UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

YURI GOMES LIMA

TEORIA ERGÓDICA:
UMA INTRODUÇÃO E APLICAÇÕES À TEORIA DOS NÚMEROS

FORTALEZA
2008

YURI GOMES LIMA

TEORIA ERGÓDICA:
UMA INTRODUÇÃO E APLICAÇÕES À TEORIA DOS NÚMEROS

Dissertação apresentada ao Programa de Pós-graduação em Matemática do Departamento de Matemática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Mestre em Matemática. Área de concentração: Teoria Ergódica.

Orientador: Prof. Dr. Abdênago Alves de Barros
Coorientador: Prof. Dr. Krerley Irraciel Martins Oliveira.

FORTALEZA
2008

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

L711t Lima, Yuri Gomes.
Teoria Ergódica : uma introdução e aplicações à teoria dos números / Yuri Gomes Lima. – 2008.
87 f. : il.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Bioquímica, Fortaleza, 2008.
Orientação: Prof. Dr. Abdênago Alves de Barros.
Coorientação: Prof. Dr. Krerley Irraciel Martins Oliveira.

1. Teorema de Birkhoff. 2. Teorema de Van der Waerden. 3. Teorema de Szemerédi. 4. Teorema de Green-Tao. I. Título.

CDD 572

YURI GOMES LIMA

TEORIA ERGÓDICA:
UMA INTRODUÇÃO E APLICAÇÕES À TEORIA DOS NÚMEROS

Dissertação apresentada ao Programa de Pós-graduação em Matemática do Departamento de Matemática da Universidade Federal do Ceará, como parte dos requisitos necessários para a obtenção do título de Mestre em Matemática. Área de concentração: Teoria Ergódica.

Aprovoda em: 23/06/2008.

BANCA EXAMINADORA

Prof. Dr. Abdênago Alves de Barros (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Antonio Caminha Muniz Neto
Universidade Federal do Ceará (UFC)

Prof. Dr. Augusto Armando de Castro Junior
Universidade Federal da Bahia (UFBA)

Aos meus pais Glaucia, Raul e Gilberto e irmãos Patrick e Gisele.

AGRADECIMENTOS

O início da minha carreira acadêmica aconteceu nas Olimpíadas de Matemática, na qual fui incentivado pelo professor Francisco Martins de Paiva (Max) e treinado pelos professores Antonio Caminha Muniz Neto e Onofre Campos da Silva Farias. É uma honra tê-los, hoje em dia, como amigos. Chegou o vestibular e eu não tinha o convencimento suficiente para me tornar matemático. Passei no Instituto Militar de Engenharia e, em três meses, descobri, mais uma vez, que meu futuro era na Matemática. Agradeço a meu orientador Abdênago Alves de Barros, que propiciou as medidas burocráticas necessárias para que eu voltasse do Rio de Janeiro e ingressasse na Universidade Federal do Ceará.

Durante a graduação, tive a ajuda de vários professores, dentre os quais destaco: Abdênago Barros, Antonio Caminha, Francisco Pimentel, João Lucas Barbosa, José Fábio Montenegro, Levi Lopes de Lima, Luquésio Petrola, Plácido Andrade e Romildo da Silva; e colegas: Bruno Holanda, Carlos Augusto, Cícero Thiago, Davi Máximo, Elivaldo Macedo, Emanuel Carneiro, Carpegiani Borges, Gleydson Ricarte, Jefferson Leite, Jobson de Queiroz, Juscelino Silva, Marcelo Melo, Marcos Melo, Paulo Alexandre, Samuel Barbosa, Thiago Barros e Tony Souza.

Agradeço o suporte financeiro dado pelo Conselho Nacional de Pesquisa e Desenvolvimento, CNPq, e à secretária da Pós-Graduação Andréa Costa Dantas.

Em 2007, iniciei o doutorado no Instituto Nacional de Matemática Pura e Aplicada, IMPA, apoiado pelos pesquisadores Carlos Gustavo Moreira e Marcelo Viana, esse último possibilitando a interação com meu co-orientador, Krerley Oliveira, a quem também estendo meus agradecimentos, tanto pela valiosa orientação, mesmo à distância, quanto pelas dicas da vida.

Fora do âmbito matemático, eu gostaria de agradecer a André Almeida, Bruno Cals, Caio Henrique, Clarissa Lima, João Marcelo, José Ricardo, Natália Diniz, Renan Furtado e Rodrigo Nobre. Por fim, claro, agradeço a meus pais Glaucia, Raul e Gilberto, irmãos Gisele e Patrick, tias Fatinha e Gardênia e primos Gabriela, Harley, Igor, Mardey e Rebeca, que me dão a força e o suporte psicológico necessários para que eu enfrente os desafios da vida e compartilhe seus bons momentos.

De manhã escureço
De dia tardo
De tarde anoiteço
De noite ardo.

A oeste a morte
Contra quem vivo
Do sul cativo
O este é meu norte.

Outros que contem
Passo por passo:
Eu morro ontem

Nasço amanhã
Ando onde há espaço:
- Meu tempo é quando."

"Com as lágrimas do tempo
E a cal do meu dia
Eu fiz o cimento
Da minha poesia.

E na perspectiva
Da vida futura
Ergui em carne viva
Sua arquitetura.

Não sei bem se é casa
Se é torre ou se é templo:
(Um templo com Deus.)

Mas é grande e clara
Pertence ao seu tempo
- Entrai, irmãos meus!"

Poética I e Poética II, de Vinícius de Moraes.

RESUMO

Em 2006, Ben Green e Terence Tao provaram que existem progressões aritméticas de tamanho arbitrário formadas somente por primos. As idéias aplicadas na prova não foram da Teoria dos Números e sim adaptações de idéias da Teoria Ergódica e Análise Harmônica. Consequentemente, a Teoria Ergódica voltou a chamar a atenção da comunidade científica, devido à sua grande interação com outras áreas. Esse trabalho desenvolve essa perspectiva, partindo de ferramentas básicas da Teoria Ergódica a aplicações à Teoria dos Números. O Capítulo 1 trata da Teoria da Medida e fixa notações para o que segue, além de discutir exemplos. No Capítulo 2, provamos a principal ferramenta para o que segue, que é o Teorema Ergódica de Birkhoff. Ele possibilita obtermos várias propriedades interessantes da expansão decimal e da representação em frações contínuas dos números reais. O Capítulo 3 trata da interação da Teoria Ergódica com a Teoria dos Números, cujo principal teorema aqui demonstrado é o de Van der Waerden. Esse é um resultado puramente combinatório. Aqui, ele será obtido por intermédio do Teorema Múltiplo de Birkhoff. Por fim, fazemos um "sketch" dos principais resultados da área.

Palavras-chave: Teorema de Birkhoff. Teorema de Van der Waerden. Teorema de Szemerédi. Teorema de Green-Tao.

ABSTRACT

In 2006, Ben Green and Terence Tao proved that the prime numbers contain arbitrarily large arithmetic progressions. The ideas of the proof come from Ergodic Theory and Harmonic Analysis and, since then, Ergodic Theory has attracted attention from the scientific community, because of its large interaction with other fields of Mathematics. This work develops this perspective, beginning from basic Ergodic Theory and arising at applications to Number Theory. Chapter 1 develops the necessary results from Measure Theory, introduces the basic examples and fixes notations. In Chapter 2, we prove the main tool to what follows, which is Birkhoff's Ergodic Theorem. As a consequence of this theorem, we obtain many properties about the decimal expansion and continued fractions of the real numbers. Chapter 3 relates Ergodic Theory and Number Theory, and the main result here presented is Van der Waerden's Theorem. This is a purely combinatorial result. Here, it will be proved as a consequences of the Multiple Birkhoff Theorem. In the end of this work, we sketch the main results in the area.

Keywords: Birkhoff's Theorem. Van der Waerden's Theorem. Szemerédi's Theorem. Green-Tao's Theorem.

SUMÁRIO

1 INTRODUÇÃO

1.1 Preliminares

Considere um número real α. Vamos analisar o comportamento distribucional do conjunto

$$G_\alpha = \{m + n\alpha \,|\, m, n \in \mathbb{Z}\}.$$

Quando α é racional, digamos

$$\alpha = \frac{p}{q} \ , \quad \mathrm{mdc}(p, q) = 1,$$

então

$$m + n\alpha = m + n \cdot \frac{p}{q} = \frac{mq + np}{q} \cdot$$

Como p e q são primos entre si, temos do Teorema de Bézout que a expressão $mq + np$ assume qualquer valor inteiro e portanto

$$G_\alpha = \left\{ \cdots, -\frac{2}{q}, -\frac{1}{q}, 0, \frac{1}{q}, \frac{2}{q}, \cdots \right\}$$

é um conjunto discreto. Quando α é irracional, ocorre algo diferente.

Teorema 1.1 (Kronecker) *Se α é irracional, então G_α é denso em $\mathbb{R}$.*

Prova. Note que G_α é um subgrupo aditivo de $\mathbb{R}$, isto é,

$$g_1, g_2 \in G_\alpha \implies g_1 + g_2 \, , \, g_1 - g_2 \in G_\alpha.$$

Lema 1.1 *Se G é um subgrupo aditivo de $\mathbb{R}$, então uma das situações ocorre:*
(i) G é denso em $\mathbb{R}$.
(ii) Existe $g_0 \in G$ tal que $G = \{pg_0 \,|\, p \in \mathbb{Z}\}$.

Demonstração. G é denso em $\mathbb{R}$ se e somente se 0 é ponto de acumulação de G. Tome

$$g_0 = \inf(G \cap (0, +\infty)).$$

Se $g_0 = 0$, então 0 é ponto de acumulação de G e portanto ocorre (i). Suponha $g_0 > 0$. Afirmamos que $g_0 \in G$. Por absurdo, se $g_0 \notin G$, existem $g_1, g_2 \in G$ tais que

$$g_0 < g_1 < g_2 < 2g_0$$

e daí $g_2 - g_1$ é um elemento de $G \cap (0, +\infty)$ menor que g_0, absurdo. Assim, $g_0 \in G$, o que estabelece a inclusão

$$G \supseteq \{pg_0 \,|\, p \in \mathbb{Z}\}.$$

Reciprocamente, se existe $g \in G$ tal que

$$pg_0 < g < (p+1)g_0, \quad p \in \mathbb{Z},$$

então

$$g - pg_0 \in G \cap (0, +\infty) \cap (0, g_0),$$

novamente contrariando a minimalidade de g_0. ■

Para concluir a prova do Teorema 1.1, basta observar que G_α não satisfaz (ii). De fato, suponha que $g_0 = m_0 + n_0\alpha$ satisfaça

$$G = \{pg_0 \,|\, p \in \mathbb{Z}\}.$$

Como $1, \alpha \in G$, temos $m_0 n_0 \neq 0$. Por outro lado, existe $p_0 \in \mathbb{Z}$ tal que

$$-m_0 + n_0\alpha = p_0(m_0 + n_0\alpha) \implies \begin{cases} -m_0 &= p_0 m_0 \\ n_0 &= p_0 n_0 \end{cases} \implies \begin{cases} p_0 &= -1 \\ p_0 &= 1 \end{cases},$$

absurdo. ■

Perguntas naturais surgem:

- Como é a distribuição dessa sequência?
- Existem regiões em que ela se concentra mais do que em outras?
- Como é a distribuição dessa sequência?
- Existem regiões em que ela se concentra mais do que em outras?

A partir daqui, consideramos as partes fracionárias

$$\{n\alpha - \lfloor n\alpha \rfloor \,|\, n \in \mathbb{Z}\}$$

dos elementos de G_α, ou seja, analisamos a sequência módulo 1. Por simplicidade, continuaremos utilizando a mesma notação.

Definição 1.1 *Dizemos que uma sequência $(x_n)_{n\in\mathbb{N}} \subset \mathbb{S}^1$ é equidistribuída se, para quaisquer $0 \leq a < b \leq 1$, temos*

$$\lim_{m\to+\infty} \frac{|\{1 \leq i \leq m\,;\, x_i \in [a,b]\}|}{m} = b - a.$$

Teorema 1.2 (Hardy-Littlewood) *Se α é irracional, então $(n\alpha)_{n\in\mathbb{Z}}$ é uma sequência equidistribuída.*

Prova. Vamos introduzir notações. Dado um intervalo $I = [a, b]$, sejam

$$\lambda(I) = b - a$$

a medida de Lebesgue em $\mathbb{S}^1$ e

$$\tau(I) = \lim_{m\to+\infty} \frac{|\{1 \le i \le m\,;\, i\alpha \in [a, b]\}|}{m}\,.$$

Queremos mostrar que $\tau(I) = \lambda(I)$. Note que, para cada $k \in \mathbb{Z}$, temos

$$\begin{aligned}
\tau(I + k\alpha) &= \lim_{m\to+\infty} \frac{|\{1 \le i \le m\,;\, i\alpha \in [a + k\alpha, b + k\alpha]\}|}{m} \\
&= \lim_{m\to+\infty} \frac{|\{1 \le i \le m\,;\, (i-k)\alpha \in [a, b]\}|}{m} \\
&= \lim_{m\to+\infty} \frac{|\{1-k \le i \le m-k\,;\, i\alpha \in [a, b]\}|}{m} \\
&= \lim_{m\to+\infty} \frac{|\{1 \le i \le m\,;\, i\alpha \in [a, b]\}|}{m} \\
&= \tau(I).
\end{aligned}$$

Suponha que I satisfaça a desigualdade

$$\lambda(I) < \frac{1}{n}\,\cdot$$

Pelo Teorema 1.1, existem inteiros $k_1 = 0, k_2, \ldots, k_n$ tais que

$$(I + k_i\alpha) \cap (I + k_j\alpha) \neq \emptyset,\ \forall\, 1 \le i < j \le n.$$

Para isso, aproxime b por $a + k_2\alpha$, $b + k_2\alpha$ por $a + k_3\alpha, \ldots, b + k_{n-1}\alpha$ por $a + k_n\alpha$. Como

$n\lambda(I) < 1$, o argumento pode ser feito de modo que os intervalos são disjuntos. Daí,

$$\begin{aligned}
\tau\left(\bigcup_{i=1}^{n}(I+k_i\alpha)\right) &\leq \tau(\mathbb{S}^1) \\
\sum_{i=1}^{n}\tau(I+k_i\alpha) &\leq 1 \\
n\tau(I) &\leq 1 \\
\tau(I) &\leq \frac{1}{n}.
\end{aligned}$$

Agora tome $I \subseteq \mathbb{S}^1$ qualquer e frações $m_k/n_k > \lambda(I)$ tais que

$$\lim_{k\to+\infty}\frac{m_k}{n_k} = \lambda(I).$$

Para cada $k \geq 1$, particione I em m_k intervalos $J_1^k, \ldots, J_{m_k}^k$, cada um deles com comprimento menor que $1/n_k$. Assim,

$$\begin{aligned}
\tau(I) &= \sum_{l=1}^{m_k}\tau(J_l^k) \\
&\leq \sum_{l=1}^{m_k}\frac{1}{n_k} \\
&= \frac{m_k}{n_k}.
\end{aligned}$$

Passando o limite, obtemos

$$\tau(I) \leq \lambda(I).$$

Mas então, como

$$1 = \tau(I) + \tau(I^c) \leq \lambda(I) + \lambda(I^c) = 1,$$

obtemos a igualdade

$$\tau(I) = \lambda(I),$$

o que conclui a prova. ■

E se tomarmos sequências não-lineares do tipo

$$\alpha, 2^k\alpha, \ldots, n^k\alpha, \ldots$$

onde k é um inteiro positivo? Essas sequências são equidistribuídas ? Novamente, a resposta é afirmativa, conforme o teorema abaixo, cuja demonstração foge da meta da dissertação.

Teorema 1.3 (Weyl, 1916) *Se*

$$P(x) = a_k x^k + \cdots + a_1 x + a_0$$

é um polinômio de coeficientes reais, com ao menos um deles irracional, então a sequência

$$(P(n))_{n \in \mathbb{N}} \subseteq \mathbb{S}^1$$

é equidistribuída.

Os resultados acima são da Teoria dos Números e suas provas clássicas usaram métodos da área. Por outro lado, eles também podem ser provados usando idéias da Teoria Ergódica. É isso que faremos no capítulo 3. Para tanto, vamos desenvolver os resultados básicos, que são: Teorema de Recorrência de Poincaré, Teorema Ergódico Médio e Teorema Ergódico de Birkhoff.

1.2 Medidas Invariantes

Sejam $(X, \mathcal{B}, \mu)$ um espaço de probabilidade e $T : (X, \mathcal{B}, \mu) \to (X, \mathcal{B}, \mu)$ uma transformação mensurável, isto é, tal que

$$T^{-1}(A) \in \mathcal{B}, \ \forall\, A \in \mathcal{B}.$$

Definição 1.2 *Dizemos que μ é uma probabilidade invariante por T se*

$$\mu(T^{-1}A) = \mu(A), \quad \forall\, A \in \mathcal{B}.$$

Diremos também que μ é T-invariante ou que T é μ-invariante.

Em geral, é difícil mostrar diretamente que uma medida é invariante, em particular pela existência de conjuntos mensuráveis "estranhos". Felizmente, existe um critério que simplifica essa verificação e essa será a maneira canônica de provar a invariância de medidas. Utilizaremos o seguinte resultado clássico da Teoria da Medida.

Teorema 1.4 (Extensão de Hahn-Kolmogorov) *Sejam X um conjunto, $\mathcal{B}_0$ uma álgebra de subconjuntos de X e $\mu_0 : \mathcal{B}_0 \to [0,1]$ uma medida. Então, se $\mathcal{B}$ é a σ-álgebra gerada por $\mathcal{B}_0$, existe uma e só uma medida $\mu : \mathcal{B} \to [0,1]$ tal que*

$$\mu|_{\mathcal{B}_0} = \mu_0.$$

Proposição 1.1 *Dada uma transformação mensurável* $T : (X, \mathcal{B}, \mu) \to (X, \mathcal{B}, \mu)$*, então* μ *é invariante por* T *se e somente se*

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

para todo A *em uma subálgebra geradora de* $\mathcal{B}$*.*

Prova. A ida é clara. Para a volta, seja $\mathcal{B}_0$ uma subálgebra geradora de $\mathcal{B}$ tal que

$$\mu(T^{-1}A) = \mu(A), \ \forall\, A \in \mathcal{B}_0.$$

Pelo Teorema 1.4, a medida $\tilde{\mu} : \mathcal{B} \to [0, 1]$ definida por

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

coincide com μ e portanto μ é T-invariante. ■

Quando X é um espaço topológico, existe uma σ-álgebra natural $\mathcal{B}$, chamada de σ-álgebra de Borel. Ela é gerada pela subálgebra $\mathcal{B}_0$, formada pelas uniões finitas de abertos de X. Cada elemento de $\mathcal{B}$ é chamado de *boreliano*. Pela Proposição 1.1, uma medida só será invariante por uma transformação agindo em X se a medida da pré-imagem de cada aberto for igual à sua medida.

Em geral, quando estiverem subentendidas, diremos apenas que $T : X \to X$ é uma transformação invariante, sem explicitar a σ-álgebra e medida consideradas.

Dada uma transformação $T : X \to X$, sempre existe medida μ invariante por T? No caso afirmativo, quais são suas propriedades?

Sejam $\mathcal{M}(X)$ o conjunto das medidas de probabilidade compatíveis com a σ-álgebra de X e $T_* : \mathcal{M}(X) \to \mathcal{M}(X)$ o operador definido por

$$(T_*\mu)(A) \doteq \mu(T^{-1}A), \ \forall\, \mu \in \mathcal{M}(X), \ \forall\, A \subseteq X \text{ mensurável.}$$

Então o conjunto $\mathcal{M}_T(X)$ das medidas de probabilidade T-invariantes é igual ao conjunto dos pontos fixos de T_*:

$$\mathcal{M}_T(X) = \mathrm{Fix}(T_*)\,.$$

Teorema 1.5 *Se* $T : X \to X$ *é uma transformação contínua, onde* X *é um espaço métrico compacto munido da* σ*-álgebra de Borel, então* $\mathcal{M}_T(X) \neq \emptyset$*.*

Prova. Ver (5), Seção 1.8. ■

Finalizamos essa seção com um resultado que será muito útil adiante.

Proposição 1.2 (Mudança de Variáveis) *Se $T : X \to X$ é mensurável, $f : X \to \mathbb{C}$ é integrável e $\mu \in \mathcal{M}(X)$, então*

$$\int_X (f \circ T) d\mu = \int_X f d(T_* \mu).$$

Prova. A prova consiste em três partes.

1. Se $f = \chi_A$, A mensurável, então $f \circ T = \chi_{T^{-1}A}$, donde

$$\int_X (f \circ T) d\mu = \int_X \chi_{T^{-1}A} d\mu = \mu(T^{-1}A) = \int_X \chi_A d(T_*\mu) = \int_X f d(T_*\mu).$$

Por linearidade, o resultado também vale para combinações lineares de funções características (funções-escada).

2. Se f é não-negativa, consideramos uma sequência crescente (f_n) de funções-escada que aproxima f. Pelo Teorema da Convergência Dominada, temos

$$\int_X (f \circ T) d\mu = \lim_{n\to\infty} \int_X (f_n \circ T) d\mu = \lim_{n\to\infty} \int_X f_n d(T_*\mu) = \int_X f d(T_*\mu).$$

3. Dada f qualquer, escrevemos $f = f^+ - f^-$, onde f^+ e f^- são não-negativas, e aplicamos o caso anterior para f^+ e f^-. ■

1.3 Exemplos de Medidas Invariantes

1.3.1 *Deltas de Dirac*

Considere uma transformação $T : X \to X$ e x_0 um ponto fixo de T. Então o Delta de Dirac em x_0, definido por

$$\begin{aligned} \delta_{x_0}(A) &= 1\,, \quad \text{se } x_0 \in A \\ &= 0\,, \quad \text{se } x_0 \notin A, \end{aligned}$$

é uma probabilidade invariante por T. Mais geralmente, se x_0 é um ponto periódico de T, digamos

$$T^{n_0} x_0 = x_0,$$

então sua órbita suporta uma medida invariante, dada pela média dos Deltas de Dirac ao longo de sua trajetória:

$$\mu = \frac{\delta_{x_0} + \delta_{Tx_0} + \cdots + \delta_{T^{n_0-1}x_0}}{n_0}.$$

Essas medidas só assumem valores positivos em conjuntos $A \subseteq X$ tais que

$$A \cap \{x_0, Tx_0, \ldots, T^{n_0-1}x_0\} \neq \emptyset.$$

Em uma σ-álgebra com "muitos" elementos, poucos serão os conjuntos de medida positiva. Em várias situações, seria bom termos medidas invariantes mais distribuídas no espaço.

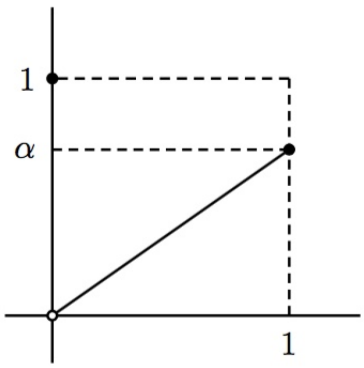
1.3.2 *Transformações sem Medidas Invariantes*

Vamos mostrar que as hipóteses do Teorema 1.5, compacidade do espaço e continuidade da transformação, são necessárias para sua validade, exibindo dois exemplos que não têm medidas invariantes.

Dado $\alpha \in (0,1)$, considere $T_\alpha : [0,1] \to [0,1]$ definida por

$$\begin{aligned} T_\alpha x &= \alpha x\,, \quad \text{se } x \neq 0 \\ &= 1\ \ , \quad \text{se } x = 0, \end{aligned}$$

cuja imagem é o intervalo $(0,\alpha] \cup \{1\}$.



Suponha $\mu \in \mathcal{M}_{T_\alpha}([0,1])$ e tome um intervalo $I \subseteq [0,1]$. Queremos analisar o valor de $\mu(I)$. Para isso, vamos dividir em três casos.

1. $I \subseteq (\alpha, 1)$:

$$\mu(I) = \mu({T_\alpha}^{-1} I) = \mu(\emptyset) = 0.$$

2. $I \subseteq (0, \alpha]$: a restrição

$$T_\alpha|_{(0,1]} : (0, 1] \to (0, \alpha]$$

é uma bijeção e daí

$$\mu(T_\alpha I) = \mu({T_\alpha}^{-1} T_\alpha I) = \mu(I).$$

Como $T_\alpha I \subseteq (0, \alpha]$, podemos repetir o argumento e concluir que

$$\mu(I) = \mu(T_\alpha I) = \mu({T_\alpha}^2 I) = \cdots$$

Por outro lado, se I é suficientemente pequeno, então

$$I, T_\alpha I, {T_\alpha}^2 I, \ldots$$

são dois a dois disjuntos e portanto

$$\begin{aligned} \sum_{k \geq 0} \mu({T_\alpha}^k I) &= \mu\left(\bigcup_{k \geq 0} {T_\alpha}^k I\right) \\ &\leq 1 \end{aligned}$$

$$\begin{aligned} \Longrightarrow \qquad \sum_{k \geq 0} \mu(I) &\leq 1 \\ \Longrightarrow \qquad \mu(I) &= 0. \end{aligned}$$

3. $I = \{1\}$:

$$\mu(I) = \mu({T_\alpha}^{-2} I) = 0.$$

Assim, $\mu([0, 1]) = 0$, contrariando o fato de μ ser uma probabilidade. Veja que, com apenas uma descontinuidade, as medidas invariantes desaparecem, o que mostra a sensibilidade de existência dessas medidas.

O segundo exemplo retira a hipótese de compacidade. Dado $\alpha \in \mathbb{R}$, considere $R_\alpha : \mathbb{R} \to \mathbb{R}$ definida por

$$R_\alpha x = x + \alpha, \quad \forall\, x \in \mathbb{R}.$$

Seja $\mu \in \mathcal{M}_{R_\alpha}(\mathbb{R})$. Se I é um intervalo de comprimento menor que α, então

$$I, {R_\alpha}^{-1} I, {R_\alpha}^{-2} I, \ldots$$

são intervalos dois a dois disjuntos e daí, pelo mesmo argumento utilizado acima, obtemos

$$\mu(I) = 0.$$

Como I é arbitrário, concluímos que $\mathcal{M}_{R_\alpha}(\mathbb{R}) = \emptyset$.

Observe que a medida de Lebesgue λ é invariante por R_α, mas não é uma probabilidade. Por outro lado, podemos induzir R_α no círculo $\mathbb{S}^1$, chamada de *rotação*, e aí λ será uma probabilidade invariante. Quando $\alpha \in \mathbb{Q}$, todo ponto é periódico e daí, pela Subseção anterior, cada um deles está associado a uma medida invariante. Quando $\alpha \notin \mathbb{Q}$, é possível mostrar, utilizando a mesma idéia da prova do Teorema 1.2, que λ é a única probabilidade invariante por R_α.

1.3.3 *Expansão Decimal*

Uma das áreas da Teoria dos Números que mais vem sendo pesquisada (seja pela necessidade de maior precisão nos algoritmos computacionais ou simplesmente pela curiosidade dos matemáticos) é a busca de propriedades interessantes dos números reais, em particular de aproximações cada vez mais apuradas de números relevantes. Por exemplo, de tempos em tempos são anunciadas aproximações maiores de π. Atualmente, a melhor delas possui cerca de um trilhão de casas decimais (sic!).

Para melhor entender essas propriedades, vamos modelar a dinâmica da expansão decimal. Considere $\mathbb{R}$ com a σ-álgebra de Borel e tome $x \in \mathbb{R}$. Como calculamos a primeira casa decimal a_1 de x? É simples: divida $[0,1)$ em dez intervalos

$$A_k = \left[\frac{k}{10}, \frac{k+1}{10}\right), \quad k = 0, 1, \ldots, 9.$$

Então

$$x = a_0, a_1 \cdots \iff x - \lfloor x \rfloor \in A_{a_1}.$$

Tudo bem. Já temos um método de identificar a primeira casa decimal de x. E as outras? Introduzimos uma transformação $f : [0,1) \to [0,1)$ que “continua”o processo, definida por

$$f(x) = 10x - \lfloor 10x \rfloor.$$

Ela “continua”o processo porque se

$$x = 0, a_1 a_2 \cdots ,$$

então

$$f(x) = 0, a_2 a_3 \cdots$$

e daí encontramos a_2 identificando em qual dos intervalos $A_0, \ldots, A_9$ o número $f(x)$ está. Procedendo dessa maneira, obtemos

$$x = 0, a_1 a_2 \cdots a_n \cdots \iff f^n(x) \in A_{a_n},\ \forall\, n \geq 1.$$

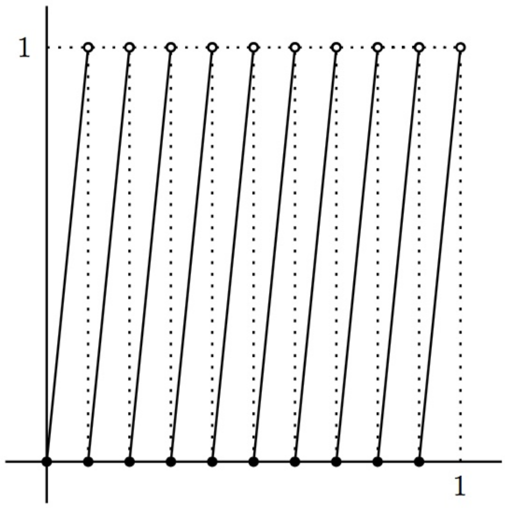


Gráfico de f.

A transformação f possui muitas medidas invariantes. De fato, o conjunto dos pontos periódicos de f é denso em $[0,1)$ e, como visto anteriormente, cada um deles suporta uma medida invariante definida por

$$\mu_x = \frac{\delta_x + \delta_{f(x)} + \cdots + \delta_{f^{n-1}(x)}}{n},$$

onde $f^n(x) = x$ e δ_y representa o delta de Dirac no ponto $y \in [0,1)$. Tais medidas, porém, não enxergam nenhum outro ponto e portanto nada dizem a respeito da generalidade das órbitas.

A medida de Lebesgue λ é f-invariante, pois a imagem inversa de qualquer intervalo $I \subseteq [0,1)$ é formada por dez intervalos, cada um medindo $1/10$ do comprimento de I. Pela Proposição 1.1, isso garante a invariância de λ. Agora sim: essa é uma medida não-atômica que enxerga todo o intervalo $[0,1)$.

1.3.4 *Transformação de Gauss*

Vamos mostrar que todo $x \in \mathbb{R}$ tem uma expansão em frações contínuas

$$x \;=\; a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}} \;\doteq\; [a_0; a_1, a_2, \ldots]\ , \tag{1}$$

possivelmente finita, onde $a_0 \in \mathbb{Z}$ e $a_i \geq 1$, para todo $i \geq 1$ em que a_i está bem-definido.

Se a expansão acima é finita, então x é racional. Por outro lado, se x é racional, o Algoritmo de Euclides garante que a sequência $a_1, a_2, \ldots$ de inteiros positivos é estritamente decrescente e portanto finita. Assim,

Proposição 1.3 *A expansão em frações contínuas de $x \in \mathbb{R}$ é finita se e somente se x é racional.*

E quando $x \notin \mathbb{R}\backslash\mathbb{Q}$: como obter $(a_n)_{n\geq 0}$? É claro que

$$a_0 = \lfloor x \rfloor$$

e daí

$$x = a_0 + b_0, \quad b_0 \in [0,1)\backslash\mathbb{Q}.$$

Então

$$a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}} = \frac{1}{x - a_0} = \frac{1}{b_0}\ .$$

Como

$$\cfrac{1}{a_2 + \cfrac{1}{\ddots}} \leq 1,$$

devemos ter

$$a_1 = \left\lfloor \frac{1}{b_0} \right\rfloor$$

e portanto

$$x = a_0 + \frac{1}{a_1 + b_1}\ , \quad b_1 \in [0,1)\backslash\mathbb{Q}.$$

Em geral, se já temos definidos $a_0, \ldots, a_n$ e

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n + b_n}}}\ , \quad b_n \in [0,1)\backslash\mathbb{Q},$$

tomamos

$$a_{n+1} = \left\lfloor \frac{1}{b_n} \right\rfloor .$$

Dessa forma,

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n + \cfrac{1}{a_{n+1} + b_{n+1}}}}} \quad , \quad b_{n+1} \in [0,1)\backslash\mathbb{Q}.$$

Mostraremos mais adiante que se denotamos a fração

$$[a_0; a_1, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}} \quad ,$$

chamada de n-ésima *convergente* de x, então

$$x = \lim_{n \to +\infty} [a_0; a_1, \ldots, a_n], \tag{2}$$

ou seja, x é igual ao limite de suas convergentes, o que justifica a notação adotada em (1).

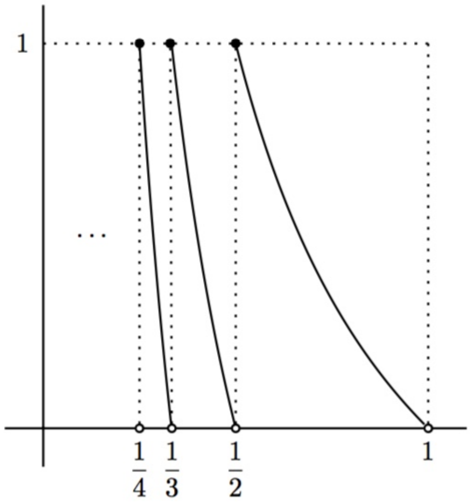
Assim, a dinâmica das frações contínuas é determinada pela sequência $(b_n)_{n \geq 0}$, cujo caráter recursivo é dado por

$$b_n = \frac{1}{a_{n+1} + b_{n+1}} \implies b_{n+1} = \frac{1}{b_n} - \left\lfloor \frac{1}{b_n} \right\rfloor .$$

Definição 1.3 *A função $G : [0,1) \to [0,1)$ definida por*

$$\begin{aligned} G(x) &= \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor \quad , \quad x \neq 0 \\ &= 0 \qquad\qquad\quad , \quad x = 0 \end{aligned}$$

é chamada transformação de Gauss e seu gráfico é dado pela figura abaixo.



Logo,

$$x = \lfloor x \rfloor + \cfrac{1}{\left\lfloor \dfrac{1}{G(x)} \right\rfloor + \cfrac{1}{\left\lfloor \dfrac{1}{G^2(x)} \right\rfloor + \cfrac{1}{\ddots}}} \quad ,$$

onde as reticências indicam que a soma se estende enquanto $G^n(x) > 0$.

Daí, se $J_k = \left(\dfrac{1}{k+1} \, , \, \dfrac{1}{k} \right)$, então

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}} \qquad \Longleftrightarrow \qquad G^n(x) \in J_{a_n}, \ \forall \, n \geq 1 \, ; \, G^n(x) > 0. \tag{3}$$

$$\tag{4}$$

A importância das frações contínuas na Teoria dos Números reside no seguinte fato: se as convergentes são escritas na forma irredutível

$$[a_0; a_1 \ldots, a_n] = \frac{p_n}{q_n} \quad , \quad \mathrm{mdc}(p_n, q_n) = 1,$$

então, dentre todas as frações com denominador no máximo q_n, essa é a que melhor aproxima x. Noutras palavras: se $a, b \in \mathbb{Z}$, $0 < b \leq q_n$, então

$$\left| x - \frac{p_n}{q_n} \right| \leq \left| x - \frac{a}{b} \right| \, .$$

A medida de Lebesgue λ não é invariante por G. De fato, se $0 \leq a < b < 1$, então

$$G^{-1}([a, b)) = \bigcup_{k \geq 1} \left(\frac{1}{k+b} \, , \, \frac{1}{k+a} \right]$$

e daí

$$\begin{aligned}
\lambda(G^{-1}([a,b))) &= \sum_{k\geq 1}\lambda\left(\left(\frac{1}{k+b}\ ,\ \frac{1}{k+a}\right]\right)\\
&= \sum_{k\geq 1}\left(\frac{1}{k+a}-\frac{1}{k+b}\right)\\
&= (b-a)\cdot\sum_{n\geq 1}\frac{1}{(k+a)(k+b)}\ ,
\end{aligned}$$

que em geral difere de $b-a$ (tome, por exemplo, $a=0$ e $b=1/2$). Felizmente, G possui uma probabilidade invariante μ, equivalente a λ, definida em cada boreliano $A\subseteq[0,1)$ por

$$\mu(A)=\frac{1}{\log 2}\int_A\frac{1}{1+x}dx.$$

Essa medida foi encontrada por Gauss em 1800 e por isso é chamada de *medida de Gauss*. Ninguém sabe como Gauss a achou e sua descoberta é ainda mais notável se observarmos que a Teoria Ergódica e a Teoria Moderna de Probabilidade desenvolveram-se apenas um século depois!

Provemos a G-invariância de μ. Temos

$$\begin{aligned}
\mu(G^{-1}([a,b))) &= \sum_{k\geq 1}\mu\left(\left(\frac{1}{k+b}\ ,\ \frac{1}{k+a}\right]\right)\\
&= \sum_{k\geq 1}\frac{1}{\log 2}\int_{\frac{1}{k+b}}^{\frac{1}{k+a}}\frac{1}{1+x}dx\\
&= \frac{1}{\log 2}\sum_{k\geq 1}\left[\log\left(1+\frac{1}{k+a}\right)-\log\left(1+\frac{1}{k+b}\right)\right]\\
&= \frac{1}{\log 2}\left[\sum_{k\geq 1}(\log(k+1+a)-\log(k+a))\right.\\
&\qquad\left.-\sum_{k\geq 1}(\log(k+1+b)-\log(k+b))\right]\\
&= \frac{1}{\log 2}\left[\log(1+b)-\log(1+a)\right]\\
&= \mu([a,b)).
\end{aligned}$$

Note que

$$\frac{1}{\log 2}\int_A\frac{1}{2}dx\leq\frac{1}{\log 2}\int_A\frac{1}{1+x}dx\leq\frac{1}{\log 2}\int_A dx$$

e portanto

$$\frac{1}{2\log 2} \cdot \lambda(A) \ \leq \ \mu(A) \ \leq \ \frac{1}{\log 2} \cdot \lambda(A). \tag{5}$$

Assim, se $\mu(A) = 0$, o lado esquerdo de (5) nos dá que $\lambda(A) = 0$ e, reciprocamente, se $\lambda(A) = 0$, o lado direito de (5) dá que $\mu(A) = 0$. Isso mostra que as medidas μ e λ são equivalentes: elas possuem os mesmos conjuntos de medida nula. Para propriedades q.t.p., elas podem ser consideradas iguais, pois os resultados obtidas a partir de uma implicam resultados para a outra e vice-versa.

1.3.5 *Difeomorfismos Lineares no Toro*

Considere uma matriz A de dimensão 2×2, entradas inteiras e determinante 1. Podemos ver A como uma transformação linear

$$\begin{array}{rcccl} A & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & & x & \longmapsto & Ax \end{array}$$

Essa transformação preserva o reticulado de coordenadas inteiras, isto é,

$$Ax \in \mathbb{Z}^2, \ \forall\, x \in \mathbb{Z}^2.$$

Assim, se $\sim$ é a relação de equivalência definida por

$$x \sim y \quad \Longleftrightarrow \quad x - y \in \mathbb{Z}^2$$

e

$$\mathbb{T}^2 = \mathbb{R}^2 / \sim$$

é o toro bidimensional, quociente de $\mathbb{R}^2$ por $\sim$, então A induz uma transformação

$$\widetilde{A} : \mathbb{T}^2 \longrightarrow \mathbb{T}^2.$$

Como A^{-1} também tem entradas inteiras, o mesmo argumento nos dá a transformação

$$\widetilde{A}^{-1} : \mathbb{T}^2 \longrightarrow \mathbb{T}^2,$$

inversa de $\tilde{A}$.

Proposição 1.4 *Sejam $X \subseteq \mathbb{R}^n$ munido da σ-álgebra de Borel e $f : X \to X$ um difeo-*

morfismo tal que

$$\det(Jf(x)) = \pm 1,\ \forall\, x \in X,$$

onde Jf *denota o Jacobiano de* f*. Então a medida de Lebesgue* λ_n *em* $\mathbb{R}^n$ *é invariante por* f*.*

Prova. Como f é invertível, a afirmação da Proposição equivale a

$$\lambda_n(f(A)) = \lambda_n(A),\ \forall\, A \subseteq X.$$

Pela Proposição 1.1, a igualdade acima precisa ser verificada apenas para abertos. Pelo Teorema de Mudança de Variáveis, se $A \subseteq X$ é aberto, então

$$\begin{aligned}
\lambda_n(f(A)) &= \int_{f(A)} d\lambda_n \\
&= \int_A |\det(Jf(x))| d\lambda_n \\
&= \int_A d\lambda_n \\
&= \lambda_n(A),
\end{aligned}$$

o que conclui a prova. ■

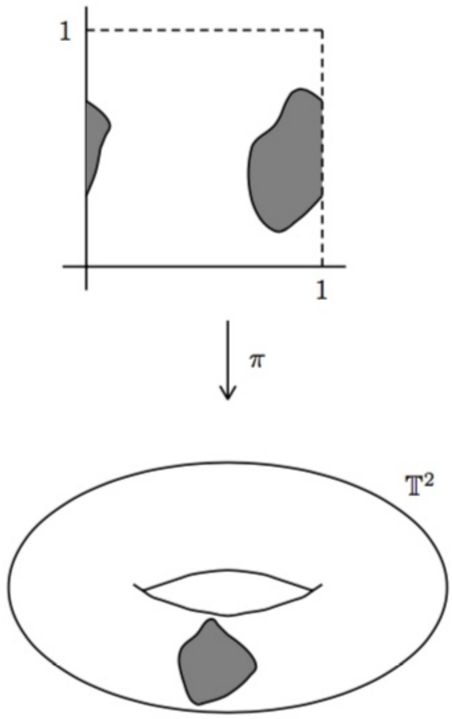
Logo, a medida de Lebesgue $\lambda_2 = \lambda \times \lambda$ em $\mathbb{R}^2$ é invariante por A. Daí, se $\pi : \mathbb{R}^2 \to \mathbb{T}^2$ é a projeção canônica de $\sim$, podemos induzir uma medida $\widetilde{\lambda_2}$ em $\mathbb{T}^2$ invariante por $\tilde{A}$, definida por

$$\widetilde{\lambda_2}(A) = \lambda_2(\pi^{-1}(A) \cap [0,1]^2).$$

O caso que mais aparece na literatura é

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

que induz a transformação $\widetilde{A}$ chamada de *difeomorfismo de Thom-Anosov*. Seus autovalores têm módulo diferente de 1 e portanto existe uma direção de expansão e outra de contração. Quando isso ocorre, dizemos que o difeomorfismo é de Anosov.



1.4 O Teorema de Recorrência de Poincaré

Definição 1.4 *Dadas uma transformação mensurável $T : X \to X$ e um subconjunto A de X, dizemos que um ponto $x \in A$ é recorrente com respeito a A se existir $n > 0$ tal que $T^n x \in A$.*

Teorema 1.6 (Recorrência de Poincaré) *Sejam $T : X \to X$ uma transformação mensurável e $\mu \in \mathcal{M}_T(X)$. Se $A \subseteq X$ é um conjunto mensurável com medida positiva, então quase todo ponto de A é recorrente com respeito a A.*

Prova. Um ponto $x \in A$ não retorna para A se e somente se não pertence à união

$$\bigcup_{k \geq 1} T^{-k} A.$$

Assim, o conjunto dos pontos não-recorrentes com respeito a A é igual a

$$B = A \backslash \bigcup_{k \geq 1} T^{-k} A.$$

Note que

$$T^{-n} B = T^{-n} A \backslash \bigcup_{k \geq n} T^{-k} A \subseteq T^{-n} A,$$

e esse último é disjunto de B. Daí, $B, T^{-1}B, T^{-2}B, \ldots$ são dois a dois disjuntos e portanto, como $\mu(X) < +\infty$, concluímos que $\mu(B) = 0$. ■

Corolário 1.1 *Nas condições do Teorema 1.6, quase todo ponto de A retorna para A infinitas vezes.*

Demonstração. Seja

$$B_k = \{x \in A \,|\, x \text{ possui exatamente } k \text{ iterados em } A\}.$$

Queremos mostrar que $\mu(\cup_{k\geq 1} B_k) = 0$, o que é o mesmo que $\mu(B_k) = 0$, para todo $k \geq 1$. Por absurdo, se $\mu(B_k) > 0$ para algum $k \geq 1$, o teorema anterior garante a existência de $x \in B_k$ e $n \geq 1$ tais que

$$T^n x \in B_k,$$

e daí $T^n x$ possui exatamete k iterados em A. Ora, mas então x possui $k+1$ iterados em A, contrariando o fato de x estar em B_k. ■

Corolário 1.2 *Nas condições do Teorema 1.6, existe um inteiro $n \geq 1$ tal que*

$$\mu(A \cap T^{-n}A) > 0.$$

Demonstração. Como

$$\begin{aligned}\sum_{n\geq 1} \mu(A \cap T^{-n}A) &= \mu\left(A \cap \bigcup_{n\geq 1} T^{-n}A\right) \\ &= \mu(A \backslash B) \\ &= \mu(A)\end{aligned}$$

alguma das parcelas $\mu(A \cap T^{-n}A)$ é positiva. ■

Vamos aplicar os resultados acima a exemplos da seção anterior.

Proposição 1.5 *Quase todo número contendo o algarismo 7 em sua representação decimal contém esse algarismo infinitas vezes.*

Prova. Com a notação da Subseção Expansão Decimal, temos, do Teorema de Recorrência de Poincaré, que λ-q.t.p. $x \in A_7$ retorna infinitas vezes para A_7 o que, noutras palavras, é o que queremos. ■

Observação. O mesmo vale para uma combinação de algarismos $a_k a_{k+1} \cdots a_n$ de tamanho arbitrário.

Assim, o conjunto

$$B = \{x \in [0,1) \,|\, 7 \text{ aparece na representação decimal de } x\}$$

possui um subconjunto C tal que $\lambda(B \backslash C) = 0$ e todo $x \in C$ contém 7 em sua representação decimal infinitas vezes. Se mostrarmos que B tem medida total, então o mesmo ocorrerá para C. Para isso, considere, para cada natural n, o conjunto

$$B_n = \{x = 0.a_1 a_2 \cdots \,|\, n \text{ é o menor natural } k \text{ tal que } a_k = 7\}.$$

Então B é a união disjunta dos B_n e

◦ $B_1 = A_7$.

◦ B_2 é a união de 9 intervalos, cada um medindo $1/100$.

Por indução, B_n é a união de 9^{n-1} intervalos, cada um medindo 10^{-n}. Assim,

$$\begin{aligned} \lambda(B) &= \sum_{n \geq 1} \lambda(B_n) \\ &= \frac{1}{10} + \frac{9}{100} + \cdots + \frac{9^{n-1}}{10^n} + \cdots \\ &= \frac{1}{10}\left[\frac{9}{10} + \frac{9^2}{10^2} + \cdots\right] \\ &= \frac{1}{10} \cdot \frac{1}{1 - \frac{9}{10}} \\ &= 1. \end{aligned}$$

Isso prova o

Corolário 1.3 *Quase todo número possui o algarismo 7 em sua representação decimal infinitas vezes.*

Cabe agora perguntar com qual frequência essa repetição ocorre. E qual das situações mais ocorre: um número ter uma distribuição homogênea dos algarismos ou não. Isso será respondido no próximo capítulo, quando mostraremos que λ é uma medida ergódica para f.

Temos um resultado semelhante para a transformação de Gauss. Abaixo, utilizamos a notação da Subseção correspondente.

Proposição 1.6 *Dado um inteiro positivo* $a \geq 1$*,* λ*-q.t.p.* $x \in \mathbb{R}$ *tal que* $a_{n_0}(x) = a$ *para algum* $n_0 \geq 1$ *satisfaz* $a_n(x) = a$ *para infinitos índices* $n \geq 1$*.*

Prova. De (3), $a_{n_0}(x) = a$ se e somente se $x \in G^{-n_0}(J_a)$. Pelo Teorema de Recorrência de Poincaré, μ-q.t.p. $x \in G^{-n_0}(J_a)$ retorna para $G^{-n_0}(J_a)$ infinitas vezes. Para cada um desses retornos, temos

$$G^m(x) \in G^{-n_0}(J_a) \implies x \in G^{-(n_0+m)}(J_a) \implies a_{n_0+m}(x) = a,$$

provando que a aparece infinitas vezes na expansão em frações contínuas de x. Como μ e λ são equivalentes, a mesma conclusão vale para λ. ■

2 TEOREMAS ERGÓDICOS

Nesse capítulo, discutiremos dois resultados importantes da Teoria Ergódica. Sabemos que, pelo Corolário 1.3, existem conjuntos $C_0, C_1, \ldots, C_9$ tais que

$$\lambda(C_0) = \lambda(C_1) = \cdots = \lambda(C_9) = 1$$

e que todo número em C_k possui o algarismo k em sua representação decimal infinitas vezes. Assim, a interseção

$$C = \bigcap_{0 \leq k \leq 9} C_k$$

é um conjunto de medida total com respeito a Lebesgue e a representação decimal de qualquer dos seus elementos possui repetições infinitas de todos os algarismos, o que nos dá uma boa resposta qualitativa ao problema. A pergunta natural que surge é: quantitativamente, o que ocorre? Mais especificamente, gostaríamos de saber se as frequências desses algarismos são todas iguais. Isso de fato ocorre, conforme veremos nas próximas seções. A ferramenta usada será o Teorema Ergódico de Birkhoff, provado em 1931, que dá condições para uma medida ser *ergódica*, conceito que será explicado mais adiante.

Historicamente, a idéia de ergodicidade surgiu da *Hipótese Ergódica de Boltzmann*. Era esperado que qualquer ponto representando o estado de um sistema físico visitasse toda a sua hipersuperfície de energia constante do espaço de fase de uma maneira equidistribuída. Boltzmann conjecturou mais: que a órbita desse ponto passava por todos os outros pontos da hipersuperfície. Não se demorou a perceber, com o advento da Teoria da Medida por Carathéodory, que isso era impossível, visto que a órbita de um ponto, sendo enumerável, tem medida nula e o mesmo não ocorre em geral com as hipersuperfícies. Foi então proposta uma hipótese mais fraca, chamada de *Hipótese Quasi-Ergódica*, de que cada órbita é densa em sua hipersuperfície de nível com respeito à função energia do sistema. Infelizmente, essa condição não mais garante a meta inicial de que as médias temporal e espacial coincidissem.

Definiremos o que é uma medida ergódica e daremos condições simples equivalentes entre si para que uma medida seja ergódica. Após estabelecidos esses resultados, faremos algumas aplicações aos exemplos do capítulo anterior.

2.1 O Teorema Ergódico Médio

No que se segue, escreveremos L^p para representar o espaço vetorial normado $L^p(X, \mathcal{B}, \mu)$ das funções mensuráveis $f : X \to \mathbb{C}$ que satisfazem

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

Teorema 2.1 (Ergódico Médio) *Sejam $T : X \to X$ uma transformação mensurável e $\mu \in \mathcal{M}_T(X)$. Para cada $f \in L^2$, existe uma função $\tilde{f} \in L^2$ para a qual a sequência*

$$S_n(f) = \frac{1}{n}\sum_{k=1}^{n-1} f \circ T^k$$

converge para $\tilde{f}$ na norma L^2.

Noutras palavras, a média temporal por f da trajetória por T de qualquer ponto de X se estabiliza.

A prova consiste em analisar o operador $U : L^2 \to L^2$ definido por

$$Uf = f \circ T.$$

Note que, da μ-invariância de T e da proposição 1.2, temos

$$\begin{aligned}
\|Uf\|_2{}^2 &= \|f \circ T\|_2{}^2 \\
&= \int (f \circ T)^2 d\mu \\
&= \int (f^2 \circ T) d\mu \\
&= \int f^2 d(T_*\mu) \\
&= \int f^2 d\mu \\
&= \|f\|_2{}^2,
\end{aligned}$$

e portanto U é um operador unitário (bem-definido) em L^2. Em termos de U, o teorema acima é equivalente à existência em L^2 do limite

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} U^k f = \tilde{f}.$$

Provaremos mais: qualquer contração fraca em um espaço de Hilbert satisfaz a igualdade acima. Como L^2, dotado do produto interno

$$(f,g) \doteq \int f\bar{g}d\mu, \quad \forall\, f,g \in L^2,$$

é de Hilbert e U é unitário, o resultado seguirá.

Sejam $\mathcal{H}$ um espaço de Hilbert e $U:\mathcal{H}\to\mathcal{H}$ um operador linear. Considere o conjunto

$$\mathcal{M} = \{f \in \mathcal{H} \,|\, Uf = f\},$$

formado pelos pontos fixos de U. $\mathcal{M}$ é um subespaço vetorial de $\mathcal{H}$, pois se $f,g\in\mathcal{M}$ e α é um escalar, então

$$\begin{aligned} U(f+g) &= Uf + Ug \\ U(\alpha\cdot f) &= \alpha\cdot Uf. \end{aligned}$$

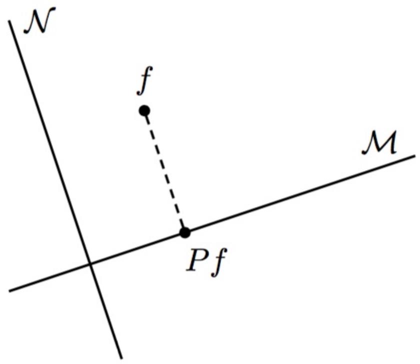
Teorema 2.2 *Sejam $\mathcal{H}$ um espaço de Hilbert e $U:\mathcal{H}\to\mathcal{H}$ um operador linear tal que*

$$\|Uf\| \leq \|f\|\,,\ \forall\, f\in\mathcal{H}.$$

Se $P:\mathcal{H}\to\mathcal{M}$ é a projeção ortogonal de $\mathcal{H}$ em $\mathcal{M}$, então

$$S_n(f) = \frac{1}{n}\sum_{i=0}^{n-1} U^k f \xrightarrow{L^2} Pf,\ \forall\, f\in L^2.$$

Prova.



Daqui em diante, sempre que não houver confusão, escreveremos S_n para representar $S_n(f)$. Seja $(\cdot,\cdot)$ o produto interno de $\mathcal{H}$. Note que U^* também é uma contração fraca, pois

$$\|U^*f\|^2 = (U^*f, U^*f) = (UU^*f, f) \leq \|UU^*f\|\cdot\|f\| \leq \|U^*f\|\cdot\|f\|\,,$$

e daí

$$\|U^*f\| \leq \|f\|\,,\ \forall\, f\in\mathcal{H}.$$

Em particular,

$$Uf = f \iff U^*f = f, \tag{6}$$

uma vez que, se $Uf = f$, então

$$\begin{aligned}
\|f - U^*f\|^2 &= (f - U^*f, f - U^*f) \\
&= (f,f) - (f, U^*f) - (U^*f, f) + (U^*f, U^*f) \\
&= (f,f) - (Uf, f) - (f, Uf) + (U^*f, U^*f) \\
&\leq (f,f) - 2(f,f) + (f,f) \\
&= 0,
\end{aligned}$$

e o mesmo vale se supusermos que $U^*f = f$.

Seja $\mathcal{N} \subseteq \mathcal{H}$ o subespaço fechado gerado pelos elementos da forma $g - Ug$, $g \in \mathcal{H}$. Afirmamos que $\mathcal{N}^\perp = \mathcal{M}$. Para isso, seja $f \in \mathcal{N}^\perp$. Temos

$$0 = (f, g - Ug) = (f,g) - (f, Ug) = (f,g) - (U^*f, g) = (f - U^*f, g).$$

Como $g \in \mathcal{H}$ é arbitrário, temos $U^*f = f$ e de 6 segue que $Uf = f$. Assim, $\mathcal{N}^\perp \subseteq \mathcal{M}$. Reciprocamente, se $f \in \mathcal{M}$, então $U^*f = f$, donde

$$(f, g - Ug) = (f - U^*f, g) = 0 \implies f \in \mathcal{N}^\perp,$$

o que estabelece a inclusão reversa.

A prova da convergência se divide em três casos:

Caso 1. $f = g - Ug$, para algum $g \in \mathcal{H}$.

Temos $Pf = 0$ e

$$\begin{aligned}
\|S_n\| &= \frac{1}{n}\left\|\sum_{k=0}^{n-1}(U^k g - U^{k+1}g)\right\| \\
&= \frac{1}{n}\|g - U^n g\| \\
&\leq \frac{1}{n}(\|g\| + \|U^n g\|) \\
&\leq \frac{2}{n}\|g\|,
\end{aligned}$$

e essa última expressão converge a zero quando $n \to \infty$.

Caso 2. $f \in \mathcal{N}$.

Novamente, $Pf = 0$. Seja $f_i = g_i - Ug_i$, $i \geq 1$, uma sequência convergindo para f. Então

$$\|S_n(f)\| \;=\; \|S_n(f - f_i) + S_n(f_i)\| \;\leq\; \|S_n(f - f_i)\| + \|S_n(f_i)\|\,.$$

Dado $\varepsilon > 0$, tome $i \geq 1$ tal que $\|f - f_i\| < \varepsilon/2$ e $n_0 \geq 1$ tal que $\|S_n(f_i)\| < \varepsilon/2$ para todo $n \geq n_0$. Daí,

$$\|S_n(f - f_i)\| \;\leq\; \frac{1}{n}\sum_{k=0}^{n-1}\left\|U^k(f - f_i)\right\| \;\leq\; \frac{1}{n}\sum_{k=0}^{n-1}\|f - f_i\| \;<\; \frac{\varepsilon}{2}\,,$$

donde

$$\|S_n(f)\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \quad \forall\, n \geq n_0.$$

Caso 3. $f \in \mathcal{H}$ arbitrário.

Como $S_n(Pf) = Pf$ e $f - Pf \in \mathcal{N}$, temos

$$\|S_n(f) - Pf\| = \|S_n(f - Pf)\| \longrightarrow 0$$

pelo caso anterior, o que conclui a prova. ∎

2.2 O Teorema Ergódico de Birkhoff para Transformações

O Teorema Ergódico Médio pode ser refinado para funções em L^1, porém com argumentos mais elaborados, uma vez que L^1 não é um espaço de Hilbert. Esse é o Teorema Ergódico de Birkhoff.

Quando não houver risco de confusão, denotaremos por $S_n(x)$ a função $S_n(f)$ aplicada em x. Defina $f^* : X \to \mathbb{C}$ por

$$f^*(x) = \sup_{n \geq 1} S_n(x).$$

É óbvio que f não é necessariamente positiva quando f^* é. Vale, porém, o seguinte

Teorema 2.3 (Ergódico Maximal) *Se $f \in L^1$, então*

$$\int_{\{f^*>0\}} f d\mu \geq 0.$$

Prova. Para que $f^*(x)$ seja positivo, é necessária e suficiente a existência de um natural mínimo $n \geq 1$ para o qual

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

Então $\{f^* > 0\}$ é a união disjunta dos conjuntos

$$\begin{aligned}
B_1 &= \{x \in X \mid f(x) > 0\} \\
B_2 &= \{x \in X \mid f(x) \leq 0, f(x) + f(Tx) > 0\} \\
&\vdots \\
B_n &= \{x \in X \mid f(x) \leq 0, \ldots, f(x) + \cdots + f(T^{n-2}x) \leq 0, \\
&\qquad f(x) + \cdots + f(T^{n-1}x) > 0\} \\
&\vdots
\end{aligned}$$

O teorema estará provado se mostrarmos que

$$\int_{B_1 \cup \cdots \cup B_n} f d\mu \geq 0,$$

pois daí uma aplicação do Teorema da Convergência Dominada à família (f_n) definida por $f_n = \chi_{B_1 \cup \cdots \cup B_n} f$ garantirá que

$$\int_{\{f^*>0\}} f d\mu = \lim_{n \to \infty} \left(\int_{B_1 \cup \cdots \cup B_n} f d\mu \right) \geq 0.$$

O problema é que não se tem necessariamente a desigualdade $\int_{B_k} f d\mu \geq 0$. O que faremos, então, é decompor $B_1 \cup \cdots \cup B_n$ em uma união disjunta de conjuntos $C_1, \ldots, C_n$, onde cada C_k é a união de iterados disjuntos de um subconjunto de B_k, digamos

$$C_k = B'_k \cup TB'_k \cup \cdots \cup T^{k-1}B'_k, \quad B'_k \subseteq B_k,$$

pois daí

$$\begin{aligned}
\int_{C_k} f d\mu &= \sum_{i=0}^{k-1} \int_{T^i B'_k} f d\mu \\
&= \sum_{i=0}^{k-1} \int_{T^i B'_k} f d((T^i)_* \mu) \\
&= \sum_{i=0}^{k-1} \int_{B'_k} (f \circ T^i) d\mu \\
&= \int_{B'_k} (f + f \circ T + \cdots + f \circ T^{k-1}) d\mu \\
&\geq 0,
\end{aligned}$$

uma vez que o integrando acima é positivo em $B_k \supseteq B'_k$.

Antes de construirmos a decomposição, vejamos duas propriedades da família (B_k).

Afirmação 1. Dados $1 \leq i < j$, tem-se

$$T^i B_j \subseteq B_1 \cup \cdots \cup B_{j-i}.$$

De fato, se $x \in B_j$, então

$$\begin{aligned}
f(x) + \cdots + f(T^{i-1}x) &\leq 0 \\
f(x) + \cdots + f(T^{i-1}x) + f(T^i x) + \cdots + f(T^{j-1}x) &> 0,
\end{aligned}$$

e portanto

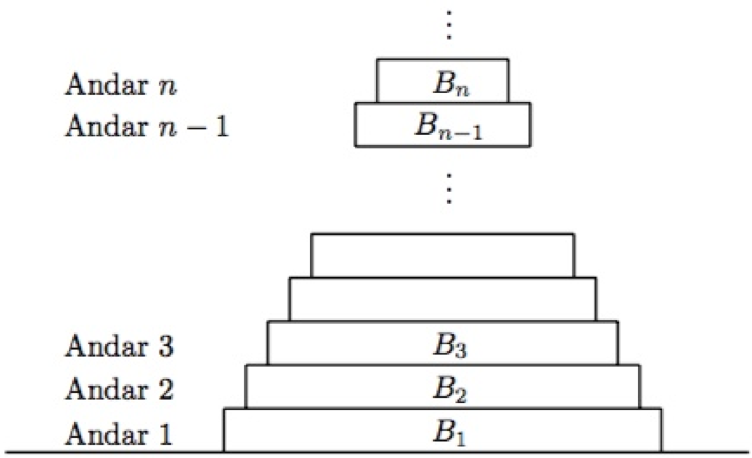
$$\begin{aligned}
f(T^i x) + f(T^{i+1}x) + \cdots + f(T^{j-1}x) &> 0 \\
\Longrightarrow \quad f(T^i x) + f(T T^i x) + \cdots + f(T^{j-i-1} T^i x) &> 0,
\end{aligned}$$

ou seja, $T^i x \in B_k$, para algum $1 \leq k \leq j - i$.

Afirmação 2. Os conjuntos $B_k, TB_k, \ldots, T^{k-1}B_k$ são dois a dois disjuntos.

Por contradição, se existissem $0 \leq i < j \leq k-1$ tais que $T^i B_k \cap T^j B_k \neq \emptyset$, teríamos $B_k \cap T^{j-i} B_k \neq \emptyset$, o que contraria a Afirmação 1, pois $T^{j-i} B_k \subseteq B_1 \cup \cdots \cup B_{k-j+i}$, e essa última união é disjunta de B_k.

Agora estamos prontos para construir a decomposição procurada. Daremos uma interpretação geométrica da dinâmica dos conjuntos $B_1, B_2, \ldots$ para clarificar a idéia.



Considere uma torre na qual o andar j representa o conjunto B_j. A aplicação T é o elevador da torre no seguinte sentido: em cada movimento, o elevador desce pelo menos um andar. Dessa forma, T^iB_j está contido na união dos andares $1, 2, \ldots, j-i$, que é a afirmação 1. O elevador também não passa pelo mesmo lugar duas vezes, que é a afirmação 2. Com essa analogia, tome

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

Note que C_n está contido em $B_1 \cup \cdots \cup B_n$. Defina

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

isto é, B'_{n-1} é a parte do $(n-1)$-ésimo andar que não foi coberta por C_n. Daí, C_{n-1} está contido em $B_1 \cup \cdots \cup B_{n-1}$ e $C_{n-1} \cup C_n \supseteq B_{n-1} \cup B_n$. Procedendo dessa maneira, definimos

$$B'_k = B_k \backslash (C_{k+1} \cup \cdots \cup C_n) \ , \quad C_k = B'_k \cup TB'_k \cup \cdots \cup T^{k-1}B'_k, \quad 1 \leq k \leq n.$$

Os conjuntos $C_1, \ldots, C_n$ cobrem $B_1 \cup \cdots \cup B_n$ e, de fato, são disjuntos dois a dois, uma vez que se tivéssemos $C_i \cap C_j \neq \emptyset$, então um iterado de B'_j intersectaria B'_i ou vice-versa, o que contraria a definição de $B'_1, \ldots, B'_n$. Isso conclui a prova. ∎

Corolário 2.1

$$\int_{\{f^* \geq 0\}} f d\mu \geq 0.$$

Demonstração. Dado $\varepsilon > 0$, temos $(f+\varepsilon)^* = f^* + \varepsilon$, e daí

$$\begin{aligned} 0 &\leq \int_{\{(f+\varepsilon)^*>0\}} (f+\varepsilon)d\mu \\ &= \int_{\{f^*>-\varepsilon\}} f d\mu + \varepsilon \cdot \mu(\{f^* > -\varepsilon\}) \\ \Longrightarrow \quad \int_{\{f^*>-\varepsilon\}} f d\mu &\geq -\varepsilon. \end{aligned}$$

O resultado segue fazendo $\varepsilon_n = 1/n \to 0$ e aplicando o Teorema da Convergência Dominada à sequência de funções $(\chi_{\{f^*>-1/n\}} f)_{n\geq 1}$. ■

Corolário 2.2 *Dado $\alpha \in \mathbb{R}$, tem-se*

$$\int_{\{f^*>\alpha\}} f d\mu \geq \alpha \cdot \mu(\{f^* > \alpha\}).$$

Demonstração. Seja $g = f - \alpha$. Então

$$\begin{aligned} 0 &\leq \int_{\{g^*>0\}} g d\mu \\ &= \int_{\{f^*>\alpha\}} (f-\alpha) d\mu \\ &= \int_{\{f^*>\alpha\}} f d\mu - \alpha \cdot \mu(\{f^* > \alpha\}) \\ \Longrightarrow \quad \int_{\{f^*>\alpha\}} f d\mu &\geq \alpha \cdot \mu(\{f^* > \alpha\}). \end{aligned}$$

■

Definição 2.1 *Dizemos que uma função $f : X \to \mathbb{C}$ é invariante com respeito a T ou simplesmente invariante se*

$$f(Tx) = f(x), \quad \textit{para } \mu\textit{-q.t.p. } x \in X.$$

Diremos que f é constante se o for em um conjunto de medida total.

Teorema 2.4 (Ergódico de Birkhoff) *Sejam $T : X \to X$ uma transformação mensurável e $\mu \in \mathcal{M}_T(X)$. Para cada $f \in L^1$, temos:*

(i) $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) \doteq \tilde{f}(x)$ *existe q.t.p. com respeito a μ.*

(ii) $\tilde{f}$ *é invariante com respeito a T.*

(iii) $\tilde{f} \in L^1$, *com* $\left\|\tilde{f}\right\|_1 \leq \|f\|_1$.

(iv) $\frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \to \tilde{f}$ *em* L^1.

(v) Se A é um subconjunto mensurável de X tal que $T^{-1}A = A$, então

$$\int_A f d\mu = \int_A \tilde{f} d\mu.$$

Prova. (i) Defina, para cada $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$, o conjunto

$$E_{\alpha,\beta} = \{x \in X \mid \liminf_{n\to\infty} S_n(x) < \alpha < \beta < \limsup_{n\to\infty} S_n(x)\}.$$

Então $E_{\alpha,\beta}$ é um subconjunto T-invariante de $\{f^* > \beta\}$, pois

$$\liminf_{n\to\infty} S_n(Tx) = \liminf_{n\to\infty} S_n(x) \quad \text{e} \quad \limsup_{n\to\infty} S_n(Tx) = \limsup_{n\to\infty} S_n(x).$$

Como

$$\{x \in X \mid S_n(x) \text{ não converge}\} \subseteq \bigcup_{\alpha,\beta\in\mathbb{Q}} E_{\alpha,\beta},$$

é suficiente mostrar que $\mu(E_{\alpha,\beta}) = 0$. Isso decorrerá de duas aplicações do Teorema Ergódico Maximal. A primeira é clara, uma vez que $E_{\alpha,\beta}$ é um conjunto invariante tal que $f^* > \beta$, e daí

$$\int_{E_{\alpha,\beta}} f d\mu \;\geq\; \beta \cdot \mu(E_{\alpha,\beta}). \tag{7}$$

Para a outra, note que $E_{\alpha,\beta} \subseteq \{(-f)^* > -\alpha\}$, pois se $x \in E_{\alpha,\beta}$, então existe $n \geq 1$ para o qual

$$\frac{1}{n}(f(x) + \cdots + f(T^{n-1}x)) < \alpha,$$

ou seja

$$(-f)^*(x) \geq \frac{1}{n}((-f)(x) + \cdots + (-f)(T^{n-1}x)) > -\alpha.$$

Assim,

$$\begin{aligned} \int_{E_{\alpha,\beta}} (-f) d\mu &\geq -\alpha \cdot \mu(E_{\alpha,\beta}) \\ \Longrightarrow \qquad \alpha \cdot \mu(E_{\alpha,\beta}) &\geq \int_{E_{\alpha,\beta}} f d\mu, \end{aligned}$$

o que, juntamente com a desigualdade 7, nos dá que

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

Como $\beta > \alpha$, concluímos que $\mu(E_{\alpha,\beta}) = 0$.

(ii) Claro.

(iii) Pelo Lema de Fatou, temos

$$\begin{aligned}
\int |\tilde{f}| d\mu &= \int \left| \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \right| d\mu \\
&= \int \lim_{n\to\infty} \left| \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \right| d\mu \\
&\leq \liminf_{n\to\infty} \int \left| \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k \right| d\mu \\
&\leq \liminf_{n\to\infty} \int \frac{1}{n} \sum_{k=0}^{n-1} \left| f \circ T^k \right| d\mu \\
&= \liminf_{n\to\infty} \frac{1}{n} (\|f\|_1 + \|f \circ T\|_1 + \cdots + \left\| f \circ T^{n-1} \right\|_1) \\
&= \liminf_{n\to\infty} \|f\|_1 \\
&= \|f\|_1 \,,
\end{aligned}$$

onde na penúltima igualdade aplicamos a Proposição 1.2.

(iv) Se f é limitada, então

$$|S_n(x)| \leq \frac{1}{n} \sum_{k=0}^{n-1} |f(T^k x)| \leq \|f\|_\infty$$

e daí, do Teorema da Convergência Dominada, obtemos que $S_n \to \tilde{f}$ em L^1.

Se f não é limitada, tome $g \in L^\infty$ tal que $\|g - f\|_1 < \varepsilon/3$. Temos

$$\begin{aligned}
\left\| S_n(f) - \tilde{f} \right\|_1 &\leq \|S_n(f-g)\|_1 + \|S_n(g) - \tilde{g}\|_1 + \left\| \widetilde{g-f} \right\|_1 \\
&\leq \|f-g\|_1 + \|S_n(g) - \tilde{g}\|_1 + \|g-f\|_1 \,.
\end{aligned}$$

Assim, se $n_0 \geq 1$ é tal que

$$\|S_n(g) - \tilde{g}\|_1 < \frac{\varepsilon}{3}\,, \quad \forall\, n \geq n_0,$$

obtemos

$$\begin{aligned}
\left\| S_n(f) - \tilde{f} \right\|_1 &\leq \|f-g\|_1 + \|S_n(g) - \tilde{g}\|_1 + \|g-f\|_1 \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \\
&= \varepsilon,
\end{aligned}$$

isto é,

$$\left\|S_n(f) - \tilde{f}\right\|_1 < \varepsilon.$$

(v)

$$\begin{aligned}
\left|\int_A f d\mu - \int_A \tilde{f} d\mu\right| &= \left|\int_A S_n d\mu - \int_A \tilde{f} d\mu\right| \\
&\leq \int_A |S_n - \tilde{f}| d\mu \\
&= \left\|S_n - \tilde{f}\right\|_{L^1(A)},
\end{aligned}$$

e essa última norma converge para zero, pelo item anterior. ■

Quando T é invertível, podemos falar nas médias temporais para frente e para trás:

$$\begin{aligned}
\tilde{f}^+(x) &= \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n} (f \circ T^k)(x) \\
\tilde{f}^-(x) &= \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n} (f \circ T^{-k})(x).
\end{aligned}$$

É claro que ambas são invariantes por f. Além disso, elas são iguais, conforme o

Corolário 2.3 *Se, nas condições do Teorema 2.4, T é invertível, então $\tilde{f}^+(x)$ e $\tilde{f}^-(x)$ existem e coincidem em μ-quase todo ponto $x \in X$.*

Demonstração. A existência de $\tilde{f}^+, \tilde{f}^- \in L^1$ decorre do Teorema Ergódico de Birkhoff aplicado a T e T^{-1}. Considere, para cada $\varepsilon > 0$, o conjunto T-invariante

$$A_\varepsilon^+ = \{x \in X \mid \tilde{f}^+(x) \geq \tilde{f}^-(x) + \varepsilon\}.$$

Então

$$\int_{A_\varepsilon^+} \tilde{f}^+ d\mu \geq \int_{A_\varepsilon^+} \tilde{f}^- d\mu + \varepsilon \cdot \mu(A_\varepsilon^+). \tag{8}$$

Por outro lado, do Teorema 2.4, item (v), temos

$$\int_{A_\varepsilon^+} \tilde{f}^+ d\mu = \int_{A_\varepsilon^+} f d\mu = \int_{A_\varepsilon^+} \tilde{f}^- d\mu.$$

Substituindo a igualdade acima em (8), concluímos que

$$\begin{aligned} \int_{A_\varepsilon^+} \tilde{f}^+ d\mu &\geq \int_{A_\varepsilon^+} \tilde{f}^- d\mu + \varepsilon \cdot \mu(A_\varepsilon^+) \\ \Longrightarrow \qquad \mu(A_\varepsilon^+) &= 0. \end{aligned}$$

Analogamente, o conjunto

$$A_\varepsilon^- = \{x \in X \mid \tilde{f}^-(x) \geq \tilde{f}^+(x) + \varepsilon\}$$

tem medida nula e portanto o mesmo vale para o conjunto

$$\{x \in X \mid \tilde{f}^+(x) \neq \tilde{f}^-(x)\} = \bigcup_{n \geq 1} \left(A_{1/n}^+ \cup A_{1/n}^- \right).$$

■

2.3 Ergodicidade

As idéias qualitativas até aqui investigadas nos levam a uma abordagem quantitativa das transformações: pelo Teorema de Recorrência de Poincaré, sob a ação de uma transformação que preserva uma medida finita, quase todo ponto de um conjunto de medida positiva retorna a ele infinitas vezes. Queremos investigar as frequências de retorno.

Mais especificamente, dados $A \subseteq X$ mensurável e $x \in X$, sejam

$$\tau_n(A, x) \doteq \frac{|\{0 \leq i \leq n-1 \mid T^i x \in A\}|}{n}$$

o tempo médio de permanência de x em A nos n primeiros iterados e

$$\tau(A, x) \doteq \lim_{n \to \infty} \tau_n(A, x)$$

o *tempo médio de permanência de x em A*. A hipótese ergódica de Boltzmann afirmava que

$$\tau(A, x) = \mu(A)$$

para μ-quase todo estado inicial $x \in X$, isto é, que, na média, as trajetórias de "todos" os pontos eram bem distribuídas.

Em termos do Teorema Ergódico de Birkhoff, a hipótese ergódica de Boltzmann pode ser formulada pelas igualdades

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

para μ-quase todo ponto $x \in X$.

Sejam $T : X \to X$, onde X é um espaço de medida, e $\mu \in \mathcal{M}_T(X)$.

Definição 2.2 *Um conjunto mensurável $A \subseteq X$ é invariante com respeito a T se*

$$T^{-1}A = A.$$

Quando não houver confusão, diremos simplesmente que A é invariante.

Existem casos em que um conjunto não é invariante mas $T^{-1}A$ e A se relacionam por uma inclusão, digamos $T^{-1}A \subseteq A$. Obviamente, A não é invariante, mas tudo funciona como se fosse, pois ele difere de um conjunto genuinamente invariante por um conjunto de medida nula, como mostraremos abaixo.

Pela invariância da medida, temos

$$\mu(A \backslash T^{-1}A) = \mu(A) - \mu(T^{-1}A) = 0$$

e o mesmo vale para os seus iterados:

$$\mu(T^{-k}A \backslash T^{-(k+1)}A) = 0, \ \forall\, k \geq 0.$$

Como

$$A \supseteq T^{-1}A \supseteq \cdots \supseteq T^{-n}A \supseteq T^{-n-1}A \supseteq \cdots,$$

o conjunto

$$\widetilde{A} = \bigcap_{k\geq 0} T^{-k}A$$

é T-invariante. Mais ainda: ele difere de A por um conjunto de medida nula, pois

$$\begin{aligned} A\backslash\widetilde{A} &= A\backslash\bigcap_{k\geq 0} T^{-k}A \\ &= \bigcup_{k\geq 0}(A\backslash T^{-k-1}A) \\ &= \bigcup_{k\geq 0}(T^{-k}A\backslash T^{-k-1}A). \end{aligned}$$

Assim, A herda todas as propriedades métricas provenientes da invariância de $\tilde{A}$. O mesmo vale se tivermos $A \subseteq T^{-1}A$. Nesse caso, a cadeia

$$A \subseteq T^{-1}A \subseteq \cdots \subseteq T^{-n}A \subseteq T^{-n-1}A \subseteq \cdots$$

nos induz a definir

$$\widetilde{A} = \bigcup_{k\geq 0} T^{-k}A.$$

Daí,

$$\begin{aligned} \widetilde{A} &= \bigcup_{k\geq 0} T^{-k}A \\ &= A \cup (T^{-1}A\backslash A) \cup (T^{-2}A\backslash T^{-1}A) \cup \cdots, \end{aligned}$$

que difere de A por uma união enumerável de conjuntos de medida nula.

Com essas observações, conjuntos "quase"-invariantes são, em termos de medida, invariantes.

Definição 2.3 *Sejam $T : X \to X$ mensurável e $\mu \in \mathcal{M}_T(X)$. Dizemos que μ é uma medida ergódica se todo conjunto invariante por T tiver medida 0 ou 1.*

Teorema 2.5 *Dadas $T : X \to X$ mensurável e $\mu \in \mathcal{M}_T(X)$, são equivalentes:*

(i) μ é uma medida ergódica.

(ii) Toda função invariante por T é constante.

(iii) A média temporal de qualquer $f \in L^1$ é igual à sua média espacial em μ-q.t.p. $x \in X$:

$$\tilde{f}(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(T^k x) = \int_X f d\mu\,, \quad \mu\text{-q.t.p. } x \in X.$$

(iv) Para qualquer $A \subseteq X$ mensurável, temos

$$\tau(A,x) = \mu(A), \quad \mu\text{-q.t.p. } x \in X.$$

Prova. (i) $\Longrightarrow$ (ii). Seja $r \in \mathbb{R}$. O conjunto mensurável

$$A_r = \{x \in X \mid f(x) > r\}$$

é invariante, devido à invariância de f. Assim, $\mu(A_r) = 0$ ou $\mu(A_r) = 1$, ou seja, f é constante.

(ii) $\Longrightarrow$ (iii). Pelo Teorema Ergódico de Birkhoff (item (ii)), $\tilde{f}$ é invariante e portanto constante. Ainda do referido teorema (item (v)), obtemos que

$$\tilde{f}(x) = \int_X \tilde{f} d\mu = \int_X f d\mu \quad \Longrightarrow \quad \tilde{f}(x) = \int_X f d\mu.$$

(iii) $\Longrightarrow$ (iv). Óbvio.

(iv) $\Longrightarrow$ (i). Se A é invariante, então $\chi_A \circ T = \chi_{T^{-1}A} = \chi_A$, e daí

$$\chi_A = \widetilde{\chi}_A = \mu(A),$$

o que por sua vez mostra que $\mu(A) = 0$ ou 1. ■

Vale observar que, quando T é invertível, então

$$\tilde{f}^+(x) = \int_X f d\mu = \tilde{f}^-(x)\,, \quad \mu\text{-q.t.p. } x \in X$$

para toda medida ergódica μ.

2.4 Exemplos de Medidas Ergódicas

2.4.1 *Rotação Irracional*

Vamos mostrar que, se α é irracional, então a medida de Lebesgue λ é ergódica para a rotação

$$\begin{array}{rcccl} R_\alpha & : & \mathbb{S}^1 & \longrightarrow & \mathbb{S}^1 \\ & & x & \longmapsto & x + \alpha \end{array}$$

Seja $f \in L^2$ uma função invariante por R_α. Se

$$f(x) = \sum_{-\infty}^{+\infty} a_n \cdot e^{2\pi i n x}$$

é a sua série de Fourier, então

$$f(x+\alpha) = \sum_{-\infty}^{+\infty} a_n \cdot e^{2\pi i n(x+\alpha)} = \sum_{-\infty}^{+\infty} a_n e^{2\pi i n\alpha} \cdot e^{2\pi i n x}$$

e daí

$$a_n = a_n e^{2\pi i n\alpha}, \quad \forall\, n \in \mathbb{Z}.$$

Como α é irracional, temos

$$e^{2\pi i n\alpha} \neq 1, \ \forall\, n \neq 0,$$

e portanto $f = a_0$.

2.4.2 *Expansão Decimal*

Seja $f : [0,1) \to [0,1)$ a transformação de expansão decimal, definida na Seção 1.3. Vamos mostrar que λ é ergódica com respeito a f. Para isso, seja $A \subseteq [0,1)$ um boreliano f-invariante tal que $\lambda(A) > 0$.

Teorema 2.6 (Derivação de Lebesgue) *Se $A \subseteq \mathbb{R}^n$ é um boreliano, então para λ-q.t.p. $x \in \mathbb{R}^n$ o limite*

$$\lim_{r \to 0} \frac{\lambda(B_r(x) \cap A)}{\lambda(B_r(x))}$$

existe e vale 0 se $x \notin A$ e 1 se $x \in A$.

Observação. $B_r(x)$ denota a bola euclidiana de centro $x \in \mathbb{R}^n$ e raio $r > 0$.

Daí, existem $x \in A$ e intervalos $I_k \subseteq [0,1)$ contendo x tais que

$$\lim_{k \to +\infty} \frac{\lambda(I_k \cap A)}{\lambda(I_k)} = 1. \tag{9}$$

Por outro lado, f tem derivada constante e maior que 1, o que garante dois fatos:

(i) Dado um intervalo $I \subseteq [0,1)$, existe $n \geq 0$ tal que

$$f^n(I) = [0,1).$$

(ii) Dados intervalos $I, J \subseteq [0,1)$, temos

$$\frac{\lambda(f^n(I))}{\lambda(f^n(J))} = \frac{\lambda(I)}{\lambda(J)}, \ \forall\, n \geq 0.$$

Façamos algumas observações. Por f ter derivada maior que 1, um intervalo I eventualmente cobre $[0, 1)$. (ii), por outro lado, nos diz que os intervalos crescem numa proporção constante. Esse é um caso particular do que chamamos de *distorção limitada*, situação em que a medida de Lebesgue λ (que possivelmente não é f-invariante) tem um comportamento controlado nos quocientes acima. Aplicaremos essa idéia novamente no próximo exemplo.

Voltemos ao que queremos: por aproximação, (ii) vale para quaisquer borelianos $I, J \subseteq [0, 1)$. Tomando, para cada $k \geq 1$, um inteiro $n_k \geq 0$ tal que $f^{n_k}(I_k) = [0, 1)$, obtemos

$$\begin{aligned} \frac{\lambda(I_k \cap A)}{\lambda(I_k)} &= \frac{\lambda(f^{n_k}(I_k) \cap f^{n_k}(A))}{\lambda(f^{n_k}(I_k))} \\ &= \frac{\lambda([0, 1) \cap A)}{\lambda([0, 1))} \\ &= \lambda(A). \end{aligned}$$

Tomando o limite quando $k \to +\infty$, concluímos de (9) que $\lambda(A) = 1$.

Definição 2.4 *Dizemos que $x \in \mathbb{R}$ é balanceado se todo dígito $0, 1, \ldots, 9$ aparece com a mesma frequência em sua expansão decimal.*

Em termos da notação já introduzida, x é balanceado se e somente se

$$\tau(A_k, x) = \frac{1}{10}\,, \ \forall\, 0 \leq k \leq 9.$$

É fácil dar exemplos de números balanceados. Por outro lado, é difícil dizer se um dado número é balanceado ou não. Por exemplo, até hoje não se sabe se π é balanceado ou não.

Proposição 2.1 *O conjunto dos números $x \in \mathbb{R}$ não-balanceados tem medida de Lebesgue nula.*

Prova. Podemos desconsiderar a parte inteira de x e supor que $x \in [0, 1)$. Como λ é ergódica para f, o conjunto

$$B_k = \left\{ x \in [0, 1) \,\middle|\, \tau(A_k, x) = \frac{1}{10} \right\}$$

é um boreliano de medida de Lebesgue total e portanto o mesmo vale para o conjunto

$$B = B_0 \cap B_1 \cap \cdots \cap B_9$$

dos números balanceados. ■

2.4.3 *Transformação de Gauss*

Vamos mostrar que a medida

$$\mu(A) = \frac{1}{\log 2} \int_A \frac{1}{1+x} dx \ , \quad \forall A \in \mathcal{B},$$

é ergódica para a transformação de Gauss G. Como μ é equivalente à medida de Lebesgue λ, queremos mostrar que todo conjunto $A \subseteq [0,1)$ tal que $G^{-1}(A) = A$ satisfaz

$$\lambda(A) = 0 \text{ ou } 1.$$

Para isso, vamos analisar mais a fundo a dinâmica da expansão em frações contínuas e das convergentes. Utilizaremos a notação

$$[a_1, \ldots, a_n] \doteq \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}} \, .$$

Seja $\mathrm{SL}(2,\mathbb{Z})$ o conjunto das matrizes 2×2 de entradas inteiras e determinante ± 1. Para cada

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

em $\mathrm{SL}(2,\mathbb{Z})$, considere a transformação $\Phi(A) : \mathbb{R} \cup \{\infty\} \to \mathbb{R} \cup \{\infty\}$ definida por

$$\Phi(A)(x) = \frac{ax+b}{cx+d} \ , \ \forall\, x \in \mathbb{R} \cup \{\infty\}.$$

Essas são as chamadas *transformações de Möbius* e nos serão úteis para representar as convergentes.

É fácil verificar que $\Phi(A) \circ \Phi(B) = \Phi(AB)$, ou seja, a aplicação acima definida é um homomorfismo de grupos entre $\mathrm{SL}(2,\mathbb{Z})$, munida da operação de multiplicação de matrizes, e o espaço de endomorfismos de $\mathbb{R} \cup \{\infty\}$, munido da operação de composição de funções. A partir de agora, por simplicidade de notação, A e $\Phi(A)$ serão ambos denotados por A, ficando subentendido quando estamos considerando a matriz ou a transformação.

Por exemplo, se estivermos tratando das transformações, então AB representará a composição $\Phi(A) \circ \Phi(B)$.

Dado $x = [a_1, a_2, \ldots]$, defina, para cada $n \geq 1$, matrizes

$$A_n = \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix} \quad \text{e} \quad M_n = A_1 \cdots A_n = \begin{pmatrix} r_n & p_n \\ s_n & q_n \end{pmatrix},$$

onde $r_n, s_n, p_n, q_n \in \mathbb{Z}$. Então

$$\begin{array}{rcl} A_n(0) & = & \dfrac{1}{a_n} \\ A_{n-1}A_n(0) & = & A_{n-1}\left(\dfrac{1}{a_n}\right) \quad = \quad \dfrac{1}{a_{n-1} + \dfrac{1}{a_n}} \\ & \vdots & \\ M_n(0) & = & [a_1, \ldots, a_n] \end{array}$$

e portanto

$$[a_1, \ldots, a_n] = \frac{p_n}{q_n} \, .$$

O legal é que a representação acima já é a irredutível! Mais ainda, r_n, s_n, p_n, q_n gozam de várias propriedades interessantes, conforme veremos abaixo.

Pela recursividade

$$\begin{pmatrix} r_n & p_n \\ s_n & q_n \end{pmatrix} = \begin{pmatrix} r_{n-1} & p_{n-1} \\ s_{n-1} & q_{n-1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_n \end{pmatrix},$$

temos, para todo $n \geq 2$, as igualdades

$$(*) \left\{ \begin{array}{rcl} r_n & = & p_{n-1} \\ s_n & = & q_{n-1} \\ p_n & = & r_{n-1} + a_n p_{n-1} \\ q_n & = & s_{n-1} + a_n q_{n-1} \, . \end{array} \right.$$

As duas primeiras igualdades nos dão

$$M_n = \begin{pmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{pmatrix}$$

e daí

$$p_{n-1}q_n - p_n q_{n-1} = \det(M_n) = \prod_{i=1}^{n} \det(A_i) = (-1)^n.$$

Em particular, $\mathrm{mdc}(p_n, q_n) = 1$. Além disso, se definirmos

$$p_{-1} = 1 \ , \ p_0 = 0 \ , \ q_{-1} = 0 \ , \ q_0 = 1,$$

e substituirmos as duas primeiras relações de $(*)$ nas duas últimas, obtemos

$$(**) \ \begin{cases} p_n &= p_{n-1}a_n + p_{n-2} \\ q_n &= q_{n-1}a_n + q_{n-2} \end{cases} , \quad \forall\, n \geq 1.$$

Ótimo, já sabemos representar as convergentes em função das transformações de Möbius. Pelo mesmo argumento utilizado acima, mostramos que se

$$A_n^* = \begin{pmatrix} 0 & 1 \\ 1 & a_n + b_n \end{pmatrix},$$

onde $b_n = G^n(x)$, então

$$x = (M_{n-1}A_n^*)(0),$$

ou seja,

$$\begin{aligned} x &= \begin{pmatrix} p_{n-2} & p_{n-1} \\ q_{n-2} & q_{n-1} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & a_n + b_n \end{pmatrix} (0) \\ &= \begin{pmatrix} p_{n-1} & p_{n-2} + p_{n-1}(a_n + b_n) \\ q_{n-1} & q_{n-2} + q_{n-1}(a_n + b_n) \end{pmatrix} (0) \\ &= \frac{p_{n-2} + p_{n-1}(a_n + b_n)}{q_{n-2} + q_{n-1}(a_n + b_n)} \\ &= \frac{p_{n-1}b_n + p_n}{q_{n-1}b_n + q_n}, \end{aligned}$$

onde na última igualdade utilizamos $(**)$.

Proposição 2.2 *Na notação acima, temos*

$$x = \lim_{n \to +\infty} \frac{p_n}{q_n} \cdot$$

Prova. O resultado é claro se $x \in \mathbb{Q}$. Se $x \notin \mathbb{Q}$, existem infinitos índices $n \geq 1$ tais que $a_n > 0$. Daí

$$\lim_{n \to +\infty} p_n = \lim_{n \to +\infty} q_n = +\infty$$

e portanto, como $b_n \in (0,1)$, obtemos

$$\begin{aligned} \left| x - \frac{p_n}{q_n} \right| &= \left| \frac{p_{n-1}b_n + p_n}{q_{n-1}b_n + q_n} - \frac{p_n}{q_n} \right| \\ &= \frac{b_n \left| p_{n-1}q_n - p_n q_{n-1} \right|}{q_n(q_{n-1}b_n + q_n)} \\ &= \frac{b_n}{q_n(q_{n-1}b_n + q_n)} \\ &< \frac{1}{{q_n}^2} \ , \end{aligned}$$

estabelecendo o resultado. ■

Considere agora, para cada $n \geq 1$, a partição de $[0,1)$ definida por

$$\mathcal{P}_n = \{ J_{(a_1,\ldots,a_n)} = ([a_1,\ldots,a_n+1],[a_1,\ldots,a_n]) \mid (a_1,\ldots,a_n) \in \mathbb{Z}^n \}.$$

Cada um desses intervalos representa o conjunto dos números em $[0,1)$ cuja n-ésima convergente é igual a $[a_1,\ldots,a_n]$. A aplicação

$$G_{(a_1,\ldots,a_n)} = G^n|_{J_{(a_1,\ldots,a_n)}} : J_{(a_1,\ldots,a_n)} \to (0,1)$$

é um difeomorfismo e portanto

$$\begin{aligned} G^{-n}((a,b)) \cap J_{(a_1,\ldots,a_n)} &= ({G_{(a_1,\ldots,a_n)}}^{-1}(a), {G_{(a_1,\ldots,a_n)}}^{-1}(b)), \quad \text{se } n \text{ é par} \\ &= ({G_{(a_1,\ldots,a_n)}}^{-1}(b), {G_{(a_1,\ldots,a_n)}}^{-1}(a)), \quad \text{se } n \text{ é ímpar,} \end{aligned} \tag{10}$$

pois G reverte a orientação. Por outro lado,

$$G([a_1,\ldots,a_n+\alpha]) = [a_2,\ldots,a_n+\alpha],$$

de modo que

$${G_{(a_1,\ldots,a_n)}}^{-1}(a) = [a_1,\ldots,a_n+a] = \frac{p_{n-1}a + p_n}{q_{n-1}a + q_n} \cdot \tag{11}$$

Assim, $G^{-n}((a,b)) \cap J_{(a_1,\ldots,a_n)}$ é um intervalo aberto com extremos

$$\frac{p_{n-1}a + p_n}{q_{n-1}a + q_n} \quad \text{e} \quad \frac{p_{n-1}b + p_n}{q_{n-1}b + q_n} \cdot$$

A partir dessa caracterização, temos a

Proposição 2.3 *Nas condições acima, valem as desigualdades*

$$\frac{1}{2}\cdot\lambda((a,b)) \;<\; \frac{\lambda\left(G^{-n}((a,b))\cap J_{(a_1,\ldots,a_n)}\right)}{\lambda(J_{(a_1,\ldots,a_n)})} \;<\; 2\cdot\lambda((a,b)).$$

Prova. De (10) e (11), temos

$$\begin{aligned}
\lambda\left(G^{-n}((a,b))\cap J_{(a_1,\ldots,a_n)}\right) &= \left|\frac{p_{n-1}a+p_n}{q_{n-1}a+q_n}-\frac{p_{n-1}b+p_n}{q_{n-1}b+q_n}\right| \\
&= \frac{|(b-a)(p_nq_{n-1}-p_{n-1}q_n)|}{(q_{n-1}a+q_n)(q_{n-1}b+q_n)} \\
&= \frac{b-a}{(q_{n-1}a+q_n)(q_{n-1}b+q_n)}\,.
\end{aligned}$$

$J_{(a_1,\ldots,a_n)}$ é um intervalo aberto de extremos $M_n(0)$ e $M_n(1)$, cujo comprimento é igual a

$$\begin{aligned}
\lambda(J_{(a_1,\ldots,a_n)}) &= \left|\frac{p_{n-1}+p_n}{q_{n-1}+q_n}-\frac{p_n}{q_n}\right| \\
&= \frac{|p_{n-1}q_n+p_nq_n-p_nq_{n-1}-p_nq_n|}{q_n(q_{n-1}+q_n)} \\
&= \frac{1}{q_n(q_{n-1}+q_n)}\,.
\end{aligned}$$

Assim,

$$\begin{aligned}
\frac{\lambda\left(G^{-n}((a,b))\cap J_{(a_1,\ldots,a_n)}\right)}{\lambda((a,b))\cdot\lambda(J_{(a_1,\ldots,a_n)})} &= \frac{\dfrac{b-a}{(q_{n-1}a+q_n)(q_{n-1}b+q_n)}}{(b-a)\cdot\dfrac{1}{q_n(q_{n-1}+q_n)}} \\
&= \frac{q_n(q_{n-1}+q_n)}{(q_{n-1}a+q_n)(q_{n-1}b+q_n)}
\end{aligned}$$

e essa última fração satisfaz

$$\frac{1}{2}<\frac{q_n}{q_{n-1}+q_n}<\frac{q_n(q_{n-1}+q_n)}{(q_{n-1}a+q_n)(q_{n-1}b+q_n)}<\frac{q_{n-1}+q_n}{q_n}<2,$$

pois $0<a,b<1$. ∎

Corolário 2.4 *Para todo boreliano $A \subseteq [0,1)$, temos*

$$\frac{\lambda(A)}{2} < \frac{\lambda\left(G^{-n}(A) \cap J_{(a_1,\ldots,a_n)}\right)}{\lambda(J_{(a_1,\ldots,a_n)})} < 2\lambda(A). \tag{12}$$

Demonstração. Da Proposição anterior e da aditividade de λ, as desigualdades valem quando A é a união disjunta de intervalos. No caso geral, aproximamos A por uniões disjuntas A_n de intervalos, aplicamos o resultado para cada A_n e tomamos o limite. ■

Proposição 2.4 *A medida de Gauss μ é ergódica para a transformação de Gauss G.*

Prova. Como afirmado no início da subseção, queremos mostrar que se $G^{-1}(A) = A$, então

$$\lambda(A) = 0 \text{ ou } 1.$$

Por absurdo, suponha que $0 < \lambda(A) < 1$. Pelo Teorema de Derivação de Lebesgue, existe um ponto

$$x = [a_1, a_2, \ldots] \in A$$

de densidade positiva, isto é,

$$\lim_{\varepsilon \to 0} \frac{\lambda(A \cap (x - \varepsilon, x + \varepsilon))}{2\varepsilon} = 1.$$

Em particular,

$$\begin{aligned}
& \lim_{n \to +\infty} \frac{\lambda(A \cap J_{(a_1,\ldots,a_n)})}{\lambda(J_{(a_1,\ldots,a_n)})} &=& \ 1 \\
\Longrightarrow \quad & \lim_{n \to +\infty} \frac{\lambda(A^c \cap J_{(a_1,\ldots,a_n)})}{\lambda(J_{(a_1,\ldots,a_n)})} &=& \ 0 \\
\Longrightarrow \quad & \lim_{n \to +\infty} \frac{\lambda(G^{-n}(A^c) \cap J_{(a_1,\ldots,a_n)})}{\lambda(J_{(a_1,\ldots,a_n)})} &=& \ 0.
\end{aligned}$$

Mas, pelo Corolário anterior,

$$\begin{aligned}
\frac{\lambda(G^{-n}(A^c) \cap J_{(a_1,\ldots,a_n)})}{\lambda(J_{(a_1,\ldots,a_n)})} &> \frac{\lambda(A^c)}{2} \\
\lim_{n \to +\infty} \frac{\lambda(G^{-n}(A^c) \cap J_{(a_1,\ldots,a_n)})}{\lambda(J_{(a_1,\ldots,a_n)})} &> \frac{\lambda(A^c)}{2}
\end{aligned}$$

e essa última fração é positiva, absurdo. Isso conclui a prova. ■

Da mesma maneira feita na Proposição 2.1, temos o

Corolário 2.5 *Para quase todo número $x \in \mathbb{R}$, a frequência de todo natural k na representação em frações contínuas de x é igual a*

$$\frac{1}{\log 2} \cdot \log\left[\frac{(k+1)^2}{k(k+2)}\right].$$

Demonstração. Novamente, podemos desconsiderar a parte inteira de x e supor que $x \in [0,1)$. Para cada natural k, temos

$$\begin{aligned}
\mu(J_k) &= \frac{1}{\log 2}\int_{\frac{1}{k+1}}^{\frac{1}{k}} \frac{1}{1+x}dx \\
&= \frac{1}{\log 2}\left[\log\left(1+\frac{1}{k}\right) - \log\left(1+\frac{1}{k+1}\right)\right] \\
&= \frac{1}{\log 2} \cdot \log\left[\frac{(k+1)^2}{k(k+2)}\right].
\end{aligned}$$

Portanto, como μ é ergódica para G, o conjunto

$$B_k = \left\{x \in [0,1) \,\middle|\, \tau(J_k, x) = \mu(J_k)\right\}$$

é um boreliano de medida total com respeito a μ. Como μ é equivalente a λ, cada B_k tem medida total com respeito a Lebesgue e daí o mesmo vale para a interseção

$$B = \bigcap_{k \geq 1} B_k.$$

Como todo elemento de B satisfaz as condições requeridas, o corolário está demonstrado.$\blacksquare$

2.4.4 *A Função Shift*

Dado um inteiro positivo k, considere o conjunto

$$\Sigma_k = \prod_{n \in \mathbb{Z}} \{1, 2, \ldots, k\},$$

formado por todas as sequências bilaterais (x_n) tais que $x_n \in \{1, 2, \ldots, k\}$ para todo n. Vamos introduzir uma métrica em Σ_k da seguinte maneira: dadas duas sequências $x = (x_n)$ e $y = (y_n)$, seja

$$\begin{aligned}
N(x,y) &= \min\{N \geq 1 \,|\, x_N \neq y_N \text{ ou } x_{-N} \neq y_{-N}\} \\
&= \max\{N \geq 0 \,|\, x_n = y_n, \forall\, |n| < N\},
\end{aligned}$$

$N(x,y) \in \mathbb{N} \cup \{\infty\}$. Noutras palavras, $N(x,y)$ representa o primeiro natural N para o qual algum dos extremos dos intervalos $(x_{-N}, \ldots, x_0, \ldots, x_N)$ e $(y_{-N}, \ldots, y_0, \ldots, y_N)$ são distintos. Observe que

$$N(x,y) = \infty \iff x = y.$$

Proposição 2.5 *A função $d : \Sigma_k \times \Sigma_k \to \mathbb{R}$ definida por*

$$d(x,y) = \left(\frac{1}{2}\right)^{N(x,y)}$$

é uma métrica em Σ_k.

Prova. A simetria e a reflexividade são claras. Ademais, a própria definição e a observação acima implicam que $d(x,y) \geq 0$, com igualdade se e somente se $x = y$. Resta mostrar que d satisfaz a desigualdade triangular.

Sejam $x, y, z \in \Sigma_k$. Note que $N(x,y) \geq N(x,z)$ ou $N(x,y) \geq N(z,y)$. De fato, se esse não é o caso, então $N = N(x,y) < N(x,z), N(z,y)$, e daí

$$\begin{array}{rcccccccc} x_N & = & z_N & = & y_N & \Longrightarrow & x_N & = & y_N \\ x_{-N} & = & z_{-N} & = & y_{-N} & \Longrightarrow & x_{-N} & = & y_{-N}, \end{array}$$

contrariando a maximalidade de N. Logo, podemos supor, sem perda de generalidade, que $N \geq N(x,z)$, e daí

$$d(x,y) = \left(\frac{1}{2}\right)^{N} \leq \left(\frac{1}{2}\right)^{N(x,z)} = d(x,z) \leq d(x,z) + d(z,y).$$

■

Podemos, então, falar em sequências convergentes de Σ_k.

Proposição 2.6 *O espaço métrico Σ_k é sequencialmente compacto.*

Prova. Seja $\left(x^{(m)}\right)_{m\geq 1}$ uma sequência em Σ_k, digamos

$$\left(x^{(m)}\right) = \left(x_n^{(m)}\right)_{n\in\mathbb{Z}}.$$

Construiremos uma subsequência $\left(x^{(m_i)}\right)$ convergente.

Como cada coordenada de elementos de Σ_k pertence ao conjunto finito $\{1, 2, \ldots, k\}$, todo conjunto infinito Y de Σ_k tem a seguinte propriedade: dados inteiros $r < s$, existe uma sequência $(i_r, \ldots, i_s)$ de inteiros pertencentes a $\{1, 2, \ldots, k\}$ que são as coordenadas $r, \ldots, s$ de infinitos elementos de Y. Isso garante que podemos construir indutivamente uma subsequência $\left(x^{(m_i)}\right)$ de $\left(x^{(m)}\right)$ satisfazendo

$$\begin{array}{rcl} x^{(m_j)}_{-(i-1)} & = & x_{-(i-1)} \\ & \vdots & \\ x^{(m_j)}_0 & = & x_0 \\ & \vdots & \\ x^{(m_j)}_{i-1} & = & x_{i-1} \end{array}$$

para qualquer $j \geq i$ (ou seja, todas as coordenadas 0 são iguais, todas as coordenadas $-1, 0$ e 1 são iguais a partir de m_2 e assim por diante). Dessa forma, se $x = (x_n) \in \Sigma_k$, a subsequência $\left(x^{(m_j)}\right)$ converge para x. De fato, dado qualquer $\varepsilon > 0$, existe $N > 0$ tal que $2^N \varepsilon > 1$, e daí, qualquer que seja $i \geq N$, valem as desigualdades

$$N\left(x^{(m_i)}, x\right) \geq N \implies d\left(x^{(m_i)}, x\right) \leq \left(\frac{1}{2}\right)^N < \varepsilon.$$

■

Em particular, Σ_k é um espaço métrico compacto.

Definição 2.5 *(Função Shift) A função shift é a aplicação $\sigma : \Sigma_k \to \Sigma_k$ que leva uma sequência bilateral x em outra sequência bilateral $\sigma(x)$ tal que*

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

ou seja, σ desloca as sequências uma posição à esquerda (o ponto e vírgula marca a posição zero):

$$\sigma : (\ldots, x_{-2}, x_{-1}, x_0; x_1, x_2, \ldots) \longmapsto (\ldots, x_{-1}, x_0, x_1; x_2, x_3, \ldots).$$

Proposição 2.7 *σ é um homeomorfismo.*

Prova. É claro que σ é uma bijeção, cuja inversa é o deslocamento à direita

$$\sigma^{-1} : (\ldots, x_{-2}, x_{-1}, x_0; x_1, x_2, \ldots) \longmapsto (\ldots, x_{-3}, x_{-2}, x_{-1}; x_0, x_1, \ldots).$$

Vamos mostrar que σ e σ^{-1} são Lipschitzianas. Dados $x, y \in \Sigma_k$, temos

$$N(\sigma(x), \sigma(y)) \geq N(x, y) - 1,$$

e portanto

$$d(\sigma(x), \sigma(y)) = \left(\frac{1}{2}\right)^{N(\sigma(x),\sigma(y))} \leq \left(\frac{1}{2}\right)^{N(x,y)-1} = 2d(x, y).$$

Pelas mesmas razões, $d(\sigma^{-1}x, \sigma^{-1}y) \leq 2d(x, y)$. ■

De maneira análoga, podem ser definidos

$$\Sigma_k^+ = \text{ espaço das sequências unilaterais com coordenadas em } \{1, 2, \ldots, k\}$$

e o shift unilateral

$$\sigma^+ : \Sigma_k^+ \longrightarrow \Sigma_k^+.$$

Eles têm as mesmas propriedades de Σ_k e σ_k, exceto pelo fato de σ^+ não possuir inversa, uma vez que a imagem inversa de qualquer sequência unilateral tem k elementos.

Vamos introduzir uma medida em Σ_k. A σ-álgebra considerada é a de Borel $\mathcal{B}$, gerada pelas bolas abertas de d. Tais bolas são conjuntos do tipo

$$[r, s; a_r, \ldots, a_s] \doteq \{x = (x_i) \in \Sigma_k \,|\, x_r = a_r, \ldots, x_s = a_s\},$$

chamados de *cilindros*.

Dado um vetor de probabilidade $p = (p_1, \ldots, p_k)$, defina $\mu \in \mathcal{M}(\Sigma_k)$ pondo

$$\mu([r, s; a_r, \ldots, a_s]) \doteq p_{a_r} \cdots p_{a_s}$$

e estendendo-a a $\mathcal{B}$ pelo Teorema 1.4. Tal medida independe das posições inicial e final dos cilindros e portanto é invariante pelo shift σ, ou seja:

$$\mu \in \mathcal{M}_\sigma(\Sigma_k).$$

Mais ainda: μ é uma medida de memória finita , isto é, se C_1 e C_2 são cilindros, digamos

$$\begin{aligned} C_1 &= [r, s; a_r, \ldots, a_s] \\ C_2 &= [t, u; b_t, \ldots, b_u]\,, \end{aligned}$$

então

$$\mu(\sigma^{-n}C_1 \cap C_2) = \mu(C_1)\mu(C_2)$$

para n suficientemente grande. De fato, se $r+n>u$, então

$$\sigma^{-n}C_1 \cap C_2 = \bigcup_{\substack{1\le c_i\le k\\ u+1\le i\le r+n-1}} [t, s+n; b_t, \ldots, b_u, c_{u+1}, \ldots, c_{r+n-1}, a_r, \ldots, a_s],$$

cuja medida é igual a

$$\begin{aligned}
\mu(\sigma^{-n}C_1 \cap C_2) &= \sum_{\substack{1\le c_i\le m\\ u+1\le i\le r+n-1}} p_{b_t}\cdots p_{b_u} p_{c_{u+1}}\cdots p_{c_{r+n-1}} p_{a_r}\cdots p_{a_s} \\
&= p_{a_r}\cdots p_{a_s} p_{b_t}\cdots p_{b_u} \sum_{\substack{1\le c_i\le m\\ u+1\le i\le r+n-1}} p_{c_{u+1}}\cdots p_{c_{r+n-1}} \\
&= p_{a_r}\cdots p_{a_s} p_{b_t}\cdots p_{b_u} (p_1+\cdots+p_k)^{r+n-u-1} \\
&= \mu(C_1)\mu(C_2).
\end{aligned}$$

Vamos mostrar que μ é ergódica para σ. Seja A um conjunto σ-invariante. Como os cilindros geram $\mathcal{B}$, dado $\varepsilon>0$, existe uma união finita de cilindros disjuntos C tal que

$$\mu(A\Delta C) < \varepsilon.$$

Então

$$\begin{aligned}
\mu(\sigma^{-n}C\Delta C) &\le \mu(\sigma^{-n}A\Delta\sigma^{-n}C) + \mu(\sigma^{-n}A\Delta A) + \mu(A\Delta C) \\
&= \mu(A\Delta C) + \mu(A\Delta A) + \mu(A\Delta C) \\
&= 2\mu(A\Delta C) \\
&< 2\varepsilon.
\end{aligned}$$

Por outro lado, para n grande temos

$$\begin{aligned}
\mu(\sigma^{-n}C\Delta C) &= \mu(\sigma^{-n}C \cap C^c) + \mu(\sigma^{-n}C^c \cap C) \\
&= 2\mu(C)\mu(C^c) \\
&= 2\mu(C)(1-\mu(C))
\end{aligned}$$

e portanto

$$\mu(C)(1-\mu(C)) < \varepsilon.$$

Como $\varepsilon>0$ é arbitrário, concluímos que

$$\mu(A)(1-\mu(A)) = 0 \quad \Longrightarrow \quad \mu(A) = 0 \text{ ou } 1.$$

2.5 Unicidade Ergódica

Definição 2.6 *Dados um espaço topológico X e uma transformação $T : X \to X$ mensurável com respeito à σ-álgebra de Borel, dizemos que T é unicamente ergódica se $\mathcal{M}_T(X)$ for um conjunto unitário.*

Lema 2.1 *Se uma transformação $T : X \to X$ é unicamente ergódica, então sua única medida invariante é ergódica.*

Demonstração. Sejam $\mathcal{M}_T(X) = \{\mu\}$ e $A \subseteq X$ invariante por T. Se $\mu(A) > 0$, definimos a medida de probabilidade restrita ao conjunto A por

$$\mu_A(B) \doteq \frac{\mu(A \cap B)}{\mu(A)}.$$

Tal medida é invariante por T, pois

$$\mu_A(T^{-1}B) = \frac{\mu(A \cap T^{-1}B)}{\mu(A)} = \frac{\mu(T^{-1}(A \cap B))}{\mu(A)} = \frac{\mu(A \cap B)}{\mu(A)} = \mu_A(B).$$

Pela unicidade ergódica, temos a igualdade $\mu_A = \mu$. Em particular,

$$\mu(A) = \mu_A(A) = 1.$$

■

Proposição 2.8 *Sejam X um espaço topológico e $T : X \to X$ uma transformação mensurável com respeito à σ-álgebra de Borel. São equivalentes:*

(i) T é unicamente ergódica.

(ii) Para toda $f : X \to \mathbb{C}$ contínua e $x \in X$, o limite

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

existe e independe de x.

(iii) Para toda $f : X \to \mathbb{C}$ contínua, a sequência de funções

$$f_n = \frac{1}{n} \sum_{k=1}^{n-1} f \circ T^k$$

converge uniformemente em X para uma função constante.

Prova. Seja $C^0(X) = \{f : X \to \mathbb{C} \,|\, f \text{ é contínua}\}$. É claro que se μ é a única medida invariante por T, então a função constante para a qual f_n converge uniformemente é $\tilde{f}(x) = \int f d\mu$.

(i) $\Longrightarrow$ (iii). Por absurdo, suponha que exista $f \in C^0(X)$ tal que f_n não converge uniformemente em X para $\tilde{f}$. Então existem $\varepsilon > 0$, $n_i \to \infty$ e $x_i \in X$ tais que

$$\left| f_{n_i}(x_i) - \int f d\mu \right| \geq \varepsilon. \tag{13}$$

Para cada i, seja $\nu_i \in \mathcal{M}(X)$ a medida associada (pelo Teorema de Riesz-Markov) ao funcional linear positivo $\Phi_i : C^0(X) \to \mathbb{C}$ definido por

$$\Phi_i(g) \doteq \frac{1}{n_i} \sum_{k=0}^{n_i - 1} g(T^k x_i).$$

Pela compacidade de $\mathcal{M}(X)$ (veja (5), seção 1.8), podemos supor que $\nu_i \to \nu$. Afirmamos que $\nu \in \mathcal{M}_T(X)$. De fato, dada $g \in C^0(X)$, temos

$$\begin{aligned}
\int (g \circ T) d\nu &= \lim_{i\to\infty} \int (g \circ T) d\nu_i \\
&= \lim_{i\to\infty} \Phi_i(g \circ T) \\
&= \lim_{i\to\infty} \frac{1}{n_i} \sum_{k=0}^{n_i - 1} g(T^{k+1} x_i) \\
&= \lim_{i\to\infty} \left[\frac{1}{n_i} \sum_{k=0}^{n_i - 1} g(T^k x_i) + \frac{1}{n_i} (g(T^{n_i} x_i) - g(x_i)) \right] \\
&= \lim_{i\to\infty} \frac{1}{n_i} \sum_{k=0}^{n_i - 1} g(T^k x_i) \\
&= \lim_{i\to\infty} \Phi_i(g) \\
&= \lim_{i\to\infty} \int g d\nu_i \\
&= \int g d\nu.
\end{aligned}$$

Por outro lado, de 13, temos

$$\left| \int f d\nu - \int f d\mu \right| = \lim_{i\to\infty} \left| \int f d\nu_i - \int f d\mu \right| = \lim_{i\to\infty} \left| f_{n_i}(x_i) - \int f d\mu \right| \geq \varepsilon$$

e portanto $\nu \neq \mu$, contrariando a unicidade ergódica de T.

(iii) $\Longrightarrow$ (ii). Óbvio.

(ii) $\Longrightarrow$ (i). Sejam $\mu, \nu \in \mathcal{M}_T(X)$ e $f \in C^0(X)$. Utilizando a hipótese e o Teorema Ergódico de Birkhoff,

$$\lim_{n\to\infty} f_n(x) = \int f d\mu, \ \forall\, x \in X.$$

Como o mesmo vale para ν, temos

$$\int f d\mu = \int f d\nu$$

e portanto $\mu = \nu$. ■

2.6 O Teorema Ergódico de Birkhoff para Campos

Definição 2.7 *Um fluxo em um espaço de probabilidade $(X, \mathcal{B}, \mu)$ é uma transformação mensurável $\varphi : \mathbb{R} \times X \to X$ tal que, se $\varphi_t : X \to X$ é definido por*

$$\varphi_t(x) = \varphi(t, x),$$

então
(i) $\varphi_0 = \mathrm{Id}$, onde $\mathrm{Id} : X \to X$ é a identidade, e
(ii) $\varphi_{t+s} = \varphi_t \circ \varphi_s$, para quaisquer $t, s \in \mathbb{R}$.
Dizemos que μ é invariante por φ quando

$$\mu(\varphi_t A) = \mu(A), \qquad \forall\, t \in \mathbb{R}, \ \ \forall\, A \in \mathcal{B}$$

e denotamos por $\mathcal{M}_\varphi(X)$ o conjunto de todas as probabilidades definidas em $\mathcal{B}$ e invariantes por φ.

Cada φ_t é uma transformação mensurável invertível, com inversa igual a φ_{-t}.

Dada $f \in L^1$, as médias temporais de f pelo fluxo φ são definidas por

$$\begin{aligned} \tilde{f}^+(x) &= \lim_{t\to+\infty} \frac{1}{t} \int_0^t (f \circ \varphi_s)(x) ds \\ \tilde{f}^-(x) &= \lim_{t\to+\infty} \frac{1}{t} \int_0^t (f \circ \varphi_{-s})(x) ds\,, \end{aligned}$$

quando tais limites existem. Assim como no caso discreto, eles existem para quase todo ponto $x \in X$ com respeito a μ, conforme veremos abaixo. Antes, precisamos do

Lema 2.2 *Se $(a_n)_{n\geq 1}$ é uma sequência tal que existe*

$$\lim_{n\to+\infty} \frac{a_n}{n},$$

então

$$\lim_{n\to+\infty} \frac{a_{n+1}-a_n}{n} = 0.$$

Demonstração.

$$\begin{aligned}
\lim_{n\to+\infty} \frac{a_{n+1}-a_n}{n} &= \lim_{n\to+\infty} \left(\frac{a_{n+1}}{n+1}\cdot\frac{n+1}{n}\right) - \lim_{n\to+\infty}\frac{a_n}{n} \\
&= \lim_{n\to+\infty} \frac{a_{n+1}}{n+1} - \lim_{n\to+\infty}\frac{a_n}{n} \\
&= 0.
\end{aligned}$$

■

Teorema 2.7 (Ergódico de Birkhoff para Fluxos) *Sejam $(X, \mathcal{B}, \mu)$ um espaço de probabilidade e $\varphi : \mathbb{R}\times X \to X$ um fluxo invariante por μ. Dada $f \in L^1$, os limites*

$$\tilde{f}^+(x) = \lim_{t\to+\infty} \frac{1}{t}\int_0^t (f\circ\varphi_s)(x)ds \tag{14}$$

$$\tilde{f}^-(x) = \lim_{t\to+\infty} \frac{1}{t}\int_0^t (f\circ\varphi_{-s})(x)ds \tag{15}$$

existem e são iguais para quase todo ponto $x \in X$ com respeito a μ.

Prova.

Caso 1. $f : X \to [0, +\infty]$.

Defina a função mensurável $g : \mathbb{R}\times X \to [0, +\infty]$ por

$$g(s, x) = (f\circ\varphi_s)(x).$$

Sejam também $g_s : X \to [0, +\infty]$ e $g^x : [0, 1] \to [0, +\infty]$ dadas por

$$g_s(x) = g(s, x) = g^x(s).$$

Como cada φ_s é uma bijeção μ-invariante, temos

$$\int_{[0,1]} \left(\int_X g_s d\mu \right) ds = \int_{[0,1]} \left(\int_X f d\mu \right) ds = \int_X f d\mu < +\infty$$

e portanto, do Teorema de Fubini, a função $F : X \to [0, +\infty]$ definida por

$$F(x) = \int_0^1 (f \circ \varphi_s)(x) ds$$

está em L^1.

Vamos aplicar a versão discreta do Teorema de Bikhoff. Sendo $T = \varphi_1$, temos

$$\begin{aligned}
\sum_{i=0}^{n-1} (F \circ T^i)(x) &= \sum_{i=0}^{n-1} \int_0^1 (f \circ \varphi_s)(\varphi_i(x)) ds \\
&= \sum_{i=0}^{n-1} \int_i^{i+1} (f \circ \varphi_s)(x) ds \\
&= \int_0^n (f \circ \varphi_s)(x) ds.
\end{aligned}$$

Logo, o limite

$$\tilde{F}(x) = \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} (F \circ T^i)(x) = \lim_{n \to +\infty} \frac{1}{n} \int_0^n (f \circ \varphi_s)(x) ds$$

existe para um conjunto de medida total $A \subseteq X$. Isso é quase o que queremos. Na verdade, todo $x \in A$ possui limite em (14). De fato, do Lema 2.2, temos

$$\begin{aligned}
\lim_{n \to +\infty} \frac{1}{n} \left(\int_0^{n+1} (f \circ \varphi_s)(x) ds - \int_0^n (f \circ \varphi_s)(x) ds \right) &= 0 \\
\Longrightarrow \qquad \lim_{n \to +\infty} \frac{1}{n} \int_n^{n+1} (f \circ \varphi_s)(x) ds &= 0 \\
\Longrightarrow \qquad \lim_{t \to +\infty} \frac{1}{t} \int_{\lfloor t \rfloor}^{t} (f \circ \varphi_s)(x) ds &= 0
\end{aligned}$$

e daí

$$\begin{aligned}
\lim_{t\to+\infty}\frac{1}{t}\int_0^t (f\circ\varphi_s)(x)ds &= \lim_{t\to+\infty}\frac{1}{t}\int_0^{\lfloor t\rfloor}(f\circ\varphi_s)(x)ds + \\
&\quad \lim_{t\to+\infty}\frac{1}{t}\int_{\lfloor t\rfloor}^t (f\circ\varphi_s)(x)ds \\
&= \lim_{t\to+\infty}\frac{1}{\lfloor t\rfloor}\int_0^{\lfloor t\rfloor}(f\circ\varphi_s)(x)ds \\
&= \tilde{F}(x),
\end{aligned}$$

provando que $\tilde{f}^+(x)$ existe para μ-quase todo ponto $x \in X$. O mesmo vale para $\tilde{f}^-(x)$.

Caso 2. $f : X \to \mathbb{R}$.

Aplique o Caso 1 às partes positiva e negativa de f.

Caso 3. $f : X \to \mathbb{C}$.

Aplique o Caso 2 às partes real e imaginária de f.

Por fim, como T é invertível, as médias de Birkhoff de F por T e T^{-1}

$$\begin{aligned}
\tilde{F}^+(x) &= \lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}(F\circ T^i)(x) \\
\tilde{F}^-(x) &= \lim_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}(F\circ T^{-i})(x)
\end{aligned}$$

coincidem em μ-quase todo ponto (Corolário 2.3) e portanto o mesmo vale para $\tilde{f}^+$ e $\tilde{f}^-$. ■

Definição 2.8 *Um conjunto mensurável $A \subseteq X$ é invariante com respeito a φ se*

$$\varphi_t(A) = A,\ \forall\, t \in \mathbb{R}.$$

Definição 2.9 *Dizemos que μ é uma medida ergódica para φ se todo conjunto invariante por φ tiver medida* 0 *ou* 1.

Definição 2.10 *Dizemos que uma função $f \in L^1$ é invariante com respeito a φ se*

$$f\circ\varphi_t = f,\ \forall\, t \in \mathbb{R}.$$

Assim como no caso discreto, existem caracterizações equivalentes para a ergodicidade.

Proposição 2.9 *Sejam $(X, \mathcal{B}, \mu)$ um espaço de probabilidade e $\varphi : \mathbb{R} \times X \to X$ um fluxo mensurável, invariante por μ. São equivalentes:*

(i) μ é uma medida ergódica para φ.

(ii) Toda função $f \in L^1$ invariante por φ é constante em μ-q.t.p. $x \in X$.

(iii) As médias temporais positiva e negativa de qualquer $f \in L^1$ são iguais à média espacial de f em μ-q.t.p. $x \in X$:

$$\tilde{f}^+(x) = \tilde{f}^-(x) = \int_X f d\mu\,, \quad \mu\text{-q.t.p. } x \in X.$$

Prova. Análoga à da Proposição 2.5. ∎

As equivalências para unicidade ergódica também existem.

Proposição 2.10 *Sejam X um espaço topológico e $\varphi : \mathbb{R} \times X \to X$ uma transformação mensurável com respeito à σ-álgebra de Borel. Então φ é unicamente ergódica se e somente se, para toda $f : X \to \mathbb{C}$ contínua e $x \in X$, o limite*

$$\lim_{t \to +\infty} \frac{1}{t} \int_0^t (f \circ \varphi_s)(x) ds$$

existe e independe de x.

Prova. Novamente, análoga à da Proposição 2.8. ∎

3 TEORIA COMBINATÓRIA DOS NÚMEROS

3.1 Dinâmica Topológica

Como em todas as áreas da matemática, um problema pode (e deve) ser abordado sob várias perspectivas: topologicamente, contínua e diferenciavelmente, algebricamente, etc. Isso também ocorre na teoria ergódica: informações topológicas são obtidas a partir de resultados ergódicos. Nesses casos, a σ-álgebra considerada é a gerada pelos abertos de um espaço topológico. Quando tais abertos têm medida positiva, as propriedades métricas, como ergodicidade, traduzem-se em propriedades topológicas, como transitividade.

Nessa seção, sempre que falamos em espaço métrico ficarão subentendidas duas propriedades: compacidade e existência de base enumerável. A última condição equivale à existência de um conjunto enumerável denso em X.

Sejam (X, d) um espaço métrico e

$$\text{Homeo}(X) = \{T : X \to X \,|\, T \text{ é homeomorfismo de } X\}.$$

Agindo em X, considere um grupo $G \subseteq \text{Homeo}(X)$.

Os casos de interesse aqui são $G = \mathbb{Z}$ e $G = \mathbb{R}$. No primeiro caso, se T é um gerador de G, dizemos que o par (X, T) é uma *cascata* (termo introduzido por Anosov). No segundo, dizemos que (X, G) é um *fluxo*.

Dado $x \in X$, definimos a *órbita* de x por G da maneira usual:

$$\mathcal{O}(x) = \{gx \,|\, g \in G\}.$$

Para cascatas, temos

$$\mathcal{O}(x) = \{\ldots, T^{-2}x, T^{-1}x, x, Tx, T^2x, \ldots\}.$$

Definição 3.1 *Dado $T \in \text{Homeo}(X)$, dizemos que F é um conjunto minimal se:*
(i) F é um fechado não-vazio invariante por T.
(ii) F não possui subconjunto fechado próprio satisfazendo (i).
Se X é minimal, dizemos que (X, T) é uma cascata minimal.

Proposição 3.1 *Seja $T \in \text{Homeo}(X)$. Um subconjunto $F \subseteq X$ é minimal se e somente se a órbita de todo ponto de F é densa em F.*

Prova. ($\Longrightarrow$) Se $x \in F$, então $\overline{\mathcal{O}(x)}$ é um subconjunto fechado T-invariante de F que, pela minimalidade, deve ser igual a F.

($\Longleftarrow$) Seja F' um subconjunto não-vazio fechado e T-invariante de F. Tomando $x \in F'$, temos

$$\mathcal{O}(x) \subseteq F' \quad \Longrightarrow \quad F = \overline{\mathcal{O}(x)} \subseteq F' \subseteq F$$

e assim $F' = F$. ■

Proposição 3.2 *Toda cascata (X, T) possui um conjunto minimal.*

Prova. Vamos aplicar o Lema de Zorn à família

$$\mathcal{F} = \{F \subseteq X \mid F \text{ é fechado não-vazio e } TF = F\}.$$

Inicialmente, note que, como X é compacto e T é um homeomorfismo, X pertence a $\mathcal{F}$ e portanto $\mathcal{F} \neq \emptyset$. Resta mostrar que $\mathcal{F}$ é indutivo. Para isso, seja

$$F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq F_{n+1} \supseteq \cdots$$

uma cadeia descendente de $\mathcal{F}$. Então

$$F = \bigcap_{n \geq 1} F_n$$

é fechado e satisfaz

$$TF = T\left(\bigcap_{n \geq 1} F_n\right) = \bigcap_{n \geq 1} TF_i = \bigcap_{n \geq 1} F_i = F,$$

ou seja, $F \in \mathcal{F}$. Dessa forma, $\mathcal{F}$ possui um elemento minimal não-vazio. ■

A proposição acima tem uma conseqüência muito interessante, provada por Birkhoff.

Definição 3.2 *Dada uma cascata (X, T), dizemos que $x \in X$ é recorrente se*

$$\liminf_{n \to \infty} d(x, T^n x) = 0.$$

O conjunto dos pontos recorrentes de (X, T) será denotado por $\mathrm{Rec}(X, T)$*, ou simplesmente* $\mathrm{Rec}(T)$*, quando o espaço considerado estiver subentendido.*

Teorema 3.1 (Recorrência de Birkhoff, 1927) *O conjunto dos pontos recorrentes de uma cascata (X, T) é não-vazio.*

Prova. Se $F \subseteq X$ é minimal, então qualquer $x \in F$ é recorrente. ■

A condição de minimalidade pode ser enfraquecida.

Definição 3.3 *Uma cascata (X, T) é transitiva se a órbita de algum ponto $x \in X$ é densa em X.*

Proposição 3.3 *Dada uma cascata (X, T), são equivalentes:*
(i) (X, T) é transitiva.
(ii) Todo aberto não-vazio T-invariante é denso em X.
(iii) Dados $U, V \subseteq X$ abertos não-vazios, existe um inteiro n tal que

$$T^n U \cap V \neq \emptyset\,.$$

(iv) O conjunto dos pontos com órbita densa é residual.

Observação. Um conjunto é *residual* se for igual à interseção de abertos densos.

Prova. (i) $\Longrightarrow$ (ii). Seja $x \in X$ tal que $\overline{\mathcal{O}(x)} = X$. Seja $U \neq \emptyset$ um aberto T-invariante. Dado $y \in U$, tome $\varepsilon > 0$ tal que $B(y, \varepsilon) \subseteq U$. Daí, existe $n \in \mathbb{Z}$ tal que $T^n x \in B(y, \varepsilon)$, de modo que

$$T^n x \in U \Longrightarrow x \in U,$$

e portanto U é denso em X.

(ii) $\Longrightarrow$ (iii). Por absurdo, se $T^n U \cap V = \emptyset$, $\forall\, n \in \mathbb{Z}$, então

$$W = \bigcup_{n \in \mathbb{Z}} T^n U$$

é um aberto T-invariante que não é denso em X, pois $\overline{W} \subseteq X \backslash V$.

(iii) $\Longrightarrow$ (iv). Tome um conjunto $\{x_m \,|\, m \geq 1\}$ denso em X. Dado $x \in X$, sua órbita é densa em X se e somente se

$$
\begin{aligned}
& & x_m \in \overline{\mathcal{O}(x)}, \quad \forall\, m \geq 1 \\
&\iff & B\left(x_m, \tfrac{1}{k}\right) \cap \mathcal{O}(x) \neq \emptyset, \quad \forall\, k, m \geq 1 \\
&\iff & x \in \textstyle\bigcup_{n\in\mathbb{Z}} T^n B\left(x_m, \tfrac{1}{k}\right), \quad \forall\, k, m \geq 1 \\
&\iff & x \in \textstyle\bigcap_{m\geq 1} \bigcap_{k\geq 1} \bigcup_{n\in\mathbb{Z}} T^n B\left(x_m, \tfrac{1}{k}\right),
\end{aligned}
$$

e portanto

$$\{x \in X \,|\, \overline{\mathcal{O}(x)} = X\} = \bigcap_{m\geq 1} \bigcap_{k\geq 1} \bigcup_{n\in\mathbb{Z}} T^n B(x_m, 1/k),$$

formado pela interseção dos abertos densos $\cup_{n\in\mathbb{Z}} B(x_m, 1/k)$.

(iv) $\Longrightarrow$ (i). Óbvio. ■

Definição 3.4 *Sejam X um espaço métrico e $\mu \in \mathcal{M}(X)$ uma probabilidade nos borelianos de X. O suporte de μ é igual a*

$$\operatorname{supp}(\mu) \doteq \{x \in X \,|\, \mu(B(x, \varepsilon)) > 0, \ \forall\, \varepsilon > 0\}.$$

Da própria definição, temos a seguinte equivalência:

$$\mu \text{ é positiva nos abertos de } X \iff \operatorname{supp}(\mu) = X.$$

Proposição 3.4 *Sejam (X, T) uma cascata e $\mu \in \mathcal{M}_T(X)$ tal que* $\operatorname{supp}(\mu) = X$*. Se μ é ergódica, então (X, T) é transitiva.*

Prova. Dados abertos não-vazios U, V de X, o conjunto

$$W = \bigcup_{n\in\mathbb{Z}} T^n U$$

é invariante por T e tem medida positiva. Pela ergodicidade, ele tem medida total e, como $\mu(V) > 0$, concluímos que

$$\mu(W \cap V) > 0 \implies W \cap V \neq \emptyset,$$

o que estabece a validade do item (iii) da proposição anterior. ■

3.2 Teorema de Van der Waerden

Em 1927, Van der Waerden publicou o seguinte resultado, generalizando uma conjectura de Baudet proposta anos antes.

Teorema 3.2 (Van der Waerden, 1927) *Se $\mathbb{N} = C_1 \cup \cdots \cup C_r$ é uma partição dos naturais, então algum C_j possui progressões aritméticas de tamanho arbitrariamente grande.*

O termo partição se refere à seguinte

Definição 3.5 *Uma partição de um conjunto X é uma decomposição de X em subconjuntos disjuntos dois a dois.*

Como é de se esperar, a prova original de Van der Waerden é puramente combinatória e pode ser encontrada em (7). Vamos prová-lo utilizando ferramentas ergódicas. Ele decorrerá do resultado abaixo, a ser provado na próxima seção.

Teorema 3.3 (Múltiplo de Birkhoff) *Sejam X um espaço métrico compacto e $T_1, \ldots, T_k \in \mathrm{Homeo}(X)$ comutativos dois a dois. Então algum $x \in X$ é múltiplo recorrente, isto é: existe uma seqüência $n_j \to \infty$ tal que*

$$T_i^{n_j} x \to x, \quad i = 1, \ldots, k.$$

Prova do Teorema 3.2. Claramente, basta mostrar que, para cada $n > 0$, existe $j(n) \in \{1, \ldots, r\}$ tal que $C_{j(n)}$ contém uma progressão aritmética de tamanho n. De fato, como a quantidade de índices é finita, algum deles se repetirá infinitas vezes e esse índice nos dará o conjunto procurado.

Cada elemento de $x \in \Sigma_r$ define uma partição de $\mathbb{Z}$ e portanto também de $\mathbb{N}$. Reciprocamente, uma partição de $\mathbb{N}$ define um elemento de Σ_r^+, que pode ser associado a um elemento de Σ_r, simplesmente completando as coordenadas de índice negativo de qualquer maneira. Seja $x \in \Sigma_r$ associado à partição $\mathbb{N} = C_1 \cup \cdots \cup C_r$. Então C_j contém a progressão aritmética $a, a+k, \ldots, a+(n-1)k$ se e somente se

$$x_a = x_{a+k} = \cdots = x_{a+(n-1)k} = j.$$

Como $x_m = (\sigma^m(x))_0$, isso equivale a

$$\begin{aligned} d(\sigma^{a+ik}(x), \sigma^a(x)) &< \frac{1}{2}\,, \quad i = 0, 1, \ldots, n-1 \\ \Longleftrightarrow \qquad d(\sigma^{ik}(y), y) &< \frac{1}{2}\,, \quad i = 0, 1, \ldots, n-1, \end{aligned}$$

onde $y = \sigma^a(x)$. Considere o conjunto σ-invariante

$$A = \overline{\mathcal{O}(x)} \subseteq \Sigma_r$$

e as aplicações $\sigma_1, \ldots, \sigma_{n-1} \in \text{Homeo}(A)$ comutativas duas a duas, restrições de $\sigma, \ldots, \sigma^{n-1}$ ao conjunto A. Pelo Teorema 3.3, existe $z \in A$ múltiplo recorrente para $\sigma_1, \ldots, \sigma_{n-1}$. Em particular, existe $k \geq 1$ tal que

$$\begin{aligned} d({\sigma_i}^k(z), z) &< \frac{1}{2}\,, \quad i = 0, 1, \ldots, n-1, \\ \Longrightarrow \qquad d(\sigma^{ik}(z), z) &< \frac{1}{2}\,, \quad i = 0, 1, \ldots, n-1. \end{aligned}$$

Pela continuidade de σ, as mesmas desigualdades valem para um elemento $y = \sigma^a(x)$ próximo de z. ∎

3.3 Prova do Teorema Múltiplo de Birkhoff

Definição 3.6 *Dizemos que uma cascata (X, T) é homogênea se existe um grupo G de homeomorfismos de X que comutam com T tal que (X, G) é minimal.*

A noção de minimalidade com respeito a um grupo qualquer G é análoga à de cascatas:

$$\overline{\{gx \mid g \in G\}} = X, \ \forall\, x \in X.$$

Definição 3.7 *Dizemos que $A \subseteq X$ é um conjunto homogêneo se existe um grupo G de homeomorfismos de X que comutam com T tal que $GA = A$ e (A, G) é minimal.*

Em particular, A é fechado. Note que A não precisa ser invariante por T. Esse é o ponto positivo da definição. Veremos agora uma proposição que é o coração da prova do Teorema Múltiplo de Birkhoff.

Proposição 3.5 *Sejam (X, T) uma cascata e $A \subseteq X$ homogêneo. Suponha que, para* *cada $\varepsilon > 0$, existam $x, y \in A$ e $n \geq 1$ tais que*

$$d(T^n x, y) < \varepsilon.$$

Então, para cada $\varepsilon > 0$, existem $z \in X$ e $n \geq 1$ tais que

$$d(T^n z, z) < \varepsilon.$$

Prova. Sejam G o grupo associado à homogeneidade de A e $\varepsilon > 0$. A órbita por G de todo ponto $x \in A$ é densa em A e portanto fica tão próxima de qualquer outro ponto $y \in A$ quanto se queira.

Vamos mostrar que é possível obter uma boa aproximação dessas órbitas em A tomando uma quantidade finita de elementos de G. Mais especificamente, mostraremos que existem $g_1, \ldots, g_k \in G$ para os quais

$$\min_{1 \leq i \leq k} d(g_i x, y) < \frac{\varepsilon}{2}, \ \forall\, x, y \in A. \tag{16}$$

Vejamos como fazer isso: pela compacidade de A, é possível cobrí-lo com uma quantidade finita $V_1, \ldots, V_r$ de abertos, cada um deles com diâmetro menor que $\varepsilon/2$. Para cada $j \in \{1, \ldots, r\}$, a família

$$\{g^{-1}V_j \,|\, g \in G\}$$

forma uma cobertura aberta de A, pois se $x \in A$, então existe $g \in G$ tal que $gx \in V_j$, ou seja, $x \in g^{-1}V_j$. Existe, portanto, uma subcobertura finita

$$A = {g_{j1}}^{-1}V_j \cup \cdots \cup {g_{js_j}}^{-1}V_j.$$

Afirmamos que $\{g_{ji} \,|\, 1 \leq j \leq r, 1 \leq i \leq s_j\}$ satisfaz o que queremos. De fato, se $x, y \in A$, então $y \in V_j$, para algum j, e $x \in {g_{ji}}^{-1}V_j$, para algum $i \in \{1, \ldots, s_j\}$. Assim,

$$d(g_{ji}x, y) < \frac{\varepsilon}{2}\,.$$

Renomeando os g_{ji}, obtemos (16).

Afirmamos agora que, dado $y \in A$, existem $x \in A$ e $n \geq 1$ tais que

$$d(T^n x, y) < \varepsilon.$$

Sejam $\delta > 0$ tal que

$$d(z,w) < \delta \implies d(g_i z, g_i w) < \frac{\varepsilon}{2}\,,\ i = 1,\ldots,k$$

e $x_0, y_0 \in A$, $n \geq 1$ tais que

$$d(T^n x_0, y_0) < \delta.$$

Tomando $i \in \{1,\ldots,r\}$ tal que

$$d(g_i y_0, y) < \frac{\varepsilon}{2}$$

e $x = g_i x_0$, temos

$$\begin{aligned} d(T^n x, y) &= d(g_i T^n x_0, y) \\ &\leq d(g_i T^n x_0, g_i y_0) + d(g_i y_0, y) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon, \end{aligned}$$

o que prova a afirmação.

Provemos, por fim, a proposição. Sejam $x_0, x_1 \in A$ e $n_1 \geq 1$ tais que

$$d(T^{n_1} x_1, x_0) < \varepsilon_1 = \frac{\varepsilon}{2}\,.$$

Por continuidade, existe $\varepsilon_2 < \varepsilon_1$ tal que

$$d(x, x_1) < \varepsilon_2 \implies d(T^{n_1} x, x_0) < \varepsilon_1.$$

Tome $x_2 \in A$ e $n_2 \geq 1$ tais que

$$d(T^{n_2} x_2, x_1) < \varepsilon_2\,.$$

Procedendo por indução, se

$$d(T^{n_i} x_i, x_{i-1}) < \varepsilon_i\,,$$

tome $\varepsilon_{i+1} < \varepsilon_i$ tal que

$$d(x, x_i) < \varepsilon_{i+1} \implies d(T^{n_i} x, x_{i-1}) < \varepsilon_i$$

e $x_{i+1} \in A$, $n_{i+1} \geq 1$ tais que

$$d(T^{n_{i+1}} x_{i+1}, x_i) < \varepsilon_{i+1}\,.$$

Assim, se $i < j$, então

$$d(T^{n_j+\cdots+n_{i+1}}x_j, x_i) < \varepsilon_{i+1}\,. \tag{17}$$

Mas i, j podem ser escolhidos de modo que $d(x_i, x_j) < \varepsilon_1$ e portanto, tomando $n = n_j + \cdots + n_{i+1}$, segue de (17) que

$$d(T^n x_j, x_j) \leq d(T^n x_j, x_i) + d(x_i, x_j) < \varepsilon_{i+1} + \varepsilon_1 \leq 2\varepsilon_1 = \varepsilon.$$

■

Proposição 3.6 *Sob as mesmas condições da proposição anterior, existe $x \in A$ recorrente para T.*

Prova. Defina $f : A \to \mathbb{R}$ por

$$f(x) \doteq \inf_{n\in\mathbb{N}} d(T^n x, x).$$

Da proposição anterior, f assume valores arbitrariamente próximos de zero. Além disso, f é semicontínua superiormente (veja a Seção 3.5). De fato, dados $x \in A$ e $\varepsilon > 0$, tome $n \geq 1$ tal que

$$d(T^n x, x) < f(x) + \varepsilon.$$

Pela continuidade de T^n, existe $\delta > 0$ tal que

$$d(x, y) < \delta \implies d(T^n y, y) < f(x) + \varepsilon.$$

Daí

$$f(y) < f(x) + \varepsilon,\ \forall\, y \in B(x, \delta)$$

e como $\varepsilon > 0$ é arbitrário, concluímos que

$$\limsup_{y\to x} f(y) \leq f(x).$$

Da Proposição 3.9, o conjunto C_f dos pontos de continuidade de f é residual (em particular não-vazio). Vamos mostrar que $f|_{C_f} = 0$. Isso concluirá a prova.

Por absurdo, suponha que $f(x_0) > 0$ para algum $x_0 \in C_f$. Tome uma vizinhança V de x_0 e $\delta > 0$ tais que

$$f(x) > \delta\,,\ \forall\, x \in V. \tag{18}$$

Vamos utilizar a homegeneidade de A para mostrar que uma desigualdade similar à acima vale em A.

Seja G o grupo associado à homogeneidade de A. Pela compacidade de A e minimalidade de (A, G), existem $g_1, \ldots, g_k \in G$ tais que

$$A \subseteq \bigcup_{i=1}^{k} {g_i}^{-1} V.$$

Tome $\eta > 0$ tal que

$$d(x, y) < \eta \implies d(g_i x, g_i y) < \delta, \ i = 1, \ldots, k. \tag{19}$$

Vamos mostrar que $f(x) \geq \eta$, $\forall\, x \in A$. Se $f(x) < \eta$, existe $n \geq 1$ tal que

$$d(T^n x, x) < \eta$$

e daí, tomando i tal que $y = g_i x \in V$, segue de (19) que

$$d(T^n y, y) = d(g_i T^n x, g_i x) < \delta,$$

o que contraria (18). ■

Prova do Teorema 3.3. Procedemos por indução. Para $k = 1$, o resultado é garantido pelo Teorema 3.1. Suponha o resultado válido para $k - 1 \geq 1$ e considere $T_1, \ldots, T_k \in \text{Homeo}(X)$ comutativos dois a dois. Seja $G \subseteq \text{Homeo}(X)$ o grupo gerado por $T_1, \ldots, T_k$. Podemos supor que (X, G) é minimal. Se esse não é o caso, restringimos X a um fechado A invariante por G para o qual (A, G) é minimal (que existe por uma aplicação do Lema de Zorn análoga à da Proposição 3.2).

Seja $\Delta = \{(x, \ldots, x) \,|\, x \in X\} \subseteq X^k$ a diagonal e considere a transformação produto $T = T_1 \times \cdots \times T_k$. Queremos mostrar que existe $x^* \in \Delta$ recorrente para T. Para isso, é suficiente verificar as hipóteses da Proposição 3.5:

I. Δ é homogêneo para (X^k, T).

II. Para cada $\varepsilon > 0$, existem $x^*, y^* \in \Delta$ e $n \geq 1$ tais que

$$d(T^n x^*, y^*) < \varepsilon.$$

A ação de G pode ser induzida em X^k da maneira canônica, associando $g \in G$

à transformação $\tilde{g} = g \times \cdots \times g$. Dessa maneira, se $\tilde{G} = \{\tilde{g} \mid g \in G\}$, então (X, G) e $(\Delta, \tilde{G})$ são isomorfos e portanto $(\Delta, \tilde{G})$ é minimal. Em particular, Δ é um conjunto homogêneo para (X^k, T), o que prova I.

Para provar II, sejam $R_i = T_i {T_k}^{-1}$, $i = 1, \ldots, k-1$. Pela hipótese de indução, existem $x \in X$ e $n \geq 1$ tais que

$$d({R_i}^n x, x) < \varepsilon, \ i = 1, \ldots, k-1.$$

Tomando

$$x^* = ({T_k}^{-n} x, \ldots, {T_k}^{-n} x) \quad \text{e} \quad y^* = (x, \ldots, x),$$

obtemos

$$\begin{aligned} d(T^n x^*, y^*) &= d(({T_1}^n {T_k}^{-n} x, \ldots, {T_{k-1}}^n {T_k}^{-n} x, x), (x, \ldots, x, x)) \\ &= d(({R_1}^n x, \ldots, {R_{k-1}}^n x, x), (x, \ldots, x, x)) \\ &< \varepsilon\,, \end{aligned}$$

o que conclui a prova do teorema. ■

3.4 Recorrência Múltipla e Aplicações

Nessa seção, vamos discutir como se desenvolveu a integração entre a Teoria Combinatória dos Números e a Teoria Ergódica, principalmente os resultados que tratam da existência de progressões aritméticas em subconjuntos de inteiros. Vale ressaltar que essa interação culminou com um teorema de Ben Green e Terence Tao provando a existência de progressões aritméticas arbitrariamente grandes formadas somente por primos e que contribuiu para que o último recebesse a Medalha Fields em 2006.

Definição 3.8 *Dizemos que um conjunto C contido em $\mathbb{N}$ ou $\mathbb{Z}$ é aritmético se, para cada $n > 0$, existem inteiros a, r tais que*

$$a, a+r, \ldots, a+(n-1)r \in C.$$

Como vimos na Seção 3.2, Van der Waerden provou em 1927 que toda partição de $\mathbb{N}$ possui algum conjunto aritmético. Mais tarde, Erdös e Turán conjecturaram que todo conjunto "gordo" o suficiente em $\mathbb{N}$ é aritmético. Essa condição de ser "gordo" é expressa pela definição abaixo. No que segue, denotaremos o conjunto $\{1, \cdots, n\}$ por $[1, n]$.

Definição 3.9 *A densidade superior ou simplesmente densidade de um conjunto A de* *inteiros positivos é igual a*

$$d(A) \doteq \limsup_{n\to\infty} \frac{|A \cap [1,n]|}{n}\,.$$

Em 1975, Szemerédi respondeu afirmativamente a conjectura de Erdös-Turán.

Teorema 3.4 (Szemerédi) *Todo conjunto de inteiros positivos com densidade positiva é aritmético.*

Mais uma vez, a prova original do teorema acima é combinatória. Observe que o Teorema de Van der Waerden decorre do acima, pois a função densidade é aditiva, isto é, se A e B são disjuntos, então

$$d(A \cup B) = d(A) + d(B).$$

Assim, dada uma partição $\mathbb{N} = C_1 \cup \cdots \cup C_r$, temos

$$d(C_1) + \cdots + d(C_r) = 1$$

e portanto algum C_j tem densidade positiva.

Em 1977, Harry Furstenberg publicou em (2) uma versão ergódica do Teorema de Szemerédi e abriu novas portas para a abordagem de problemas semelhantes, unindo duas áreas em princípio distantes: a Teoria Combinatória dos Números e a Teoria Ergódica. Ele provou, entre outros resultados, uma generalização do Corolário 1.2.

Teorema 3.5 (Recorrência Múltipla de Poincaré) *Sejam $T : X \to X$ uma transformação mensurável e $\mu \in \mathcal{M}_T(X)$. Se $A \subseteq X$ é mensurável com $\mu(A) > 0$ e $k \geq 1$ é inteiro, então existe um inteiro $n \geq 1$ tal que*

$$\mu(A \cap T^{-n}A \cap \cdots \cap T^{-(k-1)n}A) > 0.$$

Em particular, a interseção acima é não-vazia. Quando (X,T) é uma cascata minimal, essa conclusão decorre do Teorema Múltiplo de Birkhoff e nos dá a versão topológica do teorema anterior.

Teorema 3.6 *Seja (X,T) uma cascata minimal. Se $U \subseteq X$ é aberto e $k \geq 1$ é inteiro, então, para algum inteiro $n \geq 1$, temos*

$$U \cap T^{-n}U \cap \cdots \cap T^{-(k-1)n}U \neq \emptyset.$$

Prova. Tome $T_i = T^i$, $i = 0, \dots, k-1$, e seja $x \in X$ múltiplo recorrente para $T_0, \dots, T_{k-1}$, com $n_j \to \infty$ tal que $T_i^{n_j} x \to x$, isto é:

$$T^{in_j} x \to x, \quad i = 0, \dots, k-1. \tag{20}$$

Pela minimalidade de (X, T), existe $m \in \mathbb{Z}$ tal que $y = T^m x \in U$. É claro que, de (20), temos

$$T^{in_j} y \to y, \quad i = 0, \dots, k-1.$$

Assim, se j é arbitrariamente grande, então

$$T^{in_j} y \in U, \quad i = 0, \dots, k-1,$$

isto é,

$$y \in U \cap T^{-n_j} U \cap \cdots \cap T^{-(k-1)n_j} U.$$

■

O Teorema Múltiplo de Birkhoff, juntamente com o Teorema de Recorrência Múltipla de Poincaré, sugeriram a Furstenberg e Katznelson que uma versão ergódica para transformações comutativas pudesse existir, confirmada pelo

Teorema 3.7 *Sejam $T_1, \dots, T_k : X \to X$ transformações mensuráveis comutativas duas a duas e $\mu \in \mathcal{M}(X)$ invariante por $T_1, \dots, T_k$. Se $A \subseteq X$ é um conjunto mensurável tal que $\mu(A) > 0$, então*

$$\liminf_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu({T_1}^{-j} A \cap \cdots \cap {T_k}^{-j} A) > 0.$$

A prova desse resultado se encontra em (3). Em particular, existe $j \geq 1$ tal que

$$\mu({T_1}^{-j} A \cap \cdots \cap {T_k}^{-j} A) > 0$$

e portanto o Teorema 3.7 generaliza o Teorema 3.5.

Do teorema acima decorre a versão multidimensional do Teorema de Szemerédi. Para tal, precisamos da noção de densidade positiva em dimensões superiores.

Definição 3.10 *Dizemos que um conjunto $A \subseteq \mathbb{Z}^r$ tem densidade positiva se existirem cubos r-dimensionais*

$$[a_n^1, b_n^1) \times \cdots \times [a_n^r, b_n^r)$$

tais que $b_n^i - a_n^i \to \infty$ *para cada* $i = 1, \ldots, r$ *e*

$$\lim_{n\to\infty} \frac{|A \cap [a_n^1, b_n^1) \times \cdots \times [a_n^r, b_n^r)|}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)} > 0.$$

Teorema 3.8 (Multidimensional de Szemerédi) *Se* $A \subseteq \mathbb{Z}^r$ *tem densidade positiva e* $F \subseteq \mathbb{Z}^r$ *é finito, então existem* $b \in \mathbb{Z}^r$ *e* $d \in \mathbb{Z}$ *tais que*

$$b + dF \subset A.$$

Noutras palavras, A contém uma configuração semelhante a F.

Prova. É suficiente mostrar o resultado para $F_k = [1, k]^r$, $k \in \mathbb{N}$.

Considere o espaço $\Omega_r = \{0, 1\}^{\mathbb{Z}^r}$ com a topologia produto. Todo elemento de Ω_r pode ser visto como uma função $\omega : \mathbb{Z}^r \to \{0, 1\}$, igual à função característica do conjunto

$$W = \{(n_1, \ldots, n_r) \in \mathbb{Z}^r \,|\, \omega(n_1, \ldots, n_r) = 1\}.$$

Ω_r possui r tranformações naturais $\sigma_1, \ldots, \sigma_r$, comutativas entre si, que são os shifts nas coordenadas. Mais especificamente, $\sigma_i : \Omega_r \to \Omega_r$ é definido por

$$(\sigma_i\omega)(n_1, \ldots, n_{i-1}, n_i, n_{i+1}, \ldots, n_r) = \omega(n_1, \ldots, n_{i-1}, n_i + 1, n_{i+1}, \ldots, n_r).$$

Como no caso unidimensional, cada σ_i é um homeomorfismo de Ω_r.

Sejam $G \subseteq \text{Homeo}(\Omega_r)$ o grupo gerado por $\sigma_1, \ldots, \sigma_r$ e

$$X = \overline{\{g\chi_A \,|\, g \in G\}}$$

o fecho da órbita de χ_A por G. Seja também

$$S = \{\omega \in X \,|\, \omega(0, \ldots, 0) = 1\}.$$

Então

$$(n_1, \ldots, n_r) \in A \iff {\sigma_1}^{n_1} \cdots {\sigma_r}^{n_r}\chi_A \in S.$$

Pronto. Construímos as estruturas necessárias. Vamos agora usar as hipóteses do teorema. Sejam $[a_n^1, b_n^1) \times \cdots \times [a_n^r, b_n^r)$ cubos tais que $b_n^i - a_n^i \to \infty$, $i = 1, \ldots, r$, e

$$\lim_{n\to\infty} \frac{|A \cap [a_n^1, b_n^1) \times \cdots \times [a_n^r, b_n^r)|}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)} > 0. \tag{21}$$

Esses cubos nos ajudarão a construir uma medida $\mu \in \mathcal{M}(\Omega_r)$, invariante por $\sigma_1, \ldots, \sigma_r$, tal que $\mu(S) > 0$, da seguinte maneira: para cada $n \geq 1$, considere $\mu_n \in \mathcal{M}(\Omega_r)$ associada ao funcional linear unitário $\Phi_n : C^0(\Omega_r) \to \mathbb{C}$ definido por

$$\Phi_n(f) = \frac{\displaystyle\sum_{a_n^j \leq m_j < b_n^j} f({\sigma_1}^{m_1} \cdots {\sigma_r}^{m_r} \chi_A)}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)}.$$

Podemos restringir $(\mu_n)_{n \geq 1}$ e supor que $\mu_n \to \mu$ na topologia fraca*. Como

$$\begin{aligned}
\Phi_n(\chi_S) &= \frac{\displaystyle\sum_{a_n^j \leq m_j < b_n^j} \chi_S({\sigma_1}^{m_1} \cdots {\sigma_r}^{m_r} \chi_A)}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)} \\
&= \frac{\displaystyle\sum_{a_n^j \leq m_j < b_n^j} \chi_A(m_1, \ldots, m_r)}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)} \\
&= \frac{|A \cap [a_n^1, b_n^1) \times \cdots \times [a_n^r, b_n^r)|}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)},
\end{aligned}$$

segue de (21) que $\mu(S) > 0$. Além disso, μ é invariante por $\sigma_1, \ldots, \sigma_r$. De fato,

$$\begin{aligned}
\Phi_n(f \circ \sigma_i) &= \frac{\displaystyle\sum_{a_n^j \leq m_j < b_n^j} f({\sigma_1}^{m_1} \cdots {\sigma_i}^{m_i+1} \cdots {\sigma_r}^{m_r} \chi_A)}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)} \\
&= \frac{\displaystyle\sum_{a_n^j \leq m_j < b_n^j} f({\sigma_1}^{m_1} \cdots {\sigma_r}^{m_r} \chi_A)}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)} \\
&\quad + \frac{\displaystyle\sum_{\substack{a_n^j \leq m_j < b_n^j \\ m_i = b_n^i}} f({\sigma_1}^{m_1} \cdots {\sigma_r}^{m_r} \chi_A)}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)} \\
&\quad - \frac{\displaystyle\sum_{\substack{a_n^j \leq m_j < b_n^j \\ m_i = a_n^i}} f({\sigma_1}^{m_1} \cdots {\sigma_r}^{m_r} \chi_A)}{(b_n^1 - a_n^1) \cdots (b_n^r - a_n^r)}
\end{aligned}$$

e as duas últimas frações da última expressão convergem para zero, pois são majoradas por

$$\frac{\|f\|_\infty \displaystyle\prod_{j \neq i} (b_n^j - a_n^j)}{\displaystyle\prod (b_n^j - a_n^j)} = \frac{\|f\|_\infty}{b_n^i - a_n^i} \longrightarrow 0.$$

Pelo Teorema 3.7 aplicado ao conjunto de transformações

$$\{{\sigma_1}^{k_1}\cdots{\sigma_r}^{k_r}\,|\,0\le k_i\le k,\ i=1,\ldots,r\},$$

existe $d\ge 1$ tal que

$$\mu\left(\bigcap_{0\le k_i\le k}({\sigma_1}^{k_1}\cdots{\sigma_r}^{k_r})^{-d}S\right)>0.$$

Denote a interseção acima por E. Pela definição de μ, a mesma desigualdade vale para μ_n se n é suficientemente grande. Mas então algum transladado ${\sigma_1}^{b_1}\cdots{\sigma_r}^{b_r}\chi_A$ de χ_A está em E, ou seja

$$\begin{aligned}
{\sigma_1}^{b_1}\cdots{\sigma_r}^{b_r}\chi_A &\in \bigcap_{0\le k_i\le k}({\sigma_1}^{k_1}\cdots{\sigma_r}^{k_r})^{-d}S\\
\Longrightarrow\quad {\sigma_1}^{b_1+dk_1}\cdots{\sigma_r}^{b_r+dk_r}\chi_A &\in S\,,\quad 0\le k_i\le k\\
\Longrightarrow\quad (b_1+dk_1,\ldots,b_r+dk_r) &\in A\,,\quad 0\le k_i\le k
\end{aligned}$$

e portanto, tomando $b=(b_1,\ldots,b_r)$, concluímos que $b+dF_k\subset A$. ■

Para finalizar a seção, mencionemos uma conjectura que generaliza o Teorema de Szemerédi.

Conjectura. (Erdös-Turán) Se $A\subseteq\mathbb{N}$ é tal que

$$\sum_{n\in A}\frac{1}{n}=\infty,$$

então A é aritmético.

Ninguém sabe até hoje sua veracidade, nem mesmo se A contém progressões de tamanho três. Uma resposta afirmativa garantiria a existência de progressões aritméticas formadas somente por primos de tamanho arbitrariamente grande.

Ben Green e Terence Tao provaram, seguindo outra direção, que o conjunto dos números primos é, de fato, aritmético. A prova baseia-se em adaptações do Teorema Ergódico de Szemerédi para medidas pseudo-aleatórias e de resultados da análise harmônica. O leitor interessado pode encontrá-la em (7).

3.5 Fatos utilizados sobre funções semicontínuas

Nessa seção, que pode ser considerada um apêndice, estabelecemos alguns fatos básicos sobre funções semicontínuas que foram utilizados na prova da Proposição 3.6. Sejam X um espaço métrico e $f : X \to \mathbb{R}$.

Definição 3.11 *Dizemos que f é semicontínua superiormente em $x \in X$ se*

$$f(x) \geq \limsup_{y \to x} f(y).$$

Dizemos que f é uma função semicontínua superiormente se for semicontínua superiormente em todo $x \in X$.

De maneira análoga, podemos definir semicontinuidade inferior, pedindo que $-f$ seja semicontínua superiormente:

$$\begin{array}{rrcl} & (-f)(x) & \geq & \limsup_{y \to x}(-f)(y) \\ \Longleftrightarrow & f(x) & \leq & \liminf_{y \to x} f(y). \end{array}$$

Lema 3.1 *$f : X \to \mathbb{R}$ é contínua em $x \in X$ se e somente se for semicontínua inferior e superiormente em x.*

Demonstração. ($\Longrightarrow$) Óbvio.

($\Longleftarrow$) Temos

$$f(x) \geq \limsup_{y \to x} f(y) \geq \liminf_{y \to x} f(y) \geq f(x)$$

e portanto

$$\lim_{y \to x} f(y) = f(x).$$

$\blacksquare$

Proposição 3.7 *Sejam $f : X \to \mathbb{R}$ semicontínua superiormente e $c \in \mathbb{R}$. Então*

$$\{x \in X \mid f(x) < c\}$$

é um conjunto aberto de X.

Prova. Se $f(x) < c$, então

$$\limsup_{y \to x} f(y) < c.$$

Assim, $f(y) < c$ em uma bola aberta $B(x, \varepsilon)$ e portanto

$$B(x, \varepsilon) \subseteq \{x \in X \mid f(x) < c\}.$$

$\blacksquare$

Proposição 3.8 *Sejam X um espaço métrico compacto e $f : X \to \mathbb{R}$ semicontínua superiormente. Então f admite um máximo em X.*

Prova. A prova consiste em duas partes:

I. f é limitada: para cada $x \in X$, existe $\varepsilon_x > 0$ tal que

$$f(y) \leq f(x) + 1, \ \forall\, y \in B(x, \varepsilon_x).$$

Pela compacidade de X, existe uma subcobertura finita

$$X = B(x_1, \varepsilon_{x_1}) \cup \cdots \cup B(x_n, \varepsilon_{x_n})$$

e portanto

$$f(x) \leq \max_{1 \leq i \leq n} f(x_i) + 1, \quad \forall\, x \in X.$$

II. f atinge o supremo: sejam $M = \sup_{x \in X} f(x)$ e

$$A_c \doteq \{x \in X \mid f(x) < c\}.$$

Por absurdo, se $f(x) < M$ para todo $x \in X$, então

$$X = \bigcup_{k \in \mathbb{N}} A_{M-1/k}$$

é uma cobertura encaixante ascendente aberta de X. Pela compacidade, existe $k_0 \in \mathbb{N}$ tal que $X = A_{M-1/k_0}$, contrariando a definição de M. $\blacksquare$

Vale observar que o mesmo não vale para o ínfimo, existindo exemplos em que f é ilimitada inferiormente.

Por fim, vejamos o resultado de interesse

Proposição 3.9 *Se $f : X \to \mathbb{R}$ é semicontínua superiormente, então o conjunto D_f dos pontos de descontinuidade de f é de Primeira Categoria, isto é, está contido na união enumerável de fechados com interior vazio.*

4 CONCLUSÃO

A Teoria Ergódica é uma área que estuda as propriedades estatísticas qualitativas de sistemas. Nesse texto, foram expostos alguns resultados básicos da área, dentre eles o Teorema de Recorrência de Poincaré e o Teorema Ergódico de Birkhoff, que possibilitam entender partes dessas propriedades estatísticas qualitativas, a dizer: a existência de retornos arbitrariamente próximos às condições iniciais, e a existência de limites das médias temporais de observáveis (quando o sistema é ergódico, as médias temporais convergem para a média espacial). Essas propriedades foram apresentadas por meio de vários exemplos, muitos deles provenientes da Teoria dos Números.

Uma das facetas da Teoria Ergódica é a sua interação com outros ramos da Matemática. Nesse texto, expusemos sua interação com a Teoria dos Números, com foco nos desenvolvimentos obtidos a partir de 1977, quando Furstenberg identificou uma ligação entre a Teoria dos Números e teoremas ergódicos. A relação entre Teoria Ergódica e Teoria dos Números é bastante prolífica, e desde a descoberta de Furstenberg tem presenciado inúmeros avanços. O mais recente deles é o Teorema de Green e Tao, que prova que o conjunto dos números primos contém progressões aritméticas finitas de tamanho arbitrário. Aqui, o principal resultado que exemplifica essa interação é o Teorema de Van der Waerden.

REFERÊNCIAS

[1] FURSTENBERG, H. *Recurrence in Ergodic Theory and Combinatorial Number Theory.* New Jersey: Princeton University Press, 1981.

[2] FURSTENBERG, H. Ergodic behavior of diagonal measures and a theorem of Szemerédi on Arithmetic Progressions. *Journal d'Analyse Math.* v. 31, p. 204-256, 1977.

[3] FURSTENBERG, H., KATZNELSON, Y. An ergodic Szemerédi theorem for commuting transformations. *Journal d'Analyse Math.* v. 34, p. 275-291, 1978.

[4] PETERSEN, K. *Ergodic Theory.* Cambridge: Cambridge University Press, 1983.

[5] MAÑÉ, R. *Teoria Ergódica.* Rio de Janeiro: IMPA, 1983. Projeto Euclides.

[6] LIMA, E. *Espaços Métricos.* Rio de Janeiro: IMPA, 2003. Projeto Euclides.

[7] ARBIETO, A., MATHEUS, C., MOREIRA, C.G. *Aspectos Ergódicos da Teoria dos Números.* Rio de Janeiro: IMPA, 2007. Publicações Matemáticas.

[8] PARRY, W. *Topics in Ergodic Theory.* Cambridge: Cambridge University Press, 1981.

[9] HALMOS, P. *Lectures on Ergodic Theory.* Tóquio: The Mathematical Society of Japan, 1956.

[10] OLIVEIRA, K. *Um primeiro curso em Teoria Ergódica e Aplicações.* Rio de Janeiro: IMPA, 2005. Publicações Matemáticas.