UNIVERSIDADE ESTADUAL DE
CAMPINAS

Instituto de Matemática, Estatística e
Computação Científica

ENRIQUE FERNANDO LÓPEZ AGILA

REGULARIDADE DO MÁXIMO EXPOENTE DE LYAPUNOV

Campinas

2019

Enrique Fernando López Agila

REGULARIDADE DO MÁXIMO EXPOENTE DE LYAPUNOV

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Matemática Aplicada.

Orientador: Christian da Silva Rodrigues

Este exemplar corresponde à versão final da Dissertação defendida pelo aluno Enrique Fernando López Agila e orientada pelo Prof. Dr. Christian da Silva Rodrigues .

Campinas

2019

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

L881r López Agila, Enrique Fernando, 1989-
Regularidade do máximo expoente de Lyapunov / Enrique Fernando López Agila. – Campinas, SP : [s.n.], 2019.

Orientador: Christian da Silva Rodrigues.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Lyapunov, Expoentes de. 2. Pertubação (Matemática). 3. Teoria ergódica. I. Rodrigues, Christian da Silva, 1978-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Regularity of the maximal exponent of Lyapunov
Palavras-chave em inglês:
Lyapunov exponents
Perturbation (Mathematics)
Ergodic theory
Área de concentração: Matemática Aplicada
Titulação: Mestre em Matemática Aplicada
Banca examinadora:
Christian da Silva Rodrigues [Orientador]
Pedro José Catuogno
Katrin Grit Gelfert
Data de defesa: 13-03-2019
Programa de Pós-Graduação: Matemática Aplicada

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0002-2840-8572
- Currículo Lattes do autor: http://lattes.cnpq.br/3770142756607966

Dissertação de Mestrado defendida em 13 de março de 2019 e aprovada

pela banca examinadora composta pelos Profs. Drs.

Prof(a). Dr(a). CHRISTIAN DA SILVA RODRIGUES

Prof(a). Dr(a). PEDRO JOSÉ CATUOGNO

Prof(a). Dr(a). KATRIN GRIT GELFERT

A Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação do Instituto de Matemática, Estatística e Computação Científica.

Aos meus pais...

Agradecimentos

Agradeço a Deus pela fidelidade que me mostra todos os dias, a minha família, ao Prof. Dr. Christian S. Rodrigues, pela orientação e a amizade que me deu durante o mestrado. Agradeço aos professores da banca, Prof. Dra. Katrin Grit Gelfert, Prof. Dr. Pedro José Catuogno e Prof Dr. Yuri Dimitrov Bozhkov, por me ajudar com as dúvidas e pelas valiosas observações que contribuíram no meu trabalho. Também estou profundamente agradecido ao Instituto de Matemática, Estatística e Computação Científica em geral, pela contribuição na minha formação matemática. Finalmente, a meus pais e irmãos, dos quais recebo um imenso apoio e carinho apesar da distância. Por fim, o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Neste trabalho estudaremos o máximo expoente de Lyapunov. Começaremos desenvolvendo certos resultados do análise e Teoria Ergódica, posteriormente fixaremos uma medida de probabilidade μ sobre $GL(m, \mathbb{R})$ e faremos uma filtração de $\mathbb{R}^m$ que nos levará a formar uma partição de $\mathbb{R}^m$, onde a cada elemento da partição associaremos um expoente de Lyapunov diferente, tudo isto com a finalidade de estudar a estabilidade do máximo expoente de Lyapunov. Em concreto queremos estabelecer uma condição suficiente em μ para que o máximo expoente de Lyapunov seja estável para pequenas perturbações sobre μ, i.e, seja $(\mu_n)_{n\in\mathbb{N}}$ uma sequência de medidas de probabilidade sobre $GL(m, \mathbb{R})$ e se temos que tal sequência converge fracamente a μ ($\mu_n \rightharpoonup \mu$ quando $n \longrightarrow \infty$), então $\beta(\mu_n) \longrightarrow \beta(\mu)$, onde $\beta(\mu_n)$ denota o máximo expoente de Lyapunov associado a μ_n. Esta dissertação está baseada no artigo de Furstenberg e Kifer "Random matrix products and measures on projective spaces".

Palavras-chave: máximo expoente de Lyapunov; condição suficiente e pequenas perturbações.

Abstract

In this work we study the maximal Lyapunov's exponent. We begin by developing certain results of Ergodic Theory and Analysis. Then we fix a probability measure μ on $GL(m, \mathbb{R})$ and make a partition of $\mathbb{R}^m$, where to each element of the partition we associate a different Lyapunov's exponent. We do that in order to establish a sufficient condition in μ so that the maximum Lyapunov's exponent is stable for small perturbations on μ, i.e., let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $GL(m, \mathbb{R})$ such that it converges weakly to μ ($\mu_n \rightharpoonup \mu$ when $n \longrightarrow \infty$), then $\beta(\mu_n) \longrightarrow \beta(\mu)$, where $\beta(\mu_n)$ denotes the maximum Lyapunov's exponent associated to μ_n. This dissertation is based on the seminal paper of Furstenberg and Kifer "Random matrix products and measures on projective spaces".

Keywords: maximum Lyapunov's exponent; sufficient condition; small perturbations.

Sumário

Introdução

São poucas e difíceis as equações diferenciais que podemos encontrar soluções analíticas. Ainda, que encontramos a solução da equação diferencial, a expressão da solução pode ser muito complicada, não ajudando muito a entender o que vai acontecer.
No final do século XIX, Henri Poincaré fez uma proposta, pensar nas equações diferencias de uma maneira diferente, o que hoje conhecemos como a teoria qualitativa das equações diferenciais. Em poucas palavras significa o seguinte: parar de focar toda a energia em tentar resolver a equação pela razão mencionada e, pelo contrário, concentrar todo nosso esforço em descrever o comportamento da solução, mesmo sem conhecer a expressão analítica da solução.
Seja a equação diferencial

$$\begin{aligned} x(0) &= x_0, \\ x^{'} &= F(x,t), \end{aligned} \tag{1}$$

com $F(x,t) \in \mathcal{C}^1$, com respeito à primeira variável e $x \in \mathbb{R}^m$.
Um aspecto qualitativo importante de (1) é o estudo da sensibilidade com respeito à condição inicial. O matemático russo Alexander M. Lyapunov desenvolveu um método de como medir tal sensibilidade, considerando a aproximação de Taylor de F com respeito à primeira variável em x_0 e provou certos resultados conclusivos sobre a sensibilidade de (1) quando a parte linear da aproximação não depende do tempo. Tentou ainda responder para o caso mais geral, quando a parte linear depende do tempo e isto o levou ao estudo do seguinte limite

$$\beta = \lim_{t\to\infty} \sup \frac{1}{t} \log\|f^t\|,$$

onde $f^t \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$ é o fluxo da equação diferencial ordinária. Este número, β, foi chamado o máximo expoente de Lyapunov. Com isso, veio o interesse do estudo dos expoentes de Lyapunov por si só.
Vamos ilustrar um exemplo simples do cálculo do máximo expoente de Lyapunov. Sejam $A_1, A_2, \cdots, A_n$ matrizes reais invertíveis de $m \times m$ e seja $p_1, p_2. \cdots, p_n$ números positivos tal que $p_1 + \cdots + p_n = 1$. Considere

$$B^{(n)} = B_{n-1} \cdots B_1 B_0, n \geq 1,$$

onde B_j é a variável aleatória identicamente distribuída, tal que

$$\textit{a probabilidade que } \{B_j = A_j\} \text{ é } \textit{igual a } p_i,$$

para todo $j \geq 0$ e $i = 1, \cdots, n$. Vamos a calcular o limite $\lim_{n\to\infty} \frac{1}{n} \log\|B^{(n)}\|$.
Considere o caso que $n = 2$, $p_1 = p_2 = \frac{1}{2}$ e

$$A_1 = \begin{pmatrix} 3 & 0 \\ 0 & 3^{-1} \end{pmatrix} \text{ e}$$

$$A_2 = \begin{pmatrix} 2^{-1} & 0 \\ 0 & 2 \end{pmatrix}.$$

Então

$$B^{(n)} = \begin{pmatrix} 3^{n_1}2^{-n_2} & 0 \\ 0 & 3^{-n_1}2^{n_2} \end{pmatrix}, onde$$

n_1 e n_2 são iguais ao número de vezes que A_1 e A_2 foram escolhidas respectivamente, e $n_1 + n_2 = n$.
Pela lei dos grandes números temos que

$$\frac{n_1}{n} = \frac{n_2}{n} = \frac{1}{2}.$$

Então o máximo expoente de Lyapunov é $\frac{1}{2} \log\left(\frac{3}{2}\right)$.

Fazemos uma resenha histórica dos resultados relevantes sobre o estudo do máximo expoente de Lyapunov. Em 1960 H. Furstenberg e H. Kesten provaram que, com certas condições específicas, o máximo expoente de Lyapunov existe e é igual a uma constante com probabilidade 1 (veja [FK60]). Este resultado pode ser visto como uma versão da Lei dos Grandes Números para o produto de matrizes aleatórias identicamente distribuídas.
Em 1982 Yuri Kifer e Eric Slud deram uma condição suficiente para a qual o máximo expoente de Lyapunov é estável para pequenas perturbações: que o subgrupo fechado de menor tamanho sobre $GL(m, \mathbb{R})$ com suporte na medida μ não tenha subespaço não trivial $L \subset \mathbb{R}^m$ que seja μ-invariante (veja [KS82]).
Em 1983 Furstenberg e Kifer estenderam o resultado anterior para o caso no qual o subgrupo fechado de menor tamanho sobre $GL(m, \mathbb{R})$ com suporte na medida μ tenha, no máximo, um subespaço não trivial $L \subset \mathbb{R}^m$ μ-invariante. Este é o resultado o qual detalhamos nesta dissertação (veja [FK83]).
No Capítulo I, introduzimos o necessário para o desenvolvimento da dissertação, sendo isto, definições e resultados de Análise Real, Teoria Ergódica, Probabilidade e provamos algumas propriedades do Operador de Markov.
No Capítulo II, começamos a estudar algumas relações com respeito ao maior expoente de Lyapunov. Em seguida, fazemos uma filtração de $\mathbb{R}^m$ para o Produto de Matrizes Aleatórias e provamos o resultado mais importante do capítulo que é o Teorema 2.2.16.
Os resultados principais estão no Capítulo III, onde efetivamente estudamos a estabilidade

do expoente de Lyapunov para pequenas perturbações na distribuição μ sobre $GL(m,\mathbb{R})$. Devido aos resultados prévios a abordagem é direita e acaba com o Teorema Principal 3.1.4. Por último, mostramos com um exemplo que o Teorema 3.1.4 não pode ser estendido para o caso em que o suporte da medida tenha dois subespaços de $\mathbb{R}^m$ μ-invariantes.

1 Preliminares

Neste primeiro capítulo apresentaremos alguns conceitos e resultados básicos que serão necessários ao desenvolvimento do trabalho. Serão assumidos conhecimentos elementares sobre medida, Probabilidade, Integração, Teoria Ergódica e Análise Funcional.

1.1 Algumas definições e resultados de Análise Real.

Nesta dissertação trabalhamos com espaços métricos completos munidos da sigma álgebra de Borel. O espaço das medidas de probabilidade será denotado como $\mathcal{M}_1$.
Seja M um espaço topológico, o espaço das funções contínuas de M a $\mathbb{R}$ é denotado como $\mathcal{C}(M)$, i.e., $\mathcal{C}(M) = \{f : M \longrightarrow \mathbb{R} | \ f \text{ é } \textit{contínua}\}$.
Uma definição presente durante todo o trabalho é:

Definição 1.1.1. Uma sequência $(\mu_n)_{n\geq 1}$ de probabilidades em um espaço mensurável $(E, \mathcal{F})$ converge fracamente a uma medida de probabilidade μ sobre $(E, \mathcal{F})$ se para toda função real f contínua e limitada tivermos:

$$\lim_n \int_E f(x)d\mu_n(x) = \int_E f(x)d\mu(x).$$

A convergência em probabilidade será denotada como $\mu_n \rightharpoonup \mu$ (convergência en fraca estrela).

O seguinte resultado permitirá caracterizar o espaço de medidas de probabilidades como um conjunto compacto na topologia fraca.

Teorema 1.1.2 (Teorema de Banach-Alaoglu sequencial). *Seja X um espaço métrico compacto. Então B', a bola unitária do dual é sequencialmente compacta na topologia fraca estrela.*

Demonstração. Veja [Bre10, Chap03]. □

O seguinte lema será usado na prova do Lema 1.2.9.

Lema 1.1.3 (Lema do Kronecker). *Seja $(b_n)_{n\in\mathbb{N}}$ uma sequência de números reais positivos crescente com $\lim_{n\to\infty} b_n = \infty$ e seja $(x_n)_{n\in\mathbb{N}}$ uma sequência de números reais, tal que:*

$$\sum_{n=0}^{\infty} x_n < \infty.$$

Então

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{k=0}^{n} b_k x_k = 0.$$

Demonstração. Veja [Wil91, Chap12]. □

O próximo resultado permitirá caracterizar o espaço das medidas de probabilidade como o espaço de funcionais positivos com norma 1.

Teorema 1.1.4 (Teorema de Representação de Riesz). *Seja X um espaço métrico compacto. Então para cada funcional linear limitado F sobre $\mathcal{C}(M)$ corresponde uma única medida regular de Borel com sinal ν em X tal que*

$$F(f) = \int f d\nu,$$

para toda f em $\mathcal{C}(M)$. Além disso, $\|F\| = |\nu|(X)$.

Demonstração. Veja [RF88, Chap14]. □

Agora temos todas a ferramentas para caracterizar o espaço das medidas de probabilidade.

Teorema 1.1.5. *Considere o espaço vetorial $E = (\mathcal{C}(M), \|.\|_\infty)$, onde M é um espaço métrico compacto. O espaço das medidas de probabilidade sobre M ($\mathcal{M}_1(M)$) é um espaço métrico compacto e convexo na topologia fraca estrela.*

Demonstração. Como M é compacto segue que $C(M)$ é separável [Bre10, Chap03]. Portanto, pelo Teorema de Banach-Alaoglu sequencial (1.1.2), a bola unitária do dual, denotada como B', é metrizável e compacta na topologia fraca estrela. Note que pode-se considerar toda medida finita como um funcional limitado e pelo Teorema (1.1.4), todo funcional limitado em $\mathcal{C}(M)$ pode ser representado como uma medida finita. Portanto, $\mathcal{M}_1(M)$ pode ser caracterizado como os funcionais lineares positivos com norma 1. Segue que $\mathcal{M}_1(M) \subset B'$. Por outro lado este conjunto munido como a topologia fraca estrela é de Hausdorff, basta provar que $\mathcal{M}_1(M)$ é fechado na topologia fraca estrela, isto é, se $\nu \in \overline{\mathcal{M}_1(M)}$, então $\nu \in \mathcal{M}_1(M)$.
Como $\nu \in \overline{\mathcal{M}_1(M)}$, existe $(\nu_k)_{k\in\mathbb{N}} \subset \mathcal{M}_1(M)$, tal que para toda $f \in \mathcal{C}(M)$:

$$\lim_{n\to\infty} \int f d\nu_n = \int f d\nu. \tag{1.1}$$

Seja $f \geq 0$, então $\displaystyle\int f d\nu_n \geq 0$, para todo $n \in \mathbb{N}$, pois ν_n são medidas positivas. Então

$$\int f d\nu = \lim_{n\to\infty} \int f d\nu_n \geq 0, \tag{1.2}$$

ou seja, ν é positiva.
Se $f = 1$, $\displaystyle\int 1 d\nu_n = 1$, para todo $n \in \mathbb{N}$, pois ν_n são medidas de probabilidades e por (1.1) temos que $\displaystyle\int 1 d\nu = 1$.
Portanto, $\nu \in \mathcal{M}_1(M)$. □

A seguinte observação e os dois próximos teoremas são usados como parte da prova do Lema 1.2.10.
Seja E um espaço vetorial real normado, seja L um subespaço de E. Definimos a relação de equivalência $\sim$ em E por:

$$f' \sim g' \Leftrightarrow f' - g' \in L$$

É claro que $(E/L, +, \cdot)$ é um espaço vetorial. Defina:

$$\begin{aligned} \|.\|_{\sim} : E/L &\longrightarrow \mathbb{R} \\ [f'] &\longmapsto \|[f']\|_{\sim} := \inf_{g' \in L} \|f' - g'\|_{\infty} \end{aligned}$$

Observação 1.1.6. Se L é fechado, a função $\|.\|_{\sim}$ definida acima é norma do espaço vetorial $(E/L, +, \cdot)$ e, se E é um espaço Banach, então $(E/L, +, \cdot)$ munido da norma $\|.\|_{\sim}$ será de Banach.

Considerando L como um subespaço fechado próprio de E. Definimos o operador (Projeção)

$$\begin{aligned} P_{pr} : E &\longrightarrow E/L \\ f' &\longmapsto [f'], \end{aligned}$$

e, nestas condições, o operador P_{pr} é linear, contínuo e com norma 1.

Teorema 1.1.7. *Seja L um subespaço fechado e próprio de E. Sejam T_1 e T_2 elementos dos duais de E e E/L respectivamente. Se L pertence ao núcleo de T_1, então $\|T_1\| = \|T_2\|$.*

Demonstração. Seja $x \in E$. Como $\|P_{pr}\| = 1$ e, T_2 e P_{pr} são operadores limitados, temos que

$$|T_1(x)| = |T_2 \circ P_{pr}(x)| \leq \|T_2\| \|P_{pr}(x)\| \leq \|T_2\| \|x\|.$$

Portanto, $\|T_1\| \leq \|T_2\|$.
Agora, considere $x \in E$ e $y \in L$. Como y está no núcleo de T_1, $|T_2[x]| = ||T_2 \circ P_{pr}(x)| = |T_1(x)| = |T_1(x+y)| \leq \|T_1\| \|x+y\|$ é válido para todo $y \in L$. Tomando o ínfimo dos $y \in L$, concluímos que $|T_2[x]| \leq \|T_1\| \|x\|_{\sim}$. Portanto, $\|T_2\| \leq \|T_1\|$ e concluímos a prova. □

Teorema 1.1.8 (**Teorema de Hahn-Banach para espaços normados**). *Seja X um espaço vetorial normado e $x_0 \neq 0$ um elemento de X. Então existe um funcional linear limitado $\hat{f}$ sobre X tal que*

$$\|\hat{f}\| = 1 \qquad e \qquad \hat{f}(x_0) = \|x_0\|$$

Demonstração. Veja [Kre78, Chap04]. □

Vamos trabalhar com conjuntos convexos e precisamos da definição do ponto extremal e o famoso resultado de "Krein Milman"sobre conjuntos convexos e compactos.

Definição 1.1.9. *Um ponto extremal* de um conjunto A é um ponto $x \in A$, com a propriedade que se $x = \theta y + (1-\theta)z$ com $y, z \in A$ e $\theta \in [0,1]$, então $y = x$ e $z = x$. O conjunto dos pontos extremais de A será denotado $extA$.

Teorema 1.1.10 (**Teorema de Krein-Milman**)**.** *Seja K um subconjunto não vazio, compacto e convexo de um conjunto X que é um espaço vetorial localmente compacto. Então $extK$ é não vazio e todo elemento de K pode ser escrito como soma convexa de seus extremos.*

Demonstração. Veja [Sch86, Chap05]. □

Outro tópico importante para o desenvolvimento da dissertação é o conceito de medida de convolução estacionária. As duas definições seguintes nos darão a linguagem necessária para trabalhar com isso.

Definição 1.1.11. Seja G um grupo topológico. Dizemos que G age sobre um espaço topológico B se existe uma função contínua

$$\begin{aligned} \Theta : & G \times B \longrightarrow B \\ & (g, x) \longmapsto g.x, \end{aligned}$$

tal que $(g_1 g_2).x = g_1.(g_2 x)$ para todo $g_1, g_2 \in G$ e $x \in B$.

Definição 1.1.12. Sejam μ e ν medidas de probabilidade em G e B, respectivamente, onde G é grupo topológico agindo sobre o espaço topológico B. Então o produto de convolução é definido por:

$$(\mu * \nu)(E) := \int \nu(g^{-1}E) d\mu(g),$$

para todo $E \subset B$ mensurável.

Definição 1.1.13. Dizemos que ν é μ-estacionária se $\mu * \nu = \nu$.

1.2 Algumas definições e resultados de probabilidade.

Dados dois espaços mensuráveis $(X, \mathcal{F})$ e $(Y, \mathcal{G})$, dizemos que uma função $f : (X, \mathcal{F}) \longrightarrow (Y, \mathcal{G})$ é mensurável se ela preservar a estrutura de mensurabilidade no sentido de imagens inversas de subconjuntos mensuráveis em Y são subconjuntos mensuráveis em X.
Em teoria de probabilidade, funções mensuráveis recebem o nome especial de *variáveis aleatórias.*

Definição 1.2.1 (Processo Estocástico). Dizemos que

$$\begin{aligned} X : T \times \Omega &\longrightarrow M \\ (t, w) &\longmapsto X_t(w) \end{aligned}$$

é um processo estocástico se para todo $t \in T$, X_t é uma função mensurável. Quando T é contável, dizemos que o processo é discreto, caso contrário, dizemos que é contínuo.

Para $w \in \Omega$ fixo, $t \longmapsto X_t(w)$ é a trajetória de w e para t fixo, $w \longmapsto X_t(w)$ é o estado do processo no momento t.
Observações: Nesta dissertação só se trabalha com processos estocásticos discretos ($T = \mathbb{N}$), e $\Omega = M$, onde M é um espaço métrico compacto. O conjunto das medidas de probabilidade sobre M será denotado como $\mathcal{M}_1(M)$ (posteriormente faremos uma extensão de M para um espaço métrico completo).
Agora, definiremos um operador que agirá entre os espaços de medida de probabilidade de M. Este será o operador adjunto do operador de Markov.
Definimos a dinâmica aleatória Λ, como uma aplicação contínua que associa a cada ponto de M uma medida de probabilidade:

$$\begin{aligned} \Lambda : M &\longrightarrow \mathcal{M}_1(M) \\ x &\longmapsto \mu_x. \end{aligned}$$

Sua continuidade significa que,

$$se \;\; x_n \longrightarrow x \;\; temos\ que \;\; \mu_{x_n} \rightharpoonup \mu_x. \tag{1.3}$$

Agora definimos o **operador de Markov**:

$$\begin{aligned} \mathcal{T} : \mathcal{C}(M, \mathbb{R}) &\longrightarrow \mathcal{C}(M, \mathbb{R}) \\ f &\longmapsto \mathcal{T} f, \end{aligned} \tag{1.4}$$

com $\mathcal{T} f(x) = \displaystyle\int f(y) d\mu_x(y)$.

Afirmação: o operador de Markov está bem definido, isto é $\mathcal{T} f \in \mathcal{C}(M)$ para toda $f \in \mathcal{C}(M)$. É suficiente provar que se $x_n \longrightarrow x$ e $f \in \mathcal{C}(M)$, então $\mathcal{T} f(x_n) \longrightarrow \mathcal{T} f(x)$. Mas (1.3) implica que $\displaystyle\lim_{n\to\infty} \int f d\nu_{x_n} = \int f d\nu_x$, que pela definição de $\mathcal{T}$ é equivalente a $\displaystyle\lim_{n\to\infty} \mathcal{T} f(x_n) = \mathcal{T} f(x)$. Portanto, $\mathcal{T} f \in \mathcal{C}(M)$.

Observação 1.2.2. É claro que o operador de Markov é um operador positivo, i.e., se $f \geq 0$ temos que $\mathcal{T} f \geq 0$.

O operador adjunto de (1.4) é

$$\begin{aligned} \mathcal{T}^* : \mathcal{C}(M)' &\longrightarrow \mathcal{C}(M)' \\ L &\longmapsto \mathcal{T}^* L, \end{aligned} \tag{1.5}$$

com $\mathcal{T}^*L(g) = L(\mathcal{T}(g))$ para toda $g \in \mathcal{C}(M)$.

Observação 1.2.3. Pelo teorema de Representação de Riesz, L pode ser representado como

$$\begin{aligned} L : \mathcal{C}(M) &\longrightarrow \mathbb{R} \\ g &\longmapsto \int g(x) d\mu_L(x), \end{aligned}$$

O operador (1.5) restrito à $\mathcal{M}_1(M)$ é chamado **o operador adjunto de Markov** que por simplicidade será denotado $\mathcal{T}^*$ e

$$\begin{aligned} \mathcal{T}^* : \mathcal{M}_1(M) &\longrightarrow \mathcal{M}_1(M) \\ \nu &\longmapsto \mathcal{T}^* \nu, \end{aligned}$$

com $\mathcal{T}^*\nu(E) = \mathcal{T}^*\nu(\mathcal{X}_E) = \displaystyle\int \mathcal{T}\mathcal{X}_E(y) d\nu(y)$, (observação 1.2.3).

Afirmação: $\mathcal{T}^*$ está bem definido, i.e., se $\nu \in \mathcal{M}_1(M)$ temos que $\mathcal{T}^*\nu \in \mathcal{M}_1(M)$.
Como $\mathcal{T}$ é um operador positivo temos que $\mathcal{T}^*\nu$ é uma medida positiva, resta provar que é de probabilidade. De fato,

$$\mathcal{T}^*\nu(M) = \int \mathcal{T}\mathcal{X}_M(y) d\nu(y) = \int \left(\int \mathcal{X}_M(z) d\nu_y(z) \right) d\nu(y) = \int 1 \, d\nu(y) = 1.$$

Portanto, $\mathcal{T}^*$ está bem definida.

Podemos fazer uma interpretação dinâmica do operador de Markov. Como um caso especial, considere $f = \mathcal{X}_E$, onde $E \subset M$ mensurável, pensemos em $x \in M$ como a posição de uma partícula, então $\mathcal{T}\mathcal{X}_E(x)$ pode ser interpretado como a probabilidade que após uma unidade de tempo a partícula esteja no conjunto E, além disso $\mu_x^t(E) = \displaystyle\int \cdots \int \chi_E(y_t) d\mu_{y_{t-1}}(y_t) \cdots d\mu_{y_1}(y_2) d\mu_x(y_1)$ pode ser interpretado como a probabilidade de que após t unidades de tempo a partícula esteja no conjunto E.

Os pontos fixos associados ao operador adjunto de Markov, ou seja, os pontos para os quais $\mathcal{T}^*\nu = \nu$, são chamados medidas estacionárias associadas a $\mathcal{T}$.
As medidas ν que são pontos fixos do operador de Markov diz que a média da probabilidade que para qualquer $x \in M$ esteja em um conjunto mensurável coincide com a medida do conjunto.

Teorema 1.2.4. *Seja M um espaço métrico compacto e $\mathcal{T}$ um operador de Markov. Então o conjunto de medidas estacionárias associadas a $\mathcal{T}$ é não vazio.*

Demonstração. Seja $\nu \in \mathcal{M}_1(M)$, considere $\mathcal{T}^{k*}\nu = \overbrace{\mathcal{T}^* \circ \cdots \mathcal{T}^* \circ (\mathcal{T}^*\nu)}^{k \text{ vezes}}$. Defina:

$$\nu_n = \frac{1}{n+1}\sum_{k=0}^{n} \mathcal{T}^{k*}\nu$$

É claro que $\nu_n \in \mathcal{M}_1(M)$ para todo $n \in \mathbb{N}$. Como o espaço $\mathcal{M}_1(M)$ é sequencialmente compacto pelo Teorema 1.1.2, existe uma subsequência $(\nu_{n_k})_{n_k \in \mathbb{N}}$ e $\nu' \in \mathcal{M}_1(M)$ tal que $\nu_{n_k} \rightharpoonup \nu'$.

Afirmação: ν' é estacionária.

Seja $f \in \mathcal{C}(M)$ e dado $\varepsilon > 0$, para n_k suficientemente grande temos que:

$$\left|\frac{1}{n_k+1}\sum_{j=0}^{n_k}\int \mathcal{T}^j \circ f d\nu - \int f d\nu'\right| < \frac{\varepsilon}{2} \tag{1.6}$$

Como $\mathcal{T}$ é uma aplicação contínua, seu adjunto $\mathcal{T}^*$ é contínuo. Então:

$$\mathcal{T}^*\nu' = \mathcal{T}^*\left(\lim_{k\to\infty}\frac{1}{n_k+1}\sum_{j=0}^{n_k}\mathcal{T}^{j*}\nu\right) = \lim_{k\to\infty}\frac{1}{n_k+1}\sum_{j=1}^{n_k+1}\mathcal{T}^{j*}\nu \tag{1.7}$$

Agora observe que

$$\begin{aligned}
&\left|\frac{1}{n_k+1}\sum_{j=0}^{n_k}\int f d(\mathcal{T}^{j*}\nu) - \frac{1}{n_k+1}\sum_{j=1}^{n_k+1}\int f(\mathcal{T}^{j*}\nu)\right| \\
&\qquad = \left|\frac{1}{n_k+1}\sum_{j=0}^{n_k}\left(\int \mathcal{T}^j \circ f d\nu - \int \mathcal{T}^{j+1}\circ f\, d\nu\right)\right| \\
&\qquad = \frac{1}{n_k+1}\left|\int f\, d\nu - \int \mathcal{T}^{n_k+1}\circ f\, d\nu\right| \\
&\qquad \leq \frac{1}{n_k+1}\left(\left|\int f\, d\nu\right| + \left|\int \cdots \int f(y_{n_k+2})\, d\nu_{y_{n_k+1}}(y_{n_k+2})\cdots d\nu_{y_1}(y_2) d\nu_x(y_1)\right|\right) \\
&\qquad \leq \frac{2}{n_k+1}\max_{y\in M}|f(y)|.
\end{aligned}$$

Esta última expressão é menor que $\varepsilon/2$ para n_k suficientemente grande. Juntando este fato com (1.6), concluímos que

$$\left|\frac{1}{n_k+1}\sum_{j=1}^{n_k+1}\int \mathcal{T}^j \circ f\, d\nu - \int f\, d\nu'\right| < \varepsilon,$$

para todo n_k suficientemente grande. Isto significa que

$$\frac{1}{n_k+1}\sum_{j=1}^{n_k+1}\mathcal{T}^{j*}\nu \rightharpoonup \nu',$$

quando $n_k \longrightarrow \infty$. Mas (1.7) significa que esta mesma sequência converge para $\mathcal{T}^*\nu'$. Por unicidade do limite na convergência fraca, segue que $\mathcal{T}^*\nu' = \nu'$. $\square$

Uma interpretação dinâmica de uma medida de Markov estacionária é, suponha que existe $x \in M$ tal que $\mu_x = \nu$, onde ν é uma medida de Markov estacionária, então consideremos uma partícula que está em x e queremos saber a probabilidade que depois de k passos a partícula está em um conjunto mensurável E, está é exatamente o volumem de E com respeito a ν.

O seguinte resultado técnico sobre operador de Markov é utilizado para provar o Lema 1.2.9.

Lema 1.2.5. *Seja $g \in \mathcal{C}(M)$, então $\|\mathcal{T}g\|_{L_2} \leq \max_{y \in M}|g(y)|$.*

Demonstração.

$$\begin{aligned}
\|\mathcal{T}g\|_{L_2}^2 &= \int |\mathcal{T}g(x)|^2 dP(x) \\
&= \int \left| \int g(y) d\mu_x(y) \right|^2 dP(x), \ \textit{pela desigualdade de Hölder temos} \\
&\leq \int \left\{ \left[\int g(y)^2 d\mu_x(y) \right] \left[\int 1 d\mu_x(y) \right] \right\} dP(x) \\
&\leq \iint \max_{y \in M} g(y)^2 d\mu_x(y)\ dP(x) \\
&= \left(\max_{y \in M} |g(y)| \right)^2
\end{aligned}$$

□

Um tipo de processo estocástico que nos é de muito interesse é o processo de Markov. Imagine que há um processo estocástico $(Z_n)_{n \in \mathbb{N}}$ e que o problema é saber qual é a probabilidade de que $Z_{n+1} \in E$ dada a trajetória finita $Z_1, Z_2, \cdots Z_n$ e $E \subset M$ mensurável. O tipo de processo estocástico no qual esta probabilidade não depende da trajetória, mas depende apenas de seu último estado é conhecido como o processo de Markov.

Definição 1.2.6 (Processo de Markov). Um processo estocástico $(Z_n)_{n \in \mathbb{N}}$ é dito de Markov se:

$$P\{Z_{n+1} \in E | Z_n, \cdots, Z_1\} = P\{Z_{n+1} \in E | Z_n\} = \mu_{Z_n}(E) = (\mathcal{T}\mathcal{X}_E)(Z_n),$$

para todo $E \subset M$ mensurável.

Outro tipo de processo estocástico que será mencionado na dissertação é chamado processo de Martingale, mas, antes de defini-lo é necessário definir uma filtração: seja $(M, \mathcal{A}, P)$ um espaço de probabilidade completo. Uma filtração $\mathbb{F} = \{\mathcal{A}_n,\ n \in \mathbb{N}\}$ é uma sequência de sub-σ-álgebras de $\mathcal{A}$, satisfazendo $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \cdots \subset \mathcal{A}$. Um processo estocástico $(Z_n)_{n \in \mathbb{N}}$ é dito *adaptado à filtração* $\mathbb{F}$ se Z_n é $\mathcal{A}_n$ mensurável.

Definição 1.2.7 (Processo de Martingale). Dada uma filtração $\mathbb{F}$, um processo estocástico $(Z_n)_{n\in\mathbb{N}}$ é de Martingale se:
1.- $(Z_n)_n$ é adaptado à $\mathbb{F}$,
2.- $\mathbb{E}[|Z_n|] < \infty$, para todo $n \in \mathbb{N}$ e
3.- $\mathbb{E}[Z_n|\mathcal{A}_{n-1}] = Z_{n-1}$, para todo $n \in \mathbb{N}$.

Teorema 1.2.8. *Seja $(Z_n)_{n\in\mathbb{N}}$ um processo estocástico de Martingale, tal que $Z_n \in L_2$ para toda $n \in \mathbb{N}$. Então se*

$$\sum_{k=1}^{\infty} \mathbb{E}(Z_{k+1} - Z_k)^2 < \infty,$$

temos que $\lim_{n\to\infty} Z_n = Z_\infty$ para quase todo ponto em L_2.

Demonstração. Veja [Wil91, chap12]. □

Os dois seguintes resultados são necessários para a prova do Lema Central 1.2.14.
O primeiro lema nos diz que as médias temporais ao longo de um processo de Markov de qualquer função contínua coincidem com a média temporal da imagem de tal função pelo operador de Markov.

Lema 1.2.9. *Seja $(Z_n)_{n\in\mathbb{N}}$ um processo de Markov em M e seja $g \in \mathcal{C}(M)$, então para $f = g - \mathcal{T}g$ temos com probabilidade um:*

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} f(Z_k) = 0. \tag{1.8}$$

Demonstração. Considere $W_{n+1} = \sum_{k=0}^{n} \frac{\mathcal{T}g(Z_k) - g(Z_{k+1})}{k+1}$. Temos que:

$$\begin{aligned}
\mathbb{E}(W_{k+1} \mid Z_0,\ Z_1, \cdots, Z_n) &= \mathbb{E}\left(\sum_{k=0}^{n} \frac{\mathcal{T}g(Z_k) - g(Z_{k+1})}{k+1} \mid Z_0,\ Z_1, \cdots, Z_n\right) \\
&= \sum_{k=0}^{n} \mathbb{E}\left(\frac{\mathcal{T}g(Z_k) - g(Z_{k+1})}{k+1} \mid Z_0,\ Z_1, \cdots, Z_n\right)
\end{aligned}$$

como Z_j é constante pra $j = 0, 1, \cdots n$, temos que

$$\begin{aligned}
\sum_{k=0}^{n} \mathbb{E}\left(\frac{\mathcal{T}g(Z_k) - g(Z_{k+1})}{k+1} \mid Z_0,\ Z_1, \cdots, Z_n\right) &= \sum_{k=0}^{n-1} \frac{\mathcal{T}g(Z_k) - g(Z_{k+1})}{k+1} \\
&+ \frac{1}{n+1}\mathbb{E}(\mathcal{T}g(Z_n) - g(Z_{n+1}) \mid Z_0,\ Z_1, \cdots, Z_n) \\
&= W_n + \frac{1}{n+1}\mathcal{T}(g(Z_n)) - \frac{1}{n+1}\mathbb{E}(g(Z_{n+1}) \mid Z_0,\ Z_1, \cdots, Z_n),
\end{aligned}$$

como $(Z_n)_{n\in\mathbb{N}}$ é um processo de Markov temos que

$$\mathbb{E}(W_{k+1} \mid Z_0,\ Z_1, \cdots, Z_n) = W_n + \frac{1}{n+1}\mathcal{T}(g(Z_n)) - \frac{1}{n+1}\mathcal{T}(g(Z_n)) = W_n,$$

então, $(W_n)_{n\in\mathbb{N}}$ é um processo de martingale.
Afirmação: $\sum_{k>1}\|W_{k+1}-W_k\|^2_{L^2}<\infty$.
Pela desigualdade triangular e pelo Lema 1.2.5 temos:

$$\begin{aligned}\sum_{k>1}\|W_{k+1}-W_k\|^2_{L^2} &= \sum_{k\geq 1}\left\|\frac{Tg(Z_k)-g(Z_{k+1})}{k+1}\right\|^2_{L^2}\\ &\leq \sum_{k\geq 1}\left(\frac{\|Tg(Z_k)\|_{L^2}+\|g(Z_{k+1})\|_{L^2}}{k+1}\right)^2\\ &\leq \sum_{k\geq 1}\left(\frac{2\max_{y\in M}|g(y)|}{k+1}\right)^2\\ &< \infty\end{aligned}$$

Portanto, pelo Teorema 1.2.8, temos que $\lim_{n\to\infty} W_n$ existe em um conjunto de medida total. Escolhendo a sequência $(k)_{k\in\mathbb{N}}$ e pelo Lema do Kronecker 1.1.3 temos que

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}(k+1)\frac{\mathcal{T}g(Z_k)-g(Z_{k+1})}{k+1}=\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}(\mathcal{T}g(Z_k)-g(Z_{k+1}))=0 \quad (1.9)$$

Por outro lado, observamos que:

$$\frac{1}{n+1}\sum_{k=0}^{n}\mathcal{T}g(Z_k)-g(Z_{k+1})-\frac{1}{n+1}\sum_{k=0}^{n}f(Z_k)=\frac{1}{n+1}(g(Z_{n+1})-g(Z_0))$$

Agora, já que g é continua e M é compacto, segue a conclusão do lema. □

O seguinte resultado nos permitirá escrever de forma conveniente qualquer função contínua.

Lema 1.2.10. *Seja $f\in\mathcal{C}(M)$ não negativa. Então, para todo $\varepsilon>0$, podemos escrever*

$$f=\mathcal{T}g-g+h, \quad (1.10)$$

com $g,h\in\mathcal{C}(M)$, e h que satisfaz:

$$\|h\|_\infty\leq\sup\left\{\int fd\nu \mid \nu\in\mathcal{M}_1(M)\ com\ \mathcal{T}^*\nu=\nu\right\}+\varepsilon. \quad (1.11)$$

Demonstração. Observe que pelo Teorema 1.2.4 o lado direito de (1.11) é não vazio e, portanto, o cálculo do sup tem sentido.
Como M é um espaço métrico compacto, o espaço $E=(\mathcal{C}(M),\|.\|_\infty)$ é de Banach. Lembre-se que, se $g\in\mathcal{C}(M)$, então $\mathcal{T}g\in\mathcal{C}(M)$. Sendo assim, $\mathcal{L}=\{\mathcal{T}g-g,\ onde\ g\in\mathcal{C}(M)\}$ é subconjunto de E.
Consideraremos dois casos: o caso trivial, quando $f\in\overline{\mathcal{L}}$, e o caso contrário.
Caso I: Seja $d(f,\mathcal{L})=\inf_{g\in\mathcal{C}(M)}\|f-(\mathcal{T}g-g)\|_\infty=0$.
Então existe $(g_n)_{n\in\mathbb{N}}\subset E$ tal que, dado $\epsilon>0$, existe $n_0\in\mathbb{N}$:

$$\|f-\mathcal{T}g_n-g_n\|_\infty<\epsilon \qquad \forall n>n_0. \quad (1.12)$$

Tomando $h = f - \mathcal{T}g_n - g_n$ para algum n que satisfaça (1.12), segue a conclusão do lema.
Caso II: Seja

$$d(f, L) = \inf_{g \in \mathcal{C}(M)} \|f - (\mathcal{T}g - g)\|_\infty = \delta > 0. \tag{1.13}$$

Considere $\overline{\mathcal{L}}$ como o fecho de $\mathcal{L}$ com respeito à norma infinita. Veja que $d(f, \mathcal{L}) = d(f, \overline{\mathcal{L}}) = \delta$. Como $\overline{\mathcal{L}}$ é fechado e E é de Banach, pela Observação 1.1.6 segue que $E/\overline{\mathcal{L}}$ munido da norma $\|.\|_\sim$ é de Banach.
Pelo teorema de Hahn-Banach 1.1.8, temos que existe um funcional linear continuo sobre o espaço $(E/\overline{\mathcal{L}}, \|.\|_\sim)$, denotado como $\mathcal{R}_1$, tal que

$$\mathcal{R}_1([f]) = \delta \qquad \text{e} \qquad \|\mathcal{R}_1\| = 1, \tag{1.14}$$

e note que, como caso particular, temos que $\mathcal{R}_1([g']) = 0$ para todo $g' \in \overline{\mathcal{L}}$, pois a classe $[g'] = [0]$.
Definimos o funcional linear $\mathcal{R}_2 := \mathcal{R}_1 \circ P_{pr}$, que é continuo, pois é composição de dois operadores contínuos. Pelo Teorema 1.1.7 o operador $\mathcal{R}_2$ tem a mesma norma do operador $\mathcal{R}_1$, portanto $\mathcal{R}_2$ tem norma 1.
Agora, pelo Teorema 1.1.4, existe uma única medida regular de Borel λ com sinal em M tal que, para todo $g \in \mathcal{C}(M)$, temos:

$$\mathcal{R}_2(g) = \int g \, d\lambda, \ \textit{com} \ \|\lambda\| = \|\mathcal{R}_2\| = 1. \tag{1.15}$$

Observação 1.2.11. Para todo $g \in \mathcal{C}(M)$, temos que $\mathcal{T}g - g \in \mathcal{L}$. Portanto, $\mathcal{R}_2(\mathcal{T}g - g) = \mathcal{R}_2([\mathcal{T}g - g]_{E/\overline{\mathcal{L}}}) = \mathcal{R}_2([0]_{E/\overline{\mathcal{L}}}) = 0$. Daqui, segue por (1.15) que $\int g \, d\lambda = \int \mathcal{T}g \, d\lambda = \int g \, d(\mathcal{T}^*\lambda)$, isto é, a probabilidade de transição de Markov é invariante $(\lambda = \mathcal{T}^*\lambda)$.

Observação 1.2.12. Por (1.13) e (1.14) temos que $\delta = \mathcal{R}_2(f) = \int f \, d\lambda$. Por outro lado, λ pode ser escrito como a subtração de duas medidas positivas [Bar95, cap08], isto é $\lambda = \lambda^+ - \lambda^-$ e, pela observação 1.2.11, temos que $\lambda = \mathcal{T}^*\lambda^+ - \mathcal{T}^*\lambda^- = \lambda^+ - \lambda^-$. Como sabemos que o operador de Markov é um operador positivo, temos que $\mathcal{T}^*\lambda^+$ e $\mathcal{T}^*\lambda^-$ são medidas positivas, isto implica que $\mathcal{T}^*\lambda^+ = \lambda^+$ e $\mathcal{T}^*\lambda^- = \lambda^-$ [Bar95, cap08].

Observação 1.2.13. Sabemos que $1 = \|\lambda\| = \|\lambda^+\| + \|\lambda^-\|$ e $\|\lambda^+\| > 0$, pois para f não negativa do enunciado temos que $\int f \, d\lambda = \delta > 0$, então $0 < \|\lambda^+\| \leq 1$, defina $\nu_0 = \dfrac{\lambda^+}{\|\lambda^+\|}$. É claro que $\mathcal{T}^*\nu_0 = \nu_0$. Por outro lado $\delta = \int f \, d\lambda = \int f \, d\lambda^+ - \int f \, d\lambda^- \leq \|\lambda^+\| \int f \, d\nu_0 \leq \int f \, d\nu_0$.

Pela definição de δ, dado $\varepsilon > 0$, existe $Tg - g \in \mathcal{L}$ tal que

$$\|f - (\mathcal{T}g - g)\|_\infty \leq \delta + \varepsilon$$

Definindo $h = f - (\mathcal{T}g - g)$, pelas observações 1.2.12 e 1.2.13, temos que

$$\|h\|_\infty \leq \delta + \varepsilon \leq \left\{ \int f \, d\nu_0 \mid \textit{com} \ \mathcal{T}^*\nu = \nu \ \textit{e} \ \|\nu\| = 1 \right\} + \varepsilon$$

e, portanto,

$$\|h\|_\infty \leq \sup\left\{\int f d\nu \mid \nu \in \mathcal{M}_1(M) \text{ } com \text{ } \mathcal{T}^*\nu = \nu\right\} + \varepsilon.$$

□

Agora temos todas as ferramentas para provar o Lema Central, que nos diz que as médias temporais de uma função contínua ao longo de um processo de Markov estão limitadas pela média espacial da função quando consideramos apenas as medidas estacionárias pelo operador de Markov.

Teorema 1.2.14 (**Lema Central**). *Seja Λ uma dinâmica aleatória com o correspondente operador de Markov $\mathcal{T}$, seja $(Z_n)_{n\in\mathbb{N}}$ um processo de Markov sobre M e seja $f \in \mathcal{C}(M)$. Temos com probabilidade um*

$$\lim_{n\to\infty} \sup \frac{1}{n+1}\sum_{k=0}^{n} f(Z_k) \leq \sup\left\{\int f d\nu \mid \nu \in \mathcal{M}_1(M) \text{ } com \text{ } \mathcal{T}^*\nu = \nu\right\}. \tag{1.16}$$

Demonstração. Caso I. Considere $f \in \mathcal{C}(M)$ não negativa.
Dado $\varepsilon > 0$, considere $g, h \in \mathcal{C}(M)$ como no Lema 1.2.10. Segue dos lemas 1.2.9 e 1.2.10 que:

$$\begin{aligned}
\lim_{n\to\infty} \sup \frac{1}{n+1}\sum_{k=0}^{n} f(Z_k) &= \lim_{n\to\infty} \sup \frac{1}{n+1}\sum_{k=0}^{n} (\mathcal{T}g - g + h)(Z_k) \\
&\leq \lim_{n\to\infty} \sup \frac{1}{n+1}\sum_{k=0}^{n} (\mathcal{T}g - g)(Z_k) + \lim_{n\to\infty} \sup \frac{1}{n+1}\sum_{k=0}^{n} h(Z_k) \\
&= \lim_{n\to\infty} \sup \frac{1}{n+1}\sum_{k=0}^{n} h(Z_k) \\
&\leq \|h\|_\infty \\
&\leq \sup\left\{\int f d\nu \mid \nu \in \mathcal{M}_1(M) \text{ } com \text{ } \mathcal{T}^*\nu = \nu\right\} + \varepsilon
\end{aligned}$$

Como ε é arbitrário temos a conclusão do enunciado para f não negativa.
Caso II. Seja $f \in \mathcal{C}(M)$ não necessariamente não negativa.
Como e M é compacto existe uma constante positiva $K \in \mathbb{R}$ tal que $f(x) \leq K$ para todo $x \in M$. Defina:

$$\begin{aligned}
F : M &\longrightarrow \mathbb{R} \\
x &\longmapsto K + f(x)
\end{aligned}$$

É claro que F é não negativa e contínua, então, pelo Caso I, tem-se:

$$\lim_{n\to\infty} \sup \frac{1}{n+1}\sum_{j=0}^{n} F(Z_j) \leq \sup\left\{\int F d\nu \mid \nu \in \mathcal{M}_1(M) \text{ } com \text{ } \mathcal{T}^*\nu = \nu\right\}. \tag{1.17}$$

Observe que

$$\limsup_{n\to\infty} \frac{1}{n+1}\sum_{j=0}^{n} F(Z_j) = \limsup_{n\to\infty} \frac{1}{n+1}\sum_{j=0}^{n}(K+f(Z_j)) = K + \limsup_{n\to\infty} \frac{1}{n+1}\sum_{j=0}^{n} f(Z_j). \tag{1.18}$$

Por outro lado,

$$\sup\left\{\int F d\nu \mid \nu \in \mathcal{M}_1(M) \text{ com } \mathcal{T}^*\nu = \nu\right\} \tag{1.19}$$

$$= \sup\left\{\int (f(x)+K)d\nu(x) \mid \nu \in \mathcal{M}_1(M) \text{ com } \mathcal{T}^*\nu = \nu\right\}$$

$$= \sup\left\{K + \int f(x)d\nu(x) \mid \nu \in \mathcal{M}_1(M) \text{ com } \mathcal{T}^*\nu = \nu\right\}$$

$$= K + \sup\left\{\int f d\nu \mid \nu \in \mathcal{M}_1(M) \text{ com } \mathcal{T}^*\nu = \nu\right\}. \tag{1.20}$$

Juntando (1.17), (1.18) *e* (1.19), segue a conclusão do teorema. □

O seguinte resultado técnico nos dá uma condição suficiente e necessária para a convergência em probabilidade sobre um espaço produto verificando apenas que seus marginais convergem em probabilidade.

Teorema 1.2.15. *Sejam X e Y espaços métricos separáveis e $(\nu_n)_{n\in\mathbb{N}}$ e $(\mu_n)_{n\in\mathbb{N}}$ sequências de medidas de probabilidade sobre X e Y, respectivamente, com $\nu_n \rightharpoonup \nu$ e $\mu_n \rightharpoonup \mu$. Então, o produto das medidas $\mu_n \otimes \nu_n \rightharpoonup \nu \otimes \mu$ quando $n \longrightarrow \infty$ no espaço $X \times Y$.*

Demonstração. Veja [Par05, Cap 03]. □

1.3 O máximo exponente de Lyapunov.

O objetivo desta seção é definir nosso objeto de estudo, o maior exponente de Lyapunov, e também lembrar o resultado de Fustenberg e Kesten sobre tal expoente.

Definição 1.3.1. Considere uma sequência $(X_n)_{n\in\mathbb{N}}$ de variáveis aleatórias independentes e identicamente distribuídas com valores em $GL(m,\mathbb{R})$ (o grupo das matrizes reais invertíveis $m\times m$), com distribuição μ. Se $\mathbb{E}[\log^+(X_1)] < \infty$, o maior expoente de Lyapunov associado a μ é o elemento $\gamma \in \mathbb{R}\cup\{-\infty\}$ definido por

$$\gamma = \lim_{n} \frac{1}{n}\mathbb{E}[\log\|X_n X_{n-1}\cdots X_1\|]$$

O Teorema de Furstenberg e Kesten traz um resultado que pode ser interpretado como a Lei dos Grandes Números para produtos de matrizes aleatórias independentes e identicamente distribuídas.

Teorema 1.3.2 (Teorema de Furstenberg e Kesten). *Seja $(X_n)_{n\in\mathbb{N}}$ uma sequência de matrizes aleatórias independentes e identicamente distribuídas em $GL(m,\mathbb{R})$ com distribuição comum μ. Se $\mathbb{E}[\log^+\|X_1\|] < \infty$, então*

$$\lim_n \frac{1}{n}\log\|X_n X_{n-1}\cdots X_1\| = \gamma$$

existe para $\mu^{\mathbb{N}}$ quase todo ponto .

Demonstração. Veja [FK60]. □

1.4 Algumas definições e resultados de Teoria Ergódica.

Nesta subseção vamos enunciar algumas definições e resultados sobre Teoria Ergódica. Uma definição que nos seguirá durante toda a dissertação é o conceito de medida invariante por uma aplicação.

Definição 1.4.1. Seja $f : M \longrightarrow M$ uma transformação mensurável e μ uma medida de probabilidade. Se diz que μ é invariante pela transformação f se

$$\mu(B) = \mu(f^{-1}(B))$$

para todo conjunto mensurável $B \subset M$.

Uma questão interessante associada a uma medida invariante por uma aplicação abordada no Teorema Ergódico de Birkhoff é a relação entre as médias temporais e as médias espaciais.

Teorema 1.4.2 (Teorema Ergódico de Birkhoff). *Seja $f : M \longrightarrow M$ uma transformação mensurável e μ uma medida de probabilidade invariante por f. Dada qualquer função integrável $\varphi : M \longrightarrow \mathbb{R}$, o limite*

$$\overline{\varphi}(x) = \lim_n \frac{1}{n}\sum_{j=0}^{n-1}\varphi(f^j(x))$$

existe em μ-quase todo ponto $x \in M$. Além disso, a função $\overline{\varphi}$ definida desta forma é integrável e satisfaz

$$\int \overline{\varphi}(x)d\mu(x) = \int \varphi(x)d\mu(x).$$

Demonstração. Veja [OV14, Chap03]. □

Um caso especial sobre as médias temporais é quando são contantes para quase todo ponto. Primeiramente, definimos o conceito do conjunto invariante para depois enunciar uma condição necessária e suficiente para que as médias temporais sejam constantes.

Definição 1.4.3. Seja $f : M \longrightarrow M$ uma transformação mensurável e μ uma medida de probabilidade invariante por f. Dizemos que um conjunto mensurável $B \subset M$ é invariante se

$$\mu(B \triangle f^{-1}(B)) = 0,$$

onde $\triangle$ denota diferença simétrica.

Definição 1.4.4. Dizemos que $(f,\ \mu)$ é um sistema ergódico se o tempo médio de visita a qualquer conjunto mensurável coincide, em μ-quase todo ponto, com a medida desse conjunto, isto é

$$\lim_n \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{X}_B(f^j(x)) = \mu(B),$$

para μ-quase todo ponto e para todo conjunto mensurável $B \subset M$.

Teorema 1.4.5. *Seja $f : M \longrightarrow M$ uma transformação mensurável e μ uma medida de probabilidade invariante por f. Se para todo subconjunto invariante A tem-se $\mu(A) = 0$ ou $\mu(A) = 1$, então, para toda função integrável $\varphi : M \longrightarrow \mathbb{R}$, temos*

$$\lim_n \frac{1}{n}\sum_{j=0}^{n-1} \varphi(f^j(x)) = \int \varphi d\mu,$$

para μ-quase todo ponto.

Demonstração. Veja [OV14, Cap04]. □

Os resultados seguintes serão muito úteis para o próximo capítulo, permitindo relacionar os expoentes de Lyapunov com médias espaciais.

Seja $GL(m,\mathbb{R})$ o espaço das matrizes $m \times m$ invertíveis e seja $\mathbb{R}P^{m-1}$ o espaço projetivo real de $\mathbb{R}^m$. Considere os espaços de probabilidade $(GL(m,\mathbb{R}), \mathcal{A}, \mu)$ e $(\mathbb{R}P^{m-1}, \mathcal{C}, \nu)$.
Seja $(X_n)_{n\in\mathbb{N}}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas com valores em $GL(m,\mathbb{R})$ e distribuição comum μ. Considere o espaço de probabilidade $(GL(m,\mathbb{R})^{\mathbb{N}}, \mathcal{B}, \mu^{\mathbb{N}})$, onde $\mathcal{B}$ é a σ-álgebra produto das σ-álgebras de Borel sobre $GL(m,\mathbb{R})$. Agora, considere o espaço de probabilidade $(GL^{\mathbb{N}} \times \mathbb{R}P^{m-1}, \mathcal{B} \times \mathcal{C}, \mu^{\mathbb{N}} \otimes \nu)$, onde a medida ν é μ-estacionária.
Observação: Quando no contexto seja claro denotaremos $GL(m,\mathbb{R}) = GL$
Definimos a aplicação

$$\begin{aligned} T : GL^{\mathbb{N}} \times \mathbb{R}P^{m-1} &\longrightarrow GL^{\mathbb{N}} \times \mathbb{R}P^{m-1} \\ ((X_n)_{n\in\mathbb{N}}, u) &\longmapsto ((X_{n+1})_{n\in\mathbb{N}}, X_1 u) \end{aligned} \tag{1.21}$$

que será muito útil posteriormente. Uma das coisas interessantes desta aplicação é que preserva medida.

Teorema 1.4.6. *Seja $(T, \mu^{\mathbb{N}} \otimes \nu)$, onde T é definida como (1.21). Então, T preserva medida.*

Demonstração. Seja $A_n \subset GL$ mensurável para todo $n \in \mathbb{N}$ e $B \subset \mathbb{R}P^{m-1}$ mensurável. Defina $A = \prod_{j=1}^{\infty} A_j \times B$. Provaremos que

$$\mu^{\mathbb{N}} \otimes \nu(T^{-1}(A)) = \mu^{\mathbb{N}} \otimes \nu(A).$$

Então, por definição, temos que:

$$\begin{aligned}
&\mu^{\mathbb{N}} \otimes \nu(T^{-1}(A)) \\
&\qquad = \mu^{\mathbb{N}} \otimes \nu(\{((X_n)_{n\in\mathbb{N}}, u) \in GL \times \mathbb{R}P^{m-1} \mid X_2 \in A_1, X_3 \in A_2, \cdots, X_1 u \in B\}).
\end{aligned}$$

Como $\mu * \nu(B) = \nu(B)$, tem-se:

$$\begin{aligned}
&\mu^{\mathbb{N}} \otimes \nu(T^{-1}(A)) \\
&\qquad = \mu^{\mathbb{N}} \otimes \nu(\{((X_n)_{n\in\mathbb{N}}, u) \in GL \times \mathbb{R}P^{m-1} \mid X_2 \in A_1, X_3 \in A_2, \cdots, u \in B\}).
\end{aligned}$$

Como $(X_n)_{n\in\mathbb{N}}$ é i.i.d., tem-se:

$$\begin{aligned}
&\mu^{\mathbb{N}} \otimes \nu(T^{-1}(A)) \\
&\qquad = \mu^{\mathbb{N}} \otimes \nu(\{((X_n)_{n\in\mathbb{N}}, u) \in GL \times \mathbb{R}P^{m-1} \mid X_1 \in A_1, X_2 \in A_2, \cdots, u \in B\}) \\
&\qquad = \mu^{\mathbb{N}} \otimes \nu(A).
\end{aligned}$$

□

Agora, definimos uma aplicação F tal que as médias temporais da dinâmica T coincidam com os expoentes de Lyapunov:

$$\begin{aligned}
F : GL^{\mathbb{N}} \times \mathbb{R}P^{m-1} &\longrightarrow \mathbb{R} \\
((X_n)_{n\in\mathbb{N}}, [u]) &\longmapsto \log\left(\frac{\|X_1 \hat{u}\|}{\|\hat{u}\|}\right)
\end{aligned} \tag{1.22}$$

É fácil ver que F está bem definido.
Além disso,

$$\begin{aligned}
\frac{1}{n+1}\sum_{k=0}^{n} F\left(T^k((X_n)_{n\in\mathbb{N}}, [u])\right) &= \frac{1}{n+1}\sum_{k=0}^{n} F\left((X_{n+k})_{n\in\mathbb{N}}, X_k X_{k-1}\cdots X_1[u]\right) \\
&= \frac{1}{n+1}\sum_{k=0}^{n} \log\frac{\|X_{k+1}X_k\ldots X_1[u]\|}{\|X_k\ldots X_1[u]\|} \\
&= \frac{1}{n+1}\left(\log\|X_n\cdots X_1[u]\| - \log\|[u]\|\right)
\end{aligned} \tag{1.23}$$

Agora aplicaremos o Teorema Ergódico de Birkhoff a F para obter uma relação muito útil posteriormente.

Teorema 1.4.7. *Seja $F \in L^1(GL^{\mathbb{N}} \times \mathbb{R}P^{m-1}, \mu^{\mathbb{N}} \otimes \nu)$ definida como (1.22), com ν sendo μ-estacionária. Então*

$$\widehat{\rho}\left((X_n)_{n\in\mathbb{N}}\,[\widehat{u}]\right) = \lim_n \frac{1}{n+1}\log\|X_n X_{n-1}\cdots X_1\hat{u}\| \tag{1.24}$$

existe em quase todos os pontos de $\mu^{\mathbb{N}} \otimes \nu$ e, além disso,

$$\int \widehat{\rho}\, d(\mu^{\mathbb{N}} \otimes \nu) = \int_{\mathbb{R}P^{m-1}} \int_{GL} \log \frac{\|X\widehat{u}\|}{\|\widehat{u}\|} d\mu(X) d\nu(\mu). \tag{1.25}$$

Demonstração. Note que pelo Teorema 1.4.6, $\mu^{\mathbb{N}} \otimes \nu$ é T-invariante e como, por hipótese, F é integrável, podemos aplicar o Teorema Ergódico de Birkhoff. Segue que

$$\lim_{n\to\infty} \frac{1}{n+1}\sum_{k=0}^{n} F\left(T^k((X_n)_{n\in\mathbb{N}}, [u])\right) \tag{1.26}$$

existe em q.t.p. de $\mu^{\mathbb{N}} \otimes \nu$. Além disso,

$$\int \lim_{n\to\infty} \frac{1}{n+1}\sum_{k=0}^{n} F\left(T^k((X_l)_{l\in\mathbb{N}}, [u])\right) d(\mu^{\mathbb{N}} \otimes \nu)((X_n)_{n\in\mathbb{N}}, [u]) = \int F d(\mu^{\mathbb{N}} \otimes \nu). \tag{1.27}$$

Pela definição de F tem-se que

$$\int F d(\mu^{\mathbb{N}} \otimes \nu) = \int_{\mathbb{R}P^{m-1}} \int_{GL} \log \frac{\|X\widehat{u}\|}{\|\widehat{u}\|} d\mu(X) d\nu(\mu). \tag{1.28}$$

Por (1.23) se tem que

$$\lim_{n\to\infty} \frac{1}{n+1}\sum_{k=0}^{n} F\left(T^k((X_l)_{l\in\mathbb{N}}, [u])\right) = \lim_n \frac{1}{n+1}\log\|X_n X_{n-1}\cdots X_1\hat{u}\|. \tag{1.29}$$

Juntando (1.26), (1.27), (1.28) e (1.29), obtemos (1.24) e (1.25). □

Outra aplicação com um papel importante no estudo dos expoentes de Lyapunov é a aplicação Shift:

Definição 1.4.8 (Aplicação Shift). Seja $(X_n)_{n\in\mathbb{N}}$ uma sequência de variáveis aleatórias i.i.d. com valores em $GL(m,\mathbb{R})$ e com distribuição comum μ. A aplicação Shift ($\overline{T}$) é definida como

$$\begin{aligned} \overline{T} : (GL(m,\mathbb{R})^{\mathbb{N}}, \mathcal{B}, \mu^{\mathbb{N}}) &\longrightarrow (GL(m,\mathbb{R})^{\mathbb{N}}, \mathcal{B}, \mu^{\mathbb{N}}) \\ (g_n)_{n\in\mathbb{N}} &\longmapsto (g_{n+1})_{n\in\mathbb{N}}. \end{aligned}$$

De forma similar ao Teorema (1.4.6), temos:

Teorema 1.4.9. *$\mu^{\mathbb{N}}$ é invariante pela aplicação Shift.*

Demonstração. A prova é similar ao Teorema 1.4.6. Seja $A_j \subset GL$ mensurável para todo $j \in \mathbb{N}$. Defina $A = \Pi_{j\in\mathbb{N}} A_j$. Provamos que

$$\mu^{\mathbb{N}}(\overline{T}^{-1}(A)) = \mu^{\mathbb{N}}(A).$$

Então, por definição, temos que:

$$\mu^{\mathbb{N}}(\overline{T}^{-1}(A)) = \mu^{\mathbb{N}}((g_n)_{n\in\mathbb{N}} \subset GL^{\mathbb{N}} : g_2 \in A_1, g_3 \in A_2, \cdots)$$

e, pelo fato que $(X_n)_{n\in\mathbb{N}}$ é i.i.d., temos

$$\begin{aligned}\mu^{\mathbb{N}}(\overline{T}^{-1}(A)) &= \mu^{\mathbb{N}}\{(g_n)_{n\in\mathbb{N}} \subset GL^{\mathbb{N}} : g_1 \in A_1, g_2 \in A_2, \cdots\}\\ &= \mu^{\mathbb{N}}(A).\end{aligned}$$

□

Temos ainda um resultado mais interessante sobre a aplicação Shift, que nos servirá para estender um resultado restrito a um compacto para um conjunto mais geral.

Teorema 1.4.10. *O par $(\overline{T}, \mu^{\mathbb{N}})$ é ergódico.*

Demonstração. Seja $A = \Pi_{j\in\mathbb{N}} A_j$ um conjunto invariante. com $A_j \subset GL$ mensurável para todo $j \in \mathbb{N}$ tal que

$$\mu^{\mathbb{N}}(A) = \prod_{j=1}^{\infty} \mu(A_j) > 0.$$

Temos que provar que $\mu^{\mathbb{N}}(A) = 1$.
Para todo $k > 1$ temos que:

$$\overline{T}^{-1}(A) \setminus A \supset \{(X_n)_{n\in\mathbb{N}} \subset GL^{\mathbb{N}} : X_2 \in A_1, X_3 \in A_2, \cdots, X_k \notin A_k\} := A^*.$$

Pela invariância de A e o fato de que $(X_n)_{n\in\mathbb{N}}$ é i.i.d. temos

$$\mu^{\mathbb{N}}(A^*) = \prod_{j=1, j\neq k}^{\infty} \mu(A_j).\mu(A_{k-1} \setminus A_k) = 0,$$

o que acontece se, e somente se, $\mu(A_{k-1} \setminus A_k) = 0$. Portanto

$$A_1 \subset A_2 \subset A_3 \subset \cdots \tag{1.30}$$

Por outro lado, tem-se

$$\overline{T}^{-1}(A) \setminus A \supset \{(g_n)_{n\in\mathbb{N}} \subset GL^{\mathbb{N}} : g_2 \in A_1, g_3 \in A_2, \cdots, g_1 \notin A_1\} = A^{**}.$$

Novamente, pela invariância de A e o fato que $(X_n)_{n\in\mathbb{N}}$ é i.i.d., temos

$$\mu^{\mathbb{N}}(A^*) = \prod_{j=2}^{\infty} \mu(A_j).\mu(A_1^c) = 0$$

o qual acontece se, e somente se, $\mu(A_1^c) = 0$. Portanto, $\mu(A_1) = 1$. Por este fato e (1.30) temos que $\mu^{\mathbb{N}}(A) = 1$. □

2 Matrizes Aleatórias

Dada uma medida de probabilidade μ sobre $GL(m,\mathbb{R})$, vamos estudar os expoentes de Lyapunov associados a μ. Começamos nossa análise considerando o caso em que μ tem suporte compacto, fato que será substituído por uma condição técnica mais fraca. Posteriormente, faremos uma filtração de $\mathbb{R}^m$, a qual nos levará a fazer uma partição de $\mathbb{R}^m$ tal que, a cada classe, associaremos um diferente expoente de Lyapunov. Tal resultado desempenha um papel importante no estudo da estabilidade do máximo expoente de Lyapunov.
Uma observação importante é o fato que os expoentes de Lyapunov só dependem da direção dos vetores de $\mathbb{R}^m$, pelo qual é intuitivo pensar em trabalhar no espaço projetivo.

2.1 Caminhos Aleatórios e Lei dos Grandes Números

O objetivo desta seção é construir uma dinâmica aleatória, um processo de Markov e uma função tudo esto escolhido convenientemente de forma tal que as médias temporais coincidam com os expoentes de Lyapunov, para aplicar o Lema Central e relacionar os expoente de Lyapunov com as médias espaciais sobre medidas ν μ-estacionárias.
Considere os espaços de probabilidade $(GL(m,\mathbb{R}),\mathcal{A},\mu)$ e $(\mathbb{R}P^{m-1},\mathcal{C},\nu)$. Consideremos, inicialmente, que o suporte $Q\subset GL(m,\mathbb{R})$ de μ seja compacto.
Definamos a ação:

$$\begin{aligned}\Theta : GL(m,\mathbb{R})\times\mathbb{R}P^{m-1} &\longrightarrow \mathbb{R}P^{m-1}\\ (g,x) &\longmapsto gx.\end{aligned}$$

É claro que Θ é contínuo.
Definamos $M = Q\times\mathbb{R}P^{m-1}$, munido da σ-álgebra produto.
Há uma maneira natural de definir um caminho aleatório em M:

$$\begin{aligned}\Lambda : M &\longrightarrow \mathcal{M}_1(M)\\ (g,x) &\longmapsto \lambda_{(g,x)},\end{aligned} \tag{2.1}$$

com $\lambda_{(g,x)}(A) = \int_{\mathbb{R}P^{m-1}}\int_{GL}\mathcal{X}_A(g,gx)d\mu(g)d\delta_{x_0}(x)$, para todo $A\subset M$ mensurável.
Observações: Como Q e $\mathbb{R}P^{m-1}$ são compactos, então $M = Q\times\mathbb{R}P^{m-1}$ é compacto na topologia produto e para $f\in\mathcal{C}(M)$ temos que $f\circ\Theta\in\mathcal{C}(M)$. Então $f\circ\Theta$ é uniformemente continua em M, segue que dado $\varepsilon>0$ existe $\delta>0$ tal que:

$$d_M((g',x')-(g_0,x_0))<\delta \Longrightarrow |f(g',g'x')-f(g_0,g_0x_0)|<\varepsilon. \tag{2.2}$$

Afirmação: a aplicação Λ é contínua.
Temos que provar que dados $f_1, f_2, \cdots, f_k \in \mathcal{C}(M)$, $(g_0, x_0) \in M$ e $\varepsilon > 0$, existe $\delta > 0$ tal que

$$d_M((g', x'), (g_0, x_0)) < \delta \Longrightarrow \left| \int f_i d\lambda_{(g',x')} - \int f_i d\lambda_{(g_0,x_0)} \right| < \epsilon, \qquad \forall i = 1, \cdots, n. \tag{2.3}$$

Pela definição de $\lambda(g', x')$, $\lambda(g_0, x_0)$ e (2.2), temos que existe $\delta_i > 0$ tal que

$$\begin{aligned}
\left| \int f_i d\lambda_{(g',x')} - \int f_i d\lambda_{(g_0,x_0)} \right| &= \left| \iint f_i(g, gx) d\mu(g) d\delta_{x'}(x) - \iint f_i(g, gx) d\mu(g) d\delta_{x_0}(x) \right| \\
&= \left| \int f_i(g, gx') d\mu(g) - \int f_i(g, gx_0) d\mu(g) \right| \\
&\leq \int |f_i(g, gx') - f_i(g, gx_0)| d\mu(g) \\
&\leq \int \epsilon d\mu(g) = \epsilon.
\end{aligned}$$

Tomando $\delta = \min\{\delta_1, \cdots, \delta_n\}$, obtemos (2.3). Portanto, Λ é continua. $\square$
Agora definimos o respectivo operador de Markov de Λ por

$$\mathcal{T} f(g, x_0) = \int f(g, gx) d(\mu \times \delta_{x_0})(g, x) = \int f(g, gx_0) d\mu(g). \tag{2.4}$$

Seja $(X_n)_{n \in \mathbb{N}}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas com valores em $GL(m, \mathbb{R})$ e com distribuição comum μ. Definimos uma sequência de valores variáveis em M:

$$\begin{aligned}
Z_0 = (e, u), \quad Z_1 = (X_1, X_1 u), \quad Z_2 = (X_2, X_2 X_1 u) \\
Z_n = (X_n, X_n X_{n-1} \cdots X_1 u),
\end{aligned} \tag{2.5}$$

com e como o elemento neutro do grupo $GL(m, \mathbb{R})$, i.e., $e = I$.
Note que, tal processo é de Markov. De fato, seja $f : M \longrightarrow \mathbb{R}$ uma função integrável, temos que:

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

Em particular, para $f = \mathcal{X}_A$, com $A \subset M$ mensurável, temos por (2.4) que a probabilidade de $Z_n \in A$ é $P(Z_n \in A) = \lambda_{Z_{n-1}}(A)$. Portanto, $(Z_n)_{n \in \mathbb{N}}$ é um processo de Markov.

Agora vamos relacionar as medidas $\nu \in \mathcal{M}_1(\mathbb{R}P^{m-1})$ que são μ-estacionárias, com as medidas $\lambda \in \mathcal{M}_1(GL \times \mathbb{R}P^{m-1})$ que são de Markov estacionárias, i.e., $\mathcal{T}^* \lambda = \lambda$, e com ν sendo como a medida de projeção de λ sobre o espaço $\mathbb{R}P^{m-1}$.

Suponhamos que $\lambda = \mu \otimes \nu$ seja uma medida de probabilidade em M satisfazendo $\mathcal{T}^*\lambda = \lambda$, então

$$\begin{aligned}
\int f d\lambda &= \int_M \mathcal{T} f d\lambda \\
&= \int_{\mathbb{R}P^{m-1}} \int_{GL} \mathcal{T} f(g,x) d\mu(g) d\nu(x) \\
&= \int_{\mathbb{R}P^{m-1}} \int_{GL} \int_{GL} f(g', g'x) d\mu(g') d\mu(g) d\nu(x) \\
&= \int_{\mathbb{R}P^{m-1}} \int_{GL} f(g, gx) d\mu(g) d\nu(x).
\end{aligned}$$

Seja ν a medida de projeção de λ sobre $\mathbb{R}P^{m-1}$. Se $f(g,x) = \varphi(x)$, temos que

$$\begin{aligned}
\int \varphi d\nu &= \int \varphi(x) d\nu(x) \\
&= \iint f(g,x) d\mu(g) \nu(x) \\
&= \iint f(g,gx) d\mu(g) d\nu(x) \\
&= \int \varphi(gx) d\mu(g) d\nu(x).
\end{aligned}$$

Portanto, $\mu * \nu = \nu$.

Por outro lado, seja Y uma variável aleatória com valores em $\mathbb{R}P^{m-1}$ e X uma variável aleatória com valores em $GL(m, \mathbb{R})$ independente de Y e sejam μ e ν as distribuições de X e Y, respectivamente. Então a distribuição do produto das variáveis aleatórias XY é $\mu * \nu$.

De fato, seja $A \subset \mathbb{R}P^{m-1}$ mensurável, temos que:

$$\begin{aligned}
P(XY \in A) &= P\left(\bigcup_{g \in G} [X = g][Y \in g^{-1}A] \right) \\
&= \int P(Y \in g^{-1}A | X = g) d\mu(g) \\
&= \int P(Y \in g^{-1}A) d\mu(g) \\
&= \int \nu(g^{-1}A) d\mu(g) = (\mu * \nu)(A). \qquad (2.6)
\end{aligned}$$

Na direção inversa, se temos que ν é μ-estacionária e definimos

$$\int f d\lambda := \iint f(g, gx) d\mu(g) d\nu(x) \qquad (2.7)$$

para toda $f \in \mathcal{C}(M)$.

Então

$$\begin{aligned}
\int f d(\mathcal{T}^*\lambda) &= \int \mathcal{T} f d\lambda \\
&= \iint \mathcal{T} f(g, gx) d\mu(g) d\nu(x) \\
&= \iiint f(g', g'gx) d\mu(g') d\mu(g) d\nu(g) \\
&= \iint f(g', g'x) d\mu(g') d\nu(x) \\
&= \int f d\lambda.
\end{aligned}$$

Como f é arbitrária, segue que $\mathcal{T}^*\lambda = \lambda$.
Portanto, obtemos uma correspondência um a um entre $\lambda \in \mathcal{M}_1(M)$, com $\mathcal{T}^*\lambda = \lambda$ e $\nu \in \mathcal{M}_1(\mathbb{R}P^{m-1})$, com $\mu * \nu = \nu$.
Agora definiremos uma função de forma tal que os valores assintóticos da média temporal coincidam com os expoentes de Lyapunov.
Sejam $[u] \in \mathbb{R}P^{m-1}$ e $\hat{u} \in \mathbb{R}^m \backslash \{0\}$ um vetor fixo, tal que a classe $[\hat{u}] = [u]$. Veja que pela relação $\frac{1}{n+1}\log\|X_nX_{n-1}\cdots X_1\frac{\hat{u}}{\|\hat{u}\|}\| = \frac{1}{n+1}\log\|X_nX_{n-1}\cdots X_1\hat{u}\| + \frac{1}{n+1}\log\|\hat{u}\|$ o estudo do comportamento assintótico de $\lim_{n\to\infty}\frac{1}{n+1}\log\|X_nX_{n-1}\cdots X_1\hat{u}\|$ não depende do representante da classe.
Definimos:

$$\begin{aligned} \rho: \quad M \quad &\longrightarrow \quad \mathbb{R} \\ (g,[u]) &\longrightarrow \log\left(\frac{\|\hat{u}\|}{\|g^{-1}\hat{u}\|}\right), \end{aligned} \tag{2.8}$$

com $\hat{u} \in \mathbb{R}^m\backslash\{0\}$ e é tal que $[\hat{u}] = [u]$, é claro que ρ não depende da classe $[u]$. Portanto, ρ está bem definida e é contínua. Note que:

$$\begin{aligned} \frac{1}{n+1}\sum_{k=0}^{n}\rho(Z_k) &= \frac{1}{n+1}[\rho(e,\hat{u}) + \rho(X_1, X_1\hat{u}) + \cdots + \rho(X_n, X_nX_{n-1}\cdots X_1\hat{u})] \\ &= \frac{1}{n+1}[(\log\|X_1\hat{u}\| - \log\|\hat{u}\|) + (\log\|X_2X_1\hat{u}\| - \log\|X_1\hat{u}\|) + \cdots \\ &+ (\log\|X_nX_{n-1}\cdots X_1\hat{u}\| - \log\|X_{n-1}X_{n-2}\cdots X_1\hat{u}\|)] \\ &= \frac{1}{n+1}[\log\|X_nX_{n-1}\cdots X_1\hat{u}\| - \log\|\hat{u}\|]. \end{aligned}$$

Então temos que

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\rho(Z_k) = \lim_{n\to\infty}\frac{1}{n+1}\log\|X_nX_{n-1}\cdots X_1\hat{u}\|. \tag{2.9}$$

Agora, podemos aplicar o Lema Central (Teorema 1.2.14) como caso especial a função ρ e obter o resultado mais importante desta seção.

Teorema 2.1.1. *Sejam μ uma medida de probabilidade com suporte compacto em $GL(m,\mathbb{R})$ e $(X_n)_{n\in\mathbb{N}}$ um processo estocástico i.i.d. com valores em $GL(m,\mathbb{R})$, com distribuição comum μ. Então para todo $v \in \mathbb{R}^m$ temos, com probabilidade um,*

$$\begin{aligned} &\lim_{n\to\infty}\sup\frac{1}{n}\log\|X_nX_{n-1}\cdots X_1 v\| \\ &\leq \sup\left\{\iint\log\frac{\|g\hat{u}\|}{\|\hat{u}\|}d\mu(g)d\nu(u) \mid \nu \in \mathcal{M}_1(\mathbb{R}P^{m-1}) \text{ com } \mu * \nu = \nu\right\}. \end{aligned} \tag{2.10}$$

Além disso, se para toda medida μ tal que $\mu * \nu = \nu$ a expressão

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

toma um só valor β, então, com probabilidade um,

$$\lim_{n\to\infty} \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1\| = \beta, \tag{2.11}$$

com β como o máximo expoente de Lyapunov.

Demonstração. Primeiramente, provaremos (2.10).
Como o suporte da medida Q é compacto, temos que $M = Q \times \mathbb{R}P^{m-1}$ é compacto, podemos aplicar o Lema Central 1.2.14 ao processo de Markov definido em (2.5) e à função definida em (2.8), então

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} \log(Z_k)$$
$$\leq \sup\left\{ \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g) d\nu(u) \mid \nu \in \mathcal{M}_1(\mathbb{R}P^{m-1}) \ com \ \mu * \nu = \nu \right\}.$$

Por (2.9)

$$\lim_{n\to\infty} \sup \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\|$$
$$\leq \sup\left\{ \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g) d\nu(u) \mid \nu \in \mathcal{M}_1(\mathbb{R}P^{m-1}) \ com \ \mu * \nu = \nu \right\}.$$

Com o qual acaba a prova de (2.10).
Agora, provaremos (2.11)
Afirmação 1:

$$\lim_{n\to\infty} \sup \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1\| \leq \beta(\mu). \tag{2.12}$$

Veja que, se $\{v_1, v_2, \cdots, v_m\}$ é uma base de $\mathbb{R}^m$, podemos definir a norma $\|.\|_1$ em $M^{m\times m}$ como $\|g\|_1 = \max_{j=1,\cdots,m} \|gv_j\|$ para todo $g \in M^{m\times m}$. Seja $\|.\|_2$ outra norma em $M^{m\times m}$. Sabemos que em todo espaço vetorial de dimensão finita as normas são equivalentes, de modo que existem $\alpha,\ \beta \in \mathbb{R}^+$ tal que

$$\alpha\|.\|_2 \leq \|.\|_1 \leq \beta\|.\|_2,$$

e então

$$\alpha^{\frac{1}{n}} \leq \left(\frac{\|.\|_1}{\|.\|_2}\right)^{\frac{1}{n}} \leq \beta^{\frac{1}{n}}$$

é válido para todo $n \in \mathbb{N}$. Sendo assim,

$$\lim_{n\to\infty} \alpha^{\frac{1}{n}} \leq \lim_{n\to\infty} \left(\frac{\|.\|_1}{\|.\|_2}\right)^{\frac{1}{n}} \leq \lim_{n\to\infty} \beta^{\frac{1}{n}}$$

$$1 \leq \lim_{n\to\infty} \left(\frac{\|.\|_1}{\|.\|_2}\right)^{\frac{1}{n}} \leq 1 \tag{2.13}$$

$$\lim_{n\to\infty}\left(\frac{\|.\|_1}{\|.\|_2}\right)^{\frac{1}{n}}=1.$$

Então por (2.10) e (2.13), concluímos a prova da Afirmação 1.
Afirmação 2:

$$\lim_{n\to\infty}\inf\frac{1}{n}\log\|X_nX_{n-1}\cdots X_1\|\geq\beta(\mu). \tag{2.14}$$

Sejam ν uma medida μ-estacionária, U_0 uma variável aleatória com valores em $\mathbb{R}P^{m-1}$ e distribuição ν e $(X_n)_{n\in\mathbb{N}}$ um processo estocástico com valores em $GL(m,\mathbb{R})$, i.i.d. e com distribuição μ. Definamos $Z_n^\nu=(X_n,X_nX_{n-1}\cdots X_1U_0)$. Note que, em analogia a (2.5) Z_n^ν é um processo de Markov, correspondente ao operador de Markov $\mathcal{T}$. Seja a variável aleatória $U_n=X_nX_{n-1}\cdots X_1U_0$ que pode ser definida indutivamente como $U_n=X_nU_{n-1}$, com X_n e U_{n-1} independentes.
Veja que, por (2.6), temos que a distribuição de U_1 é $\mu*\nu$ e, como ν é uma medida μ-estacionária, a distribuição de U_1 é ν. De forma análoga. Para $n>1$, temos que a distribuição de U_n é ν, uma vez que $U_n=X_nU_{n-1}$. Portanto $Z_n^\nu=(X_n,U_n)$ é um processo de Markov estacionário.
Definindo λ como (2.7), temos que o operador de Markov associado à dinâmica (2.1) é estacionário, isto é, $\lambda=\mathcal{T}^*\lambda$. De fato, seja A mensurável em M, então:

$$\begin{aligned}\mathcal{T}^*\lambda(A)&=\iint\mathcal{T}\mathcal{X}_A(g,x)d\mu(g)d\nu(x)\\&=\iiint\mathcal{X}_A(g',g'x)d\mu(g')d\mu(g)d\nu(x)\\&=\iint\mathcal{X}_A(g',g'x)d\mu(g')d\nu(x)\\&=\mathbb{E}(\mathcal{X}_A(Z_2^v))\\&=\mathbb{E}(\mathcal{X}_A(Z_1^v))\\&=\lambda(A).\end{aligned}$$

E pelo Teorema Ergódico de Birkhoff, temos que

$$\lim_{n\to\infty}\frac{1}{n}\log\|X_nX_{n-1}\cdots X_1U_0\|=\hat{\rho}, \tag{2.15}$$

existe para quase todo ponto com respeito a $\mu^{\mathbb{N}}\bigotimes\nu$, onde $\hat{\rho}$ é uma variável aleatória satisfazendo

$$\begin{aligned}\mathbb{E}[\hat{\rho}]&=\mathbb{E}(\rho(X_1,X_1U_0))\\&=\iint\rho(g,g\hat{u})d\mu(g)d\nu(\hat{u})\\&=\iint\log\left(\frac{\|g\hat{u}\|}{\|\hat{u}\|}\right)d\mu(g)d\nu(\hat{u}) \qquad (2.16)\\&=\beta(\mu).\end{aligned}$$

Este resultado pode ser obtido de forma análoga pelo Teorema 1.4.7.
Por outro lado, veja que sempre existe um conjunto de medida diferente de zero tal que $\hat{\rho} \geq \mathbb{E}(\hat{\rho})$ e, pela definição de norma, temos que para todo $n \in \mathbb{N}$

$$\log\|X_n X_{n-1} \cdots X_1\| \geq \log\|X_n X_{n-1} \cdots X_1 \hat{U}_0\| - \log\|\hat{U}_0\|.$$

Logo,

$$\liminf_{n\to\infty} \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1\| \geq \liminf_{n\to\infty} \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 \hat{U}_0\| = \hat{\rho}$$

Portanto, existe um conjunto de medida diferente de zero tal que

$$\liminf_{n} \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1\| = \lim_{n} \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1\| \geq E(\hat{\rho}). \tag{2.17}$$

Como o $\liminf$ é mensurável com respeito a σ-álgebra trivial de $(X_n)_{n\in\mathbb{N}}$, pela lei zero-um de Kolmogorov temos que a desigualdade (2.17) se mantém com probabilidade um e a expressão da esquerda de (2.17) não depende da medida ν escolhida. Então concluímos a afirmação dois.
Assim, segue que

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log\|X_n \cdots X_1\| &\leq \beta(\mu) \\ &\leq \liminf_{n\to\infty} \frac{1}{n} \log\|X_n \cdots X_1\| \leq \limsup_{n\to\infty} \frac{1}{n} \log\|X_n \cdots X_1\|\end{aligned}$$

$\square$

Observação 2.1.2. Por (2.13), podemos escolher convenientemente a norma com a qual trabalharemos nos cálculos dos expoentes de Lyapunov.

Agora, queremos estender os resultados do Teorema 2.1.1 para o caso em que o suporte da medida não seja compacto. Esta hipóteses será substituída pela hipótese técnica:

$$\int [\log^+\|g\| + \log^+\|g^{-1}\|] d\mu(g) < \infty, \tag{2.18}$$

com $\log^+\|g\| = \max\{\log\|g\|, 0\}$

Teorema 2.1.3. *Sejam μ uma medida de probabilidade em $GL(m, \mathbb{R})$ que satisfaça (2.18), $(X_n)_{n\in\mathbb{N}}$ um processo estocástico i.i.d. com valores em $GL(m, \mathbb{R})$ e com distribuição comum μ. Então para todo $v \in \mathbb{R}^m$ temos, com probabilidade um,*

$$\begin{aligned}&\limsup_{n\to\infty} \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\| \qquad (2.19)\\ &\leq \sup\left\{ \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g) d\nu(u) \mid \nu \in \mathcal{M}_1(\mathbb{R}P^{m-1}) \text{ com } \mu * \nu = \nu \right\}\end{aligned}$$

Além disso, se para toda medida μ tal que $\mu * \nu = \nu$ a expressão,

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

toma um só valor β, então com probabilidade um

$$\lim_{n\to\infty} \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1\| = \beta, \tag{2.20}$$

onde β é o máximo expoente de Lyapunov.

Demonstração. Sabemos que $GL(m,\mathbb{R})$ é localmente compacto. Denotamos como $\hat{G}$ sua compacidade local. Definamos $M = \hat{G} \times \mathbb{R}P^{m-1}$, o operador de Markov fica definido como

$$\mathcal{T} f(g,x) = \int f(g', g'x) d\mu(g'),$$

para toda $g \in \hat{G}$. Para cada $T \in \mathbb{R}^+$, definimos:

$$\begin{aligned} \rho_T : A_T \times \mathbb{R}P^{m-1} &\longrightarrow \mathbb{R} \\ (g,x) &\longrightarrow \log\left(\frac{\|g\hat{x}\|}{\|\hat{x}\|}\right), \end{aligned}$$

com $A_T = \{g \in GL(m,\mathbb{R}), \|g\|, \|g^{-1}\| \le T\}$.
Estendemos continuamente para todo M, com a mesma limitação.
Considere $\rho : G \times \mathbb{R}P^{m-1} \longrightarrow M$, como antes.
Afirmação:

$$\left|\frac{1}{n+1}\sum_{k=0}^{n}(\rho(Z_k) - \rho_T(Z_k))\right| \le \frac{1}{n+1}\sum_{k=0}^{n}(\log^+\|X_k\| + \log^+\|X_k^{-1}\| + \log T)\mathcal{X}_{B_T}(X_k), \tag{2.21}$$

com $B_T = \{g \in GL(m,\mathbb{R}); \max\{\|g\|, \|g^{-1}\|\} > T\}$.
É suficiente provar, que para todo $k \in \mathbb{N}$,

$$|\rho(Z_k) - \rho_T(Z_k)| \le (\log^+\|X_k\| + \log^+\|X_k^{-1}\| + \log T)\mathcal{X}_{B_T}(X_k). \tag{2.22}$$

Vamos separar a demostração dessa desigualdade em dois casos.
Caso I: Se $g \notin B_T$.
Pela definição de ρ_T temos $\rho = \rho_T$, e o lado esquerdo de (2.22) é zero. Por outro lado, como $\mathcal{X}_T(g) = 0$, temos que o lado direito de (2.22) é zero. Portanto, cumpre-se (2.22).
Caso II: Se $g \in B_T$.
Veja que, por definição, $\rho_T \le T$. Então temos que

$$|\rho(Z_k) - \rho_T(Z_k)| \le |\rho(Z_k)| + |\rho_T(Z_k)| \le |\rho(Z_k)| + T.$$

Por outro lado, como

$$|\rho(Z_k)| = |\rho(X_k, X_k U_{k-1})| \le |\log\|X_k\||.$$

Sendo assim, é suficiente provar que

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

Se $\|X_k\| \geq 1$, $|\log\|X_k\|| = \log\|X_k\| = \log^+\|X_k\| \leq \log^+\|X_k\| + \log^+\|X_k^{-1}\|$.
Se $\|X_k\| \leq 1$, $\|X_k^{-1}\| \geq 1$ e $\|X_k^{-1}\| \geq \dfrac{1}{\|X_k\|}$. Então $\log\left(\dfrac{1}{\|X_k\|}\right) = |\log\|X_k\|| \leq \log\|X_k^{-1}\| = \log^+\|X_k^{-1}\| \leq \log^+\|X_k\| + \log^+\|X_k^{-1}\|$.
Portanto, a desigualdade (2.22) é válida.
Agora, consideremos

$$\begin{aligned} h : GL^{\mathbb{N}} &\longrightarrow \mathbb{R} \\ (X_n)_{n\in\mathbb{N}} &\longrightarrow (\log^+\|X_1\| + \log^+\|X_1^{-1}\| + \log T)\mathcal{X}_{B_T}(X_1). \end{aligned}$$

Como h é integrável e como a aplicação Shift é relativamente ergódica a $\mu^{\mathbb{N}}$ (veja o Teorema 1.4.10), então as médias espaciais coincidem com as médias temporais, com probabilidade um. Assim,

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n+1}\sum_{k=1}^{n} h(\overline{T}^k(X_j)_{j\in\mathbb{N}}) &= \frac{1}{n+1}\sum_{k=1}^{n}(\log^+\|X_k\| + \log^+\|X_k^{-1}\| + \log T)\mathcal{X}_{B_T}(X_k)) \\ &= \int h d\mu^{\mathbb{N}} \\ &= \int_{B_T} (\log^+\|g\| + \log^+\|g^{-1}\|)d\mu(g) + \log T\mu(B_T) \\ &\leq 2\int_{B_T} (\log^+\|g\| + \log^+\|g^{-1}\|)d\mu(g). \end{aligned}$$

Por (2.18) temos que $B_T \longrightarrow 0$ quando $T \longrightarrow \infty$ e, como $(\log^+\|g\| + \log^+\|g^{-1}\|)$ é μ integrável, temos que.

$$\lim_{T\to\infty} \int_{B_T} (\log^+\|g\| + \log^+\|g^{-1}\|)d\mu(g) = 0.$$

Portanto, para T grande, o comportamento assintótico de $\|X_n X_{n-1}\cdots X_1 v\|$ pode ser estimado pela integral de ρ_T com respeito à medida λ satisfazendo $\mathcal{T}^*\lambda = \lambda$.
Comparando ρ com ρ_T temos

$$\begin{aligned} \int \rho d\lambda - \int \rho_T d\lambda &= \int_{B_T} (\rho - \rho_T)d\lambda \\ &\leq 2\int_{B_T} (\log^+\|g\| + \log^+\|g^{-1}\|)d\mu(g), \end{aligned}$$

que converge para zero quando $T \longrightarrow \infty$. □

2.2 Filtração de $\mathbb{R}^m$ para o Produto de Matrizes Aleatórias.

2.2.1 Propriedades do conjunto de medidas de probabilidade de convolução estacionárias

Nesta seção, fixemos uma medida de probabilidade μ em $GL(m, \mathbb{R})$ satisfazendo

$$\int (\log^+\|g\| + \log^+\|g^{-1}\|)d\mu(g) < \infty. \tag{2.23}$$

Para cada medida $\nu \in \mathcal{M}_1(\mathbb{R}P^{m-1})$ μ-estacionária, denotemos

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

Portanto $\beta(\mu)$ será

$$\beta(\mu) = sup\{\alpha(\mu) \mid \nu \in \mathcal{M}_1(P\mathbb{R}^{m-1}) \; com \; \mu * \nu = \nu\}.$$

O conjunto das medidas de probabilidade em $\mathbb{R}P^{m-1}$ tal que $\mu * \nu = \nu$, será denotado como $\mathcal{N}$.
Os próximos resultados vão na direção de descrever propriedades do Conjunto $\mathcal{N}$, tem um papel importante na próxima seção.
A primeira característica é que o conjunto $\mathcal{N}$ é não vazio.

Teorema 2.2.1. *Seja μ uma medida de probabilidade sobre $GL(m, \mathbb{R})$. O conjunto das medidas de probabilidade sobre $\mathbb{R}P^{m-1}$ tal que $\mu * \nu = \nu$ é não vazio.*

Demonstração. Considere a medida de probabilidade $\eta = \mu \otimes \nu$ sobre $GL \times \mathbb{R}P^{m-1}$ com $\nu \in \mathcal{M}_1(\mathbb{R}P^{m-1})$.
Definimos:

$$\lambda_n = \frac{1}{n+1} \sum_{k=0}^{n} \mathcal{T}^{*k}\eta.$$

Como $\mathcal{T}^{*k}\eta$ é uma medida de probabilidade para todo $k \in \mathbb{N}$ e o espaço das medidas de probabilidade é um conjunto convexo, segue que λ_n é uma medida de probabilidade sobre $GL \times \mathbb{R}P^{m-1}$ para todo $n \in \mathbb{N}$.
Seja μ_k a medida de projeção de λ_n sobre $GL(m, \mathbb{R})$, i.e. se

$$\begin{aligned} \pi_1 : GL \times \mathbb{R}P^{m-1} &\longrightarrow GL \\ (g, x) &\longmapsto g, \end{aligned}$$

$\mu_k(A) = \lambda_k(\pi_1^{-1}(A))$.
Afirmação: $\mu_k = \mu$ para todo $k \in \mathbb{N}$.

Seja $A \subset GL$ mensurável. Temos que:

$$
\begin{aligned}
\mu_n(A) =& \lambda_n(\pi_1^{-1}(A)) \\
=& \frac{1}{n+1}\sum_{k=0}^{n} \mathcal{T}^{*k}\eta(\pi_1^{-1}(A)) \\
=& \frac{1}{n+1}\sum_{k=0}^{n} \iint \mathcal{X}_{A\times\mathbb{R}P^{m-1}}(g,u)(\mathcal{T}^{k*}(\mu\otimes\nu)(g,u) \\
=& \frac{1}{n+1}\sum_{k=0}^{n} \iint \mathcal{T}^{k}\mathcal{X}_{A\times\mathbb{R}P^{m-1}}(g,u)\mu(g)\nu(u) \\
=& \frac{1}{n+1}\sum_{k=0}^{n} \iint \cdots \int \left[\int \mathcal{X}_{A\times\mathbb{R}P^{m-1}}(g_k, g_k g_{k-1}\cdots g_1 u)d\mu(g_k)\right]\mu(g_{k-1})\cdots\mu(g_1)\mu(g)\nu(u) \\
=& \frac{1}{n+1}\sum_{k=0}^{n} \iint \cdots \int \left[\int \mathcal{X}_{A}(g_k)d\mu(g_k)\right]\mu(g_{k-1})\cdots\mu(g_1)\mu(g)\nu(u) \\
=& \frac{1}{n+1}\sum_{k=0}^{n} \iint \cdots \int \mu(A)\mu(g_{k-1})\cdots\mu(g_1)\mu(g)\nu(u) \\
=& \mu(A).
\end{aligned}
$$

De forma análoga, definimos ν_k como a medida de projeção de λ_k sobre $\mathbb{R}P^{m-1}$. Assim, temos que $(\nu_n)_{n\in\mathbb{N}}$ é uma sequência de medidas de probabilidade sobre $\mathbb{R}P^{m-1}$. Então pelo fato que $\mathcal{M}_1(\mathbb{R}P^{m-1})$ é sequencialmente compacto (Teorema 1.1.5) temos que existe $(\nu_{n_k})_{n_k\in\mathbb{N}}$ e ν^* tal que $\nu_{n_k} \rightharpoonup \nu^*$.
Como as medidas marginais convergem fracamente e $GL \times \mathbb{R}P^{m-1}$ é separável, segue pelo Teorema 1.2.15 que $\lambda_{n_k} \rightharpoonup \mu * \nu^*$ e, por um argumento similar ao usado na prova do Teorema 1.2.4, temos que $\mathcal{T}^*\lambda = \lambda$, o qual implica que $\mu * \nu = \nu$. Assim, $\mathcal{N}$ é não vazio. □

Agora provaremos que o conjunto $\mathcal{N}$ é convexo e compacto.

Teorema 2.2.2. *Seja μ uma medida de probabilidade sobre $GL(m,\mathbb{R})$. O conjunto das medidas de probabilidade sobre $\mathbb{R}P^{m-1}$, tal que $\mu*\nu = \nu$, é convexo e compacto na topologia fraca estrela.*

Demonstração. Afirmação: $\mathcal{N}$ é convexo.
Sejam $\nu_1,\ \nu_2 \in \mathcal{N}$, $t \in [0,1]$. Provaremos que $t\nu_1 + (1-t)\nu_2 \in \mathcal{N}$.
É claro que $t\nu_1 + (1-t)\nu_2 \in \mathcal{M}_1(\mathbb{R}P^{m-1})$. Por outro lado, temos que

$$\mu * (t\nu_1 + (1-t)\nu_2) = t\mu * \nu_1 + (1-t)\mu * \nu_2 = t\nu_1 + (1-t)\nu_2.$$

Portanto, $\mathcal{N}$ é convexo.
Afirmação: $\mathcal{N}$ é compacto.
Como $\mathcal{N}$ é subconjunto de um conjunto compacto e como a topologia fraca estrela é de Hausdorff, é suficiente provar que $\mathcal{N}$ é fechado, i.e., se $\overline{\nu} \in \overline{\mathcal{N}}$, então $\overline{\nu} \in \mathcal{N}$.

É claro que, $\overline{\nu}$ é de probabilidade, então é suficiente provar que ν é μ-estacionária. Como $\overline{\nu} \in \mathcal{N}$, existe $(\nu_n)_{n\in N} \subset \mathcal{N}$ tal que $\nu_n \rightharpoonup \overline{\nu}$. Precisamos provar que

$$\int f(\hat{u})d\overline{\nu}(\hat{u}) = \iint f(g\hat{u})d\mu(g)d\overline{\nu}(\hat{u}).$$

Pela desigualdade triangular e pelo fato que $(\nu_n)_{n\in\mathbb{N}}$ são μ-estacionárias,

$$\begin{aligned}\left|\int fd\overline{\nu} - \int fd(\mu * \overline{\nu})\right| &\leq \left|\int fd\overline{\nu} - \int fd\nu_n\right| + \left|\int fd\nu_n - \int fd(\mu * \overline{\nu})\right| \\ &= \left|\int fd\overline{\nu} - \int fd\nu_n\right| + \left|\int fd(\mu * \nu_n) - \int fd(\mu * \overline{\nu})\right| \\ &= \left|\int fd\overline{\nu} - \int fd\nu_n\right| + \left|\int \left(\int f(g\hat{u})d\mu(g)\right) d(\overline{\nu} - \nu_n)\right|.\end{aligned}$$

Agora, suponha que a função

$$\begin{aligned} L : \mathbb{R}P^{m-1} &\longrightarrow \mathbb{R} \\ \hat{u} &\longmapsto \int f(g\hat{u})d\mu(g), \end{aligned}$$

seja contínua.
Então como $\nu_n \rightharpoonup \overline{\nu}$, temos que, dado $\varepsilon > 0$, existem n_{01}, n_{02} naturais tais que

$$\begin{aligned} \left|\int fd\overline{\nu} - \int fd\nu_n\right| &< \frac{\varepsilon}{2}, && \forall n \geq n_{01} \\ \left|\iint f(g\hat{u})d\mu(g)d\overline{\nu}(\hat{u}) - \iint f(g\hat{u})d\mu(g)d\nu_n(\hat{u})\right| &< \frac{\varepsilon}{2}, && \forall n \geq n_{02}. \end{aligned}$$

Tomando $n_0 = \max\{n_{01}, n_{02}\}$, temos que

$$\left|\int fd\overline{\nu} - \int fd(\mu * \overline{\nu})\right| < \varepsilon$$

e, como ε é arbitrário, segue que $\overline{\nu}$ é μ-estacionária.
Agora, provaremos que L é contínua. Sejam $(\hat{u_n})_{n\in N} \subset \mathbb{R}P^{m-1}$ e $\hat{u} \in \mathbb{R}P^{m-1}$, tal que $\lim_{n\to\infty} \hat{u_n} = \hat{u}$. Então é suficiente provar que $\lim_{n\to\infty} L(\hat{u_n}) = L(\hat{u})$.
Defina

$$\begin{aligned} f_n : GL(m, \mathbb{R}) &\longrightarrow \mathbb{R} \\ g &\longmapsto f(g\hat{u_n}). \end{aligned}$$

É claro que $\lim_{n\to\infty} f_n(g) = f(g\hat{u})$ para toda $g \in GL(m, \mathbb{R})$ e existe um número positivo $M < \infty$, tal que $|f_n| \leq M$. Como μ tem medida finita, podemos aplicar o Teorema da Convergência Dominada e então

$$\lim_n \int f(g\hat{u_n})d\mu(g) = \int f(g\hat{u})d\mu(g).$$

□

Dada a aplicação

$$\begin{aligned} T : GL^{\mathbb{N}} \times \mathbb{R}P^{m-1} &\longrightarrow GL^{\mathbb{N}} \times \mathbb{R}P^{m-1} \\ ((g_n)_{n\in N}, u) &\longmapsto ((g_{n+1})_{n\in N}, g_1 u), \end{aligned} \tag{2.24}$$

quando $\nu \in \mathcal{N}$, temos que $\mu^{\mathbb{N}} \otimes \nu$ é invariante por T (Teorema 1.4.6).
Se temos que o sistema $(T, \mu^{\mathbb{N}} \otimes \nu)$ é ergódico, então temos, com probabilidade um, que

$$\lim_{n\to\infty} \frac{1}{n+1} \log\|X_n X_{n-1} \cdots X_1 v\| = \iint \log \frac{\|gu\|}{\|u\|} d\mu(g) d\nu(g) = \alpha(\nu) \tag{2.25}$$

Portanto, podemos formular o seguinte lema.

Lema 2.2.3. *Se $\nu \in \mathcal{N}$ é tal que o sistema $(T, \mu^{\mathbb{N}} \otimes \nu)$ é ergódico, então para ν-quase toda direção u, se $v \in \mathbb{R}^m$ for um vetor na direção de u, então:*

$$\frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\| \longrightarrow \alpha(\nu).$$

Agora, provaremos uma característica muito importante quando o sistema $(T, \mu^{\mathbb{N}} \otimes \nu)$ não for ergódico, que, de fato, é uma equivalência. Este resultado será relevante para caracterizar os pontos extremais de $\mathcal{N}$.

Lema 2.2.4. *Sejam μ e ν medidas de probabilidade em $GL(m, \mathbb{R})$ e $\mathbb{R}P^{m-1}$, respectivamente, tais que ν seja μ-estacionária. Se $(T, \mu^{\mathbb{N}} \otimes \nu)$ não for ergódico, existe um conjunto $B \subset \mathbb{R}P^{m-1}$ mensurável tal que $gB \subset B$ para μ-q.t.p., com $0 < \nu(B) < 1$.*

Demonstração. Como $(T, \mu^{\mathbb{N}} \otimes \nu)$ não é ergódico, existe $A \subset GL^{\mathbb{N}} \times \mathbb{R}P^{m-1}$ mensurável com $0 < \mu^{\mathbb{N}} \otimes \nu(A) < 1$ tal que

$$\mu^N \otimes \nu(A \triangle T^{-1}A) = 0. \tag{2.26}$$

Considere $A = \prod_{j=1}^{\infty} A_j \times B$, com cada $A_i \subset GL$ e $B \subset \mathbb{R}P^{m-1}$ mensuráveis. Como os (X_n) são i.i.d., temos que

$$\mu^{\mathbb{N}} \otimes \nu(A) = \left(\prod_{j=1}^{\infty} \mu(A_j) \right) \nu(B). \tag{2.27}$$

Então, é suficiente provar que:
(a) $gB \subset B$ para μ-q.t.p.;
(b) $\mu(A_j) = 1$ para todo $j = 1, \cdots, k$.
Prova de (a).
É claro que

$$A \triangle T^{-1}A \supset A \backslash T^{-1}(A) \supset \{((g_n)_{n\in\mathbb{N}}, u) | g_1 \in A_1, \cdots, g_k \in A_k, \cdots, \hat{u} \in B \text{ e } g_1 u \notin B\} = A^*.$$

Por (2.26) e pelo fato que $(X_n)_{n\in\mathbb{R}}$ é uma sequência i.i.d., temos que

$$\mu^{\mathbb{N}}\otimes\nu(A^*)=\left(\prod_{j=2}^{\infty}\mu(A_j)\right)\int_{GL}\int_{B}\mathcal{X}_{B^c}(g\hat{u})d\nu(\hat{u})d\mu(g)=0,$$

o qual acontece se, e somente se,

$$\int_{GL}\int_{B}\mathcal{X}_{B^c}(g\hat{u})d\nu(\hat{u})d\mu(g)=0,$$

o que é equivalente a $gB\subset B$ para μ-q.t.p.
De fato, se $gB\subset B$ para μ-q.t.p., então $\int_{GL}\int_{B}\mathcal{X}_{B^c}(g\hat{u})d\nu(\hat{u})d\mu(g)=0$.
Agora para o sentido contrário: por absurdo, suponha que existe $D\subset GL$ mensurável, tal que, para todo $g\in D$, $gB\subset B^c$, com $\mu(D)>0$. Então

$$\begin{aligned}\int_{GL}\int_{B}\mathcal{X}_{B^c}(g\hat{u})d\nu(\hat{u})d\mu(g)&\geq\int_{D}\int_{B}\mathcal{X}_{B^c}(g\hat{u})d\nu(\hat{u})d\mu(g)\\&=\int_{D}\int_{B}1d\nu(\hat{u})d\mu(g)\\&=\mu(D)\nu(B)>0,\end{aligned}$$

o que é absurdo. Assim, concluímos (a).
Prova de (b).
É suficiente provar que:
(b_1) $A_1\subset A_2\subset\cdots\subset A_k\subset A_{k+1}\subset\cdots$ para μ-q.t.p.;
(b_2) $\mu(A_1)=1$
Prova de (b_1). Veja que, para qualquer número natural $j\geq 2$, temos que

$$T^{-1}A\setminus A\supset\{((g_n)_{n\in N},u)|g_2\in A_1,g_3\in A_2,\cdots,g_{k+1}\in A_k,\cdots,g_1u\in B,g_j\notin A_j\}=A^{j*}.$$

Então, por (2.26) e o fato que $(X_n)_{n\in\mathbb{R}}$ é uma sequência i.i.d., temos

$$\mu^{\mathbb{N}}\otimes\nu(A^{j*})=\left(\prod_{k=1,k\neq j-1}^{\infty}\mu(A_k)\right)\nu(B)\mu(A_{j-1}\setminus A_j)=0,$$

o que acontece se, e somente se, $A_{j-1}\subset A_j$ para μ-q.t.p.
Para a prova de (b_2). Por absurdo, suponha que $\mu(A_1^c)>0$. Veja que

$$T^{-1}A\setminus A\supset\{((g_n)_{n\in\mathbb{N}},u)|g_2\in A_1,g_3\in A_2,\cdots,g_{k+1}\in A_k,\cdots,g_1u\in B,g_1\notin A_1\}=A^{1*}.$$

Então, novamente por (2.26) e o fato que $(X_n)_{n\in\mathbb{R}}$ é uma sequência i.i.d., temos

$$\mu^{\mathbb{N}}\otimes\nu(A^{1*})=\left(\prod_{j=1}^{\infty}\mu(A_j)\right)\int_{A_1^c}\int_{\mathbb{R}P^{m-1}}\mathcal{X}_{B^c}(g\hat{u})d\nu(\hat{u})d\mu(g)=0,$$

o que acontece se, e somente se, $\int_{A_1^c}\int_{\mathbb{R}P^{m-1}}\mathcal{X}_{B^c}(g\hat{u})d\nu(\hat{u})d\mu(g)=0$. Mas por (b_1), temos que $gB\subset B$ para μ-q.t.p., o que implica que $B^c\subset(gB)^c=gB^c$ para μ-q.t.p e então

$$\begin{aligned}\int_{A_1^c}\int_{\mathbb{R}P^{m-1}}\mathcal{X}_{B^c}(g\hat{u})d\nu(\hat{u})d\mu(g)&\geq\int_{A_1^c}\int_{B}\mathcal{X}_{B^c}(g\hat{u})d\nu(\hat{u})d\mu(g)\\&=\int_{A_1^c}\int_{\mathbb{R}P^{m-1}}1d\nu(\hat{u})d\mu(g)\\&=\mu(A_1^c)\nu(B)>0\end{aligned}$$

o que é absurdo. Portanto $\mu(A_1^c) = 0$, o que implica que $\mu(A_1) = 1$. $\square$

Agora podemos caracterizar os pontos extremais de $\mathcal{N}$ como as medidas ν que fazem com que o sistema $(T, \mu^{\mathbb{N}} \otimes \nu)$ seja ergódico.

Teorema 2.2.5. *Seja μ uma medida de probabilidade fixa em $GL(m, \mathbb{R})$, satisfazendo (2.23). A medida $\nu \in \mathcal{N}$ faz $(T, \mu^{\mathbb{N}} \otimes \nu)$ ergódico se, e somente se, ν é extremal em $\mathcal{N}$.*

Demonstração. Provar a parte "se", é equivalente a mostrar que, se ν não for extremal em $\mathcal{N}$, então $(T, \mu^{\mathbb{N}} \otimes \nu)$ não será ergódico.
Por absurdo, suponha que $(T, \mu^{\mathbb{N}} \otimes \nu)$ seja ergódico e $\nu = t\nu_1 + (1-t)\nu_2$, onde ν_1 e ν_2 estão em $\mathcal{N}$ e são extremais, com $\nu_1 \neq \nu_2$. É claro que ν_1 é absolutamente contínua com relação a ν. Portanto $\mu^{\mathbb{N}} \otimes \nu_1$ e absolutamente contínua com relação a $\mu^{\mathbb{N}} \otimes \nu$. Seja $A \subset GL^{\mathbb{N}} \times \mathbb{R}P^{m-1}$ mensurável. Então a média temporal

$$\overline{\mathcal{X}_A}(w) = \lim_n \frac{1}{n+1} \sum_{j=0}^{n} \mathcal{X}_A(w)$$

é constante para $\mu^{\mathbb{N}} \otimes \nu$-q.t.p. Como $\nu_1 << \nu$, temos que $\overline{\mathcal{X}_A}$ é constante para $\mu^{\mathbb{N}} \otimes \nu_1$-q.t.p. e, pelo teorema de Birkhoff, temos que

$$\mu^{\mathbb{N}} \otimes \nu_1(A) = \int \overline{\mathcal{X}_A} d(\mu^{\mathbb{N}} \otimes \nu_1) = \mu^{\mathbb{N}} \otimes \nu(A).$$

Como A é arbitrário, segue que $\nu_1 = \nu$ e, de forma análoga, temos que $\nu_2 = \nu$, então $\nu_1 = \nu_2$. Absurdo.
Vamos provar a recíproca. Por absurdo, suponha que $(T, \mu^{\mathbb{N}} \otimes \nu)$ não seja ergódico. Então, pelo Lema 2.2.4, existe $B \subset \mathbb{R}P^{m-1}$ mensurável com $0 < \nu(B) < 1$ tal que $gB \subset B$ para μ-q.t.p.. Defina as medidas de probabilidade em $\mathbb{R}P^{m-1}$ como

$$\nu_1(A) = \frac{\nu(A \cap B)}{\nu(B)}, \text{ e} \tag{2.28}$$

$$\nu_2(A) = \frac{\nu(A \cap B^c)}{\nu(B^c)}, \tag{2.29}$$

para todo $A \subset \mathbb{R}P^{m-1}$ mensurável.
É claro que ν_1 e ν_2 então bem definidas, pois $1 > \nu(B) > 0$ e, por definição, são medidas de probabilidade.
Afirmação: ν_1 e ν_2 são μ-estacionárias.

Seja $A \subset \mathbb{R}P^{m-1}$ mensurável, então pelo fato que $gB \subset B$ e ν é μ-estacionárias,

$$\begin{aligned}
\mu * \nu_1(A) := & \int_{GL} \nu_1(g^{-1}A)d\mu(g) \\
= & \frac{1}{\nu(B)} \int_{GL} \int_B \mathcal{X}_A(g\hat{u})d\nu(\hat{u})d\mu(g) \\
= & \frac{1}{\nu(B)} \int_{GL} \int_B \mathcal{X}_{A\cap B}(g\hat{u})d\nu(\hat{u})d\mu(g) \\
= & \frac{1}{\nu(B)} \int_{GL} \int_{\mathbb{R}P^{m-1}} \mathcal{X}_{A\cap B}(g\hat{u})d\nu(\hat{u})d\mu(g) \\
= & \frac{1}{\nu(B)} \mu * \nu(A \cap B) \\
= & \frac{\nu(A \cap B)}{\nu(B)} \\
= & \frac{\nu_1(A)}{\nu(B)} \nu(B) \\
= & \nu_1(A).
\end{aligned}$$

Portanto, ν_1 é μ-estacionária. De forma análoga, provamos que ν_2 é μ-estacionária. Tome $t = \nu(B)$. Então $\nu = t\nu_1 + (1-t)\nu_2$, isto é, ν é soma convexa de medidas em $\mathcal{N}$. Mas, por hipótese, ν não é soma convexa. Absurdo. Então, concluímos que $(T, \mu^N \otimes \nu)$ é ergódica. □

2.2.2 Filtração de $\mathbb{R}^m$.

Queremos fazer uma filtração de $\mathbb{R}^m$, o que nos levará a fazer uma partição de $\mathbb{R}^m$, na qual, a cada classe, associaremos um expoente de Lyapunov diferente.
Antes, faremos uma extensão do Teorema 2.1.3 para o caso em que existe uma medida de convolução estacionária ν tal que $\alpha(\nu) < \beta(\mu)$. Porém, serão necessários alguns resultados prévios.
Com o seguinte resultado pretendemos restringir o conjunto das medidas $\nu \in \mathcal{N}$ que nos ajudem no cálculo do máximo expoente de Lyapunov.

Lema 2.2.6. *Seja μ uma medida de probabilidade sobre $GL(m, \mathbb{R})$, satisfazendo (2.23). Então*

$$\beta(\mu) = \sup\{\alpha(\nu) | \nu \in \mathcal{N}, \ \nu \textit{ extremal}\}.$$

Demonstração. Defina:

$$\begin{aligned}
J : \mathcal{N} \longrightarrow & \mathbb{R} \\
\nu \longrightarrow & \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g)d\nu(u).
\end{aligned}$$

Lembre-se que: $|\log\|g\|| \leq \log^+\|g\| + \log^+\|g^{-1}\|$ que, junto com (2.23), fazem que J seja limitado. Por outro lado, como $\mathcal{N}$ é sequencialmente compacto, existe $\nu \in \mathcal{N}$ tal que $\|J\| = |J(\nu)|$.
Pelo Teorema de Krein-Milman 1.1.10 ν pode ser escrita como soma convexa de medidas extremais em $\mathcal{N}$, sem perdida de generalidade suponha que se pode escrever como suma convexa de duas medidas, então

$$\nu = t\nu_1 + (1-t)\nu_2,$$

com $t \in [0,1]$ e ν_1 e ν_2 extremais.
Mas, pela linealidade de J, temos que $J(\nu) = tJ(\nu_1) + (1-t)J(\nu_2)$. Então $|J(\nu)| \leq |J(\nu_1)|$ ou $|J(\nu)| \leq |J(\nu_2)|$. Portanto, a norma é atingida em uma medida extremal. $\square$

Suponhamos que, para toda medida extremal $\nu \in \mathcal{N}$, temos que $\alpha(\nu) = \beta$ para um β fixo. Então $\beta = \beta(\mu)$ e, pelo Teorema ergódico de Birkhoff (2.1.1), para quase todo $v \in \mathbb{R}^m$, temos que:

$$\lim_{n\to\infty} \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\| = \beta(\mu). \tag{2.30}$$

Se (2.30) não for válido para quase todo $v \in \mathbb{R}^m$, então existe uma medida extremal ν com $\alpha(\nu) < \beta(\mu)$. Como ν é extremal, o processo $(T, \mu^{\mathbb{N}} \otimes \nu)$ é ergódico e temos que o Lema 2.2.3 é válido. Portanto, para quase todo $v \in \mathbb{R}^m$:

$$\lim_{n\to\infty} \sup \frac{1}{n} \|X_n X_{n-1} \cdots X_1 v\| = \alpha(\nu).$$

Seja L subconjunto de $\mathbb{R}^m$, denotamos como $\overline{L}$ o correspondente conjunto de pontos em $\mathbb{R}P^{m-1}$, isto é, o conjunto de todas as direções representadas em L.

Lema 2.2.7. *Seja L o conjunto de vetores v tais que*

$$\lim_{n\to\infty} \sup \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\| \leq \alpha(\nu). \tag{2.31}$$

Então, L é um subespaço de $\mathbb{R}^m$.

Demonstração. $i)$ Se $u \in L$ e $t \in \mathbb{R}$ então $tu \in L$.
Se $t = 0$, trivial.
Para $t \neq 0$.
É suficiente notar que, para todo $n \in \mathbb{N}$, temos

$$\frac{1}{n} \log\|X_n \cdots X_1(tu)\| = \frac{1}{n} \log\|X_n \cdots X_1 u\| + \frac{1}{n} \log|t|.$$

Tomando o limite quando $n \longrightarrow \infty$ temos

$$\lim_{n\to\infty} \frac{1}{n} \log\|X_n \cdots X_1(tu)\| = \lim_{n\to\infty} \frac{1}{n} \log\|X_n \cdots X_1 u\| \leq \alpha(\nu).$$

ii) Se $u_1, u_2 \in L$, então $u_1 + u_2 \in L$.
Veja que, para todo $n \in \mathbb{N}$, temos que

$$\frac{1}{n}\log\|X_n \cdots X_1(u_1+u_2)\| \leq \frac{1}{n}\log\|X_n \cdots X_1 u_1\| + \frac{1}{n}\log\|X_n \cdots X_1 u_2\| \tag{2.32}$$

e

$$\limsup_n \frac{1}{n}\log(\|X_n \cdots X_1 u_1\| + \|X_n \cdots X_1 u_2\|) = \max\{\limsup_n \frac{1}{n}\log\|X_n \cdots X_1 u_1\|, \limsup_n \frac{1}{n}\log\|X_n \cdots X_1 u_2\|\}. \tag{2.33}$$

Novamente, tomando o limite quando $n \longrightarrow \infty$ e juntando (2.32) e (2.33) temos

$$\lim_{n\to\infty} \frac{1}{n}\log\|X_n \cdots X_1(u_1+u_2)\| \leq \alpha(\nu).$$

□

Note que pelo Lema 2.2.3 temos que a media espacial é igual à medias espacias com probabilidade um, então temos que $\nu(\overline{L}) = 1$ e, como $\alpha(\nu) < \beta(\mu)$, pelo Teorema 2.1.1, L é um subespaço próprio de $\mathbb{R}^m$. Se L_ν denota o subespaço minimal tal que $\nu(\overline{L_\nu}) = 1$, então $\overline{L_\nu} \subset \overline{L}$. Portanto, $\overline{L_\nu}$ é subespaço próprio de $\mathbb{R}^m$ e (2.31) é válido para todo $v \in \overline{L_\nu}$.

Definição 2.2.8. Seja $L \subset \mathbb{R}^m$ e μ uma medida de probabilidade sobre $GL(m,\mathbb{R})$, dizemos que L é *μ-quase invariante*[1] se $gL = L$ para μ-q.t.p.

Lema 2.2.9. *Sejam $\nu \in \mathcal{N}$ extremal e $\overline{L_\nu}$ o subespaço minimal de $\mathbb{R}^m$ tal que $\nu(\overline{L_\nu}) = 1$. Então L_ν é μ-quase invariante.*

Demonstração. Como $\nu = \mu * \nu$,

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

Portanto $\nu(g^{-1}\overline{L_\nu}) = 1$ para μ- q.t.p.. Como $\overline{L_\nu}$ é minimal, temos que $g^{-1}\overline{L_\nu} \supset \overline{L_\nu}$. Logo $gL_\nu = L_\nu$ em um conjunto de medida total. □

Podemos formular de forma imediata:

Teorema 2.2.10. *Seja μ uma medida de probabilidade sobre $GL(m,\mathbb{R})$ que satisfaça a condição* (2.23).
Então, para todo $v \in \mathbb{R}^m$ com probabilidade um, ou

$$\lim_n \frac{1}{n}\log\|X_n \cdots X_1 v\| = \beta(\mu), \tag{2.34}$$

[1] Aqui preferimos destoar da terminologia no artigo original e usar o termo quase invariante para ficar mais claro

ou existe algum subespaço próprio $L \subset \mathbb{R}^m$, μ-quase invariante tal que, para todo $v \in L$,

$$\lim_n \frac{1}{n} \log \|X_n \cdots X_1 v\| \leq \alpha, \tag{2.35}$$

onde $\alpha < \beta(\mu)$, com probabilidade um.

Se L é um subespaço μ-quase invariante, então para μ-q.t.p. g, temos uma transformação $g|_L : L \to L$. Seja $B = \{v_1, v_2 \cdots, v_k\}$ uma base de L. Estendemos a uma base de $\mathbb{R}^m$ $\overline{B} = \{v_1, v_2, \cdots, v_k, v_{k+1}, \cdots, v_m\}$. Lembrando a relação de mudança de base temos que

$$g = \overline{B}\overline{g}\overline{B}^{-1},$$

onde $\overline{g} = ([gv_1]_{\overline{B}}\ [gv_2]_{\overline{B}} \cdots [gv_m]_{\overline{B}})$, pela quase invariância de L se tem

$$gv_i = \eta_1^i v_1 + \cdots + \eta_k^i v_k,$$

para todo $i = 1, 2, \cdots, k$. Então $\overline{g}$ tem a forma

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

Observe que

$$\begin{aligned} g_n g_{n-1} \cdots g_1 &= (\overline{B}\overline{g}_n\overline{B}^{-1})(\overline{B}\overline{g}_{n-1}\overline{B}^{-1}) \cdots (\overline{B}\overline{g}_1\overline{B}^{-1}) \\ &= \overline{B}\overline{g}_n\overline{g}_{n-1} \cdots \overline{g}_1\overline{B}^{-1} \end{aligned}$$

O qual implica que os expoentes de Lyapunov ficam invariantes com respeito a mudança de base, sem perdida de generalidade escreveremos $g = \overline{g}$.
Podemos associar duas taxas de crescimentos, uma com respeito à direção dos vetores contidos em L e outra no complemento de L.

$$\begin{aligned} \beta(\mu_L) &= \lim_{n\to\infty} \frac{1}{n} \log \|g_{11}^n \cdots g_{11}^1\| \\ \beta(\mu_{\mathbb{R}^m/L}) &= \lim_{n\to\infty} \frac{1}{n} \log \|g_{22}^n \cdots g_{22}^1\| \end{aligned}$$

Observação 2.2.11. É claro que pelo Lema 2.2.7, $\beta(\mu_{L'+L''}) = \max\{\beta(\mu_{L'}), \beta(\mu_{L''})\}$, se L' e L'' são subespaços μ-quase invariantes. Segue que, se existe algum subespaço com $\beta(\mu_L) < \beta(\mu)$, então existe um único subespaço maximal com uma taxa de crescimento estritamente menor a $\beta(\mu)$. Denotamos este subespaço por L_1.

Vejamos um exemplo de como calcular L_1

Exemplo 2.2.12. Seja μ uma medida de probabilidade sobre $GL(m,\mathbb{R})$, tal que o suporte de μ esteja contido, por simplicidade nas matrizes diagonais

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

Sejam $\gamma_i = \int \log|g_{ii}|d\mu(g)$, para $i = 1,2,3$ com $\gamma_1 > \gamma_2, \gamma_3$. Então pela lei dos grandes números temos que

$$\beta(\mu) = \max\{\gamma_1, \gamma_2, \gamma_3\} = \gamma_1.$$

Portanto o máximo subespaço μ-quase invariante que tem associado taxa de crescimento menor a γ_1 é

$$L_1 = \left\{ \begin{pmatrix} 0 \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid y, z \in \mathbb{R} \right\}$$

O seguinte Lema nos dá a relação entre $\beta(\mu)$, $\beta(\mu_{L_1})$ e $\beta(\mu_{\mathbb{R}^m/L})$.

Lema 2.2.13. *Seja L_1 um subespaço μ-quase invariante de $\mathbb{R}^m$. Então*

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

Demonstração. Escolhemos uma base de $\mathbb{R}^m$ cujos vetores iniciais formam uma base do espaço L. As matrizes g no suporte de μ tem a forma

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

com g_{ij} como submatrizes, onde g_{11} corresponde à restrição de g a L e g_{22} à ação de g em $\mathbb{R}^m/L$. Tomando o produto $X_nX_{n-1}\cdots X_1$ de matrizes com esta forma e identificando $\beta(\mu)$ como a taxa de crescimento do produto aleatório, vemos imediatamente que

$$\beta(\mu_L) \leq \beta(\mu), \qquad \beta(\mu_{\mathbb{R}^m}/L) \leq \beta(\mu). \tag{2.36}$$

Agora, suponha que essas duas desigualdades sejam estritas. Considere o produto

$$(X_{2n}X_{2n-1}\cdots X_n+1)(X_nX_{n-1}\cdots X_1) = \begin{pmatrix} A'_n & B'_n \\ 0 & C'_n \end{pmatrix} \begin{pmatrix} A''_n & B''_n \\ 0 & C''_n \end{pmatrix}$$

Dado $\varepsilon > 0$, existe $n_{01}, n_{02},\ n_{03} \in \mathbb{N}$ tal que temos, com probabilidade um,

$$-\varepsilon + \beta(\mu_L) < \frac{1}{n}\log\|X_nX_{n-1}\cdots X_1\| < \varepsilon + \beta(\mu_L), \qquad \forall n \geq n_{01}$$

$$-\varepsilon + \beta(\mu_{\mathbb{R}^m/L}) < \frac{1}{n}\log\|X_nX_{n-1}\cdots X_1\| < \varepsilon + \beta(\mu_{\mathbb{R}^m/L}), \qquad \forall n \geq n_{02}$$

$$-\varepsilon + \beta(\mu) < \frac{1}{n}log\|X_nX_{n-1}\cdots X_1\| < \varepsilon + \beta(\mu), \qquad \forall n \geq n_{03}$$

Tomando $n_0 = \max\{n_{01}, n_{02}, n_{03}\}$, então para $n \geq n_0$ temos que

$$\begin{aligned}\|A_n'\|, \|A_n''\| &\leq e^{(1+\varepsilon)n\beta(\mu_L)},\\ \|B_n'\|, \|B_n''\| &\leq e^{(1+\varepsilon)n\beta(\mu)},\\ \|C_n'\|, \|C_n''\| &\leq e^{(1+\varepsilon)n\beta(\mu_{\mathbb{R}^m/L})}.\end{aligned}$$

Tomando a norma infinito temos que

$$\begin{aligned}\|X_{2n}X_{2n-1}\cdots X_1\|_\infty &\leq \|A_n'A_n''\|_\infty + \|A_n'B_n'' + B_n'C_n''\|_\infty + \|C_n'C_n''\|_\infty\\ &\leq \|A_n'\|_\infty\|A_n''\|_\infty + \|A_n'B_n''\|_\infty + \|B_n'C_n''\|_\infty + \|C_n'C_n''\|_\infty\\ &\leq \|A_n'\|_\infty\|A_n''\|_\infty + \|A_n'\|_\infty\|B_n''\|_\infty + \|B_n'\|_\infty\|C_n''\|_\infty + \|C_n'\|_\infty\|C_n''\|_\infty\end{aligned}$$

Portanto com probabilidade um, para toda $n \geq n_0$, temos que

$$\begin{aligned}&\|X_{2n}X_{2n-1}\cdots X_1\| \leq\\ &\qquad e^{(1+\varepsilon)2n\beta(\mu_L)} + e^{(1+\varepsilon)2n\beta(\mu_{\mathbb{R}^m/L})} + e^{(1+\varepsilon)n(\beta(\mu)+\beta(\mu_L))} + e^{(1+\varepsilon)n(\beta(\mu)+\beta(\mu_{\mathbb{R}^m/L}))}\end{aligned} \tag{2.37}$$

Mas, também temos que

$$e^{(1+\varepsilon)2n\beta(\mu)} \leq \|X_{2n}X_{2n-1}\cdots X_1\| \tag{2.38}$$

Sem perda de generalidade, suponha que $\beta(\mu_{\mathbb{R}^m/L}) \leq \beta(\mu_L)$. Então juntando (2.37) e (2.38), temos, com probabilidade um, que para toda $n \geq n_0$

$$e^{(1+\varepsilon)2n\beta(\mu)} \leq 2e^{(1+\varepsilon)n\beta(\mu_L)} + 2e^{(1+\varepsilon)n(\beta(\mu)+\beta(\mu_L))},$$

a qual é equivalente a

$$1 \leq 2e^{2n[(1+\varepsilon)\beta(\mu_L)-(1-\varepsilon)\beta(\mu)]} + 2e^{n[(1+\varepsilon)\beta(\mu_L)-(3-\varepsilon)\beta(\mu)]}. \tag{2.39}$$

Agora, tomamos um ε que faz o expoente da direita da desigualdade (2.39) ser negativo. De fato, note que tomando

$$\varepsilon < \frac{\beta(\mu)-\beta(\mu_L)}{\beta(\mu)+\beta(\mu_L)},$$

conseguimos que o expoente seja negativo. Então para n suficientemente grande temos que:

$$2e^{2n[(1+\varepsilon)\beta(\mu_L)-(1-\varepsilon)\beta(\mu)]} + 2e^{n[(1+\varepsilon)\beta(\mu_L)-(3-\varepsilon)\beta(\mu)]} < \frac{1}{2}$$

O qual contradiz (2.39). Absurdo. Portanto, as duas desigualdades em (2.36) não podem ser estritas $\square$

Agora, sabemos que, por construção, $\beta(\mu_{L_1}) < \beta(\mu)$. Então, pelo teorema anterior, $\beta(\mu_{\mathbb{R}/L}) = \beta(\mu)$. Entretanto o próximo resultado nos diz que, para qualquer $v \in \mathbb{R}/L$, a média espacial é constante e igual a $\beta(\mu)$.

Lema 2.2.14. *Com L_1 definido como acima (L_1 pode ser $\{0\}$), a medida $\mu_{\mathbb{R}^m/L}$ tem a propriedade que, com probabilidade um,*

$$\lim_n \frac{1}{n} \log\|\overline{X}_n \overline{X}_{n-1} \cdots \overline{X}_1 z\| = \beta(\mu)$$

Para cada $[z] \in \mathbb{R}^m/L_1$, com $[z] \notin [0]$, denotamos por $\overline{X_n}$ a transformação induzida em $\mathbb{R}^m/L_1$.

Demonstração. Podemos escrever, com um abuso inofensivo de notação,

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

na qual Y_n denota a restrição de X_n a L_1. Pelo Lema 2.2.13 e como $\beta(\mu_{L_1}) < \beta(\mu)$, temos que $\beta(\mu_{\mathbb{R}^m/L_1}) = \beta(\mu)$. Agora, aplicamos o Teorema 2.2.10 a $\mu_{\mathbb{R}^m/L_1}$. Se a afirmação do Lema 2.2.14 não estivesse certa, $\mathbb{R}^m/L_1$ conteria um subespaço próprio $\overline{L}_2 \subsetneq \mathbb{R}^m/L_1$ tal que a taxa de crescimento $\beta(\mu_{\overline{L}_2}) < \beta(\mu_{\mathbb{R}^m/L_1}) = \beta(\mu)$. Podemos escrever

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

onde as matrizes aleatórias de $\overline{X_n'}$ tem taxa de crescimento estritamente menor que $\beta(\mu)$. Mas, também, a taxa de crescimento de Y_n é menor, pois $\beta(\mu_{L_1}) < \beta(\mu)$. O mesmo é verdade para

$$\begin{pmatrix} Y_n & Z_n' \\ 0 & \overline{X}_n' \end{pmatrix}$$

de acordo com o Lema 2.2.13. Mas isto contradiz o fato que L_1 é subespaço maximal μ-quase invariante que tem taxa de crescimento estritamente menor que $\beta(\mu)$. Isto prova o Lema 2.2.14

□

Isto nos leva a formular

Corolário 2.2.15. *Para todo $v \in \mathbb{R}^m$, com $v \notin L_1$,*

$$\lim_n \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\|$$

existe com probabilidade um e é igual a $\beta(\mu)$.

Demonstração. Como $v \notin L_1$, temos que $[v] \in \mathbb{R}^m \setminus L_1$, com $[v] \notin [0]$. Então o resultado é consequência direta do Lema 2.2.14.

□

Agora podemos fazer a filtração de $\mathbb{R}^m$.

Teorema 2.2.16. *Seja μ uma medida de probabilidade em $GL(m,\mathbb{R})$ satisfazendo* (2.18)*. Então, existe uma sequência de subespaços*

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

e uma sequência de valores $\beta(\mu) = \beta^{(0)}(\mu) > \beta^{(')}(\mu) > \beta^{('')}(\mu) > \cdots > \beta^{(r)}(\mu)$ tais que, se $v \in L_i \setminus L_{i+1}$, temos, com probabilidade um,

$$\lim_n \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\| = \beta^i(\mu).$$

Demonstração. Pelos Lemas 2.2.13 e 2.2.14 e o Corolário 2.2.15 temos que $\mathbb{R}^m = L_0 \supset L_1$ com taxas de crescimento $\beta(\mu) = \beta^0(\mu) > \beta^1(\mu)$ tais que para todo $v \in L_0 \setminus L_1$ temos que

$$\lim_{n\to\infty} \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\| = \beta^0(\mu),$$

na qual L_1 é um subespaço μ-invariante.
Fazendo a mesma análise para L_1 como fizemos para $\mathbb{R}^m$. Poderíamos encontramos um subespaço próprio $L_2 \subset L_1$, que contém todos os subespaços μ-invariantes com $\beta(\mu_{L_2}) < \beta(\mu_{L_1})$ e, para todo $v \in L_2 \setminus L_1$, teremos

$$\lim_n \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\| = \beta(\mu_{L_1})$$

Repetimos o mesmo procedimento até que não existam subespaços próprios com taxa crescimento menor. Assim, concluímos a prova do teorema. □

O seguinte lema é útil para a prova de um dos teoremas principais da dissertação, o lema fala sobre a existência de medidas extremais em $\mathcal{N}$, a qual tem suporte no complemento de $\overline{L_1}$.

Lema 2.2.17. *Sejam μ uma medida de probabilidade em $GL(m,\mathbb{R})$ e L_1 um subespaço próprio de $\mathbb{R}^m$ μ-quase invariante. Então existe uma medida ν extremal em $\mathcal{N}$, tal que $\nu(\overline{L_1}) = 0$.*

Demonstração. Pela compacidade sequencial de $\mathcal{N}$ e pelo Lema (2.2.6) temos que existe um $\nu \in \mathcal{N}$ extremal tal que

$$\beta(\mu) = \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g) d\nu(u). \tag{2.40}$$

Como as medidas extremais são ergódicas, então as médias temporais e as médias espaciais coincidem, com probabilidade um, isto é

$$\lim_n \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\| = \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g) d\nu(u). \tag{2.41}$$

Para $v \in L_1$ temos que

$$\lim_n \frac{1}{n} \log\|X_n X_{n-1} \cdots X_1 v\| < \beta(\mu). \tag{2.42}$$

Juntando (2.40), (2.41) e (2.42) concluímos que $\nu(\overline{L_1}) = 0$.

□

Agora fazemos uma caracterização das médias espaciais das medidas $\nu \in \mathcal{N}$ tais que $\nu(\overline{L_1}) = 0$.

Corolário 2.2.18. *Se existe uma medida $v \in \mathcal{N}$ satisfazendo $\nu(\overline{L_1}) = 0$, então:*

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

Demonstração. Como $\mu^{\mathbb{N}} \otimes \nu$ é invariante por T, segue de imediato do Teorema 2.2.16 e do Teorema ergódico de Birkhoff 1.4.2. □

3 Perturbação do Produto de Matrizes Aleatórias

Neste capítulo estudaremos a estabilidade do máximo expoente de Lyapunov, daremos uma condição suficiente na qual o máximo expoente de Lyapunov seja estável a pequenas perturbações (Teorema 3.1.4), uma das hipóteses é que o suporte da distribuição μ tenha apenas um subespaço μ-quase invariante de $\mathbb{R}^m$, com um Exemplo 3.1.3 mostramos que esta hipóteses não pode ser enfraquecida.

3.1 Estabilidade do Máximo Expoente de Lyapunov.

Definição 3.1.1. Dizemos que o máximo expoente de Lyapunov é estável para pequenas perturbações no caso em que:

$$\textit{se } \mu_k \rightharpoonup \mu \textit{ então } \beta(\mu_k) \longrightarrow \beta(\mu).$$

Lembremos que para um subespaço $L \subset \mathbb{R}^m$ que seja μ-quase invariante, $\beta(\mu_L)$ representa a máxima taxa de crescimento do produto de matrizes aleatórias nas direções de vetores que estejam em L.

Teorema 3.1.2. *Suponha que para todo subespaço μ-quase invariante $L \neq \{0\}$ de $\mathbb{R}^m$ nós temos $\beta(\mu_L) = \beta(\mu)$. Seja $\{\mu_k\}$ uma sequência de medidas de probabilidade em $GL(m, \mathbb{R})$ satisfazendo*

$$\int_{\|g\|>T} \log^+ \|g\| d\mu_k(g) + \int_{\|g^{-1}\|>T} \log^+ \|g^{-1}\| d\mu_k(g) \longrightarrow 0 \tag{3.1}$$

quando $T \longrightarrow \infty$ uniformemente em k. Se $\mu_k \rightharpoonup \mu$, então $\lim_k \beta(\mu_k) = \beta(\mu)$.

Demonstração. Pelo Lema 2.2.17, para cada medida μ_k existe uma medida ν_k de probabilidade sobre $\mathbb{R}P^{m-1}$ com $\mu_k * \nu_k = \nu_k$ e tal que

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

Como $\mathcal{M}_1(\mathbb{R}P^{m-1})$ é sequencialmente compacto, existe uma subsequência (ν_{k_n}) e $\tilde{\nu} \in \mathcal{M}_1(\mathbb{R}P^{m-1})$ tal que a subsequência converge em probabilidade a $\tilde{\nu}$. Sem perda de generalidade denotaremos $\nu_{k_n} = \nu_k$.
Afirmação $\mu * \tilde{\nu} = \tilde{\nu}$.
Seja $f \in \mathcal{C}(\mathbb{R}P^{m-1})$. Temos que a aplicação L definida como

$$\begin{aligned} L : \mathbb{R}P^{m-1} &\longrightarrow \mathbb{R} \\ u &\longmapsto \int_{GL} f(gu) d\mu(g) \end{aligned}$$

é contínua. Isto foi parte da prova do Teorema 2.2.2.

Por outro lado, como f é contínua e tem suporte compacto, existe H tal que o valor absoluto de f está limitado por H. Então, pela desigualdade triangular, pelo fato que $\mu_k * \nu_k = \nu_k$ e por $\nu_k \rightharpoonup \tilde{\nu}$ e $\mu_k \rightharpoonup \mu$, dado $\varepsilon > 0$, existe $k_0 \in N$ tal que, para $k \geq k_0$,

$$\begin{aligned}
&|(\mu * \tilde{\nu} - \tilde{\nu})f| \\
&\quad \leq |(\mu * \tilde{\nu})(f) - (\mu * \nu_k)(f)| + |(\mu * \nu_k)(f) - (\mu_k * \nu_k)(f)| + |(\nu_k)(f) - \tilde{\nu}(f)| \\
&\quad \leq \left| \int \left(\int f(gu) d\mu(g) \right) d(\tilde{\nu} - \nu_k)(g) \right| + H[(\mu_k - \mu)^+(GL) + (\mu_k - \mu)^-(GL)] + \frac{\varepsilon}{4} \\
&\quad \leq \varepsilon.
\end{aligned}$$

Como ε é arbitrário temos a conclusão da afirmação.

Por hipótese temos que $\beta(\mu_L) = \beta(\mu)$ para cada $L \subset \mathbb{R}^m$ subespaço μ invariante , então a filtração $\{L_i\}$ correspondente a μ tem que ser trivial. Portanto não existe subespaço próprio μ-quase invariante. Assim para toda medida $\nu \in \mathcal{N}$ extremal, $\alpha(\nu) = \beta(\mu)$ e como para qualquer outra medida não extremal pode ser escrita como suma convexa de medidas extremais podemos concluir que também vale para qualquer medida $\nu \in \mathcal{N}$. Em particular,

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

Agora definimos $B_n = \{g \in GL; \frac{1}{T} \leq \|g\| \leq T\}$. Pelo fato que $\mu_k * \nu_k \rightharpoonup \mu * \tilde{\nu}$ temos que

$$\lim_{k\to\infty} \int_{B_n} \int_{\mathbb{R}P^{m-1}} \log \frac{\|gu\|}{\|u\|} d\mu_k(g) d\nu_k(u) = \int_{B_n} \int_{\mathbb{R}P^{m-1}} \log \frac{\|gu\|}{\|u\|} d\mu(g) d\tilde{\nu}(u).$$

Por outro lado,

$$\begin{aligned}
&\int_{GL} \int_{\mathbb{R}P^{m-1}} \log \frac{\|gu\|}{\|u\|} d\mu_k(g) d\nu_k(u) \\
&= \int_{B_n} \int_{\mathbb{R}P^{m-1}} \log \frac{\|gu\|}{\|u\|} d\mu_k(g) d\nu_k(u) + \int_{B_n^c} \int_{\mathbb{R}P^{m-1}} \log \frac{\|gu\|}{\|u\|} d\mu_k(g) d\nu_k(u) \\
&\leq \int_{B_n} \int_{\mathbb{R}P^{m-1}} \log \frac{\|gu\|}{\|u\|} d\mu_k(g) d\nu_k(u) + \int_{\|g\|>n} \log^+ \|g\| d\mu_k(g) + \int_{\|g^{-1}\|>n} \log^+ \|g^{-1}\| d\mu_k(g)
\end{aligned}$$

Por (3.1) temos que dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$ a soma dos dois últimos termos da última expressão é menor que ε, para todo $k \in \mathbb{N}$ e também quando $\mu_k = \mu$. Portanto,

$$\lim_k \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu_k(g) d\nu_k(u) = \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g) d\tilde{\nu}(u)$$

Isto prova que $\beta(\mu_k) \longrightarrow \beta(\mu)$ e completa a prova do teorema.

□

Este teorema também nos dá uma maneira prática de calcular o maior expoente de Lyapunov quando não existe subespaço próprio μ-quase invariante $L \subset \mathbb{R}^m$.

Exemplo 3.1.3. Seja a medida de probabilidade μ_ε em $GL(M,\mathbb{R})$ definida como:

$$\mu_\varepsilon = (1-\varepsilon)\delta_A + \varepsilon\delta_J,$$

onde

$$A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \text{ e } J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Com $a > b > 1$.
Veja que não existem $J \in \mathbb{R}^2 \setminus \{0\}$ e $\zeta \in \mathbb{R} \setminus \{0\}$ tal que $XJ = \zeta J$. Portanto, não existe subespaço próprio de $\mathbb{R}^2$ que seja μ-quase invariante.
Definamos uma medida de probabilidade em $\mathbb{R}P^1$,

$$\nu = \frac{1}{2}(\delta_{[e]} + \delta_{[f)}),$$

onde

$$[e] = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ e } [f] = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Pode-se provar que ν é μ_ε-estacionária. Então pelo teorema 3.1.2 temos que

$$\begin{aligned} \beta(\mu_\epsilon) &= \iint \log \frac{\|g\hat{u}\|}{\|\hat{u}\|} d\mu(g) d\nu(u) \\ &= (1-\epsilon)\int \log \frac{\|A\hat{u}\|}{\|\hat{u}\|} d\nu(u) + \epsilon \int \log \frac{\|J\hat{u}\|}{\|\hat{u}\|} d\nu(u) \\ &= \frac{1}{2}(1-\epsilon)(\log(a) + \log(b)). \end{aligned}$$

Agora vamos a enunciar e provar o resultado principal da dissertação. O exemplo de acima tem relevância pois nos mostrará que a hipóteses de que quando o suporte da medida tem associado mais de um subespaço μ-quase invariante de $\mathbb{R}^m$ não necessariamente vamos a poder garantir estabilidade.

Teorema 3.1.4. *Sejam μ uma medida de probabilidade em $GL(m,\mathbb{R})$ e G_μ o menor subgrupo fechado de $GL(m,\mathbb{R})$ com suporte na medida μ. Se G_μ tem a propriedade de que existe no máximo um subespaço não trivial $L \subset \mathbb{R}^m$ o qual é μ-quase invariante, $\mu_k \rightharpoonup \mu$ e $(\mu_k)_{k\in\mathbb{N}}$ satisfazem (3.1), então $\beta(\mu_k) \longrightarrow \beta(\mu)$.*

Demonstração. Seja L o único subespaço próprio μ-quase invariante. Decompomos as matrizes $g \in G_\mu$ da forma

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

com g_{11} como a restrição de g a L e g_{22} como a transformação induzida em $\mathbb{R}^m/L$. Pelo Teorema 3.1.2, se $\beta(\mu_L) = \beta(\mu)$ então temos estabilidade. Por outro lado, se $\beta(\mu_L) <$

$\beta(\mu)$, então, pelo Lema 2.2.13 temos $\beta(\mu_{\mathbb{R}^m/L}) = \beta(\mu)$. Agora consideremos as matrizes transpostas

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

com a correspondente medida μ^T. Existe um único subespaço não trivial μ^T-quase invariante, que será denotado por L^+, onde a restrição de $'g$ a este subespaço é $g22^T$. Agora, se consideramos $\beta(\mu)$ como a taxa de crescimento das matrizes aleatórias, é claro que $\beta(\mu) = \beta(\mu^T)$ e, pela mesma razão, $\beta(\mu_{\mathbb{R}^m/L}) = \beta(\mu^T_{L^+})$. Como $\beta(\mu_{\mathbb{R}^m/L}) = \beta(\mu)$, segue que, para o subespaço não trivial μ^T-quase invariante L^+,

$$\beta(\mu^T_L) = \beta(\mu^T).$$

Novamente, aplicando o Teorema 3.1.2, concluímos que μ^T é estável. Mas isto é equivalente à estabilidade de μ. Isto prova o teorema.

$\square$

Uma variação do Teorema 3.1.4 é o seguinte resultado.

Teorema 3.1.5. *Sejam μ uma medida de probabilidade em $GL(m, \mathbb{R})$ e G_μ o menor subgrupo fechado de $GL(m, \mathbb{R})$ com suporte na medida μ. Seja $V \subset \mathbb{R}^m$ um subespaço μ-quase invariante por G_μ tal que V está contido em qualquer outro subespaço μ-quase invariante. Suponha que $\beta(\mu_V) = \beta(\mu)$. Então sempre que $\mu_k \rightharpoonup \mu$ em probabilidade e satisfazendo (3.1), temos que $\beta(\mu_k) \longrightarrow \beta(\mu)$.*

Um exemplo relevante é quando G_μ consiste em todas as matrizes da forma

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

Seja $\gamma_i = \displaystyle\int \log|g_{ii}| d\mu(g)$ para $i = 1, 2, 3$. Afirmamos que se γ_1 ou γ_3 é o maior dos três números, então teremos estabilidade no sentido dos teoremas precedentes.
Veja que $\beta(\mu) = \max\{\gamma_1, \gamma_2, \gamma_3\}$. Suponha, primeiramente que γ_1 é o maior. Existem dois subespaços μ-quase invariantes, sendo estes

$$L_1 = \left\{ \begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^3 \mid x \in \mathbb{R} \right\}$$

$$L_2 = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \in \mathbb{R}^3 \mid x, y \in \mathbb{R} \right\},$$

e com $\beta(\mu_{L_1})$ e $\beta(\mu_{L_2})$ iguais a $\beta(\mu)$. Então as condições do Teorema 3.1.2 são cumpridas para μ e portanto, temos estabilidade.

Agora, suponha que γ_3 é o maior. Então os únicos subespaços μ-quase invariantes de $\mathbb{R}^m$ com respeito á medida transposta μ^T são

$$L_1^T = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \in \mathbb{R}^3 \mid z \in \mathbb{R} \right\},$$

$$L_2^T = \left\{ \begin{pmatrix} 0 \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid y, z \in \mathbb{R} \right\}.$$

Agora, como $\beta(\mu) = \beta(\mu^T)$, segue que $\beta(\mu^T_{L_1^T})$ e $\beta(\mu^T_{L_2^T})$ são iguais a $\beta(\mu^T)$. Então as condições do Teorema 3.1.2 são cumpridas para μ^T e, portanto, temos estabilidade para μ^T o que é equivalente a estabilidade para μ.

Uma pergunta interessante é se é possível encontrar uma medida de probabilidade μ em $GL(m, \mathbb{R})$ com as mesmas condições do Teorema 3.1.4, mas com mais de um subespaço quase invariante. Isto é equivalente a perguntar nos se é possível estender o Teorema 3.1.4, no sentido de ter dois subespaços próprios μ-quase invariantes. A resposta é não é possível. Considere o caso $m = 2$. Há três possibilidades no que diz respeito aos subespaços quase invariantes para G_μ: ou não há subespaço não-trivial, ou há apenas um, ou existem dois subespaços invariantes. Nos dois primeiros casos teremos estabilidade de $\beta(\mu)$ pelo Teorema 3.1.4, mas no terceiro caso, não necessariamente teremos estabilidade de $\beta(\mu)$.

Por Exemplo, definamos $\mu_\epsilon = (1-\epsilon)\delta_A + \epsilon\delta_J$ como no Exemplo 3.1.3 e $\mu = \delta_A$. Veja que existem dois subespaços próprios de $\mathbb{R}^2$ que são μ-quase invariantes. É claro que $\mu_\epsilon \rightharpoonup \mu$ quando $\epsilon \longrightarrow 0$. Lembremos que $\beta(\mu_\epsilon) = \dfrac{1}{2}(1-\epsilon)\log(ab)$, mas por outro lado temos que $\beta(\mu) = \log(a)$, isto é $\beta(\mu_\epsilon) \nrightarrow \beta(\mu)$ quando $\epsilon \longrightarrow 0$. Portanto, não temos estabilidade para $\beta(\mu)$.

Referências

[Bar95] Robert Gardner Bartle. *The elements of integration and Lebesgue measure*, volume 27. Wiley Online Library, 1995.

[Bre10] Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Springer Science & Business Media, 2010.

[FK60] Harry Furstenberg and Harry Kesten. Products of random matrices. *The Annals of Mathematical Statistics*, 31(2):457–469, 1960.

[FK83] Harry Furstenberg and Yuri Kifer. Random matrix products and measures on projective spaces. *Israel Journal of Mathematics*, 46(1-2):12–32, 1983.

[Kre78] Erwin Otto Kreyszig. *Introductory functional analysis with applications*, volume 1. Wiley New York, 1978.

[KS82] Yuri Kifer and Eric Slud. Perturbations of random matrix products in a reducible case. *Ergodic Theory and Dynamical Systems*, 2(3-4):367–382, 1982.

[OV14] Krerley Oliveira and Marcelo Viana. *Fundamentos da teoria ergódica.* IMPA, Brazil, 2014.

[Par05] Kalyanapuram Rangachari Parthasarathy. *Probability measures on metric spaces*, volume 352. American Mathematical Soc., 2005.

[RF88] Halsey Royden and Patrick Fitzpatrick. *Real analysis*, volume 32. Macmillan New York, 1988.

[Ruf92] Paulo Régis Ruffino. Produto de matrizes aleátorias. Master's thesis, Instituto de Matemática, Estatística e Computação Científica-UNICAMP, 1992.

[Sch86] Martin Schechter. *A Course in Functional Analysis (John B. Conway)*, volume 28. Society for Industrial and Applied Mathematics, 1986.

[Wil91] David Williams. *Probability with martingales.* Cambridge university press, 1991.