O Teorema Ergódico de Birkhoff e Aplicações

L. Cioletti

Março de 2026

Resumo

Nestas notas, apresentamos uma demonstração detalhada do Teorema Ergódico de Birkhoff e algumas de suas aplicações. O texto é focado na convergência das médias temporais, tanto no sentido da norma L^p quanto em quase todo ponto, para o contexto de dinâmicas que preservam medida em espaços de medida σ-finitos.

A prova da convergência em quase todo ponto segue o roteiro clássico baseado nas desigualdades maximais. Para fornecer uma intuição geométrica da convergência em norma, incluímos também a demonstração do Teorema Ergódico de von Neumann em L^2, que aborda o problema de um ponto de vista um pouco mais geométrico usando a teoria de espaços de Hilbert. Como aplicações, demonstramos o Teorema de Borel sobre a equidistribuição dos dígitos na expansão binária de um número real típico e estabelecemos, para dinâmicas que preservam medidas de probabilidade, uma caracterização da ergodicidade em termos da convergência no sentido de Cesàro de $\mu(T^{-n}(A) \cap B)$ para $\mu(A)\mu(B)$.

Ressaltamos que este é um texto de caráter expositivo cujo principal objetivo é apresentar, de forma didática, clara e rigorosa, esse importante resultado a estudantes que estão iniciando seus estudos em Teoria Ergódica. A exposição é baseada em referências clássicas como [4, 5].

1. Introdução e Algumas Notações Básicas

O primeiro grande resultado em Teoria Ergódica foi demonstrado em 1931 por G. D. Birkhoff. Este resultado é conhecido hoje em dia como Teorema Ergódico de Birkhoff. Neste texto vamos apresentá-lo no contexto de transformações que preservam medida em um espaço de medida σ-finita.

Antes de prosseguir, lembramos que uma medida σ-finita definida sobre um espaço mensurável $(X, \mathscr{B})$ é uma função de conjuntos $\mu : \mathscr{B} \to \mathbb{R}^+ \cup \{\infty\}$ satisfazendo

- $\mu(\varnothing) = 0$;
- $\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \displaystyle\sum_{n=1}^{\infty} \mu(B_n)$ se $B_n \cap B_m = \varnothing$, para todo $m \neq n$;
- existe uma coleção enumerável $\{A_n\}_{n\in\mathbb{N}}$ de elementos de $\mathscr{B}$ com $\mu(A_n) < \infty$, para todo $n \in \mathbb{N}$ e $\bigcup_{n\in\mathbb{N}} A_n = X$.

A medida de Lebesgue em $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ fornece um exemplo muito importante de medida σ-finita. Naturalmente, toda medida de probabilidade é σ-finita.

Ao longo deste texto, utilizamos as seguintes notações e convenções. A tripla ordenada $(X, \mathscr{B}, \mu)$ denota um espaço de medida σ-finita. Se $T : X \to X$ denota uma função definida em X e tomando valores em X e $n \in \mathbb{N} \cup \{0\}$ um inteiro não-negativo, denotamos por T^n a n-ésima iterada de T (com a convenção $T^0 \equiv \mathrm{Id}$). Como de costume, usamos a notação $T^{-n}(E) \equiv \{x \in X : T^n(x) \in E\}$ para a imagem inversa, por T^n, do conjunto $E \subseteq X$.

Definição 1 (Transformações que Preservam Medida)**.** Sejam $(X, \mathscr{B}, \mu)$ um espaço de medida σ-finita. Dizemos que uma transformação $T : X \to X$ é mensurável se, para todo $B \in \mathscr{B}$, temos que $T^{-1}(B) \in \mathscr{B}$. Se $T : X \to X$ é uma transformação mensurável e satisfaz

$$\mu(T^{-1}(B)) = \mu(B), \qquad \forall B \in \mathscr{B},$$

então dizemos que T é uma transformação que preserva medida.

Exemplo 2. Sejam $\mathbb{N}$ o conjunto dos números inteiros positivos e $X \equiv \{0,1\}^{\mathbb{N}}$ o espaço produto equipado com sua σ-álgebra produto usual, que será denotada por $\mathscr{B}$. Seja ν uma medida de probabilidade arbitrária definida sobre o conjunto das partes de $\{0,1\}$ e considere a medida produto definida sobre $(X, \mathscr{B})$ e dada por

$$\mu \equiv \prod_{j \in \mathbb{N}} \nu_j, \qquad \text{onde } \nu_j = \nu \quad \forall j \in \mathbb{N}.$$

Considere a transformação $T : X \to X$, denominada descolamento para a esquerda (ou simplesmente "shift"), definida para cada $x \equiv (x_1, x_2, x_3, \ldots) \in X$ por

$$T(x_1, x_2, x_3, \ldots) = (x_2, x_3, \ldots).$$

Afirmamos que T é uma transformação que preserva medida no espaço de probabilidade $(X, \mathscr{B}, \mu)$. Para provar essa afirmação, começaremos mostrando que T é uma transformação mensurável. Para isso, é suficiente mostrar que a imagem inversa de um conjunto cilíndrico em X é também um conjunto cilíndrico. Já que a σ-álgebra produto $\mathscr{B}$ é gerada pelos conjuntos cilíndricos de X da forma

$$[a_1, a_2, \ldots, a_n] \equiv \{x \in X : x_j = a_j, j = 1, \ldots, n\},$$

onde $(a_1, \ldots, a_n)$ é um elemento fixado de $\{0,1\}^n$, a mensurabilidade de T segue diretamente da igualdade abaixo:

$$T^{-1}([a_1, a_2, \ldots, a_n]) = [0, a_1, a_2, \ldots, a_n] \cup [1, a_1, a_2, \ldots, a_n]. \tag{1}$$

Agora vamos mostrar que a transformação T preserva medida no espaço de probabilidade $(X, \mathscr{B}, \mu)$. Primeiro, observamos que segue diretamente da definição da medida produto que

$$\mu([a_1, a_2, \ldots, a_n]) = \nu(\{a_1\}) \cdot \ldots \cdot \nu(\{a_n\}).$$

Agora, usando as duas igualdades anteriores e o fato que a união que aparece em (1) é uma união de conjuntos disjuntos, podemos concluir que vale a seguinte igualdade

$$
\begin{aligned}
\mu\left(T^{-1}([a_1,a_2,\ldots,a_n])\right) &= \mu\Big([0,a_1,a_2,\ldots,a_n]\cup[1,a_1,a_2,\ldots,a_n]\Big)\\
&= \mu\Big([0,a_1,a_2,\ldots,a_n]\Big)+\mu\Big([1,a_1,a_2,\ldots,a_n]\Big)\\
&= \nu(\{0\})\nu(\{a_1\})\cdot\ldots\cdot\nu(\{a_n\})+\nu(\{1\})\nu(\{a_1\})\cdot\ldots\cdot\nu(\{a_n\})\\
&= \Big(\nu(\{0\})+\nu(\{1\})\Big)\cdot\nu(\{a_1\})\cdot\ldots\cdot\nu(\{a_n\})\\
&= \nu(\{a_1\})\cdot\ldots\cdot\nu(\{a_n\})\\
&= \mu\Big([a_1,a_2,\ldots,a_n]\Big),
\end{aligned}
$$

para todo conjunto cilíndrico da forma $[a_1, a_2, \ldots, a_n]$. Já que a coleção de tais conjuntos cilíndricos forma um π-sistema que gera a σ-álgebra produto, segue do Teorema π–λ de Dynkin que $\mu \circ T^{-1} = \mu$. Consequentemente, concluímos que T é uma transformação que preserva medida no espaço de probabilidade $(X, \mathscr{B}, \mu)$.

O leitor mais atento deve ter observado que a transformação T deste exemplo age preservando medida em uma quantidade infinita de espaços de probabilidade distintos, já que temos infinitas possibilidades de escolhas distintas para a medida ν que define a medida produto considerada neste exemplo.

Definição 3 (**Transformação Ergódica**)**.** Seja $(X, \mathscr{B}, \mu)$ um espaço de medida σ-finito e $T : X \to X$ uma transformação mensurável. Dizemos que T é ergódica se para todo $E \in \mathscr{B}$ satisfazendo $T^{-1}(E) = E$, temos $\mu(E) = 0$ ou $\mu(E^c) = 0$.

Uma observação importante é que quando estamos trabalhando com transformações que preservam medida no contexto de espaços de probabilidade é mais comum definir transformações ergódicas como sendo aquelas satisfazendo: para todo $E \in \mathscr{B}$ tal que $T^{-1}(E) = E$, temos $\mu(E) = 0$ ou $\mu(E) = 1$. É claro que a condição que aparece na Definição 3 coincide com esta que acabamos de apresentar, quando $\mu(X) = 1$. Em alguns dos exemplos destas notas vamos trabalhar com espaços de probabilidade e eventualmente usar esta outra maneira de caraterizar transformações ergódicas.

Sejam T uma transformação que preserva medida definida sobre um espaço de probabilidade $(X, \mathscr{B}, \mu)$ e $f \in L^1(X, \mathscr{B}, \mu)$. Definimos a *média temporal* de f em $x \in X$ pela seguinte expressão

$$
\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \qquad \text{se o limite existir.}
$$

Às vezes, chamamos de *média espacial* (ou média no espaço de fase) de uma função $f : X \to \mathbb{R} \cup \{\pm\infty\}$ a seguinte integral de Lebesgue

$$
\int_X f\, d\mu, \qquad \text{caso exista.}
$$

Uma das importantes conclusões do Teorema Ergódico de Birkhoff (Teorema 4) é que para cada $f \in L^1(X, \mathscr{B}, \mu)$ fixada, essas médias são iguais μ-qtp se T é ergódica. Reciprocamente, se para toda $f \in L^1(X, \mathscr{B}, \mu)$ as médias espacial e temporal de f coincidem, μ-q.t.p, então T é uma aplicação ergódica sobre $(X, \mathscr{B}, \mu)$.

2. O Enunciado do Teorema Ergódico de Birkhoff

Teorema 4 (Teorema Ergódico de Birkhoff)**.** Sejam $(X, \mathscr{B}, \mu)$ um espaço de medida σ-finito, $T\colon (X, \mathscr{B}, \mu) \to (X, \mathscr{B}, \mu)$ uma transformação que preserva medida e $f \in L^1(X, \mathscr{B}, \mu)$. Então existe uma função mensurável $f^* \in L^1(X, \mathscr{B}, \mu)$ tal que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = f^*(x) \qquad \mu - \text{qtp.} \tag{2}$$

A função $f^* : X \to \mathbb{R} \cup \{\pm\infty\}$ é $\mathscr{B}$-mensurável e satisfaz a identidade $f^* \circ T = f^*$ μ-qtp.

Além do mais, para todo $W \in \mathscr{B}$, satisfazendo

$$T^{-1}(W) = W \quad \text{e} \quad \mu(W) < +\infty \qquad \text{temos que} \qquad \int_W f^*\, d\mu = \int_W f\, d\mu.$$

Observação. No caso em que $(X, \mathscr{B}, \mu)$ é um espaço de medida finito e T é ergódica, podemos tirar as seguintes conclusões. Primeiro, do Teorema Ergódico de Birkhoff segue que para cada $f \in L^1(X, \mathscr{B}, \mu)$ temos $f^* \circ T = f^*$, μ-qtp. Fazendo uma simples adaptação do Teorema 1.6 (página 28) em [5], considerando a medida de probabilidade $\mathscr{B} \ni E \longmapsto \frac{\mu(E)}{\mu(X)}$ e levando em conta a observação anterior, podemos garantir que f^* é constante μ-qtp. Já que estamos assumindo que $\mu(X) < \infty$, segue da última conclusão do Teorema Ergódico de Birkhoff que

$$f^* = \frac{1}{\mu(X)} \int_X f\, d\mu \qquad \mu - \text{qtp.}$$

No caso particular em que $(X, \mathscr{B}, \mu)$ é um espaço de probabilidade e T é ergódica, temos diretamente da observação anterior, para todo $f \in L^1(X, \mathscr{B}, \mu)$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int_X f\, d\mu, \qquad \mu - \text{qtp.}$$

Apresentaremos a demonstração do Teorema Ergódico de Birkhoff na Seção 5.2. Antes vamos mostrar algumas de suas aplicações.

3. Aplicações do Teorema Ergódico de Birkhoff

3.1. Frequência Assintótica de Visitas a Conjuntos Mensuráveis

Seja $(X, \mathscr{B}, \mu)$ um espaço de probabilidade e seja $T\colon X \to X$ uma transformação que preserva medida. Dado $E \in \mathscr{B}$, podemos perguntar: com que frequência os elementos da órbita de um ponto $x \in X$, isto é, $\{x, T(x), T^2(x), \dots\}$ "entram" no conjunto E? Mais precisamente, podemos determinar algum dos seguintes valores:

$$\liminf_{n\to\infty} \frac{1}{n} \operatorname{Card}(\{x, T(x), \dots T^{n-1}(x)\} \cap E) \quad \text{ou} \quad \limsup_{n\to\infty} \frac{1}{n} \operatorname{Card}(\{x, T(x), \dots T^{n-1}(x)\} \cap E).$$

Claramente, temos $T^i(x) \in E$ se, e somente se, $\chi_E(T^i(x)) = 1$, de modo que o número de elementos de $\{x, T(x), \dots, T^{n-1}(x)\}$ que pertencem ao conjunto E, pode ser dado pela seguinte expressão

$$\operatorname{Card}(\{x, T(x), \dots T^{n-1}(x)\} \cap E) = \sum_{k=0}^{n-1} \chi_E(T^k(x)).$$

Observe também que a quantidade relativa de elementos da órbita de $x \in X$ que pertencem ao conjunto E é dada pela seguinte expressão

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_E(T^i(x)).$$

Já que $(X, \mathscr{B}, \mu)$ é um espaço de probabilidade, então $\chi_E \in L^1(X, \mathscr{B}, \mu)$ assim se assumimos que a aplicação T é ergódica, então segue do Teorema Ergódico de Birkhoff que

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_E(T^i(x)) \xrightarrow{n\to\infty} \int_X \chi_E^*\, d\mu = \int_X \chi_E\, d\mu = \mu(E), \qquad \mu - \text{qtp}.$$

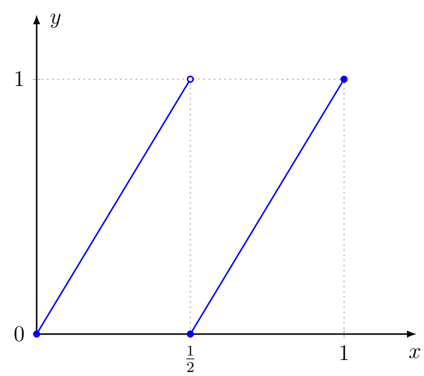
Logo, a órbita de quase todo ponto de $x \in X$, visita o conjunto E com frequência relativa assintótica igual à $\mu(E)$.

3.2. Frequência dos Dígitos na Expansão Binária

A seguir, ilustramos como o Teorema ergódico pode ser aplicado para obter alguns resultados mais simples em Teoria dos Números. Primeiro vamos precisar de um resultado auxiliar cujo enunciado é apresentado abaixo.

Considere o espaço de probabilidade $([0,1), \mathscr{B}, \lambda)$, onde $\mathscr{B}$ denota a σ-álgebra de Borel de $[0,1)$ e λ a medida de Lebesgue sobre o intervalo semi-aberto $[0,1)$. Considere a transformação $T\colon [0,1) \to [0,1)$ dada por

$$T(x) \equiv 2x \bmod 1$$



Mostramos em [1] que a transformação T é uma transformação que age em $([0,1), \mathscr{B}, \lambda)$ preservando a medida λ. A seguir, mostramos um fato ainda mais forte que T é, de fato, uma transformação ergódica.

De fato, seja $f \in L^2([0,1), \mathscr{B}, \lambda)$, e suponha que $(f \circ T)(x) = f(x)$ λ-qtp. Como toda função $f \in L^2([0,1), \mathscr{B}, \lambda)$ admite uma representação em série de Fourier, temos

$$f = \sum_{n\in\mathbb{Z}} \hat{f}(n)\, \gamma_n,$$

onde

$$\gamma_n(x) = e^{2\pi i n x} \qquad \text{e} \qquad \hat{f}(n) = \int_0^1 f(x)\, e^{-2\pi i n x}\, d\lambda(x)$$

Já que $T(x) = 2x \bmod 1$ e $f = f \circ T$ λ-qtp, segue diretamente do Teorema da Mudança de variáveis que

$$\begin{aligned}
\hat{f}(2n) = \widehat{(f \circ T)}(2n) &= \int_0^1 f(T(x))\, e^{-2\pi i 2nx}\, d\lambda(x) \\
&= \int_0^{\frac{1}{2}} f(2x)\, e^{-2\pi i 2nx}\, d\lambda(x) + \int_{\frac{1}{2}}^1 f(2x-1)\, e^{-2\pi i 2nx}\, d\lambda(x) \\
&= \int_0^1 f(x)\, e^{-2\pi i nx}\, d\lambda(x) \\
&= \hat{f}(n), \qquad \forall n \in \mathbb{Z}.
\end{aligned}$$

Iterando a igualdade acima, podemos garantir que $\hat{f}(n) = \hat{f}(2^k n)$, para todo $k \geqslant 1$. Pela Identidade de Parseval, sabemos que $\hat{f}(m) \to 0$ quando $|m| \to \infty$. Desta forma, podemos concluir que os coeficientes

$$\hat{f}(n) = \lim_{k\to\infty} \hat{f}(2^k n) = 0, \qquad \forall n \in \mathbb{Z} \setminus \{0\}.$$

Portanto $f = \hat{f}(0)$ em $L^2([0,1), \mathscr{B}, \lambda)$ e consequentemente f é constante λ-qtp. Aplicando novamente o Teorema 1.6 (página 28) em [5] concluímos finalmente que T é ergódica.

Teorema 5 (Teorema de Borel sobre Números Normais)**.** Considere o espaço de probabilidade $([0,1), \mathscr{B}([0,1)), \lambda)$, onde λ denota a medida de Lebesgue em $[0,1)$. Para cada $x \in [0,1)$ seja $\{d_k(x)\}_{k\in\mathbb{N}}$ a sequência dos dígitos de x na base 2, isto é, $d_k(x) \in \{0,1\}$, para todo $k \in \mathbb{N}$ e vale a seguinte igualdade

$$x = \sum_{k=1}^{\infty} \frac{d_k(x)}{2^k} = \frac{d_1(x)}{2} + \frac{d_2(x)}{2^2} + \frac{d_3(x)}{2^3} + \ldots, \qquad \forall x \in [0,1).$$

Então quase todo número em $[0,1)$, com respeito à medida de Lebesgue, é normal na base 2, isto é, para λ-quase todo ponto $x \in [0,1)$, a frequência dos dígitos 0 ou 1 na expansão binária de x é igual a $\frac{1}{2}$. Mais precisamente, fixado $j \in \{0,1\}$ (conjunto dos dígitos) existe o limite abaixo e vale a igualdade

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{Card}\Big(\{k \in \{1,2,\ldots,n\} : d_k(x) = j\}\Big) = \frac{1}{2}, \qquad \lambda - \text{qtp}.$$

Prova. Seja $T\colon [0,1) \to [0,1)$ definida por $T(x) = 2x \bmod 1$. Sabemos que T preserva a medida de Lebesgue λ e como mostrado acima esta transformação é ergódica em $([0,1), \mathscr{B}, \lambda)$. Outro fato bem-conhecido é que existe um conjunto $Y \subseteq [0,1)$ tal que cada ponto $x \in Y$ possui uma única expansão binária. Além do mais, podemos verificar que Y tem complementar enumerável, portanto $\lambda(Y) = 1$.

Também tem probabilidade um o conjunto de todos pontos que possuem expansão binárias únicas e todos seus iterados também possuem expansão binária única. De fato, considere o conjunto

$$Z \equiv \bigcup_{n=0}^{\infty} T^{-n}([0,1) \setminus Y).$$

Pela σ-aditividade da medida de Lebesgue e pelo fato de T ser uma transformação que preserva esta medida, então podemos verificar que $\lambda(Z) = 0$. De fato,

$$\lambda(Z) = \lambda\left(\bigcup_{n=0}^{\infty} T^{-n}([0,1) \setminus Y)\right) \leqslant \sum_{n=0}^{\infty} \lambda(T^{-n}([0,1) \setminus Y)) = \sum_{n=0}^{\infty} \lambda([0,1) \setminus Y) = 0.$$

Note que $x \in [0,1) \setminus Z$ se, e somente se, $x \in Y$ e todo ponto da órbita de x também está em Y, isto é, $T^n(x) \in Y$, para todo $n \in \mathbb{N} \cup \{0\}$.

Já que o ponto $x = 1/2$ não possui expansão binária única, temos que $(1/2) \in [0,1) \setminus Y$. Além do mais, este é o único ponto onde a aplicação T não é contínua. Logo para todo ponto $x \in [0,1) \setminus Z$ da forma $x = a_1/2 + a_2/2^2 + \ldots$, temos que $x \neq (1/2)$ e da definição de T que

$$T(x) = T\left(\frac{a_1}{2} + \frac{a_2}{2^2} + \frac{a_3}{2^3} + \ldots\right) = \frac{a_2}{2} + \frac{a_3}{2^2} + \ldots \qquad \text{e} \qquad T(x) \in Y.$$

Considere a função $f : [0,1) \to \mathbb{R}$ dada por

$$f(x) \equiv \chi_{[\frac{1}{2},1)}(x), \qquad \forall x \in [0,1).$$

Pelas observações acima, pela definição de f e pelas propriedades elementares de séries geométricas temos para todo $x \in [0,1) \setminus Z$ que

$$f(T^i(x)) = f\left(\frac{a_{i+1}}{2} + \frac{a_{i+2}}{2^2} + \cdots\right) = \begin{cases} 1, & \text{se } a_{i+1} = 1; \\ 0, & \text{se } a_{i+1} = 0. \end{cases}$$

Logo, se $x \in [0,1) \setminus Z$, o número de 1's nos primeiros n dígitos da expansão diádica de x é dado pela seguinte expressão

$$\operatorname{Card}\Big(\{k \in \{1,2,\ldots,n\} : d_k(x) = 1\}\Big) = \sum_{i=0}^{n-1} f(T^i(x)).$$

Dividindo ambos os lados dessa igualdade por n e aplicando o Teorema Ergódico de Birkhoff, obtemos

$$\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) \xrightarrow{n\to\infty} \int_{[0,1)} \chi_{[\frac{1}{2},1)}\, d\lambda = \frac{1}{2} \qquad \lambda - \text{qtp}.$$

Como Z é um conjunto de medida de Lebesgue nula, segue da expressão acima que a frequência de 1's na expansão binária de quase todo ponto em $[0,1)$ é $\frac{1}{2}$. Analogamente, mostramos que a frequência dos dígitos 0's é igual a $\frac{1}{2}$. ■

O Teorema Ergódico de Birkhoff também pode ser aplicado para obter outros resultados em teoria dos números. Alguns são obtidos em [2, 3].

3.3. Convergência das Médias Temporais em L^p

Nesta subseção vamos mostrar como obter a convergência das médias temporais no sentido L^p, a partir do Teorema 4 que garante a convergência das médias temporais em quase todo ponto.

Corolário 6 (Teorema Ergódico L^p de Von Neumann)**.** Sejam $1 \leqslant p < \infty$, $T : X \to X$ uma transformação definida sobre um espaço probabilidade $(X, \mathscr{B}, \mu)$ que preserva medida. Para cada $f \in L^p(X, \mathscr{B}, \mu)$, existe $f^* \in L^p(X, \mathscr{B}, \mu)$ satisfazendo $f^* \circ T = f^*$ μ-qtp e

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i - f^* \right\|_p \ = \ 0.$$

Prova. Se $g : X \to \mathbb{R}$ é uma função limitada e mensurável, então $g \in L^p(X, \mathscr{B}, \mu)$ para todo $p \in [1, +\infty]$. Pelo Teorema Ergódico de Birkhoff, temos que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} g(T^i x) = g^*(x) \qquad \mu - \text{qtp}.$$

Claramente, $g^* \in L^\infty(X, \mathscr{B}, \mu)$ e consequentemente $g^* \in L^p(X, \mathscr{B}, \mu)$. Além disso, segue da continuidade da aplicação $x \longmapsto x^p$ que

$$\lim_{n\to\infty} \left| \frac{1}{n} \sum_{i=0}^{n-1} g(T^i x) - g^*(x) \right|^p = 0, \qquad \mu - \text{qtp}.$$

Da desigualdade triangular, e da monotonicidade do mapa $x \longmapsto x^p$ em $[0, +\infty)$ segue que para μ-qtp $x \in X$ temos

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} g(T^i x) - g^*(x) \right|^p \leqslant \left| \frac{1}{n} \sum_{i=0}^{n-1} |g(T^i x)| + |g^*(x)| \right|^p \leqslant |\|g\|_\infty + \|g\|_\infty|^p = 2^p \, \|g\|_\infty^p.$$

Portanto, segue das observações acima e do Teorema da Convergência Dominada, que

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{i=0}^{n-1} g \circ T^i - g^* \right\|_p = 0. \tag{3}$$

Por questão de simplicidade, dada uma função genérica $g \colon X \to \mathbb{R} \cup \{-\infty, +\infty\}$ vamos denotar por

$$S_n(g, T)(x) \equiv \frac{1}{n} \sum_{i=0}^{n-1} g(T^i(x)), \qquad \forall x \in X.$$

Da igualdade (3) segue que $\{S_n(g, T)\}_{n \in \mathbb{Z}_+}$ é uma sequência de Cauchy em $L^p(X, \mathscr{B}, \mu)$. Portanto, dado $\varepsilon > 0$, existe algum número natural $N(\varepsilon, g)$ tal que, para todo $n \geqslant N(\varepsilon, g)$ e $k \in \mathbb{N}$, temos

$$\left\| \frac{1}{n} \sum_{i=0}^{n-1} g \circ T^i - \frac{1}{n+k} \sum_{i=0}^{n+k-1} g \circ T^i \right\|_p < \varepsilon. \tag{4}$$

O próximo passo é mostrar que se $f \in L^p(X, \mathscr{B}, \mu)$, então $\{S_n(f,T)\}_{n\in\mathbb{Z}_+}$ é uma sequência de Cauchy em $L^p(X, \mathscr{B}, \mu)$. Primeiro, observamos que segue diretamente da desigualdade triangular que para quaisquer $f, g \in L^p(X, \mathscr{B}, \mu)$

$$\|S_n(f-g,T)\|_p \leqslant \|f-g\|_p \qquad \forall n \in \mathbb{Z}_+. \tag{5}$$

Dado $\varepsilon > 0$, sabemos que, usando truncamentos, segue de uma aplicação do Teorema da Convergência Dominada de Lebesgue que podemos encontrar uma função $g \in L^\infty(X, \mathscr{B}, \mu)$ tal que $\|f-g\|_p < \varepsilon/4$. Aplicando novamente a desigualdade triangular e em seguida, as desigualdade (3) e (5) temos a seguinte estimativa

$$\begin{aligned}\|S_n(f,T) - S_{n+k}(f,T)\|_p &\leqslant \|S_n(f,T) - S_n(g,T)\|_p + \|S_n(g,T) - S_{n+k}(g,T)\|_p \\ &\quad + \|S_{n+k}(g,T) - S_{n+k}(f,T)\|_p \\ &\leqslant \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon, \qquad \forall n \geqslant N(\varepsilon/2, g) \text{ e } k > 0\end{aligned}$$

Portanto, $\{S_n(f,T)\}_{n\in\mathbb{Z}_+}$ é uma sequência de Cauchy em $L^p(X, \mathscr{B}, \mu)$ e, consequentemente, $\|S_n(f,T) - L(f)\|_p \to 0$, para alguma função $L(f) \in L^p(X, \mathscr{B}, \mu)$.

Para finalizar precisamos mostrar que $L(f) = f^*$ (a função limite fornecida pelo Teorema Ergódico de Birkhoff). De fato, observe inicialmente que como μ é uma medida de probabilidade, então segue da desigualdade de Lyapunov que para todo $p \in [0, \infty]$, temos $f \in L^p(X, \mathscr{B}, \mu) \Longrightarrow f \in L^1(X, \mathscr{B}, \mu)$. Desta forma para qualquer que seja $f \in L^p(X, \mathscr{B}, \mu)$ podemos aplicar o Teorema Ergódico de Birkhoff, e assegurar que existe algum subconjunto $Z(f) \equiv Z \subseteq X$ com $\mu(Z) = 0$ tal que $S_n(f,T)(x) \to f^*(x)$, quando $n \to \infty$, para todo $x \in X \setminus Z$. Por outro lado, como $\{S_n(f,T)\}_{n\in\mathbb{Z}_+}$ converge para $L(f)$ em $L^p(X, \mathscr{B}, \mu)$ sabemos que existe uma subsequência $S_{n_k}(f,T)$ que converge μ-qtp para $L(f)$, ou seja, para algum subconjunto $W \subset X$ de medida μ-nula temos que

$$\lim_{k\to\infty} S_{n_k}(f,T)(x) = Lf(x), \qquad \forall x \in X \setminus W.$$

Desta forma temos para todo $x \in X \setminus (Z \cup W)$ temos

$$f^*(x) = \lim_{k\to\infty} S_{n_k}(f,T)(x) = L(f)(x).$$

O que mostra que $L(f) = f^*$ μ-qtp. ■

3.4. Ergodicidade e Independência Assintótica no Sentido de Cesàro

O próximo corolário fornece outra forma da definir a noção de ergodicidade. Este resultado ilustra o poder do teorema ergódico, pois ao supor que T é ergódica em $(X, \mathscr{B}, \mu)$, podemos concluir do Teorema 1.5(iv) da referência [5] que para todo $A, B \in \mathscr{B}$ com $\mu(A) > 0$, $\mu(B) > 0$, existe algum $n \geqslant 1$ com $\mu(T^{-n}A \cap B) > 0$. Além disso, segue do próximo resultado que a sequência $\mu(T^{-n}A \cap B)$ converge, no sentido de Cesàro, para $\mu(A)\mu(B)$.

Corolário 7. Sejam $(X, \mathscr{B}, \mu)$ um espaço de probabilidade e $T\colon X \to X$ uma transformação que preserva medida. Então T é ergódica se, e somente se, para todo $A, B \in \mathscr{B}$,

$$\frac{1}{n}\sum_{i=0}^{n-1} \mu(T^{-i}A \cap B) \xrightarrow{n\to\infty} \mu(A)\,\mu(B).$$

Prova. Suponha que T seja uma transformação ergódica. Já que $(X, \mathscr{B}, \mu)$ é um espaço de probabilidade então podemos aplicar o teorema ergódico à $f = \chi_A$ para afirmar que

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_A(T^i(x)) \xrightarrow{n\to\infty} \mu(A) \quad \mu-\text{qtp}.$$

Multiplicando por χ_B, ambos lados da igualdade acima ficamos com

$$\frac{1}{n}\sum_{i=0}^{n-1} \chi_A(T^i(x))\,\chi_B(x) \xrightarrow{n\to\infty} \mu(A)\,\chi_B(x) \qquad \mu-\text{qtp}.$$

Aplicando o Teorema da Convergência Dominada, temos

$$\frac{1}{n}\sum_{i=0}^{n-1} \mu(T^{-i}A\cap B) \xrightarrow{n\to\infty} \mu(A)\,\mu(B).$$

Reciprocamente, suponha que para quaisquer subconjuntos $A, B \in \mathscr{B}$, fixados que a sequência $\mu(T^{-n}A \cap B)$ converge no sentido de Cesàro para $\mu(A)\mu(B)$. Seja $E \in \mathscr{B}$ tal que $T^{-1}E = E$. Tomando $A = B = E$, então segue da hipótese que

$$\frac{1}{n}\sum_{i=0}^{n-1} \mu(E) \xrightarrow{n\to\infty} \mu(E)^2.$$

Logo $\mu(E) = \mu(E)^2$ e, portanto, $\mu(E) = 0$ ou 1. ■

4. O Teorema Ergódico de von Neumann

Vamos começar esta seção obtendo uma versão abstrata do Teorema Ergódico para Médias em Espaços de Hilbert. Em seguida, mostramos como obter como corolário deste teorema o Teorema Ergódico de von Neumann.

Teorema 8 (Teorema Ergódico para Médias em Espaços de Hilbert)**.** Seja $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ um espaço de Hilbert e $U : \mathcal{H} \to \mathcal{H}$ um operador linear satisfazendo $\|Uf\| \leqslant \|f\|$, para todo $f \in \mathcal{H}$. Denote por $\mathcal{M}$ o subespaço fechado de $\mathcal{H}$ definido por $\mathcal{M} = \{f \in \mathcal{H} : Uf = f\}$. Seja $P : \mathcal{H} \to \mathcal{M}$ a projeção ortogonal de $\mathcal{H}$ sobre $\mathcal{M}$. Então para cada $f \in \mathcal{H}$,

$$\lim_{n\to\infty} \left\| \frac{1}{n}\sum_{k=0}^{n-1} U^k f - Pf \right\| = 0.$$

Prova. Seja $\mathcal{N}$ o fecho, com respeito a norma $\|\cdot\|$, do subespaço $\{g - Ug : g \in \mathcal{H}\}$, isto é,

$$\mathcal{N} \equiv \overline{\{g - Ug : g \in \mathcal{H}\}}.$$

Afirmamos que $\mathcal{N}$ e $\mathcal{M}$ são subespaços ortogonais complementares de $\mathcal{H}$. Para isto é suficiente mostrar que $\mathcal{N}^\perp = \mathcal{M}$, onde

$$\mathcal{N}^\perp = \{h \in \mathcal{H} : \langle h, n\rangle = 0, \quad \forall n \in \mathcal{N}\}.$$

Seja $h \in \mathcal{N}^\perp$, então $\langle h, g - Ug\rangle = 0$, para todo $g \in \mathcal{H}$ e portanto

$$0 = \langle h, g\rangle - \langle h, Ug\rangle = \langle h, g\rangle - \langle U^*h, g\rangle = \langle h - U^*h, g\rangle, \qquad \forall g \in \mathcal{H}.$$

Como a igualdade acima é válida para todo $g \in \mathcal{H}$, podemos tomar $g = h - U^*h$ e concluir da igualdade acima que $\|h - U^*h\|^2 = 0$ e consequentemente que $h = U^*h$. Vamos mostrar a seguir que vale também a seguinte igualdade $Uh = h$. Para isto basta observar que segue da propriedade de bilinearidade do produtos interno, da definição do operador adjunto, da igualdade $h = U^*h$ e da hipótese de contratividade fraca do operador U que vale a seguinte desigualdade

$$\begin{aligned}
\|Uh - h\|^2 &= \langle Uh - h, Uh - h\rangle \\
&= \|Uh\|^2 - \langle h, Uh\rangle - \langle Uh, h\rangle + \|h\|^2 \\
&= \|Uh\|^2 - \langle U^*h, h\rangle - \langle h, U^*h\rangle + \|h\|^2 \\
&\leqslant \|h\|^2 - \langle h, h\rangle - \langle h, h\rangle + \|h\|^2 \\
&= 0.
\end{aligned}$$

O que mostra que $Uh = h$ e portanto que $h \in \mathcal{M}$. Deste fato segue que $\mathcal{N}^\perp \subseteq \mathcal{M}$.

Reciprocamente, suponha que $h \in \mathcal{M}$. Neste caso sabemos pela definição de $\mathcal{M}$ que $Uh = h$. Nosso próximo passo é mostrar que esta igualdade implica $U^*h = h$. A ideia é usar um argumento análogo ao empregado acima, mas para isto vamos precisar mostrar primeiro que U^* também é uma contração fraca. De fato, segue das propriedades do operador adjunto e da desigualdade de Cauchy-Schwarz que temos

$$\|U^*f\|^2 = |\langle U^*f, U^*f\rangle| = |\langle UU^*f, f\rangle| \leqslant \|UU^*f\| \cdot \|f\| \leqslant \|U^*f\| \cdot \|f\|, \qquad \forall f \in \mathcal{H}.$$

Se $\|U^*f\| = 0$, então obviamente temos $\|U^*f\| \leqslant \|f\|$. Por outro lado, se $\|U^*f\| \neq 0$, podemos dividir ambos lados da desigualdade acima por $\|U^*f\|$, ficando com $\|U^*f\| \leqslant \|f\|$. O que mostra que U^* também é uma contração fraca. De posse deste fato e lembrando que $Uh = h$, podemos verificar (como feito acima) que

$$\begin{aligned}
\|U^*h - h\|^2 &= \langle U^*h - h, U^*h - h\rangle \\
&= \|U^*h\|^2 - \langle h, U^*h\rangle - \langle U^*h, h\rangle + \|h\|^2 \\
&= \|U^*h\|^2 - \langle h, Uh\rangle - \langle Uh, h\rangle + \|h\|^2 \\
&\leqslant \|h\|^2 - \langle h, h\rangle - \langle h, h\rangle + \|h\|^2 \\
&= 0.
\end{aligned}$$

Estabelecida esta igualdade podemos concluir que para todo $g \in \mathcal{H}$,

$$\langle h, g - Ug\rangle = \langle h, g\rangle - \langle h, Ug\rangle = \langle h - U^*h, g\rangle = 0.$$

Como g na igualdade acima é arbitrária segue da definição de $\mathcal{N}$ que $h \in \mathcal{N}^\perp$. Estabelecendo assim a inclusão $\mathcal{M} \subseteq \mathcal{N}^\perp$. Como já havíamos provado a inclusão reversa, concluímos finalmente que $\mathcal{N}^\perp = \mathcal{M}$. Já que $\mathcal{N}$ é um subespaço fechado de $\mathcal{H}$ podemos afirmar que vale a seguinte igualdade $\mathcal{M}^\perp = (\mathcal{N}^\perp)^\perp = \mathcal{N}$. Como $\mathcal{M}$ também é um subespaço fechado de $\mathcal{H}$, podemos concluir então que

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp = \mathcal{M} \oplus \mathcal{N} \qquad \Longrightarrow \qquad \text{se } f \in \mathcal{H}, \text{ então } f = Pf + f_0. \tag{6}$$

Próximo passo é mostrar que

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{k=0}^{n-1} U^k f \right\| = 0, \qquad \forall f \in \mathcal{N}. \tag{7}$$

Vamos mostrar o fato acima primeiro no caso em que $f = g - Ug$ para algum elemento $g \in \mathcal{H}$. Neste caso, temos por um argumento telescópico e pela desigualdade triangular que

$$\frac{1}{n}\sum_{k=0}^{n-1} U^k f = \frac{1}{n}(g - U^n g), \qquad \Longrightarrow \qquad \left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k f\right\| \leqslant \frac{2}{n}\|g\| \xrightarrow{n\to\infty} 0.$$

No caso geral, se $f \in \mathcal{N}$ é um elemento genérico, sabemos que existe alguma sequência $\{g_i\}_{i\in\mathbb{N}}$ de elementos de $\mathcal{H}$ tal que $f_i = g_i - Ug_i$, e $\|f_i - f\| \to 0$, quando $i \to \infty$. Desta forma para cada $i \in \mathbb{N}$ fixado temos da desigualdade triangular que

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k f\right\| \leqslant \left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k (f - f_i)\right\| + \left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k f_i\right\|$$

Dado $\varepsilon > 0$, podemos escolher $i \in \mathbb{N}$ tal que $\|f - f_i\| < \varepsilon/2$. Como o índice i está fixado podemos usar o fato provado acima para garantir que existe algum $n_0 \in \mathbb{N}$ tal que para todo $n \geqslant n_0$ temos

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

Desta forma segue das desigualdades acima que para cada $\varepsilon > 0$ podemos encontrar $n_0 \in \mathbb{N}$ tal que para todo $n \geqslant n_0$ temos

$$\left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k f\right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

O que encerra a prova de (7).

Agora, dado um vetor genérico $f \in \mathcal{H}$, sabemos de (6) que existe algum $f_0 \in \mathcal{N}$ tal que $f = Pf + f_0$. Como $Pf \in \mathcal{M}$ temos que $U^k(Pf) = Pf$, para todo $k \in \mathbb{N}$ e assim temos

$$\begin{aligned}\left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k f - Pf\right\| &= \left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k (Pf + f_0) - Pf\right\| \\ &= \left\|\frac{1}{n}\sum_{k=0}^{n-1} U^k f_0\right\| \xrightarrow{n\to\infty} 0.\end{aligned}$$

O que encerra a prova do Teorema. ■

Teorema 9 (Teorema Ergódico von Neumann - 1932)**.** Sejam $(X, \mathscr{B}, \mu)$ um espaço de medida σ-finito, $T\colon X \to X$ uma transformação que preserva medida e $f \in L^2(X, \mathscr{B}, \mu)$. Então existe uma função $f^* \in L^2(X, \mathscr{B}, \mu)$ tal que

$$\lim_{n\to\infty} \left\|\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k - f^*\right\|_2 = 0.$$

Prova. A ideia é mostrar que o operador de Koopman $U_T : L^2(\mu) \to L^2(\mu)$ dado por $L^2(\mu) \ni f \longmapsto U_T f \equiv f \circ T \in L^2(\mu)$ satisfaz as hipóteses do Teorema 8. De fato, já que estamos assumindo que $T : X \to X$ preserva medida, então o operador de Koopman descrito acima está bem definido (nas μ-classes de equivalência) e além do mais satisfaz

$$\|U_T f\|_2^2 = \int_X |U_T f|^2 \, d\mu = \int_X |f \circ T|^2 \, d\mu = \int_X |f|^2 \, d\mu = \|f\|_2^2.$$

Portanto, U_T satisfaz a hipótese de contração fraca $\|U_T f\| \leqslant \|f\|$ exigida no Teorema 8. Portanto, da definição do operador de Koopman e da conclusão do Teorema 8 segue que existe um elemento $f^* \in L^2(\mu)$ tal que

$$\lim_{n\to\infty} \left\| \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k - f^* \right\| = \lim_{n\to\infty} \left\| \frac{1}{n} \sum_{k=0}^{n-1} U_T^k f - f^* \right\| = 0.$$

O que encerra a prova do teorema. ■

5. Prova do Teorema Ergódico de Birkhoff

5.1. Desigualdades Maximais

Passamos agora para os preparativos da prova do Teorema 4. Se $T\colon X \to X$ é uma transformação que preserva medida em $(X, \mathscr{B}, \mu)$, então podemos definir o operador de Koopman U_T para cada elemento de $L^1(\mu)$ por $U_T(f) = f \circ T$. Já que μ é T-invariante temos $U_T(L^1(\mu)) \subseteq L^1(\mu)$, $\quad U_T L^1_{\mathbb{R}}(\mu) \subset L^1_{\mathbb{R}}(\mu)$ e $\|U_T f\|_1 = \|f\|_1$ para todo $f \in L^1(\mu)$.

Um dos ingredientes principais da prova do teorema ergódico é o seguinte resultado, que aplicaremos ao operador U_T. Antes, lembramos que um operador linear $U\colon L^1_{\mathbb{R}}(\mu) \to L^1_{\mathbb{R}}(\mu)$ é dito *positivo* se, para toda $f \geqslant 0$, temos $Uf \geqslant 0$.

Teorema 10 (Teorema Ergódico Maximal)**.** Seja $U\colon L^1_{\mathbb{R}}(\mu) \to L^1_{\mathbb{R}}(\mu)$ um operador linear positivo com $\|U\|_{\mathrm{op}} \leqslant 1$. Seja $N \in \mathbb{N}$ e $f \in L^1_{\mathbb{R}}(\mu)$. Defina

- $f_0 \equiv 0$;
- $f_n \equiv f + Uf + U^2 f + \cdots + U^{n-1} f$, para $n \geqslant 1$; e
- $F_N = \max\limits_{0 \leqslant n \leqslant N} f_n$.

Então

$$\int_{\{x\colon F_N(x) > 0\}} f \, d\mu \geqslant 0.$$

Prova (devida a A. Garsia). Primeiro, como $f_0 \equiv 0$, temos claramente $F_N \geqslant 0$ e portanto $|F_N| = F_N$. Além do mais, segue da desigualdade triangular que

$$\begin{aligned} |F_N| = F_N &\leqslant \max_{0 \leqslant n \leqslant N} |f_n| \\ &\leqslant \max_{0 \leqslant n \leqslant N} \sum_{i=0}^{n-1} |U^i f| \\ &\leqslant \sum_{i=0}^{N} |U^i f|. \end{aligned}$$

Logo $F_N \in L^1_{\mathbb{R}}(\mu)$. Por definição, para cada $0 \leqslant n \leqslant N$, temos $F_N \geqslant f_n$. Logo, $UF_N \geqslant Uf_n$, pela positividade do operador U. Como $Uf_n = f_{n+1} - f$, temos $UF_N + f \geqslant f_{n+1}$. Segue que

$$
\begin{aligned}
UF_N(x) + f(x) &\geqslant \max_{1\leqslant n\leqslant N} f_n(x) \\
&= \max_{0\leqslant n\leqslant N} f_n(x) \quad \text{quando } F_N(x) > 0 \\
&= F_N(x).
\end{aligned}
$$

Assim, $f \geqslant F_N - UF_N$ em $A = \{x : F_N(x) > 0\}$, de modo que

$$
\begin{aligned}
\int_A f\,d\mu &\geqslant \int_A F_N\,d\mu - \int_A UF_N\,d\mu \\
&= \int_X F_N\,d\mu - \int_A UF_N\,d\mu && \text{(pois } F_N = 0 \text{ em } X \setminus A) \\
&\geqslant \int_X F_N\,d\mu - \int_X UF_N\,d\mu && \text{(pois } F_N \geqslant 0 \text{ e portanto } UF_N \geqslant 0) \\
&\geqslant 0 && \text{(pois } \|U\| \leqslant 1). \qquad \blacksquare
\end{aligned}
$$

Corolário 11. Seja $T\colon X \to X$ uma transformação que preserva medida em um espaço de medida σ-finito. Se $g \in L^1_{\mathbb{R}}(\mu)$ e

$$
B_\alpha = \left\{ x \in X : \sup_{n\geqslant 1} \frac{1}{n} \sum_{i=0}^{n-1} g(T^i(x)) > \alpha \right\},
$$

então

$$
\int_{B_\alpha \cap A} g\,d\mu \geqslant \alpha\,\mu(B_\alpha \cap A)
$$

sempre que $T^{-1}A = A$ e $\mu(A) < \infty$.

Prova. Assuma, inicialmente, que $\mu(X) < \infty$. Vamos considerar primeiro o caso $A = X$. Defina $f \equiv g - \alpha$. Então

$$
B_\alpha = \left\{ x \in X : \sup_{n\geqslant 1} \frac{1}{n} \sum_{i=0}^{n-1} g(T^i(x)) > \alpha \right\} = \bigcup_{N=0}^{\infty} \{x : F_N(x) > 0\}.
$$

Aplicando o Teorema Ergódico Maximal (Teorema 10) com U sendo o operador de Koopman U_T obtemos a seguinte desigualdade $\int_{B_\alpha} f\,d\mu \geqslant 0$. Segue desta desigualdade e da definição de f que $\int_{B_\alpha} g\,d\mu \geqslant \alpha\,\mu(B_\alpha)$.

Agora vamos considerar o caso em que $\mu(X) = +\infty$. Seja $A \in \mathscr{B}$ um conjunto de medida finita, satisfazendo $T^{-1}(A) = A$. Considere o espaço de medida $(A, \mathscr{B} \cap A, \nu)$, onde $\nu(E) = \mu(A \cap E)$, para todo $E \in \mathscr{B} \cap A$. Como estamos assumindo que A é T-invariante temos que $T|_A$ é uma transformação $(\mathscr{B} \cap A)$-mensurável que leva A em A e que preserva a medida em $(A, \mathscr{B} \cap A, \nu)$. Usando a desigualdade obtida no parágrafo anterior e a definição da medida ν obtemos finalmente que

$$
\int_{A \cap B_\alpha} g\,d\mu \geqslant \alpha\,\mu(A \cap B_\alpha).
$$

O que encerra a prova do corolário. $\blacksquare$

5.2. Prova do Teorema Ergódico de Birkhoff

A prova segue o roteiro descrito abaixo e será dividida em 5 partes. Os argumentos das Partes 1, 2 e 3 estabelecem resultados que são válidos em quaisquer espaços de medida σ-finitos. A afirmação provada na Parte 4 é restrita à espaços de medida finita. Na Parte 5, estendemos parcialmente, o resultado da Parte 4 para espaços gerais de medida σ-finitos. A notação $\mathscr{I}$ que aparece abaixo se refere a σ-álgebra dos subconjuntos de X que são T-invariantes, isto é, a coleção $\mathscr{I} \equiv \{W \subseteq X : W \in \mathscr{B} \text{ e } T^{-1}(W) = W\}$.

1 T-invariância de f^* e f_*

Definimos

$$f^*(x) \equiv \limsup_{n\ to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$$

$$f_*(x) \equiv \liminf_{n\ to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x)$$

Provamos:

$$f^* \circ T = f^* \quad \text{e} \quad f_* \circ T = f_* \ \mu-\text{qtp}.$$

2 Convergência $f^* = f_*$

Definimos $E_{\alpha,\beta}(f) \equiv E_{\alpha,\beta}$ por

$$E_{\alpha,\beta} \equiv \{x : f_*(x) < \beta \text{ e } \alpha < f^*(x)\}.$$

Consideramos a decomposição

$$\{f_* < f^*\} = \bigcup_{\beta<\alpha;\alpha,\beta\in\mathbb{Q}} E_{\alpha,\beta}$$

Provamos que $f^* = f_*$ mostrando que $\mu(E_{\alpha,\beta}) = 0$ para todo par $\beta < \alpha$ racional.

3 Integrabilidade de f^*

Verificamos que a sequência

$$g_n = \left|\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^i\right|$$

satisfaz $\int g_n \, d\mu \leqslant \int |f| \, d\mu$ e usamos o Lema de Fatou para mostrar

$$\int |f^*| \, d\mu \leqslant \int |f| \, d\mu < \infty.$$

4 Igualdade das Integrais em X

Consideramos o caso $\mu(X) < +\infty$ e mostramos que

$$\int_X f^* \, d\mu = \int_X f \, d\mu$$

5 Igualdade das Integrais em $\mathscr{I}$

Mostramos que para toda $f \in L^1(\mu)$ e $W \in \mathscr{B}$ satisfazendo $T^{-1}(W) = W$ e $\mu(W) < +\infty$

$$\int_W f^* \, d\mu = \int_W f \, d\mu$$

Prova. Antes de passar para os argumentos observamos que é suficiente provar o teorema para funções $f \in L^1_{\mathbb{R}}(\mu)$. Já que f é uma função à valores reais, então podemos afirmar que estão bem definidas e são $\mathscr{B}$-mensuráveis as seguintes funções:

$$f^*(x) \equiv \limsup_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) \qquad \text{e} \qquad f_*(x) \equiv \liminf_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x).$$

<u>Parte 1.</u> Seja $(X, \mathscr{B}, \mu)$ um espaço de medida σ-finito e $f \in L^1(X, \mathscr{B}, \mu)$. Afirmamos que valem as seguintes igualdades:

$$f^* \circ T = f^* \qquad \text{e} \qquad f_* \circ T = f_*.$$

De fato, para cada $n \in \mathbb{Z}_+$ e $x \in X$ seja

$$a_n(x) \equiv \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x).$$

Então segue diretamente da definição de $a_n(x)$ que

$$\frac{n+1}{n} a_{n+1}(x) = \frac{f(x)}{n} + a_n(Tx).$$

Como estamos assumindo que $f \in L^1(X, \mathscr{B}, \mu)$, então podemos afirmar que existe algum conjunto $\mathscr{B}$-mensurável $Z_1 \subseteq X$ tal que $|f(x)| < +\infty$ para todo $x \in X \setminus Z_1$. Portanto segue das propriedades elementares de $\liminf$ e $\limsup$, da igualdade acima e da definição de $f_*(x)$ que

$$\begin{aligned} f_*(x) &= \left(\lim_{n\to\infty} \frac{n+1}{n}\right) \cdot \liminf_{n\to\infty} (\, a_{n+1}(x)) \\ &= \liminf_{n\to\infty} \left(\frac{n+1}{n}\, a_{n+1}(x)\right) \\ &= \liminf_{n\to\infty} \left(\frac{f(x)}{n} + a_n(Tx)\right) \\ &= f_*(T(x)), \qquad \forall x \in X \setminus Z_1. \end{aligned}$$

Argumento análogo substituindo $\liminf$ por $\limsup$ mostra que $f^* = f^* \circ T$ μ-qtp.

Parte 2. Para cada par de números reais α, β fixado, defina

$$E_{\alpha,\beta}(f) \equiv \{x \in X : f_*(x) < \beta \text{ e } \alpha < f^*(x)\}.$$

Por questão de simplicidade, ocasionalmente também vamos usar a notação mais curta $E_{\alpha,\beta}$ para nos referir ao conjunto $E_{\alpha,\beta}(f)$.

Um dos principais objetivos nesta parte da prova é mostrar que $\mu(E_{\alpha,\beta}) < \infty$, para quaisquer $\beta < \alpha$. Vamos estabelecer isto considerando separadamente os casos em que $\alpha > 0$ e $\alpha \leqslant 0$.

Considere inicialmente que $\alpha > 0$. Por hipótese, $(X, \mathscr{B}, \mu)$ é um espaço de medida σ-finito. Portanto podemos garantir que existe algum $C \in \mathscr{B}$ satisfazendo $C \subseteq E_{\alpha,\beta}$ e $\mu(C) < \infty$.

Seja $C \in \mathscr{B}$ um subconjunto arbitrário de $E_{\alpha,\beta}$ de medida finita. Considere a função auxiliar $h : X \to [-\infty, +\infty]$ dada por $h = f - \alpha\, \chi_C$. Por hipótese, $f \in L^1(\mu)$ e como $\mu(C) < +\infty$ segue que $\chi_C \in L^1(\mu)$ e logo $h \in L^1(\mu)$. Aplicando o Teorema Ergódico Maximal (Teorema 10) à função h com o operador $U : L^1(\mu) \to L^1(\mu)$, sendo o operador de Koopman U_T temos:

$$\int_{\{x:\, H_N(x)>0\}} (f - \alpha\, \chi_C)\, d\mu \geqslant 0, \qquad \forall N \geqslant 1, \tag{8}$$

onde função H_N é a função maximal, do Teorema 10, associada à função h.

Argumentando exatamente como no caso $\mu(X) < +\infty$, temos as seguintes relações

$$C \subseteq E_{\alpha,\beta} \subseteq \bigcup_{N=0}^{\infty} \{x : H_N(x) > 0\}.$$

Como $H_N \equiv \max\{h_n : 0 \leqslant n \leqslant N\}$, temos $\{x : H_N(x) > 0\} \subseteq \{x : H_{N+1}(x) > 0\}$. Portanto segue do Teorema da Convergência Dominada e (8) que existe o limite

$$\int_{\bigcup\limits_{N\in\mathbb{N}} \{x:\, H_N(x)>0\}} (f - \alpha\, \chi_C)\, d\mu = \lim_{N\to\infty} \int_{\{x:\, H_N(x)>0\}} (f - \alpha\, \chi_C)\, d\mu \geqslant 0. \tag{9}$$

Usando a desigualdade acima e que $C \subseteq E_{\alpha,\beta} \subseteq \bigcup_{N=0}^{\infty}\{x : H_N(x) > 0\}$, temos

$$\alpha\mu(C) = \alpha \int_{\bigcup\limits_{N\in\mathbb{N}}\{x\colon H_N(x)>0\}} \chi_C \, d\mu \;\leqslant\; \int_{\bigcup\limits_{N\in\mathbb{N}}\{x\colon H_N(x)>0\}} f \, d\mu$$

$$\leqslant \int_{\bigcup\limits_{N\in\mathbb{N}}\{x\colon H_N(x)>0\}} |f| \, d\mu \leqslant \int_X |f| \, d\mu.$$

Como estamos assumindo que $\alpha > 0$, segue das desigualdades acima que

$$\mu(C) \leqslant \frac{1}{\alpha}\int_X |f|\, d\mu, \qquad \forall C \subseteq E_{\alpha,\beta} \text{ satisfazendo } \mu(C) < +\infty. \tag{10}$$

Como $(X, \mathscr{B}, \mu)$ é um espaço de medida σ-finito, existe uma sequência de mensuráveis monótona não-decrescente $\{F_n\}_{n\in\mathbb{N}}$ tal que $\cup_{n=1}^{\infty} F_n = X$ e $\mu(F_n) < +\infty$. Da propriedade de monotonicidade da medida, segue que $F_n \cap E_{\alpha,\beta}$ é um subconjunto de medida finita de $E_{\alpha,\beta}$ para todo $n \in \mathbb{N}$. É claro que a sequência de conjuntos $\{F_n \cap E_{\alpha,\beta}\}_{n\in\mathbb{N}}$ é uma sequência de mensuráveis monótona não-decrescente que converge para $E_{\alpha,\beta}$. Portanto segue da desigualdade (10) e da continuidade da medida e da hipótese que $f \in L^1(\mu)$ que

$$\mu(E_{\alpha,\beta}) = \lim_{n\to\infty} \mu(F_n \cap E_{\alpha,\beta}) \leqslant \frac{1}{\alpha}\int_X |f|\, d\mu < +\infty. \tag{11}$$

O que encerra a prova de que $\mu(E_{\alpha,\beta}) < +\infty$, para o caso $\alpha > 0$.

Vamos considerar agora o caso em que $\alpha \leqslant 0$. Como na prova do teorema só precisamos considerar conjuntos $E_{\alpha,\beta}$'s, com $\beta < \alpha$, então, neste caso, temos necessariamente que $\beta < 0$.

Uma observação muito útil é que se substituímos f, α, β por $-f, -\beta, -\alpha$, respectivamente, temos das igualdades $(-f)^* = -f_*$ e $(-f)_* = -f^*$ e da definição de $E_{\alpha,\beta}(f)$ que

$$\begin{aligned} E_{-\beta,-\alpha}(-f) &= \{x \in X : (-f)_*(x) < -\alpha \text{ e } -\beta < (-f)^*(x)\} \\ &= \{x \in X : -f^*(x) < -\alpha \text{ e } -\beta < -f_*(x)\} \\ &= \{x \in X : f_*(x) < \beta \text{ e } \alpha < f^*(x)\} \\ &= E_{\alpha,\beta}(f). \end{aligned} \tag{12}$$

Aplicando diretamente o resultado obtido em (11) para a função $(-f)$ no lugar de f e $(-\beta)$ no lugar de α, concluímos que $\mu(E_{-\beta,-\alpha}(-f)) < +\infty$. Mas como mostrado acima o conjunto $E_{-\beta,-\alpha}(-f) = E_{\alpha,\beta}(f)$. Logo $\mu(E_{\alpha,\beta}) = \mu(E_{-\beta,-\alpha}(-f)) < \infty$, qualquer que seja $\alpha \leqslant 0$. Este argumento encerra a prova de que $\mu(E_{\alpha,\beta}) < +\infty$, para quaisquer $\beta < \alpha$.

Para demonstrar a existência do limite em (2) temos que mostrar que $f^* = f_*$ μ-qtp e para encerrar a prova do teorema devemos provar que $f^* \in L^1(\mu)$ e também que as integrais de f e f^*, com respeito à μ, coincidem.

Já que vale a seguinte igualdade

$$\{x \in X : f_*(x) < f^*(x)\} = \bigcup_{\substack{\beta<\alpha \\ \alpha,\beta\in\mathbb{Q}}} E_{\alpha,\beta}(f), \tag{13}$$

para mostrar a validade de (2) basta mostrar que $\mu(E_{\alpha,\beta}(f)) = 0$ para todo par de racionais $\beta < \alpha$, pois disso seguirá que $f^* = f_*$ μ-qtp.

Claramente $T^{-1}E_{\alpha,\beta}(f) = E_{\alpha,\beta}(f)$, e se definimos

$$B_\alpha(f) = \left\{x \in X : \sup_{n\geqslant 1} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) > \alpha\right\}, \tag{14}$$

então temos diretamente da definição de $\limsup$ que $E_{\alpha,\beta}(f) \cap B_\alpha(f) = E_{\alpha,\beta}(f)$. Já que provamos que $\mu(E_{\alpha,\beta}) < +\infty$, então podemos aplicar o Corolário 11 para garantir que

$$\int_{E_{\alpha,\beta}(f)} f\,d\mu = \int_{E_{\alpha,\beta}(f)\cap B_\alpha(f)} f\,d\mu \geqslant \alpha\,\mu(E_{\alpha,\beta}(f) \cap B_\alpha(f)) = \alpha\,\mu(E_{\alpha,\beta}(f)).$$

Portanto

$$\int_{E_{\alpha,\beta}(f)} f\,d\mu \geqslant \alpha\,\mu(E_{\alpha,\beta}(f)). \tag{15}$$

Usando a a desigualdade (15) com $(-f)$ no lugar de f, $(-\beta)$ no lugar de α e $-\alpha$ no lugar de β e também a identidade (12), concluímos que

$$\int_{E_{-\beta,-\alpha}(-f)} (-f)\,d\mu \geqslant (-\beta)\mu(E_{-\beta,-\alpha}(f)) \Longrightarrow \int_{E_{\alpha,\beta}(f)} f\,d\mu \leqslant \beta\,\mu(E_{\alpha,\beta}). \tag{16}$$

Portanto segue das desigualdades (15) e (16) que $\alpha\,\mu(E_{\alpha,\beta}) \leqslant \beta\,\mu(E_{\alpha,\beta})$. Lembrando que $\mu(E_{\alpha,\beta}) < +\infty$ e que $\beta < \alpha$, segue da desigualdade acima que $\mu(E_{\alpha,\beta}) = 0$. Usando este fato em (13) e a σ-subaditividade da medida concluímos finalmente que

$$f^* = f_* \quad \mu - \text{qtp} \tag{17}$$

e consequentemente que existe o limite

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) = f^*(x) \qquad \mu - \text{qtp}. \tag{18}$$

Parte 3. Para mostrar que $f^* \in L^1(\mu)$. Vamos usar o Lema de Fatou. Considere a sequência de funções mensuráveis não-negativas

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

Note que segue da desigualdade triangular e de μ ser T-invariante que

$$\int_X g_n\,d\mu \leqslant \int_X |f|\,d\mu.$$

Então segue da convergência μ-qtp obtida em (18), do Lema de Fatou e da desigualdade acima que vale a seguinte desigualdade

$$\begin{aligned}
\int_X |f^*|\, d\mu &= \int_X \lim_{n\to\infty} \left| \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \right| d\mu \\
&= \int_X \liminf_{n\to\infty} \left| \frac{1}{n}\sum_{i=0}^{n-1} f(T^i(x)) \right| d\mu \\
&\leqslant \liminf_{n\to\infty} \int_X g_n\, d\mu \\
&\leqslant \int_X |f|\, d\mu < +\infty \qquad\qquad \Longrightarrow \qquad f^* \in L^1(\mu).
\end{aligned}$$

<u>Parte 4.</u> Nesta parte da prova vamos assumir que $\mu(X) < +\infty$ e vamos mostrar que

$$\int_X f^*\, d\mu = \int_X f\, d\mu.$$

Defina, para cada $k \in \mathbb{Z}$ e $n \geqslant 1$,

$$D_k^n \equiv \left\{ x \in X : \frac{k}{n} \leqslant f^*(x) < \frac{k+1}{n} \right\}.$$

Note que para todo $y \in D_k^n$ temos

$$\begin{aligned}
\frac{k}{n} \leqslant f^*(y) < \frac{k+1}{n} \qquad \Longrightarrow \qquad \frac{k}{n} &\leqslant \limsup_{m\to\infty} \frac{1}{m}\sum_{i=0}^{m-1} f(T^i y) \\
&= \inf_{m\in\mathbb{N}} \sup_{j\geqslant m} \frac{1}{j}\sum_{i=0}^{j-1} f(T^i y) \\
&\leqslant \sup_{m\in\mathbb{N}} \frac{1}{m}\sum_{i=0}^{m-1} f(T^i y)
\end{aligned}$$

Desta forma para qualquer $\varepsilon > 0$ fixado, independentemente das escolhas de k e n, temos das desigualdades acima a seguintes desigualdades

$$\left(\frac{k}{n} - \varepsilon\right) < \frac{k}{n} \leqslant \sup_{m\in\mathbb{N}} \frac{1}{m}\sum_{i=0}^{m-1} f(T^i y), \qquad \forall y \in D_k^n.$$

Deste fato e da definição dos conjuntos B_α's dada em (14) temos

$$y \in D_k^n \quad \Longrightarrow \quad y \in \left\{ x \in X : \sup_{m\geqslant 1} \frac{1}{m}\sum_{i=0}^{m-1} f(T^i x) > \frac{k}{n} - \varepsilon \right\} \equiv B_{\frac{k}{n}-\varepsilon}.$$

O que mostra que $D_k^n \subseteq B_{\frac{k}{n}-\varepsilon}$. Portanto $D_k^n \cap B_{\frac{k}{n}-\varepsilon} = D_k^n$. Usando o fato demonstrado anteriormente que f^* é T-invariante, isto é, $f^* \circ T(x) = f^*(x)$ para todo $x \in X$ temos que

$T^{-1}(D_k^n) = D_k^n$. De fato,

$$D_k^n \equiv (f^*)^{-1}\left(\left[\frac{k}{n}, \frac{k+1}{n}\right)\right) = (f^* \circ T)^{-1}\left(\left[\frac{k}{n}, \frac{k+1}{n}\right)\right) = T^{-1}\left((f^*)^{-1}\left(\left[\frac{k}{n}, \frac{k+1}{n}\right)\right)\right)$$
$$= T^{-1}(D_k^n).$$

Como estamos assumindo que $\mu(X) < +\infty$, segue que $\mu(D_k^n) < +\infty$. Dessa finitude da medida de D_k^n e da igualdade obtida acima, $T^{-1}(D_k^n) = D_k^n$, podemos afirmar que todas as hipóteses do Corolário 11 são satisfeitas. Aplicando este corolário e lembrando do fato provado acima $D_k^n \cap B_{\frac{k}{n}-\varepsilon} = D_k^n$, obtemos a seguinte estimativa:

$$\int_{D_k^n} f \, d\mu = \int_{B_{\frac{k}{n}-\varepsilon} \cap D_k^n} f \, d\mu \;\geqslant\; \left(\frac{k}{n} - \varepsilon\right) \mu(D_k^n).$$

Como a desigualdade acima é verdadeira para todo $\varepsilon > 0$, independentemente de k e n, podemos concluir que

$$\frac{k}{n}\, \mu(D_k^n) \leqslant \int_{D_k^n} f \, d\mu.$$

Usando a definição de D_k^n, para estimar f^* superiormente, e a desigualdade acima obtemos

$$\int_{D_k^n} f^* \, d\mu \leqslant \frac{k+1}{n}\, \mu(D_k^n) \leqslant \frac{1}{n}\, \mu(D_k^n) + \int_{D_k^n} f \, d\mu.$$

Somando sobre $k \in \mathbb{Z}$, ambos lados da desigualdade acima e usando que a família $\{D_k^n\}_{k\in\mathbb{Z}}$ é formada por conjuntos dois-a-dois disjuntos e o Teorema da Convergência Dominada, que pode ser aplicado pois $f, f^* \in L^1(\mu)$, obtemos a seguinte desigualdade

$$\begin{aligned}
\int_X f^* \, d\mu &= \int_{\bigcup\limits_{k\in\mathbb{Z}} D_k^n} f^* \, d\mu \\
&= \sum_{k\in\mathbb{Z}} \int_{D_k^n} f^* \, d\mu \\
&\leqslant \sum_{k\in\mathbb{Z}} \left(\frac{1}{n}\, \mu(D_k^n) + \int_{D_k^n} f \, d\mu\right) \\
&= \frac{\mu(X)}{n} + \int_X f \, d\mu, \qquad\qquad \forall n \in \mathbb{N}.
\end{aligned}$$

Usando mais uma vez a hipótese $\mu(X) < +\infty$, concluímos da desigualdade acima que

$$\int_X f^* \, d\mu \leqslant \int_X f \, d\mu. \tag{19}$$

Repetindo o argumento acima para a função $(-f)$ no lugar de f, ficamos com

$$\int_X (-f)^* \, d\mu \leqslant \int_X -f \, d\mu \implies -\int_X f_* \, d\mu \leqslant -\int_X f \, d\mu \implies \int_X f \, d\mu \leqslant \int_X f_* \, d\mu. \tag{20}$$

Por outro lado, provamos em (17) que $f_* = f^*$ μ-qtp. Desta observação e das desigualdades (19) e (20) concluímos, finalmente que

$$\int_X f^* \, d\mu = \int_X f \, d\mu. \tag{21}$$

<u>Parte 5.</u> Seja $(X, \mathscr{B}, \mu)$ um espaço de medida σ-finito arbitrário. Para encerrar a prova do teorema resta mostrar que

$$\int_W f^* \, d\mu = \int_W f \, d\mu, \qquad \forall f \in L^1(X, \mathscr{B}, \mu)$$

e para todo $W \in \mathscr{B}$ satisfazendo $T^{-1}(W) = W$ e $\mu(W) < +\infty$. Para isto consideramos a transformação $S \equiv T|_W : W \to W$ definida sobre o espaço de medida $(W, \mathscr{F}, \nu)$, onde a σ-álgebra $\mathscr{F} \equiv \mathscr{B} \cap W$ e a medida $\nu(A) \equiv \mu(A \cap W)$, para todo $A \in \mathscr{F}$.

Primeiro, observe que a transformação $S : W \to W$ preserva a medida ν. De fato, seja $B \subseteq W$ tal que $B \in \mathscr{F}$. É claro que, neste caso, $T^{-1}(B) = S^{-1}(B)$ e como estamos assumindo que $T^{-1}(W) = W$ temos portanto que

$$\begin{aligned}
\nu(S^{-1}(B)) = \mu(S^{-1}(B) \cap W) &= \mu(T^{-1}(B) \cap W) \\
&= \mu(T^{-1}(B) \cap T^{-1}(W)) \\
&= \mu(T^{-1}(B \cap W)) = \mu(B \cap W) = \nu(B).
\end{aligned}$$

Dada $f \in L^1(X, \mathscr{B}, \mu)$ considere a função $\varphi : W \to \mathbb{R} \cup \{\pm\infty\}$ definida da seguinte maneira $\varphi \equiv f|_W \cdot \chi_W$. Embora o fator χ_W na definição desta função seja supérfluo decidimos acrescentá-lo por questão de clareza.

Observe que podemos verificar usando funções simples e as propriedades elementares da integral de Lebesgue que

$$\int_W \varphi \, d\nu = \int_X f \cdot \chi_W \, d\mu \qquad \text{e} \qquad \int_W |\varphi| \, d\nu = \int_X |f| \cdot \chi_W \, d\mu < +\infty.$$

Outra observação simples e importante é que para todo $x \in W$ temos $S^i(x) = T^i(x)$, $\chi_W(T^i x) = 1$ e consequentemente

$$\varphi^*(x) = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \varphi(S^i x) = \limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_W(T^i x) f(T^i x) = f^*(x).$$

Como $(W, \mathscr{F}, \nu)$ é um espaço de medida finita e a transformação $S : W \to W$ preserva medida, segue da conclusão da Parte 4 e das duas últimas identidades estabelecidas que

$$\int_W f \, d\mu = \int_W \varphi \, d\nu = \int_W \varphi^* \, d\nu = \int_W f^* \, d\mu. \tag{22}$$

O que encerra a demonstração do teorema. ■

Referências

[1] G. Andrade and L. Cioletti. O teorema da recorrência de poincaré. 2026. MAT-UnB. Lecture Notes in Ergodic Theory.

[2] V. I. Arnold and A. Avez. *Ergodic problems of classical mechanics*. W. A. Benjamin, Inc., New York-Amsterdam, 1968. Translated from the French by A. Avez.

[3] P. Billingsley. *Ergodic theory and information.* John Wiley & Sons, Inc., New York-London-Sydney, 1965.

[4] Karl Petersen. *Ergodic Theory*, volume 2 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1983.

[5] P. Walters. *An introduction to ergodic theory.*, volume 7 of *Graduate Texts in Mathematics.* Springer-Verlag, New York-Berlin., 1982.