



UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
Campus de São José do Rio Preto

Nilton Cesar Moretti Junior

Dinâmica, Combinatória e Ergodicidade

São José do Rio Preto

2017

Nilton Cesar Moretti Junior

Dinâmica, Combinatória e Ergodicidade

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Pós-Graduação em Matemática, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto

Orientador: Profº Drº Ali Messaoudi

Coorientador: Danilo Antonio Caprio

Financiadores: CAPES e CNPQ

São José do Rio Preto, 2017

Moretti Júnior, Nilton Cesar.

Dinâmica, combinatória e ergodicidade /Nilton Cesar Moretti Júnior. -- São José do Rio Preto, 2017
52 f. : il.

Orientador: Ali Messaoudi

Coorientador: Danilo Antonio Caprio

Dissertação (mestrado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Instituto de Biociências, Letras e Ciências Exatas

1. Matemática. 2. Teoria dos números. 3. Teoria dos sistemas dinâmicos. 4. Teoria ergódica. 5. Weyl, Problemas de. 6. Combinações (Matemática) I. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. II. Título.

CDU – 511

Ficha catalográfica elaborada pela Biblioteca do IBILCE
UNESP - Câmpus de São José do Rio Preto

Nilton Cesar Moretti Junior

Dinâmica, Combinatória e Ergodicidade

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre em Matemática, junto ao Programa de Pós-Graduação em Matemática, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto

Orientador: Profº Drº Ali Messaoudi
Coorientador: Danilo Antonio Caprio

Banca Examinadora

Prof. Dr. Ali Messaoudi

UNESP/ Câmpus de São José do Rio Preto

Orientador

Prof Dr. Danilo Antonio Caprio

UNESP/Câmpus de São José do Rio Preto

Prof. Dr. Paulo Ricardo da Silva

UNESP/Câmpus de São José do Rio Preto

Prof. Dr. José Régis Azevedo Varão Filho

UNICAMP

Financiadoras: CNPQ e CAPES

São José do Rio Preto

30/agosto/2017

Agradecimentos

Agradeço a Deus por tudo durante este trabalho. Agradeço aos meus pais, Nilton e Marilani, por esse apoio imenso que eles me deram.

Também agradeço ao Prof. Dr. Ali Messaoudi pela grande orientação, não só pelas dicas, conselhos e paciência, mas também pela inspiração que ele me deu, por toda a sua empolgação em problemas de matemática, além de todo o amadurecimento que eu tive com tudo isso.

Obrigado, Prof. Dr. Danilo Antonio Caprio, pela coorientação e também pela grande paciência comigo.

Também quero agradecer ao Prof. Dr. Paulo Ricardo da Silva pelos 3 anos e meio de orientação que me fez amadurecer muito também na matemática e pela disponibilidade da minha banca.

Agradeço também ao Prof. Dr. José Régis Azevedo Varão Filho, pela disponibilidade em fazer parte da banca avaliadora.

Aos professores de matemática do departamento por todo o seu tempo gasto e pelo conhecimento passado.

Aos meus amigos da pós que me ajudaram, me apoiaram e me fizeram companhia durante todo esse tempo Paula Akari Fukushima, Eduardo Machado (estes dois em especial por me ajudarem na elaboração dos slides da apresentação), Fabrício Rissão, Melka Carolina, Diana Rodriguez, Samuel Santos, Antonio Miálíchi, Bruna Flório, Mayara Braz, Luan, Talita Thomazi, Ezequiel Silva, Marcos Adriano, Nicolas Iago, Beatriz Melo.

Também um agradecimento muito grande ao apoio dos meus alunos, que melhoram muito o meu humor durante esse período: Bruna Gomes, Valeria Izidoro, Pedro Henrique Moraes, Larissa Mazi, Tasmin Pereda, Amanda Ribeiro, Fabrício Pizelli, Poliana Basilio, Thalissa Ribeiro, Gabriel Silveira entre outros.

Agradeço também a CNPQ e pela CAPES pelo financiamento dado neste tempo.

Agradeço a todos que me apoiaram durante esse tempo todo e me ajudaram durante o processo. Peço desculpas se eu não lembrei de colocar alguém.

Resumo

Neste trabalho estudamos vários resultados relacionados com sistemas dinâmicos, teoria dos números e combinatória. Em particular, provamos os teoremas de Van Der Waerden, Szemerédi, Koksma e Weyl.

Palavras-chave: Teoria Ergódica, Sistemas Dinâmicos, Van der Waerden, Szemerédi, Koksma, Weyl

Abstract

In this work we study several results connected with dynamical systems, number theory and combinatorics. In particular, we prove Van Der Waerden, Szemerédi, Koksma and Weyl's theorems.

Keywords: Ergodic theory, Dynamical systems, Van der Waerden, Szemerédi, Koksma, Weyl.

Sumário

1	Sistemas Dinâmicos	8
2	Teoria Ergódica	16
3	Teoremas de Van Der Waerden e de Szemerédi	27
3.1	Prova do Teorema de Van der Waerden usando Sistemas Dinâmicos	31
3.2	Teorema de Szemerédi	32
4	Equidistribuição	34
A	Prova do Teorema da Recorrência Múltipla de Furstenberg em casos particulares	49

Introdução

O objetivo deste trabalho é analisar alguns tópicos de Sistemas Dinâmicos sob aspectos combinatórios e da teoria dos números. Um sistema dinâmico é um par (X, T) , onde X é um conjunto não vazio e $T : X \rightarrow X$ é uma aplicação. Se X é um espaço topológico e T uma aplicação contínua, então dizemos que (X, T) é um sistema dinâmico topológico. Se $(X, \mathcal{A}, \mu)$ é um espaço de medida e T é uma aplicação μ -mensurável, então (X, T) é chamado de sistema dinâmico mensurável.

Existe uma relação forte entre a área de sistemas dinâmicos, como Teoria dos Números, Análise Funcional, Teoria da Medida, entre outras. Em particular, existem teoremas clássicos de teoria dos números que podem ser provados usando sistemas dinâmicos. Alguns exemplos são os teoremas de Van Der Waerden [1],[5], Szemerédi [1],[6], Weyl [2] e Koksma.

Dividiremos nosso trabalho em quatro capítulos e um apêndice. No primeiro capítulo, começaremos com algumas noções de sistemas dinâmicos. Em particular, trataremos sobre sistemas dinâmicos minimais [3] e transitividade topológica [4].

O segundo capítulo tratará sobre teoria ergódica. Daremos alguns conceitos básicos sobre a teoria, além de enunciarmos os teoremas de Poincaré e Birkhoff.

No terceiro capítulo trataremos do teorema de Van Der Waerden. Este teorema diz, dados $k, m \geq 1$, existe N tal que qualquer coloração com m cores de $\{1, 2, \dots, N\}$ contém uma progressão aritmética de comprimento k monocromática. Apresentaremos duas provas pra este teorema: : uma combinatória e outra usando sistemas dinâmicos. Para esta última, usaremos o teorema da recorrência múltipla ergódica de Furstenberg.

Ainda neste capítulo, provaremos o teorema de Szemerédi, que diz que se $A \subset \mathbb{N}$ tem densidade positiva, então A possui progressões aritméticas de comprimento arbitrariamente grande.

No quarto capítulo, falaremos sobre equidistribuição. Seja $(x_n)_{n \geq 0}$ uma sequência em $[0, 1[$. Dizemos que $(x_n)_{n \geq 0}$ é equidistribuída se para quaisquer $0 \leq a < b < 1$, a sequência

$$N(a, b, n) = \#\{m \leq n : x_m \in [a, b[\}.$$

é assintoticamente equivalente à $n(b - a)$ quando n tende a infinito. Mais geralmente, se $(y_n)_{n \geq 0}$ é uma sequência de números reais, dizemos que $(y_n)_{n \geq 0}$ é equidistribuída módulo 1 se a sequência $(\{y_n\})_{n \geq 0}$ é equidistribuída, onde $\{y\}$ representa a parte fracionária de y . Provaremos o Critério de Weyl e os teoremas de Koksma e de Weyl.

Finalmente, no apêndice, provaremos o teorema de Furstenberg em casos particulares e daremos uma ideia de como ele pode ser generalizado.

Capítulo 1

Sistemas Dinâmicos

Neste capítulo veremos conceitos sobre sistemas dinâmicos topológicos, como a minimalidade e a transitividade topológica.

Definição 1. *Um sistema dinâmico é um par (X, T) , onde X é um conjunto não vazio e $T : X \rightarrow X$ é uma aplicação. Se X é um espaço topológico e T uma aplicação contínua, então dizemos que (X, T) é um sistema dinâmico topológico. Se $(X, \mathcal{A}, \mu)$ é um espaço de medida e T é uma aplicação μ -mensurável, então (X, T) é chamado de sistema dinâmico mensurável.*

Se $x \in X$, definimos a órbita de x sob T como sendo o conjunto

$$O(x) := \{T^n(x) : n \in \mathbb{N}\}, .$$

onde $T^0(x) = x$ e $T^n(x) = T(T^{n-1}(x))$, para todo inteiro $n \geq 1$.

Definição 2. *Sejam X um conjunto não vazio e $T : X \rightarrow X$ uma aplicação. Dizemos que um subconjunto E de X é T -invariante quando $T(E) \subset E$.*

Definição 3. *Dizemos que um sistema dinâmico topológico (X, T) é minimal quando os únicos conjuntos T -invariantes fechados são $\emptyset$ e X .*

O resultado a seguir nos ajuda a caracterizar sistemas dinâmicos minimais.

Proposição 4. *Um sistema dinâmico (X, T) é minimal se, e somente se, a órbita $O(x)$ é densa em X , para todo $x \in X$.*

Demonstração. Seja $x \in X$. É fácil verificar que $T(O(x)) \subset O(x)$. Desta forma,

$$\overline{T(O(x))} \subset \overline{O(x)}.$$

Porém, como T é contínua, $T(\overline{O(x)}) \subset \overline{T(O(x))}$. Assim, $T(\overline{O(x)}) \subset \overline{O(x)}$, ou seja, $\overline{O(x)}$ é T -invariante. Como (X, T) é minimal e $\overline{O(x)} \neq \emptyset$, segue que $\overline{O(x)} = X$.

Reciprocamente, vamos supor, por absurdo, que (X, T) não é minimal, isto é, existe $E \subset X$ fechado tal que E é T -invariante e diferente de $\emptyset$ e X . Sejam $x \in E$ e $y \in X - E$. Como $T(E) \subset E$, segue que $T^n(E) \subset E$ para todo n natural. Assim, $O(x) \subset E$. Desta forma, $(X - E) \cap O(x) = \emptyset$. Como E é fechado, segue que $X - E$ é aberto. Logo, como $y \in X - E$, segue que $y \notin \overline{O(x)}$. $\square$

Definição 5. Um sistema dinâmico topológico (X, T) é topologicamente transitivo quando existe uma órbita densa.

Proposição 6. (Teorema da Transitividade de Birkhoff) Sejam X um espaço métrico, separável e $T : X \rightarrow X$ uma aplicação. As seguintes afirmações são equivalentes:

- (i) T é topologicamente transitiva.
- (ii) Para quaisquer U e V abertos não vazios contidos em X , existe $n \in \mathbb{N}$ tal que $T^n(U) \cap V \neq \emptyset$.

Demonstração. $((i) \Rightarrow (ii))$

Considere x um ponto cuja órbita é densa e sejam U e V abertos não vazios. Temos $O(x) \cap U \neq \emptyset$, isto é, existe $p \in \mathbb{N}$ tal que $T^p(x) \in U$. Além disso,

$$O(T^{p+1}(x)) = O(x) \setminus \{x, T(x), \dots, T^p(x)\}.$$

Como estamos tirando uma quantidade finita de pontos de um conjunto denso, segue que $O(T^{p+1}(x))$ é denso, ou seja, $\overline{O(T^{p+1}(x))} = X$. Assim, $O(T^{p+1}(x)) \cap V \neq \emptyset$, isto é, existe $q > p$ tal que $T^q(x) \in V$. Se tomarmos $n = q - p$, temos $T^n(T^p(x)) \in T^n(U) \cap V$.

$((ii) \Rightarrow (i))$

Reciprocamente, como X é um espaço separável, ele possui uma base enumerável de abertos, digamos $(V_j)_{j \geq 0}$. A órbita de x é densa se, e somente se, para todo $j \geq 0$ existe

$n \geq 0$ inteiro tal que $T^n(x) \in V_j$. Assim, se D for o conjunto dos pontos que possuem órbita densa em X , segue que

$$D = \bigcap_{j \in \mathbb{N}} \bigcup_{n \geq 0} T^{-n}(V_j).$$

Por outro lado, (ii) implica que o conjunto $\bigcup_{n \geq 0} T^{-n}(V_j)$ é denso de X . Logo, pelo teorema de Baire, D é denso em X . Em particular, D não é vazio. $\square$

Na demonstração da proposição anterior, provamos o seguinte resultado.

Corolário 7. *Se (X, T) é topologicamente transitivo e X é separável, então o conjunto dos pontos cuja órbita é densa e residual (ou seja, é a intersecção enumerável de abertos densos em X).*

Exemplo 1.0.1. *Considere $A = \{0, 1, 2, \dots, d\}$ um alfabeto finito e $X = A^{\mathbb{N}^*}$ o conjunto de todas as sequências cujos termos estão em A . X pode ser identificado com o conjunto das palavras infinitas sobre A . A transformação shift é a aplicação $T : X \rightarrow X$ definida por*

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots),$$

para todo $(x_1, x_2, x_3, \dots) \in X$. Podemos ver o conjunto X como um espaço métrico. De fato, se $x = (x_1, x_2, \dots), y = (y_1, y_2, \dots) \in X$, vamos definir a seguinte função $d : X \times X \rightarrow \mathbb{R}_+$ por

$$d(x, y) = \begin{cases} 2^{-\min \{i; x_i \neq y_i\}}, & \text{se } x \neq y \\ 0, & \text{se } x = y. \end{cases}$$

Por exemplo, se $A = \{a, b, c\}$ o alfabeto, então $d((a, a, a, \dots), (a, a, b, \dots)) = 2^{-3}$

Proposição 8. *A função d definida anteriormente é uma métrica.*

Demonstração. Basta provarmos a desigualdade triangular. Considere $x, y, z \in X$. Vamos dividir em casos:

- Se $x \neq y$ e $y = z$, então

$$d(x, z) + d(z, y) = d(x, z) + 0 = d(x, y) \geq d(x, y).$$

- Se $x \neq y$ e $x = z$, então o procedimento é totalmente análogo.
- Se $x = y$ e $y \neq z$, então

$$d(x, y) = d(x, x) = 0 \leq d(x, z) + d(z, y).$$

- Se $x \neq y$, $y \neq z$ e $x \neq z$, consideraremos $i_0 := \min \{i; x_i \neq y_i\}$, $i_1 := \min \{i; x_i \neq z_i\}$ e $i_2 := \min \{i; y_i \neq z_i\}$. Neste caso, provar a desigualdade triangular é equivalente a mostrar que

$$2^{-i_0} \leq 2^{-i_1} + 2^{-i_2}.$$

Como $x_{i_0} \neq y_{i_0}$, segue que $z_{i_0} \neq x_{i_0}$ ou $z_{i_0} \neq y_{i_0}$. Assim,

$$i_1 \leq i_0 \text{ ou } i_2 \leq i_0 \Rightarrow 2^{-i_0} \leq 2^{-i_1} \text{ ou } 2^{-i_0} \leq 2^{-i_2} \Rightarrow 2^{-i_0} \leq 2^{-i_1} + 2^{-i_2}.$$

□

Proposição 9. *A transformação shift $T : X \rightarrow X$ é contínua com relação a métrica d definida anteriormente.*

Demonstração. Considere $x = x_0x_1\dots$ e $y = y_0y_1\dots$ em X . Se $x = y$, então $d(x, y) = d(T(x), T(y)) = 0$. Vamos supor que $x \neq y$ e $d(x, y) = 2^{-(n+1)}$, com n natural. Então $x_0x_1\dots x_n = y_0y_1\dots y_n$ e $x_{n+1} \neq y_{n+1}$. Observemos que $T(x) = x_1x_2\dots$ e $T(y) = y_1y_2\dots$. Então $d(T(x), T(y)) = 2^{-n}$. Desta maneira, quaisquer $x, y \in X$,

$$d(T(x), T(y)) = 2d(x, y)$$

de onde segue que T é contínua.

□

Definição 10. *Os cilindros são conjuntos da forma*

$$C := [a_1, \dots, a_n] = \{(x_i)_{i \in \mathbb{N}} : x_i = a_i, \text{ para todo } 1 \leq i \leq n\},$$

onde $a_1, \dots, a_n \in A = \{0, 1, \dots, d\}$.

Proposição 11. *Os cilindros são abertos e fechados.*

Demonstração. Considere o cilindro $C = [a_1, \dots, a_n]$, com $a_1, \dots, a_n \in A$. Sejam $x = (x_1, x_2, \dots) \in C$ e $\varepsilon = 2^{-(n+1)}$. Então $B(x, \varepsilon) = \{(x_1, x_2, \dots, x_n, y_{n+1}, y_{n+2}, \dots); y_i \in A, \text{ para todo } i \geq n+1\} \subset C$. Logo, C é aberto. Agora, vamos provar que C é fechado. Observemos que $(x_i)_{i \in \mathbb{N}} \in X - C$ se, e somente se, $x_i \neq a_i$ para pelo menos um $i \in \{1, 2, \dots, n\}$. Consideremos $x \in X - C$, $k_0 := \min \{i \in \{1, \dots, n\}; x_i \neq a_i\}$ e $\varepsilon = 2^{-k_0}$. Seja $y \in B(x, \varepsilon)$. Se $y = x$, então $y \in X - C$. Caso contrário,

$$d(x, y) < \varepsilon \implies 2^{-\min \{i; x_i \neq y_i\}} < 2^{-k_0} \implies \min \{i; x_i \neq y_i\} > k_0.$$

Assim $x_i = y_i$ para $1 \leq i \leq k_0$. Em particular, $x_{k_0} = y_{k_0}$. Como $x_{k_0} \neq a_{k_0}$, segue que $y_{k_0} \neq a_{k_0}$. Logo, $y \in X - C$ e assim $B(x, \varepsilon) \subset X - C$. Portanto C é fechado. $\square$

Proposição 12. *Os cilindros geram τ_d , que é a topologia gerada pela métrica d .*

Demonstração. Sejam G um aberto não vazio de X e $x \in G$. Então existe $r > 0$ tal que $B(x, r) \subset G$. Considere p inteiro tal que $2^{-p} < r$. Desta forma, $B(x, 2^{-p}) \subset B(x, r) \subset G$. Provaremos que o cilindro $C = [x_1, \dots, x_p]$ é tal que $x \in C \subset G$. Claramente $x \in C$. Dado $y \in C$, com $y \neq x$, $\min \{i; x_i \neq y_i\} \geq p+1$. Assim,

$$d(x, y) = 2^{-\min \{i; x_i \neq y_i\}} \leq e^{-(p+1)} \leq e^{-p} < r,$$

ou seja, $y \in B(x, r) \subset G$. Logo, $C \subset G$. $\square$

Proposição 13. *Todo aberto pode ser escrito como união enumerável de cilindros.*

Demonstração. Seja $G \neq \emptyset$ um aberto de X . Pela proposição anterior, para todo $x \in G$, existe um cilindro C_x tal que $x \in C_x \subset G$. Assim

$$G = \bigcup_{x \in G} \{x\} \subset \bigcup_{x \in G} C_x \subset G,$$

de onde segue que

$$G = \bigcup_{x \in G} C_x.$$

Por outro lado, o conjunto dos cilindros é igual a $I = \bigcup_{n \geq 1} I_n$, onde $I_n = \{[a_1, \dots, a_n] : a_i \in A\}$. Assim, I é enumerável. $\square$

Proposição 14. *A topologia induzida pela distância d é igual à topologia produto das topologias discretas sobre A .*

Demonstração. Sejam τ_d e τ a topologia induzida por d e a topologia produto das topologias discretas sobre A , respectivamente.

Considere $C = [u_1, \dots, u_n]$ um cilindro que é um aberto básico de τ_d . Podemos escrever

$$C = \{u_1\} \times \dots \times \{u_n\} \times A \times A \times \dots$$

de onde segue que $C \in \tau$.

Assim, $\tau_d \subset \tau$. Além disso, para n natural, tomemos

$$O_n := \prod_{i \in \mathbb{N}^*} A_i,$$

onde $A_i = A$, se $i > n$ e $A_i \subset A$, se $1 \leq i \leq n$. O_n é um aberto básico de τ . Provemos que $O_n \in \tau_d$. Com efeito, seja $x = (x_1 x_2 \dots) \in O_n$. Então, $x \in C = [x_1, \dots, x_n] \subset O_n$. Logo, $\tau \subset \tau_d$. $\square$

Como o conjunto dos cilindros é uma base da topologia dada pela métrica d , deduzimos o seguinte resultado.

Corolário 15. *A σ -álgebra gerada pelos abertos é igual à σ -álgebra gerada pelos cilindros.*

Observação 1. *Pelo teorema de Tychonoff, deduzimos que $A^{\mathbb{N}}$ é compacto.*

Seja $K_3 \subset [0, 1]$ o conjunto de Cantor clássico, isto é,

$$K_3 = \bigcap_{n=0}^{\infty} A_n,$$

onde $A_0 = [0, 1]$, $A_1 = A_0 \setminus]\frac{1}{3}, \frac{2}{3}[$, $A_2 = A_1 \setminus (]\frac{1}{9}, \frac{2}{9}[\cup]\frac{7}{9}, \frac{8}{9}[)$. É um resultado conhecido que K_3 é conjunto dos elementos x de $[0, 1]$, cujo desenvolvimento na base 3 é

$$x = \sum_{n=1}^{\infty} a_n 3^{-n},$$

onde a_n é igual a 0 ou 2, para todo $n \geq 1$.

Proposição 16. *K_3 é homeomorfo a $X = \{0, 1\}^{\mathbb{N}}$ munido da topologia produto das topologias discretas.*

Demonstração. Consideremos a aplicação $\varphi : \{0, 2\}^{\mathbb{N}} \rightarrow K$ definida por

$$\varphi((a_n)_{n \geq 0}) = \sum_{n=0}^{\infty} a_n 3^{-n-1}.$$

Mostraremos que φ é injetiva. Consideremos $x = x'$ com desenvolvimentos na base 3 distintos, dados por $x = \sum_{i=0}^{\infty} a_i 3^{-i-1}$ e $x' = \sum_{i=0}^{\infty} a'_i 3^{-i-1}$ onde $a_i, a'_i \in \{0, 2\}$. Seja $i_0 := \min \{i; a_i \neq a'_i\} < \infty$. Logo,

$$\begin{aligned} |x - x'| &= \left| \sum_{i=0}^{\infty} a_i 3^{-i-1} - \sum_{i=0}^{\infty} a'_i 3^{-i-1} \right| = \left| \sum_{i=0}^{\infty} (a_i - a'_i) 3^{-i-1} \right| \geq \\ &\geq |a_{i_0} - a'_{i_0}| 3^{-i_0-1} - \sum_{i \geq i_0+1} |a_i - a'_i| 3^{-i-1} \geq 2 \cdot 3^{-i_0-1} - 2 \sum_{i > i_0} 3^{-i-1} = 2 \cdot 3^{-i_0-1} - 3^{-i_0-1} = 3^{-i_0-1} > 0. \end{aligned}$$

Assim, $x \neq x'$, o que é uma contradição. Por outro lado, φ é sobrejetora. Logo, φ é bijetora. Além disso, φ é contínua. Como $\{0, 2\}^{\mathbb{N}}$ é compacto, segue que φ^{-1} é contínua. $\square$

Proposição 17. $\{0, 1, 2, \dots, d\}^{\mathbb{N}}$ é homeomorfo a K_3 .

Demonstração. Consideremos a aplicação T de $\{0, 1, \dots, d\}$ no conjunto de todas as palavras finitas formadas pelos elementos de $\{0, 1\}$, tal que $T(0) = 0, T(1) = 10, T(2) = 1^2 0 = 110, \dots, T(s-2) = 1^{s-2} 0, T(s-1) = 1^{s-1} 0$ onde $1^r 0 = \underbrace{1 \cdots 1}_r 0$. Dada esta aplicação, por concatenação, podemos construir $T_{\infty} : \{0, 1, \dots, d\}^{\mathbb{N}} \rightarrow \{0, 1\}^{\mathbb{N}}$ dada por

$$T_{\infty}(a_0, a_1, \dots, a_n, \dots) = (T(a_0), T(a_1), \dots).$$

Verificaremos que T_{∞} é um homeomorfismo.

(i) T_{∞} é contínua.

De fato, se $d((a_k)_k, (a'_k)_k) < 2^{-m}$, segue que $(a_k)_k$ e $(a'_k)_k$ coincidem até a m -ésima posição. Assim, $(T_{\infty}(a_k))_{\infty}$ e $(T_{\infty}(a'_k))_{\infty}$ também. Desta forma $d((T_{\infty}(a_k))_k, (T_{\infty}(a'_k))_k) < 2^{-m}$ e com isso, concluímos a continuidade de T_{∞} em $\{0, 1\}^{\mathbb{N}}$

(ii) T_{∞} é bijetiva.

Exibiremos uma função U_{∞} que será inversa da T_{∞} . Considere $(b_n)_n$ uma sequência. Construiremos $a_n = U_{\infty}(b_n)$ por recorrência. Seja k o número 1's que aparecem no começo

de (b_n) (repare que k pode ser zero ou infinito). Se $k < s - 1$, tomaremos $a_0 = k$ e $n_0 = k + 1$. O procedimento continuará na sequência $(b_{n_0}, b_{n_0+1}, \dots)$. Caso contrário, colocaremos $a_0 = s - 1$, $n_0 = s - 1$ e continuamos o mesmo para a sequência $(b_{n_0}, b_{n_0+1}, \dots)$. Construímos U_∞ de modo que $T_\infty(U_\infty(b_n)) = b_n$. Com efeito, após a primeira etapa, a sequência $(T(a_0), b_{n_0}, b_{n_0+1}, \dots)$ coincide com a sequência $(b_n)_n$. Reciprocamente, se $(a_n)_n$ é uma sequência de $\{0, 1, \dots, s-1\}^\mathbb{N}$, mostraremos que $U_\infty(T_\infty(a_n)) = (a_n)_n$. Se $a_0 < s - 1$, a quantidade de 1 do começo da sequência $T_\infty(a_n)$ é a_0 , seguido de um 0. Assim $U_\infty(T_\infty(a_n))$ começa por a_0 . O processo continuará por $(T(a_1), T(a_2), \dots)$. No caso onde $a_0 = s - 1$, a sequência $T_\infty((a_n)_n)$ começa com $s - 1$ números 1, de modo que $U_\infty(T_\infty(a_n)_n)$ também comece por a_0 . Basta repetirmos, agora, o método para a sequência. Assim, U_∞ é a inversa de T_∞ .

iii) U_∞ é contínua por causa da compacidade de $\{0, 1, \dots, d\}$. Portanto, T_∞ é um homeomorfismo. $\square$

Proposição 18. *A transformação shift é topologicamente transitiva.*

Demonstração. Seja C_k o conjunto de todas as palavras de tamanho k sobre o alfabeto $\{0, 1, \dots, d\}$. Assim, $C_k = \{B_1^{(k)}, B_2^{(k)}, \dots, B_{d^k}^{(k)}\}$. Definamos $u \in A^\mathbb{N}$ da seguinte maneira:

$$u = B_1^{(1)} B_2^{(1)} \dots B_d^{(1)} B_1^{(2)} B_2^{(2)} \dots B_{d^2}^{(2)} \dots$$

Mostraremos que u possui uma órbita densa. Sejam $x \in A^\mathbb{N}$ e $\varepsilon > 0$. Consideremos $n \in \mathbb{N}$ tal que $2^{-n} < \varepsilon$. Seja t tal que $B_t^{(n)}$ coincida com as n primeiras letras de x . Sabemos que existe k tal que $T^k(u) = B_t^{(n)} B_{t+1}^{(n)} \dots$. Então $d(T^k(u), x) < 2^{-n} < \varepsilon$. Logo, u tem órbita densa. $\square$

Capítulo 2

Teoria Ergódica

Neste capítulo definiremos alguns conceitos da teoria ergódica, como medida invariante, medida de Bernoulli e transformações ergódicas (junto com algumas de suas equivalências). Provaremos também o teorema de Birkhoff.

Definição 19. *Seja $(X, \mathcal{A}, \mu)$ um espaço de medida e $f : X \rightarrow X$ uma transformação mensurável. Dizemos que μ é invariante por f quando*

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

para todo conjunto mensurável E .

Exemplo 2.0.1. *Sejam $X = \{0, 1, \dots, d\}^{\mathbb{N}^*}$ e $p_1, \dots, p_d \in]0, 1[$ tais que $\sum_{i=1}^d p_i = 1$, definimos a chamada medida de Bernoulli em $\mathcal{A}$ por*

$$\mu([a_1, \dots, a_d]) = p_{a_1} \dots p_{a_d},$$

para quaisquer $a, a_1, a_2, \dots, a_n \in \{0, 1, \dots, d\}$.

Teorema 20. *(Teorema de Poincaré) Seja (X, T) um sistema dinâmico mensurável e μ uma medida invariante finita. Considere $E \subset X$ um conjunto mensurável com medida positiva. Então para μ -quase todo ponto $x \in E$, existe $n \geq 1$ natural tal que $T^n(x) \in E$.*

Proposição 21. *A medida de Bernoulli é invariante pela aplicação shift.*

Demonstração. Seja $a_1, \dots, a_n \in \{0, 1, \dots, d\}^{\mathbb{N}^*}$. Então

$$\begin{aligned}\mu(T^{-1}([a_1, \dots, a_n])) &= \mu\left(\bigcup_{i=0}^d [i; \dots, a_n]\right) = \\ \sum_{i=0}^d \mu([i; a_1, \dots, a_n]) &= \sum_{i=0}^d p_i = \mu([a_1, \dots, a_n]).\end{aligned}$$

□

Definição 22. *Seja $(X, \mathcal{A}, \mu)$ um espaço de medida. Suponha que $\mu(X) = 1$. Uma transformação que preserva medida é chamada de ergódica quando somente os elementos $B \in \mathcal{A}$ com $T^{-1}(B) = B$ satisfazem $\mu(B) = 0$ ou $\mu(B) = 1$.*

Mas nem sempre é fácil provarmos que uma função é ergódica usando as definições. Entretanto, existem várias equivalências aqui.

Teorema 23. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade e $T : X \rightarrow X$ uma transformação que preserva medida. As seguintes afirmações são equivalentes:*

- (i) T é ergódica
- (ii) Se B é um elemento de $\mathcal{A}$ tal que $\mu(T^{-1}(B) \triangle B) = 0$, então $\mu(B) = 0$ ou $\mu(B) = 1$.
- (iii) Para qualquer $A \in \mathcal{A}$ vale que

$$\mu\left(\bigcup_{n=1}^{\infty} T^{-n}(A)\right) = 1.$$

- (iv) Para quaisquer $A, B \in \mathcal{A}$, com $\mu(A) > 0$, $\mu(B) > 0$, existe $n > 0$ tal que

$$\mu\left(\bigcup_{n=1}^{\infty} T^{-n}(A) \cap B\right) > 0.$$

Demonstração. ((i) $\Rightarrow$ (ii)) Considere $B \in \mathcal{A}$ tal que $\mu(T^{-1}(B) \triangle B) = 0$. Queremos concluir que $\mu(B) = 0$ ou $\mu(B) = 1$. Construiremos $B_1 \in \mathcal{A}$ tal que $T(B_1) = B_1$ e $\mu(B_1 \triangle B) = 0$. Como T é ergódica, poderemos concluir que $\mu(B_1) = 0$ ou $\mu(B_1) = 1$. Observe que

$$\mu(B_1 \triangle B) = 0 \Rightarrow \mu(B \cup B_1) - \mu(B \cap B_1) = 0 \Rightarrow \mu(B \cup B_1) = \mu(B \cap B_1).$$

Como $B_1 \subset B_1 \cup B$, se $\mu(B_1) = 1$, então $\mu(B \cup B_1) = 1$. Além disso, $\mu(B \cap B_1) = 1$. Desta forma,

$$\mu(B \cup B_1) = \mu(B) + \mu(B_1) - \mu(B \cap B_1) \Rightarrow 1 = \mu(B) + 1 - 1 \Rightarrow \mu(B) = 1.$$

Além disso, como $B_1 \cap B \subset B_1$, segue que se $\mu(B_1) = 0$, então $\mu(B_1 \cap B) = 0$, de onde segue que $\mu(B \cup B_1) = 1$. Desta forma,

$$\mu(B \cup B_1) = \mu(B) + \mu(B_1) - \mu(B \cap B_1) \Rightarrow 0 = \mu(B) + 0 - 0 \Rightarrow \mu(B) = 0.$$

Ou seja, se encontrarmos B_1 , podemos resolver o problema. Aqui, escolheremos

$$B_1 := \bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} T^{-i}(B).$$

Precisaremos mostrar que $T^{-1}(B_1) = B_1$ e $\mu(B_1 \triangle B) = 0$. Com efeito,

$$\mu(B_1 \triangle B) = \mu\left(\bigcap_{n=0}^{\infty} \bigcup_{i=n}^{\infty} T^{-i}(B) \triangle B\right).$$

Por outro lado, para cada $n \geq 0$, temos

$$\mu\left(\bigcup_{i=n}^{\infty} T^{-i}(B) \triangle B\right) = \mu\left(\bigcup_{i=n}^{\infty} (T^{-i}(B) \triangle B)\right) \leq \sum_{i=n}^{\infty} \mu(T^{-i}(B) \triangle B).$$

Como

$$T^{-i}(B) \triangle B \subset \bigcup_{j=0}^{i-1} T^{-(j+1)}(B) \triangle T^{-j}(B) = \bigcup_{j=0}^{i-1} T^{-j}(T^{-1}(B) \triangle B)$$

e T é invariante, segue que

$$\begin{aligned} \mu(T^{-i}(B) \triangle B) &\leq \mu\left(\bigcup_{j=0}^{i-1} T^{-j}(T^{-1}(B) \triangle B)\right) \leq \sum_{j=0}^{i-1} \mu(T^{-j}(T^{-1}(B) \triangle B)) = \\ &\leq \sum_{j=0}^{i-1} \mu(T^{-1}(B) \triangle B) = i\mu(T^{-1}(B) \triangle B) = 0. \end{aligned}$$

Com isso, $\mu(T^{-i}(B) \triangle B) = 0$, para todo $i \geq 0$. Desta forma,

$$\mu\left(\bigcup_{i=1}^n (B \triangle T^{-i}(B))\right) \leq \sum_{i=1}^n \mu(B \triangle T^{-i}(B)) = 0.$$

Como $\bigcup_{i=n}^{\infty} (T^{-i}(B) \triangle B) \supset \bigcup_{i=n+1}^{\infty} (T^{-i}(B) \triangle B)$, segue que $\mu(B_1 \triangle B) = 0$ e assim $\mu(B) = \mu(B_1)$. Desde que T é ergódica segue que $\mu(B_1) = 0$ ou $\mu(B_1) = 1$, ou seja, $\mu(B) = 0$

ou $\mu(B) = 1$.

((ii) $\Rightarrow$ (iii)) Seja $A \in \mathcal{A}$ com $\mu(A) > 0$. Além disso, considere $A_1 := \bigcup_{n=1}^{\infty} T^{-n}(A)$. Basta provarmos que $\mu(T^{-1}(A) \triangle A) = 0$. Sabemos que T preserva medida. Então $\mu(T^{-1}(A_1)) = \mu(A_1)$. Além disso,

$$T^{-1}(A_1) = T^{-1}\left(\bigcup_{n=1}^{\infty} T^{-n}(A)\right) = \bigcup_{n=1}^{\infty} T^{-1}(T^{-n}(A)) = \bigcup_{n=2}^{\infty} T^{-n}(A) \subset \bigcup_{n=1}^{\infty} T^{-n}(A) = A_1.$$

Com isso, $T^{-1}(A_1) \cup A_1 = A_1$ e $T^{-1}(A_1) \cap A_1 = T^{-1}(A_1)$. Daí

$$\mu(T^{-1}(A) \triangle A) = \mu(T^{-1}(A_1) \cup A_1) - \mu(T^{-1}(A_1) \cap A_1) = \mu(A_1) - \mu(T^{-1}(A_1)) = 0.$$

Pela hipótese, $\mu(A_1) = 0$ ou $\mu(A_1) = 1$. Basta provarmos que $\mu(A_1) > 0$. Com efeito,

$$T^{-1}(A) \subset \bigcup_{n=1}^{\infty} T^{-n}(A) = A_1$$

Daí $\mu(A_1) \geq \mu(T^{-1}(A)) = \mu(A) > 0$.

((iii) $\Rightarrow$ (iv)) Sejam $A, B \in \mathcal{A}$, com $\mu(A), \mu(B) > 0$. Por (iii),

$$\mu\left(\bigcup_{n=1}^{\infty} T^{-n}(A)\right) = 1.$$

Observemos que

$$\mu\left(\bigcup_{n=1}^{\infty} T^{-n}(A) \cup B\right) = \mu(B) + \mu\left(\bigcup_{n=1}^{\infty} T^{-n}(A)\right) - \mu\left(\bigcup_{n=1}^{\infty} T^{-n}(A) \cap B\right),$$

ou seja,

$$1 = \mu(B) + 1 - \mu\left(\bigcup_{n=1}^{\infty} T^{-n}(A) \cap B\right).$$

Logo,

$$\mu\left(\bigcup_{n=1}^{\infty} T^{-n}(A) \cap B\right) = \mu(B) > 0.$$

((iv) $\Rightarrow$ (i)) Considere $B \in \mathcal{A}$ com $T^{-1}(B) = B$. Suponha, por absurdo que $0 < \mu(B) < 1$. Então,

$$0 = \mu(\emptyset) = \mu(B \cap (X - B)) = \mu(T^{-n}(B) \cap (X - B)).$$

Isto contradiz (iv). Desta forma, $\mu(B) = 0$ ou $\mu(B) = 1$. □

Teorema 24. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade e $T : X \rightarrow X$ uma transformação que preserva medida. São equivalentes:*

- (i) *T é ergódica.*
- (ii) *Para toda f mensurável, com $(f \circ T)(x) = f(x)$, para todo $x \in X$, f é constante μ -qtp.*
- (iii) *Para toda f mensurável, com $(f \circ T)(x) = f(x)$ μ -qtp, temos f constante μ -qtp.*
- (iv) *Para toda $f \in L^2(\mu)$ com $(f \circ T)(x) = f(x)$, f é constante μ -qtp.*
- (v) *Para toda $f \in L^2(\mu)$ com $(f \circ T)(x) = f(x)$ μ -qtp, f é constante μ -qtp.*

Demonstração. (iii) $\Rightarrow$ (ii) Se $(f \circ T)(x) = f(x)$, então $(f \circ T)(x) = f(x)$ μ -qtp, e pelo item (iii) f é constante μ -qtp.

(ii) $\Rightarrow$ (v) Se $f \in L^2(\mu)$, então f é mensurável. Além disso, se $(f \circ T)(x) = f(x)$, então $(f \circ T)(x) = f(x)$, μ -qtp. Pelo item (ii), f é constante μ -qtp.

(v) $\Rightarrow$ (iv) Se $(f \circ T)(x) = f(x)$, então $(f \circ T)(x) = f(x)$ μ -qtp. Pelo item (v), f é constante μ -qtp.

(i) $\Rightarrow$ (iii) Sejam T uma aplicação ergódica e f mensurável, com $(f \circ T)(x) = f(x)$ μ -qtp. Vamos supor que f assume valores em $\mathbb{R}$, pois se assumisse valores em $\mathbb{C}$, bastaria analisar as partes real e imaginária. Para cada $k, n \in \mathbb{N}$, seja $X(k, n)$ o conjunto definido por

$$X(k, n) := \{x : \frac{k}{2^n} \leq f(x) < \frac{k+1}{2^n}\} = f^{-1}([\frac{k}{2^n}, \frac{k+1}{2^n})).$$

Observe que

$$T^{-1}(X(k, n)) \triangle X(k, n) \subset \{x : (f \circ T)(x) \neq f(x)\}.$$

Assim, $\mu(T^{-1}(X(k, n)) \triangle X(k, n)) = 0$. Como T é ergódica, pelo teorema 23, $\mu(X(k, n)) = 0$ ou 1. Porém para cada n , a união $\bigcup_{k \in \mathbb{Z}} X(k, n)$ é disjunta. Assim, para cada n natural, existe um único K_n tal que $\mu((K_n, n)) = 1$. Daí, seja Y o conjunto definido por:

$$Y := \bigcap_{n=1}^{\infty} X(K_n, n).$$

Provaremos que $\mu(Y) = 1$ e f é constante em Y . De fato,

$$Y^c = \left(\bigcap_{n=1}^{\infty} X(K_n, n) \right)^c = \bigcup_{n=1}^{\infty} X(K_n, n)^c.$$

Como $\mu((K_n, n)) = 1$, segue que $\mu((K_n, n)^c) = 0$. Então,

$$\mu(Y^c) = \mu\left(\bigcup_{n=1}^{\infty} X(K_n, n)^c\right) \leq \sum_{n=1}^{\infty} \mu(X(K_n, n)^c) = 0.$$

Logo, $\mu(Y) = 1$. Por outro lado, dados $x, y \in X$, temos

$$x, y \in Y \Rightarrow x, y \in \bigcap_{n=1}^{\infty} X(K_n, n) \Rightarrow x, y \in X(K_n, n), \forall n \geq 1 \Rightarrow x, y \in f^{-1}\left(\left[\frac{K_n}{2^n}, \frac{K_n+1}{2^n}\right)\right), \forall n \geq 1 \Rightarrow$$

$$f(x), f(y) \in \left[\frac{K_n}{2^n}, \frac{K_n+1}{2^n}\right), \forall n \geq 1 \Rightarrow \frac{K_n}{2^n} \leq f(x), f(y) \leq \frac{K_n+1}{2^n} \forall n \geq 1.$$

$$\Rightarrow -\frac{1}{2^n} \leq f(x) - f(y) \leq \frac{1}{2^n}, \forall n \geq 1,$$

Logo, $f(x) = f(y)$, ou seja, f é constante em Y .

(iv) $\Rightarrow$ (i) Afirmamos que, se $T^{-1}(B) = B$, então $\chi_B \in L^2(\mu)$. Com efeito,

$$\int \chi_B^2 dx = \int \chi_B \cdot \chi_B dx = \int \chi_{B \cap B} dx = \int \chi_B dx = \mu(B) < \infty.$$

Provemos que $(\chi_B \circ T)(x) = \chi_B(x)$. Com efeito,

$$(\chi_B \circ T)(x) = \chi_B(T(x)) = \chi_{T^{-1}(B)}(x) = \chi_B(x).$$

Logo, pelo item (iv), χ_B é constante μ -qtp. □

Teorema 25. (Teorema Ergódico de Birkhoff) *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade e $T : X \rightarrow X$ uma transformação mensurável que preserva a medida μ . Então, dada qualquer função μ -integrável $f : X \rightarrow \mathbb{R}$, o limite*

$$\tilde{f}(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \quad (2.1)$$

existe para μ -quase todo ponto $x \in X$. Além disso, $\tilde{f}$ é uma função integrável com

$$\int_X \tilde{f} d\mu = \int_X f d\mu \quad e \quad \tilde{f}(T(x)) = \tilde{f}(x), \text{ em } \mu\text{-quase todo ponto } x \in X.$$

Demonstração. Começamos por provar a existência (em μ -qtp) do limite (2.1), que é o ingrediente fundamental deste enunciado. Como toda função f pode ser escrita como diferença de duas funções não negativas

$$f = \max\{f, 0\} - \max\{-f, 0\},$$

podemos, sem restrição, supor $f \geq 0$. Consideremos as funções

$$\bar{f}(x) = \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \quad \text{e} \quad \underline{f}(x) = \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)).$$

Assim, basta mostrar que

$$\int_X \bar{f} d\mu \leq \int_X f d\mu \leq \int_X \underline{f} d\mu. \quad (2.2)$$

De fato, a desigualdade (2.2) implica que

$$\int_X (\bar{f} - \underline{f}) d\mu \leq 0 \text{ e então,}$$

como $(\bar{f} - \underline{f}) \geq 0$, segue que $\bar{f} = \underline{f}$ em μ -qtp, que é precisamente o que queremos concluir.

Para provarmos a primeira desigualdade em (2.2), introduzimos $\bar{f}_K = \min\{\bar{f}, K\}$, onde $K \in \mathbb{N}$ é um inteiro fixado (grande). Também, fixamos $\epsilon > 0$ (pequeno) e para cada $x \in X$ definamos

$$t(x) = \min \left\{ n \geq 1 : \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \geq \bar{f}_K(x) - \epsilon \right\}. \quad (2.3)$$

Note que, da definição de $\limsup$, $t(x)$ sempre existe, já que $\bar{f} \geq \bar{f}_K$. Note ainda que (2.3) implica

$$\sum_{j=0}^{t(x)-1} f(T^j(x)) \geq t(x) (\bar{f}_K(x) - \epsilon). \quad (2.4)$$

Finalmente, escolhemos $M \in \mathbb{N}$ grande tal que o conjunto

$$E = \{x \in X : t(x) > M\}$$

tenha $\mu(E) \leq \frac{\epsilon}{K}$. Note que isso é possível pois a família de subconjuntos

$$E_n = \{x \in X : t(x) = n\}$$

forma uma partição do espaço X , logo tem-se

$$\sum_{n=1}^{\infty} \mu(E_n) = \mu(X) = 1$$

e, portanto,

$$\mu(E) = \sum_{n>M} \mu(E_n) \leq \frac{\epsilon}{K} \quad (2.5)$$

desde que M seja suficientemente grande.

Agora, para cada $x \in X$ e $n \geq 1$ definimos sequências x_i (de iterados de x) e t_i (de inteiros positivos), da seguinte maneira:

1. Tomamos $x_0 = x$.
2. Definimos t_i e x_{i+1} da seguinte maneira:
 - a) Se $t(x_i) \leq M$ então tomamos $t_i = t(x_i)$ e $x_{i+1} = T^{t_i}(x_i)$.
 - b) Se $t(x_i) > M$ então tomamos $t_i = 1$ e $x_{i+1} = T(x_i)$.
3. Terminamos quando encontramos x_s tal que $t_0 + t_1 + \dots + t_{s-1} + t_s \geq n$.

Com isso, temos

$$t_0 + t_1 + \dots + t_{s-1} \geq n - M. \quad (2.6)$$

Em a), usando (2.4) e sendo $\mathcal{X}_E$ a função característica de E , obtemos

$$\sum_{j=0}^{t_i-1} (f + K\mathcal{X}_E)(T^j(x_i)) \geq \sum_{j=0}^{t_i-1} f(T^j(x_i)) \geq t_i(\bar{f}_K(x_i) - \epsilon) = t_i(\bar{f}_K(x) - \epsilon), \quad (2.7)$$

onde a última igualdade segue de que $\bar{f}_K(T(y)) = \bar{f}_K(y)$ para todo y . Por outro lado, no caso b) temos

$$\sum_{j=0}^{t_i-1} (f + K\mathcal{X}_E)(T^j(x_i)) = f(x_i) + K \geq K \geq t_i(\bar{f}_K(x) + \epsilon). \quad (2.8)$$

Então, usando (2.6), (2.7) e (2.8) (vamos usar $\tau = t_0 + t_1 + \dots + t_{s-1}$ para simplificar a notação) temos

$$\begin{aligned} \sum_{j=0}^{n-1} (f + K\mathcal{X}_E)(T^j(x)) &= \sum_{i=0}^{s-1} \left(\sum_{j=0}^{t_i-1} (f + K\mathcal{X}_E)(T^j(x_i)) \right) + \sum_{j=\tau}^{n-1} (f + K\mathcal{X}_E)(T^j(x)) \geq \\ &\geq \left(\sum_{i=0}^{s-1} t_i \right) (\bar{f}_K(x) - \epsilon) + 0 \geq (n - M)(\bar{f}_K(x) - \epsilon). \end{aligned} \quad (2.9)$$

Uma vez que a medida μ é T -invariante, tem-se

$$\int_X (f + K\mathcal{X}_E) \circ T^j d\mu = \int_X (f + K\mathcal{X}_E) d\mu, \text{ para todo } j.$$

Assim, integrando o primeiro e o último membro de (2.9) obtemos

$$n \left(\int_X f d\mu + K\mu(E) \right) \geq (n - M) \left(\int_X \bar{f}_K d\mu - \epsilon \right).$$

Dividindo ambos os termos por n , usando (2.5) e fazendo $n \rightarrow \infty$, deduzimos que

$$\int_X f d\mu + \epsilon \geq \int_X \bar{f}_K d\mu - \epsilon, \text{ para todo } \epsilon > 0 \text{ e, portanto, } \int_X f d\mu \geq \int_X \bar{f}_K d\mu.$$

Então, fazendo $K \rightarrow \infty$ e observando que $\bar{f}_K$ converge monotonamente para $\bar{f}$, obtemos

$$\int_X f d\mu \geq \int_X \bar{f} d\mu,$$

como pretendíamos.

A segunda desigualdade em (2.2) é provada de modo muito semelhante. Apenas indicarei as diferenças com relação ao caso anterior. Considere

$$t(x) = \min \left\{ n \geq 1 : \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) \leq \underline{f}(x) + \epsilon \right\}. \quad (2.10)$$

(comparar com (2.3)) e defina x_i e t_i como foi feito anteriormente. Tome E como antes e M fixado de tal modo que

$$\int_E f d\mu \leq \epsilon. \quad (2.11)$$

Aplicando em $(f - f\mathcal{X}_E)$ o mesmo tipo de cálculo que usamos em (2.7), (2.8) e (2.9) para a função $(f + K\mathcal{X}_E)$, obtemos

$$\begin{aligned}
\sum_{j=0}^{n-1} (f - f\chi_E)(T^j(x)) &= \sum_{i=0}^{s-1} \left(\sum_{j=0}^{t_i-1} (f - f\chi_E)(T^j(x_i)) \right) + \sum_{j=\tau}^{n-1} (f - f\chi_E)(T^j(x)) \leq \\
&\leq \left(\sum_{i=0}^{s-1} t_i \right) (\underline{f}(x) + \epsilon) + \sum_{j=\tau}^{n-1} f(T^j(x)) \leq n(\underline{f}(x) + \epsilon) + \sum_{j=\tau}^{n-1} f(T^j(x)). \tag{2.12}
\end{aligned}$$

(comparar com (2.9)). Integrando e sabendo que $n - \tau \leq M$, por (2.6), temos

$$n \left(\int_X f d\mu - \int_E f d\mu \right) \leq n \left(\int_X \underline{f} d\mu + \epsilon \right) + M \int_X f d\mu.$$

Dividindo ambos os lados por n , fazendo $n \rightarrow \infty$ e usando (2.11), temos

$$\int_X d\mu \leq \int_X \underline{f} d\mu + \epsilon + \epsilon, \text{ para todo } \epsilon > 0 \text{ e, portanto, } \int_X f d\mu \leq \int_X \underline{f} d\mu.$$

Com isso, encerramos a demonstração da relação (2.2).

Provemos agora que $\tilde{f} \circ T = \tilde{f}$. De fato, seja $f : X \rightarrow \mathbb{R}$ uma função integrável. Daí, por (2.1), temos

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

Como f é integrável, segue que f é limitada para μ -quase todo ponto $x \in X$ e, portanto, o último limite é igual a zero, provando assim tal igualdade. $\square$

Proposição 26. *Sejam $T : X \rightarrow X$ uma transformação mensurável e μ uma probabilidade invariante sob T . As seguintes condições são equivalentes:*

$$(i) \quad \tilde{f}(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)) = \int_X f d\mu,$$

para toda função μ -integrável f e para μ -quase todo $x \in X$.

(ii) μ é ergódica.

Demonstração. ((i) $\Rightarrow$ (ii))

Seja A um conjunto T -invariante e considere $f = \chi_A$. Da hipótese (i) temos que

$$\tilde{f}(x) = \int_X f d\mu = \mu(A)$$

para quase todo ponto $x \in X$. Por outro lado, como A é T -invariante, temos que $x \in A$ se, e somente se, $T(x) \in A$. Isso implica que $f(T^j(x)) = f(x)$ para todo $j \geq 0$ e para todo $x \in X$. Portanto,

$$\tilde{f}(x) = f(x) = \chi_A(x)$$

para todo $x \in X$. Como a função característica só toma valores 0 e 1, essas duas igualdades implica que $\mu(A) = 0$ ou $\mu(A) = 1$, como afirmado em (ii).

((ii) $\Rightarrow$ (i))

Seja

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

Podemos ver que $\tilde{f} \circ T = \tilde{f}$. Como μ é ergódica, segue que $\tilde{f}(x)$ é constante μ -qtp. Por outro lado,

$$\int f(x) d\mu = \int \tilde{f}(x) d\mu,$$

que é constante. Desta forma,

$$\int \tilde{f}(x) d\mu = \tilde{f}(x)$$

para μ -quase todo $x \in X$. □

Proposição 27. *Se T é a aplicação shift, então (T, μ) é ergódico.*

Capítulo 3

Teoremas de Van Der Waerden e de Szemerédi

Neste capítulo daremos duas provas do teorema de Van Der Waerden, uma combinatória e outra envolvendo sistemas dinâmicos. Para esta última, usaremos o teorema da recorrência múltipla ergódica de Furstenberg.

Teorema 28. (*Van Der Waerden*) *Se colorirmos os inteiros positivos usando k cores, existem progressões aritméticas de comprimento arbitrário cujos termos são todos da mesma cor. Em outras palavras, se $\mathbb{N} = \bigcup_{i=1}^k A_i$, então para qualquer $l \geq 1$, existem $x_0, r \in \mathbb{N}$ e $i \in \{1, 2, \dots, k\}$ tais que $x_0, x_0 + r, x_0 + 2r, \dots, x_0 + (l-1)r \in A_i$.*

Vamos dar uma prova combinatória deste teorema.

Observação 2. *Vamos usar a notação $a + [0, k)r$ para representarmos a progressão aritmética $a, a + r, a + 2r, \dots, a + (k-1)r$.*

Observação 3. *Suponha que vamos colorir os números naturais de 1 a N e vamos colorí-los usando m cores. Sejam $1, 2, 3, \dots, m$ as cores. Para cada número de 1 a N , associamos uma cor de 1 a m . Podemos pensar nisso como sendo uma função $c : \{1, 2, \dots, N\} \rightarrow \{1, 2, \dots, m\}$.*

Definição 29. *Sejam $k \geq 1$, $d \geq 0$, $r_1, r_2, \dots, r_d$ e a um inteiro tal que $1 \leq a \leq N$. Um ventilador é uma d -upla de progressões aritméticas $(a + [0, k)r_1, a + [0, k)r_2, \dots, a + [0, k)r_d)$.*

Observação 4. Na definição anterior, k é o raio, d o grau, a é o ponto base e as progressões aritméticas $a + [1, k)r_i$ são chamadas de pás do ventilador, para $1 \leq i \leq d$.

Definição 30. Seja $c : \{1, 2, \dots, N\} \rightarrow \{1, 2, \dots, m\}$ uma maneira de colorir os números. Uma pá $a + [1, k)r_i$ é chamada de monocromática quando $c(a + jr_i) = c(a + lr_i)$ para quaisquer $1 \leq j \leq l \leq k - 1$. Além disso, vamos dizer que um ventilador é policromático se seu ponto base e suas pás são monocromáticas, isto é, existem cores $c_0, c_1, \dots, c_d$ distintas tais que $c(a) = c_0$ e $c(a + jr_i) = c_i$ para $1 \leq j \leq k - 1$ e $1 \leq i \leq d$.

Teorema 31. Dados $k, m \geq 1$, existe N tal que qualquer coloração com m cores de $\{1, 2, \dots, N\}$ contém uma progressão aritmética de comprimento k monocromática.

Demonstração. Vamos usar duas induções. Começamos com a indução em k . O caso $k = 1$ é verdadeiro. Com efeito, é só tomarmos $N = 1$, pois ao pintarmos o conjunto $\{1\}$, teremos uma progressão aritmética formada pelo 1 de comprimento 1 monocromática.

Vamos supor que o teorema é válido para $k - 1$, onde $k \geq 2$.

Afirmção: dado $d \geq 0$, existe N tal que qualquer coloração com m cores de $\{1, 2, \dots, N\}$ contém uma progressão aritmética monocromática ou um ventilador policromático de raio k e grau d .

Observe que o caso $d = m$ é equivalente a dizer que existe N tal que qualquer coloração com m cores de $\{1, 2, \dots, N\}$ contém uma progressão aritmética monocromática ou um ventilador policromático de raio k e grau d . Porém não existe um ventilador policromático de grau m . Com efeito, se ele existisse, você precisaria de $m + 1$ cores, afinal m delas para pintar as pás e uma delas para pintar a base. Mas só temos m cores. Com isso, só vai poder acontecer que ao colorirmos $1, 2, \dots, N$ com m cores, aparecerá uma progressão aritmética monocromática, que é justamente o que queremos provar.

Começamos provando o caso $d = 0$. Neste caso, basta tomarmos $N = 1$, pois colorindo o conjunto 1, teremos um ventilador de grau 0 (basta colocarmos o 1 na base).

Vamos supor que a afirmação seja válida para $d - 1$, com $d \geq 1$.

Daí, existe N_1 tal que qualquer coloração com m cores de $\{1, 2, \dots, N_1\}$ contém uma progressão aritmética monocromática ou um ventilador policromático de raio k e grau $d - 1$.

Além disso, este resultado vale para quaisquer conjuntos que possuem N_1 elementos, inclusive para os conjuntos da forma $\{bkN_1 + 1, bkN_1 + 2, \dots, bkN_1 + N_1\}$, com b natural. Sejam $N_2 > 0$ e $b \in \{1, 2, \dots, N_2\}$.

Assim, teremos N_2 conjuntos e cada um deles com N_1 elementos. Desta forma, podemos usar nossa hipótese de indução em cada um deles. Por ela, qualquer coloração com m cores de cada um destes N_2 conjuntos contém uma progressão aritmética monocromática ou um ventilador policromático de raio k e grau $d - 1$. Se para algum b , encontrarmos a progressão aritmética de comprimento k , então o problema está resolvido. Provemos que é sempre possível encontrá-la.

Suponha, por absurdo, que para todo $b \in \{1, 2, \dots, N_2\}$ encontramos um ventilador policromático de raio k e grau $d - 1$. Vamos fixar b . Dizer que encontramos um ventilador policromático de grau $d - 1$, é o mesmo que dizer que encontramos $d - 1$ progressões aritméticas de comprimento $k - 1$ no conjunto $\{bkN_1 + 1, bkN_1 + 2, \dots, bkN_1 + N_1\}$ que são monocromáticas.

Vamos transformar estas progressões aritméticas em um novo ventilador com um grau a mais. Para este b fixo, tomemos os números $a(b), r_1(b), \dots, r_{d-1}(b) \in \{1, 2, \dots, N_1\}$ tais que, associado ao número b , encontramos os ventiladores

$$(bkN_1 + a(b) + [0, k)r_1(b), bkN_1 + a(b) + [0, k)r_2(b), \dots, bkN_1 + a(b) + [0, k)r_{d-1}(b)).$$

Além disso, considere $c_0(b), c_1(b), \dots, c_{d-1}(b)$ cores distintas tais que

$$c(bkN_1 + a(b)) = c_0(b) \tag{3.1}$$

e

$$c(bkN_1 + a(b) + jr_i(b)) = c_i(b). \tag{3.2}$$

Para $j = 1, 2, \dots, k - 1$ e $i = 1, 2, \dots, d - 1$. As equações (3.1) e (3.2) serão chamadas de primeira e segunda propriedades do ventilador gerado por b .

Em particular, podemos pensar na aplicação $\alpha : \{1, 2, \dots, N_2\} \rightarrow \{1, 2, \dots, N_1\}^d \times \{1, 2, \dots, m\}^d$ tal que $\alpha(b) = (a(b), r_1(b), \dots, r_{d-1}(b), c_0(b), c_1(b), \dots, c_{d-1}(b))$. como sendo uma coloração com $N_1^d m^d$ cores.

Estamos supondo, por hipótese de indução, que o resultado vale para $k - 1$, isto é, existe N_2 tal que qualquer coloração com $N_1^d m^d$ cores de $\{1, 2, \dots, N_2\}$ possui uma progressão

aritmética de comprimento $k - 1$ monocromática, digamos, $b + [0, k - 1)s$, com alguma cor da forma $(a, r_1, r_2, \dots, r_{d-1})$.

Vamos construir este ventilador com um grau a mais. Considere $b_0 = (b - s)kN_1 + a \in \{1, 2, \dots, N\}$. Nosso ventilador será:

$$(b_0 + [0, k)skN_1, b_0 + [0, k)(skN_1 + r_1), \dots, b_0 + [0, k)(skN_1 + r_{d-1})).$$

Observe que ele possui raio k , grau d e ponto base b_0 . Porém, para terminarmos, precisamos verificar que ele é policromático.

Para isto, mostraremos que

$$c(b_0 + jskN_1) = c_0(b)$$

e

$$c(b_0 + j(skN_1 + r_t)) = c_t$$

para $1 \leq j \leq k - 1$ e $1 \leq t \leq d$.

De fato,

$$c(b_0 + jskN_1) = c((b - s)kN_1 + a + jskN_1) = c((b - s + js)kN_1 + a) = c((b + (j - 1)s)kN_1 + a).$$

Mas pela primeira propriedade do ventilador para $b + (j - 1)s$,

$$c((b + (j - 1)s)kN_1 + a(b)) = c_0(b + (j - 1)s).$$

Como a progressão aritmética $b + [0, k - 1)s$ é monocromática, segue que $c_0(b + (j - 1)s) = c_0(b)$. Além disso, se $1 \leq j \leq k - 1$ e $1 \leq t \leq d$, então

$$\begin{aligned} c(b_0 + j(skN_1 + r_t)) &= c((b - s)kN_1 + a + j(skN_1 + r_t)) = \\ &= c((b - s)kN_1 + a + jskN_1 + jr_t) = c((b - s + js)kN_1 + a + jr_t). \end{aligned}$$

Se usarmos a segunda propriedade do ventilador para $b + (j - 1)s$:

$$c((b + (j - 1)s)kN_1 + a + jr_t) = c_t(b + (j - 1)s) = c_t.$$

Se o ponto base b_0 tem a mesma cor que uma de suas pás, então encontramos uma progressão aritmética de tamanho k . Caso contrário, encontramos um ventilador policromático de raio k e grau d . $\square$

3.1 Prova do Teorema de Van der Waerden usando Sistemas Dinâmicos

Podemos provar o resultado anterior usando Teoria Ergódica. Para isto, precisamos do Teorema da Recorrência Múltipla Ergódica de Furstenberg:

Teorema 32. (*Recorrência Múltipla Ergódica de Furstenberg*). *Sejam $T : X \rightarrow X$ μ -invariante, $k \geq 3$ e $\mu(A) > 0$. Então existe N tal que:*

$$\mu(A \cap T^{-N}(A) \cap T^{-2N}(A) \cap \dots \cap T^{-(k-1)N}(A)) > 0.$$

Demonstração. (Teorema de Van der Waerden) Consideraremos $A = \{c_1, c_2, \dots, c_k\}$ o conjunto das cores. Este será o nosso "alfabeto". Assim, nossa transformação shift pode ser escrita como $T : A^{\mathbb{N}} \rightarrow A^{\mathbb{N}}$. Considere também $z = (z_1, z_2, \dots) \in A^{\mathbb{N}}$ uma maneira de colorir $\mathbb{N}$.

Observe que

$$\begin{aligned} d(T^m(x), T^l(y)) < 1 &\Leftrightarrow d((x_{m+1}, x_{m+2}, x_{m+3}, \dots), (y_{l+1}, y_{l+2}, y_{l+3}, \dots)) < 1 \Leftrightarrow \\ &\Leftrightarrow x_{m+1} = y_{l+1}. \end{aligned}$$

Precisamos encontrar uma progressão aritmética monocromática, isto é, $m, m+n, m+2n, \dots, m+kn$ de tal forma que $z_m = z_{m+n} = z_{m+2n} = \dots = z_{m+kn}$. Para isto, basta mostrarmos que existe $m \in \mathbb{N}$ tal que

$$d(T^{m-1}(z), T^{m-1+in}(z)) < 1$$

para $i = 1, 2, \dots, k$. Com efeito, basta usarmos o Teorema da Recorrência Múltipla Ergódica de Furstenberg e considerarmos

$$d(T^{m-1}(z), T^{m-1+in}(z)) = d(T^{m-1}(z), T^{in}(T^{m-1}(z))).$$

□

3.2 Teorema de Szemerédi

Definiremos aqui a noção de densidade e provaremos o teorema de Szemerédi.

Definição 33. *Seja $A \subset \mathbb{N}$. A densidade de A é definida por*

$$d(A) := \limsup_{N \rightarrow \infty} \frac{|[1, N] \cap A|}{N}.$$

Exemplo 3.2.1. *Seja P o conjunto dos números pares. Então $|[1, N] \cap P|$ é igual a $\frac{N-1}{2}$, se N é ímpar e $\frac{N}{2}$, se N é par. Logo, a densidade do conjunto dos números pares é $\frac{1}{2}$.*

Teorema 34. (Szemerédi) *Se $A \subset \mathbb{N}$ tem densidade positiva, então A possui infinitas progressões aritméticas de comprimento arbitrariamente grande, isto é, para todo p , existem $x_0 \in \mathbb{N}$ e $m \in \mathbb{N}$ tais que $x_0, x_0 + m, \dots, x_0 + (p-1)m \in A$.*

Demonstração. Vamos usar o teorema da recorrência múltipla ergódica de Furstenberg. Consideremos $X = \{0, 1\}^{\mathbb{N}}$, $T : X \rightarrow X$ a aplicação shift, $(x_n) = \chi_A(n)$. Para todo $k \geq 1$, seja μ_k a medida definida por

$$\mu_k = \frac{1}{k} \sum_{j=0}^{k-1} \delta_{T^j(x)},$$

onde $\delta_{T^j(x)}(A)$ é igual a 1, se $T^j(x) \in A$ e 0, caso contrário. Esta função é chamada de medida de Dirac. Por outro lado, como μ_k são medidas de probabilidade e usando o teorema de Banach-Alaoglu, podemos supor que $\mu := \lim_{n_k} \mu_{n_k}$ é uma medida invariante para T . Tome também $Y = \{(y_n) : y_1 = 1\}$. Provaremos que $\mu(Y) > 0$ para podermos usar o teorema de Furstenberg.

$$\mu(Y) = \lim \mu_k(Y) = \lim \frac{1}{k} \sum_{j=0}^{k-1} \delta_{T^j(x)}(Y) = \lim \frac{1}{k} \sum_{j=0}^{k-1} \chi_Y(T^j(x)). \quad (3.3)$$

Alguns termos deste somatório dão 0, enquanto outros dão 1. Vamos nos preocupar com aqueles que resultam 1. Notemos que

$$\chi_Y(T^j(x)) = 1 \iff T^j(x) \in Y \iff (x_{j+1}, x_{j+2}, \dots) \in Y \iff x_{j+1} = 1 \iff$$

$$\chi_A(j+1) = 1 \iff j+1 \in A.$$

Desta forma,

$$\sum_{j=0}^{k-1} \chi_Y(T^j(x)) = \sum_{j+1 \in A, j=0,1,\dots,k-1} \chi_Y(T^j(x)) = |A \cap [1, k]|.$$

Logo, a equação (3.3) pode ser escrita como

$$\mu(Y) = \lim \frac{1}{k} |A \cap [1, k]| = d(A) > 0.$$

Pelo teorema da recorrência de Furstenberg existe N tal que

$$\mu(Y \cap T^{-N}(Y) \cap T^{-2N}(Y) \cap \dots \cap T^{-(k-1)N}(Y)) > 0.$$

Em particular, existem infinitos $z = (z_1, z_2, \dots)$ tais que $z \in Y \cap T^{-N}(Y) \cap T^{-2N}(Y) \cap \dots \cap T^{-(k-1)N}(Y)$. Notemos que

$$z \in Y \iff z_1 = 1 \iff \chi_A(1) = 1 \iff 1 \in A.$$

Além disso, para $j = 1, 2, \dots, k-1$:

$$z \in T^{-jn}(Y) \iff T^{jN}(z) \in Y \iff (z_{jN+1}, z_{jN+2}, \dots) \in Y \iff z_{jN+1} = 1$$

$$\chi_A(jN+1) = 1 \iff jN+1 \in A.$$

Logo, $1, 1+N, 1+2N, \dots, 1+(k-1)N \in A$. Isto prova o teorema de Szemerédi. $\square$

Observação 5. Em 2004, os matemáticos Ben Green e Terence Tao provaram que existem infinitas progressões aritméticas de comprimento arbitrário formadas somente de primos [7].

Capítulo 4

Equidistribuição

Neste capítulo, definiremos sequências equidistribuídas. Provaremos os teoremas de Weyl e de Koksma.

Definição 35. *Seja $(x_n)_{n \geq 0}$ uma sequência em $[0, 1[$. Dizemos que $(x_n)_{n \geq 0}$ é equidistribuída se para quaisquer $0 \leq a < b < 1$, a sequência*

$$N(a, b, n) = \#\{m \leq n : x_m \in [a, b]\}.$$

é assintoticamente equivalente à $n(b - a)$ quando n tende a infinito, isto é,

$$\lim_{n \rightarrow \infty} \frac{N(a, b, n)}{n} = b - a.$$

Mais geralmente, se $(y_n)_{n \geq 0}$ é uma sequência de números reais, dizemos que $(y_n)_{n \geq 0}$ é equidistribuída módulo 1 se a sequência $(\{y_n\})_{n \geq 0}$ é equidistribuída, onde $\{y\}$ representa a parte fracionária de y .

Proposição 36. *Se a sequência $(x_n)_{n \geq 0}$ é equidistribuída, então $\{x_n : n \in \mathbb{N}\}$ é densa em $[0, 1[$.*

Demonstração. Seja $[a, b[\subset [0, 1[$, onde $a < b$. Como

$$\lim_{n \rightarrow \infty} \frac{N(a, b, n)}{n} = b - a,$$

temos $N(a, b, n) \neq 0$. □

Proposição 37. *A sequência $(x_n)_{n \geq 0}$ é equidistribuída módulo 1 se, e somente se, para toda função f integrável à Riemann de período 1 vale a seguinte igualdade*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(x_k) = \lim_{n \rightarrow \infty} S_n(f) = \int_0^1 f(t) dt. \quad (4.1)$$

Demonstração. Suponha, sem perda de generalidade, que a sequência $(x_n)_n$ pertence ao intervalo $[0, 1[$. Como $f : [0, 1] \rightarrow \mathbb{R}$ é integrável à Riemann, então para todo $\varepsilon > 0$, existem duas funções escadas φ e γ tais que $\varphi \leq f \leq \gamma$ e $\int_0^1 (\gamma - \varphi) \leq \varepsilon$. Temos a integral de f é o limite de $\int_0^1 \gamma$ quando ε tende a 0. Suponha que $S_n(f)$ converge para $\int_0^1 f$ para toda função f integrável à Riemann e que $0 \leq a < b < 1$. Consideremos a função escada que vale 1 em $[a, b[$ e 0 no complementar deste intervalo. Então, $N(a, b, n)$ converge, por hipótese, para $\int_a^b 1 = b - a$, pois a função escada é integrável à Riemann. Assim, $(x_n)_n$ é equidistribuída.

Reciprocamente, seja φ uma função escada de $[0, 1[$ em $\mathbb{R}$. Assim, existem uma partição $\{0 = a_0 < a_1 < \dots < a_m = 1\}$ e números reais $c_1, c_2, \dots, c_m$ tais que $f(x) = c_i$ se $a_{i-1} < x < a_i$. Observe que, se $a = b$, então $N(a, b, n)$ está bem definido e é igual a $o(n)$ quando n tende para infinito. Com efeito, tomemos $b = a + \varepsilon$, onde $\varepsilon > 0$ é arbitrário. Então

$$N(a, a, n) \leq N(a, b, n) \leq \varepsilon n + o(n) \leq 2\varepsilon n,$$

se n é suficientemente grande. Note que

$$S_n(f) = n^{-1} \left(\sum_{k=1}^m c_k N(a_{k-1}, a_k, n) - \sum_{k=0}^m f(a_k) N(a_k, a_k, n) \right).$$

Em particular, $S_n(f)$ converge para $\sum_{k=1}^m (a_k - a_{k-1}) c_k = \int_0^1 f$.

Se f é uma função integrável à Riemann qualquer, sejam $\varepsilon > 0$ e φ, γ funções escadas tais que $\varphi \leq f \leq \gamma$ e $\int_0^1 (\gamma - \varphi) \leq \varepsilon$. Então $S_n(\varphi) \leq S_n(f) \leq S_n(\gamma)$. Se n é suficientemente grande, $S_n(\gamma) \leq \int_0^1 (\gamma + \varepsilon)$ e $S_n(\varphi) \geq \int_0^1 (\varphi - \varepsilon)$. Então $|S_n(f) - \int_0^1 f| \leq 2\varepsilon$. Desta forma, podemos concluir que $S_n(f)$ converge para $\int_0^1 f$. $\square$

Proposição 38. *A sequência $(x_n)_{n \geq 0}$ é equidistribuída se, e somente se, a igualdade (4.1) vale para toda função contínua de período 1.*

Demonstração. Como toda função contínua é integrável à Riemann, a condição é evidentemente necessária. Mostremos que ela também é suficiente. De fato, seja φ uma função que vale 0 em $[0, x[$ e 1 em $[x, 1]$. Sejam ε e φ_ε uma função que vale 0 em $[0, x[$, $\frac{x-\varepsilon}{\varepsilon}$ em $[x, x+\varepsilon[$ e 1 em $[x+\varepsilon, 1]$. A função φ_ε é contínua e $\varphi_\varepsilon \leq \varphi$. Assim, $S_n(\varphi) \geq S_n(\varphi_\varepsilon)$ que tende para $\int_0^1 \varphi_\varepsilon = 1 - x - \varepsilon$. Em particular, $S_n(\varphi) \geq 1 - x - \varepsilon$ para n suficientemente grande.

Provamos também que $S_n(\varphi) \leq 1 - x + \varepsilon$ para n suficientemente grande, o que nos permite concluir que $S_n(\varphi)$ converge para $1 - x = \int_0^1 \varphi$. Assim, $S_n(\varphi)$ converge para $\int_0^1 \varphi$, para toda função em escada. Em particular para a função $\varphi = \chi_{[a,b]}$.

Logo, podemos concluir que $(x_n)_n$ é equidistribuída. $\square$

Teorema 39. (*Critério de Weyl*) A sequência $(x_n)_{n \geq 0}$ é equidistribuída se, e somente se, para todo $m \in \mathbb{Z}^*$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^n e^{2im\pi x_k} = 0,$$

Demonstração. Se aplicarmos a proposição anterior para as funções $e^{2i\pi m x}$ para $m \in \mathbb{Z}$, obtemos a condição enunciada, que é necessária.

Para a recíproca, sabemos que toda função contínua f pode ser aproximada uniformemente em $[0, 1[$ por polinômios trigonométricos. Seja T um polinômio tal que $|T(x) - f(x)| \leq \varepsilon$. Então, $|S_n(T) - S_n(f)| \leq \varepsilon$ e $S_n(T)$ converge para $\int_0^1 T$. Como $|\int_0^1 T - \int_0^1 f| \leq \varepsilon$, para ε arbitrariamente pequeno, $S_n(f)$ converge para $\int_0^1 f$. $\square$

Corolário 40. Seja θ irracional, então a sequência $(n\theta)_{n \geq 0}$ é equidistribuída módulo 1

Demonstração. Seja $m \in \mathbb{Z}$, temos $x_k = k\theta$, $k \geq 0$. Calculemos $S_n(e^{2i\pi m k \theta})$.

$$S_n = \frac{1}{n} \sum_{k=1}^n e^{2i\pi m k \theta} = \frac{1 - e^{2i\pi(n+1)\theta}}{n(1 - e^{2i\pi n \theta})},$$

pois $e^{2i\pi m \theta}$ é não nulo. Podemos ver que

$$|S_n| \leq \frac{2}{n|1 - e^{2i\pi n \theta}|} \rightarrow 0.$$

$\square$

Teorema 41. (Koksma) Para quase todo $x > 1$, a sequência $(x^n)_n$ é equidistribuída módulo 1 (relativamente à medida de Lebesgue).

Lema 42. Considere $(f_n)_{n \geq 1}$ uma sequência de funções integráveis de $[a, b]$ em $\mathbb{R}$. Se a série

$$\sum_{n=1}^{\infty} \frac{1}{n} \int_a^b \left| \frac{1}{n} \sum_{k=0}^n f_k \right|^2 dx$$

converge, então

$$\left| \frac{1}{n} \sum_{k=0}^n f_k \right|^2$$

converge para 0, para quase todo x em $[a, b]$ (relativamente à medida de Lebesgue).

Demonstração. Consideremos

$$F_N := \frac{1}{N} \sum_{n < N} f_n$$

e A_N o conjunto de todos os valores de x em $[a, b]$ tais que $|F_N| \geq \varepsilon$. Então

$$\varepsilon^2 \lambda(A_N) \leq \int_a^b |F_N|^2 dx,$$

onde λ é a medida de Lebesgue. A série

$$\sum_{N=1}^{\infty} \frac{\lambda(A_N)}{N}$$

converge, por hipótese. Afirmamos que existe uma subsequência N_k tal que $\lambda(A_{N_k})$ converge para zero. De fato, se isto não ocorresse, então $\lambda(A_n) \geq \alpha > 0$ para todo n e assim

$$\sum_{N=1}^{\infty} \frac{\lambda(A_N)}{N} \geq \sum_{N=1}^{\infty} \frac{\alpha}{N},$$

o que implicaria que a série $\sum_{N=1}^{\infty} \frac{\lambda(A_N)}{N}$ diverge. Assim, para quase todo x , $|F_{N_k}(x)| < \varepsilon$ para k suficientemente grande (repare que k depende de x). Porém a sequência (f_n) é limitada, isto é, existe uma constante C tal que $|f_n(x)| < C$, para todo n natural e $x \in [a, b]$. Se $N_k \leq N \leq N_{k+1}$, então

$$\begin{aligned} \frac{1}{N} \sum_{n < N} f_n(x) - \frac{1}{N_k} \sum_{n < N_k} f_n(x) &= \frac{1}{N} \sum_{n < N_k} f_n(x) + \frac{1}{N} \sum_{N_k \leq n < N} f_n(x) - \frac{1}{N_k} \sum_{n < N_k} f_n(x) = \\ &= \frac{1}{N} \sum_{N_k \leq n < N} f_n(x) + \left(\frac{1}{N} - \frac{1}{N_k} \right) \sum_{n < N_k} f_n(x). \end{aligned}$$

Por consequência,

$$\begin{aligned}
& |F_N(x) - F_{N_k}(x)| = \\
& = \left| \frac{1}{N} \sum_{N_k \leq n < N} f_n(x) + \left(\frac{1}{N} - \frac{1}{N_k} \right) \sum_{n < N_k} f_n(x) \right| \leq \frac{1}{N} \left| \sum_{N_k \leq n < N} f_n(x) \right| + \left| \frac{1}{N} - \frac{1}{N_k} \right| \sum_{n < N_k} |f_n(x)| \leq \\
& \leq \frac{1}{N} \sum_{N_k \leq n < N} |f_n(x)| + \left(\frac{1}{N_k} - \frac{1}{N} \right) \sum_{n < N_k} |f_n(x)| \leq \frac{1}{N} C(N - N_k) + \left(\frac{N - N_k}{NN_k} \right) \varepsilon N_k = \\
& = \frac{C(N - N_k)}{N} + \varepsilon \frac{(N - N_k)}{N}.
\end{aligned}$$

Fixemos uma sequência estritamente crescente de inteiros M_k tal que $\frac{M_{k+1}}{M_k} < 1 + \varepsilon$. Consideremos, em seguida, inteiros N_k entre M_k e M_{k+1} tais que $\mu(A_n)$ tende a 0. Assim, se N_k é escolhido de forma a minimizar $\frac{\lambda(A_N)}{N}$ no intervalo $[M_k, M_{k+1}]$, então

$$\sum_{M_k \leq N \leq M_{k+1}} \frac{\lambda(A_N)}{N} \geq (M_{k+1} - M_k) \frac{\lambda(A_{N_k})}{N_k}.$$

Concluimos que é necessário reduzir $\frac{M_{k+1} - M_k}{N_k}$ e a única maneira possível é $\frac{M_{k+1}}{M_k} - 1$. Assim, segue que

$$M_{k+1} = \lfloor (1 + \varepsilon)M_k \rfloor + 1.$$

Daí, a escolha de N_k implica que a sequência $\mu(A_N)$ tende a zero, que permite recuperar o raciocínio anterior. Assim, $N - N_k \leq M_{k+2} - M_k$. Porém $M_{k+2} \leq (1 + \varepsilon)M_{k+1} \leq (1 + \varepsilon)^2 M_k$, de onde segue que

$$M_{k+2} - M_k \leq [(1 + \varepsilon)^2 - 1]M_k \leq 2C\varepsilon N.$$

Resulta que, para $N \rightarrow \infty$ e para quase todo $x \in [a, b]$, nós temos $|F_N(x)| \leq C'\varepsilon$. Como uma união enumerável de conjuntos de medida nula ainda é um conjunto de medida nula, podemos concluir que $F_N(x)$ converge quase sempre para 0. $\square$

Demonstração. (teorema de Koksma) Pelo critério de Weyl, é suficiente provarmos que para m inteiro positivo, as somas

$$\frac{1}{n} \sum_{k=1}^n e^{2i\pi m x^k}$$

tende para 0 quase sempre. Pelo lema anterior, basta provarmos que a série

$$\sum_n \frac{1}{n} \int_a^b \left| \frac{1}{n} \sum_{k=0}^n e^{2i\pi m x^k} \right|^2 dx$$

converge para quaisquer $b \geq a \geq 1$.

Observemos que

$$\begin{aligned} & \left| \sum_{k=1}^n e^{2i\pi m x^k} \right|^2 = \\ &= \sum_{k=1}^n \sum_{l=1}^n e^{2i\pi m(x^k - x^l)} = \sum_{k=l} e^{2i\pi m(x^k - x^l)} + \sum_{k>l} e^{2i\pi m(x^k - x^l)} + \sum_{k<l} e^{2i\pi m(x^k - x^l)} = \\ &= n + 2 \sum_{k<l} e^{2i\pi m(x^k - x^l)} = n + 2 \sum_{k<l} \cos(2\pi m(x^k - x^l)). \end{aligned}$$

Como a função $x \mapsto x^l - x^k$ é estritamente crescente em $[a, b] \subset [1, +\infty[$, ela possui uma inversa que denotaremos por φ . Assim,

$$\int_a^b \cos(2\pi m(x^l - x^k)) dx = \int_{a^l - a^k}^{b^l - b^k} \varphi'(t) \cos(2\pi m t) dt.$$

Notemos que

$$\varphi(t)^l - \varphi(t)^k = t \implies l \cdot \varphi(t)^{l-1} \varphi'(t) - m \varphi(t)^{m-1} \varphi'(t) = 1 \implies \varphi'(t) [l \varphi(t)^{l-1} - m \varphi(t)^{m-1}] = 1.$$

Como $x \mapsto l x^{l-1} - k x^{k-1}$ é positiva e crescente em $[1, +\infty[$, φ' é decrescente, positiva e majorada por $\frac{1}{l-k}$. Assim,

$$l \varphi(t)^{l-1} - m \varphi(t)^{m-1} \geq l - k \iff \frac{1}{l \varphi(t)^{l-1} - m \varphi(t)^{m-1}} \geq \frac{1}{l - k} \iff \varphi'(t) \geq \frac{1}{l - k}.$$

Desta maneira,

$$\left| \int_a^b \cos(2\pi m(x^l - x^k)) dx \right| = \left| \int_{a^l - a^k}^{b^l - b^k} \varphi'(t) \cos(2\pi m t) dt \right| \leq \frac{1}{l - k} [b^l - b^k - (a^l - a^k)] \leq \frac{1}{l - k}.$$

Logo, se tomarmos $C = b - a$, a série que queremos estudar é majorada por:

$$\begin{aligned} & \sum_n \frac{1}{n} \int_a^b \left| \frac{1}{n} \sum_{k=0}^n e^{2i\pi m x^k} \right|^2 dx = \sum_n \frac{1}{n} \int_a^b \left| \frac{1}{n} (n + 2 \sum_{k<l} \cos(2\pi m(x^k - x^l))) \right| dx \leq \\ & \leq \sum_n \frac{1}{n} \int_a^b \left| 1 + \frac{2}{n} \sum_{k<l} \cos(2\pi m(x^k - x^l)) \right| dx \leq \sum_n \frac{1}{n} \int_a^b 1 + \left| \frac{2}{n} \sum_{k<l} \cos(2\pi m(x^k - x^l)) \right| dx = \\ & = \sum_n \frac{1}{n} \left(\int_a^b 1 dx + \int_a^b \left| \frac{2}{n} \sum_{k<l} \cos(2\pi m(x^k - x^l)) \right| dx \right) = \sum_{n=1}^{\infty} \frac{1}{n^3} (Cn + \sum_{1 \leq k < l \leq n} \frac{1}{l - k}) \leq \\ & \leq \sum_{n=1}^{\infty} \frac{1}{n^3} (Cn + \sum_{d=1}^n \frac{n - d}{d}) \leq C \sum_{n=1}^{\infty} \frac{1}{n^2} + \frac{\ln x}{x} < +\infty. \end{aligned}$$

Por consequência, como n é um inteiro fixo, verificamos o critério de Weyl para todo $x > 1$. Como a intersecção enumerável de conjuntos enumeráveis com medida total ainda é total, a sequência $(x^n)_n$ é equidistribuída módulo 1 para quase todo $x > 1$. $\square$

Exemplo 4.0.1. A sequência $((\sqrt{2} + 1)^n)_n$ não é equidistribuída módulo 1. De fato, pelo Binômio de Newton,

$$\begin{aligned} (\sqrt{2} + 1)^n + (-\sqrt{2} + 1)^n &= \sum_{k=0}^n \binom{n}{k} (\sqrt{2})^k \cdot 1^{n-k} + \sum_{k=0}^n \binom{n}{k} (-\sqrt{2})^k \cdot 1^{n-k} = \\ &= \sum_{k=0}^n \binom{n}{k} [(\sqrt{2})^k + (-\sqrt{2})^k] = \sum_{k=0}^n \binom{n}{k} (\sqrt{2})^k [1 + (-1)^k]. \end{aligned}$$

Porém $1 + (-1)^k = 2$, se k é par e $1 + (-1)^k = 0$, se k é ímpar. Assim, o somatório anterior pode ser escrito como

$$\sum_{0 \leq 2k \leq n} \binom{n}{2k} 2^{k+1},$$

que é um número inteiro. Assim, a classe de $(\sqrt{2} + 1)^n$ módulo 1 é a mesma que $(-1)^{n+1}(\sqrt{2} - 1)^n$, que tende para 0, pois $\sqrt{2} - 1 < 1$. Em particular, a sequência $(\sqrt{2} + 1)^n$ não pode ser equidistribuída módulo 1.

Definição 43. Um número $\beta > 1$ é chamado número de Pisot se ele é algébrico e todos os seus conjugados (no sentido de Galois) possuem módulo menor do que 1.

Teorema 44. Se β é um Número de Pisot, então (β^n) não é equidistribuída módulo 1.

Demonstração. Suponha que o polinômio minimal de β possui grau d e que seus conjugados sejam $\beta_2, \beta_3, \dots, \beta_{d-1}$. Então $k_n = \beta^n + \beta_2^n + \beta_3^n + \dots + \beta_{d-1}^n$ é inteiro, para todo n natural. Como $\beta_2, \beta_3, \dots, \beta_{d-1}$ possuem módulo menor do que 1, pela definição, segue que $\beta_2^n + \beta_3^n + \dots + \beta_{d-1}^n$ tende a zero, quando n tende a infinito. Assim, $\{\beta_n\}$ tende a zero, de onde segue que ela não pode ser equidistribuída. $\square$

Observação 6. O problema de encontrar um x tal que $(\{x^n\})_n$ é equidistribuída ainda se encontra em aberto.

Consideremos $P(x) = a_0 + a_1x + \dots + a_dx^d$ um polinômio de grau d , com $d \geq 1$. A parte fracionária de um número real x será $\{x\} = x - [x]$. Consideraremos a sequência $z_n = \{P(n)\}$, ou seja, $z_n \in \mathbb{R}/\mathbb{Z} \simeq S^1$.

Teorema 45. (Weyl) *Se algum dos coeficientes $a_1, a_2, \dots, a_d$ é irracional, então a sequência $(\{P(n)\})_{n \geq 0}$ é equidistribuída*

Provaremos alguns resultados antes. Considere o toro d dimensional

$$\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d = S^1 \times \dots \times S^1,$$

munido da medida de Lebesgue. Tomemos a transformação $f : \mathbb{T}^d \rightarrow \mathbb{T}^d$ dada por $f(\theta_1, \theta_2, \dots, \theta_d) = (\theta_1 + \alpha, \theta_2 + \theta_1, \dots, \theta_d + \theta_{d-1})$ onde α é um número irracional que escolheremos depois.

Proposição 46. *A transformação f definida anteriormente é ergódica com relação à medida de Lebesgue no toro $\mathbb{T}^d$, para qualquer escolha de α .*

Demonstração. Seja $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$ uma função em $L^2(\lambda)$. Vamos supor que $(\varphi \circ f)(x) = \varphi(x)$ λ -qtp. Concluiremos que φ é constante λ -qtp. Como $\varphi \in L^2(\lambda)$,

$$\varphi(\theta) = \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i n \cdot \theta},$$

onde $\theta = (\theta_1, \theta_2, \dots, \theta_d)$, $n = (n_1, n_2, \dots, n_d)$ e $n \cdot \theta$ representa o produto escalar entre n e θ , ou seja, $n \cdot \theta = n_1\theta_1 + n_2\theta_2 + \dots + n_d\theta_d$ e

$$\sum_{n \in \mathbb{Z}^d} |a_n|^2 = \int |\varphi(\theta)|^2 d\theta. \quad (4.2)$$

Mostraremos que os termos se anulam quase sempre. Notemos que

$$\begin{aligned} \varphi(f(\theta)) &= \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i n \cdot f(\theta)} = \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i (n_1, n_2, \dots, n_d) \cdot f(\theta_1, \theta_2, \dots, \theta_d)} = \\ &= \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i n_1 \alpha} e^{2\pi i ((n_1 + n_2)\theta_1 + (n_2 + n_3)\theta_2 + \dots + (n_d + n_{d-1})\theta_{d-1} + n_d \theta_d)} = \\ &= \sum_{n \in \mathbb{Z}^d} a_n e^{2\pi i n_1 \alpha} e^{2\pi i L(n) \cdot \theta}. \end{aligned}$$

onde $L(n) = (n_1 + n_2, n_2 + n_3, \dots, n_{d-1} + n_d)$. Comparemos $\varphi \circ f$ com φ . Se igualarmos os termos que multiplicam $e^{2\pi i L(n) \cdot \theta}$:

$$a_n e^{2\pi i n_1 \alpha} = a_{L(n)} \Rightarrow |a_n| = |a_{L(n)}|.$$

Se usarmos a equação (4.2), temos que existe somente um número finito de termos a_n com valor absoluto diferente de zero. Com efeito, $\varphi \in L^2(\lambda)$. Logo existe k tal que $L^k(n) = n$, de onde segue que

$$a_n e^{2i\pi m\alpha} = a_{L^k(n)} = a_n \Rightarrow a_n = 0$$

Por isso, quase todos os termos da série de Fourier se anulam, exceto, possivelmente, o termo constante. Com isso, f é ergódica. $\square$

Proposição 47. *Seja f a transformação definida anteriormente. Então a medida de Lebesgue é a sua única medida invariante. Assim, f é unicamente ergódica, a menos de uma constante.*

Demonstração. A prova será feita por indução em d que é tanto o grau do polinômio P que é igual à dimensão do toro $\mathbb{T}^d$.

- $d = 1$

Neste caso, $P(x) = a_0 + a_1x$. Vamos supor a_1 irracional. Consideraremos aqui a transformação $f : S^1 \rightarrow S^1$ dada por $f(\theta) = \theta + \alpha_1 \pmod{\mathbb{Z}}$ como uma rotação. É fácil verificar que a medida de Lebesgue λ é invariante por toda rotação R_θ . Resta provarmos que λ é a única medida invariante.

Afirmção: para qualquer função $\varphi : S^1 \rightarrow \mathbb{R}$ integrável, existe c_φ tal que

$$\varphi_n = \frac{1}{n} \sum_{j=0}^{n-1} (\varphi \circ R_\theta^j)$$

converge para c_φ para todo $x \in S^1$.

De fato, como R_θ é ergódica,

$$\lim_{n \rightarrow \infty} \varphi_n = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} (\varphi \circ R_\theta^j) = \int \varphi d\lambda,$$

λ -qtp. Por outro lado,

$$\lim_{n \rightarrow \infty} \varphi_n(x) = \lim_{n \rightarrow \infty} \varphi_n(R_\theta(x)).$$

Logo,

$$\lim_{n \rightarrow \infty} \varphi_n(y) = c_\varphi,$$

para todo $y \in O(x)$. Portanto, $\varphi_n \rightarrow c_\varphi$ em um conjunto denso de valores de $x \in \mathbb{T}^d$. Mostremos que (φ_n) é equicontínua. Seja d a distância induzida no toro $\mathbb{T}^d$ pela norma usual em $\mathbb{R}^d$: basta tomarmos o mínimo das distâncias entre seus representantes em $\mathbb{R}^d$. Observemos que a rotação R_θ preserva medida, isto é,

$$d(R_\theta(x), R_\theta(y)) = d(x, y).$$

Como φ é contínua para todo $\varepsilon > 0$, existe $\delta > 0$ tal que

$$d(x, y) = d(R_\theta^j(x), R_\theta^j(y)) < \delta \Rightarrow |\varphi(R_\theta^j(x)) - \varphi(R_\theta^j(y))| < \varepsilon,$$

para todo $j \geq 0$. Então

$$\begin{aligned} d(x, y) < \delta &\Rightarrow |\varphi_n(x) - \varphi_n(y)| = \left| \frac{1}{n} \sum_{j=0}^{n-1} (\varphi \circ R_\theta^j)(x) - \frac{1}{n} \sum_{j=0}^{n-1} (\varphi \circ R_\theta^j)(y) \right| = \\ &= \frac{1}{n} \left| \sum_{j=0}^{n-1} (\varphi \circ R_\theta^j)(x) - \sum_{j=0}^{n-1} (\varphi \circ R_\theta^j)(y) \right| = \frac{1}{n} \left| \sum_{j=0}^{n-1} (\varphi \circ R_\theta^j)(x) - (\varphi \circ R_\theta^j)(y) \right| \\ &\leq \frac{1}{n} \sum_{j=0}^{n-1} |(\varphi \circ R_\theta^j)(x) - (\varphi \circ R_\theta^j)(y)| < \varepsilon. \end{aligned}$$

Como ε não depende de n , podemos concluir que φ_n é equicontínua.

Seja ν outra medida invariante por R_θ tal que $\lambda(S^1) = \nu(S^1)$. Seja também C um boreliano de S^1 . Pelo teorema de Birkhoff (25),

$$\lambda(C) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_C(R_\theta^j(x)) = \nu(C).$$

Logo, λ é unicamente ergódica.

- Vamos supor que o resultado é válido para $d - 1$. Provaremos que ele vale para d . Escrevamos $\mathbb{T}^d = \mathbb{T}^{d-1} \times S^1$. Consideremos $f : \mathbb{T}^{d-1} \times S^1 \rightarrow \mathbb{T}^{d-1} \times S^1$ dada por $f(\theta_0, \eta) = (f_0(\theta_0), \eta + \theta_{d-1})$, onde $\theta_0 = (\theta_1, \theta_2, \dots, \theta_{d-1})$ e $f_0(\theta_0) = (\theta_1 + \alpha, \theta_2 + \theta_1, \dots, \theta_{d-1} + \theta_{d-2})$.

Representaremos a projeção por $\pi : \mathbb{T}^d = \mathbb{T}^{d-1} \times S^1 \rightarrow \mathbb{T}^{d-1}$ dada por $\pi(\theta) = \pi(\theta_0, \eta) = \theta_0$, ou seja, pegaremos as $d - 1$ primeiras coordenadas. Por hipótese de indução, a transformação $f_0 : \mathbb{T}^{d-1} \rightarrow \mathbb{T}^{d-1}$ é unicamente ergódica.

Afirmação: se μ é uma medida invariante por f , então a projeção $\pi_*\mu$ coincide com a medida de Lebesgue λ_0 em $\mathbb{T}^{d-1}$, onde $f_*\nu(E) := \nu(f^{-1}(E))$. No caso, $\pi_*\mu(E) := \mu(\pi^{-1}(E))$.

Por hipótese de indução, f_0 é unicamente ergódica e m_0 é a única medida invariante de f_0 . Provemos que $\pi_*\mu$ é uma probabilidade invariante de f_0 . Provemos então que

$$(\pi_*\mu)(f_0^{-1}(E)) = (\pi_*\mu)(E),$$

para isto, é suficiente mostrarmos que

$$\mu(\pi^{-1}(f_0^{-1}(E))) = \mu(\pi^{-1}(E)).$$

Por hipótese, μ é uma probabilidade invariante por f , isto é,

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

Porém, observe que $(\pi \circ f) = (f_0 \circ \pi)$, pois

$$\begin{aligned} (\pi \circ f)(\theta_0, \eta) &= \pi(f(\theta_0, \eta)) = \pi(f_0(\theta_0), \eta + \theta_{d-1}) = \\ &= \pi(\theta_1 + \alpha, \theta_2 + \theta_1, \dots, \theta_{d-1} + \theta_{d-2}, \eta + \theta_{d-1}) = \\ &= (\theta_1 + \alpha, \theta_2 + \theta_1, \dots, \theta_{d-1} + \theta_{d-2}) = f_0(\theta_0) = f_0(\pi(\theta)) = (f_0 \circ \pi)(\theta). \end{aligned}$$

Notemos que $(f \circ g)^{-1}(X) = (g^{-1} \circ f^{-1})(X)$. Com isso, como μ é invariante por f :

$$\mu(\pi^{-1}(f_0^{-1}(E))) = \mu((f_0 \circ \pi)^{-1}(E)) = \mu((\pi \circ f)^{-1}(E)) = \mu(f^{-1}(\pi^{-1}(E))) = \mu(\pi^{-1}(E)).$$

Logo, $\pi_*\mu$ é uma probabilidade invariante de f_0 e assim $\pi_*\mu$ coincide com λ_0 . Consideraremos μ uma medida invariante por f . Provaremos que $\mu = \lambda$. Considere G_μ o conjunto dos pontos $\theta \in \mathbb{T}^d$ tais que

$$\int \tilde{\varphi}(\theta) d\mu(\theta) = \int \varphi(\theta) d\mu(\theta),$$

para toda $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$. Pelo teorema de Birkhoff 25, G_μ tem medida total. Seja $G_0(\mu)$ o conjunto dos $\theta_0 \in \mathbb{T}^{d-1}$ tais que $G_0(\mu) = \pi(G_\mu)$. Observe que, se $\theta_0 \in \pi(G_\mu)$

se, e somente se, as primeiras $d - 1$ coordenadas de $\tilde{\theta}$ é θ_0 , para algum $\tilde{\theta} \in G_\mu$.
Notemos que

$$G_\mu \subset \pi^{-1}(\pi(G_\mu)) = \pi^{-1}(G_0(\mu)).$$

Como $\pi(G_\mu) = 1$, segue que $\mu(\pi^{-1}(G_0(\mu))) = 1$. Com isso, pelo que já vimos

$$1 = \mu(\pi^{-1}(G_0(\mu))) = (\pi_*\mu)(G_0(\mu)) = \lambda_0(G_0(\mu)).$$

Repare que, na igualdade anterior, μ pode ser qualquer medida invariante de f . Em particular, ela vale para $\mu = \lambda$.

$$\lambda_0(G_0(\lambda)) = \lambda(\pi^{-1}(G_0(\lambda))) = 1.$$

Mostraremos que $G_0(\mu) \cap G_0(\lambda)$ é diferente de vazio. Para isto, mostraremos algo mais forte: $\lambda(G_0(\mu) \cap G_0(\lambda)) = 1$. Com efeito,

$$(G_0(\mu) \cap G_0(\lambda))^c = G_0(\mu)^c \cup G_0(\lambda)^c.$$

Assim,

$$\lambda((G_0(\mu) \cap G_0(\lambda))^c) = \lambda(G_0(\mu)^c \cup G_0(\lambda)^c) \leq \lambda(G_0(\mu)^c) + \lambda(G_0(\lambda)^c) = 0.$$

De onde segue que

$$\lambda((G_0(\mu) \cap G_0(\lambda))^c) = 0.$$

Assim, $\lambda(G_0(\mu) \cap G_0(\lambda)) = 1$. Considere $\theta_0 \in G_0(\mu) \cap G_0(\lambda)$. Mostraremos que se $\theta_0 \in G_0(\lambda)$, então $\{\theta_0\} \times S^1 \subset G(\lambda)$. De fato, a medida λ é invariante por toda a transformação da forma $R_\beta : \mathbb{T}^{d-1} \times S^1 \rightarrow \mathbb{T}^{d-1} \times S^1$ dada por $R(\zeta, \eta) = (\zeta, \eta + \beta)$. Lembremos que $G_0(\lambda) = \pi(G_\lambda)$. Então

$$\theta_0 \in G_0(\lambda) \Rightarrow \theta_0 \in \pi(G_\lambda) \Rightarrow \theta_0 = \pi(\theta); \theta \in G_\lambda \Rightarrow \exists \eta \in S^1; (\theta_0, \eta) \in G_n \Rightarrow$$

$$\Rightarrow \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(\theta, \eta)) = \int \varphi d\lambda, \forall \varphi : \mathbb{T}^d \rightarrow \mathbb{R} \text{ contínua.}$$

Queremos provar que todo ponto de $\{\theta_0\} \times S^1$ pertence a G_λ . Observe que qualquer ponto de $\{\theta_0\} \times S^1$ pode ser escrito como $(\theta_0, \eta + \beta) = R_\beta(\theta_0, \eta)$ para algum $\beta \in S$.

Mostraremos que $(\theta_0, \eta + \beta) \in G_\lambda$, isto é, para qualquer função contínua $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$, vale que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(\theta, \eta + \beta)) = \int \varphi d\lambda.$$

Para $\tau_0 = (\tau_1, \tau_2, \dots, \tau_{d-1})$,

$$\begin{aligned} f(R_\beta(\tau_0, \eta)) &= f(\tau_0, \eta + \beta) = f(\tau_1, \tau_2, \dots, \tau_{d-1}, \eta + \beta) = (\tau_1 + \alpha, \tau_2 + \tau_1, \dots, \tau_{d-1} + \zeta + \beta) = \\ &= R_\beta(f(\tau_0, \zeta)). \end{aligned}$$

Logo $(f \circ R_\beta) = (R_\beta \circ f)$. Por indução, podemos provar que $(f^n \circ R_\beta) = (R_\beta \circ f^n)$.

Assim, como λ é invariante por R_β :

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(\theta_0, \eta + \beta)) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi((f^k \circ R_\beta)(\theta_0, \eta)) = \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi((R_\beta \circ f^k)(\theta_0, \eta)) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} (\varphi \circ R_\beta)(f^k(\theta_0, \eta)) = \\ &= \int (\varphi \circ R_\beta) d\lambda = \int \varphi d\lambda. \end{aligned}$$

Desta maneira, $(\theta_0, \eta + \beta) \in G_m$ para todo $\beta \in S^1$. Com isso, $\{\theta + 0\} \times S^1 \subset G_\lambda$.

Pela definição de G_μ , existe θ tal que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(\theta)) = \int \varphi d\mu$$

e

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(f^k(\theta)) = \int \varphi d\lambda.$$

Com isso,

$$\int \varphi d\mu = \int \varphi d\lambda.$$

Logo, $\mu = \lambda$.

□

Como f é unicamente ergódica, então

$$\lim_{N \rightarrow \infty} \frac{1}{N} \# \{0 \leq i \leq N-1 : f^i(\theta) \in O\} = \mu(O).$$

Corolário 48. *A órbita de todo ponto $\theta \in \mathbb{T}^d$ é equidistribuída no toro.*

Antes de provarmos o teorema de Weyl, ainda mostraremos dois lemas.

Lema 49. *Se considerarmos $p_d(x) := P(x)$ e $p_{j-1}(x) := p_j(x+1) - p_j(x)$, para $j = 2, 3, \dots, d$, então*

(i) O polinômio $p_j(x)$ tem grau j , para todo $1 \leq j \leq d$.

(ii) $p_1(x) = \alpha x + \beta$, onde $\alpha = a_d \cdot d!$ é irracional.

Demonstração. (i) Repare que o resultado é válido para $j = d$. Com efeito, $p_d(x) = P(x)$. Vamos supor que o resultado é verdadeiro para algum j e mostrar que vale para $j - 1$. Como o grau de $p_j(x)$ é j , segue que

$$p_j(x) = a_{j,j}x^j + a_{j,j-1}x^{j-1} + \dots + a_{1,j}x + a_{0,j}.$$

Então

$$p_{j-1}(x) = p_j(x+1) - p_j(x) = a_{j,j}x^j + (ja_{j-1,j} + a_{j-1,j})x^{j-1} + r(x) - a_{j,j}x^j - a_{j-1,j}x^{j-1} + s(x),$$

onde $r(x)$ e $s(x)$ possuem grau $j - 2$. Com isso,

$$p_{j-1}(x) = ja_{j-1,j}x^{j-1} + t(x),$$

com $t(x)$ um polinômio de grau menor ou igual a $j - 2$. Daí, $p_{j-1}(x)$ possui grau $j - 1$. (ii) Basta notarmos que o coeficiente líder de $p_d(x)$ é da_d , enquanto o de $p_{d-1}(x)$ é $d(d-1)a_d$ e assim por diante. $\square$

Lema 50. *Para todo n não negativo, vale*

$$f^n(p_1(0), p_2(0), \dots, p_d(0)) = (p_1(n), p_2(n), \dots, p_d(n)).$$

Demonstração. Vamos usar indução em n . Para $n = 0$, o resultado é claramente verdadeiro. Vamos supor que a igualdade é válida para $n - 1$, isto é

$$f^{n-1}(p_1(0), p_2(0), \dots, p_d(0)) = (p_1(n-1), p_2(n-1), \dots, p_d(n-1)).$$

Então

$$f^n(p_1(0), p_2(0), \dots, p_d(0)) = f(f^{n-1}(p_1(0), p_2(0), \dots, p_d(0))) =$$

$$\begin{aligned}
&= (p_1(n-1) + \alpha, p_2(n-1) + p_1(n-1), \dots, p_d(n-1) + p_{d-1}(n-1)) = \\
&= (p_1(n), p_2(n), \dots, p_d(n)).
\end{aligned}$$

□

Demonstração. (teorema de Weyl) Seja $\varphi : S^1 \rightarrow \mathbb{R}$ uma função contínua qualquer. Considere $\psi : \mathbb{T}^d \rightarrow \mathbb{R}$ definida por

$$\psi(\theta_1, \theta_2, \dots, \theta_d) = \varphi(\theta_d).$$

Fixemos $\theta = (p_1(0), p_2(0), \dots, p_d(0))$. Então, como já vimos que o resultado vale em $\mathbb{T}^d$:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(z_n) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \psi(f^n(\theta)) = \int \psi \, dm = \int \varphi \, dx.$$

□

Apêndice A

Prova do Teorema da Recorrência Múltipla de Furstenberg em casos particulares

Demonstraremos apenas alguns casos particulares e, com isso, daremos a ideia da demonstração geral.

Exemplo A.0.1. *Começaremos mostrando para um Sistema de Bernoulli, que definiremos a seguir. Aqui, cada um dos espaços $(X_i, \mathcal{A}_i, \mu_i)$ será $(X, \mathcal{A}, \nu)$. Assim, $\nu(X) = 1$ e*

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

Sejam $A = \{1, 2, 3, \dots, d\}$ um conjunto finito, $\mathcal{A} = P(X) = \{Y | Y \subset X\}$, $X = A^{\mathbb{N}}$ e T o shift neste espaço. Vamos definir uma probabilidade em X . Para isto, basta tomarmos uma probabilidade em A e depois definirmos uma medida produto para termos uma probabilidade em X . Sejam também $p_1, p_2, \dots, p_d$ números não negativos cuja soma é 1. Observe que, neste caso, todo cilindro pode ser escrito como união de cilindros elementares. Por isso, no caso em que estamos, podemos considerar apenas esses.

Se $C = [m; a_m, a_{m+1}, \dots, a_n]$ (repare que, neste caso, $1 \leq a_m, a_{m+1}, \dots, a_n \leq d$), então

definimos a medida de Bernoulli como sendo

$$\mu(C) = p_{a_m} \cdot \dots \cdot p_{a_n}.$$

A medida de Bernoulli é invariante pela aplicação shift. Vamos mostrar que o teorema de Furstenberg vale para os cilindros.

Com efeito, se $C_0, C_1, \dots, C_k$ são cilindros, para n suficientemente grande, as coordenadas que definem $T^{-nl}(C_l)$ são todas disjuntas.

Desta forma,

$$\begin{aligned} \mu(C_0 \cap T^{-n}(C_1) \cap T^{-2n}(C_2) \cap \dots \cap T^{-kn}(C_k)) &= \\ \mu(C_0)\mu(T^{-n}(C_1))\mu(T^{-2n}(C_2)) \dots \mu(T^{-kn}(C_k)) &> 0 \Rightarrow \\ \mu(C_0)\mu(C_1)\mu(C_2) \dots \mu(C_k) &> 0. \end{aligned}$$

Exemplo A.0.2. No caso em que o sistema é periódico, o resultado é trivial.

Exemplo A.0.3. Provemos que o teorema de Furstenberg vale para sistemas quasi-periódicos, isto é, estamos considerando $X = S^1 \simeq \mathbb{R}/\mathbb{Z}$, μ a medida de Lebesgue no círculo e $T(x) = x + \alpha(\text{mod } 1)$, para algum α . Seja A um conjunto mensurável com medida positiva. Observe que, a função

$$\int \chi_A(x + y) d\mu(x)$$

é contínua em y .

Desta forma, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que se $|y| < \delta$, então

$$\mu(A \cap (A - y)) < \mu(A) - \varepsilon.$$

Desta forma, podemos provar que

$$\mu(A \cap (A - y) \cap (A - 2y) \cap \dots \cap (A - ky)) > \mu(A) - (k + 1)\varepsilon.$$

Vamos escolher $\varepsilon < \frac{\mu(A)}{k+1}$. Para ele, existe um certo δ . Assim, seja

$$D_\delta = \{n \geq 1 : n\alpha \in (-\delta, \delta)(\text{mod } 1)\}.$$

Logo, se $n \in D_\delta$, então para a aplicação T definida anteriormente

$$\mu(A \cap T^{-n}(A) \cap T^{-2n}(A) \cap \dots \cap T^{-kn}(A)) > 0.$$

O primeiro exemplo é um caso particular de sistemas fracamente misturadores, isto é, sistemas tais que para quaisquer A e B mensuráveis:

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

Podemos adaptar a ideia do exemplo e provar o Teorema de Furstenberg para sistemas fracamente misturadores.

Já o segundo exemplo é um caso particular de sistemas compactos, isto é, tais que, para qualquer função f , o conjunto $\overline{\{f, T(f), T^2(f), \dots, T^n(f), \dots\}}$ é compacto.

Podemos provar que o teorema de Furstenberg vale também para este caso. A vantagem de fazermos isto, é que existe um resultado chamado teorema da estrutura. Podemos fazer uma decomposição do sistema dinâmico em partes fracamente misturadoras e compactas. Não demonstraremos o resultado aqui, mas a partir disto, podemos demonstrar o Teorema de Furstenberg para o caso geral.

Referências Bibliográficas

- [1] Moreira, Carlos Gustavo Tamm de Araújo; Arbieto, Alexander ; Matheus, Carlos . Aspectos Ergódicos da Teoria dos Números. Rio de Janeiro: [s.n.], 2007. 124 p.
- [2] Oliveira, Krerley; Viana, Marcelo. Fundamentos da Teoria Ergódica. Rio de Janeiro: [s.n.], 2014. 556 p.
- [3] Dutra, Aline Gobbi. Sistemas Dinâmicos e Substituições. 2007. 59 p. Dissertação (Mestrado em Matemática)- Instituto de Biologia, Letras e Ciências Exatas da Universidade Estadual Paulista, São José do Rio Preto, 2007.
- [4] Walters, Peter. An Introduction to Ergodic Theory. New York: Springer-Verlag New York, 1982. 250 p.
- [5] Van Der Waerden, B. L. "How the Proof of Baudet's Conjecture Was Found." Studies in Pure Mathematics (Presented to Richard Rado). London: Academic Press, pp. 251-260, 1971.
- [6] Szemerédi, Endre (1975). "On sets of integers containing no k elements in arithmetic progression"(PDF). Acta Arithmetica. 27: 199–245. MR 0369312. Zbl 0303.10056.
- [7] B. Green and T. Tao: The primes contain arbitrarily long arithmetic progressions. Ann. Math. 167, 481-547 (2008)