Universidade Federal do Piauí
Centro de Ciências da Natureza
Pós-Graduação em Matemática
Mestrado em Matemática

Teoremas Ergódicos e Aplicações

Jefferson Victor de Sousa Galvão

Teresina - 2023

Jefferson Victor de Sousa Galvão

Dissertação de Mestrado:

Teoremas Ergódicos e Aplicações

Dissertação submetida à Coordenação do Programa de Pós-Graduação em Matemática, da Universidade Federal do Piauí, como requisito parcial para obtenção do grau de Mestre em Matemática.

Orientador:
Prof. Dr. Ítalo Dowell Lira Melo

Teresina - 2023

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PIAUÍ
CENTRO DE CIÊNCIAS DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

Teoremas Ergódicos e Aplicações

Jefferson Victor de Sousa Galvão

Esta Dissertação foi submetida como parte dos requisitos necessários à obtenção do grau de **Mestre em Matemática**, outorgado pela Universidade Federal do Piauí.

A citação de qualquer trecho deste trabalho é permitida, desde que seja feita em conformidade com as normas da ética científica.

Dissertação aprovada em 14 de Julho de 2023.

Banca Examinadora:

Ítalo Dowell Lira Melo

Prof. Dr. Ítalo Dowell Lira Melo - Orientador

Sandoel de Brito Vieira

Prof. Dr. Sandoel de Brito Vieira (UFPI)

Paulo Alexandre Araújo Sousa

Prof. Dr. Paulo Alexandre Araújo Sousa (UFPI)

Prof. Dr. Mauricio José Poletti Merlo (UFC)

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Sistema de Bibliotecas UFPI - SIBi/UFPI
Biblioteca Setorial do CCN

G182t Galvão, Jefferson Victor de Sousa.
Teoremas ergódicos e aplicações / Jefferson Victor de Sousa Galvão. – 2023.
88 f.

Dissertação (Mestrado) - Universidade Federal do Piauí, Programa de Pós-Graduação em Matemática, Teresina, 2023.
"Orientador: Prof. Dr. Ítalo Dowell Lira Melo."

1. Teoria ergódica 2. Teorema de Weyl. 3. Projeções alternadas. I. Melo, Jefferson Victor de Sousa. II. Titulo.

CDD 515

Bibliotecária: Caryne Maria da Silva Gomes - CRB3/1461

Dedicatória

Aos sonhos que foram interrompidos precocemente, dedico este trabalho ao meu primo querido, Arthur Galvão. Sua memória viverá eternamente em nossos corações, e cada conquista minha será uma homenagem ao brilho que ele trouxe às nossas vidas.

Com saudade,

Jefferson Galvão.

Agradecimentos

Neste momento tão especial em minha vida acadêmica, é com imensa gratidão que desejo expressar meus sinceros agradecimentos a todas as pessoas que tornaram possível esta conquista.

Em primeiro lugar, agradeço a Deus por me abençoar com saúde e inteligência, proporcionando-me as condições necessárias para enfrentar os desafios e alcançar meus objetivos.

À minha amada mãe, Joselma, e ao meu pai, Gerson, sou imensamente grato pelo amor incondicional, pelo apoio incansável em meus estudos e por serem exemplos inspiradores em minha formação como pessoa. Vocês são a base sólida em que construí minha trajetória e meu caráter.

Meu sincero agradecimento ao meu tio Carlos André e à minha querida vó Raimunda por estenderem suas mãos e me oferecerem auxílio nos momentos em que precisei. Suas palavras de encorajamento e apoio foram fundamentais em minha caminhada.

À minha irmã Gabryella, agradeço por estar sempre ao meu lado, oferecendo seu apoio e carinho incondicional durante esta etapa tão desafiadora. Sua presença fez toda a diferença.

Aos respeitados professores da UFPI, em especial aos professores Kelson e Ítalo, quero expressar minha profunda gratidão por compartilharem seus conhecimentos e por desempenharem um papel essencial em minha formação acadêmica. Suas aulas e orientações foram fundamentais para o meu crescimento intelectual.

Também agradeço aos meus amigos de graduação e mestrado, Sillas, Gabriel, Suerlan, Paulo Sergio, Paulo Henrique, Raquel, Fauster, Tiffany, Jonatas, Danilo Santos, Danilo Oliveira, Rafael Douglas, Erisvaldo, Ana Júlia, João Victor, João Victor Carvalho, Honório, Emanuelly, Luzivania, Dieme, José, Vinicius, Eduardo, José Vitor, Dario, João Vinicios, Ruan Diego, Vitor Gabriel, Raylson, Brenner, Samuel e Isaque, pelos inúmeros momentos de descontração, estudo sério, diversão e café compartilhados ao longo dessa jornada. Suas amizades foram um apoio inestimável que tornou essa caminhada mais significativa.

Cada um de vocês, de alguma forma, contribuiu para o meu crescimento e enriqueceu a minha vida acadêmica e pessoal. É impossível listar todos os momentos bons que compartilhamos, mas cada lembrança será guardada com carinho em meu coração.

Mais uma vez, a todos que fazem parte desta conquista, meu profundo agradecimento. Sem vocês, nada disso teria sido possível.

Muito obrigado!

Agradeço CAPES pelo apoio financeiro.

"A ambição universal dos homens é viver colhendo o que nunca plantaram.".

Adam Smith.

Resumo

Estudamos neste trabalho diversos teoremas ergódicos que desempenham um papel importante na Teoria Ergódica, entre eles, o teorema ergódico de Birkhoff, o teorema ergódico maximal e o teorema ergódico subaditivo de Kingmann. Além disso, aplicaremos tais resultados para demonstrar o teorema de Weyl e estudar o conjunto das sequências quase-normais que recentemente foi utilizado para demonstrar uma conjectura relacionada ao método das projeções alternadas.

Palavras-chave: Teoremas ergódicos, Teorema ergódico de Birkhoff, Ergodicidade, Teorema de Weyl, Sequências quase-normais, Projeções alternadas.

Abstract

In this work we study several ergodic theorems that play an important role in Ergodic Theory, among them, Birkhoff's ergodic theorem, maximal ergodic theorem and Kingmann's subadditive ergodic theorem. Furthermore, we will apply these results to prove Weyl's theorem and study the set of the quasi-normal sequences that was recently used to prove a conjecture related to the method of alternating projections.

Keywords: Ergodic Theorems, Birkhoff Ergodic Theorem, Ergodicity, Weyl's Theorem, Quasi-Normal Sequences, Alternating Projections.

Sumário

Introdução

A Teoria Ergódica estuda sistemas dinâmicos munidos de uma medida invariante. As origens da Teoria Ergódica remetem à mecânica estatística, com uma tentativa de aplicar a teoria da probabilidade à sistemas mecânicos conservativos. A palavra “ergódico” foi criada no século XIX pelo físico austríaco L. Boltzmann, um dos fundadores da teoria cinética dos gases. Boltzmannn estava interessado nos sistemas descritos por fluxos Hamiltonianos. Com o tempo, tornou-se usual chamar hipótese ergódica a afirmação de que as médias temporais e espaciais são iguais. Sistemas para os quais vale essa igualdade foram chamados ergódicos. Em 1890, Henri Poincaré obteve um importante resultado na direção de entender o comportamento de órbitas de um sistema dinâmico discreto, este resultado ficou conhecido como Teorema de Recorrência de Poincaré, que afirma que se uma medida finita é invariante por uma transformação $T : X \to X$, então quase todo ponto de um conjunto mensurável com medida positiva retorna infinitamente para ele. Desta forma, é natural pensar como é o comportamento médio dessas órbitas por meio de uma função. Dados $x \in X$, $E \subset X$ e $n \in \mathbb{N}$ podemos definir o número de iterados de x que visitam E até o tempo $(n-1)$ por $\tau_n(E, x) = \#\left\{j \in \{0, 1, ..., n-1\} : T^j(x) \in E\right\}$. Quando o limite

$$\lim_{n\to\infty} \frac{\tau_n(E, x)}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_E(T^j(x)) \tag{1}$$

existe, ele nos fornece a frequência com que a órbita de x visita o conjunto E. O Teorema de Recorrência de Poincaré afirma que $\tau_n(E, x) \to \infty$ quando $n \to \infty$ quase sempre, porém ele não nos dá nenhuma informação sobre o limite (1).

Outro matemático que estudou as médias temporais e espaciais foi Hermann Weyl. Em 1910, Weyl provou que para todo número irracional ξ a sequência $x_n = n\xi \pmod 1$ é equidistribuída em $[0, 1]$, ou seja, para todo $0 \leqslant a \leqslant b \leqslant 1$, temos

$$\lim_{n\to\infty} \frac{\#\{1 \leqslant j \leqslant n : x_j \in (a, b)\}}{n} = b - a. \tag{2}$$

Porém foi em 1916 que Weyl em [20] publicou um paper que deu origem á teoria da distribuição uniforme. Em 1929, Koopmam começou a investigar grupos de transformações que preservam medidas, ou seja, transformações que levam cada conjunto em outro que mede o mesmo. Os teoremas ergódicos de von Neumann e Birkhoff foram provados nos anos trinta do século passado e forneceram informações mais precisas sobre as médias temporais, tais resultados desempenharam um importante papel no desenvolvimento da Teoria Ergódica.

Devido a grande importância dos teoremas ergódicos no desenvolvimento da Teoria Ergódica e a sua grande aplicabilidade em diversas áreas da matemática, neste trabalho estudaremos diversos teoremas ergódicos, entre eles, o teorema ergódico de von Neumann, o teorema ergódico de Birkhoff, o teorema ergódico maximal e teorema ergódico de Kingmann. No Capítulo 2, apresentaremos a demonstração destes teoremas e as relações existententes entre eles, por exemplo, o teorema ergódico de Birkhoff pode ser demonstrado como consequência do teorema ergódico maximal. No Capítulo 3 abordaremos o conceito de ergodicidade que possui grande utilidade em vários ramos da matemática. No Capítulo 4, aplicaremos o teorema de Birkhoff para demonstrar o teorema de Weyl e para estudar o conjunto das sequências quase-normais que recentemente foi utilizado para demonstrar uma conjectura relacionada ao método das projeções alternadas.

Capítulo 1

Noções Preliminares

Neste capítulo apresentaremos algumas definições e resultados que serão utilizados ao longo deste trabalho. Parte substancial deste capítulo encontra-se em [18]. Veja [5], [8] e [4].

1.1 Espaços de medida

1.1.1 Espaços mensuráveis

Dado um subconjunto $A \subset X$ denotaremos por A^c o complementar $X \setminus A$ do conjunto A em relação a X.

Definição 1.1. *Uma álgebra de X é uma família $\mathcal{B}$ de subconjuntos de X que é fechada para as operações elementares de conjuntos e contém o conjunto vazio, isto é, tal que*

- $\emptyset \in \mathcal{B}$ *;*
- $A \in \mathcal{B}$ *implica* $A^c \in \mathcal{B}$*;*
- $A \in \mathcal{B}$ *e* $B \in \mathcal{B}$ *implica* $A \cup B \in \mathcal{B}$.

Observe que $A \cap B = (A^c \cup B^c)^c$ também está em $\mathcal{B}$ para quaisquer $A, B \in \mathcal{B}$. Além disso, a união e interseção finita de quaisquer elementos de $\mathcal{B}$ também é um elemento de $\mathcal{B}$.

Definição 1.2. *Uma álgebra diz-se uma σ-álgebra de subconjuntos de X se também for fechada para as uniões enumeráveis:*

- $A_j \in \mathcal{B}$ *para* $j = 1, ..., n, ...$ *implica* $\bigcup_{j=1}^{\infty} A_j \in \mathcal{B}$.

Observe que uma σ-álgebra $\mathcal{B}$ também é fechada para as interseções enumeráveis, de fato, se $A_j \in \mathcal{B}, \forall j \in \mathbb{N}$, então $\bigcap_{j=1}^{\infty} A_j = (\bigcup_{j=1}^{\infty} A_j^c)^c$ também está em $\mathcal{B}$.

Definição 1.3. *Um espaço mensurável é uma dupla* $(X, \mathcal{B})$*, onde X é um conjunto e* $\mathcal{B}$ *é uma* σ*-álgebra de subconjuntos de* X*. Os elementos de* $\mathcal{B}$ *são chamados de conjuntos mensuráveis do espaço.*

A seguir, temos a seguinte Proposição:

Proposição 1.1. *Considere uma família* $\{\mathcal{B}_i : i \in \mathcal{I}\}$ *de* σ*-álgebras, onde* $\mathcal{I}$ *é um conjunto qualquer. Então a interseção* $\mathcal{B} = \bigcap_{i \in \mathcal{I}} \mathcal{B}_i$ *também é uma* σ*-álgebra.*

Definição 1.4. *A* σ*-álgebra gerada por uma família* $\mathcal{E}$ *de subconjuntos de* X *é a menor* σ*-álgebra que contém a família* $\mathcal{E}$*, ou seja, é a interseção de todas as* σ*-álgebras que contêm* $\mathcal{E}$*. Denotaremos tal* σ*-álgebra por* $\sigma(\mathcal{E})$*.*

Definição 1.5. *A* σ*-álgebra de Borel de um espaço topológico é a* σ*-álgebra gerada pela topologia* τ*, isto é, a menor* σ*-álgebra que contém todos os subconjuntos abertos. Neste caso, os conjuntos mensuráveis recebem o nome de borelianos.*

1.1.2 Espaços de medida

Definição 1.6. *Seja* $(X, \mathcal{B})$ *um espaço mensurável. Uma medida em* $(X, \mathcal{B})$ *é uma função* $\mu : \mathcal{B} \to [0, +\infty]$ *tal que* $\mu(\emptyset) = 0$ *e*

$$\mu(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu(A_j),$$

para qualquer família enumerável de conjuntos $A_j \in \mathcal{B}$ *disjuntos dois-a-dois. Esta última propriedade é chamada de* σ*-aditividade. A tripla* $(X, \mathcal{B}, \mu)$ *é chamada espaço de medida. Quando* $\mu(X) < \infty$ *dizemos que* μ *é uma medida finita. No caso particular em que* $\mu(X) = 1$*, dizemos que* μ *é uma probabilidade. Neste caso* $(X, \mathcal{B}, \mu)$ *é chamado de espaço de probabilidade.*

Definição 1.7. *Dizemos que uma medida* μ *é* σ*-finita se existe uma família enumerável de subconjuntos mensuráveis* $A_1, ..., A_n, ...$ *de* X *tal que :* $\mu(A_i) < \infty$*, para todo* i*, e*

$$X = \bigcup_{i=1}^{\infty} A_i.$$

Definição 1.8. *Dizemos que um conjunto mensurável* A*, pertencente á* σ*-álgebra* $\mathcal{B}$ *tem medida nula se* $\mu(A) = 0$.

Uma propriedade importante de conjuntos de medida nula é que a união enumerável de conjuntos de medida nula também é um conjunto de medida nula. No próximo resultado enunciamos um teorema de extensão de medidas.

Teorema 1.1. *Seja* $\mathcal{A}$ *uma álgebra de subconjuntos de* X *e seja* $\mu_0 : \mathcal{A} \to [0, +\infty]$ *uma função* σ*-aditiva com* $\mu_0(X) < \infty$*. Então existe uma única medida* μ *definida na* σ*-álgebra* $\mathcal{B}$ *gerada por* $\mathcal{A}$ *que é uma extensão de* μ_0*, ou seja, tal que* $\mu(A) = \mu_0(A)$ *para todo* $A \in \mathcal{A}$.

Definição 1.9. *Dizemos que uma família não vazia* $\mathcal{C}$ *de subconjuntos de* X *é uma classe monótona, se* $\mathcal{C}$ *contém* X *e é fechada para as uniões e interseções enumeráveis monótonas, ou seja:*

- *dados subconjuntos* $A_1 \subset A_2 \subset \ldots$ *em* $\mathcal{C}$ *, então* $\bigcup_{n \geqslant 1} A_n \in \mathcal{C}$ *e*
- *dados subconjuntos* $A_1 \supset A_2 \supset \ldots$ *em* $\mathcal{C}$ *, então* $\bigcap_{n \geqslant 1} A_n \in \mathcal{C}$.

Teorema 1.2. *A menor classe monótona que contém uma álgebra* $\mathcal{A}$ *coincide com a* σ*-álgebra gerada por* $\mathcal{A}$.

Teorema 1.3. *(Aproximação) Seja* $(X, \mathcal{B}, \mu)$ *um espaço de probabilidade e seja* $\mathcal{A}$ *uma álgebra que gera a* σ*-álgebra* $\mathcal{B}$*. Então para todo* $\varepsilon > 0$ *e todo* $B \in \mathcal{B}$ *existe* $A \in \mathcal{A}$ *tal que* $\mu(A \Delta B) < \varepsilon$*, onde:*

$$A \Delta B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B).$$

O resultado acima nos diz que todo elemento da σ-álgebra gerada por uma álgebra pode ser aproximado por um elemento da álgebra.

Definição 1.10. *Um espaço de medida diz-se completo se todo subconjunto de um conjunto mensurável com medida nula também é mensurável.*

1.1.3 Aplicações mensuráveis

Definição 1.11. *Dados espaços mensuráveis* $(X, \mathcal{B})$ *e* $(Y, \mathcal{C})$*, dizemos que uma aplicação* $f : X \to Y$ *é mensurável se* $f^{-1}(C) \in \mathcal{B}$*, para todo* $C \in \mathcal{C}$.

Proposição 1.2. *Sejam* $f, g : X \to [-\infty, +\infty]$ *funções mensuráveis e sejam* $a, b \in \mathbb{R}$. *Então também são mensuráveis as seguintes funções:*

$$(af + bg)(x) = af(x) + bg(x) \qquad e \qquad (f \cdot g)(x) = f(x) \cdot g(x).$$

Além disso, se $f_n : X \to [-\infty, +\infty]$ *é uma sequência de funções mensuráveis, também são mensuráveis as seguintes funções:*

$$s(x) = \sup\{f_n(x) : n \geqslant 1\} \qquad e \qquad i(x) = \inf\{f_n(x) : n \geqslant 1\},$$

$$f^*(x) = \limsup_{n \to \infty} f_n(x) \qquad e \qquad f_*(x) = \liminf_{n \to \infty} f_n(x).$$

Em particular, se $f(x) = \lim_{n \to \infty} f_n(x)$ *existe, então* f *é mensurável.*

Definição 1.12. *Dizemos que uma função* $s : X \to \mathbb{R}$ *é simples se existem constantes* $\alpha_1, ..., \alpha_k \in \mathbb{R}$ *e conjuntos mensuráveis* $A_1, ..., A_K \in \mathcal{B}$ *disjuntos dois a dois tais que*

$$s = \sum_{j=1}^{k} \alpha_j \mathcal{X}_{A_j},$$

onde $\mathcal{X}_A$ *é a função característica do conjunto* A.

Observação 1.1. *Toda função simples é mensurável.*

Proposição 1.3. *Seja* $f : X \to [-\infty, +\infty]$ *uma função mensurável. Então existe uma sequência* (s_n) *de funções simples tal que* $|s_n(x)| \leqslant |f(x)|$ *para todo* n *e*

$$\lim_{n \to \infty} s_n(x) = f(x), \forall x \in X.$$

Se f *toma valores em* $\mathbb{R}$, *podemos tomar* s_n *com valores em* $\mathbb{R}$. *Se* f *é limitada, a sequência pode ser escolhida tal que a convergência seja uniforme. Se* f *é não negativa, podemos tomar* $0 \leqslant s_1 \leqslant s_2 \leqslant ... \leqslant f$.

1.2 Integração em espaços de medida

Nesta seção definiremos a integral de Lebesgue de uma função em relação a uma medida.

1.2.1 Integral de Lebesgue

Seja $(X, \mathcal{B}, \mu)$ um espaço de medida.

Definição 1.13. *Seja* $s = \sum_{j=1}^{k} \alpha_j \mathcal{X}_{A_j}$ *uma função simples. Então a integral de* s *em relação á medida* μ *é dada por:*

$$\int s\, d\mu = \sum_{j=1}^{k} \alpha_j \mu(A_j).$$

Agora definiremos a integral de Lebesgue para funções não negativas.

Definição 1.14. *Seja* $f : X \to [0, +\infty]$ *uma função mensurável não negativa. Então*

$$\int f\, d\mu = \lim_{n\to\infty} \int s_n\, d\mu,$$

onde $s_1 \leqslant s_2 \leqslant \ldots$ *é uma sequência não decrescente de funções simples tal que* $\lim\limits_{n\to\infty} s_n(x) = f(x)$ *para todo* $x \in X$.

Para estender a definição de integral a qualquer função mensurável, observemos que dada uma função $f : X \to [-\infty, +\infty]$ sempre podemos escrever $f = f^+ - f^-$ com

$$f^+ = \max\{f(x), 0\} \quad \text{e} \quad f^- = \max\{-f(x), 0\}.$$

Definição 1.15. *Seja* $f : X \to [-\infty, +\infty]$ *uma função mensurável. Então,*

$$\int f\, d\mu = \int f^+\, d\mu - \int f^-\, d\mu,$$

desde que pelo menos uma das integrais do lado direito seja finita.

Definição 1.16. *Dizemos que uma função* $f : X \to [-\infty, +\infty]$ *é integrável se for mensurável e sua integral for um número real. Denotamos o conjunto das funções integráveis por* $L^1(X, \mathcal{B}, \mu)$ *ou, mais simplesmente, por* $L^1(\mu)$.

Da definição 1.15 segue que $|f| \in L^1(\mu)$ se, e somente se, $f \in L^1(\mu)$.

Definição 1.17. *Dada uma função mensurável* $f : X \to [-\infty, +\infty]$ *e um conjunto mensurável* E *definimos a integral de* f *sobre* E *por*

$$\int_E f\, d\mu = \int f \cdot \mathcal{X}_E\, d\mu.$$

Proposição 1.4. *O conjunto* $L^1(\mu)$ *das funções reais integráveis é um espaço vetorial real e a aplicação* $I : L^1(\mu) \to \mathbb{R}$ *dada por* $I(f) = \int f \, d\mu$ *é um funcional linear positivo. Além disso,* $|\int f \, d\mu| \leqslant \int |f| \, d\mu$ *se* $f \in L^1(\mu)$*. Em particular,* I *é um funcional linear limitado.*

Definição 1.18. *Dizemos que uma propriedade é válida em* μ*-quase todo ponto se é válida em todo o* X *exceto, possivelmente, num conjunto de medida nula. Analogamente, dizemos que duas funções* f *e* g *são iguais em* μ*-quase todo ponto se existe uma conjunto mensurável* N *com* $\mu(N) = 0$ *tal que* $f(x) = g(x)$ *para todo* $x \in X \backslash N$*. Neste caso, supondo que as funções sejam integráveis, as suas integrais coincidem.*

Lema 1.1. *(Borel-Cantelli) Sejam* (E_n) *uma família enumerável de conjuntos mensuráveis e* F *o conjunto dos pontos que pertencem a* E_n*, para infinitos valores de* n*, ou seja,*

$$F = \limsup_n E_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n.$$

Se $\sum_{n \in \mathbb{N}} \mu(E_n) < \infty$*, então* $\mu(F) = 0$.

1.2.2 Teoremas de convergência

Nesta subseção mencionamos importantes resultados de convergência de funções que serão utilizados na demonstração dos principais resultados desta dissertação.

Teorema 1.4. *(Convergência monótona) Seja* $f_n : X \to [-\infty, +\infty]$ *uma sequência não-decrescente de funções mensuráveis não-negativas e seja* $f : X \to [-\infty, +\infty]$ *a função definida por* $f(x) = \lim_{n\to\infty} f_n(x)$*. Então*

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

Teorema 1.5. *(Lema de Fatou) Seja* $f_n : X \to [0, +\infty]$ *uma sequência de funções mensuráveis não-negativas. Então, a função* $f : X \to [-\infty, +\infty]$ *definida por* $f(x) = \liminf_{n\to\infty} f_n(x)$ *é mensurável e vale*

$$\int \liminf_{n\to\infty} f_n(x) \, d\mu \leqslant \liminf_{n\to\infty} \int f_n \, d\mu.$$

Teorema 1.6. *(Convergência dominada) Seja* $f_n : X \to \mathbb{R}$ *uma sequência de funções mensuráveis e suponha que existe uma função integrável* g *tal que* $|f_n(x)| \leqslant |g(x)|$ *para*

μ-quase todo x em X. Suponha também que a sequência (f_n) converge em μ-quase todo ponto para uma função $f : X \to \mathbb{R}$. Então f é integrável e vale :

$$\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu.$$

Teorema 1.7. *(Egorov) Seja $(X, \mathcal{A}, \mu)$ um espaço de medida com $\mu(X) < \infty$. Seja (f_n) uma sequência de funções mensuráveis, convergindo em μ-q.t.p. $x \in X$ para $f : X \to \mathbb{R}$. Então, dado qualquer $\delta > 0$, existe $E \in \mathcal{A}$ com $\mu(E^c) \leqslant \delta$ e tal que $(f_n|_E)$ converge uniformemente.*

1.2.3 Produto de medidas

Sejam $(X_k, \mathcal{A}_k, \mu_k)$ espaços de medida finita, para $k = 1, ..., n$. Podemos tornar o produto cartesiano $X_1 \times \cdots \times X_n$ um espaço de medida da seguinte forma, considere em $X_1 \times \cdots \times X_n$ a σ-álgebra gerada pela família de todos os conjuntos da forma $A_1 \times \cdots \times A_n$ com $A_j \in \mathcal{A}_j$. Ela é chamada σ-álgebra produto e é representada por $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$.

Teorema 1.8. *Existe uma única medida μ em $(X_1 \times \cdots \times X_n, \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n)$ tal que $\mu(A_1 \times \cdots \times A_n) = \mu_1(A_1) \cdots \mu_n(A_n)$ para todo $A_1 \in \mathcal{A}_1, \cdots, A_n \in \mathcal{A}_n$. Em particular, μ é uma medida finita.*

1.2.4 Derivação de medidas

Definição 1.19. *Sejam μ e ν duas medidas num mesmo espaço mensurável $(X, \mathcal{B})$. Dizemos que ν é absolutamente contínua em relação a μ se todo conjunto mensurável E que satisfaz $\mu(E) = 0$ também satisfaz $\nu(E) = 0$. Nesse caso escrevemos $\nu \ll \mu$.*

Teorema 1.9. *(Radon-Nikodym) Se μ e ν são medidas finitas tais que $\nu \ll \mu$, então existe uma função mensurável $\rho : X \to [0, +\infty]$ tal que $\nu = \rho\mu$, ou seja, tal que*

$$\int \varphi \, d\nu = \int \varphi\rho \, d\mu$$

para toda função mensurável limitada $\varphi : X \to \mathbb{R}$. Em particular, $\nu(E) = \int_E \rho \, d\mu$, para todo conjunto mensurável $E \subset X$.

Notação: Em geral dizemos que $\dfrac{d\nu}{d\mu} = \rho$.

1.3 Medidas invariantes

Definição 1.20. *Sejam* $(M, \mathcal{B}, \mu)$ *um espaço de medida e* $T : M \to M$ *uma transformação mensurável. A medida* μ *é invariante por* T *se*

$$\mu(E) = \mu(T^{-1}(E))$$

para todo conjunto mensurável $E \subset M$*. Neste caso, também dizemos que* T *preserva* μ*.*

Proposição 1.5. *Seja* $T : M \to M$ *uma transformação mensurável e* μ *uma medida finita em* M*. Então* μ *é invariante por* T *se e somente se para toda função* $\phi \in L^1(\mu)$ *vale que* $\phi \circ T \in L^1(\mu)$ *e*

$$\int_M \phi \, d\mu = \int_M (\phi \circ T) \, d\mu \tag{1.1}$$

Demonstração: Primeiramente, assuma (1.1) para toda $\phi : M \to \mathbb{R}$ integrável. Dado um conjunto mensurável $B \subset M$, considere a função $\phi = \mathcal{X}_B$. Como $\int_M |\phi| \, d\mu = \mu(B) < \infty$ segue que $\mathcal{X}_B \in L^1(\mu)$.

Afirmo que $\mathcal{X}_B \circ T = \mathcal{X}_{T^{-1}(B)}$. De fato, para todo $x \in M$, observe que se $\mathcal{X}_{T^{-1}(B)}(x) = 1$ então $x \in T^{-1}(B)$. Em particular, $T(x) \in B$ e $\mathcal{X}_B \circ T(x) = 1$. De modo análogo se $\mathcal{X}_{T^{-1}(B)}(x) = 0$ então $\mathcal{X}_B \circ T(x) = 0$. Portanto,

$$\mu(B) = \int_M \mathcal{X}_B \, d\mu = \int_M \mathcal{X}_B \circ T \, d\mu = \int_M \mathcal{X}_{T^{-1}(B)} \, d\mu = \mu(T^{-1}(B))$$

Isto prova que μ é invariante por T. Agora assuma que μ é invariante por T, então temos que

$$\int_M \mathcal{X}_B \, d\mu = \int_M \mathcal{X}_B \circ T \, d\mu$$

e $\mathcal{X}_B \circ T \in L^1(\mu)$. Assim, (1.1) é válida para funções características.

Se s é uma função simples, isto é, s é uma combinação linear finita de funções características, usando a linearidade da integral, é fácil ver que

$$\int_M (s \circ T) \, d\mu = \int_M s \, d\mu.$$

Considere uma função arbitrária $\phi \in L^1(\mu)$. Podemos escrever $\phi = \phi^+ - \phi^-$, onde ϕ^+ e ϕ^- é a parte positiva e a parte negativa de ϕ, respectivamente. Como $\phi^+, \phi^- \geqslant 0$, é suficiente provar a igualdade (1.1) para funções $\phi \in L^1(\mu)$ tal que $\phi \geqslant 0$. Para isso, seja (s_n) uma sequência de funções simples tal que

$$0 \leqslant s_n \leqslant s_{n+1} \leqslant \phi \text{ para todo } n \in \mathbb{N},$$

tal que $s_n \to \phi$ pontualmente. Do teorema da convergência monótona segue que

$$\int_M s_n \, d\mu \longrightarrow \int_M \phi \, d\mu$$

quando $n \to \infty$. Como $s_n \circ T \nearrow \phi \circ T$ quando $n \to \infty$, do teorema da convergência monótona segue que

$$\begin{aligned} \int_M (\phi \circ T) \, d\mu &= \lim_{n\to\infty} \int_M (s_n \circ T) \, d\mu \\ &= \lim_{n\to\infty} \int_M s_n \, d\mu = \int_M \phi \, d\mu \end{aligned}$$

e isto completa a prova da Proposição. ■

O próximo resultado nos diz que para provar que uma medida é invariante basta analisar os conjuntos numa álgebra que gera a σ-álgebra.

Lema 1.2. *Sejam* $T : M \to M$ *uma transformação mensurável e* μ *uma medida finita em* M*. Suponha que existe uma álgebra* $\mathcal{A}$ *de subconjuntos mensuráveis de* M *tal que* $\mathcal{A}$ *gera a* σ*-álgebra* $\mathcal{B}$ *de* M *e* $\mu(E) = \mu(T^{-1}(E))$ *para todo* $E \in \mathcal{A}$*. Então o mesmo vale para todo conjunto* $E \in \mathcal{B}$*, isto é, a medida* μ *é invariante por* T*.*

1.4 Espaços $L^p(\mu)$

1.4.1 Espaços $L^p(\mu)$ com $1 \leqslant p < \infty$

Definição 1.21. *Dado qualquer* $p \in [1, \infty)$*, dizemos que* $f : X \to \mathbb{C}$ *é uma função p-integrável com relação a* μ *se a função* $|f|^p$ *é integrável em relação a* μ*.*

Definição 1.22. *Denotamos por* $L^p(\mu)$ *o conjunto das funções complexas p-integráveis com relação a* μ*, módulo a relação de equivalência que identifica quaisquer funções que são iguais em* μ*-quase todo ponto. Para cada função* $f \in L^p(\mu)$*, definimos a norma* L^p *de* f*:*

$$\|f\|_p = \left(\int |f|^p \, d\mu \right)^{\frac{1}{p}}.$$

O próximo resultado garante que $L^p(\mu)$ é um espaço de Banach.

Teorema 1.10. *O conjunto* $L^p(\mu)$ *é um espaço vetorial complexo. Além disso,* $\|\cdot\|_p$ *é uma norma em* $L^p(\mu)$ *e essa norma é completa.*

Teorema 1.11. *(Desigualdade de Minkowski) Sejam* $f, g \in L^p(\mu)$*. Então:*

$$\left(\int |f+g|^p \, d\mu \right)^{\frac{1}{p}} \leqslant \left(\int |f|^p \, d\mu \right)^{\frac{1}{p}} + \left(\int |g|^p \, d\mu \right)^{\frac{1}{p}}.$$

1.4.2 Produto interno em $L^2(\mu)$

O caso $p = 2$ é um caso especial pois a norma $\|\cdot\|_2$ definida anteriormente provém de um produto interno, a saber:

$$\langle f, g \rangle = \int f\overline{g} d\mu.$$

Segue das propriedades de integral que esta expressão realmente define um produto interno em $L^2(\mu)$. Este produto se relaciona com a norma $\|\cdot\|_2$ por:

$$\|f\|_2 = (\langle f, f \rangle)^{1/2}.$$

Teorema 1.12. *(Desigualdade de Cauchy-Schwarz) Dadas $f, g \in L^2(\mu)$, então $f\overline{g} \in L^1(\mu)$ e vale a desigualdade:*

$$\left| \int f\overline{g} d\mu \right| \leqslant \int |f\overline{g}| d\mu \leqslant \left(\int |f|^2 d\mu \right)^{1/2} \left(\int |g|^2 d\mu \right)^{1/2}.$$

Teorema 1.13. *(Desigualdade de Hölder). Dado $1 < p < \infty$ considere q definido pela relação $\frac{1}{p} + \frac{1}{q} = 1$. Então, para toda $f \in L^p(\mu)$ e $g \in L^q(\mu)$ temos que $f\overline{g} \in L^1(\mu)$ e vale a desigualdade:*

$$\int |f\overline{g}| d\mu \leqslant \left(\int |f|^p d\mu \right)^{1/p} \left(\int |g|^q d\mu \right)^{1/q}.$$

1.5 Espaços de Hilbert

Um espaço de Hilbert é um espaço vetorial munido de produto interno cuja norma $\|\cdot\|$ proveniente deste produto interno é completa, ou seja, relativamente a $\|\cdot\|$ toda sequência de Cauchy é convergente. Como vimos anteriormente $L^2(\mu)$ é um espaço de Hilbert.

1.5.1 Operador de Koopman

Sejam $(M, \mathcal{B})$ um espaço mensurável, $T : M \to M$ uma transformação mensurável e μ uma medida invariante por T. O *operador de Koopman* é o operador linear $U_T : L^1(\mu) \to L^1(\mu)$, definido por $U_T(\phi) = \phi \circ T$.

Note que U_T está bem definido e é uma isometria, isto é, ele preserva a norma do espaço $L^1(\mu)$. De fato,

$$\|U_T(\phi)\|_1 = \int |U_T(\phi)| \, d\mu = \int |\phi| \circ T \, d\mu = \int |\phi| \, d\mu = \|\phi\|_1,$$

a terceira igualdade é verdadeira uma vez que μ é invariante. Além disso, U_T é um operador linear positivo, $U_T \geqslant 0$ em μ-quase todo ponto sempre que $\phi \geqslant 0$ em μ-quase todo ponto. Disto temos a seguinte Proposição.

Proposição 1.6. *O operador* $U_T : L^1(\mu) \to L^1(\mu)$ *induzido por* T *é linear, positivo e uma isometria.*

1.5.2 Isometrias em espaços de Hilbert

Sejam H um espaço de Hilbert e F um subespaço fechado de H. Então

$$H = F \oplus F^{\perp},$$

onde $F^{\perp} = \{w \in H; \langle v, w \rangle = 0, \forall v \in F\}$ é o complementar ortogonal de F. A projeção $P_F : H \to F$ associada à decomposição $H = F \oplus F^{\perp}$ é chamada projeção ortogonal sobre F. Ela está unicamente caracterizada por

$$\|x - P_F(x)\| = \min\{\|x - v\|; v \in F\}.$$

Observe que $P_F(v) = v$ para todo $v \in F$, e por consequência, $P_F^2 = P_F$.

O operador adjunto $U^* : H \to H$ de um operador linear contínuo $U : H \to H$ está definido pela relação

$$\langle U^* u, v \rangle = \langle u, Uv \rangle \text{ para todo } u, v \in H.$$

O operador U diz-se uma isometria se ele preserva produto interno, ou seja,

$$\langle Uu, Uv \rangle = \langle u, v \rangle \text{ para todo } u, v \in H.$$

Isso é equivalente a dizer que U preserva a norma de H. Outra condição equivalente é $U^*U = id$. De fato,

$$\langle Uu, Uv \rangle = \langle u, v \rangle \quad \forall u, v \in H \qquad \Leftrightarrow \qquad \langle U^*Uu, v \rangle = \langle u, v \rangle \quad \forall u, v \in H.$$

Definição 1.23. *O conjunto dos vetores invariantes de um operador linear contínuo* $U : H \to H$ *é o subespaço vetorial definido por,*

$$I(U) = \{v \in H : Uv = v\}.$$

Observe que $I(U)$ é um subespaço vetorial fechado, uma vez que U é contínuo. Quando U é uma isometria, temos o seguinte lema :

Lema 1.3. *Se* $U : H \to H$ *é uma isometria, então* $Uv = v$ *se, e somente se,* $U^*v = v$.

Demonstração: Como $U^*Uv = v$, supondo $Uv = v$ temos,

$$\langle U^*v, u\rangle = \langle U^*Uv, u\rangle = \langle v, u\rangle \text{ , para todo } u \in H,$$

assim $U^*v = v$. Agora suponha que $U^*v = v$. Então $\langle Uv, v\rangle = \langle v, U^*v\rangle = \langle v, v\rangle = \|v\|^2$. Logo, usando que U preserva a norma de H,

$$\|Uv - v\|^2 = \langle Uv - v, Uv - v\rangle = \|Uv\|^2 - \langle Uv, v\rangle - \langle v, Uv\rangle + \|v\|^2 = 0.$$

Portanto, $Uv = v$. ∎

Capítulo 2

Teoremas ergódicos

Neste capítulo apresentamos os principais resultados desta dissertação.

2.1 Teorema ergódico de von Neumann

John von Neumann (1903-1957) foi uma matemático húngaro, naturalizado estadunidense, que obteve contribuições em áreas como teoria dos conjuntos, análise funcional, teoria ergódica, mecânica quântica e várias outras áreas de estudo importantes. Von Neumann publicou em 1932 um paper com o resultado que ficou conhecido como teorema ergódico médio. Este resultado foi um dos primeiros a descrever o comportamento geométrico de transformações que preservam medidas. A demonstração do teorema de von Neumann, que apresentamos aqui, pode ser encontrada em [18].

Teorema 2.1. *(von Neumann) Seja* $U : H \to H$ *uma isometria num espaço de Hilbert* H*, e seja* P *a projeção ortogonal sobre o subespaço* $I(U)$ *dos vetores invariantes por* U*. Então,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} U^j v = Pv\,, \quad \textit{para todo } v \in H.$$

Demonstração: Seja $L(U)$ o conjunto dos vetores $w \in H$ da forma $w = Uu - u$ para algum $u \in H$ e seja $\overline{L(U)}$ o seu fecho. Afirmamos que

$$I\,(U) = \overline{L(U)}^{\perp}.$$

Começaremos mostrando que $I(U) \subset \overline{L(U)}^{\perp}$. Primeiramente, considere vetores $v \in I(U)$ e $w \in L(U)$. Pelo Lema 1.3 temos que $v \in I(U^*)$, ou seja, $U^*v = v$ e $w = Uu - u$, para algum $u \in H$. Note que

$$\begin{aligned}\langle v, w\rangle &= \langle v, Uu - u\rangle \\ &= \langle v, Uu\rangle - \langle v, u\rangle \\ &= \langle U^* v, u\rangle - \langle v, u\rangle \\ &= 0.\end{aligned}$$

Logo $\langle v, w\rangle = 0$, para qualquer $w \in L(U)$. Agora, considere $w \in \overline{L(U)}$, existe uma sequência (w_k) sequência em $L(U)$ tal que $w_k \to w$. Daí,

$$\langle v, w\rangle = \langle v, \lim_{k\to\infty} w_k\rangle = \lim_{k\to\infty} \langle v, w_k\rangle = 0.$$

Logo $v \in \overline{L(U)}^{\perp}$ e provamos a primeira inclusão.

Agora, iremos mostrar que $\overline{L(U)}^{\perp} \subset I(U)$. Dado $v \in \overline{L(U)}^{\perp}$, temos que $\langle v, w\rangle = 0$, para todo $w \in L(U)$. Escrevendo $w = Uu - u$ temos que $\langle v, Uu\rangle = \langle v, u\rangle$. Daí,

$$\langle U^* v, u\rangle = \langle v, u\rangle.$$

A igualdade acima vale para todo $u \in H$ assim $U^* v = v$. Do Lema 1.3 segue que $v \in I(U)$. Assim temos que $\overline{L(U)}^{\perp} = I(U)$. Em particular, podemos escrever

$$H = I(U) \oplus \overline{L(U)}.$$

Agora provaremos o resultado nos casos particulares onde $v \in I(U)$ ou $v \in \overline{L(U)}$. Suponha que $v \in I(U)$. Observe que $Pv = v$. Por outro lado,

$$\frac{1}{n}\sum_{j=0}^{n-1} U^j v = \frac{1}{n}\sum_{j=0}^{n-1} v = v,$$

para todo n. Logo esta sequência converge para v quando $n \to \infty$. E isto prova o resultado neste caso. Em seguida suponha que $v \in L(U)$. Então, por definição, existe $u \in H$ tal que $v = Uu - u$. Daí,

$$\frac{1}{n}\sum_{j=0}^{n-1} U^j v = \frac{1}{n}\sum_{j=0}^{n-1} (U^{j+1}u - U^j u) = \frac{1}{n}(U^n u - u).$$

Por outro lado como U é isometria, temos que

$$\|U^n u - u\| \leqslant \|U^n u\| + \|u\| = 2\|u\|.$$

Isto mostra que,

$$\left\|\frac{1}{n}\sum_{j=0}^{n-1}U^j v\right\| \leqslant \frac{2}{n}\|u\|.$$

Disto segue que

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}U^j v = 0, \text{ para todo } v \in L(U).$$

Mais em geral, suponha que $v \in \overline{L(U)}$. Então existem $v_k \in L(U)$ convergindo para v quando $k \to \infty$.

Como U é isometria, observe que

$$\begin{aligned}
\left\|\frac{1}{n}\sum_{j=0}^{n-1}U^j v - \frac{1}{n}\sum_{j=0}^{n-1}U^j v_k\right\| \quad &= \left\|\frac{1}{n}\sum_{j=0}^{n-1}U^j(v-v_k)\right\| \\
&\leqslant \frac{1}{n}\sum_{j=0}^{n-1}\|U^j(v-v_k)\| \\
&\leqslant \|v-v_k\|.
\end{aligned}$$

Dado $\varepsilon > 0$ existe $k_0 \in \mathbb{N}$ tal que $\|v - v_{k_0}\| < \varepsilon/2$. Por outro lado, $v_{k_0} \in L(U)$ assim existe n_0 tal que para todo $n > n_0$ temos

$$\left\|\frac{1}{n}\sum_{j=0}^{n-1}U^j v_{k_0}\right\| < \frac{\varepsilon}{2}.$$

Para $n > n_0$, observe que

$$\begin{aligned}
\left\|\frac{1}{n}\sum_{j=0}^{n-1}U^j v\right\| \quad &= \left\|\frac{1}{n}\sum_{j=0}^{n-1}U^j v - \frac{1}{n}\sum_{j=0}^{n-1}U^j v_{k_0} + \frac{1}{n}\sum_{j=0}^{n-1}U^j v_{k_0}\right\| \\
&\leqslant \left\|\frac{1}{n}\sum_{j=0}^{n-1}U^j v - \frac{1}{n}\sum_{j=0}^{n-1}U^j v_{k_0}\right\| + \left\|\frac{1}{n}\sum_{j=0}^{n-1}U^j v_{k_0}\right\| \\
&\leqslant \|v-v_{k_0}\| + \left\|\frac{1}{n}\sum_{j=0}^{n-1}U^j v_{k_0}\right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Portanto,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}U^j v = 0, \text{ para todo } v \in \overline{L(U)}.$$

Para o caso geral onde v é um vetor qualquer de H podemos escrever $v = Pv + w$, onde $Pv \in I(U)$ e $w \in \overline{L(U)}$ uma vez que $H = I(U) \oplus \overline{L(U)}$. Daí,

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} U^j v &= \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} U^j(Pv + w) \\ &= \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} U^j Pv + \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} U^j w \\ &= Pv.\end{aligned}$$

Concluindo assim a prova do resultado. ■

2.1.1 Teorema ergódico de von Neumann em $L^2(\mu)$

Sejam $(X, \mathcal{B}, \mu)$ um espaço de probabilidade e $T : X \to X$ uma transformação mensurável que preserva μ em X. Uma função mensurável $f : X \to \mathbb{R}$ é dita invariante se $f \circ T = f$ em μ-quase todo ponto. O seguinte resultado é um caso particular do teorema ergódico de von Neumann.

Teorema 2.2. *Para qualquer $f \in L^2(\mu)$, seja $\widetilde{f}$ a projeção ortogonal de f no subespaço das funções invariantes. Então a sequência*

$$\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j$$

converge para $\widetilde{f}$ no espaço $L^2(\mu)$. Se T é invertivel, então a sequência

$$\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^{-j}$$

também converge para $\widetilde{f}$ em $L^2(\mu)$.

Demonstração: Dada uma transformação $T : X \to X$ que preserva uma medida finita μ. Seja $U = U_T : L^2(\mu) \to L^2(\mu)$ o operador de Koopman. Note que $g \in I(U)$ se e somente se, $g \circ T = g$ em μ-quase todo ponto. Seja $\widetilde{f}$ a projeção de f em $I(U)$, Pelo Teorema 2.1, temos que

$$\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j \longrightarrow \widetilde{f}$$

em $L^2(\mu)$, e isso prova a primeira afirmação.

Agora, suponha que T é invertível, considere $\tilde{U} = U_{T^{-1}}$, assim temos que $\tilde{U}(g) = g \circ T^{-1} = U_T^{-1}$. Assim, analogamente ao que foi feito, a segunda sequência converge para a projeção ortogonal de f no espaço $I(U_T^{-1})$. Veja que se $g \in I(U_T^{-1})$ então $g \circ T^{-1} = g$ em μ-q.t.p. Daí, $g = g \circ T$, ou seja, $g \in I(U_T)$. De modo análogo, $I(U_T) \subset I(U_T^{-1})$ logo $I(U_T^{-1}) = I(U_T)$. O limite da sequência

$$\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^{-j}$$

em $L^2(\mu)$ é a mesma função $\widetilde{f}$ que obtivemos antes. ■

2.2 Teorema ergódico de Birkhoff

George David Birkhoff (1884-1944) foi um matemático estadunidense, que teve trabalhos importantes em teoria algébrica dos grafos, sistemas dinâmicos, mecânica quântica. Uma de suas maiores contribuições foi o seu teorema ergódico publicado em 1931. Demonstraremos o teorema ergódico de Birkhoff seguindo a prova que pode ser encontrada em [2]. Apresentaremos agora a noção de expectativa condicional de uma função que será útil na demonstração.

Definição 2.1. *Sejam $\mathcal{A}$ e $\mathcal{B}$ duas σ-álgebras de um conjunto X. Diremos que $\mathcal{B}$ é uma σ-subálgebra de $\mathcal{A}$ se $\mathcal{B} \subset \mathcal{A}$.*

Proposição 2.1. *Seja $(X, \mathcal{A}, \mu)$ um espaço de medida finita e seja $\mathcal{F} \subset \mathcal{A}$ uma σ-subálgebra. Para toda função $\mathcal{A}$-mensurável $f \in L^1(\mu)$, sempre existe uma função $\mathcal{F}$-mensurável $f_{\mathcal{F}} \in L^1(\mu)$ tal que*

$$\int_A f_{\mathcal{F}}\, d\mu = \int_A f\, d\mu \textit{ para todo } A \in \mathcal{F}.$$

Demonstração: Defina a medida finita ν em $\mathcal{F}$ por

$$\nu(A) = \int_A f\, d\mu \text{ para todo } A \in \mathcal{F}.$$

Note que se $\mu(A) = 0$ para $A \in \mathcal{F}$ então $\nu(A) = 0$. Em outras palavras, ν é absolutamente contínua com respeito a restrição de μ à σ-álgebra $\mathcal{F}$. Portanto, pelo Teorema de Radon-Nikodym, existe uma função $\mathcal{F}$- mensurável $f_{\mathcal{F}} \in L^1(\mu)$ tal que

$$\nu(A) = \int_A f_{\mathcal{F}}\, d\mu \text{ para todo } A \in \mathcal{F}$$

e obtemos o resultado. ■

Definição 2.2. *Toda função* $f_{\mathcal{F}}$ *como na Proposição anterior é chamada de esperança condicional de* f *com respeito à* $\mathcal{F}$.

Exemplo 2.1. *Sejam* $(X, \mathcal{A}, \mu)$ *um espaço de medida finita e* $T : X \to X$ *uma transformação* $\mathcal{A}$*-mensurável. Considere a* σ*-subálgebra de conjuntos* T*-invariantes:*

$$\mathcal{F} = \{A \in \mathcal{A}; T^{-1}A = A\}.$$

Seja $f : X \to \mathbb{R}$ *uma função* $\mathcal{A}$*-mensurável. Vamos mostrar que* f *é* $\mathcal{F}$*-mensurável se e somente se os conjuntos* $f^{-1}\{\alpha\}$ *são* T*-invariantes para todo* $\alpha \in \mathbb{R}$. *De fato,* f *é* $\mathcal{F}$*-mensurável se e somente se* $f^{-1}B$ *é* T*-invariante para todo conjunto mensurável* $B \subset \mathbb{R}$, *ou seja, se e somente se,*

$$f^{-1}B = T^{-1}(f^{-1}B) = (f \circ T)^{-1}B, \tag{2.1}$$

para todo conjunto mensurável $B \subset \mathbb{R}$. *Como* $\{\alpha\} \subset \mathbb{R}$ *é mensurável para cada* $\alpha \in \mathbb{R}$, *segue de (2.1) que*

$$(f \circ T)^{-1}\alpha = f^{-1}\alpha \tag{2.2}$$

para todo $\alpha \in \mathbb{R}$. *Agora, assumindo que (2.2) é válido para todo* $\alpha \in \mathbb{R}$, *temos*

$$\begin{aligned} (f \circ T)^{-1}B \quad &= (f \circ T)^{-1} \bigcup_{\alpha \in B} \{\alpha\} \\ &= \bigcup_{\alpha \in B} (f \circ T)^{-1}\alpha \\ &= \bigcup_{\alpha \in B} f^{-1}\alpha \\ &= f^{-1} \bigcup_{\alpha \in B} \{\alpha\} = f^{-1}B, \end{aligned}$$

para todo conjunto $B \in \mathbb{R}$. *Isto implica que* f *é* $\mathcal{F}$*-mensurável. Em particular, a expectativa condicional* $f_{\mathcal{F}}$ *de uma função* $\mathcal{A}$*-mensurável* $f \in L^1(\mu)$ *é uma função* T*-invariante.*

Teorema 2.3. *(Birkhoff) Sejam* $T : X \to X$ *uma transformação mensurável e* μ *uma probabilidade invariante por* T. *Dada qualquer função integrável* $f : X \to \mathbb{R}$, *o limite*

$$\tilde{f}(x) = \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

existe em μ*-q.t.p.* $x \in X$. *Além disso, a função* $\tilde{f}$ *é integrável e*

$$\int_X \tilde{f}\, d\mu = \int_X f\, d\mu.$$

Demonstração: Seja $\psi : X \to \mathbb{R}$ uma função integrável. Para cada $n \in \mathbb{N}$, considere a função

$$\psi_n = \max\left\{\sum_{j=0}^{l-1} \psi \circ T^j ; 1 \leqslant l \leqslant n\right\}.$$

Claramente, $\psi_{n+1} \geqslant \psi_n$ para todo n. Como

$$|\psi_n| \leqslant \sum_{j=0}^{n-1} |\psi \circ T^j|$$

e μ é T-invariante, segue da Proposição 1.5 que

$$\begin{aligned} \int_X |\psi_n| \, d\mu & \leqslant \sum_{j=0}^{n-1} \int_X |\psi \circ T^j| \, d\mu \\ & = n \int_X |\psi| \, d\mu < \infty. \end{aligned}$$

Portanto, (ψ_n) é uma sequência não-decrescente de funções μ-integráveis. Além disso,

$$\begin{aligned} \psi_{n+1} & = \max\left\{\sum_{j=0}^{l-1} \psi \circ T^j ; 1 \leqslant l \leqslant n+1\right\} \\ & = \psi + \max\left\{0, \sum_{j=1}^{l} \psi \circ T^j ; 1 \leqslant l \leqslant n\right\} \qquad (2.3) \\ & = \psi + \max\{0, \psi_n \circ T\}. \end{aligned}$$

Observe que $\lim\limits_{n\to\infty} \psi_{n+1}(x) = +\infty$, se e somente se , $\lim\limits_{n\to\infty} \psi_n(T(x)) = +\infty$. Portanto, o conjunto

$$A = \left\{x \in X; \lim_{n\to\infty} \psi_n(x) = +\infty\right\}$$

é T-invariante. Assim, segue de (2.3) que

$$\begin{aligned} \psi_{n+1} - \psi_n \circ T & = \psi + \max\{0, \psi_n \circ T\} - \psi_n \circ T \\ & = \psi - \min\{0, \psi_n \circ T\}, \qquad (2.4) \end{aligned}$$

logo $\psi_{n+1} - \psi_n \circ T \searrow \psi$ em A, quando $n \to \infty$. Além disso, de (2.4) temos que

$$0 \leqslant \psi_{n+1} - \psi_n \circ T - \psi \leqslant \psi_2 - \psi_1 \circ T - \psi$$

para todo $n \in \mathbb{N}$. Como $\psi_2 - \psi_1 \circ T - \psi$ é μ-integrável, segue do teorema da convergência dominada que

$$\int_A (\psi_{n+1} - \psi_n \circ T) \, d\mu \to \int_A \psi \, d\mu,$$

quando $n \to \infty$. Por outro lado, da Proposição 1.5 temos que

$$\begin{aligned} 0 &\leqslant \int_A (\psi_{n+1} - \psi_n)\, d\mu \\ &= \int_A (\psi_{n+1} - \psi_n \circ T)\, d\mu. \end{aligned}$$

Portanto,

$$\int_A \psi \, d\mu \geqslant 0. \tag{2.5}$$

Agora, considere a σ-subálgebra $\mathcal{F}$ dos conjuntos T-invariantes. Dado $\varepsilon > 0$, considere a função

$$\psi = f - f_{\mathcal{F}} - \varepsilon$$

onde $f_{\mathcal{F}}$ é a esperança condicional de f. Da Proposição 2.1, como A é T-invariante, ou seja, $A \in \mathcal{F}$ temos que

$$\int_A \psi \, d\mu = \int_A (f - f_{\mathcal{F}} - \varepsilon)\, d\mu = -\varepsilon \cdot \mu(A).$$

De (2.5) podemos concluir que $\mu(A) = 0$. Então,

$$\lim_{n\to\infty} \psi_n(x) < +\infty$$

para μ-quase todo ponto $x \in X$. Note que

$$\frac{1}{n}\sum_{j=0}^{n-1} \psi \circ T^j \leqslant \frac{\psi_n}{n}.$$

Daí,

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \psi \circ T^j \leqslant 0$$

para μ-quase todo ponto em X. Por outro lado, $f_{\mathcal{F}}$ é T-invariante assim

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j \leqslant \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f_{\mathcal{F}} \circ T^j + \varepsilon = f_{\mathcal{F}} + \varepsilon \tag{2.6}$$

em μ-quase todo ponto em X. Agora, note que $(-f)_{\mathcal{F}} = -f_{\mathcal{F}}$. Portanto, substituindo f por $-f$ em (2.6), temos

$$\begin{aligned} -\liminf_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j &= \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} (-f) \circ T^j \\ &\leqslant -f_{\mathcal{F}} + \varepsilon. \end{aligned}$$

Logo,

$$f_{\mathcal{F}} - \varepsilon \leqslant \liminf_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j \tag{2.7}$$

em μ-quase todo ponto. Por (2.6) e (2.7) existe um conjunto $X_\varepsilon \subset X$ com medida total, ou seja, $\mu(X_\varepsilon) = \mu(X)$ tal que em X_ε vale

$$f_{\mathcal{F}} - \varepsilon \leqslant \liminf_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j \leqslant \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j \leqslant f_{\mathcal{F}} + \varepsilon.$$

Como ε é arbitrário, tomando $\varepsilon = \frac{1}{k}$, com $k \in \mathbb{N}$ existe um conjunto $X_{1/k} \subset X$, com medida total tal que em $X_{1/k}$ vale

$$f_{\mathcal{F}} - \frac{1}{k} \leqslant \liminf_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j \leqslant \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j \leqslant f_{\mathcal{F}} + \frac{1}{k}.$$

Uma vez que $(X_{1/k})^c$ tem medida nula, e a união enumerável de conjuntos de medida nula tem medida nula segue que

$$D = \bigcap_{k=1}^{\infty} X_{1/k} = \left(\bigcup_{k=1}^{\infty} (X_{1/k})^C\right)^c$$

tem medida total. Assim para todo $x \in D$ temos que

$$\tilde{f}(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j(x) = f_{\mathcal{F}}(x). \tag{2.8}$$

Como $\mu(D) = \mu(X)$ e o número $f_{\mathcal{F}}(x)$ está bem definido para μ-quase todo ponto $x \in X$ temos que $\widetilde{f} = f_{\mathcal{F}}$ em μ-quase todo ponto. Por outro lado, $f_{\mathcal{F}} \in L^1(\mu)$ assim $\widetilde{f} \in L^1(\mu)$. Da Proposição 2.1 temos que

$$\int_X \widetilde{f}\, d\mu = \int_X f_{\mathcal{F}}\, d\mu = \int_X f\, d\mu,$$

o que completa a prova do resultado. ■

Corolário 2.1. *Suponha que* M *é um espaço métrico compacto e* $T : M \to M$ *é uma aplicação mensurável. Então existe um conjunto mensurável* $G \subset M$ *com* $\mu(G) = 1$ *tal que*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j(x) = \widetilde{f}(x), \tag{2.9}$$

para todo $x \in G$ *e toda função contínua* $f : M \to \mathbb{R}$.

Demonstração: Pelo Teorema Ergódico de Birkhoff, para cada função contínua f existe um conjunto $G(f) \subset M$ com $\mu(G(f)) = 1$ tal que (2.9) é válido para todo $x \in G(f)$. Considere $C^0(M)$ como sendo o espaço das funções contínuas $f : M \to \mathbb{R}$. Como M é compacto, temos que $C^0(M)$ admite algum subconjunto $\{f_k : k \in \mathbb{N}\}$ enumerável denso. Tomemos

$$G = \bigcap_{k=1}^{\infty} G(f_k).$$

É claro que $\mu(G) = 1$. Portanto basta provar que (2.9) vale para toda função contínua f sempre que $x \in G$. Isso pode ser feito da seguinte maneira. Dado $f \in C^0(M)$ e qualquer $\varepsilon > 0$, tomemos $k \in \mathbb{N}$ tal que

$$\|f - f_k\| = \sup\{|f(x) - f_k(x)| : x \in M\} \leqslant \varepsilon.$$

Então, dado qualquer ponto $x \in G$,

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(x) \leqslant \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_k \circ T^j(x) + \varepsilon = \widetilde{f_k}(x) + \varepsilon,$$

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(x) \geqslant \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_k \circ T^j(x) - \varepsilon = \widetilde{f_k}(x) - \varepsilon.$$

Isto implica que

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(x) - \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(x) \leqslant 2\varepsilon.$$

Como ε é arbitrário, segue que o limite $\widetilde{f}(x)$ existe, conforme afirmado.

■

2.2.1 Relação entre o teorema ergódico de Birkhoff e o teorema ergódico de Von Neumann

Nesta seção mostraremos que o teorema ergódico de Von Neumann pode ser obtido como uma consequência do teorema ergódico de Birkhoff. Uma importante vantagem do Teorema ergódico de Birkhoff é a sua formulação em termos de convergência em μ-quase todo ponto, o que neste contexto é uma propriedade mais forte do que a convergência em $L^2(\mu)$.

Considere uma função $f \in L^2(\mu)$, observe que $f \in L^1(\mu)$ pois X é um espaço de medida finito. Seja $\widetilde{f}$ sua média temporal. Primeiramente, vamos mostrar que $\widetilde{f} \in L^2(\mu)$ e que sua norma satisfaz $\|\widetilde{f}\|_2 \leqslant \|f\|_2$. De fato, temos que

$$|\widetilde{f}| \leqslant \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} |f \circ T^j|.$$

Daí,

$$|\widetilde{f}|^2 \leqslant \lim_{n\to\infty} \left(\frac{1}{n}\sum_{j=0}^{n-1} |f \circ T^j|\right)^2.$$

Então, pelo lema de Fatou, temos que

$$\left[\int |\widetilde{f}|^2\, d\mu\right]^{\frac{1}{2}} \leqslant \liminf_{n\to\infty} \left[\int \left(\frac{1}{n}\sum_{j=0}^{n-1} |f \circ T^j|\right)^2 d\mu\right]^{\frac{1}{2}}.$$

Usando a desiguladade de Minkowski, podemos majorar a desigualdade do lado direito. Assim,

$$\left[\int \left(\frac{1}{n}\sum_{j=0}^{n-1} |f \circ T^j|\right)^2 d\mu\right]^{\frac{1}{2}} \leqslant \frac{1}{n}\sum_{j=0}^{n-1} \left[\int |f \circ T^j|^2\, d\mu\right]^{\frac{1}{2}}.$$

Como μ é invariante por T, da Proposição 1.5 segue que

$$\frac{1}{n}\sum_{j=0}^{n-1} \left[\int |f \circ T^j|^2\, d\mu\right]^{\frac{1}{2}} = \left[\int |f|^2\, d\mu\right]^{\frac{1}{2}}.$$

Portanto,

$$\|\widetilde{f}\|_2 = \left[\int |\widetilde{f}|^2\, d\mu\right]^{\frac{1}{2}} \leqslant \left[\int |f|^2\, d\mu\right]^{\frac{1}{2}} = \|f\|_2.$$

Uma vez que $f \in L^2(\mu)$, temos $\|\widetilde{f}\|_2 \leqslant \|f\|_2 < \infty$ assim $\widetilde{f} \in L^2(\mu)$. Agora, resta mostrar que a sequência

$$\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j$$

converge para $\widetilde{f}$ em $L^2(\mu)$. Vamos começar supondo que a função f é limitada, isto é, existe $C > 0$ tal que $|f| \leqslant C$, então $|\widetilde{f}| \leqslant C$ e

$$\left|\frac{1}{n}\sum_{j=0}^{n-1} f \circ T^j\right| \leqslant C, \text{ para todo } n.$$

Segue do teorema da convergência dominada que

$$\begin{aligned}
\lim_{n\to\infty}\left\|\frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j - \widetilde{f}\right\|_2^2 &= \lim_{n\to\infty}\int\left(\frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j - \widetilde{f}\right)^2 d\mu\\
&= \int\left(\lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j - \widetilde{f}\right)^2 d\mu\\
&= 0.
\end{aligned}$$

Portanto,

$$\frac{1}{n}\sum_{j=0}^{n-1} f\circ T^j \longrightarrow \widetilde{f}$$

em $L^2(\mu)$.

Agora, vamos retirar a condição de limitação de f. Defina, para cada $k\in\mathbb{N}$,

$$f_k(x)=\begin{cases} f(x) \text{ , se } |(f(x)|\leqslant k\\ 0\text{, caso contrário}\end{cases}.$$

Claramente, a sequência (f_k) converge para f em $L^2(\mu)$, e para cada k, f_k é limitada. Seja $\widetilde{f}_k$ a média temporal de f_k. Dado $\varepsilon>0$, fixe k_0 tal que

$$\|f-f_{k_0}\|_2<\frac{\varepsilon}{3}.$$

Da Proposição 1.5 temos que

$$\begin{aligned}
\|(f-f_{k_0})\circ T^j\|_2 &= \left(\int|(f-f_{k_0})|\circ T^j|^2 d\mu\right)^{\frac{1}{2}}\\
&= \left(\int|f-f_{k_0}|^2 d\mu\right)^{\frac{1}{2}}\\
&= \|f-f_{k_0}\|_2,
\end{aligned}$$

para todo $j\geqslant 0$. Logo, para todo $n\geqslant 1$ temos

$$\begin{aligned}
\left\|\frac{1}{n}\sum_{j=0}^{n-1}(f-f_{k_0})\circ T^j\right\|_2 &\leqslant \frac{1}{n}\sum_{j=0}^{n-1}\|(f-f_{k_0})\circ T^j\|_2\\
&\leqslant \|f-f_{k_0}\|_2\\
&< \frac{\varepsilon}{3}.
\end{aligned}$$

Note também que $\widetilde{f}-\widetilde{f_{k_0}}$ é a média temporal de $f-f_{k_0}$. Como $f-f_{k_0}\in L^2(\mu)$ segue da observação feita no início da demonstração que $\|\widetilde{f}-\widetilde{f_{k_0}}\|_2\leqslant\|f-f_{k_0}\|_2$. Por outro lado, f_{k_0} é limitada assim existe $n_0\in\mathbb{N}$ tal que se $n>n_0$ então

$$\left\| \frac{1}{n} \sum_{j=0}^{n-1} f_{k_0} \circ T^j - \widetilde{f}_{k_0} \right\|_2 < \frac{\varepsilon}{3}.$$

Assim, para todo $n > n_0$ temos

$$\begin{aligned}
\left\| \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j - \widetilde{f} \right\|_2 &= \left\| \frac{1}{n} \sum_{j=0}^{n-1} (f - f_{k_0}) \circ T^j + \frac{1}{n} \sum_{j=0}^{n-1} f_{k_0} \circ T^j - \widetilde{f}_{k_0} + \widetilde{f}_{k_0} - \widetilde{f} \right\|_2 \\
&\leqslant \left\| \frac{1}{n} \sum_{j=0}^{n-1} (f - f_{k_0}) \circ T^j \right\|_2 + \left\| \frac{1}{n} \sum_{j=0}^{n-1} f_{k_0} \circ T^j - \widetilde{f}_{k_0} \right\|_2 + \left\| \widetilde{f}_{k_0} - \widetilde{f} \right\|_2 \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

Portanto a sequência

$$\frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j$$

converge para $\widetilde{f}$ em $L^2(\mu)$. Isto completa a prova do teorema ergódico de Von Neumann via teorema ergódico de Birkhoff. Como consequência, se T é invertível, temos

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j = \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^{-j}.$$

2.3 Teorema ergódico de Riesz

Neste tópico apresentaremos o teorema ergódico médio de Riesz que foi provado em [16] no ano de 1938. Frigyes Riesz (1880-1956) foi um matemático húngaro que teve um trabalho fundamental em análise funcional e seu trabalho teve várias aplicações importantes na física. Seu teorema ergódico é uma generalização do Teorema ergódico de von Neumann para espaços de Banach. A demonstração do teorema ergódico médio de Riesz que será feita neste tópico pode ser encontrada em [14].

Definição 2.3. *Um espaço normado X é dito uniformemente convexo se, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que, se x e y são vetores unitários em X,*

$$\|x - y\| \geqslant \varepsilon \Longrightarrow \left\| \frac{x + y}{2} \right\| \leqslant 1 - \delta.$$

Definição 2.4. *Seja X um espaço vetorial. Se $Z \subset X$, definimos por $\mathrm{co}(Z)$ a envoltória convexa de Z, ou seja, o conjunto de todas as combinações convexas finitas de elementos de Z. Denotaremos por $\overline{\mathrm{co}}(Z)$ o fecho da envoltória convexa de Z em X.*

Teorema 2.4. *(Riesz) Seja* X *um espaço de Banach uniformemente convexo. Considere* $T : X \to X$ *um operador linear tal que*

$$\|Tx\| \leqslant \|x\|, \text{ para todo } x \in X.$$

Então para todo $x \in X$*, o limite*

$$p_x = \lim_{n\to\infty} \frac{1}{n+1} \sum_{j=0}^{n} T^j x$$

existe. Além disso, o operador $P : X \to X$ *definido por* $Px = p_x$ *é a projeção contínua do vetor* x *no espaço* $M = \{y \in X : Ty = y\}$.

Demonstração: Fixe $x \in X$ e considere o conjunto

$$C = \overline{\text{co}}(\{x, Tx, T^2x, ..., T^n x, ...\}).$$

Observe que C é um conjunto convexo fechado e não-vazio, segue da convexidade uniforme de X que existe um único $p_x \in C$ tal que

$$\delta = \|p_x\| = \inf\{\|z\| : z \in C\}.$$

Tome $\varepsilon > 0$. Da definição de C existe $z = \sum_{j=0}^{m} \alpha_j T^j x$, onde $\alpha_0, \alpha_1, \cdots, \alpha_m$ são constantes não-negativas tais que $\sum_{j=0}^{m} \alpha_j = 1$ e $\|p_x - z\| < \varepsilon$. Por outro lado, para cada $n \in \mathbb{N}$,

$$\begin{aligned} \left\|\frac{z + Tz + ... + T^n z}{n+1}\right\| &\leqslant \frac{\|z\| + \|Tz\| + ... + \|T^n z\|}{n+1} \\ &\leqslant \|z\| \\ &\leqslant \|z - p_x\| + \|p_x\| \\ &< \delta + \varepsilon. \end{aligned}$$

Note que

$$\begin{aligned} z + Tz + \cdots + T^n z &= (\alpha_0 x + \cdots + \alpha_m T^m x) + (\alpha_0 Tx + \cdots + \alpha_m T^{m+1} x) \\ &\quad + \cdots + (\alpha_0 T^n x + \cdots + \alpha_m T^{m+n} x). \end{aligned}$$

Agora suponha que $n > m$. Como $(\alpha_0 + \alpha_1 + ... + \alpha_m) = 1$ podemos escrever

$$\begin{aligned}
z + Tz + \cdots + T^n z &= x + Tx + \cdots + T^{m-1}x + \\
&\quad +(\alpha_0 - 1)x + (\alpha_0 + \alpha_1 - 1)Tx + (\alpha_0 + ... + \alpha_{m-1} - 1)T^{m-1}x \\
&\quad +(\alpha_0 + \alpha_1 + ... + \alpha_m)T^m x + \cdots + (\alpha_0 + \alpha_1 + \cdots + \alpha_m)T^n x \\
&\quad +(\alpha_1 + \cdots + \alpha_m)T^{1+n}x + \cdots + (\alpha_m)T^{m+n}x \\
&= x + Tx + \cdots + T^n x + r,
\end{aligned}$$

onde

$$\begin{aligned}
r = \quad & (\alpha_0 - 1)x + (\alpha_0 + \alpha_1 - 1)Tx + (\alpha_0 + \cdots + \alpha_{m-1} - 1)T^{m-1}x \\
& +(\alpha_1 + \cdots + \alpha_m)T^{1+n}x + \cdots + (\alpha_m)T^{m+n}x.
\end{aligned}$$

Portanto,

$$\frac{1}{n+1}\sum_{j=0}^{n} T^j x = \frac{1}{n+1}\sum_{j=0}^{n} T^j z - \frac{r}{n+1}.$$

Veja que

$$\begin{aligned}
\|r\| \quad & \leqslant |\alpha_0 - 1|\|x\| + |\alpha_0 + \alpha_1 - 1|\|Tx\| + |\alpha_0 + \cdots + \alpha_{m-1} - 1|\|T^{m-1}x\| \\
& +|\alpha_1 + \cdots + \alpha_m|\|T^{1+n}x\| + \cdots + |\alpha_m|\|T^{m+n}x\| \\
& \leqslant \|x\| + \|Tx\| + \cdots + \|T^{m-1}x\| + \|T^{1+n}x\| + \cdots + \|T^{m+n}x\| \\
& \leqslant 2m\|x\|.
\end{aligned}$$

Daí,

$$\left\|\frac{r}{n+1}\right\| \leqslant \frac{2m\|x\|}{n+1}.$$

Tomando n suficientemente grande tal que $2m\|x\| < \varepsilon(n+1)$ temos que

$$\left\|\frac{1}{n+1}\sum_{j=0}^{n} T^j x\right\| \leqslant \left\|\frac{1}{n+1}\sum_{j=0}^{n} T^j z\right\| + \left\|\frac{r}{n+1}\right\| < \delta + 2\varepsilon.$$

Por outro lado, temos que

$$\delta \leqslant \left\|\frac{x}{n+1} + \frac{Tx}{n+1} + ... + \frac{T^n x}{n+1}\right\| = \left\|\frac{1}{n+1}\sum_{j=0}^{n} T^j x\right\|.$$

Daí,

$$\delta \leqslant \liminf_{n\to\infty}\left\|\frac{1}{n+1}\sum_{j=0}^{n} T^j x\right\| \leqslant \limsup_{n\to\infty}\left\|\frac{1}{n+1}\sum_{j=0}^{n} T^j x\right\| < \delta + 2\varepsilon,$$

como ε é arbitrário, concluimos que

$$\lim_{n\to\infty}\left\|\frac{1}{n+1}\sum_{j=0}^{n} T^j x\right\| = \delta.$$

Isso mostra que $\dfrac{1}{n+1}\displaystyle\sum_{j=0}^{n} T^j x$ é uma sequência minimizante em C. Da convexidade uniforme de X segue que

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{j=0}^{n} T^j x = p_x.$$

Agora, só resta mostrar que o operador $P(x) = p_x$ é uma projeção contínua em M. De fato, é claro que se $x \in M$ então $p_x = x$. Em geral,

$$\begin{aligned} Tp_x \quad &= T\left(\lim_{n\to\infty}\frac{1}{n+1}\sum_{j=0}^{n} T^j x\right) \\ &= \lim_{n\to\infty}\frac{1}{n+1}\sum_{j=1}^{n+1} T^j x \\ &= \lim_{n\to\infty}\frac{1}{n+1}\sum_{j=0}^{n} T^j x + \lim_{n\to\infty}\frac{T^{n+1}x - x}{n+1} = p_x, \end{aligned}$$

ou seja, $p_x \in M$, assim $p_{(p_x)} = p_x$. Portanto,

$$P^2 x = PPx = P(p_x) = p_x = Px.$$

Disto temos que $P : X \to X$ é uma projeção em M, a continuidade de P segue do fato de que

$$\|p_x\| \leqslant \|x\|.$$

■

2.4 Teorema ergódico maximal

O teorema ergódico maximal é um resultado que foi provado de maneira independente por Wiener e por Yosida e Kakutani no ano de 1939, veja [21]. Mais adiante veremos que o teorema ergódico de Birkhoff pode ser demonstrado utilizando o teorema ergódico maximal. Para auxiliar na demonstração do teorema ergódico maximal, provaremos o lema abaixo devido a A. Garsia que pode ser encontrado em [6].

Lema 2.1. *Seja* $U : L^1(\mu) \to L^1(\mu)$ *um operador linear positivo satisfazendo* $\|U\|_1 \leqslant 1$. *Dado* $N \in \mathbb{N}$ *e* $f \in L^1(\mu)$ *defina*

$$f_0 = 0, f_1 = f, f_n = \sum_{j=0}^{n-1} U^j f, \text{com } 1 \leqslant n \leqslant N \text{ e } \Psi_N = \max_{0 \leqslant i \leqslant N} f_i.$$

Então $\displaystyle\int_{A_N} f \, d\mu \geqslant 0$*, onde* $A_N = \{x \in M; \Psi_N(x) > 0\}$.

Demonstração: Observe que $|\Psi_N| \leqslant \sum_{i=0}^{N} |f_i|$. Em, particular $\Psi_N \in L^1(\mu)$. Como $f_0 = 0$, temos por definição que $\Psi_N \geqslant 0$. Uma vez que U é positivo temos que

$$U\Psi_N \geqslant U(0) = 0.$$

Dado $k \in \mathbb{N}$ tal que $1 \leqslant k \leqslant N - 1$, temos que

$$\begin{aligned} Uf_k &= U\big(\sum_{j=0}^{k-1} U^j f\big) \\ &= \sum_{j=0}^{k-1} U(U^j f) \\ &= \sum_{j=1}^{k} U^j f \\ &= f_{k+1} - f. \end{aligned}$$

Como $\Psi_N \geqslant f_k$, temos que $U\Psi_N \geqslant Uf_k = f_{k+1} - f$ assim $U\Psi_N + f \geqslant f_{k+1}$. Esta desigualdade vale para todo k entre 1 e $N - 1$. Daí,

$$U\Psi_N + f \geqslant \max\{f_1, f_2, ..., f_N\}, \forall x \in M. \tag{2.10}$$

Para $x \in A_N$, temos que $\Psi_N(x) > 0$. Daí,

$$U\Psi_N + f \geqslant \max\{f_1, ..., f_N\} > 0$$

em A_N. Como $\Psi_N \in L^1(\mu)$ e $U\Psi_N + f \in L^1(\mu)$. De (2.10) segue que

$$\begin{aligned} \int_{A_N} (U\Psi_N + f) \, d\mu &\geqslant \int_{A_N} \max\{f_1, ..., f_N\} \, d\mu \\ &= \int_{A_N} \Psi_N \, d\mu. \end{aligned}$$

Portanto,

$$\int_{A_N} f \, d\mu \geqslant \int_{A_N} \Psi_N d\mu - \int_{A_N} U\Psi_N \, d\mu.$$

Por outro lado, veja que $A_N^c = \{x \in M : \Psi_N(x) = 0\}$, pois $\Psi_N \geqslant 0$. Daí,

$$\begin{aligned}\int_M \Psi_N \, d\mu &= \int_{A_N} \Psi_N \, d\mu + \int_{A_N^c} \Psi_N \, d\mu \\ &= \int_{A_N} \Psi_N \, d\mu.\end{aligned}$$

Como $\|U\|_1 \leqslant 1$, $\Psi_N \geqslant 0$ e $U\Psi_N \geqslant 0$ temos que

$$\int_M U\Psi_N \, d\mu \leqslant \int_M \Psi_N \, d\mu.$$

Das desigualdades acima temos que

$$\begin{aligned}\int_{A_N} f \, d\mu &\geqslant \int_M \Psi_N \, d\mu - \int_{A_N} U\Psi_N \, d\mu \\ &\geqslant \int_M \Psi_N \, d\mu - \int_M U\Psi_N \, d\mu \\ &\geqslant 0.\end{aligned}$$

■

Teorema 2.5. *(Maximal) Seja* $T : M \to M$ *uma transformação mensurável e* μ *uma probabilidade invariante por* T. *Dada* $f : M \to \mathbb{R}$ *integrável, defina*

$$f^*(x) = \sup_{n \geqslant 1} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

e

$$D = \{x \in M; f^*(x) > 0\}.$$

Nestas condições, temos que $\displaystyle\int_D f\, d\mu \geqslant 0$.

Demonstração: Seja $U_T : L^1(\mu) \to L^1(\mu)$ o operador de Koopman de T. Da Proposição 1.6 temos que U_T é linear, isométrico e positivo. Em particular, $\|U_T\|_1 \leqslant 1$. Com a mesma notação do Lema 2.1, veja que se $x \in D$ então existe $n \geqslant 1$ tal que

$$\frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) = \frac{1}{n} \sum_{j=0}^{n-1} U_T^j(f(x)) > 0.$$

Logo $D = \displaystyle\bigcup_{n=1}^{\infty} A_n$. Perceba também que se $x \in A_n$ então $\Psi_{n+1}(x) \geqslant \Psi_n(x) > 0$ assim $x \in A_{n+1}$. Portanto $A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots$ e $A_j \nearrow D$. Daí,

$$\int_D f\, d\mu = \lim_{n \to \infty} \int_{A_n} f\, d\mu.$$

Agora, note que para todo $x \in M$, temos que $f(x) \cdot \mathcal{X}_{A_n}(x) \longrightarrow f(x) \cdot \mathcal{X}_D(x)$ quando $n \to \infty$. Como

$$|f(x) \cdot \mathcal{X}_{A_j}(x)| \leqslant |f(x)|, \forall x \in M$$

e $|f| \in L^1(\mu)$, segue do teorema da convergência dominada e do Lema 2.1 que

$$0 \leqslant \int_{A_j} f\, d\mu = \int_M f \cdot \mathcal{X}_{A_j}\, d\mu \longrightarrow \int_M f \cdot \mathcal{X}_D\, d\mu = \int_D f\, d\mu$$

quando $j \to \infty$. Portanto,

$$\int_D f\, d\mu \geqslant 0$$

como queríamos demonstrar. ∎

Corolário 2.2. *Seja* $T : X \to X$ *uma aplicação mensurável* μ*-invariante. Dada* $f \in L^1(\mu)$ *e* $\alpha \in \mathbb{R}$, *defina*

$$f^*(x) = \sup_{n} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

e

$$D_\alpha = \{x \in X; f^*(x) > \alpha\}$$

então temos que

$$\int_{D_\alpha} f d\mu \geqslant \alpha \cdot \mu(D_\alpha).$$

Demonstração: Defina $g(x) = f(x) - \alpha$, observe que

$$\begin{aligned} g^*(x) &= \sup_n \frac{1}{n} \sum_{j=0}^{n-1} g(T^j(x)) \\ &= \sup_n \frac{1}{n} \sum_{j=0}^{n-1} (f(T^j(x)) - \alpha) \\ &= f^*(x) - \alpha. \end{aligned}$$

Logo,

$$D_g = \{x \in X; g^*(x) > 0\} = \{x \in X; f^*(x) > \alpha\} = D_\alpha.$$

Como $g \in L^1(\mu)$, pelo teorema ergódico maximal temos

$$\begin{aligned} 0 &\leqslant \int_{D_g} g \, d\mu \\ &= \int_{D_\alpha} (f - \alpha) \, d\mu \\ &= \int_{D_\alpha} f \, d\mu - \alpha\mu(D_\alpha). \end{aligned}$$

Portanto,

$$\int_{D_\alpha} f d\mu \geqslant \alpha \cdot \mu(D_\alpha).$$

■

Corolário 2.3. *Seja* $T : X \to X$ *uma aplicação mensurável* μ*-invariante. Dada* $f \in L^1(\mu)$ *e* $\alpha \in \mathbb{R}$*, defina*

$$f_*(x) = \inf_n \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

e

$$C_\alpha = \{x \in X; f_*(x) < \alpha\}$$

então temos que

$$\int_{C_\alpha} f d\mu \leqslant \alpha \cdot \mu(C_\alpha).$$

Demonstração: Veja que $f_*(x) < \alpha$ se, e somente se, $g^*(x) > -\alpha$ onde $g = -f$. Em particular, $D_{-\alpha} = C_\alpha$. Pelo Corolário 2.2 temos que

$$\int_{D_{-\alpha}} g \, d\mu \geqslant -\alpha \cdot \mu(D_{-\alpha}).$$

Daí,

$$\int_{C_\alpha} f \, d\mu \leqslant \alpha \cdot \mu(C_\alpha).$$

■

2.4.1 Demonstração alternativa do teorema ergódico maximal.

A demonstração abaixo pode ser encontrada em [15].

Demonstração: Primeiramente, veja que o conjunto onde $f^* > 0$ pode ser escrito como a união disjunta dos conjuntos abaixos

$$B_1 = \{x \in M : f(x) > 0\},$$

$$B_2 = \{x \in M : f(x) \leqslant 0, f(x) + f(T(x)) > 0\},$$

$$\vdots$$

$$B_n = \left\{ x \in M : f(x) \leqslant 0, f(x) + f(T(x)) \leqslant 0, ..., \sum_{j=0}^{n-2} f(T^j(x)) \leqslant 0, \sum_{j=0}^{n-1} f(T^j(x)) > 0 \right\},$$

$$\vdots$$

Se mostrarmos que

$$\int_{B_1 \cup ... \cup B_n} f \, d\mu \geqslant 0 \text{ , para todo } n,$$

então segue do teorema da convergência dominada, aplicado a sequência $\mathcal{X}_{B_1 \cup ... \cup B_n} \cdot f$ que

$$\int_{[f^*(x) > 0]} f \, d\mu \geqslant 0.$$

Fixemos n, a ideia agora é quebrar $B_1 \cup ... \cup B_n$ em uma união disjunta de diferentes partes, na qual a integral de f sobre cada uma destas partes é não-negativa. Estas partes serão escritas como $B'_k \cup TB'_k \cup ... \cup T^{k-1}B'_k$, antes de construir estes conjuntos, faremos algumas observações:

1. $T^k B_n \subset B_1 \cup ... \cup B_{n-k}$ para $k = 1, ..., n-1$, isto acontece pois se $x \in B_n$, então

$$f(x) + f(T(x)) + ... + f(T^{k-1}(x)) \leqslant 0.$$

enquanto

$$f(x) + f(T(x)) + ... + f(T^{k-1}(x)) + f(T^k(x)) + ... + f(T^{n-1}(x)) > 0.$$

Então devemos ter que

$$f(T^k(x)) + ... + f(T^{n-1}(x)) > 0$$

isto é,

$$f(T^k(x)) + f(T(T^k(x)) + ... + f(T^{n-k-1}(T^k(x)) > 0,$$

ou seja, $T^k(x) \in B_1 \cup ... \cup B_{n-k}$

2. Os conjuntos $B_n, TB_n, ..., T^{n-1}B_n$ são dois a dois disjuntos. Isto acontece pois se $T^iB_n \cap T^jB_n \neq \emptyset$ para $0 \leqslant i < j \leqslant n-1$, então $B_n \cap T^{j-i}B_n \neq \emptyset$, contradizendo a observação acima.

3. Se tomarmos

$$B'_n = B_n \quad , \quad C_n = B_n \cup TB_n \cup ... \cup T^{n-1}B_n \quad ,$$

$$B'_{n-1} = B_{n-1} \setminus C_n \quad , \quad C_{n-1} = B'_{n-1} \cup TB'_{n-1} \cup ... \cup T^{n-2}B'_{n-1}$$

$$\vdots$$

$$B'_1 = B_1 \setminus (C_2 \cup ... \cup C_n) \quad , \quad C_1 = B'_1$$

então as colunas $C_1, ..., C_n$ são duas a duas disjuntas e os níveis $B'_k, TB'_k, ..., T^{k-1}B'_k$ dentro de cada coluna C_k também são dois a dois disjuntos. Assim a seguinte figura é uma representação correta de $B_1 \cup ... \cup B_n$:

$$\begin{array}{ccccc}
T^{n-1}B'_n & & & & \\
\vdots & T^{n-2}B'_{n-1} & & & \\
\vdots & \vdots & T^{n-3}B'_{n-2} & & \\
\vdots & \vdots & \vdots & & \\
TB'_n & TB'_{n-1} & TB'_{n-2} & & \\
B'_n = B_n & B'_{n-1} = B_{n-1} \setminus C_n & B'_{n-2} = B_{n-2} \setminus (C_n \cup C_{n-1}) & \dots & B'_1 \\
C_n & C_{n-1} & C_{n-2} & & C_1
\end{array}$$

A observação (3) decorre de

a) As partes estão em $B_1 \cup ... \cup B_n$, por (1).

b) Cada base é disjunta de todas as colunas à esquerda, por definição.

c) Cada base é disjunta das colunas à direita, por definição e por (1) (para os niveis mais altos, já que as imagens da base à direita estão contidas nos $B_j's$ à sua direita).

Assim, se duas colunas se cruzassem, poderiamos aplicar T^{-1} até que uma coluna cruzasse uma base ou outra coluna, e vimos que isso nao acontece, pois são disjuntas à esquerda, por (b).

Agora, é uma questão simples de fazer a estimativa

$$\begin{aligned}\int_{B_1\cup\ldots\cup B_n} f\,d\mu = \int_{C_1\cup\ldots\cup C_n} f\,d\mu &= \sum_{k=1}^{n}\int_{C_k} f\,d\mu\\ &= \sum_{k=1}^{n}\int_{B'_k\cup TB'_k\cup\ldots\cup T^{k-1}B'_k} f\,d\mu\\ &= \sum_{k=1}^{n}\int_{B'_k}(f+fT+\ldots+fT^{k-1})\,d\mu \geqslant 0\end{aligned}$$

pois por definição, $B'_k \subset B_k$, logo, $f + fT + \ldots + fT^{k-1} > 0$. Assim, para cada $n \in \mathbb{N}$,

$$\int_{B_1\cup\ldots\cup B_n} f d\mu \geqslant 0.$$

Isto conclui a prova do teorema. ∎

2.4.2 Relação entre o teorema ergódico maximal e o teorema ergódico de Birkhoff

Nesta seção mostraremos que o teorema ergódico de Birkhoff pode ser obtido como uma consequência do teorema ergódico maximal.

Dado $x \in X$, defina

$$f_1(x) = \liminf_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f(T^j(x))$$

e

$$f_2(x) = \limsup_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} f(T^j(x)).$$

Pra mostrar que $\widetilde{f}$ existe quase sempre, devemos mostrar que $f_1 = f_2$ em μ-q.t.p. $x \in X$. Por definição, temos que $f_1(x) \leqslant f_2(x)$, se $f_1(x) < f_2(x)$, então existem $\alpha, \beta \in \mathbb{Q}$ tais que $f_1(x) < \alpha < \beta < f_2(x)$.

Dados $\alpha, \beta \in \mathbb{Q}$, com $\alpha < \beta$, defina

$$E_{\alpha,\beta} = \{x \in X; f_1(x) < \alpha < \beta < f_2(x)\}.$$

Desta forma,

$$E = \bigcup_{\alpha<\beta} E_{\alpha,\beta},$$

com $\alpha, \beta \in \mathbb{Q}$ é o conjunto dos pontos $x \in X$ tal que o limite $\widetilde{f}(x)$ não existe. Como $\mathbb{Q} \times \mathbb{Q}$ é enumerável, basta mostrar que cada $E_{\alpha,\beta}$ tem medida nula para concluir que E tem medida nula e portanto $\widetilde{f}(x)$ existe em μ-quase todo ponto $x \in X$.

Primeiramente, note que $E_{\alpha,\beta}$ é invariante por T, ou seja, se $x \in E_{\alpha,\beta}$ então $T(x) \in E_{\alpha,\beta}$. De fato, veja que

$$\begin{aligned} f_1(T(x)) &= \liminf_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} f(T^j(x)) \\ &= \liminf_{n\to\infty} \frac{1}{n} \left(\sum_{j=0}^{n} f(T^j(x)) - f(x) \right) \\ &= \liminf_{n\to\infty} \frac{n+1}{n} \left(\frac{1}{n+1} \sum_{j=0}^{n} f(T^j(x)) - \frac{f(x)}{n+1} \right) \\ &= f_1(x), \end{aligned}$$

e o mesmo acontece para $f_2(x)$, ou seja, $f_2(T(x)) = f_2(x)$. Consequentemente se $x \in E_{\alpha,\beta}$, temos que $f_1(x) < \alpha < \beta < f_2(x)$. Então $f_1(T(x)) < \alpha < \beta < f_2(T(x))$ e portanto $T(x) \in E_{\alpha,\beta}$.

Agora, podemos olhar para as restrições $T : E_{\alpha,\beta} \to E_{\alpha,\beta}$ e $f : E_{\alpha,\beta} \to \mathbb{R}$. Se $x \in E_{\alpha,\beta}$, temos que

$$\beta < \sup_n \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) = f^*(x).$$

Observe que $E_{\alpha,\beta} = D_\beta = \{y \in E_{\alpha,\beta} : f^*(y) > \beta\}$ na restrição de f a $E_{\alpha,\beta}$. Do Corolário 2.2 do teorema ergódico maximal segue que

$$\int_{D_\beta} f \, d\mu \geqslant \beta \cdot \mu(D_\beta),$$

ou seja,

$$\int_{E_{\alpha,\beta}} f \, d\mu \geqslant \beta \cdot \mu(E_{\alpha,\beta}). \tag{2.11}$$

Por outro lado, veja que se $x \in E_{\alpha,\beta}$ então

$$\inf_n \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) < \alpha.$$

De modo anáologo, segue do Corolário 2.3 que

$$\int_{E_{\alpha,\beta}} f \, d\mu \leqslant \alpha \cdot \mu(E_{\alpha,\beta}). \tag{2.12}$$

De (2.11) e (2.12) segue que

$$\beta \cdot \mu(E_{\alpha,\beta}) \leqslant \alpha \cdot \mu(E_{\alpha,\beta}).$$

Como $\mu(E_{\alpha,\beta}) \geqslant 0$ e $\alpha < \beta$, a última desigualdade só ocorre se $\mu(E_{\alpha,\beta}) = 0$. Portanto, o limite

$$\widetilde{f}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)),$$

existe em μ-q.t.p. $x \in X$.

Mostraremos agora que $\widetilde{f}$ é integrável, ou seja, que $\int_X |\widetilde{f}| \, d\mu < \infty$. Observemos que

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \right| \leqslant \frac{1}{n} \sum_{j=0}^{n-1} |f|(T^j(x)).$$

Do lema de Fatou, temos

$$\int_X \liminf_{n\to\infty} \left| \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \right| d\mu \leqslant \liminf_{n\to\infty} \int_X \left| \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \right| d\mu.$$

Daí,

$$\int_X \liminf_{n\to\infty} \left| \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \right| d\mu \leqslant \liminf_{n\to\infty} \int_X \frac{1}{n} \sum_{j=0}^{n-1} |f|(T^j(x)) \, d\mu.$$

Portanto,

$$\begin{aligned} \int_X |\widetilde{f}| \, d\mu &\leqslant \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int_X |f| \circ T^j \, d\mu \\ &\leqslant \int_X |f| \, d\mu, \end{aligned}$$

uma vez que μ é invariante por T, assim $\widetilde{f}$ é integrável. Finalmente, mostraremos que

$$\int_X \widetilde{f} \, d\mu = \int_X f \, d\mu,$$

para isso, dados $n \in \mathbb{N}$ e $k \in \mathbb{Z}$, considere o conjunto

$$A_{n,k} = \left\{ x \in X; \frac{k}{2^n} \leqslant \widetilde{f}(x) < \frac{k+1}{2^n} \right\}.$$

Observe que

$$\frac{k}{2^n}\cdot\mu(A_{n,k})\leqslant\int_{A_{n,k}}\widetilde{f}\,d\mu\leqslant\frac{(k+1)}{2^n}\cdot\mu(A_{n,k}).$$

Note também que $\widetilde{f}(T(x))=\widetilde{f}(x)$. De fato,

$$\begin{aligned}\widetilde{f}(T(x)) &= \lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}f(T^j(T(x)))\\ &= \lim_{n\to\infty}\frac{1}{n}\left(\sum_{j=0}^{n}f(T^j(x))-f(x)\right)\\ &= \lim_{n\to\infty}\frac{n+1}{n}\left(\frac{1}{n+1}\sum_{j=0}^{n}f(T^j(x))-\frac{f(x)}{n+1}\right)\\ &= \widetilde{f}(x),\end{aligned}$$

logo $A_{n,k}$ é invariante por T.

Considere as restrições $T : A_{n,k}\to A_{n,k}$ e $f : A_{n,k}\to\mathbb{R}$. Dado $\varepsilon>0$ e $x\in A_{n,k}$, existe $N\in\mathbb{N}$ tal que

$$\frac{k}{2^n}-\varepsilon<\frac{1}{N}\sum_{j=0}^{N-1}f(T^j(x)).$$

Daí,

$$\frac{k}{2^n}-\varepsilon<f^*(x).$$

Seja $\theta_\varepsilon=\frac{k}{2^n}-\varepsilon$, pelo Corolário 2.2 temos que

$$\int_{D_{\theta_\varepsilon}}f\,d\mu\geqslant\theta_\varepsilon\cdot\mu(D_{\theta_\varepsilon}).$$

Da mesma maneira que foi feita com os conjuntos $E_{\alpha,\beta}$, olhando na restrição, temos que $D_{\theta_\varepsilon}=A_{n,k}$ assim

$$\int_{A_{n,k}}f\,d\mu\geqslant\left(\frac{k}{2^n}-\varepsilon\right)\cdot\mu(A_{n,k}).$$

Fazendo $\varepsilon\to 0^+$, obtemos

$$\int_{A_{n,k}}f\,d\mu\geqslant\frac{k}{2^n}\cdot\mu(A_{n,k}).$$

De modo análogo, usando o Corolário 2.3 temos que

$$\int_{A_{n,k}}f\,d\mu\leqslant\frac{(k+1)}{2^n}\cdot\mu(A_{n,k}).$$

Das desigualdades acima temos que

$$\frac{k}{2^n}\cdot\mu(A_{n,k})\leqslant\int_{A_{n,k}}f\,d\mu\leqslant\frac{(k+1)}{2^n}\cdot\mu(A_{n,k}).$$

Portanto,

$$\left|\int_{A_{n,k}} \widetilde{f}\, d\mu - \int_{A_{n,k}} f\, d\mu\right| \leqslant \frac{1}{2^n} \cdot \mu(A_{n,k}).$$

Para n fixado, podemos escrever X como a união disjunta abaixo,

$$X = \bigcup_{k\in\mathbb{Z}} A_{n,k}.$$

Daí,

$$\begin{aligned}
\left|\int_X \widetilde{f}\, d\mu - \int_X f\, d\mu\right| &= \left|\sum_{k\in\mathbb{Z}} \int_{A_{n,k}} \widetilde{f}\, d\mu - \sum_{k\in\mathbb{Z}} \int_{A_{n,k}} f\, d\mu\right| \\
&= \left|\sum_{k\in\mathbb{Z}} \left(\int_{A_{n,k}} \widetilde{f}\, d\mu - \int_{A_{n,k}} f\, d\mu\right)\right| \\
&\leqslant \sum_{k\in\mathbb{Z}} \left|\int_{A_{n,k}} \widetilde{f}\, d\mu - \int_{A_{n,k}} f\, d\mu\right| \\
&= \sum_{k\in\mathbb{Z}} \frac{1}{2^n} \mu(A_{n,k}) \\
&= \frac{1}{2^n} \sum_{k\in\mathbb{Z}} \mu(A_{n,k}) \\
&= \frac{1}{2^n},
\end{aligned}$$

fazendo $n \to \infty$, obtemos que $\int_X \widetilde{f}\, d\mu = \int_X f\, d\mu$, o que conclui a prova do teorema.

2.5 Teorema ergódico subaditivo

Mostraremos aqui o teorema ergódico subaditivo em tempo discreto, que é devido à Kingman, veja [9]. A demonstração aqui apresentada é baseada no livro de Viana e Oliveira [18], que por sua vez é devida à Artur Ávila e Jairo Bochi, veja [1].

Ao longo deste tópico, considere fixado um espaço de probabilidade $(X, \mathcal{A}, \mu)$ e uma transformação mensurável $T : X \to X$ que preserva a medida μ.

Definição 2.5. *Uma sequência* $(a_n)_n$ *em* $\mathbb{R} \cup \{-\infty\}$ *é dita subaditiva se vale* $a_{m+n} \leqslant a_m + a_n$ *para todo* $m, n \geqslant 1$.

Definição 2.6. *Uma sequência de funções* $\mathcal{A}$*-mensuráveis* $f_n : X \to \mathbb{R}, n \geqslant 1$ *é dita subaditiva para uma transformação* $T : X \to X$ *se vale* $f_{m+n} \leqslant f_m + f_n \circ T^m$ *para todo* $m, n \geqslant 1$.

Teorema 2.6. *(Kingman) Seja* $f_n : X \to \mathbb{R}$, $n \geqslant 1$ *uma sequência de funções mensuráveis tal que* $f_1^+ \in L^1(\mu)$ *e*

$$f_{m+n} \leqslant f_m + f_n \circ T^m \text{ , para todo } m, n \geqslant 1$$

então a sequência $(f_n/n)_n$ *converge em* μ*-quase todo ponto para uma função mensurável* $f : X \to \mathbb{R} \cup \{-\infty\}$. *Além disso,* $f^+ \in L^1(\mu)$ *e*

$$\int f d\mu = \lim_{n\to\infty} \frac{1}{n} \int f_n \, d\mu = \inf_n \frac{1}{n} \int f_n \, d\mu \in [-\infty, +\infty).$$

Lema 2.2. *Seja* $f : X \to \mathbb{R}$ *uma função integrável. Então*

$$\lim_{n\to\infty} \frac{f \circ T^n(x)}{n} = 0$$

em μ*-quase todo ponto* $x \in X$.

Demonstração: Dado $\varepsilon > 0$ e $n \in \mathbb{N}$, defina o conjunto

$$E_n := \{x \in X; \frac{|f \circ T_n|(x)}{n} \geqslant \varepsilon\} = \{x \in X; |f \circ T_n|(x) \geqslant \varepsilon n\}.$$

considere tambem

$$F_\varepsilon := \limsup_{n\to\infty} E_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n.$$

assim, se mostrarmos que $\mu(F_\varepsilon) = 0$, temos que o conjunto dos valores $x \in X$ tal que $x \in E_n$ para infinitos n tem medida nula, assim teremos que

$$\lim_{n\to\infty} \frac{f \circ T^n(x)}{n} = 0$$

em μ-q.t.p. x em X.

Observe que, como T preserva a medida μ, temos que

$$\sum_{n=1}^{\infty} \mu\left(\{|f \circ T^n| \geqslant \varepsilon n\}\right) = \sum_{n=1}^{\infty} \mu\left(\{|f| \geqslant \varepsilon n\}\right) = \sum_{n=1}^{\infty} \mu\left(\left\{|\frac{f}{\varepsilon}| \geqslant n\right\}\right)$$

para cada $n \in \mathbb{N}$, podemos considerar os valores de x tais que $k \leqslant |\frac{f}{\varepsilon}| \leqslant k+1$ onde $k \geqslant n$ assim temos que

$$\sum_{n=1}^{\infty} \mu\left(\left\{|\frac{f}{\varepsilon}| \geqslant n\right\}\right) = \sum_{n=1}^{\infty} \left(\sum_{k=n}^{\infty} \mu\left(\left\{k \leqslant |\frac{f}{\varepsilon}| < k+1\right\}\right)\right)$$

seja

$$C_k := \left\{k \leqslant |\frac{f}{\varepsilon}| < k+1\right\},$$

veja que

$$\begin{aligned}\sum_{n=1}^{\infty}\left(\sum_{k=n}^{\infty}\mu(C_k)\right) &= \sum_{k=1}^{\infty}\mu(C_k)+\sum_{k=2}^{\infty}\mu(C_k)+\sum_{k=3}^{\infty}\mu(C_k)+\sum_{k=4}^{\infty}\mu(C_k)+\cdots\\ &= \sum_{k=1}^{\infty}k\cdot\mu(C_k)\end{aligned}$$

por outro lado, veja que $\bigcup_{k=1}^{\infty}C_k$ é uma união disjunta, dai temos que

$$\begin{aligned}\int_X \frac{|f|}{\varepsilon}|\,d\mu \geqslant \int_{\bigcup_{k=1}^{\infty}C_k}\frac{|f|}{\varepsilon}\,d\mu &= \sum_{k=1}^{\infty}\int_{C_k}\frac{|f|}{\varepsilon}\,d\mu\\ &\geqslant \sum_{k=1}^{\infty}\int_{C_k}k\,d\mu = \sum_{k=1}^{\infty}k\mu(C_k)\end{aligned}$$

assim, como f é integrável, temos que

$$\sum_{n=1}^{\infty}\mu(\{|f\circ T^n|\geqslant \varepsilon n\}) = \sum_{k=1}^{\infty}k\mu(C_k)\leqslant \int_X\frac{|f|}{\varepsilon}|\,d\mu<\infty$$

concluimos que $\sum_{n=1}^{\infty}\mu(E_n)$ é finito e, portanto, pelo lema de Borel-Cantelli segue que $\mu(F_\varepsilon)=0$. ∎

Lema 2.3. *(Fekete) Se* (a_n) *é uma sequência subaditiva então*

$$\lim_{n\to\infty}\frac{a_n}{n}=\inf_n\frac{a_n}{n}\in[-\infty,+\infty).$$

Demonstração: Se existe $m\in\mathbb{N}$ tal que $a_m=-\infty$ então como (a_n) é subaditiva, para todo $n\geqslant m$, vale $a_n=-\infty$, e nada há de ser feito, pois

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

Suponha que $a_n\in\mathbb{R}$ para todo n. Considere

$$L=\inf_n\frac{a_n}{n}\in[-\infty,+\infty)$$

e fixemos $\beta\in\mathbb{R}$ tal que $\beta>L$. Assim, por definição de ínfimo, existe $k\in\mathbb{N}, k\geqslant 1$ tal que

$$\frac{a_k}{k}\leqslant\beta.$$

para todo $n>k$, podemos escrever $n=qk+r$ com $q\geqslant 1$ e $0\leqslant r\leqslant k$. Da subaditividade segue que

$$\begin{aligned}a_n\leqslant a_{kq}+a_r &\leqslant a_k+a_k+\ldots+a_k+a_r\\ &\leqslant qa_k+a_r\\ &\leqslant qa_k+\alpha\end{aligned}$$

onde $\alpha = \max\{a_i; 1 \leqslant i \leqslant k\}$. Logo,

$$\frac{a_n}{n} \leqslant \frac{qa_k + \alpha}{n} \leqslant \frac{qk}{n} \cdot \frac{a_k}{k} + \frac{\alpha}{n}$$

veja que

$$\lim_{n\to\infty} \frac{qk}{n} = \lim_{n} \frac{n-r}{n} = 1 \text{ e } \lim_{n} \frac{\alpha}{n} = 0$$

logo

$$\lim_{n\to\infty} \frac{a_n}{n} \leqslant \frac{a_k}{k}$$

assim, como $\frac{a_k}{k} \leqslant \beta$, para n suficientemente grande temos que

$$L \leqslant \frac{a_n}{n} \leqslant \beta$$

portanto, como β é arbitrário, fazendo $\beta \to L^+$, concluimos que

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

■

Dada uma sequência de funções $f_n : X \to \mathbb{R}, n \geqslant 1$ mensuráveis tal que $f_1^+ \in L^1(\mu)$ e (f_n) é uma sequência subaditiva, temos que

$$f_n \leqslant f_1 + f_1 \circ T + ... + f_1 \circ T^{n-1}$$

Note que a desigualdade permanece se trocarmos f_n e f_1 por f_n^+ e f_1^+. Assim como $f_1^+ \in L^1(\mu)$ segue que $f_n^+ \in L^1(\mu)$ para todo n. Além disso, como (f_n) é subaditiva, e T preserva μ, para $n \geqslant 1$ temos que a sequência

$$a_n = \int f_n \, d\mu$$

é subaditiva em $\mathbb{R} \cup \{-\infty\}$. Logo pelo lema 2.3 temos que o limite

$$L = \lim_{n\to\infty} \frac{1}{n} \int f_n \, d\mu = \inf_{n} \frac{1}{n} \int f_n \, d\mu \in [-\infty, \infty)$$

existe.

Defina $f_- : X \to [-\infty, +\infty]$ e $f_+ : X \to [-\infty, +\infty]$ dadas por:

$$f_-(x) = \liminf_{n\to\infty} \frac{f_n(x)}{n}$$

e

$$f_+(x) = \limsup_{n\to\infty} \frac{f_n(x)}{n},$$

observe que f_- e f_+ são T- invariantes, ou seja,

$$f_-(T(x)) = f_-(x) \text{ e } f_+(T(x)) = f_+(x)$$

para μ-quase todo ponto $x \in X$. De fato, como (f_n) é subaditiva, segue que $f_{1+n} \leqslant f_1 + f_n \circ T$ para todo n, assim temos

$$\frac{f_{1+n}(x)}{n} \leqslant \frac{f_1(x)}{n} + \frac{f_n \circ T(x)}{n}$$

$$\Downarrow$$

$$\liminf_{n\to\infty} \frac{f_-(x)}{n} \leqslant \liminf_{n\to\infty} \frac{f_- \circ T(x)}{n}$$

$$\Downarrow$$

$$f_- \leqslant f_- \circ T$$

então

$$\int f_- \, d\mu \leqslant \int f_- \circ T \, d\mu$$

de onde devemos ter a igualdade, pois, caso contrário, como f_- é integrável e T é invariante por μ teriamos

$$\int f_- \, d\mu < \int f_- \circ T \, d\mu = \int f_- \, d\mu$$

um absurdo. Portanto, como

$$\int f_- \, d\mu \leqslant \int f_- \circ T \, d\mu,$$

temos que $f_- = f_- \circ T$ em μ-q.t.p. $x \in X$.

Veja que $f_-(x) \leqslant f_+(x)$ para todo $x \in X$ logo

$$\int f_- \, d\mu \leqslant \int f_+ \, d\mu$$

supondo que (f_n) é limitada por baixo, iremos mostrar que

$$\int f_- \, d\mu \geqslant L \geqslant \int f_+ \, d\mu \tag{2.13}$$

assim $f_- = f_+$ em μ-quase todo ponto e

$$\int f_- \, d\mu = \int f_+ \, d\mu = L$$

ou seja, se (f_n) for limitada, o teorema estará demonstrado. Assim, sob hipótese de limitação de (f_n), mostraremos que vale (2.13), posteriormente, retiraremos esta hipótese utilizando uma método de truncagem. Mas primeiro faremos os preparativos para o lema fundamental na demonstração de (2.13).

Suponha que $f_- > -\infty$. Fixado $\varepsilon > 0$, para cada $k \in \mathbb{N}$, definimos o conjunto

$$E_k := \left\{ x \in X; \frac{f_j(x)}{j} \leqslant f_-(x) + \varepsilon \text{ para algum } j \in \{1, ..., k\} \right\}$$

veja que se $x \in E_k$ existe $j \in \{1, ..., k\}$ tal que

$$\frac{f_j(x)}{j} \leqslant f_-(x) + \varepsilon$$

assim $j \in \{1, ..., k, k+1\}$ e portanto $E_k \subset E_{k+1}$ para todo $k \in \mathbb{N}$. Por outro lado, da definição de $f_-(x)$, para todo x existe k tal que $x \in E_k$, logo

$$X = \bigcup_k E_k.$$

Assim, temos que $E_k \nearrow X$, logo, segue que $\mu(E_k) \nearrow \mu(X) = 1$. Daí, definimos

$$\psi_k(x) = \begin{cases} f_-(x) + \varepsilon \ \texttt{se} \ x \in E_k \\ f_1(x) \ \texttt{se} \ x \in E_k^c \end{cases} .$$

assim, para todo k, se $x \in E_k^c$, temos $\psi_k(x) = f_1(x) > f_-(x) + \varepsilon$ (por definição de E_k). Temos que a sequência $(\psi_k(x))$ é não-crescente, de fato, dado $x \in X$, temos que:

- $x \in E_k \subset E_{k+1} \Rightarrow \psi_k(x) = f_1(x) + \varepsilon = \psi_{k+1}(x)$

- $x \in E_k^c = E_{k+1}^c \cup (E_{k+1} - E_k)$, ou seja:

$$\psi_k(x) = f_1(x) = \psi_{k+1}(x)$$

ou

$$\psi_k(x) = f_1(x) > f_-(x) + \varepsilon = \psi_{k+1}(x).$$

Logo

$$\psi_{k+1}(x) \leqslant \psi_k(x) \text{ para todo } k.$$

Note também que $\psi_k(x) \to f_-(x) + \varepsilon$, para todo x, pois dado $x \in X$, como $X = \bigcup_k E_k$, existe k_0 suficientemente grande tal que $x \in E_{k_0}$, assim $\psi_k(x) = f_-(x) + \varepsilon$ para todo $k \geqslant K_0$. Assim, pelo teorema da convergência monótona, temos que

$$\int \psi_k \, d\mu \to \int (f_- + \varepsilon) \, d\mu, \ \texttt{quando} \ k \to \infty.$$

Com isso, temos que o passo crucial da demonstração do teorema é a seguinte estimativa:

Lema 2.4. *(lema fundamental) Dados n, k inteiros tais que $n > k \geqslant 1$, vale a desigualdade*

$$f_n(x) \leqslant \sum_{i=0}^{n-k-1} \psi_k(T^i(x)) + \sum_{i=n-k}^{n-1} max\{\psi_k, f_1\}(T^i(x))$$

para μ-quase todo ponto $x \in X$.

Demonstração: Pelo que acabamos de mostrar f_- é T-invariante, assim, tome $x \in X$ tal que $f_-(x) = f_-(T^j(x))$ (esta igualdade vale para μ-quase todo ponto e $j \geqslant 1$). Considere a sequência (possivelmente finita) de números inteiros

$$m_0 \leqslant n_1 < m_1 \leqslant n_2 < m_2 \leqslant ...$$

definida indutivamente da seguinte forma:

1. tome $m_0 = 0$ considere n_1 o menor inteiro maior ou igual a zero tal que $T_{n_1}(x) \in E_k$, daí, por definição de E_k, existe m_1 tal que $1 \leqslant m_1 - n_1 < k$ e

$$\frac{f_{m_1-n_1}(T^{n_1}(x))}{m_1 - n_1} \leqslant f_-(T^{n_1}(x)) + \varepsilon$$

2. para $j \geqslant 1$, seja n_j o menor inteiro maior ou igual a m_{j-1} tal que $T_{n_j}(x) \in E_k$, assim existe m_j tal que $1 \leqslant m_j - n_j < k$ e que

$$\frac{f_{m_j-n_j}(T^{n_j}(x))}{m_j - n_j} \leqslant f_-(T^{n_j}(x)) + \varepsilon \tag{2.14}$$

concluindo assim a definição da sequência. Agora, dado $n \geqslant K$, seja $l \leqslant 0$ o maior número inteiro tal que $m_l \leqslant n$. Pela subaditividade de f, teremos

$$\begin{aligned} f_{n_j-m_{j-1}}(T^{m_{j-1}}(x)) \quad & \leqslant f_1(T^{m_{j-1}}(x)) + f_1(T^{m_{j-1}+1}(x)) + \\ & +... + f_1(T^{n_{j-1}}(x)) = \sum_{i=m_{j-1}}^{n_j-1} f_1(T^i(x)) \end{aligned}$$

válido para todo $j \in \{1, ..., l\}$ tal que $m_{j-1} \neq n_j$, e para $f_{n-m_l}(T^{m_l}(x))$. Dai, tomando $I = \bigcup_{j=1}^{l}[m_{j-1}, n_j) \cup [m_l, n)$, temos que

$$f_n(x) \leqslant \sum_{i \in I} f_1(T^i(x)) + \sum_{j=1}^{l} f_{m_j-n_j}(T^{n_j}(x)) \tag{2.15}$$

Analisando a trajetória do ponto, temos duas opções, ou $T^n(x) \in E_k^c$ ou $T^n(x) \in E_k$, neste segundo caso existe $n_{l+1} < n$. Tomando

$$i \in \bigcup_{j=1}^{l} [m_{j-1}, n_j) \cup [m_l, \min\{n_{l+1}, n\})$$

temos que $T^i(x) \in E_k^c$ e portanto

$$f_1(T^i(x)) = \psi_k(T^i(x))).$$

Como f_- é T-invariante, temos que f_- é constante em órbitas e como $\psi_k \geqslant f_- + \varepsilon$, pela relação (2.14), teremos

$$f_{m_j - n_j}(T^{n_j}(x)) \leqslant \sum_{i=n_j}^{m_j - 1} (f_-(T^i(x)) + \varepsilon) \leqslant \sum_{i=n_j}^{m_j - 1} \psi_k(T^i(x))$$

para todo $j \in \{1, ..., l\}$. Assim da relação (2.15), resulta em :

$$f_n(x) \leqslant \sum_{i=0}^{\min\{n_{l+1}, n\} - 1} \psi_k(T^i(x)) + \sum_{i=n_{l+1}}^{n-1} f_1(T^i(x))$$

portanto, como $n_{l+1} > n - k$, concluimos a demonstração do lema. ■

Á fim de mostrar (2.13), provaremos o seguinte lema:

Lema 2.5. $\displaystyle\int f_- \, d\mu = L$

Demonstração: Primeiramente, suponha que f_n/n está uniformemente limitada por baixo (logo após retiraremos esta hipótese), ou seja, existe $k \in \mathbb{N}$ fixado tal que $-k \leqslant \dfrac{f_n}{n}$, para todo n. Temos que $\left(\dfrac{f_n}{n} + k\right)$ é uma sequência de funções não-negativas e, aplicando o lema de Fatou, temos

$$\int \liminf_{n\to\infty} \left(\frac{f_n}{n} + k\right) d\mu \leqslant \liminf_{n\to\infty} \int \left(\frac{f_n}{n} + k\right) d\mu$$

$$\Downarrow$$

$$\int f_- \, d\mu \leqslant \liminf_{n\to\infty} \int \frac{f_n}{n} \, d\mu = \lim_{n\to\infty} \int \frac{f_n}{n} \, d\mu = L$$

assim, temos que f_- é integrável e sua integral é menor ou igual a L. Para provar a outra desigualdade, note que o lema fundamental e a invariância de μ nos dá que

$$\frac{1}{n} \int f_n \, d\mu \leqslant \frac{1}{n} \int \left(\sum_{i=0}^{n-k-1} \psi_k(T^i(x)) + \sum_{i=n-k}^{n-1} \max\{\psi_k, f_1\}(T^i(x)) \right)$$

$$\Downarrow$$

$$\frac{1}{n}\int f_n\, d\mu \leqslant \frac{(n-k)}{n}\int \Psi_k\, d\mu + \frac{k}{n}\int \max\{\Psi_k, f_1\}\, d\mu \tag{2.16}$$

Note que

$$\max\{\Psi_k, f_1\} \leqslant \max\{f_- + \varepsilon, f_1^+\},$$

na qual $\max\{f_- + \varepsilon, f_1^+\}$ é integrável, pois $f_- + \varepsilon$ e f_1^+ são integráveis, logo

$$\limsup_{n\to\infty} \frac{k}{n}\int \max\{\Psi_k, f_1\}\, d\mu \leqslant \limsup_{n\to\infty} \frac{k}{n}\int \max\{f_- + \varepsilon, f_1^+\}\, d\mu = 0$$

ou seja,

$$\frac{1}{n}\int f_n\, d\mu \leqslant \frac{(n-k)}{n}\int \Psi_k\, d\mu \text{ , para todo n}$$

logo, fazendo $n \to \infty$, temos

$$L \leqslant \int \Psi_k\, d\mu$$

para todo k, assim temos

$$L \leqslant \lim_{k\to\infty}\int \Psi_k\, d\mu = \int (f_- + \varepsilon)\, d\mu$$

portanto, fazendo $\varepsilon \to 0^+$, obtemos que $L \leqslant \int f_-\, d\mu$ e estará provado o lema.

Agora, provaremos o lema retirando a hipótese de ser uniformemente limitada por baixo. Procederemos da seguinte maneira, para cada $k > 0$ defina as funções

$$f_n^k = \max\{f_n, -kn\} \text{ e } f_-^k = \max\{f_-, -k\}$$

veja que

$$\begin{aligned} f_{n+m}^k(x) = \max\{f_{n+m}(x), -k(n+m)\} \quad &\leqslant \max\{f_n(x) + f_m(T^n(x)), (-kn) + (-km)\} \\ &\leqslant \max\{f_n(x), -kn\} + \max\{f_m(T^n(x)), -km\} \\ &= f_n^k(x) + f_m^k(T^n(x)) \end{aligned}$$

assim, temos que a sequência $(f_n^k)_k$ é subaditiva e é fácil ver que a parte positiva de $f_1^k = \max\{f_1, -k\}$ é integrável, ou seja, $(f_n^k)_n$ satisfaz as hipóteses do teorema subaditivo. Além disso

$$f_-^k = \max\{f_-, -k\} = \liminf_{n\to\infty}\left(\max\left\{\frac{f_n}{n}, \frac{-kn}{n}\right\}\right) = \liminf_{n\to\infty}\left(\frac{f_n^k}{n}\right)$$

e também,

$$\frac{f_n^k}{n} = \max\left\{\frac{f_n}{n}, -k\right\} \geqslant -k, \text{ para todo } n$$

logo f_n^k é limitada inferiormente, pelo argumento apresentado anteriormente neste lema, temos

$$\int f_-^k \, d\mu = \inf_n \frac{1}{n} \int f_n^k \, d\mu \tag{2.17}$$

note também que quando $k \to \infty$, então $f_n^k \to f_n$ e $f_n^{k+1} = \max\{f_n, -(k+1)n\} \leqslant \max\{f_n, -kn\} = f_n^k$, assim a sequência $(f_n^k)_k$ converge monotóticamente para f_n, e da mesma forma, $(f_-^k)_k$ converge monotóticamente para f_-, assim, pelo teorema da convergência monótona, temos

$$\int f_n \, d\mu = \inf_k \int f_n^k \, d\mu \ \text{ e } \ \int f_- \, d\mu = \inf_k \int f_-^k \, d\mu \tag{2.18}$$

portanto, das relações (2.17) e (2.18) obtemos

$$\int f_- \, d\mu = \inf_k \int f_-^k \, d\mu = \inf_k \left(\inf_n \frac{1}{n} \int f_n^k \, d\mu \right) = \inf_n \frac{1}{n} \int f_n \, d\mu = L$$

o que completa a demonstração do lema. ■

Ainda supondo que $\inf_n f_n$ seja finito para todo n, vamos mostrar que $\int f_+ \, d\mu \leqslant L$, para isso, precisamos do seguinte resultado auxiliar:

Lema 2.6. *Dado* $k \in \mathbb{N}$, *temos*

$$\limsup_{n\to\infty} \frac{f_{kn}}{n} = k \cdot \limsup_{n\to\infty} \frac{f_n}{n}.$$

Demonstração: Veja que (f_{kn}/kn) é uma subsequência de (f_n/n), assim, temos que

$$\limsup_{n\to\infty} \frac{f_{kn}}{n} = k \cdot \left(\limsup_{n\to\infty} \frac{f_{kn}}{kn} \right) \leqslant k \cdot \limsup_{n\to\infty} \frac{f_n}{n}.$$

Para mostrar a outra desigualdade, usando o algoritmo de Euclides, podemos escrever $n = kq + r$, com $r \in \{1, ..., k\}$. Definindo $\phi = \max\{f_1^+, ..., f_k^+\}$, pela subaditividade de $(f_n)_n$, temos

$$f_n \leqslant f_{kq} + f_r \circ T^{kq} \leqslant f_{kq} + \phi \circ T^{kq}$$

veja que,

$$\lim_{n\to\infty} \frac{q}{n} = \frac{1}{k} \cdot \lim_{n\to\infty} \frac{kq}{n} = \frac{1}{k} \cdot \lim_{n\to\infty} \frac{n-r}{n} = \frac{1}{k}$$

além disso, como $f_1^+ \in L^1(\mu)$, e pela subaditividade de (f_n), temos que $f_i^+ \in L^1(\mu)$ para todo $i \in \{1, ..., k\}$, logo $\phi \in L^1(\mu)$ e pelo lema 2.2, temos que $((\phi \circ T^n)/n)$ converge para zero em μ-quase todo ponto, assim, temos que

$$\limsup_{n\to\infty} \frac{1}{n} f_n \leqslant \limsup_{n\to\infty} \frac{1}{n} f_{kq} + \limsup_{n\to\infty} \frac{1}{n} \phi \circ T^{kq}$$

como $q = (n-r)/k$, temos que $n \to \infty$ se, e somente se $q \to \infty$, dai temos

$$\limsup_{n\to\infty} \frac{1}{n} f_{kq} = \limsup_{q\to\infty} \frac{q}{n}\frac{1}{q} f_{kq} = \frac{1}{k} \cdot \limsup_{q\to\infty} \frac{1}{q} f_{kq}$$

e também,

$$\limsup_{n\to\infty} \frac{1}{n} \phi \circ T^{kq} = \limsup_{n\to\infty} \frac{1}{n} \phi \circ T^{n-r} = \limsup_{n\to\infty} \frac{1}{n} \phi \circ T^{n}(T^{-r}(x)) = 0$$

em μ-quase todo ponto. Portanto, temos que

$$\limsup_{n\to\infty} \frac{1}{n} f_n \leqslant \frac{1}{k} \cdot \limsup_{q\to\infty} \frac{1}{q} f_{kq} \Rightarrow k \cdot \limsup_{n\to\infty} \frac{1}{n} f_n \leqslant \limsup_{n\to\infty} \frac{1}{n} f_{kn}$$

completanto a demonstração do lema. ∎

Lema 2.7. *Suponha que* $\inf_n f_n > -\infty$. *Então* $\int f_+ \, d\mu \leqslant L$.

Demonstração: Para cada $k \in \mathbb{N}$ fixado e $n \geqslant 1$, definimos

$$\theta_n = -\sum_{j=0}^{n-1} f_k \circ T^{jk}$$

dai, para todo n, temos

$$\int \theta_n \, d\mu = -\sum_{j=0}^{n-1} \int f_k \circ T^{jk} \, d\mu = -\sum_{j=0}^{n-1} \int f_k \, d\mu = -n \int f_k \, d\mu \tag{2.19}$$

pois T preserva μ. Pela subaditividade de (f_n), temos

$$\begin{aligned} f_{kn} + \theta_n = f_{kn} - \sum_{j=0}^{n-1} f_k \circ T^{jk} \quad & \leqslant f_{(n-1)k} + f \circ T^{(n-1)k} - \sum_{j=0}^{n-1} f_k \circ T^{jk} \\ & \leqslant \ldots \\ & \leqslant \sum_{j=0}^{n-1} f_k \circ T^{jk} - \sum_{j=0}^{n-1} f_k \circ T^{jk} \leqslant 0 \end{aligned}$$

e portanto $\theta_n \leqslant -f_{kn}$ para todo n. Usando o lema 2.6, teremos

$$\theta_- = \liminf_{n\to\infty} \frac{\theta_n}{n} \leqslant \limsup_{n\to\infty} \frac{\theta_n}{n} \leqslant -\limsup_{n\to\infty} \frac{f_{kn}}{n} = -k \cdot \limsup_{n} \frac{f_n}{n} = -kf_+$$

logo

$$\int \theta_- \, d\mu \leqslant -k \int f_+ \, d\mu. \tag{2.20}$$

Observe que (θ_n) é aditiva, ou seja, $\theta_{m+n} = \theta_m + \theta_n \circ T^{km}$ para todo $m, n \geqslant 1$, de fato,

$$\begin{aligned}
\theta_{m+n} = -\sum_{j=0}^{m+n-1} f_k \circ T^{jk} \quad &= -\sum_{j=0}^{m-1} f_k \circ T^{jk} - \sum_{j=m}^{m+n-1} f_k \circ T^{jk} \\
&= \theta_m - \sum_{j=0}^{n-1} f_k \circ T^{(j+m)k} \\
&= \theta_m - \sum_{j=0}^{n-1} f_k \circ T^{jk}(T^{km}) \\
&= \theta_m + \theta_n \circ T^{km}
\end{aligned}$$

note também que $\theta_1^+ = (-f_k)^+$ é limitada (pois f_k é limitada inferiormente), e assim θ_1^+ é integrável. Aplicando o lema 2.5, temos

$$\int \theta_- \, d\mu = \inf_n \int \frac{\theta_n}{n} \, d\mu$$

pela igualdade (2.19), teremos

$$\int \theta_- \, d\mu = \inf_n -\frac{n}{n} \int f_k \, d\mu = -\int f_k \, d\mu \tag{2.21}$$

portanto, das relações (2.20) e (2.21) obtemos

$$\int f_+ \, d\mu \leqslant \frac{1}{k} \int f_k \, d\mu$$

assim,

$$\int f_+ \, d\mu \leqslant \inf_k \frac{1}{k} \int f_k \, d\mu = L$$

e o lema está provado. ■

Com os lemas 2.5 e 2.7, a relação (2.13) é satisfeita, e portanto o teorema está provado sob a condição de que $\inf_n f_n > -\infty$. Para o caso geral, definimos,

$$f_n^k = \max\{f_n, -kn\}\ ,\ f_-^k = \max\{f_-, -k\} \text{ e } f_+^k = \max\{f_+, -k\}$$

para cada $k \in \mathbb{N}$. Analogamente ao que já foi feito, mostramos que $(f_n^k)_n$ é subaditiva e $(f_1^k)^+ \in L^1(\mu)$ e como $\inf_n f_n^k > -k$, obtemos que $f_-^k = f_+^k$ em μ-quase todo ponto para todo k. Fazendo $k \to \infty$, é fácil ver que $f_-^k \to f_-$ e $f_+^k \to f_+$ e, portanto, no geral, temos $f_- = f_+$ em μ-quase todo ponto e a demonstração do teorema está completa.

Corolário 2.4. *Teorema subaditivo ⇒ teorema de Birkhoff*

Demonstração: Considere a seguinte sequência de somas orbitais :

$$f_n = \sum_{j=0}^{n-1} f \circ T^j,\ n \geqslant 1$$

temos que (f_n) é aditiva, de fato

$$\begin{aligned} f_{n+m} = \sum_{j=0}^{n+m-1} f \circ T^j \quad &= \sum_{j=0}^{n-1} f \circ T^j + \sum_{j=n}^{n+m-1} f \circ T^j \\ &= \sum_{j=0}^{n-1} f \circ T^j + \sum_{j=0}^{m-1} f \circ T^j(T^n) \\ &= f_n + f_m \circ T^n \end{aligned}$$

em particular, (f_n) é subaditiva, e como $f_1^+ = f^+ \in L^1(\mu)$, pelo teorema subaditivo, temos que existe o limite de (f_n/n) em μ-quase todo ponto para uma função $\widetilde{f}$, e além disso como μ é T- invariante, pela proposição 1.5 temos,

$$\begin{aligned} \int \widetilde{f} d\mu \quad &= \lim_{n\to\infty} \frac{1}{n} \int \sum_{j=0}^{n-1} f \circ T^j \, d\mu \\ &= \lim_{n\to\infty} \frac{n}{n} \int f \, d\mu = \int f \, d\mu \end{aligned}$$

disto decorre o teorema de Birkhoff. ∎

Capítulo 3

Ergodicidade

Neste capítulo estudaremos o conceito de ergodicidade que será muito importante para as aplicações que serão apresentadas no próximo capítulo. Parte substancial deste capítulo encontra-se em [18].

3.1 Definição e propriedades

No capítulo anterior foram apresentados resultados importantes da Teoria Ergódica. Para adentrar no conceito de ergodicidade, primeiro precisamos definir o conceito de tempo médio de visita.

Definição 3.1. *Sejam* $E \subset M$ *um conjunto mensurável com medida positiva e* $T : M \to M$ *uma transformação mensurável invariante por* μ. *Chamamos de tempo médio de visita de* x *a* E *o valor de*

$$\tau(E, x) = \lim_{n\to\infty} \frac{1}{n} \#\{0 \leqslant j < n; T^j(x) \in E\},$$

onde $\#$ *denota a cardinalidade do conjunto.*

Assim o tempo médio de visita calcula a média dos iterados de x que "visitam" o conjunto E. Veja que ao usar a função característica do conjunto E podemos redefinir a expressão anterior por

$$\tau(E, x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_E \circ T^j(x).$$

Como a função caracteristica é integrável, o teorema ergódico de Birkhoff nos diz que o tempo médio de visita $\tau(E, x)$ está bem definido para quase todo ponto x. Ao longo deste capítulo vamos supor que μ é uma medida de probabilidade invariante pela transformação mensurável $T : M \to M$.

Definição 3.2. *O sistema (T, μ) é dito ergódico se dado qualquer conjunto mensurável E, temos que $\tau(E, x) = \mu(E)$ para μ-quase todo ponto $x \in M$, ou seja, o tempo médio de visita a qualquer conjunto mensurável coincide, em μ-quase todo ponto, com a medida desse conjunto. Neste caso, dizemos que μ é ergódica para T.*

Definição 3.3. *Dizemos que uma função mensurável $f : M \to \mathbb{R}$ é invariante por T se $f = f \circ T$ em μ-quase todo ponto. Neste caso, dizemos que f é constante na trajetória de T, ou simplesmente, f é constante em órbitas.*

Definição 3.4. *Dizemos que um conjunto mensurável $B \subset M$ é invariante se ele difere de sua pré-imagem $T^{-1}(B)$ por um conjunto de medida nula, ou seja,*

$$\mu(B \Delta T^{-1}(B)) = 0.$$

Proposição 3.1. *Seja μ uma probabilidade invariante de uma transformação mensurável $T : M \to M$. As seguintes condições são equivalentes:*

a) Para todo conjunto mensurável $B \subset M$, tem-se $\tau(B, x) = \mu(B)$ em μ-quase todo ponto x.

b) Para todo conjunto mensurável $B \subset M$, a função $\tau(B, \cdot)$ é constante em μ-quase todo ponto.

c) Para toda função integrável $f : M \to \mathbb{R}$, tem-se que $\widetilde{f}(x) = \int f \, d\mu$ para μ-quase todo ponto.

d) Para toda função integrável $f : M \to \mathbb{R}$, a média temporal $\widetilde{f} : M \to \mathbb{R}$ é constante para μ-quase todo ponto.

e) Para toda função integrável invariante $g : M \to \mathbb{R}$ tem-se que $g(x) = \int g \, d\mu$ para μ-quase todo ponto.

f) Toda função integrável invariante $g : M \to \mathbb{R}$ é constante em μ-quase todo ponto.

g) Para todo subconjunto invariante $A \subset M$ tem-se $\mu(A) = 0$ ou $\mu(A) = 1$.

Demonstração: (a) ⇒ (b), (c) ⇒ (d) e (e) ⇒ (f) são imediatos. Já, (e) ⇒ (c) e (f) ⇒ (d) decorrem do fato de que a média temporal é invariante. De fato, para $\mu - q.t.p$ $x \in M$ temos,

$$\begin{aligned}
\widetilde{f}(T(x)) &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(T(x))) \\
&= \lim_{n\to\infty} \left[\frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) + \frac{f(T^n)}{n} - \frac{f(x)}{n} \right] \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) + \lim_{n\to\infty} \frac{f(T^n(x))}{n} - \lim_{n\to\infty} \frac{f(x)}{n} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \\
&= \widetilde{f}(x).
\end{aligned}$$

(c) ⇒ (a) Seja $B \subset M$ um conjunto mensurável. Como a função característica do conjunto B é integrável, temos que

$$\begin{aligned}
\tau(B, x) &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_B \circ T^j(x) \\
&= \int \mathcal{X}_B \, d\mu \\
&= \mu(B).
\end{aligned}$$

(d) ⇒ (b) Decorre do fato de que o tempo médio de visita é a média temporal da função caracteristica de B como visto acima.

(b) ⇒ (g) Seja A um conjunto invariante, ou seja, $\mu(A \Delta T^{-1}(A)) = 0$. Observe que $\mathcal{X}_A = \mathcal{X}_{T^{-1}(A)} = \mathcal{X}_A \circ T$, em μ-q.t.p. $x \in M$. Logo $\mathcal{X}_A$ é invariante e

$$\tau(A, x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_A \circ T^j(x) = 1,$$

para μ-q.t.p. $x \in A$. De modo anáologo $\tau(A, x) = 0$ para μ-q.t.p. $x \in A^c$. Por hipótese, $\tau(A, \cdot)$ é constante em μ-quase todo ponto assim $\mu(A) = 1$ ou $\mu(A) = 0$.

(g) ⇒ (e) Seja g uma função integrável invariante. Então, para todo $c \in \mathbb{R}$, o conjunto

$$B_c = \{x \in M; g(x) \leqslant c\}$$

é invariante. Por hipótese, temos que $\mu(B_c) = 1$ ou $\mu(B_c) = 0$. Por outro lado, observe que a função $c \mapsto \mu(B_c)$ é não decrescente, pois se $c_1 < c_2$ então $B_{c_1} \subset B_{c_2}$ assim

$\mu(B_{c_1}) \leqslant \mu(B_{c_2})$. Note que existe $\bar{c}$ tal que tal $\mu(B_c) = 0$ se $c < \bar{c}$ e $\mu(B_c) = 1$ se $c \geqslant \bar{c}$. Em particular, $\mu(B_{\bar{c}}) = 1$, logo $g = \bar{c}$ em μ-quase todo ponto. Daí,

$$\int_M g \, d\mu = \bar{c} \int_M 1 \, d\mu = \bar{c} \cdot \mu(M) = \bar{c}.$$

Portanto, $g = \displaystyle\int_M g \, d\mu$ como queríamos demonstrar. ∎

A seguir, temos uma caracterização da propriedade de ergodicidade por meio do operador de Koopman $U_T(f) = f \circ T$

Proposição 3.2. *Seja μ uma probabilidade invariante por uma transformação mensurável $T : M \to M$. As seguintes condições são equivalentes :*

a) (T, μ) *é ergódico*

b) Para qualquer par de conjuntos mensuráveis A e B vale

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}(A) \cap B) = \mu(A) \cdot \mu(B).$$

c) Para quaisquer funções $\phi \in L^p(\mu)$ e $\psi \in L^q(\mu)$ com $\frac{1}{p} + \frac{1}{q} = 1$, vale:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int (U_T^j \phi) \cdot \psi \, d\mu = \int \phi \, d\mu \cdot \int \psi \, d\mu.$$

Demonstração: (c) $\Rightarrow$ (b) Sejam A e B conjuntos mensuráveis. Assumindo (c), tome $\phi = \mathcal{X}_A$ e $\psi = \mathcal{X}_B$. Daí, temos

$$\begin{aligned}
\mu(A) \cdot \mu(B) &= \int \mathcal{X}_A \, d\mu \cdot \int \mathcal{X}_B \, d\mu \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int U_T^j(\mathcal{X}_A) \cdot \mathcal{X}_B \, d\mu \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int (\mathcal{X}_A \circ T^j) \cdot \mathcal{X}_B \, d\mu \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int \mathcal{X}_{T^{-j}(A)} \cdot \mathcal{X}_B \, d\mu \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int \mathcal{X}_{(T^{-j}(A) \cap B)} \, d\mu \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}(A) \cap B).
\end{aligned}$$

(b) $\Rightarrow$ (a) Seja A um conjunto invariante, tomando $A = B$ em (b), temos que

$$(\mu(A))^2 = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j}(A) \cap A) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(A) = \mu(A).$$

Desta forma, $\mu(A) = 0$ ou $\mu(A) = 1$. Segue da Proposição 3.1 que (T, μ) é ergódico.

(a) $\Rightarrow$ (c) Sejam $\phi \in L^p(\mu)$ e $\psi \in L^q(\mu)$. Como o sistema é ergódico segue da Proposição 3.1 e do Teorema Ergódico de Birkhoff que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} U_T^j \phi = \widetilde{\phi} = \int \phi \, d\mu,$$

para μ-q.t.p. $x \in M$. Primeiramente, suponha que ϕ é limitada, ou seja, existe $b > 0$ tal que $|\phi| \leqslant b$. Assim, para todo $n \in \mathbb{N}$ temos que

$$\left| \left(\frac{1}{n} \sum_{j=0}^{n-1} U_T^j \phi \right) \psi \right| \leqslant b|\psi|.$$

Como $b|\psi| \in L^1(\mu)$, podemos usar o Teorema da Convergência Dominada para concluir que

$$\lim_{n\to\infty} \int \left(\frac{1}{n} \sum_{j=0}^{n-1} U_T^j \phi \right) \psi \, d\mu = \int \phi \, d\mu \cdot \int \psi \, d\mu$$

e isto prova (c) sob a condição de limitação de ϕ. Para remover esta condição, dado $k \geqslant 1$, defina

$$\phi_k(x) = \begin{cases} k, & \text{se } \phi(x) > k \\ \phi(x), & \text{se } \phi(x) \in [-k, k] \\ -k, & \text{se } \phi(x) < -k \end{cases}$$

Dado $\varepsilon > 0$, pelo argumento anterior, para todo $k \geqslant 1$, temos que

$$\left| \int \left(\frac{1}{n} \sum_{j=0}^{n-1} U_T^j \phi_k \right) \psi \, d\mu - \int \phi_k \, d\mu \cdot \int \psi \, d\mu \right| < \frac{\varepsilon}{3} \tag{3.1}$$

para todo n suficientemente grande, onde n depende de k. Por outro lado, veja que $\|\phi_k - \phi\|_p \to 0$ quando $k \to \infty$. De fato:

1. Se $p = \infty$, temos que $\phi_k = \phi$, para todo $k \geqslant \|\phi\|_\infty$

2. Se $p < \infty$, veja que $\phi(x) = \lim\limits_{k\to\infty} \phi_k(x)$ e $|\phi_k(x) - \phi(x)| \leqslant 2|\phi(x)|$ para todo x. Pelo Teorema da Convergência Dominada temos que $\|\phi_k - \phi\|_p \to 0$.

Da desiguladade de Hölder, temos

$$\begin{aligned}
\left|\int (\phi_k - \phi)\cdot\left(\int \psi\, d\mu\right) d\mu\right| &\leqslant \left|\left(\int |\phi_k - \phi|^p d\mu\right)^{\frac{1}{p}} \cdot \left(\int \left(\left|\int \psi\, d\mu\right|\right)^q \, d\mu\right)^{\frac{1}{q}}\right| \\
&= \|\phi_k - \phi\|_p \cdot \left|\int \psi\, d\mu\right| < \frac{\varepsilon}{3} \qquad (3.2)
\end{aligned}$$

para todo k suficientemente grande. Da mesma forma, pela desigualdade de Hölder, temos

$$\begin{aligned}
\left|\int \frac{1}{n}\sum_{j=0}^{n-1} U_T^j(\phi_k - \phi)\cdot \psi\, d\mu\right| &\leqslant \frac{1}{n}\sum_{j=0}^{n-1}\left|\int U_T^j(\phi_K - \phi)\cdot\psi\, d\mu\right| \\
&\leqslant \frac{1}{n}\sum_{j=0}^{n-1}\left(\int U_T^j|\phi_k - \phi|^p\, d\mu\right)^{\frac{1}{p}} \cdot \left(\int |\psi|^q\, d\mu\right)^{\frac{1}{q}} \\
&= \frac{1}{n}\sum_{j=0}^{n-1} \|U_T^j(\phi_k - \phi)\|_p \cdot \|\psi\|_q \\
&= \|\phi_k - \phi\|_p \cdot \|\psi\|_q < \frac{\varepsilon}{3} \qquad (3.3)
\end{aligned}$$

para todo k suficientemente, veja que a desigualdade acima não depende de n. Fixe k de tal forma que valham as desigualdades (3.2) e (3.3) e tome n suficientemente grande de modo que valha (3.1), veja que agora n depende de k. Daí, temos que

$$\begin{aligned}
\left|\int\left(\frac{1}{n}\sum_{j=0}^{n-1}U_T^j\phi\right)\psi\, d\mu - \int\phi\, d\mu\int\psi\, d\mu\right| &\leqslant \left|\int\left(\frac{1}{n}\sum_{j=0}^{n-1}U_T^j\phi_k\right)\psi\, d\mu - \int\phi_k\, d\mu\int\psi\, d\mu\right| \\
&\quad + \left|\int(\phi_k - \phi)\, d\mu\int\psi\, d\mu\right| \\
&\quad + \left|\int\frac{1}{n}\sum_{j=0}^{n-1}U_T^j(\phi - \phi_k)\psi\, d\mu\right| \\
&< \varepsilon,
\end{aligned}$$

para todo n suficientemente grande. Desta forma, vale (c) e a proposição está demonstrada. ■

3.2 Deslocamentos de Bernoulli.

Nesta seção estudaremos um exemplo de sistema ergódico bastante interessante. Considere o espaço produto $\Sigma = X^{\mathbb{N}}$, munido da σ-álgebra produto $\mathcal{B} = \mathcal{C}^{\mathbb{N}}$ e da medida produto $\mu = \nu^{\mathbb{N}}$. Desta forma, Σ é o conjunto de todas as sequências $(x_n)_{n\in\mathbb{N}}$ onde $x_n \in X$ para todo $n \in \mathbb{N}$. Temos também, que $\mathcal{B}$ é a σ-ágebra gerada pelos cilindros

$$[m; A_m, ..., A_n] = \{(x_i)_{i\in\mathbb{N}}; x_i \in A_i \text{ para } m \leqslant i \leqslant n\}$$

com $m \leqslant n$ e $A_i \in \mathcal{C}$. Além disso, μ é caracterizada por

$$\mu([m; A_m, ..., A_n]) = \prod_{i=m}^{n} \nu(A_i).$$

Definição 3.5. *O deslocamento de Bernoulli é a aplicação $\sigma : \Sigma \to \Sigma$ definida por*

$$\sigma((x_n)_n) = (x_{n+1})_n$$

ou seja σ envia a sequência $(x_0, x_1, ..., x_n, ...)$ na sequência $(x_1, ..., x_n, ...)$.

Proposição 3.3. *A medida produto é invariante para σ.*

Demonstração: Primeiro, note que $\sigma^{-1}([m; A_m, ..., A_n]) = [m + 1; A_m, ..., A_n]$. De fato, veja que $(x_0, x_1, ..., x_m, x_{m+1}, ...) \in \sigma^{-1}([m; A_m, ..., A_n])$ se, e somente se,

$$\sigma((x_0, x_1, ..., x_m, x_{m+1}, ...)) \in [m; A_m, ..., A_n],$$

ou seja, $(x_1, ..., x_m, x_{m+1}, ...) \in [m; A_m, ..., A_n]$. Observe que $x_{m+1} \in A_m, x_{m+2} \in A_{m+1}, ..., x_{n+1} \in A_n$ assim $(x_0, x_1, ..., x_m, x_{m+1}, ...) \in [m + 1; A_m, ..., A_n]$. Daí,

$$\begin{aligned}
\mu(\sigma^{-1}([m; A_m, ..., A_n])) &= \mu([m + 1; A_m, ..., A_n]) \\
&= \nu(A_m) \cdot ... \cdot \nu(A_n) \\
&= \mu([m; A_m, ..., A_n]).
\end{aligned}$$

Segue do Lema 1.2 que μ é invariante para σ. ■

Agora provaremos um lema que nos auxiliará na demonstração de que o deslocamento de Bernoulli (σ, μ) é ergódico.

Lema 3.1. *Se B e C são uniões finitas de cilindros dois a dois disjuntos, então tem-se*

$$\mu(B \cap \sigma^{-j}(C)) = \mu(B)\mu(\sigma^{-j}(c)) = \mu(B)\mu(C)$$

para j suficientemente grande.

Demonstração: Dividiremos a demonstração em dois casos:

Caso 1: Suponha que B e C são ambos cilindros, ou seja $B = [k : B_k, ..., B_l]$ e $C = [m; C_m, ..., C_n]$.

Analogamente ao que foi feito na proposição anterior, para cada j, temos

$$\sigma^{-j}(C) = [m + j; C_m, ..., C_n].$$

Agora considere j suficientemente grande tal que $m + j > l$. Assim, temos

$$\begin{aligned} B \cap \sigma^{-j}(C) \quad &= \{(x_n)_n; x_k \in B_k, ..., x_l \in B_l, x_{m+j} \in C_m, ..., x_{n+j} \in C_n\} \\ &= [k; B_k, ..., B_l, X, ..., X, C_m, ..., C_n] \end{aligned}$$

onde X aparece $m + j - (l + 1)$ vezes. Daí,

$$\begin{aligned} \mu(B \cap \sigma^{-j}(C)) \quad &= \mu([k; B_k, ..., B_l, X, ..., X, C_m, ..., C_n]) \\ &= \nu(B_k) \cdot ... \cdot \nu(B_l) \cdot \nu(X) \cdot ... \cdot \nu(X) \cdot \nu(C_m) \cdot ... \cdot \nu(C_n) \\ &= \nu(B_k) \cdot ... \cdot \nu(B_l) \cdot \nu(C_m) \cdot ... \cdot \nu(C_n) \\ &= \mu(B) \cdot \mu(C). \end{aligned}$$

Caso 2: B e C são uniões finitas de cilindros disjuntos dois a dois, ou seja,

$$B = \sum_{i=1}^{m} B_i \text{ e } C = \sum_{k=1}^{n} C_k,$$

onde cada B_i e cada C_k são cilindros.

Note que

$$\sigma^{-j}(C) = \sigma^{-j}\left(\sum_{k=1}^{n} C_k\right) = \sum_{k=1}^{n} \sigma^{-j}(C_k).$$

O Caso 2 segue do Caso 1 e do fato de μ ser finitamente aditiva. ∎

Proposição 3.4. *Todo deslocamento de Bernoulli* (σ, μ) *é ergódico.*

Demonstração: Basta mostrar que para todo conjunto invariante A, tem-se $\mu(A) = 0$ ou $\mu(A) = 1$. Suponhamos inicialmente que o conjunto invariante A é uma união finita de cilindros disjuntos. Pelo lema anterior, temos que $\mu(A \cap \sigma^{-j}(A)) = (\mu(A))^2$ para j suficientemente grande, e como A é invariante temos $\mu(A \cap \sigma^{-j}(A)) = \mu(A)$. Daí, $\mu(A) = (\mu(A))^2$ logo $\mu(A) = 0$ ou $\mu(A) = 1$.

Para o caso geral, considere a álgebra $\mathcal{B}_0$ das uniões finitas de cilindros disjuntos, note que a álgebra $\mathcal{B}_0$ gera a σ-álgebra $\mathcal{B}$. Seja A um conjunto invariante mensurável qualquer.

Dado $\varepsilon > 0$, pelo Teorema 1.3 existe $B \in \mathcal{B}_0$ tal que $\mu(A\Delta B) < \dfrac{\varepsilon}{4}$. Como $B \in \mathcal{B}_0$, usando o lema anterior, podemos tomar j tal que

$$\mu(B \cap \sigma^{-j}(B)) = \mu(B) \cdot \mu(\sigma^{-j}(B)) = (\mu(B))^2.$$

Afirmo que

$$(A \cap \sigma^{-j}(A))\Delta(B \cap \sigma^{-j}(B)) \subset (A\Delta B) \cup (\sigma^{-j}(A)\Delta\sigma^{-j}(B)) = (A\Delta B) \cup (\sigma^{-j}(A\Delta B)).$$

De fato, se $x \in (A \cap \sigma^{-j}(A))\Delta(B \cap \sigma^{-j}(B))$ então $x \in (A \cap \sigma^{-j}(A)) \setminus (B \cap \sigma^{-j}(B))$ ou $x \in (B \cap \sigma^{-j}(B)) \setminus (A \cap \sigma^{-j}(A))$. Daí, temos as seguintes possibilidades:

1. Se $x \in (A\cap\sigma^{-j}(A))\setminus(B\cap\sigma^{-j}(B))$ então $x \in (A\cap\sigma^{-j}(A))$ e $x \notin B$ ou $x \in (A\cap\sigma^{-j}(A))$ e $x \notin \sigma^{-j}(B)$. Logo, $x \in (A\Delta B)$ ou $x \in (\sigma^{-j}(A)\Delta\sigma^{-j}(B))$.

2. Se $x \in (B \cap \sigma^{-j}(B)) \setminus (A \cap \sigma^{-j}(A))$, a argumentação é análoga.

Veja que se C e D são conjuntos mensuráveis então

$$\begin{aligned}
|\mu(C) - \mu(D)| &= |\mu(C) - \mu(C \cap D) - (\mu(D) - \mu(C \cap D))| \\
&= |\mu(C \setminus D) - \mu(D \setminus C)| \\
&\leqslant \mu(C \setminus D) + \mu(D \setminus C) \\
&= \mu(C \setminus D \cup D \setminus C) \\
&= \mu(C\Delta D).
\end{aligned}$$

Daí, substituindo $C = A \cap \sigma^{-j}(A)$ e $D = B \cap \sigma^{-j}(B)$, temos que

$$\begin{aligned}
|\mu(A \cap \sigma^{-j}(A)) - \mu(B \cap \sigma^{-j}(B))| &\leqslant \mu((A \cap \sigma^{-j}(A))\Delta(B \cap \sigma^{-j}(B))) \\
&\leqslant \mu((A\Delta B) \cup (\sigma^{-j}(A\Delta B))) \\
&\leqslant \mu((A\Delta B)) + \mu((\sigma^{-j}(A\Delta B))) \\
&= 2\mu((A\Delta B)) \\
&< \frac{\varepsilon}{2}.
\end{aligned}$$

Por outro lado, temos que

$$|\mu(A)^2 - \mu(B)^2| = |\mu(A) + \mu(B)| \cdot |\mu(A) - \mu(B)| \leqslant 2 \cdot |\mu(A) - \mu(B)| < \frac{\varepsilon}{2}.$$

Como A é invariante temos que $\mu(A \cap \sigma^{-j}(A)) = \mu(A)$. Disto segue que

$$\begin{aligned} |\mu(A)^2 - \mu(A)| &\leqslant |\mu(A)^2 - \mu(B)^2| + |\mu(B)^2 - \mu(A)| \\ &< \frac{\varepsilon}{2} + |\mu(B \cap \sigma^{-j}(B)) - \mu(A \cap \sigma^{-j}(A))| \\ &< \varepsilon. \end{aligned}$$

Fazendo $\varepsilon \to 0^+$, temos que $\mu(A)^2 = \mu(A)$. Portanto, $\mu(A) = 0$ ou $\mu(A) = 1$.

$\blacksquare$

Capítulo 4

Aplicações

Neste capítulo, apresentaremos alguns resultados cujas demonstrações tem o conceito de ergodicidade e os teoremas ergódicos como elementos fundamentais.

4.1 Sequências quase-normais

Neste tópico apresentaremos alguns resultados obtidos em [13]. Neste trabalho, os autores estudaram condições necessárias e suficientes para a convergência forte do método de projeções alternadas em espaços de Hadamard. Em particular, eles introduziram o conceito de sequências quase-normais e utilizaram o teorema ergódico de Birkhoff para mostrar que o conjunto das sequências quase normais possui medida total com respeito a medida de Bernoulli. Esta abordagem utilizando técnicas de Teoria Ergódica os permitiu responder um problema em aberto relacionado ao método das projeções alternadas, que foi introduzido por von Neumann em 1933 para resolver o problema de viabilidade convexa, ou seja, encontrar um ponto na interseção de conjuntos convexos.

Seja $N \geqslant 2$ um inteiro e suponha que $C_1, ..., C_N$ são conjuntos convexos e fechados de um espaço de Hadamard $\mathcal{H}$ cuja interseção é não-vazia. Denotaremos por P_j a projeção sobre C_j. Agora considere uma sequência (j_n) tomando valores em $I_N = \{1, 2, ..., N\}$, escolha um ponto $x_0 \in \mathcal{H}$ e defina a sequência (x_n) por

$$x_n = P_{j_n}(x_{n-1}),$$

para todo $n \in \mathbb{N}$. Ao longo desta seção, denotaremos por (j_n) uma sequência tomando valores em $I_N = \{1, 2, \ldots, N\}$ e por Σ o conjunto de todas essas sequências, ou seja, $\Sigma = \{1, 2, ..., N\}^{\mathbb{N}}$.

Definição 4.1. *Para cada* $j \in \{1, ..., N\}$*, considere a sequência crescente de números naturais* (k_n) *tal que* $j_{k_n} = j$*. Dizemos que uma sequência* (j_n) *é* ***quase-periódica*** *se* $I(j) = \sup_n(k_{n+1} - k_n)$ *é finito para todo* j*.*

Definição 4.2. *Dizemos que uma sequência* (j_n) *é* ***quase-normal*** *se existe um número positivo* L *e uma sequência de blocos disjuntos* $(\mathcal{R}_k)$ *de elementos consecutivos de* (j_n) *com* L *termos, onde cada bloco* $\mathcal{R}_k$ *possui todos os índices de* $\{1, ..., N\}$ *e uma função* $f : \mathbb{N} \to (0, \infty)$ *com* $\lim\limits_{n_k \to \infty} f(n_k) = +\infty$ *tal que*

$$\sum_{k=1}^{\infty} \frac{1}{n_k \cdot f(n_k)} = +\infty,$$

onde j_{n_k} *é o primeiro elemento do bloco* $(\mathcal{R}_k)$ *e* (n_k) *é uma sequência crescente.*

Proposição 4.1. *Suponha que existam um número natural* L *e uma sequência de blocos disjuntos* $(\mathcal{R}_k)$ *de elementos consecutivos de* (j_n) *com* L *termos, onde cada bloco* $\mathcal{R}_k$ *possui todos os elementos de* $\{1, ..., N\}$*. Para cada* $k \in \mathbb{N}$*, seja* $\mathcal{S}_k$ *o bloco formado pelos elementos entre* $\mathcal{R}_{k-1}$ *e* $\mathcal{R}_k$*, que eventualmente pode ser vazio. Assim, a sequência* (j_n) *pode ser vista da forma:*

$$\mathcal{S}_1\mathcal{R}_1\mathcal{S}_2\mathcal{R}_2, ..., \mathcal{S}_k\mathcal{R}_k, ...$$

Seja $|\mathcal{S}_k|$ *o número de elementos do bloco* $\mathcal{S}_k$ *e* $c > 0$ *uma constante. Se*

$$\sum_{i=1}^{k} |\mathcal{S}_i| \leqslant c \cdot k, \text{ para todo } k \in \mathbb{N}, \tag{4.1}$$

então a sequência (j_n) *é quase-normal.*

Demonstração: Veja [13]. ∎

Corolário 4.1. *Toda sequência quase-periódica é quase-normal.*

Demonstração: Veja [13]. ∎

Exemplo 4.1. *Considere a sequência abaixo, onde temos blocos consecutivos* $\mathcal{R}_k = \{123\}$ *intercalados por blocos da forma* $S = \{1212...12\}$*.*

$$\underbrace{123123\cdots123}_{10 \text{ blocos}}12\underbrace{123123\cdots123}_{10^2 \text{ blocos}}1212\underbrace{123123\cdots123}_{10^3 \text{ blocos}}121212\underbrace{123123\cdots123}_{10^4 \text{ blocos}}\cdots$$

Para cada $l \in \mathbb{N}$, *seja* $k_l = \dfrac{10^{l+1}-1}{9}$. *Usando a notação da Proposição 4.1, note que* $\mathcal{S}_i \neq \emptyset$ *se, e somente se,* $i = k_l$ *para algum* $l \in \mathbb{N}$ *e além disso* $|\mathcal{S}_{k_l}| = 2l$. *Dado* $k \in \mathbb{N}$ *tal que* $k_l \leqslant k < k_{l+1}$, *temos*

$$\sum_{i=1}^{k} |\mathcal{S}_i| = \sum_{j=1}^{l} |\mathcal{S}_{k_j}| = \sum_{j=1}^{l} 2j = l(l+1) \leqslant \frac{10^{l+1}-1}{9} \leqslant k.$$

Pela Proposição 4.1, a sequência (j_n) *é quase-normal. Por outro lado, esta sequência não é quase-periódica, pois para cada* $n \in \mathbb{N}$ *temos blocos* $1212\cdots12$ *com* $2n$ *elementos distintos de 3.*

Sejam $\mathcal{A}$ o conjunto das sequências quase-normais, $\mathcal{Q}$ o conjunto das sequências quase-periódicas e $\mathcal{B}$ o conjunto das sequências (j_n) que satisfazem (4.1). Dos resultados acima segue que $\mathcal{Q} \subset \mathcal{B} \subset \mathcal{A}$.

Utilizando o teorema ergódico de Birkhoff e a ergodicidade do deslocamento de Bernoulli com respeito a medida de Bernoulli foi provado em [13] que os conjuntos $\mathcal{A}$ e $\mathcal{B}$ possuem medida total e que o conjunto $\mathcal{Q}$ possui medida nula. A seguir apresentamos a demonstração destes resultados.

Proposição 4.2. $\mathcal{P}(\mathcal{A}) = 1$, *onde* $\mathcal{P}$ *é a medida de Bernoulli em* Σ.

Demonstração: Sejam $\mathcal{P}$ a medida de Bernoulli em Σ e $\beta : \Sigma \to \Sigma$ o deslocamento de Bernoulli. Do teorema ergódico de Birkhoff e da ergodicidade da medida de Bernoulli, existe um subconjunto $\Omega \subset \Sigma$ com $\mathcal{P}(\Omega) = 1$ tal que para cada $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n, ...) \in \Omega$, tem-se

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_E(\beta^j(\alpha)) = \mathcal{P}(E), \tag{4.2}$$

onde E é o cilindro $[1; 1, 2, ..., N]$ e $\mathcal{P}(E) = p_1 \cdot p_2 \cdot ... \cdot p_N < 1$. Pela desigualdade das médias aritmética e geométrica segue que

$$0 < a = p_1 \cdot p_2 \cdot ... \cdot p_N \leqslant \left(\frac{p_1 + p_2 + ... + p_N}{N}\right)^N = \frac{1}{N^N}.$$

Para $\alpha \in \Omega$, observe que $\beta^{j-1}(\alpha) \in [1; 1, 2, ..., N]$ se, e somente se, $\alpha_j = 1, \alpha_{j+1} = 2, ..., \alpha_{j+n-1} = N$. Considere a sequência crescente (n_k), onde $\beta^{n_k-1}(\alpha) \in [1; 1, 2, ..., N]$ e a sequência de blocos $(\mathcal{R}_k)$ com N termos cujo primeiro elemento é j_{n_k}. Segue de (4.2), que existe m tal que se $n > m$ então

$$\frac{a}{2} < \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_E(\beta^j(\alpha)) < \frac{3a}{2}. \tag{4.3}$$

Observe que se $j = n_l - 1$ para algum l, então $\mathcal{X}_E(\beta^j(\alpha)) = 1$ e que $\mathcal{X}_E(\beta^j(\alpha)) = 0$ caso contrário. Portanto,

$$\sum_{j=0}^{n_k-1} \mathcal{X}_E(\beta^j(\alpha)) = k.$$

Tome $n_{k_1} > m$. De (4.3) segue que para $k \geqslant k_1$,

$$\frac{a}{2} < \frac{k}{n_k} < \frac{3a}{2}.$$

Agora, considere a função $f : \mathbb{N} \to (0, \infty)$, onde $f(n) = \log\left(n + \frac{1}{a}\right)$. Tem-se

$$\begin{aligned}\sum_{k=1}^{\infty} \frac{1}{n_k \cdot f(n_k)} &= \sum_{k=1}^{k_l-1} \frac{1}{n_k \cdot f(n_k)} + \sum_{k=k_l}^{\infty} \frac{1}{n_k \cdot f(n_k)} \\ &\geqslant \sum_{k=1}^{k_l-1} \frac{1}{n_k \cdot f(n_k)} + \sum_{k=k_l}^{\infty} \frac{a}{2k\log(2k+1) - 2k\log a}. \\ &= \infty.\end{aligned}$$

Isto prova que $\Omega \subset \mathcal{A}$. Em particular, $\mathcal{P}(\mathcal{A}) = 1$. ■

Agora, mostraremos que o conjunto $\mathcal{B}$ das sequências (j_n) que satisfazem (4.1) também tem medida total.

Proposição 4.3. $\mathcal{P}(\mathcal{B}) = 1$, *onde $\mathcal{P}$ é a medida de Bernoulli em Σ.*

Demonstração: Considere o cilindro $E = [1; 1, 2, ..., N]$, cuja medida de Bernoulli é

$$\mathcal{P}([1; 1, 2, ..., N]) = p_1 \cdot p_2 \cdot ... \cdot p_N = a > 0.$$

Como a medida de Bernoulli é ergódica em relação ao deslocamento de Bernoulli $\beta : \Sigma \to \Sigma$ segue do teorema ergódico de Birkhoff que existe um subconjunto $\Omega \subset \Sigma$ com $\mathcal{P}(\Omega) = 1$ tal que se $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n, ...) \in \Omega$, então

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_E(\beta^j(\alpha)) = \mathcal{P}(E). \tag{4.4}$$

Em particular, existe uma uma sequência crescente (n_k), onde $\beta^{n_k-1}(\alpha) \in [1; 1, 2, ..., N]$ para todo $k \in \mathbb{N}$. Por definição do deslocamento de Bernoulli temos

$$\mathcal{R}_k := (\alpha_{n_k}, \alpha_{n_k+1}, ..., \alpha_{n_k+N-1}) = (1, 2, ..., N).$$

Segue de (4.4) que

$$a = \lim_{n_k\to\infty} \frac{1}{n_k} \# \left\{1 \leqslant j \leqslant n_k : \beta^{j-1}(\alpha) \in [1; 1, 2, ..., N]\right\} = \lim_{k\to\infty} \frac{k}{n_k}.$$

Tome $k_0 \in \mathbb{N}$ tal que $k > k_0$ implica

$$\frac{a}{2} < \frac{k}{n_k} < \frac{3a}{2}.$$

Sendo $n_k = 1 + (k-1)L + \sum_{i=1}^{k} |\mathcal{S}_i|$, para $k > k_0$, temos

$$\begin{aligned}
\frac{a}{2} < \frac{k}{n_k} < \frac{3a}{2} \quad & \Rightarrow \frac{2}{3a} < \frac{n_k}{k} < \frac{2}{3a} \\
& \Rightarrow n_k < \frac{2k}{3a} \\
& \Rightarrow 1 + (k-1)L + \sum_{i=1}^{k} |\mathcal{S}_i| < \frac{2k}{3a} \\
& \Rightarrow \frac{Lk}{2} + \sum_{i=1}^{k} |\mathcal{S}_i| < \frac{2k}{3a} \\
& \Rightarrow \sum_{i=1}^{k} |\mathcal{S}_i| < \left(\frac{2}{3a} - \frac{L}{2}\right) k.
\end{aligned}$$

Tomando

$$c = \max\left\{\sum_{i=1}^{k_0} |\mathcal{S}_i|, \frac{2}{3a} - \frac{L}{2}\right\},$$

obtemos

$$\sum_{i=1}^{k} |\mathcal{S}_i| \leqslant c \cdot k$$

para todo $k \in \mathbb{N}$. Como $\Omega \subset \mathcal{B}$, então $\mathcal{P}(\mathcal{B}) = 1$. ∎

Proposição 4.4. $\mathcal{P}(\mathcal{Q}) = 0$*, onde* $\mathcal{P}$ *é a medida de Bernoulli em* Σ*.*

Demonstração: Denote por Q_k o conjunto das sequências quase-periódicas (j_n) tal que $\max_{1 \leqslant j \leqslant N} I(j) \leqslant k$. Vamos mostrar que $\mathcal{P}(Q_k) = 0$. De fato, considere o cilindro $E_k = [1; 1, 1, ..., 1]$ (onde 1 se repete 2k-vezes), como visto anteriormente, pelo teorema de Birkhoff e a ergodicidade da medida $\mathcal{P}$, existe um subconjunto $\Omega_k \subset \Sigma$ com $\mathcal{P}(\Omega_k) = 1$ tal que se $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n, ...) \in \Omega_k$, então

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_{E_k}(\beta^j(\alpha)) = \mathcal{P}(E_k) > 0.$$

Em particular, se $\alpha \in \Omega_k$, então existe j tal que $\beta^{j-1}(\alpha) \in E_k$ e $\alpha_j = \alpha_{j+1} = \alpha_{j+2k-1} = 1$. Isto implica que $\alpha \notin Q_k$. Portanto, $\Omega_k \cap Q_k$ é vazio, assim $\mathcal{P}(Q_k) = 0$. Como $\mathcal{Q} = \bigcup_{k=1}^{\infty} Q_k$ segue que

$$0 \leqslant \mathcal{P}(\mathcal{Q}) \leqslant \sum_{k=1}^{\infty} \mathcal{P}(Q_k) = 0.$$

Isto completa a demonstração. ∎

4.2 Teorema de Weyl

Neste tópico apresentaremos um teorema de Hermann Weyl provado em [20] que trata da distribuição de valores de polinômios aplicados no conjunto dos números inteiros. Seguiremos a demonstração apresentada em [18]. Seja $P : \mathbb{R} \to \mathbb{R}$ uma função polinomial qualquer de coeficientes reais de grau $d \geqslant 1$:

$$P(x) = a_0 + a_1 x + a_2 x^2 + ... + a_d x^d.$$

Seja $P_* : \mathbb{R} \to S^1$ a composição de P com a projeção canônica $\mathbb{R} \to S^1 = \mathbb{R}/\mathbb{Z}$. Defina, para todo $n \geqslant 1$,

$$z_n = P_*(n).$$

Assim podemos pensar em z_n como sendo a parte fracionária de $P(n)$.

Definição 4.3. *Dizemos que uma sequência (x_n) em S^1 é equidistribuída se para qualquer função contínua $\varphi : S^1 \to \mathbb{R}$ tem-se*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \varphi(x_j) = \int \varphi(x)\, dx.$$

Isto equivale a dizer que, para todo segmento $I \subset S^1$, a fração dos termos da sequência que estão em I é igual ao comprimento $m(I)$ do segmento, onde m é a medida de lebesgue.

Teorema 4.1. *(Weyl) Se algum dos coeficientes $a_1, ..., a_d$ é irracional, então a sequência $z_n = P_*(n), n \in \mathbb{N}$ é equidistribuída.*

Para demonstrar esse teorema, faremos alguns preparativos inicialmente. Considere a transformação $f : \mathbb{T}^d \to \mathbb{T}^d$ definida no toro d-dimensional $\mathbb{T}^d$ da seguinte maneira,

$$f(\theta_1, \theta_2, ..., \theta_d) = (\theta_1 + \alpha, \theta_2 + \theta_1, ..., \theta_d + \theta_{d-1})$$

onde α é um número irracional. Note que f é invertível, e sua inversa é dada por:

$$f^{-1}(\theta_1, \theta_2, ..., \theta_d) = (\theta_1 - \alpha, \theta_2 - \theta_1 + \alpha, ..., \theta_d - \theta_{d-1} + ... + (-1)^{d-1}\theta_1 + (-1)^d \alpha).$$

Temos também que o determinante da matriz derivada de f é 1, e isso garante que f preserva a medida de Lebesgue no toro.

Proposição 4.5. *A medida de Lebesgue em $\mathbb{T}^d$ é ergódica para f.*

Demonstração: Iremos usar um argumento de expansão em Série de Fourier. Seja $\varphi : \mathbb{T}^d \to \mathbb{R}$ uma função em $L^2(m)$. Escrevemos

$$\varphi(\theta) = \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i n \cdot \theta}$$

onde $\theta = (\theta_1, \theta_2, ..., \theta_d)$, $n = (n_1, ..., n_d)$ e $n \cdot \theta = n_1\theta_1 + ... + n_d\theta_d$. A norma de φ em $L^2(m)$ é dada por

$$\sum_{n \in \mathbb{Z}^d} |a_n|^2 = \int |\varphi(\theta)|^2 d\theta_1 \cdots d\theta_d < \infty. \tag{4.5}$$

Observe que

$$\begin{aligned} \varphi(f(\theta)) \quad &= \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i (n_1(\theta_1 + \alpha) + ... + n_d(\theta_d + \theta d - 1))} \\ &= \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i n_1 \alpha} e^{2\pi i L(n) \cdot \theta}, \end{aligned}$$

onde $L(n) = (n_1 + n_2, n_2 + n_3, ..., n_{d-1} + n_d, n_d)$. Suponha que φ é invariante, ou seja, que $\varphi \circ f = \varphi$ me quase todo ponto. Teremos então

$$a_n e^{2\pi i n_1 \alpha} = a_{L(n)} \tag{4.6}$$

para todo $n \in \mathbb{Z}^d$. Isto implica que a_n e $a_{L(n)}$ têm o mesmo valor absoluto. Por outro lado, pela relação (4.5) temos que existe no máximo um número finito de termos com um dado valor absoluto não-nulo. Concluímos que $a_n = 0$ para todo $n \in \mathbb{Z}^d$ cuja órbita $L^j(n), j \in \mathbb{Z}$ seja infinita. Analisando a expressão de L temos que $a_n = 0$ exceto, possivelmente, se $n_2 = ... = n_d = 0$. Por outro lado, para $n = (n_1, 0, ..., 0)$, tem- se que $L(n) = n$ e , portanto, a relação (4.6) torna-se

$$a_n = a_n e^{2\pi i n_1 \alpha}.$$

Como α é irracional, o último é diferente de 1 sempre que n_1 é não-nulo. Portanto esta relação dá que $a_n = 0$ também para $n = (n_1, 0, ..., 0)$ com $n_1 \neq 0$. Deste modo, mostramos que se φ é uma função invariante, então todos os termos de sua expansão em série de Fourier se anulam exceto, possivelmente o termo constante. Isto mostra que φ é constante em quase todo ponto, e isso prova que a medida de Lebesgue é ergódica para f. ∎

Definição 4.4. *Dado $f : X \to X$ e qualquer medida μ em X denota-se por $f_*\mu$ e chama-se iterado (ou imagem) de μ por f a medida definida por $f_*\mu(B) = \mu(f^{-1}(B))$ para cada conjunto mensurável $B \subset X$. Note que μ é invariante pela transformação f se, e somente se, $f_*\mu = \mu$.*

Definição 4.5. *Dizemos que uma transformação $f : X \to X$ é unicamente ergódica se admite exatamente uma medida de probabilidade invariante.*

O próximo passo para a demonstração do teorema de Weyl é o seguinte resultado:

Proposição 4.6. *A transformação f é unicamente ergódica: a medida de Lebesgue no toro é a sua única probabilidade invariante.*

Demonstração: A demonstração será por indução sobre o grau d do polinômio P. Para $d = 1$, a função polinomial resume-se a $P(x) = a_0 + a_1 x$ e a transformação f é dada por

$$f : S^1 \to S^1, \qquad f(\theta) = \theta + a_1$$

Por hipótese, o coeficiente a_1 é irracional, assim, esta transformação admite uma única probabilidade invariante, que é a medida de Lebesgue m. Portanto o caso $d = 1$ está feito, agora só precisamos explicar como o caso de grau d pode ser deduzido do caso de grau $d - 1$. Para isso, escrevemos $\mathbb{T}^d = \mathbb{T}^{d-1} \times S^1$ e

$$f : \mathbb{T}^{d-1} \times S^1 \to \mathbb{T}^{d-1} \times S^1, \qquad f(\theta_0, \eta) = (f_0(\theta_0), \eta + \theta_{d-1}),$$

onde $\theta_0 = (\theta_1, ..., \theta_{d-1})$ e $f_0(\theta_0) = (\theta_1 + \alpha, \theta_2 + \theta_1, ..., \theta_{d-1} + \theta_{d-2})$. Por indução, a transformação

$$f_0 : \mathbb{T}^{d-1} \to \mathbb{T}^{d-1}$$

é unicamente ergódica. Representamos por $\pi : \mathbb{T}^d \to \mathbb{T}^{d-1}$ a projeção $\pi(\theta) = \theta_0$.

Lema 4.1. *Se μ é uma probabilidade invariante por f, então a projeção $\pi_*\mu$ coincide com a medida de Lebesgue m_0 em $\mathbb{T}^{d-1}$.*

Demonstração: Veja [18]. ■

Agora suponhamos que μ, além de invariante, também é ergódica para f. Por ergodicidade e pelo corolário 2.1 do Teorema Ergódico de Birkhoff, o conjunto $G(\mu)$ dos pontos $\theta \in \mathbb{T}^d$ tais que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(\theta)) = \int \varphi \, d\mu \text{ para toda função contínua } \varphi : \mathbb{T}^d \to \mathbb{R} \tag{4.7}$$

tem medida total. Seja $G_0(\mu)$ o conjunto dos $\theta_0 \in \mathbb{T}^{d-1}$ tais que $G(\mu)$ intersecta $\{\theta_0\}\times S^1$, ou seja $G_0(\mu) = \pi(G(\mu))$. É claro que $\pi^{-1}(G_0(\mu))$ contém $G(\mu)$ e portanto, tem medida total. Logo pelo lema 4.1,

$$m_0(G_0(\mu)) = \mu(\pi^{-1}(G_0(\mu))) = 1. \tag{4.8}$$

Pelas mesmas razões, esta relação também vale para a medida de Lebesgue:

$$m_0(G_0(m)) = m(\pi^{-1}(G_0(m))) = 1 \tag{4.9}$$

Uma consequência direta das igualdades (4.8) e (4.9) é que a intersecção de $G_0(\mu)$ e $G_0(m)$ tem medida m_0 total. Logo, em particular, estes conjuntos não podem ser disjuntos. Seja θ_0 um ponto qualquer na intersecção. Por definição $G(\mu)$ intersecta $\{\theta_0\}\times S^1$, mas temos o seguinte resultado:

Lema 4.2. *Se* $\theta_0 \in G_0(m)$*, então* $\{\theta_0\}\times S^1$ *está contido em* $G(m)$.

Demonstração: Veja [18]. ∎

Segue do que foi dito até agora que $G(\mu)$ e $G(m)$ se intersectam em algum ponto de $\{\theta_0\}\times S^1$. Tendo em vista a relação (4.7), isto implica que as duas medidas têm a mesma integral para cada função contínua. Segue do fato de que se duas probabilidades em M são tais que para toda função $\varphi : M \to \mathbb{R}$ Lipschitz limitada a integral com respeito ambas probabilidades coincidem então estas probabilidades são iguais, que $\mu = m$ (veja mais em [18]), como queríamos demonstrar.

∎

Corolário 4.2. *A órbita de todo ponto* $\theta \in \mathbb{T}^d$ *é equidistribuida no toro* $\mathbb{T}^d$*, ou seja, para toda função contínua* $\psi : \mathbb{T}^d \to \mathbb{R}$ *tem-se*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \psi(f^j(\theta)) = \int \psi \, dm.$$

Demonstração: Segue diretamente das Proposições 4.5 e 4.6. ∎

Para completar a demonstração do teorema de Weyl, introduzimos as funções polinomiais $p_1, ..., p_d$ definidas por $p_d(x) = P(x)$ e $p_{j-1}(x) = p_j(x+1) - p_j(x)$ para $j = 2, ..., d$.

Lema 4.3. *O polinômio* $p_j(x)$ *tem grau* j*, para todo* $1 \leqslant j \leqslant d$*. Além disso,* $p_1(x) = \alpha x + \beta$*, com* $\alpha = d!a_d$.

Demonstração: Ver [18]. ∎

Lema 4.4. *Para todo* $n \geqslant 0$,

$$f^n(p_1(0), p_2(0), ..., p_d(0)) = (p_1(n), p_2(n), ..., p_d(n)).$$

Demonstração: Veja [18]. ∎

Finalmente, estamos prontos para provar que a sequência $z_n = P_*(n), n \in \mathbb{N}$ é equidistribuida. Vamos tratar dois casos separadamente.

Suponha que a_d é irracional. Então o número α no Lema 4.3 é irracional e, portanto, os resultados da Proposição acima são válidos para a transformação $f : \mathbb{T}^d \to \mathbb{T}^d$. Seja $\varphi : S^1 \to \mathbb{R}$ uma função contínua qualquer. Considere $\psi : \mathbb{T}^d \to \mathbb{R}$ definida por

$$\psi(\theta_1, ..., \theta_d) = \varphi(\theta_d).$$

Fixemos $\theta = (p_1(0), p_2(0), ..., p_d(0))$, usando o Lema 4.4 e o Corolário 4.2, obtemos que

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(z_j) = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \psi(f^j(\theta)) = \int \psi \, dm = \int \varphi \, dx.$$

Isto termina a demonstração do teorema de Weyl no caso em que a_d é irracional.

Agora, suponha que a_d é racional, digamos $a_d = p/q$ com $p \in \mathbb{Z}$ e $q \in \mathbb{N}$. Podemos escrever z_n como uma soma

$$z_n = x_n + y_n, \qquad x_n = a_d n^d \qquad \text{e} \qquad y_n = Q_*(n)$$

onde $Q(x) = a_0 + a_1 x + ... + a_{d-1}x^{d-1}$ e $Q_* : \mathbb{R} \to S^1$ é dada por $Q_* = \pi \circ Q$. Observe, em primeiro lugar, que

$$x_{n+q} - x_n = \frac{p}{q}(n+q)^d - \frac{p}{q}n^d$$

é um número inteiro, para todo $n \in \mathbb{N}$. Isto significa que a sequência x_n é periódica de período q no círculo $\mathbb{R}/Z$. Em particular, ela toma q valores distintos. Observe também que, como a_d é racional, a hipótese do teorema implica que algum dos coeficientes $a_1, a_2, ..., a_{d-1}$ de Q é irracional. Logo por indução no grau, temos que y_n é equidistribuida. Mais do que isso, as subsequências

$$y_{(qn+r)} = Q_*(qn+r), \qquad n \in \mathbb{Z}$$

são equidistribuidas para todo $r \in \{0, 1, ..., q-1\}$. Na verdade, estas subsequências podem ser escritas como $y_{nq+r} = Q_*^{(r)}(n)$ para algum polinômio $Q^{(r)}$ que também tem grau $(d-1)$, e portanto a hipótese de indução se aplica sobre elas também. Destas duas observações segue que toda subsequência z_{qn+r} é equidistribuída. Consequentemente, z_n, $n \in \mathbb{Z}$ também é equidistribuida. Isto completa a demonstração do teorema.

Referências Bibliográficas

[1] AVILA, Artur; BOCHI, Jairo. On the subadditive ergodic theorem. preprint, 2009.

[2] BARREIRA, Luis. Ergodic theory, hyperbolic dynamics and dimension theory. dimension, v. 37, n. 37DXX, p. 37C45, 2012.

[3] BERGELSON, Vitaly. Some Historical Remarks and Modern Questions Around the Ergodic Theorem (Dynamics of Complex Systems). v. 1404, p. 1-11, 2004.

[4] BOTELHO, Geraldo; PELLEGRINO, Daniel; TEIXEIRA, Eduardo. Fundamentos de análise funcional. SBM, 2012.

[5] FOLLAND, Gerald B. Real analysis: modern techniques and their applications. John Wiley Sons, 1999.

[6] GARSIA, Adriano M. A simple proof of E. Hopf's maximal ergodic theorem. Journal of Mathematics and Mechanics, v. 14, n. 3, p. 381-382, 1965.

[7] HUTCHINSON, John E. Fractals and self similarity. Indiana University Mathematics Journal, v. 30, n. 5, p. 713-747, 1981.

[8] JÚNIOR, AA de C. Curso de Teoria da Medida. Projeto Euclides, 2004.

[9] KINGMAN, John FC. The ergodic theory of subadditive stochastic processes. Journal of the Royal Statistical Society: Series B (Methodological), v. 30, n. 3, p. 499-510, 1968.

[10] KRAVCHENKO, Aleksei Stanislavovich. Completeness of the space of separable measures in the Kantorovich-Rubinshtein metric. Siberian Mathematical Journal, v. 47, p. 68-76, 2006.

[11] MANÉ, Ricardo. Introdução à teoria ergódica. Instituto de Matemática Pura e Aplicada, 1983.

[12] MELO, Ítalo. On P-weakly hyperbolic iterated function systems. Bulletin of the Brazilian Mathematical Society, New Series, v. 48, n. 4, p. 717-732, 2017.

[13] MELO, Ítalo Dowell Lira; DA CRUZ NETO, João Xavier; DE BRITO, José Márcio Machado. Strong Convergence of Alternating Projections. Journal of Optimization Theory and Applications, v. 194, n. 1, p. 306-324, 2022.

[14] PATA, Vittorino et al. Fixed point theorems and applications. Cham: Springer, 2019.

[15] PETERSEN, Karl E. Ergodic theory. Cambridge university press, 1989.

[16] RIESZ, Frederick. Some mean ergodic theorems. Journal of the London Mathematical Society, v. 1, n. 4, p. 274-278, 1938.

[17] SAKAI, Makoto. Strong convergence of infinite products of orthogonal projections in Hilbert space. Applicable Analysis, v. 59, n. 1-4, p. 109-120, 1995.

[18] VIANA, Marcelo; OLIVEIRA, Krerley. Foundations of ergodic theory. Cambridge University Press, 2016.

[19] WALTERS, Peter. Ergodic theory—introductory lectures. Springer, 2007.

[20] WEYL, Hermann. Über die gleichverteilung von zahlen mod. eins. Mathematische Annalen, v. 77, n. 3, p. 313-352, 1916.

[21] YOSIDA, Kôsaku; KAKUTANI, Shizuo. Birkhoff's ergodic theorem and the maximal ergodic theorem. Proceedings of the Imperial Academy, v. 15, n. 6, p. 165-168, 1939.