V BIENAL DA SOCIEDADE BRASILEIRA DE MATEMÁTICA

Uma Introdução à Teoria Ergódica

Carlos Bocker - UFPA
Krerley Oliveira - UFAL
Marcelo Viana - IMPA

18 a 22 de Outubro de 2010
Universidade Federal da Paraíba - João Pessoa - PB

Conteúdo

Prefácio

Em termos simples, a Teoria Ergódica é a disciplina matemática que estuda sistemas dinâmicos munidos de medidas invariantes. Começaremos por dar as definições precisas destas noções e algumas das principais motivações para o seu estudo. No final deste prefácio faremos alguns comentários sobre a história desta disciplina.

Sistemas dinâmicos

Há várias definições, mais ou menos gerais, do que é um sistema dinâmico. Nós nos restringiremos a dois modelos principais. O primeiro deles, ao qual nos referiremos na maior parte do tempo, são as transformações $f : M \to M$ em algum espaço métrico ou topológico M. Heuristicamente, pensamos em f como associando a cada estado $x \in M$ do sistema o estado $f(x) \in M$ em que o sistema se encontrará uma unidade de tempo depois. Trata-se portanto de um modelo de dinâmica com tempo discreto.

Também consideraremos fluxos, que são modelos de sistemas dinâmicos com tempo contínuo. Lembre que um fluxo em M é uma família $f^t : M \to M$, $t \in \mathbb{R}$ de transformações satisfazendo

$$f^0 = \text{ identidade} \quad \text{e} \quad f^t \circ f^s = f^{t+s} \text{ para todo } t, s \in \mathbb{R}. \tag{1}$$

Fluxos aparecem, por exemplo, associados a equações diferenciais: tome como f^t a transformação que associa a cada ponto x o valor no tempo t da solução da equação que passa por x no tempo zero.

Num caso e no outro, sempre iremos supor que o sistema dinâmico é pelo menos mensurável: na maior parte dos casos será até contínuo, ou mesmo diferenciável.

Medidas invariantes

Sempre consideraremos medidas μ definida na σ-álgebra de Borel do espaço M. Dizemos que μ é uma probabilidade se $\mu(M) = 1$. Na maior parte dos casos trataremos com medidas finitas, isto é, tais que $\mu(M) < \infty$. Neste caso sempre podemos transformar μ numa probabilidade ν: para isso basta definir

$$\nu(E) = \frac{\mu(E)}{\mu(M)} \quad \text{para cada conjunto mensurável } E \subset M.$$

Em geral, uma medida μ diz-se invariante pela transformação f se

$$\mu(E) = \mu(f^{-1}(E)) \quad \text{para todo conjunto mensurável } E \subset M. \tag{2}$$

Heuristicamente, isto significa que a probabilidade de um ponto estar num dado conjunto e a probabilidade de que a sua imagem esteja nesse conjunto são iguais. Note que a definição (2) faz sentido, uma vez que a pré-imagem de um conjunto mensurável por uma transformação mensurável ainda é um conjunto mensurável.

No caso de fluxos, substituímos a relação (2) por

$$\mu(E) = \mu(f^{-t}(E)) \quad \text{para todo mensurável } E \subset M \text{ e todo } t \in \mathbb{R}. \tag{3}$$

Por que estudar medidas invariantes ?

Como em todo ramo da Matemática, parte importante da motivação é intrínseca e estética: estas estruturas matemáticas têm propriedades profundas e surpreendentes que conduzem à demonstração de belíssimos teoremas. Igualmente fascinante, idéias e resultados da Teoria Ergódica se aplicam em outras áreas da Matemática que a priori nada têm de probabilístico, por exemplo a Combinatória e a Teoria dos Números.

Outra razão para este estudo é que muitos fenômenos importantes na Natureza e nas ciências experimentais são modelados por sistemas dinâmicos que deixam invariante alguma medida interessante. O exemplo mais importante, historicamente, veio da Física: sistemas hamiltonianos, que descrevem a evolução de sistemas conservativos na mecânica newtoniana, correspondem a fluxos que preservam uma medida natural, a medida de Liouville. Aliás veremos que sistemas dinâmicos muito gerais possuem medidas invariantes.

Ainda outra motivação fundamental para que nos interessemos por medidas invariantes é que o seu estudo pode conduzir a informação importante sobre o comportamento dinâmico do sistema, que dificilmente poderia ser obtida de outro modo. O Teorema de Recorrência de Poincaré é uma excelente ilustração do que acabamos de dizer: ele afirma que a órbita de quase todo ponto, relativamente a qualquer medida invariante finita, regressa arbitrariamente perto do ponto inicial.

Breve apresentação histórica

A palavra *ergódico* é a concatenação de duas palavras gregas, *ergos* = trabalho e *odos* = caminho, e foi introduzida pelo físico L. Boltzmann, no século XIX, no seu trabalho sobre a teoria cinética dos gases. Os sistemas em que L. Boltzmann, J. C. Maxwell, J. C. Gibbs, os principais fundadores da teoria cinética, estavam interessados são descritos por um fluxo hamiltoniano, ou seja, uma equação diferencial da forma

$$\left(\frac{dq_1}{dt}, \dots, \frac{dq_n}{dt}, \frac{dp_1}{dt}, \dots, \frac{dp_n}{dt}\right) = \left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \dots, -\frac{\partial H}{\partial q_n}\right).$$

Boltzmann acreditava que as órbitas típicas do fluxo preenchem toda a superfície de energia $H^{-1}(c)$ que as contém. A partir desta *hipótese ergódica*, ele deduziu que as médias temporais de grandezas observáveis (funções) ao longo de órbitas típicas coincidem com as respectivas médias espaciais na superfície de energia, um fato crucial para a sua formulação da teoria cinética.

De fato, esta hipótese é claramente falsa e, com o tempo, tornou-se usual chamar hipótese ergódica a sua consequência de que as médias temporais e espaciais são iguais. Sistemas para os quais vale esta igualdade foram chamados *ergódicos*. E pode dizer-se que uma boa parte da Teoria Ergódica, tal como ela se desenvolveu ao longo do século XX, foi motivada pelo problema de decidir se a maioria dos sistemas hamiltonianos, especialmente aqueles que aparecem na teoria cinética dos gases, são ergódicos ou não.

Um avanço fundamental ocorreu nos anos trinta, quando os matemáticos J. von Neumann e G. D. Birkhoff provaram que médias temporais existem para quase toda órbita. No entanto, em meados dos anos cinquenta, o grande matemático russo A. N. Kolmogorov observou que muitos sistemas hamiltonianos *não são* ergódicos. Este resultado espectacular foi muito expandido por V. Arnold e por J. Moser, no que veio a ser chamado teoria KAM em homenagem aos três.

Por outro lado, ainda nos anos trinta, E. Hopf tinha dado os primeiros exemplos importantes de sistemas hamiltonianos ergódicos, os fluxos geodésicos de superfícies com curvatura negativa. O seu resultado foi generalizado por D. Anosov, nos anos sessenta, para variedades de qualquer dimensão. De fato, Anosov tratou uma classe bem mais geral de sistemas, tanto com tempo contínuo como com tempo discreto, que são chamados sistemas de Anosov, ou globalmente hiperbólicos. Uma classe ainda mais ampla de sistemas, chamados uniformemente hiperbólicos, foi introduzida por S. Smale, e constituiu um importante foco da teoria dos Sistemas Dinâmicos ao longo das últimas décadas.

Nos anos setenta, Ya. Sinai desenvolveu a teoria das medidas de Gibbs dos sistemas de Anosov, conservativos ou dissipativos, que foi logo em seguida estendida por D. Ruelle e por R. Bowen para sistemas uniformemente hiperbólicos, constituindo uma das maiores realizações da teoria ergódica diferenciável. Não podemos deixar de mencionar, nesta breve lista de contribuições fundamentais, a introdução da noção de *entropia* por Kolmogorov e Sinai no final dos anos cinquenta, e a demonstração, por D. Ornstein cerca de dez anos depois, de que a entropia é um invariante completo para deslocamentos ("shifts") de Bernoulli.

Organização deste livro

Boa parte dos avanços alcançados em Sistemas Dinâmicos e Teoria Ergódica nas últimas décadas está ligados ao objetivo maior de estender para grande generalidade o escopo dos resultados que vimos mencionando, almejando obter uma descrição global e coerente do modo como sistemas dinâmicos evoluem.

Neste livro buscamos apresentar ao leitor, de forma direta e sucinta, os principais conceitos e resultados que permitem ter acesso direto aos avanços mais recentes nesta área da Matemática, visando tanto uma carreira de pesquisa quanto a efetiva utilização destas ferramentas nas diversas áreas de aplicações práticas.

O texto se estende ao longo de 4 capítulos, que podem ser estruturados do seguinte modo. Os Capítulos 2 a 4 formam uma espécie de ciclo básico, no qual apresentamos as noções e resultados fundamentais da Teoria Ergódica, bem como os principais exemplos que motivam a teoria. Os exemplos são introduzidos gradativamente, buscando para cada um o contexto que melhor realça a sua relevância.

Krerley Oliveira [1] e Marcelo Viana [2]

[1]Departamento de Matemática, Universidade Federal de Alagoas, Campus A. C. Simões s/n, 57072-090 Maceió, Brasil. krerley@mat.ufal.br.

[2]IMPA, Estrada D. Castorina 110, 22460-320 Rio de Janeiro, Brasil viana@impa.br.

Capítulo 1

Noções Básicas de Teoria da Medida

Neste capítulo vamos apresentar apenas alguns elementos básicos de Teoria da Medida que são úteis para o que segue. Para um estudo mais profundo do tema consultar os livros de Castro [Cas04], Fernandez [Fer02] ou Rudin [Rud87].

1.1 Espaços mensuráveis

Uma noção básica na Teoria da Medida é a noção de *álgebra*, e por extensão, a noção de σ-*álgebra* de subconjuntos. Começamos por introduzi-las e estudar algumas de suas propriedades. Em seguida definimos espaços mensuráveis e apresentamos uma técnica de construção de σ-álgebras. Dado um subconjunto $A \subset X$ denotaremos por A^c o complementar $X \setminus A$ do conjunto A em relação à X.

Definição 1.1. Uma *álgebra* de X é uma família $\mathcal{B}$ de subconjuntos de X que é fechada para as operações elementares de conjuntos e contém X. Isto é:

i) $X \in \mathcal{B}$

ii) $A \in \mathcal{B}$ implica $A^c \in \mathcal{B}$

iii) $A \in \mathcal{B}$ e $B \in \mathcal{B}$ implica $A \cap B \in \mathcal{B}$.

Observe que $A \cup B = (A^c \cap B^c)^c$ e $A \setminus B = A \cap B^c$ também estão em $\mathcal{B}$, quaisquer que sejam $A, B \in \mathcal{B}$. Além disso, por associatividade, a união e a intersecção de qualquer número finito de elementos de $\mathcal{B}$ também estão em $\mathcal{B}$.

Observação 1.2. Se na definição 1.1 substituirmos a condição **ii)** pela condição

ii') $A \in \mathcal{B}$ implica $A^c = \cup_{j=1}^{k} C_j$ com $C_j \in \mathcal{B}$ e dois a dois disjuntos.

dizemos que $\mathcal{B}$ é uma *sub-álgebra*.

Os dois exemplos abaixo ilustra bem a diferença entre sub-álgebra e álgebra:

Exemplo 1.3.

1. A família dos intervalos da reta é uma sub-álgebra, mas não é uma álgebra, pois o complementar de um intervalo nem sempre é um intervalo.
2. A família das uniões finitas de intervalos da reta é uma álgebra.

Definição 1.4. Uma álgebra diz-se uma *σ-álgebra* de subconjuntos de X se também for fechada para uniões enumeráveis:

- $A_j \in \mathcal{B}$ para $j = 1, 2, \ldots$ implica $\displaystyle\bigcup_{j=1}^{\infty} A_j \in \mathcal{B}$.

Observação 1.5. Uma σ-álgebra $\mathcal{B}$ também é fechada para intersecções enumeráveis: se $A_j \in \mathcal{B}$ para $j = 1, 2, \ldots, n, \ldots$ então $\displaystyle\bigcap_{j=1}^{\infty} A_j = \Big(\bigcup_{j=1}^{\infty} A_j^c\Big)^c \in \mathcal{B}$.

Definição 1.6. Um *espaço mensurável* é uma dupla $(X, \mathcal{B})$ onde X é um conjunto e $\mathcal{B}$ é uma σ-álgebra de subconjuntos de X. Os elementos de $\mathcal{B}$ são chamados *conjuntos mensuráveis*.

Em seguida apresentamos alguns exemplos de σ-álgebras.

Exemplo 1.7. Seja X um conjunto qualquer.

1. Denotemos por 2^X a família de todos os subconjuntos de X. Então $\mathcal{B} = 2^X$ é claramente uma σ-álgebra.
2. $\mathcal{B} = \{\emptyset, X\}$ é também uma σ-álgebra.

Note que se $\mathcal{B}$ é uma álgebra de X então $\{\emptyset, X\} \subset \mathcal{B} \subset 2^X$. Portanto $\{\emptyset, X\}$ é a menor álgebra e 2^X é a maior álgebra de X.

Proposição 1.8. *Considere uma família não-vazia $\{\mathcal{B}_i : i \in \mathcal{I}\}$ qualquer de σ-álgebras ($\mathcal{I}$ é um conjunto qualquer, que serve apenas para indexar os elementos da família). Então a intersecção*

$$\mathcal{B} = \bigcap_{i \in \mathcal{I}} \mathcal{B}_i$$

é também uma σ-álgebra.

Demonstração. Exercício 1.1.

□

Agora, dado um conjunto qualquer $\mathcal{E}$ de subconjuntos de X, podemos aplicar a Proposição 1.8 à família de todas as σ-álgebras que contém $\mathcal{E}$. Note que esta família é não vazia, uma vez que contém a σ-álgebra 2^X, pelo menos. De acordo com a observação anterior, a intersecção de todas estas σ-álgebras é também uma σ-álgebra, e é claro que contém $\mathcal{E}$. Além disso, do modo que é construída, ela está contida em todas as σ-álgebras que contêm $\mathcal{E}$. Portanto é a menor σ-álgebra que contém $\mathcal{E}$.

Definição 1.9. A *σ-álgebra gerada* por uma família $\mathcal{E}$ de subconjuntos de X é a menor σ-álgebra que contém a família $\mathcal{E}$.

No caso em que X vem munido da estrutura de espaço topológico, há uma escolha natural para $\mathcal{E}$, a saber, o conjunto dos subconjuntos abertos. Isto nos conduz à noção de *σ-álgebra de Borel*.

Definição 1.10. Seja (X, τ) um espaço topológico, ou seja, X um conjunto e τ a família dos subconjuntos abertos de X. Então a *σ-álgebra de Borel* de X é a σ-álgebra gerada por τ, isto é, a menor σ-álgebra que contém todos os subconjuntos abertos. Neste caso, os conjuntos mensuráveis recebem o nome de *borelianos*.

Um caso particular especial ocorre quando consideramos o espaço métrico $X = \mathbb{R}^n$ munido com a distância euclidiana. Os borelianos formam uma grande gama de conjuntos, que por definição inclui todos os abertos e fechados, já que estes são os complementares dos abertos. Além disso, uniões ou intersecções de conjuntos abertos ou fechados ainda são borelianos. Podemos nos perguntar sobre os conjuntos *não*-borelianos. Um exemplo está construído no Exercício 1.9.

1.2 Espaços de medida

Agora introduzimos o conceito de medida e analisamos algumas das suas propriedades fundamentais. Em seguida apresentamos alguns resultados sobre construção de medidas. Finalmente, analisamos duas importantes classes de medidas: medidas de Lebesgue em espaços euclideanos e medidas produto em espaço de sequências.

Definição 1.11. Uma *medida* num espaço mensurável $(X, \mathcal{B})$ é uma função $\mu : \mathcal{B} \to [0, +\infty]$ que satisfaz:

1. $\mu(\emptyset) = 0$;
2. $\mu(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu(A_j)$ para quaisquer $A_j \in \mathcal{B}$ disjuntos dois-a-dois.

A tripla $(X, \mathcal{B}, \mu)$ é chamada *espaço de medida*.

A segunda propriedade na definição de medida é chamada de *σ-aditividade*. Dizemos que uma função $\mu : \mathcal{B} \to [0, +\infty]$ é *finitamente aditiva* se:

$$\mu(\bigcup_{j=1}^{N} A_j) = \sum_{j=1}^{N} \mu(A_j)$$

para qualquer família finita $A_1, \ldots, A_N \in \mathcal{B}$ de subconjuntos disjuntos dois-a-dois. Note que, em particular, toda medida é finitamente aditiva.

Exemplo 1.12. Seja X um conjunto e consideremos a σ-álgebra $\mathcal{B} = 2^X$. Dado qualquer $p \in X$, consideremos a função $\delta_p : 2^X \to [0, +\infty]$ definida por:

$$\delta_p(A) = \begin{cases} 1, & \text{se } p \in A \\ 0, & \text{se } p \notin A \end{cases} .$$

A medida δ_p é usualmente designada por *delta de Dirac* no ponto p.

Quando vale $\mu(X) < \infty$ dizemos que μ é uma medida *finita* e se $\mu(X) = 1$ dizemos que μ é uma *probabilidade*. Neste último caso, $(X, \mathcal{B}, \mu)$ é um *espaço de probabilidade*.

Definição 1.13. Diremos que uma medida é σ-finita se existir uma sequência de subconjuntos $A_1, A_2, \ldots$ de X satisfazendo $\mu(A_i) < \infty$ e

$$X = \bigcup_{i=1}^{\infty} A_i.$$

Em seguida apresentamos um resultado muito útil na construção de medidas.

Teorema 1.14 (Extensão)**.** *Seja $\mathcal{B}_0$ uma álgebra de subconjuntos de X e seja $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ uma função σ-aditiva com $\mu(\emptyset) = 0$ e $\mu(X) < \infty$. Então existe uma única medida definida na σ-álgebra $\mathcal{B}$ gerada por $\mathcal{B}_0$ que é uma extensão de μ_0, isto é, para todo elemento $B \in \mathcal{B}_0$ temos que $\mu_0(B) = \mu(B)$.*

Observação 1.15. O Teorema 1.14 se aplica com as mesmas conclusões quando a medida μ_0 em questão é apenas σ-finita. Além disso, se μ_0 é apenas finitamente aditiva ainda assim existe uma medida que estende μ_0 à $\mathcal{B}$. Porém, neste caso não podemos garantir que tal extensão é única.

Em geral, ao tentarmos mostrar que uma função definida na σ-álgebra é uma medida o mais difícil é verificar a σ-aditividade. O critério mais usado para esse efeito é expresso no seguinte resultado. A sua demonstração é proposta como Exercício 1.12.

Teorema 1.16 (Continuidade no vazio)**.** *Seja $\mathcal{B}$ uma σ-álgebra e seja $\mu : \mathcal{B} \to [0, +\infty]$ uma função finitamente aditiva com $\mu(M) < \infty$. Suponha que*

$$\lim_{n\to\infty} \mu\Big(\bigcap_{j=1}^{n} A_j\Big) = 0 \tag{1.1}$$

para toda sequência $A_1 \supset \cdots \supset A_j \supset \cdots$ de conjuntos mensuráveis tal que $\bigcap_{j=1}^{\infty} A_j = \emptyset$. Então μ é σ-aditiva.

O resultado seguinte nos diz que todo elemento B da σ-álgebra gerada por uma álgebra é aproximado por algum elemento B_0 da álgebra, no sentido em que a medida da diferença simétrica $B \Delta B_0 = B \setminus B_0 \cup B_0 \setminus B$ é pequena.

Teorema 1.17 (Aproximação)**.** *Seja $(X, \mathcal{B}, \mu)$ um espaço de probabilidade e seja $\mathcal{B}_0$ uma álgebra que gera a σ-álgebra $\mathcal{B}$. Então para todo $\varepsilon > 0$ e todo $B \in \mathcal{B}$ existe $B_0 \in \mathcal{B}_0$ tal que $\mu(B \Delta B_0) < \varepsilon$.*

1.2.1 Medida de Lebesgue

A medida de Lebesgue corresponde ao que entendemos por volume de subconjuntos de $\mathbb{R}^d$. Para construí-la, recorremos ao Teorema de Extensão 1.14. Consideremos $M = [0,1]$ e seja $\mathcal{B}_0$ a família de todos os subconjuntos da forma $B = I_1 \cup \cdots \cup I_N$ onde $I_1, \ldots, I_N$ são intervalos disjuntos dois-a-dois. É fácil ver que $\mathcal{B}_0$ é uma álgebra de subconjuntos de M. Além disso, temos uma função $\mu_0 : \mathcal{B}_0 \to [0,1]$ definida nesta álgebra por

$$\mu_0\big(I_1 \cup \cdots \cup I_N\big) = |I_1| + \cdots + |I_N|,$$

onde $|I_j|$ representa o comprimento de I_j. Note que $\mu_0(M) = 1$. No Exercício 1.5 pedimos para que você mostre que μ_0 é uma função σ-aditiva.

Note que a σ-álgebra $\mathcal{B}$ gerada por $\mathcal{B}_0$ coincide com a σ-álgebra de Borel de M, já que todo aberto pode ser escrito como união enumerável de intervalos abertos disjuntos dois-a-dois. Pelo Teorema 1.14, existe uma única probabilidade μ definida em $\mathcal{B}$ que é uma extensão de μ_0. Chamamos μ de *medida de Lebesgue* em $[0,1]$.

Mais geralmente, definimos medida de Lebesgue μ no cubo $M = [0,1]^d$ de qualquer dimensão $d \geq 1$ da seguinte maneira: chamamos retângulo em M qualquer subconjunto da forma $R = I_1 \times \cdots \times I_d$ onde os I_j são intervalos, e definimos

$$\mu_0(R) = |I_1| \times \cdots \times |I_d|.$$

Em seguida, consideramos a álgebra $\mathcal{B}_0$ dos subconjuntos de $[0,1]^d$ da forma $B = R_1 \cup \cdots \cup R_N$, onde $R_1, \ldots, R_N$ são retângulos disjuntos dois-a-dois, e definimos

$$\mu_0(B) = \mu_0(R_1) + \cdots \mu_0(R_N)$$

para todo B nessa álgebra. A medida de Lebesgue em $M = [0,1]^d$ é a extensão de μ_0 à σ-álgebra gerada por $\mathcal{B}_0$, que coincide com a σ-álgebra de Borel de M. Para definir a medida de Lebesgue num espaço euclidiano $\mathbb{R}^d$ decompomos o espaço em cubos de lado unitário

$$\mathbb{R}^d = \bigcup_{m_1 \in \mathbb{Z}} \cdots \bigcup_{m_d \in \mathbb{Z}} [m_1, m_1 + 1) \times \cdots \times [m_d, m_d + 1).$$

Para cada cubo $C = [m_1, m_1 + 1) \times \cdots \times [m_d, m_d + 1)$ podemos construir uma medida $\mu_{m_1, m_2, \ldots, m_d}$ colocando

$$\mu_{m_1, m_2, \ldots, m_d}(A) = \mu_0\big(T_{m_1, \ldots, m_d}(A)\big),$$

onde $T_{m_1, \ldots, m_d}(x) = x - (m_1, \ldots, m_d)$ é a translação que leva o vetor $(m_1, m_2, \ldots, m_d)$ em 0. Finalmente, definimos para cada subconjunto mensurável E

$$\mu(E) = \sum_{m_1 \in \mathbb{Z}} \cdots \sum_{m_d \in \mathbb{Z}} \mu_{m_1, \ldots, m_d}\big(E \cap [m_1, m_1 + 1) \times \cdots \times [m_d, m_d + 1)\big).$$

Note que μ não é uma medida finita, mas é uma medida σ-finita.

Exemplo 1.18 (Medida de Volume em S^1)**.** Considere a aplicação sobrejetora $\gamma : [0,1] \to S^1$ definida por:

$$\gamma(t) = e^{2\pi i t}.$$

A *medida de Lebesgue* em S^1 é a medida μ definida por $\mu(A) = 2\pi m(\gamma^{-1}(A))$. Assim segue-se, or exemplo, que a medida de Lebesgue de um arco de círculo corresponde ao seu comprimento.

Observe que com esta definição, a medida de A é igual a medida de $R_\alpha(A)$, onde $R_\alpha : S^1 \to S^1$ denota a rotação de ângulo α. Na verdade, módulo multiplicação por um número positivo, μ é a única medida que satisfaz essa condição para todo α (veja o Exercício 1.15).

Exemplo 1.19. Seja $\phi : [0,1] \to \mathbb{R}$ uma função contínua e positiva. Defina a medida μ_ϕ num intervalo $[a,b] \subset [0,1]$ por:

$$\mu_\phi([a,b]) = \int_a^b \phi(x)\,dx.$$

Observe que se consideramos a álgebra $\mathcal{B}_0$ das uniões finitas de intervalos de $[0,1]$ e um elemento $A = I_1 \cup \cdots \cup I_k$ em $\mathcal{B}_0$, definimos $\mu_\phi(A) = \sum_{j=1}^k \mu_\phi(I_j)$.

As propriedades básicas da integral nos dizem que μ_ϕ é finitamente aditiva. Deixamos para o leitor a tarefa de mostrar que a medida μ_ϕ é σ-aditiva na álgebra formada pelas uniões finitas de intervalos. Além disso, $\mu_\phi(\emptyset) = 0$ e $\mu_\phi([0,1]) < \infty$ já que ϕ é contínua, portanto limitada. Com o auxílio do Teorema 1.14 podemos estender μ_ϕ para toda σ-álgebra dos borelianos de $[0,1]$.

Observe que a medida μ_ϕ tem a seguinte propriedade especial: se um conjunto $A \subset [0,1]$ tem medida de Lebesgue zero então $\mu_\phi(A) = 0$. Essa propriedade chama-se continuidade absoluta (com respeito à Lebesgue).

Exemplo 1.20. Vamos agora exibir uma medida que, apesar de ser positiva em qualquer aberto, não é absolutamente contínua com respeito a medida de Lebesgue. Para isso, considere uma enumeração $\{r_1, r_2, \dots\}$ do conjunto $\mathbb{Q}$ dos racionais. Defina μ por:

$$\mu(A) = \sum_{r_i \in A} \frac{1}{2^i}.$$

Observe que a medida de qualquer aberto da reta é positiva, pois necessariamente A contém algum r_i. Apesar disso, a medida de $\mathbb{Q}$ é

$$\mu(\mathbb{Q}) = \sum_{r_i \in \mathbb{Q}} \frac{1}{2^i} = 1.$$

Em particular, μ não é absolutamente contínua com respeito a medida de Lebesgue.

O exemplo anterior nos motiva a definir o *suporte* de uma medida:

Definição 1.21. Seja $(X, \mathcal{B}, \mu)$ um espaço de medida e X um espaço topológico separável. O suporte $supp(\mu)$ da medida μ é o fecho do conjunto formado pelos pontos $x \in X$ tais que para qualquer vizinhança aberta V_x contendo x, temos que $\mu(V_x) > 0$.

Destacamos que, por definição, o suporte de uma medida é um conjunto fechado. No Exemplo 1.20 acima, o suporte da medida μ é a reta inteira, apesar de $\mu(\mathbb{Q}) = 1$. A próxima proposição nos garante que o suporte de uma medida de probabilidade é sempre um conjunto não-vazio.

Proposição 1.22. *Seja X um espaço topológico separável e μ uma medida de probabilidade em X. Então, o suporte $supp(\mu)$ é não-vazio.*

Demonstração. De fato, se $supp(\mu)$ é vazio, então para cada ponto x de X podemos encontrar um aberto V_x com $x \in V_x$ de modo que $\mu(V_x) = 0$. Como X é separável, existe uma base enumerável de abertos $\tau = \{A_1, A_2, \dots\}$ de modo que para cada V_x podemos escolher $A_{i(x)} \in \tau$ satisfazendo $x \in A_{i(x)} \subset V_x$. Consequentemente:

$$X = \bigcup_{x \in X} V_x = \bigcup_{x \in X} A_{i(x)}.$$

Isso nos permite concluir que

$$1 = \mu(X) = \mu(\bigcup_{x \in X} A_{i(x)}) \leq \sum_{i=1}^{\infty} \mu(A_i) = 0,$$

o que é um absurdo. Logo $supp(\mu)$ não pode ser vazio.

□

1.2.2 Medida produto no espaço das sequências

Consideremos os espaços de probabilidade $(X_i, \mathcal{B}_i, \mu_i)$, com $i \in \mathbb{Z}$. Vamos construir uma probabilidade μ no conjunto

$$\Sigma = \prod_{i=-\infty}^{\infty} X_i$$

das sequências bilaterais $(x_i)_{i=-\infty}^{\infty}$ com $x_i \in X_i$ para cada i. Mais precisamente, a medida μ será definida na *σ-álgebra produto* $\mathcal{B}$ das σ-álgebras $\mathcal{B}_i$, que é caracterizada do seguinte modo: dados inteiros $m \leq n$ e conjuntos $A_j \in \mathcal{B}_j$ para $m \leq j \leq n$, consideremos

$$[m; A_m, \dots, A_n] = \{(x_i)_{i \in \mathbb{Z}} : x_j \in A_j \text{ para } m \leq j \leq n\}.$$

Estes subconjuntos de Σ são chamados *cilindros*. A família $\mathcal{B}_0$ das uniões finitas de cilindros disjuntos dois-a-dois é uma álgebra. Por definição, a σ-álgebra produto $\mathcal{B}$ é a σ-álgebra gerada por $\mathcal{B}_0$.

Para construir a medida μ procedemos do seguinte modo: primeiramente, consideramos a aplicação μ_0 definida na família dos cilindros por

$$\mu_0([m; A_m, \dots, A_n]) = \prod_{j=m}^{n} \mu_j(A_j).$$

Em seguida estendemos μ_0 à álgebra $\mathcal{B}_0$, estipulando que a imagem de qualquer união finita de cilindros disjuntos dois-a-dois é igual à soma das imagens dos cilindros. Esta extensão está bem definida e é σ-aditiva. Então, recorrendo ao Teorema 1.14, obtemos uma medida de probabilidade μ em $(\Sigma, \mathcal{B})$ que estende μ_0.

Definição 1.23. O espaço de probabilidade $(\Sigma, \mathcal{B}, \mu)$ construído acima é designado *produto direto* dos espaços $(X_i, \mathcal{B}_i, \mu_i)$.

Observação 1.24. Uma observação importante é que podemos proceder de modo inteiramente análogo para obter uma probabilidade μ no conjunto

$$\Sigma^+ = \prod_{i=0}^{\infty} X_i$$

das sequências *unilaterais* $(x_i)_{i=0}^{\infty}$ com $x_i \in X_i$ para cada i.

Existe um caso particular importante, que corresponde à situação onde os espaços $(X_i, \mathcal{B}_i, \mu_i)$ são todos iguais a um dado $(X, \mathcal{C}, \nu)$, em que $X = \{1, \dots, d\}$ é um conjunto finito e $\mathcal{C} = 2^X$ é a σ-álgebra de todos os subconjuntos de X. Neste caso basta considerar apenas cilindros elementares, isto tais que cada A_j consiste de um único ponto de X. De fato, todo cilindro é uma união finita disjunta de tais cilindros elementares. Obtemos então subconjuntos de Σ da forma

$$[m; a_m, \dots, a_n] = \{(x_i)_{i=-\infty}^{\infty} \in \Sigma : x_m = a_m, \dots, x_n = a_n\}$$

onde $a_j \in \{1, \dots, d\}$. A medida μ é designada *medida de Bernoulli* definida por ν e é caracterizada por $\mu([m; a_m, \dots, a_n]) = \nu(\{a_m\}) \cdots \nu(\{a_n\})$.

margemjuntar tudo de shifts de bernoulli aqui, incluindo metrizacao

1.3 Funções Mensuráveis

Nesta seção introduzimos a noção de *função mensurável*. Assim como as funções contínuas desempenham um papel crucial em Topologia, já que elas preservam a noção de conjunto aberto, as funções mensuráveis são fundamentais na Teoria da Medida, pois preservam a noção de conjunto mensurável. Denotaremos por $\overline{\mathbb{R}}$ o conjunto $\mathbb{R} \cup \{-\infty, +\infty\}$.

Definição 1.25. Uma função $f : X \to \overline{\mathbb{R}}$ diz-se *mensurável* se $f^{-1}((c, +\infty]) \in \mathcal{B}$ para todo $c \in \mathbb{R}$.

Pode-se mostrar que uma função f nas condições anteriores é mensurável se a pré-imagem de um boreliano da reta pertence a σ-algebra $\mathcal{B}$. Além disso,

Proposição 1.26. *Seja $f : X \to Y$ uma função qualquer e $\mathcal{C}$ uma σ-álgebra de Y. Se definimos $\mathcal{B} = \{f^{-1}(C) \subset X; C \in \mathcal{C}\}$, então $\mathcal{B}$ é uma σ-álgebra de X.*

Demonstração. Faça o exercício 1.8 $\square$

A proposição anterior nos garante que dada uma função $f : X \to \overline{\mathbb{R}}$ qualquer, sempre podemos munir X com uma σ-álgebra adequada de modo que f seja mensurável. Isso nos mostra a generalidade da definição de função mensurável que demos anteriormente. Entretanto, estaremos interessados principalmente quando o espaço mensurável $(X, \mathcal{B})$ em questão é um espaço topológico munido da σ-álgebra de Borel. Nesta situação, podemos garantir que

Exemplo 1.27. Se X é um espaço topológico e $\mathcal{B}$ a σ-álgebra de Borel associada então toda função contínua $f : X \to \mathbb{R}$ é mensurável.

Exemplo 1.28. Considere a função $\mathcal{X}_{\mathbb{Q}} : [0,1] \to \mathbb{R}$ definida por:

$$\mathcal{X}_{\mathbb{Q}}(x) = \begin{cases} 1, & \text{se } x \in \mathbb{Q}; \\ 0, & \text{caso contrário.} \end{cases}$$

Observe que $\mathcal{X}_{\mathbb{Q}}$ é mensurável, já que para todo conjunto $A \subset \mathbb{R}$

$$\mathcal{X}_{\mathbb{Q}}^{-1}(A) \in \{\emptyset, \mathbb{Q}, \mathbb{R} \setminus \mathbb{Q}, \mathbb{R}\}.$$

Essa função é chamada de *função característica* do conjunto $\mathbb{Q}$. Em geral, dado um conjunto $B \subset X$ podemos definir uma função $\mathcal{X}_B : X \to \mathbb{R}$ chamada *função característica* do conjunto B por:

$$\chi_B(x) = \begin{cases} 1, & \text{se } x \in B; \\ 0, & \text{caso contrário.} \end{cases}$$

Observe que $\mathcal{X}_B$ é mensurável se, e somente se, B o for: de fato, $\mathcal{X}_B^{-1}(A) \in \{\emptyset, B, X \setminus B, X\}$ para qualquer $A \subset \mathbb{R}$.

Entre as propriedades mais simples das funções mensuráveis temos:

Proposição 1.29. *Sejam $f, g : X \to \overline{\mathbb{R}}$ funções mensuráveis e $a, b \in \mathbb{R}$. Então também são mensuráveis as seguintes funções:*

$$(af + bg)(x) = af(x) + bg(x) \quad e \quad (f \cdot g)(x) = f(x) \cdot g(x).$$

Além disso, se $f_n : X \to \overline{\mathbb{R}}$ é uma sequência de funções mensuráveis, são mensuráveis também as seguintes funções:

- $s(x) = \sup\{f_n(x); n \geq 1\}$,
- $i(x) = \inf\{f_n(x); n \geq 1\}$,
- $f^*(x) = \limsup f_n(x)$,
- $f_*(x) = \liminf f_n(x)$. *Em particular, se $f(x) = \lim f(x)$ existe, f é mensurável.*

Uma classe particular de funções mensuráveis são aquelas formadas por combinações de funções características.

Definição 1.30. Dizemos que uma função $s : M \to \mathbb{R}$ é *simples* se existem constantes $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ e conjuntos $A_1, \ldots, A_k \in \mathcal{B}$ disjuntos dois-a-dois tais que

$$s = \sum_{j=1}^{k} \alpha_k \mathcal{X}_{A_k}, \tag{1.2}$$

onde $\mathcal{X}_A$ é a função característica do conjunto A.

Um teorema importante que será utilizado na seção seguinte afirma que toda função mensurável é de fato o limite de uma sequência de funções simples. Mais precisamente:

Teorema 1.31. *Seja $f : X \to \overline{\mathbb{R}}$ uma função mensurável. Então existe uma sequência $s_1,\ s_2,\ \ldots$ de funções simples mensuráveis tal que*

$$\lim_{k\to\infty} s_k(x) = f(x) \text{ para todo } x \in X.$$

Se $f \geq 0$ então a sequência pode ser escolhida de modo que $0 \leq s_1 \leq s_2 \leq \cdots$.

1.4 Integração em espaços de medida

Nesta seção definimos a noção de integral de Lebesgue de uma função em relação a uma medida. Uma das motivações para essa nova definição de integral é considerar uma classe mais ampla de funções que possam ser integradas. Em particular, gostaríamos de poder integrar as funções características de conjuntos mensuráveis, o que não é possível em geral, se utilizamos a integral de Riemann. Assim, a Integral de Lebesgue vem como uma generalização natural da definição de integral de Riemann e consiste numa das noções mais importantes da Teoria da Medida. Ao longo desta seção $(X, \mathcal{B}, \mu)$ será sempre um espaço de medida.

Vamos definir a noção de integral de Lebesgue em etapas.A primeira etapa é definir integral de uma função simples.

Definição 1.32. Seja $s = \sum_{j=1}^{k} \alpha_k \mathcal{X}_{A_k}$ uma função simples. Então a *integral* de s em relação a medida μ é dada por:

$$\int s\ d\mu = \sum_{j=1}^{k} \alpha_k \mu(A_k).$$

É fácil verificar que esta definição é coerente: se duas combinações lineares de funções características definem uma mesma função simples, os valores das integrais obtidos a partir das duas combinações coincidem. O próximo passo é definir integral de uma função mensurável qualquer. Para isso, trataremos primeiro do caso da função ser não-negativa. A idéia é definir a integral de uma função não-negativa como o limite das integrais de funções simples que a aproximam, utilizando o Teorema 1.31:

Definição 1.33. Seja $f : X \to [0, \infty]$ uma função mensurável não-negativa. Então
$$\int f d\mu = \lim_{n\to\infty} \int s_n d\mu,$$
onde $s_1 \leq s_2 \leq \dots$ é qualquer sequência de funções simples crescentes para f, isto é, $\lim_{n\to\infty} s_n(x) = f(x)$ para todo $x \in X$.

Podemos verificar que o valor da integral não depende da escolha da sequência de funções simples crescendo para f. Como consequência, a definição de integral de uma função não-negativa é coerente.

Para estender a definição de integral a qualquer função mensurável, observemos que dada uma função $f : X \to \overline{\mathbb{R}}$ sempre podemos escrever $f = f^+ - f^-$ onde $f^+(x) = \max\{f(x), 0\}$ e $f^-(x) = \max\{-f(x), 0\}$ são não-negativas. Pelo Item 2 da Proposição 1.29 segue-se que f^+ e f^- são mensuráveis se, e só se, f é mensurável.

Definição 1.34. Seja $f : X \to \overline{\mathbb{R}}$ uma função mensurável. Então
$$\int f\, d\mu = \int f^+\, d\mu - \int f^-\, d\mu,$$
desde que alguma das integrais do lado direito seja finita.

Definição 1.35. Dizemos que uma propriedade é válida *em μ-quase todo ponto* se é válida em todo o X exceto, possivelmente, num conjunto de medida nula. Quando isto ocorrer, escreveremos abreviadamente que a propriedade é válida em *μ-q.t.p.*

Por exemplo, dizemos que duas funções f, g são iguais em μ-quase todo ponto se existe um conjunto mensurável N com $\mu(N) = 0$ tal que $f(x) = g(x)$ para todo $x \in X \setminus N$.

Definição 1.36. Dizemos que uma função é *integrável* se for mensurável e sua integral for um número real. Denotamos o conjunto das funções integráveis por $L^1(X, \mathcal{B}, \mu)$ ou, mais simplesmente, por $L^1(\mu)$. Aqui estamos identificando duas funções que coincidem em q.t.p.

Restrito ao conjunto $L^1(\mu)$ das funções integráveis, o processo de integração é linear:

Proposição 1.37. *O funcional $I : L^1(\mu) \to \mathbb{R}$ definido por*
$$I(f) = \int f\, d\mu$$
é um funcional linear positivo. Isto é, se $a, b \in \mathbb{R}$ e $f, g : M \to \mathbb{R}$ integráveis. Então:

(a) *A função $af + bg$ é integrável. Além disso,*
$$\int af + bg\, d\mu = a \int f\, d\mu + b \int g\, d\mu.$$

(b) *Se $f(x) \leq g(x)$ para um conjunto de pontos x com medida total, então*

$$\int f\, d\mu \leq \int g\, d\mu.$$

Em particular,

$$\left| \int f\, d\mu \right| \leq \int |f|\, d\mu.$$

Dada uma função mensurável $f : X \to \mathbb{R}$ e um conjunto mensurável E definimos a *integral de f sobre E* por

$$\int_E f d\mu = \int f \mathcal{X}_E\, d\mu,$$

onde $\mathcal{X}_E$ é a função característica do conjunto E.

Exemplo 1.38. Sejam $x_1, \ldots, x_m \in X$ e $p_1, \ldots, p_m > 0$ com $p_1 + \cdots + p_m = 1$. Consideremos a medida de probabilidade $\mu : 2^X \to [0, 1]$ dada por:

$$\mu(A) = \sum_{x_i \in A} p_i.$$

Notemos que $\mu = \sum_{i=1}^m p_i \delta_{x_i}$, onde δ_{x_i} é a medida delta de Dirac em x_i. Neste caso temos que se f é uma função integrável então

$$\int f\, d\mu = \sum_{i=1}^m f(x_i) p_i.$$

1.5 Exercícios

1.1. Seja X um conjunto e $(\mathcal{B}_i)_{i \in \mathcal{I}}$ uma família de σ-álgebras de subconjuntos de X. Mostre que

$$\mathcal{B} = \bigcap_{i \in \mathcal{I}} \mathcal{B}_i$$

é uma σ-álgebra.

1.2. Seja X um conjunto e considere a família de conjuntos

$$\mathcal{B}_0 = \{A \subset X : A \text{ é finito ou } A^c \text{ é finito}\}.$$

Mostre que $\mathcal{B}_0$ é uma álgebra. Além disso, $\mathcal{B}_0$ é uma σ-álgebra se e somente se o conjunto X é finito.

1.3. Seja X um conjunto e considere a seguinte família de conjuntos

$$\mathcal{B}_1 = \{A \subset X : A \text{ é finito ou enumerável ou } A^c \text{ é finito ou enumerável}\}.$$

Mostre que $\mathcal{B}_1$ é uma σ-álgebra. De fato, $\mathcal{B}_1$ é a σ-álgebra gerada pela álgebra $\mathcal{B}_0$ do Exercício 1.2.

1.4. Sejam μ_n, μ medidas em M com $\lim \mu_n = \mu$ e $A \subset M$ um conjunto mensurável. Então

- Se A é aberto mod 0, $\liminf \mu_n(A) = \mu(A)$.
- Se A é fechado mod 0, $\limsup \mu_n(A) = \mu(A)$.
- Em particular, se $\mu(\partial A) = 0$, então $\lim \mu_n(A) = \mu(A)$.

1.5. Seja $\mathcal{B}_0$ a coleção dos conjuntos que se escrevem como união finita de intervalos disjuntos.

1. Prove que $\mathcal{B}_0$ é uma álgebra de subconjuntos de M.

   Defina uma função $\mu_0 : \mathcal{B}_0 \to [0,1]$ definida nesta álgebra por

   $$\mu_0\big(I_1 \cup \cdots \cup I_N\big) = |I_1| + \cdots + |I_N|\,,$$

   onde $|I_j|$ representa o comprimento de I_j.

2. Mostre que μ_0 é uma função σ-aditiva.

1.6. O *limite superior* (limsup) de uma sequência de conjuntos $E_n \in X$ é definido como o conjunto $\limsup E_n$ formado pelos pontos $x \in X$ tais que $x \in E_n$ para infinitos valores de n. Analogamente, o *limite inferior* (liminf) é o conjunto $\liminf E_n$ dos pontos $x \in X$ tais que existe n_0 tal que $x \in E_n$ para todo $n \geq n_0$. Nestas condições, mostre que:

(a) $\liminf E_n = \bigcup_{n\geq 1} \bigcap_{m \geq n} E_n$;

(b) $\limsup E_n = \bigcap_{n\geq 1} \bigcup_{m \geq n} E_n$.

Conclua daí que $\liminf E_n \subset \limsup E_n$.

1.7. Seja $\mathcal{E}$ uma família de subconjuntos de um conjunto X. Mostre que existe a menor *álgebra* $\mathcal{B}_0$ que contém $\mathcal{E}$. Que relação existe entre $\mathcal{B}_0$ e a σ-álgebra $\mathcal{B}$ gerada por $\mathcal{E}$?

1.8. Seja $f : X \to Y$ uma função qualquer e $\mathcal{C}$ uma σ-álgebra de Y. Se definimos $\mathcal{B} = \{f^{-1}(C) \subset X; C \in \mathcal{C}\}$, então $\mathcal{B}$ é uma σ-álgebra.

1.9. O objetivo desse exercício é exibir um subconjunto da Reta que não é boreliano. Para isso, considere um número irracional α qualquer e defina uma relação na reta tomando $x, y \in \mathbb{R}$ e definindo $x \sim y$ se :

$$x - y = n + m\alpha, \text{ para alguns } n, m \in \mathbb{Z}.$$

(a) Mostre que $\sim$ é uma relação de equivalência (isto é, é reflexiva, simétrica e transitiva).

A classe de equivalência do ponto x_0 é o conjunto de todos os pontos $y \in \mathbb{R}$ de modo que $x_0 \sim y$.

(b) Verifique que duas classes de equivalência ou coincidem ou são disjuntas. Mostre também que o conjunto das classes de equivalência particionam a reta.

Considere E_0 um conjunto que possua exatamente um elemento de cada classe de equivalência (pode-se provar que tal conjunto existe utilizando o axioma da escolha).

(c) Mostre que que E_0 *não* é boreliano. (Dica: assuma o contrário e chegue a um absurdo utilizando a σ-aditividade da medida de Lebesgue).

1.10. Seja $(X, \mathcal{B}, \mu)$ um espaço de medida. Mostre que se A_1,A_2, ... estão em $\mathcal{B}$ então

$$\mu(\bigcup_{j=1}^{\infty} A_j) \leq \sum_{j=1}^{\infty} \mu(A_j).$$

1.11. Seja $\mathcal{B} = 2^X$ e considere $\mu : 2^X \to [0, +\infty]$ definido por:

$$\mu(A) = \begin{cases} \#A & , \text{ se } A \text{ é finito} \\ \infty & \text{ se } A \text{ é infinito} \end{cases}.$$

Mostre que μ é uma medida. Esta medida é designada *medida de contagem*.

1.12. Demonstre o Teorema 1.16.

Dica: Dados quaisquer conjuntos disjuntos dois a dois $B_1, \ldots, B_n, \ldots$ em $\mathcal{B}_0$ tais $B = \bigcup_{j=1}^{\infty} B_j$ também está em $\mathcal{B}_0$, defina $C_j = B_1 \cup \cdots \cup B_j$ para cada $j \geq 1$. Verifique que os conjuntos $A_j = B \setminus C_j$ satisfazem a hipótese (1.1) no Teorema 1.16.

1.13. Seja $(X, \mathcal{B})$ um espaço mensurável.

(a) Mostre que se $\mu : \mathcal{B} \to [0, +\infty]$ é uma medida então

$$\mu(\bigcup_{j=1}^{\infty} A_j) = \lim_{j \to \infty} \mu(A_j).$$

para qualquer sequência crescente $A_1 \subset A_2 \subset \cdots A_n \subset \cdots$ de elementos de $\mathcal{B}$.

(b) Reciprocamente, mostre que se $\mu_0 : \mathcal{B} \to [0, +\infty]$ é uma função finitamente aditiva que satisfaz a condição do item anterior então μ_0 é σ-aditiva.

1.14. Seja $(X, \mathcal{B})$ um espaço mensurável, onde o conjunto X é não-enumerável e a σ-álgebra $\mathcal{B}$ é definida como no Exercício 1.3. Mostre que $\mu : \mathcal{B} \to [0, +\infty]$ definida por:

$$\mu(A) = \begin{cases} 0 & \text{se } A \text{ é finito ou enumerável} \\ 1 & \text{se } A^c \text{ é finito ou enumerável} \end{cases}$$

é uma medida de probabilidade.

1.15. Prove que se μ é a medida de Lebesgue em S^1 definida no Exemplo 1.18, a medida de A é igual a medida de $R_\alpha(A)$ para qualquer boreliano A de S^1, onde $R_\alpha : S^1 \to S^1$ denota a rotação de ângulo α. Prove também que , módulo multiplicação por um número positivo, μ é a única medida que satisfaz essa condição para todo α.

1.16. Se X é um espaço topológico e $\mathcal{B}$ a σ-álgebra de Borel associada. Prove que toda função $f : X \to \mathbb{R}$ contínua é mensurável. Dê exemplo de uma função mensurável que não é contínua em nenhum ponto.

Dica: Prove que o conjunto $\mathcal{C} = \{C \subset \mathbb{R}; f^{-1}(C) \in \mathcal{B}\}$ é uma σ-algebra de X que contém os abertos.

1.17. Seja $f : X \to \mathbb{R}$. Mostre que se $f^{-1}(-\infty, c) \in \mathcal{B}$ para todo $c \in \mathbb{R}$ então f é mensurável.

Dica: Mostre que a família $\mathcal{C} = \{A \subset \mathbb{R} : f^{-1}(A) \in \mathcal{B}\}$ é uma σ-álgebra e contém todos os subconjuntos abertos.

1.18. Prove o Teorema 1.31.

Dica: Trate primeiro o caso onde f é não-negativa.

1.19. Seja $T : X \to X$ uma função mensurável e ν uma medida. Defina $T_*\nu(A) = \nu(T^{-1}(A))$. Mostre que $T_*\nu$ é uma medida.

1.20. Sejam f e g funções mensuráveis. Mostre que f é integrável se e somente se $|f|$ é integrável e, nesse caso,

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu.$$

Além disso, se f é integrável e $|f| \geq |g|$ então g é integrável.

1.21. Seja E um conjunto mensurável com $\mu(E) = 0$. Mostre que $\int_E f \, d\mu = 0$ para qualquer função mensurável f.

1.22. Mostre que a é um ponto de densidade do conjunto A se e só se

$$\lim_{\varepsilon \to 0} \inf \left\{ \frac{\mu(B \cap A)}{\mu(B)} : B \text{ bola contida em } B(a, \varepsilon) \text{ e contendo } a \right\} = 1$$

Capítulo 2

Medidas Invariantes e Recorrência

2.1 Preliminares

Sejam $f : M \to M$ uma transformação mensurável e μ uma medida sobre M. Dizemos que μ é invariante pela transformação f (ou que f preserva a medida μ) se

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

para todo conjunto mensurável $E \subset M$.

É claro que está subentendido que M está dotada de uma σ-álgebra $\mathcal{A}$, a princípio, arbitrária. Mas, quando M for um espaço topológico, a σ-álgebra considerada será a de Borel, isto é, a σ-álgebra gerada pelos abertos de M.

Uma maneira equivalente de dizer que uma medida é invariante por uma dada transformação é dada pela seguinte

Proposição 2.1. *Seja $f : M \to M$ uma transformação e μ uma medida. Então f preserva μ se, e somente se, para toda função μ integrável $\phi : M \to \mathbb{R}$ vale:*

$$\int \phi \, d\mu = \int \phi \circ f \, d\mu. \tag{2.1}$$

Demonstração. A prova desta proposição utiliza um argumento bem clássico em Teoria da Medida, que consiste em mostrar que a Equação 2.1 é válida para toda função integrável dividindo a demonstração em partes:

- primeiro mostramos que a Equação 2.1 vale quando ϕ é uma função característica;
- Em seguida mostramos que a Equação 2.1 vale se ϕ é uma função simples;
- Utilizando um argumento de aproximação e a definição de integral, mostramos que a Equação 2.1 vale para toda função integrável positiva. Finalmente, concluímos que a Equação 2.1 é valida para toda função integrável.

Assim, assuma que f preserva a medida μ. É claro que toda função característica é integrável, uma vez que μ é probabilidade. Se ϕ é função característica de algum conjunto, digamos $\phi = \chi_A$, é imediato verificar que $\mu(f^{-1}(A)) = \int \phi \circ f \, d\mu$, já que $\chi_{f^{-1}(A)} = \phi \circ f$. Assim, fica provado que $\int \phi \, d\mu = \int \phi \circ \ f \, d\mu$, quando ϕ é uma função característica. Observe que segue diretamente da linearidade da integral que se ϕ é uma função simples, então a igualdade na Equação 2.1 ainda vale.

Finalmente, se ϕ é uma função integrável positiva qualquer, pela definição de integral

$$\int \phi \, d\mu = \lim_{n\to\infty} \int \phi_n \, d\mu,$$

onde ϕ_n é uma sequência de funções simples positivas crescendo para ϕ. Por outro lado, $\phi_n \circ f$ é uma sequência de funções simples crescendo para $\phi \circ f$. Logo,

$$\int \phi \circ f \, d\mu = \lim_{n\to\infty} \int \phi_n \circ f \, d\mu.$$

Como $\int \phi_n \, d\mu = \int \phi_n \circ f \, d\mu$, tomando o limite em ambos os lados, vem que

$$\int \phi \, d\mu = \int \phi \circ f \, d\mu.$$

A conclusão para uma função integrável f qualquer segue facilmente, observando-se que $f = f^+ - f^-$, onde f^+ e f^- são funções positivas.

A recíproca é imediata, desde que dado um boreliano A, tomando $\phi = \chi_A$, então

$$\mu(A) = \mu(f^{-1}(A)) \Leftrightarrow \int \phi \, d\mu = \int \phi \circ f \, d\mu.$$

□

Na próxima seção vamos mostrar o Teorema de Recorrência de Poincaré e nas demais seções daremos vários exemplos de medidas invariantes conectando-as com o Teorema de recorrência de Poincaré.

2.2 Teorema de Recorrência de Poincaré

No final do século XIX, Poincaré estava estudando o movimento dos corpos celestes, tais como planetas e cometas. Esse movimento é descrito por certas equações diferenciais que resultam da Lei de Gravitação de Newton. Poincaré mostrou que quando o número n de corpos é maior ou igual a 3 essas equações não podem ser resolvidas analiticamente, quer dizer, não é possível escrever fórmulas explícitas para as suas soluções. No lugar disso, ele propôs que se buscasse uma descrição qualitativa da evolução do sistema que não precisasse de tais fórmulas explícitas.

Nessa direção, Poincaré observou que para quase todo estado inicial dos n corpos, ou seja quase todo valor das posições e velocidades iniciais, a solução da equação diferencial regressa arbitrariamente perto desse estado inicial. Mais do que isso, ele apontou que essa propriedade não é exclusiva dos sistemas da Mecânica Celeste: ela depende apenas do fato de que o fluxo solução das respectivas equações diferenciais admite uma medida invariante. Este fato havia sido provado anteriormente por Liouville, para sistemas da Mecânica muito mais gerais, em que há conservação da energia. Nós iremos discuti-lo na Seção 2.3.5.

Numa linguagem moderna, esse teorema de Poincaré afirma que, relativamente a qualquer medida invariante *finita* do sistema dinâmico, quase todo ponto x é recorrente: existem tempos $t_j \to \infty$ tais que $f^{t_j}(x) \to x$. De fato, daremos duas versões deste enunciado, a primeira de natureza mensurável (Seção 2.2.1) e a segunda de natureza topológica (Seção 2.2.2).

Posteriormente, na Seção **??**, comentaremos sobre algumas generalizações destes resultados. Nas demais seções do presente capítulo apresentamos alguns exemplos de sistemas com medidas invariantes que nos permitarão ilustrar algumas consequências simples do Teorema de Recorrência, na Teoria dos Números (Seções 2.3.1 e 2.3.2) e em Sistemas Dinâmicos (Seção **??**).

2.2.1 Versão mensurável

Teorema 2.2. *Seja $f : M \to M$ uma transformação mensurável e μ uma medida invariante finita. Seja $E \subset M$ qualquer conjunto mensurável com $\mu(E) > 0$. Então, μ-quase todo ponto $x \in E$ tem algum iterado $f^n(x)$, $n \geq 1$, que também está em E.*

Em outras palavras, o teorema afirma que quase todo ponto de E regressa a E no futuro. Antes mesmo de demonstrar este fato, podemos mostrar que ele implica outro aparentemente mais forte: quase todo ponto de E regressa a E infinitas vezes:

Corolário 2.3. *Nas condições do Teorema 2.2, para μ-quase todo ponto $x \in E$ existem infinitos valores de $n \geq 1$ tais que $f^n(x)$ está em E.*

Demonstração. Sejam F o conjunto dos pontos $x \in E$ que regressam a E apenas um número finito de vezes e E^0 o conjunto dos pontos $x \in E$ que nunca regressam a E.

Note que $F \subset \cup_{k=0}^{\infty} f^{-k}(E^0)$. Portanto, como $\mu(E^0) = 0$ e μ é invariante, temos:

$$\mu(F) \leq \mu(\cup_{k=0}^{\infty} f^{-k}(E^0)) \leq \sum_{k=0}^{\infty} \mu(f^{-k}(E^0)) = \sum_{k=0}^{\infty} \mu(E^0) = 0$$

Ou seja, $\mu(F) = 0$. E o corolário está provado. $\square$

Vamos agora dar a

Demonstração do Teorema 2.2. Representemos por E^0 o conjunto dos pontos $x \in E$ que nunca regressam a E. O nosso objetivo é provar que E^0 tem medida nula. Para isso, começamos por afirmar que as suas pré-imagens $f^{-n}(E^0)$ são disjuntas duas-a-duas. De fato, suponhamos que existem $m > n \geq 1$ tais que $f^{-m}(E^0)$ intersecta $f^{-n}(E^0)$. Seja x um ponto na intersecção e seja $y = f^n(x)$. Então $y \in E^0$ e $f^{m-n}(y) = f^m(x) \in E^0$, que está contido em E. Isto quer dizer que y volta pelo menos uma vez a E, o que contradiz a definição de E^0. Esta contradição, prova que as pré-imagens são disjuntas duas-a-duas, como afirmamos.

Isto implica que

$$\mu\big(\bigcup_{n=0}^{\infty} f^{-n}(E^0)\big) = \sum_{n=0}^{\infty} \mu(f^{-n}(E^0)) = \sum_{n=0}^{\infty} \mu(E^0).$$

Na última igualdade usamos a hipótese de que μ é invariante, que implica que $\mu(f^{-n}(E^0)) = \mu(E^0)$ para todo $n \geq 1$. Como supomos que a medida é finita, a expressão do lado esquerdo é finita. Por outro lado, à direita temos uma soma de infinitos termos, todos iguais. O único jeito desta soma ser finita é que as parcelas sejam nulas. Portanto, devemos ter $\mu(E^0) = 0$, tal como foi afirmado. □

2.2.2 Versão topológica

Dizemos que um ponto $x \in M$ é *recorrente* para uma transformação $f : M \to M$ se, para toda vizinhança U de x, existe algum iterado $f^n(x)$ que está em U. A definição para fluxos é análoga, apenas nesse caso o tempo n é um número real.

Na formulação topológica do Teorema de Recorrência supomos que o espaço M admite uma base enumerável de abertos, ou seja, uma família enumerável $\{U_k : k \in \mathbb{N}\}$ de abertos tal que todo aberto de M pode ser escrito como união de elementos U_k dessa família. Esta hipótese é satisfeita na maioria dos exemplos interessantes.

Teorema 2.4. *Suponhamos que M admite uma base enumerável de abertos. Seja $f : M \to M$ uma transformação mensurável e μ uma medida invariante finita. Então, μ-quase todo ponto $x \in M$ é recorrente para f.*

Demonstração. Para cada k representamos por U_k^0 o conjunto dos pontos $x \in U_k$ que nunca regressam a U_k. De acordo com o Teorema 2.2, todo U_k^0 tem medida nula. Consequentemente, a união enumerável

$$\tilde{U} = \bigcup_{k \in \mathbb{N}} U_k^0$$

tem medida nula. Portanto, para demonstrar o teorema será suficiente que mostremos que todo ponto x que não está em $\tilde{U}$ é recorrente. Isso é fácil, como vamos ver.

Seja $x \in M \setminus \tilde{U}$ e seja U uma vizinhança qualquer de x. A definição de base de abertos implica que existe algum $k \in \mathbb{N}$ tal que $x \in U_k$ e $U_k \subset U$. Como x não está em $\tilde{U}$, também $x \notin U_k^0$. Em outras palavras, x tem algum iterado $f^n(x), n \geq 1$ que está em U_k. Em particular, $f^n(x)$ também está em U. Como a vizinhança U é arbitrária, isto prova que x é um ponto recorrente, como havíamos afirmado. □

Observe que as conclusões dos Teoremas 2.2 e 2.4 não são verdadeiras, em geral, se omitirmos a hipótese de que a medida μ é finita. O exemplo mais simples é o seguinte:

Exemplo 2.5. Seja $f : \mathbb{R} \to \mathbb{R}$ a translação de 1 unidade, isto é, $f(x) = x + 1$ para todo $x \in \mathbb{R}$. É fácil verificar que f deixa invariante a medida de Lebesgue em $\mathbb{R}$ (que é infinita). Por outro lado nenhum ponto é recorrente para f. Portanto, pelo Teorema de Recorrência, f não pode admitir nenhuma medida invariante finita.

No entanto, é possível estender estes enunciados para certos casos de medidas infinitas: veja o Exercício 2.2.

Na seções seguintes vamos descrever alguns exemplos simples de medidas invariantes por transformações ou por fluxos, que nos ajudam a interpretar o significado do Teorema de Recorrência de Poincaré.

2.3 Exemplos de Medidas Invariantes

2.3.1 Expansão decimal

O nosso primeiro exemplo é a transformação definida no intervalo $[0,1]$ do seguinte modo

$$f : [0,1] \to [0,1], \quad f(x) = 10x - [10x]$$

onde $[10x]$ representa o maior inteiro menor ou igual a $10x$. Em outras palavras, f associa a cada $x \in [0,1]$ a parte fracionária de $10x$. O gráfico da transformação f está descrito na Figura 2.1.

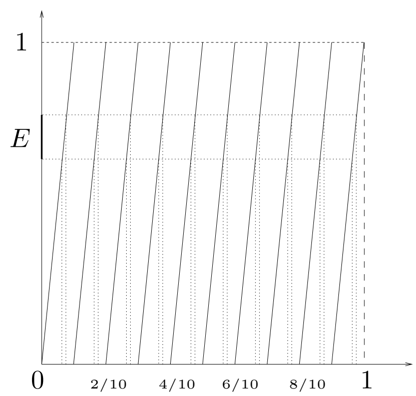

Figura 2.1: Transformação parte fracionária de $10x$

Afirmamos que a medida de Lebesgue μ no intervalo é invariante pela transformação f, isto é, satisfaz a condição

$$\mu(E) = \mu(f^{-1}(E)) \quad \text{para todo conjunto mensurável } E \subset M. \tag{2.2}$$

Comecemos por supor que E é um intervalo. Então, como ilustrado na Figura 2.1, a pré-imagem $f^{-1}(E)$ consiste de dez intervalos, cada um deles dez vezes mais curto do que E. Logo, a medida de Lebesgue de $f^{-1}(E)$ é igual à medida de Lebesgue de E. Isto mostra que (2.2) é satisfeita no caso de intervalos. Por outro lado, a família dos intervalos gera a σ-álgebra de Borel de $[0,1]$. Portanto, para concluir a demonstração basta usar o seguinte fato geral:

Lema 2.6. *Seja $f : M \to M$ uma transformação mensurável e μ uma medida finita em M. Suponha que existe uma sub-álgebra geradora $\mathcal{I}$ da σ-álgebra de M tal que $\mu(E) = \mu(f^{-1}(E))$ para todo $E \in \mathcal{I}$. Então o mesmo vale para todo conjunto mensurável E, isto é, a medida μ é invariante por f.*

Demonstração. Faça o Exercício 2.5. □

Agora vamos explicar como, a partir do fato de que a medida de Lebesgue é invariante pela transformação f, podemos obter conclusões interessantes e não-triviais usando o Teorema de Recorrência de Poincaré. Como observamos no capítulo anterior, f tem uma expressão muito simples em termos de expansões decimais: se x é dado por

$$x = 0, a_0 a_1 a_2 a_3 \cdots$$

com $a_i \in \{0,1,2,3,4,5,6,7,8,9\}$, então a sua imagem é dada por

$$f(x) = 0, a_1 a_2 a_3 \cdots .$$

Com isso, fica muito fácil escrever a expressão do iterado n-ésimo, para qualquer $n \geq 1$:

$$f^n(x) = 0, a_n a_{n+1} a_{n+2} \cdots \tag{2.3}$$

Agora, seja E o subconjunto dos $x \in [0,1]$ cuja expansão decimal começa com o dígito 7, ou seja, tais que $a_0 = 7$. De acordo com o Corolário 2.3, quase todo elemento de E tem infinitos iterados que também estão em E. Levando em conta a expressão (2.3), isto quer dizer que existem infinitos valores de n tais que $a_n = 7$. Portanto, provamos que *quase todo número x cuja expansão decimal começa por* 7 *tem infinitos dígitos iguais a* 7!

Claro que no lugar de 7 podemos considerar qualquer outro dígito. Além disso, também podemos considerar blocos de dígitos mais complicados (veja os Exercícios 2.6–2.7). Mais tarde provaremos resultados mais fortes: para quase todo número $x \in [0,1]$, todo dígito aparece com frequência $1/10$ na sua expansão decimal. O enunciado preciso aparecerá na Proposição 4.2, que será obtida a partir do teorema ergódico de Birkhoff.

2.3.2 Transformação de Gauss

A transformação de Gauss $G : (0,1] \to [0,1]$ é definida por $G(x) =$ parte fracionária de $1/x$, ou seja,

$$G(x) = \frac{1}{x} - \left[\frac{1}{x}\right].$$

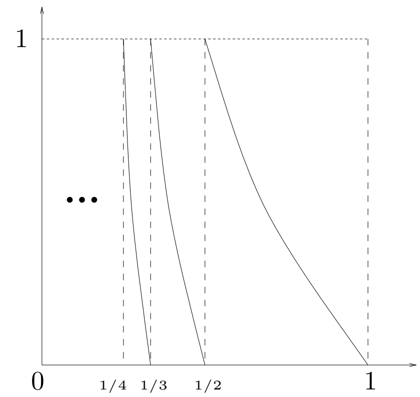

Figura 2.2: Transformação de Gauss

O gráfico de G pode ser esboçado facilmente, a partir da seguinte observação.

- Se $x \in (1/2, 1]$ então $1/x \in [1, 2)$ e portanto a sua parte inteira $[1/x]$ é igual a 1. Isto quer dizer que neste intervalo a transformação é dada por $G(x) = (1/x) - 1$.
- Mais geralmente, se $x \in (1/(k+1), 1/k)$ para algum $k \in \mathbb{N}$ então a parte inteira de $1/x$ é igual a k, e tem-se $G(x) = 1/x - k$. Veja também a Figura 2.2.

Note que G não está definida no ponto $x = 0$. Além disso, $G(1/k) = 0$ para todo $k \in \mathbb{N}$ e portanto o segundo iterado $G^2(1/k)$ não está definido nestes pontos (e o terceiro iterado não está definido nas suas pré-imagens, etc). Isto quer dizer, a rigor, que G não é um sistema dinâmico segundo a definição que demos antes. No entanto, isto não coloca nenhum problema para o que pretendemos fazer. De fato, todos os iterados estão bem definidos no conjunto dos números irracionais: basta observar que a imagem de um irracional também é irracional. Isto é suficiente para os nossos objetivos porque sempre tratamos de propriedades que valem para quase todo ponto, e o conjunto dos números irracionais tem medida de Lebesgue total no intervalo.

O que torna esta transformação interessante do ponto de vista ergódico é que G admite uma probabilidade invariante que é equivalente à medida de Lebesgue no intervalo. De fato, considere a medida definida por

$$\mu(E) = \int_E \frac{c}{1+x}\, dx \quad \text{para cada mensurável } E \subset [0,1]$$

onde c é uma constante positiva. Note que a integral está bem definida, já que a função integranda é contínua no intervalo $[0, 1]$. Note também que

$$\frac{c}{2}\, m(E) \leq \mu(E) \leq c m(E) \quad \text{para todo mensurável } E \subset [0,1].$$

Em particular, μ é de fato equivalente à medida de Lebesgue m: as duas medidas têm os mesmos conjuntos com medida nula.

Proposição 2.7. *A medida μ é invariante por G. Além disso, se escolhermos $c = 1/log2$ então μ é uma probabilidade.*

Demonstração. Vamos usar o critério dado pelo exercício 2.10: a medida μ é invariante por G se tivermos

$$\sum_{x \in g^{-1}(y)} \frac{\rho(x)}{|G'(x)|} = \rho(y) \qquad \text{onde } \rho(x) = \frac{c}{1+x} \tag{2.4}$$

para todo $y \in [0,1]$. Comece por observar que cada y tem exatamente uma pré-imagem x_k em cada intervalo $(1/(k+1), 1/k]$, dada por

$$G(x_k) = \frac{1}{x_k} - k = y \quad \Leftrightarrow \quad x_k = \frac{1}{y+k}\,.$$

Note também que $G'(x) = (1/x)' = -1/x^2$. Portanto, (2.4) se reescreve como

$$\sum_{k=1}^{\infty} \frac{cx_k^2}{1+x_k} = \frac{c}{1+y} \quad \Leftrightarrow \quad \sum_{k=1}^{\infty} \frac{c}{(y+k)(y+k+1)} = \frac{c}{1+y} \tag{2.5}$$

Para verificar que esta igualdade é realmente satisfeita, observe que

$$\frac{1}{(y+k)(y+k+1)} = \frac{1}{y+k} - \frac{1}{y+k+1}\,.$$

Isto quer dizer que a última soma em (2.5) pode ser escrita na forma telescópica: todos os termos, exceto o primeiro, aparecem duas vezes, com sinais contrários, e portanto se cancelam. Logo a soma é igual ao primeiro termo, que é precisamente o que se afirma em (2.5). Isto prova a invariância.

Finalmente, usando a primitiva $c\log(1+x)$ da função $\rho(x)$ vemos que

$$\mu([0,1]) = \int_0^1 \frac{c}{1+x}\,dx = c\log 2.$$

Logo, escolhendo $c = 1/\log 2$ obtemos que μ é uma probabilidade. □

A transformação de Gauss tem um papel muito importante em teoria dos números, devido à sua relação com o processo de expansão dos números em fração contínua. Recordemos do que se trata.

Dado um número $x_0 \in (0,1)$, seja

$$a_1 = \left[\frac{1}{x_0}\right] \qquad \text{e} \qquad x_1 = \frac{1}{x_0} - a_1 = G(x_0).$$

Note que a_1 é um número natural, $x_1 \in [0,1)$ e tem-se

$$x_0 = \frac{1}{a_1 + x_1}\,.$$

Agora, supondo que x_1 seja diferente de zero, podemos repetir o processo, definindo

$$a_2 = \left[\frac{1}{x_1}\right] \qquad \text{e} \qquad x_2 = \frac{1}{x_1} - a_2 = G(x_1).$$

Então

$$x_1 = \frac{1}{a_1 + x_2} \quad \text{portanto} \quad x_0 = \cfrac{1}{a_1 + \cfrac{1}{a_2 + x_2}}.$$

Por recorrência, para cada $n \geq 1$ tal que $x_{n-1} \in (0,1)$ se define

$$a_n = \left[\frac{1}{x_{n-1}}\right] \qquad \text{e} \qquad x_n = \frac{1}{x_{n-1}} - a_n = G(x_{n-1})$$

e tem-se

$$x_0 = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n + x_n}}}}. \tag{2.6}$$

Não é difícil mostrar (verifique!) que a sequência

$$z_n = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}}.$$

converge para x_0 quando $n \to \infty$, e é usual traduzir este fato escrevendo

$$x_0 = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n + \cfrac{1}{\cdots}}}}}, \tag{2.7}$$

que é chamada *expansão em fração contínua* de x_0.

Note que a sequência z_n consiste de números racionais. De fato se mostra que estes são os números racionais que melhor aproximam o número x_0, no sentido de que z_n está mais próximo de x_0 do que qualquer outro número racional com denominador menor ou igual que o denominador de z_n (escrito em forma irredutível). Observe também que para obter (2.7) supusemos que $x_n \in (0,1)$ para todo $n \in \mathbb{N}$. Se encontramos algum $x_n = 0$, o processo pára nesse momento e consideramos (2.6) a expansão em fração contínua de x_0. Claro que este último caso ocorre somente se x_0 é um número racional.

Estas idéias de Teoria Ergódica podem ser usadas para obter conclusões não triviais em Teoria dos Números. Por exemplo (veja o Exercício 2.8), para quase todo número $x_0 \in (1/8, 1/7)$ o número 7 aparece infinitas vezes na sua expansão em fração contínua, isto é, tem-se $a_n = 7$ para infinitos valores de $n \in \mathbb{N}$.

De fato, mais tarde provaremos um fato muito mais preciso: para quase todo $x_0 \in (0, 1)$ o número 7 aparece com frequência

$$\frac{1}{\log 2} \log \frac{64}{63}$$

na sua expansão em fração contínua. Tente intuir desde já de onde vem este número!

2.3.3 Rotações

Um exemplo bem natural de sistema dinâmico preservando uma medida são as *rotações*. Uma rotação de ângulo θ é uma função

$$R_\theta : S^1 \to S^1, \quad R_\theta(x) = x + \theta \mod \mathbb{Z}.$$

Observe que como a derivada de R_θ é identicamente igual a 1, temos que R_θ preserva a medida de Lebesgue de S^1. Aqui destacamos que a dinâmica de R_θ possui dois comportamentos bem distintos:

- Se θ é racional, digamos $\theta = p/q$ com $p, q \in \mathbb{Z}$, temos que $R_\theta^q(x) = x + q\theta = x + p = x \mod \mathbb{Z}$. Como consequência disso, todo ponto $x \in S^1$ é periódico de período no máximo q;
- Se θ é irracional, pode ser mostrado que $\mathcal{O}(x) = \{R_\theta^n; n \in \mathbb{N}\}$ é um conjunto denso em S^1. Vamos fazer isso na proposição a seguir:

Proposição 2.8. *Se θ é irracional, então $\mathcal{O}(x) = \{R_\theta^n; n \in \mathbb{N}\}$ é um conjunto denso em S^1.*

Demonstração. Afirmamos que o conjunto $\mathcal{D} = \{m + n\theta; m \in \mathbb{Z}, n \in \mathbb{N}\}$ é denso em $\mathbb{R}$. De fato, afirmamos que para todo natural N, podemos escolher p, q inteiros tais que

$$|q\theta - p| < \frac{1}{N}.$$

Como θ é irracional, temos que $a = |q\theta - p| > 0$. Sem perda de generalidade, assuma que $a = q\theta - p$ (o outro caso é análogo). Podemos subdividir a reta em intervalos de comprimento a e concluir que existe um inteiro M tal que $Ma \le r < (M+1)a$. Logo,

$$|r - Ma| < |r - M(q\theta - p)| = |r - \big((Mq)\theta + (-Mp)\big)| < \frac{1}{N}.$$

Tomando $n = Mq$ e $m = -Mp$, temos o que queríamos.

Agora para mostrarmos que $\mathcal{O}(x)$ é denso em S^1, observe que dado $y \in [0, 1]$ e $\epsilon > 0$, podemos tomar $r = y - x$ e escolher m, n tais que $|m + n\theta - (y - x)| < \epsilon$. Isso equivale a dizer que $d(R_\theta^n(x), y) < \epsilon$ em S^1, como queríamos demonstrar. $\square$

A proposição anterior terá várias implicações interessantes no estudo das medidas invariantes de R_θ. Entre outras coisas, veremos na Proposição **??** que se θ é irracional a medida de Lebesgue é a única medida que é preservada por R_θ. Além disso, veremos que as órbitas de R_θ se distribuem de modo uniforme em S^1. Isso ficará mais claro no Capítulo 4.

2.3.4 Deslocamentos ("shifts") de Bernoulli

Estes sistemas modelam sequências de experimentos aleatórios em que o resultado de cada experimento é independente dos demais. Supõe-se que em cada experimento há um número finito de resultados possíveis, designados por $1, 2, \dots, d$, com probabilidades $p(1), p(2), \dots, p(d)$ de ocorrerem, sendo

$$p(1) + p(2) + \cdots + p(d) = 1\,.$$

O conjunto M das sequências $\underline{\alpha} = (\alpha_n)_{n\in\mathbb{Z}}$ com cada $\alpha_n \in \{1, 2, \dots, d\}$ contém os possíveis resultados da sequência de experimentos. Chamam-se *cilindros* os subconjuntos da forma

$$[k, l; a_k\,, \dots, a_l] = \{\underline{\alpha} \in M : \alpha_k = a_k\,, \dots, \alpha_l = a_l\}$$

onde $k, l \in \mathbb{Z}$, com $k \le l$, e cada $a_j \in \{1, 2, \dots, d\}$. Definimos

$$\mu([k, l; a_k\,, \dots, a_l]) = p(a_k) \cdots p(a_l) \tag{2.8}$$

Heuristicamente, isto significa que a probabilidade do evento composto

$$\alpha_k = a_k \quad \text{e} \quad \alpha_{k+1} = a_{k+1} \quad \text{e} \quad \cdots \quad \text{e} \quad \alpha_l = a_l$$

é o produto das probabilidades de cada um deles. Isto traduz, precisamente, que os resultados sucessivos são independentes entre si.

Consideramos em M a σ-álgebra $\mathcal{B}$ gerada pelos cilindros. A família $\mathcal{B}_0$ das uniões disjuntas finitas dos cilindros é uma álgebra (por convenção, M é um cilindro e $\mu(M) = 1$). Estendemos μ de modo a que seja finitamente aditiva: se $E \in \mathcal{B}_0$ é a união disjunta de cilindros $C_1\,, \dots, C_N$, definimos

$$\mu(E) = \mu(C_1) + \cdots + \mu(C_N).$$

Verifica-se que esta função μ é, de fato, σ-aditiva em $\mathcal{B}_0$; por exemplo, isso pode ser feito usando o Teorema 1.16. Portanto existe uma única probabilidade na σ-álgebra $\mathcal{B}$ gerada por $\mathcal{B}_0$ que é uma extensão de μ, isto é, que coincide com ela restrita a $\mathcal{B}_0$. Chamamos essa probabilidade *medida de Bernoulli* definida por $p(1), p(2), \dots, p(d)$ e, para não complicar desnecessariamente a notação, a representamos também por μ.

No espaço M consideramos a transformação deslocamento ("shift") à esquerda

$$f : M \to M \quad f\big((\alpha_n)_{n\in\mathbb{Z}}\big) = (\alpha_{n+1})_{n\in\mathbb{Z}}$$

que corresponde a fazer uma translação no tempo. Observe que a medida de Bernoulli é invariante por essa transformação. De fato, se $E = [k, l; a_k, \dots, a_l]$ então $f^{-1}(E) = [k+1, l+1; a_k, \dots, a_l]$ e a definição (4.3) dá que

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

neste caso. Como a família dos cilindros gera a σ-álgebra $\mathcal{B}$, isto juntamente com o Lema 2.6, prova que a medida μ é invariante para f.

2.3.5 Sistemas conservativos

Seja U um aberto em algum espaço euclidiano $\mathbb{R}^d$, $d \geq 1$ e seja $f : U \to U$ um difeomorfismo de classe C^1. Isto quer dizer que f é uma bijeção e tanto ele quanto a sua inversa são deriváveis com derivada contínua.

Representaremos por vol a medida de Lebesgue, ou volume, em $\mathbb{R}^k$. Em outras palavras,

$$\operatorname{vol}(B) = \int_B dx_1 \dots dx_d \quad \text{e} \quad \int_B \varphi \, d\operatorname{vol} = \int_B \varphi(x_1, \dots, x_d) \, dx_1 \dots dx_d$$

para qualquer conjunto mensurável B e qualquer função integrável φ.

A fórmula de mudança de variáveis afirma que, para qualquer conjunto mensurável $B \subset U$,

$$\operatorname{vol}(f(B)) = \int_B |\det Df| \, d\operatorname{vol} \tag{2.9}$$

Daqui se deduz facilmente

Lema 2.9. *Um difeomorfismo $f : M \to M$ de classe C^1 deixa invariante o volume se e somente se o valor absoluto $|\det Df|$ do seu jacobiano é constante igual a* 1.

Demonstração. Suponha primeiro que o valor absoluto do jacobiano é igual 1 em todo ponto. Considere um conjunto mensurável E e seja $B = f^{-1}(E)$. A fórmula (2.9) dá que

$$\operatorname{vol}(E) = \int_B 1 \, d\operatorname{vol} = \operatorname{vol}(B) = \operatorname{vol}(f^{-1}(E)).$$

Isto significa que f deixa invariante o volume e, portanto, provamos a parte "se"do enunciado.

Para provar a parte "somente se", suponha que $|\det Df|$ fosse maior que 1 em algum ponto x. Então, como o jacobiano é contínuo, existiria uma vizinhança U de x e algum número $\sigma > 1$ tais que

$$|\det Df(y)| \geq \sigma \quad \text{para todo } y \in U.$$

Então a fórmula (2.9) aplicada a $B = U$ daria

$$\operatorname{vol}(f(U)) \geq \int_U \sigma \, d\operatorname{vol} \geq \sigma \operatorname{vol}(U).$$

Denotando $E = f(U)$, isto implica que $\mathrm{vol}(E) > \mathrm{vol}(f^{-1}(E))$ e, portanto, f não deixa invariante o volume. Do mesmo modo se mostra que se o valor absoluto do jacobiano é menor que 1 em algum ponto então f não deixa invariante o volume. □

Os Exercícios 2.9–2.10 estendem este lema para transformações não necessariamente invertíveis e também para uma classe mais ampla de medidas. As suas conclusões nos serão úteis mais tarde.

Agora vamos considerar o caso de fluxos $f^t : U \to U$, $t \in \mathbb{R}$. Suporemos que o fluxo é de classe C^1. Claro que o Lema 2.9 se aplica neste contexto: o fluxo deixa invariante o volume se e somente se

$$\det Df^t(x) = 1 \quad \text{para todo } x \in U \text{ e todo } t \in \mathbb{R}. \tag{2.10}$$

Façamos duas observações simples antes de prosseguirmos. A primeira é que segue da definição de fluxo que todo f^t é invertível (um difeomorfismo, neste caso): a sua inversa é f^{-t}. A segunda observação é que o jacobiano de f^t é sempre positivo. Isso é claro quando $t = 0$ porque, outra vez por definição de fluxo, f^0 é a identidade. Segue que o mesmo é verdade para todo $t \in \mathbb{R}$, porque o jacobiano varia continuamente com t e, como acabamos de ver, nunca se anula.

Embora a resposta que acabamos de dar esteja inteiramente correta, ela não é muito útil na prática porque em geral não temos uma expressão explícita para f^t, e portanto não é claro como verificar a condição (2.10). Felizmente, existe uma expressão razoavelmente explícita para o jacobiano, de que iremos falar em seguida, que pode ser usada em muitas situações interessantes.

Suponhamos que o fluxo f^t corresponde às trajetórias de um campo de vetores $F : U \to U$ de classe C^1, quer dizer $f^t(x)$ é o valor no tempo t da solução da equação diferencial

$$\frac{dx}{dt} = F(x) \tag{2.11}$$

(quando tratando de equações diferenciáveis sempre suporemos que as suas soluções estão definidas para todo tempo). A fórmula de Liouville exprime o jacobiano de f^t em termos do divergente $\operatorname{div} F$ do campo de vetores F:

$$\det Df^t(x) = \exp\Big(\int_0^t \operatorname{div} F(f^s(x))\, ds\Big).$$

Lembre que o divergente de um campo de vetores F é o traço da sua matriz jacobiana, isto é

$$\operatorname{div} F = \frac{\partial F_1}{\partial x_1} + \cdots + \frac{\partial F_d}{\partial x_d}. \tag{2.12}$$

Combinando esta fórmula com (2.10) obtemos

Lema 2.10. *O fluxo f^t associado a um campo de vetores F de classe C^1 deixa invariante o volume se e somente se o divergente de F é identicamente nulo.*

O Exercício 2.11 é uma aplicação deste fato no caso, muito importante, de fluxos hamiltonianos.

Na próxima seção, vamos enunciar um teorema que garante a existência de medidas invariantes em grande generalidade.

2.4 Teorema de Existência de Medidas Invariantes

Teorema 2.11. *Seja $f : M \to M$ uma transformação contínua num espaço métrico compacto. Então existe pelo menos uma probabilidade invariante por f. O mesmo resultado vale para fluxos.*

A demonstração foge ao âmbito deste curso e será omitida. Veremos a seguir algus exemplos que mostram que compassidade e continuidade são indispensáveis para a validez do teorema.

2.4.1 Alguns exemplos simples

Considere $f : (0,1] \to (0,1]$ dada por $f(x) = x/2$. Suponha que f admite alguma probabilidade invariante (o objetivo é mostrar que isso não acontece). Pelo Teorema de Recorrência 2.4, relativamente a essa probabilidade quase todo ponto de $(0,1]$ é recorrente. Mas é imediato que não existe nenhum ponto recorrente: a órbita de qualquer $x \in (0,1]$ converge para zero e, em particular, não acumula no ponto inicial x. Isto mostra que f é um exemplo de transformação contínua num espaço não compacto que não admite nenhuma medida probabilidade invariante.

Modificando um pouco o exemplo, podemos mostrar que o mesmo fenômeno pode ocorrer em espaços compactos, se a transformação não é contínua. Considere $f : [0,1] \to [0,1]$ dada por $f(x) = x/2$ se $x \neq 0$ e $f(0) = 1$. Pela mesma razão que antes, nenhum ponto $x \in (0,1]$ é recorrente. Portanto, se existe alguma probabilidade invariante μ ela tem dar peso total ao único ponto recorrente que é $x = 0$. Em outras palavras, μ precisa ser a medida de Dirac δ_0 suportada em zero, que é definida por

$$\delta_0(E) = 1 \text{ se } 0 \in E \quad \text{e} \quad \delta_0(E) = 0 \text{ se } 0 \notin E.$$

Mas a medida δ_0 *não é* invariante por f: tomando $E = \{0\}$ temos que E tem medida 1 mas a sua pré-imagem $f^{-1}(E)$ é o conjunto vazio, que tem medida nula. Portanto, esta transformação também não tem nenhuma probabilidade invariante.

2.5 Exercícios

2.1. Mostre que o seguinte enunciado é equivalente ao Teorema 2.2, isto é, qualquer um dos dois pode ser deduzido a partir do outro: Seja $f : M \to M$

uma transformação mensurável e μ uma medida invariante finita. Seja $E \subset M$ qualquer conjunto mensurável com $\mu(E) > 0$. Então existe $N \geq 1$ e um conjunto $D \subset E$ com medida positiva, tal que $f^N(x) \in E$ para todo ponto $x \in D$.

2.2. Suponha que $f : M \to M$ é invertível e que μ é uma medida σ-finita invariante por f. Mostre que, dado qualquer conjunto mensurável $E \subset M$ com $\mu(E) > 0$, quase todo ponto $x \in E$ ou regressa a E ou "vai para infinito".

Dica: Considere o conjunto $E^{0,k}$ dos pontos $x \in E$ que nunca regressam a E e têm um número infinito de iterados em M_k. Comece por mostrar que os seus iterados $f^n(E^{0,k})$ são dois-a-dois disjuntos. Usando que $\mu(M_k)$ é finito, deduza que $\mu(E^{k,0}) = 0$ para todo k.

2.3. Seja $f : M \to M$ uma transformação não necessariamente invertível, μ uma probabilidade invariante e $D \subset M$ um conjunto com medida positiva. Defina $D_0 = D$ e, para cada $n = 1, 2, \ldots$, $D_n = f^{-1}(D_{n-1}) \setminus D$ e $E_n = f^{-1}(D_{n-1}) \cap D$. Mostre que $\mu(D_n)$ converge para zero quando $n \to \infty$.

Dica: Começar pelo caso em que f é invertível. Em geral, aplicar o teorema de recorrência 'a extensão natural de f.

2.4. Seja $f : M \to M$ uma transformação não necessariamente invertível, μ uma probabilidade invariante e $D \subset M$ um conjunto com medida positiva. Prove que quase todo ponto de D passa uma fração positiva do tempo em D:

$$\limsup_{n\to\infty} \frac{1}{n} \#\{0 \leq j \leq n-1 : f^j(x) \in D\} > 0$$

para μ-quase todo ponto $x \in D$. (Dica: considere o subconjunto dos pontos onde o $\limsup = 0$ e use o Teorema da Convergencia Dominada.) Dá para substituir $\limsup$ por $\liminf$ no enunciado ?

2.5. Demonstre o Lema 2.6. Dica: mostre que a família de todos os conjuntos E tais que $\mu(E) = \mu(f^{-1}(E))$ é uma σ-álgebra.

2.6. Prove que, para quase todo número $x \in [0, 1]$ cuja expansão decimal contém o bloco 617 (por exemplo $x = 0,3375\mathbf{617}264\cdots$), esse bloco aparece infinitas vezes na expansão.

2.7. Prove que o dígito 7 aparece infinitas vezes na expansão decimal de quase todo número $x \in [0, 1]$. Dica: Comece por mostrar que o conjunto dos números cuja expansão decimal *nunca* exibe o dígito 7 tem medida nula.

2.8. Para (Lebesgue) quase todo número $x_0 \in (1/8, 1/7)$ o número 7 aparece infinitas vezes na sua expansão em fração contínua, isto é, tem-se $a_n = 7$ para infinitos valores de $n \in \mathbb{N}$.

2.9. Suponha que $f : U \to U$ é um difeomorfismo local (isto é: o seu jacobiano é não nulo em todo ponto) de classe C^1 tal que todo ponto tem um número finito de pré-imagens. Mostre que f deixa invariante o volume se e somente se

$$\sum_{x\in f^{-1}(y)} \frac{1}{|\det Df(x)|} = 1 \quad \text{para todo } y \in U.$$

2.10. Dada uma função $\rho : U \to [0, \infty)$, denotamos por $\mu = \rho \operatorname{vol}$ a medida definida por $\mu(E) = \int_E \rho \, d \operatorname{vol}$. Suponha que $f : U \to U$ é um difeomorfismo local de classe C^1 tal que todo ponto tem um número finito de pré-imagens e que ρ é uma função contínua. Mostre que f deixa invariante a medida $\mu = \rho \operatorname{vol}$ se e somente se

$$\sum_{x \in f^{-1}(y)} \frac{\rho(x)}{|\det Df(x)|} = \rho(y) \quad \text{para todo } y \in U.$$

Em particular, no caso em que f é invertível, f deixa invariante a medida μ se e somente se $\rho(x) = \rho(f(x))|\det Df(x)|$ para todo $x \in U$.

2.11. Seja U um aberto de $\mathbb{R}^{2d}$ e $H : U \to \mathbb{R}$ uma função de classe C^2. Denotamos as variáveis em $\mathbb{R}^{2d}$ por $(p_1, q_1, \ldots, \ldots, p_d, q_d)$. O campo de vetores hamiltoniano associado a H é definido por

$$F(p_1\,, \ldots, p_d\,, q_1\,, \ldots, q_d) = \left(\frac{\partial H}{\partial q_1}\,, \ldots, \frac{\partial H}{\partial q_d}\,, -\frac{\partial H}{\partial p_1}\,, \ldots, -\frac{\partial H}{\partial p_d} \right).$$

Verifique que o fluxo definido por F preserva o volume.

2.12. Considere a sequência $1, 2, 4, 8, \ldots, a_n = 2^n, \ldots$. Mostre que dado um dígito $i \in \{0, \ldots, 9\}$, existe uma quantidade infinita de valores n tal que a_n começa com este dígito.

2.13. Se $\theta_1, \theta_2, \ldots, \theta_d$ são racionalmente independentes então colocando $\Theta = (\theta_1, \theta_2, \ldots, \theta_d)$, prove que $\mathcal{O}(x) = \{R_\Theta^n; n \in \mathbb{N}\}$ é um conjunto denso em Π^n.

2.14. Considere a sequência $1, 2, 4, 8, \ldots, a_n = 2^n, \ldots$. Mostre que dado um dígito $i \in 0, \ldots, 9$, existe uma quantidade infinita de valores n tal que a_n começa com este dígito.

2.15. Mostre que o deslocamento σ definido na Secção 4.2.2 é transitivo e que o conjunto de suas órbitas periódicas é denso.

Capítulo 3

Teorema Ergódico de Birkhoff

O teorema fundamental da Teoria Ergódica afirma que, para qualquer subconjunto mensurável e para quase todo ponto, existe um *tempo médio de permanência* da órbita do ponto nesse conjunto. Este resultado é devido a von Neumann, que provou um enunciado mais fraco, e sobretudo a Birkhoff, que o provou na forma definitiva que iremos estudar.

Em muitos casos, esse tempo médio de permanência é precisamente igual à medida do subconjunto, ou seja, órbitas típicas passam em cada subconjunto um tempo que é exatamente igual à "importância" que a probabilidade invariante atribui ao conjunto. Isto é o que se chama de *ergodicidade*, uma propriedade que remonta a Boltzmann, e que estudaremos mais tarde.

3.1 Enunciados e comentários

Comecemos por explicar o que entendemos por tempo médio de permanência de uma órbita num conjunto. Dado $x \in M$ e um conjunto mensurável $E \subset M$, vamos tomar um certo número (grande) de iterados iniciais da órbita de x e vamos considerar a fração desses iterados que estão em E:

$$\tau_n(E,x) = \frac{1}{n}\#\big\{j \in \{0,1,\ldots,n-1\} : f^j(x) \in E\big\}.$$

Observe que isto é o mesmo que

$$\tau_n(E,x) = \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{X}_E(f^j(x)),$$

onde $\mathcal{X}_E$ designa a função característica do conjunto E, isto é, $\mathcal{X}_E(x) = 1$ se $x \in E$ e $\mathcal{X}_E(x) = 0$ caso contrário.

Em seguida, fazemos n ir para infinito e chamamos *tempo médio de permanência* da órbita de x em E o limite destas frações:

$$\tau(E,x) = \lim_{n\to\infty} \tau_n(E,x).$$

Em geral, este limite pode não existir. Iremos ver um exemplo desse fato daqui a pouco. No entanto, o teorema ergódico afirma que, relativamente a qualquer probabilidade invariante, o limite realmente existe para quase todo ponto:

Teorema 3.1. *Seja $f : M \to M$ uma transformação mensurável e μ uma probabilidade invariante por f. Dado qualquer conjunto mensurável $E \subset M$, o tempo médio de permanência $\tau(E,x)$ existe em μ-quase todo ponto $x \in M$. Além disso,*

$$\int \tau(E,x)\, d\mu(x) = \mu(E).$$

Antes de passarmos à demonstração deste resultado notável, e a algumas das suas aplicações, vamos fazer alguns comentários relacionados. O primeiro deles é que se $\tau(E,x)$ existe para um certo ponto $x \in M$ então

$$\tau(E,f(x)) = \tau(E,x). \tag{3.1}$$

De fato, por definição,

$$\begin{aligned}
\tau(E,f(x)) &= \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathcal{X}_E(f^j(x)) \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{X}_E(f^j(x)) - \frac{1}{n}\big[\mathcal{X}_E(x) - \mathcal{X}_E(f^n(x))\big] \\
&= \tau(E,x) + \lim_{n\to\infty} \frac{1}{n}\big[\mathcal{X}_E(x) - \mathcal{X}_E(f^n(x))\big]
\end{aligned}$$

Como a função característica é limitada, o último limite é igual a zero. Isto prova a igualdade (3.1).

O teorema ergódico pode ser enunciado de modo um pouco mais geral:

Teorema 3.2. *Seja $f : M \to M$ uma transformação mensurável e μ uma probabilidade invariante por f. Dada qualquer função integrável $\varphi : M \to \mathbb{R}$ o limite*

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

existe em μ-quase todo ponto $x \in M$. Além disso,

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

Observe que o Teorema 3.1 é o caso particular $\varphi =$ função característica $\mathcal{X}_E$ do conjunto E.

3.2 Demonstração do Teorema 3.1

A estratégia da prova é a seguinte. Seja $E \subset M$ um conjunto mensurável qualquer. Para cada $x \in M$, definimos

$$\overline{\tau}(E, x) = \limsup \frac{1}{n} \# \Big\{ j \in \{0, \ldots, n-1\} : f^j(x) \in E \Big\}$$
$$\underline{\tau}(E, x) = \liminf \frac{1}{n} \# \Big\{ j \in \{0, \ldots, n-1\} : f^j(x) \in E \Big\}.$$

Note que, para todo $x \in M$,

$$\overline{\tau}(E, f(x)) = \overline{\tau}(E, x) \quad \text{e} \quad \underline{\tau}(E, f(x)) = \underline{\tau}(E, x) \tag{3.2}$$

A justificação é análoga à da relação (3.1).

O principal passo da demonstração consiste em mostrar que

$$\overline{\tau}(E, x) = \underline{\tau}(E, x) \quad \text{para } \mu\text{-quase todo ponto } x. \tag{3.3}$$

É claro que $\overline{\tau}(E, x)$ é sempre maior ou igual que $\underline{\tau}(E, x)$. Portanto, para mostrar (3.3) será suficiente que provemos

$$\int \overline{\tau}(E, x)\, d\mu(x) \le \mu(E) \le \int \underline{\tau}(E, x)\, d\mu(x). \tag{3.4}$$

Vamos provar a primeira desigualdade em (3.4). A segunda segue de um argumento inteiramente análogo [1].

Fixemos qualquer $\varepsilon > 0$. Por definição de lim sup, para cada $x \in M$ existem inteiros $t \ge 1$ tais que

$$\frac{1}{t} \# \Big\{ j \in \{0, \ldots, t-1\} : f^j(x) \in E \Big\} \ge \overline{\tau}(E, x) - \varepsilon. \tag{3.5}$$

Representaremos por $t(x)$ o menor inteiro com esta propriedade. Para tornar a demonstração mais transparente, consideraremos primeiro o caso particular em que a função $x \mapsto t(x)$ é limitada, isto é,

Caso particular: Existe $T \in \mathbb{N}$ tal que $t(x) \le T$ para todo $x \in M$.

Dado qualquer $x \in M$, definimos uma sequência $x_0, x_1, \ldots, x_s$ de pontos em M e uma sequência $t_0, t_1, \ldots, t_s$ de números naturais, do seguinte modo:

1. Primeiramente, tomamos $x_0 = x$.
2. Supondo que x_i já foi definido, tomamos $t_i = t(x_i)$ e $x_{i+1} = f^{t_i}(x_i)$.
3. Terminamos quando encontramos x_s tal que $t_0 + t_1 + \cdots + t_{s-1} + t_s \ge n$.

[1] Alternativamente, a segunda desigualdade pode ser deduzida da primeira, aplicada ao complementar E^c, observando que $\mu(E) = 1 - \mu(E^c)$ e $\underline{\tau}(E, x) = 1 - \overline{\tau}(E^c, x)$.

Note que todo x_i é iterado do ponto x: de fato $x_i = f^{t_0+\cdots+t_{i-1}}(x)$. Aplicando (3.2) concluímos que $\overline{\tau}(E, x_i) = \overline{\tau}(E, x)$ para todo i. A definição de $t(x_i)$ implica que, dos t_i primeiros iterados de x_i, pelo menos

$$t_i\big(\overline{\tau}(E, x_i) - \varepsilon\big) = t_i\big(\overline{\tau}(E, x) - \varepsilon\big) \tag{3.6}$$

estão em E. Isto vale para cada $i = 0, 1, \ldots, s-1$. Portanto, pelo menos

$$(t_0 + t_1 + \cdots + t_{s-1})(\overline{\tau}(E, x) - \varepsilon)$$

dos n primeiros iterados de x, estão em E. Além disso, a última regra na definição das nossas sequências implica que

$$t_0 + t_1 + \cdots + t_{s-1} \geq n - t_s \geq n - T.$$

Deste modo, mostramos que pelo menos $(n - T)(\overline{\tau}(E, x) - \varepsilon)$ dos n primeiros iterados de x estão em E. Em outras palavras,

$$\sum_{j=0}^{n-1} \mathcal{X}_E(f^j(x)) \geq (n - T)(\overline{\tau}(E, x) - \varepsilon) \tag{3.7}$$

para todo $x \in M$ e todo $n \geq 1$. Integrando a relação (3.7), obtemos que

$$\sum_{j=0}^{n-1} \int \mathcal{X}_E(f^j(x))\, d\mu(x) \geq (n - T) \int \overline{\tau}(E, x)\, d\mu(x) - (n - T)\varepsilon.$$

Todas as parcelas no membro da esquerda são iguais a $\mu(E)$, uma vez que a probabilidade μ é invariante por f. Portanto, esta desigualdade pode ser escrita como

$$n\mu(E) \geq (n - T) \int \overline{\tau}(E, x)\, d\mu(x) - (n - T)\varepsilon.$$

Dividindo os dois termos por n e fazendo n ir para infinito, concluímos que

$$\mu(E) \geq \int \overline{\tau}(E, x)\, d\mu(x) - \varepsilon$$

Como $\varepsilon > 0$ é qualquer, isto implica a primeira desigualdade em (3.4). Isto termina a demonstração neste caso.

Caso geral: Vamos indicar as modificações que devem ser feitas relativamente ao caso particular.

Dado $\varepsilon > 0$, começamos por fixar $T \geq 1$ suficientemente grande, de modo que a medida do conjunto

$$B = \{y \in M : t(y) > T\}$$

seja menor que ε. Em seguida, na definição das sequências substituímos a regra 2 por

2a. Se $t(x_i) \le T$, tomamos $t_i = t(x_i)$ e $x_{i+1} = f^{t_i}(x_i)$.

2b. Se $t(x_i) > T$, tomamos $t_i = 1$ e $x_{i+1} = f(x_i)$.

As regras 1 e 3 permanecem inalteradas. A estimativa referente a (3.6) continua válida, *para os valores de i aos quais se aplica a regra 2a:*

$$\sum_{j=0}^{t_i-1} \mathcal{X}_E(f^j(x_i)) \ge t_i\big(\overline{\tau}(E,x) - \varepsilon\big).$$

É claro que esta desigualdade implica a seguinte:

$$\sum_{j=0}^{t_i-1} \mathcal{X}_E(f^j(x_i)) \ge t_i\big(\overline{\tau}(E,x) - \varepsilon\big) - \sum_{j=0}^{t_i-1} \mathcal{X}_B(f^j(x_i)). \tag{3.8}$$

A vantagem é que (3.8) *é válida também para os valores de i aos quais se aplica a regra 2b.* De fato, nesse caso tem-se $t_i = 1$, o membro da esquerda é maior ou igual que zero e o membro da direita é menor que zero, uma vez que $\overline{\tau}(E,x)$ é sempre menor ou igual que 1. Isso significa que, no lugar de (3.7), tem-se

$$\sum_{j=0}^{n-1} \mathcal{X}_E(f^j(x)) \ge (n-T)(\overline{\tau}(E,x) - \varepsilon) - \sum_{j=0}^{n-1} \mathcal{X}_B(f^j(x)).$$

Integrando, como fizemos anteriormente, obtemos

$$n\mu(E) \ge (n-T)\int \overline{\tau}(E,x)\, d\mu(x) - (n-T)\varepsilon - n\mu(B).$$

Dividindo por n e fazendo $n \to \infty$, deduzimos que (lembre que $\mu(B) < \varepsilon$)

$$\mu(E) \ge \int \overline{\tau}(E,x)\, d\mu(x) - \varepsilon - \mu(B) \ge \int \overline{\tau}(E,x)\, d\mu(x) - 2\varepsilon.$$

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

$$\mu(E) \ge \int \overline{\tau}(E,x)\, d\mu(x).$$

Isto completa a demonstração do Teorema 3.1.

3.3 Exercícios

3.1. Considere a transformação $f : M \to M$, $f(x) = 10x - [10x]$ introduzida na Seção 2.3.1. Considere

$$x = 0,3355333355555555553333333333333335\ldots.$$

Ou seja: a expansão decimal de x consiste de blocos de 3s e 5s, alternados, cada bloco (exceto o segundo) com duas vezes mais dígitos que o anterior. Considere também $E = [0,3,\ 0,4)$. Mostre que

$$\tau_2(E,x) = 1, \quad \tau_8 = \frac{3}{4}, \quad \ldots \quad \tau_{2^{2k-1}}(E,x) \to \frac{2}{3}\,,$$

enquanto que

$$\tau_4(E,x) = \frac{1}{2}, \quad \tau_{16} = \frac{3}{8}, \quad \ldots \quad \tau_{2^{2k}}(E,x) \to \frac{1}{3}\,,$$

e portanto o tempo médio de permanência da órbita de x em E não existe.

3.2. Mostre que, para qualquer função integrável φ, a média temporal $\tilde{\varphi}$ satisfaz $\tilde{\varphi} \circ f = \tilde{\varphi}$ em μ-quase todo ponto.

3.3. Sejam $f : M \to M$ uma transformação mensurável e μ uma probabilidade f-invariante. Mostre que, se A é um conjunto f-invariante então $\int_A \phi\, d\mu = \int_A \widetilde{\phi}\, d\mu$ para toda ϕ em $L_1(\mu)$.

3.4. Sejam $f : M \to M$ uma transformação mensurável, μ uma probabilidade f-invariante e ϕ uma função mensurável.

1. Prove que se ϕ pertence a L_p então sua média de Birkhoff $\widetilde{\phi}$ também está em L_p.

2. Prove que se ϕ está em L_p então $\frac{1}{n}\sum_{j=0}^{n-1} \phi \circ f^j$ converge a $\widetilde{\phi}$ em L_p. *Dica*: Mostre primeiro para ϕ em L_∞, depois usando que as funções de L_∞ são densas em L_p conclua.

3.5. Sejam $f : M \to M$ uma transformação mensurável e μ uma probabilidade f-invariante. Prove que se ϕ é uma função em $L_2(\mu)$ então $\int (\widetilde{\phi} - \phi)\widetilde{\phi}\, d\mu = 0$.

Capítulo 4

Ergodicidade

4.1 Definição

Uma transformação $f : M \to M$ diz-se *ergódica* para uma probabilidade invariante μ (também dizemos que a medida μ é ergódica para f, ou que o sistema (f, μ) é ergódico) se as médias temporais dadas pelo Teorema de Birkhoff 3.2 coincidem em quase todo ponto com as respectivas médias espaciais:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \int \varphi \, d\mu \,,$$

para toda função μ-integrável $\varphi : M \to \mathbb{R}$ e μ-quase todo $x \in M$.

Na próxima proposição vamos reescrever esta condição de várias maneiras equivalentes, para ajudar a entender o seu significado. Um conjunto mensurável $A \subset M$ diz-se *invariante* se $f^{-1}(A) = A$. Uma função mensurável $\psi : M \to \mathbb{R}$ diz-se *invariante* se $\psi \circ f = \psi$.

Proposição 4.1. *Seja μ uma probabilidade invariante de uma transformação $f : M \to M$ mensurável. As seguintes condições são equivalentes:*

1. *O sistema (f, μ) é ergódico.*
2. *Para todo subconjunto invariante A tem-se $\mu(A) = 0$ ou $\mu(A) = 1$.*
3. *Toda função invariante ψ é constante num conjunto de medida total.*

Demonstração. **(1) implica (2):** Considere $\varphi = \mathcal{X}_A$. Por um lado, a hipótese (1) significa que

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

para quase todo $x \in M$. Por outro lado, como A é invariante, temos que $x \in A$ se e somente se $f(x) \in A$. Isto implica que $\varphi(f^j(x)) = \varphi(x)$ para todo $j \geq 0$ e para todo x. Portanto,

$$\tilde{\varphi}(x) = \varphi(x) = \mathcal{X}_A(x)$$

para todo $x \in M$. Como a função característica só toma os valores 0 e 1, estas duas igualdades implicam que $\mu(A) = 0$ ou $\mu(A) = 1$, como é afirmado em (2).

(2) implica (3): Seja ψ uma função invariante qualquer. Então, a pré-imagem $\psi^{-1}(I)$ de qualquer intervalo $I \subset \mathbb{R}$ é um conjunto invariante. Portanto, pela hipótese (2), essa pré-imagem tem medida zero ou um. Como o intervalo I é qualquer, isto prova que ψ é constante num conjunto com probabilidade μ total.

(3) implica (1): Seja φ uma função integrável qualquer. Como vimos no exercício 3.2, a média temporal $\tilde{\varphi}$ é uma função invariante. Logo, pela hipótese (3), $\tilde{\varphi}$ é constante em quase todo ponto. Então, usando o teorema ergódico,

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

em quase todo ponto. Isto é, o sistema é ergódico. $\square$

4.2 Exemplos

Nesta seção descrevemos diversos exemplos de sistemas ergódicos e interpretaremos o que esta propriedade significa em termos da dinâmica das órbitas.

4.2.1 Expansão decimal

Considere a transformação $f : [0,1] \to [0,1]$, $f(x) = 10x - [10x]$ da seção 2.3.1. Afirmamos que f é ergódica para a medida de Lebesgue μ. Tendo em vista a proposição 4.1, para mostrar isto só temos que provar que se A é um conjunto invariante com medida positiva então A tem medida total.

Suponhamos então que A é invariante e $\mu(A) > 0$. O ingrediente principal é o teorema de derivação **??**. No nosso caso, como estamos tratando com subconjuntos de $\mathbb{R}$, a condição (**??**) torna-se

$$\lim_{\varepsilon \to 0} \inf \Big\{ \frac{\mu(I \cap A)}{\mu(I)} : I \subset (a - \varepsilon, a + \varepsilon) \text{ intervalo contendo } a \Big\} = 1\,. \tag{4.1}$$

Fixemos um ponto de densidade $a \in A$ qualquer. Consideremos a sequência de intervalos

$$I_k = \Big(\frac{m_k}{10^k}, \frac{m_k + 1}{10^k}\Big), \quad m_k \in \mathbb{Z}, k \in \mathbb{N},$$

que contém o ponto a. Como a é um ponto de densidade de A, a propriedade (4.1) implica que

$$\frac{\mu(I_k \cap A)}{\mu(I_k)} \to 1 \quad \text{quando } k \to \infty.$$

Observe também que cada f^k é uma bijeção *afim* de I_k sobre o intervalo $(0,1)$. Isso tem a seguinte conseqüência, que é crucial para o nosso argumento:

$$\frac{\mu(f^k(E_1))}{\mu(f^k(E_2))} = \frac{\mu(E_1)}{\mu(E_2)} \tag{4.2}$$

para quaisquer subconjuntos mensuráveis E_1 e E_2 de I_k. Aplicando este fato a $E_1 = I_k \cap A$ e $E_2 = I_k$ obtemos que

$$\frac{\mu\big(f^k(I_k \cap A)\big)}{\mu\big((0,1)\big)} = \frac{\mu(I_k \cap A)}{\mu(I_k)}\,.$$

Claro que $\mu\big((0,1)\big) = 1$. Além disso, como estamos supondo que A é invariante, $f^k(I_k \cap A)$ está contido em A. Deste modo obtemos que

$$\mu(A) \geq \frac{\mu(I_k \cap A)}{\mu(I_k)}\,.$$

Como a sequência do lado direito converge para 1, segue que $\mu(A) = 1$, como queríamos demonstrar. Ficou provado que a transformação f é ergódica para a medida de Lebesgue μ.

Em seguida vamos dar uma aplicação deste fato no contexto da Teoria dos Números. Dizemos que um número $x \in \mathbb{R}$ é *balanceado* se todo dígito aparece com a mesma freqüência, $1/10$, na sua expansão decimal. É fácil dar exemplos de números balanceados. Mas em geral é muito difícil decidir se um dado número irracional é balanceado ou não. Por exemplo, não é sabido até hoje se o número π é balanceado.

No entanto, a conclusão da seção anterior nos permite deduzir que quase todo número é balanceado:

Proposição 4.2. *O conjunto dos números $x \in \mathbb{R}$ não balanceados tem medida de Lebesgue nula.*

Demonstração. Como o fato de ser balanceado é independente da parte inteira do número, só precisamos mostrar que quase todo $x \in [0,1]$ é balanceado. Considere $f : [0,1] \to [0,1]$ definida por $f(x) = 10x - [10x]$. Para cada dígito $j \in \{0, 1, \ldots, 9\}$ considere o intervalo $E_j = [j/10, (j+1)/10)$. Recorde que se $x = 0, a_0 a_1 \cdots a_k a_{k+1} \cdots$ então $f^k(x) = 0, a_k a_{k+1} \cdots$. Portanto, $f^k(x) \in E_j$ se e somente se o k-ésimo dígito da expansão decimal de x é igual a j. Conseqüentemente, o tempo médio de permanência $\tau(E_j, x)$ é exatamente a freqüência do dígito j na expansão decimal de x. Usando o teorema ergódico e o fato de que a transformação é ergódica para a medida de Lebesgue μ, concluímos que para cada $j \in \{0, 1, \ldots, 9\}$ existe um subconjunto B_j de M com $\mu(B_j) = 1$ tal que

$$\tau(E_j, x) = \mu(E_j) = \frac{1}{10} \quad \text{para todo } x \in B_j.$$

Então $B = B_0 \cap B_1 \cap \cdots \cap B_9$ também tem $\mu(B) = 1$, e todo número $x \in B$ é balanceado. $\square$

4.2.2 Deslocamentos de Bernoulli

Os deslocamentos ("shifts") de Bernoulli modelam sequências de experimentos aleatórios em que o resultado de cada experimento é independente dos demais.

Supõe-se que em cada experimento há um número finito de resultados possíveis, designados por $1, 2, \dots, d$, com probabilidades $p(1), p(2), \dots, p(d)$ de ocorrerem, sendo

$$p(1) + p(2) + \cdots + p(d) = 1\,.$$

O conjunto M das sequências $\underline{\alpha} = (\alpha_n)_{n\in\mathbb{Z}}$ com cada $\alpha_n \in \{1, 2, \dots, d\}$ contém os possíveis resultados da sequência de experimentos. Chamam-se *cilindros* os subconjuntos da forma

$$[k, l; a_k\,, \dots, a_l] = \{\underline{\alpha} \in M : \alpha_k = a_k\,, \dots, \alpha_l = a_l\}$$

onde $k, l \in \mathbb{Z}$, com $k \le l$, e cada $a_j \in \{1, 2, \dots, d\}$. Definimos

$$\mu([k, l; a_k\,, \dots, a_l]) = p(a_k) \cdots p(a_l) \tag{4.3}$$

Heuristicamente, isto significa que a probabilidade do evento composto

$$\alpha_k = a_k \quad \text{e} \quad \alpha_{k+1} = a_{k+1} \quad \text{e} \quad \cdots \quad \text{e} \quad \alpha_l = a_l$$

é o produto das probabilidades de cada um deles. Isto traduz, precisamente, que os resultados sucessivos são independentes entre si.

Consideramos em M a σ-álgebra $\mathcal{B}$ gerada pelos cilindros. A família $\mathcal{B}_0$ das uniões disjuntas finitas dos cilindros é uma álgebra (por convenção, M é um cilindro e $\mu(M) = 1$). Estendemos μ de modo a que seja finitamente aditiva: se $E \in \mathcal{B}_0$ é a união disjunta de cilindros $C_1\,, \dots, C_N$, definimos

$$\mu(E) = \mu(C_1) + \cdots + \mu(C_N).$$

Verifica-se que esta função μ é, de fato, σ-aditiva em $\mathcal{B}_0$; por exemplo, isso pode ser feito usando o Teorema 1.16. Portanto existe uma única probabilidade na σ-álgebra $\mathcal{B}$ gerada por $\mathcal{B}_0$ que é uma extensão de μ, isto é, que coincide com ela restrita a $\mathcal{B}_0$. Chamamos essa probabilidade *medida de Bernoulli* definida por $p(1), p(2), \dots, p(d)$ e, para não complicar desnecessariamente a notação, a representamos também por μ.

No espaço M consideramos a transformação deslocamento ("shift") à esquerda

$$f : M \to M \quad f\big((\alpha_n)_{n\in\mathbb{Z}}\big) = (\alpha_{n+1})_{n\in\mathbb{Z}}$$

que corresponde a fazer uma translação no tempo. Observe que a medida de Bernoulli é invariante por essa transformação. De fato, se $E = [k, l; a_k\,, \dots, a_l]$ então $f^{-1}(E) = [k+1, l+1; a_k\,, \dots, a_l]$ e a definição (4.3) dá que

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

neste caso. Como a família dos cilindros gera a σ-álgebra $\mathcal{B}$, isto juntamente com o Lema 2.6, prova que a medida μ é invariante para f.

Em seguida mostraremos que as medidas de Bernoulli são ergódicas. Para isso, a seguinte propriedade das medidas de Bernoulli vai ser útil :

Lema 4.3. *Se A e B são elementos da álgebra $\mathcal{B}_0$, isto é, uniões finitas de cilindros disjuntos, então tem-se*

$$\mu(A \cap f^{-m}(B)) = \mu(A)\mu(f^{-m}(B)) = \mu(A)\mu(B),$$

para todo m suficientemente grande.

Demonstração. Expliquemos porque esta propriedade é verdadeira quando A e B são cilindros, $A = [k, l; a_k, \ldots, a_l]$ e $B = [u, v; b_u, \ldots, b_v]$. Para cada m tem-se $f^{-m}(B) = [u+m, v+m; b_u, \ldots, b_v]$. Escolhendo m suficientemente grande garantimos que $u + m > l$ e, então,

$$\begin{aligned} A \cap f^{-m}(B) &= \{\underline{\alpha} : \alpha_k = a_k, \ldots, \alpha_l = a_l, \alpha_{u+m} = b_u, \ldots, \alpha_{v+m} = b_v\} \\ &= \bigcup [k, v+m; a_k, \ldots, a_l, c_{l+1}, \ldots, c_{u+m-1}, b_u, \ldots, b_v], \end{aligned}$$

onde a união é sobre todos os valores possíveis de $c_{l+1}, \ldots, c_{u+m-1}$. Usando (4.3), concluímos que $\mu(A \cap f^m(B)) = \mu(A)\mu(B)$. Isto prova o lema quando os conjuntos envolvidos são cilindros. O caso geral segue pelo fato de μ ser finitamente aditiva. □

Proposição 4.4. *Seja $f : M \to M$ um deslocamento e μ uma medida de Bernoulli em M, como antes. Então o sistema (f, μ) é ergódico.*

Demonstração. Seja A um conjunto mensurável invariante qualquer. Queremos mostrar que $\mu(A) = 0$ ou $\mu(A) = 1$. Para tornar a idéia da prova mais clara, comecemos por um caso particular: suponhamos que A está na álgebra $\mathcal{B}_0$ das uniões finitas de cilindros disjuntos dois-a-dois. Nesse caso podemos aplicar o lema anterior, com $B = A$. Concluímos que $\mu(A \cap f^{-m}(A)) = \mu(A)^2$ sempre que tomemos m suficientemente grande. Mas, como A é invariante, $f^{-m}(A) = A$ para todo m. Então a igualdade anterior quer dizer que $\mu(A) = \mu(A)^2$, o que só pode acontecer se $\mu(A) = 0$ ou $\mu(A) = 1$.

Agora vamos fazer a prova quando $A \in \mathcal{B}$ é um conjunto invariante qualquer. A idéia é aproximar A por elementos da álgebra $\mathcal{B}_0$, usando o Teorema de Aproximação 1.17: dado qualquer $\varepsilon > 0$ existe $A_0 \in \mathcal{B}_0$ tal que $\mu(A \Delta A_0) < \varepsilon$. Escolha m como no caso anterior, de modo que

$$\mu(A_0 \cap f^{-m}(A_0)) = \mu(A_0)\mu(f^{-m}(A_0)) = \mu(A_0)^2. \tag{4.4}$$

Observe que

$$\begin{aligned} (A \cap f^{-m}(A))\Delta(A_0 \cap f^{-m}(A_0)) &\subset (A\Delta A_0) \cup (f^{-m}(A)\Delta f^{-m}(A_0)) \\ &\subset (A\Delta A_0) \cup f^{-m}(A\Delta A_0). \end{aligned}$$

Isto, junto com o fato de que μ é invariante por f, implica que

$$\left|\mu(A \cap f^{-m}(A)) - \mu(A_0 \cap f^{-m}(A_0))\right| \le 2\mu(A\Delta A_0) < 2\varepsilon. \tag{4.5}$$

Além disso,

$$\left|\mu(A)^2 - \mu(A_0)^2\right| \le \left|(\mu(A) + \mu(A_0))(\mu(A) - \mu(A_0))\right| \le 2\left|\mu(A) - \mu(A_0)\right| < 2\varepsilon. \tag{4.6}$$

Juntando as relações (4.4), (4.5), (4.6), concluímos que $|\mu(A) - \mu(A)^2| < 4\varepsilon$. Como ε é arbitrário, deduzimos que $\mu(A) = \mu(A)^2$ e então, do mesmo modo que antes, concluímos que $\mu(A) = 0$ ou $\mu(A) = 1$. □

4.2.3 Rotações do círculo

Um exemplo bem natural de sistema dinâmico preservando uma medida são as *rotações*. Uma rotação de ângulo θ é uma função

$$R_\theta : S^1 \to S^1, \quad R_\theta(x) = x + \theta \mod \mathbb{Z}.$$

Observe que como a derivada de R_θ é identicamente igual a 1, temos que R_θ preserva a medida de Lebesgue de S^1. Aqui destacamos que a dinâmica de R_θ possui dois comportamentos bem distintos:

- Se θ é racional, digamos $\theta = p/q$ com $p, q \in \mathbb{Z}$, temos que $R_\theta^q(x) = x + q\theta = x + p = x \mod \mathbb{Z}$. Como consequência disso, todo ponto $x \in S^1$ é periódico de período no máximo q;
- Se θ é irracional, pode ser mostrado que $\mathcal{O}(x) = \{R_\theta^n; n \in \mathbb{N}\}$ é um conjunto denso em S^1. Vamos fazer isso na proposição a seguir:

Proposição 4.5. *Se θ é irracional, então $\mathcal{O}(x) = \{R_\theta^n; n \in \mathbb{N}\}$ é um conjunto denso em S^1.*

Demonstração. Afirmamos que o conjunto $\mathcal{D} = \{m + n\theta; m \in \mathbb{Z}, n \in \mathbb{N}\}$ é denso em $\mathbb{R}$. De fato, para todo natural N, podemos escolher p, q inteiros tais que

$$|q\theta - p| < \frac{1}{N}.$$

Como θ é irracional, temos que $a = |q\theta - p| > 0$. Sem perda de generalidade, assuma que $a = q\theta - p$ (o outro caso é análogo). Podemos subdividir a reta em intervalos de comprimento a e concluir que existe um inteiro M tal que $Ma \le r < (M+1)a$. Logo,

$$|r - Ma| < |r - M(q\theta - p)| = |r - \big((Mq)\theta + (-Mp)\big)| < \frac{1}{N}.$$

Tomando $n = Mq$ e $m = -Mp$, temos o que queríamos.

Agora para mostrarmos que $\mathcal{O}(x)$ é denso em S^1, observe que dado $y \in [0,1]$ e $\epsilon > 0$, podemos tomar $r = y - x$ e escolher m, n tais que $|m + n\theta - (y - x)| < \epsilon$. Isso equivale a dizer que $d(R_\theta^n(x), y) < \epsilon$ em S^1, como queríamos demonstrar. □

A proposição anterior terá várias implicações interessantes no estudo das medidas invariantes de R_θ. Entre outras coisas, veremos na Proposição **??** que se θ é irracional esta medida é a única medida que é preservada por R_θ. Além disso, veremos que as órbitas de R_θ se distribuem de modo uniforme em S^1.

O comportamento dinâmico e ergódico de R_α depende muito da natureza de α, como vamos ver. Dizemos que a rotação é *irracional* se o número $\alpha/(2\pi)$ é irracional, e dizemos que a rotação é racional no caso contrário.

A recíproca é muito mais interessante:

Proposição 4.6. *Se R_α é rotação irracional então R_α é ergódica para a medida de Lebesgue.*

Vamos mencionar duas demonstrações diferentes deste fato. A primeira, que detalharemos a seguir, usa fatos simples de análise de Fourier. A segunda, que deixaremos como exercício, é baseada num argumento de ponto de densidade semelhante ao que usamos no caso da expansão decimal.

Seja μ a medida de Lebesgue no círculo. Chama-se $L^2(\mu)$ o espaço das funções [1] mensuráveis $\psi : S^1 \to \mathbb{C}$ cujo quadrado é integrável:

$$\int |\psi|^2 \, d\mu < \infty.$$

É claro que este espaço contém todas as funções mensuráveis limitadas e, em particular, todas as funções características de conjuntos mensuráveis. Outro fato de que necessitamos é que a família de funções $\{\phi_k(z) = z^k : k \in \mathbb{Z}\}$ é uma base (de Hilbert) desse espaço: dada qualquer $\varphi \in L^2(\mu)$ existe uma única sequência $(c_k)_{k\in\mathbb{Z}}$ de números complexos tais que

$$\varphi(z) = \sum_{k\in\mathbb{Z}} c_k z^k \quad \text{para quase todo } z \in S^1.$$

Demonstração. Pela proposição 4.1, basta mostrar que a função característica de qualquer conjunto mensurável, digamos $\varphi = \mathcal{X}_A$, que é invariante é constante em μ-quase todo ponto. Note que $\varphi \in L^2(\mu)$, assim, usando a expansão de Fourier $\varphi(z) = \sum_{k\in\mathbb{Z}} c_k z^k$, a condição de ser invariante $\varphi \circ R_\alpha = \varphi$ se escreve

$$\sum_{k\in\mathbb{Z}} c_k e^{ki\alpha} z^k = \sum_{k\in\mathbb{Z}} c_k z^k$$

Por unicidade dos coeficientes da expansão em série de Fourier, obtemos que

$$c_k(e^{ki\alpha} - 1) = 0 \quad \text{para todo } k \in \mathbb{Z}.$$

A hipótese de que a rotação é irracional significa que $e^{ki\alpha} - 1 \neq 0$ para todo $k \neq 0$, e portanto, $c_k = 0$ para todo $k \neq 0$. Ou seja, $\varphi(z) = c_0$ para μ-quase todo $z \in S^1$, como queríamos provar. □

De fato as rotações irracionais satisfazem uma propriedade muito mais forte do que ergodicidade: elas são *unicamente ergódicas*, o que quer dizer que têm uma única probabilidade invariante (que é a medida de Lebesgue, claro).

[1] Quando lidamos com $L^2(\mu)$ sempre identificamos funções que diferem apenas num conjunto de medida nula.

Observação 4.7. A noção de rotação irracional se estende para dimensões maiores. Dado qualquer $d \geq 1$ chamamos *d-toro* o produto $\mathbb{T}^d = S^1 \times \cdots \times S^1$ do círculo por si mesmo d vezes. A rotação de ângulo $\alpha = (\alpha_1, \ldots, \alpha_d)$ é a aplicação $R_\alpha : \mathbb{T}^d \to \mathbb{T}^d$, $R_\alpha(z_1, \ldots, z_d) = (e^{i\alpha_1} z_1, \ldots, e^{i\alpha_d} z_d)$. A rotação é irracional se os números $\alpha_j/(2\pi)$ são incomensuráveis:

$$m_0 + m_1 \frac{\alpha_1}{2\pi} + \cdots + m_d \frac{\alpha_d}{2\pi} = 0 \quad \Rightarrow \quad m_0 = m_1 = \cdots = m_d = 0,$$

quaisquer que sejam os inteiros $m_0, m_1, \ldots, m_d$. Usando uma versão multidimensional das idéias anteriores, se prova que uma rotação é ergódica se e somente se ela é irracional.

4.2.4 Transformação de Gauss

Como vimos na Seção 2.3.2, a transformação de Gauss $G(x) = 1/x - [1/x]$ admite uma probabilidade invariante que é equivalente à medida de Lebesgue, nomeadamente,

$$\mu(E) = \frac{1}{\log 2} \int_E \frac{dx}{1+x}$$

Temos também que o sistema (G, μ) é ergódico. Este fato pode ser demonstrado pelo mesmo tipo de argumento que usamos na seção 4.2.1. Vamos esboçar o argumento neste caso, explicando qual é a principal dificuldade adicional.

Seja A um conjunto invariante com medida positiva. Em primeiro lugar, continua sendo verdade que para quase todo ponto $a \in [0,1]$ existe uma sequência de intervalos I_k contendo a e tais que f^k envia I_k bijetivamente e diferenciavelmente sobre $(0,1)$. O diâmetro desses intervalos converge para zero. Logo, tomando para a um ponto de densidade qualquer de A, temos que

$$\frac{\mu(I_k \cap A)}{\mu(I_k)} \to 1 \quad \text{quando } k \to +\infty. \tag{4.7}$$

Por outro lado embora f^k seja uma bijeção restrita a cada I_k, ela não é afim. Por essa razão não temos o análogo da relação (4.2) neste caso. Esta dificuldade é contornada através do seguinte resultado, que é um exemplo de controle de *distorção*: é muito importante notar que a constante K é independente de k, I_k, E_1, e E_2.

Lema 4.8. *Existe uma constante $K > 1$ tal que para todo $k \geq 1$, todo intervalo I_k tal que G restrita a I_k é uma bijeção diferenciável, tem-se*

$$\frac{\mu(f^k(E_1))}{\mu(f^k(E_2))} \leq K \frac{\mu(E_1)}{\mu(E_2)}$$

para quaisquer subconjuntos mensuráveis E_1 e E_2 de I_k.

Antes de demonstrarmos o Lema 4.8, explicamos como a ergodicidade de (G, μ) pode ser obtida a partir dele. Observe que $f^k(I_k \cap A^c) = A^c$, porque o conjunto A é invariante. Lembre também que $f^k(I_k) = (0,1)$, que tem medida total. Tomando $E_1 = I_k \cap A^c$ e $E_2 = I_k$ no lema 4.8, concluímos que

$$\mu(A^c) \leq \frac{\mu(f^k(I_k \cap A^c)}{\mu(A^c)} \leq K\frac{\mu(I_k \cap A^c)}{\mu(I_k} .$$

De acordo com (4.7), a expressão do lado direito converge para zero quando $k \to \infty$. Logo $\mu(A^c) = 0$, como queríamos demonstrar.

Daremos agora a prova do Lema 4.8.

Prova do Lema 4.8. Usaremos os seguintes fatos sobre a transformação f que podem ser facilmente verificados pelo leitor:

1. Para todo $x \in (0,1)$ vale que $|f'(x)| > 1$ e $|(f^2)'(x)| \geq 4$.
2. Existe $C_1 > 0$ tal que $|\frac{f''(x)}{f'(x)}| < C_1$.

Observe que a partir do item (1) acima, podemos mostrar que se $x, y \in I_k$ então

$$|f^i(x) - f^i(y)| \leq \frac{1}{2^{k-i}}|f^k(x) - f^k(y)| \text{ se } i = 0,1,2,\ldots k. \tag{4.8}$$

Observe também que se $x, y \in I_k$ temos que

$$|\log\frac{(f^k)'(x)}{(f^k)'(y)}| \leq \sum_{i=0}^{k-1} |\log f'(f^i(x)) - \log f'(f^i(y))|.$$

O item (2) nos garante que a função $x \to \log f'(x)$ tem derivada limitada por C, logo pelo Teorema do Valor Médio temos que $|\log f'(a) - \log f'(b)| \leq C_1|a-b|$. Aplicando este fato na desigualdade acima e observando a equação 4.8:

$$|\log\frac{(f^k)'(x)}{(f^k)'(y)}| \leq C_1\sum_{i=0}^{k-1}|f^i(x) - f^i(y)| \leq \sum_{i=0}^{k-1}\frac{1}{2^{k-i}}C_1|f^k(x) - f^k(y)| \leq C_2,$$

onde C_2 é uma constante propriamente escolhida. Logo, tomando $K = \exp C_2$, vem que para todos $x, y \in I_k$ vale:

$$\frac{(f^k)'(x)}{(f^k)'(y)} < C_3.$$

Note que a constante C_3 escolhida não depende de k nem de I_k. Observe ainda que se $A \subset [0,1]$ é um conjunto mensurável, então

$$\frac{1}{2\log 2}m(A) \leq \mu(A) \leq \frac{1}{\log 2}m(A),$$

onde m representa a medida de Lebesgue de [0,1]. Assim, para concluir a prova do Lema 4.8, basta observar que se E_1 e E_2 são subconjuntos mensuráveis de

I_k, então:

$$\frac{\mu(f^k(E_1))}{\mu(f^k(E_2))} = 2(\log 2)^2 \frac{m(f^k(E_1))}{m(f^k(E_2))} \leq \frac{\int_{E_1}(f^k)'(x)\,dm}{\int_{E_2}(f^k)'(y)\,dm} \leq$$
$$2(\log 2)^2(C_3)^2\frac{m(E_1)}{m(E_2)} \leq 4(\log 2)^4 C_3 \frac{\mu(E_1)}{\mu(E_2)}.$$

Assim, basta tomar $K = 4(\log 2)^4(C_3)^2$ e o lema está provado. □

4.3 Exercícios

4.1. A rotação R_α é racional se e somente se $e^{\alpha i}$ é uma raiz da unidade, isto é, se existe $k \neq 0$ tal que $e^{ki\alpha} = 1$.

4.2. Se R_α é rotação racional então R_α não é ergódica para a medida de Lebesgue.

No exercício a seguir propomos outra demonstração para a proposição 4.6:

4.3. Suponha que R_α é uma rotação irracional.

1. Mostre que a órbita $\{R_\alpha^n(z) : n \in \mathbb{Z}\}$ de todo $z \in S^1$ é densa em S^1.
2. Seja A um conjunto invariante com medida positiva. Mostre que *nenhum* ponto de S^1 é ponto de densidade de A^c. Conclua que $\mu(A) = 1$.

   Dica: considere um ponto de densidade de A e use o item (1).

4.4. Suponha que R_α é uma rotação irracional.

1. Seja $\varphi : S^1 \to \mathbb{R}$ uma função contínua qualquer. Mostre que

   $$\tilde{\varphi}(x) = \lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(R_\alpha^j(z))$$

   existe em *todo* ponto e, de fato, o limite é uniforme. Justifique que $\tilde{\varphi}$ é constante em todo ponto.

   Dica: Verifique que a sequência do lado direito é equicontínua e use o teorema de Ascoli-Arzelá.
2. Deduza que R_α tem uma única probabilidade invariante.

4.5. Mostre que o deslocamento σ definido na Secção 4.2.2 é transitivo e que o conjunto de suas órbitas periódicas é denso.

4.6 (Teorema de Kac)**.** Seja μ uma medida ergódica para uma transformação $f : M \to M$ e A um conjunto com $\mu(A) > 0$. Considere $n_A : A \to \mathbb{N} \cup \{+\infty\}$ como o menor número $n_A(x) > 0$ tal que $f^{n_A(x)}(x) \in A$. Caso este número não exista, definimos $n_A(x) = +\infty$.

1. Mostre que n_A é integrável com respeito a μ.
2. Mostre que se $\mu_A(B) = \frac{\mu(A\cap B)}{\mu(A)}$ então:

$$\int_A n_A(x)\, d\mu_A(x) = \frac{1}{\mu(A)}.$$

4.7. Seja $f : M \to M$ definida no espaço topológico M tal que existe alguma medida ergódica μ tal que para todo A aberto, $\mu(A) > 0$. Mostre que f e transitiva e a órbita de μ-quase todo ponto é densa.

Bibliografia

[Cas04] A. A. Castro. *Teoria da medida*. Projeto Euclides. IMPA, 2004.

[Fer02] R. Fernandez. *Introdução à teoria da medida*. Projeto Euclides. IMPA, 2002.

[Rud87] W. Rudin. *Real and complex analysis*. McGraw-Hill, 3 edition, 1987.