrt{2k+1}} v_{2k+1}.$$

Then

$$\sum_{i=1}^{\infty} \langle v_i, \mathcal{T}v_i \rangle = \sum_{i=1}^{\infty} (-1)^i \frac{1}{\sqrt{i}} < \infty.$$

However, since

$$\sum_{i=1}^{\infty} \langle v_i, \mathcal{T}^*\mathcal{T}v_i \rangle = \sum_{i=1}^{\infty} \frac{1}{i} = \infty,$$

$\mathcal{T}$ is not a trace class operator.

(4) We can observe that

$$\mathrm{tr}(\mathcal{L}) \leq \sum_{i=1}^{\infty} |\rho_i|, \tag{5.40}$$

which is analogous to (5.34). Moreover, for each trace class operator $\mathcal{L}$ with singular values $\{\rho_i(\mathcal{L})\}_{i=1}^{\infty}$, we can define a norm by

$$\|\mathcal{L}\|_1 := \sum_{i=1}^{\infty} |\rho_i(\mathcal{L})|.$$

Then the set $\mathcal{S}_1 := \{\mathcal{L} : \|\mathcal{L}\|_1 < \infty\}$ consisting of trace class operators form a Banach space (see e.g. [28] or [52]). That is to say, the definition of a trace given by (5.39) is indeed a generalisation for the trace of a finite rank operator or of a matrix, in the sense that, if $\{F_n\}$ is a sequence of finite rank operators converging to $\mathcal{L}$ in the norm of $\|\cdot\|_1$, then $\mathrm{tr}(F_n) \to \mathrm{tr}(\mathcal{L})$ as $n \to \infty$. Equation (5.40) also implies that $\mathrm{tr} : (\mathcal{S}_1, \|\cdot\|_1) \to \mathbb{R}$ is a bounded functional.

It is worth noticing that in this dissertation, the spaces we are interested in, such as $\mathcal{A}_\infty(U)$ or $\mathscr{H}(U)$ are not Hilbert spaces. Therefore we cannot apply the theory above directly to our problem. However, if we denote by $\mathcal{B}(\mathcal{H})$ the set of bounded linear operators on the Hilbert space $\mathcal{H}$, operators which can be written in the form of finite sums $\sum_{i=1}^N \rho_i u_i \otimes v_i$ belong to the set $\mathcal{H} \otimes \mathcal{H}^*$. And trace-class operators can be regarded as belonging to $\mathcal{H}\tilde{\otimes}\mathcal{H}^*$, the completion of $\mathcal{H} \otimes \mathcal{H}^*$ under the norm topology. Now let $\mathcal{B}(E, F)$ denote the set of operators from space E to F. Can we give a similar definition by studying the subspace $F\tilde{\otimes}E^*$ of $\mathcal{B}(E, F)$?

In [22], Grothendieck studied the possibility of generalising trace-class operators on Hilbert spaces to nuclear operators on topological vector spaces (and particularly, on Banach spaces). In the following we choose to state only the version for Banach spaces as this is what will be mostly used in later sections. Readers interested in the further generalisations to topological vector spaces can refer to [53] or [22].

Definition 5.4.3. Let B_1, B_2 be Banach spaces equipped with norms $\|\cdot\|_{B_1}$ and $\|\cdot\|_{B_2}$ respectively, and denote by B_1' the strong dual of B_1, i.e., the dual space of $(B_1, \|\cdot\|_{B_1})$. Then $\mathcal{L} : B_1 \to B_2$ is called a **nuclear operator** if it can be represented as

$$\mathcal{L} = \sum_{i=1}^{\infty} \rho_i u_i \otimes f_i, \tag{5.41}$$

where $u_i \in B_2, f_i \in B_1'$ with $\|u_i\|_{B_2} = \|f_i\|_{B_1'} = 1$, and $\sum_{i=1}^\infty |\rho_i| < \infty$, in the sense that for any $w \in B_1$, we have

$$\mathcal{L}w = \sum_{i=1}^{\infty} \rho_i f_i(w) u_i.$$

Moreover, $\mathcal{L}$ is said to be **nuclear of order p** if $\mathcal{L}$ admits a representation (5.41) such that $\sum_{i=1}^\infty |\rho_i|^p < \infty$, and **nuclear of order 0** if this holds for any $p > 0$.

Remark 5.4.4. (1) An alternate definition is, $\mathcal{L}$ is said to be nuclear from B_1 to B_2 if it admits

a representation

$$\mathcal{L} = \sum_{i=1}^{\infty} u_i \otimes f_i, \tag{5.42}$$

where $u_i \in B_2$, $f_i \in B_1'$, and

$$\sum_{i=1}^{\infty} \|u_i\|_{B_2} \|f_i\|_{B_1'} < \infty.$$

It is not hard to see that these two definitions are equivalent. The advantage of introducing this definition is that it suggests the definition of a norm $\|\cdot\|_1$ on nuclear operators given by

$$\|\mathcal{L}\|_1 = \inf \sum_{i=1}^{\infty} \|u_i\|_{B_2} \|f_i\|_{B_1'},$$

where the infimum is taken over all the representations (5.42) of $\mathcal{L}$. We can define

$$\mathcal{S}_1 := \{\mathcal{L} : \|\mathcal{L}\|_1 < \infty\}, \tag{5.43}$$

and show that $\mathcal{S}_1$ is an operator algebra. See e.g, [28]. Section V.1.

(2) It is obvious that for each $0 \leq p \leq 1$, nuclear operators on a separable Hilbert space $\mathcal{H}$ of order p form a linear space. This fact is also true for nuclear operators on a Banach space. See e.g. [47] 1.7.

(3) The concept of the trace (resp. of the determinant) on finite rank operators can be extended by defining it to be a continuous extension of the usual trace (resp. determinant) from the algebra of finite rank operators to $\mathcal{S}_1$ (see. e.g. [28] Chapter II). As a consequence, the following trace-determinant formula can be extended to $\mathcal{S}_1$ as well.

$$\det(I - z\mathcal{L}) = \exp\left(-\sum_{n=1}^{\infty} \frac{\mathrm{tr}(\mathcal{L}^n)}{n} z^n\right). \tag{5.44}$$

(4) A locally convex space is said to have the **approximation property** if its identity map can be approximated, uniformly on precompact sets, by continuous linear maps of finite rank ([53] pp.108). For example, the space $\mathscr{H}(D)$ of analytic function on a domain D of $\mathbb{C}^d$ has the approximation property. See e.g. [53] III. 9.2.

(5) It is proved in [22] that if $B_1 = B_2 = B$ is a Banach space with the approximation property and $\mathcal{L}$ is a nuclear operator of order $2/3$ on B. Then

$$\mathrm{tr}(\mathcal{L}) = \sum_j \lambda_j(\mathcal{L}), \tag{5.45}$$

$$\det(I + \mathcal{L}) = \prod_j (1 + \lambda_j(\mathcal{L})), \tag{5.46}$$

where $\{\lambda_j(\mathcal{L})\}$ is the set of eigenvalues of $\mathcal{L}$ with multiplicities taken into account. (An English-language proof can be found in [28] Section V.3.)

(6) Let B_0, B_1, B_2, B_3 be Banach spaces, let $\mathcal{L} : B_1 \to B_2$ be a nuclear operator, and $\alpha : B_0 \to B_1, \beta : B_2 \to B_3$ be continuous linear mappings. Then $\mathcal{L}\alpha, \beta\mathcal{L}$ (and therefore $\beta\mathcal{L}\alpha$) are all nuclear. In fact, if $\mathcal{L}$ admits a representation of the form (5.41), then

$$\beta\mathcal{L} = \sum_{i=1}^{\infty} \rho_i(\beta u_i) \otimes f_i.$$

Moreover, it is obvious that if $\mathcal{L}$ is nuclear of order p then $\beta\mathcal{L}$ is also nuclear of order p ($p \geq 0$). Similarly, we can have

$$\mathcal{L}\alpha = \sum_{i=1}^{\infty} \rho_i u_i \otimes (\alpha^* f_i),$$

where $\alpha^* : B_1' \to B_0'$ is the adjoint operator of α. This result can also be generalised to locally convex spaces. (See e.g. [53] III.7. or [22] II. pp.9)

In the end, we give the definition of a nuclear space. Here we choose not to get into too much detail on the definition of a nuclear map on a locally convex space as it would be a digression. We give the necessary results for this dissertation in Proposition 5.4.6.

Definition 5.4.5. We say a locally convex space E is **nuclear** if every continuous linear map of E into any Banach space is nuclear.

Proposition 5.4.6. *(a) If E is nuclear, then every bounded linear map from E into a quasi-complete locally convex space (particularly a Banach space) is nuclear of order 0. (A topological vector space is said to be **quasi-complete** if every bounded, closed subset is complete.)*

(b) Let D be a domain of $\mathbb{C}^d$, then the set $\mathscr{H}(D)$ of complex analytic functions on D is a nuclear space.

Proof. The proof of (a) can be found in [22] II. Corollary 2. pp 61. And the proof of (b) can be found in [53] III. 8.4 or [22] II. pp.56. □

5.5 Nuclearity of Transfer Operators

In this section, we prove that weighted transfer operators are nuclear of order 0 on $\mathcal{A}_\infty(U)$ or $\mathscr{H}(U)$ (where U is chosen according to Lemma 5.1.3), and give a formula for the traces of weighted transfer operators.

By Remark 5.4.4 (2), without loss of generality we only need to consider operators (on $\mathcal{A}_\infty(U)$ or $\mathscr{H}(U)$) of the form

$$(\mathcal{L}w)(z) = \varphi(z)w(\psi(z)), \tag{5.47}$$

where ψ is a holomorphic function mapping U strictly inside U. We call them **simple transfer operators**.

We first prove that simple transfer operators are all nuclear of order zero. This result is given by Ruelle [49].

Theorem 5.5.1. *(Ruelle) Let U be a bounded open subset of $\mathbb{C}^d$ and ψ a holomorphic function mapping U strictly inside U. Let $\mathcal{L}$ be an operator on $\mathcal{A}_\infty(U)$ or $\mathscr{H}(U)$ given by*

$$(\mathcal{L}f)(z) = \varphi(z)(f \circ \psi(z)).$$

Then $\mathcal{L}$ is nuclear of order zero.

Proof. By Proposition 5.4.6, $\mathscr{H}(U)$ is a nuclear space. Since $\mathscr{H}(U)$ is also quasi-complete, we see that $\mathcal{L} : \mathscr{H}(U) \to \mathscr{H}(U)$, as a bounded linear map from a nuclear space into a quasi-complete locally convex space, has to be nuclear of order zero.

On the other hand, since ψ maps U strictly inside U, $\mathcal{L} : \mathscr{H}(U) \to \mathscr{H}(U)$ maps $\mathscr{H}(U)$ to $\mathcal{A}_\infty(U)$ (with a slight abuse of notation, we still denote it by $\mathcal{L}$). As $\mathcal{A}_\infty(U)$ is complete, by Proposition 5.4.6, $\mathcal{L} : \mathscr{H}(U) \to \mathcal{A}_\infty(U)$ is nuclear of order zero. Write $\iota : \mathcal{A}_\infty(U) \to \mathscr{H}(U)$ as the identity embedding operator. Then by Remark 5.4.4 (6), $\mathcal{L} \circ \iota : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$ is also nuclear of order zero. (We will also denote this by $\mathcal{L}$ in following for convenience.) □

The property that ψ maps U strictly inside U can be considered a sort of contraction property. This is essential, as it guarantees the existence and uniqueness of a fixed point of ψ.

Theorem 5.5.2. *Let U be a bounded domain of $\mathbb{C}^d$, and ψ be a holomorphic function mapping U strictly inside U. Then ψ has exactly one fixed point t in U. Moreover, $|\det D\psi(t)| < 1$, and all eigenvalues of $D\psi(t)$ have modulus strictly less than 1.*

Proof. If all the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_d$ of $D\psi(t)$ have modulus strictly less than 1, then $|\det D\psi(t)| = |\lambda_1 \lambda_2 \cdots \lambda_d| < 1$. The rest of the proof can be found in [49] Lemma 1. □

Remark 5.5.3. This result is obvious for positive matrices. The contraction property of positive matrices can be visually seen as suggested in Section 3.2, or follows from (5.8). And the existence and uniqueness of fixed points of positive matrices are guaranteed by the Perron-Frobenius theorem for positive matrices.

By Remark 5.4.4 (5), in order to understand the trace and the determinant function of $\mathcal{L} : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$, we only need to find the eigenvalues of $\mathcal{L}$. This problem will be solved by Theorem 5.5.4. Our proof is constructed based on the one due to Mayer ([39] Lemma 7.10), which gives all the eigenvalues of simple transfer operators in $d = 1$.

Theorem 5.5.4. *Let $\mathcal{L}$ be a simple transfer operator on $\mathcal{A}_\infty(U)$ given by* (5.47)*. Assume ψ maps the bounded domain $U \subset \mathbb{C}^d$ strictly inside U. Then the eigenvalues of $\mathcal{L}$ are exactly*

$$\varphi(t)\lambda_1^{n_1} \cdots \lambda_d^{n_d}, \quad n_1, \ldots, n_d \geq 0$$

(multiplicities are taken into account if some of them coincide), where t is the fixed point of ψ (whose existence and uniqueness are given by Theorem 5.5.2) and $\{\lambda_i\}_{1 \leq i \leq d}$ are eigenvalues (with multiplicities) of the linear map $(D\psi)(t)$.

Proof. Consider the functional equation

$$\varphi(z) f(\psi(z)) = \lambda f(z), \tag{5.48}$$

where $f \in \mathcal{A}_\infty(U)$. Let $z = t$ in (5.48), then we have $\varphi(t)f(t) = \lambda f(t)$. Thus either $\lambda = \varphi(t)$ or $f(t) = 0$.

If $f(t) = 0$, differentiating both sides of (5.48) gives

$$\mathrm{d}\varphi(z) \cdot f(\psi(z)) + \varphi(z) \cdot (D\psi)(z)\, \mathrm{d}f(\psi(z)) = \lambda\, \mathrm{d}f(z), \tag{5.49}$$

where $\mathrm{d}f, \mathrm{d}\varphi$ are the total derivatives of f and φ respectively. Letting $z = t$ in (5.49) shows that $\varphi(t) \cdot (D\psi)(t)\, \mathrm{d}f(t) = \lambda\, \mathrm{d}f(t)$. Then either $\mathrm{d}f(t) = 0$ or λ is an eigenvalue of $\varphi(t) \cdot (D\psi)(t)$.

If $f(t) = 0$ and $\mathrm{d}f(t) = 0$, we can again differentiate (5.49) and obtain,

$$\begin{aligned} \lambda\, \mathrm{d}^2 f = (\mathrm{d}^2\varphi)\left(I \otimes (f \circ \psi)\right) + 2(\mathrm{d}\varphi)(\mathrm{d}f \circ \psi)(D\psi) \\ + \varphi(\mathrm{d}^2 f \circ \psi)(D\psi \otimes D\psi) + \varphi(\mathrm{d}f \circ \psi)D^2\psi, \end{aligned}$$

where we omit z for convenience. For an explanation of how tensor products are used to compute high derivatives of composite maps, see e.g. [13].

Letting $z = t$ gives

$$\lambda \, \mathrm{d}^2 f(t) = \varphi(t) \left(\mathrm{d}^2 f(t) \right) \left(D\psi(t) \otimes D\psi(t) \right).$$

Since the eigenvalues of $D\psi(t) \otimes D\psi(t)$ are $\{\lambda_i \lambda_j : 1 \leq i \leq j \leq d\}$, therefore either λ is of the form $\varphi(t)\lambda_i\lambda_j,\ 1 \leq i, j \leq d$ or $\mathrm{d}^2 f(t) = 0$.

Similarly, if $f(t) = \mathrm{d}f(t) = \mathrm{d}^2 f(t) = 0$, we can deduce that either $\mathrm{d}^3 f(t) = 0$ or λ is of the form $\varphi(t)\lambda_i\lambda_j\lambda_k,\ 1 \leq i, j, k \leq d$. Inductively, we can see the eigenvalues of $\mathcal{L}$ form a subset of $\{\varphi(t)\lambda_1^{n_1} \cdots \lambda_d^{n_d} : n_1, \ldots, n_d \geq 0\}$.

Moreover, we can see if $\varphi(t) = 0$ and $\lambda \neq 0$, then letting $z = t$ in (5.48), (5.49) inductively, we have $f(t) = \mathrm{d}f(t) = \cdots = \mathrm{d}^n f(t) = 0$ for each $n \geq 1$, hence, $f(t) = 0$. That is to say, if $\varphi(t) = 0$, $\mathcal{L}$ has no non-zero eigenvalues, i.e., the proof is complete automatically. In the rest of the proof, we will assume $\varphi(t) \neq 0$.

Now we show that each $\varphi(t)\lambda_1^{n_1} \cdots \lambda_d^{n_d},\ n_1, \ldots, n_d \geq 0$ is indeed an eigenvalue.

Take $g \in \mathcal{A}_\infty(U)$ such that $g(t) \neq 0$ and assume $h \in \mathcal{A}_\infty(U)$ is such a function satisfying

$$(\mathcal{L}h)(z) - \varphi(t)h(z) = g(z). \tag{5.50}$$

If $\varphi(t) \neq 0$ is not an eigenvalue of $\mathcal{L}$, then $\mathcal{L} - \varphi(t)I$ has a bounded inverse and $h(z) = (\mathcal{L} - \varphi(t)I)^{-1} g(z)$ would be a solution of (5.50). However, letting $z = t$ in (5.50) yields a contradiction, since the left-hand side is 0 while $g(t) \neq 0$. Therefore $\varphi(t)$ is an eigenvalue of $\mathcal{L}$.

Take $g_1, \ldots, g_d \in \mathcal{A}_\infty(U)$ such that

$$g_i(t) = 0, \text{ for } 1 \leq i \leq d, \quad \text{and} \quad \frac{\partial g_i}{\partial z_j}(t) = \delta_{ij},\ 1 \leq i, j \leq d. \tag{5.51}$$

If $\varphi(t)\lambda_l$ is not a eigenvalue of $\mathcal{L}$, then similarly there exists $f_1, \ldots, f_d \in \mathcal{A}_\infty(U)$ such that for each $1 \leq i \leq d$, we have

$$(\mathcal{L}f_i)(z) - \varphi(t)\lambda_l f_i(z) = g_i(z). \tag{5.52}$$

Letting $z = t$ in (5.52) shows that $f_i(t) = 0$ for $1 \leq i \leq d$, since $\varphi(t) \neq 0$ and $|\lambda_l| < 1$ by Theorem 5.5.2.

Differentiating (5.52) and letting $z = t$ shows that

$$\varphi(t) \, \mathrm{d}f_i(t)(D\psi(t) - \lambda_l I) = \mathrm{d}g_i(t).$$

By (5.51), we obtain that

$$\varphi(t) \left(\mathrm{d}f_1, \ldots, \mathrm{d}f_d \right) (D\psi(t) - \lambda_l I) = I,$$

which contradicts the fact that $D\psi(t) - \lambda_l I$ is not invertible. Thus $\varphi(t)\lambda_l,\ 1 \leq l \leq d$ are eigenvalues of $\mathcal{L}$.

In general, we can follow the same procedure, by taking functions

$$g_{i_1\dots i_k} \in \mathcal{A}_\infty(U),\ 1 \leq i_1, \dots, i_k \leq d$$

satisfying

$$\mathrm{d}^p g_{i_1\dots i_k}(t) = 0, \qquad 0 \leq p \leq k-1, 1 \leq i_1, \dots, i_k \leq d, \tag{5.53}$$

$$\frac{\partial^k g_{i_1\cdots i_k}}{\partial z_{j_1} \cdots \partial z_{j_k}} = \delta_{i_1 j_i} \otimes \cdots \otimes \delta_{i_k j_k}, \qquad 1 \leq i_1, j_1, \dots, i_k, j_k \leq d. \tag{5.54}$$

If $\varphi(t)\lambda_{l_1} \cdots \lambda_{l_k}$ is not an eigenvalue of $\mathcal{L}$, assume each $f_{i_1\dots i_k} \in \mathcal{A}_\infty(U)$ is a solution of

$$(\mathcal{L} - \varphi(t)\lambda_{l_1} \cdots \lambda_{l_k} I) f_{i_1\cdots i_k}(z) = g_{i_1\cdots i_k}(z). \tag{5.55}$$

Differentiating (5.55) k times and letting $z = t$ yields a contradiction.

We conclude by proving that the multiplicities of these eigenvalues are correct.

Assume $\omega_{l_1\cdots l_k}(z) \in \mathcal{A}_\infty(U)$ is the eigenfunction corresponding to the eigenvalue $\varphi(t)\lambda_{l_1} \cdots \lambda_{l_k}$, i.e.,

$$(\mathcal{L}\omega_{l_1\cdots l_k})(z) = \varphi(t)\lambda_{l_1} \cdots \lambda_{l_k} \omega_{l_1\cdots l_k}(z).$$

By the first part of the proof, we know that $\omega_{l_1\cdots l_k}$ satisfies (5.53) and (5.54). Therefore by the second part of the proof, there does not exist a function $f \in \mathcal{A}_\infty(U)$ such that

$$(\mathcal{L} - \varphi(t)\lambda_{l_1} \cdots \lambda_{l_k}) f = \omega_{l_1\cdots l_k}.$$

Thus for each eigenvalue $\varphi(t)\lambda_{l_1} \cdots \lambda_{l_k}$, one multiplicity should be taken into account .

If, furthermore,

$$\varphi(t)\lambda_1^{m_1} \cdots \lambda_d^{m_d} = \varphi(t)\lambda_1^{n_1} \cdots \lambda_d^{n_d}$$

for two different integer sequences (m_i) and (n_i), we discuss the following two cases. When $\sum m_i \neq \sum n_i$, by the first part of the proof, we know the corresponding eigenfunctions have different differentiation properties. Therefore multiplicities should be taken into account. When $\sum m_i = \sum n_i =: N$, they correspond to the multiple eigenvalues of $\otimes^N D\psi(t)$, for which multiplicities should also be taken into account. □

Corollary 5.5.5. *Let $\mathcal{L}$ be an operator on $\mathcal{A}_\infty(U)$ with the form of* (5.47)*. Then the trace of $\mathcal{L}$ can be given by*

$$\mathrm{tr}(\mathcal{L}) = \frac{\varphi(t)}{\det\big(I - D\psi(t)\big)},$$

where t is the unique fixed point of ψ.

Proof. By Theorem 5.5.1 and (5.45), the trace of $\mathcal{L}$ can be given by summing over all the eigenvalues of $\mathcal{L}$. By Theorem 5.5.4, we have

$$\operatorname{tr}(\mathcal{L}) = \sum_{n_1,\dots,n_{d-1}\geq 0} \varphi(t)\lambda_1^{n_1}\cdots\lambda_{d-1}^{n_{d-1}} = \frac{\varphi(t)}{(1-\lambda_1)\cdots(1-\lambda_{d-1})} = \frac{\varphi(t)}{\det\big(1-D\psi(t)\big)}.$$

□

5.6 Fredholm Theory

In this section, we introduce another result in which we treat operators as infinite matrices.

In the usual matrix setting, let $K = (K_{ij})$ be a $d \times d$ matrix, then we know that K acts on a d-dimensional vector $w \in \mathbb{R}^d$ by

$$(Kw)_i = \sum_{j=1}^{d} K_{ij} w_j,$$

and the trace of K can be given by $\operatorname{tr}(K) = \sum_{j=1}^{d} K_{jj}$.

Now if an operator $\mathcal{L}$ acts on some function space $\mathcal{G}([a,b])$ on $[a,b]$, in the sense that there exists a function $K : [a,b] \times [a,b] \to \mathbb{R}$ such that for any $w \in \mathcal{G}([a,b])$

$$(\mathcal{L}w)(x) = \int_a^b K(x,y)w(y)\,\mathrm{d}y, \tag{5.56}$$

we informally regard $\big(w(x)\big)_{x\in[a,b]}$ as a vector indexed by x and $\big(K(x,y)\big)_{x,y\in[a,b]}$ as an infinite matrix. Then we want to know whether we can analogously compute the trace by "summing" over the "diagonal entries" as

$$\operatorname{tr}(\mathcal{L}) = \int_a^b K(x,x)\,\mathrm{d}x. \tag{5.57}$$

Operators in the form of (5.56) are usually called **integral operators** on $\mathcal{G}([a,b])$, and $K(s,t)$ on $[a,b]\times[a,b]$ is called a **kernel function**.

The discussion of this problem was firstly brought up by Fredholm [14], and he confirmed this statement for the case of continuous kernels by giving the following Theorem 5.6.1. (See e.g. [28] Chapter 6, or [51] for an English-language proof.)

Theorem 5.6.1. *(Fredholm) Let $K(x,y) : [a,b] \times [a,b] \to \mathbb{R}$ be a continuous function. For $m \geq 1$, we can define*

$$a_m = \frac{1}{m!} \int_{[a,b]^m} \det \begin{pmatrix} K(t_1,t_1) & \cdots & K(t_1,t_m) \\ \vdots & \ddots & \vdots \\ K(t_m,t_1) & \cdots & K(t_m,t_m) \end{pmatrix} \mathrm{d}t_1 \cdots \mathrm{d}t_m.$$

Then the power series

$$1 + \sum_{m=1}^{\infty} (-1)^m a_m z^m$$

*converges uniformly to an entire function. And this entire function is exactly the determinant function $\det(I - z\mathcal{L})$ of the integral operator $\mathcal{L}$ on $\mathcal{C}[a,b]$ with the kernel function K. This determinant function $\det(I - z\mathcal{L})$ will be called the **Fredholm determinant** of $\mathcal{L}$.*

Remark 5.6.2. (1) On the one hand, by Theorem 5.6.1,

$$a_1 = \int_a^b K(t,t)\,\mathrm{d}t.$$

On the other hand, as an integral operator $\mathcal{L}$ with a continuous kernel K is nuclear (see e.g. [28] V.2.), by expanding $\det(I - z\mathcal{L})$, we have $a_1 = \mathrm{tr}(\mathcal{L})$. This confirms (5.57)

(2) Now if we choose $\mathcal{G}([a,b]) = L^2[a,b]$ with the natural inner product

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,\mathrm{d}x,$$

and assume $\mathcal{L}$ is a nuclear operator on $L^2[a,b]$ with the form

$$\mathcal{L} = \sum_{i=1}^{\infty} \rho_i u_i \otimes v_i,$$

where $u_i, v_i \in L^2[a,b]$ and $\|u_i\|_2 = \|v_i\|_2 = 1$ such that for any $w \in L^2[a,b]$,

$$(\mathcal{L}w)(x) = \sum_{i=1}^{\infty} \rho_i \langle v_i, w \rangle u_i.$$

Now if we define

$$K(x,y) = \sum_{i=1}^{\infty} \rho_i u_i(x) v_i(y),$$

where the series converges with respect to $\|\cdot\|_2$. It is obvious that the nuclear operator $\mathcal{L}$ is an integral operator with the kernel $K(x,y)$ given above. Sometimes one would simply write $K = \sum_{i=1}^{\infty} \rho_i u_i \otimes v_i$, not distinguishing K from $\mathcal{L}$.

(3) Grothendieck [23] generalised Fredholm's theory to nuclear operators on locally convex spaces. Note this becomes natural after identifying the kernel function with the Schmidt representation of the operator as above. We will give Grothendieck's version of the Fredholm theorem in Theorem 5.6.3.

(4) (5.57) can also be used to informally compute the trace of simple transfer operators. In fact, given a simple transfer operator of the form $(\mathcal{L}w)(x) = \varphi(x)w(\psi(x))$. If we define $K(x,y) = \varphi(x)\delta(y-\psi(x))$, where δ is the Dirac function, then $\mathcal{L}$ can be seen as an integral operator with the kernel function K. Now using a local chart near the fixed point, the trace can be given by

$$\operatorname{tr}(\mathcal{L}) = \int K(x,x)\,\mathrm{d}x = \frac{\varphi(t)}{\det(I - D\psi(t))},$$

according to distribution theory, where t is the unique fixed point of ψ. Although the kernel function K here is not continuous, the fact that simple transfer operators are nuclear of order zero (by Theorem 5.5.1) suggests that K can be approximated by continuous kernel functions uniformly. This informal approach is also suggested in [50] Section 1.9.

Theorem 5.6.3. *Let $\mathcal{L}$ be a nuclear operator on a Banach space with the Schmidt representation*

$$\mathcal{L} = \sum_{n=1}^{\infty} \rho_n u_n \otimes f_n,$$

Then the Fredholm determinant

$$\det(I - z\mathcal{L}) := \exp\left(-\sum_{n=1}^{\infty} \frac{\operatorname{tr}(\mathcal{L}^n)}{n} z^n\right)$$

is an entire function.

Moreover, if we let $\det(I - z\mathcal{L}) = 1 + \sum_{n=1}^{\infty}(-1)^n a_n z^n$ be the power series expansion of its Fredholm determinant, then

$$a_n = \sum_{k_1 < \cdots < k_n} \rho_{k_1} \cdots \rho_{k_n} \det(f_{k_i}(u_{k_j}))_{1 \le i,j \le n}.$$

Remark 5.6.4. The proof can be found in [23]. With the same reason as in Remark 5.6.2 (1), we can see if $\mathcal{L}$ is a nuclear operator with the Schmidt representation

$$\mathcal{L} = \sum_{n=1}^{\infty} \rho_n u_n \otimes f_n,$$

then by Theorem 5.6.3 its trace can be given by

$$tr(\mathcal{L}) = \sum_{n=1}^{\infty} \rho_n f_n(u_n).$$

We end this section with a discussion on the convergence rate of the coefficients of the determinant function $\det(I - z\mathcal{L})$ for transfer operators. The following result is first given by [15].

Theorem 5.6.5. *Let $\{M_i\}_{i\in I}$ be a finite set of invertible positive matrices. Choose a proper domain $U \subset C^{d-1}$ and define the operator $\mathcal{L} : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$ by*

$$(\mathcal{L}w)(z) = \sum_{i\in I} \varphi_i(z) w(\hat{M}_i z),$$

where $\varphi_i \in \mathcal{A}_\infty(U)$. Then

$$\det(I - z\mathcal{L}) = 1 + \sum_{n=1}^{\infty} (-1)^n d_n z^n$$

is an entire function with $d_n = O(\delta^{n^{1+1/(d-1)}})$ for some $0 < \delta < 1$.

The proof of the theorem above can be found in [29]. Instead, we give a much weaker conclusion for simple transfer operators by a direct application of previous results.

Theorem 5.6.6. *Let U be a domain in $\mathbb{C}^d$, let $\mathcal{L} : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$ be a simple transfer operator with the form*

$$(\mathcal{L}w)(z) = \varphi(z) w(\psi(z)),$$

where ψ is a holomorphic map sending U strictly inside U. Then

$$\det(I - z\mathcal{L}) = 1 + \sum_{n=1}^{\infty} (-1)^n d_n z^n$$

is an entire function with $d_n = O(\delta^{n^{1+1/d}})$ for some $0 < \delta < 1$.

Proof. By Theorem 5.5.4, we know the eigenvalues of $\mathcal{L}$ are explicitly $\varphi(t)\lambda_1^{n_1}\cdots\lambda_d^{n_d}$ with $n_1,\ldots,n_d \geq 0$ where $|\lambda_1| \geq \cdots \geq |\lambda_d|$ are eigenvalues of $D\psi(t)$ and t is the unique fixed point of ψ.

By Theorem 5.5.2, we know $|\lambda_1| < 1$. Define $(\mu_n)_n$ to be a non-increasing rearrangement of $(\lambda_1^{n_1}\cdots\lambda_d^{n_d})_{n_1,\ldots,n_d}$. Since $n_1+\cdots+n_d \leq N$ has $\binom{N+d}{d}$ non-negative integer solutions for $(n_1,\ldots,n_d)$, we can obtain $\mu_n \leq \lambda_1^N$, where $n = \binom{N+d}{d} = O(N^d)$. Therefore

$$\mu_n = O(\lambda_1^{n^{1/d}}).$$

Since $\mathcal{L}$ is nuclear of order 0, by (5.46) and Theorem 5.6.3, $\det(I - z\mathcal{L})$ is an entire function and we can write

$$\det(I - z\mathcal{L}) = \prod_{n=1}^{\infty}(1-\mu_n z).$$

By comparing the coefficients of $\det(I - z\mathcal{L})$, we have

$$d_m = \sum_{k_1<k_2<\cdots<k_m} \mu_{k_1}\cdots\mu_{k_m}.$$

Choose $\delta = \lambda_1$, now the proof is complete by the lemma below given by [15] pp. 506.

Lemma 5.6.7. *If (μ_n) is a positive sequence satisfying $\mu_n = O(\delta^{n^{1/d}})$ for some $0 < \delta < 1$. Then*

$$d_m := \sum_{k_1<k_2<\cdots<k_m} \mu_{k_1}\cdots\mu_{k_m} = O(\delta^{m^{1+1/d}}).$$

□

Remark 5.6.8. In [29], Theorem 5.6.5 is proved by showing $\rho_n = O(\delta^{n^{1+1/(d-1)}})$ for some $0 < \delta < 1$, where $\{\rho_n\}_{n\geq 1}$ are the coefficients of the Schmidt representation (also known as singular values) of the transfer operator. Now by the lemma above, we have

$$\sum_{k_1<\cdots<k_n} \rho_{k_1}\cdots\rho_{k_n} = O(\delta^{n^{1+1/(d-1)}}).$$

Applying Fredholm theoy Theorem 5.6.3, we know

$$d_n := \sum_{k_1<\cdots<k_n} \rho_{k_1}\cdots\rho_{k_n}\det(f_{k_i}(u_{k_j}))_{1\leq i,j\leq n},$$

if $\mathcal{L}$ admits the representation $\mathcal{L} = \sum_{n\geq 1} \rho_n u_n \otimes f_n$. In the end, the conclusion follows by applying Hadamard's inequality

$$|\det(f_{k_i}(u_{k_j}))_{1\leq i,j\leq n}| = O(n^{n/2}).$$

5.7 Algorithm for the Lyapunov Exponent

Now we are ready to give a complete description of Pollicott's algorithm for the computation of the Lyapunov exponent.

Recall from Section 5.3 that, given a finite set of positive and invertible matrices $\{M_i\}_{1\leq i\leq k}$ with a probability vector $(p_i)_{1\leq i\leq k}$, a family of parametrised transfer operators $\mathcal{L}_t$ on $\mathcal{A}_\infty(U)$ can be defined by

$$(\mathcal{L}_t w)(z) = \sum_{i=1}^{k} p_i e^{t\log\|M_i z\|} w(\hat{M}_i z);$$

when $|t|$ is sufficiently small, $\mathcal{L}_t$ possesses a simple maximal positive top eigenvalue $\beta(t)$. And by Theorem 5.3.3 the Lyapunov exponent associated with this problem equals $\beta'(0)$.

If, for each $1 \leq i \leq k$, we define $\mathcal{M}_i : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$ by

$$(\mathcal{M}_i w)(z) = p_i e^{t\log\|M_i z\|} w(\hat{M}_i z).$$

Then by Theorem 5.5.1, we know $\mathcal{M}_i$ is nuclear of order 0. By Remark 5.4.4 (2), $\mathcal{L}_t = \sum_{i=1}^{k} \mathcal{M}_i$ is also nuclear of order 0. Therefore the **determinant function** $\det(I - z\mathcal{L}_t)$ is well-defined and $\beta(t)$ is a solution of the determinant function by (5.46).

By the trace-determinant formula (5.44), we can obtain

$$\det(I - z\mathcal{L}) = \exp\left(-\sum_{n=1}^{\infty} \frac{z^n}{n} \operatorname{tr}(\mathcal{L}_t^n)\right). \tag{5.58}$$

Therefore in order to compute $\det(I - z\mathcal{L})$, we want to compute $\operatorname{tr}(\mathcal{L}_t^n)$ for each $n \geq 1$.

Now by Corollary 5.5.5 the trace of $\mathcal{M}_i$ can be computed by

$$\operatorname{tr}(\mathcal{M}_i) = \frac{p_i e^{t\log\|M_i s_i\|}}{\det(I - D\hat{M}_i(s_i))} = \frac{p_i \lambda_i^t}{\det(I - D\hat{M}_i(s_i))},$$

where s_i is the fixed point of $\hat{M}_i$, and λ_i is the eigenvalue of M_i corresponding to s_i. Therefore

$$\operatorname{tr}(\mathcal{L}_t) = \sum_{i=1}^{k} \frac{p_i \lambda_i^t}{\det(I - D\hat{M}_i(s_i))}.$$

Moreover, we can see

$$\mathcal{L}_t^n = \left(\sum_{i=1}^k \mathcal{M}_i\right)^n = \sum_{|\underline{i}|=n} \mathcal{M}_{i_n}\mathcal{M}_{i_{n-1}}\cdots\mathcal{M}_{i_1}.$$

And

$$(\mathcal{M}_{i_n}\mathcal{M}_{i_{n-1}}\cdots\mathcal{M}_{i_1}w)(z) = p_{\underline{i}}e^{t\log\|S_{\underline{i}}z\|}w(\hat{S}_{\underline{i}}z),$$

where for $\underline{i} = (i_1, \ldots, i_n)$, $p_{\underline{i}}$ denotes $p_{i_1}p_{i_2}\cdots p_{i_n}$, $S_{\underline{i}}$ denotes $M_{i_n}M_{i_{n-1}}\cdots M_{i_1}$ and $\hat{S}_{\underline{i}}$ denotes $\hat{M}_{i_n}\hat{M}_{i_{n-1}}\cdots\hat{M}_{i_1}$.

Similarly, denote by $s_{\underline{i}}$ the fixed point of $\hat{S}_{\underline{i}}$, and $\lambda_{\underline{i}}$ the top eigenvalue of $S_{\underline{i}}$. Then we have

$$\mathrm{tr}(\mathcal{L}_t^n) = \sum_{|\underline{i}|=n} \frac{p_{\underline{i}}\lambda_{\underline{i}}^t}{\det(I - D\hat{S}_{\underline{i}}(s_{\underline{i}}))}. \tag{5.59}$$

With this formula for traces, by Fredholm theory, we can expand (5.58) as

$$d(z,t) := \det(I - z\mathcal{L}_t) = 1 + \sum_{n=1}^{\infty} a_n(t)z^n. \tag{5.60}$$

On the other hand, by (5.46), we have

$$d(z,t) = \det(I - z\mathcal{L}_t) = (1 - z\beta(t))r(z,t), \tag{5.61}$$

where $r(z,t) := \prod_{j\geq 2}(1 - z\lambda_j(\mathcal{L}_t))$ and $\{\lambda_j(\mathcal{L}_t)\}_{j\geq 2}$ denote the other eigenvalues of $\mathcal{L}_t$ apart from the top eigenvalue $\beta(t)$.

Differentiating (5.61) with respect to t at $t = 0$ on both sides gives

$$\left.\frac{\partial}{\partial t}\right|_{t=0} d(z,t) = -z\beta'(0)r(z,0) + (1-z)\left.\frac{\partial}{\partial t}\right|_{t=0} r(z,t), \tag{5.62}$$

Letting $z = 1$ in (5.62) gives

$$\left.\frac{\partial}{\partial t}\right|_{t=0} d(1,t) = -\beta'(0)r(1,0). \tag{5.63}$$

Then we differentiate (5.61) with respect to z at $z = 1$ on both sides. We can get

$$\left.\frac{\partial}{\partial z}\right|_{z=1} d(z,t) = -\beta(t)r(1,t) + (1-\beta(t))\left.\frac{\partial}{\partial z}\right|_{z=1} r(z,t) \tag{5.64}$$

Letting $t = 0$ in (5.64) gives

$$\left.\frac{\partial}{\partial z}\right|_{z=1} d(z,0) = -r(1,0) \tag{5.65}$$

Note by (5.61), $r(1,0) \neq 0$, since otherwise 1 would not be a simple eigenvalue of $\mathcal{L}_0$, contradicting Theorem 5.2.4.

By (5.63) and (5.65), we get

$$\beta'(0) = \frac{\left.\frac{\partial}{\partial t}\right|_{t=0} d(1,t)}{\left.\frac{\partial}{\partial z}\right|_{z=1} d(z,0)}.$$

As the determinant function is an entire function, we can express the equation above using the expansion (5.60) and obtain

$$\beta'(0) = \frac{\sum_{n=1}^{\infty} a_n'(0)}{\sum_{n=1}^{\infty} n a_n(0)}. \tag{5.66}$$

Now for $p \geq 1$, if we truncate the series (5.60) and use the first $p+1$ terms as the p-th estimate of the determinant function, consequently, we can obtain the p-th estimate of the Lyapunov exponent

$$\gamma^{(p)} = \frac{\sum_{n=1}^{p} a_n'(0)}{\sum_{n=1}^{p} n a_n(0)}. \tag{5.67}$$

Moreover, by (5.58), we can explicitly give $a_n(t)$ by

$$a_n(t) = \sum_{1 \leq l \leq n} \frac{(-1)^l}{l!} \sum_{n_1 + \cdots + n_l = n} \frac{\mathrm{tr}(\mathcal{L}_t^{n_1}) \cdots \mathrm{tr}(\mathcal{L}_t^{n_l})}{n_1 \cdots n_l}. \tag{5.68}$$

And we can summarise the results above and give the following algorithm.

Theorem 5.7.1. *Given a finite set $\{M_i\}_{1 \leq i \leq k}$ of positive and invertible matrices and a probability vector $(p_i)_{1 \leq i \leq k}$. Then the Lyapunov exponent γ associated with $\{M_i\}_{1 \leq i \leq k}$ and $(p_i)_{1 \leq i \leq k}$ can be given by* (5.66)*, where $a_n(t)$ is given by* (5.68) *and the trace $\mathrm{tr}(\mathcal{L}_t^n)$ for each $n \geq 1$ is given by* (5.59)*. The p-th estimate of γ can be obtained from* (5.67)*,*

Remark 5.7.2. (1) While computing the trace $\mathrm{tr}(\mathcal{L}_t^n)$, we need to calculate the denominator $\det(I - D\hat{S}_{\underline{i}}(s_{\underline{i}}))$ of (5.59). The formula given in Pollicott's original paper [48] can only be applied to 2×2 matrices. A more general formula will be given in Section 5.8.

(2) In Pollicott's original paper [48], the weighted transfer operators $\mathcal{L}_{j,t}$ with weights $(e^{\varphi_i^{(j)}})_{1\le i\le k}$ are defined for each $1 \le j \le k$, where

$$\varphi_i^{(j)}(x) = t \log \frac{\|A_j A_i x\|}{\|A_i x\|}, \quad \text{for each } 1 \le j \le k.$$

We can see now it is not an optimal choice of weight functions.

(3) The following theorem on the rate of convergence for this algorithm is a direct application of Theorem 5.6.5.

Theorem 5.7.3. *For any $\epsilon > 0$ there exists $K_0 > 0$ and $C > 0$ such that $|a_n(t)| \le K_0 \exp(-Cn^{1+\frac{1}{d-1}})$ whenever $|t| < \epsilon$. And consequently, $|\gamma^{(p)} - \gamma| = O(\delta^{p^{1+1/(d-1)}})$ for some $0 < \delta < 1$.*

5.8 Modified Lemma

The goal of this section is to provide a general formula to compute the denominator $\det(I - D\hat{S}_{\underline{i}}(s_{\underline{i}}))$ of (5.59), where for each sequence $\underline{i} = (i_1, i_2, \ldots, i_n)$, $\hat{S}_{\underline{i}}$ denotes $\hat{M}_{i_n}\hat{M}_{i_{n-1}}\cdots\hat{M}_{i_1}$ and $s_{\underline{i}}$ denotes the fixed point of $\hat{S}_{\underline{i}}$.

In [48], Lemma 4.2 states the following identity

$$\det(I - D\hat{S}_{\underline{i}}(s_{\underline{i}})) = \left(1 - \frac{\det S_{\underline{i}}}{\lambda_{\underline{i}}^2}\right),$$

where $\lambda_{\underline{i}}$ denotes the top eigenvalue of $S_{\underline{i}} := M_{i_n}M_{i_{n-1}}\cdots M_{i_1}$. However, we will see that this equation is only valid for 2×2 matrices with determinant 1. A general formula is given below with no assumptions on determinants.

Lemma 5.8.1. *(Modified Lemma) For any $d \times d$ positive matrix A, we have*

$$\det(I - D\hat{A}(x)) = \prod_{i=2}^{d}\left(1 - \frac{\lambda_i}{\lambda_1}\right),$$

where λ_1 and x denote the maximal simple positive eigenvalue and the associated eigenvector of A, and $\lambda_2, \cdots, \lambda_d$ are the other eigenvalues.

Proof. Let $(1 \;\; t_2 \;\; \cdots \;\; t_d)^T$ be a coordinate system around the point $\bar{x} = (1 \;\; x_2 \;\; \cdots \;\; x_d)^T$, denote $\sim$ to be the equivalence relation on $\mathbb{R}^d \setminus \{0\}$ given by $\alpha \sim \beta$ if and only if there exists

$\lambda \neq 0$ such that $\alpha = \lambda\beta$.

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1d} \\ a_{21} & a_{22} & \cdots & a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{pmatrix} \begin{pmatrix} 1 \\ t_2 \\ \vdots \\ t_d \end{pmatrix} = \begin{pmatrix} a_{11} + \sum_{j=2}^{d} a_{1j}t_j \\ a_{21} + \sum_{j=2}^{d} a_{2j}t_j \\ \vdots \\ a_{d1} + \sum_{j=2}^{d} a_{dj}t_j \end{pmatrix} \sim \begin{pmatrix} 1 \\ b_2 \\ \vdots \\ b_d \end{pmatrix}$$

where $b_k = \dfrac{a_{k1} + \sum_{j=2}^{d} a_{kj}t_j}{a_{11} + \sum_{j=2}^{d} a_{1j}t_j}, \quad k = 2, 3, \cdots, d$. Therefore,

$$\frac{\partial b_k}{\partial t_i} = \frac{a_{ki}(a_{11} + \sum_{j=2}^{d} a_{1j}t_j) - a_{1i}(a_{k1} + \sum_{j=2}^{d} a_{kj}t_j)}{(a_{11} + \sum_{j=2}^{d} a_{1j}t_j)^2}.$$

At the point $\bar{x}$, the derivative is $\left.\dfrac{\partial b_k}{\partial t_i}\right|_{\bar{x}} = \dfrac{a_{ki} - a_{1i}x_k}{\lambda_1}, \quad k, i = 2, 3, \cdots, d.$

$$(D\hat{A}(x))_{i,j} = \frac{a_{i+1,j+1} - a_{1,j+1}x_{i+1}}{\lambda_1}, \quad i, j = 1, 2, \cdots, d-1.$$

So,

$$\det(I - D\hat{A}(x)) = \lambda_1^{-(d-1)} \det \begin{pmatrix} \lambda_1 - a_{22} + a_{12}x_2 & -a_{23} + a_{13}x_2 & \cdots & -a_{2d} + a_{1d}x_2 \\ -a_{32} + a_{12}x_3 & \lambda_1 - a_{33} + a_{13}x_3 & \cdots & -a_{3d} + a_{1d}x_3 \\ \vdots & \vdots & \ddots & \vdots \\ -a_{d2} + a_{12}x_d & -a_{d3} + a_{13}x_d & \cdots & \lambda_1 - a_{dd} + a_{1d}x_d \end{pmatrix}$$

Since $a_{11} + \sum_{j=2}^{d} a_{1j}x_j = \lambda_1$, we have

$$\begin{aligned} \prod_{i=1}^{d}(\lambda - \lambda_i) &= \det \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1d} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{d1} & -a_{d2} & \cdots & \lambda - a_{dd} \end{pmatrix} \\ &= \det \begin{pmatrix} \lambda - \lambda_1 + \sum a_{1j}x_j & -a_{12} & \cdots & -a_{1d} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{d1} & -a_{d2} & \cdots & \lambda - a_{dd} \end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
&= \det \begin{pmatrix} \lambda - \lambda_1 & -a_{12} & \cdots & -a_{1d} \\ -a_{21} + \lambda x_2 - \sum a_{2j} x_j & \lambda - a_{22} & \cdots & -a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{d1} + \lambda x_d - \sum a_{dj} x_j & -a_{d2} & \cdots & \lambda - a_{dd} \end{pmatrix} \\
&= \det \begin{pmatrix} \lambda - \lambda_1 & -a_{12} & \cdots & -a_{1d} \\ (\lambda - \lambda_1) x_2 & \lambda - a_{22} & \cdots & -a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ (\lambda - \lambda_1) x_d & -a_{d2} & \cdots & \lambda - a_{dd} \end{pmatrix} \\
&= (\lambda - \lambda_1)^d \det \begin{pmatrix} 1 & -a_{12} & \cdots & -a_{1d} \\ x_2 & \lambda - a_{22} & \cdots & -a_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_d & -a_{d2} & \cdots & \lambda - a_{dd} \end{pmatrix} \\
&= (\lambda - \lambda_1) \det \begin{pmatrix} 1 & -a_{12} & \cdots & -a_{1d} \\ 0 & \lambda - a_{22} + a_{12} x_2 & \cdots & -a_{2d} + a_{1d} x_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & -a_{d2} + a_{12} x_d & \cdots & \lambda - a_{dd} + a_{1d} x_d \end{pmatrix},
\end{aligned}
$$

where the third equality follows by multiplying the i-th column by x_i and adding it to the first column for each $2 \leq i \leq d$, the fourth follows by noticing $a_{k1} + \sum_{j=2}^{d} a_{kj} x_j = \lambda_1 x_k$ for each $2 \leq k \leq d$, and the last by multiplying the first row by $-x_i$ and adding it to the i-th row for each $2 \leq i \leq d$.

Therefore

$$
\begin{aligned}
\det(I - D\hat{A}(x)) &= \frac{1}{\lambda_1^{d-1}} \frac{(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_d)}{\lambda - \lambda_1} \bigg|_{\lambda = \lambda_1} \\
&= \prod_{i=2}^{d} \left(1 - \frac{\lambda_i}{\lambda_1} \right).
\end{aligned}
$$

□

When $d = 2$ and $\lambda_1 \lambda_2 = \det(A) = 1$, we have

$$
\det\left(I - D\hat{A}(x) \right) = 1 - \frac{\lambda_2}{\lambda_1},
$$

which coincides with Pollicott's formula. In Section 5.9, we will use numerical experiments to confirm these statements.

5.9 Numerical Results

In this section, we provide numerical results for two purposes: (1) to compare different methods on computing the Lyapunov exponent; (2) to illustrate the advantages of Lemma 5.8.1 for 3×3 matrices.

Comparison among Different Methods

Consider two randomly chosen positive matrices

$$M_1 = \begin{pmatrix} 6.264538 & 8.356286 \\ 1.836433 & 15.952808 \end{pmatrix} \quad \text{and} \quad M_2 = \begin{pmatrix} 3.295078 & 4.874291 \\ 8.204684 & 7.383247 \end{pmatrix}$$

with the probability vector $(1/2 \quad 1/2)$.

n	$\frac{1}{n}\mathbb{E}[\log \|A_{i_n} \cdots A_{i_1}\|]$
1	2.848496
2	2.775680
4	2.732259
6	2.717080
8	2.709486
10	2.704929

Table 5.1: Approximations to γ using complete enumeration

n	$\frac{1}{n}\log \|A_{i_n} \cdots A_{i_1}\|$
100	2.711835
500	2.687567
1,000	2.693547
5,000	2.688680
10,000	2.687495
50,000	2.685902

Table 5.2: Approximations to γ using a randomly chosen sequence (Monte-Carlo Method)

n	$\gamma^{(n)}$
1	2.66830404262989562270
2	2.68471843157167988636
3	2.68649862000410255547
4	2.68668232964356912262
5	2.68670112853547760068
6	2.68670304868715748015
7	2.68670324472386790049
8	2.68670326473559128999

Table 5.3: Cycle-expansion method

n	$\gamma^{(n)}$
1	2.68495107481152750495
2	2.68677207328942779796
3	2.68670230467707460420
4	2.68670327105702488879
5	2.68670326700637840389
6	2.68670326701056927376
7	2.68670326701057460284
8	2.68670326701057593510

Table 5.4: Pollicott's method

Approximations in Table 5.1 is based on the defintion of the Lyapunov exponent (Definition 2.1.2) where $\|\cdot\|$ is the matrix norm induced by l_1-norm. Table 5.2 illustrates Furstenberg-Kesten Theorem (Theorem 2.1.13) and is a Monte-Carlo method to compute γ. For each n, we

randomly choose n i.i.d. samples $\{A_{i_k}\}_{1\le k\le n}$ of random matrices and compute $n^{-1}\log\|A_{i_n}\cdots A_{i_1}\|$ directly.

$\gamma^{(n)}$ for cycle expansion (see (1.5)) in Table 5.3 is given by (1.5), and $\gamma^{(n)}$ for Pollicott's algorithm in Table 5.4 is given by (5.67). We can see that Pollicott's algorithm converges to γ faster than the cycle-expansion method for the first 8 estimates.

Modified Lemma

In R, after setting seed to be 6, we randomly generate two 3×3 matrices M_1, M_2 by choosing each entry independently from a uniform distribution and normalise them so that $\det M_1 = \det M_2 = 1$. The probability vector is chosen to be $\begin{pmatrix}1/2 & 1/2\end{pmatrix}$.

We present in Table 5 numerical estimates using Lemma 4.2 in the original article [48] and Table 6 using the modified lemma.

n	$\gamma^{(n)}$
1	1.21636399319921384432
2	1.20833806968087298372
3	1.20898945643705424402
4	1.20916678285678469429
5	1.20912549861442730759
6	1.20912795102400316694
7	1.20912849027657287593
8	1.20912835316815714393

Table 5.5: Old lemma

n	$\gamma^{(n)}$
1	1.21826749514180576739
2	1.20905546034860567062
3	1.20907056882226981998
4	1.20912928255354135132
5	1.20912836100351284152
6	1.20912836338173157458
7	1.20912836344152019308
8	1.20912836343907681425

Table 5.6: Modified lemma

Meanwhile, using (1.5) of the cycle-expansion method we can get

$$\gamma^{(8)} = 1.20912840717672231960.$$

It is not hard to see that the algorithm converges faster in the first 8 estimates with the modified lemma .

Chapter 6

Generalizations with Transfer Operators

In this chapter, we follow Pollicott's idea and use transfer operators to study two problems: the Markovian products of positive matrices and the Markovian products of general matrices (from a strongly irreducible and contracting matrix set). The difficulties in these settings are to come up with the right definitions for transfer operators and the function spaces they shall act on, and to prove the corresponding spectral properties.

The results of Section 6.1 are new. We construct the Markovian transfer operator to be a "matrix" of simple transfer operators, and study its action on the space of the direct sums of $\mathcal{A}_\infty(U)$. The spectral properties for standard and parametrised transfer operators are given by Theorem 6.1.1 and Theorem 6.1.2 respectively. The trace formula can be found in Theorem 6.1.5 as well as the algorithm for computing the Lyapunov exponent in Theorem 6.1.6.

The ideas in Section 6.2 are inspired by the works on transfer operators for symbolic dynamics [44] and those on i.i.d. products of general matrices [35]. We define our function space $G^+_{\theta,\alpha}$ on the Cartesian product $\Omega^+ \times \mathcal{X}$ as well as the transfer operators acting on it in Section 6.2.1. The boundedness of transfer operators and the Lasota-Yorke inequalities are proved in Section 6.2.2. In Section 6.2.3, we give the spectral properties applying the perturbation theorem from [32]. We remark here that the operators defined in this section are mostly not of trace-class, therefore an algorithm for computing the Lyapunov exponent is not given.

6.1 Markovian Products of Positive Matrices

Given a finite set $\{M_i\}_{1\leq i\leq k}$ of invertible and positive matrices, a $k \times k$ stochastic matrix $P = (p_{ij})_{1\leq i,j\leq k}$ and an initial probability vector $\mathbf{p_0}$. Then on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we

can construct a Markov chain $A(\omega)$ of random matrices such that

$$\mathbb{P}(A(\sigma\omega) = M_j | A(\omega) = M_i) = p_{ij}, \quad 1 \leq i, j \leq k.$$

In words, if the current matrix is M_i, then we have the probability p_{ij} to choose M_j to be the next matrix. Here as suggested in Remark 4.2.4 we will always assume P is strictly positive. Denote $A_n(\omega) := A(\sigma^n\omega),\ n \geq 1$.

By the introduction of Section 4.3, we know the Lyapunov exponent γ associated with this problem, defined by

$$\gamma := \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[\log \|\varphi(n,\omega)\|],$$

exists, where $\varphi(n,\omega) := A_n(\omega)A_{n-1}(\omega)\cdots A_1(\omega)$.

In this section, our goal is to generalise Pollicott's algorithm for Markovian products of positive matrices.

6.1.1 Markovian Transfer Operators

First, we can choose the domain $U \subset C_+^{d-1}$ according to Corollary 5.1.4 such that each $\hat{M}_i$ maps U strictly inside U for each $1 \leq i \leq k$ and let $V = \operatorname{Re} U$.

Now for each pair (i,j) such that $1 \leq i, j \leq k$, we can define an operator $\mathcal{M}_{ij} : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$ by

$$(\mathcal{M}_{ij}w)(z) = p_{ij}w(\hat{M}_j z). \tag{6.1}$$

Denote by $\mathcal{A}_\infty^{\oplus k}(U)$ the direct sum of k copies of $\mathcal{A}_\infty(U)$. Let $\mathbf{q} = (q_1, \ldots, q_k)$ be the stationary distribution for P. For each $\mathbf{w} = (w_1, \ldots, w_k) \in \mathcal{A}_\infty^{\oplus k}(U)$, define $\|\mathbf{w}\| = \sum_{i=1}^k q_i\|w_i\|$. Then $(\mathcal{A}_\infty^{\oplus k}(U), \|\cdot\|)$ is a Banach space.

For each $1 \leq i \leq k$, define $\mathcal{E}_i : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty^{\oplus k}(U)$ by sending $w \in \mathcal{A}_\infty(U)$ to an element in $\mathcal{A}_\infty^{\oplus k}(U)$ with the i-th entry w and 0 elsewhere; define $\mathcal{P}_i : \mathcal{A}_\infty^{\oplus k}(U) \to \mathcal{A}_\infty(U)$ to be the natural projection to the i-th entry.

Now define $\mathcal{M} : \mathcal{A}_\infty^{\oplus k}(U) \to \mathcal{A}_\infty^{\oplus k}(U)$ by

$$\mathcal{M} = \sum_{1\leq i,j\leq k} \mathcal{E}_i\mathcal{M}_{ij}\mathcal{P}_j, \tag{6.2}$$

where $\mathcal{M}_{ij}$ is given by (6.1).

We remark here that $\mathcal{M}_{ij}$ is also a well-defined operator on $\mathcal{C}^\omega(V)$ or $\mathcal{C}(\overline{V}, \mathbb{R})$. Thus if we denote by $\mathcal{C}^\omega(V)^{\oplus k}$ (by $\mathcal{C}(\overline{V}, \mathbb{R})^{\oplus k}$ resp.) the direct sum of k copies of $\mathcal{C}^\omega(V)$ (of $\mathcal{C}(\overline{V}, \mathbb{R})$

resp.), then we can see $\mathcal{M}$ also acts on $\mathcal{C}^\omega(V)^{\oplus k}$ (on $\mathcal{C}(\overline{V},\mathbb{R})^{\oplus k}$ resp.). The norm on $\mathcal{C}(\overline{V},\mathbb{R})^{\oplus k}$ is given by

$$|\mathbf{f}|_\infty = \sum_{i=1}^{k} q_i |f_i|_\infty, \quad \mathbf{f} = (f_1, \ldots, f_k) \in \mathcal{C}(\overline{V},\mathbb{R})^{\oplus k}.$$

Now we prove the following theorem.

Theorem 6.1.1. *Let $\mathcal{M}$ be an operator on $\mathcal{A}_\infty^{\oplus k}(U)$ defined as* (6.2)*. Then*

(a) $\mathcal{M}$ is a bounded operator with $\|\mathcal{M}\| = 1$;

(b) The function $\mathbf{1}^{\oplus k} := \mathbf{1} \oplus \cdots \oplus \mathbf{1} \in \mathcal{A}_\infty^{\oplus k}(U)$ is the eigenfunction of $\mathcal{M}$ corresponding to the eigenvalue 1. This is the maximum eigenvalue, and it is simple;

(c) The rest of the spectrum of $\mathcal{M}$ lies in the open ball $\{z \in \mathbb{C} : |z| < 1\}$;

(d) $\mathcal{M}$ is nuclear of order zero;

(e) There exist k unique bounded positive linear functionals $\nu_1, \ldots, \nu_k$ on $\mathcal{A}_\infty(U)$, such that $\|\nu_i\| = 1$ for each $1 \le i \le k$, and that $\mathcal{M}^*\nu = \nu$ where $\nu = \oplus_{i=1}^k q_i \nu_i$ and $\mathbf{q} = (q_1, \ldots, q_k)$ is the stationary distribution for P;

(f) For any k functions $w_1, \ldots, w_k \in \mathcal{A}_\infty(U)$, denoting $\mathbf{w} = (w_1, \ldots, w_k)$, we have

$$\mathcal{M}^n \mathbf{w} \to (\nu_1(w_1), \ldots, \nu_k(w_k)), \quad \text{in the uniform topology as } n \to \infty,$$

where $\nu_1, \ldots, \nu_k$ are the linear functionals given in (e).

Moreover, when $\mathcal{M}$ is regarded as acting on $\mathcal{C}(\overline{V},\mathbb{R})^{\oplus k}$, then

(g) There exist k unique probability measures $\nu_1^*, \ldots, \nu_k^*$ on $\overline{V}$, such that

$$\int \mathcal{M}\mathbf{f} \, \mathrm{d}\nu^* = \int \mathbf{f} \, \mathrm{d}\nu^*, \quad \mathbf{f} = (f_1, \ldots, f_k) \in \mathcal{C}(\overline{V},\mathbb{R})^{\oplus k}, \tag{6.3}$$

where $\nu^* = \oplus_{i=1}^k q_i \nu_i^*$ and $\mathbf{q} = (q_1, \ldots, q_k)$ is the stationary distribution for P;

(h) For any $\mathbf{f} = (f_1, \ldots, f_k) \in \mathcal{C}(\overline{V},\mathbb{R})^{\oplus k}$,

$$\mathcal{M}^n \mathbf{f} \to \left(\int f_1 \, \mathrm{d}\nu_1^*, \ldots, \int f_k \, \mathrm{d}\nu_k^* \right), \quad \text{in uniform topology as } n \to \infty,$$

where $\nu_1^*, \ldots, \nu_k^*$ are the probability measures given in (g).

Proof. Part 1. For any $\mathbf{w} \in \mathcal{A}_\infty^{\oplus k}(U)$, we have

$$\|\mathcal{M}\mathbf{w}\| = \sum_{i=1}^k q_i \left\| \sum_{j=1}^k \mathcal{M}_{ij} w_j \right\| \leq \sum_{i=1}^k q_i \left| \sum_{j=1}^k p_{ij} w_j(\hat{M}_j z) \right|$$

$$\leq \sum_{i=1}^k \sum_{j=1}^k q_i p_{ij} \|w_j\| = \|\mathbf{w}\|.$$

Thus (a) is proved since $\|\mathcal{M}\mathbf{1}^{\oplus k}\| = \|\mathbf{1}^{\oplus k}\| = 1$.

Part 2. It is obvious that $\mathbf{1}^{\oplus k}$ is an eigenfunction corresponding to the eigenvalue 1. We prove that it is simple by discussing the following two cases.

Case (i). Assume $\mathbf{w}$ is another eigenfunction corresponding to the eigenvalue 1. Then since $\mathcal{M}\mathbf{w} = \mathbf{w}$, we have, for each $1 \leq i \leq k$,

$$w_i(z) = \sum_{j=1}^k (\mathcal{M}_{ij} w_j)(z) = \sum_{j=1}^k p_{ij} w_j(\hat{M}_j z).$$

Choose a compact set K such that $\bigcup_{i=1}^k \hat{M}_i U \subset K \subset U$. Assume $|w_i(z)|$ attains its supremum at $\zeta_i \in \partial U$, note $\hat{M}_j \zeta_i \in K$, we have

$$\|w_i\| = |w_i(\zeta_i)| = \left| \sum_{j=1}^k p_{ij} w_j(\hat{M}_j \zeta_i) \right| \leq \sum_{j \neq i} p_{ij} \|w_j\| + p_{ii} \sup_{z \in K} |w_i(z)| \leq \sum_{j=1}^k p_{ij} \|w_j\|. \quad (6.4)$$

Then

$$\|\mathbf{w}\| = \sum_{i=1}^k q_i \|w_i\| \leq \sum_{i,j=1}^k q_i p_{ij} \|w_j\| = \|\mathbf{w}\|.$$

Therefore the inequalities in (6.4) must be equalities. Then by the maximum modulus principle, each w_i has to be a constant function. Then $\mathbf{w}$ is a right eigenvector of P corresponding to the eigenvalue 1, which has to be $\mathbf{1}^{\oplus k}$.

Case (ii). If the eigenvalue 1 of $\mathcal{M}$ admits a Jordan block. More specifically, there exists some $\mathbf{g} \in \mathcal{A}_\infty^{\oplus k}(U)$ such that $\mathcal{M}\mathbf{g} = \mathbf{1}^{\oplus k} + \mathbf{g}$. This implies that $\mathcal{M}^n \mathbf{g} = n\mathbf{1}^{\oplus k} + \mathbf{g}$ for $n \geq 0$, which is contradictory to the fact that $\|\mathcal{M}^n \mathbf{g}\|$ is uniformly bounded by $\|\mathbf{g}\|$ by part 1.

Part 3. By Theorem 5.5.1, each $\mathcal{M}_{ij}$ is nuclear of order zero. By Remark 5.4.4 (6), each $\mathcal{E}_i \mathcal{M}_{ij} \mathcal{P}_j$ is nuclear of order zero. By Remark 5.4.4 (2), $\mathcal{M}$ is nuclear of order zero. This proves (d).

Part 4. As $\mathcal{M}$ is nuclear, it is compact. Therefore by Schauder's theorem, there exists $\nu \in [\mathcal{A}_\infty^{\oplus k}(U)]' = [(\mathcal{A}_\infty(U))']^{\oplus k}$ such that $\mathcal{M}^*\nu = \nu$ and $\nu(\mathbf{1}^{\oplus k}) = 1$. Write $\nu = \oplus_{i=1}^k r_i\nu_i$ such that $\nu_i \in (\mathcal{A}_\infty(U))'$, $\nu_i(\mathbf{1}) = 1$ for each $1 \leq i \leq k$, and $\sum_{i=1}^k r_i = 1$. For any $\mathbf{w} \in \mathcal{A}_\infty^{\oplus k}(U)$, we have

$$\nu(\mathbf{w}) = \sum_{i=1}^k r_i\nu_i(w_i).$$

Moreover, if we let $\mathbf{w}$ be a vector with $\mathbf{1}$ in the j-th entry and 0 elsewhere in the equation $\nu(\mathbf{w}) = \nu(\mathcal{M}\mathbf{w})$, we can obtain

$$r_j = \sum_{l=1}^k r_l\nu_l(\mathcal{M}_{lj}\mathbf{1}) = \sum_{l=1}^k r_l p_{lj}\nu_l(\mathbf{1}) = \sum_{l=1}^k r_l p_{lj}.$$

This shows that $\mathbf{r} = (r_1, \ldots, r_k)$ is a left eigenvector of P associated with the eigenvalue 1. Thus $\mathbf{r} = \mathbf{q}$. And

$$\nu(\mathbf{w}) = \sum_{i=1}^k q_i\nu_i(w_i). \tag{6.5}$$

Part 5. Note for $1 \leq j \leq k$ and $\mathbf{w} = (w_1, \ldots, w_k) \in \mathcal{A}_\infty^{\oplus k}(U)$,

$$\|\mathcal{P}_j\mathbf{w}\| = \|w_j\| \leq q_j^{-1}\|\mathbf{w}\|. \tag{6.6}$$

Therefore, by part 1 of the proof, we have

$$\|\mathcal{P}_j\mathcal{M}^n\mathbf{w}\| \leq q_j^{-1}\|\mathcal{M}^n\mathbf{w}\| \leq q_j^{-1}\|\mathbf{w}\|.$$

That is to say, $\{\mathcal{P}_j\mathcal{M}^n\mathbf{w}\}$ is uniformly bounded by $q_j^{-1}\|\mathbf{w}\|$.

Then applying Lemma 5.2.2 and following the same arguments as in part 2 of Theorem 5.2.4, we can show $\{\mathcal{P}_j\mathcal{M}^n\mathbf{w}\}_{n\geq 1}$ is also equicontinuous. Hence by applying the Arzelà-Ascoli theorem to $\{\mathcal{P}_j\mathcal{M}^n\mathbf{w}\}_{n\geq 1}$ for each $1 \leq j \leq k$ and choosing further subsequence of $(\mathcal{M}^n\mathbf{w})_{n\geq 1}$ for each coordinate, there exists a convergent subsequence $(\mathcal{M}^{n_\alpha}\mathbf{w})_\alpha$. We write $\mathbf{w}^* = (w_1^*, \ldots, w_k^*)$ as the limit point of $(\mathcal{M}^{n_\alpha}\mathbf{w})_\alpha$.

Note by part 1 of the proof, for any $\mathbf{w} \in \mathcal{A}_\infty^{\oplus k}(U)$, we have $\|\mathcal{M}\mathbf{w}\| \leq \|\mathbf{w}\|$. On the one hand, we have for each $n \geq 1$, $\|\mathcal{M}^n\mathbf{w}^*\| \leq \|\mathbf{w}^*\|$. On the other hand, for $n \geq 1, m \geq 0$, $\|\mathbf{w}^*\| \leq \|\mathcal{M}^{n+m}\mathbf{w}\|$, which implies that $\|\mathbf{w}^*\| \leq \|\mathcal{M}^n\mathbf{w}^*\|$. Hence for each $n \geq 1$,

$$\|\mathbf{w}^*\| = \|\mathcal{M}^n\mathbf{w}^*\|$$

In particular, when $n = 1$ assume $|w_i^*|$ attains its supremum at $\zeta_0^{(i)} \in \partial U$ and $|(\mathcal{M}\mathbf{w}^*)_i|$ at $\zeta^{(i)} \in \partial U$. Then

$$\|\mathbf{w}^*\| = \sum_{i=1}^{k} q_i |w_i^*(\zeta_0^{(i)})| = \sum_{i=1}^{k} q_i |(\mathcal{M}\mathbf{w}^*)_i(\zeta^{(i)})|$$

$$\leq \sum_{1\leq i,j\leq k} q_i p_{ij} |w_j^*(\hat{M}_j \zeta^{(i)})| \leq \sum_{j=1}^{k} q_j |w_j^*(\zeta_0^{(j)})| \leq \|\mathbf{w}^*\|.$$

Therefore the inequalities above must be equalities, and consequently $w_i^*(z)$ is a constant function on U for each $1 \leq i \leq k$, by the maximum modulus principle.

Now by part 4, we have for any $\mathbf{w} \in \mathcal{A}_\infty^{\oplus k}(U)$,

$$\nu(\mathbf{w}) = \lim_{\alpha\to\infty} \nu(\mathcal{M}^{n_\alpha}\mathbf{w}) = \nu(\mathbf{w}^*) = \sum_{i=1}^{k} q_i w_i^* \nu_i(\mathbf{1}) = \sum_{i=1}^{k} q_i w_i^*. \tag{6.7}$$

Comparing (6.7) with (6.5) gives $w_i^* = \nu_i(w_i)$.

Applying the same arguments to each subsequence of $(\mathcal{M}^n\mathbf{w})$, we know there exists a further subsequence which uniformly converges to the same limit $\mathbf{w}^*$. Therefore as $n \to \infty$, in the uniform topology we have

$$\mathcal{M}^n\mathbf{w} \to (\nu_1(w_1), \ldots, \nu_k(w_k)). \tag{6.8}$$

Thus (f) is proved. Moreover, as the evaluations of each ν_i $(1 \leq i \leq k)$ on $\mathcal{A}_\infty(U)$ are completely determined by (6.8), each ν_i $(1 \leq i \leq k)$ is unique.

Notice if we let $\mathbf{w}_i$ be a vector with w_i in the i-th entry and 0 elsewhere, then

$$|\nu_i(w_i)| = \lim_{n\to\infty} |(\mathcal{M}^n\mathbf{w}_i)_i| \leq \lim_{n\to\infty} q_i^{-1}\|\mathcal{M}^n\mathbf{w}_i\| \leq q_i^{-1}\|\mathbf{w}_i\| = \|w_i\|.$$

Together with $\nu_i(\mathbf{1}) = 1$, we have $\|\nu_i\| = 1$ for each $1 \leq i \leq k$.

Part 6. Now we turn to the uniqueness and the construction of a probability measure satisfying (6.3). Let ν^* be any probability measure satisfying (6.3), then by (6.8), for any $\mathbf{w} \in \mathcal{C}^\omega(V)^{\oplus k} \cap \mathcal{A}_\infty^{\oplus k}(U)$,

$$\nu^*(\mathbf{w}) = \lim_{n\to\infty} \nu^*(\mathcal{M}^n\mathbf{w}) = \nu^*(\nu(\mathbf{w})) = \nu(\mathbf{w}).$$

As $\mathcal{C}^\omega(V)^{\oplus k} \cap \mathcal{A}_\infty^{\oplus k}(U)$ is dense in $\mathcal{C}(\overline{V},\mathbb{R})^{\oplus k}$ with respect to the topology induced by the norm $|\cdot|_\infty$ on $\mathcal{C}(\overline{V},\mathbb{R})^{\oplus k}$, the evaluations of ν^* are fully determined, therefore a probability measure ν^* satisfying (6.3) is unique.

By (6.5), $\nu = \oplus_{i=1}^k q_i \nu_i$. For $1 \le i \le k$, denote by $\tilde{\nu}_i$ the restriction of ν_i on $\mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$. Then for each $f \in \mathcal{C}(\overline{V},\mathbb{R})$, let $f^{(\alpha)}$ be a sequence in $\mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$ converging to f in the uniform topology, i.e., $|f^{(\alpha)} - f|_\infty \to 0$ as $\alpha \to \infty$, define

$$\nu_i^*(f) := \lim_{\alpha\to\infty} \tilde{\nu}_i(f^{(\alpha)}), \tag{6.9}$$

and $\nu^* = \oplus_{i=1}^k q_i \nu_i^*$. Equivalently, denote by $\tilde{\nu}$ the restriction of ν on $\mathcal{C}^\omega(V) \cap \mathcal{A}_\infty(U)$, if $\mathbf{f}^{(\alpha)}$ is a sequence in $\mathcal{C}^\omega(V)^{\oplus k} \cap \mathcal{A}_\infty^{\oplus k}(U)$ such that $|\mathbf{f}^{(\alpha)} - \mathbf{f}|_\infty \to 0$ as $\alpha \to \infty$ for $\mathbf{f} \in \mathcal{C}(\overline{V},\mathbb{R})^{\oplus k}$, then

$$\nu^*(\mathbf{f}^{(\alpha)}) = \lim_{\alpha\to\infty} \tilde{\nu}(\mathbf{f}^{(\alpha)}). \tag{6.10}$$

Similarly to part 5 in the proof of Theorem 5.2.4, we can see the extension is well-defined by noticing, if $\mathbf{f}^{(\alpha)} = (f_1^{(\alpha)}, \dots, f_k^{(\alpha)})$ is a sequence in $\mathcal{C}^\omega(V)^{\oplus k} \cap \mathcal{A}_\infty^{\oplus k}(U)$ such that $|\mathbf{f}^{(\alpha)}|_\infty \to 0$ as $\alpha \to \infty$, then for each $1 \le j \le k$,

$$q_j |(\mathcal{M}^n \mathbf{f}^{(\alpha)})_j| \le |\mathcal{M}^n \mathbf{f}^{(\alpha)}|_\infty \le |\mathbf{f}^{(\alpha)}|_\infty \to 0, \quad (\alpha \to \infty).$$

Therefore, for each $1 \le j \le k$

$$\lim_{\alpha\to\infty} |\nu_j^*(f_j^{(\alpha)})| = \lim_{\alpha\to\infty} \lim_{n\to\infty} |(\mathcal{M}^n \mathbf{f}^{(\alpha)})_j| = \lim_{n\to\infty} \lim_{\alpha\to\infty} |(\mathcal{M}^n \mathbf{f}^{(\alpha)})_j| = 0.$$

This shows that the limit given by (6.9) always exists and does not depend on the choice of the sequence of functions.

Now if $f \in \mathcal{C}(\overline{V},\mathbb{R})$ is a non-negative function, for each $1 \le i \le k$, let $\mathbf{f}^{(\alpha)} \in \mathcal{C}^\omega(V)^{\oplus k} \cap \mathcal{A}_\infty^{\oplus k}(U)$ be a sequence such that each $\mathbf{f}^{(\alpha)}$ has $f^{(\alpha)}$ in the i-th entry and 0 elsewhere and that $|f^{(\alpha)} - f|_\infty \to 0$. Then

$$\nu_i^*(f) = \lim_{\alpha\to\infty} (\mathcal{M}^n \mathbf{f}^{(\alpha)})_i \ge 0.$$

Therefore ν_i^* $(1 \le i \le k)$ is a positive linear functional on $\mathcal{C}(\overline{V},\mathbb{R})$ with $\nu_i^*(\mathbf{1}) = 1$. By the Riesz representation theorem, we know there exist probability measures (which we still denote by ν_i^*, $1 \le i \le k$) on $\overline{V}$ such that

$$\nu_i^*(f) = \int f \, \mathrm{d}\nu_i^*, \quad f \in \mathcal{C}(\overline{V},\mathbb{R}).$$

Part 7. In this part, we prove that the construction of ν^* above satisfies the condition given by (g) and (h). In fact, for any $\mathbf{f} \in \mathcal{C}(\overline{V}, \mathbb{R})^{\oplus k}$, let $\mathbf{f}^{(\alpha)} \in \mathcal{C}^\omega(V)^{\oplus k} \cap \mathcal{A}_\infty^{\oplus k}(U)$ be a sequence satisfying $|\mathbf{f}^{(\alpha)} - \mathbf{f}|_\infty \to 0$ as $\alpha \to \infty$. Then we have

$$\begin{aligned}
&|\nu^*(\mathbf{f}) - \nu^*(\mathcal{M}\mathbf{f})| \\
\leq &|\nu^*(\mathbf{f}) - \nu^*(\mathbf{f}^{(\alpha)})| + |\nu^*(\mathbf{f}^{(\alpha)}) - \nu^*(\mathcal{M}\mathbf{f}^{(\alpha)})| + |\nu^*(\mathcal{M}\mathbf{f}^{(\alpha)}) - \nu^*(\mathcal{M}\mathbf{f})| \\
\leq &|\nu^*(\mathbf{f}) - \tilde{\nu}(\mathbf{f}^{(\alpha)})| + |\nu(\mathbf{f}^{(\alpha)}) - \nu(\mathcal{M}\mathbf{f}^{(\alpha)})| + |\tilde{\nu}(\mathcal{M}\mathbf{f}^{(\alpha)}) - \nu^*(\mathcal{M}\mathbf{f})| \\
\to &0 \quad (\alpha \to \infty),
\end{aligned}$$

by applying (6.10) and noticing

$$|\mathcal{M}\mathbf{f}^{(\alpha)} - \mathcal{M}\mathbf{f}|_\infty \leq |\mathbf{f}^{(\alpha)} - \mathbf{f}|_\infty \to 0, \quad \alpha \to \infty.$$

That is to say, $\nu^*(\mathcal{M}\mathbf{f}) = \nu^*(\mathbf{f})$ for $\mathbf{f} \in \mathcal{C}^\omega(V)^{\oplus k} \cap \mathcal{A}_\infty^{\oplus k}(U)$.

Moreover, for $\mathbf{f} = (f_1, \ldots, f_k) \in \mathcal{C}(\overline{V}, \mathbb{R})^{\oplus k}$ and a sequence $\mathbf{f}^{(\alpha)} = (f_1^{(\alpha)}, \ldots, f_k^{(\alpha)}) \in \mathcal{C}^\omega(V)^{\oplus k} \cap \mathcal{A}_\infty^{\oplus k}(U)$ with $|\mathbf{f}^{(\alpha)} - \mathbf{f}|_\infty \to 0$ as $\alpha \to \infty$, letting $n \to \infty$ in the following inequality

$$\begin{aligned}
&\left|(\mathcal{M}^n\mathbf{f})_i - \int f_i \, \mathrm{d}\nu_i^*\right| \\
\leq &\left|(\mathcal{M}^n\mathbf{f})_i - (\mathcal{M}^n\mathbf{f}^{(\alpha)})_i\right| + \left|(\mathcal{M}^n\mathbf{f}^{(\alpha)})_i - \int f_i^{(\alpha)} \, \mathrm{d}\nu_i^*\right| + \left|\int f_i^{(\alpha)} \, \mathrm{d}\nu_i^* - \int f_i \, \mathrm{d}\nu_i^*\right| \\
\leq &q_i^{-1}|\mathbf{f} - \mathbf{f}^{(\alpha)}|_\infty + \left|(\mathcal{M}^n\mathbf{f}^{(\alpha)})_i - \int f_i^{(\alpha)} \, \mathrm{d}\nu_i^*\right| + \left|\int f_i^{(\alpha)} \, \mathrm{d}\nu_i^* - \int f_i \, \mathrm{d}\nu_i^*\right|.
\end{aligned}$$

and applying (6.8) shows that

$$\limsup_{n\to\infty} \left|(\mathcal{M}^n\mathbf{f})_i - \int f_i \, \mathrm{d}\nu_i^*\right| \leq q_i^{-1}|\mathbf{f} - \mathbf{f}^{(\alpha)}|_\infty + \left|\int f_i^{(\alpha)} \, \mathrm{d}\nu_i^* - \int f_i \, \mathrm{d}\nu_i^*\right|$$

Then letting $\alpha \to \infty$ implies $|(\mathcal{M}^n\mathbf{f})_i - \nu_i^*(f_i)|_\infty \to 0$ by (6.9). This proves (h).

Part 8. In the end, we prove (c). Define

$$\Gamma := \{\mathbf{w} \in \mathcal{C}^\omega(V)^{\oplus k} \cap \mathcal{A}_\infty^{\oplus k}(U) : \|\mathbf{w}\| \leq 1, \nu(\mathbf{w}) = 0\}, \quad \mathcal{M}\Gamma := \{\mathcal{M}\mathbf{w} : \mathbf{w} \in \Gamma\} \subset \Gamma.$$

As was shown in part 5, $\mathcal{M}\Gamma$ is relatively compact. Then for some $0 < \delta < 1$, there exists $N > 0$ such that $\|\mathcal{M}^N\mathbf{g}\| \leq \delta < 1$ for any $\mathbf{g} \in \mathcal{M}\Gamma$. That is to say $\|\mathcal{M}^{N+1}\mathbf{w}\| \leq \delta < 1$ for any $\mathbf{w} \in \Gamma$. Thus $\|(\mathcal{M}|_\Gamma)^{N+1}\|^{1/(N+1)} \leq \delta^{1/(N+1)}$, i.e., the spectral radius of $\mathcal{M}|_\Gamma$ is strictly smaller than 1. This completes the proof.

□

6.1.2 Parametrised Markovian Transfer Operators

Following on from the approach of Section 5.2 and Section 5.3, we will now define the parametrised transfer operators associated with the problem of Markovian products of positive matrices and study their spectral properties.

First of all, for each $1 \leq i, j \leq k$, define $\mathcal{L}_{ij} : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$ by

$$(\mathcal{L}_{ij} w)(z) = p_{ij} e^{t \log \|M_j z\|} w(\hat{M}_j z), \tag{6.11}$$

as well as $\mathcal{L}_t : \mathcal{A}_\infty^{\oplus k}(U) \to \mathcal{A}_\infty^{\oplus k}(U)$ by

$$\mathcal{L}_t = \sum_{1 \leq i,j \leq k} \mathcal{E}_i \mathcal{L}_{ij} \mathcal{P}_j, \tag{6.12}$$

where $\mathcal{E}_i$ and $\mathcal{P}_j$ are the same as defined in the previous section. We remark here the symbol $\|\cdot\|$ in (6.11) is defined in (5.28) as a complex extension of the ℓ_2-norm to an analytic function.

Theorem 6.1.2. *Let $\mathcal{L}_t : \mathcal{A}_\infty^{\oplus k}(U) \to \mathcal{A}_\infty^{\oplus k}(U)$ be defined as in* (6.12) *for $t \in \mathbb{R}$. Then when $|t|$ is sufficiently small,*

(a) $\mathcal{L}_t$ is a bounded linear operator on $\mathcal{A}_\infty^{\oplus k}(U)$;

(b) $\mathcal{L}_t$ possesses a maximal simple top eigenvalue $\beta(t) > 0$ and the corresponding eigenfunction $\mathbf{h} = (h_1, \ldots, h_k)$ can be chosen such that each h_i is strictly positive on $\overline{V}$;

(c) The rest of the spectrum lies in the open ball $\{z \in \mathbb{C} : |z| < \beta(t)\}$;

(d) $\mathcal{L}_t$ is nuclear of order zero;

(e) There exist k unique bounded positive linear functionals $\nu_1, \ldots, \nu_k$ on $\mathcal{A}_\infty(U)$ such that

$$\nu(\mathcal{L}_t \mathbf{w}) = \beta(t) \nu(\mathbf{w}), \quad \mathbf{w} \in \mathcal{A}_\infty^{\oplus k}(U),$$

where $\nu = \oplus_{i=1}^k q_i \nu_i$ and $\mathbf{q} = (q_1, \ldots, q_k)$ is the stationary distribution for P;

(f) For any $\mathbf{w} = (w_1, \ldots, w_k) \in \mathcal{A}_\infty^{\oplus k}(U)$, we have

$$\frac{1}{\beta(t)^n} \mathcal{L}_t^n \mathbf{w} \to \big(h_1 \nu_1(w_1), \ldots, h_k \nu_k(w_k)\big), \quad \text{uniformly as } n \to \infty,$$

where $\mathbf{h}$ is the eigenfunction given in (b) such that $\nu(\mathbf{h}) = 1$.

Moreover, when $\mathcal{L}_t$ is regarded as acting on $\mathcal{C}(\overline{V}, \mathbb{R})^{\oplus k}$, then

(g) There exist k unique probability measures $\nu_1^*, \ldots, \nu_k^*$ on $\overline{V}$ such that for each $\mathbf{f} \in \mathcal{C}(\overline{V}, \mathbb{R})^{\oplus k}$,

$$\int \mathcal{L}_t \mathbf{f} \, \mathrm{d}\nu^* = \beta(t) \int \mathbf{f} \, \mathrm{d}\nu^*,$$

where $\nu^* = \oplus_{i=1}^k q_i \nu_i^*$;

(h) For any $\mathbf{f} = (f_1, \ldots, f_k) \in \mathcal{C}(\overline{V}, \mathbb{R})^{\oplus k}$, we have

$$\frac{1}{\beta(t)^n} \mathcal{L}_t^n \mathbf{f} \to \left(h_1(x) \int f_1 \, \mathrm{d}\nu_1^*, \ldots, h_k(x) \int f_k \nu_k^* \right), \quad \textit{uniformly as } n \to \infty,$$

where $\mathbf{h}$ is the eigenfunction given in (b) (restricted on $\overline{V}$) satisfying $\int \mathbf{h} \, \mathrm{d}\nu^* = 1$.

Proof. Part 1. For each $\mathbf{w} \in \mathcal{A}_\infty^{\oplus k}(U)$, we have

$$\|\mathcal{L}_t \mathbf{w}\| = \sum_{i=1}^k q_i \left\| \sum_{j=1}^k \mathcal{L}_{ij} w_j \right\| \leq \sum_{i,j=1}^k q_i p_{ij} e^{|t|\eta(M_j)} \|w_j\|,$$

where $\eta(M) := \sup_{z \in U} |\log \|Mz\||$. Thus

$$\|\mathcal{L}_t \mathbf{w}\| \leq \sum_{j=1}^k q_j e^{|t|\eta(M_j)} \|w_j\| \leq \max_j e^{|t|\eta(M_j)} \|\mathbf{w}\|.$$

Part 2. In this part, we will follow a similar route as in part 1 of the proof of Theorem 5.3.2 to construct certain compact subsets of $\mathcal{A}_\infty^{\oplus k}(U)$, where the eigenfunction $\mathbf{h}$ locates.

From now on, we write $\varphi_j(z) = \log \|M_j z\|$, then

$$\mathcal{P}_i \mathcal{L}_t \mathbf{w} = \sum_{j=1}^k \mathcal{L}_{ij} \mathcal{P}_j \mathbf{w} = \sum_{j=1}^k p_{ij} e^{t\varphi_j(z)} w_j(\hat{M}_j z).$$

Note it has a very similar form with a weighted transfer operator defined by Definition 5.3.1. Therefore by the same proof as in part 1 of Theorem 5.3.2, we have the following inequality analogous to (5.20),

$$\left| (\mathcal{P}_i \mathcal{L}_t) \left(\mathbf{w} + \frac{\mathbf{1}^{\oplus k}}{n} \right) (z) - (\mathcal{P}_i \mathcal{L}_t) \left(\mathbf{w} + \frac{\mathbf{1}^{\oplus k}}{n} \right) (z') \right| \leq D' \|z - z'\|, \tag{6.13}$$

for any $\mathbf{w} \in \mathcal{A}_\infty^{\oplus k}(U)$ with $\|w_j\| \le 1$ for each $1 \le j \le k$, $z, z' \in U$ sufficiently close and some constant $D' > 0$ not depending on $\mathbf{w}$ or $1 \le i \le k$.

Moreover, if w is non-negative when restricted on $\overline{V}$, then similarly to (5.21)

$$\sup_{z\in U}\left|(\mathcal{P}_i\mathcal{L}_t)\left(\mathbf{w}+\frac{\mathbf{1}^{\oplus k}}{n}\right)\right| \ge \frac{1}{n}\sum_{j=1}^{k} p_{ij}e^{-|t\varphi_j(x)|_\infty} \ge \frac{1}{n}\min_j e^{-|t\varphi_j|_\infty} =: D_n,$$

for some $D_n > 0$ not depending on $\mathbf{w}$ or $1 \le i \le k$. And therefore

$$\left\|\mathcal{L}_t\left(\mathbf{w}+\frac{\mathbf{1}^{\oplus k}}{n}\right)\right\| = \sum_{i=1}^{k} q_i \left\|(\mathcal{P}_i\mathcal{L}_t)\left(\mathbf{w}+\frac{\mathbf{1}^{\oplus k}}{n}\right)\right\| \ge D_n. \tag{6.14}$$

Define $\Theta := \max_{1\le i\le k} q_i^{-1} < \infty$ since the transition matrix $P = (p_{ij})$ is strictly positive. And define

$$\Lambda_n := \{\mathbf{w} = (w_1, \dots, w_k) \in \mathcal{A}_\infty^{\oplus k}(U) : \|w_i\| \le 1, \|\partial w_i/\partial z_j\| \le D'/\Theta D_n, 2 \le j \le d,$$
$$\text{and } w_i \text{ is (real and) non-negative on } \overline{V} \text{ for each } 1 \le i \le k.\}$$

Then for $\mathbf{w} \in \Lambda_n$, we have $\|\mathbf{w}\| = \sum_{i=1}^n q_i\|w_i\| \le 1$, and when $z, z' \in U$ are sufficiently close,

$$\|\mathbf{w}(z) - \mathbf{w}(z')\| := \sum_{i=1}^{k} q_i|w_i(z) - w_i(z')| \le (D'/D_n)\|z - z'\|.$$

Therefore Λ_n is compact by the Arzelà-Ascoli theorem.

Now we can define a map L_n on Λ_n by

$$L_n(\mathbf{w}) = \frac{\mathcal{L}_t(\mathbf{w} + \mathbf{1}^{\oplus k}/n)}{\Theta\|\mathcal{L}_t(\mathbf{w} + \mathbf{1}^{\oplus k}/n)\|}.$$

Here we can see L_n maps Λ_n into Λ_n by (6.6) and noticing

$$\|\mathcal{P}_i L_n(\mathbf{w})\| = \frac{\|\mathcal{P}_i\mathcal{L}_t(\mathbf{w} + \mathbf{1}^{\oplus k}/n)\|}{\Theta\|\mathcal{L}_t(\mathbf{w} + \mathbf{1}^{\oplus k}/n)\|} \le \frac{q_i^{-1}}{\Theta} \le 1,$$

as well as

$$|\mathcal{P}_i L_n(\mathbf{w})(z) - \mathcal{P}_i L_n(\mathbf{w})(z')| \le (D'/\Theta D_n)\|z - z'\|$$

for $z, z' \in U$ sufficiently close, by (6.13) and (6.14).

Since Λ_n is obviously convex, by the Schauder-Tychonoff fixed point theorem, there exists a fixed point $\mathbf{h}^{(n)} \in \Lambda_n$ such that $L_n(\mathbf{h}^{(n)}) = \mathbf{h}^{(n)}$. In other words,

$$\mathcal{L}_t\left(\mathbf{h}^{(n)} + \frac{\mathbf{1}^{\oplus k}}{n}\right) = \beta_n \mathbf{h}^{(n)}, \tag{6.15}$$

where

$$\beta_n = \Theta \|\mathcal{L}_t(\mathbf{h}^{(n)} + \mathbf{1}^{\oplus k}/n)\| \geq 0. \tag{6.16}$$

Write $\mathbf{h}^{(n)} = (h_1^{(n)}, \ldots, h_k^{(n)}) \in \Lambda_n$. As $h_i^{(n)}$ is non-negative on $\overline{V}$, for each $1 \leq i \leq k$,

$$\begin{aligned}
\beta_n h_i^{(n)} &= \sum_{j=1}^k p_{ij} e^{t\varphi_j(x)} \left(h_j^{(n)}(\hat{M}_j x) + \frac{1}{n}\right) \\
&\geq \sum_{j=1}^k p_{ij} e^{t\varphi_j(x)} \left(\inf_{x\in V} h_j^{(n)} + \frac{1}{n}\right), \qquad (6.17) \\
&\geq \frac{1}{n}\sum_{j=1}^k p_{ij} e^{t\varphi_j(x)} \geq \frac{1}{n} \min_{1\leq j\leq k} e^{-|t\varphi_j|_\infty}, \quad x \in \overline{V}.
\end{aligned}$$

Therefore $\inf_{x\in V} h_i^{(n)} > 0$ for each $1 \leq i \leq k$. Moreover, (6.17) implies

$$\beta_n \inf_{x\in V} h_i^{(n)} \geq \sum_{j=1}^k p_{ij} e^{t\varphi_j(x)} \left(\inf_{x\in V} h_j^{(n)} + \frac{1}{n}\right),$$

and therefore

$$\begin{aligned}
\beta_n \sum_{i=1}^k q_i \inf_{x\in V} h_i^{(n)} &\geq \sum_{i,j=1}^k q_i p_{ij} e^{t\varphi_j(x)} \left(\inf_{x\in V} h_j^{(n)} + \frac{1}{n}\right) \\
&\geq \sum_{j=1}^k q_j e^{t\varphi_j(x)} \left(\inf_{x\in V} h_j^{(n)} + \frac{1}{n}\right) \\
&\geq \sum_{j=1}^k q_j \left(\inf_{x\in V} h_j^{(n)} + \frac{1}{n}\right) \min_{1\leq l\leq k} e^{-|t\varphi_l|_\infty}.
\end{aligned}$$

As $\sum_{i=1}^k q_i \inf_{x\in V} h_i^{(n)} > 0$, we have

$$\beta_n \geq \frac{\sum_{i=1}^k q_i \inf_{x\in V} h_i^{(n)} + 1/n}{\sum_{i=1}^k q_i \inf_{x\in V} h_i^{(n)}} \min_{1\leq j\leq k} e^{-|t\varphi_j|_\infty} \geq \min_{1\leq j\leq k} e^{-|t\varphi_j|_\infty} =: D'' > 0, \tag{6.18}$$

similarly to (5.24).

Notice $D'' > 0$ does not depend on n. Therefore we can also define

$$\Lambda'' := \{\mathbf{w} \in \mathcal{A}_\infty^{\oplus k}(U) : \|w_i\| \le 1, \|\partial w_i/\partial z_j\| \le D'/D'', 2 \le j \le d, \\ \text{and } w_i \text{ is (real and) non-negative on } \overline{V} \text{ for each } 1 \le i \le k\}.$$

Then by (6.13) and (6.16), we have for each $n \ge 1$, $\mathbf{h}^{(n)} \in \Lambda''$. As Λ'' is compact by the Arzelà-Ascoli theorem, we know there exists a subsequence $(\mathbf{h}^{(n_\alpha)})_\alpha$ in Λ'' uniformly converging to $\mathbf{h} \in \Lambda''$ as $\alpha \to \infty$, then

$$\mathcal{L}_t \mathbf{h} = \beta(t)\mathbf{h},$$

where

$$\beta(t) = \Theta\|\mathcal{L}_t \mathbf{h}\| \ge D'' > 0$$

by (6.15), (6.16), (6.18) and continuity. Moreover, as $\mathbf{h} \in \Lambda''$, each component of $\mathbf{h}$ is non-negative on $\overline{V}$.

Part 3. Write $\mathbf{h} = (h_1, \dots, h_k)$, now we prove for each $1 \le i \le k$, h_i is strictly positive on $\overline{V}$. Denote by $\mathcal{Z}_i$ the set of zeros of function $h_i, 1 \le i \le k$. Now assume $\mathcal{Z}_{i_0} \ne \emptyset$ for some $1 \le i_0 \le k$ and $x_0 \in \mathcal{Z}_{i_0}$. Then

$$0 = \beta(t)h_{i_0}(x_0) = \sum_{j=1}^{k} p_{i_0 j} e^{t\varphi_j(x)} h_j(\hat{M}_j x_0).$$

This shows that $h_j(\hat{M}_j x_0) = 0$ for each $1 \le j \le k$. By induction, for any $x_0 \in \mathcal{Z}_{i_0}$ for some i_0,

$$\{\hat{M}_j \hat{M}_{i_n} \cdots \hat{M}_{i_1} x_0 : n \ge 1, 1 \le i_1, \dots, i_n \le k\} \subset \mathcal{Z}_j,$$

for any $1 \le j \le k$.

Note as $\mathbf{h} \in \Lambda''$, we have $\|\partial h_i/\partial z_j\| \le D'/D'' =: D'''$ for $1 \le i \le k, 2 \le j \le d$. And by Proposition 5.1.8 (d), we can choose $N_1 > 0, \kappa > 0$ and $0 < \delta < 1$ such that when $n > N_1$,

$$\mathbb{E}\left[\left\|\hat{M}(\sigma^n\omega)\cdots\hat{M}(\omega)x - \hat{M}(\sigma^n\omega)\cdots\hat{M}(\omega)y\right\|\right] \le \kappa\delta^n d_H(x,y).$$

Denote $\rho := \max_{1\le i\le k} |\varphi_i(x)|_\infty$. For any $x \in V, \epsilon > 0, 1 \le j \le k$, choose $y \in \bigcup_{1\le j\le k} \mathcal{Z}_j$ and $|t| < -\log\delta/\rho$, there exists $N = \max\{N_1, \log(\epsilon/\kappa D''')/(\log\delta + |t|\rho)\}$ such that when $n > N$,

$$|h_j(x)| \le \sum_{i_1,\dots,i_n} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} e^{|t|n\rho} |h_{i_n}(\hat{M}_{i_n} \cdots \hat{M}_{i_1} x)|$$

$$= \sum_{i_1,\dots,i_n} p_{i_0 i_1} p_{i_1 i_2} \cdots p_{i_{n-1} i_n} e^{|t|n\rho} |h_{i_n}(\hat{M}_{i_n} \cdots \hat{M}_{i_1} x) - h_{i_n}(\hat{M}_{i_n} \cdots \hat{M}_{i_1} y)|$$

$$\leq e^{|t|n\rho} D''' \mathbb{E}\left[\left\|\hat{M}_{i_n} \cdots \hat{M}_{i_1} x - \hat{M}_{i_n} \cdots \hat{M}_{i_1} y\right\|\right]$$

$$\leq \kappa D''' e^{|t|n\rho} \delta^n d_H(x,y) < \epsilon d_H(x,y).$$

As $\epsilon > 0, x \in V$ are arbitrary, $h_j(x)$ is the constant function 0 on $\overline{V}$ for each $1 \leq j \leq k$, which can be uniquely analytically extended to the constant zero function on U. This contradicts $\beta(t) = \Theta\|\mathcal{L}_t \mathbf{h}\| > 0$.

Part 4. Now we prove the uniqueness of the eigenvector $\mathbf{h}$ (up to complex scalars) corresponding to the eigenvalue $\beta(t)$. Assume there is another non-zero eigenfunction $\mathbf{g} = (g_1, \dots, g_k) \in \mathcal{A}_\infty^{\oplus k}(U)$ such that $\mathcal{L}_t \mathbf{g} = \beta(t)\mathbf{g}$. Therefore on $\overline{V}$ we have

$$\mathcal{L}_t(\operatorname{Re}\mathbf{g}) = \beta(t)\operatorname{Re}\mathbf{g}, \quad \mathcal{L}_t(\operatorname{Im}\mathbf{g}) = \beta(t)\operatorname{Im}\mathbf{g}.$$

Since for each $1 \leq i \leq k$, h_i is strictly positive on $\overline{V}$, let

$$u := \inf_{1 \leq i \leq k,\, x \in V} \frac{(\operatorname{Re} g_i)(x)}{h_i(x)}.$$

Assume this infimum is attained at $1 \leq i_0 \leq k, x_0 \in V$. Then we have $(\operatorname{Re} g_i)(x) - u h_i(x) \geq 0$ for any $1 \leq i \leq k, x \in V$ and $(\operatorname{Re} g_{i_0})(x_0) - u h_{i_0}(x_0) = 0$. With the same proof as part 4, we have $(\operatorname{Re} g_j)(x) - u h_j(x)$ is constantly zero on V. Similarly, there exists $v \in \mathbb{R}$ such that $(\operatorname{Im} g_j)(x) = v h_j(x)$ for each $1 \leq j \leq k, x \in V$. Therefore $\mathbf{g} = (u + \mathrm{i}v)\mathbf{h}$ on V and therefore on U.

Part 5. By Theorem 5.5.1, $\mathcal{L}_{ij}$ is nuclear of order zero. Then $\mathcal{E}_i \mathcal{L}_{ij} \mathcal{P}_j$ is nuclear of order zero by Remark 5.4.4 (6). And $\mathcal{L}$ is nuclear of order zero by Remark 5.4.4 (2).

Part 6. With a similar approach as given by part 4 of the proof of Theorem 5.3.2, we can prove $\beta(t)$ and $\mathbf{h}$ are continuous at $t = 0$. This together with Theorem 6.1.1 implies that when $|t|$ is sufficiently small, h_j does not take values in $\{0\} \cup \mathbb{R}_-$ for each $1 \leq j \leq k$, i.e., $\log h_j$ is analytic on U. Therefore we can define

$$\phi_{ij}(t,z) := t \log \|M_j z\| + \log h_j(\hat{M}_j z) - \log h_i(z) - \log \beta(t) \in \mathcal{A}_\infty(U), \quad 1 \leq i, j \leq k.$$

Then we can consider transfer operators $\tilde{\mathcal{L}}_{ij} : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U)$ given by

$$(\tilde{\mathcal{L}}_{ij} w)(z) = p_{ij} e^{\phi_{ij}(t,z)} w(\hat{M}_j z).$$

And define $\tilde{\mathcal{L}}_t = \sum_{i,j} \mathcal{E}_i \tilde{\mathcal{L}}_{ij} \mathcal{P}_j$. Observe that

$$\sum_{j=1}^{k} p_{ij} e^{\phi_{ij}(t,z)} = \frac{1}{\beta(t)} \sum_{j=1}^{k} p_{ij} e^{t \log \|M_j z\|} \frac{h_j(\hat{M}_j z)}{h_i(z)} = 1.$$

Moreover, as h_j $(1 \leq j \leq k)$ is strictly positive when restricted on $\overline{V}$ by part 3. Therefore $\phi_{ij}(t,x)$ is also well-defined on $\overline{V}$ and its restriction on V in fact lies in $\mathcal{C}^\omega(V)$.

We can observe that $\tilde{\mathcal{L}}_t = \beta(t)^{-1} \mathscr{M}_{\mathbf{h}}^{-1} \mathcal{L}_t \mathscr{M}_{\mathbf{h}}$, where $\mathscr{M}_{\mathbf{h}} : \mathcal{A}_\infty^{\oplus k}(U) \to \mathcal{A}_\infty^{\oplus k}(U)$ is the multiplication operator sending $\mathbf{w} = (w_1, \ldots, w_k)$ to $(h_1 w_1, \ldots, h_k w_k)$. Thus the spectrum of $\tilde{\mathcal{L}}_t$ is the spectrum of $\mathcal{L}_t$ scaled by $\beta(t)^{-1}$. Therefore $\mathbf{1}^{\oplus k}$ is the eigenfunction of $\tilde{\mathcal{L}}_t$ corresponding to its maximal top eigenvalue 1. We can use the same proof of Theorem 6.1.1 to prove the rest of the statements for $\tilde{\mathcal{L}}_t$. Thus they are also true for $\mathcal{L}_t$.

□

Now we keep following the same procedure and give a result analogous to Theorem 5.3.3.

Theorem 6.1.3. *Given a problem of Markovian products of positive matrices (i.e., a finite matrix set, a $k \times k$ stochastic matrix and an initial probability vector), let $\mathcal{L}_t$ be the associated operator on $\mathcal{A}_\infty^{\oplus k}(U)$ defined as in* (6.12) *and $\beta(t)$ be its maximal simple top eigenvalue given by Theorem 6.1.2 (b). Then $\beta'(0)$ equals the Lyapunov exponent corresponding to this problem of random matrix products.*

Proof. For each $1 \leq i \leq k$, define

$$\mathcal{L}_i^{(0)} : \mathcal{A}_\infty(U) \to \mathcal{A}_\infty(U),$$
$$(\mathcal{L}_i^{(0)} w)(z) = p_i e^{t \log \|M_i z\|} w(\hat{M}_i z),$$

where $\mathbf{p}_0 = (p_1, \ldots, p_k)$ is the initial probability vector. Let

$$\mathcal{L}^{(0)} := \sum_{i=1}^{k} \mathcal{L}_i^{(0)} \mathcal{P}_i : \mathcal{A}_\infty^{\oplus k}(U) \to \mathcal{A}_\infty(U).$$

Now we have for each $w \in \mathcal{A}_\infty^{\oplus k}(U)$,

$$(\mathcal{L}^{(0)} \mathcal{L}_t^n \mathbf{w})(z) = \sum_{i_0, i_1, \ldots, i_n} (\mathcal{L}_{i_0}^{(0)} \mathcal{L}_{i_0 i_1} \mathcal{L}_{i_1 i_2} \cdots \mathcal{L}_{i_{n-1} i_n} w_{i_n})(z)$$
$$= \sum_{i_0, i_1, \ldots, i_n} p_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} e^{t \log \|M_{i_n} \cdots M_{i_0} z\|} w_{i_n}(\hat{M}_{i_n} \cdots \hat{M}_{i_0} z).$$

Therefore taking $\mathbf{1}^{\oplus k}(V)$ on V, we have $\mathcal{L}^{(0)}\mathcal{L}_t^n\mathbf{1}^{\oplus k} = \mathbb{E}[e^{t\log\|\varphi(n,\omega)x\|}]$.

Write $\mathcal{L}_t = \beta(t)Q(t) + R(t)$, where $Q(t)$ is the projection given by $\mathbf{w} \to \beta(t)\nu(\mathbf{w})$, and $R(t) = \mathcal{L}_t - \beta(t)Q(t)$ has spectral radius strictly smaller than $\beta(t)$. Note $Q(t)R(t) = R(t)Q(t) = 0$. Then we have

$$\mathbb{E}[e^{t\log\|\varphi(n,\omega)x\|}] = \beta(t)^n\mathcal{L}^{(0)}Q(t)\mathbf{1}^{\oplus k} + R(t)^n\mathbf{1}^{\oplus k}$$

Similar to Theorem 5.3.3, let γ be the Lyapunov exponent associated with this problem, then

$$\gamma = \lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\log\|\varphi(n,\omega)\|] = \beta'(0).$$

□

6.1.3 The Trace, the Determinant and the Algorithm

By Theorem 6.1.3 our problem has been converted to the computation of $\beta(t)$, the top eigenvalue of the Markovian transfer operator $\mathcal{L}_t$. Note by Theorem 6.1.2 (d), $\mathcal{L}$ is a nuclear operator of order zero. Thus by Fredholm theory, the Fredholm determinant $\det(I - z\mathcal{L}_t)$ is an entire function with respect to z. And according to the trace-determinant formula (see (5.44)),

$$\det(I - z\mathcal{L}) = \exp\left(-\sum_{n=1}^{\infty}\frac{\operatorname{tr}(\mathcal{L}_t^n)}{n}z^n\right),$$

we expect to compute $\operatorname{tr}(\mathcal{L}_t^n)$ for each $n \geq 1$ so that an explicit expansion of $\det(I - z\mathcal{L}_t)$ can be given.

Lemma 6.1.4. *Let $\mathcal{L}$ be a nuclear operator on $\mathcal{A}_\infty(U)$. Then*

$$tr(\mathcal{E}_i\mathcal{L}\mathcal{P}_j) = \delta_{ij}\operatorname{tr}(\mathcal{L}),$$

where δ_{ij} is the Kronecker symbol which takes 1 when $i = j$ and 0 otherwise.

Proof. Assume $\mathcal{L}$ admits a Schmidt representation $\mathcal{L} = \sum_{n=1}^{\infty}\rho_n u_n \otimes f_n$ with the trace given by

$$\operatorname{tr}(\mathcal{L}) = \sum_{n=1}^{\infty}\rho_n f_n(u_n),$$

by Remark 5.6.4. Then

$$\mathcal{E}_i\mathcal{L}\mathcal{P}_j = \sum_{n=1}^{\infty}\rho_n(\mathcal{E}_i u_n)\otimes(\mathcal{P}_j^* f_n).$$

Again, by Remark 5.6.4, we have

$$\operatorname{tr}(\mathcal{E}_i\mathcal{L}\mathcal{P}_j) = \sum_{n=1}^{\infty} \rho_n(\mathcal{P}_j^* f_n)(\mathcal{E}_i u_n) = \sum_{n=1}^{\infty} \rho_n f_n(\mathcal{P}_j\mathcal{E}_i u_n) = \delta_{ij} \operatorname{tr}(\mathcal{L}).$$

□

Now the trace formula follows immediately.

Theorem 6.1.5. *Let $\mathcal{L}_t$ be the operator on $\mathcal{A}_\infty^{\oplus k}(U)$ given by* (6.12)*, then*

$$\operatorname{tr}(\mathcal{L}_t) = \sum_{i=1}^{k} \operatorname{tr}(\mathcal{L}_{ii}) = \frac{p_{ii}\lambda_i^t}{\det(I - D\hat{M}_i(s_i))}. \tag{6.19}$$

Moreover,

$$\operatorname{tr}(\mathcal{L}_t^n) = \sum_{|\underline{i}|=n} \operatorname{tr}(\mathcal{L}_{i_1 i_2} \cdots \mathcal{L}_{i_{n-1} i_n} \mathcal{L}_{i_n i_1}) = \sum_{|\underline{i}|=n} \frac{p_{\underline{i}}^* \lambda_{\underline{i}}^t}{\det(I - D\hat{S}_{\underline{i}}(s_{\underline{i}}))}, \tag{6.20}$$

*where for each n-length sequence $\underline{i} = (i_1, \ldots, i_n)$, $p_{\underline{i}}^*$ denotes the **cyclic** probability $p_{i_1 i_2} \cdots p_{i_{n-1} i_n} p_{i_n i_1}$, $\hat{S}_{\underline{i}}$ denotes the product $\hat{M}_{i_n} \cdots \hat{M}_{i_1}$, $s_{\underline{i}}$ denotes the unique fixed point of $\hat{S}_{\underline{i}}$ and $\lambda_{\underline{i}}$ denotes the top eigenvalue of $S_{\underline{i}} := M_{i_n} \cdots M_{i_1}$.*

Proof. The first equality of (6.19) follows directly from the definition (6.12) and Lemma 6.1.4, and the second from Corollary 5.5.5.

For (6.20), we notice

$$\mathcal{L}_t^n = \left(\sum_{i,j=1}^{k} \mathcal{E}_i \mathcal{L}_{ij} \mathcal{P}_j \right)^n = \sum_{i_0, i_1, \ldots, i_n} \mathcal{E}_{i_0} \mathcal{L}_{i_0 i_1} \cdots \mathcal{L}_{i_{n-1} i_n} \mathcal{P}_{i_n},$$

since $\mathcal{P}_j\mathcal{E}_i = \delta_{ij} \operatorname{id}_{\mathcal{A}_\infty(V)}$. Therefore by Lemma 6.1.4,

$$\operatorname{tr}(\mathcal{L}_t^n) = \sum_{i_1, \ldots, i_n} \operatorname{tr}(\mathcal{L}_{i_n i_1} \mathcal{L}_{i_1 i_2} \cdots \mathcal{L}_{i_{n-1} i_n}).$$

Note

$$(\mathcal{L}_{i_n i_1} \mathcal{L}_{i_1 i_2} \cdots \mathcal{L}_{i_{n-1} i_n} w)(x) = p_{\underline{i}}^* e^{t \log \|S_{\underline{i}} x\|} w(\hat{S}_{\underline{i}} x),$$

thus the rest follows from Corollary 5.5.5. □

Now we are ready to give the following algorithm.

Theorem 6.1.6. *Given a finite set $\{M_i\}_{1\leq i\leq k}$ of positive and invertible matrices, a $k\times k$ positive stochastic matrix $P=(p_{ij})_{1\leq i,j\leq k}$ and an initial probability vector $\mathbf{p}_0$. On a probability space $(\Omega,\mathcal{F},\mathbb{P})$, define a Markov chain $A(\omega)$ of random matrices with $\mathbb{P}(A(\sigma\omega)=M_j|A(\omega)=M_i)=p_{ij}$ for any $1\leq i,j\leq k$. Let γ be the Lyapunov exponent associated with this random matrix product problem. Then we have*

$$\gamma=\frac{\sum_{n=1}^{\infty}a_n'(0)}{\sum_{n=1}^{\infty}na_n(0)},$$

where for each $n\geq 1$, $a_n(t)$ is given by

$$a_n(t)=\sum_{1\leq l\leq n}\frac{(-1)^l}{l!}\sum_{n_1+\cdots+n_l=n}\frac{\mathrm{tr}(\mathcal{L}_t^{n_1})\cdots\mathrm{tr}(\mathcal{L}_t^{n_l})}{n_1\cdots n_l},$$

and $\mathrm{tr}(\mathcal{L}_t^m)$ is given by Theorem 6.1.5, for each $m\geq 1$. Moreover, the p-th estimate of γ can be given by

$$\gamma^{(p)}=\frac{\sum_{n=1}^{p}a_n'(0)}{\sum_{n=1}^{p}na_n'(0)}.$$

Remark 6.1.7. Having worked to establish the spectral properties of Markovian transfer operators and the trace formula, the rest of the work is exactly the same as described in Section 5.7. We can also notice this algorithm contains Theorem 5.7.1 as its special case. With the same reason as for Pollicott's algorithm, we know this estimation also has $|\gamma^{(p)}-\gamma|=O(\delta^{n^{1+1/(d-1)}})$ for some $0<\delta<1$.

Remark 6.1.8. An area for further exploration is to generalise the results in this section to Markovian shifts over countable (or uncountable) symbols. Moreover, in the definition (6.12), we construct $\mathcal{L}$ to be a $k\times k$ matrix of simple transfer operators. In fact, matrices of operators are intensively studied in operator K-theory. And $\mathcal{L}_t$ and $\mathcal{M}$ are in fact homotopic. We expect that there are deeper connections to be drawn between these two areas.

6.2 Markovian Products of General Matrices

We now turn to a more general setting. Instead of being a set of positive matrices, assume $\{M_i\}_{1\leq i\leq k}$ is a finite set of invertible matrices satisfying strong irreducibility and the contracting property. In this section, we will find a proper function space on which we can define transfer operators and prove the corresponding spectral properties.

6.2.1 Function Space $G^+_{\theta,\alpha}$

Assume $\mathcal{K} = \{1, 2, \ldots, k\}$ and define the one-sided shift space $\Omega^+ = \mathcal{K}^{\mathbb{N}}$ over k symbols. Let $0 < \theta < 1$, define the metric d_θ on Ω^+ by $d_\theta(\omega, \tilde{\omega}) = \theta^N$ where N is the largest integer such that $\omega_i = \tilde{\omega}_i, 0 \le i < N$.

Denote the real projective space $\mathbb{R}P^{d-1}$ $(d \ge 2)$ by $\mathcal{X}$. Define the metric $d_{\mathcal{X}}$ on the real projective space by $d_{\mathcal{X}}(x, y) = (1 - \langle x, y\rangle^2)^{1/2}$, where $x, y \in \mathbb{R}P^{d-1}$ are two unit vectors.

Let $0 < \theta < 1, \alpha > 0$. Define $G^+_{\theta,\alpha}$ to be a function space on $\Omega^+ \times \mathcal{X}$ such that for any $f \in G^+_{\theta,\alpha}$,

(a) uniformly for each $x \in \mathcal{X}$, $f(\omega, x)$ is Lipschitz with respect to ω under the metric d_θ;

(b) uniformly for each $\omega \in \Omega^+$, $f(\omega, x)$ is α-Hölder with respect to x under the metric $d_{\mathcal{X}}$ on $\mathcal{X}$.

In other words, (a) there exists a finite constant $C_1 > 0$ (not depending on x) such that for each $x \in \mathcal{X}$, we have

$$|f(\omega, x) - f(\tilde{\omega}, x)| \le C_1 d_\theta(\omega, \tilde{\omega});$$

(b) there exists a finite constant $C_2 > 0$ (not depending on ω) such that for each $\omega \in \Omega^+$, we have

$$|f(\omega, x_1) - f(\omega, x_2)| \le C_2 d^\alpha_{\mathcal{X}}(x_1, x_2).$$

Now for any $(\omega, x_1), (\tilde{\omega}, x_2) \in \Omega^+ \times \mathcal{X}$, if $f \in G^+_{\theta,\alpha}$ we have

$$\begin{aligned} |f(\omega, x_1) - f(\tilde{\omega}, x_2)| &\le |f(\tilde{\omega}, x_2) - f(\omega, x_2)| + |f(\omega, x_1) - f(\omega, x_2)| \\ &\le C_1 d_\theta(\omega, \tilde{\omega}) + C_2 d^\alpha_{\mathcal{X}}(x_1, x_2) \\ &\le \max\{C_1, C_2\}(d^\alpha_{\mathcal{X}}(x_1, x_2) + d_\theta(\omega, \tilde{\omega})). \end{aligned}$$

Thus we can define

$$|f|_{\theta,\alpha} = \sup_{\substack{\omega \ne \tilde{\omega} \in \Omega^+ \\ \text{or} \\ x_1 \ne x_2 \in \mathcal{X}}} \frac{|f(\omega, x_1) - f(\tilde{\omega}, x_2)|}{d^\alpha_{\mathcal{X}}(x_1, x_2) + d_\theta(\omega, \tilde{\omega})}.$$

Conversely, any function f on $\Omega^+ \times \mathcal{X}$ such that $|f|_{\theta,\alpha} < \infty$ obviously lies in $G^+_{\theta,\alpha}$.

Lemma 6.2.1. *Given $M \in GL(d, \mathbb{R})$ and define*

$$\ell(M) := \sup\{\log^+ \|M\|, \log^+ \|M^{-1}\|\}.$$

Then

$$d_{\mathcal{X}}(M \cdot x_1, M \cdot x_2) \leq e^{4\ell(M)} d_{\mathcal{X}}(x_1, x_2),$$

for any $x_1, x_2 \in \mathbb{R}P^{d-1}$.

Proof. By Lemma III.5.4 in [4], we have

$$\left|\log \| \wedge^p M\|\right| \leq p\, \ell(M);$$

$$\left|\log \| \wedge^p Mu\|\right| \leq p\, \ell(M)\|u\|, \tag{6.21}$$

for any $u \in \wedge^p \mathbb{R}^d$, where $1 \leq p < d$ is an integer. Therefore

$$\frac{d_{\mathcal{X}}(M \cdot x_1, M \cdot x_2)}{d_{\mathcal{X}}(x_1, x_2)} = \frac{\|Mx_1 \wedge Mx_2\|}{\|Mx_1\|\|Mx_2\|} \frac{\|x_1\|\|x_2\|}{\|x_1 \wedge x_2\|} \leq e^{4\ell(M)}.$$

□

Proposition 6.2.2. $(G^+_{\theta,\alpha}, \|\cdot\|_{\theta,\alpha})$ *is a Banach space, where* $\|\cdot\|_{\theta,\alpha} = |\cdot|_\infty + |\cdot|_{\theta,\alpha}$.

Proof. Since the sum of a norm and a seminorm is a norm, we only need to show completeness.

Let $\{f_n\}_{n\geq 1}$ be a Cauchy sequence with respect to $\|\cdot\|_{\theta,\alpha}$. Then $\{\|f_n\|_{\theta,\alpha} : n \geq 1\}$ is bounded. Consequently, $\{f_n\}_{n\geq 1}$ is uniformly bounded (since it is bounded under the sup-norm) and equicontinuous (since it is bounded under $|\cdot|_{\theta,\alpha}$). As $\Omega^+ \times \mathcal{X}$ is a compact space, by the Arzelà-Ascoli theorem, $\{f_n\}_{n\geq 1}$ has a limit point f in the $|\cdot|_\infty$-topology. We want to show that f_n converges to the same limit in the $|\cdot|_{\theta,\alpha}$-topology.

For any $\epsilon > 0$, there exists $N > 0$, such that when $m, n > N$, $\|f_m - f_n\|_{\theta,\alpha} < \epsilon$. So $|f_m - f_n|_{\theta,\alpha} < \epsilon$. Fixing two points (ω, x) and $(\tilde{\omega}, y)$, then

$$\begin{aligned} |(f_m(\omega, x) - f_n(\omega, x)) - (f_m(\tilde{\omega}, y) - f_n(\tilde{\omega}, y)| &\leq |f_m - f_n|_{\theta,\alpha} \left(d_\theta(\omega, \tilde{\omega}) + d^\alpha_{\mathcal{X}}(x, y)\right) \\ &< \epsilon \left(d_\theta(\omega, \tilde{\omega}) + d^\alpha_{\mathcal{X}}(x, y)\right). \end{aligned}$$

As f_n converges to f uniformly through a subsequence, we have

$$|(f_m(\omega, x) - f(\omega, x)) - (f_m(\tilde{\omega}, y) - f(\tilde{\omega}, y)| \leq \epsilon \left(d_\theta(\omega, \tilde{\omega}) + d^\alpha_{\mathcal{X}}(x, y)\right).$$

So $|f_m - f|_{\theta,\alpha} \leq \epsilon$ for $m \geq N$. This completes the proof. □

6.2.2 Markovian Transfer Operators on $G^+_{\theta,\alpha}$

Given a random matrix $M(\omega)$ on the Markovian shift space $(\Omega^+, \mathcal{F}^+, \mathbb{P})$ over k symbols induced by the $k \times k$ positive stochastic matrix $P = (p_{ij})_{1\le i,j\le k}$, where

$$p_{ij} = \mathbb{P}(\omega_0 = i | \omega_1 = j), \quad 1 \le i, j \le k,$$

and the initial probability vector $\mathbf{p}_0$, to imitate (6.11), we want to define a family of parametrised transfer operators $\mathcal{L}_t$ on $G^+_{\theta,\alpha}$ by

$$(\mathcal{L}_t w)(\omega, x) = \sum_{\sigma\omega' = \omega} p(\omega, \omega') e^{t \log \|M(\omega')x\|} w\left(\omega', M(\omega') \cdot x\right), \tag{6.22}$$

where for any $\xi, \eta \in \Omega^+$, $p(\xi, \eta)$ denote the entry $p_{\xi_0\eta_0}$ from the positive stochastic matrix P.

If (6.22) gives a well-defined operator on $G^+_{\theta,\alpha}$, then similar to Section 6.1, we can get the n-th power $\mathcal{L}^n_t$ of $\mathcal{L}_t$,

$$(\mathcal{L}^n_t w)(\omega, x) = \sum_{\sigma^n\omega^{(n)} = \omega} P(n, \omega^{(n)}) e^{t \log \|\psi(n,\omega^{(n)})x\|} w(\omega^{(n)}, \psi(n, \omega^{(n)}) \cdot x), \tag{6.23}$$

where for any given $\omega^{(n)} \in \Omega^+, n \ge 1$, we write

$$\omega^{(k)} := \sigma^{n-k}\omega^{(n)}, \quad \text{for } 1 \le k \le n, \quad \omega' = \omega^{(1)}, \omega = \omega^{(0)}; \tag{6.24}$$

$$P(n, \omega) := p(\sigma^n\omega, \sigma^{n-1}\omega) p(\sigma^{n-1}\omega, \sigma^{n-2}\omega) \cdots p(\sigma\omega, \omega); \tag{6.25}$$

$$\psi(n, \omega) := M(\omega) M(\sigma\omega) \cdots M(\sigma^{n-1}\omega). \tag{6.26}$$

The goal of this section and the next is to prove that when $|t|$ is sufficiently small, $\mathcal{L}_t$ possesses a maximal simple top eigenvalue $\beta(t)$ with the rest of the spectrum lying strictly inside the open ball $\{z \in \mathbb{C} : |z| < \beta(t)\}$. This would guarantee a decomposition $\mathcal{L}_t = \beta(t)Q(t) + R(t)$, where $Q(t)$ is a one-dimensional projection and $R(t)$ has spectral radius strictly smaller than $\beta(t)$.

Moreover, if we define on $G^+_{\theta,\alpha}$

$$(\mathcal{L}^{(0)} w)(\omega, x) = \sum_{\sigma\omega' = \omega} p(\omega') e^{t \log \|M(\omega')x\|} w(M(\omega') \cdot x),$$

where $p(\omega) = p_{\omega_0}$ and $(p_1, \ldots, p_k) = \mathbf{p}_0$ denotes the initial probability vector. We obtain

$$\mathcal{L}^{(0)} \mathcal{L}^n_t \mathbf{1} = \mathbb{E}[e^{t \log \|\psi(n+1,\omega^{(n)})x\|}], \tag{6.27}$$

then following the proof of Theorem 5.3.3, we can see $\beta'(0)$ equals the Lyapunov exponent associated to this problem of Markovian random matrix products.

We start this section with defining of a more general type of transfer operators, namely the weighted Markovian transfer operators, and prove their boundedness afterwards.

Given a complex weight function g on $\Omega^+ \times \mathcal{X}$, define

$$(\mathcal{L}_g w)(\omega, x) = \sum_{\sigma\omega'=\omega} p(\omega, \omega') e^{g(\omega', M(\omega')\cdot x)} w(\omega', M(\omega')\cdot x). \tag{6.28}$$

Unless otherwise indicated, we will assume $g \in G^+_{\theta,\alpha}$ with $g = u + \mathrm{i}v$, where u, v are real functions. If $\mathcal{L}_u \mathbf{1} = 1$, we say $\mathcal{L}_g$ is **normalised**.

Theorem 6.2.3. *Let $(\Omega^+, \mathcal{F}^+, \mathbb{P})$ be a one-sided Markovian shift space over k symbols, and $M(\omega) : \Omega^+ \to GL(d, \mathbb{R})$ depend only on the first coordinate ω_0. Then $\mathcal{L}_g$ defined by* (6.28) *is a bounded operator on $(G^+_{\theta,\alpha}, \|\cdot\|_{\theta,\alpha})$.*

Proof. We first notice that if $g \in G^+_{\theta,\alpha}$, then $e^g \in G^+_{\theta,\alpha}$, i.e., $|e^g|_{\theta,\alpha} < \infty$. This follows from

$$|e^{g(\omega,x)} - e^{g(\tilde{\omega},y)}| \le e^{|g|_\infty} |g(\omega, x) - g(\tilde{\omega}, y)|.$$

We also have $|\mathcal{L}_g|_\infty \le |e^g|_\infty < \infty$ since

$$|\mathcal{L}_g w|_\infty \le \sum_{\sigma\omega'=\omega} p(\omega, \omega') |e^g|_\infty |w|_\infty = |e^g|_\infty |w|_\infty. \tag{6.29}$$

Moreover, for any $\omega, \tilde{\omega} \in \Omega^+$ and $x \in \mathcal{X}$, let $i\omega$ denote the sequence satisfying $(i\omega)_0 = i$ and $\sigma(i\omega) = \omega$, now we have

$$\begin{aligned}
&|(\mathcal{L}_g w)(\omega, x) - (\mathcal{L}_g w)(\tilde{\omega}, x)| \\
\le &\sum_{i=1}^{k} \left| p(\omega, i\omega) e^{g(\omega, M_i\cdot x)} w(i\omega, M_i \cdot x) - p(\tilde{\omega}, i\tilde{\omega}) e^{g(i\tilde{\omega}, M_i\cdot x)} w(i\tilde{\omega}, M_i \cdot x) \right| \\
\le &\sum_{i=1}^{k} \left| p(\omega, i\omega) - p(\tilde{\omega}, i\tilde{\omega}) \right| |e^g|_\infty |w|_\infty + \sum_{i=1}^{k} p(\tilde{\omega}, i\tilde{\omega}) \left| e^{g(i\omega, M_i\cdot x)} - e^{g(i\tilde{\omega}, M_i\cdot x)} \right| |w|_\infty \\
&\qquad + \sum_{i=1}^{k} p(\tilde{\omega}, i\tilde{\omega}) |e^g|_\infty \left| w(i\omega, M_i \cdot x) - w(i\tilde{\omega}, M_i \cdot x) \right| \\
\le &\left(2|e^g|_\infty |w|_\infty + \theta |e^g|_{\theta,\alpha} |w|_\infty + |e^g|_\infty \, \theta |w|_{\theta,\alpha} \right) \, d_\theta(\omega, \tilde{\omega}).
\end{aligned}$$

And for any $\omega \in \Omega^+$ and $x, y \in \mathcal{X}$,

$$
\begin{aligned}
&|(\mathcal{L}_g w)(\omega, x) - (\mathcal{L}_g w)(\omega, y)| \\
\leq &\sum_{\sigma\omega'=\omega} p(\omega, \omega') \left(|e^{g(\omega', M(\omega')\cdot x)} - e^{g(\omega', M(\omega')\cdot y)}||w|_\infty + |w(M(\omega')\cdot x) - w(M(\omega')\cdot y)||e^g|_\infty \right) \\
\leq &\sum_{\sigma\omega'=\omega} p(\omega, \omega') \left(|e^g|_{\theta,\alpha}|w|_\infty + |w|_{\theta,\alpha}|e^g|_\infty \right) d^\alpha_{\mathcal{X}}(M(\omega')\cdot x, M(\omega')\cdot y) \\
\leq &\sum_{\sigma\omega'=\omega} p(\omega, \omega') \left(|e^g|_{\theta,\alpha}|w|_\infty + |e^g|_\infty|w|_{\theta,\alpha} \right) e^{4\alpha\ell(M(\omega'))} d^\alpha_{\mathcal{X}}(x, y) \qquad \text{(by Lemma 6.2.1)} \\
\leq &e^{4\alpha K} \left(|e^g|_{\theta,\alpha}|w|_\infty + |e^g|_\infty|w|_{\theta,\alpha} \right) d^\alpha_{\mathcal{X}}(x, y),
\end{aligned}
$$

if we denote $K := \max_{1\leq i\leq k} \ell(M_i) < \infty$. This shows

$$
|\mathcal{L}_g w|_{\theta,\alpha} \leq \left[2|e^g|_\infty + (\theta + e^{4\alpha K})|e^g|_{\theta,\alpha} \right] |w|_\infty + (\theta + e^{4\alpha K})|e^g|_\infty |w|_{\theta,\alpha}. \tag{6.30}
$$

Therefore $\mathcal{L}_g$ is a bounded operator on $(G^+_{\theta,\alpha}, \|\cdot\|_{\theta,\alpha})$ by (6.29) and (6.30). □

Note if $g(\omega, x) = -t \log \|M(\omega)^{-1}x\|$ lies in $G^+_{\theta,\alpha}$, as

$$
g(\omega', M(\omega')\cdot x) = -t \log \left\| M(\omega')^{-1} \frac{M(\omega')x}{\|M(\omega')x\|} \right\| = t \log \|M(\omega')x\|,
$$

$\mathcal{L}_g$ would coincide with the parametrised Markovian transfer operators given by (6.22). The following lemma in [4] (see Chapter V Lemma 4.2) can be used to validate this choice of g.

Lemma 6.2.4. *For $t \in \mathbb{C}$ and $0 < \alpha \leq 1$, there exists $c_1, c_2 > 0$ such that for any $M \in GL(d, \mathbb{R})$,*

(a) $$\sup_{x\neq y\in\mathcal{X}} \frac{\left|\log\|Mx\| - \log\|My\|\right|}{d^\alpha_{\mathcal{X}}(x, y)} \leq c_1\, \ell(M) e^{2\alpha\ell(M)}.$$

(b) $$\sup_{x\neq y\in\mathcal{X}} \frac{\left|e^{t\log\|Mx\|} - e^{t\log\|My\|}\right|}{d^\alpha_{\mathcal{X}}(x, y)} \leq c_2\, e^{[(1+\alpha)|\operatorname{Re} t| + 2\alpha]\ell(M)}.$$

In fact, if we denote $K := \max_{1\leq i\leq k} \ell(M_i)$, then by Lemma 6.2.4 (a), for $\omega \in \Omega^+$ and $x, y \in \mathcal{X}$,

$$
|g(\omega, x) - g(\omega, y)| \leq c_1 |\operatorname{Re} t| K e^{2\alpha K} d^\alpha_{\mathcal{X}}(x, y).
$$

And for $\omega, \tilde{\omega} \in \Omega^+, x \in \mathcal{X}$,

$$|g(\omega, x) - g(\tilde{\omega}, x)| \leq 2|\operatorname{Re} t| K d_\theta(\omega, \tilde{\omega}).$$

These two inequalities show that $|g|_{\theta,\alpha} < \infty$ and therefore we can give the following corollary of Theorem 6.2.3.

Corollary 6.2.5. *Let $(\Omega^+, \mathcal{F}^+, \mathbb{P})$ be a one-sided Markovian shift space over k symbols, and $M(\omega) : \Omega^+ \to GL(d, \mathbb{R})$ depend only on the first coordinate ω_0. Then $\mathcal{L}_t$ defined by* (6.22) *is a bounded operator on $(G^+_{\theta,\alpha}, \|\cdot\|_{\theta,\alpha})$.*

Remark 6.2.6. We remark here that, under the assumption of Theorem 6.2.3, $\mathcal{L}_g$ and $\mathcal{L}_t$ can also be regarded as bounded operators acting on the space of bounded Borel functions on $\Omega^+ \times \mathcal{X}$ by (6.29).

The advantages of studying weighted transfer operators of the form (6.28) will be clear if we define a skew product Φ on $\Omega^+ \times \mathcal{X}$ by

$$\Phi(\omega, x) = (\sigma\omega, M(\omega)^{-1} \cdot x).$$

Clearly Φ is the skew product induced by the RDS φ determined by $\varphi(1, \omega) = M(\omega)^{-1}$ by Definition 4.2.7, And $\mathcal{L}_g$ can be simplified as

$$(\mathcal{L}_g w)(\omega, x) = \sum_{(\omega', x') \in \Phi^{-1}(\omega, x)} p(\omega, \omega') e^{g(\omega', x')} w(\omega', x')$$

By Proposition 4.3.1, there exists an invariant measure ν on $\Omega^+ \times \mathcal{X}$ for Φ. Moreover, we can prove the following theorem.

Theorem 6.2.7. *Let $\mathcal{L}_0$ be the parametrised transfer operator $\mathcal{L}_t$ defined by* (6.22) *when $t = 0$, let ν be a probability measure on $\Omega^+ \times \mathcal{X}$. Then the following statements are equivalent:*

(a) ν is invariant with respect to Φ;

(b) For any $w \in G^+_{\theta,\alpha}$, we have

$$\int \mathcal{L}_0 w \, \nu(\mathrm{d}\omega, \mathrm{d}x) = \int w \, \nu(\mathrm{d}\omega, \mathrm{d}x) \tag{6.31}$$

Proof. Suppose ν is invariant for Φ. We can define a two-sided shift space $(\Omega, \mathcal{F}, \mathbb{P})$ over k symbols, as well as an RDS $\tilde{\varphi}$ over $(\Omega, \mathcal{F}, \mathbb{P}) \times \mathcal{X}$ given by $\tilde{\varphi}(1, \omega) = M(\omega)^{-1}$ and $\tilde{\varphi}(-1, \omega) =$

$M(\sigma^{-1}\omega)$ for $\omega \in \Omega$. Then we can define the skew product $\tilde{\Phi}$ on $\Omega \times \mathcal{X}$ by $\Phi^n(\omega, x) = (\sigma^n\omega, \tilde{\varphi}(n, \omega) \cdot x)$, for $n \in \mathbb{Z}$.

By [1] Theorem 1.7.3, there exists a one-to-one correspondence between invariant measures for one-sided and two-sided time. Therefore, there exists an invariant measure $\tilde{\nu}$ for $\tilde{\Phi}$ such that,

$$\nu_\omega = \mathbb{E}[\tilde{\nu}.|\mathcal{F}^+]_\omega,$$

where ν_ω denotes the part of the factorisation $\nu(\mathrm{d}\omega, \mathrm{d}x) = \nu_\omega(\mathrm{d}x)\mathbb{P}(\mathrm{d}\omega)$. Note $\tilde{\nu}$ is also an invariant measure for $\tilde{\Phi}^{-1}$ as $\tilde{\Phi}^{-1}\tilde{\nu} = \tilde{\Phi}^{-1}\tilde{\Phi}\tilde{\nu} = \tilde{\nu}$.

Now we have, for any $w \in G^+_{\theta,\alpha}$,

$$\begin{aligned}
\int w\, \nu(\mathrm{d}\omega, \mathrm{d}x) &= \mathbb{E}\left[\int w\tilde{\nu}(\mathrm{d}\omega, \mathrm{d}x)\middle|\mathcal{F}^+\right] \\
&= \mathbb{E}\left[\int w \circ \tilde{\Phi}^{-1}(\omega, x)\tilde{\nu}(\mathrm{d}\omega, \mathrm{d}x)\middle|\mathcal{F}^+\right] \\
&= \mathbb{E}\left[\int w(\sigma^{-1}\omega, M(\sigma^{-1}\omega) \cdot x)\tilde{\nu}(\mathrm{d}\omega, \mathrm{d}x)\middle|\mathcal{F}^+\right] \\
&= \sum_{i=1}^{k} p(\omega, i\omega)\mathbb{E}\left[\int w(i\omega, M_i \cdot x)\tilde{\nu}(\mathrm{d}\omega, \mathrm{d}x)\middle|\mathcal{F}^+\right] \\
&= \sum_{i=1}^{k} p(\omega, i\omega)\int w(i\omega, M_i \cdot x)\nu(\mathrm{d}\omega, \mathrm{d}x) \\
&= \int \mathcal{L}_0 w\, \nu(\mathrm{d}\omega, \mathrm{d}x).
\end{aligned}$$

Therefore (b) is proved.

Conversely, if (b) is satisfied. Let $w = f \circ \Phi$ in (6.31) for some $f \in G^+_{\theta,\alpha}$. As $\mathcal{L}_0(f \circ \Phi) = f$, we have

$$\int f\, \nu(\mathrm{d}\omega, \mathrm{d}x) = \int f \circ \Phi\, \nu(\mathrm{d}\omega, \mathrm{d}x).$$

That is to say, ν is invariant for Φ. □

In the end, we will give the $|\cdot|_{\theta,\alpha}$-bounds for the power $\mathcal{L}_g^n$ when n is sufficiently large in Theorem 6.2.10 and Corollary 6.2.11. This type of bounds is also known as **Lasota-Yorke inequalities** [34]. The importance of such inequalities for proving quasi-compactness (stated

below in Theorem 6.2.12)[1] — and hence the crucial fact (stated below in Theorem 6.2.17) was emphasized by [26]. Applications to transfer operators can also be found in [61].

We start with a lemma which is an easy corollary of Proposition 4.3.8.

Lemma 6.2.8. *Let $(\Omega^+, \mathcal{F}^+, \mathbb{P})$ be a one-sided Markovian shift space over k symbols, and $M(\omega) : \Omega^+ \to GL(d, \mathbb{R})$ depend only on the first coordinate ω_0. Assume the matrix set $\{M(\omega) : \omega \in \Omega^+\}$ is strongly irreducible and contracting, then there exists $0 < \delta < 1$ such that when n is sufficiently large, we have*

$$\sup_{x \neq y} \mathbb{E}[d_{\mathcal{X}}^{\alpha}(\psi(n, \omega) \cdot x, \psi(n, \omega) \cdot y)] < \delta^n d_{\mathcal{X}}^{\alpha}(x, y),$$

where $\psi(n, \omega)$ is given by (6.26)*.*

Remark 6.2.9. We shall notice that strong irreducibility and contracting property are both algebraic properties of a matrix set, which does not depend on the choice of the probability measure on the sample space. However, in the proof of Proposition 2.3.15, the choice of α_0 depends on the probability measure. We need $\alpha_0 < \beta$ for which $\mathbb{E}[e^{\beta \ell(M(\omega))}] < \infty$. Fortunately, the latter condition follows automatically for any $\beta > 0$ as in our setting the cardinality of the matrix set is finite.

Theorem 6.2.10. *Let $(\Omega^+, \mathcal{F}^+, \mathbb{P})$ be a one-sided Markovian shift space over k symbols, and $M(\omega) : \Omega^+ \to GL(d, \mathbb{R})$ depend only on the first coordinate ω_0. Assume the matrix set $\{M(\omega) : \omega \in \Omega^+\}$ is strongly irreducible and contracting, and let $\mathcal{L}_g$ be a normalised weighted Markovian transfer operators defined by* (6.28) *satisfying that the function $g(\cdot, x)$ on Ω^+ depends only on the first coordinate ω_0. Denote*

$$G(\omega^{(n)}, x) := \sum_{i=1}^{n} g\left(\omega^{(i)}, M(\omega^{(i-1)}) \cdots M(\omega') \cdot x\right),$$

and assume

$$\left| e^{G(\omega^{(n)}, x)} - e^{G(\omega^{(n)}, y)} \right| \leq H(n) d_{\mathcal{X}}^{\alpha}(x, y), \tag{6.32}$$

for $H(n)$ depending only on n. Then there exist constants $\theta_0 > 0$ and $0 < \delta_0 < 1$, such that when $0 < \theta < \theta_0$ and n is sufficiently large, we have

$$|\mathcal{L}_g^n w|_{\theta, \alpha} \leq (H(n) + 2|e^g|_{\infty}^n)|w|_{\infty} + \delta_0^n |w|_{\theta, \alpha}.$$

[1] **Quasi-compactness**: a bounded linear operator T acting on a Banach space is quasi-compact if $\|T^n - K\| < \rho(T)^n$ for some $n > 0$ and some compact linear operator $K \neq 0$.

Proof. For $\omega, \tilde{\omega} \in \Omega^+$ and $x \in \mathcal{X}$, as the function $g(\cdot, x)$ on Ω^+ only depends on the first coordinate ω_0, we have

$$
\begin{aligned}
&|(\mathcal{L}_g^n w)(\omega, x) - (\mathcal{L}_g^n w)(\tilde{\omega}, x)| \\
\leq &|e^G|_\infty \left| \sum_{\sigma^n \omega^{(n)} = \omega} P(n, \omega^{(n)}) w(\omega^{(n)}, \psi(n, \omega^{(n)}) \cdot x) - \sum_{\sigma^n \tilde{\omega}^{(n)} = \tilde{\omega}} P(n, \tilde{\omega}^{(n)}) w(\tilde{\omega}^{(n)}, \psi(n, \tilde{\omega}^{(n)}) \cdot x) \right|.
\end{aligned}
$$

Note each $\omega^{(n)}$ satisfying $\sigma^n \omega^{(n)} = \omega$ is of the form $(i_n \dots i_1 \omega)$ for $1 \leq i_s \leq k$ $(1 \leq s \leq n)$. For each $\omega^{(n)}, \tilde{\omega}^{(n)}$ of the form $\omega^{(n)} = (i_n \dots i_1 \omega), \tilde{\omega}^{(n)} = (i_n \dots i_1 \tilde{\omega})$, by (6.25) and (6.26), we have

$$
\begin{aligned}
\psi(n, \omega^{(n)}) &= \psi(n, \tilde{\omega}^{(n)}), \\
P(n, \omega^{(n)}) &= P(n-1, \omega^{(n)}) p(\omega, i_1 \omega), \\
P(n, \tilde{\omega}^{(n)}) &= P(n-1, \omega^{(n)}) p(\tilde{\omega}, i_1 \tilde{\omega})
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&|(\mathcal{L}_g^n w)(\omega, x) - (\mathcal{L}_g^n w)(\tilde{\omega}, x)| \\
\leq &|e^G|_\infty \sum_{i=1}^k \sum_{\sigma^{n-1} \omega^{(n)} = (i\omega)} P(n-1, \omega^{(n)}) |p(\omega, i\omega) - p(\tilde{\omega}, i\tilde{\omega})| |w|_\infty \qquad (6.33) \\
&+ |e^G|_\infty \left| w(\omega^{(n)}, \psi(n, \omega^{(n)}) \cdot x) - w(\tilde{\omega}^{(n)}, \psi(n, \omega^{(n)}) \cdot x) \right| \qquad (6.34) \\
\leq &|e^g|_\infty^n (2|w|_\infty + \theta^n |w|_{\theta,\alpha}) d_\theta(\omega, \tilde{\omega}),
\end{aligned}
$$

since $\sum_{i=1}^k |p(\omega, i\omega) - p(\tilde{\omega}, i\tilde{\omega})| \leq 2 d_\theta(\omega, \tilde{\omega})$ and $\sum_{\sigma^{n-1} \omega^{(n)} = (i\omega)} P(n-1, \omega^{(n)}) = 1$.

Then fix $\omega \in \Omega^+$ and $x, y \in \mathcal{X}$, we have

$$
\begin{aligned}
&|(\mathcal{L}_g^n w)(\omega, x) - (\mathcal{L}_g^n w)(\omega, y)| \\
\leq &\sum_{\sigma^n \omega^{(n)} = \omega} P(n, \omega^{(n)}) \left| e^{G(\omega^{(n)}, x)} - e^{G(\omega^{(n)}, y)} \right| |w|_\infty \\
&+ \sum_{\sigma^n \omega^{(n)} = \omega} \left| w(\omega^{(n)}, \psi(n, \omega^{(n)}) \cdot x) - w(\omega^{(n)}, \psi(n, \omega^{(n)}) \cdot y) \right| \left| P(n, \omega^{(n)}) e^{G(\omega^{(n)}, x)} \right| \\
\leq &H(n) |w|_\infty d_{\mathcal{X}}^\alpha(x, y) + |w|_{\theta,\alpha} \mathbb{E}_{pe^g} \left[d_{\mathcal{X}}^\alpha \left(\psi(n, \omega^{(n)}) \cdot x, \psi(n, \omega^{(n)}) \cdot y \right) \right] \\
\leq &\left(H(n) |w|_\infty + \delta^n |w|_{\theta,\alpha} \right) d_{\mathcal{X}}^\alpha(x, y),
\end{aligned}
$$

by Lemma 6.2.8 and Remark 6.2.9, where $\mathbb{E}_{pe^g}$ denotes the expectation with respect to the probability measure given by $p(\omega,\omega')e^{g(\omega,x)}$, note $\mathcal{L}_g$ being normalised implies that

$$1 = \mathcal{L}_g\mathbf{1} = \sum_{\sigma\omega'=\omega} p(\omega,\omega')e^{g(\omega',M(\omega')\cdot x)}.$$

Therefore when n is sufficiently large we would have

$$|\mathcal{L}_g^n w|_{\theta,\alpha} \leq (H(n) + 2|e^g|_\infty^n)|w|_\infty + (\theta^n|e^g|_\infty^n + \delta^n)|w|_{\theta,\alpha}.$$

Now choose $\theta_0 = \delta/|e^g|_\infty$, then when $0 < \theta < \theta_0$, we have

$$\theta^n|e^g|_\infty^n + \delta^n < 2\delta^n < \delta_0^n$$

for some $0 < \delta_0 < 1$ when n is sufficiently large. □

Corollary 6.2.11. *Let $\mathcal{L}_g$ be a normalised Markovian transfer operator satisfying the assumptions of Theorem 6.2.10, then there exists $\theta_0 > 0$ such that when $0 < \theta < \theta_0$, there exists an increasing sequence $(n_k)_{k\geq 1}$ of integers and $0 < \delta < 1$, such that*

$$|\mathcal{L}_g^{n_k} w|_{\theta,\alpha} \leq C|w|_\infty + \delta^{n_k}|w|_{\theta,\alpha},$$

for any $w \in G^+_{\theta,\alpha}$, where $C > 0$ is a constant.

Proof. By Theorem 6.2.10, there exists some fixed $m_0 \geq 1$ such that

$$|\mathcal{L}_g^{m_0} w|_{\theta,\alpha} \leq C_{m_0}|w|_\infty + \delta^{m_0}|w|_{\theta,\alpha}$$

where $C_{m_0} = H(m_0) + 2|e^g|_\infty^{m_0} > 0$. Now for any integer $m \geq 1$ satisfying the conclusion of Theorem 6.2.10 we have

$$|\mathcal{L}_g^{m+m_0} w|_{\theta,\alpha} \leq C_m|\mathcal{L}_g^{m_0} w|_\infty + \delta^m|\mathcal{L}_g^{m_0} w|_{\theta,\alpha} \leq C_m|w|_\infty + \delta^m C_{m_0}|w|_\infty + \delta^{m+m_0}|w|_{\theta,\alpha}.$$

Therefore we can take $C_{m+m_0} = C_m + \delta^m C_{m_0}$. Thus

$$C_{km_0} = C_{m_0}(1 + \delta^{m_0} + \cdots + \delta^{(k-1)m_0}) < C_{m_0}/(1 - \delta^{m_0}).$$

That is to say, if we take $n_k = km_0$, then

$$|\mathcal{L}_g^{km_0} w|_{\theta,\alpha} \leq \frac{C_{m_0}}{1-\delta^{m_0}}|w|_\infty + \delta^{km_0}|w|_{\theta,\alpha}.$$

□

6.2.3 Spectral Properties

We start with the spectral properties of $\mathcal{L}_t$ when $t = 0$. Here we define an operator $\mathcal{Q}$ on $G^+_{\theta,\alpha}$ by

$$(\mathcal{Q}w)(\omega, x) = \int w \ \mathrm{d}\nu, \tag{6.35}$$

where ν is the invariant measure for the RDS φ characterised by $\varphi(1, \omega) = M(\omega)^{-1}$. By Theorem 6.2.7, we know ν is also an eigenmeasure of $\mathcal{L}_0$.

Theorem 6.2.12. *Let $(\Omega^+, \mathcal{F}^+, \mathbb{P})$ be a one-sided Markovian shift space over k symbols, and $M(\omega) : \Omega^+ \to GL(d, \mathbb{R})$ depend only on the first coordinate ω_0. Assume the matrix set $\{M(\omega) : \omega \in \Omega^+\}$ is strongly irreducible and contracting. Let $\mathcal{L}_0$ be the parametrised Markovian transfer operator $\mathcal{L}_t$ defined by* (6.22) *acting on $G^+_{\theta,\alpha}$ when $t = 0$ and $\mathcal{Q}$ be defined as in* (6.35)*. Then there exists $\theta_0 > 0$ such that when $0 < \theta < \theta_0$ and n sufficiently large, we have*

$$\|\mathcal{L}_0^n - \mathcal{Q}\|^{1/n}_{\theta,\alpha} < 1.$$

Proof. Note $\mathcal{L}_0$ equals the weighted transfer operator $\mathcal{L}_g$ when $g = 0$. Therefore the condition (6.32) is automatically satisfied with $H(n) = 0$. Fix $w \in G^+_{\theta,\alpha}$, then by Theorem 6.2.10 we know there exists $\theta_0 > 0$ and $0 < \delta < 1$ such that when n is sufficiently large, we have

$$|\mathcal{L}_0^n w|_{\theta,\alpha} \le 2|w|_\infty + \delta^n |w|_{\theta,\alpha}. \tag{6.36}$$

This shows that $\{\mathcal{L}_0^n w\}$ is equicontinuous. Therefore by the Arzelà-Ascoli theorem, $\{\mathcal{L}_0^n w : n \ge 1\}$ is relatively compact with respect to the uniform topology. Assume w^* is the limit point of a subsequence $\{\mathcal{L}_0^{n_j} w : j \ge 1\}$.

Note for $\omega, \tilde{\omega} \in \Omega^+, x, y \in \mathcal{X}$ with $\omega_0 = \tilde{\omega}_0$, by the same proof as that of Theorem 6.2.10, we have

$$|\mathcal{L}_0^n(\omega, x) - \mathcal{L}_0^n(\tilde{\omega}, y)| \le \theta^n |w|_{\theta,\alpha} d_\theta(\omega, \tilde{\omega}) + \delta^n |w|_{\theta,\alpha} d^\alpha_{\mathcal{X}}(x, y),$$

by noticing the term (6.33) vanishes because $\omega_0 = \tilde{\omega}_0$. Letting $n = n_j \to 0$ $(j \to \infty)$ shows that $w^*(\omega, x)$ depends only on ω_0. Denote $w^*(\omega, x) = w^*_{\omega_0}$.

As $\mathcal{L}_0$ is normalised,

$$\sup w \ge \sup \mathcal{L}_0 w \ge \cdots \ge \sup \mathcal{L}_0^n w \ge \ldots$$

This implies that $\sup w^* = \sup \mathcal{L}_0 w^*$. Since $\Omega^+ \times \mathcal{X}$ is compact, we assume $\sup w^*$ is attained at $(\omega, x) \in \Omega^+ \times \mathcal{X}$ and $\sup \mathcal{L}_0 w^*$ at $(\tilde{\omega}, y)$, thus

$$w^*(\omega, x) = (\mathcal{L}_0 w^*)(\tilde{\omega}, y) = \sum_{i=1}^{k} p(\tilde{\omega}, i\tilde{\omega}) w^*(i\tilde{\omega}, M_i \cdot y).$$

As w^* only depends on ω_0, we have

$$\sup w^* = \sum_{i=1}^{k} p(\tilde{\omega}, i\tilde{\omega}) w_i^* \leq \sup w^*.$$

This shows that $w_i^* = \sup w^*$ for each $1 \leq i \leq k$, and w^* is a constant function on $\Omega^+ \times \mathcal{X}$.

Moreover, as ν is an invariant measure, we have $\mathcal{Q}\mathcal{L}_0 = \mathcal{Q}$. This shows that

$$w^* = \mathcal{Q}w^* = \lim_{j \to \infty} \mathcal{Q}\mathcal{L}_0^{n_j} w = \mathcal{Q}w.$$

By applying the same arguments to each subsequence of $\{\mathcal{L}_0^n w : n \geq 1\}$, we can obtain a further subsequence converging to the same limit $\mathcal{Q}w$. This shows that as $n \to \infty$, $\mathcal{L}_0^n w$ uniformly converges to $\mathcal{Q}w$.

Now for $n, m \geq 0$, by (6.36) we have

$$\begin{aligned}
|(\mathcal{L}_0^{n+m} - \mathcal{Q})(w)|_{\theta,\alpha} &= |\mathcal{L}_0^n (\mathcal{L}_0^m - \mathcal{Q})(w)|_{\theta,\alpha} \\
&\leq 2|(\mathcal{L}_0^m - \mathcal{Q})(w)|_\infty + \delta^n |(\mathcal{L}_0^m - \mathcal{Q})(w)|_{\theta,\alpha} \\
&= 2|(\mathcal{L}_0^m - \mathcal{Q})(w)|_\infty + \delta^n |\mathcal{L}_0^m w|_{\theta,\alpha} \qquad (6.37) \\
&\leq 2|(\mathcal{L}_0^m - \mathcal{Q})(w)|_\infty + 2\delta^n |w|_\infty + \delta^{n+m} |w|_{\theta,\alpha}, \qquad (6.38)
\end{aligned}$$

where the equality in (6.37) follows from

$$\left|(\mathcal{L}_0^n w - \mathcal{Q}w)(\omega, x) - (\mathcal{L}_0^n w - \mathcal{Q}w)(\tilde{\omega}, y)\right| = |(\mathcal{L}_0^n w)(\omega, x) - (\mathcal{L}_0^n w)(\tilde{\omega}, y)|,$$

which implies $|\mathcal{L}_0^n w - \mathcal{Q}w|_{\theta,\alpha} = |\mathcal{L}_0^n w|_{\theta,\alpha}$.

Define $\Gamma = \{w \in G_{\theta,\alpha}^+ : \|w\|_{\theta,\alpha} \leq 1\}$, which is compact in the uniform topology by the Arzelà-Ascoli theorem. Therefore for any $\epsilon > 0$, there exists an integer M (not depending on $w \in \Gamma$), such that when $m > M$, we have $|\mathcal{L}_0^m w - \mathcal{Q}w|_\infty < \epsilon$. Choose an integer $N > 0$ such that when $n > N$, $\delta^n < \epsilon$. Then for any $w \in \Gamma$, by (6.38), we have

$$\|(\mathcal{L}_0^{n+m} - \mathcal{Q})(w)\|_{\theta,\alpha} = |(\mathcal{L}_0^{n+m} - \mathcal{Q})(w)|_\infty + |(\mathcal{L}_0^{n+m} - \mathcal{Q})(w)|_{\theta,\alpha}$$

$$\leq \epsilon + (2\epsilon + 2\epsilon|w|_\infty + \epsilon|w|_{\theta,\alpha}) \leq 6\epsilon.$$

Now the conclusion follows by choosing $\epsilon < 1/6$.

□

Remark 6.2.13. $\mathcal{L}_0$ obviously possesses an eigenvalue 1. This theorem suggests that 1 is in fact a simple maximal top eigenvalue. More precisely, $\mathcal{L}_0$ can be written as $\mathcal{Q} + \mathcal{R}$ where $\mathcal{Q}$ is the one-dimensional projection given by (6.35) and $\mathcal{R} = \mathcal{L}_0 - \mathcal{Q}$ has spectral radius strictly smaller than 1 with $\mathcal{Q}\mathcal{R} = \mathcal{R}\mathcal{Q} = 0$, since the spectral radius of $\mathcal{R}$ is given by $\inf\{\|\mathcal{R}^n\|_{\theta,\alpha}^{1/n} : n \geq 1\}$ and $\mathcal{R}^n = \mathcal{L}_0^n - \mathcal{Q}$.

We now turn to the corresponding Lasota-Yorke inequality for $\mathcal{L}_t$.

Lemma 6.2.14. *For any $n \geq 1, \omega \in \Omega^+$, we have $\ell(\psi(n,\omega)) \leq nK$, where $K := \max_{1\leq i\leq k} \ell(M_i)$.*

Proof. We notice that for any $A, B \in GL(d,\mathbb{R})$, $\ell(AB) \leq \ell(A) + \ell(B)$ follows immediately from the following two inequalities.

$$\log^+ \|AB\| \leq \log^+ \|A\| + \log^+ \|B\| \leq \ell(A) + \ell(B),$$

$$\log^+ \|(AB)^{-1}\| \leq \log^+ \|B^{-1}\| + \log^+ \|A^{-1}\| \leq \ell(A) + \ell(B).$$

□

Theorem 6.2.15. *Let $(\Omega^+, \mathcal{F}^+, \mathbb{P})$ be a one-sided Markovian shift space over k symbols, and $M(\omega) : \Omega^+ \to GL(d,\mathbb{R})$ depend only on the first coordinate ω_0. Assume the matrix set $\{M(\omega) : \omega \in \Omega^+\}$ is strongly irreducible and contracting. Let t be purely imaginary and $\mathcal{L}_t$ be the parametrised Makovian transfer operators defined by* (6.22) *acting on $G_{\theta,\alpha}^+$. Then there exists $\theta_0 > 0$ such that when $0 < \theta < \theta_0$, there exists an increasing sequence $(n_k)_{k\geq 1}$, such that*

$$|\mathcal{L}_t^{n_k} w|_{\theta,\alpha} \leq C|w|_\infty + \delta_0^{n_k}|w|_{\theta,\alpha}$$

for constants $C > 0$, $0 < \delta_0 < 1$ and any $k \geq 1$.

Proof. Notice first that when t is purely imaginary, $\mathcal{L}_t$ is normalised. Now by Corollary 6.2.11, we only need to verify inequality (6.32). Note $\mathcal{L}_t$ equals $\mathcal{L}_g$ when

$$g(\omega, x) = -t \log \|M(\omega)^{-1}x\|$$

by the discussion after Theorem 6.2.3. Since $M(\omega)$ only depends on the first coordinate ω_0, we know that g (as a function of Ω^+) also only depends on the first coordinate. Moreover,

$$G(\omega^{(n)}, x) = t \log \|M(\omega^{(n)}) \cdots M(\omega')x\|.$$

Then by Lemma 6.2.4 (b), we have

$$\begin{aligned}|e^{G(\omega^{(n)},x)} - e^{G(\omega^{(n)},y)}| &\leq c_2 e^{[(1+\alpha)|\operatorname{Re} t|+2\alpha]\ell(\psi(n,\omega^{(n)}))} d_{\mathcal{X}}^{\alpha}(x,y) \\ &\leq c_2 e^{[(1+\alpha)|\operatorname{Re} t|+2\alpha]nK} d_{\mathcal{X}}^{\alpha}(x,y),\end{aligned}$$

by Lemma 6.2.14. Therefore (6.32) is satisfied with $H(n) = c_2 e^{2\alpha nK}$ as $\operatorname{Re} t = 0$. □

Now we give the following perturbation theorem from [32] and apply it to conclude our final result Theorem 6.2.17 of this chapter.

Theorem 6.2.16. *Let $(B, \|\cdot\|)$ be a Banach space with a second norm $|\cdot|$ such that $|\cdot| \leq \|\cdot\|$ (we do not require $(B, |\cdot|)$ to be complete). Let $\mathcal{P}_\epsilon : B \to B$ be a family of bounded linear operators for some parameter ϵ in a set containing 0. Suppose*

(a) the inclusion $\iota : (B, \|\cdot\|) \hookrightarrow (B, |\cdot|)$ is compact;

(b) there exists a set $\mathcal{E}$ containing 0, for which $\mathcal{P}_\epsilon$ $(\epsilon \in \mathcal{E})$ satisfies the uniform Lasota-Yorke inequality, i.e., there exists $n_0 > 0, 0 < \delta < 1$ and $C > 0$ such that

$$\|\mathcal{P}_\epsilon^{n_0} w\| \leq \delta \|w\| + C|w|, \quad \forall \epsilon \in \mathcal{E};$$

(c) there is a monotone upper semi-continuous function ψ such that $\psi(\epsilon) \to 0$ as $\epsilon \to 0$ and $\sup_{w \in B} |(\mathcal{P}_\epsilon - \mathcal{P}_0)w|/\|w\| \leq \psi(\epsilon)$.

Suppose $\mathcal{P}_0$ has a simple maximal eigenvalue 1 with corresponding eigenprojection ν. Then there exists $\epsilon_0 > 0$ such that when $|\epsilon| < \epsilon_0$ in $\mathcal{E}$, $\mathcal{P}_\epsilon$ has a simple maximal eigenvalue with a corresponding eigenprojection ν_ϵ.

Theorem 6.2.17. *Let $\mathcal{L}_t$ be the parametrised transfer operator on $G_{\theta,\alpha}^+$ defined as in* (6.22) *with t purely imaginary. Then $\mathcal{L}_t$ can be decomposed as*

$$\mathcal{L}_t = \beta(t)\mathcal{Q}(t) + \mathcal{R}(t),$$

where $\mathcal{Q}(t)$ is a one-dimensional projection, $\mathcal{R}(t)$ has the spectral radius strictly smaller than $\beta(t)$ with $\mathcal{Q}(t)\mathcal{R}(t) = \mathcal{R}(t)\mathcal{Q}(t) = 0$. And $\beta(0) = 1, \mathcal{Q}(0) = \mathcal{Q}$ given by (6.35).

Proof. We apply Theorem 6.2.16 with $|\cdot| = |\cdot|_\infty$ and $\|\cdot\| = \|\cdot\|_{\theta,\alpha}$. Now we verify the three conditions.

For (a), assume $\{w_n\}$ to be a bounded sequence in $(G_{\theta,\alpha}^+, \|\cdot\|_{\theta,\alpha})$ with $\|w\|_{\theta,\alpha} \leq 1$, then they are uniformly bounded (since $|w_n|_\infty \leq 1$) and equicontinuous (since $|w_n|_{\theta,\alpha} \leq 1$). Then by the Arzelà-Ascoli theorem, $\{w_n\}$ is relatively compact in the uniform topology.

(b) is satisfied by Theorem 6.2.15. Now since for any $h \in \mathbb{R}$, we have

$$|e^{ih} - 1|^2 = 2 - 2\cos h = 4\sin^2(h/2) \leq h^2,$$

therefore when t is purely imaginary,

$$|\mathcal{L}_t w - \mathcal{L}_0 w| \leq |t| K |w|_\infty \leq |t| K \|w\|_{\theta,\alpha}.$$

Now if we take $\psi(t) = |t|K$. It is obvious that $\psi(t)$ is monotone upper semi-continuous, and $\psi(t) \to 0$ as $|t| \to 0$.

The fact that $\mathcal{L}_0$ has a simple maximal eigenvalue 1 is guaranteed by Theorem 6.2.12. Therefore the results follows by applying Theorem 6.2.16. □

Remark 6.2.18. As before, by the general perturbation theory on [43] Section VII.6 and (6.27), we can deduce that $\beta(t)$ is analytic and that $\beta'(0)$ equals the top Lyapunov exponent associated to this problem of Markovian random matrix products.

Chapter 7

Algebras of Transfer Operators

7.1 Overview

We have seen in Chapter 5 that, in order to compute the Lyapunov exponent for products of i.i.d. matrices, we can define the corresponding transfer operators on some function space $\mathcal{G}$ with its form

$$\mathcal{L}_t f = \sum_{i=1}^{n} p_i e^{t \log \|M_i x\|} f(M_i \cdot x)$$

for any $f \in \mathcal{G}$, where $\{M_i\}_{1 \leq i \leq k}$ is the matrix set, p_i the probability to choose i-th matrix. If $\{M_i\}_{1 \leq i \leq k}$ are all positive, by setting $\mathcal{G} = \mathcal{A}_\infty(U)$, $\mathcal{L}_t$ is nuclear of order zero, and therefore we can use the trace-determinant formula

$$\det(\mathcal{I} - z\mathcal{L}_t) = \exp\left(-\sum_{n=1}^{\infty} \frac{\operatorname{tr} \mathcal{L}_t^n}{n} z^n\right) \tag{7.1}$$

to compute the top eigenvalue $\beta(t)$ of $\mathcal{L}_t$ when $|t|$ is sufficiently small, and the Lyapunov exponent λ can be obtained by $\lambda = \beta'(0)$.

In Section 6.2, we see that we can generalise the algorithm of Pollicott by assuming the matrix set is strongly irreducible and contracting. We show that we still have Theorem 6.2.17 so that the Lyapunov exponent still can be computed by finding the top eigenvalue of $\mathcal{L}_t$. However, there is a problem in generalising (7.1) as in most of the cases the trace of $\mathcal{L}_t$ can not be computed in the same way.

For example, let A be a $d \times d$ positive matrix, let $\hat{A}$ be the map defined in Section 5.1, let $\lambda_1, \ldots, \lambda_d$ be the eigenvalues of A (counting multiplicity) $\lambda_1 > |\lambda_2| \geq \cdots \geq |\lambda_d|$, x_1 be the normalised eigenvector corresponding to λ_1 and $D\hat{A}(x_1)$ be the differential of $\hat{A}$ at x_1. Then by

Lemma 5.8.1

$$\det\left(I - D\hat{A}(x_1)\right) = \prod_{i=2}^{d}\left(1 - \frac{\lambda_i}{\lambda_1}\right).$$

And if we denote by $\mathcal{L}_A$ the operator $(\mathcal{L}_A w)(x) = w(\hat{A}x)$ on $\mathcal{A}_\infty(U)$, where U is an open domain chosen according to Lemma 5.1.3 on which $\tilde{A}$ is analytic, then we get the trace formula by Corollary 5.5.5

$$\mathrm{tr}(\mathcal{L}_A) = \frac{1}{\det(I - D\hat{A}(x_1))}.$$

We can notice that the validity of this formula relies on the fact that λ_1 is strictly greater than any other eigenvalues (which is guaranteed in the positive matrix setting), which makes it hard to generalise Pollicott's algorithm even to non-negative matrices.

For example, let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ which has an eigenvalue 1 with multiplicity 2, we follow the same procedure as in the proof of Lemma 5.8.1 and choose a coordinate system $x(t) = (1 \quad t)^T$ around the eigenvector $x_1 = (1 \quad 0)^T$ then when t is sufficiently close to 0, we have

$$\hat{A}x(t) = \begin{pmatrix} 1+t \\ t \end{pmatrix} \sim \begin{pmatrix} 1 \\ t/(1+t) \end{pmatrix},$$

which gives $D\hat{A}(x_1) = D\hat{A}|_{t=0} = 1$ and $\det(I - D\hat{A}) = 0$.

We might also expect to generalise the algorithm to the matrix set $M = \{A_1, \cdots, A_n\}$ where $A_1, \ldots, A_n$ all have a dominant eigenvalue. However, we also need to compute the traces of $\mathrm{tr}(\mathcal{L}_{A_i A_j}), \mathrm{tr}(\mathcal{L}_{A_i A_j A_k}), \ldots$, and the property of "possessing a dominant eigenvalue" in general cannot be preserved under multiplication. A trivial example would be

$$A_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \text{but} \quad A_1 A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

A less trivial one could be obtained by a simple modification

$$A_1 = \begin{pmatrix} 2 & 1 \\ 0 & 1/2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1/2 & -1 \\ 0 & 2 \end{pmatrix}, \quad \text{but} \quad A_1 A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus we cannot simply apply the old formulas and expect to generalise automatically.

However, a possible method arises if we notice the trace of an operator can be defined abstractly. We can borrow the definition from theory of operator algebras, a trace ϕ on a C*-algebra $\mathcal{A}$ is a linear map to $\mathbb{R}_{\geq 0}$ defined on the positive part $\mathcal{A}^+$ (and on the whole $\mathcal{A}$ when ϕ is finite at the unit of $\mathcal{A}$), which satisfies: $\phi(aa^*) = \phi(a^*a)$ for any $a \in \mathcal{A}$.

According to [60] it was firstly proved by von Neumann that any operator algebras on a separable Hilbert space can be decomposed into the direct sum of its building blocks called factors. In [41], Murray and von Neumann showed that there are three types of factors, designated type I, type II and type III factors.

Type I factors are those which possess a minimal projection. For example, in $M_d(\mathbb{R})$, each elementary matrix $E_{ii}, 1 \leq i \leq d$ is a minimal projection. By defining the trace on the minimal projection to be 1, we can define the trace for the whole factor. Our usual definition for a trace is of this type.

Type II and type III factors do not possess a minimal projection. Murray and von Neumann constructed a concrete example of a type II factor in [41], and in [59] they proved that there always exists a unique (up to a scaling constant) faithful trace on a type II factor. If the trace is finite the factor is called a type II_1 factor, otherwise it is a type II_∞ factor.

For type III factors, although it was already defined in the first paper of the series by Murray and von Neumann in 1936 [41], its existence was first proved by von Neumann in 1940 [59]. As there is no finite trace on a type III factor, we are not particularly interested in it in this thesis.

In this chapter, we borrow the new definition of trace to apply to transfer operators. Before doing so, we introduce the necessary terminologies and tools for a C*-algebra (Section 7.2), and those for a more specific kind of C*-algebra, namely the von Neumann algebras (Section 7.3). Then we give some intuitive description of type I factors (Section 7.4), type II_1 factors (Section 7.5) and how to define traces on them. In Section 7.6, we introduce a widely used construction for von Neumann algebras introduced by von Neumann in [41], the crossed products, and explain how to identify its type.

Our new work starts in Section 7.7. After reviewing some properties of a transfer operator, defining a basic transfer operator and a multiplication operator, we consider the operator algebra generated by transfer operators and prove that they can form a von Neumann algebra, which turns out to be a crossed product. In Section 7.8, we borrow the definition of a determinant function Δ from [16] and prove that for an operator a in a von Neumann algebra, the definition of $d(\lambda) = \Delta(a - \lambda e)$ can be extended to the top eigenvalue of a and we give a series expansion for $d(\lambda)$.

Finally in Section 7.9, we discuss the limitations of this method, as well as prospects for further development.

Throughout this chapter, we will denote the algebraic dual of the Banach space B by B^*,

and the topological dual of B by B'.

7.2 C*-Algebras

7.2.1 Definition of C*-Algebras

Definition 7.2.1. Let $\mathcal{A}$ be an associative algebra (over the complex numbers) equipped with norm $\|\cdot\|$. We say $\mathcal{A}$ is a **Banach algebra** if it is complete and satisfies $\|ab\| \leq \|a\|\|b\|$ for any $a, b \in \mathcal{A}$.

Moreover, it is called a **Banach involutive algebra** (or **Banach *-algebra**) if there exists a map $* : \mathcal{A} \to \mathcal{A}$ such that for any $a, b \in \mathcal{A}, \lambda, \mu \in \mathbb{C}$

(1) $(\lambda a + \mu b)^* = \bar{\lambda} a^* + \bar{\mu} b^*$;

(2) $a^{**} = a$;

(3) $(ab)^* = b^* a^*$.

Such a map $*$ is called the **involution**.

A **C*-algebra** is a Banach *-algebra which satisfies $\|a^* a\| = \|a\|^2$ for any $a \in \mathcal{A}$.

Example 4. A typical example of a C*-algebra is the algebra $B(\mathcal{H})$ of bounded operators on a Hilbert space $\mathcal{H}$. The norm is induced by the inner product of $\mathcal{H}$, i.e. for $\mathcal{L} \in B(\mathcal{H})$

$$\|\mathcal{L}\| = \sup_{0 \neq x \in \mathcal{H}} \frac{\langle \mathcal{L}x, x \rangle}{\langle x, x \rangle}.$$

The adjoint map is defined by the usual Hermitian adjoint, i.e. it is uniquely determined by the equation $\langle \mathcal{L}^* x, y \rangle = \langle x, \mathcal{L} y \rangle, \forall x, y \in \mathcal{H}$. It is not hard to verify the axioms of the adjoint map and the fact that $\|\mathcal{L}^*\| = \|\mathcal{L}\|$. Moreover, the inequality

$$\|\mathcal{L}x\|^2 = \langle \mathcal{L}^* \mathcal{L} x, x \rangle \leq \|\mathcal{L}^* \mathcal{L}\| \|x\|^2 \leq \|\mathcal{L}\|^2 \|x\|^2$$

implies that $\|\mathcal{L}^* \mathcal{L}\| = \|\mathcal{L}\|^2$. Thus $B(\mathcal{H})$ is a C*-algebra.

Example 5. Given a C*-algebra $\mathcal{A}$, we say $e \in \mathcal{A}$ is an **identity** (or a **unit**) if for any $a \in \mathcal{A}$, we have $ae = ea = a$. A unit is crucial in defining invertibility, spectrum and so on. However, there does not always exist a unit in a C*-algebra. For example, let $\mathcal{A} = C_0((0, 1))$, the set of complex-valued continuous functions on the open interval $(0, 1)$ which vanish at the boundary. Equip $\mathcal{A}$ with the supremum norm and define the involution to be the complex conjugation. We can see that $\mathcal{A}$ is actually a C*-algebra. But there does not exist a unit in $\mathcal{A}$.

We say C*-algebra $\mathcal{A}$ is **unital** if there exists a unit in $\mathcal{A}$. However, we can always add a unit to $\mathcal{A}$ if it does not possess one. Define $\mathcal{A}_e = \{a + \lambda e : a \in \mathcal{A}, \lambda \in \mathbb{C}\}$, norm on $\mathcal{A}_e$ to be $\|a + \lambda e\| = \|a\| + |\lambda|$, involution to be $(a + \lambda e)^* = a^* + \bar{\lambda} e$ and multiplication to be $(a + \lambda e)(b + \mu e) = ab + \lambda b + \mu a + \lambda\mu$. Then we can show that $\mathcal{A}_e$ is a unital C*-algebra (See the proof in e.g. [52] 1.1.7). We call this process **unification**.

In our example of $\mathcal{A} = C_0((0,1))$, we can let $e(x)$ be the constant function on $(0,1)$. Note for $a(x) \in \mathcal{A}$, we must have $\lim_{x\to 0} a(x) = \lim_{x\to 1} a(x) = 0$, while $\lim_{x\to 0} e(x) = \lim_{x\to 1} e(x) = 1$, i.e. $a + \lambda e$ must reach the same value on the boundary. Therefore after unification, we can get $\mathcal{A}_e$ is the set of continuous functions on $[0,1]/0 \sim 1 \simeq S^1$. $\mathcal{A}_e = C(S^1)$ is obviously a unital C*-algebra.

Now we can define the spectrum of an element in a C*-algebra. Without loss of generality, we assume here $\mathcal{A}$ is a C*-algebra with a unit e. Then we say an element $a \in \mathcal{A}$ is **invertible** if there exists $b \in \mathcal{A}$ such that $ab = ba = e$. And for $a \in \mathcal{A}$, the **spectrum** of a in $\mathcal{A}$ is the set

$$\operatorname{spec}_{\mathcal{A}} a = \{\lambda \in \mathbb{C} : a - \lambda e \text{ is not invertible in } \mathcal{A}\}.$$

When $\mathcal{A}$ does not have a unit, define $\operatorname{spec}_{\mathcal{A}} a = \operatorname{spec}_{\mathcal{A}_e} a$.

We remark here, the spectrum of an element actually does not depend on the choice of the algebra $\mathcal{A}$ in which it lies as long as $\mathcal{A}$ is unital. This result can be found in [31]. So in the following, we will write $\operatorname{spec} a$ for short.

The **spectral radius** $\rho(a)$ of $a \in \mathcal{A}$ is defined to be

$$\rho(a) = \max\{|\lambda| : \lambda \in \operatorname{spec} a\}.$$

An analogous result of spectral radius as in operator theory is $\rho(a) = \lim_{n\to\infty} \|a^n\|^{1/n}$. (The proof can be found in [31], or we shall see that a C*-algebra is not very different from an operator algebra by representation theory.)

Definition 7.2.2. Given a C*-algebra $\mathcal{A}$, we say an element $a \in \mathcal{A}$ is **normal** if $a^*a = aa^*$; **self-adjoint** if $a^* = a$; it is **unitary** if $a^*a = e$ (consider $\mathcal{A}_e$ if $\mathcal{A}$ is not unital). If there exists $b \in \mathcal{A}$ such that $a = b^*b$, we say a is **positive**. We denote by $\mathcal{A}^+$ the set of all the positive elements in $\mathcal{A}$ and call it the **positive portion** of $\mathcal{A}$. The positive portion forms a cone, i.e. it is closed under convex linear combination (see [52] Theorem 1.4.2), therefore it is also called the positive cone. A self-adjoint element p which satisfies $p^2 = p$ is called a **projection**. A projection is obviously positive by definition. An element u is a **partial isometry** if u^*u is a projection.

We can define a partial order $\leq$ on the positive portion $\mathcal{A}^+$. We say $a \leq b$ in $\mathcal{A}^+$ if and only if $b - a$ is positive. Similarly, for projections p_1 and p_2, we say $p_1 \leq p_2$ if $p_2 - p_1$ is positive. A **minimal projection** is a projection p such that if $q \leq p$ then $q = 0$ or p.

We can also define an equivalence relation on the set of projections. We say projection p is **equivalent** to projection q and write $p \sim q$ if there exists a partial isometry u such that $u^*u = p$ and $uu^* = q$.

Definition 7.2.3. Given a C*-algebra $\mathcal{A}$, and let $S \subset \mathcal{A}$ be a subset. We say S is **self-adjoint** if for any $a \in S$, we have $a^* \in S$. We say S is a **C*-subalgebra** of $\mathcal{A}$ if S is a self-adjoint closed subalgebra of $\mathcal{A}$.

Given a C*-algebra $\mathcal{A}$, and let $S \subset \mathcal{A}$ be a subset, the **C*-subalgebra generated by** the set S is the smallest C*-subalgebra which contains $S \cup S^*$, where $S^* = \{x^* : x \in S\}$. We write it as $\mathcal{A}\langle S\rangle$. We usually use $\mathcal{A}\langle a\rangle$ to denote the C*-subalgebra generated by a and e.

The **ideal** of a C*-algebra is just the ideal of the algebra.

7.2.2 Gelfand Representation Theorem

The Gelfand representation theorem is crucial for understanding the commutative C*-algebras, as well as the self-adjoint elements in a C*-algebra, as a self-adjoint element generates a commutative C*-subalgebra. And sometimes in order to understand a non-necessarily self-adjoint element $a \in \mathcal{A}$, we start with considering a^*a in order to apply the Gelfand Representation Theorem. We refer to [31] for a complete introduction. Here we only introduce a simple version.

Definition 7.2.4. Let $\mathcal{A}_1, \mathcal{A}_2$ be C*-algebras, then we say Φ is a ***-homomorphism** if it is an algebra homomorphism from $\mathcal{A}_1$ to $\mathcal{A}_2$ satisfying $\Phi(a^*) = \Phi(a)^*$ for each $a \in \mathcal{A}_1$. Moreover, if Φ is a ***-isomorphism** if it is bijective.

Theorem 7.2.5. *(Gelfand Representation Theorem)*
*Let a be a normal element in a C*-algebra $\mathcal{A}$. Then a generates a commutative C*-subalgebra $\mathcal{A}\langle a\rangle$ of $\mathcal{A}$. Let $K(a) = \operatorname{spec} a \cup \{0\}$ (it is a compact space). Denote by $C_0(K(a))$ the C*-algebra of complex-valued continuous functions on K which vanish at 0. Then there exists an isometric *-isomorphism from $\mathcal{A}\langle a\rangle$ to $C_0(K(a))$ which maps the identity to the constant function 1_K, and a to the function $\lambda \mapsto \lambda$.*

Proof. A complete proof can be found in [31]. We only state the general idea and some important implications here.

We say φ is a **multiplicative linear functional** on a C*-algebra $\mathcal{A}$, if φ is a linear functional satisfying $\varphi(ab) = \varphi(a)\varphi(b)$ for any $a, b \in \mathcal{A}$. Denote by $\Omega(\mathcal{A})$ the set of all non-zero

multiplicative linear functional on $\mathcal{A}$. Endow it with the weakest topology such that all the evaluation maps: $\hat{x} : \varphi \mapsto \varphi(x)$ are continuous. We call this the **Gelfand topology** . Note $\Omega(\mathcal{A}) = \Omega(\mathcal{A}_e) \cup \{\varphi_\infty\}$, where $\varphi_\infty(x + \lambda e) = \lambda$ for $x \in \mathcal{A}, \lambda \in \mathbb{C}$. Under the Gelfand topology, $\Omega(\mathcal{A})$ is always locally compact and compact only if $\mathcal{A}$ is unital. Thus $\Omega(\mathcal{A}_e)$ can be considered as the one-point compactification of $\Omega(\mathcal{A})$. (See e.g. [31] Theorem 2.2.3)

We might as well assume $\mathcal{A}$ is unital. Note the evaluation map $\hat{x}$ only takes values in $\operatorname{spec} x \cup \{0\}$ because $\varphi(x - \lambda e) = \varphi(x) - \lambda$, thus $x - \lambda e$ is invertible would indicate that $\varphi(x) - \lambda \neq 0$. Therefore evaluation map $\hat{x}$ establish a connection between $\Omega(\mathcal{A})$ and the spectrum of x. In fact, this is a homeomorphism because $\hat{x}$ is continuous by the definition of the Gelfand topology and $\hat{x}$ maps a compact space onto a Hausdorff space. (The proof that $\hat{x}(\Omega(\mathcal{A})) = \operatorname{spec} x$ can be found in [31] Theorem 2.2.5)

Therefore we can define an isometric *-isomorphism

$$\psi : C(\operatorname{spec} a) \to C(\Omega(\mathcal{A}\langle a\rangle)), \quad g \mapsto g \circ \hat{a}.$$

The general Gelfand representation claims that for a commutative C*-algebra $\mathcal{A}$, there exists an isometric *-isomorphism Φ from $\mathcal{A}$ to $C(\Omega(\mathcal{A}))$ by $x \mapsto \hat{x}$.

Then

$$\mathcal{A}\langle a\rangle \xrightarrow{\Phi} C(\Omega(\mathcal{A}\langle a\rangle)) \xrightarrow{\psi^{-1}} C(\operatorname{spec} a)$$

is the isometric *-isomorphism we are looking for.

□

We introduce some applications or implications of the Gelfand representation theorem.

(a) Let a be an element of C*-algebra $\mathcal{A}$, then the followings are equivalent: (1) $a \geq 0$; (2) $\hat{a} \geq 0$; (3) $\operatorname{spec} a \subseteq [0, \infty)$. In fact, (1 $\implies$ 2) for any $\varphi \in \Omega(\mathcal{A})$, if $a = b^*b$ is positive, then $\varphi(a) = \varphi(b^*b) = |\varphi(b)|^2 \geq 0$. (2 $\implies$ 3) Note $\{\varphi(a) : \varphi \in \Omega(\mathcal{A})\} \subseteq \operatorname{spec} a \cup \{0\}$. (3 $\implies$ 1) We can consider the C*-subalgebra $\mathcal{A}\langle a\rangle$. Let $K = \operatorname{spec} a \cup \{0\} \subseteq [0, \infty)$, let Ψ be the isometric *-isomorphism from $\mathcal{A}\langle a\rangle$ to $C_0(K)$. Then $f : \lambda \mapsto \sqrt{\lambda}$ is a well-defined continuous function on K. Therefore $\Psi^{-1}(f) \in \mathcal{A}\langle a\rangle$. We call this the **square root** of a.

The square root of a^*a is often denoted by $|a|$, and called the **absolute value** of a. If a is invertible, let $u = a|a|^{-1}$, then we have

$$u^*u = |a|^{-1}a^*a|a|^{-1} = e,$$

$$uu^* = a|a|^{-2}a^* = a(a^*a)^{-1}a^* = e,$$

thus u is a unitary. We say $a = u|a|$ is the **polar decomposition** of a.

(b) For a self-adjoint element a in a C*-algebra $\mathcal{A}$, $\mathcal{A}\langle a\rangle$ is a commutative C*-subalgebra. Let Ψ be the Gelfand isomorphism, then as Ψ is isometric, we have $\|a\| = |\Psi(a)|_\infty$, where $\Psi(a)$ is the function $\lambda \mapsto \lambda$ on $\operatorname{spec} a$. This implies that $\|a\| = \rho(a)$. This result can also be obtained by $\rho(a) = \lim_{n\to\infty} \|a^n\|^{1/n}$ and the definition of C*-algebra that $\|a^*a\| = \|a\|^2$.

(c) When $\mathcal{A}$ is unital, we would have $a^*a \leq \|a^*a\|e$ for any $a \in \mathcal{A}$. In fact, we can notice for any multiplicative linear function $\varphi \in \Omega(\mathcal{A})$, we have $\varphi(\|a^*a\|e - a^*a) = \|a^*a\| - |\varphi(a^*a)| \geq 0$, as $|\varphi(a^*a)| \leq \rho(a^*a) = \|a^*a\|$.

7.2.3 Positive Linear Functionals and *-Representations

Remark 7.2.6. Let φ be a linear functional of a C*-algebra $\mathcal{A}$, then we say φ is **positive** if $\varphi(a^*a) \geq 0$ for any $a \in \mathcal{A}$. A positive linear functional works very similar to an inner product. In fact, we can define a conjugate bilinear functional $(\cdot,\cdot)$ by $(x, y) = \varphi(y^*x)$. And by the Cauchy-Schwartz inequality for $(\cdot,\cdot)$, we have

$$|\varphi(y^*x)|^2 \leq \varphi(y^*y)\varphi(x^*x).$$

Remark 7.2.7. Now let $\mathcal{A}$ be a C*-algebra, then we say the pair $\{\pi, \mathcal{H}\}$ is a ***-representation** of $\mathcal{A}$ if π is a *-homomorphism from $\mathcal{A}$ to $B(\mathcal{H})$, where $\mathcal{H}$ is a Hilbert space. The *-representation is called **faithful** if $\pi(a) = 0$ implies that $a = 0$.

We introduce here the standard method constructing a *-representation from a positive linear functional which is called the GNS construction (named after for Israel Gelfand, Mark Naimark, and Irving Segal [21][54]). Later in Section 7.7, we will apply this procedure in order to make the algebra generated by transfer operators represented on a Hilbert space.

Let φ be a positive linear functional of C*-algebra $\mathcal{A}$. Then define a conjugate bilinear functional on $\mathcal{A}$ by $(x, y) = \varphi(y^*x)$. Define $\mathcal{I}_\varphi = \{x \in \mathcal{A} : \varphi(x^*x) = 0\}$, then $\mathcal{I}_\varphi$ is an ideal of $\mathcal{A}$ by the Cauchy-Schwartz inequality. In fact, if $x \in \mathcal{I}_\varphi, y \in \mathcal{A}$, then

$$|\varphi(y^*x)| \leq \varphi(y^*y)^{1/2}\varphi(x^*x)^{1/2} = 0.$$

Let $\mathcal{A}_\varphi = \mathcal{A}/\mathcal{I}_\varphi$ (the quotient $\mathcal{A}_\varphi$ is still a C*-algebra, see [9] Proposition 1.8.2), and denote x_φ the equivalence class $x + \mathcal{I}_\varphi$ and introduce the conjugate bilinear functional on $\mathcal{A}_\varphi$ by $(x_\varphi, y_\varphi) = \varphi(y^*x)$, note this does not depend on the choice of the representatives x, y. We can see here $(\cdot,\cdot)$ is an inner product, $\mathcal{A}_\varphi$ is a pre-Hilbert space. Let $\mathcal{H}_\varphi$ be the completion of $\mathcal{A}_\varphi$ with respect to the inner product. Then $\mathcal{H}_\varphi$ is a Hilbert space. We denote by $\|\cdot\|_\varphi$ the norm induced by the inner product $(\cdot,\cdot)$ in order to distinguish from the norm $\|\cdot\|$ on $\mathcal{A}$.

Now we define a linear operator $\mathcal{T}_a$ on $\mathcal{A}_\varphi$ by $\mathcal{T}_a x_\varphi = (ax)_\varphi$. Note

$$\|\mathcal{T}_a x_\varphi\|_\varphi^2 = \varphi(x^*a^*ax) \leq \|a^*a\|\varphi(x^*x) = \|a\|^2\|x\|_\varphi^2,$$

i.e. $\|\mathcal{T}_a\|_\varphi \leq \|a\|$, therefore $\mathcal{T}_a$ is bounded. As $\mathcal{A}_\varphi$ is dense in $\mathcal{H}_\varphi$, we can extend $\mathcal{T}_a$ to $\mathcal{H}_\varphi$ and denote it by $\widetilde{\mathcal{T}_a}$. Therefore we obtain a linear map π_φ from $\mathcal{A}$ to $B(\mathcal{H}_\varphi)$ by $a \mapsto \widetilde{\mathcal{T}_a}$.

As $\pi_\varphi(ab)x_\varphi = (abx)_\varphi = \pi_\varphi(a)(bx)_\varphi$ implies that $\pi_\varphi(ab) = \pi_\varphi(a)\pi_\varphi(b)$; and

$$(\pi_\varphi(a)x_\varphi, y_\varphi) = \varphi(y^*ax) = (x_\varphi, (a^*y)_\varphi) = (x_\varphi, \pi_\varphi(a^*)y_\varphi)$$

shows that $\pi_\varphi(a^*) = \pi_\varphi(a)^*$, π_φ is a *-homomorphism. Now we have shown that $\{\pi_\varphi, \mathcal{H}_\varphi\}$ is a *-representation induced by the positive linear functional φ.

7.3 Von Neumann Algebras

7.3.1 Definition of von Neumann Algebras

Definition 7.3.1. Let $\mathcal{A}$ be a *-subalgebra of $B(\mathcal{H})$ for a Hilbert space $\mathcal{H}$. Define the **commutant** $\mathcal{A}'$ of $\mathcal{A}$ to be the set of bounded operators which commute with all elements in $\mathcal{A}$. Denote $\mathcal{A}'' = (\mathcal{A}')'$ and call it the **bicommutant** of $\mathcal{A}$. We say $\mathcal{A}$ is a **von Neumann algebra** on $\mathcal{H}$ if $\mathcal{A} = \mathcal{A}''$.

Given a self-adjoint set S of bounded linear operators, S'' is a von Neumann algebra, and we call it the **von Neumann algebra generated by** S.

Definition 7.3.2. A **W*-algebra** $\mathcal{W}$ is a C*-algebra which admits a faithful *-representation $\{\pi, \mathcal{H}\}$ such that $\pi(\mathcal{W})$ is a von Neumann algebra on the Hilbert space $\mathcal{H}$.

An alternate definition is: $\mathcal{W}$ is a W*-algebra if it is a C*-algebra and is the dual of a Banach space $\mathcal{W}_*$. We refer [52] [56] for the proof of equivalence of these definitions. By definition, $B(\mathcal{H})$ itself is a W*-algebra, thus the collection of W*-algebras is not empty. Moreover, the set cannot be the same as the collection of C*-algebras. The bicommutant condition of a von Neumann algebra $\mathcal{A}$ guarantees that the identity operator satisfies $\mathrm{id}_\mathcal{H} \in \mathcal{A}'' = \mathcal{A}$. Therefore there always exists a unit in each W*-algebra. So obviously a non-unital C*-algebra such as $C[(0,1)]$ cannot be a W*-algebra.

Remark 7.3.3. (Topologies on Von Neumann Algebras)
There are several types of topologies commonly discussed on the von Neumann algebra $B(\mathcal{H})$, where $\mathcal{H}$ is a Hilbert space. Their names have changed significantly along with the development of operator algebras and topological vector spaces. Here we clarify the nomenclature, pointing out where different names have been applied to what is actually the same topology.

(a) Norm topology: as the name suggests, the topology induced by the norm $\|\cdot\|$ on $B(\mathcal{H})$.

(b) Strong operator topology (SOT): for any $x \in \mathcal{H}$, $\mathcal{T} \mapsto \|\mathcal{T}x\|$ defines a seminorm. Then the collection of all these seminorms induces a Hausdorff locally convex topology on $B(\mathcal{H})$. In other words, a subbasis of neighbourhoods around $\mathcal{S} \in B(\mathcal{H})$ under the strong operator topology can be given by

$$\{U_s(\mathcal{S}, x, \epsilon) : x \in \mathcal{H}, \epsilon > 0\}, \text{ where } U_s(\mathcal{S}; x, \epsilon) = \{\mathcal{T} : \|(\mathcal{T} - \mathcal{S})x\| < \epsilon\}.$$

(c) Weak operator topology (WOT): for $x, y \in \mathcal{H}$, $\mathcal{T} \mapsto |\langle \mathcal{T}x, y\rangle|$ defines a seminorm. Similarly, the collection of these seminorms generates a topology on $B(\mathcal{H})$, which has a subbasis of neighbourhoods around $\mathcal{S} \in B(\mathcal{H})$ of the form

$$\{U_w(\mathcal{S}; x, y, \epsilon) : x, y \in \mathcal{H}, \epsilon > 0\},$$

where

$$U(\mathcal{S}; x, y, \epsilon) = \{\mathcal{T} \in B(\mathcal{H}) : |\langle (\mathcal{T} - \mathcal{S})x, y\rangle| < \epsilon\}.$$

(d) Ultrastrong operator topology (uSOT): the seminorms $\mathcal{T} \mapsto \sum_{n=1}^{\infty} \|\mathcal{T}x_n\|^2$, for any sequence (x_n) such that $\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$, generate a Haudorff locally convex topology on $B(\mathcal{H})$, namely the ultrastrong operator topology. Since

$$\sum_{n=1}^{\infty} \|\mathcal{T}x_n\|^2 \leq \|\mathcal{T}\|(\sum_{n=1}^{\infty} \|x_n\|^2),$$

we see the uSOT is stronger than the SOT but weaker than the norm topology.

(e) Ultraweak operator topology (uWOT): Given any sequences (x_n) and (y_n) such that $\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$ and $\sum_{n=1}^{\infty} \|y_n\|^2 < \infty$, we can define a seminorm

$$\mathcal{T} \mapsto \sum_{n=1}^{\infty} \langle \mathcal{T}x_n, y_n\rangle.$$

The collection of all these forms of seminorms generate the uWOT as above.

Remark 7.3.4. These names above are used in the classic textbook [10], however, although uWOT on $B(\mathcal{H})$ is equivalent to the topology $\sigma(B(\mathcal{H}), B(\mathcal{H})_*)$, the latter is called σ-topology in [56] and weak (operator) topology in [52]. There is no ambiguity in the latter two books because they do not discuss the WOT defined in [10].

We have a more general result for the polar decomposition of a W*-algebra. More specifically, we do not assume the invertibility anymore, but we have to weaken the assumption of u being a unitary.

Proposition 7.3.5. *Given a W*-algebra $\mathcal{W}$, for any $a \in \mathcal{W}$, we can write $a = u|a|$, where u is a partial isometry and $|a|$ is the absolute value of a. This decomposition is unique and called the* ***polar decomposition*** *of a.*

See e.g. [52] 1.12.1. for the proof.

7.3.2 Projections

Now we have almost all the tools to study the projections of a W*-algebra. Projections are much more crucial in the W*-algebras than in C*-algebras because we can always find enough projections from a W*-algebra but sometimes none in a C*-algebra. A nice example is $\mathcal{A} = C_0((0,1))$, the C*-algebra of continuous functions on $(0,1)$ which vanish at boundary. There exist no non-zero projections in $\mathcal{A}$. However, a W*-algebra is completely generated by its projections (See Part I, Chapter 1, Corollary 2 of [10]).

We recall the definition of equivalence of two projections p, q. $p \sim q$ if and only if there exists a partial isometry u such that $u^*u = p, uu^* = q$. Now let assume the polar decomposition of u is $u = v|u|$, then since $u^*u = |u|^2$, we have

$$q = uu^* = v|u|^2v^* = vpv^*. \tag{7.2}$$

In the context of a von Neumann algebra $\mathcal{W}$ on a Hilbert space $\mathcal{H}$, let p, q be the projections from $\mathcal{H}$ to $p(\mathcal{H})$ and $q(\mathcal{H})$ respectively. Assume $p \sim q$ via the partial isometry u. Denote the **support** of u by $s(u)$ (i.e. $s(u)$ is the smallest projection $s \in \mathcal{W}$ such that $us = u$), then u is an isometry from $s(u)(\mathcal{H})$ to $u(\mathcal{H})$. We denote $E = s(u)(\mathcal{H}), F = u(\mathcal{H})$ and call E the **initial subspace** of u and F the **final subspace**. It is not hard to see the support of u^* is F and u^* is an isometry from F to E. Note $u^*u = \mathrm{id}_E, uu^* = \mathrm{id}_F$, thus $E = p(\mathcal{H}), F = q(\mathcal{H})$. Therefore equation (7.2) shows that two projections to subspaces E and F respectively are equivalent if and only if we can find an isometry $v \in \mathcal{W}$ that maps E to F. Based on this, we have some corollaries.

(1) If p is a minimal projection and $p \sim q$, then q is a minimal projection. To see this, assume $p \sim q$ via u, and there exists a proper projection q_0 such that $q \neq q_0 \leq q$. By (7.2) there exists a partial isometry v such that $q = vpv^*$. Then $p \neq v^*q_0v \leq p$, contradicting the minimality of p.

(2) If p, q are minimal projections, then either $p = q$ or p and q are orthogonal: if p, q are not orthogonal, then $0 \neq pq \leq p$, thus $p = pq \leq q$.

(3) If $p \sim q$ then $\dim p(\mathcal{H}) = \dim q(\mathcal{H})$. This follows from (7.2), since v is an isometry from E to F.

7.3.3 Standard Traces

In linear algebra, the trace of a matrix is defined to be the sum of all the diagonal entries. And since a linear transformation on a finite-dimensional linear space $\mathcal{H}$ can be represented as a matrix after fixing a basis, we can define the trace of such a linear transformation to be the trace of the representing matrix. It is well-defined as we can show that the definition does not depend on the choice of the basis.

We have seen in Definition 5.4.1 that this concept can be generalised for trace-class operators on a separable Hilbert space. And a **standard trace** of a trace-class operator $\mathcal{T}$ is given by

$$\mathrm{tr}(\mathcal{T}) = \sum_{i=1}^{\infty} \langle \xi_i, \mathcal{T}\xi_i \rangle,$$

where $\{\xi_i\}$ is a orthonormal basis of $\mathcal{H}$. We denote by $T(\mathcal{H})$ the set of trace class operators, which is a Banach space, cf. [52] pp.36.

When $\mathcal{H}$ is finite-dimensional, let $\mathcal{W} = B(\mathcal{H})$, and consider the matrix representations of elements in $\mathcal{W}$. A linear functional on $\mathcal{W}$ can be written as

$$(M_{ij})_{1 \le i,j \le n} \mapsto \sum_{1 \le i,j \le n} a_{ij} M_{ij},$$

for $a_{ij} \in \mathbb{C}$. If we denote $A = (a_{ji})$ as a matrix, then we actually can write this in the form

$$\phi_A : M \mapsto \mathrm{tr}(AM).$$

Here we can see every matrix $A \in M_n(\mathbb{C})$ corresponds to a unique bounded linear functional $\phi_A : M \mapsto \mathrm{tr}(AM)$ and vice versa. That is to say $B(\mathcal{H})' \simeq B(\mathcal{H})$.

In the infinite-dimensional case, given $\mathcal{S} \in B(\mathcal{H})$,we want to similarly define

$$\phi_{\mathcal{S}} : \mathcal{T} \mapsto \mathrm{tr}(\mathcal{S}\mathcal{T}).$$

A natural question is whether $\mathcal{S}\mathcal{T}$ is of trace class.

However, we observe that if $\mathcal{S}$ is of trace class $\mathcal{S}\mathcal{T}$ is always of trace class for any $\mathcal{T} \in B(\mathcal{H})$, since

$$|\phi_{\mathcal{S}}(\mathcal{T})| = |\,\mathrm{tr}(\mathcal{S}\mathcal{T})| = \left| \sum_{i=1}^{\infty} \langle \mathcal{S}\mathcal{T}\xi_i, \xi_i \rangle \right| \le \|\mathcal{T}\| \sum_{i=1}^{\infty} \left| \langle \mathcal{S}\xi_i, \xi_i \rangle \right| = \mathrm{tr}(|\mathcal{S}|)\|\mathcal{T}\| < \infty.$$

We remark here that a deeper result can be found in [10] Chapter 6 Proposition 1, all the trace-finite elements form an ideal. This result will also be used in Section 7.8 to define the determinant.

Also notice that $\phi : \mathcal{S} \mapsto \phi_{\mathcal{S}}$ from $T(\mathcal{H})$ to $B(\mathcal{H})'$ is linear. In fact, $T(\mathcal{H})$ can be identified with $B(\mathcal{H})_*$, the predual of $B(\mathcal{H})$ (see e.g. Sakai [52] Theorem 1.15.3).

7.4 Type I Factors

As Von Neumann [59] showed that every von Neumann algebra on a separable Hilbert space can be uniquely decomposed as a direct sum of three types of factors, the study of operator algebras can usually be reduced to the study of factors.

Definition 7.4.1. We say a von Neumann algebra $\mathcal{W}$ is a **factor** if $\mathcal{W}' \cap \mathcal{W} = \mathbb{C}e$, where $\mathcal{W}'$ is the commutant of $\mathcal{W}$ and $e \in \mathcal{W}$ is the unit.

Type I factor is very closed to our understanding of a matrix. It is quite obvious that the set $M_n(\mathbb{C})$ of all $n \times n$ matrices over the complex field $\mathbb{C}$ forms a von Neumann algebra. Note the commutant of $M_n(\mathbb{C})$ is $\mathbb{C} \cdot I_n$, where I_n is the $n \times n$ identity matrix, thus $M_n(\mathbb{C})$ is a factor. If we denote by E_{ij} the elementary matrix with 1 at entry (i, j) but 0 elsewhere, we can see $M_n(\mathbb{C})$ possesses a minimal projection E_{11}. (Note all the minimal projections have the same dimension.) Here we say $M_n(\mathbb{C})$ is a type I_n factor.

Note if n is divisible by $m > 1$, write $n = tm$, then the matrices in $M_n(\mathbb{C})$ can be written in the form of a $t \times t$ matrices of blocks of $m \times m$ matrices like this

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1t} \\ A_{21} & A_{22} & \cdots & A_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ A_{t1} & A_{t2} & \cdots & A_{tt} \end{pmatrix}, \quad \text{where } A_{ij} \in M_m(\mathbb{C});$$

in other words, $M_n(\mathbb{C}) \simeq M_t(\mathbb{C}) \otimes M_m(\mathbb{C})$. In this way, we can construct a von Neumann algebra whose minimal projection is of dimension $m > 1$.

Now we replace $M_m(\mathbb{C})$ by I_m and let $\mathcal{W} = M_t(\mathbb{C}) \otimes I_m$. We can see $\mathcal{W}' = I_t \otimes M_m(\mathbb{C})$ (the result that $(\mathcal{A}_1 \otimes \mathcal{A}_2)' = \mathcal{A}_1' \otimes \mathcal{A}_2'$ can be found in [10] Part I, Chapter 6, Prop 14), therefore $\mathcal{W}$ is a factor. And $E_{11} \otimes I_m$ is a minimal projection of $\mathcal{W}$ which has dimension m. Also $I_t \otimes E_{11}$ is a minimal projection of $\mathcal{W}'$ which has dimension t. The pair of integers (t, m) fully determine the von Neumann algebra factor $\mathcal{W}$. We say $\mathcal{W} = M_t(\mathbb{C}) \otimes I_m$ is of type I_t with multiplicity m.

Here we note $M_n(\mathbb{C}) \otimes I_m$ is isomorphic with $M_n(\mathbb{C})$ by the mapping $\Phi : a \otimes I_m \mapsto a$. In order to distinguish between the two algebras, we introduce the idea of spatial isomorphism.

Given two von Neumann algebras $\mathcal{W}_1, \mathcal{W}_2$ on Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. We say Φ is a **spatial isomorphism** from $\mathcal{W}_1$ to $\mathcal{W}_2$ if there exists an isometric isomorphism $u : \mathcal{H}_1 \to \mathcal{H}_2$ such that $\Phi : a \mapsto uau^{-1}$ is a *-isomorphism.

Definition 7.4.2. Given a von Neumann algebra factor $\mathcal{W}$, if it is isomorphic with $B(\mathcal{H})$ for a Hilbert space $\mathcal{H}$ and $\dim \mathcal{H} = n(1 \leq n \leq \infty)$, then we say $\mathcal{W}$ is of **type I**; if we want to indicate the dimension, we say $\mathcal{W}$ is of **type $\mathbf{I}_n$**. Moreover, if it is spatially isomorphic with $B(\mathcal{H}) \otimes I_{\mathcal{K}}$ where $\mathcal{K}$ is a Hilbert space with dimension m, then we say $\mathcal{W}$ is of type I_n with **multiplicity** m.

We can see from the discussion above that if a factor is of type I then it contains a minimal projection. The converse is also true. Assume factor $\mathcal{W}$ contains a minimal projection p, then we can find orthogonal projections $p_1, p_2, \ldots, p_n$ such that $p_i \sim p$ for each i, and $p_1 + p_2 + \cdots + p_n = e$ where e is the unit of $\mathcal{W}$. (Here we use a result in [10] Part III Chapter 1 that in a factor, we can always compare two projections , in the sense that given two projections p, q, we can always map one of them inside another via (7.2), therefore we can transfer p into $e - p$ and get p_2. We immediately know p and p_2 are orthogonal. Then transfer p_2 into $e - p - p_2$ and get p_3 and so on, until we cannot proceed any longer. If what has been left is smaller than p, then we get a contradiction to the minimality of p.) Note $p\mathcal{W}p$ contains no proper projections as p is minimal, and since a von Neumann algebra can be generated by its projections (see e.g. [52] 1.11 for this conclusion), $p\mathcal{W}p$ is generated by $\{0, p\}$, therefore $p\mathcal{W}p \simeq \mathbb{C}p$. Thus any element $a \in \mathcal{W}$ can be written in the form

$$a = (\sum_i p_i)a(\sum_j p_j) = \sum_{i,j} p_i a p_j.$$

Note all the $p_i a p_i$ are in the form of $\lambda_i p_i$ for some $\lambda_i \in \mathbb{C}$. As we know $p_j = v_{ij} p_i v_{ij}^*$ for some partial isometry $v_{ij} \in \mathcal{W}$ by (7.2), then

$$p_i a p_j = p_i a v_{ij} p_i v_{ij}^* = \lambda_{ij} p_i v_{ij}^*.$$

That is to say $p_i a p_j$ is map which simply transfers $p_j(\mathcal{H})$ to $p_i(\mathcal{H})$ isometrically for each i, j, and multiplies by a scalar λ_{ij}. Therefore $\mathcal{W}$ is spatially isometric to a type I_n factor.

Proposition 7.4.3. *(Murray, von Neumann [41]) A von Neumann algebra factor is of type I if and only if it contains a minimal projection.*

With similar arguments as in finite-dimensional space, we can show that the definition does not depend on the choice of the orthonormal basis either, thus it is well-defined.

7.5 Abstract Traces and Type $\mathbf{II}_1$ Factors

We can easily see some basic properties of a standard trace: (1) tr is a linear functional on $B(\mathcal{H})$; (2) When $\mathcal{T} = \mathcal{S}^*\mathcal{S}$ is a positive operator, we have $\mathrm{tr}(\mathcal{T}) \geq 0$ and equality holds if and only if $\mathcal{T} = 0$; (3) for any $\mathcal{S}, \mathcal{T} \in B(\mathcal{H})$, we have $\mathrm{tr}(\mathcal{S}\mathcal{T}) = \mathrm{tr}(\mathcal{T}\mathcal{S})$.

Now we use these properties to characterise an abstract trace as follows:

Definition 7.5.1. Let $\mathcal{W}$ be a von Neumann algebra and $\mathcal{W}^+$ denotes the positive cone of $\mathcal{W}$. We say a function $\phi : \mathcal{W}^+ \to \mathbb{R}_{\geq 0}$ is an **abstract trace** on $\mathcal{W}^+$ if it satisfies the following properties:

(a) $\phi(\lambda\mathcal{S} + \mu\mathcal{T}) = \lambda\phi(\mathcal{S}) + \mu\phi(\mathcal{T})$, for $\mathcal{S}, \mathcal{T} \in \mathcal{A}^+$ and $\lambda, \mu \geq 0$;

(b) $\phi(\mathcal{T}^*\mathcal{T}) = \phi(\mathcal{T}\mathcal{T}^*)$ for any $\mathcal{T} \in \mathcal{W}$.

Moreover, we say ϕ is **faithful** if $\phi(\mathcal{S}) = 0$ for some $\mathcal{S} \in \mathcal{W}^+$ implies that $\mathcal{S} = 0$; we say ϕ is **finite** if $\phi(\mathcal{S}) < \infty$ for each $\mathcal{S} \in \mathcal{W}^+$. We say ϕ is **normal** if for any bounded increasing net $\{\mathcal{T}_\alpha\}_\alpha$ in $\mathcal{W}^+$, we have $\phi(\sup_\alpha \mathcal{T}_\alpha) = \sup_\alpha \phi(\mathcal{T}_\alpha)$ (note this is equivalent to ultraweakly continuous, see e.g [10] pp.57, pp.96 or [30] Chapter 7).

Now we introduce a type of factor on which we can define a very different kind of trace. We start with an example of group actions. Let G be a countable group and list the elements of $G = \{g_1, g_2, \ldots\}$. Denote by $l^2(G)$ the Hilbert space of squared-summable real sequences indexed by G. Note we can write the element of $l^2(G)$ in the form $a(g)v_g$, where $a : G \to \mathbb{R}$ such that $\sum_{g\in G} |a(g)|^2 < \infty$, $\{v_g\}_{g\in G}$ is a basis of $l^2(G)$.

Given $g \in G$, define the operator u_g on the basis of $l^2(G)$ by $u_g(v_h) = v_{gh}$ for each $h \in G$. Thus

$$u_g\Big(\sum_h a(h)v_h\Big) = \sum_h a(h)v_{gh} = \sum_h a(g^{-1}h)v_h.$$

We can get that u_g acts on $L^2(G, \mathbb{R})$ by

$$(u_g a)(h) = a(g^{-1}h), \text{ for } a \in L^2(G, \mathbb{R}).$$

Then $\mathcal{W} = \{u_g : g \in G\}''$ is a von Neumann algebra.

Denote by e the identity in group G. For $\beta = \sum b_g u_g \in \mathcal{W}$, we have $\sum b_g u_g(v_e) = \sum b_g v_g \in L^2(G, \mathbb{R})$, thus $\sum |b_g|^2 < \infty$. On other other hand, if β commutes with u_h, then

$$(\sum b_g u_g)u_h = u_h(\sum b_g u_g),$$

which shows that $b_{gh^{-1}} = b_{h^{-1}g}$, i.e., b is constant on conjugacy classes.

Now, if G is a free group generated by at least 2 elements, there are infinitely many elements in each conjugate class, which implies that b has to be identically 0. Thus $\mathcal{W} \cap \mathcal{W}'$ is trivial, and $\mathcal{W}$ is therefore a factor.

There is no finite normal trace on the factor $\mathcal{W}$, as $\mathcal{H}$ is of infinite-dimensional and $id_{\mathcal{H}} \in \mathcal{W}$ cannot have a finite normal trace value. But at the same time we can also notice the matrix representation of each operator in $\mathcal{W}$ has the the form $(b_{gh^{-1}})_{g,h\in G}$, i.e. the matrix is constant on every diagonal, like this

$$\begin{pmatrix} b(e) & b(x) & b(y) & \cdots \\ b(x^{-1}) & b(e) & b(x) & \cdots \\ b(y^{-1}) & b(x^{-1}) & b(e) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Now we imagine we can normalise the matrix by the dimension, and choose $b(e)$ as its trace value. Define a map $\phi : \sum b(g)u_g \mapsto b(e)$. We shall see that ϕ is actually an abstract trace.

ϕ is obviously linear. Moreover, letting $\alpha = \sum c(g)u_g$, we have $\alpha^*\alpha = \sum_{g,h} |c(g)|c(h)u_{g^{-1}h}$ and $\alpha\alpha^* = \sum_{g,h} c(g)|c(h)|u_{gh^{-1}}$, thus $\phi(\alpha^*\alpha) = \sum_g |c(g)|^2 = \phi(\alpha\alpha^*)$.

If $\phi(\alpha^*\alpha) = 0$, then $\sum |c(g)|^2 = 0$, thus c has to be constantly 0, which shows that ϕ is faithful.

Proposition 7.5.2. *(Murray, von Neumann [42]) A von Neumann algebra factor $\mathcal{W}$ is of type II if and only if there exists a unique faithful, finite normalised trace on $\mathcal{W}$.*

7.6 Crossed Products

Given a von Neumann algebra $\mathcal{W}$, some obvious ways to construct new von Neumann algebras from $\mathcal{W}$ include taking subalgebras, quotient algebras, direct sums and tensor products; crossed products have also been on the list of fundamental constructions since they were first used by Murray and von Neumann in [41] to construct a type II factor. In fact, we have already given an example of a crossed product while introducing the type II_1 factors.

Given a von Neumann algebra $\mathcal{W}$ and a locally group G, the idea of the crossed product $\mathcal{W} \rtimes G$ is to construct the set of sums of the form $\sum_g a_g u_g$, where $a_g \in \mathcal{W}$. We can see in the example of last section, we have constructed $\mathbb{C} \rtimes G$ with $a_g \in \mathbb{C}$ and G a free group. Different from when $\mathcal{W} = \mathbb{C}$, in a general scenario, we need a representation $\{\pi, \mathcal{H}\}$ of $\mathcal{W}$ such that $\{u_g : g \in G\}$ acts on the same Hilbert space $\mathcal{H}$ so that the expression $a_g u_g := \pi(a_g)u_g$ is

well-defined. Moreover, in order to compute the product $\left(\sum a_g u_g\right)\left(\sum b_h u_h\right)$, we need a rule in order to compute $u_g \pi(b_h)$.

Definition 7.6.1. Given a von Neumann algebra $\mathcal{W}$ and a group G, let α be a homomorphism from G to $\mathrm{Aut}(\mathcal{W})$, $\{\pi, \mathcal{H}\}$ be a *-representation of $\mathcal{W}$ and $g \mapsto u_g$ be a unitary representation of G on the same Hilbert space $\mathcal{H}$ such that

$$u_g \pi(b) u_g^* = \pi(\alpha(g)b), \quad \text{for } g \in G, b \in \mathcal{W}.$$

Then the **crossed product** $\mathcal{W} \rtimes_\alpha G$ is the von Neumann algebra on $\mathcal{H}$ generated by $\pi(\mathcal{W}) \cup \{u_g : g \in G\}$.

In order to make $\mathcal{W} \rtimes_\alpha G$ well-defined, we need the result that given $\mathcal{W}, G, \alpha$, such a pair of representations $\{\pi, u\}$ on a Hilbert space $\mathcal{H}$ exists. This can be found in [57] Chapter X Theorem 1.7 or [30] Chapter 11.

Now we have the following rules: (we omit the notation π for the sake of simplicity)

$$(\sum_g a_g u_g)^* = \sum_g u_g^* a_g^* u_g u_g^* = \sum_g \alpha(g^{-1})(a_g^*) u_{g^{-1}};$$

$$(\sum_g a_g u_g)(\sum_h b_h u_h) = \sum_{g,h} a_g u_g b_h u_g^* u_g u_h = \sum_{g,h} a_g \alpha(g)(b_h) u_{gh}.$$

We shall see that the W*-algebra generated by transfer operators are actually a crossed product in Section 7.7. Now we consider how to determine whether a crossed product is a factor, and if so, which type it is.

By [58] [30], for a type II_1 crossed product $R = \mathcal{W} \rtimes \Gamma$ of a von Neumann algebra $\mathcal{W}$ which equipped with a trace tr and a locally compact group Γ, for any $\sum a_g u_g \in R$, there exists a unique faithful trace ϕ defined by

$$\phi\left(\sum_g a_g u_g\right) = \mathrm{tr}(a_e).$$

Now the following is to determine when a crossed product is type II_1 factor.

Definition 7.6.2. Let G be a group acting on a von Neumann algebra $\mathcal{W}$. We say the action is **transitive** if for any $a, b \in \mathcal{W}$, there exists some $g \in G$ such that $g.a = b$; we say it is **ergodic** if the fixed points for all elements of G are $\mathbb{C} \cdot \mathrm{id}$.

Theorem 7.6.3. *(Murray, von Neumann) Given $\{G, X, \mu\}$ with G discrete and countable, $R = L^\infty(X, \mu) \rtimes G$ has the following properties:*

(a) R is a factor if and only if $\{G, X, \mu\}$ is ergodic;

Assume the ergodicity of $\{G, X, \mu\}$. Then we have:

(b) R is of type I if and only if $\{G, X, \mu\}$ is transitive;

(c) R is of type II_1 if and only if G is infinite and there exists a finite invariant measure on X equivalent to μ.

The proof can be found in [58] Chapter XIII.

7.7 Transfer Operator Algebras

We write $\mathcal{X}$ as the projective space $\mathbb{R}P^{d-1}$. Denote by F_α the set of functions on $\mathcal{X}$ which are α-Hölder continuous, i.e., if $f \in F_\alpha$, there exists a positive constant K such that for any $x, y \in \mathcal{X}$, we have $|f(x) - f(y)| \leq K d(x, y)^\alpha$. Here

$$d(x, y) = [1 - \langle x, y\rangle^2]^{1/2}$$

is the sine of the angle between direction x and y.

For any $f(x) \in F_\alpha$, define

$$|f|_\alpha = \sup_{\forall x \neq y} \frac{|f(x) - f(y)|}{d(x, y)^\alpha},$$

and $\|f\|_\alpha = |f|_\infty + |f|_\alpha$. Then $(F_\alpha, \|\cdot\|_\alpha)$ is a Banach space.

From [4] Chapter V, we have the following result:

Theorem 7.7.1. *Let $\{A_1, A_2, \ldots, A_n\}$ be a strongly irreducible and contracting set of matrices in $GL(d, \mathbb{R})$, then we can define an operator*

$$\begin{aligned} \mathcal{L}_t : \quad & C(\mathcal{X}) \to C(\mathcal{X}) \\ & w \mapsto \sum_{i=1}^{n} p_i e^{t \log \|A_i x\|} w(A_i \cdot x), \end{aligned}$$

such that there exists $\epsilon > 0$, and when $|t| < \epsilon$,

(a) $\mathcal{L}_t$ has a simple maximal eigenvalue $\beta(t) > 0$;

(b) $\mathcal{L}_t$ *restricted to* F_α *is a bounded linear operator on* F_α*. The remainder of the spectrum of* $\mathcal{L}_t : F_\alpha \to F_\alpha$ *lies in* $\{z \in \mathbb{C} : |z| < \beta(t)\}$*;*

(c) $\beta(t)$ *is differentiable with respect to* t*;*

(d) $\beta'(0)$ *equals the Lyapunov exponent.*

Now we want to use operator algebras to study transfer operators.

First we notice there is an inner product $\langle \cdot, \cdot \rangle$ on $C(\mathcal{X})$ defined by

$$\langle u(x), v(x) \rangle = \int_{\mathcal{X}} u(x)\overline{v(x)}\,\mathrm{d}x,$$

for $u(x), v(x) \in C(\mathcal{X})$. And denote $\|u(x)\|_2 = \langle u(x), u(x) \rangle^{1/2}$.

Definition 7.7.2. We say $\mathcal{L}_A$ is a **basic transfer operator** induced by matrix A, if

$$(\mathcal{L}_A w)(x) = |\det D\tilde{A}(x)|^{1/2} w(A \cdot x),$$

for any $w \in C(\mathcal{X})$.

We say $\mathcal{M}_h$ (induced by a function $h \in C(\mathcal{X})$) acting on $(C(\mathcal{X}), \|\cdot\|_2)$ is a **multiplication operator** by h if

$$(\mathcal{M}_h w)(x) = h(x)w(x), \quad \forall w \in C(\mathcal{X}).$$

We can easily see that $\mathcal{L}_A$ is an isometry under the inner product on $C(\mathcal{X})$, and that $\|\mathcal{M}_h w\|_2^2 \leq |h|_\infty^2 \|w\|_2^2$, thus $\|\mathcal{M}_h\| \leq |h|_\infty < \infty$.

Lemma 7.7.3. *(Multiplication Rule)*
Given $A, B \in GL(d, \mathbb{R})$ *and* $h_1, h_2 \in C(\mathcal{X})$*, then on* $C(\mathcal{X})$ *we have*

(a) $\mathcal{M}_{h_1}\mathcal{M}_{h_2} = \mathcal{M}_{h_1 h_2}$;

(b) $\mathcal{L}_A\mathcal{L}_B = \mathcal{L}_{BA}$;

(c) $\mathcal{L}_A\mathcal{M}_h = \mathcal{M}_{h\circ\tilde{A}}\mathcal{L}_A$.

Proof. The computation is rather direct. Since $\widetilde{BA} = \tilde{B} \circ \tilde{A}$, thus by the chain rule

$$D\widetilde{BA}(x) = D\tilde{B}(A \cdot x) \cdot D\tilde{A}(x).$$

Therefore for any $w \in C(\mathcal{X})$, we have

$$(\mathcal{L}_A\mathcal{L}_B w)(x) = |\det D\tilde{A}(x)|^{1/2} (\mathcal{L}_B w)(A \cdot x)$$

$$= |\det D\tilde{A}(x)|^{1/2} |\det D\tilde{B}(A\cdot x)|^{1/2} w(BA\cdot x)$$
$$= |\det D\widetilde{BA}(x)|^{1/2} w(BA\cdot x) = (\mathcal{L}_{BA}w)(x),$$

which shows that $\mathcal{L}_A\mathcal{L}_B = \mathcal{L}_{BA}$. Similarly

$$(\mathcal{L}_A\mathcal{M}_h w)(x) = (\mathcal{L}_A hw)(x) = |\det D\tilde{A}(x)|^{1/2} h(A\cdot x) w(A\cdot x)$$
$$= \mathcal{M}_{h\circ\tilde{A}} |\det D\tilde{A}(x)|^{1/2} w(A\cdot x) = (\mathcal{M}_{h\circ\tilde{A}}\mathcal{L}_A w)(x),$$

therefore $\mathcal{L}_A\mathcal{M}_h = \mathcal{M}_{h\circ\tilde{A}}\mathcal{L}_A$ on $C(\mathcal{X})$. □

Remark 7.7.4. The derivation of (b) might seem confusing, since we need to distinguish between $w(A\cdot x)$ as a function of x and $w(y)|_{y=A\cdot x}$ as a function of y which takes value at $A\cdot x$. When we compute $(\mathcal{L}_A\mathcal{L}_B w)(x)$, there are two ways we can do it: (1) $(\mathcal{L}_A(\mathcal{L}_B w))(x)$; (2) $\mathcal{L}_A((\mathcal{L}_B w)(x))$.

In the first way, we firstly see $\mathcal{L}_B w$ as a single function and apply $\mathcal{L}_A$, therefore we get $|\det D\tilde{A}(x)|^{1/2}(\mathcal{L}_B w)(A\cdot x) = |\det D\tilde{A}(x)|^{1/2}(\mathcal{L}_B w)(y)|_{y=A\cdot x}$. Note here y is the true variable of w before applying $\mathcal{L}_B$. Thus after expanding the expression of $(\mathcal{L}_B w)(y)$ we obtain $|\det D\tilde{A}(x)|^{1/2}|\det D\tilde{B}(y)|^{1/2} w(B\cdot y)|_{y=A\cdot x}$, which gives the result we want.

In the second way, we apply $\mathcal{L}_B$ to w first and get $\mathcal{L}_A(|\det D\tilde{B}(x)|^{1/2} w(B\cdot x))$. Now denote $u(x) = |\det D\tilde{B}(x)|^{1/2} w(B\cdot x)$, then we get

$$(\mathcal{L}_A u)(x) = |\det D\tilde{A}(x)|^{1/2} u(A\cdot x).$$

Replacing $u(x)$ by the definition, we can get the same result.

Lemma 7.7.5. *(Involution Rule)*
Given $A \in GL(d,\mathbb{R})$ and $h \in C(\mathcal{X})$, then on $C(\mathcal{X})$, define the involution $*$ of operators induced by the inner product on $C(\mathcal{X})$, we have

(a) $\mathcal{M}_h^* = \mathcal{M}_{\bar{h}}$, *where $\bar{h}$ is the complex conjugation of function h;*

(b) $\mathcal{L}_A^* = \mathcal{L}_{A^{-1}}$.

Proof. For any $u(x), v(x) \in C(\mathcal{X})$, we have

$$\langle(\mathcal{M}_h u)(x), v(x)\rangle = \int_{\mathcal{X}} h(x)u(x)\overline{v(x)}\,\mathrm{d}x$$
$$= \int_{\mathcal{X}} u(x)\overline{\overline{h(x)}v(x)}\,\mathrm{d}x$$

$$= \langle u(x), (\mathcal{M}_{\bar{h}} v)(x) \rangle.$$

Similarly,

$$\begin{aligned}
\langle (\mathcal{L}_A u)(x), v(x) \rangle &= \int_{\mathcal{X}} |\det D\tilde{A}(x)|^{1/2} u(A \cdot x) \overline{v(x)} \, \mathrm{d}x \\
&= \int_{\mathcal{X}} |\det D\tilde{A}(A^{-1} \cdot y)|^{1/2} u(y) \overline{v(A^{-1} \cdot y)} \, \mathrm{d}(A^{-1} \cdot y) \quad \text{[change of variable]} \\
&= \int_{\mathcal{X}} |\det D\tilde{A}(A^{-1} \cdot y)|^{1/2} |\det D\widetilde{A^{-1}}(y)| u(y) \overline{v(A^{-1} \cdot y)} \, \mathrm{d}y \\
&= \int_{\mathcal{X}} |\det D\widetilde{A^{-1}}(y)|^{1/2} u(y) \overline{v(A^{-1} \cdot y)} \, \mathrm{d}y \quad \text{[chain rule]} \\
&= \langle u(y), (\mathcal{L}_{A^{-1}} v)(y) \rangle,
\end{aligned}$$

where we use the chain rule for $\tilde{A} \circ \widetilde{A^{-1}} = \mathrm{id}_{\mathcal{X}}$ and get $D\tilde{A}(A^{-1} \cdot y) \cdot D\widetilde{A^{-1}}(y) = \mathrm{id}$. □

Now given a set of invertible matrices S, say $S^{-1} = \{A^{-1} \in GL(d, \mathbb{R}) : A \in S\}$, we denote by $\mathcal{TO}(S)$ the algebra generated by $\{\mathcal{L}_A, \mathcal{M}_h : A \in S \cup S^{-1}, h \in L^\infty(\mathcal{X})\}$. By the multiplication rule 7.7.3 and the involution rule 7.7.5, we can see $\mathcal{TO}(S)$ is closed under multiplication and involution. Thus $\mathcal{TO}(S)$ is a subalgebra of $B(C(\mathcal{X}))$, the set of bounded linear operators on $C(\mathcal{X})$.

As $\|\cdot\|_2$ is an operator norm induced by the inner product on $C(\mathcal{X})$, we can use the same arguments as in Example 4 to deduce that $\overline{\mathcal{TO}(S)}$ is a C*-algebra, where $\overline{\mathcal{TO}(S)}$ is the closure of $\mathcal{TO}(S)$ with respect to the norm topology. And $\overline{\mathcal{TO}(S)}^w$ is a W*-algebra where $\overline{\mathcal{TO}(S)}^w$ is the closure of $\mathcal{TO}(S)$ with respect to the weak topology.

Now with the GNS construction introduced in Section 7.2.3, we know that $\overline{\mathcal{TO}(S)}$ and $\overline{\mathcal{TO}(S)}^w$ can both be represented on a Hilbert space.

Theorem 7.7.6. *If an element in $\mathcal{TO}(S)$ can be written as $M_h \mathcal{L}_A$, then this expression is unique. Moreover, the convex hull of $\{M_h \mathcal{L}_A : h \in L^\infty(\mathcal{X}), A \in S\}$ is dense in both $\overline{\mathcal{TO}(S)}$ and $\overline{\mathcal{TO}(S)}^w$.*

Proof. The contribution of the determinant in the definition of a basic transfer operator can be absorbed into the multiplication operator. Without loss of generality, we may assume that the determinant is always 1.

(1) Assume $M_{h_1} \mathcal{L}_{A_1} = M_{h_2} \mathcal{L}_{A_2}$. Applying both to the constant function 1, we have $h_1 = h_2$, and thus $\mathcal{L}_{A_1} = \mathcal{L}_{A_2}$. Applying both to the function x, we have $A_1 \cdot x = A_2 \cdot x$ for any x

in the projective space. Write $A_1 = (\alpha_1\ \alpha_2\ \cdots \alpha_n)$ and $A_2 = (\beta_1\ \beta_2\ \cdots\ \beta_n)$. Then choose $x = e_i$ gives $\alpha_i/\|\alpha_i\| = \beta_i/\|\beta_i\|$. Write $\beta_i = k_i\alpha_i$. Now choose $x = e_1 + e_2$, yielding

$$\frac{\alpha_1 + \alpha_2}{\|\alpha_1 + \alpha_2\|} = \frac{k_1\alpha_1 + k_2\alpha_2}{\|k_1\alpha_1 + k_2\alpha_2\|},$$

then let $x = k_1e_1 + k_2e_2$ and follow the same procedure, we can inductively prove

$$\frac{\alpha_1 + \alpha_2}{\|\alpha_1 + \alpha_2\|} = \frac{k_1^n\alpha_1 + k_2^n\alpha_2}{\|k_1^n\alpha_1 + k_2^n\alpha_2\|}.$$

Assume $k_1 \neq k_2$, and without loss of generality $k_1 < k_2$, we see the latter expression converges to $\alpha_2/\|\alpha_2\|$ as $n \to \infty$. Then it shows α_1 and α_2 are in the same direction, which contradicts the fact that $\det A_1 = 1$. Therefore $k_1 = k_2$, and similarly we can derive that $A_1 = kA_2$ for some k. And since $\det A_1 = \det A_2 = 1$, we have $k^d = 1$. $k = -1$ is obviously not a solution, therefore $A_1 = A_2$.

(2) In order to see the convex hull of $\{M_h\mathcal{L}_A\}$ is dense, we only need to see the follow facts:

(a) $(M_{h_1}\mathcal{L}_{A_1})(M_{h_2}\mathcal{L}_{A_2}) = M_{h_1}M_{h_2\circ\bar{A}_1}\mathcal{L}_{A_1}\mathcal{L}_{A_2} = M_{h_1\cdot h_2\circ\bar{A}_1}\mathcal{L}_{A_2A_1}$. Thus, by the uniqueness of the expression, we see $(\sum_{h,A} M_h\mathcal{L}_A)(\sum_{h',A'} M_{h'}\mathcal{L}_{A'})$ can be uniquely written in the canonical form;

(b) If $A_1, \cdots, A_s$ are different matrices, $h_1, \ldots, h_s$ are non-zero functions, and $k_1M_{h_1}\mathcal{L}_{A_1} + \cdots + k_sM_{h_s}\mathcal{L}_{A_s} = 0$, then $k_1 = \cdots = k_s = 0$.

(c) $M_{h_1}\mathcal{L}_A + M_{h_2}\mathcal{L}_A = M_{h_1+h_2}\mathcal{L}_A$.

Thus given any expression, we can use rule (a) to write it as a sum of terms which have the form $M_h\mathcal{L}_A$; by rule (c), we can combine all the M_h which share the same $\mathcal{L}_A$; finally by rule (b), we know the final expression is canonical and unique.

In order to prove the less obvious (b), we choose independent functions $w_k(x) = x^k$, then we have

$$\begin{pmatrix} k_1h_1 & k_2h_2 & \cdots k_sh_s \end{pmatrix} \begin{pmatrix} A_1\cdot x & (A_1\cdot x)^2 & \cdots & (A_1\cdot x)^{s-1} \\ A_2\cdot x & (A_2\cdot x)^2 & \cdots & (A_2\cdot x)^{s-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_s\cdot x & (A_s\cdot x)^s & \cdots & (A_s\cdot x)^{s-1} \end{pmatrix} = O$$

The Vandermonde matrix is invertible as we have proved $A_1\cdot x, \ldots, A_s\cdot x$ are mutually different if all the matrices are all distinct functions. Thus we have $k_1h_1 = k_2h_2 = \cdots = k_sh_s = 0$. Now this proof is finished. □

This result reminds us that, if we denote by $\Gamma(S)$ the group generated by S, then $\overline{\mathcal{TO}(S)}^w = L^\infty(\mathcal{X}) \rtimes \Gamma(S)$. In fact, $u : A \to \mathcal{L}_A$ for $A \in \Gamma(S)$ is a unitary representation for $\Gamma(S)$ on $L^2(\mathcal{X})$ and $\pi : h \mapsto \mathcal{M}_h$ is a *-representation of the von Neumann algebra $L^\infty(\mathcal{X})$ on the same space.

7.8 Determinant Function

Definition 7.8.1. Let $\mathcal{W}$ be a von Neumann algebra and ϕ is a faithful finite trace on $\mathcal{W}^+$ such that $\phi(e)$=1. Given an invertible element $a \in \mathcal{W}$, let $|a|$ be its absolute value, and define the **determinant function** to be

$$\Delta(a) = \exp(\phi(\log|a|)).$$

Remark 7.8.2. Usually trace is defined on $\mathcal{W}^+$, but the domain can be extended to an ideal $\mathfrak{m}$ on which ϕ is finite (see [10] Chapter 6, Proposition 1). In this definition of determinant, as $\phi(e) = 1$, $e \in \mathfrak{m}$, we instantly have that ϕ is finite everywhere on $\mathcal{W}$.

Then we have the following properties of such a determinant function. (They are originally proved in [16] [17], we also refer [10] Part I.6.11 for a complete proof of the first four statements.)

Theorem 7.8.3. *The function Δ defined above has the following properties:*

(a) $\Delta(\lambda e) = |\lambda|$*;*

(b) $\Delta(a^*) = \Delta(a) = \Delta(a^*a)^{1/2}$*, for $a \in \mathcal{W}$ invertible;*

(c) $\Delta(ab) = \Delta(a)\Delta(b)$ *for $a, b \in \mathcal{W}$ invertible;*

(d) Let $a, b \in \mathcal{W}$ with b self-adjoint, then we have

$$\Delta(\exp a^* \exp b \exp a) = \exp \phi(a^* + b + a).$$

(e) Δ is continuous on the set of invertible elements.

Now our idea here is to define a function $d(\lambda) = \Delta(a - \lambda e)$. Note $d(\lambda)$ is only well-defined on $\mathbb{C} \setminus \operatorname{spec} a$. So we want to extend its definition and hopefully it would reach 0 on $\operatorname{spec} a$. There is some discussion in [17] on how to extend Δ to non-invertible elements but it is too general to be useful and also different from extending $d(\lambda)$, as what we want is more specific: our ultimate goal is to calculate the simple top eigenvalue of a transfer operator. Therefore, if the spectral radius of the transfer operator is β, we only need to extend $d(\lambda)$ from $|\lambda| > \beta$ to the single point β. Here is our main result.

Theorem 7.8.4. *Let $\mathcal{W}$ be a von Neumann algebra and ϕ a faithful finite trace on $\mathcal{W}^+$ such that $\phi(e) = 1$. Let $a \in \mathcal{W}$, assume the spectral radius of a is β, and define*

$$d_N(\lambda) = |\lambda| \exp\Big(-\sum_{n=1}^{N} \frac{\phi(a^n) + \phi(a^{*n})}{2n\lambda^n} \Big),$$

then $d_N(\lambda) \to \Delta(a - \lambda e)$ in the ultraweak topology for any $\lambda > \beta$ as $N \to \infty$. Moreover, if a has β as its top eigenvalue and the remainder of the spectrum lies in $\{z \in \mathbb{C} : |z| < \beta\}$, then $\lim_{N\to\infty} d_N(\beta) = 0$.

Assume the spectral radius of a is β and choose $\lambda > \beta$, then $\rho(a/\lambda) < 1$. Find a proper $r \in \mathbb{R}$ satisfying $\beta/\lambda < r < 1$. Define $\Gamma_r = \{z \in \mathbb{C} : |z| = r\}$, then $\xi e - a/\lambda$ is invertible for $\xi \in \Gamma_r$ and $\log(1 - \xi)$ is holomorphic in the open disc

$$D_r = \{z \in \mathbb{C} : |z| < r\}.$$

Then by functional calculus

$$\log(e - a/\lambda) = \frac{1}{2\pi i} \int_{\Gamma_r} \log(1 - \xi)(\xi e - a/\lambda)^{-1} \, \mathrm{d}\xi$$

As the Laurent series of $\log(1 - \xi)$ in D_r is

$$\log(1 - \xi) = -\xi - \frac{\xi^2}{2} - \frac{\xi^3}{3} - \cdots$$

We can define the n-th truncated sum

$$S_N(\xi) = -\sum_{n=1}^{N} \frac{\xi^n}{n},$$

and we have the error estimation

$$\begin{aligned} |\log(1 - \xi) - S_N(\xi)| &\le \sum_{n=N+1}^{\infty} \frac{|\xi|^n}{n} = \sum_{n=1}^{\infty} \frac{|\xi|^{N+n}}{N+n} \\ &< |\xi|^N \sum_{n=1}^{\infty} \frac{|\xi|^n}{n} = |\xi|^N |\log(1 - |\xi|)| < r^N |\log(1 - r)|. \end{aligned}$$

That is to say, for a given r, $S_N(\xi)$ converges to $\log(1 - \xi)$ uniformly on the open disc D_r as $N \to \infty$.

Now we denote

$$S_N(a/\lambda) = \frac{1}{2\pi i}\int_{\Gamma_r} S_N(\xi)(\xi e - a/\lambda)^{-1}\,\mathrm{d}\xi = -\sum_{n=1}^{N}\frac{(a/\lambda)^n}{n}.$$

Lemma 7.8.5. *Assume $T \in B(\mathcal{H})$ is a bounded linear operator on a Hilbert space $\mathcal{H}$ with spectral radius β. Then the resolvent $R(\xi, T) = (\xi e - T)^{-1}$ is bounded on $\mathbb{C} \setminus D_r$ for any $r > \beta$, where $D_r = \{z \in \mathbb{C} : |z| < r\}$.*

Proof. Note when $|\xi| > \|T\| \geq \rho(T)$

$$(\xi e - T)^{-1} = \xi^{-1}\Big(e - \frac{T}{\xi}\Big)^{-1} = \xi^{-1}\sum_{n=0}^{\infty}\Big(\frac{T}{\xi}\Big)^n,$$

thus

$$\|(\xi e - T)^{-1}\| \leq |\xi|^{-1}\sum_{n=0}^{\infty}\frac{\|T\|^n}{|\xi|^n} = \frac{|\xi|^{-1}}{1 - \|T\|/|\xi|} = \frac{1}{|\xi| - \|T\|} \to 0$$

as $|\xi| \to \infty$. On the other hand, we denote $f(\xi) = \|R(\xi, T)\|$, then

$$\begin{aligned} |f(\xi) - f(\eta)| &\leq \|(\xi e - T)^{-1} - (\eta e - T)^{-1}\| \\ &= \|(\xi e - T)^{-1}(\eta e - T - \xi e + T)(\eta e - T)^{-1}\| \\ &= |\eta - \xi|\|(\xi e - T)^{-1}\|\|(\eta e - T)^{-1}\| = |\eta - \xi| f(\xi) f(\eta). \end{aligned}$$

Fix $\eta \in \mathbb{C} \setminus \operatorname{spec} T$ and we have $f(\eta) < \infty$, then when $|\xi - \eta| < f(\eta)^{-1}$, we have

$$\frac{f(\eta)}{1 + |\eta - \xi| f(\eta)} \leq f(\xi) \leq \frac{f(\eta)}{1 - |\eta - \xi| f(\eta)}.$$

This gives $\lim_{\xi \to \eta} f(\xi) = f(\eta)$, i.e., $f(\xi)$ is a continuous function on $\mathbb{C} \setminus \operatorname{spec} T$ which vanishes at infinity. This implies that f is bounded on $\mathbb{C} \setminus D_r$ fo any $r > \beta$. □

Now if we denote by K the bound of $\|(\xi e - a/\lambda)^{-1}\|$, we have

$$\begin{aligned} \|\log(e - a/\lambda) - S_N(a/\lambda)\| &\leq \frac{1}{2\pi} K \int_{\Gamma_r} |\log(1 - \xi) - S_N(\xi)|\,\mathrm{d}\xi \\ &\leq K r^{N+1} |\log(1 - r)|. \end{aligned}$$

That is to say $S_N(a/\lambda)$ converges to $\log(e - a/\lambda)$ in the norm topology as $N \to \infty$.

Now for a general a, by Theorem 7.8.3

$$\Delta(a-\lambda e)=\Delta(\lambda e)\Delta(e-a/\lambda)=|\lambda|\Delta((e-a/\lambda)^*(e-a/\lambda))^{1/2}.$$

Denote $b=\log(e-a/\lambda)$, then $\exp b=e-a/\lambda$ and $(e-a/\lambda)^*=(\exp b)^*=\exp b^*$. Thus

$$\Delta(a-\lambda e)=|\lambda|\Delta(\exp b^*\exp b)^{1/2}=|\lambda|\exp\Big(\phi\Big(\frac{b+b^*}{2}\Big)\Big).$$

Now we use $S_N(a/\lambda)$ to approximate b, and define

$$\begin{aligned}d_N(\lambda)&=|\lambda|\exp\Big(\phi\Big(\frac{S_N(a/\lambda)+S_N(a/\lambda)^*}{2}\Big)\Big)\\&=|\lambda|\exp\Big(-\sum_{n=1}^{N}\frac{\phi(a^n)+\phi(a^{*n})}{2n\lambda^n}\Big)\end{aligned}$$

As the map the involution map $x\mapsto x^*$ is ultraweakly continuous by Theorem 1.7.8 in [52], we have $d_N(\lambda)$ converges to $\Delta(a-\lambda e)$ in the ultraweak topology as $N\to\infty$.

For the second statement, as a a an isolated top eigenvalue β, we can write

$$a=\beta p+r,$$

where p is a one-dimensional projection on the eigenvector corresponding to β, $\|r\|<\beta_0<\beta$ and $pr=rp=0$. Now as $a^n=\beta^n p+r^n$, we have

$$\sum_{n=1}^{\infty}\frac{\phi(a^n)}{n\beta^n}=\sum_{n=1}^{\infty}\frac{\phi(p)}{n}+\sum_{n=1}^{\infty}\frac{\phi(r^n)}{n\beta^n}.$$

Note $|\phi(r^n)|\le\|\phi\|\|r\|^n<\|\phi\|\beta_0^n$, thus

$$\left|\sum_{n=1}^{\infty}\frac{\phi(r^n)}{n\beta^n}\right|\le\sum_{n=1}^{\infty}\|\phi\|\left(\frac{\beta_0}{\beta}\right)^n<\infty,$$

however $\phi(p)\neq 0$ as ϕ is faithful, so

$$\sum_{n=1}^{\infty}\frac{\phi(p)}{n}=\infty,$$

and

$$\sum_{n=1}^{\infty}\frac{\phi(a^n)}{n\beta^n}=\infty,$$

similarly,

$$\sum_{n=1}^{\infty} \frac{\phi(a^{*n})}{n\beta^n} = \infty,$$

thus the conclusion follows.

7.9 Discussion

We remark here that a complete algorithm generalising Pollicott's for a probability measure on a general matrix set is very challenging and has not been achieved in this thesis, and that further work is required in order to have a usable algorithm.

Ideally, if we want to use this method to compute the Lyapunov exponent, we shall start with determining whether the action is free, ergodic or transitive, then by Theorem 7.6.3 we will determine which type of factor the crossed product is. We define the corresponding trace for the factor then apply Theorem 7.8.4 to get an algorithm analogous to the one by Pollicott.

This method at its current form, in fact, turns out to be not very efficient in computing the Lyapunov exponent numerically. For example, we need to use Theorem 7.6.3 to determine which type of factor our problem is, which is typically quite challenging.

However, regarding transfer operators as part of an operator algebra; identifying the transfer operator algebra with a crossed product; and proving the continuity of a determinant function after extending its domain in order to compute the top eigenvalue are still interesting on their own.

Moreover, there are two possible improvements for this method. The first one is considering a general von Neumann algebra instead of just factors. Since a von Neumann algebra on a separable Hilbert space can be decomposed into a direct sum of factors, given a transfer operator algebra, we can decompose it into factors first, define the trace for each factor and obtain a trace for the whole algebra.

The second one is combining the operator algebra theory with semigroup theory. In Section 3.1, we have shown that a completely simple semigroup (the semigroup formed by limit points) can be written in the form of $E \times G \times F$, where G is a locally compact group and $FE \subset G$. Here we have a more natural group which arises. Note by the multiplication rule (b), for $s = egf$, we would have $\mathcal{L}_s = \mathcal{L}_f \mathcal{L}_g \mathcal{L}_e$, therefore, instead of considering the group generated by the matrix set, we can try $\{\mathcal{L}_g : g \in G\}$.

Appendix A

Code

A.1 Pollicott's Algorithm in R

```
############################################################
#
#  PART2: POLLICOTT'S METHOD
#
############################################################

set.seed(1)
d <- 2
A1 <- matrix(abs(rnorm(d^2)),d,d)
A2 <- matrix(abs(rnorm(d^2)),2,2)
n <- 3 #nth approximation is lambda_n
Alist <- list(A1,A2)
norm.vec <- function(x) sqrt(sum(x^2))
#denom <- function(A) 1-1/eigen(A)$values[1]^2 #By lemma 4.2
denom <- function(A){
eig <- eigen(A)$values
eig1 <- 1-eig[-1]/eig[1]
return(prod(eig1))
}#New formula

############################################################
#tr.part() is used to caculate some terms in equation (4.1)
#input: string i
#output: denominator and sum(f_j) in (4.1)
############################################################
tr.part <- function(i){
```

```
if(length(i)==1) {
A <- Alist[[i]]
eigva.A <- eigen(A)$values[1]
eigvec.A <- eigen(A)$vectors[,1]
denom <- denom(A) #By lemma 4.2
sum.f1 <- log(norm.vec(A1%*%eigvec.A))-log(norm.vec(eigvec.A))
sum.f2 <- log(norm.vec(A2%*%eigvec.A))-log(norm.vec(eigvec.A))
result <- c(denom,sum.f1,sum.f2)
return(result)
}
else
prod.A <- Reduce('%*%',Alist[i]) #product wrt string i
eigva.prod.A <- eigen(prod.A)$values[1] #top eigenvalue of prod.A
eigvec.prod.A <- eigen(prod.A)$vectors[,1] #top eigenvector of prod.A
denom <- denom(A)
sum.f1 <- 0
sum.f2 <- 0
for(r in 1:length(i)){
i <- i[c(2:length(i),1)]
prod.A <- Reduce('%*%',Alist[i]) #products wrt string sigma_i
eigva.prod.A <- eigen(prod.A)$values[1]
x <- eigen(prod.A)$vectors[,1]
f1 <- log(norm.vec(A1%*%x))-log(norm.vec(x))
f2 <- log(norm.vec(A2%*%x))-log(norm.vec(x))
sum.f1 <- sum.f1 + f1
sum.f2 <- sum.f2 + f2
}
result <- c(denom,sum.f1,sum.f2)
return(result)
}

##############################################################
#string.gen() generates all the strings with length m
#input: length m
#ouput: all possible strings in the form of string.gen(m)[k,]
##############################################################
string.gen <- function(m){
i <- c(1,2)
ii <- replicate(m,list(i))
result <- expand.grid(ii)
```

```
return(result)
}

############################################################
#tr.L() calculates tr(L) when t=0 based on the first equation on p218;
#Notice when t=0 tr.L() does not depend on j
#input: length m
############################################################
tr.L <- function(m){
result <- 0
for (k in 1:2^m){
i <- string.gen(m)[k,]
term <- 1/(tr.part(as.numeric(i))[1]*2^m)
result <- result + term
}
return(result)
}

############################################################
#d.tr.L() calculates d(tr.L)/dt at t=0
#Notice d.tr.L relies on j
#input: j=1 or 2, length m
############################################################
d.tr.L <- function(j,m){
result <- 0
for(k in 1:2^m){
i <- string.gen(m)[k,]
term <- tr.part(as.numeric(i))[j+1]/(tr.part(as.numeric(i))[1]*2^m)
result <- result + term
}
return(result)
}

############################################################
#sum.index() finds all the integer solutions for x_1+...+x_k=m
#input: k,m
#output: ith solution is in the form of sumindex(k,m)[,i]
############################################################
sum.index <- function(k,m){
if(k==1) return(matrix(m))
else
```

```
s <- combn(1:(m-1),k-1) #every combination generates a solution
n <- length(s[1,])
sollist <- NULL
for(i in 1:n){
sol <- NULL
sol[1] <- s[1,i]
if(k==2) {sol[2] <- m - sol[1]}
else{
for(j in 2:(k-1)){sol[j] <- s[j,i]-s[j-1,i]}
sol[k] <- m-s[j,i]}
sollist <- cbind(sollist,sol)
}
return(sollist)
}

############################################################
#tr.L.m makes A() neat (tr.L.m=tr.L/m)
############################################################
tr.L.m <- function(m){
result <- NULL
for (i in 1:length(m)){
result[i] <- tr.L(m[i])/m[i]}
return(result)
}

#########################################################
#q.tr.L makes C() neat (q.tr.L=d.tr.L/tr.L)
#########################################################
q.tr.L <- function(j,m){
result <- NULL
for (i in 1:length(m)){
result[i] <- d.tr.L(j,m[i])/tr.L(m[i])}
return(result)
}

############################################################
#A() calculates a_m^j at t=0 based on first equation on p220
#Notice A() does not depend on j
#input: m
############################################################
A <- function(m){
```

```
term.sum <- 0
for(k in 1:m){
term.subsum <- 0
for(i in 1:length(sum.index(k,m)[1,])){
term <- prod(tr.L.m(sum.index(k,m)[,i]))
term.subsum <- term.subsum + term
}
term.sum <- term.sum + (-1)^k*term.subsum/factorial(k)
}
return(term.sum)
}

B <- function(m){
#if(m==1) return(0)
#else
#return(m*A(m-1))
return(m*A(m))
} #the definitions of B() are different in section 1 and lemma5.2
#I think m*A(m) should be right

#############################################################
#C() calculates c_m^j based on lemma 5.2
#############################################################
C <- function(j,m){
term.sum <- 0
for(k in 1:m){
term.subsum <- 0
for(i in 1:length(sum.index(k,m)[1,])){
term <- sum(q.tr.L(j,sum.index(k,m)[,i]))*prod(tr.L.m(sum.index(k,m)[,i]))
term.subsum <- term.subsum + term
}
term.sum <- term.sum + (-1)^k*term.subsum/factorial(k)
}
return(term.sum)
}

#################################################################
#lambda(n) calculates nth approximation of lambda based on theorem 1.1
#################################################################
lambda <- function(n)
{
c1.sum <- 0
```

```
c2.sum <- 0
b.sum <- 0
for(i in 1:n) c1.sum <- c1.sum + C(1,i)
for(i in 1:n) c2.sum <- c2.sum + C(2,i)
for(i in 1:n) b.sum <- b.sum + B(i)
lambda.n <- (c1.sum/b.sum + c2.sum/b.sum)/2
return(lambda.n)
}


sprintf("%.20f",lambda(n))
```

Bibliography

[1] L Arnold. *Random dynamical systems*. Springer monographs in mathematics. Springer, Berlin ; London, 1998.

[2] R. Bellman. Limit theorems for non-commutative operations. i. *Duke Mathematical Journal*, 21(3):491–500, 1954.

[3] Garrett Birkhoff. Extensions of Jentzsch's theorem. *Transactions of the American Mathematical Society*, 85(1):219–227, 1957.

[4] Philippe Bougerol and Jean Lacroix. *Products of random matrices with applications to Schrödinger operators*. Progress in probability and statistics ; v. 8. Birkhäuser, Boston, 1985.

[5] Rufus Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Lecture notes in mathematics (Springer-Verlag) ; 470. Springer-Verlag, Berlin, 1975.

[6] P. J. Bushell. Hilbert's metric and positive contraction mappings in a Banach space. *Arch. Rational Mech. Anal.*, 52(4):330–338, 1973.

[7] Alfred Hoblitzelle Clifford and Gordon Bamford Preston. *The algebraic theory of semigroups*, volume 7. American Mathematical Soc., 1961.

[8] B Derrida, K Mecheri, and JL Pichard. Lyapounov exponents of products of random matrices: weak disorder expansion.-application to localisation. *Journal de Physique*, 48(5):733–740, 1987.

[9] Jacques Dixmier. *C^*-algebras*. North-Holland Publishing Co., Amsterdam, 1977. Translated from the French by Francis Jellett, North-Holland Mathematical Library, Vol. 15.

[10] Jacques Dixmier. *von Neumann algebras*, volume 27 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1981. With a preface by E. C. Lance, Translated from the second French edition by F. Jellett.

[11] Richard Durrett. *Probability : theory and examples*. Cambridge series on statistical and probabilistic mathematics. Cambridge, 4th edition edition, 2010.

[12] Michael Fekete. Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten. *Mathematische Zeitschrift*, 17(1):228–249, 1923.

[13] L. E. Fraenkel. Formulae for high derivatives of composite functions. *Mathematical Proceedings of the Cambridge Philosophical Society*, 83, 3 1978.

[14] Ivar Fredholm. Sur une classe d'équations fonctionnelles. *Acta Mathematica*, 27, 1903.

[15] David Fried. The zeta functions of Ruelle and Selberg. I. In *Annales scientifiques de l'Ecole normale supérieure*, volume 19, pages 491–517. Elsevier, 1986.

[16] Bent Fuglede and Richard V Kadison. On determinants and a property of the trace in finite factors. *Proceedings of the National Academy of Sciences*, 37(7):425–431, 1951.

[17] Bent Fuglede and Richard V Kadison. Determinant theory in finite factors. *Annals of Mathematics*, (3):520–530, 1952.

[18] H. Furstenberg. Noncommuting random products. *Transactions of the American Mathematical Society*, 108(3):377–428, 1963.

[19] H. Furstenberg and H. Kesten. Products of random matrices. *The Annals of Mathematical Statistics*, 31(2):457–469, June 1960.

[20] H. Furstenberg and Y. Kifer. Random matrix products and measures on projective spaces. *Israel J. Math.*, 46(1-2):12–32, 1983.

[21] Israel Gelfand and Mark Neumark. On the imbedding of normed rings into the ring of operators in Hilbert space. *Rec. Math. [Mat. Sbornik] N. S.*, 12(2):197–217, 1943.

[22] Alexandre Grothendieck. *Produits tensoriels topologiques et espaces nucléaires*, volume 16. American Mathematical Soc., 1955.

[23] Alexandre Grothendieck. La théorie de Fredholm. *Bull. Soc. Math. France*, 84(31):384, 1956.

[24] Y Guivarc'h and A Raugi. Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence. *Probability Theory and Related Fields*, 69(2):187–242, 1985.

[25] Yves Guivarc'h. Exposants caracteristiques des produits de matrices aléatoires en dépendance markovienne. In *Probability Measures on Groups VII*, pages 161–181. Springer, 1984.

[26] Hubert Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proceedings of the American Mathematical Society*, 118(2):627–634, 1993.

[27] Göran Högnäs and Arunava Mukherjea. *Probability measures on semigroups: convolution products, random walks, and random matrices*. SIAM, 2011.

[28] Nahum Krupnik Israel Gohberg, Seymour Goldberg. *Traces and Determinants of Linear Operators*. Operator theory, advances and applications 116. Birkhäuser Verlag, 1st edition, 2000.

[29] Oliver Jenkinson and Mark Pollicott. Calculating Hausdorff dimension of Julia sets and Kleinian limit sets. *American Journal of Mathematics*, 124(3):495–545, 2002.

[30] Vaughan F.R. Jones. Lecture Notes: Von Neumann Algebras.

[31] Eberhard Kaniuth. *A course in commutative Banach algebras*. Graduate texts in mathematics ; 246. Springer, New York, 2009.

[32] Gerhard Keller and Carlangelo Liverani. Stability of the spectrum for transfer operators. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 28(1):141–152, 1999.

[33] John Frank Charles Kingman. Subadditive ergodic theory. *The annals of Probability*, 1(6):883–899, 1973.

[34] Andrzej Lasota and James A Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Transactions of the American Mathematical Society*, 186:481–488, 1973.

[35] Emile Le Page. Théorèmes limites pour les produits de matrices aléatoires. *Probability Measures on Groups*, pages 258–303, 1982.

[36] John M. Lee. *Introduction to smooth manifolds*. Graduate texts in mathematics ; 218. Springer, New York ; London, 2002.

[37] Ronnie Mainieri. Cycle expansion for the Lyapunov exponent of a product of random matrices. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 2(1):91–97, 1992.

[38] Jens Marklof, Yves Tourigny, and Lech Wołowski. Explicit invariant measures for products of random matrices. *Transactions of the American Mathematical Society*, 360(7):3391–3427, 2008.

[39] Dieter Mayer. Continued fractions and related transformations. *Conference:Ergodic theory, symbolic dynamics, and hyperbolic spaces*, 1989.

[40] Arunava Mukherjea. *Topics in products of random matrices*. Tata Institute of Fundamental Research lectures on mathematics ; 87. Narosa, New Delhi ; London, 2000.

[41] Francis J Murray and John von Neumann. On rings of operators. *Annals of Mathematics*, 37(1):116–229, 1936.

[42] Francis J Murray and John von Neumann. On rings of operators. ii. *Transactions of the American Mathematical Society*, 41(2):208–248, 1937.

[43] Jacob T. Schwartz Nelson Dunford. *Linear Operators, Part I: General Theory*. Wiley-Interscience, 1988.

[44] William Parry and Mark Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics*. Astérisque, 187-188. Société mathématique de France, Paris, 1990.

[45] Kalyanapuram Rangachari Parthasarathy. *Probability measures on metric spaces*, volume 352. American Mathematical Soc., 2005.

[46] Karl Endel Petersen. *Ergodic theory*. Cambridge studies in advanced mathematics ; 2. Cambridge University Press, Cambridge [Cambridgeshire] ; New York, 1983.

[47] Albrecht Pietsch. *Eigenvalues and s-Numbers*. Cambridge Studies in Advanced Mathematics 13. Cambridge University Press, 1987.

[48] Mark Pollicott. Maximal Lyapunov exponents for random matrix products. *Inventiones mathematicae*, 181(1):209–226, July 2010.

[49] David Ruelle. Zeta-functions for expanding maps and Anosov flows. *Inventiones mathematicae*, 34(3):231–242, October 1976.

[50] David Ruelle. *Dynamical zeta functions for piecewise monotone maps of the interval*. CRM monograph series, 4. American Mathematical Society, Providence, R.I., 1994.

[51] Anthony Francis Ruston. *Fredholm theory in Banach spaces*. Cambridge Tracts in Mathematics. Cambridge University Press, 1986.

[52] Shôichirô Sakai. *C^*-algebras and W^*-algebras*. Classics in Mathematics. Springer-Verlag, Berlin, 1998. Reprint of the 1971 edition.

[53] Helmut H Schaefer. *Topological vector spaces*. Graduate texts in mathematics ; 3. Springer-Verlag, New York, 4th corr. printing. edition, 1980.

[54] Irving E Segal. Irreducible representations of operator algebras. *Bulletin of the American Mathematical Society*, 53(2):73–88, 1947.

[55] Michael Shub, Albert Fathi, Rémi Langevin, and Joseph Christy. *Global stability of dynamical systems*. Springer-Verlag, New York ; London, 1987.

[56] M. Takesaki. *Theory of operator algebras. I*, volume 124 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2002. Reprint of the first (1979) edition, Operator Algebras and Non-commutative Geometry, 5.

[57] M. Takesaki. *Theory of operator algebras. II*, volume 125 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2003. Operator Algebras and Non-commutative Geometry, 6.

[58] M. Takesaki. *Theory of operator algebras. III*, volume 127 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2003. Operator Algebras and Non-commutative Geometry, 8.

[59] John von Neumann. On rings of operators iii. *Annals of Mathematics*, 41(1):94–161, 1940.

[60] John von Neumann. On rings of operators. reduction theory. *Annals of Mathematics*, 50(2):401–485, 1949.

[61] Charles Walkden. Transfer operators for contractive markov systems and stochastic stability of the invariant measure. *Dynamical Systems*, 28(1):34–47, 2013.

Index

*-homomorphism, 172
*-isomorphism, 172

absolute value, 173
abstract trace, 181
additive cocycle, 26
admissible sequence, 66
approximation property, 114

Banach *-algebra, 170
Banach algebra, 170
basic transfer operator, 185
bicommutant, 175

C*-algebra, 170
C*-subalgebra, 172
cocycle, 67
commutant, 175
completely simple, 52
cone, 87
contracting, 38
contracting RDS, 69
convolution, 22
crossed product, 183

determinant function, 125, 189
dimension of a cone, 87

equivalent projections, 172
exterior power, 31

factor, 179
faithful *-representation, 174
faithful trace, 181
final subspace, 177
finite expectation condition, 21
finite trace, 181
Fredholm determinant, 121
full shift, 66

Gelfand topology, 173
generated semigroup, 34
generated von Neumann algebra, 175

Hilbert-Schmidt operator, 112

i.i.d., 66
ideal, 172
ideal of a semigroup, 52
idempotent, 52
index of a matrix set, 38
index, RDS, 69
initial subspace, 177
integral operators, 120
invariant measure, 22
invariant measure (for LRDS), 69
invertible, 171
invertible RDS, 69
involution, 170
irreducible, 34
irreducible on the cone, 87

irreducible RDS, 69

kernel function, 120

Lasota-Yorke inequalities, 157
linear cocycle, 67
linear RDS (LRDS), 67
Lyapunov exponent, 21
Lyapunov spectrum, 33

Markovian, 66
matrix action, 22
minimal projection, 172
multiplication operator, 185
multiplicative linear functional, 172
multiplicity of type I, 180

natural actions, 22
normal element, 171
normal trace, 181
normalise actions, 22
normalised cocyle, 68
normalised RDS (NRDS), 68
normalised transfer operators, 154
nuclear of order 0, 113
nuclear of order p, 113
nuclear operator, 113
nuclear space, 115

one-sided shift space, 66

parametrised transfer operators, 99
partial isometry, 171
polar decomposition, 173, 177
positive element, 171
positive linear functional, 174
positive portion, 171
primitive, 52
probability vector, 79
projection, 171
proper measure, 36

quasi-complete, 115

real projective space, 21

self-adjoint element, 171
self-adjoint set, 172
semigroup action, 22
shift map, 65
shift of finite type, 66
shift over k symbols, 66
shift space, 65
simple semigroup, 52
simple transfer operators, 116
singular value decomposition, 30
singular value of operators, 111
singular values, 30
skew product, 69
spatial isomorphism, 180
spatially invertible, 69
spectral radius, 171
spectrum, 171
square root, 173
standard trace, 178
stationary, 66
strong operator topology, SOT, 176
strongly irreducible, 34
strongly irreducible on the cone, 87
strongly irreducible RDS, 69
subadditive inequality, 21
subcone, 87
support of an operator, 177
support of measure, 34

tightness, 24
topological semigroup, 22

trace-class operator, 112
transfer operator, 91
transitive action, 183
two-sided shift space, 66
type I factor, 180

ultrastrong operator topology, uSOT, 176
ultraweak operator topology, uWOT, 176
unification, 171
unit, 170
unital, 171
unitary element, 171

von Neumann algebra, 175

W*-algebra, 175
weak operator topology, WOT, 176
weighted transfer operators, 99