

Relatório Final de Iniciação Científica - FAPESP

Teoria Ergódica e Dinâmica em Espaços Homogêneos

Processo: #2016/22585-9.

Aluno: Pedro Griguol de Mattos.

Orientador: Prof. Dr. José Régis Azevedo Varão Filho.

Período do relatório: 01/01/2017 a 31/12/2017.

Vigência da bolsa: 01/01/2017 a 31/12/2017.

Sumário

1	Relato das Atividades do Relatório	3
2	Introdução	4
3	Preliminares de Teoria da Medida	6
4	Tranformações que Preservam Medida	7
5	Recorrência e Transformação Induzida	10
6	Ergodicidade	15
7	Média Ergódica	20
8	Sistemas Misturadores	21
9	O Plano Hiperbólico	27
10	Ação de Grupos no Plano Hiperbólico	28
11	Geodésicas e Fluxos no Plano Hiperbólico	33
12	Os Fluxos Geodésico e Horociclo	35
	12.1 Fluxo Geodésico	35
	12.2 Fluxo Horociclo	37
13	Conclusão	38

1 Relato das Atividades do Relatório

O período vigente deste projeto de iniciação científica é de 01 de janeiro de 2017 a 31 de dezembro de 2017. Nesse período de trabalho, realizei encontros semanais com o professor responsável. Esses encontros tinham como objetivo a apresentação dos tópicos estudados, bem como a indicação por parte do professor de como proceder aos estudos. No segundo semestre de 2016, cursei a disciplina Análise Real do catálogo de mestrado da Unicamp, a qual me foi necessária para realizar este estudo. No primeiro semestre de 2017, cursei 3 disciplinas do catálogo de mestrado da Matemática da Unicamp — Análise no \mathbb{R}^n , Equações Diferenciais Ordinárias, Topologia Geral.

Neste segundo semestre de 2017, cursei as disciplinas de Introdução às Equações Diferenciais Parciais e Geometria Diferencial. A segunda foi principalmente importante para introduzir noções geométricas em variedades diferenciáveis, que também estudei neste relatório através do exemplo do plano hiperbólico.

2 Introdução

Este trabalho consiste no estudo de dinâmica em espaços homogêneos, considerando modelos matemáticos de geometria não-euclidiana como plano de Poincaré, e sua relação com a área de sistemas dinâmicos através da Teoria Ergódica. Faz-se necessário o entendimento de conceitos básicos da Teoria da Medida para que o conceito de ergodicidade possa ser introduzida devida e proveitosamente no estudo de sistemas dinâmicos. A primeira parte do trabalho, de que consiste este primeiro relatório, está voltada ao estudo dessa teoria matemática necessária para a análise e compreensão dos modelos matemáticos que serão estudados à frente.

Na seção §3, alguns conceitos da Teoria da Medida são definidos e alguns teoremas enunciados. As demonstrações não são apresentadas, já que não são o foco deste estudo e já estudei essa teoria em disciplina cursada no semestre passado na UNICAMP. Na seção §4, o primeiro conceito importante é apresentado, o de transformações que preservam a medida em um espaço de probabilidade. Algumas propriedades e consequências das definições são apresentadas, e alguns exemplos são analisados – a saber, a rotação e expansão no círculo e o shift de Bernoulli. Na seção §5, são apresentados os conceitos de recorrência e de transformação induzida. O teorema de recorrência de Poincaré é um dos pilares da teoria ergódica, um resultado que motivou o estudo dessa teoria. Quanto à transformação induzida, alguns teoremas são apresentados, e a principal construção é o arranha-céu de Kakutani. Na seção §6, definimos o conceito de ergodicidade e apresentamos uma proposição que mostra quatro propriedades que equivalem à definição de ergodicidade. Novamente, a rotação e expansão no círculo e o shift de Bernoulli são analisados, dessa vez quanto à ergodicidade. A seção seguinte enuncia alguns resultados que não são demonstrados, mas cuja referência é o livro [2]. As demonstrações foram omitidas porque entram em conceitos de análise funcional, os quais não são foco deste estudo. Os teoremas enunciados são o teorema da média ergódica e o teorema de Birkoff. A seção §8 trata do conceito de sistemas misturadores. Enunciamos um teorema sobre ergodicidade que dialoga com as definições de sistema misturador e fracamente misturador. Algumas relações entre esses conceitos são apresentadas numa proposição e, por fim, demonstramos uma equivalência de definições sobre misturador e fracamente misturador.

Na seção §9 introduzimos o plano hiperbólico \mathbb{H} , o objeto de estudo das seções seguintes. O plano hiperbólico é estudado como um exemplo de uma variedade diferencial, de fato uma variedade riemanniana. São introduzidos conceitos chave como os planos e o fibrado tangentes e, em cada plano tangente, um produto interno, formando uma métrica riemanniana, embora isso não seja comentado a fundo. A partir dessa métrica, introduzimos uma distância no espaço a partir da integral de curvas em \mathbb{H} . Na seção seguinte, §10, um conceito muito importante é trabalhado, a ação de grupos. Uma ação no espaço hiperbólico nos permitirá estudar melhor as curvas

geodésicas nesse espaço, e mostrar que elas nos dão curvas com a menor distância no sentido definido na seção anterior. Alguns grupos importantes são analisados, como $PSL_2(\mathbb{R})$ e PSO_2 , fora algumas definições e proposições gerais sobre ações de grupo. Mostramos que a ação de $PSL_2(\mathbb{R})$ é isométrica e que sua ação induzida no espaço tangente pela derivada também. Geodésicas são calculadas na seção §11. Calculamos as geodésicas verticais mais simples e usamos a isometria das ações de grupo para transportá-las para quaisquer outros pontos do espaço e, assim, calcular qualquer geodésica. O conceito de fluxo também é introduzido nessa seção, usando o fato de que a ação de grupo já foi definida, e são estudados dois fluxos importantes: o fluxo geodésico, que segue as geodésicas, e o fluxo horociclo, que segue as variedades estável e instável do espaço.

3 Preliminares de Teoria da Medida

O estudo da teoria ergódica presuppõe o entendimento de teoria da medida e, em especial, de espaços de probabilidade. Sendo assim, definiremos nesta seção alguns conceitos essenciais para o desenvolvimento da teoria e enunciaremos alguns teoremas principais. As demonstrações serão omitidas, mas o leitor interessado poderá consultar as referências oferecidas.

O conceito mais básico a ser considerado é o de um espaço de medida. A ideia de um espaço de medida é que temos um conjunto, o espaço, em que podemos medir o tamanho de subconjuntos desse espaço através de uma função chamada *medida*, que satisfaz algumas propriedades intuitivas. O exemplo motivador é a reta real, em que sabemos dizer qual o tamanho de um intervalo $[a, b]$ — esse tamanho, chamado comprimento do intervalo, é $b - a$. Sabemos, ainda, medir o tamanho de dois intervalos disjuntos, basta somar os respectivos comprimentos. Evidentemente, não empregamos o termo comprimento nesse caso, pois o conjunto se trata da união de intervalos distintos, mas a mesma ideia de tamanho ainda vale. No entanto, não é sempre possível medir o tamanho de qualquer subconjunto de um espaço. O exemplo disso é menos trivial. Por essa razão, devemos considerar, também, num espaço de medida, um conjunto que consiste nos subconjuntos do espaço que podem ser medidos. Esses conjuntos são chamados *conjuntos mensuráveis*. Baseados nessas definições, definimos os seguintes conceitos formalmente.

Definição 3.1. Seja X um conjunto não vazio. Uma *sigma-álgebra* \mathcal{X} sobre X é um subconjunto de $\mathcal{P}(X)$ que satisfaz

1. O vazio é mensurável.

$$\emptyset \in \mathcal{X}$$

2. Complementar de mensurável é mensurável.

$$A \in \mathcal{X} \Rightarrow A^c \in \mathcal{X}$$

3. União enumerável de mensuráveis é mensurável.

$$(A_i)_{i \in \mathbb{N}} \subseteq \mathcal{X} \Rightarrow \bigcup_{i \in \mathbb{N}} A_i \in \mathcal{X}$$

Um par (X, \mathcal{X}) é um espaço mensurável.

Definição 3.2. Seja (X, \mathcal{X}) um espaço mensurável. Uma *medida* sobre X é uma função $\mu : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ que satisfaz

1. $\mu(\emptyset) = 0$;
2. $\forall E \in \mathcal{X} \quad \mu(E) \geq 0$;
3. Se $(E_i)_{i \in \mathbb{N}} \in \mathcal{X}$ é uma família de conjuntos disjuntos dois a dois, então

$$\mu \left(\bigcup_{i \in \mathbb{N}} E_i \right) = \sum_{i \in \mathbb{N}} \mu(E_i).$$

Definição 3.3. Um *espaço de medida* é uma tripla (X, \mathcal{X}, μ) , em que (X, \mathcal{X}) é um espaço mensurável e $\mu : \mathcal{X} \rightarrow \overline{\mathbb{R}}$ é uma medida sobre (X, \mathcal{X}) .

Um *espaço de probabilidade* é um espaço de medida em que $\mu(X) = 1$.

Enunciamos a seguir dois teoremas importantes de integração, cuja demonstração pode ser encontrada em [1].

Teorema 3.1 (Teorema da Convergência Monótona). *Seja $(f_n)_{n \in \mathbb{N}}$ uma sequência monótona crescente de funções positivas mensuráveis que convergem para f . Então*

$$\int f d\mu = \lim \int f_n d\mu.$$

Teorema 3.2 (Teorema da Convergência Dominada). *Seja $(f_n)_{n \in \mathbb{N}}$ uma sequência de funções integráveis que converge em quase todo ponto para uma função mensurável f . Se existe uma função integrável g tal que $|f_n| \leq g$ para todo n , então f é integrável e*

$$\int f d\mu = \lim \int f_n d\mu.$$

4 Transformações que Preservam Medida

Como é costume na Matemática, o estudo de alguma estrutura é facilitado e expandido ao definirmos mapas que preservam estrutura entre modelos dessa estrutura. Na Álgebra, esses mapas são chamados de homomorfismos, ou transformações lineares na Álgebra Linear. Na Topologia e Análise, há a noção de mapas contínuos, diferenciáveis e analíticos. Os exemplos não se esgotam. Ainda mais que isso, usando esses mapas podemos definir quando o estudo de uma estrutura é análogo ao de alguma outra, o que significa que há um mapa que preserva estrutura entre elas. Daí vêm grupos, anéis, corpos e espaços vetoriais isomorfos, bem como espaços homeomorfos e difeomorfos. Nesta seção, a noção de mapas que preservam medida será definida, bem como a de sistemas de medida isomorfos.

Todo o estudo feito a seguir depende do conceito de uma transformação em um espaço de probabilidade que, de certa forma, preserva a medida dos conjuntos em que age. A definição de preservar medida poderia ser, naturalmente, que a imagem de um conjunto tem a mesma medida que ele próprio. Isso, porém, acaba sendo uma propriedade não tão boa – de fato, não tão natural. Assim como quando se define uma função contínua como preservando imagem inversa de abertos como aberto, aqui a definição de uma transformação que preserva medida será análoga. Essa é a definição que de fato preserva a estrutura do espaço como queremos. Em alguns sistemas, ficará evidente, a imagem e a imagem inversa são preservadas, mas em outros isso não ocorre.

Definição 4.1. Sejam (X, \mathcal{X}, μ) e (Y, \mathcal{Y}, ν) espaços de probabilidade. Uma *função mensurável* de X em Y é uma função $T : X \rightarrow Y$ que satisfaz

$$\forall A \in \mathcal{Y} \quad T^{-1}(A) \in \mathcal{X}.$$

Uma função que *preserva medida* é uma função mensurável que satisfaz

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

Além disso, se T^{-1} está definida para quase todo ponto e é mensurável, então T é dita uma função invertível que preserva medida.

Definição 4.2. Um *sistema que preserva medida* é uma quádrupla (X, \mathcal{X}, μ, T) em que (X, \mathcal{X}, μ) é um espaço de probabilidade e $T : X \rightarrow X$ preserva medida. Nesse caso, μ é dita *T -invariante* e T uma *transformação que preserva medida*.

Proposição 4.1. *Seja (X, \mathcal{X}, μ) um espaço de probabilidade e $T : X \rightarrow X$. A transformação T preserva a medida μ se, e somente se, para toda $f \in \mathcal{L}^\infty$,*

$$\int f d\mu = \int f \circ T d\mu$$

Demonstração. Se vale a igualdade de integrais do teorema, então segue que, para qualquer $M \in \mathcal{X}$, temos

$$\mu(M) = \int \chi_M d\mu = \int \chi_M \circ T d\mu = \mu(T^{-1}(M)),$$

o que mostra que T preserva a medida μ . A recíproca é simples de ser demonstrada mas não vamos demonstrá-la aqui pois exige mais conceitos de teoria da medida, e trabalhar esses conceitos não é o foco deste estudo. Conferir [2] para a demonstração. ■

Com o intuito de entender sistema bem conhecidos sob essa nova propriedade, mostramos a seguir que os sistemas de rotação e expansão no toro preservam medida. É importante notar que, no caso da rotação, a medida do iterado $R_\alpha([a, b])$ também é preservada, mas no caso da expansão, o iterado $E_m([a, b])$ tem medida m vezes maior, o que de certa forma justifica a definição de preservar medida.

Proposição 4.2. 1. A rotação no círculo, definida por

$$\begin{aligned} R_\alpha : \mathbb{T} &\rightarrow \mathbb{T} \\ x &\mapsto x + \alpha \pmod{1}, \end{aligned}$$

preserva a medida de Lebesgue.

2. A expansão no círculo, definida por

$$\begin{aligned} E_m : \mathbb{T} &\rightarrow \mathbb{T} \\ x &\mapsto mx \pmod{1}, \end{aligned}$$

preserva a medida de Lebesgue.

Demonstração. 1. É suficiente mostrar que a medida é preservada para intervalos. Mas isso é simples, pois dado $[a, b] \in \mathbb{T}$, como $R_\alpha^{-1}([a, b]) = [a - \alpha, b - \alpha]$, segue que

$$\mu(R_\alpha^{-1}([a, b])) = (b - \alpha) - (a - \alpha) = b - a = \mu([a, b]).$$

2. Como no caso anterior, é suficiente mostrar para intervalos. Dado $[a, b] \in \mathbb{T}$, como $E_m^{-1}([a, b]) = \bigsqcup_{i=0}^{m-1} [\frac{a+i}{m}, \frac{b+i}{m}]$, segue que

$$\mu(E_m^{-1}([a, b])) = \sum_{i=0}^{m-1} \frac{b+i}{m} - \frac{a+i}{m} = m \frac{b-a}{m} = \mu([a, b]).$$

■

Agora que já definimos os mapas que preservam medida, definimos a seguir o conceito de sistemas de medida isomorfos.

Definição 4.3. Sejam (X, \mathcal{X}, μ, T) e (Y, \mathcal{Y}, ν, S) sistemas que preservam medida. Os sistemas são *isomorfos* se existem $M \in \mathcal{X}$ e $N \in \mathcal{Y}$ tais que $\mu(M) = \nu(N) = 1$, $T(M) \subseteq M$, $S(N) \subseteq N$ e um mapa $\phi : X \rightarrow Y$ invertível que preserve medida tal que, para todo $x \in M$

$$\phi \circ T(x) = S \circ \phi(x).$$

Um sistema importante é o *sistema de Bernoulli* $(X, \mathcal{B}, \mu, \sigma)$, definido a seguir. Seja $X := \{1, \dots, n\}^{\mathbb{Z}}$, \mathcal{B} sua σ -álgebra de Borel, $p = (p_1, \dots, p_n) \in \mathbb{R}^+$ tal que $\sum p_i = 1$, $\mu_p(i) := p_i$, $\mu := \prod_{\mathbb{Z}} \mu_p$ e

$$\begin{aligned} \sigma : X &\rightarrow X \\ (x_n)_{n \in \mathbb{N}} &\mapsto (x_{n+1})_{n \in \mathbb{N}}. \end{aligned}$$

Nesse sistema, são importantes conjuntos chamados cilindros. De fato, eles geram a σ -álgebra de Borel. Para isso, dados $C \subseteq \mathbb{Z}$ finito e $a : C \rightarrow \{1, \dots, n\}$, definimos o *cilindro*

$$C(a) := \{x \in X : \forall j \in C \quad x_j = a(j)\}.$$

Nesse caso, $\mu(C(a)) = \prod_{i \in C} p_{a(i)}$. O sistema $(X, \mathcal{B}, \mu, \sigma)$ preserva medida, e a demonstração pode ser conferida em [2]. Usaremos isso na proposição a cerca do sistema de Bernoulli simplificado a seguir, que pode ser colocado em isomorfismo de medida com a expansão E_2 no círculo.

Proposição 4.3. *Seja o sistema de Bernoulli $(\Sigma, \mathcal{B}, \mu, \sigma)$, em que $\Sigma := \{0, 1\}^{\mathbb{N}}$, \mathcal{B} é a σ -álgebra de Borel de Σ , σ é o shift de Bernoulli e μ é a medida produto $\mu := \prod_{\mathbb{N}} \nu$, em que ν é a medida em $\{0, 1\}$ definida por $\nu(\{0\}) = \nu(\{1\}) = \frac{1}{2}$. Então o mapa*

$$\begin{aligned} \phi : \Sigma &\rightarrow \mathbb{T} \\ (x_n)_{n \in \mathbb{N}} &\mapsto \sum_{n \in \mathbb{N}} \frac{x_n}{2^{n+1}} \end{aligned}$$

é um isomorfismo entre o sistema de Bernoulli e a expansão E_2 no círculo.

Demonstração. Primeiro notemos que o sistema de Bernoulli deve preservar medida pelo comentário anterior. Assim, notemos que, para $(x_n)_{n \in \mathbb{N}}$,

$$\phi \circ \sigma((x_n)) = \phi((x_{n+1})) = \sum_{n \in \mathbb{N}} \frac{x_{n+1}}{2^{n+1}} = 2 \sum_{n \in \mathbb{N}} \frac{x_{n+1}}{2^{n+2}} = E_2 \circ \phi(x_n).$$

A inversa de ϕ é definida a menos de o conjunto de racionais diádicos, que tem medida nula, e então segue o resultado. ■

5 Recorrência e Transformação Induzida

Teorema 5.1 (Teorema da Recorrência de Poincaré). *Seja (X, \mathcal{X}, μ, T) um sistema que preserva medida e $E \subseteq X$ um conjunto mensurável. Então existe $F \subseteq E$ tal que $\mu(F) = \mu(E)$ satisfazendo que, para todo $x \in F$, existem inteiros $0 < n_1 < n_2 < \dots$ tais que, para todo $i \in \mathbb{N}^*$, $T^{n_i}(x) \in E$.*

Demonstração. Consideremos o conjunto $A := \{x \in E : \forall n \in \mathbb{N} \ T^n(x) \notin E\}$ dos pontos de E que não retornam a E em nenhuma iteração. Então temos que

$$A = E \cap \bigcap_{i \in \mathbb{N}^*} T^{-i}(E^c) = T^{-1}(E^c) \cap T^{-2}(E^c) \cap \dots \quad (1)$$

Para ver isso, tomemos $x \in A$. Então $x \in E$ e, para todo $n \in \mathbb{N}^*$, $T^n(x) \notin E$, ou $T^n(x) \in E^c$, o que implica $x \in T^{-n}(E^c)$. Por outro lado, se x é elemento da interseção acima, então $x \in E$ e, para todo $n \in \mathbb{N}^*$, $x \in T^{-n}(E^c)$, ou $T^n(x) \in E^c$, logo $T^n(x) \notin E$. Sendo assim, segue que A é mensurável, pois E o é e T preserva medida. Notemos então que os conjuntos

$$T^{-n}(A) = T^{-n}(E) \cap \bigcap_{i \in \mathbb{N}^*} T^{-n-i}(E^c) = T^{-n-1}(E^c) \cap T^{-n-2}(E^c) \cap \dots \quad (2)$$

são disjuntos para todo $n \in \mathbb{N}^*$, pois, dados $n, m \in \mathbb{N}$, suponhamos $n < m$. Então temos que $T^{-n}(A) \subseteq T^{-n}(E)$ e $T^{-m}(A) \subseteq T^{-m}(E^c) = (T^{-n}(E))^c$. Ainda, como T preserva medida, todos têm medida $\mu(A)$. Isso implica que $\mu(A) = 0$. Concluimos que $F_1 := E \setminus A$ é um conjunto conjunto $F_1 \subseteq E$ tal que $\mu(F_1) = \mu(E)$ e todo ponto de F_1 retorna ao menos uma vez para E . Definimos, analogamente, usando as transformações T^2, T^3, \dots , conjuntos F_2, F_3, \dots tais que $\mu(F_n) = \mu(E)$ e todo ponto de F_n retornando para E sob T^n . Assim, o conjunto

$$F := \bigcap_{n \in \mathbb{N}^*} F_n \quad (3)$$

satisfaz $F \subseteq E$, $\mu(F) = \mu(E)$ e todo ponto de F retorna infinitas vezes para E . ■

Pelo teorema de Recorrência de Poincaré, sabemos que, num espaço que preserva medida, para qualquer conjunto mensurável existe um subconjunto desse conjunto com mesma medida e tal que todo elemento retornam para o conjunto original infinitas vezes. Sendo assim, podemos definir qual é o menor tempo de retorno para quase todo ponto de um conjunto mensurável. Ainda, podemos a partir disso definir uma transformação que leva um ponto do conjunto mensurável no primeiro iterado seu que volta para o conjunto. As definições abaixo se ocupam disso.

Definição 5.1. Sejam (X, \mathcal{X}, μ, T) um sistema que preserva medida e $A \in \mathcal{X}$. Para quase todo ponto $x \in A$, o *menor tempo de retorno* de x para A sob T é

$$r_A(x) := \inf_{n \geq 1} \{n : T^n(x) \in A\}.$$

Definição 5.2. Sejam (X, \mathcal{X}, μ, T) um sistema que preserva medida e $A \in \mathcal{X}$. A *transformação induzida* T_A por T em A é definida para quase todo ponto $x \in A$ como

$$T_A(x) := T^{r_A(x)}(x).$$

Proposição 5.2. *As transformações r_A e T_A são mensuráveis.*

Demonstração. Consideremos os conjuntos

$$A_n := \{x \in A : r_A(x) = n\}$$

dos pontos de A cujo primeiro retorno por T é n . Esses conjuntos são mensuráveis, já que, para todo n , vale que

$$A_n = A \cap T^{-n}(A) \setminus \bigcup_{i=1}^{n-1} A_i.$$

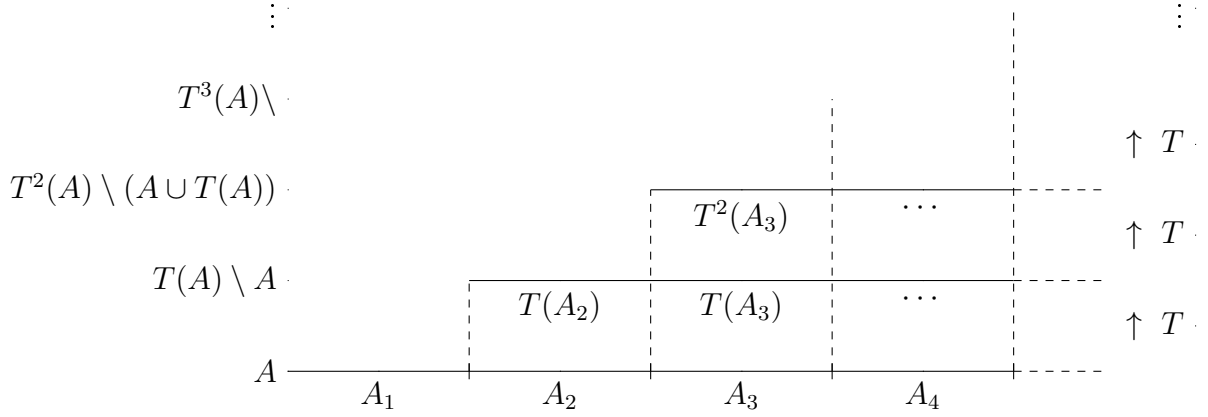
Isso ocorre porque, para um ponto ter o menor retorno n , ele não deve estar em nenhum dos A_i anteriores e deve estar em $T^{-n}(A)$, bem como em A . Agora, também é mensurável o conjunto

$$T^n(A_n) = A \cap T^n(A) \setminus \bigcup_{i=1}^{n-1} T^i(A_i),$$

e segue a proposição. ■

O Aranha-Céu de Kakutani

Uma construção importantíssima para facilitar o entendimento dessas definições acima e dos próximos teoremas é a chamada *arranha-céu de Kakutani*. A ideia é representar graficamente todos os conjuntos A_n e seus iterados. Representamos o conjunto A como uma reta dividida em segmentos seguidos que representam cada A_n . Acima dessa reta, usamos uma nova reta segmentada para representar $T(A) \setminus A$, de modo que cada segmento acima de A_n é a imagem de A_n , e o mesmo processo é repetido infinitamente. Cada uma dessas retas é chamada de um andar, e cada um dos conjuntos de iterado de um dos A_n é a torre desse respectivo conjunto. No andar n , a torre n é terminada, pois, por definição, a imagem desse conjunto está novamente em A , a reta mais embaixo, também chamado de chão do arranha-céu. Sendo assim, a n -ésima torre sempre tem n andares. A transformação induzida definida acima pode ser pensada como a função que leva um ponto até o último andar de sua torre e, então, desce com ele de volta para o chão.



Lema 5.3. *Sejam (X, \mathcal{X}, μ, T) um sistema que preserva medida e $A \in \mathcal{X}$, e definamos $\mu_A := \frac{1}{\mu(A)}\mu$ e $\mathcal{X}|_A := \{M \cap A : M \in \mathcal{X}\}$. Então o sistema $(A, \mathcal{X}|_A, \mu_A, T_A)$ preserva medida. Ainda, se T é ergódica com respeito a μ , então T_A o é com respeito a μ_A .*

Demonstração. Sejam $M \subseteq A$ mensurável e A_n os conjuntos definidos na proposição anterior. Então $M = \bigsqcup_{n \geq 1} M \cap A_n$, o que implica que

$$\mu_A(M) = \frac{1}{\mu(A)} \sum_{n \geq 1} \mu(M \cap A_n).$$

Ainda,

$$T_A(M) = \bigsqcup_{n \geq 1} T_A(M \cap A_n) = \bigsqcup_{n \geq 1} T^n(M \cap A_n).$$

Então

$$\begin{aligned} \mu_A(T_A(M)) &= \frac{1}{\mu(A)} \sum_{n \geq 1} \mu(T^n(M \cap A_n)) \\ &= \frac{1}{\mu(A)} \sum_{n \geq 1} \mu(M \cap A_n) \\ &= \mu(M), \end{aligned}$$

já que T preserva medida.

Agora, vamos provar a proposição sobre ergodicidade pela contrapositiva. Se T_A não é ergódica, existe $M \subseteq A$ mensurável e T_A invariante com $0 < \mu(M) < \mu(A)$. Então

$$\bigcup_{n \geq 1} \bigcup_{j=0}^{n-1} T^j(M \cap A_n)$$

é invariante, mas não é o conjunto todo, o que implica que T não é ergódica. ■

Teorema 5.4 (Kac). *Sejam (X, \mathcal{X}, μ, T) um sistema ergódico que preserva medida e $A \in \mathcal{X}$ tal que $\mu(A) > 0$. Então*

$$\int_A r_A d\mu = 1$$

(o que é equivalente a dizer que o tempo de retorno esperado para A é $\frac{1}{\mu(A)}$).

Demonstração. Sejam A_n os conjuntos definidos na proposição anterior. Notemos que, para todo n ,

$$A_n, T(A_n), \dots, T^{n-1}(A_n)$$

são disjuntos e têm medida $\mu(A_n)$, pois T preserva medida. Por ergodicidade, sabemos que, se $\mu(A) > 0$, então $\mu(\bigcup_{n \geq 1} T^{-n}(A)) = 1$. Usando isso para T^{-1} , segue que

$$\mu\left(\bigcup_{n \geq 1} T^n(A)\right) = 1,$$

e então que

$$1 = \mu\left(\bigcup_{n \geq 1} T^n(A)\right) = \sum_{n \geq 1} \mu(T^n(A)) = \sum_{n \geq 1} \mu(A) = \int_A r_A d\mu,$$

pelo teorema da convergência monótona, pois r_A é o limite da sequência de funções $\sum_{k=1}^n k \chi_{A_n}$ quando $n \rightarrow \infty$. ■

Para o lema a seguir, introduzimos antes a ideia de uma medidã atômica. A intuição é que ua medida é não atômica se ele não detecta "átomos", ou seja, a medida de um ponto do conjunto é zero.

Definição 5.3. Uma medida *não atômica* é uma medida de um sistema (X, \mathcal{X}, μ) que satisfaz

$$\forall x \in X \quad \mu(\{x\}) = 0.$$

A ideia do lema a seguir é a seguinte. Dado um sistema nas condições adequadas, podemos, para todo ε e todo n , escolher conjunto mensurável tal que seus $n - 1$ iterados são disjuntos dele e a medida desses n conjuntos é quase tudo, a menos de ε .

Lema 5.5 (Kakutani-Rokhlin). *Sejam (X, \mathcal{X}, μ, T) um sistema ergódico invertível que preserva medida e μ não atômica. Então, para todo $n \geq 1$ e $\varepsilon > 0$, existe $M \in \mathcal{X}$ tal que*

$$M, T(M), \dots, T^{n-1}(M)$$

são disjuntos e

$$\mu\left(\bigsqcup_{i=0}^{n-1} T^i(M)\right) > 1 - \varepsilon.$$

Demonstração. Seja $A \in \mathcal{X}$ tal que $0 < \mu(A) < \frac{\varepsilon}{n}$. Podemos escolher tal conjunto porque μ é não atômica. Então, usando os conjuntos A_n definidos na proposição anterior, X pode ser escrito como

$$X = \bigsqcup_{k \geq 1} \bigsqcup_{i=0}^{k-1} T^i(A_k).$$

Consideremos o conjunto

$$M := \bigsqcup_{k \geq n} \bigsqcup_{j=0}^{\lfloor \frac{k}{n} \rfloor - 1} T^{jn}(A_k).$$

Esse conjunto é a união de todos A_k , para $K \geq n$, junto a todos n -ésimos iterados acima no arranha-céu de Kakutani, mas não pegando o teto de cada torre. Por construção, os conjuntos

$$M, T(M), \dots, T^{n-1}(M)$$

são disjuntos, pois a cada iterado de M anda-se em cada torre $T^i(A_k)$ um andar, e isso é repetido somente até chegar antes do n -ésimo andar. Notando que, para cada k e cada i , $\mu(T^i(A_k)) = \mu(A_k)$, segue então que a união desses n conjuntos cobre todo o X a menos de, em cada torre, no máximo n andares, o que tem medida menor que

$$n \sum_{k=1}^{\infty} \mu(A_k) \leq n\mu(A) < \varepsilon. \quad \blacksquare$$

6 Ergodicidade

Nesta seção, introduziremos o conceito de ergodicidade. A definição de transformação em sistemas ergódicos virá seguida de uma proposição que mostra várias equivalências de propriedades que caracterizam a ergodicidade. Essas equivalências são importantes não só como resultados teóricos interessantes e importantes ferramentas da teoria, mas também como embasamento intuitivo da ideia de ergodicidade.

Definição 6.1. Seja (X, \mathcal{X}, μ) um espaço de probabilidade. Uma transformação *ergódica* é uma transformação T que preserva a medida μ e, para todo $M \in \mathcal{X}$,

$$T^{-1}(M) = M \Rightarrow \mu(M) = 0 \quad \text{ou} \quad \mu(M) = 1.$$

Um *sistema ergódico* é um sistema (X, \mathcal{X}, μ, T) em que T é ergódica.

A propriedade do conjunto mensurável M na definição de ergodicidade é chamada *invariância estrita*. A definição diz que um sistema ergódico é aquele cujos conjuntos

estritamente invariantes têm medida zero ou um. Nas equivalências a seguir, usamos o operador diferença simétrica. Ele representa a parte da união de dois conjuntos que não pertence a ambos. Esse operador é útil para falar, por exemplo, que a porção que não está na união de dois conjuntos tem medida nula; ou seja, para falar que os conjuntos são quase os mesmos. Abaixo, definimos esse operador e, em seguida, enunciamos a proposição.

Definição 6.2. A *diferença simétrica* entre conjuntos A e B é

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

Proposição 6.1. Seja (X, \mathcal{X}, μ) um sistema que preserva medida. São equivalentes

1. T é ergódica.
2. Para todo $M \in \mathcal{X}$, $\mu(T^{-1}(M) \Delta M) = 0$ implica

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

3. Para todo $M \in \mathcal{X}$, $\mu(M) > 0$ implica

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

4. Para todos $M, N \in \mathcal{X}$, $\mu(M)\mu(N) > 0$ implica que existe $n \in \mathbb{N}^*$ tal que

$$\mu(T^{-n}(M) \cap N) = 1.$$

5. Seja $f : X \rightarrow \mathbb{C}$ mensurável. Se $f \circ T = f$ quase sempre, então f é quase sempre constante.

Demonstração. (1) \Rightarrow (2). Seja $M \in \mathcal{X}$ um conjunto que satisfaz $\mu(T^{-1}(M) \Delta M) = 0$. A demonstração consiste em construir um conjunto N que tenha a mesma medida de M e seja estritamente invariante, para que assim a hipótese de ergodicidade possa ser usada. Sendo assim, consideremos

$$N := \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} T^{-n}(M).$$

Notemos que, para todo $m \in \mathbb{N}$,

$$M \Delta \bigcup_{n=m}^{\infty} T^{-n}(M) \subseteq \bigcup_{n=m}^{\infty} M \Delta T^{-n}(M),$$

pois, se x é elemento do primeiro, então é elemento de M mas não da união dos $T^{-n}(M)$ ou é elemento da união mas não de M . No primeiro caso, ele não é elemento de nenhum dos $T^{-n}(M)$, logo é elemento de todos $M \Delta T^{-n}(M)$ e, portanto, da união; no segundo caso, é elemento de algum $T^{-n}(M)$ e, portanto, de algum $M \Delta T^{-n}(M)$, logo da união. Ainda, para todo $n \in \mathbb{N}^*$ vale

$$M \Delta T^{-n}(M) \subseteq \bigcup_{i=0}^{n-1} T^{-i}(M) \Delta T^{-(i+1)}(M).$$

Para ver isso, seja $x \in M \Delta T^{-n}(M)$. Suponhamos que $x \in M$. Então não pertence a $T^{-n}(M)$. Então, se $x \notin T^{-1}(M)$, temos que $x \in M \Delta T^{-1}(M)$ e, se $x \in T^{-1}(M)$, repetimos o processo com $T^{-2}(M)$ chegando, no máximo, até a $x \in T^{-(n-1)}(M)$, o que implicará $x \in T^{-(n-1)}(M) \Delta T^{-n}(M)$; em todos os casos, x pertence à união acima. Se $x \in T^{-n}(M)$, o mesmo processo se aplica, mas na ordem inversa. Isso implica que $\mu(M \Delta T^{-n}(M)) = 0$, pois o termo à direita tem medida zero.

Definamos $N_m := \bigcup_{n=m}^{\infty} T^{-n}(M)$, e temos que $N = \bigcap_{m=0}^{\infty} N_m$ e, claramente,

$$N_0 \supseteq N_1 \supseteq N_2 \supseteq \dots$$

Mas de $\mu(M \cap N_m) = 0$, então, segue que $\mu(N \Delta M) = 0$, logo que $\mu(N) = \mu(M)$. Por fim, notando que

$$T^{-1}(N) = \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} T^{-(n+1)}(M) = \bigcap_{m=0}^{\infty} \bigcup_{n=m+1}^{\infty} T^{-n}(M) = N,$$

segue que N é invariante, o que implica, por ergodicidade, que $\mu(N)$ é 0 ou 1; logo, o mesmo vale para M .

(2) \Rightarrow (3). Seja $M \in \mathcal{X}$ com $\mu(M) > 0$ e definamos

$$N := \bigcup_{n=1}^{\infty} T^{-n}(M).$$

Segue que $T^{-1}(N) \subseteq N$. Mas $\mu(T^{-1}(N)) = \mu(N)$, logo $\mu(N \Delta T^{-1}(N)) = 0$, o que implica, por (2), que $\mu(N)$ é 0 ou 1. Mas $\mu(N) = 0$ não pode ocorrer, pois $T^{-1}(M) \subseteq N$ temos que $\mu(N) = 1$.

(3) \Rightarrow (4). Sejam $M, N \in \mathcal{X}$ com $\mu(M)\mu(N) > 0$. Por (3), sabemos que

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

e disso segue que

$$0 < \mu(N) < \mu\left(\bigcup_{n=1}^{\infty} M \cap T^{-n}(M)\right) \leq \sum_{n=1}^{\infty} \mu(M \cap T^{-n}(M)).$$

Logo, existe $n \in \mathbb{N}^*$ tal que $\mu(M \cap T^{-n}(M)) > 0$.

(4) \Rightarrow (1). Seja $M \in \mathcal{X}$ tal que $T^{-1}(M) = M$. Então segue que

$$0 = \mu(M \cap M^{\complement}) = \mu(T^{-n}(M) \cap M^{\complement})$$

para todo $n \in \mathbb{N}^*$, o que implica por (4) que $\mu(M) = 0$ ou $\mu(M^{\complement})$, logo $\mu(M)$ é 0 ou 1, e a ergodicidade está provada. A demonstração da equivalência de (5) pode ser conferida em [2]. \blacksquare

A propriedade do conjunto mensurável M no segundo item da proposição acima é chamada *quase invariância*. A proposição diz que também é uma definição válida de ergodicidade considerar aqueles sistemas cujos conjuntos quase invariantes têm medida zero ou um.

Proposição 6.2. 1. A rotação R_α no círculo é ergódica se, e somente se, α é irracional.

2. O shift de Bernoulli é ergódico.

3. A expansão no círculo é ergódica.

Demonstração. 1. Se $\alpha \in \mathbb{Q}$, escrevemos $\alpha = \frac{p}{q}$ fração reduzida. Então, tomando um conjunto mensurável $A \subseteq \mathbb{T}$ tal que $0 < \mu(A) < \frac{1}{q}$, segue que o conjunto

$$M := A \cup R_\alpha(A) \cup R_\alpha^2(A) \cap \dots \cup R_\alpha^{q-1}(A)$$

é invariante sob R_α , pois $R_\alpha^q = id_{\mathbb{T}}$. No entanto, $0 < \mu(M) < 1$, o que mostra que R_α não é ergódica.

Se $\alpha \notin \mathbb{Q}$, então a rotação é densa em \mathbb{T} . Sendo assim, seja $M \subseteq \mathbb{T}$ invariante sob R_α . Podemos escolher, para todo ε uma função f contínua no toro tal que $\|f - \chi_M\|_1 < \varepsilon$ e da invariância de M segue que $\|f \circ R_\alpha^n - f\|_1 < 2\varepsilon$. Da continuidade de f , segue que, para todo $t \in \mathbb{R}$, $\|f \circ R_t - f\|_1 \leq 2\varepsilon$. Agora, como a medida é invariante por rotação, usando o teorema de Fubini e a desigualdade triangular para integrais temos que

$$\begin{aligned} \left\| f - \int f(t) dt \right\|_1 &= \int \left| \int (f(x) - f(x+t)) dt \right| dx \\ &\leq \int \int |f(x) - f(x+t)| dx dt \\ &\leq 2\varepsilon. \end{aligned}$$

Mas então

$$\|\chi_M - \mu(M)\|_1 \leq \|\chi_m - f\|_1 + \left\| f - \int f(t)dt \right\|_1 + \left\| \int f(t)dt - \mu(M) \right\|_1 < 4\varepsilon.$$

Concluimos, portanto, que χ_M é constante, pois isso vale para qualquer ε . Mas então $\mu(M)$ é 0 ou 1, logo R_α é ergódica.

2. Seja $N \in \mathcal{X}$ invariante. Para qualquer $\varepsilon > 0$, existe uma união finita de cilindros M tais que $\mu(M \Delta N) < \varepsilon$, o que implica $|\mu(M) - \mu(N)| < \varepsilon$. O conjunto M pode ser escrito como

$$M = \{x \in X : x_{-m} \cdots x_m \in F\}$$

para algum m e $F \subseteq \{0, 1, \dots, n\}^{\{-m, \dots, m\}}$. Assim, para todo $k > 2m$, temos que

$$\sigma^{-k}(M) = \{x \in X : x_{k-m} \cdots x_{k+m} \in F\}.$$

Como esse dois conjuntos são independentes, segue que

$$\mu(\sigma^{-k}(M) \setminus M) = \mu(\sigma^{-k}(M) \cap M^c) = \mu(\sigma^{-k}(M)) \mu(M^c) = \mu(M) \mu(M^c).$$

Como N é invariante, temos que $\mu(N \Delta \sigma^{-1}(M))$, e então

$$\mu(\sigma^{-k}(M) \Delta N) = \mu(\sigma^{-k}(M) \Delta \sigma^{-k}(N)) = \mu(M \Delta N) < \varepsilon$$

logo $\mu(\sigma^{-k}(M) \Delta M) < 2\varepsilon$ e

$$\mu(\sigma^{-k}(M) \setminus M) + \mu(M \setminus \sigma^{-k}(M)) = \mu(\sigma^{-k}(M) \Delta M) < 2\varepsilon.$$

Finalmente, obtemos que

$$\begin{aligned} \mu(N) \mu(N^c) &< (\mu(M) + \varepsilon)(\mu(M^c) + \varepsilon) \\ &= \mu(M) \mu(M^c) + \varepsilon \mu(M) + \varepsilon \mu(M^c) + \varepsilon^2 \\ &< 2\varepsilon + \varepsilon + \varepsilon^2. \end{aligned}$$

Como isso vale para qualquer ε , segue que $\mu(N) \mu(N^c) = 0$; ou seja, que $\mu(N)$ é 0 ou 1, logo o sistema é ergódico.

3. Como os sistemas de Bernoulli Σ e a expansão E_2 são isomorfos por medida, segue que a expansão é ergódica. ■

7 Média Ergódica

Nesta seção, são apresentados alguns resultados importantes da definição de ergodicidade. No entanto, as demonstrações não serão apresentadas, e podem ser encontradas em [2], devido ao fato de que muitos dos resultados dependem da teoria da Análise Funcional e o objetivo desse estudo de iniciação não entra muito nessa área, mas sim na parte de teoria da medida, sendo que já cursei a disciplina de Teoria da Medida, mas não a de Análise Funcional. Ainda assim, essa seção foi incluída por sua importância e por motivo de completeza. Os teoremas ergódicos abaixo são parte importante da teoria. O primeiro analisa o comportamento das médias ergódicas de uma função em um sistema que preserva medida. Essas médias, para funções em L^2 , convergem para a projeção da função no subespaço de L^2 das funções invariantes sob ação de T .

Teorema 7.1 (Teorema da Média Ergódica). *Seja (X, \mathcal{X}, μ, T) um sistema que preserva medida e P_T a projeção ortogonal no subespaço $I := \{g \in L^2 : g \circ T = g\} \subseteq L^2$. Então, para toda $f \in L^2$, em L^2 ,*

$$\frac{1}{m} \sum_{n=0}^{m-1} f \circ T^n \longrightarrow P_T(f)$$

quando $m \longrightarrow \infty$.

A quantidade $\frac{1}{m} \sum_{n=0}^{m-1} f \circ T^n$ é conhecida como a m -ésima *média ergódica* de f no sistema acima.

Teorema 7.2 (Birkoff). *Seja (X, \mathcal{X}, μ, T) um sistema que preserva medida. Se $f \in \mathcal{L}^1$, então*

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

converge em quase todo ponto e em L^1 para uma função T -invariante $f^ \in \mathcal{L}^1$ e*

$$\int f^* d\mu = \int f d\mu.$$

Ainda, se T é ergódica, então

$$f^*(x) = \int f d\mu$$

para quase todo ponto.

8 Sistemas Misturadores

Nesta seção consideraremos duas noções de um sistema misturador: a de *misturador* e de *fracamente misturador*. Um sistema misturador também é conhecido como fortemente misturador para distinguí-lo dos fracamente misturadores. Existem, de fato, outras noções de misturador, que generalizam a ideia de misturador, como *misturador de ordem k* (cf. [2]), mas aqui consideraremos apenas essas duas. Antes de introduzir a definição, vale ressaltar uma outra definição equivalente de sistema ergódico, enunciada abaixo, cuja demonstração pode ser achada em [2]. Essa classificação é uma consequência do teorema da média ergódica.

Proposição 8.1. *Um sistema (X, \mathcal{X}, μ, T) é ergódico se, e somente se, para todos $M, N \in \mathcal{X}$,*

$$\frac{1}{n} \sum_{i=0}^{n-1} \mu(M \cap T^{-i}(N)) \longrightarrow \mu(M)\mu(N)$$

quando $n \longrightarrow \infty$.

A definição a seguir, de um sistema misturador, destaca a seguinte ideia. Em alguns sistema, se tomamos dois subconjuntos mensuráveis quaisquer dos espaços, notamos que sempre ocorre que um deles é independente da n -ésima imagem inversa do outro quando n é grande. O sentido de independência empregado é que, em espaços de probabilidade, eventos são pensados como conjuntos mensuráveis e dois eventos E, F são independentes quando $\mu(E \cap F) = \mu(E)\mu(F)$; ou seja, a probabilidade dos dois ocorrerem é o produto das probabilidades de cada um deles ocorrer separadamente. Pensando em conjuntos, e agora considerando uma ideia geométrica que motiva o nome *misturador*, essa ideia é equivalente a pensar que os conjuntos, quando iterados, estão muito bem separados no sistema, misturados, de modo que se fixamos um deles, o primeiro, a medida das imagens inversas do segundo que estão nesse primeiro conjunto, dividida pela medida desse primeiro, é igual à medida de do segundo; ou seja, a quantidade do segundo que está no primeiro é proporcionalmente igual à quantidade desse segundo no espaço todo. A definição a seguir formaliza o que foi aqui explicado.

Definição 8.1. Um sistema *misturador* é um sistema (X, \mathcal{X}, μ, T) que preserva medida tal que, para todos $M, N \in \mathcal{X}$,

$$\mu(M \cap T^{-n}N) \longrightarrow \mu(M)\mu(N)$$

quando $n \longrightarrow \infty$

Proposição 8.2. *A rotação R_α no círculo não é misturadora.*

Demonstração. Tomando uma sequência $n_i \rightarrow \infty$ tal que $n_i \alpha \pmod{1} \rightarrow 0$. No caso de α irracional isso é possível pois a rotação é densa, e no caso de racional, basta tomar uma sequência constante igual a zero. Tomando os conjuntos $M = N = [0, \frac{1}{2}]$, segue que, quando $i \rightarrow \infty$,

$$\mu(M \cap R_\alpha^{-n_i}(N)) = \mu(M \cap R_{-n_i \alpha}(M)) \rightarrow \mu(M) = \frac{1}{2} \neq \left(\frac{1}{2}\right)^2 = \mu(M)\mu(N).$$

■

Embora o conceito de sistema misturador como definido acima é mais natural, a noção levemente modificada abaixo é, muitas vezes, mais útil e simples de se aplicar. Esse é o conceito de fracamente misturador.

Definição 8.2. Um sistema *fracamente misturador* é um sistema (X, \mathcal{X}, μ, T) tal que, para todos $M, N \in \mathcal{X}$,

$$\frac{1}{n} \sum_{i=0}^{n-1} |\mu(M \cap T^{-i}(N)) - \mu(M)\mu(N)| \rightarrow 0$$

quando $n \rightarrow \infty$

É interessante comentar algo sobre os diferentes tipos de convergência aqui considerados. Além da convergência tradicional, $a_n \rightarrow a$, existem outras noções de convergência de uma sequência. Uma delas é convergência de Cesàro, que considera a convergência da sequência $\frac{1}{n} \sum_{i=0}^{n-1} a_n$ das n -ésimas médias da sequência original. Pelas definições e proposições anteriores, considerando a sequência $a_i := \mu(M \cap T^{-i}(N)) - \mu(M)\mu(N)$, temos que um sistema é *ergódico* se, e somente se, a_n converge a 0 pela convergência de Cesàro, é *misturador* se, e somente se, a_n converge a 0 no sentido tradicional, e é *fracamente misturador* se $|a_n|$ converge a 0 pela convergência de Cesàro. Assim, como, para qualquer sequência a_n que converge a 0, vale que $|a_n|$ converge a 0 por Cesàro e, então, a_n converge a 0 por Cesàro, segue a proposição abaixo.

Proposição 8.3. *Seja (X, \mathcal{X}, μ, T) um sistema que preserva medida. Então*

1. *Se o sistema é misturador, é fracamente misturador.*
2. *Se o sistema é fracamente misturador, é ergódico.*

Demonstração. Basta considerar a sequência $a_i := \mu(M \cap T^{-i}(N)) - \mu(M)\mu(N)$ e os comentários do parágrafo anterior. ■

Definição 8.3. A *densidade* de um conjunto $C \subseteq \mathbb{N}$ é o número

$$d(C) := \lim_{n \rightarrow \infty} \frac{1}{n} |\{k \in C : 0 \leq k < n\}| = \frac{1}{n} |(C \cap [0, n])|$$

quando o limite existir.

Antes de demonstrar um teorema importante a seguir com várias equivalências envolvendo a propriedade de ergodicidade e mistura fraca, demonstraremos um lema sobre convergência de sequências.

Lema 8.4. *Seja $(a_n)_{n \in \mathbb{N}}$ uma sequência limitada de números reais positivos. São equivalentes as seguintes noções de convergência.*

1. $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} a_i = 0$.
2. *Existe um conjunto $C \subseteq \mathbb{N}$ com densidade zero tal que $a_n \rightarrow 0$ para $n \rightarrow \infty$, $n \notin C$.*
3. $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} a_i^2 = 0$.

Demonstração. Vamos mostrar que (1) implica (2). Definamos $C_k := \{j \in \mathbb{N} : a_j > \frac{1}{k}\}$. Então temos que

$$C_1 \subseteq C_2 \subseteq C_3 \subseteq \dots, \quad (4)$$

pois, se $x \in C_k$, então $x > \frac{1}{k} > \frac{1}{k+1}$, logo $x \in C_{k+1}$. Então temos que, para todo $k \geq 1$,

$$\frac{1}{k} |C_k \cap [0, n)| = \sum_{i \in C_k \cap [0, n)} \frac{1}{k} < \sum_{i \in C_k \cap [0, n)} a_i \leq \sum_{i=0}^n a_i$$

e então segue que

$$\frac{1}{n} |C_k \cap [0, n)| \leq k \frac{1}{n} \sum_{i=0}^n a_i \rightarrow 0$$

quando $n \rightarrow \infty$. Logo, para todo $k \geq 1$, C_k tem densidade 0. Como cada C_k tem densidade 0, podemos escolher inteiros $0 < l_1 < l_2 < \dots$ tais que, para todo $k \geq 1$ e para todo $n \geq l_k$,

$$\frac{1}{n} |C_k \cap [0, n)| \leq \frac{1}{k}. \quad (5)$$

Basta escolher l_1 e tomar um l_2 suficientemente grande e assim por diante. Definimos então o conjunto

$$C := \bigcup_{k \in \mathbb{N}} (C_k \cap [l_k, l_{k+1})). \quad (6)$$

Vamos mostrar que a_i converge a 0 para $n \notin C$ e que C tem densidade 0. Primeiro, notamos que $C_k \cap [l_k, \infty) \subseteq C$, pois, se $x \in C_k$, então está em todo C_p para $p \geq k$, e, então, se estiver em $[l_p, l_{p+1})$, estará em $C_p \cap [l_p, l_{p+1})$, logo em C . Assim, para todo $n \geq l_k$, então se $n \notin C$, temos $n \notin C_k$ e, portanto, $a_n \leq \frac{1}{k}$, logo $a_n \rightarrow 0$. Agora, para mostrar que C tem densidade zero, notemos que, para todo $n \in [l+k, l_{k+1})$ temos que $C \cap [0, n) \subseteq C_k \cap [0, n)$, pois se $x \in C \cap [0, n)$, está em algum $[l_p, l_{p+1})$ e, por consequência, em $C_k \cap [0, n)$, já que todo C_p está contido no C_k para $p \leq k$. Isso implica que

$$\frac{1}{n}|C \cap [0, n)| \leq \frac{1}{n}|C_k \cap [0, n)| \leq \frac{1}{k} \quad (7)$$

e, portanto, C tem densidade zero.

Agora, vamos mostrar que (2) implica (1). Como a sequência é limitada, seja $L > 0$ tal que $a_n \leq L$ para todo $n \geq 1$. Como a_n converge para zero fora de C e o conjunto C tem densidade zero, podemos escolher, para cada $k \geq 1$, $n_k \in \mathbb{N}$ tal que, para todo $n \notin C$, $n \geq n_k$, temos que $a_n < \frac{1}{k}$ e, para todo $n \geq n_k$, $\frac{1}{n}|C \cap [0, n)| \leq \frac{1}{k}$. Sendo assim, para todo $n \geq kn_k$, temos que

$$\begin{aligned} \frac{1}{n} \sum_{i=0}^{n-1} a_i &= \frac{1}{n} \left(\sum_{i=0}^{n_k-1} a_i + \sum_{i \in C \cap [n_k, n)} a_i + \sum_{i \in C^c \cap [n_k, n)} a_i \right) \\ &< \frac{1}{n} \left(Ln_k + L|C \cap [0, n)| + n \frac{1}{k} \right) \\ &\leq \frac{2L+1}{k}, \end{aligned}$$

o que mostra que a média parcial converge para 0.

Para mostrar a equivalência de (3), basta notar que pela equivalência de (1) e (2), e como $a_n \rightarrow 0$ se, e somente se, $a_n^2 \rightarrow 0$, pois a sequência é positiva, temos a equivalência de (1) e (3). \blacksquare

Teorema 8.5. *Seja (X, \mathcal{X}, μ, T) um sistema que preserva medida. São equivalentes*

1. T é fracamente misturadora.
2. $T \times T$ é ergódica com respeito a $\mu \times \mu$.
3. $T \times T$ é fracamente misturadora com respeito a $\mu \times \mu$.
4. Para todo sistema ergódico (Y, \mathcal{Y}, ν, S) que preserva medida, o sistema $(X \times Y, \mathcal{X} \otimes \mathcal{Y}, \mu \times \nu, T \times S)$ é ergódico.
5. Para todos $M, N \in \mathcal{X}$, existe conjunto $C \subseteq \mathbb{N}$ com densidade zero tal que $\mu(M \cap T^{-n}(N)) \rightarrow \mu(M)\mu(N)$ quando $n \rightarrow \infty$ e $n \notin C$.

6. Para todos $M, N \in \mathcal{X}$,

$$\frac{1}{n} \sum_{i=0}^{n-1} |\mu(M \cap T^{-i}(N)) - \mu(M)\mu(N)|^2 \longrightarrow 0$$

quando $n \longrightarrow \infty$.

Demonstração. A equivalência de (1),(5) e (6) é consequência do lema anterior tomando como a_i a sequência $|\mu(M \cap T^{-i}(N)) - \mu(M)\mu(N)|$. A demonstração das demais equivalências seguirá o seguinte esquema: (5) \Rightarrow (3) \Rightarrow (1) \Rightarrow (4) \Rightarrow (2) \Rightarrow (6) \Rightarrow (5).

(5) \Rightarrow (3). Sejam $M_1, M_2, N_1, N_2 \in \mathcal{X}$, existem conjuntos $C_1, C_2 \subseteq \mathbb{N}$ com densidade 0 tais que

$$\begin{aligned} \mu(M_1 \cap T^{-n}(N_1)) &\longrightarrow \mu(M_1)\mu(N_1) \quad \text{para } n \notin C_1 \\ \mu(M_2 \cap T^{-n}(N_2)) &\longrightarrow \mu(M_2)\mu(N_2) \quad \text{para } n \notin C_2. \end{aligned}$$

Seja $C := C_1 \cup C_2$. Então C tem densidade zero, pois $C \cap [0, n) = (C_1 \cap [0, n)) \cup (C_2 \cap [0, n))$, o que implica

$$\frac{1}{n} |C \cap [0, n)| \leq \frac{1}{n} (|C_1 \cap [0, n)| + |C_2 \cap [0, n)|) \longrightarrow 0.$$

Assim, temos que, para todo $n \notin C$,

$$\begin{aligned} &((\mu \times \mu)(M_1 \times M_2) \cap (T \times T)^{-n}(N_1 \times N_2) - (\mu \times \mu)(M_1 \times M_2)(\mu \times \mu)(N_1 \times N_2)) \\ &= (\mu(M_1 \cap T^{-n}(N_1))\mu(M_2 \cap T^{-n}(N_2)) - \mu(M_1)\mu(M_2)\mu(N_1)\mu(N_2)) \longrightarrow 0. \end{aligned}$$

Isso mostra que $T \times T$ é fracamente misturadora, pois os retângulos geram $\mathcal{X} \otimes \mathcal{X}$.

(3) \Rightarrow (1). Sejam $M, N \in \mathcal{X}$. Então, para $M \times X, N \times X \in \mathcal{X} \otimes \mathcal{X}$ e um conjunto com densidade zero $C \subseteq \mathbb{N}$, e para todo $n \notin C$,

$$\begin{aligned} 0 &= \lim_{n \notin C} ((\mu \times \mu)(M \times X) \cap (T \times T)^{-n}(N \times X) - (\mu \times \mu)(M \times X)(\mu \times \mu)(N \times X)) \\ &= \lim_{n \notin C} (\mu(M \cap T^{-n}(N))\mu(X \cap T^{-n}(X)) - \mu(M)\mu(X)\mu(N)\mu(X)) \\ &= \lim_{n \notin C} (\mu(M \cap T^{-n}(N)) - \mu(M)\mu(N)), \end{aligned}$$

pois $\mu(X) = 1$ e T é invariante, o que implica que T é fracamente misturadora.

(1) \Rightarrow (4). Consideremos um sistema ergódico (Y, \mathcal{Y}, ν, S) que preserva medida. Sejam $M_1, N_1 \in \mathcal{X}$ e $M_2, N_2 \in \mathcal{Y}$. Então

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} (\mu \times \nu)(M_1 \times M_2 \cap (T \times S)^{-i}(N_1 \times N_2)) \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(M_1 \cap T^{-i}(N_1)) \nu(M_2 \cap S^{-i}(N_2)) \\
&= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(M_1) \mu(M_2) \nu(M_2 \cap S^{-i}(N_2)) \\
&\quad + \frac{1}{n} \sum_{i=0}^{n-1} (\mu(M_1 \cap T^{-i}(N_1)) - \mu(M_1) \mu(M_2)) \nu(M_2 \cap S^{-i}(N_2)).
\end{aligned}$$

Mas a primeira parcela dessa soma final é $\mu(M_1) \mu(N_1) \nu(M_1) \nu(N_2)$, pois, pela ergodicidade de ν , temos que $\sum_{i=0}^{n-1} \nu(M_2 \cap S^{-i}(N_2)) \rightarrow \nu(M_2) \nu(N_2)$, e a segunda parcela da soma é zero, pois $\nu(M_2 \cap S^{-i}(N_2)) \leq 1$ é limitado e T é fracamente misturadora. Como $\mu(M_1) \mu(N_1) \nu(M_1) \nu(N_2) = (\mu \times \nu)(M_1 \times M_2) (\mu \times \nu)(N_1 \times N_2)$, temos o resultado procurado.

(4) \Rightarrow (2). Consideremos o sistema (Y, \mathcal{Y}, ν, S) , com $Y = \{y\}$ e $S = id_Y$, que é ergódico e preserva medida. Nesse caso, $T \times S$ é isomórfico a T e, então, por (4), T é ergódico. Usando (4) novamente, e tomando $Y = X$, temos que $T \times T$ é ergódico.

(2) \Rightarrow (6). Sejam $M, N \in \mathcal{X}$. Pela ergodicidade de $T \times T$, temos que

$$\begin{aligned}
& \frac{1}{n} \sum_{i=0}^{n-1} \mu(M \cap T^{-i}(N)) \\
&= \frac{1}{n} \sum_{i=0}^{n-1} ((\mu \times \mu)(M \times X \cap (T \times T)^{-i}(N \times X))) \\
&\rightarrow (\mu \times \mu)(M \times X) (\mu \times \mu)(N \times X) = \mu(M) \mu(N)
\end{aligned}$$

e, analogamente,

$$\frac{1}{n} \sum_{i=0}^{n-1} \mu(M \times T^{-i}(N))^2 \rightarrow \mu(M)^2 \mu(N)^2.$$

Disso, segue que

$$\begin{aligned}
& \frac{1}{n} \sum_{i=0}^{n-1} (\mu(M \cap T^{-i}(N)) - \mu(M)\mu(N))^2 \\
&= \frac{1}{n} \sum_{i=0}^{n-1} \mu(M \cap T^{-i}(N))^2 - 2\mu(M)\mu(N) \frac{1}{n} \sum_{i=0}^{n-1} \mu(M \cap T^{-i}(N)) + \mu(M)^2 \mu(N)^2 \frac{1}{n} \sum_{i=0}^{n-1} 1 \\
&\rightarrow \mu(M)^2 \mu(N)^2 - 2\mu(M)\mu(N)\mu(M)\mu(N) + \mu(M)^2 \mu(N)^2 = 0.
\end{aligned}$$

Isso completa a demonstração do teorema, pois já sabemos que (6) \Rightarrow (5). ■

9 O Plano Hiperbólico

Nesta seção, introduzimos um modelo do plano hiperbólico, o meio plano de Poincaré. Primeiro, consideramos um subconjunto dos números complexos. Em geral, identificaremos os números complexos com o plano real no sentido de imaginar sua representação gráfica como um plano, e nos referiremos a nomenclaturas usuais de \mathbb{R}^2 , como a ‘retas verticais’, ‘retas horizontais’ etc. Denotamos a parte real de um número complex por \Re e a parte imaginária por \Im . O *plano hiperbólico* é definido como

$$\mathbb{H} := \{z \in \mathbb{C} \mid \Im(z) > 0\}.$$

O plano hiperbólico, no entanto, não consiste somente de um conjunto, mas também de uma estrutura que permite falarmos de geometria. Em cada ponto do plano hiperbólico há um espaço tangente, que será indentificado como o conjunto dos números complexos naquele ponto. Esse é conjunto de todos os vetores tangentes a curvas passando naquele ponto. Além disso, em cada espaço tangente existe um produto interno entre os vetores tangentes nesse ponto. Esse produto interno está relacionado ao produto interno usual em \mathbb{C} , mas não é o mesmo. Ele depende do ponto e essa dependência que gerará toda a riqueza geométrica do plano hiperbólico. Antes de definirmos rigorosamente esse produto interno, primeiro definimos o *fibrado tangente* de \mathbb{H} como a união disjunta de todos os espaços tangentes

$$T\mathbb{H} := \bigsqcup_{z \in \mathbb{H}} \mathbb{C} = \mathbb{H} \times \mathbb{C}.$$

Cada espaço tangente será denotado $T_z\mathbb{H}$ e seus vetores (z, v) . Em cada $z \in \mathbb{H}$, definimos o *produto interno* em $T_z\mathbb{H}$ como

$$\langle (z, v_0), (z, v_1) \rangle_z := \frac{1}{\Im(z)^2} \langle v_0, v_1 \rangle,$$

em que $v_1, v_2 \in T_z\mathbb{H}$ e $\langle \cdot, \cdot \rangle$ é o produto interno usual de \mathbb{C} . Essa definição faz com que, quanto mais longe está z da reta horizontal $\Im = 0$, menor é o valor do produto interno no espaço tangente a z . A partir desse produto interno, podemos definir uma distância em \mathbb{H} , mas antes é preciso comentar sobre funções e diferenciabilidade em \mathbb{H} . Consideraremos uma função em \mathbb{H} como uma função em \mathbb{C} , de modo que as derivadas também são as derivadas de uma função complexa. Sendo assim, definimos o *comprimento* de uma curva $\gamma :]a, b[\rightarrow \mathbb{H}$ em \mathbb{H} por

$$C(\gamma) := \int_a^b \|\gamma'(t)\|_z dt$$

em que $\gamma'(t) \in T_{\gamma(t)}\mathbb{H}$ e $\|\cdot\|_z$ é a norma induzida pelo produto interno $\langle \cdot, \cdot \rangle_z$. A partir do comprimento de curvas, dados dois pontos $z_0, z_1 \in \mathbb{H}$, a *distância* entre eles é

$$d(z_0, z_1) := \inf C(\gamma),$$

em que γ é qualquer curva contínua por partes em \mathbb{H} iniciada em z_0 e terminada em z_1 . Quando a integral diverge, a distância é entendida como ‘infinita’. Ainda não será possível exibir quais são as curvas que minimizam a distância entre dois pontos quaisquer, nem mesmo se elas existem. Mais à frente, faremos isso usando ferramentas — como ação de grupos, por exemplo — que facilitarão o entendimento e estudo da teoria.

10 Ação de Grupos no Plano Hiperbólico

Para estudar fluxos melhor a dinâmica no plano hiperbólico, introduzimos alguns grupos e o conceito de ação de grupos.

Definição 10.1. Sejam X um conjunto e G um grupo. Uma *ação* de G em X é uma função

$$\begin{aligned} A : G \times X &\rightarrow X \\ (g, x) &\mapsto g \cdot x \end{aligned}$$

que satisfaz

1. (Identidade) Para todo $x \in X$,

$$e \cdot x = x;$$

2. (Compatibilidade) Para todos $g_0, g_1 \in G$ e $x \in X$,

$$(g_1 g_0) \cdot x = g_1 \cdot (g_0 \cdot x).$$

Um grupo G age em o conjunto X se, e somente se, existe uma ação de G em X . Em geral, exigimos que a ação seja um morfismo compatível com o espaço, no sentido em que se X é um espaço topológico, então deve ser uma função contínua etc.

Por G ser um grupo, temos que ao fixarmos um elemento agindo, a função obtida é uma bijeção, como mostra a simples proposição a seguir.

Proposição 10.1. *Sejam X um conjunto, G um grupo, A uma ação de G em X e $g \in G$. Então a função*

$$\begin{aligned} g \cdot : X &\rightarrow X \\ x &\mapsto g \cdot x \end{aligned}$$

é uma bijeção. Equivalentemente, $g \cdot \in \mathfrak{S}_X$.

Proposição 10.2. *Sejam X um conjunto, G um grupo que age em X . Então a função*

$$\begin{aligned} h : G &\rightarrow \mathfrak{S}_X \\ g &\mapsto g \cdot : X \rightarrow X \\ &x \mapsto g \cdot x. \end{aligned}$$

é um homomorfismo de grupos de G para \mathfrak{S}_X , o grupo simétrico de X (conjunto das bijeções de X para X).

Demonstração. Sejam $g_1, g_0 \in G$ e $x \in X$. Então

$$h(g_1 g_0)(x) = (g_1 g_0) \cdot x = g_1 \cdot (g_0 \cdot x) = h(g_1)(h(g_0)(x)) = h(g_1) \circ h(g_0)(x),$$

logo $h(g_1 g_0) = h(g_1) \circ h(g_0)$. ■

Definição 10.2. Sejam X um conjunto, G um grupo que age em X e $x \in X$. A órbita de x sob G é o conjunto

$$G \cdot x := \{g \cdot x \mid g \in G\}.$$

Os grupos que estudaremos são o grupo das matrizes 2×2 reais com determinante unitário, o grupo linear especial

$$SL_2(\mathbb{R}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R}, \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = 1 \right\},$$

e o grupo projetivo de $SL_2(\mathbb{R})$, que resulta de identificar as matrizes Id e $-Id$ e quocientar por essa relação de equivalência, que é denotado

$$PSL_2(\mathbb{R}) := SL_2(\mathbb{R}) / \{Id, -Id\}.$$

Vamos mostrar que $PSL_2(\mathbb{R})$ age em \mathbb{H} por

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}.$$

Notemos que, se $cz + d = 0$ não ocorre, pois, como $z \neq 0$, teríamos que $c = d = 0$, o que é absurdo pois o determinante é unitário, ou $z = -\frac{d}{c}$, que é um número real, o que também não ocorre pois $z \in \mathbb{H}$ sempre tem parte imaginária estritamente positiva. Ainda, é fácil notar que $g \cdot z = \frac{az+b}{cz+d} \in \mathbb{H}$, pois a ação é contínua, logo leva conexo em conexo, e como os números reais, a fronteira de \mathbb{H} , são levados neles mesmos, basta notar que algum elemento de \mathbb{H} é levado em \mathbb{H} , o que é simples. Uma relação importante de se ter antes de continuar, e que também mostra que $g \cdot z \in \mathbb{H}$, é a seguinte:

$$\begin{aligned} \Im \left(\frac{az + b}{cz + d} \right) &= \Im \left(\frac{a(x + yi) + b}{c(x + yi) + d} \right) \\ &= \Im \left(\frac{(ax + b) + (ay)i}{(cx + d) + (cy)i} \right) \\ &= \Im \left(\frac{(ax + b + ayi)(cx + d - cyi)}{(cx + d)^2 + (cy)^2} \right) \\ &= \frac{y}{\|cz + d\|^2}. \end{aligned}$$

Agora, mostremos que essa relação estabelece de fato uma ação. Primeiro, tomando $z \in \mathbb{H}$, temos que

$$Id \cdot z = \frac{1z + 0}{0z + 1} = z.$$

Segundo, sejam $g_1 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix}$ e $g_0 = \begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix}$ elementos de $SL_2(\mathbb{R})$. Então, para todo $z \in \mathbb{H}$, vale que $g_1 g_0 \cdot z = g_1 \cdot (g_0 \cdot z)$, pois

$$g_1 g_0 = \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix} = \begin{pmatrix} a_1 a_0 + b_1 c_0 & a_1 b_0 + b_1 d_0 \\ c_1 a_0 + d_1 c_0 & c_1 b_0 + d_1 d_0 \end{pmatrix}$$

e

$$\begin{aligned} g_1 \cdot (g_0 \cdot z) &= g_1 \cdot \frac{a_0 z + b_0}{c_0 z + d_0} \\ &= \frac{a_1 \left(\frac{a_0 z + b_0}{c_0 z + d_0} \right) + b_1}{c_1 \left(\frac{a_0 z + b_0}{c_0 z + d_0} \right) + d_1} \\ &= \frac{(a_1 a_0 + b_1 c_0)z + (a_1 b_0 + b_1 d_0)}{(c_1 a_0 + d_1 c_0)z + (c_1 b_0 + d_1 d_0)}. \end{aligned}$$

No entanto, notamos também que $-Id = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ também age trivialmente em z , pois

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \cdot z = \frac{-1z + 0}{0z - 1} = z$$

e segue que $PSL_2(\mathbb{R})$ age em \mathbb{H} .

Vamos, agora, provar uma primeira proposição que mostra como o estudo da ação desse grupo em \mathbb{H} interage com a geometria do espaço. Mas antes alguns comentários sobre a derivada das ações de elementos do grupo. Se pensamos, para cada $g \in PSL_2(\mathbb{R})$, na função complexa $g : \mathbb{C} \rightarrow \mathbb{C}$, dada por $g(z) = g \cdot z$, ela é uma função diferenciável e

$$g'(z) = \left(\frac{az + b}{cz + d} \right)' = \frac{(cz + d)a - (az + b)c}{(cz + d)^2} = \frac{1}{(cz + d)^2},$$

pois $ad - bc = 1$ é o determinante de g . Sendo assim, definimos a derivada de g por

$$Dg : T\mathbb{H} \rightarrow T\mathbb{H}$$

$$(z, v) \mapsto (g(z), g'(z)v) = \left(\frac{az + b}{cz + d}, \frac{v}{(cz + d)^2} \right).$$

Essa função é linear para um $z \in \mathbb{H}$ fixo: $(Dg)_z : T_z\mathbb{H} \rightarrow T_{g(z)}\mathbb{H}$. Pela regra da cadeia, obtemos que (1) $DId(z, v) = (z, v)$ e, a partir da relação $(g_1g_0) \cdot z = g_1 \cdot (g_0 \cdot z)$, (2) $(g_1g_0)' \cdot z = g_1' \cdot (g_0' \cdot z)$, o que mostra que a derivada também nos dá uma ação de grupo.

Proposição 10.3. *A ação de $PSL_2(\mathbb{R})$ em \mathbb{H} é isométrica; ou seja, para todos $z_0, z_1 \in \mathbb{H}$ e todo $g \in PSL_2(\mathbb{R})$,*

$$d(g \cdot z_0, g \cdot z_1) = d(z_0, z_1).$$

Ainda, a ação de Dg em $T\mathbb{H}$ preserva o produto interno em cada espaço tangente.

Demonstração. Para demonstrar a primeira afirmação, demonstraremos a segunda. Pela construção da distância em \mathbb{H} , segue que se o produto interno em cada espaço tangente é preservado, então a distância em \mathbb{H} também será preservada. Para isso, sejam $(z, v), (z, w) \in T_z\mathbb{H}$. Então $(Dg)_zv, (Dg)_zw \in T_{g(z)}\mathbb{H}$ e segue que, usando a relação $\Im(g \cdot z) = \Im(z) / \|cz + d\|^2$ obtida anteriormente,

$$\begin{aligned} \langle (Dg)_zv, (Dg)_zw \rangle_{g(z)} &= \frac{1}{\Im(g \cdot z)} \left\langle \frac{v}{\|cz + d\|^2}, \frac{w}{\|cz + d\|^2} \right\rangle \\ &= \frac{1}{\Im(z)} \langle v, w \rangle \\ &= \langle v, w \rangle_z. \end{aligned}$$

Assim mostramos que o produto interno é preservado e, portanto, que a distância também o é. ■

Mostraremos agora que a ação de $PSL_2(\mathbb{R})$ é transitiva. Para isso, esclarecemos a definição de uma ação transitiva.

Definição 10.3. Sejam X um conjunto e G um grupo. Uma ação *transitiva* de G em X é uma ação de G em X tal que, para todos $z_0, z_1 \in \mathbb{H}$, existe $g \in G$ tal que $g \cdot z_0 = z_1$.

Proposição 10.4. A ação de $PSL_2(\mathbb{R})$ em \mathbb{H} é transitiva.

Demonstração. Para mostrar isso, mostraremos que i pode ser levado em qualquer $z = x + yi$, com $y > 0$. Isso implicará que, se $g_0 \cdot i = z_0$ e $g_1 \cdot i = z_1$, então $g_0 g_1^{-1} \cdot z_1 = z_0$. Para isso, tomemos a matriz $\begin{pmatrix} \sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix}$, que está $PSL_2(\mathbb{R})$ pois tem determinante igual a 1 e entradas reais. Claramente,

$$\begin{pmatrix} \sqrt{y} & x/\sqrt{y} \\ 0 & 1/\sqrt{y} \end{pmatrix} \cdot i = \frac{\sqrt{y}i + x/\sqrt{y}}{1/\sqrt{y}} = x + yi.$$

■

Acharemos agora um subgrupo de $PSL_2(\mathbb{R})$ chamado subgrupo estabilizador de i . Para isso, brevemente definimos antes o conceito.

Definição 10.4. Sejam X um conjunto, G um grupo que age em X e $x \in X$. O *subgrupo estabilizador* (ou *grupo de isotropia*) de G com respeito a x é

$$G_x := \{g \in G \mid g \cdot x = x\}.$$

Definimos ainda o grupo ortogonal especial

$$SO_2 := \left\{ \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}$$

e o grupo ortogonal especial projetivo $PSO_2 := SO_2 / \{Id, -Id\}$.

Proposição 10.5. O subgrupo estabilizador de $PSL_2(\mathbb{R})$ em i é

$$PSL_2(\mathbb{R})_i = PSO_2.$$

Demonstração. Se $g \cdot i = i$, então $\|d + ci\| = 1$, pois $\Im(g \cdot z) = \Im(z) / \|cz + d\|^2$. Então existe $\theta \in \mathbb{R}$ com $d = \cos \theta$. $c = \pm \sin \theta$, logo

$$g \cdot i = \frac{ai + b}{\sin \theta i + \cos \theta} = i \Leftrightarrow ai + b = \cos \theta i - \sin \theta,$$

então $a = \cos \theta$ e $b = -\sin \theta$. ■

A última proposição dessa seção é sobre a ação derivada de $PSL_2(\mathbb{R})$ em $\mathbb{T}\mathbb{H}$ e em $\mathbb{T}^1\mathbb{H}$. Como mostrado na proposição, existe uma ação em $\mathbb{T}\mathbb{H}$ a partir da derivada da ação em \mathbb{H} . Como essa ação preserva o produto interno, ela está bem definida em $\mathbb{T}^1\mathbb{H}$ também. Essa ação é simplesmente transitiva, como definimos a seguir.

Definição 10.5. Sejam X um conjunto e G um grupo. Uma ação *simplesmente transitiva* de G em X é uma ação transitiva de G em X que também é *livre*, ou seja, para todos $g \in G$, se existe $x \in X$ tal que $g \cdot x = x$, então $g = Id$.

Proposição 10.6. A ação de $PSL_2(\mathbb{R})$ em $\mathbb{T}^1\mathbb{H}$ é simplesmente transitiva.

Demonstração. A demonstração é simples e pode ser achada em [2], p.281. ■

11 Geodésicas e Fluxos no Plano Hiperbólico

Podemos agora achar as geodésicas do plano hiperbólico, que foram definidas na seção §9, a partir da ação de $PSL_2(\mathbb{R})$ em $|Hp$. Para isso, primeiro mostraremos que as geodésicas entre pontos imaginários puros são retas com uma parametrização específica, e então levaremos essas retas em outras geodésicas entre pontos quaisquer através da ação do grupo. Para isso, provamos a próxima proposição.

Proposição 11.1. Sejam $z_0 = y_0i$ e $z_1 = y_1i$, com $0 < y_0 < y_1$. Então

$$d(z_0, z_1) = \log y_1 - \log y_0$$

e

$$\begin{aligned} \phi : [0, 1] &\rightarrow \mathbb{H} \\ t &\mapsto y_0 \left(\frac{y_1}{y_0} \right)^t i \end{aligned}$$

é a geodésica de z_0 a z_1 com velocidade constante $\log y_1 - \log y_0$.

Demonstração. A velocidade de ϕ é claramente constante igual a $\log y_1 - \log y_0$. Portanto

$$d(z_0, z_1) \leq \log y_1 - \log y_0.$$

Seja $\eta : [0, 1] \rightarrow \mathbb{H}$ caminho de z_0 a z_1 , sendo $\eta(t) = \eta_x(t) + \eta_y(t)i$. Então

$$\begin{aligned} C(\eta) &= \int_0^1 \frac{\|\eta'(t)\|}{\eta_y(t)} dt \\ &\geq \int_0^1 \frac{\|\eta'_y(t)\|}{\eta_y(t)} dt \\ &\geq \int_0^1 \frac{\eta'_y(t)}{\eta_y(t)} dt \\ &= \log y_1 - \log y_0, \end{aligned}$$

sendo que a primeira desigualdade é uma igualdade se, e somente se, $\eta'_x = \eta_x = 0$, e a segunda desigualdade é uma igualdade se, e somente se, $\eta'_y \geq 0$. Isso mostra que ϕ é a geodésica procurada e que seu comprimento é $\log y_1 - \log y_0$. ■

As geodésicas com velocidade constante igual a 1 são, portanto, parametrizadas por

$$\begin{aligned} \phi : [0, d(z_0, z_1)] &\rightarrow \mathbb{H} \\ t &\mapsto y_0 e^{ti}. \end{aligned}$$

Como sabemos, cada ação de $g \in PSL_2(\mathbb{R})$ leva geodésicas em geodésicas, pois é uma isometria. A seguir, provamos que a volta também é válida.

Proposição 11.2. *Sejam $z_0, z_1 \in \mathbb{H}$. Existe um único caminho $\phi : [0, d(z_0, z_1)] \rightarrow \mathbb{H}$ com $\phi(0) = z_0$, $\phi(d(z_0, z_1)) = z_1$ e $\phi' = 1$. Ainda, existe única isometria $g \in PSL_2(\mathbb{R})$ tal que $\phi(t) = g(e^t i)$.*

Demonstração. Existe $\tilde{g} \in PSL_2(\mathbb{R})$ com $\tilde{g}^{-1}(z_0) = i$. Ainda, sabemos que $h \in PSO_2$ fixa i . Portanto, escolhendo um h que maximiza $\Im(PSO_2 \tilde{g}^{-1}(z_1))$, definimos $g := h\tilde{g}^{-1}$. Segue que g é único pois $g^{-1}g_1$ vai de i a $y_1 i$, logo $y \mapsto g^{-1}g_1 e^{ti}$, com velocidade constante igual a 1, logo esse mapa é $t \mapsto e^{ti}$. Derivando em $t = 0$,

$$Dg^{-1}g_1(i, i) = (i, i),$$

logo $g^{-1}g_1 = Id$. ■

Uma transformação $g \in PSL_2(\mathbb{R})$ leva uma curva no eixo $\{yi \mid y > 0\}$ em outra reta vertical ou em um semi-círculo

$$(x - f)^2 + y^2 = r^2$$

com centro em $\partial\mathbb{H}$. Para confirmar isso, basta ver que $z \mapsto az + b$ leva o eixo $\{yi \mid y > 0\}$ na reta $\Re(z) = b$, e leva $(x - f)^2 + y^2 = r^2$ em $(x - (af + b))^2 + y^2 = a^2 r^2$,

Ainda, a transformação $z \mapsto \frac{1}{z}$ leva cada um dos casos no outro, pois se olharmos em coordenadas polares, temos $(r, \theta) \mapsto (\frac{1}{r}, \pi - \theta)$, e tanto retas como círculos de centro real são da forma

$$\alpha r^2 + \beta r \cos \phi + \gamma = 0$$

com $(\alpha, \beta, \gamma) \neq 0$. Como qualquer $g \in PSL_2(\mathbb{R})$ é gerado por $z \mapsto az + b$ e $z \mapsto \frac{1}{z}$, temos o que procurávamos.

12 Os Fluxos Geodésico e Horociclo

Denotaremos o espaço dos vetores tangentes a \mathbb{H} unitários como

$$T^1\mathbb{H} := \{(z, v \in T\mathbb{H}) \mid \|v\|_z = 1\}.$$

Geodésicas são curvas com velocidade constante e por isso podemos estudar as geodésicas com velocidade constante unitária, de modo que fluxos geodésicos serão definidos usando esse espaço agora definido. Antes de comentar sobre os fluxos geodésico e horociclo, definimos o que é um fluxo.

Definição 12.1. Seja X um conjunto. Um *fluxo* em X é uma ação $\varphi : (\mathbb{R}, +) \curvearrowright X$ do grupo aditivo dos números reais em X . Denota-se $\varphi^t(x)$ para a ação de $t \in \mathbb{R}$ em $x \in X$. Em geral, quando X tem uma estrutura adicional exige-se que o fluxo preserve essa estrutura.

12.1 Fluxo Geodésico

O fluxo geodésico é o fluxo em $X = T^1\mathbb{H}$ que segue as geodésicas em \mathbb{H} a velocidade unitária. Dado $(z, v) \in T^1\mathbb{H}$, existe uma única geodésica que sai de z com direção v , e seguindo essa geodésica por tempo t temos o fluxo geodésico. Para cada $t \in \mathbb{R}$, definimos o fluxo $\varphi = G$ do seguinte modo. Definimos $G^t(i, i) = (e^t i, e^t i)$, pois a geodésica vertical no eixo imaginário é essa em tempo t . Note que $(e^t i, e^t i) \in T^1\mathbb{H}$, já que

$$\|e^t i\|_{e^t i}^2 = \frac{1}{(e^t)^2} \langle e^t i, e^t i \rangle = 1.$$

Para definirmos o fluxo em todo $(z, v) \in T^1\mathbb{H}$, definimos primeiro $a_t := \begin{pmatrix} e^{-t/2} & 0 \\ 0 & e^{t/2} \end{pmatrix}$ e notamos que

$$D(a_t^{-1})(i, i) = D \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} (i, i) = \left(\frac{e^{t/2} i}{e^{-t/2}}, \frac{i}{(e^{-t/2})^2} \right) = (e^t i, e^t i) = G^t(i, i).$$

Agora, para $(z, v) \in \mathbb{T}^1\mathbb{H}$, tomamos $g \in PSL_2(\mathbb{R})$ tal que $(z, v) = Dg(i, i)$, que existe pois a ação derivada é transitiva, e definimos

$$G^t(z, v) := Dg(G^t(i, i)) = Dg(Da_t^{-1}(i, i)) = D(ga_t^{-1})(i, i).$$

Definindo, por fim, $R_{a_t}(g) = ga_t^{-1}$ para qualquer $g \in PSL_2(\mathbb{R})$, temos a definição final.

Definição 12.2. O *fluxo geodésico* em $\mathbb{T}^1\mathbb{H}$ é o fluxo definido por

$$\begin{aligned} G^t : \mathbb{T}^1\mathbb{H} &\rightarrow \mathbb{T}^1\mathbb{H} \\ (z, v) &\mapsto D(R_{a_t}(g))(i, i), \end{aligned}$$

para $t \in \mathbb{R}$ e $g \in PSL_2(\mathbb{R})$ tal que $Dg(i, i) = (z, v)$.

Mostremos que isso é uma ação. Primeiro, mostramos a identidade

$$G^0(z, v) = D(R_{a_0}(g))(i, i) = D(gId)(i, i) = Dg(i, i) = (z, v).$$

Agora, notemos de

$$a_t a_s = \begin{pmatrix} e^{-t/2} & 0 \\ 0 & e^{t/2} \end{pmatrix} \begin{pmatrix} e^{-s/2} & 0 \\ 0 & e^{s/2} \end{pmatrix} = \begin{pmatrix} e^{-(t+s)/2} & 0 \\ 0 & e^{(t+s)/2} \end{pmatrix} = a_{t+s}$$

segue que $R_{a_{t+s}} = R_{a_t} R_{a_s}$. Assim, temos a compatibilidade

$$\begin{aligned} G^{t+s}(z, v) &= D(R_{a_{t+s}}(g))(i, i) \\ &= D(R_{a_t} R_{a_s}(g))(i, i) \\ &= G^t(D(R_{a_s}(g))(i, i)) \\ &= G^t(G^s(z, v)). \end{aligned}$$

A notação $R_{a_t}(g)$ indica uma ação de \mathbb{R} em $PSL_2(\mathbb{R})$. Claramente vale a identidade e já mostramos acima a compatibilidade. O fluxo geodésico é, de fato, a ação derivada dessa ação. Podemos expressar explicitamente o fluxo geodésico escrevendo $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, sendo $Dg(i, i) = (z, v)$, e calculando

$$ga_t^{-1} = \begin{pmatrix} ae^{t/2} & be^{-t/2} \\ ce^{t/2} & de^{-t/2} \end{pmatrix}.$$

Segue assim que

$$\begin{aligned} G^t(z, v) &= D(R_{a_t}(g))(i, i) \\ &= \left(\frac{ae^{t/2}i + be^{-t/2}}{ce^{t/2}i + de^{-t/2}}, \frac{i}{(ce^{t/2}i + de^{-t/2})^2} \right) \\ &= \left(\frac{ae^t i + b}{ce^t i + d}, \frac{e^t i}{(ce^t i + d)^2} \right). \end{aligned}$$

12.2 Fluxo Horociclo

O fluxo horociclo está relacionado às variedades estáveis e instáveis de \mathbb{H} . Para definir precisamente esses conceitos precisamos definir uma métrica em $T^1\mathbb{H}$, mas não faremos isso neste trabalho. No entanto, algumas ideias de como o fluxo horociclo é serão passadas. Primeiro consideramos os pontos da forma $(s + i, i)$, para $s \in \mathbb{R}$. As geodésicas por eles são retas verticais paralelas dadas por

$$\begin{aligned} G^t(i, i) &= (e^t i, e^t i) \\ G^t(s + i, i) &= (s + e^t i, e^t i) \end{aligned}$$

para $t \in \mathbb{R}$. Sendo assim, a distância hiperbólica entre esses pontos não ultrapassa $\frac{|s|}{e^t}$, já que essa é a distância horizontal entre (i, i) e $(s + i, i)$. Isso sugere que esses pontos se encontram quando $t \rightarrow \infty$, ou seja, que o limite da distância entre os dois é zero, mas como comentado acima essa noção não será definida rigorosamente aqui. De qualquer modo, consideraremos a ação derivada de $\varphi^s = u^-(s)$, sendo

$$u^-(s) := \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}.$$

Essas matrizes formam um subgrupo de $PSL_2(\mathbb{R})$. Note que $Du^-(s)(i, i) = (s + i, i)$, e isso nos dá todos os vetores $(z, v) \in T^1\mathbb{H}$ com $\Im(z) = 1$ e $v = i$, que é a variedade estável em (i, i) , os pontos (z, v) tais que suas órbitas sob o fluxo geodésico convergem. Similarmente ao fluxo geodésico, estendemos o fluxo horociclo pela ação derivada para todo $(z, v) \in T^1\mathbb{H}$, definido

$$R_{u^-(s)}(g) := gu^-(-s).$$

Definição 12.3. O *fluxo horociclo estável* em $T^1\mathbb{H}$ é o fluxo definido por

$$\begin{aligned} u^-(s) : T^1\mathbb{H} &\rightarrow T^1\mathbb{H} \\ (z, v) &\mapsto D(R_{u^-(s)}(g))(i, i), \end{aligned}$$

para $s \in \mathbb{R}$ e $g \in PSL_2(\mathbb{R})$ tal que $Dg(i, i) = (z, v)$.

Analogamente, pode-se verificar que o subgrupo das matrizes

$$u^-(s) := \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix}$$

dá as variedades instáveis e o fluxo geodésico instável.

13 Conclusão

Nesta segunda etapa da Iniciação Científica, o estudo foi focado em um único espaço: o semiplano superior de \mathbb{C} , um modelo do plano hiperbólico. Através desse exemplo, foram introduzidos os conceitos de variedades diferenciais, planos e fibrado tangente, métricas em variedades diferenciais (variedades riemannianas); o conceito de ação de grupo, incluindo o de ação contínua/suave (que por sua vez exige a definição de um grupo topológico/de Lie, que não foram tratados neste trabalho) e conceitos relacionados, como órbitas de grupo, subgrupos estabilizadores; e por fim o conceito de fluxo, que é essencial no estudo de sistemas dinâmicos. Todos os conceitos introduzidos foram voltados a esse exemplo do plano hiperbólico, mas serão explorados a fundo e de modo mais genérico no curso de metrado, nas disciplinas de topologia diferencial e geometria riemanniana, principalmente, e por isso contribuíram para a formação matemática geral do aluno.

Referências

- [1] Bartle, R. G. *The Elements of Integrations*. John Wiley and Sons, Inc. 1966 7
- [2] Einsiedler, Manfred, and Ward, Thomas. *Ergodic Theory with a view towards Number Theory*. Springer-Verlag London Limited 2011 4, 8, 10, 18, 20, 21, 33
- [3] Bekka, M. Bachir, and Mayer, Matthias *Ergodic Theory and Topological Dynamics of Group Actions on Homogeneous Spaces*. Cambridge University Press 2000