Os sistemas mais caóticos conhecidos pela humanidade

Prof. Dr. Gabriel Ponce
UNICAMP
www.ime.unicamp.br/∼gaponce
gaponce@ime.unicamp.br

Prof. Dr. Régis Varão
UNICAMP
www.ime.unicamp.br/∼regisvarao
regisvarao@ime.unicamp.br

October 19, 2015

Contents

Prefácio

Escrever estas notas foi para nós um grande desafio. Nosso objetivo é apresentar um tema interessante e que chegue o mais próximo da pesquisa em Matemática, com o adicional de ser acessível ao máximo de alunos possíveis. Essa não é uma tarefa fácil, com certeza realizá-la em um minicurso parece ainda mais difícil. Mas a verdade é que estamos muito felizes com o trabalho que fizemos e conseguimos cumprir nossa tarefa dentro dos limites possíveis. Estas notas possuem mais matemática do que seremos capaz de passar no minicurso, o que é proposital. Queremos incentivar os jovens que se iniciam na vida acadêmica a conhecerem a teoria que apresentamos aqui (Teoria de Ornstein). Esperamos portanto que o aluno interessado retorne a essas notas mais a frente, que se junte com outros alunos para estudá-las. E com prazer seria bom reencontrá-los para conversar sobre o assunto novamente.

Aos jovens que iniciam sua vida acadêmica podemos garantir-lhes que o tema dessas notas é um tema clássico (leia-se também importante) e muito ativo em Sistemas Dinâmicos. Aos estudantes interessados em ir mais a fundo no tema podem ficar tranquilos, estão se deparando com matemática de qualidade. Não podemos garantir que essas notas estão ao nível da qualidade do tema, mas nos esforçamos em deixá-las bem escrita. Não conhecemos textos acessíveis a alunos de fim de graduação e início de mestrado que possam iniciar seus estudos nesta teoria de classificação (Kolmogorov-Bernoulli). Esperamos servir de base para iniciar o mais cedo possível jovens nesta área.

Estas notas tem por objetivo provar o Teorema 3.3.1. O Teorema usa alguns termos que vamos introduzir ao longo do texto, mas não há problema nenhum, e encorajamos que assim o faça, de ler o enunciado assim:

Teorema Principal: Seja $A : \mathbb{T}^2 \to \mathbb{T}^2$ onde A é uma matriz sem auto valor com norma 1, então A é isomorfo a um shift de Bernoulli.

Isto quer dizer que olhar esse sistema $A : \mathbb{T}^2 \to \mathbb{T}^2$ é o mesmo que olhar para um sistema muito "mais simples" (veja Definição 1.1.2). Há um jogo que temos que envolver o conceito de medida que você vai conhecer ao longo do texto.

No capítulo §1 iniciamos nossos estudos em Teoria Ergódica. Neste capítulo apresentado os shifts, estudamos um pouco da sua dinâmica, vemos o conceito de medida, estudamos propriedades. No próximo capítulo, §2, apresentamos o conceito base mais importante em Teoria Ergódica que chama-se *ergodicidade*. Por fim apresentamos o capítulo §3 o mais difícil dessas notas. Nada vem de graça e a dificuldade do capítulo reside justamente em oferecer ideias importantes e úteis de se trabalhar em Teoria Ergódica. Deixamos o capítulo o mais auto contido possível, a menos de alguns resultados que são muito avançados (e.g. Teoremas 3.2.12, 3.2.13) podemos dizer que estas notas são auto contidas.

Gabriel Ponce & Régis Varão
Outubro 2015, Campinas.

Chapter 1

Teoria Ergódica

Nesta seção apresentamos o modelo base do nosso estudo, os shifts (Definição 1.1.2). Esses sistemas possuem uma descrição muito simples e por outro lado possuem grande grau de complexidade (e.g. Proposição 1.1.2). Mas o principal sistema que estamos interessados em estudar é o Shift de Bernoulli (vide §1.1.1) que é o mapa shift com uma medida muito natural associada a ele.

1.1 Shifts

Seja $\Sigma_k = \{1, \ldots, k\}^{\mathbb{Z}}$. E para duas sequências $\alpha, \beta \subset \Sigma_k$ definamos o número

$$n(\alpha, \beta) := max\{l \in \mathbb{N} : \alpha(i) = \beta(i), |i| \leq l\}.$$

Exemplo 1.1.1. Considere sequências em Σ_3 definidas por

$$\alpha := \{\ldots 0, 0, 2, 2, 0, 0.2, 2, 0, 0, 2, 2 \ldots\} \subset \Sigma_3$$

e

$$\beta := \{\ldots 0, 0, 2, 2, 0, 0.2, 2, 0, 0, 2, 2 \ldots\} \subset \Sigma_3.$$

Note que o ponto "." serve para indicar a coordenada zero-ésima da sequência. Isto é $\alpha = (\alpha_i)_{i \in \mathbb{Z}}$ e $\alpha_0 = 2$, $\alpha_1 = 2$. Note que $\beta_0 = 2$, $\beta_{-1} = 0$ etc. E usamos $:=$ ao invés de $=$ porque que estamos definindo os objetos α e β.

Definição 1.1.2. O mapa shift é definido como

$$\begin{array}{rcl} \sigma : \Sigma_k & \to & \Sigma_k \\ \alpha & \mapsto & \sigma(\alpha) \\ \{\ldots, \alpha_{-1}.\alpha_0, \alpha_1, \alpha_2, \ldots\} & \mapsto & \{\ldots, \alpha_{-1}, \alpha_0.\alpha_1, \alpha_2, \ldots\} \end{array}$$

com $\sigma(i) = \alpha(i + 1)$.

Vejamos um exemplo simples

Exemplo 1.1.3. Considere $\alpha = \{\ldots 0, 0, 2, 2, 0, 0.2, 2, 0, 0, 2, 2\ldots\}$.

$$\sigma(\alpha) = \{\ldots 0, 0, 2, 2, 0, 0, 2.2, 0, 0, 2, 2\ldots\};$$

$$\sigma \circ \sigma(\alpha) = \{\ldots 0, 0, 2, 2, 0, 0, 2, 2.0, 0, 2, 2\ldots\};$$

$$\sigma \circ \sigma \circ \sigma(\alpha) = \{\ldots 0, 0, 2, 2, 0, 0, 2, 2, 0.0, 2, 2\ldots\}.$$

Denotamos também por $\sigma^n(\alpha) := \sigma \circ \ldots \circ \sigma(\alpha)$, isto é a composição de σ n-vezes. De fato usaremos esta notação mais a frente também para outros mapas, ou seja se tivermos $f : X \to X$.

Definição 1.1.4. Definimos a função distância em Σ_k por:

$$\begin{aligned} d : \Sigma_k \times \Sigma_k &\to \mathbb{R} \\ d(\alpha, \beta) &= \begin{cases} 0 \text{ se } \alpha = \beta \\ \left(\frac{1}{2}\right)^n, \ n = max\{a \in \mathbb{N} : \alpha(i) = \beta(i), |i| \leq a\} \end{cases} \end{aligned}$$

A definição de função distância deve ficar clara para você. A função distância está medindo o quão igual são duas sequências em torno da entrada zero. Vamos ver uns exemplos.

Exemplo 1.1.5. Considere a sequência $\alpha = \{\ldots, 1, 2, 3.1, 2, 3, 1, 2, 3, \ldots\} \subset \Sigma_5$

- $d(\alpha, \beta) = (1/2)$ se $\beta = \{\ldots, 1, 2, 3.1, 5, 3, 1, 2, 3, \ldots\}$;
- $d(\alpha, \beta) = (1/2)^2$ se $\beta = \{\ldots, 1, 4, 3.1, 2, 3, 5, 2, 3, \ldots\}$;
- $d(\alpha, \beta) = (1/2)^3$ se $\beta = \{\ldots, 5, 2, 3.1, 2, 3, 1, 2, 4, \ldots\}$.

Proposição 1.1.1. *Sendo d definido como acima, então valem*

a) d é uma distância em Σ_k;

b) (Σ_k, d) é um espaço métrico compacto;

c) o mapa shift $\sigma : \Sigma_k \to \Sigma_k$ é contínuo

Proof. a) Da definição de d segue diretamente que

- $d \geq 0$;
- $d(\alpha, \beta) = 0 \Leftrightarrow \alpha = \beta$
- $d(\alpha, \beta) = d(\beta, \alpha)$

para ver que d é uma distância falta verificar que $d(\alpha,\beta) \leq d(\alpha,\gamma)+d(\gamma,\beta)$. Suponha sem perda de generalidade que $d(\alpha,\gamma) \leq d(\gamma,\beta)\}$. Se $d(\alpha,\beta) = 0$ não há nada a se fazer, suponha então que $d(\alpha,\beta) = (1/2)^m$, para algum $m \in \mathbb{N}$. Suponha que $d(\gamma,\beta) < d(\alpha,\beta)$. Assim tanto tanto α quanto β devem coincidir com γ nas entras i com $|i| \leq m+1$ e portanto α e β deveriam coincidir também nessas entradas o que dá um absurdo na definição de d.

b) Para checar a compacidade basta ver que toda sequência $\{\omega_k\}_k$ possui subsequência convergente. Definamos uma subsequência $\{k(1,i)\}$ tal que $\{\omega_{k(1,i)}\}_{i\in\mathbb{Z}}$ possui todos os elementos da entrada 0 iguais. Agora definamos indutivamente a sequência $\{k_i^{n+1}\} \subset \{k_i^n\}$ tal que todos os elementos de $\{\omega_k(n+1,i)\}_k$ coincidam nas entradas $l \in \mathbb{Z}$ com $|l| \leq k+1$.

Definamos assim a seguinte subsequência

$$\{\omega_{k(n,n)}\}_{n\in\mathbb{Z}}$$

Note que $\omega_{k(n,n)}$ coincide com $\omega_{k(n+1,n+1)}$ nas entradas l tais que $|l| \leq n$. É fácil ver que esta sequência é convergente.

c) Exercício. $\square$

Definição 1.1.6. Dizemos que um subconjunto $C \subset \Sigma_k$ é um cilindro se existem $a_0, \ldots, a_l \in \{1, \ldots, k\}$ e $i_0 \in \mathbb{Z}$ tais que:

$$C = \{\{\alpha_i\}_{i\in\mathbb{Z}} \,|\alpha_{i_0} = a_0, \alpha_{i_1} = a_1, \ldots \alpha_{i_l} = a_l\ \}$$

Uma **base de abertos** ou simplesmente uma base num espaço topológico X é uma coleção $\mathcal{B}$ de subconjuntos abertos de X que chamaremos de **abertos básicos** ou vizinhanças fundamentais com a seguinte propriedade: Todo subconjunto aberto $A \subset X$ se exprime como reunião $A = \cup_\lambda B_\lambda$ de abertos $B_\lambda \in \mathcal{B}$. Assumimos os seguintes fatos topológicos:

Fato Topológico 1: Seja X um espaço topológico. Uma coleção $\mathcal{B}$ de abertos de X constitui uma base em X se, e somente se, para todo aberto $A \subset X$ e cada $x \in A$ existe $B_x \in \mathcal{B}$ tal que $x \in B_x \subset A$.

Fato Topológico 2: Seja $\mathcal{B}$ uma coleção de subconjuntos de um conjunto X. Para que $\mathcal{B}$ seja base de uma topologia em X é necessário e suficiente que se cumpram as condições abaixo.

1. Para todo $x \in X$, existe $B \in \mathcal{B}$ tal que $x \in \mathcal{B}$;

2. Se $x \in B_1 \cap B_2$ onde $B_1, B_2 \in \mathcal{B}$ então exite $B \in \mathcal{B}$ tal que $x \in B \subset B_1 \cap B_2$.

Fato Topológico 3: E no nosso caso não é difícil de provar que os cilindros formam uma base para para a topologia gerada pela distância d definida acima.

Definição 1.1.7. Dizemos que α é um ponot periódico para σ se existe n tal que $\sigma^n(\alpha) = \alpha$. Ou em geral, dizemos que p é um ponto periódico para um mapa $f : X \to X$ se $f^n(p) = p$.

Definição 1.1.8. Dizemos que um ponto $\alpha \in \Sigma_k$ é transitivo se dado qualquer cilindro $C \subset \Sigma_k$ existe um n_0 tal que $\sigma^{n_0}(\alpha) \in C$.

Uma outra maneira de dizer que $\alpha \in \Sigma_k$ é transitivo é o mesmo que dizer que: dados $\epsilon > 0$ e $\beta \in \Sigma_k$ existe um n_0 tal que $d(\sigma^{n_0}(\alpha), \beta) < \epsilon$ (Exercício). Ou ainda, poderíamos dizer que a órbita do ponto α é densa em Σ_k.

Definição 1.1.9. Dizemos que um conjunto $S \subset \Sigma_k$ é denso em Σ_k se: dados $\epsilon > 0$ e $\beta \in \Sigma_k$ existe um n_0 tal que $d(\sigma^{n_0}(\alpha), \beta) < \epsilon$.

Proposição 1.1.2. *O shift $\sigma : \Sigma_k \to \Sigma_k$ satisfaz*

- *Os pontos periódicos são densos;*
- *Possui órbita transitiva.*

Proof.

- Para provar que os pontos períodicos são densos precisamos provar que: dado $\beta \in \Sigma_k$ e $\epsilon > 0$ existe um ponto periódico $\alpha \in \Sigma_k$ tal que $d(\alpha, \beta) < \epsilon$.

  Logo, dado $\epsilon > 0$ precisamo encontrar um periódico que coincida muitas entradas com β para que a distância entre os dois seja menos que ϵ. Existe um n_0 para o qual $(1/2)^{n_0} < \epsilon$. Assim, basta definirmos α como $\alpha_i = \beta_i$ para $|i| \leq n_0$ e depois repetimos periodicamente.

- Exercício (Dica: Enumere todos os pontos periódicos e vá concatenando cada um.)

□

1.1.1 Teoria Ergódica

Agora queremos poder medir o "tamanho" de conjuntos. O objeto matemático que usaremos para medir o tamanho de um conjunto recebe o nome de medida, que veremos mais a frente. Por exemplo o "tamanho" no intervalo $[a, b]$ é $b - a$. Ou melhor o comprimento do intervalo $[a, b]$ é $b - a$. Vamos denotar $l(I)$ o comprimento do intervalo I. Note que naturalmente faz sentido falar do comprimento de $[a, b] \cup [c, d]$, ou seja $l([a, b] \cup [c, d]) = b - a + d - c$. Assim se tivermos um conjunto formado por intervalos disjunto naturalmente podemos considerar $l(\cup_i I_i) = \sum_i l(I_i)$. Ocorre que podemos definir uma função l definida em certos conjuntos de $\mathbb{R}$ desde que faça sentido olhá-lo como uniões de intervalos. Podemos bem definir qual o tipo de conjuntos podemos usar o mapa l de forma matematicamente rigorosa. Para isso precisamos entrar no mundo das teoria da medida. O conjunto que podemos bem definir a função l é chamado de σ álgebra e l é uma medida. Analogamente ao invés de l poderíamos considerar o volume em $\mathbb{R}^2$. Então comprimento em $\mathbb{R}$ e volume em $\mathbb{R}^2$ são exemplos de medidas. Então o que faremos a seguir é generalizar/formalizar as propriedades do comprimento e do volume.

Definição 1.1.10. Dizemos que a aplicação

$$\mu : \mathcal{B} \to \mathbb{R}$$

é uma medida na σ-álgebra $\mathcal{B}$ se

- $\mu(\emptyset) = 0$;
- $\mu(E) \geq 0, \forall E \in \mathcal{B}$;
- $\mu(\cup_{i\in\mathbb{N}} E_i) = \cup_{i\in\mathbb{N}} \mu(E_i)$ onde E_i são disjuntos e $E_i \in \mathcal{B}$.

Na definição faltou dizer o que é uma σ-álgebra. Note, a seguir, que é uma definição que faz bastante principalmente se tivermos em mente os conjuntos que podemos calcular para o volume em $\mathbb{R}^2$.

Definição 1.1.11. Uma família $\mathcal{B}$ de conjuntos de X é uma σ-álgebra se satisfaz:

- $\emptyset, X \subset \mathcal{B}$;
- Se $A \in \mathcal{B}$, então $A^c \in \mathcal{B}$;
- Se $\{A_n\}_{n\in\mathbb{N}}$ com A_n em $\mathcal{B}$ então $\cup_{n=1}^{\infty} A_n \in \mathcal{B}$.

Exemplo 1.1.12. Definimos a medida δ_0 uma medida nos reais tal que a medida de $\delta_0(A) = 1$ se $0 \in A$ e $\delta_0(A) = 0$ se $0 \notin A$.

Dizemos que a tripla $(X, \mathcal{B}, \mu)$ é um espaço de medida se X é um conjunto qualquer, $\mathcal{B}$ uma σ-algebra de X e μ uma medida. Quando $\mu(X) = 1$ chamamos a tripla $(X, \mathcal{B}, \mu)$ de espaço de probabilidade. Trabalharemos com funções $T : X \to X$ que preservam a medida μ:

Definição 1.1.13. Dizemos que $f : X \to X$ preserva a medida μ se dado $A \in \mathcal{B}$ então $\mu(f^{-1}(A)) = \mu(A)$.

Estamos enunciando num contexto mais geral porque as definições são mais gerais, caso vocês esteja vendo esses conceitos pela primeira vez pode supor sempre que $X = \mathbb{T}^2$ e não haverá nenhum prejuízo na compreensão das notas. Afinal de contas nosso Teorema principal que queremos provar é em $\mathbb{T}^2$.

Apresentamos agora um teorema muito conhecido e útil e que serve também para praticarmos um pouco as definições que apresentamos.

Teorema 1.1.14 (Recorrência de Poincaré)**.** *Seja $T : X \to X$ uma transformação mensurável que preserva uma probabilidade μ. Dado $E \in \mathcal{B}$ com $\mu(E) > 0$, então existe $F \subset E$ com $\mu(F) = \mu(E)$, tal que: se $x \in F$, existem inteiros n_i, $i \in \mathbb{N}$, satisfazendo*

$$0 < n_1 < n_2 < n_3 < \dots \; ; \; f^{n_i}(x) \in E$$

Proof. Para $N \geq 0$ definamos $E_N = \bigcup_{n=N}^{\infty} T^{-n}E$, onde $T^0E = E$. Considere agora o conjunto

$$F = E \cap \left(\bigcap_{n=0}^{\infty} E_n \right)$$

note que F é exatamente o conjuntos dos ponto de E que retornam infinitas vezes para E. Queremos provar que $\mu(F) = \mu(E)$.

Temos $T^{-1}(E_N) = E_{N+1}$, como T preserva μ, $\mu(E_N) = \mu(E_{N+1})$. Isso implica $\mu(E_0) = \mu(E_N)$, para todo N. Observe que $E_0 \supset E_1 \supset E_2 \ldots$. Assim $\mu(E_0) = \mu(\bigcap_{N=0}^{\infty} E_N)$ e por conseguinte

$$\mu(F) = \mu(E \cap (\cap_{n=0}^{\infty} E_n)) = \mu(E \cap E_0) = \mu(E)$$

A segunda igualdade vem de $\cap_{n=0}^{\infty} E_n = E_0$ quase certamente; a última, decorre de $E \subset E_0$. Como queríamos provar. $\square$

Shifts de Bernoulli

Vimos na Definição 1.1.2 o mapa shift. Queremos considerar certas medidas naturais que são invariantes pelo shift. Considere o mapa shift $\sigma : \Sigma_k \to \Sigma_k$ e $p_i \in [0, \infty)$ com $i \in \{1, \ldots, k\}$ e $\sum_{i=1}^{k} p_i = 1$. Vamos definir uma medida μ na σ-álgebra $\mathcal{C}$ que é a σ-álgebra gerada pelos cilindros (Definição 1.1.6) impondo que

$$\begin{aligned}
\mu(\{\{\alpha_i\}_{i\in\mathbb{Z}} \quad &| \quad \alpha_{i_0} = 1\}) = p_1 \\
\mu(\{\{\alpha_i\}_{i\in\mathbb{Z}} \quad &| \quad \alpha_{i_0} = 2\}) = p_2 \\
&\ldots \\
\mu(\{\{\alpha_i\}_{i\in\mathbb{Z}} \quad &| \quad \alpha_{i_0} = k\}) = p_k
\end{aligned}$$

Ocorre que esta relação define de maneira única uma medida na σ-álgebra gerada pelos cilindros. Note que usando a propriedade da medida e o fato de μ ser invariante por σ que

$$\mu\{\{\alpha_i\}_{i\in\mathbb{Z}} \,|\alpha_{i_0} = a_0, \alpha_{i_1} = a_1, \ldots \alpha_{i_l} = a_l \,\} = p_{a_0} p_{a_1} \ldots p_{a_l}$$

1.2 Medidas Invariantes

A menos que dito o contrário, M será um espaço métrico compacto.

Exemplo 1.2.1. Queremos encontrar para os próximos exemplos medidas (borelianas) invariantes.

a)

$$f : [0,1] \to [0,1]$$
$$x \mapsto x^2$$

O conjuntos dos pontos recorrêntes para f é $\{0,1\}$. Portanto o suporte de uma medida invariante está contido em $\{0,1\}$. Por isso as únicas medidas possíveis são da forma

$$\alpha\delta_0 + \beta\delta_1$$

onde $\alpha, \beta \in \mathbb{R}_+$ e δ_i é a medida de Dirac no ponto i.

b)

$$f : [0,1] \longrightarrow [0,1]$$
$$x \longmapsto x^2 \;, x \notin \{0,1\}$$
$$f(x) = 1/2 \;, x \in \{0,1\}$$

O único ponto errante é o zero. Mas a probabilidade candidata, o delta de dirac em 0, não é invariante.

$$\delta_0(f^{-1}[0,1/4]) = \delta_0((0,1/4]) = 0 \neq 1 = \delta_0([0,1/2])$$

Logo, f não possui medida invariante.

c)

$$f : (0,1) \longrightarrow (0,1)$$
$$x \longmapsto x^2$$

f não possui medida invariante.

Topologia no conjunto de probabilidades

Definição 1.2.2. Seja $(M, \mathcal{B}(M))$ um espaço mensurável, com M espaço métrico compacto e $\mathcal{B}(M)$ a sigma álgebra de Borel. Definimos o seguinte conjunto

$$\mathcal{M} = \{\textit{probabilidades em } (M, \mathcal{B}(M))\}$$

Definição 1.2.3. Dado $f : M \to M$ mensurável, denotamos por $\mathcal{M}_0$ o conjuntos das medidas $\mu \in \mathcal{M}$ que sejam f-invariantes.

Munimos o conjunto $\mathcal{M}$ da topologia menos fina que torna a seguinte aplicação, I_ϕ, contínua

$$\begin{array}{rcl} I_\phi : \mathcal{M} & \to & \mathbb{R} \\ \eta & \mapsto & \displaystyle\int \phi \, d\eta \end{array}$$

onde $\phi : M \to \mathbb{R}$ é contínua.

Uma base para esta topologia seria portanto os conjuntos da forma

$$V_\mu(f_1, \ldots, f_n; \epsilon) = \left\{ m \in \mathcal{M}(X) \mid \left\| \int f_i \, dm - \int f_i \, d\mu \right\| \leq \epsilon, \; i = 1, \ldots, n. \right\}$$

$$\forall f_i \in C^0(X), \mu \in \mathcal{M}(X), \epsilon > 0, n \in \mathbb{N}.$$

Definição 1.2.4. Chamamos a topologia de $\mathcal{M}$ dada acima, por *Topologia Fraca** ou *Topologia Fraca Estrela.*

Fixemos um subconjunto denso enumerável, $\{\phi_k \mid k \in \mathbb{N}\}$, na bola unitária de $C^0(M)$ (funções contínua de M em $\mathbb{R}$). Consideremos a seguinte função

$$\begin{array}{rcl} d : \mathcal{M} \times \mathcal{M} & \to & \mathbb{R} \\ (\eta_1, \eta_2) & \mapsto & \sum_{k=1}^{\infty} \frac{1}{2^k} \left| \int \phi_k d\eta_1 - \int \phi_k d\eta_2 \right| \end{array}$$

Proposição 1.2.1. *A função d, acima, define uma métrica no espaço $\mathcal{M}$.*

Proof. A única dificuldade está em mostrar: se $d(\eta_1, \eta_2) = 0$, então $\eta_1 = \eta_2$. Este fato segue diretamente pela unicidade no teorema de representação de Riez. Entretando damos uma solução que não usa este teorema.

Suponha, por absurdo, que $\eta_1 \neq \eta_2$. Existe B boreliano tal que $\eta_1(B) > \eta_2(B)$ (outro caso é análogo). Logo existe um compacto K tal que $\eta_1(K) > \eta_2(K) \Rightarrow \eta_1(K) = \eta_2(K) = \delta$. Pelo lema de Urysohn podemos tomar $\phi : X \to \mathbb{R}$ tal que $\phi|_K \equiv 1$, $\phi|_{V \backslash K} \in (0, 1)$ e $\phi|_{V^c} \equiv 0$, onde V é uma vizinhança de K.

Podemos fazer V tão próxima de K de forma que $\delta > \eta_2(V \backslash K)$. Com isso,

$$\int \phi d\eta_1 \geq \eta_1(K) = \eta_2(K) + \delta > \int_K \phi d\eta_2 + \int_{V \backslash K} \phi d\eta_2 = \int \phi d\eta_2$$

O cálculo acima mostra que $\int \phi d\eta_1 > \int \phi d\eta_2$. E como os ϕ_k da definição de d são densos, existe ϕ_{k_0} com $\int \phi_{k_0} d\eta_1 > \int \phi_{k_0} d\eta_2$. O que é um absurdo. □

Proposição 1.2.2. *A Topologia Fraca* e a topologia dada pela métrica d coincidem.*

Proof. i) Seja U_0 um aberto da topologia gerada por d. Podemos supor que $U_0 = B_d(\mu_0, \epsilon)$ (bola na métrica d de raio ϵ centrada em μ_0) Seja $l > 0$ tal que

$$\sum_{k=l+1}^{\infty} \frac{1}{2^k} \left| \int \phi_k d\eta_1 - \int \phi_k d\eta_2 \right| < \epsilon/2$$

para quaisquer probabilidades η_1, η_2.

Como $I_{\phi_1}, \ldots, I_{\phi_l}$ são contínuas, tome δ pequeno o suficiente de forma que, se

$$\eta \in V_0 \Rightarrow \sum_{k=1}^{l} \left| \int \phi_k d\mu_0 - \int \phi_k d\eta \right| < \epsilon/2$$

onde

$$V_0 = \bigcap_{\alpha=1}^{l} I_{\phi_\alpha}^{-1}\left(I_{\phi_\alpha}(\mu_0) - \delta, I_{\phi_\alpha}(\mu_0) + \delta\right)$$

Obtemos, $d(\mu_0, \eta) < \epsilon/2 + \epsilon/2 = \epsilon$. Então $V_0 \subset U_0$.

ii) Seja V_0 um aberto da topologia fraca estrela. Podemos supor que

$$V_0 = \bigcap_{\alpha=1}^{l} I_{\phi_\alpha}^{-1}(I_{\phi_\alpha}(\mu) - \delta, I_{\phi_\alpha}(\mu) + \delta)$$

É fácil ver que tomando uma bola suficientemente pequena, $B_d(\mu, \epsilon) \subset V_0$.

□

Proposição 1.2.3. *Uma sequência $\{\eta_k\}_{k\in\mathbb{N}} \subset \mathcal{M}$ converge a η em $\mathcal{M}$ se, e somente se, $\lim\limits_{k\to\infty} \int \phi d\eta_k = \int \phi d\eta$, $\forall \phi \in \mathcal{M}$.*

Proof. $(\Rightarrow)$: Suponha $d(\eta_i, \eta) \to 0$. Dado $\epsilon > 0$ e $\phi \in C^0(M)$, defina $k_0 \in \mathbb{N}$ de forma que $|\phi - \phi_{k_0}| < \epsilon/4$. Tome i_0 grande o suficiente para que se $i \geq i_0$, então

$$|\int \phi_{k_0} d\eta_i - \int \phi_{k_0} d\eta| < \epsilon/2$$

Para $i \geq i_0$ tem-se

$$\begin{aligned} |\int \phi d\eta_i - \int \phi d\eta| &\leq |\int \phi_i d\eta - \int \phi_k d\eta_i| + |\int \phi_k d\eta_i - \int \phi_k d\eta| \\ &+ |\int \phi_k d\eta - \int \phi d\eta| \leq \epsilon/4 + \epsilon/2 + \epsilon/4 = \epsilon \end{aligned}$$

$(\Leftarrow)$: Exercício.

□

Corolário 1.2.1. *Seja $(\eta_i)_{i\in\mathbb{N}}$ um sequência em $\mathcal{M}$ Se $(\int \phi d\eta_i)_{i\in\mathbb{N}}$ converge para toda $\phi \in C^0(M)$, então a sequência $(\eta_i)_{i\in\mathbb{N}}$ converge fracamente a uma medida.*

Proof. Defina

$$\begin{aligned} I : C^0(M) &\to \mathbb{R} \\ \phi &\mapsto \lim_{i\to\infty} \int \phi d\eta_i \end{aligned}$$

I é linear, não negativa e de norma igual a 1. Podemos aplicar o Teorema de Riez. □

Teorema 1.2.5. *Na topologia fraca estrela, $\mathcal{M}$ é compacto.*

Proof. Seja $\{\phi_k\}_{i\in\mathbb{N}}$ um conjunto denso em $C^0(M)$. Dada uma sequência $\{\mu_i\}$ em $\mathcal{M}$, mostremos que admite uma subsequência convergente.

Para $k = 1$, a sequência $\{\int \phi_1 d\mu_i\}_{i\in\mathbb{N}}$ de reais, é limitada logo admite subsequência convergente (de subíndice $\mathbb{N}'$).

Para $k = 2$, a sequência $\{\int \phi_2 d\mu_i\}_{i\in\mathbb{N}'}$ é limitada, logo admite subsequência convergente.

Utilizando o processo de diagonal de cantor obtemos uma sequência $\{m_i\}_{i\in\mathbb{N}}$ de naturais, de forma que para todo $k \in \mathbb{N}$ a sequência $\{\int \phi_k d\mu_{m_i}\}_{i\in\mathbb{N}}$ seja convergente.

Isto implica que $\{\int \phi d\mu_{m_i}\}_{i\in\mathbb{N}}$ converge para toda $\phi \in C^0(M)$, pois $\{\int \phi_k d\mu_{m_i}\}_{i\in\mathbb{N}}$ é uma sequência de cauchy, pela desigualdade abaixo

$$\begin{aligned}\left|\int \phi d\mu_{m_i} - \int \phi d\mu_{m_j}\right| \quad &\leq \quad \left|\int \phi d\mu_{m_i} - \int \phi_k d\mu_{m_i}\right| + \left|\int \phi_k d\mu_{m_i}\right. \\ &- \left.\int \phi_k d\mu_{m_j}\right| + \left|\int \phi_k d\mu_{m_j} - \int \phi d\mu_{m_j}\right|\end{aligned}$$

Vemos que $\{\int \phi d\mu_{m_i}\}_{i\in\mathbb{N}}$ é cauchy $\forall \phi \in C^0(M)$. O corolário acima implica a convergência das medidas $\{\mu_{m_i}\}$, como queríamos. □

Seja $f : M \to M$ uma função mensurável. O push-forward em $\mathcal{M}$ é a função

$$f_* : \mathcal{M} \to \mathcal{M}$$

onde $(f_*\mu)(E) = \mu(f^{-1}(E))$. Note que uma medida invariante para f é um ponto fixo de f_*.

Proposição 1.2.4. *Se $f : M \to M$ for contínua, então $f_* : \mathcal{M} \to \mathcal{M}$ é contínua.*

Proof. Seja $\eta_n \to \eta$. Então, para $\phi \in C^0(M)$

$$\int \phi \; d(f_*\eta_n) = \int \phi \circ f \; d\eta_n \to \int (\phi \circ f) \; d\eta = \int \phi \; d(f_*\eta)$$

Então $f_*\eta_n \to f_*\eta$, logo f_* é contínua. □

Exemplo 1.2.6. Seja $f : [0,1] \to [0,1]$, definida por $f(0) = 0$, $f(x) = 1$ se $x \neq 0$. Tome $\phi \in C^0([0,1])$ dada por $\phi(x) = x$. Note que $\phi \circ f = f$. Sabemos que as medidas $\delta_{1/n}$, delta de Dirac centradas em $1/n$, convergem fracamente a δ_0. Entretanto $f_*\delta_{1/n} \nrightarrow f_*\delta_0$, já que

$$\begin{aligned}\int \phi \; d(f_*\delta_{1/n}) \quad &= \quad \int \phi \circ f \; d\delta_{1/n} = \int f \; d\delta_{1/n} = 1 \\ &\neq \quad 0 = \int f \; d\delta_0 = \int \phi \; d(f_*\delta_0)\end{aligned}$$

Proposição 1.2.5. *Sejam $\eta_n, \eta \in \mathcal{M}$, $n \geq 0$ então são equivalentes:*

- $\eta_n \to \eta$, *na Topologia Fraca**;
- *Para todo fechado $F \subset M$ então lim sup $\eta_n(F) \leq \eta(F)$;*
- *Para todo aberto $U \subset M$ então lim inf $\eta_n(U) \geq \eta(U)$;*
- *Para todo $A \in \mathcal{B}(M)$ com $\eta(\partial A) = 0$ então $\eta_n(A) \to \eta(A)$.*

Teorema 1.2.7 (Krylov-Bogolubov)**.** *Seja $f : M \to M$ uma transformação contínua em um espaço métrico compacto. Então f admite uma medida de probabilidade invariante.*

Proof. Seja μ uma probabilidade qualquer em $\mathcal{M}$. Defina

$$\mu_n = \frac{1}{n} \sum_{j=0}^{n-1} f_*^j \mu$$

Sabemos que μ_n possui subsequência convergente. Denote este limite por μ_0; $\mu_{n_i} \to \mu_0$. Usando a continuidade de f_*:

$$\begin{aligned} f_*\mu_0 &= f_* \left(\lim_{i\to\infty} \frac{1}{n_i} \sum_{j=0}^{n_i-1} f_*^j \mu \right) = \lim_{i\to\infty} \left(\frac{1}{n_i} \sum_{j=0}^{n_i-1} f_*^j \mu + \frac{1}{n_i} f_*^{n_i} \mu - \frac{1}{n_i} \mu \right) \\ &= \mu_0. \end{aligned}$$

Para a última igualdade olhe as medidas integrando sobre as funções contínuas, a contribuição de $\frac{1}{n_i} f_*^{n_i} \mu - \frac{1}{n_i} \mu$ vai a zero.

□

Chapter 2

Ergodicidade

Considere $(X, \mathcal{F}, \mu)$ um espaço de medida. Se $T : X \to X$ uma função mensurável, então a n-ésima média de Birkhoff para uma função $f : X \to \mathbb{R}$ mensurável é definida como

$$S_n(f, x) = \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x))$$

Teorema 2.0.8 (Teorema Ergódico de Birkhoff, em espaço de probabilidade)**.** *Seja $T : X \to X$ função mensurável preservando uma probabilidade μ.*

a) Se $f \in L^p(X, \mu)$, $1 \leq p < \infty$, então

- $f^+(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) \in L^p(\mu)$;
- $\lim_{n\to\infty} \left\| f^+ - \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) \right\|_p = 0$;
- $f^+ \circ T = f^+$.

b) Se $f \in L^1(\mu)$, então

$$\int_X f^+ d\mu = \int_X f d\mu.$$

Teorema 2.0.9. *Seja $T : X \to X$ uma função mensurável em um espaço de probabilidade que preserva a probabilidade μ. São equivalentes:*

a) ϕ^+ é constante μ-q.t.p. $\forall \phi \in L^1$;

a') ϕ^+ é constante μ-q.t.p. $\forall \phi$ em m conjunto denso de L^1;

b) $\phi^+ = \int \phi \, d\mu$ $\forall \phi \in L^1$;

b') $\phi^+ = \int \phi \, d\mu$ $\forall \phi$ em um subconjunto denso de L^1;

c) Se $\phi : M \to \mathbb{R}$, tal que $\phi(T(x)) = \phi(x)$ μ-q.t.p. então ϕ é constante μ-q.t.p.;

d) Se $A \subset M$ tal que $T^{-1}(A) = A$, então $\mu(A) = 0$ ou $\mu(A) = 1$;

d') Se $A \subset M$ tal que $T^{-1}(A) \subset A$, então $\mu(A) = 0$ ou $\mu(A) = 1$;

d") Se $A \subset M$ tal que $T^{-1}(A) \supset A$, então $\mu(A) = 0$ ou $\mu(A) = 1$.

Proof.

- *a)* $\Rightarrow$ *c)*: $\phi \circ T = \phi$ implica $\phi^+ = \phi$, que é constante q.t.p. por hipótese.
- *c)* $\Rightarrow$ *a)*: Sabemos que $\phi^+ \circ T = \phi^+$, logo ϕ^+ é constante q.t.p..
- *c)* $\Rightarrow$ *d)*: Tome $\phi = \chi_A$. Se $T^{-1}(A) = A$ então $\phi \circ T = \phi$. Por hipótese ϕ é constante q.t.p. igual a zero ou 1. Integrando ϕ sobre X obtemos o desejado.
- *d)* $\Rightarrow$ *c)*: Dada ϕ invariante. Defina $A_c = \{x \mid \phi(x) \leq c\}$, note que A_c é invariante por T. Logo $\mu(A_c) = 0$ ou 1. Considere a função

  $$c \mapsto \mu(A_c)$$

  seja c_0 o ponto em que ocorre a descontinuidade, isto é, $c_0 = sup\{c \mid \mu(A_c) = 0\}$. Portanto,

  $$\forall c < c_0 \Rightarrow \mu(\{x|\phi(x) < c\}) = 0 \text{ e } \forall c > c_0 \Rightarrow \mu(\{x|\phi(x) > c\}) = 0$$

  Sendo

  $$\{x \mid \phi(x) \neq c_0\} = (\cup_{k=1}^{\infty} A_{c-1/k}) \cup (\cup_{k=1}^{\infty} (A_{c_0+1/k})^c)$$

  uma união enumerável de conjuntos de medida nula, então $\{x \mid \phi(x) \neq c_0\}$ tem medida nula. Por conseguinte $\phi(x) = c_0$ q.t.p..
- $a')$ $\Rightarrow$ $a)$: Pelo Teorema Ergódico definamos

  $$\begin{array}{rcl} B : L^1(\mu) & \to & L^1(\mu) \\ \phi & \mapsto & \phi^+ \end{array}$$

  que é linear e contínua dado que $||B|| \leq 1$ $(||\phi^+||_1 \leq ||\phi||_1)$. Seja $D \subset L^1(\mu)$ um conjunto denso como na hipótese. Sendo B contínua $B(\overline{D}) \subset \overline{B(D)}$. Então

  $$B(L^1(\mu)) \subset \overline{\{\text{funções constantes}\}} = \{\text{funções constantes}\}$$

$\square$

Definição 2.0.10. Dizemos que (f, μ) é ergódica se satisfaz uma das propriedades acima.

2.1 Exemplos de transformações ergódicas

Exemplo 2.1.1.

a) A rotação irracional $R_\theta : S^1 \to S^1$ é ergódica com relação a medida de Lebesgue.

   Seja $\psi : S^1 \to \mathbb{C}$ integrável. Sua série de Fourier é dada por $\psi(\alpha) = \sum_{n\in\mathbb{Z}} a_n e^{2\pi i n\alpha}$, $a_n \in \mathbb{C}$. Chequemos a propriedade $c)$ do Teorema 2.0.9. Se $\psi \circ R_\theta = \psi$, então

$$\sum_{n\in\mathbb{Z}} a_n e^{2\pi i n(\alpha+\theta)} = \sum_{n\in\mathbb{Z}} a_n e^{2\pi i n\alpha} .$$

   Por conseguinte $a_n e^{2\pi i n\theta} = a_n \; \forall n \in \mathbb{Z}$. Se $2\pi i n\theta \in 2\pi i\mathbb{Z}$, então $n = 0$ já que $\theta \in \mathbb{R}\backslash\mathbb{Q}$.

   Para $n \neq 0 \Rightarrow a_n = 0$. Logo $\psi = a_0$, e portanto constante q.t.p..

b) Seja $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n$ um vetor *racionalmente independente*, isto é $\sum_{i=1}^n a_i\alpha_i \notin \mathbb{Z}, \forall (a_1, \dots, a_n) \in \mathbb{Z}^n$. Uma rotação irracional no toro é uma aplicação da forma

$$\begin{array}{rcl} T : \mathbb{T}^n & \to & \mathbb{T}^n \\ (x_1, \dots, x_n) & \mapsto & (x_1 + \alpha_1, \dots, x_n + \alpha_n) \end{array}$$

   Analogamente ao item a), seja $\phi \in C^0(\mathbb{T}^n)$ que comuta com T, então $\phi(X) = \sum_{Z\in\mathbb{Z}^n} a_Z e^{2\pi i <Z,X>}$. E concluímos que $a_Z = 0, \forall Z \neq 0 \in \mathbb{Z}^n$.

Exemplo 2.1.2. Seja

$$\begin{array}{rcl} E_m : [0,1] & \to & [0,1] \\ x & \mapsto & mx \; (mod \; 1) \end{array}$$

E_m preserva lebesgue. Provemos que E_m é ergódica. Seja $\phi \in C^0([0,1])$ que comuta com E_m. Considere a série de Fourier de ϕ. $\phi(x) = \sum_{n\in\mathbb{Z}} a_n e^{2\pi i n x}$. Portanto $a_n = a_{mn}$. A teoria de Fourier nos fornece que $\sum_n |a_n|^2 = ||\phi||_{L^2} < \infty$. Consequentemente se $n \neq 0$, então $a_n = a_{mn} = a_{m^2 n} = \dots$ devemos ter $a_n = 0$. E portanto E_m é ergódica.

Exemplo 2.1.3. Para quase todo ponto (Lebesgue) em $[0,1)$ a frequência de $1's$ na expansão binária é $1/2$.

De fato, considere a função

$$T : [0,1) \to [0,1)$$

$$T(x) = 2x \; mod \; 1$$

então T preserva a medida de Lebesgue. Considere $x = a_1/2 + a_2/2^2 + \ldots$ que possui expansão binária única (os que não possuem são um conjunto enumerável, por isso os desconsideramos). Contamos a quantidade de $1's$ na expansão de x, definindo a funcão $f = \chi_{[1/2,1)}$ e notando que a soma $\sum_{k=1}^{n} f(T^k)$ fornece a quantidade de $1's$ que aparecem nos coeficiente $a_1, \ldots, a_n$. Portanto dividindo por n tomamos a média. O Teorema Ergódico implica

$$\frac{1}{n}\sum_{k=1}^{n} f(T^k) \to \int \chi_{[1/2,1)} = 1/2$$

Exemplo 2.1.4. Consideremos o 2-toro $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. Dada uma matriz 2×2 com entradas inteiras e determinante 1, $A : \mathbb{R}^2 \to \mathbb{R}^2$, como $A(\mathbb{Z}) = \mathbb{Z}$ então A induz naturalmente uma transformação em $\mathbb{T}^2$ que também denotaremos por $A : \mathbb{T}^2 \to \mathbb{T}^2$.

Uma vez que A tem determinante 1, A preserva a medida de volume usual em $\mathbb{R}^2$ e, portanto, a transformação induzida preserva a medida usual no 2-toro. Vamos mostrar que A é ergódica se, e somente se, A não tem autovalores de módulo igual a um.

Considere $\varphi \in L^2$ uma função A-invariante, isto é, $\varphi \circ A = \varphi$ q.t.p. Escreva a expansão em Série de Fourier de φ:

$$\varphi(x) = \sum_{n\in\mathbb{Z}^2} a_n \cdot e^{2\pi i <x,n>}.$$

Compondo φ com A temos então

$$\varphi(A(x)) = \sum_{n\in\mathbb{Z}^2} a_n \cdot e^{2\pi i <A(x),n>} = \sum_{n\in\mathbb{Z}^2} a_n \cdot e^{2\pi i <x,A^*n>}.$$

Logo, como $\varphi(A(x)) = \varphi(x)$ devemos ter

$$\sum_{n\in\mathbb{Z}^2} a_n \cdot e^{2\pi i <x,n>} = \sum_{n\in\mathbb{Z}^2} a_n \cdot e^{2\pi i <x,A^*n>}.$$

Pela unicidade dos coeficientes na expansão em série de Fourier segue que

$$a_{A^*n} = a_n$$

para todo $n \in \mathbb{Z}^2$.

Fixado qualquer $n \in \mathbb{Z}$, pela desigualdade de Bessel sabemos que

$$\sum_{i\in\mathbb{Z}} |a_{(A^*)^i n}|^2 \leq \sum_{n\in\mathbb{Z}^2} |a_n|^2 = \|\varphi\|_2^2 < +\infty.$$

Então, dado qualquer $n \in \mathbb{Z}^2 \setminus \{(0,0)\}$ ou $a_n = 0$ ou existe $i \in \mathbb{Z}$ tal que $(A^*)^i n = n$. A segunda hipótese ocorre se, e somente se, A tem autovalores unitários.

Caso não tenha então $\varphi(x) = a_{(0,0)}$ concluindo que A é ergódica.

Exercício 2.1.1. Conclua a demonstração mostrando que no caso de A ter autovalores unitários é possível construir uma função φ não constante e A-invariante.

Ao pensar no sentido de ***caos*** para uma função $f : X \to X$ com respeito a uma medida μ preservada por f, onde (X, μ) é um espaço de probabilidade, podemos dizer intuitivamente que se o comportamento de f for suficientemente imprevisível ou "bagunçado" então f tem comportamento caótico. Uma das formas de dizer que o comportamento de f é imprevisível é dizer que se conseguimos prever o comportamento de um conjunto de pontos então esse conjunto de pontos deve "representar" todo o espaço ou não deve "representar" nada.

Conseguir prever o comportamento de pontos significa fechar um conjunto de órbitas em alguma região, isto é, significa tomar um conjunto $K \subset X$ de forma que a órbita de qualquer ponto de K nunca escapa de K. Isto foi o que definimos como invariância do conjunto K, ou seja, K é chamado de invariante por um automorfismo f se

$$f(K) = K.$$

Do ponto de vista de medida dizer que um conjunto de pontos "representa" todo o espaço signifca dizer que este conjunto tem medida total. Analogamente, dizer que um conjunto de pontos não "representa" nada significa dizer que ele tem medida nula. Assim, reescrevendo o primeiro parágrafo em termos matemáticos, concluímos que uma das formas de dizer que um sistema tem comportamento caótico (com relação a alguma medida μ) é dizer: se um conjunto K é f - invariante então

$$\mu(K) = 0 \text{ ou } \mu(K) = 1.$$

Esta propriedade é exatamente o que definimos por propriedade ergódica na seção anterior. Entretanto, como o leitor pode imaginar, deve existir diversos tipos diferentes de sistemas com comportamento caótico (no sentido ergódico) e que apresentam comportamento dinâmico completamente distintos. Por exemplo, a rotação irracional e o shift de Bernoulli são ambos ergódicos, entretanto, não são topologicamente conjugados. Uma forma simples de ver isto é observando que o shift de Bernoulli admite pontos periódicos enquanto que a rotação irracional em S^1 não tem pontos periódicos.

Assim, dentre os sistemas ergódicos deve existir diferentes graus de caoticidade de acordo com o quão complexa é uma dada dinâmica. Esses diferentes graus de caoticidade com respeito à uma medida invariante da origem a uma hierarquia entre as propriedades ergódicas.

Nas próximas seções deste capítulo apresentaremos dois graus de caoticidade, ou seja, duas propriedades ergódicas, que são mais fortes do que a ergodicidade.

2.2 A propriedade de Bernoulli e o Teorema de Ornstein

Nas seções anteriores vimos quão importante são os sistemas chamados de shifts de Bernoulli. Estes sistemas admitem uma variedade de propriedades probabilísticas e também topológicas. Do ponto de vista de teoria da medida seria interessante podermos trabalhar não apenas com os shifts de Bernoulli propriamente ditos, mas sim com uma classe de transformações cujas propriedades fossem as mesmas que as propriedades dos shifts de Bernoulli.

Definição 2.2.1. Sejam (X, μ) e (Y, ν) espaços de medida. Uma aplicação mensurável $\phi : X \to Y$ que preserva medida, é dita um isomorfismo $\bmod 0$ (ou simplesmente um isomorfismo), se existir um subconjunto $Z \subset X$ tal que $\mu(Z) = 0$ e $\phi : X \setminus Z \to \phi(X \setminus Z)$ é bijetora.

Definição 2.2.2. Sejam (X, μ) e (Y, ν) espaços de medida e sejam $f : X \to X$ e $g : Y \to Y$ automorfismos mensuráveis de X e Y. Dizemos que f é isomorfo a g $\bmod 0$ (ou simplesmente diremos que f é isomorfo a g) se existir um isomorfismo $\bmod 0$

$$h : X \to Y$$

que conjuga f e g, isto é

$$g \circ h = h \circ f.$$

Definição 2.2.3. Seja (X, μ) um espaço de medida e $f : X \to X$ um automorfismo μ-invariante. Dizemos que f é um automorfismo de Bernoulli se existe um shift de Bernoulli $\sigma : \Sigma_k \to \Sigma_k$ isomorfo a f.

Exercício 2.2.1. Prove que dado um número irracional qualquer $\alpha \in (0, 1)$, a rotação irracional de ângulo α,

$$R_\alpha : S^1 \to S^1$$

não é um automorfismo de Bernoulli.

Dica: Suponha por absurdo que R_α fosse isomorfo a um shift de Bernoulli $\sigma : \Sigma_k \to \Sigma_k$. Observe que dados dois intervalos pequenos $I_1, I_2 \subset S^1$ existe uma sequência $N_i \to +\infty$ tal que para todo i temos $f^{N_i}(I_1) \cap I_2 = \emptyset$. Entretanto, dados quaisquer conjuntos mensuráveis de medida positiva $B_1, B_2 \subset \Sigma_k$, utilize o passo 2 da prova do Teorema 2.4.3 para provar que existe $N_0 \geq 1$ tal que para todo $N \geq N_0$ temos $f^N(B_1) \cap B_2 \neq \emptyset$.

Uma pergunta natural a se fazer após a definição de automorfismos de Bernoulli é: quais são os invariantes que classificam os sistemas de Bernoulli? Ou seja, existe alguma propriedade que determine quando dois shifts de Bernoulli $\sigma_1 : \Sigma_k \to \Sigma_k$ e $\sigma_2 : \Sigma_l \to \Sigma_l$ são isomorfos?

Esta pergunta foi respondida na decada de 70 pelo matemático americano Donald Ornstein no artigo [5], o qual recebeu o Prêmio Bôcher, um prêmio concedido pela sociedade americana de matemática para publicações com destaque na área de análise.

As técnicas desenvolvidas para a demonstração do Teorema de Ornstein, o qual apresentaremos a seguir, constituem o que hoje é denominado como Teoria de Ornstein.

Teorema 2.2.4 (Teorema de Ornstein [5], [7])**.** *Dados dois shifts de Bernoulli*

$$\sigma_1 : \Sigma_k \to \Sigma_k \text{ e } \sigma_2 : \Sigma_l \to \Sigma_l,$$

associados a vetores de probabilidade

$$p_1 = (w_1, w_2, ..., w_k)$$

e

$$p_2 = (u_1, u_2, ..., u_l)$$

respectivamente. Então σ_1 é isomorfo a σ_2 se, e somente se,

$$\sum_{i=1}^{k} -w_i \cdot \log w_i = \sum_{i=1}^{l} -u_i \cdot \log u_i.$$

Observação 2.2.1. Os somatórios que aparecem no Teorema de Ornstein são as entropias (veja [4]) dos respectivos shifts de Bernoulli. Por isso podemos reescrever o Teorema de Ornstein dizendo que a entropia é um invariante que classifica os shifts de Bernoulli.

2.3 Operações com partições

A caoticidade ou “bagunça” gerada por um automofismo $f : X \to X$ pode ser medida através das modificações que ele faz à uma configuração inicial ao longo do tempo. Para entender melhor este conceito, tome uma partição inicial do espaço X em um número finito de conjuntos. Agora, iteramos os conjuntos dessa partição e comparamos com a partição inicial. É de se esperar que, a menos que f for muito bem comportada, a nova partição seja bem diferente da partição inicial e, se conseguirmos comparar as duas, podemos dizer que quanto maior for a diferença das duas mais bagunçado é o sistema.

Essa é uma idéia bem brusca e nada especifica do que realmente significa medir caoticidade através da "bagunça" causada em uma configuração inicial, mas o ponto aqui é que para que possamos seguir com esse raciocínio precisaremos saber fazer operações básicas entre partições assim como fazemos operações com conjuntos. Nesta seção introduziremos as operações com partições que nos permitirão, nas próximas seções, definir a propriedade de Kolmogorov e também provar a propriedade de Bernoulli para alguns sistemas.

Definição 2.3.1. Dados dois conjuntos A, B de um espaço de medida (X, μ), dizemos que A e B são iguais módulo zero, e denotamos

$$A \overset{\circ}{=} B,$$

se tivermos $m(A\Delta B) = 0$, ou seja, a menos de um conjunto de medida nula eles são o mesmo conjunto. Onde $A\Delta B$ é a diferença simétrica dos conjuntos A e B, ou seja $A\Delta B := A \cup B \setminus A \cap B$.

Definição 2.3.2. Seja (X, μ) um espaço de medida e $\mathcal{P}$ uma família de subconjuntos de X. Diremos que a família $\mathcal{P}$ é uma partição de X se ela satisfaz:

1. $$\bigcup_{P \in \mathcal{P}} P \overset{\circ}{=} X$$

2. Se $P_1, P_2 \in P$ com $P_1 \neq P_2$ então $P_1 \cap P_2 = \emptyset$.

Definição 2.3.3 (União de partições)**.**

- Sejam α e β duas partições de X, a união de α e β, denotada por $\alpha \vee \beta$, é a partição dada por

  $$\alpha \vee \beta := \{A \cap B : A \in \alpha, B \in \beta.\}$$

- Dada uma família infinita $\{\alpha_\omega\}_{\omega \in \Omega}$ de partições de um espaço X. A união infinita das partições α_ω, denotada por

  $$\bigvee_{\omega \in \Omega} \alpha_\omega,$$

  é a menor σ-álgebra contendo todos os elementos das partições α_ω, $\omega \in \Omega$.

Exemplo 2.3.4. Considere $X = [0, 1]$. Tomemos as partições

$$\alpha = \{[0, 1/5), [1/5, 4/5), [4/5, 1]\}$$

e

$$\beta = \{[0, 1/3), [1/3, 2/3], (2/3, 1]\}.$$

Então

$$\alpha \vee \beta = \{[0, 1/5), [1/5, 1/3), [1/3, 2/3], (2/3, 4/5), [4/5, 1]\}.$$

Exercício 2.3.1. Consideremos $X = [0, 1] \times [0, 1]$. Tomemos as partições α e β dadas por

$$\alpha = \{[0, 1/2] \times [0, 1/2], [0, 1/2] \times [1/2, 1], [1/2, 1] \times [0, 1/3], [1/2, 1] \times [1/3, 1]\},$$

$$\beta = \{[0, 1/4] \times [0, 1/2], [0, 1/4] \times [1/2, 1], [1/4, 1] \times [0, 2/3], [1/4, 1] \times [2/3, 1]\}.$$

Calcule $\alpha \vee \beta$.

Exemplo 2.3.5. Considere $\mathcal{B}$ a sigma álgebra de todos os Borelianos em $M = \mathbb{T}^2$, ou seja, é a menor sigma álgebra que contém todas as bolas abertas em M.

Agora consideramos α_i a partição de M obtida ao particionar o intervalo $[0, 1]$ em i pedaços idênticos e então considerar os quadradinhos formados no cartesiano $[0, 1] \times [0, 1]$. Observe que a sequência de partições ξ_{2^i} é cada vez mais fina que a anterior e converge para a partição por pontos. Além disso,

$$\bigvee_{i=1}^{+\infty} \xi_{2^i} = \mathcal{B}.$$

Definição 2.3.6 (Partição geradora)**.** Seja $(X, \mathcal{A}, \mu)$ um espaço de medida, onde $\mathcal{A}$ é uma sigma álgebra em X. Dado um automorfismo (i.e. bijeção mensurável) μ-invariante $f : X \to X$, dizemos que uma partição α é dita geradora se

$$\bigvee_{-\infty}^{+\infty} f^n \alpha = \mathcal{A}.$$

2.4 A propriedade de Kolmogorov

A propriedade de Kolmogorov é muito menos intuitiva de se definir do que a propriedade ergódica e a propriedade de Bernoulli. A idéia é considerar sistemas que geram bagunça independente da configuração inicial dada. Em outras palavras, queremos que qualquer configuração inicial se espalhe por todo o espaço ao longo do tempo.

Então uma forma de tentar expressar esse fato em termos matemáticos seria a seguinte: tome um automorfismo que preserva medida $f : (X, \mu) \to (X, \mu)$ e uma partição finita ξ do seu espaço X. Tome qualquer conjunto mensurável $B \subset X$. Dizer que a partição ξ fica bagunçada pelo sistema f seria, a grosso modo, dizer que num futuro suficientemente distante a partição $f^N\xi$ (que é o futuro da partição ξ) se mistura muito bem com o conjunto B, ou seja, para quase todo elemento de $f^N\xi$ a medida do conjunto B pode ser aproximada pela medida condicional do B neste elemento.

A seguir apresentamos a definição formal da propriedade de Kolmogorov, porém, o leitor deve levar consigo a característica intuitiva desta propriedade: *um sistema é dito de Kolmogorov se qualquer partição finita é suficientemente bagunçada ao longo do tempo.*

Definição 2.4.1. Seja f um automorfismo de um espaço de Lebesgue (X, μ). Uma partição finita ξ de X é dita uma partição Kolmogorov (ou simplesmente uma K-partição) se ela satisfaz a seguinte propriedade: dado qualquer conjunto mensurável $B \subset X$, um real $\delta > 0$, existem $N_0 > 0$ tais que para quaisquer $N' \geq N \geq N_0$ e δ-quase todo elemento $A \in \bigvee_{k=N}^{N'} f^k\xi$ temos

$$\left| \frac{\mu(A \cap B)}{\mu(A)} - \mu(B) \right| \leq \delta.$$

Se existir uma K-partição ξ geradora então dizemos que f é um automorfismo de Kolmogorov (ou simplesmente um K-automorfismo).

Um resultado de Pinsker [8] e de Rohklin - Sinai [11] demonstra que se uma partição finita e geradora ξ for K então toda partição finita será K.

Teorema 2.4.2 (Pinsker [8], Rokhlin-Sinai [11])**.** *Seja f um K-automorfismo de um espaço de Lebesgue (X, μ). Toda partição finita ξ de X é uma K-partição.*

Exercício 2.4.1. Demonstre que a propriedade de Kolmogorov é invariante por isomorfismo, isto é: dados dois automorfismos $f : X \to X$ e $g : Y \to Y$ em espaços de medida $(X, \mathcal{A}, \mu)$, $(Y, \mathcal{D}, \nu)$. Se f é Kolmogorov e f é isomorfo a g então g é Kolmogorov.

O próximo resultado não é facilmente encontrado em livros textos de Teoria Ergódica e para tornar as notas o mais auto contida faremos a prova em detalhes.

Teorema 2.4.3 (Hierarquia Ergódica)**.**

$$\textit{Bernoulli} \subsetneq \textit{Kolmogorov} \subsetneq \textit{Ergódico}.$$

Proof. Primeiramente provaremos que:

$$\textit{Kolmogorov} \subsetneq \textit{Ergódico}.$$

Tome um automorfismo $f : X \to X$ que é Kolmogorov com respeito a uma medida de probabilidade μ em X. Seja J um conjunto f-invariante, isto é,

$$f(J) = J,$$

assuma por contradição que

$$0 < \mu(J) < 1.$$

Considere a partição finita $\xi = \{J, X \setminus J\}$. Como $f(J) = J$ então

$$\bigvee_{N}^{N'} f^k \xi = \xi$$

para todos $N' \geq N$. Pelo Teorema 2.4.2 a partição ξ é uma K-partição, então tomando $B = J$ e

$$\delta := \min\{\mu(J)/3, \mu(X \setminus J)/3\}$$

na definição de K-partição (veja Definição 2.4.1) temos que, para δ-quase todo átomo A de

$$\bigvee_{N}^{N'} f^k \xi = \xi$$

devemos ter

$$\left| \frac{\mu(A \cap J)}{\mu(A)} - \mu(J) \right| \leq \delta. \tag{2.1}$$

Como $m(J) > \delta$ então (2.1) deve ocorrer para $A = J$. Logo, substituindo $A = J$ em (2.1) temos:

$$|1 - \mu(J)| \leq \delta.$$

Portanto,

$$\mu(X \setminus J) = 1 - \mu(J) \leq \delta,$$

caindo em contradição com a escolha de δ. Portanto Kolmogorov implica a ergodicidade.

Provaremos agora que:

$$\textit{Bernoulli} \subsetneq \textit{Kolmogorov}$$

Lembremos que pela definição um automorfismo f é de Bernoulli se ele for isomorfo a algum shift de Bernoulli. Assim, pelo exercício 2.4.1, para provar que a propriedade de Bernoulli implica a propriedade de Kolmogorov basta provar que todos os shifts de Bernoulli são automorfismos de Kolmogorov.

Considere um $k \in \mathbb{N}$ arbitrário e tome o shift de Bernoulli $\sigma : \Sigma_k \to \Sigma_k$. Consideremos a partição ξ formada pelos cilindros que tem restrição apenas na primeira coordenada (ξ possui exatamente k elementos).

Passo 1: Provemos que a partição ξ é geradora, isto é,

$$\bigvee_{-\infty}^{+\infty} f^n\xi = \mathcal{C}$$

onde $\mathcal{C}$ é a σ-algebra gerada pelos cilindros.

Observe que $f^N\xi$ é exatamente a família dos cilindros com restrição na N-ésima coordenada. Consideremos um cilindro C qualquer. Então, podemos tomar $n_1, n_2, ..., n_l$ inteiros tais que C é um cilindro caracterizado por restrições nas coordenadas $n_1, ..., n_l$. Então,

$$C \in f^{n_1}(\xi) \cap f^{n_2}\xi \cap ... \cap f^{n_l}\xi.$$

Logo,

$$C \in \bigvee_{-\infty}^{+\infty} f^n\xi.$$

Da arbitrariedade na escolha do cilindro C concluímos que

$$\mathcal{C} \subset \bigvee_{-\infty}^{+\infty} f^n\xi.$$

Para obter o outro lado basta lembrar que todo elemento de $f^n\xi$ é um cilindro, portanto a σ-algebra $\bigvee_{-\infty}^{+\infty} f^n\xi$ está contida em $\mathcal{C}$. Concluímos assim que

$$\bigvee_{-\infty}^{+\infty} f^n\xi = \mathcal{C}.$$

Passo 2: Vamos provar que todo conjunto mensurável é aproximado por união finita disjunta de cilindros.

Considere um conjunto mensurável $X \subset \Sigma_k$ qualquer. Denote por $\pi_N : \Sigma_k \to \{0,1\}^{2N+1}$ a projeção sobre as coordenadas $-N, -N+1, ..., 0, ..., N-1, N$. Para cada $N \in \mathbb{N}$ denote X_N o conjunto

$$X_N = (\pi_N)^{-1}(\pi_N(X))$$

que é um conjunto da forma

$$... \times \{0,1,...,k-1\} \times \pi_N(X) \times \{0,1,...,k-1\} \times ...$$

o qual é uma união finita, disjunta, de cilindros uma vez que $\pi_N(X) \subset \{0, 1, ..., k-1\}^{2N+1}$ Assim, pela definição da medida produto μ em Σ_k sabemos que $\mu(X_N) = \mu(\pi_N(X))$. Mas observe que

$$X = \bigcap_{N \in \mathbb{N}} X_N \Rightarrow \mu(X) = \lim_{N \to +\infty} \mu(X_N).$$

Logo, dado qualquer $\varepsilon > 0$ podemos encontrar $N(\varepsilon) \in \mathbb{N}$ de forma que

$$\mu(X \Delta X_{N(\varepsilon)}) < \varepsilon$$

onde $X_{N(\varepsilon)}$ é uma união finita disjunta de cilindros.

Passo 3: Considere $B \subset \Sigma_k$ um conjunto mensurável qualquer e $\delta > 0$ um número real dado. Seja $p = (p(0), ..., p(k-1))$ o vetor de probabilidade associado ao shift em questão, denotemos por P o menor valor dentre os $p(i)$'s, isto é,

$$P = \min_{1 \leq i \leq k-1} p(i).$$

Tome

$$\varepsilon = \frac{P}{2(P+1)} \cdot \delta.$$

Pelo passo 2 podemos tomar uma coleção finita de cilindros disjuntos $C_1, ..., C_l$ de forma que

$$\mu(B \Delta \bigcup_{i=1}^{l} C_i) < \varepsilon. \tag{2.2}$$

Tome N_0 grande o bastante para que todos os cilindros C_i , $1 \leq i \leq l$, envolvam restrições até no máximo a coordenada $N_0 - 1$. Então, dados quaisquer $N' \geq N \geq N_0$, qualquer elemento

$$A \in \bigvee_{N}^{N'} f^n \xi$$

é um cilindro cujas posições de restrição são diferentes das de todos os cilindros C_i, portanto

$$\mu\left(A \cap \left(\bigcup_{i=1}^{l} C_i\right)\right) = \sum_{i=1}^{l} \mu(A \cap C_i) = \mu(A) \cdot \sum_{i=1}^{l} \mu(C_i) = \mu(A) \cdot \mu\left(\bigcup_{i=1}^{l} C_i\right).$$

Consequentemente

$$\frac{\mu\left(A \cap \left(\bigcup_{i=1}^{l} C_i\right)\right)}{\mu(A)} = \mu\left(\bigcup_{i=1}^{l} C_i\right).$$

Observe que

$$\mu\left(A\cap\left(\bigcup_{i=1}^{l}C_i\right)\right)=\mu\left(A\cap\left(\bigcup_{i=1}^{l}C_i\right)\cap B\right)+\mu\left(A\cap\left(\left(\bigcup_{i=1}^{l}C_i\right)\setminus B\right)\right)$$

$$<\mu(A\cap B)+\varepsilon.$$

Similarmente,

$$\mu(A\cap B)<\mu\left(A\cap\left(\bigcup_{i=1}^{l}C_i\right)\right)+\varepsilon.$$

Deste modo

$$\left|\mu\left(A\cap\left(\bigcup_{i=1}^{l}C_i\right)\right)-\mu(A\cap B)\right|<\varepsilon.$$

Portanto

$$\left|\frac{\mu\left(A\cap\left(\bigcup_{i=1}^{l}C_i\right)\right)}{\mu(A)}-\frac{\mu(A\cap B)}{\mu(A)}\right|<\frac{1}{\mu(A)}\cdot\varepsilon,$$

logo

$$\begin{aligned}\left|\frac{\mu(A\cap B)}{\mu(B)}-\mu(B)\right|\leq&\left|\frac{\mu(A\cap B)}{\mu(B)}-\frac{\mu\left(A\cap\left(\bigcup_{i=1}^{l}C_i\right)\right)}{\mu(A)}\right|+\\&+\left|\frac{\mu\left(A\cap\left(\bigcup_{i=1}^{l}C_i\right)\right)}{\mu(A)}-\mu\left(\bigcup_{i=1}^{l}C_i\right)\right|+\left|\mu\left(\bigcup_{i=1}^{l}C_i\right)-\mu(B)\right|\\&\leq\frac{1}{\mu(A)}\varepsilon+0+\varepsilon<\varepsilon(\frac{1}{P}+1)<\delta.\end{aligned}$$

Ou seja, dado qualquer $\delta>0$ encontramos N_0 grande o bastante de forma que para todos $N'\geq N\geq N_0$ e todo $A\in\bigvee_{N}^{N'}f^n\xi$ temos

$$\left|\frac{\mu(A\cap B)}{\mu(B)}-\mu(B)\right|<\delta.$$

Assim, ξ é de fato uma partição de Kolmogorov e, como ela é geradora, concluímos que o shift $\sigma:\Sigma_k\to\Sigma_k$ é um automorfismo de Kolmogorov, como queríamos demonstrar.

Para finalizar a demonstração do teorema devemos mostrar que a inclusão de cada classe é estrita na classe superior, ou seja, devemos apresentar um exemplo de transformação Kolmogorov que não é Bernoulli e um exemplo de transformação ergódica que não é Kolmogorov. Para não fugir ao propósito destas notas apenas citamos o livro [3] onde o leitor pode encontrar tais exemplos.

$\square$

Existem outros sistemas intermediários aos que aparecem no enunciado do Teorema, o leitor mais experiente pode conferir o capítulo 7 de [3] para buscar mais informações.

Chapter 3

Automorfismos ergódicos de $\mathbb{T}^2$ são Kolmogorov

O objetivo principal deste último capítulo é demonstrar que uma classe muito famosa de transformações ergódicas, os automorfismos ergódicos do 2-toro, são na verdade Bernoulli.

A demonstração desse fato foi dada por D. Ornstein e B. Weiss no artigo intitulado "Geodesic flows are Bernoullian" [2] publicado em 1973. O método introduzido por Ornstein-Weiss para demonstrar este fato também é utilizada em [2] para demonstrar que fluxos geodésicos em superfícies Riemanianas de curvatura negativa são Bernoulli. Posteriormente diversos autores (por exemplo [1], [9], [10], [13]) utilizaram as ferramentas introduzidas para provar que outros tipos de sistemas que são Kolmogorov são Bernoulli e, até o momento, esta ainda é a ferramenta mais poderosa para obter a propriedade de Bernoulli a partir da propriedade de Kolmogorov.

3.1 Automorfismos em grupos compactos

Para que possamos provar que automorfismos ergódicos de $\mathbb{T}^2$ são Bernoulli, primeiro iremos verificar que eles são Kolmogorov, este será o principal objetivo desta seção. Para não fugir do propósito destas notas os teoremas dessa seção serão apresentados sem demonstrações e maiores aprofundamentos.

Definição 3.1.1. G é chamado um grupo topológico se G é munido de uma topologia onde as aplicações

$$G \times G \to G$$

$$(a, b) \mapsto a \cdot b$$

$$G \to G$$

$$a \mapsto a^{-1}$$

são contínuas.

Exemplo 3.1.2. $(\mathbb{R}, +)$ é grupo topológico.

Exemplo 3.1.3. $(\mathbb{T}^2, +)$ com a topologia usual (induzida da projeção $\mathbb{R}^2 \to \mathbb{T}^2$) é grupo topológico.

Mostraremos agora alguns resultados obtidos para automorfismos ergódicos em grupos. Considere G um grupo, como usual, a terminologia "automorfismo" denotará aqui automorfismos no sentido de grupos topológicos, ou seja, uma aplicação $\varphi : G \to G$ de um grupo topológico G é dita um automorfismo de G se φ é bijetora, contínua e satisfaz

$$\varphi(a \cdot b) = \varphi(a) \cdot \varphi(b),$$

para todos $a, b \in G$. Em outras palavras, um automorfismo de G é um homomorfismo bijetor de G.

Exemplo 3.1.4. Qualquer aplicação linear bijetora $f : \mathbb{R} \to \mathbb{R}$ é um automorfismo pois é contínua e

$$f(a + b) = f(a) + f(b).$$

Exemplo 3.1.5. Seja $L : \mathbb{R}^2 \to \mathbb{R}^2$ uma aplicação linear bijetora, considere a função $A : \mathbb{T}^2 \to \mathbb{T}^2$ induzida da aplicação L. Então A é um automorfismo pois é bijetora, contínua e

$$A(x + y) = A(x) + A(y),$$

para quaisquer $x, y \in \mathbb{T}^2$.

Sempre que falamos sobre ergodicidade é necessário que haja uma medida envolvida. No caso de grupos é natural que tentemos usar medidas que de alguma forma levem em conta a operação definida no grupo e que sejam, portanto, bem comportadas com respeito à operação no grupo. Tais medidas são chamadas *medidas de Haar*. Para não fugir ao objetivo destas notas, iremos nos restringir apenas aos resultados necessários para a continuidade da leitura, não apresentando muitos resultados sobre a teoria de medidas de Haar. Como uma referência para o estudo de medidas de Haar e para a prova dos teoremas desta seção citamos [4].

Definição 3.1.6. Uma medida boreliana finita μ_G definida em G é chamada de medida de Haar se satisfaz as seguintes condições:

1. $\mu_G(U) > 0$ para todo conjunto aberto não vazio $U \subset G$,

2. para qualquer conjunto mensurável $E \subset G$ e qualquer elemento $g \in G$ temos
$$\mu_G(E \cdot g) = \mu_G(E) = \mu_G(g \cdot E),$$
isto é, μ_G é invariante pela operação do grupo à esquerda ou à direita.

Teorema 3.1.7. *Se G é um grupo topológico compacto então G admite uma única medida de probabilidade que é medida de Haar. Além disso, essa medida é invariante por qualquer automorfismo $\varphi : G \to G$.*

Exemplo 3.1.8. Considere $G = \mathbb{T}^2$ munido da operação usual de soma. A medida de Haar $\mu_{\mathbb{T}^2}$ é simplesmente a medida induzida pela medida de Lebesgue usual em $[0,1]^2$. Ou seja, considere a projeção natural $\pi : [0,1]^2 \to \mathbb{T}^2$ e *Leb* a medida de Lebesgue normalizada em $[0,1]^2$. A medida de Haar $\mu_{\mathbb{T}^2}$ é definida por
$$\mu_{\mathbb{T}^2}(E) = Leb(\pi^{-1}(E)).$$
Ao longo deste capítulo também utilizaremos a notação $m = \mu_{\mathbb{T}^2}$ quando estiver implícito que estamos trabalhando em $\mathbb{T}^2$, e também chamaremos $m = \mu_{\mathbb{T}^2}$ de medida de Lebesgue em $\mathbb{T}^2$.

O Teorema 3.1.7 nos mostra que, de fato, a medida de Haar associada a um grupo topológico compacto G é a medida "natural" para este grupo no sentido que ela é preservada por qualquer automorfismo do grupo e, portanto, podemos estudar a ergodicidade de automorfismos com respeito à essa medida de Haar.

Definição 3.1.9. Seja G um grupo topológico compacto e μ_G a medida de Haar em G que é medida de probabilidade. Diremos que um automorfismo $\varphi : G \to G$ é ergódico se ele for ergódico com respeito à medida μ_G.

Teorema 3.1.10 (V. A. Rokhlin [12], S. A. Yuzvinskii [14])**.** *Seja G um grupo topológico compacto e $\varphi : G \to G$ um automorfismo ergódico, então φ é um automorfismo Kolmogorov.*

Corolário 3.1.1. *Automorfismos ergódicos de $\mathbb{T}^2$ são Kolmogorov.*

Proof. Como vimos anteriormente, automorfismos ergódicos de $\mathbb{T}^2$ são automorfismos $A : \mathbb{T}^2 \to \mathbb{T}^2$ induzidos de matrizes 2×2 com entradas inteiras e sem autovalores não unitários. Pelo Teorema 3.1.10 segue que automorfismos ergódicos de $\mathbb{T}^2$ são, portanto, Kolmogorov. $\square$

3.2 Alguns resultados de Teoria de Ornstein

Nos anos 70, o matemático estadunidense Donald Samuel Ornstein obteve uma classificação completa para os shifts de Bernoulli. Como descrito no

Teorema 2.2.4, Ornstein provou [5] que dois shifts de Bernoulli $\sigma_1 : \Sigma_k \to \Sigma_k$ e $\sigma_2 : \Sigma_l \to \Sigma_l$, associados a vetores de probabilidade

$$p_1 = (w_1, w_2, ..., w_k)$$

e

$$p_2 = (u_1, u_2, ..., u_l)$$

respectivamente, são isomorfos se, e somente se

$$\sum_{i=1}^{k} -w_i \cdot \log w_i = \sum_{i=1}^{l} -u_i \cdot \log u_i.$$

Para isto, D. Ornstein desenvolveu uma teoria que atualmente é conhecida como Teoria de Ornstein. Um elemento central dessa teoria é o conceito de *partições Bernoulli muito fracas*, o qual introduziremos nessa seção. Por meio desse conceito, D. Ornstein construiu uma forma de garantir que um sistema é de Bernoulli sem propriamente encontrar uma conjugação desse sistema com um shift de Bernoulli, mas ao invés disso encontramos partições que estão muito próximas de ser uma partição isomorfa à partição por cilindros de um espaço de sequências. Os conceitos ficarão mais precisos durante a exposição mas é importante o leitor obter a intuição de que partições Bernoulli muito fracas são como aproximações de uma partição de Bernoulli.

Daqui em diante, M sempre denotará $\mathbb{T}^2$ e m sempre denotará a medida de Lebesgue (medida de Haar) em $\mathbb{T}^2$.

Definição 3.2.1 (Espaço de átomos)**.** Dizemos que um espaço de probabilidade (P, p) é composto de átomos se X for contável e cada elemento de X tiver medida positiva, isto é,

$$p(\{x\}) > 0, \forall x \in X.$$

Definição 3.2.2 (Espaço de Lebesgue)**.** Um espaço de probabilidade (X, μ) é dito um espaço de Lesbesgue se ele for isomorfo mod 0 ao intervalo $([0, 1], Leb)$, a um espaço contável de átomos (P, p), ou uma união disjunta de ambos.
Se (X, μ) for um espaço de Lebesgue que não tem parte atômica diremos que X é espaço de Lebesgue não atômico.

Informalmente falando, um espaço de Lesbesgue é um espaço de probabilidade consistindo de um intervalo e/ou um número contável de átomos.

Definição 3.2.3. Seja ξ uma partição finita de um espaço de medida (X, μ). Dizemos que uma propriedade ocorre para ε-quase todo elemento de ξ se a união dos elementos para os quais a propriedade falha tem medida μ no máximo ε.

3.2.1 Métrica no espaço de partições

Como discutimos acima, a idéia de se definir o que chamaremos de partições Bernoulli muito fracas é poder aproximar partições que tem as mesmas características das partições de um espaço de sequências Σ_k por cilindros. Mas afinal, o que é "aproximar" uma partição através de outras partições?

Para que possamos definir e trabalhar com este conceito é necessário introduzir uma estrutura métrica, uma distância, no espaço das partições. A partir deste ponto sempre que nos referirmos a uma partição estamos considerando que a partição é finita.

A distância usual

Dadas duas partições finitas α e β de um mesmo espaço de Lebesgue (X, μ), definimos a distância entre α e β por

$$d(\alpha, \beta) = \sum_{j=1}^{m} \mu(A_j \Delta B_j)$$

onde n é o número de elementos das partições α e β. O valor n pode ser considerado o mesmo para ambas partições pois, se α tem m elementos e β tem $n < m$ elementos, então tomamos $B_{n+1} = ... = B_m := \emptyset$.

Na definição da distância entre duas partições finitas o leitor deve observar que a ordem dos elementos da partição é crucial e influencia no cálculo da distância entre duas partições fixadas.

Exemplo 3.2.4. Considere $X = I = [0, 1]$ com a medida de Lebesgue *Leb*. Tome $\alpha = \{[0, 1/2), [1/2, 1]\}$ e $\beta = \{[0, 1/3), [1/3, 2/3), [2/3, 1]\}$. Então temos

$$d(\alpha, \beta) = Leb([0, 1/2)\Delta[0, 1/3)) + Leb([1/2, 1]\Delta[1/3, 2/3)) + Leb(\emptyset\Delta[2/3, 1]) =$$

$$= \left(\frac{1}{2} - \frac{1}{3}\right) + \left(1 - \frac{2}{3}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \frac{1}{3} = 1.$$

Exemplo 3.2.5. Consideremos agora $X = I$ e tome α a mesma partição do exemplo anterior porém tomaremos β com uma ordenação diferente. Considere

$$\beta = \{[2/3, 1], [0, 1/3), [1/3, 2/3)\}$$

Exercício 3.2.1. Verifique que no exemplo 3.2.5 temos

$$d(\alpha, \beta) = 2.$$

O exercício acima é para que fique claro que os elementos da partição estão ordenados e dependendo de como fazemos essa ordenação a distância entre as partições pode mudar.

Distribuições equivalentes

Definição 3.2.6 (Distribuições equivalentes)**.** Sejam $\xi_i = \{A_1^{(i)}, ..., A_m^{(i)}\}$ e $\eta_i = \{B_1^{(i)}, ..., B_m^{(i)}\}, 1 \leq i \leq n$ duas sequências de partições de dois espaços de medida (X, μ) e (Y, ν) respectivamente. Diremos que as sequências de partições $\{\xi_i\}$ e $\{\eta_i\}$ tem distribuições quivalentes, e escreveremos

$$\{\xi_i\}_1^n \sim \{\eta_i\}_1^n,$$

se, para cada $1 \leq k_i \leq m$ e $1 \leq i \leq n$, temos

$$\mu\left(\bigcap_{i=1}^{n} A_{k_i}^{(i)}\right) = \nu\left(\bigcap_{i=1}^{n} B_{k_i}^{(i)}\right).$$

Exercício 3.2.2. Se $\{\alpha_i\}_{i=1,...,n}$ é uma sequência de partições finitas de um espaço mensurável (X, μ) e $\theta : X \to X$ é uma transformação que preserva medida μ então as sequências de partições $\{\alpha_i\}_{i=1,...,n}$ e $\{\theta\alpha_i\}_{i=1,...,n}$ tem distribuições equivalentes, isto é,

$$\{\theta\alpha_i\}_1^n \sim \{\alpha_i\}_1^n.$$

A métrica $\bar{d}$

Definição 3.2.7. Sejam $\{\xi_i\}_1^n$ e $\{\eta_i\}_1^n$ partições de espaço de Lebesgue (X, μ) e (Y, ν) respectivamente. Dizemos que a distância entre $\{\xi_i\}_1^n$ e $\{\eta_i\}_1^n$ é menor ou igual a ε, e denotamos por

$$\bar{d}(\{\xi_i\}_1^n, \{\eta_i\}_1^n) \leq \varepsilon,$$

se existem sequências de partições $\bar{\xi}_i = \{\bar{A}_1^{(i)}, ..., \bar{A}_m^{(i)}\}$ e $\bar{\eta}_i = \{\bar{B}_1^{(i)}, ..., \bar{B}_m^{(i)}\}$ de um espaço de Lebesgue (Z, λ) tais que

$$\{\xi_i\}_1^n \sim \{\bar{\xi}_i\}_1^n \text{ e } \{\eta_i\}_1^n \sim \{\bar{\eta}_i\}_1^n,$$

e

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} \lambda(\bar{A}_j^{(i)} \triangle \bar{B}_j^{(i)}) \leq \varepsilon.$$

A definição da distância $\bar{d}$ é um tanto técnica e um pouco ausente de intuição se não estivermos em algum determinado contexto. Observe que na definição 3.2.7, as partições ξ_i não tem nenhuma relação entre si. Idem para o caso das partições η_i. Entretanto, ao longo do texto o leitor perceberá que a fim de obter informações sobre a dinâmica de uma transformação, será necessário trabalhar com iterados de uma determinada partição α fixada. Então as nossas partições ξ_i serão, na verdade, da forma $\xi_i = f^{-i}\alpha$. Façamos o mesmo para a sequência de partições η_i, isto é, escrevemos $\eta_i = f^{-i}\beta$ para uma certa partição β fixada. Tendo isto em mente e olhando novamente para a definição, podemos dizer que a distância $\bar{d}$ mede o quão distante estão duas partições α e β quando consideramos os iterados de ambas e tomamos a média das distâncias entre os iterados. Em outras palavras, $\bar{d}$ é como uma distância média ou uma distância assintótica.

Exercício 3.2.3. Considere os espaços de Lebesgue $(X = [0,1], Leb)$ e $(Y = \{1,2,3\}, \nu)$, onde a medida ν em Y é definida por:

$$\nu(\{1\}) = \nu(\{2\}) = \nu(\{3\}) = \frac{1}{3}.$$

Considere as partições $\alpha = \{[0,1/2),[1/2,1]\}$ de X e a partição $\beta = \{\{1\},\{2\},\{3\}\}$ de Y. Dado um $n \in \mathbb{N}$ qualquer, considere as sequências de partições $\{\xi_i\}_1^n$ e $\{\eta_i\}_1^n$ de X e Y respectivamente, definidas de forma constante igual a α e β respectivamente. Ou seja,

$$\xi_i = \alpha, \forall 1 \leq i \leq n,$$

$$\eta_i = \beta, \forall 1 \leq i \leq n.$$

Mostre que

$$\bar{d}(\{\xi_i\}_1^n, \{\eta_i\}_1^n) \leq 1.$$

3.2.2 Partições Bernoulli Muito Fracas e Teoremas de Ornstein

Antes de introduzirmos a definição de partições Bernoulli muito fracas, é importante ressaltar que a propriedade de ser isomorfo a um shift de Bernoulli pode ser vista de uma forma abstrata através de partições que satisfazem as propriedades de serem *geradoras* e *independentes*.

Definição 3.2.8 (Partição independente)**.** Uma partição $\alpha = \{A_1, A_2, ..., A_l\}$ de um espaço de medida (X, μ) é dita independente para um automorfismo

$f : X \to X$ μ-invariante se, para qualquer $n \in \mathbb{N}$ e para todas as escolhas de i_j's, com $-n \leq j \leq n$, temos

$$\mu\left(\bigcap_{j=-n}^{n} f^{-j}(A_{i_j})\right) = \prod_{j=-n}^{n} \mu(f^{-j}(A_{i_j}))$$

Definição 3.2.9 (Partição de Bernoulli)**.** Seja $(X, \mathcal{A}, \mu)$ um espaço de medida e $f : X \to X$ um automorfismo μ-invariante. Dizemos que uma partição α de X é uma partição de Bernoulli se α for geradora (veja a definição 2.3.6) e independente com respeito à f. Se um sistema $(X, \mathcal{A}, \mu, f)$ admitir uma partição de Bernoulli α então ele é dito um sistema de Bernoulli.

Exercício 3.2.4. Um automorfismo $f : (X, \mu) \to (X, \mu)$ admite uma partição de Bernoulli se, e somente se, ele é um automorfismo de Bernoulli, isto é, o sistema $(X, \mathcal{A}, \mu, f)$ é conjugado a um shift de Bernoulli.

Definição 3.2.10. Seja $f : X \to X$ um automorfismo de um espaço de medida (X, μ) preservando μ. Uma partição ξ de X é chamada Bernoulli muito fraca (BMF) para f se para qualquer $\varepsilon > 0$ existe $N_0 = N_0(\varepsilon)$ tal que para qualquer $N' \geq N \geq N_0, n \geq 0$, e ε-quase todo elemento $A \in \bigvee_{k=N}^{N'} f^k\xi$, temos

$$\bar{d}(\{f^{-i}\xi\}_1^n, \{f^{-i}\xi|A\}_1^n) \leq \varepsilon.$$

Enfatizamos que a partição $\xi|A$ é considerada com respeito à medida normalizada $\mu/\mu(A)$.

Teorema 3.2.11. *Uma partição de Bernoulli para um automorfismo f é BMF para f.*

Proof. Consideremos $\alpha = \{A_1, A_2, ..., A_l\}$ uma partição de Bernoulli. Tome $N' \geq N \geq 1$ quaisquer e $A \in \bigvee_N^{N'} f^k\alpha$ um elemento qualquer. Vamos provar que

$$\{f^{-i}\alpha\}_1^n \sim \{f^{-i}\alpha|A\}_1^n.$$

para qualquer n.

A sequência de partições $\{f^{-i}\alpha|A\}_1^n$ é uma sequência de partições do espaço de medida $(A, \mu/\mu(A))$. Então, queremos provar que dados quaisquer $1 \leq k_i \leq m$ e $1 \leq i \leq n$, temos

$$\mu\left(\bigcap_{i=1}^{n} A_{k_i}^{(i)}\right) = \frac{\mu\left(\bigcap_{i=1}^{n} A_{k_i}^{(i)} \cap A\right)}{\mu(A)},$$

onde $f^{-i}\alpha = \{A_1^{(i)},\ A_2^{(i)},\ ...,\ A_l^{(i)}\}$ e, consequentemente, $f^{-i}\alpha|A = \{A_1^{(i)} \cap A,\ A_2^{(i)} \cap A,\ ...,\ A_l^{(i)} \cap A\}$.

Como a partição α é de Bernoulli ela é independente. Logo pela definição de partição independente (veja 3.2.8) segue que

$$\mu\left(\bigcap_{i=1}^{n} A_{k_i}^{(i)} \cap A\right) = \mu\left(\bigcap_{i=1}^{n} A_{k_i}^{(i)}\right) \cdot \mu(A).$$

Portanto,

$$\mu\left(\bigcap_{i=1}^{n} A_{k_i}^{(i)}\right) = \frac{\mu\left(\bigcap_{i=1}^{n} A_{k_i}^{(i)} \cap A\right)}{\mu(A)},$$

como queríamos.

Ou seja, provamos que

$$\{f^{-i}\alpha\}_1^n \sim \{f^{-i}\alpha|A\}_1^n,$$

donde segue que α é BMF pela definição da distância $\bar{d}$. □

Os dois grandes Teoremas de Teoria de Ornstein que usaremos são os dois Teoremas a seguir. O primeiro teorema nos diz que se α é uma partição BMF para um automorfismo f, então f é um automorfismo de Bernoulli com respeito à σ-algebra gerada pelo join infinito de iterados da partição α. O segundo teorema complementa o primeiro, afirmando que para que um automorfismo seja de Bernoulli basta que encontremos uma sequência de partições refinadoras, cada uma delas sendo BMF, e de modo que o join infinito de todas as partições da sequência gere a σ-álgebra do sistema dado.

Teorema 3.2.12 ([7]). *Se α é uma partição* BMF *de X, então o sistema*

$$(X, \bigvee_{-\infty}^{+\infty} f^{-i}\alpha, \mu, f)$$

é Bernoulli.

Teorema 3.2.13 (veja [2], [6]). *Se $\alpha_1 \leq \alpha_2 \leq \alpha_3 \leq ...$ é uma sequência crescente de partições de um espaço de Lebesgue $(X, \mathcal{A}, \mu)$ tal que*

$$\bigvee_{i=1}^{+\infty} \bigvee_{n=-\infty}^{+\infty} f^{-n}\alpha_i = \mathcal{A}$$

e, para cada i, $(X, \bigvee_{n=-\infty}^{+\infty} f^{-n}\alpha_i, \mu, f)$ é um sistema de Bernoulli, então $(X, \mathcal{A}, \mu, f)$ é um automorfismo de Bernoulli.

3.2.3 Comparando distâncias entre partições através de seus nomes

Estimar a distância $\bar{d}$ entre duas sequências de partições não é uma tarefa muito fácil tendo em vista que a definição é bastante técnica e envolve conhecer todos os elementos dos iterados destas partições. Nesta seção apresentaremos uma forma de estimar a distância $\bar{d}$ entre duas sequências de partições de uma "forma mais dinâmica". Explicaremos superficialmente o significado intuitivo desta "forma dinâmica" a seguir.

Considere um automorfismo $f : X \to X$ e uma partição finita α de X. Dado um ponto $x \in X$, podemos associar a este ponto x uma sequência de números. Cada número nesta sequência representa o elemento da partição α que contém um determinado iterado de x. Assim, por exemplo, o número na posição 10 é exatamente o índice do elemento da partição α que contém $f^{10}(x)$. Isso nos da uma função $l(x)$, chamada de nome da partição (veja definição 3.2.14 abaixo), que associa, a cada elemento, $x \in X$ a sequência de números correspondente aos elementos de α pelos quais x passa. Mostraremos nos lemas 3.2.2 e 3.2.4 que para estimar a distância $\bar{d}$ entre duas sequências de partições, basta, num certo sentido, estimar a diferença entre seus nomes para um conjunto suficientemente grande de pontos.

Definição 3.2.14 (Nome de um ponto com respeito à uma partição)**.** Dada uma sequência de partições $\{\xi_i\}_1^n$ de um espaço de Lebesgue (X, μ), definimos a sequência de funções inteiras $l_i(x)$ pela condição $x \in A^{(i)}_{l_i(x)}$. Chamamos a sequência de funções $l_i(x)$ de o nome da sequência de partições $\{\xi_i\}_1^n$.

Definição 3.2.15. Dizemos que uma transformação $\theta : X \to Y$, entre espaços de Lebesgue (X, μ) e (Y, ν), ε-preserva medida se existe um conjunto $E \subset X$ tal que $\mu(E) \leq \varepsilon$ e para todo conjunto mensurável $A \subset X \setminus E$,

$$\left| \frac{\nu(\theta(A))}{\mu(A)} - 1 \right| \leq \varepsilon.$$

Defina a função $e : \mathbb{N} \to \{0, 1\}$ por $e(0) = 0$ e $e(n) = 1$ para todo $n > 0$.

Lema 3.2.1. *Sejam α e β duas partições finitas de X com nomes $l(x)$ e $m(x)$ respectivamente. Considere $f : X \to \{0, 1\}$ definida por*

$$f(x) = 1, \text{ se } l(x) \neq m(x)$$

$$f(x) = 0, \text{ se } l(x) = m(x).$$

Então
$$d(\alpha, \beta) = 2 \int f(x) d\mu(x).$$

Proof. Lembremos que a distância entre α e β é definida por
$$d(\alpha, \beta) = \sum_{j=1}^{m} \mu(A_j \Delta B_j)$$
onde m é o número de elementos de α e β.
Para cada $j = 1, ..., m$ temos:

- se $x \in A_j \cap B_j$ então $f(x) = 0$ pois neste caso $l(x) = j = m(x)$,
- se $x \in A_j \setminus B_j$ então $f(x) = 1$ pois neste caso $l(x) = j \neq m(x)$,
- se $x \in B_j \setminus A_j$ então $f(x) = 1$ pois neste caso $l(x) \neq j = m(x)$.

Então temos:
$$\begin{aligned}\int_{A_j} f(x) d\mu + \int_{B_j} f(x) d\mu &= \int_{A_j \setminus B_j} f(x) d\mu + 2 \cdot \int_{A_j \cap B_j} f(x) d\mu + \int_{B_j \setminus A_j} f(x) d\mu \\ &= \int_{A_j \setminus B_j} 1 d\mu + 2 \cdot 0 + \int_{B_j \setminus A_j} 1 d\mu \\ &= \mu(A_j \setminus B_j) + \mu(B_j \setminus A_j) \\ &= \mu(A_j \Delta B_j).\end{aligned}$$

Logo,
$$\begin{aligned}d(\alpha, \beta) &= \sum_{j=1}^{m} \mu(A_j \Delta B_j) = \sum_{j=1}^{m} \int_{A_j} f(x) d\mu + \sum_{j=1}^{m} \int_{B_j} f(x) d\mu \\ &= \int_{\bigcup_{j=1}^{m} A_j} f(x) d\mu + \int_{\bigcup_{j=1}^{m} B_j} f(x) d\mu = 2 \cdot \int f(x) d\mu,\end{aligned}$$
como queríamos demonstrar.

□

O seguinte lema provado por D. Ornstein e B. Weiss em [2] é o que faz possível uma abordagem geométrica para provar a propriedade BMF.

Lema 3.2.2 (Ornstein - Weiss [2]). *Sejam $\{\alpha_i\}_{i=1}^{n}$ partições de X com nomes $l_i(x)$, e $\{\beta_i\}_1^n$ partições de Y com nomes $m_i(y)$. Se existe uma aplicação mensurável $\theta : X \to Y$ e um conjunto $E \subset X$ tal que*

1. $\mu(E) \leq \varepsilon$,

2.
$$\frac{1}{n}\sum e(l_i(x) - m_i(\theta(x)) \leq \varepsilon, \quad x \in X \setminus E,$$

então
$$\bar{d}(\{\alpha_i\}_1^n, \{\beta_i\}_1^n) \leq 4\varepsilon.$$

Proof. Tome $\overline{\alpha_i} := \theta\alpha_i, \overline{\beta_i} := \beta_i$, como θ preserva medida temos

$$\{\theta\alpha_i\}_1^n \sim \{\alpha_i\}_1^n,$$

ou seja,
$$\overline{\{\alpha_i\}}_1^n \sim \{\alpha_i\}_1^n.$$

Para cada i defina

$$f_i(y) := e(l_i(\theta^{-1}(y)) - m_i(y)).$$

Se o $\overline{\alpha_i}$- nome de y é diferente do $\overline{\beta_i}$- nome de y então $f_i(y) = 1$. Se são iguais então $f_i(y) = 0$. Pelo Lema 3.2.1 segue que

$$d(\overline{\alpha}_i, \overline{\beta}_i) = 2 \cdot \int e(l_i(\theta^{-1}(y)) - m_i(y))d\mu.$$

Portanto

$$\frac{1}{n}\sum_{i=1}^{n} d(\overline{\alpha}_i, \overline{\beta}_i) = \frac{2}{n}\sum_{i=1}^{n}\int e(l_i(\theta^{-1}(y)) - m_i(y))d\mu$$
$$= 2 \cdot \int \frac{1}{n}\sum_{i=1}^{n} e(l_i(\theta^{-1}(y)) - m_i(y))d\mu$$
$$= 2\int_{X\setminus E} \frac{1}{n}\sum_{i=1}^{n} e(l_i(\theta^{-1}(y)) - m_i(y))d\mu + 2\int_E \frac{1}{n}\sum_{i=1}^{n} e(l_i(\theta^{-1}(y)) - m_i(y))d\mu$$
$$\leq 2 \cdot \int_{X\setminus E} \varepsilon d\mu + 2 \cdot \int_E 1 d\mu$$
$$\leq 2 \cdot \varepsilon + 2 \cdot \varepsilon = 4 \cdot \varepsilon.$$

Da definição de $\bar{d}$ concluímos que

$$\bar{d}(\{\alpha_i\}_1^n, \{\beta_i\}_1^n) \leq 4\varepsilon.$$

□

O Lema 3.2.2 pode ser melhorado um pouco trocando a hipótese de que θ preserva medida pela hipótese de que θ ε-preserva medida, mas para isso precisamos demostrar que em um certo sentido uma aplicação que ε-preserva medida pode ser aproximada, exceto em um conjunto muito pequeno, por uma que preserva medida. O próximo lema demonstra este fato e a prova é um argumento clássico de teoria da medida. A demonstração que apresentamos é inspirada na demonstração dada por N. Chernov e C. Haskell [1].

Lema 3.2.3 (veja [1] Lema 4.3)**.** *Sejam (X, μ) e (Y, ν) espaços de Lebesgue não atômicos. Seja $\theta : (X, \mu) \to (Y, \nu)$ uma transformação que ε- preserva medida, e $\alpha = \{A_1, ..., A_k\}$ uma partição finita de X, então existe uma aplicação $\overline{\theta} : X \to Y$ tal que*

$$\{\overline{\theta}\alpha\} \sim \{\alpha\}$$

e

$$\mu(\{x \in X : \theta x \neq \overline{\theta} x\}) \leq 8 \cdot \varepsilon.$$

Proof. Como θ ε-preserva medida, por definição podemos tomar um conjunto $E_2 \subset X$ tal que $\mu(E_2) \leq \varepsilon$ e para todo conjunto mensurável $A \subset X \backslash E_2$ temos

$$\left| \frac{\nu(\theta(A))}{\mu(A)} - 1 \right| \leq \varepsilon.$$

Para qualquer conjunto mensurável $S \subset X$ denote $\overline{S} := S \backslash E_2$. Então temos

$$\left| \nu(\theta(\overline{S})) - \mu(\overline{S}) \right| < \mu(\overline{S}) \cdot \varepsilon.$$

Logo,

$$\sum_j \nu(\theta(\overline{A}_j)) < 1 + \varepsilon, \tag{3.1}$$

e

$$\nu(\bigcup_j \theta(\overline{A}_j)) > 1 - 2 \cdot \varepsilon. \tag{3.2}$$

Consideremos $B := \theta^{-1}(\bigcup_{j \neq k}(\theta(\overline{A}_j) \cap \theta(\overline{A}_k)))$. Então temos

$$\sum_j \nu(\theta(\overline{A}_j)) - \nu(\bigcup_j \theta(\overline{A}_j)) < 1 + \varepsilon - 1 + 2 \cdot \varepsilon = 3 \cdot \varepsilon.$$

Mas note que

$$\nu(\bigcup_{j \neq k}(\theta(\overline{A}_j) \cap \theta(\overline{A}_k))) = \sum_j \nu(\theta(\overline{A}_j)) - \nu(\bigcup_j \theta(\overline{A}_j)),$$

portanto

$$\nu(\bigcup_{j \neq k}(\theta(\overline{A}_j)) < 3 \cdot \varepsilon.$$

Logo, pelo fato que θ ε-preserva medida segue que

$$\mu(\overline{B}) < 3 \cdot \varepsilon \cdot (1+\varepsilon) < 6\varepsilon.$$

Agora considere os conjuntos $\tilde{A}_j := A_j \setminus (B \cup E_2)$. Note que $\theta(\tilde{A}_j) \cap \theta(\tilde{A}_k) = \emptyset$, para todos $j \neq k$ e que

$$\mu(\bigcup_j \tilde{A}_j) = \mu(\Big(\bigcup A_j\Big) \setminus (B \cap E_2)) = 1 - \mu(B \cup E_2) > 1 - 7 \cdot \varepsilon.$$

A construção de $\overline{\theta}$ será feita da seguinte forma: primeiro modificaremos θ em uma pequena parte de cada conjunto $\tilde{A}_j$ para garantir que cada a imagem de $\tilde{A}_j$ tem a mesma medida que $\tilde{A}_j$ e de forma que a imagem de $\tilde{A}_j$ seja disjunta da imagem de $\tilde{A}_i$ para todos $i \neq j$. O conjunto $B \cup E_2$ nós mapeamos no resto de Y preservando medida.

Passo 1: Modificar θ ligeiramente nos $\tilde{A}_j$. Tome $j \in \{1, 2, ..., k\}$. Suponhamos que $\nu(\theta(\tilde{A}_j)) > \mu(\tilde{A}_j)$. Consideremos um conjunto $G_j \subset \tilde{A}_j$ que tenha a propriedade que

$$\nu(\theta(G_j)) = \mu(\tilde{A}_j).$$

Então temos

$$\mu(G_j) > (1-\varepsilon)\nu(\theta(G_j)) = (1-\varepsilon)\mu(\tilde{A}_j).$$

Tome qualquer conjunto de medida zero $P_j \subset Y$ e escolha um ponto arbitrário $p_j \in Y$. Defina $\overline{\theta}(x) := \theta(x)$ para todo $x \in G_j$ e defina $\overline{\theta}(x) := p_j$ para todo $x \in \tilde{A}_j \setminus G_j$.

Observe que na definição de $\overline{\theta}$ em $\tilde{A}_j$, a função $\overline{\theta}$ preserva a medida de $\tilde{A}_j$ e é diferente de θ em apenas um sub-conjunto $\tilde{A}_j$ de medida bem pequena.

Agora suponhamos que $\nu(\theta(\tilde{A}_j)) < \mu(\tilde{A}_j)$. Considere um conjunto $G_j \subset \tilde{A}_j$ cuja medida é maior que $(1-\varepsilon)\mu(\tilde{A}_j)$. Defina $\overline{\theta}(x) := \theta(x)$ para todo $x \in G_j$, e de forma que $\overline{\theta}(\tilde{A}_j \setminus G_j)$ tem medida $\mu(\tilde{A}_j) - \nu(\theta(G_j)$ e não intersecta a imagem de nenhum dos outros $\tilde{A}_j$'s.

Passo 2: Consideremos agora o conjunto $B \cup E_2$. Este é o único lugar onde ainda não definimos $\overline{\theta}$. Observe que, pela forma que definimos $\overline{\theta}$ até o momento temos

$$\nu\left(\overline{\theta}\left(\bigcup A_j \setminus (B \cup E_2)\right)\right) = \mu\left(\bigcup A_j \setminus (B \cup E_2)\right),$$

pois $\overline{\theta}$ preserva a medida de cada $\tilde{A}_j$ e as imagens destes são mutualmente disjuntas. Então sobrou em X o conjunto $B \cup E_2$ e em Y um conjunto também de medida $B \cup E_2$. Então podemos definir $\overline{\theta}$ em $B \cup E_2$ de forma que ela seja injetora e mande $B \cup E_2$ no conjunto que sobrou em Y.

Exercício 3.2.5. Conclua a demonstração do Lema verificando que a aplicação $\overline{\theta}$ definida acima satisfaz o enunciado do Lema.

$\square$

Lema 3.2.4 (Ornstein-Weiss, [2]). *Sejam $\{\alpha_i\}_1^n$ e $\{\beta_i\}_1^n$ duas sequências de partições de espaços de Lebesgue (X, μ) e (Y, ν) com nomes $l_i(x)$ e $m_i(x)$ respectivamente. Assuma que existe uma transformação que ε-preserva medida $\theta : X \to Y$, satisfazendo*

$$\frac{1}{n}\sum_{i=1}^{n} e(l_i(x) - m_i(\theta(x))) \leq \varepsilon$$

para todo $x \in X$ com possivel exceção para um conjunto $E \subset X$ com $\mu(E) \leq \varepsilon$. Então

$$d(\{\alpha_i\}_1^n, \{\beta_i\}_1^n) \leq 36 \cdot \varepsilon.$$

Exercício 3.2.6. Demonstre o Lema 3.2.4.

Dica: Tome $\alpha := \bigvee_1^n \alpha_i$ e aplique os Lemas 3.2.3 e 3.2.2.

3.3 Automorfismos ergódicos de $\mathbb{T}^2$ são Bernoulli

Como vimos no exemplo 2.1.4, um automorfismo ergódico do toro $f : \mathbb{T}^3 \to \mathbb{T}^3$ é induzido por uma matriz inteira 2×2 $A : \mathbb{R}^2 \to \mathbb{T}^2$ sem autovalores unitários. Então podemos considerar λ_1, λ_2 os autovalores de A onde

$$|\lambda_1| < 1 < |\lambda_2|.$$

Associados aos autovalores, temos duas direções de autoespaços, isto é, duas retas (em $\mathbb{R}^2$, passando pela origem) F^s e F^u associadas respectivamente a λ_1 e λ_2. Observe que como $|\lambda_1| < 1$ então F^s apresenta uma característica de contração, isto é, dados $x, y \in F^s$ temos:

$$d(Ax, Ay) = |\lambda_1| \cdot d(x, y).$$

O mesmo acontece para F^u com respeito à expansão de razão $|\lambda_2|$. Para cada ponto $x \in \mathbb{R}^2$ podemos definir as retas $F^s(x)$ e $F^u(x)$ por:

$$F^s(x) = x + F^s,$$

$$F^u(x) = x + F^u.$$

Seja $\pi : \mathbb{R}^2 \to \mathbb{T}^2$ a projeção natural, definimos:

$$\mathcal{F}^s(x) = \pi \circ F^s(x),$$

$$\mathcal{F}^u(x) = \pi \circ F^u(x),$$

ou seja, $\mathcal{F}^s(x)$ e $\mathcal{F}^u(x)$ são as projeções de $F^s(x)$ e $F^u(x)$ no 2-toro.

Notação: Dado um conjunto $J \subset M$, denotaremos por $\mathcal{F}^s(x) \cap J$ a componente conexa da interseção $\mathcal{F}^s(x) \cap J$ que contém o ponto x. Similar para $\mathcal{F}^u(x) \cap J$.

3.3.1 Teorema principal

Nesta seção demonstraremos o Teorema principal destas notas.

Teorema 3.3.1. *Seja $f : \mathbb{T}^2 \to \mathbb{T}^2$ um automorfismo ergódico com respeito à medida de Haar de $\mathbb{T}^2$. Então f é um automorfismo de Bernoulli.*

Rascunho da prova

Para provar que um automorfismo ergódico de $\mathbb{T}^2$ é Bernoulli tomaremos uma sequência de partições cada vez mais finas

$$\alpha_1 \leq \alpha_2 \leq ...$$

que converge para a partição por pontos, e provaremos que cada partição nesta sequência é BMF. Depois, utilizaremos os Teoremas 3.2.12 e 3.2.13 para concluir que os automorfismos em questão são de fato Bernoulli.

Antes de passarmos para a prova formal vamos descrever um rascunho de como a prova é feita. Inicialmente tomaremos uma partição α onde os elementos tem fronteiras suaves por partes (veja definição 3.3.3). Para provar que α é uma partição BMF utilizaremos, como controle, uma partição β por retângulos (veja definição 3.3.2) muito pequenos.

A estratégia utilizada para provar que α é BMF é estimar a distância $\bar{d}$ entre as sequências de partições $\{f^{-i}\alpha\}$ e $\{f^{-i}\alpha|A\}$, para muitos átomos A de um refinamento futuro de α. Esta estimativa será feita construindo uma aplicação $\theta : X \to A$ que satisfaça as condições do lema 3.2.4 e, portanto, conseguiremos a estimativa desejada.

Para construir θ inicialmente iteramos os elementos de α pelo automorfismo induzido pela matriz ergódica $A : \mathbb{R}^2 \to \mathbb{R}^2$. Como a matriz A tem uma direção de expansão e uma de contração, é de se esperar que os elementos A_i de α comecem a se esticar na direção da expansão e passem a cruzar totalmente os retângulos da partição β. Este fato é rigorosamente demonstrado no lema 3.3.1.

Uma vez que ocorre esse cruzamento, a interseção dos iterados $f^j A_i$ dos elementos de α com os retângulos de β, geram sub-retangulos. A idéia então é construir, para cada retângulo de β, uma função θ que vai do sub-retângulo ao retângulo maior e que tenha as boas propriedades exigidas no lema 3.2.4.

Uma das boas propriedades exigidas de θ é que dado um ponto x, os iterados de x e de $\theta(x)$ tenham nomes muito próximos. Em particular, se garantirmos que x e $\theta(x)$ estão muito próximos podemos esperar que seus nomes também estejam. Assim, construiremos θ de forma que $\theta(x) \in \mathcal{F}^s(x)$ pois, como $\mathcal{F}^s(x)$ tem característica contratora, os iterados de x e de $\theta(x)$ estão cada vez mais próximos. Esta construção é feita rigorosamente no lema 3.3.2.

Feito esta construção local de θ, utilizados a propriedade de Kolmogorov para grudar as funções θ definidas em cada retângulo. Utilizamos então o lema 3.2.4 para estimar a distância $\bar{d}$ entre as sequências de partições envolvidas e concluímos que α é BMF.

3.3.2 Demonstração do Teorema 3.3.1

Definição 3.3.2 (Retângulo)**.** Um conjunto $\Pi \subset \mathbb{T}^2$ é chamado de retângulo se ele satisfazer as seguintes condições:

1. Π é conexo;
2. $\overline{\Pi} = \overline{\text{Int}(\Pi)}$;
3. para quaisquer $x, y \in \Pi$, temos
$$(\mathcal{F}^s(x) \cap \Pi) \cap (\mathcal{F}^u(y) \cap \Pi) = z \in \Pi.$$

Definição 3.3.3. Seja $\alpha = \{A_1, ..., A_a\}$ uma partição de M. Dizemos que α é FS (esta abreviação vem de "fronteira suave"), se

1. a fronteira de cada átomo A_i é união finita de curvas suaves,
2. $\overline{A_i} = \overline{\text{Int}(A_i)}$.

Definição 3.3.4. Seja Π um retângulo. Dado um conjunto $E \subset \mathbb{T}^2$, dizemos que E intersecta Π de forma u-tubular se para qualquer $x \in \Pi \cap E$ temos $\mathcal{F}^u(x) \cap \Pi \subset \mathcal{E} \cap \Pi$.

Lema 3.3.1. *Suponha que α é uma partição FS, Π é um retângulo e $\delta > 0$ um real positivo dado. Existe um N_1 tal que para quaisquer $N' > N \geq N_1$ e δ-quase todo átomo $A \in \bigvee_N^{N'} f^k\alpha$ existe um subconjunto $E \subset A$ com*

1. $$\frac{m(E)}{m(A)} > 1 - \delta,$$
2. *E intersecta Π em um subconjunto u-tubular.*

Proof. Observe que pela invariância das folheações podemos observar que a u-tubularidade de conjuntos não é afetada ao aplicar f.

Denotemos por G_k as interseções não tubulares de Π com $f^k(A)$, mais especificamente:

$$G_k = \{y \in \Pi \cap f^k(A) : \mathcal{F}^u(y) \cap \Pi \not\subset \Pi \cap f^k(A)\}$$

Para estimar o tamanho de G_k é conveniente aplicar f^{-k}.

Afirmação: Para uma certa constante $C > 0$, dado qualquer $y \in f^{-k}(G_k)$ temos

$$d(y, \partial A) \leq C \cdot |\lambda_2|^{-k}.$$

prova: Consideremos $y \in f^{-k}(G_k)$ arbitrário. Como y está na interseção não tubular temos que $\mathcal{F}^u(y) \cap f^{-k}(\Pi) \not\subset f^{-k}(\Pi) \cap A$. Como $y \in A$ e A é conexo então a reta $\mathcal{F}^u(y)$ intersecta a fronteira de A em um ponto que chamaremos de z e também intersecta a fronteira de $f^{-k}(\Pi)$ em um certo ponto, que chamaremos de w, de forma que

$$d(y, w) = d(y, z) + d(z, w)$$

pois a interseção é não tubular. Como Π é um retângulo e a folheação $\mathcal{F}^u$ contrai exponencialmente com potências negativas de f, temos

$$d(y, w) \leq d(f^k(y), f^k(w)) \cdot |\lambda_2|^{-k} \leq \mathrm{diam}(\Pi) \cdot |\lambda_2|^{-k}.$$

Então basta considerar $C = \mathrm{diam}(\Pi)$, concluindo a prova da afirmação.

$\triangle$

Uma vez que a fronteira de A é suave por partes segue que:

$$m(G_k) \leq C \cdot |\lambda_2|^{-k}.$$

Considere N_1 grande o bastante para que

$$m(G) \leq \sum_{i=N_1}^{+\infty} m(G_k) \leq \delta^2,$$

onde $G := \bigcup_{N_1}^{+\infty} G_k$. Consideremos

$$\Omega := \left\{ B \in \bigvee_{N}^{N'} f^k\alpha : \frac{m(L \cap G)}{m(L)} > \delta \right\}$$

e tome

$$L = \bigcup_{B \in \Omega} B.$$

Suponha que $m(L) > \delta$. Note que, como L é união disjunta de conjuntos temos

$$m(L) = \sum_{B \in \Omega} m(B).$$

Portanto

$$\delta^2 > m(G) \geq m(G \cap L) = \sum_{B \in \Omega} m(G \cap B) > \sum_{B \in \Omega} \delta \cdot m(B) = \delta \cdot m(L) > \delta^2,$$

caindo em contradição. Ou seja, δ- quase todo átomo $B \in \bigvee_N^{N'} f^k \alpha$ intersecta G em um conjunto de medida relativa no máximo δ. Finalmente, para cada átomo $A \in \bigvee_N^{N'} f^k \alpha$ com $\alpha \notin \Omega$, tome $E = A \cap G^c$. E é tubular (pois todas as interseções não tubulares estão dentro de G) e $m(E)/m(A) > 1 - \delta$ pois para $A \notin \Omega$ temos (por definição de Ω)

$$\frac{m(A \cap G)}{m(A)} \leq \delta \Rightarrow \frac{m(A \cap G^c)}{m(A)} > 1 - \delta.$$

□

Lema 3.3.2. *Dado $\delta_1 > 0$ existe $\delta_2 > 0$ tal que se Π é um retângulo de diâmetro menor que δ_2 e E é um subconjunto u-tubular de Π, existe uma aplicação bijetora $\theta : E \to \Pi$ tal que*

1. *θ preserva medida,*
2. $$d(f^k \theta x, f^k x) < \delta_1,$$

   para todo $k \geq 0$, $x \in E$.

Proof. Se δ_2 for pequeno o bastante então o segundo item será satisfeito se garantirmos $\theta(x) \in \mathcal{F}^s(x) \cap \Pi$, pois f contrai distâncias em $\mathcal{F}^s$. Fixemos um ponto x_0 no interior de Π,

$$x_0 \in \operatorname{Int}(\Pi).$$

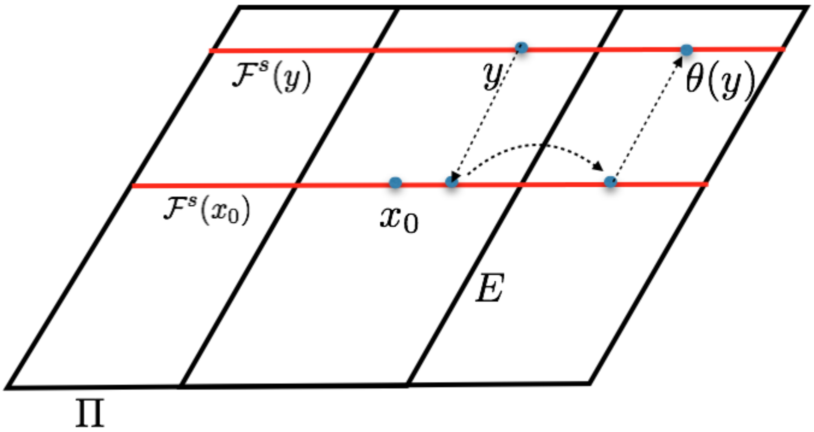
Agora consideremos os segmentos de reta $E \cap \mathcal{F}^s(x_0) \cap \Pi$ e $\mathcal{F}^s(x_0) \cap \Pi$. Defina $\theta_0 : E \cap \mathcal{F}^s(x_0) \cap \Pi \to \mathcal{F}^s(x_0) \cap \Pi$ a aplicação linear e bijetora entre essas retas, levando extremos em extremos.
Observe que essa transformação preserva a medida de Lebesgue normalizada. Agora, definiremos θ usando a θ_0 já definida.

Seja $x \in \Pi$ qualquer, existe uma aplicação natural $\pi_{x_0,x} : \mathcal{F}^s(x_0) \cap \Pi \to \mathcal{F}^s(x) \cap \Pi$ dada por

$$\pi_{x_0,x}(y) = (\mathcal{F}^u(y) \cap \Pi) \cap (\mathcal{F}^s(x_0) \cap \Pi).$$

Observe que essa transformação, que é apenas uma projeção paralela, preserva a medida de Lebesgue linear. Agora, definimos θ da seguinte forma: para cada $x \in E$, dado $y \in \mathcal{F}^s(x) \cap \Pi$ definimos

$$\theta(y) := \pi_{x_0,x} \circ \theta_0 \circ \pi_{x_0,x}^{-1}(y).$$



Observe que essa definição apenas é possível porque E é um conjunto u-tubular em Π. Pelo Teorema de Fubini, a aplicação θ definida em $E \cap \Pi$ preserva medida. $\square$

Lema 3.3.3. *Sejam $\varepsilon, \varepsilon' > 0$ e α uma partição FS. Então existe um N tal que para quaisquer $N' \geq N$ e para ε-quase todo átomo $A \in \bigvee_N^{N'} f^k\alpha$ existe um conjunto $E \subset A$, e uma função $1-1$ $\theta : E \to M$ tal que*

1. $$\frac{m(E)}{m(A)} > 1 - \varepsilon,$$

2. *θ $100 \cdot \varepsilon$-preserva medida,*

3. $$d(f^k\theta x, f^k x) < \varepsilon',$$

   para todo $k \geq 0$, $x \in E$.

Proof. Dado $\varepsilon' > 0$ fixo, pelo Lema 3.3.2 (aplicado para $\delta_1 := \varepsilon'$), existe $\delta_2 > 0$ tal que se Π é um retângulo de diâmetro menor que δ_2 e E é um subconjunto u-tubular de Π, existe uma aplicação bijetora $\theta : E \to \Pi$ satisfazendo as propriedades (1) e (2) no Lema 3.3.2. Tomemos então uma partição $\beta := \{\Pi_0, ..., \Pi_b\}$ onde $\Pi_1, ..., \Pi_b$ são retângulos com diâmetro menor ou igual a δ_2 e $m(\Pi_0) < \varepsilon/10$. Tais partições sempre existem, basta tomar, por exemplo, as partições dadas no exemplo 2.3.5.

Consideremos

$$\gamma := \varepsilon \cdot b^{-1} \cdot \min\{m(\Pi_i) : 1 \le i \le b\}.$$

Pelo Lema 3.3.1, para cada $1 \le i \le b$ podemos encontrar $N_1^i > 0$ grande o bastante de forma que, para todos $N > N' \ge N_1^i$ e para γ-quase todo átomo A de $\bigvee_N^{N'} f^k\alpha$ existe um conjunto $E \subset A$ tal que

$$\frac{m(E)}{m(A)} > 1 - \gamma > 1 - \varepsilon$$

e E intersecta Π_i de forma u-tubular. Tomando $N_1 = \max\{N_1^i : 1 \le i \le b\}$ temos que para ε-quase todo átomo A de $\bigvee_N^{N'} f^k\alpha$, para cada $1 \le i \le b$ existe um conjunto $E_i \subset A$ tal que

$$\frac{m(E_i)}{m(A)} > 1 - \gamma > 1 - \varepsilon$$

e E_i intersecta Π_i de forma u-tubular para todo $1 \le i \le b$. Defina

$$E = \bigcup (E_i \cap \Pi_i).$$

Exercício 3.3.1. Mostre que E intersecta cada Π_i de forma u-tubular, $E \cap \Pi_0 = \emptyset$ e

$$\frac{m(E)}{m(A)} > 1 - \gamma.$$

Agora, pelo Teorema 3.1.10 (e consequentemente pelo Corolário 3.1.1) sabemos que f é Kolmogorov. Logo, pelo Lema 2.4.2 segue que toda partição finita, em particular a partição α, é Kolmogorov. Pela definição de partição de Kolmogorov (ver Definição 2.4.1) segue que: para cada $1 \le i \le b$, dado $\xi_i > 0$ existe $N_0^i > 0$ tais que para quaisquer $N' > N \ge N_0^i$ e ξ_i-quase todo elemento $A \in \bigvee_{k=N}^{N'} f^k\alpha$ temos

$$\left| \frac{m(A \cap \Pi_i)}{m(A)} - m(\Pi_i) \right| \le \xi_i.$$

Tomando ξ_i pequenos o suficiente e $N_0 = \max\{N_0^i : 1 \le i \le b\}$ temos que para quaisquer $N' > N \ge N_0$ e γ-quase todo elemento $A \in \bigvee_{k=N}^{N'} f^k\alpha$ temos um conjunto $E \subset A$ intersectando cada Π_i de forma u-tubular, satisfazendo

$$\frac{m(E)}{m(A)} > 1 - \gamma$$

e

$$\sum_{i=1}^{b}\left|\frac{m(E\cap\Pi_i)}{m(E)}-m(\Pi_i)\right|\leq 20\cdot\gamma. \tag{3.3}$$

Para cada $1\leq i\leq b$ defina $E_i := E\cap\Pi_i$. Pelo Lema 3.3.2 , para cada $1\leq i\leq b$ podemos definir uma aplicação bijetora e que preserva medida

$$\theta_i : E_i \to X$$

satisfazendo

$$d(f^k(\theta(x)), f^k(x)) < \varepsilon \tag{3.4}$$

para todo $x\in E$, $k\geq 0$. Defina agora a aplicação $\theta : E\to X$ da seguinte forma

$$\theta(x)=\theta_i(x) \text{ se } x\in E_i.$$

O item (3) é claramente satisfeito devido a (3.4). O item (1) também seque como consequência da nossa escolha de E. Basta provar o item (2).

Tome $B\subset E$ um conjunto mensurável. Como $\theta_i : E_i\to\Pi_i$ preserva medida temos

$$\frac{m(\theta(B)\cap\Pi_i)}{m(\Pi_i)}=\frac{m(B\cap\Pi_i)}{m(E\cap\Pi_i)}$$

Agora note que

$$\left|\frac{m(\theta_i(B_i))m(E)}{m(B_i)}-1\right|=$$

$$=\left|\frac{m(\theta_i(B_i))m(E_i)}{m(B_i)m(\Pi_i)}-1+\frac{m(\theta_i(B_i))m(E_i)}{m(B_i)m(\Pi_i)}\cdot\left(\frac{m(E)m(\Pi_i)}{m(E_i)}-1\right)\right|$$
$$=\left|\frac{m(E)m(\Pi_i)}{m(E_i)}-1\right|.$$

Exercício 3.3.2. Conclua que θ satisfaz a propriedade (2), isto é, que θ $100\cdot\varepsilon$-preserva medida.

□

Lema 3.3.4. *Dada uma partição FS α e $\varepsilon>0$, existe N tal que para todo $N'\geq N$ e ε-quase todo átomo $A\in\bigvee_N^{N'} f^k\alpha$*

$$d(\{f^{-i}\alpha|A\}_1^n, \{f^{-i}\alpha\}_1^n)\leq\varepsilon$$

para todo $n\geq 1$. Em outras palavras, qualquer partição FS é BMF.

Proof. Fixe $\varepsilon > 0$, $0 < \varepsilon' < \varepsilon$ e escolha N_2 como no Lema 3.3.3 e escolha $N_2 := N$ como no Lema 3.3.3 e Y o conjunto exceção obtido, isto é:

$$m(Y) \leq 100 \cdot \varepsilon, Y = \text{ união de elementos de } \bigvee_{k=N}^{N'} f^k \alpha,$$

onde $N' > N \geq N_2$. Considere um elemento $A \in \bigvee_{k=N}^{N'} f^k \alpha$. Considere $E \subset A$ e $\theta : E \to M$ como construidos no Lema 3.3.3. Vamos calcular a distância $\bar{d}$ entre as sequências de partições $\{f^{-i}\alpha\}_1^n$ e $\{f^{-i}\alpha|A\}_1^n$.

Tome um ponto arbitrário $y_0 \in A$ e defina $\overline{\theta} : A \to M$ por

$$\overline{\theta}(x) = \theta(x), \text{ se } x \in E$$

$$\overline{\theta}(x) = y_0, \text{ se } x \notin E.$$

Lembremos que $m(A \setminus E) < 2\varepsilon$. Observe que se $x \in E$ então, pelo Lema 3.3.3 temos

$$d(f^k \theta x, f^k x) < \varepsilon'$$

logo, se $e(l_i(x) - m_i(\overline{\theta}(x))) = 1$ então $f^i(x)$ e $f^i(\overline{\theta}(x))$ estão em átomos diferentes de α, o que implica

$$f^i(x) \in O_{\varepsilon'}(A_{l_i(x)}).$$

Defina

$$O_{\varepsilon'} := \bigcup_{i=1}^{k} O_{\varepsilon'}(A_i).$$

Pela definição da função e, se $x \in E$ temos

$$\frac{1}{n}\sum_{i=1}^{n} e(l_i(x) - m_i(\overline{\theta}(x))) \leq \frac{1}{n}\sum_{i=1}^{n} \chi_{O_{\varepsilon'}}(f^i(x)).$$

Pelo Teorema Ergódico de Birkhoff, como f é ergódica sabemos que o lado direito converge para quase todo ponto, isto é, para quase todo ponto $x \in E$ temos

$$\frac{1}{n}\sum_{i=1}^{n} e(l_i(x) - m_i(\overline{\theta}(x))) \leq \frac{1}{n}\sum_{i=1}^{n} \chi_{O_{\varepsilon'}}(f^i(x)) \to m(O_{\varepsilon'}).$$

Agora, como α é partição onde todos os elementos tem fronteiras suaves por partes, sabemos que

$$m(O_{\varepsilon'}) \to 0$$

quando $\varepsilon' \to 0$. Logo, tomando $\varepsilon' > 0$ arbitrariamente pequeno podemos garantir que para quase todo ponto $x \in E$ temos

$$\frac{1}{n}\sum_{i=1}^{n} e(l_i(x) - m_i(\overline{\theta}(x))) \leq \varepsilon.$$

Então do Lema 3.2.4 obtemos

$$d(\{f^{-i}\alpha\}_1^n, \{f^{-i}\alpha|A\}_1^n) \leq 10000 \cdot \varepsilon.$$

Como $\varepsilon > 0$ é arbitrário concluímos que α is BMF.

$\square$

Bibliography

[1] N. Chernov and C. Haskell. Nonuniformly hyperbolic k-systems are bernoulli. *Ergodic Theory and Dynamical Systems*, 16:19–44, 1996.

[2] D.Ornstein and B.Weiss. Geodesic flows are bernoullian. *Israel Journal of Mathematics*, 14(2):Hebrew University Magnes Press—198, 1973.

[3] S. Kalikow and R. Mccutcheon. *An Outline of Ergodic Theory.* Cambridge University Press, 2010.

[4] K. Oliveira and M. Viana. *Fundamentos da teoria ergódica*, volume 1. SBM, 2014.

[5] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Advances in Math.*, (4):337–352, 1970.

[6] D. Ornstein. Imbedding bernoulli shifts in flows, contributions to ergodic theory and probability. *Lecture Notes in Math, Springer Berlin*, pages 178–218, 1970.

[7] D. Ornstein. Two bernoulli shifts with infinite entropy are isomorphic. *Advances in Mathematics*, (5):339–348, 1970.

[8] M. S. Pinsker. Dynamical systems with completely positive and zero entropy. *Dokl. Akad. Nauk SSSR (N.S.)*, 133:1025–1026, 1960.

[9] G. Ponce, A. Tahzibi, and R. Varão. Bernoulli property for partially hyperbolic diffeomorphisms on the 3-torus. in preparation.

[10] M. Ratner. Anosov flows with gibbs measures are also bernoullian. *Israel Journal of Mathematics*, 17:380–391, 1974.

[11] V. A. Rohlin and Y. Sinai. Construction and properties of invariant measurable partitions. *Doklady Akademii Nauk SSSR*, 141:pp. 1038–1041, 1961 (Russian); Soviet Mathematics, vol. 2, pp. 1611–1614, 1961 (English)., 1971.

[12] V. A. Rokhlin. Metric properties of endomorphisms of compact commutative groups. *Izv. Akad. Nauk SSSR Ser. Mat.*, 28(4):867–874, 1964.

[13] Y.Pesin. Characteristic lyapunov exponents, and smooth ergodic theory. *Russ. Math. Surv.*, 32:55–114, 1977.

[14] S. A. Yuzvinskii. Metric properties of the endomorphisms of compact groups. *Izv. Akad. Nauk SSSR Ser. Mat.*, 29(6):1295–1328, 1965.