Artur Silva Santos

Teoria Ergódica de Transformações Expansoras

unesp

1910001294

Universidade Estadual Paulista

Câmpus de São José do Rio Preto
Instituto de Biociências, Letras e Ciências Exatas

Teoria Ergódica de Transformações Expansoras

Artur Silva Santos

Orientador: Prof. Dr. Vanderlei Minori Horita

Instituto de Biociências, Letras e Ciências Exatas - UNESP
BIBLIOTECA
S. J. Rio Preto

t. 1.294/04

Dissertação apresentada ao Departamento de Matemática - IBILCE - UNESP, como parte dos requisitos para a obtenção do título de Mestre em Matemática

São José do Rio Preto - SP
Junho - 2004

517.93
S237t

10103040

21/10/2004
Doação
R$ 23,00
Autor
ND: 110/04 –
Proc. 047/04

Santos, Artur Silva.
Teoria ergódica de transformações expansoras/Artur Silva Santos - São José do Rio Preto: [s.n.], 2004.
96 f.: il.; 30cm.

Orientador: Vanderlei Minori Horita.
Dissertação (mestrado) - Universidade Estadual Paulista. Instituto de Biociências, Letras e Ciências Exatas.

1. Sistemas dinâmicos. 2. Teoria ergódica. 3. Transformações expansoras diferenciais. 4. Transformações expansoras diferenciais por partes. I. Horita, Vanderlei Minori. II. Universidade Estadual Paulista. Instituto de Biociências, Letras e Ciências Exatas. III. Título.
CDU - 517.93

Teoria Ergódica de Transformações Expansoras

Artur Silva Santos

Dissertação apresentada ao Departamento de Matemática - IBILCE - UNESP, como parte dos requisitos para a obtenção do título de Mestre em Matemática

Aprovado em 05 de junho de 2004.

BANCA EXAMINADORA

Vanderlei Minori Horita
Professor Doutor - IBILCE - UNESP
Orientador

Ali Tahzibi
Professor Doutor - ICMC - USP
1º Examinador

Cláudio Aguinaldo Buzzi
Professor Doutor - IBILCE - UNESP
2º Examinador



Dedico aos amores da minha vida, minhas filhas, Cynthya e Ana Carolina.

Resumo

Neste trabalho estudamos transformações que preservam medida e funções de variação limitada como ferramentas essenciais para o estudo de existência de medidas invariantes absolutamente contínuas para transformações expansoras em variedades compactas e conexas, bem como transformações expansoras por partes.

Palavras Chave: Teoria ergódica, Teorema de recorrência de Poincaré, Teorema ergódico de Birkhoff, medidas invariantes absolutamente contínuas, transformações expansoras e transformações expansoras por partes.

cm 1 2 3 4 5 6 7 unesp 1 12 13 14 15 16 17 18

Abstract

In this work we study transformations preserving measure and bounded variation functions and serve as essential tools for the study of the existence of absolutely continuous invariant measures for expanding maps in compact and connected manifold as well as for piecewise expanding.

Key-words: *Ergodic theory, Poincare's recurrence theorem, Birkhoff ergodic theorem, absolutely continuous invariant measures, expanding maps and piecewise expanding maps.*

Agradecimentos

Ao término deste trabalho, agradeço de coração:

- Primeiramente a Deus por tudo.
- Ao Professor Doutor Vanderlei Minori Horita pela disponibilidade na orientação e formação, principalmente, pela confiança, paciência, amizade e aprendizado.
- Ao amigo da UFMA, Professor Doutor Nivaldo Costa Muniz pelos conselhos e disponibilidade na orientação, principalmente, pela paciência e amizade.
- Ao Professor Doutor José Carmelo Interlando por ter aberto as portas deste Instituto para que eu podesse realizar mais um sonho.
- À minha família por tudo o que sou e tenho.
- Aos meus professores: Professora Doutora Aparecida Francisca, Professor Doutor Cláudio Aguinaldo Buzzi, Clotilzio Moreira dos Santos, Professora Doutora Maria Gorete Carreira Andrade, Professora Doutor Sérgio Henrique Minori Soares e Professor Doutor Waldemar Donizete Bastos pelos ensinamentos e pelo convívio pacífico e amigável.
- Aos demais professores do Departamento de Matemática deste Instituto.
- À Professora Solange dos Santos Lima pela tradução do resumo.
- Aos amigos de república pela compreensão, respeito e amizade.
- Aos amigo(a)s que sempre torceram para o meu sucesso.
- Aos colegas de Pós-Graduação pelo convívio e companheirismo.
- A todos que de alguma maneira contribuíram direta ou indiretamente para que este sonho fosse realizado.



A cura da AIDS está na Matemática. Em particular, na Teoria Ergódica.

Artur

Sumário

Introdução **2**

1 Transformações que preservam medida **5**
1.1 Teorema de recorrência de Poincaré 5
1.2 Exemplos de medidas invariantes 8
1.2.1 Expansão decimal 8
1.2.2 Difeomorfismos 12
1.2.3 Transformação de Gauss 14
1.3 Existência de medidas invariantes 17
1.4 Teorema ergódico de Birkhoff 28
1.4.1 Comentários 28
1.4.2 Enunciados 29
1.5 Ergodicidade 36
1.6 Exemplos e aplicações 39
1.6.1 Expansão decimal 39
1.6.2 Deslocamentos (shifts) de Bernoulli 41
1.6.3 Rotação irracional no círculo 44
1.6.4 Transformação de Gauss 46
1.7 Propriedades de medidas ergódicas 47

2 Transformações expansoras **51**
2.1 Lema de distorção 53
2.2 Medidas absolutamente contínuas 55
2.3 Existência de medidas ergódicas 56
2.4 Unicidade e conclusão da prova 57

3 Transformações expansoras por partes **59**
3.1 Funções de variação limitada 60
3.2 Medidas invariantes absolutamente contínuas 70

A Elementos de teoria da medida e integração **76**
A.1 Espaços mensuráveis 76
A.2 Espaços de medida 77
A.3 Extensão 78

A.4 Medida de Lebesgue 79
A.5 Funções integráveis 80
A.6 Teorema de Radon-Nikodym 83
A.7 Teorema de decomposição ergódico 84
A.8 Teorema de derivação de Lebesgue 85
A.9 Espaços L^p 85

Referências Bibliográficas **87**

Introdução

A Teoria de Sistemas Dinâmicos teve grande impulso nas últimas décadas e constitui-se em uma das principais áreas da matemática. Em termos gerais, esta teoria está interessada em descrever, para a maioria dos sistemas, como a maioria das órbitas se comportam, especialmente quando o tempo vai para o infinito (a descrição de **todas** as órbitas ou sistemas não é, em geral, um objetivo realístico, pois existem muitas formas de comportamentos excepcionais). Mais ainda, quando e em que sentido este comportamento é robusto sob pequenas pertubações do sistema (a formulação matemática da lei de evolução é quase sempre uma simplificação do processo real. Por isso, é muito importante que as conclusões obtidas dela não sejam muito específicas: elas devem permanecer válidas para leis próximas). Um panorama atual da teoria pode ser encontrado em [16], [17] e [18] e em suas referências.

A formulação matemática de um sistema dinâmico inclui dois ingredientes principais: um conjunto M (geralmente uma variedade diferenciável de dimensão $d \geq 1$), o **espaço de estados**, cujos pontos representam os possíveis estados do sistema; e uma **lei de evolução**, descrevendo como o sistema evolui de um estado a outro. Esta lei de evolução pode ser **tempo discreto**: uma transformação $f : \mathrm{M} \longrightarrow \mathrm{M}$ mapeando cada estado $x_0 \in \mathrm{M}$ a outro estado $x_1 = f(x_0)$ que o sistema estará uma unidade de tempo depois. A seqüência $x_0, x_1, \ldots$ definida por $x_{j+1} = f(x_j)$ é a **trajetória** de um **estado inicial** x_0. Se a transformação f for inversível então temos também a trajetória passada $\ldots, x_{-2}, x_{-1}, x_0$ definida por $x_{-j+1} = f(x_{-j})$. Outro modelo de evolução é o **tempo contínuo**, expresso por uma equação diferencial

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

A trajetória de um estado inicial x_0 é a solução $x(t)$, $t \in \mathbb{R}$ da equação diferencial com $x(0) = x_0$.

Dado um sistema físico $f : \mathrm{M} \longrightarrow \mathrm{M}$, por vezes nos utilizamos de uma **função observável** $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$ para obter informações deste sistema. Em geral, a função $\varphi(f^j(z))$ comporta-se ao longo do tempo de um modo **complicado e errante**, mesmo para leis de evolução simples: escolhidos dois estados iniciais próximos z e z' os valores de $\varphi(f^j(z))$ e $\varphi(f^j(z'))$, por continuidade, estão próximos para valores pequenos de j, mas depois de alguns iterados estas seqüências aparentemente não tem mais correlação.

Este comportamento **caótico** significa que a dinâmica pode ser difícil de

unesp

ser entendida do ponto de vista determinístico, e que a análise probabilística pode ser mais frutífera. Esta abordagem é feita dentro de uma nova área, a **Teoria Ergódica**.

Em termos simples, a Teoria Ergódica é a disciplina matemática que estuda sistemas dinâmicos munidos de medidas invariantes. O bjetivo da teoria é estudar as propriedades estatísticas de **sistemas determinísticos**, isto é, sistemas com lei de evolução. A idéia básica é abandonar o estudo determinístico de um estado em favor de estudo estatístico de um conjunto de estados. Em vez de procurarmos respostas para a pergunta *qual será o estado do sistema no tempo t?*, procuramos resposta de *qual é a probabilidade que no tempo t o sistema pertença a um subconjunto do espaço de estados?* Associada à última questão está a noção de probabilidade.

Neste contexto procuramos respostas para:

1) Dado um subconjunto A $\subset$ M e $z \in$ A, a trajetória de z passa por A? Com que freqüência?

2) Dada um função $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$ e $z \in \mathrm{M}$

$$\frac{1}{n}\sum_{j=0}^{n-1}\varphi\big(f^j(z)\big) \quad \text{ou} \quad \frac{1}{T}\int_0^T \varphi\big(f^t(z)\big)\, dt$$

converge quando n ou $T \to \infty$?

Nesta direção estudaremos os Teorema de Recorrência de Poincaré e o Teorema Ergódico de Birkhoff que dão respostas positivas para pontos z **típicos** e para **transformações que preservam probabilidade** (veja Capítulo 1). Estas transformações são os objetos de estudo da teoria ergódica. Os enunciados precisos destes teoremas serão dados no Capítulo 1.

O objetivo central deste trabalho é estudar a teoria ergódica de transformações expansoras. No Capítulo 2, estudaremos as transformações expansoras diferenciáveis e no Capítulo 3, faremos o estudo de transformações expansoras diferenciáveis por partes.

Capítulo 1

Transformações que preservam medida

As medidas invariantes por uma transformação assumem um papel de primordial importância na Teoria Ergódica. Estaremos particularmente interessados no caso em que $f : \mathrm{M} \longrightarrow \mathrm{M}$ é uma transformação mensurável do espaço $(\mathrm{M}, \mathcal{A}, \mu)$ de medida em si mesmo, isto é, $f^{-1}(\mathrm{A}) \in \mathcal{A}$ para todo $\mathrm{A} \in \mathcal{A}$. Dizemos que f é **invertível** se é uma bijeção e a sua inversa é também uma transformação mensurável. Sempre consideraremos μ definida na σ-álgebra de Borel do espaço M. Na maior parte dos casos trataremos com medidas finitas. Nesse caso, sempre podemos transformar μ numa probabilidade ν: para isso basta definir

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

Em geral, uma medida μ diz-se invariante pela transformação mensurável $f : \mathrm{M} \longrightarrow \mathrm{M}$ se

$$\mu(\mathrm{A}) = \mu(f^{-1}(\mathrm{A})) \quad \text{para cada conjunto mensurável } \mathrm{A} \subset \mathrm{M}.$$

Heuristicamente, isto significa que a probabilidade de um ponto estar num dado conjunto e a probabilidade de que a sua pré-imagem esteja nesse conjunto são iguais. Note que a definição faz sentido, uma vez que a pré-imagem de um conjunto mensurável por uma transformação mensurável ainda é um conjunto mensurável. Dizemos que um conjunto $\mathrm{A} \subset \mathrm{M}$ mensurável qualquer tem **medida total** se $\mu(\mathrm{A}^c) = 0$ para toda medida μ invariante.

1.1 Teorema de recorrência de Poincaré

A existência de uma probabilidade invariante μ para um dado sistema dinâmico fornece muita informação acerca do comportamento desse sistema. Um dos aspectos mais simples é a **recorrência** dos iterados do sistema. **Poincaré** mostrou que escolhendo um estado inicial ao acaso, com

probabilidade 1, isto é, quase certamente os iterados obtidos pela iteração da transformação partindo desse estado inicial devem passar uma infinidade de vezes tão perto quanto quisermos do mesmo estado inicial, se esperarmos tempo suficiente. Porém tal resultado nada afirma sobre a forma do conjunto de iterados ou a sua posição, apenas informa sobre a passagem dos iterados perto de um certo ponto **em períodos de tempo muito grandes** para um conjunto **muito grande** de estados iniciais. Claro que podemos escolher um estado inicial que não verifique esta propriedade, mas **Poincaré** mostrou que a probabilidade de tal acontecer é zero. Hoje este resultado é conhecido como **Teorema de Recorrência de Poincaré**, publicado em 1899. Enunciaremos duas versões, a primeira numa linguagem mensurável e a segunda de natureza mais topológica.

O teorema seguinte, em **versão mensurável**, afirma que **quase todo ponto** do conjunto A **regressa** a A no **futuro**.

Teorema 1.1. Sejam $f : \mathrm{M} \longrightarrow \mathrm{M}$ uma transformação mensurável e μ uma medida invariante finita. Seja $\mathrm{A} \subset \mathrm{M}$ qualquer conjunto mensurável com $\mu(\mathrm{A}) > 0$. Então, μ-quase todo ponto $x \in \mathrm{A}$ tem algum iterado $f^n(x)$, $n \geq 1$, que também está em A.

Demonstração: Consideremos o conjunto

$$\mathrm{B} = \{x \in \mathrm{A}; \quad f^n(x) \notin \mathrm{A} \ \text{ para todo } \ n \geq 1\}$$

dos pontos $x \in \mathrm{A}$ que nunca regressam a A. Note que B é mensurável, pois $\mathrm{B} = \mathrm{A} \setminus \bigcup_{j=1}^{n} f^{-j}(\mathrm{A})$. Devemos provar que B tem medida nula.

Afirmação: As pré-imagens de B são disjuntas duas-a-duas.

Prova: Suponhamos que existem $m > n \geq 1$ tais que $f^{-m}(\mathrm{B})$ intersecta $f^{-n}(\mathrm{B})$. Seja x um ponto na interseção e denotemos $y = f^n(x)$. Então, $x \in f^{-m}(\mathrm{B}) \cap f^{-n}(\mathrm{B}) \iff f^n(x) \in f^{n-m}(\mathrm{B}) \cap \mathrm{B} \iff f^n(x) \in f^{n-m}(\mathrm{B})$ e $f^n(x) \in \mathrm{B} \iff y \in f^{n-m}(\mathrm{B})$ e $y \in \mathrm{B}$. Como $\mathrm{B} \subset \mathrm{A}$, então y regressa pelo menos uma vez ao conjunto A, o que contradiz a definição de B. Portanto, esta contradição prova que as pré-imagens de B são disjuntas duas-a-duas. Sendo B mensurável, então

$$\mu\left(\bigcup_{n=1}^{\infty} f^{-n}(\mathrm{B})\right) = \sum_{n=1}^{\infty} \mu(f^{-n}(\mathrm{B})) = \sum_{n=1}^{\infty} \mu(\mathrm{B}),$$

onde a última igualdade é devido ao fato de μ ser invariante por f. Como supomos que a medida é finita, a expressão $\mu\left(\bigcup_{n=1}^{\infty} f^{-n}(\mathrm{B})\right)$ é finita. Por outro lado, temos que $\sum_{n=1}^{\infty} \mu(\mathrm{B})$ é uma soma infinita de termos todos iguais, então a única maneira desta soma ser finita é que seus termos sejam nulos. Portanto, devemos ter $\mu(\mathrm{B}) = 0$. ∎



Observação 1.2. O Teorema 1.1 implica outro teorema aparentemente mais forte, o qual afirma que **quase todo ponto** do cojunto A regressa a A **infinitas vezes**.

Corolário 1.3. Nas condições do Teorema 1.1, para μ - quase todo ponto $x \in A$, existem infinitos valores de $n \geq 1$ tais que $f^n(x)$ está em A.
Demonstração: Para cada $k \geq 0$, definimos o conjunto

$$A_k = \{x \in A;\ \text{existe exatamente } k \text{ valores de } n \geq 1 \text{ tais que } f^n(x) \in A\}$$

dos pontos $x \in A$ que regressam a A exatamente k vezes. Então, $\bigcup_{k=0}^{\infty} A_k$ é o conjunto dos pontos $x \in A$ que regressam a A apenas um número finito de vezes. Portanto, para provar o corolário, basta mostrar que $\mu(A_k) = 0$ para todo $k \geq 0$. A demonstração será por contradição.

Suponhamos que $\mu(A_k) > 0$ para algum $k \geq 0$. Então, aplicando o Teorema 1.1 com A_k no lugar de B, obtemos que quase todo ponto $x \in A_k$ tem algum iterado $f^n(x) \in A, n \geq 1$, que está em A_k. Fixemos um tal x em A_k e denotemos $y = f^n(x)$. Por definição de A_k, y tem exatamente k iterados futuros que estão em A. Como y é um iterado de x implica que x tem $k+1$ iterados futuros que estão em A. Mas isso contradiz o fato de que $x \in A_k$. Logo, esta contradição prova que A_k tem medida nula, relativamente a μ. ■

Dizemos que um ponto $x \in M$ é **recorrente** para uma transformação $f : M \longrightarrow M$ se, para toda vizinhança V de x, existe algum iterado $f^n(x)$ que está em V. Uma **base enumerável de abertos de** $\mathcal{T}$ é uma família enumerável $\mathcal{B} = \{B_k;\ k \in \mathbb{N}\}$ de abertos tais que todo aberto de $\mathcal{T}$ pode ser escrito como união de elementos B_k dessa família.

O teorema a seguir, em **versão topológica**, afirma que **quase todo ponto** de M é **recorrente**, relativamente a qualquer medida invariante **finita** do sistema dinâmico.

Teorema 1.4. Suponhamos que M admite uma base enumerável de abertos. Sejam $f : M \longrightarrow M$ uma transformação mensurável e μ uma medida invariante finita. Então, μ - quase todo ponto $x \in M$ é recorrente para f.
Demonstração: Seja $\mathcal{B} = \{V_k;\ k \in \mathbb{N}\}$ uma base enumerável de abertos. Para cada $n \in \mathbb{N}$, definimos o conjunto

$$C_k = \{x \in V_k;\ f^n(x) \notin V_k \text{ para todo } n \geq 1\}$$

dos pontos $x \in V_k$ que nunca regressam a V_k. Então, pelo Teorema 1.1, todo C_k tem medida nula. Conseqüentemente, a união enumerável $C = \bigcup_{k \in \mathbb{N}} C_k$ tem medida nula. Portanto, para demonstrar o teorema será suficiente mostrar que todo ponto $x \notin C$ é recorrente.

Sejam $x \in M \setminus C$ e V uma vizinhança qualquer de x. A definição de $\mathcal{B}$ implica que existe algum $k \in \mathbb{N}$ tal que $x \in V_k$ e $V_k \subset V$. Como $x \notin C$,



também $x \notin C_k$, isso implica que x tem algum iterado $f^n(x) \in V_k$, para algum $n \geq 1$. Em particular, $f^n(x) \in V$, para algum $n \geq 1$, pois $V_k \subset V$. Como V é arbitrário, temos que x é recorrente. ∎

Observação 1.5. As conclusões dos Teoremas 1.1 e 1.4 não são verdadeiras, em geral, se omitirmos a hipótese de que μ é finita.

Exemplo 1.6. A função translação de 1 unidade $f : \mathbb{R} \longrightarrow \mathbb{R}$, isto é, $f(x) = x + 1$ para todo $x \in \mathbb{R}$, deixa invariante a medida de Lebesgue em $\mathbb{R}$ (que é infinita) e por outro lado, nenhum ponto é recorrente para f.

1.2 Exemplos de medidas invariantes

Nesta seção apresentamos alguns exemplos simples de medidas invariantes por transformações.

1.2.1 Expansão decimal

A transformação $f : [0,1] \longrightarrow [0,1]$ definida por

$$f(x) = 10x - [10x]$$

onde $[10x]$ representa o maior inteiro menor ou igual a $10x$. Isto é, f associa a cada $x \in [0,1]$ a parte fracionária de $10x$. Veja o gráfico de f descrito na Figura 1.1.

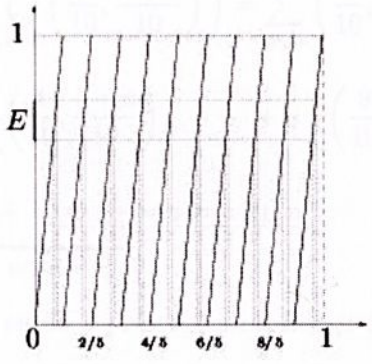


Figura 1.1: Transformação parte fracionária de $10x$

O gráfico de f foi esboçado conforme a seguir:

- Se $x \in \left[0, \frac{1}{10}\right)$ então $10x \in [0,1)$ e portanto a sua parte inteira $[10x]$ é igual a 0. Isto quer dizer que neste intervalo a transformação f é dada por $f(x) = 10x - 0$.

- Se $x \in \left[\frac{1}{10}, \frac{2}{10}\right)$ então $10x \in [1,2)$ e portanto a sua parte inteira $[10x]$ é igual a 1 e tem-se $f(x) = 10x - 1$.
- Em geral, se $x \in \left[\frac{k}{10}, \frac{k+1}{10}\right)$ para algum $0 \leq k \leq 9$ então $10x \in [k, k+1)$ e portanto a sua parte inteira $[10x]$ é igual a k. Assim, $f(x) = 10x - k$.

Afirmação: A restrição da medida de Lebesgue ao intervalo $[0,1]$ é invariante por f.
Prova: Seja o intervalo $(0,a) \subset [0,1]$ com $0 \leq a \leq 1$. Como f restrita a cada um dos intervalos $\left[\frac{k}{10}, \frac{k+1}{10}\right)$, $0 \leq k \leq 9$, é bijetora, então $f^{-1}(0)$ e $f^{-1}(a)$ são os únicos pontos em cada um dos intervalos e esses pontos são $\frac{k}{10}$ e $\frac{k+a}{10}$ respectivamente, pois

- $f\left(\frac{k}{10}\right) = k - [k] = k - k = 0 \Longrightarrow f^{-1}(0) = \bigcup_{k=0}^{9} \left\{\frac{k}{10}\right\}$.
- $f\left(\frac{k+a}{10}\right) = k + a - [k+a] = k + a - k = a \Longrightarrow f^{-1}(a) = \bigcup_{k=0}^{9} \left\{\frac{k+a}{10}\right\}$.

Então,

$$\begin{aligned}\mu\left(f^{-1}((0,a))\right) &= \mu\left(\bigcup_{k=0}^{9}\left(\frac{k}{10}, \frac{k+a}{10}\right)\right) = \sum_{k=0}^{9}\left(\frac{k}{10}, \frac{k+a}{10}\right) = \mu\left(\left(0, \frac{a}{10}\right)\right) \\ &+ \mu\left(\left(\frac{1}{10}, \frac{1+a}{10}\right)\right) + \cdots + \mu\left(\left(\frac{9}{10}, \frac{9+a}{10}\right)\right) \\ &= \underbrace{\frac{a}{10} + \cdots + \frac{a}{10}}_{10 \text{ vezes}} = a = \mu((0,a)).\end{aligned}$$

Agora, se $(a,b) \subset [0,1]$ então f preserva a medida μ desse intervalo, pois

$$\begin{aligned}\mu\left(f^{-1}((a,b))\right) &= \mu\left(f^{-1}((0,b) - (0,a))\right) = \mu\left(f^{-1}((0,b))\right) - \mu\left(f^{-1}((0,a))\right) \\ &= \mu((0,b)) - \mu((0,a)) = b - a = \mu((a,b)).\end{aligned}$$

Portanto, pelo lema a seguir, conclui-se a Afirmação.

Lema 1.7. Sejam $(\mathrm{M}, \mathcal{A}, \mu)$ um espaço de medida finita e $f : \mathrm{M} \longrightarrow \mathrm{M}$ uma transformação mensurável. Suponha que existe uma álgebra $\mathcal{C}$ que gera a σ-álgebra $\mathcal{A}$ tal que $\mu(\mathrm{A}) = \mu(f^{-1}(\mathrm{A}))$ para todo $\mathrm{A} \in \mathcal{C}$. Então, o mesmo vale para todo conjunto mensurável A, isto é, a medida μ é invariante por f.



Demonstração: Considere

$$\mathcal{B} = \left\{ A \in \mathcal{A};\ f^{-1}(A) \in \mathcal{A} \text{ e } \mu(A) = \mu(f^{-1}(A)) \right\}.$$

Devemos mostrar que $\mathcal{B} = \mathcal{A}$. Em primeiro lugar, por hipótese, a álgebra $\mathcal{C}$ está contida em $\mathcal{B}$. Como $\mathcal{A}$ é a menor σ-álgebra que contém $\mathcal{C}$, então basta mostrar que $\mathcal{B}$ é uma σ-álgebra. Para isto, basta mostrar que $\mathcal{B}$ é fechada por uniões enumeráveis e complemento.

Dado $B \in \mathcal{B}$ e como $f^{-1}(M) = M$, temos

$$\begin{aligned}\mu(f^{-1}(B^c)) &= \mu(f^{-1}(M \setminus B)) = \mu(f^{-1}(M) \setminus f^{-1}(B)) \\ &= \mu(f^{-1}(M)) - \mu(f^{-1}(B)) = \mu(M) - \mu(B) = \mu(M \setminus B) = \mu(B^c).\end{aligned}$$

Portanto, $B^c \in \mathcal{B}$.

Dados $A, B \in \mathcal{B}$, provemos que $A \setminus B \in \mathcal{B}$. Temos por definição de $\mathcal{B}$ que $A, B \in \mathcal{A}$. Como $\mathcal{A}$ é gerado por $\mathcal{C}$, basta mostrarmos para $A = \sum_{i=1}^{\infty} A_i$ e $B = \sum_{j=1}^{\infty} B_j$, onde $A_i, B_j \in \mathcal{C}$. Logo,

$$\begin{aligned}A \setminus B &= \left(\sum_{i=1}^{\infty} A_i\right) \setminus B = \sum_{i=1}^{\infty} A_i \setminus B = \sum_{i=1}^{\infty} \left(A_i \setminus \sum_{j=1}^{\infty} B_j\right) \\ &= \sum_{i=1}^{\infty} \left(\bigcap_{j=1}^{\infty} A_i \setminus B_j\right). \qquad (1.2.1)\end{aligned}$$

Assim,

$$\begin{aligned}\mu\left(f^{-1}(A \setminus B)\right) &= \mu\left(f^{-1}\left(\sum_{i=1}^{\infty}\left(\bigcap_{j=1}^{\infty} A_i \setminus B_j\right)\right)\right) \\ &= \mu\left(\sum_{i=1}^{\infty} f^{-1}\left(\bigcap_{j=1}^{\infty} A_i \setminus B_j\right)\right) = \sum_{i=1}^{\infty} \mu\left(f^{-1}\left(\bigcap_{j=1}^{\infty} A_i \setminus B_j\right)\right) \\ &= \sum_{i=1}^{\infty} \mu\left(\bigcap_{j=1}^{\infty} f^{-1}\left(A_i \setminus B_j\right)\right) = \sum_{i=1}^{\infty} \mu\left(\bigcup_{j=1}^{\infty} \left(f^{-1}\left(A_i \setminus B_j\right)\right)^c\right)^c \\ &= \sum_{i=1}^{\infty} \left(1 - \mu\left(\bigcup_{j=1}^{\infty} f^{-1}\left(A_i \setminus B_j\right)^c\right)\right).\end{aligned}$$

Seja $C_j = f^{-1}\left(A_i \setminus B_j\right)^c$, então

$$\bigcup_{j=1}^{\infty} C_j = C_1 + \sum_{k=2}^{\infty} \left(C_k \setminus \bigcup_{j=1}^{k-1} C_j\right).$$

Logo,

$$\begin{aligned}
\mu\left(\bigcup_j^{\infty} f^{-1}(A_i \setminus B_j)^c\right) &= \mu\left(f^{-1}(A_i \setminus B_j)^c\right) \\
&\quad + \sum_{k=2}^{\infty} \mu\left(f^{-1}(A_i \setminus B_k)^c \setminus \bigcup_{j=1}^{k-1} f^{-1}(A_i \setminus B_j)^c\right) \\
&= 1 - \mu(f^{-1}(A_i \setminus B_j)) \\
&\quad + \sum_{k=2}^{\infty} \mu\left(f^{-1}\left((A_i \setminus B_k) \setminus \bigcup_{j=1}^{k-1}(A_i \setminus B_j)\right)^c\right) \\
&= 1 - \mu(A_i \setminus B_j) \\
&\quad + \sum_{k=2}^{\infty}\left(1 - \mu\left((A_i \setminus B_k) \setminus \bigcup_{j=1}^{k-1}(A_i \setminus B_j)\right)\right) \\
&= \mu(A_i \setminus B_j)^c + \sum_{k=2}^{\infty} \mu\left((A_i \setminus B_k) \setminus \bigcup_{j=1}^{k-1}(A_i \setminus B_j)\right)^c \\
&= \mu(A_i \setminus B_j)^c + \mu\left(\sum_{k=2}^{\infty}\left((A_i \setminus B_k) \setminus \bigcup_{j=1}^{k-1}(A_i \setminus B_j)\right)^c\right) \\
&= \mu\left(\bigcup_{j=1}^{\infty}(A_i \setminus B_j)^c\right).
\end{aligned}$$

Portanto, usando (1.2.1), obtemos

$$\begin{aligned}
\mu(f^{-1}(A \setminus B)) = \sum_{i=1}^{\infty}\left(1 - \mu\left(\bigcup_{j=1}^{\infty}(A_i \setminus B_j)^c\right)\right) &= \sum_{i=1}^{\infty} \mu\left(\bigcup_{j=1}^{\infty}(A_i \setminus B_j)^c\right)^c \\
= \sum_{i=1}^{\infty} \mu\left(\bigcap_{j=1}^{\infty}(A_i \setminus B_j)\right) &= \mu\left(\sum_{i=1}^{\infty}\left(\bigcap_{j=1}^{\infty} A_i \setminus B_j\right)\right) \\
= \mu(A \setminus B).
\end{aligned}$$

Mostremos que $\mathcal{B}$ é fechado por união enumerável. Sejam $A_1, A_2, \cdots \in \mathcal{B}$. Definimos os conjuntos disjuntos

$$D_1 = A_1, D_2 = A_2 \setminus A_1, D_3 = A_3 \setminus (A_1 \cup A_2), \ldots, D_k = A_k \setminus \left(\bigcup_{j=1}^{k-1} A_j\right).$$

Assim, $A = \bigcup_{j=1}^{\infty} A_j = \sum_{i=1}^{\infty} D_i$. Logo,

$$\mu(f^{-1}(\mathrm{A})) = \mu\left(f^{-1}\left(\sum_{i=1}^{\infty}\mathrm{D}_i\right)\right) = \mu\left(\sum_{i=1}^{\infty} f^{-1}(\mathrm{D}_i)\right) = \sum_{i=1}^{\infty}\mu\left(f^{-1}(\mathrm{D}_i)\right)$$
$$= \sum_{i=1}^{\infty}\mu(\mathrm{D}_i) = \mu\left(\sum_{i=1}^{\infty}\mu(\mathrm{D}_i)\right) = \mu(\mathrm{A}).$$

■

1.2.2 Difeomorfismos

Seja U um aberto em algum espaço euclidiano $\mathbb{R}^d$ com $d \geq 1$. Dizemos que uma aplicação $f : \mathrm{U} \longrightarrow \mathrm{U}$ é de classe C^1 em U se f é diferenciável e $\mathrm{D}f$ é uma aplicação contínua.

Um **difeomorfismo** $f : \mathrm{U} \longrightarrow \mathrm{U}$ é uma bijeção diferenciável cuja inversa também é diferenciável. Dizemos que f é um **difeomorfismo de classe** C^1 se f e f^{-1} são aplicações de classe C^1 em U.

Uma aplicação diferenciável $f : \mathrm{U} \longrightarrow \mathrm{U}$ é um **difeomorfismo local** se, para cada $x \in \mathrm{U}$, existe uma vizinhança $\mathrm{V}_x \subset \mathrm{U}$ aberta tal que a restrição de f a V_x é um difeomorfismo sobre um aberto $\mathrm{W}_x \subset \mathrm{U}$. Dizemos que f é um **difeomorfismo local de classe** C^1 se f é um difeomorfismo local e $f|_{\mathrm{V}_x}$ e $f^{-1}|_{\mathrm{V}_x}$ são aplicações de classe C^1.

Observação 1.8. Se $f : \mathrm{U} \longrightarrow \mathbb{R}^d$ é um difeomorfismo local de U sobre $\mathrm{V} = f(\mathrm{U})$, então para todo $x \in \mathrm{U}$, $\mathrm{D}f(x) \in \mathrm{L}(\mathbb{R}^d)$ é um isomorfismo.

Em termos do **determinante jacobiano**, dizemos que $f : \mathrm{U} \longrightarrow \mathrm{U}$ é um **difeomorfismo local** se o seu determinante jacobiano é não nulo em todo ponto $x \in \mathrm{U}$.

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

$$\mathrm{vol}(\mathrm{A}) = \int_{\mathrm{A}} dx_1 \ldots dx_d \quad \text{e} \quad \int_{\mathrm{A}} \varphi \mathrm{dvol} = \int_{\mathrm{A}} \varphi(x_1, \ldots, x_d) dx_1 \ldots dx_d$$

para qualquer conjunto mensurável A e qualquer função φ integrável.

Observação 1.9. A fórmula de mudança de variáveis afirma que, para qualquer conjunto mensurável $\mathrm{A} \subset \mathrm{U}$, temos

$$\mathrm{vol}(f(\mathrm{A})) = \int_{\mathrm{A}} |\mathrm{det}\mathrm{D}f| \ \mathrm{dvol}. \tag{1.2.2}$$

Lema 1.10. Um difeomorfismo $f : \mathrm{M} \longrightarrow \mathrm{M}$ de classe C^1 deixa invariante o volume se, e somente se, o valor absoluto $|\mathrm{det}\mathrm{D}f|$ do seu determinante jacobiano é constante igual a 1.

Demonstração: Como f é de classe C^1 então o seu jacobiano é contínuo.

unesp

Suponha que $|\text{detD}f| > 1$ em algum ponto $x \in \text{M}$. Então, como o jacobiano é contínuo, existem uma vizinhança V de x e algum número $\sigma > 1$ tais que

$$|\text{detD}f(y)| \geq \sigma \quad \text{para todo } y \in \text{V}. \tag{1.2.3}$$

Substituindo $\text{A} = \text{V}$ na equação (1.2.2) e usando (1.2.3), obtemos

$$\text{vol}(f(\text{V})) = \int_{\text{V}} |\text{detD}f| \, \text{dvol} \geq \int_{\text{V}} \sigma \, \text{dvol} = \sigma \text{vol}(\text{V}) > \text{vol}(\text{V}).$$

Denotando $\text{B} = f(\text{V})$, temos que $f^{-1}(\text{B}) = \text{V}$. Como $\text{vol}(f(\text{V})) > \text{vol}(\text{V})$ então $\text{vol}(\text{B}) > \text{vol}(f^{-1}(\text{B}))$. Portanto, f não deixa invariante o volume.

De modo análogo, mostra-se que para $|\text{detD}f| < 1$ em algum ponto $x \in \text{M}$, f não deixa invariante o volume.

Para mostrar a recíproca, suponha que $|\text{detD}f| = 1$ em todo ponto $x \in \text{M}$ e considere B um conjunto mensurável. Seja $\text{A} = f^{-1}(\text{B})$, então $f(\text{A}) = \text{B}$. Assim, usando (1.2.2), obtemos

$$\text{vol}(\text{B}) = \text{vol}(f(\text{A})) = \int_{\text{A}} |\text{detD}f| \, \text{dvol} = \int_{\text{A}} 1.\text{dvol} = \text{vol}(\text{A}) = \text{vol}(f^{-1}(\text{B})).$$

Portanto, f deixa invariante o volume. ∎

Dada uma função $\rho : \text{U} \longrightarrow [0, +\infty)$, denota-se por $\mu = \rho\text{vol}$ a medida definida por

$$\mu(\text{A}) = \int_{\text{A}} \rho \, \text{dvol}. \tag{1.2.4}$$

Lema 1.11. Suponha que $f : \text{U} \longrightarrow \text{U}$ é um difeomorfismo local de classe C^1 e ρ é uma função contínua. Então, f deixa invariante a medida $\mu = \rho\text{vol}$ se, e somente se,

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

Demonstração: Suponha primeiro que $\displaystyle\sum_{x \in f^{-1}(y)} \frac{\rho(x)}{|\text{detD}f(x)|} = \rho(y)$ para todo $y \in \text{U}$ e considere A um conjunto mensurável. Sejam $\text{B}_1, \ldots, \text{B}_n$ as componentes conexas de $f^{-1}(\text{A})$. Podemos supor B_j disjuntos dois-a-dois (caso contrário, decompomos A para que isso aconteça). Desta forma, usando (1.2.4), obtemos

$$\mu(\text{A}) = \int_{\text{A}} \rho(y) \, \text{dvol} = \int_{\text{A}} \sum_{x \in f^{-1}(y)} \frac{\rho(x)}{|\text{detD}f(x)|} \, \text{dvol} = \sum_{j=1}^{n} \int_{\text{B}_j} \rho(x) \, \text{dvol}$$

$$= \sum_{j=1}^{n} \mu(\text{B}_j) = \mu(f^{-1}(\text{A})).$$

Portanto, f deixa invariante a medida $\mu = \rho\text{dvol}$.

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Como f é um difeomorfismo local o seu determinante jacobiano é não nulo e sendo f de classe C^1 o seu jacobiano é contínuo. Tendo em vista que ρ e o jacobiano de f são contínuas, segue-se que $\sum_{x\in f^{-1}(y)} \frac{\rho(x)}{|\det Df(x)|}$ é contínuo.

Para provar a recíproca, suponha que $\sum_{x\in f^{-1}(y)} \frac{\rho(x)}{|\det Df(x)|} > \rho(y)$ para algum $y \in U$. Então, existe uma vizinhança V de y tal que

$$\sum_{z\in f^{-1}(w)} \frac{\rho(z)}{|\det Df(z)|} > \rho(w) \quad \text{para todo} \quad w \in V. \tag{1.2.5}$$

Sejam $U_1, \ldots, U_n$ as componentes conexas de $f^{-1}(V)$. Podemos supor, diminuindo V se necessário, que $U_i \cap U_j = \emptyset$ se $i \neq j$. Seja $z = f^{-1}(w)$, então $f(z) = w$. Assim,

$$\begin{aligned}
\mu(f^{-1}(V)) &= \int_{f^{-1}(V)} \rho(z)\ \mathrm{dvol} = \sum_{j=1}^{n} \int_{U_j} \rho(z)\ \mathrm{dvol} \\
&= \sum_{j=1}^{n} \int_{f(U_j)} \frac{\rho(z)}{|\det Df(z)|}\ \mathrm{dvol} = \int_V \sum_{z\in f^{-1}(w)} \frac{\rho(z)}{|\det Df(z)|}\ \mathrm{dvol} \\
&> \int_V \rho(w)\ \mathrm{dvol} = \mu(V).
\end{aligned}$$

Portanto, f não deixa invariante a medida $\mu = \rho\mathrm{dvol}$.

De modo análogo, mostra-se que para $\sum_{x\in f^{-1}(y)} \frac{\rho(x)}{|\det Df(x)|} < \rho(y)$ em algum $x \in U$, f não deixa invariante a medida $\mu = \rho\mathrm{dvol}$. ∎

1.2.3 Transformação de Gauss

A transformação de Gauss $G : (0,1] \longrightarrow [0,1]$ é definida por $G(x) =$ parte fracionária de $\frac{1}{x}$, ou seja, $G(x) = \frac{1}{x} - \left[\frac{1}{x}\right]$. Veja o gráfico de G descrito na Figura 1.2.

O gráfico de G foi esboçado conforme a seguir:

- Se $x \in \left(\frac{1}{2}, 1\right]$ então $\frac{1}{x} \in [1,2)$ e portanto a sua parte inteira $\left[\frac{1}{x}\right]$ é igual a 1. Isto quer dizer que neste intervalo a transformação G é dada por $G(x) = \frac{1}{x} - 1$.
- Se $x \in \left(\frac{1}{3}, \frac{1}{2}\right]$ então $\frac{1}{x} \in [2,3)$ e portanto a sua parte inteira $\left[\frac{1}{x}\right]$ é igual a 2. Logo, $G(x) = \frac{1}{x} - 2$.

unesp

- Em geral, se $x \in \left(\frac{1}{k+1}, \frac{1}{k}\right]$ para algum $k \in \mathbb{N}$ então a sua parte inteira $\left[\frac{1}{x}\right]$ é igual a k. Portanto, $\mathrm{G}(x) = \frac{1}{x} - k$.

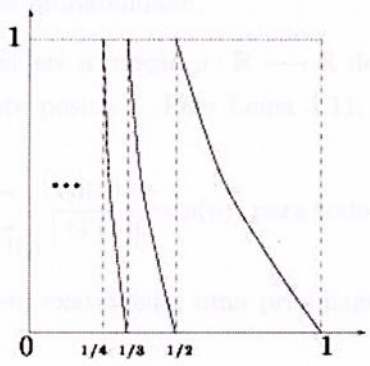


Figura 1.2: Transformações de Gauss

Note que G não está definida para $x = 0$. Além disso, $\mathrm{G}(\frac{1}{k}) = 0$ para todo $k \in \mathbb{N}$. Portanto, o segundo iterado $\mathrm{G}^2(\frac{1}{k})$ não está definido nestes pontos e os iterados G^n, $n \geq 3$, não estão definidos nas suas pré-imagens. Isto quer dizer, a rigor, que G não é um sistema dinâmico segundo a definição dada anteriormente. No entanto, isto não coloca nenhum problema para o que pretendemos fazer. Pois, 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 a transformação de Gauss interessante do ponto de vista ergódico é que ela admite uma probabilidade invariante que é equivalente à medida de Lebesgue no intervalo.

De fato: Considere a medida definida por

$$\mu(\mathrm{A}) = \int_{\mathrm{A}} \frac{\mathrm{c}}{x+1}\, dx$$

para cada mensurável $\mathrm{A} \subset [0,1]$, onde c é uma constante positiva. Note que a integral está bem definida, pois a função integranda é contínua no intervalo $[0,1]$. Tomando $\frac{\mathrm{c}}{2} \leq \frac{\mathrm{c}}{x+1} \leq \mathrm{c}$, temos

$$\int_{\mathrm{A}} \frac{\mathrm{c}}{2}\, dx \leq \int_{\mathrm{A}} \frac{\mathrm{c}}{x+1}\, dx \leq \int_{\mathrm{A}} \mathrm{c}\, dx \Longrightarrow \frac{\mathrm{c}}{2} \int_{\mathrm{A}} dx \leq \int_{\mathrm{A}} \frac{\mathrm{c}}{x+1}\, dx \leq \mathrm{c} \int_{\mathrm{A}} dx.$$

Portanto, $\frac{c}{2}\, m(A) \leq \mu(A) \leq cm(A)$ para todo mensurável $A \subset [0,1]$. Em particular, μ é equivalente à medida de Lebesgue m: as duas medidas têm os mesmos conjuntos com medida nula.

Proposição 1.12. A medida μ é invariante por G. Além disso, se escolhermos $c = \frac{1}{\log 2}$ então μ é uma probabilidade.

Demonstração: Considere a função $\rho : \mathbb{R} \longrightarrow \mathbb{R}$ definida por $\rho(x) = \frac{c}{x+1}$, onde c é uma constante positiva. Pelo Lema 1.11, μ é invariante por G se tivermos

$$\sum_{x \in G^{-1}(y)} \frac{\rho(x)}{|\, G'(x)|} = \rho(y) \text{ para todo } y. \tag{1.2.6}$$

Observe que cada y tem exatamente uma pré-imagem x_k em cada intervalo $\left(\frac{1}{k+1}, \frac{1}{k}\right]$, dada por

$$G(x_k) = \frac{1}{x_k} - k = y \Longleftrightarrow x_k = \frac{1}{y+k}. \tag{1.2.7}$$

Usando (1.2.7), temos

$$\frac{cx_k^2}{x_k+1} = \frac{c\left(\frac{1}{y+k}\right)^2}{\frac{1}{y+k}+1} = \frac{c(y+k)}{(y+k)^2(y+k+1)} = \frac{c}{(y+k)(y+k+1)}. \tag{1.2.8}$$

Note que $G'(x) = \left(\frac{1}{x}\right)' = -\frac{1}{x^2}$. Portanto, usando (1.2.8) para reescrever (1.2.6), obtemos

$$\sum_{k=1}^{\infty} \frac{cx_k^2}{x_k+1} = \frac{c}{y+1} \Longleftrightarrow \sum_{k=1}^{\infty} \frac{c}{(y+k)(y+k+1)} = \frac{c}{y+1}. \tag{1.2.9}$$

Para verificar que esta igualdade é realmente satisfeita, observe que

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

Note

$$\sum_{k=1}^{n} \left(\frac{c}{y+k} - \frac{c}{y+k+1}\right) = \left(\frac{c}{y+1} - \frac{c}{y+2}\right) + \left(\frac{c}{y+2} - \frac{c}{y+3}\right) + \dots$$
$$\dots + \left(\frac{c}{y+n-1} - \frac{c}{y+n}\right) = \frac{c}{y+1} - \frac{c}{y+n}.$$

unesp

Desta forma,

$$\sum_{k=1}^{\infty} \frac{c}{(y+k)(y+k+1)} = \lim_{n\to\infty} \sum_{k=1}^{n} \left(\frac{c}{y+k} - \frac{c}{y+k+1} \right)$$
$$= \lim_{n\to\infty} \left(\frac{c}{y+1} - \frac{c}{y+n} \right) = \frac{c}{y+1}.$$

Que é precisamente o que se afirma em (1.2.9). Isto prova a invariância. Finalmente,

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

Logo, escolhendo $c = \dfrac{1}{\log 2}$, obtemos $\mu([0,1]) = 1$. ∎

1.3 Existência de medidas invariantes

Nesta seção consideremos M um espaço métrico compacto e introduziremos uma topologia no conjunto $\mathcal{M}_1(M)$ das probabilidades sobre os borelianos de M, chamada **topologia fraca***, que será útil para provar o **Teorema da existência de medidas invariantes**.

A definição a seguir diz que duas medidas estão próximas se suas integrais estão próximas para **muitas** funções contínuas.

Definição 1.13. Dados um número $\varepsilon > 0$, uma medida $\mu \in \mathcal{M}_1(M)$ e um conjunto finito $F = \{\varphi_1, \ldots, \varphi_K\}$ de funções contínuas $\varphi_j : M \longrightarrow \mathbb{R}$. Seja

$$V(\mu, F, \varepsilon) = \left\{ \nu \in \mathcal{M}_1(M);\ \left| \int \varphi_j \, d\nu - \int \varphi_j \, d\mu \right| < \varepsilon \ \text{ para todo } \varphi_j \in F \right\}.$$

A **topologia fraca*** é definida estipulando que os conjuntos $V(\mu, F, \varepsilon)$, com F e ε variáveis, constituem uma base de vizinhanças da medida μ.

O lema seguinte ajudará a compreender melhor o significado da topologia fraca*.

Lema 1.14. Uma seqüência $(\mu_n)_{n\in\mathbb{N}}$ em $\mathcal{M}_1(M)$ converge para uma medida $\mu \in \mathcal{M}_1(M)$ na topologia fraca* se, e somente se,

$$\int \varphi \, d\mu_n \longrightarrow \int \varphi \, d\mu$$

para toda função contínua $\varphi : M \longrightarrow \mathbb{R}$.

Demonstração: Para provar a afirmação **somente se**, considere qualquer função contínua φ e forme o conjunto $F = \{\varphi\}$. Como $\mu_n \longrightarrow \mu$ em $\mathcal{M}_1(M)$, então dado qualquer $\varepsilon > 0$, existe um índice $\bar{n}$ a partir do qual μ_n está na vizinhança (V, μ, F, ε). Mas isto significa, precisamente, que,

$$\left| \int \varphi \, d\mu_n - \int \varphi \, d\mu \right| < \varepsilon$$



para todo $n \geq \bar{n}$. Em outras palavras, a seqüência $\int \varphi \, d\mu_n \longrightarrow \int \varphi \, d\mu$.

A recíproca afirma que se $\int \varphi \, d\mu_n \longrightarrow \int \varphi \, d\mu$, para toda função contínua φ, então dados quaisquer F e $\varepsilon > 0$, existe um índice $\bar{n}$ a partir do qual $\mu_n \in V(\mu, F, \varepsilon)$.

De fato: Seja $F = \{\varphi_1, \ldots, \varphi_K\}$. Por hipótese, para cada $1 \leq j \leq K$, existem $\bar{n}_j$ tais que

$$\left| \int \varphi_j \, d\mu_n - \int \varphi_j \, d\mu \right| < \varepsilon$$

para todo $n \geq \bar{n}_j$. Tomando $\bar{n} = \max\{\bar{n}_1, \ldots, \bar{n}_K\}$, temos $\mu_n \in V(\mu, F, \varepsilon)$ para $n \geq \bar{n}$. ∎

As principais propriedades da topologia fraca*, necessárias para analisar o conjunto $\mathcal{M}_1(M)$, são dadas no teorema seguinte.

Teorema 1.15. $\mathcal{M}_1(M)$ munido da topologia fraca* é metrizável e compacto.

Vamos começar por demonstrar que $\mathcal{M}_1(M)$ é metrizável, isto é, existe uma distância d que gera a topologia fraca* em $\mathcal{M}_1(M)$. Para isso, usamos o seguinte resultado:

Proposição 1.16. Se M é um espaço métrico compacto, então $C^0(M)$ tem subconjuntos enumeráveis densos.

Demonstração: Ver referência [19].

Denotemos por $C^0(M)$ o espaço das funções contínuas $\varphi : M \longrightarrow \mathbb{R}$ munido da norma da convergência uniforme:

$$\| \varphi_1 - \varphi_2 \| = \sup \{ | \varphi_1(x) - \varphi_2(x) |; \; x \in M \} .$$

Pela Proposição 1.16, podemos escolher um subconjunto enumerável $\mathcal{F} = \{\varphi_n; \; n \in \mathbb{N}\}$ denso na bola unitária do espaço $C^0(M)$. Feito isso, definimos

$$d(\mu_1, \mu_2) = \sum_{n=1}^{\infty} \frac{1}{2^n} \left| \int \varphi_n \, d\mu_1 - \int \varphi_n \, d\mu_2 \right| \tag{1.3.10}$$

para qualquer par de medidas μ_1 e μ_2.

Proposição 1.17. Seja d a expressão em 1.3.10. Então,

1. d está bem definida;
2. d é uma distância;
3. d gera a topologia fraca* em $\mathcal{M}_1(M)$.

Demonstração: Seja $B_1 = \{\varphi \in C^0(M); \sup|\varphi| \leq 1\}$ a bola unitária de $C^0(M)$.
1. Devemos mostrar que a série em (1.3.10) converge. Note que

$$\frac{1}{2^n}\left|\int \varphi_n \, d\mu_1 - \int \varphi_n \, d\mu_2\right| \leq \frac{1}{2^n}\left(\int |\varphi_n| \, d\mu_1 + \int |\varphi_n| \, d\mu_2\right)$$
$$\leq \frac{1}{2^n}\left(\int 1 \cdot d\mu_1 + \int 1 \cdot d\mu_2\right) = \frac{2}{2^n}. \qquad (1.3.11)$$

Como o termo geral da série em (1.3.10) é limitada por $\frac{2}{2^n}$, então ela converge.
2. Sejam μ_1, μ_2 e μ_3 medidas em $\mathcal{M}_1(M)$.
Afirmação 1: $d(\mu_1, \mu_2) > 0$ e $d(\mu_1, \mu_2) = 0$ se, e somente se, $\mu_1 = \mu_2$.
prova: Usando (1.3.11) e supondo $\mu_1 \neq \mu_2$, então existe $n \geq 1$ tal que $\left|\int \varphi_n \, d\mu_1 - \int \varphi_n \, d\mu_2\right| > 0$. Logo,

$$d(\mu_1, \mu_2) = \sum_{n=1}^{\infty} \frac{1}{2^n}\left|\int \varphi_n \, d\mu_1 - \int \varphi_n \, d\mu_2\right| > 0.$$

Isto prova que $d(\mu_1, \mu_2) > 0$. Supondo $\mu_1 = \mu_2$, obtemos

$$d(\mu_1, \mu_2) = \sum_{n=1}^{\infty} \frac{1}{2^n}\left|\int \varphi_n \, d\mu_1 - \int \varphi_n \, d\mu_2\right|$$
$$= \sum_{n=1}^{\infty} \frac{1}{2^n}\left|\int \varphi_n \, d\mu_1 - \int \varphi_n \, d\mu_1\right| = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot 0 = 0.$$

A hipótese $d(\mu_1, \mu_2) = 0$ significa que $\left|\int \varphi_j \, d\mu_1 - \int \varphi_j \, d\mu_2\right| = 0$. Logo,

$$\int \varphi_j \, d\mu_1 - \int \varphi_j \, d\mu_2 = 0 \Longrightarrow \int \varphi_j \, d\mu_1 = \int \varphi_j \, d\mu_2 \qquad (1.3.12)$$

para todo $\varphi_j \in \mathcal{F}$. Agora, dada qualquer $\varphi \in B_1$, podemos encontrar uma seqüência de funções de $\mathcal{F}$ convergindo uniformemente para φ. Conseqüentemente, a igualdade (1.3.12) continua valendo para φ:

$$\int \varphi \, d\mu_1 = \int \varphi \, d\mu_2 \quad \text{para toda } \varphi \in B_1. \qquad (1.3.13)$$

De fato: Suponha que $\varphi_j(x) \longrightarrow \varphi(x)$ uniformemente em M. Então, dado qualquer $\varepsilon > 0$, podemos encontrar $n_0 \in \mathbb{N}$ tal que

$$j > n_0 \Longrightarrow |\varphi_j(x) - \varphi(x)| < \frac{\varepsilon}{2} \text{ para todo } x \in M.$$

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Usando a igualdade (1.3.12), como μ_1 e μ_2 são probabilidades, obtemos

$$\begin{aligned}
\left|\int \varphi\, d\mu_1 - \int \varphi\, d\mu_2\right| &= \left|\int \varphi\, d\mu_1 - \int \varphi\, d\mu_2 + \int \varphi_j\, d\mu_1 - \int \varphi_j\, d\mu_1\right| \\
&= \left|\int \varphi\, d\mu_1 - \int \varphi\, d\mu_2 + \int \varphi_j\, d\mu_2 - \int \varphi_j\, d\mu_1\right| \\
&\leq \left|\int \varphi\, d\mu_1 - \int \varphi_j\, d\mu_1\right| + \left|\int \varphi_j\, d\mu_2 - \int \varphi\, d\mu_2\right| \\
&= \left|-\int (\varphi_j - \varphi)\, d\mu_1\right| + \left|\int (\varphi_j - \varphi)\, d\mu_2\right| \\
&\leq \int |\,\varphi_j - \varphi|\; d\mu_1 + \int |\,\varphi_j - \varphi|\; d\mu_2 \\
&< \frac{\varepsilon}{2}\int d\mu_1 + \frac{\varepsilon}{2}\int d\mu_2 = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Portanto, a igualdade (1.3.13) é verdadeira para toda função contínua φ. Isto quer dizer que $\mu_1 = \mu_2$. O que completa a prova da Afirmação 1.

Afirmação 2: $d(\mu_1, \mu_2) = d(\mu_2, \mu_1)$.

Prova: Por definição, temos

$$\begin{aligned}
d(\mu_1, \mu_2) &= \sum_{n=1}^{\infty} \frac{1}{2^n}\left|\int \varphi_n\, d\mu_1 - \int \varphi_n\, d\mu_2\right| \\
&= \sum_{n=1}^{\infty} \frac{1}{2^n}\left|-\left(\int \varphi_n\, d\mu_2 - \int \varphi_n\, d\mu_1\right)\right| \\
&= \sum_{n=1}^{\infty} \frac{1}{2^n}\left|\int \varphi_n\, d\mu_2 - \int \varphi_n\, d\mu_1\right| = d(\mu_2, \mu_1).
\end{aligned}$$

Afirmação 3: $d(\mu_1, \mu_3) \leq d(\mu_1, \mu_2) + d(\mu_2, \mu_3)$.

Prova: Por definição, temos

$$
\begin{aligned}
d(\mu_1,\mu_3) &= \sum_{n=1}^{\infty}\frac{1}{2^n}\left|\int \varphi_n\, d\mu_1 - \int \varphi_n\, d\mu_3\right| \\
&= \sum_{n=1}^{\infty}\frac{1}{2^n}\left|\int \varphi_n\, d\mu_1 - \int \varphi_n\, d\mu_2 + \int \varphi_n\, d\mu_2 - \int \varphi_n\, d\mu_3\right| \\
&\leq \sum_{n=1}^{\infty}\frac{1}{2^n}\left(\left|\int \varphi_n\, d\mu_1 - \int \varphi_n\, d\mu_2\right| + \left|\int \varphi_n\, d\mu_2 - \int \varphi_n\, d\mu_3\right|\right) \\
&\leq \sum_{n=1}^{\infty}\frac{1}{2^n}\left|\int \varphi_n\, d\mu_1 - \int \varphi_n\, d\mu_2\right| + \sum_{n=1}^{\infty}\frac{1}{2^n}\left|\int \varphi_n\, d\mu_2 - \int \varphi_n\, d\mu_3\right| \\
&= d(\mu_1,\mu_2) + d(\mu_2,\mu_3).
\end{aligned}
$$

As Afirmações 1, 2 e 3 provam que d é uma métrica

3. Devemos mostrar que toda bola $\mathrm{B}(\mu,\delta) = \{\nu \in \mathcal{M}_1(\mathrm{M});\ d(\mu,\nu) < \delta\}$ contém alguma vizinhança $\mathrm{V}(\mu,\mathrm{F},\varepsilon)$ e reciprocamente. Dado $\delta > 0$ e fixando $K \geq 1$ suficientemente grande para que

$$\sum_{n=K}^{\infty}\frac{1}{2^n} < \frac{\delta}{2} \tag{1.3.14}$$

e considere $\mathrm{F} = \{\varphi_1,\ldots,\varphi_K\}$ formado pelos primeiros K elementos de $\mathcal{F}$. Além disso, tomemos $\varepsilon = \dfrac{\delta}{2}$.

Afirmação 4: $\mathrm{V}(\mu,\mathrm{F},\varepsilon) \subset \mathrm{B}(\mu,\delta)$.

Prova: Usando (1.3.14), note que

$$\sum_{n=K+1}^{\infty}\frac{2}{2^n} = \frac{1}{2}\sum_{n=K}^{\infty}\frac{2}{2^n} = \sum_{n=K}^{\infty}\frac{1}{2^n} < \frac{\delta}{2}. \tag{1.3.15}$$

Como $\displaystyle\sum_{n=1}^{K}\frac{1}{2^n} = \frac{1}{2}+\frac{1}{4}+\cdots+\frac{1}{2^K} = \frac{1/2(1-1/2^K)}{1-1/2} = 1-\frac{1}{2^K} < 1$ e usando (1.3.15), obtemos

$$\nu \in \mathrm{V}(\mu,\mathrm{F},\varepsilon) \Longrightarrow \left|\int \varphi_n\, d\mu - \int \varphi_n\, d\nu\right| < \epsilon \text{ para todo } 1 \leq n \leq K.$$

Logo,

$$
\begin{aligned}
&\sum_{n=1}^{\infty}\frac{1}{2^n}\left|\int \varphi_n\, d\mu - \int \varphi_n\, d\nu\right| \\
&= \sum_{n=1}^{K}\frac{1}{2^n}\left|\int \varphi_n\, d\mu - \int \varphi_n\, d\nu\right| + \sum_{n=K+1}^{\infty}\frac{1}{2^n}\left|\int \varphi_n\, d\mu - \int \varphi_n\, d\nu\right| \\
&< \sum_{n=1}^{K}\frac{1}{2^n}\cdot\varepsilon + \sum_{n=K+1}^{\infty}\frac{2}{2^n} < \frac{\delta}{2}+\frac{\delta}{2} = \delta.
\end{aligned}
$$

O que prova a Afirmação 4.

Reciprocamente, dados $\mathrm{F} = \{\psi_1, \ldots, \psi_K\}$ e $\varepsilon > 0$, como $\mathcal{F}$ é denso, existe para cada $j = 1, \ldots, K$ algum φ_{n_j} em $\mathcal{F}$ tal que

$$\|\varphi_{n_j} - \psi_j\| < \frac{\varepsilon}{3}. \tag{1.3.16}$$

Fixemos $\delta = 2\varepsilon > 0$ suficientemente pequeno para que

$$2^{n_j}\delta < \frac{\varepsilon}{3} \qquad \text{para cada } j = 1, \ldots, K. \tag{1.3.17}$$

Afirmação 5: $\mathrm{B}(\mu, \delta) \subset \mathrm{V}(\mu, \mathrm{F}, \varepsilon)$.

Prova: Por definição, $\nu \in \mathrm{B}(\mu, \delta)$ implica $d(\mu, \nu) < \delta$. Para cada $j = 1, \ldots, K$, obtemos

$$\frac{1}{2^{n_j}}\left|\int \varphi_{n_j}\, d\mu - \int \varphi_{n_j}\, d\nu\right| < \delta \implies \left|\int \varphi_{n_j}\, d\mu - \int \varphi_{n_j}\, d\nu\right| < 2^{n_j}\delta$$

o que implica, combinando as desigualdades (1.3.16) e (1.3.17), que

$$\begin{aligned}\left|\int \psi_j\, d\mu - \int \psi_j\, d\nu\right| &\leq \left|\int \psi_j\, d\mu - \int \varphi_{n_j}\, d\mu\right| + \left|\int \varphi_{n_j}\, d\mu - \int \varphi_{n_j}\, d\nu\right| \\ &+ \left|\int \varphi_{n_j}\, d\nu - \int \psi_j\, d\nu\right| < \left|-\int (\varphi_{n_j} - \psi_j)\, d\mu\right| + 2^{n_j}\delta \\ &+ \left|\int (\varphi_{n_j} - \psi_j)\, d\nu\right| \leq \sup|\varphi_{n_j} - \psi_j| \int d\mu + \frac{\varepsilon}{3} \\ &+ \sup|\varphi_{n_j} - \psi_j| \int d\nu = \|\varphi_{n_j} - \psi_j\| + \frac{\varepsilon}{3} + \|\varphi_{n_j} - \psi_j\| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.\end{aligned}$$

Isto prova a Afirmação 5. ∎

Resta provar que $(\mathcal{M}_1, \text{fraca}^*)$ é um espaço compacto.

Um **operador linear** $\Phi : \mathrm{C}^0(\mathrm{M}) \longrightarrow \mathbb{R}$ é dito **positivo** se $\Phi(\varphi) > 0$ para toda função $\varphi \in \mathrm{C}^0(\mathrm{M})$ positiva em todo ponto.

O teorema seguinte afirma que as integrais são os únicos operadores lineares positivos no espaço das funções contínuas. Uma demonstração deste resultado poder encontrada na referência [19].

Teorema 1.18 (Riesz - Markov). Seja $\Phi : \mathrm{C}^0(\mathrm{M}) \longrightarrow \mathbb{R}$ qualquer operador linear positivo. Então, existe uma única medida boreliana μ em M tal que

$$\Phi(\varphi) = \int \varphi\, d\mu \qquad \text{para toda } \varphi \in \mathrm{C}^0(\mathrm{M}).$$

Observação 1.19. μ é uma probabilidade se, e somente se, $\Phi(1) = 1$.

De fato: Usando o teorema de Riesz-Markov, obtemos

$$\mu(\mathrm{M}) = \int 1\, d\mu = \Phi(1) = 1.$$

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Vamos provar que o espaço $\mathcal{M}_1(M)$ é compacto para a topologia fraca*. Como $\mathcal{M}_1(M)$ é metrizável, basta provar:

Proposição 1.20. Toda seqüência $(\mu_n)_{k\in\mathbb{N}}$ em $\mathcal{M}_1(M)$ admite alguma subseqüência que é convergente na topologia fraca*.

Demonstração: Seja $\mathcal{F} = \{\varphi_n;\ n \in \mathbb{N}\}$ um subconjunto enumerável denso na bola unitária de $C^0(M)$.

Afirmação 1: Para cada $n \in \mathbb{N}$, a seqüência de números reais $\int \varphi_n\ d\mu_k$, $k \in \mathbb{N}$, é limitada por 1.

Prova: Como $\sup|\ \varphi_n| \leq 1$ e μ_k são probabilidades, obtemos

$$\left|\int \varphi_n\ d\mu_k\right| \leq \int |\varphi_n|\ d\mu_k = \sup|\ \varphi_n| \int d\mu_k \leq 1.$$

Isto prova a Afirmação 1.

Como a seqüência $\int \varphi_n\ d\mu_k$ com $k \in \mathbb{N}$ é limitada, pelo Teorema de Bolzano-Weierstrass, para cada $n \in \mathbb{N}$, existe uma seqüência de índices $\left(k_j^n\right)_{j\in\mathbb{N}}$ tal que

$$\int \varphi_n\ d\mu_{k_j^n} \text{ converge para algum número } \Phi_n \in \mathbb{R} \text{ quando } j \longrightarrow \infty.$$

Além disso, cada seqüência de índices $\left(k_j^{n+1}\right)_{j\in\mathbb{N}}$ pode ser escolhida como subseqüência da seqüência anterior $\left(k_j^n\right)_{j\in\mathbb{N}}$. Definimos $\ell_j = k_j^j$ para cada $j \in \mathbb{N}$. Por construção, a menos de um número finito de termos, $(\ell_j)_{j\in\mathbb{N}}$ é uma subseqüência de índices de cada uma das $\left(k_j^n\right)_{j\in\mathbb{N}}$. Logo,

$$\int \varphi_n\ d\mu_{\ell_j} \longrightarrow \Phi_n \qquad \text{para todo } j \in \mathbb{N}.$$

Afirmação 2: $\Phi(\rho) = \lim_{j\to\infty} \int \rho\ d\mu_{\ell_j}$ existe para toda função $\rho \in C^0(M)$.

Prova: Suponha primeiro que ρ está na bola unitária de $C^0(M)$. Dado qualquer $\varepsilon > 0$, podemos encontrar $\varphi_n \in \mathcal{F}$ tal que $\|\ \rho - \varphi_n\| \leq \varepsilon$. Então, como μ_{ℓ_j} é uma probabilidade, obtemos

$$\begin{aligned}\left|\int \rho\ d\mu_{\ell_j} - \int \varphi_n\ d\mu_{\ell_j}\right| &= \left|\int (\rho - \varphi_n)\ d\mu_{\ell_j}\right| \leq \int |\ \rho - \varphi_n|\ d\mu_{\ell_j} \\ &= \sup|\ \rho - \varphi_n| \int d\mu_{\ell_j} = \sup|\ \rho - \varphi_n| \\ &= \|\ \rho - \varphi_n\| \leq \varepsilon \text{ para todo } j.\end{aligned}$$

Como $\int \varphi_n\ d\mu_{\ell_j} \longrightarrow \Phi_n$ para todo $j \in \mathbb{N}$, seque-se que

$$\left|\limsup_{j\to\infty} \int \rho d\mu_{\ell_j} - \liminf_{j\to\infty} \int \rho d\mu_{\ell_j}\right| \leq \left|\limsup_{j\to\infty} \int \rho d\mu_{\ell_j} - \limsup_{j\to\infty} \int \varphi_n d\mu_{\ell_j}\right|$$

unesp

$$+\left|\liminf_{j\to\infty}\int\varphi_n\,d\mu_{\ell_j}-\liminf_{j\to\infty}\int\rho\,d\mu_{\ell_j}\right| = \limsup_{j\to\infty}\left|\int\rho d\mu_{\ell_j}-\int\varphi_n d\mu_{\ell_j}\right|$$

$$+\liminf_{j\to\infty}\left|\int\rho\,d\mu_{\ell_j}-\int\varphi_n\,d\mu_{\ell_j}\right|\leq\limsup_{j\to\infty}\varepsilon+\liminf_{j\to\infty}\varepsilon=\varepsilon+\varepsilon=2\varepsilon.$$

Como ε é arbitrário, então $\limsup_{j\to\infty}\int\rho\,d\mu_{\ell_j}=\liminf_{j\to\infty}\int\rho\,d\mu_{\ell_j}$. Portanto, o $\lim_{j\to\infty}\int\rho\,d\mu_{\ell_j}$ existe. Isto prova a Afirmação 2 quando a função está na bola unitária. O caso geral reduz-se imediatamente a esse, substituindo ρ por $\frac{\rho}{\|\rho\|}$. Assim, completamos a prova da Afirmação 2.

Afirmação 3: O operador $\Phi : C^0(M) \longrightarrow \mathbb{R}$ definido por $\Phi(\rho)=\lim_{j\to\infty}\int\rho\,d\mu_{\ell_j}$ é linear e positivo: $\Phi(\rho)\geq\min\rho>0$ para toda função $\varphi\in C^0(M)$ positiva em todo ponto. Além disso, $\Phi(1)=1$.

Prova: Suponha $\rho_1,\rho_2\in C^0(M)$ positivas em todo ponto. Então, dado $c\in\mathbb{R}$ e usando o Teorema da Convergência Monótona, obtemos

$$\begin{aligned}\Phi(c\rho_1+\rho_2)&=\lim_{j\to\infty}\int(c\rho_1+\rho_2)\,d\mu_{\ell_j}=\int\lim_{j\to\infty}(c\rho_1+\rho_2)\,d\mu_{\ell_j}\\&=\int\lim_{j\to\infty}c\rho_1\,d\mu_{\ell_j}+\int\lim_{j\to\infty}\rho_2\,d\mu_{\ell_j}\\&=c\int\lim_{j\to\infty}\rho_1\,d\mu_{\ell_j}+\int\lim_{j\to\infty}\rho_2\,d\mu_{\ell_j}\\&=c\lim_{j\to\infty}\int\rho_1\,d\mu_{\ell_j}+\lim_{j\to\infty}\int\rho_2\,d\mu_{\ell_j}=c\Phi(\rho_1)+\Phi(\rho_2).\end{aligned}$$

Isto mostra que Φ é linear. Além disso, por definição e como μ_{ℓ_j} é uma probabilidades, obtemos

$$\Phi(1)=\lim_{j\to\infty}\int 1\,d\mu_{\ell_j}=\lim_{j\to\infty}\int d\mu_{\ell_j}=\lim_{j\to\infty}1=1.$$

Assim, completamos a prova da Afirmação 3.

Como Φ é linear e positivo, pelo Teorema de Riesz-Markov, existe alguma probabilidade boreliana μ em M tal que $\Phi(\rho)=\int\rho\,d\mu$ para toda função contínua $\rho\in C^0(M)$. Agora a Afirmação 2 pode ser reescrita como

$$\int\rho\,d\mu=\lim_{j\to\infty}\int\rho\,d\mu_{\ell_j}\quad\text{para toda }\rho\in C^0(M).$$

De acordo com o Lema 1.14, isto quer dizer que a subseqüência $(\mu_{\ell_j})_{j\in\mathbb{N}}$ converge para μ na topologia fraca*. Logo, $\mathcal{M}_1(M)$ é compacto. Isto completa a demonstração do teorema 1.15. ∎

Dadas $f : M \longrightarrow M$ uma transformação mensurável e μ uma medida qualquer de M. Denota-se por $f_*\mu$, e chama-se **imagem de μ por f**, a medida

$$f_*\mu : \mathcal{A} \longrightarrow [0,+\infty]$$

definida por $f_*\mu(A)=\mu(f^{-1}(A))$ para cada conjunto mensurável $A\subset M$.

unesp

Observação 1.21. μ é invariante por f se, e somente se, $f_*\mu = \mu$.
De fato: Por definição e supondo μ invariante, obtemos

$$f_*\mu(\mathrm{A}) = \mu\left(f^{-1}(\mathrm{A})\right) = \mu(\mathrm{A}).$$

Consideremos a transformação $f_* : \mathcal{M}_1(\mathrm{M}) \longrightarrow \mathcal{M}_1(\mathrm{M})$ que associa a cada $\mu \in \mathcal{M}_1(\mathrm{M})$ a medida $f_*\mu \in \mathcal{M}_1(\mathrm{M})$.

O lema seguinte afirma que se f é contínua em M então f_* é contínua na topologia fraca* de $\mathcal{M}_1(\mathrm{M})$.

Lema 1.22. A transformação $f_* : \mathcal{M}_1(\mathrm{M}) \to \mathcal{M}_1(\mathrm{M})$ é contínua relativamente à topologia fraca*.
Demonstração: Devemos mostrar que se $\mu_n \longrightarrow \mu$ em $\mathcal{M}_1(\mathrm{M})$ então $f_*\mu_n \longrightarrow f_*\mu$. Suponha que $\mu_n \longrightarrow \mu$ em $\mathcal{M}_1(\mathrm{M})$ e seja $\varphi \in \mathrm{C}^0(\mathrm{M})$ uma função qualquer. Então, combinando a observação 1.21 com o lema 1.14, obtemos

$$\int \varphi \, df_*\mu_n = \int \varphi \, d\mu_n \longrightarrow \int \varphi \, d\mu = \int \varphi \, df_*\mu.$$

Isto prova que $f_*\mu_n \longrightarrow f_*\mu$ na topologia fraca*. ∎

O teorema seguinte garante a **existência de medidas invariantes** em grande generalidade.

Teorema 1.23. Seja $f : \mathrm{M} \longrightarrow \mathrm{M}$ uma transformação contínua num espaço métrico compacto. Então, existe pelo menos uma probabilidade invariante por f.
Demonstração: Considere qualquer probabilidade ν em M: por exemplo, a medida de Dirac num ponto qualquer. Forme a seqüência de probabilidades

$$\mu_n = \frac{1}{n}\sum_{j=0}^{n-1} f_*^j\nu \tag{1.3.18}$$

onde $f_*^j\nu$ é a imagem de ν pelo iterado f^j. Pelo Teorema de Riesz-Markov, esta seqüência tem algum ponto de acumulação: existe alguma subseqüência de índices $(n_k)_{k\in\mathbb{N}}$ e alguma probabilidade $\mu \in \mathcal{M}_1(\mathrm{M})$ tais que

$$\mu = \lim_{k\to+\infty} \mu_{n_k} = \lim_{k\to+\infty} \frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^j\nu. \tag{1.3.19}$$

Agora é suficiente provar o lema seguinte.

Lema 1.24. Todo ponto de acumulação da seqüência $(\mu_n)_{n\in\mathbb{N}}$ é uma probabilidade invariante por f.
Demonstração: Seja μ um ponto de acumulação de (μ_n), isto é, $\mu = \lim_{k\to\infty} \mu_{n_k}$

para alguma subseqüência (μ_{n_k}). Note que $\sum_{j=0}^{n_k-1} f_*^j\nu = \nu + f_*\nu + \cdots + f_*^{n_k-1}\nu$.

Portanto,

$$\sum_{j=0}^{n_k-1} f_*^j\nu - \nu = f_*\nu + \cdots + f_*^{n_k-1}\nu. \tag{1.3.20}$$

Seja $A \subset M$ um conjunto mensurável qualquer. Note que

$$\begin{aligned} f_*\left(\frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^j\nu\right)(A) &= \left(\frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^j\nu\right)\left(f^{-1}(A)\right) = \frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^j\nu\left(f^{-1}(A)\right) \\ &= \frac{1}{n_k}\sum_{j=0}^{n_k-1} \nu\left(f^{-j}\left(f^{-1}(A)\right)\right) = \frac{1}{n_k}\sum_{j=0}^{n_k-1} \nu\left(f^{-j-1}(A)\right) \\ &= \frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^{j+1}\nu(A) = \left(\frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^{j+1}\nu\right)(A). \end{aligned}$$

Portanto,

$$f_*\left(\frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^j\nu\right) = \frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^{j+1}\nu. \tag{1.3.21}$$

A partir de (1.3.19), combinando o Lema 1.22 com as igualdades (1.3.21) e (1.3.20), obtemos

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

Como ν é limitada por 1, então $\lim_{k\to+\infty}\frac{1}{n_k}\nu \leq \lim_{k\to+\infty}\frac{1}{n_k} = 0$. Note que, para todo conjunto mensurável $A \subset M$ e como $\nu(f^{-n_k}(A))$ é limitada por 1, obtemos $\frac{1}{n_k}f_*^{n_k}\nu(A) = \frac{1}{n_k}\nu\left(f^{-n_k}(A)\right) \leq \frac{1}{n_k}$. Portanto, $\lim_{k\to+\infty}\frac{1}{n_k}f_*^{n_k}\nu \leq \lim_{k\to+\infty}\frac{1}{n_k} = 0$. Deste modo, usando (1.3.19), obtemos

$$f_*\mu = \lim_{k\to+\infty}\frac{1}{n_k}\sum_{j=0}^{n_k-1} f_*^j\nu = \mu.$$

Logo, μ é invariante por f. Isto completa a demonstração do teorema 1.23. ■

Os exemplos a seguir mostram que nenhuma das hipóteses do Teorema 1.23, continuidade e compacidade, podem ser omitidas.

Exemplo 1.25. A transformação $f : (0,1] \longrightarrow (0,1]$ definida por $f(x) = \frac{x}{2}$ não admite probabilidade invariante.
De fato: Suponha que f admite alguma probabilidade invariante e seja μ essa probabilidade. Então, pelo Teorema de Recorrência 1.4, μ - quase todo ponto $x \in (0,1]$ é recorrente para f. Absurdo!, pois a órbita de qualquer ponto $x \in (0,1]$ converge para zero. Em particular, não acumula no ponto inicial x. Portanto, não existe nenhum ponto recorrente para f. ■

Exemplo 1.26. A transformação $f : [0,1] \longrightarrow [0,1]$ definida por

$$f(x) = \begin{cases} \frac{x}{2}, & \text{se } x \neq 0 \\ 1, & \text{se } x = 0 \end{cases}$$

não admite probabilidade invariante.
De fato: Suponha que f admite alguma probabilidade invariante e seja μ essa probabilidade. Então, pelo Teorema de Recorrência 1.4, μ-quase todo ponto $x \in [0,1]$ é recorrente para f. Absurdo!, pois pelo exemplo 1.25, nenhum ponto $x \in (0,1]$ é recorrente para f. Portanto, se existir alguma probabilidade invariante μ ela tem de 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, a qual é definida por

$$\delta_0(\mathrm{A}) = \begin{cases} 1, & \text{se } 0 \in \mathrm{A} \\ 0, & \text{se } 0 \notin \mathrm{A} \end{cases} \tag{1.3.22}$$

Mas a medida δ_0 não é invariante por f, pois tomando $\mathrm{A} = \{0\}$, temos

$$\delta_0(\mathrm{A}) = 1 \text{ e } f^{-1}(\mathrm{A}) = \emptyset \Longrightarrow \delta_0\left(f^{-1}(\mathrm{A})\right) = 0.$$

Portanto, f não admite nenhuma probabilidade invariante. ■

O exemplo seguinte enfatiza as limitações do teorema 1.23 (que são inerentes à sua grande generalidade): as medidas que ele garante existirem podem ser bastante triviais, por exemplo, neste caso quando falamos de **quase todo ponto** estamos nos referindo apenas o ponto $x = 0$. Por isso, um objetivo importante é obter resultados mais sofisticados de existência de medidas com propriedades adicionais que tornem mais interessante, por exemplo serem equivalentes à medida de Lebesgue. Veremos um exemplo importante de um tal resultado no Capítulo 2.

Exemplo 1.27. A transformação $f : [0,1] \longrightarrow [0,1]$ definida por $f(x) = \frac{x}{2}$, para todo $x \in [0,1]$, admite alguma probabilidade invariante por f.
De fato: Como o intervalo $[0,1]$ é compacto e f é contínua em $[0,1]$, pelo Teorema 1.23, f admite alguma probabilidade invariante. Seja μ essa probabilidade, então μ tem de 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. Usando a definição (1.3.22) e tomando $\mathrm{A} = \{0\}$, obtemos

$$\delta_0(\mathrm{A}) = \delta_0\left(f^{-1}(\mathrm{A})\right) = 1.$$

Portanto, δ_0 é invariante por f. ■



1.4 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 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 na Seção 1.5.

1.4.1 Comentários

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

$$\tau_n(A, x) = \frac{1}{n} \# \left\{ j \in \{0, 1, \cdots, n-1\} ;\ f^j(x) \in A \right\}.$$

Observe que isto é o mesmo que

$$\tau_n(A, x) = \frac{1}{n} \sum_{j=0}^{n-1} \chi_A \left(f^j(x) \right)$$

onde χ_A designa a função característica do conjunto A. Isto é,

$$\chi_A(x) = \begin{cases} 1, & \text{se } x \in A \\ 0, & \text{se } x \notin A \end{cases}$$

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

$$\tau(A, x) = \lim_{n \to \infty} \tau_n(A, x) = \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_A \left(f^j(x) \right)$$

Em geral, este limite pode não existir. No entanto, o **teorema ergódico** afirma que, relativamente a qualquer probabilidade invariante, o limite realmente existe para quase todo ponto.

Exemplo 1.28. Considerando $A = [0,3,\ 0,4)$ e

$$x = 0,335533335555555533333333333333335\ldots.$$

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

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. Então, o tempo médio de permanência da órbita de x em A, da transformação $f : \mathrm{M} \longrightarrow \mathrm{M}$ definida por $f(x) = 10x - [10x]$, não existe.
De fato: Note que se $n_k = \#\{0 \leq j < k;\ \ f^j(x) \in \mathrm{A}\}$, então

$$\begin{aligned} &n_1 = 1 \\ &n_2 = n_3 = 2 \\ &n_4 = 2,\ n_5 = 3,\ n_6 = 4,\ n_7 = 5 \\ &n_8 = n_9 = \cdots = n_{15} = 6 \\ &n_{16} = 6,\ n_{17} = 7, \ldots, n_{31} = 21 \\ &n_{32} = n_{33} = \cdots = n_{63} = 22 \end{aligned}$$

Em geral, temos

$$n_{2^{2k-1}} = n_{2^{2k}} = 1 + \sum_{j=1}^{k} 2^{2(j-1)} \quad \text{para } k \geq 1.$$

Logo,

$$\begin{aligned} \tau_{2^{2k-1}}(\mathrm{A}, x) &= \frac{1 + \sum_{j=1}^{k} 2^{2(j-1)}}{2^{2k-1}} = \frac{1}{2^{2k-1}} + \sum_{j=1}^{k} 2^{-2(k-j)-1} \\ &= \frac{1}{2^{2k-1}} + \frac{1}{2} \sum_{j=0}^{k-1} 2^{-2j} \to \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{2}{3}, \end{aligned}$$

enquanto que

$$\begin{aligned} \tau_{2^{2k}}(\mathrm{A}, x) &= \frac{1 + \sum_{j=1}^{k} 2^{2(j-1)}}{2^{2k}} = \frac{1}{2^k} + \sum_{j=1}^{k} 2^{-2(k-j)-2} \\ &= \frac{1}{2^k} + \frac{1}{4} \sum_{j=0}^{k-1} 2^{-2j} \to \frac{1}{4} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{1}{3}. \end{aligned}$$

Portanto, o tempo médio de permanência da órbita de x em A não existe. ■

1.4.2 Enunciados

A seguir enunciaremos duas versões do Teorema Ergódico de Birkhoff, a primeira é o caso particular quando φ = função característica χ_{A} do conjunto A, e a segunda é o caso geral. Provaremos diretamente o caso geral.

unesp

Teorema 1.29. Sejam $f : M \longrightarrow M$ uma transformação mensurável e μ uma probabilidade invariante por f. Dado qualquer conjunto mensurável $A \subset M$, o tempo médio de permanência $\tau(A, x)$ existe para μ-quase todo ponto $x \in M$. Além disso,

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

Teorema 1.30. *Sejam $f : M \longrightarrow M$ uma transformação mensurável e μ uma* probabilidade invariante por f. Dada qualquer função integrável $\varphi : M \longrightarrow \mathbb{R}$, o limite

$$\widetilde{\varphi}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) \tag{1.4.23}$$

existe para μ-quase todo ponto $x \in M$. Além disso, $\widetilde{\varphi}$ é uma função integrável com $\int \widetilde{\varphi}\, d\mu = \int \varphi\, d\mu$ e $\widetilde{\varphi} \circ f = \widetilde{\varphi}$. Finalmente, para qualquer $1 \leq p < \infty$, se $\varphi \in L^p(\mu)$ então $\widetilde{\varphi} \in L^p(\mu)$ e têm-se $\| \widetilde{\varphi}\|_p \leq \| \varphi\|_p$.

Demonstração: Começamos por provar a existência do limite (1.4.23) para μ-quase todo ponto $x \in M$, que é o ingrediente fundamental do teorema. Como toda função pode ser escrita como diferença de duas funções não negativas

$$\varphi(x) = \text{máx}\,\{\varphi(x), 0\} - \text{máx}\,\{-\varphi(x), 0\} \quad \text{para todo } x \in M.$$

Então, sem restrição, podemos supor $\varphi \geq 0$. Para cada $x \in M$, consideremos as funções

$$\overline{\varphi}\,(x) = \limsup_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) \quad \text{e} \quad \underline{\varphi}\,(x) = \liminf_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(f^j(x)\right).$$

Agora, basta mostrar que

$$\int \overline{\varphi}\, d\mu \leq \int \varphi\, d\mu \leq \int \underline{\varphi}\, d\mu, \tag{1.4.24}$$

pois

$$\int \overline{\varphi}\, d\mu \leq \int \varphi\, d\mu \leq \int \underline{\varphi}\, d\mu \Longrightarrow \int \overline{\varphi}\, d\mu \leq \int \underline{\varphi}\, d\mu \Longrightarrow \int (\overline{\varphi} - \underline{\varphi})\, d\mu \leq 0.$$

Como $\overline{\varphi} - \underline{\varphi} \geq 0$, segue-se que $\overline{\varphi} - \underline{\varphi} = 0$ implica $\overline{\varphi} = \underline{\varphi}$ para μ-quase todo ponto $x \in M$, que é precisamente o queremos concluir.

Para provar a desigualdade $\int \overline{\varphi}\, d\mu \leq \int \varphi\, d\mu$ introduzimos

$$\overline{\varphi}_K = \min\{\overline{\varphi},\ K\},$$

onde $K \in \mathbb{N}$ é um inteiro fixado suficientemente grande. Também fixamos $\varepsilon > 0$ suficientemente pequeno e, para cada $x \in M$, definimos

$$t(x) = \min\left\{ n \geq 1;\ \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) \geq \overline{\varphi}_K(x) - \varepsilon \right\}. \tag{1.4.25}$$

unesp

Note que, por definição de lim sup e porque $\overline{\varphi} \geq \overline{\varphi}_K$, $t(x)$ sempre existe. Note ainda que (1.4.25) implica

$$\sum_{j=0}^{t(x)-1} \varphi\left(f^j(x)\right) \geq t(x)\left(\overline{\varphi}_K(x) - \varepsilon\right). \tag{1.4.26}$$

Finalmente, escolhemos $T \geq 1$ suficientemente grande tal que o conjunto

$$A = \{x \in M;\ t(x) > T\} \quad \text{tenha} \quad \mu(A) \leq \frac{\varepsilon}{K}.$$

Note que isto é sempre possível, pois a família de conjuntos

$$A_n = \{x \in M;\ t(x) = n\}$$

forma uma partição do espaço M. Logo, tem-se

$$\sum_{n=1}^{\infty} \mu(A_n) = \mu(M) = 1.$$

Portanto, tomando $T \geq 1$ suficientemente grande, obtemos

$$\mu(A) = \sum_{n>T} \mu(A_n) \leq \frac{\varepsilon}{K}. \tag{1.4.27}$$

Agora, para cada $x \in M$ e $n \geq 1$. definimos uma seqüência $x_0, x_1, \ldots, x_s$ de pontos em M e uma seqüência $t_0, t_1, \ldots, t_s$ de números inteiros positivos, do seguinte modo:

1. Tomamos $x_0 = x$.
2. Suponhamos que x_i já foi definido. Para as definições de t_i e de x_{i+1} temos duas possibilidades:
   (a) se $t(x_i) \leq T$ então tomamos $t_i = t(x_i)$ e $x_{i+1} = f^{t_i}(x_i)$.
   (b) se $t(x_i) > T$ então tomamos $t_i = 1$ e $x_{i+1} = f(x_i)$.
3. Terminamos quando encontramos x_s tal que

$$t_0 + t_1 + \cdots + t_{s-1} + t_s \geq n. \tag{1.4.28}$$

Usando $t_i = t(x_i) \leq T$, a desigualdade (1.4.28) pode ser escrita como

$$t_0 + t_1 + \cdots + t_{s-1} \geq n - t_s \geq n - T. \tag{1.4.29}$$

Note que, no caso $(2a)$, escreve-se a estimativa (1.4.26) como

$$\sum_{j=0}^{t_i-1} \varphi\left(f^j(x_i)\right) \geq t_i\left(\overline{\varphi}_K(x_i) - \varepsilon\right).$$



É claro que esta desigualdade implica a seguinte:

$$\sum_{j=0}^{t_i-1} \varphi\left(f^j(x_i)\right) \geq t_i\left(\overline{\varphi}_K(x_i) - \varepsilon\right) - \sum_{j=0}^{t_i-1} \mathrm{K}\,\chi_{\mathrm{A}}\left(f^j(x_i)\right)$$

Portanto,

$$\sum_{j=0}^{t_i-1} \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right)\left(f^j(x_i)\right) \geq t_i\left(\overline{\varphi}_K(x_i) - \varepsilon\right) = t_i\left(\overline{\varphi}_K(x) - \varepsilon\right) \qquad (1.4.30)$$

onde a última igualdade segue de que $\overline{\varphi}_K \circ f(y) = \overline{\varphi}_K(y)$ para todo y (ver abaixo prova de afirmação análoga para $\widetilde{\varphi}$).

Observe que, uma vez que a medida μ é invariante por f, tem-se

$$\int \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right) \circ f^j \, d\mu = \int \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right) d\mu \quad \text{para todo } j.$$

No caso $(2b)$, temos

$$\begin{aligned}\sum_{j=0}^{t_i-1} \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right)\left(f^j(x_i)\right) &= \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right)(x_i) = \varphi(x_i) + \mathrm{K}\,\chi_{\mathrm{A}}(x_i) \\ &= \varphi(x_i) + \mathrm{K} \geq \mathrm{K} \geq t_i\left(\overline{\varphi}_K(x) - \varepsilon\right). \qquad (1.4.31)\end{aligned}$$

Então, usando (1.4.30), (1.4.31), (1.4.29) e escrevendo $\tau = t_0 + t_1 + \cdots + t_{s-1}$ para simplificar a notação , obtemos

$$\begin{aligned}\sum_{j=0}^{n-1} \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right)\left(f^j(x)\right) &= \sum_{i=0}^{s-1} \left(\sum_{j=0}^{t_i-1} \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right)\left(f^j(x_i)\right)\right) \\ &\quad + \sum_{j=\tau}^{n-1} \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right)\left(f^j(x)\right) \geq \left(\sum_{i=0}^{s-1} t_i\right)\left(\overline{\varphi}_K(x) - \varepsilon\right) \\ &\quad + 0 \geq (n - \mathrm{T})\left(\overline{\varphi}_K(x) - \varepsilon\right). \qquad (1.4.32)\end{aligned}$$

Integrando o primeiro e último membros, como μ é invariante por f, obtemos

$$\begin{aligned}&\int \sum_{j=0}^{n-1} \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right)\left(f^j\right) d\mu \;\geq \int (n - \mathrm{T})\left(\overline{\varphi}_K - \varepsilon\right) d\mu \\ \Longrightarrow &\sum_{j=0}^{n-1} \int \left(\varphi + \mathrm{K}\chi_{\mathrm{A}}\right)\left(f^j\right) d\mu \geq (n - \mathrm{T})\left(\int \overline{\varphi}_K d\mu - \varepsilon\right) \\ \Longrightarrow &\sum_{j=0}^{n-1} \int \left(\varphi + \mathrm{K}\chi_{\mathrm{A}}\right) d\mu \geq (n - \mathrm{T})\left(\int \overline{\varphi}_K d\mu - \varepsilon\right) \\ \Longrightarrow &\, n \int \left(\varphi + \mathrm{K}\,\chi_{\mathrm{A}}\right) d\mu \geq (n - \mathrm{T})\left(\int \overline{\varphi}_K \, d\mu - \varepsilon\right).\end{aligned}$$

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Dividindo ambos os termos por n, obtemos

$$\int (\varphi + \mathrm{K}\ \chi_{\mathrm{A}})\, d\mu \geq (1 - \frac{\mathrm{T}}{n}) \left(\int \overline{\varphi}_{\kappa} d\mu - \varepsilon \right).$$

Fazendo $n \to \infty$ e usando (1.4.27), temos

$$\int (\varphi + \mathrm{K}\ \chi_{\mathrm{A}})\, d\mu \geq \int \overline{\varphi}_{\kappa} d\mu - \varepsilon \Longrightarrow \int \varphi\, d\mu + \mathrm{K}\ \mu(\mathrm{A}) \geq \int \overline{\varphi}_{\kappa} d\mu - \varepsilon$$

o que implica $\int \varphi\, d\mu \geq \int \overline{\varphi}_{\kappa} d\mu - 2\varepsilon$ para todo $\varepsilon > 0$. Portanto, como ε é arbitrário, segue-se que

$$\int \varphi\, d\mu \geq \int \overline{\varphi} d\mu.$$

A igualdade $\int \varphi\, d\mu \leq \int \underline{\varphi}\, d\mu$ é provada de modo semelhante, apenas indicamos as diferenças com relação a desigualdade anterior. Consideremos

$$t(x) = \min \left\{ n \geq 1;\ \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) \leq \underline{\varphi}(x) + \varepsilon \right\}. \tag{1.4.33}$$

Note que, por definição de lim inf, $\mathrm{t}(x)$ sempre existe. Note ainda que (1.4.33) implica

$$\sum_{j=0}^{t(x)-1} \varphi\left(f^j(x)\right) \leq \mathrm{t}(x)\left(\underline{\varphi}(x) + \varepsilon\right). \tag{1.4.34}$$

Finalmente, escolhemos $\mathrm{T} \geq 1$ fixado suficientemente grande de tal modo que o conjunto

$$\mathrm{A} = \{x \in \mathrm{M};\ \mathrm{t}(x) > T\} \quad \text{tenha} \quad \int_{\mathrm{A}} \varphi\, d\mu \leq \varepsilon. \tag{1.4.35}$$

Agora, para cada $x \in \mathrm{M}$ e $n \geq 1$, definimos x_i e t_i do mesmo modo que antes. Note que, no caso $(2a)$, a estimativa (1.4.34) continua válida:

$$\sum_{j=0}^{t_i-1} \varphi\left(f^j(x_i)\right) \leq t_i\left(\underline{\varphi}(x_i) + \varepsilon\right).$$

É claro que esta desigualdade implica a seguinte:

$$\sum_{j=0}^{t_i-1} \varphi\left(f^j(x_i)\right) \leq t_i\left(\underline{\varphi}(x_i) + \varepsilon\right) + \sum_{j=0}^{t_i-1} \varphi\chi_{\mathrm{A}}\left(f^j(x_i)\right).$$

Portanto,

$$\sum_{j=0}^{t_i-1} \left(\varphi - \varphi\chi_{\mathrm{A}}\right)\left(f^j(x_i)\right) \leq t_i\left(\underline{\varphi}(x_i) + \varepsilon\right) = t_i\left(\underline{\varphi}(x) + \varepsilon\right) \tag{1.4.36}$$

unesp

onde a última igualdade segue de que $\underline{\varphi} \circ f(y) = \underline{\varphi}(y)$ para todo y (ver abaixo prova de afirmação análoga para $\widetilde{\varphi}$).

No caso $(2b)$, temos

$$\sum_{j=0}^{t_i-1} (\varphi - \varphi\chi_{_A})\,(f^j(x_i)) = (\varphi - \varphi\chi_{_A})\,(x_i) = \varphi(x_i) - \varphi\chi_{_A}(x_i)$$
$$= \varphi(x_i) - \varphi \leq \varphi \leq t_i\,(\underline{\varphi}(x) + \varepsilon)\,. \qquad (1.4.37)$$

Então, usando (1.4.36), (1.4.37) e (1.4.29), obtemos

$$\sum_{j=0}^{n-1} (\varphi - \varphi\chi_{_A})\,(f^j(x)) = \sum_{i=0}^{s-1} \left(\sum_{j=0}^{t_i-1} (\varphi - \varphi\chi_{_A})\,(f^j(x_i)) \right)$$
$$+ \sum_{j=\tau}^{n-1} (\varphi + \varphi\chi_{_A})\,(f^j(x)) \leq \left(\sum_{i=0}^{s-1} t_i \right) (\,\underline{\varphi}(x) + \varepsilon)$$
$$+ \sum_{j=\tau}^{n-1} \varphi\,(f^j(x)) \leq (n - \mathrm{T})\,(\underline{\varphi}(x) + \varepsilon) + \sum_{j=\tau}^{n-1} \varphi\,(f^j(x))\,.$$

Integrando o primeiro e último membros, como a medida μ é invariante por f e recorde por (1.4.29) que $n - \tau \leq \mathrm{T}$, obtemos

$$\int \sum_{j=0}^{n-1} (\varphi - \varphi\chi_{_A})\,(f^j)\,d\mu \leq \int (n - \mathrm{T})\,(\underline{\varphi} + \varepsilon)\,d\mu + \int \sum_{j=\tau}^{n-1} \varphi(f^j)\,d\mu$$
$$\Longrightarrow \sum_{j=0}^{n-1} \int (\varphi - \varphi\chi_{_A})\,(f^j)\,d\mu \leq (n - \mathrm{T}) \left(\int \underline{\varphi}\,d\mu + \varepsilon \right) + \sum_{j=\tau}^{n-1} \int \varphi(f^j)\,d\mu$$
$$\Longrightarrow \sum_{j=0}^{n-1} \int (\varphi - \varphi\chi_{_A})\,(f^j)\,d\mu \leq (n - \mathrm{T}) \left(\int \underline{\varphi}\,d\mu + \varepsilon \right) + \sum_{j=\tau}^{n-1} \int \varphi\,d\mu$$
$$\Longrightarrow n \int \varphi\,d\mu - n \int_{_A} \varphi\,d\mu \leq n - \mathrm{T}) \left(\int \underline{\varphi}\,d\mu + \varepsilon \right) + \mathrm{T} \int \varphi\,d\mu$$

Dividindo ambos os termos por n, obtemos

$$\int \varphi\,d\mu - \int_{_A} \varphi\,d\mu \leq (1 - \frac{\mathrm{T}}{n}) \left(\int \underline{\varphi}\,d\mu + \varepsilon \right) + \frac{\mathrm{T}}{n} \int \varphi\,d\mu.$$

Fazendo $n \longrightarrow \infty$ e usando (1.4.35), temos

$$\int \varphi\,d\mu - \varepsilon \leq \int \underline{\varphi}\,d\mu + \varepsilon \Rightarrow \int \varphi\,d\mu \leq \int \underline{\varphi}\,d\mu + 2\varepsilon \ \text{ para todo } \varepsilon > 0.$$

Portanto, como ε é arbitrário, segue-se que

$$\int \varphi\,d\mu \leq \int \underline{\varphi}\,d\mu.$$

unesp

Isto termina a prova da existência do limite (1.4.23).

Note que $\sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) = \varphi(x) + \varphi(f(x)) + \cdots + \varphi\left(f^{n-1}(x)\right)$. Portanto,

$$\sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) - \varphi(x) = \varphi(f(x)) + \cdots + \varphi\left(f^{n-1}(x)\right). \tag{1.4.38}$$

Usando o limite (1.4.23) e a igualdade (1.4.38), obtemos

$$\begin{aligned}
\widetilde{\varphi} \circ f(x) &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(f^{j+1}(x)\right) \\
&= \lim_{n\to\infty} \frac{1}{n} \left(\varphi\left(f(x)\right) + \cdots + \varphi\left(f^{n-1}(x)\right) + \varphi\left(f^n(x)\right)\right) \\
&= \lim_{n\to\infty} \frac{1}{n} \left(\sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) - \varphi(x) + \varphi\left(f^n(x)\right)\right) \\
&= \lim_{n\to\infty} \left(\frac{1}{n} \sum_{j=0}^{n} \varphi\left(f^j(x)\right) - \frac{1}{n}\varphi(x)\right)
\end{aligned}$$

Note que $\lim_{n\to\infty} \frac{1}{n}\, \varphi(x) = 0$, pois φ é limitada. Portanto,

$$\begin{aligned}
\widetilde{\varphi} \circ f(x) &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n} \varphi\left(f^j(x)\right) = \lim_{n\to\infty} \frac{n+1}{n} \cdot \frac{1}{n+1} \sum_{j=0}^{n} \varphi\left(f^j(x)\right) \\
&= \lim_{n\to\infty} \frac{1}{n+1} \sum_{j=0}^{n} \varphi\left(f^j(x)\right) = \widetilde{\varphi}(x).
\end{aligned}$$

Isto prova que $\widetilde{\varphi} \circ f = \widetilde{\varphi}$ para μ-quase todo ponto $x \in \mathrm{M}$.

Note que, usando o limite (1.4.23), o Teorema da Convergência Monótona e o Teorema A.15, obtemos

$$\begin{aligned}
\int \widetilde{\varphi}\, d\mu &= \int \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) d\mu = \lim_{n\to\infty} \int \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) d\mu \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int \varphi\left(f^j(x)\right) d\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int \varphi\, d\mu \\
&= \lim_{n\to\infty} \frac{1}{n} \cdot n \int \varphi\, d\mu = \lim_{n\to\infty} \int \varphi\, d\mu = \int \varphi\, d\mu.
\end{aligned}$$

Isto prova que $\int \widetilde{\varphi}\, d\mu = \int \varphi\, d\mu$ para μ-quase todo ponto $x \in \mathrm{M}$.

Note que $|\, \widetilde{\varphi}(x)| \leq \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \left|\, \varphi\left(f^j(x)\right)\right|$. Então,

$$|\, \widetilde{\varphi}(x)|^p \leq \lim_{n\to\infty} \left(\frac{1}{n} \sum_{j=0}^{n-1} \left|\, \varphi\left(f^j(x)\right)\right|\right)^p. \tag{1.4.39}$$

unesp

Como cada $\frac{1}{n}\sum_{j=0}^{n-1}\varphi\left(f^j(x)\right)$ é uma função mensurável e todas elas são limitadas inferiormente pela função identicamente nula, que é integrável pelo Teorema A.12, então a seqüencia de funções $\frac{1}{n}\sum_{j=0}^{n-1}\varphi\left(f^j(x)\right)$ satisfaz as hipóteses do Lema de Fatou. Sendo $|\ \widetilde{\varphi}(x)|^p$ uma função positiva, para provar que é integrável, basta mostrar que o limite em (1.4.39) define uma função integrável, o que, pelo Lema de Fatou, se reduz a mostrar que

$$\liminf_{n\to\infty}\int\left(\frac{1}{n}\sum_{j=0}^{n-1}|\ \varphi\left(f^j(x)\right)|\right)^p d\mu<\infty.$$

Usando $\widetilde{\varphi}\circ f=\widetilde{\varphi}$, observe que

$$\int\left(\frac{1}{n}\sum_{j=0}^{n-1}|\varphi\left(f^j(x)\right)|\right)^p d\mu=\left\|\frac{1}{n}\sum_{j=0}^{n-1}|\ \varphi\left(f^j(x)\right)|\right\|^p\leq\left(\frac{1}{n}\sum_{j=0}^{n-1}\left\|\varphi\left(f^j(x)\right)\right\|_p\right)^p$$
$$=\left(\frac{1}{n}\sum_{j=0}^{n-1}\|\ \varphi\|_p\right)^p=\|\ \varphi\|_p^p.$$

Portanto, o limite em (1.4.39) define uma função integrável e usando o Teorema da Convergência Monótona, obtemos

$$\int\lim_{n\to\infty}\left(\frac{1}{n}\sum_{j=0}^{n-1}|\ \varphi\left(f^j(x)\right)|\right)^p d\mu\leq\|\ \varphi\|_p^p.$$

Isto mostra que $|\ \widetilde{\varphi}|^p$ é integrável o que implica $\widetilde{\varphi}\in L^p(\mu)$. Finalmente, dado qualquer $1\leq p<\infty$, obtemos

$$\|\ \widetilde{\varphi}\|_p^p=\int|\widetilde{\varphi}|^p\,d\mu\leq\int\lim_{n\to\infty}\left(\frac{1}{n}\sum_{j=0}^{n-1}|\ \varphi\left(f^j(x)\right)|\right)^p d\mu\leq\|\ \varphi\|_p^p.$$

Isto prova que $\|\ \widetilde{\varphi}\|_p\leq\|\ \varphi\|_p$. A prova do teorema está completa. ■

1.5 Ergodicidade

Uma transformação $f:M\longrightarrow M$ diz-se **ergódica** para uma probabilidade invariante μ (dizemos, também, 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 1.30 coincidem para μ-quase todo ponto $x\in M$ com as respectivas **médias espaciais**:

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

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

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

Um conjunto mensurável $\mathrm{A} \subset \mathrm{M}$ diz-se **invariante** por f se $f^{-1}(\mathrm{A}) = \mathrm{A}$. Dizemos que uma função mensurável $\psi : \mathrm{M} \longrightarrow \mathbb{R}$ é **invariante** por f se $\psi \circ f = \psi$.

Proposição 1.31. Sejam $f : \mathrm{M} \longrightarrow \mathrm{M}$ uma transformação mensurável e μ uma probabilidade invariante por f. As seguintes condições são equivalentes:

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

Demonstração: (1) implica (2): Considere $\varphi = \chi_{\mathrm{A}}$. Como o sistema (f, μ) é ergódico, então

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

para toda função μ-integrável $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$ e μ-quase todo ponto $x \in \mathrm{M}$. Aplicando o Teorema de Birkhoff 1.30, obtemos

$$\widetilde{\varphi}\,(x) = \int \varphi \, d\mu = \int \chi_{\mathrm{A}} \, d\mu = \int_{\mathrm{A}} d\mu = \mu(\mathrm{A}) \tag{1.5.40}$$

para μ-quase todo ponto $x \in \mathrm{M}$.

Como $\mathrm{A} \subset \mathrm{M}$ é invariante por f, então

$$x \in \mathrm{A} \Longleftrightarrow f(x) \in \mathrm{A} \Longrightarrow \varphi\left(f^j(x)\right) = \varphi(x) \tag{1.5.41}$$

para todo $j \geq 0$ e para todo $x \in \mathrm{M}$. Usando (1.5.41), note que

$$\widetilde{\varphi}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\left(f^j(x)\right) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(x) = \varphi(x) = \chi_{\mathrm{A}}(x). \tag{1.5.42}$$

Logo, $\mu(\mathrm{A}) = \widetilde{\varphi}\,(x) = \varphi(x) = \chi_{\mathrm{A}}(x)$. Portanto, como $\chi_{\mathrm{A}}(x)$ só toma valores 0 e 1, temos que $\mu(\mathrm{A}) = 0$ ou $\mu(\mathrm{A}) = 1$.

(2) implica (3): Seja $\psi : \mathrm{M} \longrightarrow \mathbb{R}$ uma função invariante por f qualquer. Então, $\psi \circ f(\mathrm{I}) = \psi(\mathrm{I})$ implica $f^{-1}(\psi^{-1}(\mathrm{I})) = \psi^{-1}(\mathrm{I})$ para qualquer intervalo $\mathrm{I} \subset \mathbb{R}$. Isto mostra que a pré-imagem $\psi^{-1}(I)$ é um conjunto invariante por f. Portanto, pela hipótese (2), $\mu(\psi^{-1}(I)) = 0$ ou $\mu(\psi^{-1}(I)) = 1$. Como I é um intervalo qualquer, temos que $\mu(\psi^{-1}(I)) = 1$ e $\mu((\psi^{-1}(I))^c) = 0$, isto prova que ψ é constante num conjunto com medida μ total.

(3) implica (1): Seja $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$ uma função integrável qualquer. Pelo Teorema de Birkhoff 1.30, a média temporal $\widetilde{\varphi}$ é uma função invariante por f.



Logo, pela hipótese (3), $\widetilde{\varphi}$ é constante em μ-quase todo ponto $x \in M$. Então, usando o Teorema de Birkhoff 1.30,

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

para μ-quase todo ponto $x \in M$. Portanto, o sistema é ergódico. ∎

Proposição 1.32. Sejam $f : M \longrightarrow M$ uma transformação mensurável e μ uma probabilidade invariante por f. As seguintes condições são equivalentes:

1. O sistema (f, μ) é ergódico.
2. Se $A, B \subset M$ então $\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu\left(f^{-j}(A) \cap B\right) = \mu(A)\mu(B)$.

Demonstração: (1) implica (2): Suponha que o sistema (f, μ) é ergódico. Aplicando o Teorema de Birkhoff 1.30, obtemos

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

onde o limite é tomado no sentido da convergência em $L^2(\mu)$. Então, usando o Teorema da Convergência Monótona, obtemos

$$\begin{aligned}
\mu(A)\mu(B) &= \int \mu(A)\chi_B\, d\mu = \int \left(\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_A \circ f^j \right) \chi_B\, d\mu \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int \left(\chi_A \circ f^j\right) \chi_B\, d\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int \chi_{f^{-j}(A)} \chi_B\, d\mu \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu\left(f^{-j}(A) \cap B\right).
\end{aligned}$$

Para a terceira igualdade na seqüência acima, note que χ_B é uma função real e o produto interno em $L^2(\mu)$ é contínuo.
(1) implica (2): Suponha que vale a segunda propriedade e seja $A \subset M$ tal que $f^{-1}(A) = A$. Então,

$$\mu(A)\mu(A) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu\left(f^{-j}(A) \cap A\right) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu(A) = \mu(A).$$

Logo, $\mu(A)^2 = \mu(A)$, o que implica $\mu(A) = 0$ ou $\mu(A) = 1$. ∎

O exemplo a seguir mostra uma interpretação do limite na segunda condição da proposição 1.32

unesp

Exemplo 1.33. Suponhamos que temos em um recipiente M com água uma mancha de tinta vermelha numa certa região A e uma mancha de tinta branca numa outra região B. Consideremos em M uma transformação f que consiste em agitar o líquido no recipiente. Se $\mu(\mathrm{B}) > 0$, dizer que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mu\left(f^{-j}(\mathrm{A}) \cap \mathrm{B}\right) = \mu(\mathrm{A})\mu(\mathrm{B})$$

equivale a dizer que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \frac{\mu\left(f^{-j}(\mathrm{A}) \cap \mathrm{B}\right)}{\mu(\mathrm{B})} = \mu(\mathrm{A}). \tag{1.5.43}$$

O quociente

$$\frac{\mu\left(f^{-j}(\mathrm{A}) \cap \mathrm{B}\right)}{\mu(\mathrm{B})}$$

representa a proporção de tinta vermelha existente na região B depois de agitar j vezes. A igualdade (1.5.43) significa que, em média, a proporção de tinta vermelha na região B se aproxima da proporção que a região A ocupa no recipiente M.

1.6 Exemplos e aplicações

Nesta seção, descreveremos diversos exemplos de sistemas ergódicos. Consideremos $\mathrm{M} = \mathbb{R}^d$ com $d \geq 1$ e m = medida de Lebesgue (volume usual).

1.6.1 Expansão decimal

Considere a transformação $f : [0,1] \longrightarrow [0,1]$ definida por

$$f(x) = 10x - [10x], \text{ para todo } x \in [0,1].$$

Afirmaçao: f é ergódica para a medida de Lebesgue m.
De fato: Mostramos na Subseção 1.2.1 que a medida de Lebesgue m restrita ao intervalo $[0,1]$ é invariante por f. Tendo em vista a Proposição 1.31, só resta provar que se $\mathrm{A} \subset [0,1]$ é um conjunto invariante por f com medida positiva então A tem medida total.

Suponha que A é invariante por f e $m(\mathrm{A}) > 0$. Então, pelo Teorema A.24, m-quase todo ponto $p \in \mathrm{A}$ é ponto de densidade de A. Como estamos tratando com subconjuntos de $\mathbb{R}$, a condição (A.8.2) torna-se

$$\liminf_{\varepsilon\to 0} \left\{ \frac{m(\mathrm{I} \cap \mathrm{A})}{m(\mathrm{I})};\ \mathrm{I} \subset (\mathrm{p}-\varepsilon, \mathrm{p}+\varepsilon) \text{ intervalo contendo } p \right\} = 1. \tag{1.6.44}$$

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Fixemos um ponto de densidade p ∈ A qualquer e considere a seqüência de intervalos

$$I_k = \left[\frac{n_k}{10^k}, \frac{n_k+1}{10^k}\right)$$

que contém o ponto p, onde k ∈ ℕ e $n_k \in \mathbb{Z}$. A propriedade (1.6.44) implica que

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

Note que cada f^k é uma **bijeção afim** de I_k sobre o intervalo $[0,1)$. Isso tem a seguinte conseqüência, que é fundamental para o nosso argumento.

$$\frac{m\left(f^k(A_1)\right)}{m\left(f^k(A_2)\right)} = \frac{m(A_1)}{m(A_2)} \tag{1.6.45}$$

para quaisquer subconjuntos mensuráveis A_1 e A_2 de I_k. Aplicando este fato a $A_1 = I_k \cap A$ e $A_2 = I_k$, obtemos

$$\frac{m\left(f^k(I_k \cap A)\right)}{m\left(f^k(I_k)\right)} = \frac{m\left(f^k(I_k \cap A)\right)}{m(0,1)} = \frac{m(I_k \cap A)}{m(I_k)}.$$

É claro que $m(0,1) = 1$. Além disso, como estamos supondo que A é invariante por f então $f^k(I_k \cap A)$ está contido em A. Deste modo, obtemos

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

Como $\dfrac{m(I_k \cap A)}{m(I_k)}$ converge para 1, segue que $m(A) = 1$. ∎

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 de $\dfrac{1}{10}$ na sua expansão decimal.

Observação 1.34.

- Em geral, é muito difícil decidir se um dado número irracional é balanceado ou não.
- Não é sabido até hoje se o número π é balanceado.
- Quase todo número é balanceado.

Proposição 1.35. O conjunto dos números $x \in \mathbb{R}$ não balanceados tem medida de Lebesgue nula.

Demonstração: Como o fato de ser balanceado ou não é independente da parte inteira do número, então só precisamos mostrar que quase todo ponto $x \in [0,1]$ é balanceado. Considere a transformação $f : [0,1] \longrightarrow [0,1]$ definida por $f(x) = 10x - [10x]$. Para cada dígito $j \in \{0,1,\ldots,9\}$ considere o intervalo $\mathrm{A}_j = \left[\frac{j}{10}, \frac{j+1}{10}\right)$. Recorde que se $x = 0, a_0 a_1 \ldots a_k a_{k+1}$ então $f^k(x) = 0, a_k a_{k+1} \ldots$. Portanto, $f^k(x) \in \mathrm{A}_j$ se, e somente se, o k-ésimo dígito da expansão decimal de x é igual a j. Conseqüentemente, o tempo médio de permanência $\tau(\mathrm{A}_j, x)$ é exatamente a freqüência do dígito j na expansão decimal de x. Usando o Teorema de Birkhoff 1.29 e o fato de que a transformação f é ergódica para a medida de Lebesgue m, concluímos que para cada $j \in \{0,1,\ldots,9\}$ existe um subconjunto B_j de M com $m(\mathrm{B}_j) = 1$ tal que

$$\tau(\mathrm{A}_j, x) = m(\mathrm{A}_j) = \frac{1}{10} \qquad \text{para todo } x \in \mathrm{B}_j.$$

Então $\mathrm{B} = \mathrm{B}_0 \cap \mathrm{B}_1 \cap \cdots \cap \mathrm{B}_9$ também tem $m(\mathrm{B}) = 1$.

De fato:

$$m\left(\bigcap_{j=0}^{9} \mathrm{B}_j\right) = m\left(\bigcup_{j=0}^{9} (\mathrm{B}_j^c)^c\right) = 1 - m\left(\bigcup_{j=0}^{9} \mathrm{B}_j^c\right) \geq 1 - \sum_{j=0}^{9} m\left(\mathrm{B}_j^c\right) = 1,$$

pois $\sum_{j=0}^{9} m\left(\mathrm{B}_j^c\right) = 0$. Portanto, $m\left(\bigcap_{j=0}^{9} \mathrm{B}_j\right) = m(\mathrm{B}) = 1$. Logo, todo número $x \in \mathrm{B}$ é balanceado. ∎

1.6.2 Deslocamentos (shifts) de Bernoulli

Estes sistemas **modelam seqüências 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 $1, 2, \ldots, d$, com probabilidade $p(1), p(2), \ldots, p(d)$ de ocorrerem, sendo

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

O conjunto M das seqüências $\underline{\alpha} = (\alpha_n)_{n\in\mathbf{Z}}$ com cada $\alpha_n \in \{0, 1, \ldots, d\}$ contém os possíveis resultados da seqüência de experimentos. Chamam-se **cilindros** os subconjuntos da forma

$$[k, \ell;\ a_k, \ldots, a_\ell] = \{\underline{\alpha} \in \mathrm{M};\ \alpha_k = a_k, \ldots, \alpha_\ell = a_\ell\},$$

onde $k, \ell \in \mathbb{Z}$ com $k \leq \ell$ e cada $a_j \in \{1, 2, \ldots, d\}$. Definimos

$$\mu([k, \ell;\ a_k, \ldots, a_\ell]) = p(a_k) \ldots p(a_\ell). \tag{1.6.46}$$

Heuristicamente, isto significa que a probabilidade do evento composto

$$\alpha_k = a_k,\ \ \alpha_{k+1} = a_{k+1},\ \ \ldots,\ \ \alpha_\ell = a_\ell$$

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

Consideremos 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 $\mu(\mathrm{M}) = 1$). Estendemos μ_0 de modo que seja finitamente aditiva: se $\mathrm{A} \in \mathcal{B}_0$ é a união disjunta de cilindros $\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_\mathrm{N}$, definimos

$$\mu(\mathrm{A}) = \mu(\mathrm{C}_1) + \mu(\mathrm{C}_2) + \cdots + \mu(\mathrm{C}_\mathrm{N}).$$

Verifica-se que esta função μ é, de fato, σ-aditiva em $\mathcal{B}_0$. Por exemplo, isso pode ser feito usando a Proposição A.5. 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 $p(1), p(2), \ldots, p(d)$. Para não complicar desnecessariamente a notação, representamos também por μ.

No espaço M consideremos a **transformação deslocamento (shift)** à esquerda $f : \mathrm{M} \longrightarrow \mathrm{M}$ definida por

$$f\left((\alpha_n)_{n\in\mathbf{Z}}\right) = (\alpha_{n+1})_{n\in\mathbf{Z}}$$

que corresponde a fazer uma translação no tempo.

Observação 1.36. A medida de Bernoulli é invariante por f.
De fato: Se $\mathrm{A} = [k, \ell;\ a_k, \ldots, a_\ell]$, então $f^{-1}(\mathrm{A}) = [k+1, \ell+1;\ a_k, \ldots, a_\ell]$. Usando (1.6.46), obtemos, neste caso

$$\mu(\mathrm{A}) = \mu\left(f^{-1}(\mathrm{A})\right).$$

Como a família dos cilindros gera a σ-álgebra $\mathcal{B}$, isto juntamente com o Lema 1.7, prova que a medida μ é invariante por f. ∎

Lema 1.37. Se A e B são elementos da álgebra $\mathcal{B}_0$, isto é, uniões finitas de cilindros disjuntos, então para n suficientemente grande tem-se

$$\mu\left(\mathrm{A} \cap f^{-n}(\mathrm{B})\right) = \mu(\mathrm{A})\mu(f^{-n}(\mathrm{B})) = \mu(\mathrm{A})\mu(\mathrm{B}).$$

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

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

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

onde a união é sobre todos os valores possíveis de $c_{\ell+1}, \ldots, c_{u+n-\ell}$. Usando (1.6.46) e como $\sum p(c_{\ell+1}) \ldots p(c_{u+n-\ell}) = 1$, note que

$$\begin{aligned} \mu\left(A \cap f^{-n}(B)\right) &= \mu\left(\bigcup [k, v+n;\ a_k, \ldots, a_\ell, c_{\ell+1}, \ldots, c_{u+n-\ell}, b_u, \ldots, b_v]\right) \\ &= \sum \mu\left([k, v+n;\ a_k, \ldots, a_\ell, c_{\ell+1}, \ldots, c_{u+n-\ell}, b_u, \ldots, b_v]\right) \\ &= \sum p(a_k) \ldots (a_\ell).p(c_{\ell+1}) \ldots p(c_{u+n-\ell}).p(b_u) \ldots p(b_v) \\ &= \sum \mu(A).p(c_{\ell+1}) \ldots p(c_{u+n-\ell}).\mu(B) \\ &= \mu(A)\mu(B) \sum p(c_{\ell+1}) \ldots p(c_{u+n-\ell}) = \mu(A)\mu(B). \end{aligned}$$

Isto prova o lema quando os conjuntos envolvidos são cilindros. O caso geral segue do fato de μ ser finitamente aditiva. ∎

Proposição 1.38. Sejam $f : M \longrightarrow M$ 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 qualquer invariante por f qualquer. Devemos mostrar que $\mu(A) = 0$ ou $\mu(A) = 1$. Para tornar a idéia da prova mais clara, comecemos por um caso particular: suponha que A está na álgebra $\mathcal{B}_0$ das uniões finitas de cilindros disjuntos dois-a-dois. Nesse caso, podemos aplicar o Lema 1.37 com B = A. Concluímos que

$$\mu\left(A \cap f^{-n}(A)\right) = \mu(A)^2 \tag{1.6.47}$$

sempre que tomemos n suficientemente grande. Como A é invariante por f, $f^{-n}(A) = A$ para todo n. Então, a igualdade (1.6.47) pode ser escrita como $\mu(A) = \mu(A)^2$, 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 por f qualquer. A idéia é aproximar A por elementos da álgebra $\mathcal{B}_0$, usando a Proposição A.6: dado qualquer $\varepsilon > 0$, existe $A_0 \in \mathcal{B}_0$ tal que $\mu(A \triangle A_0) < \varepsilon$. Escolha n suficientemente grande, de modo que

$$\mu\left(A_0 \cap f^{-n}(A_0)\right) = \mu(A_0)\mu(f^{-n}(A_0)) = \mu(A_0)^2. \tag{1.6.48}$$

Note que:

$$(A \cap f^{-n}(A)) \setminus (A_0 \cap f^{-n}(A_0)) \subset (A \setminus A_0) \cup (f^{-n}(A) \setminus f^{-n}(A_0)). \tag{1.6.49}$$



De fato: Dado $x \in (A \cap f^{-n}(A)) \setminus (A_0 \cap f^{-n}(A_0))$, então

$$x \notin A_0 \implies x \in A \setminus A_0$$
e
$$x \in A_0 \implies x \notin f^{-n}(A_0) \implies x \in f^{-n}(A) \setminus f^{-n}(A_0).$$

Portanto, $x \in (A \setminus A_0) \cup (f^{-n}(A) \setminus f^{-n}(A_0))$.
Analogamente,

$$(A_0 \cap f^{-n}(A_0)) \setminus (A \cap f^{-n}(A)) \subset (A_0 \setminus A) \cup (f^{-n}(A_0) \setminus f^{-n}(A)). \quad (1.6.50)$$

De (1.6.49) e (1.6.50), temos

$$(A \cap f^{-n}(A)) \setminus (A_0 \cap f^{-n}(A_0)) \quad \cup \quad (A_0 \cap f^{-n}(A_0)) \setminus (A \cap f^{-n}(A))$$

$$\subset (A \setminus A_0) \cup (f^{-n}(A) \setminus f^{-n}(A_0)) \quad \cup \quad (A_0 \setminus A) \cup (f^{-n}(A_0) \setminus f^{-n}(A)).$$

Portanto,

$$(A \cap f^{-n}(A)) \triangle (A_0 \cap f^{-n}(A_0)) \subset A \triangle A_0 \cup (f^{-n}(A_0) \setminus f^{-n}(A)).$$

$$\begin{aligned} \left|\mu(A \cap f^{-n}(A)) - \mu(A_0 \cap f^{-n}(A_0))\right| &\leq \mu((A \cap f^{-n}(A)) \triangle (A_0 \cap f^{-n}(A_0))) \\ &\leq \mu((A \triangle A) \cup (A_0 \cap f^{-n}(A_0))) \quad (1.6.51) \\ &= \mu(A \triangle A_0) + \mu(f^{-n}(A \triangle A_0)). \\ &\leq 2\mu(A \triangle A_0) \leq 2\varepsilon. \quad (1.6.52) \end{aligned}$$

Além disso,

$$\left|\mu(A)^2 - \mu(A_0)^2\right| \leq 2\left|\mu(A) - \mu(A_0)\right| \leq 2\varepsilon. \quad (1.6.53)$$

Juntando as relações (1.6.48), (1.6.51) e (1.6.53), concluímos que

$$\left|\mu(A) - \mu(A)^2\right| < 4\varepsilon.$$

Como ε é arbitrário, deduzimos que $\mu(A) = \mu(A)^2$ e então, do mesmo modo que antes, concluímos que $\mu(A) = 0$ ou $\mu(A) = 1$. ■

1.6.3 Rotação irracional no círculo

O círculo pode ser visto de duas maneiras: como um subconjunto do plano complexo $S^1 = \{z \in \mathbb{C};\ |z| = 1\}$, que é um grupo sob a operação de multiplicação complexa, ou como o grupo quociente $\frac{\mathbb{R}}{\mathbb{Z}}$, onde a operação de grupo é a adição módulo 1. Consideremos o primeiro caso: dado $\alpha \in \mathbb{R}$, a **rotação** de ângulo α é a multiplicação pelo número complexo $e^{\alpha i}$

$$\begin{aligned} R_\alpha : \ S^1 &\longrightarrow S^1 \\ z &\longmapsto e^{\alpha i} z \end{aligned}$$

É claro que R_α preserva o comprimento dos intervalos (segmento) de S^1. Usando o Lema 1.7 deduz-se que a medida de Lebesgue (comprimento de arco) é invariante por qualquer R_α.

O comportamento dinâmico e ergódico de R_α depende muito da natureza de α, como veremos. Dizemos que a rotação é **irracional** se o número $\frac{\alpha}{2\pi}$ é irracional, e **racional** no caso contrário.

Exemplo 1.39. Se R_α é rotação racional então R_α não é ergódico para a medida de Lebesgue, pois todo ponto $z \in S^1$ é periódico com período k para algum k inteiro positivo. Se I for um intervalo suficientemente pequeno, então os intervalos $\mathrm{I}, f(\mathrm{I}), f^{-1}(\mathrm{I}), \ldots, f^{k-1}(\mathrm{I})$ serão dois-a-dois disjuntos e sua união será um conjunto invariante por f, de medida maior do que zero e menor do que 1.

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

$$\int |\psi|^2 \, dm < \infty.$$

É claro que este espaço contém todas as funções mensuráveis limitadas. Em particular, todas as funções características de conjuntos mensuráveis.
De fato: Suponha que φ é uma função limitada qualquer. Então, existe $\mathrm{K} > 0$ tal que $|\varphi| \leq \mathrm{K}$. Seja $\varphi \in \mathrm{L}^2(m)$, então

$$\|\varphi\|_2 = \int |\psi|^2 \, dm \leq \int \mathrm{K}^2 \, dm = \mathrm{K}^2 \int dm < \infty.$$

Outro fato de que necessitamos é que a família de funções $\{\phi_k(z) = z^k; \quad z \in \mathbb{Z}\}$ seja uma base de Hilbert desse espaço: dada qualquer $\varphi \in \mathrm{L}^2(m)$ existe uma única seqüência $(c_k)_{k\in\mathbb{Z}}$ de números complexos tais que

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

A proposição a seguir será demonstrada usando fatos simples de análise de Fourier.

Proposição 1.40. Se R_α é rotação irracional então R_α é ergódica para a medida de Lebesgue.
Demonstração: Pela Proposição 1.31, basta mostrar que toda função $\varphi \in \mathrm{L}^2(m)$ que é invariante por f é constante em m-quase todo ponto $x \in \mathrm{M}$. Usando a expansão em série 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

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

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

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

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

A hipótese de que a rotação é irracional significa que $c_k(e^{k\alpha i} - 1) \neq 0$ para todo $k \neq 0$. Portanto, $c_k = 0$ para todo $k \neq 0$. Ou seja, $\varphi(z) = c_0$ para m-quase todo $z \in S^1$. ∎

Observação 1.41. 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 $\alpha = (\alpha_1, \cdots, \alpha_d)$ é a aplicação

$$\begin{aligned} R_\alpha : \ \mathbb{T}^d &\longrightarrow \mathbb{T}^d \\ z &\longmapsto (e^{i\alpha_1} z_1, \ldots, e^{i\alpha_d} z_d) \end{aligned}$$

A rotação é irracional se os números $\dfrac{\alpha_j}{2\pi}$ são incomensuráveis:

$$n_0 + n_1 \frac{\alpha_1}{2\pi} + \cdots + n_d \frac{\alpha_d}{2\pi} = 0 \Longrightarrow n_0 = n_1 = \cdots = n_d = 0,$$

quaisquer que sejam os inteiros $n_0, n_1, \ldots, n_d$. Usando uma versão multidimensional das idéias anteriores, prova-se que uma rotação é ergódica se, e somente se, ela é irracional.

1.6.4 Transformação de Gauss

Vimos na subseção 1.2.3 que a transformação $G(x) = \dfrac{1}{x} - \left[\dfrac{1}{x}\right]$ admite uma probabilidade invariante por G que é equivalente à medida de Lebesgue, nomeadamente,

$$m(\mathrm{A}) = \frac{1}{\log 2} \int_{\mathrm{A}} \frac{dx}{1+x}.$$

Temos também que o sistema (G, m) é ergódico. Este fato pode ser demonstrado pelo mesmo tipo de argumento que usamos na Subseçã 1.6.1. 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 m-quase todo ponto $a \in [0, 1]$ existe uma seqüência de intervalos I_k contendo a e tais que f^k envia I_k bijetivamente e diferencialmente sobre $(0, 1)$. O diâmetro desses intervalos converge para zero. Logo, tomando para a um ponto de densidade qualquer de A, temos

$$\frac{m(\mathrm{I}_k \cap \mathrm{A})}{m(\mathrm{I}_k)} \to 1 \quad \text{quando} \quad k \to +\infty. \tag{1.6.54}$$

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 (1.6.45) 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 $k, \mathrm{I}_k, \mathrm{A}_1$ e A_2.

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Lema 1.42. 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

$$\frac{m(f^k(A_1))}{m(f^k(A_2))} \leq K \frac{m(A_1)}{m(A_2)}$$

para quaisquer subconjuntos mensuráveis A_1 e A_2 de I_k.

Não demonstramos este lema, porque um resultado mais geral será apresentado na Proposição 2.8 no Capítulo 2. Explicamos apenas como a ergodicidade de (G, m) pode ser obtida a partir dele. Note que $f^k(I_k \cap A^c) = A^c$, pois o conjunto A é invariante. Recorde também que $f^k(I_k) = (0, 1)$ tem medida total. Tomando $A_1 = I_k \cap A^c$ e $A_2 = I_k$ no Lema 1.42, concluímos que

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

De acordo com (1.6.54), a expressão do lado direito converge para zero quando $k \to +\infty$. Portanto, $m(A^c) = 0$. ∎

1.7 Propriedades de medidas ergódicas

Nesta seção vamos fixar uma transformação $f : M \longrightarrow M$ qualquer.

O lema seguinte afirma que **probabilidades ergódicas** são **minimais** para a relação $\nu \ll \mu$.

Lema 1.43. Se ν e μ são probabilidades invariantes tais que μ é ergódica e ν é absolutamente contínua com relação a μ então $\mu = \nu$.

Demonstração: Sejam $\varphi : M \longrightarrow \mathbb{R}$ uma função mensurável limitada qualquer e

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

a sua média temporal. Como μ é invariante e ergódica por f, pela Proposição 1.31, $\widetilde{\varphi}$ é constante e

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

para μ-quase todo ponto $x \in M$. Tendo em vista que $\nu \ll \mu$ e invariante por f, pela Proposição 1.31, $\widetilde{\varphi}$ é constante e

$$\widetilde{\varphi}(x) = \int \varphi \, d\nu$$

para ν-quase todo ponto $x \in M$. Em particular,

$$\int \widetilde{\varphi} \, d\nu = \int \varphi \, d\mu.$$

Por outro lado, pelo Teorema de Birkhoff 1.30, temos

$$\int \tilde{\varphi}\, d\nu = \int \varphi\, d\nu.$$

Portanto, as integrais de φ com relação a μ e em relação a ν coincidem, qualquer que seja a função mensurável limitada φ.

Agora, consideremos $\varphi = \chi_A$. Note que

- $\int \varphi\, d\mu = \int \chi_A\, d\mu = \int_A d\mu = \mu(A).$
- $\int \varphi\, d\nu = \int \chi_A\, d\nu = \int_A d\nu = \nu(A).$

Como $\int \varphi\, d\mu = \int \varphi\, d\nu$ implica que $\mu(A) = \nu(A)$ para todo conjunto mensurável A. Isto prova que $\mu = \nu$. ∎

Seja C um conjunto convexo num espaço vetorial W. Um ponto $p \in C$ diz-se **ponto extremal** se não é possível escrevê-lo na forma $p = \lambda x - (1-\lambda)y$ com x e y em C, $x \neq y$ e $0 < \lambda < 1$.

O próximo lema afirma que uma probabilidade invariante é **ergódica** se, e somente se, é **ponto extremal** no conjunto das probabilidades invariantes.

Lema 1.44. Uma probabilidade invariante μ é ergódica se, e somente se, não é possível escrevê-la na forma

$$\mu = c_1\mu_1 + c_2\mu_2$$

com c_1, c_2 maiores que zero e μ_1, μ_2 probabilidade invariantes distintas.

Demonstração: Para provar a parte "se", suponha que μ não é ergódica. Então, existe algum conjunto mensurável A invariante por f com $0 < \mu(A) < 1$. Definimos μ_1 e μ_2 como sendo as restrições normalizadas de μ a A e ao seu complementar, respectivamente:

$$\mu_1(B) = \frac{\mu(B \cap A)}{\mu(A)} \quad \text{e} \quad \mu_2(B) = \frac{\mu(B \cap A^c)}{\mu(A^c)}.$$

Como A e A^c são conjuntos mensuráveis invariantes por f e μ é probabilidade invariante por f, μ_1 e μ_2 são também probabilidades invariantes por f.

De fato: Suponha que B é um conjunto invariante por f, então

- $\mu_1(f^{-1}(B)) = \dfrac{\mu(f^{-1}(B) \cap A)}{\mu(A)} = \dfrac{\mu(B \cap A)}{\mu(A)} = \mu_1(B).$
- $\mu_2(f^{-1}(B)) = \dfrac{\mu(f^{-1}(B) \cap A^c)}{\mu(A^c)} = \dfrac{\mu(B \cap A^c)}{\mu(A^c)} = \mu_2(B).$

Além disso,

$$\mu = \mu(\mathrm{A})\mu_1 + \mu(\mathrm{A})\mu_2.$$

Portanto, μ não é ponto extremal.

Para provar a recíproca, suponha que μ é ergódica e temos $\mu = c_1\mu_1 + c_2\mu_2$ com $c_1, c_2 > 0$. Note que $\mu(\mathrm{B}) = 0$ implica que $\mu_1(\mathrm{B}) = \mu_2(\mathrm{B}) = 0$, ou seja, μ_1 e μ_2 são absolutamente contínuas com relação a μ. Logo, pelo Lema 1.43, $\mu_1 = \mu = \mu_2$. Isto prova que μ é ponto extremal. ∎

Lema 1.45. Sejam $\mu_1, \ldots, \mu_N$ probabilidades invariantes e ergódicas, todas distintas. Então, existem subconjuntos mensuráveis $\mathrm{P}_1, \ldots, \mathrm{P}_N$ invariantes distintos tais que

$$\mu_j(\mathrm{P}_k) = \begin{cases} 1, & \text{se } j = k \\ 0, & \text{se } j \neq k \end{cases}$$

Demonstração: Fixe qualquer par j, k de números distintos em $\{1, \ldots, N\}$. Pelo Lema 1.43, a medida μ_j não pode ser absolutamente contínua com relação a μ_k. Em outras palavras, existe algum conjunto mensurável A tal que $\mu_j(\mathrm{A}) > 0$ mas $\mu_k(\mathrm{A}) = 0$. Então, como μ_j e μ_k são invariantes por f, temos

$$\mu_j\left(\bigcup_{j=0}^{\infty} f^{-j}(\mathrm{A})\right) \geq \mu_j(\mathrm{A}) > 0. \tag{1.7.55}$$

$$\mu_k\left(\bigcup_{j=0}^{\infty} f^{-j}(\mathrm{A})\right) \leq \sum_{j=0}^{\infty} \mu_k(f^{-j}(\mathrm{A})) = \sum_{j=0}^{\infty} \mu_k(\mathrm{A}) = 0. \tag{1.7.56}$$

Defina $\mathrm{P}_{j,k} = \bigcap_{n=0}^{\infty} \bigcup_{j=n}^{\infty} f^{-j}(\mathrm{A})$. Como a seqüência de conjuntos na interseção é decrescente com n, então

$$\mu_j(\mathrm{P}_{j,k}) = \lim_{n\to\infty} \mu_j\left(\bigcup_{j=n}^{\infty} f^{-j}(\mathrm{A})\right) \tag{1.7.57}$$

e analogamente, $\mu_k(\mathrm{P}_{j,k}) = \lim_{n\to\infty} \mu_k\left(\bigcup_{j=n}^{\infty} f^{-j}(\mathrm{A})\right)$. Como as medidas μ_j e μ_k são invariantes por f, e

$$\bigcup_{j=n}^{\infty} f^{-j}(\mathrm{A}) = f^{-n}\left(\bigcup_{j=0}^{\infty} f^{-j}(\mathrm{A})\right),$$

a seqüência no lado direito de (1.7.57) é constante. Usando (1.7.55) e (1.7.56), concluímos que

$$\mu_j(\mathrm{P}_{j,k}) = \mu_j\left(\bigcup_{j=0}^{\infty} f^{-j}(\mathrm{A})\right) > 0 \quad \text{e} \quad \mu_k(\mathrm{P}_{j,k}) = \mu_k\left(\bigcup_{j=0}^{\infty} f^{-j}(\mathrm{A})\right) = 0.$$

unesp

Além disso, $P_{j,k}$ é um conjunto invariante por f.
De fato:

$$f^{-1}(P_{j,k}) = f^{-1}\left(\bigcap_{n=0}^{\infty}\bigcup_{j=n}^{\infty} f^{-j}(A)\right) = \bigcap_{n=0}^{\infty} f^{-1}\left(\bigcup_{j=n}^{\infty} f^{-j}(A)\right) = \bigcap_{n=1}^{\infty}\bigcup_{j=n}^{\infty} f^{-j}(A)$$
$$= P_{j,k}.$$

Portanto, como μ é ergódica, pela Proposição 1.32, $\mu_j(P_{j,k}) = 1$. Agora defina

$$\widetilde{P}_j = \bigcap_{k\neq j} P_{j,k} \quad \text{e} \quad P_j = \widetilde{P}_j \setminus \bigcup_{k\neq j} \widetilde{P}_k.$$

Primeiro, $\mu_j(\widetilde{P}_j) = 1$ e $\mu_k(\widetilde{P}_k) = 0$ para todo $k \neq j$. Segue-se que $\mu_j(P_j) = 1$ e $\mu_k(P_k) = 0$. Além disso, os P_j são disjuntos dois-a-dois. ■

Unesp Biociências, Letras e Ciências Exatas - UNESP - S J Rio Preto - BIBLIOTECA

Capítulo 2

Transformações expansoras

Neste capítulo provamos que para qualquer transformação expansora cujo determinante jacobiano detDf é Hölder, existe uma única probabilidade invariante absolutamente contínua com relação à medida de Lebesgue. Essa probabilidade é positiva em todos os subconjuntos abertos de M, é ergódica, e a sua bacia de atração tem medida de Lebesgue total em M.

Definição 2.1. Dado $\nu > 0$, dizemos que $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$ é **ν-Hölder** se existe alguma constante $\mathrm{K} > 0$ tal que

$$|\varphi(x_1) - \varphi(x_2)| \leq \mathrm{K}|\, x_1 - x_2|^{\nu} \text{ para todo } x_1 \text{ e } x_2 \text{ em M.}$$

Definição 2.2. Sejam M um espaço métrico separável, $f : \mathrm{M} \longrightarrow \mathrm{M}$ uma transformação contínua e μ uma probabilidade invariante por f.

1. Chamamos o **suporte** de μ o conjunto $\mathrm{supp}(\mu)$ dos pontos $x \in \mathrm{M}$ tais que toda vizinhança de x tem medida positiva para μ.
2. Chamamos **bacia** de μ o conjunto $\mathrm{B}(\mu)$ dos pontos $x \in \mathrm{M}$ tais que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \int \varphi \, d\mu \text{ para toda função contínua } \varphi : \mathrm{M} \longrightarrow \mathbb{R}.$$

Definição 2.3. Sejam M uma variedade compacta e $f : \mathrm{M} \longrightarrow \mathrm{M}$ uma transformação de classe C^1. Dizemos que que f é **expansora** se existem $\sigma > 1$ e alguma métrica riemanniana $\|\cdot\|$ em M tais que

$$\| \mathrm{D}f(x)v\| \geq \sigma \| v\| \quad \text{para todo } x \in \mathrm{M} \text{ e } v \in \mathrm{T}_x\mathrm{M}. \tag{2.0.1}$$

Exemplo 2.4. Seja $F : \mathbb{R}^d \longrightarrow \mathbb{R}^d$, $d \geq 1$, um **isomorfismo linear** satisfazendo $F(\mathbb{Z}^d) \subset \mathbb{Z}^d$. Então, existe uma única transformação $f : \mathbb{T}^d \longrightarrow \mathbb{T}^d$ no **toro** d-dimensional $\mathrm{M} = \mathbb{T}^d$ tal que $f \circ \pi = \pi \circ F$, onde

$$\pi : \ \mathbb{R}^d \longrightarrow \mathbb{T}^d \text{ definida por } \pi(x_1, \ldots, x_d) = (e^{2\pi i x_1}, \ldots, e^{2\pi i x_d})$$

é a **projeção canônica**. Se todos os autovalores $\lambda_1, \ldots, \lambda_n$ de F têm norma maior que 1, então a transformação f é expansora: podemos tomar qualquer $1 < \sigma < \inf_i |\lambda_i|$ em (2.0.1).

unesp

Exemplo 2.5. Consideremos o conjunto obtido do seguinte modo:

$$\begin{aligned} A_0 &= [0,1] \\ A_1 &= \left[0,\frac{1}{3}\right] \cup \left[\frac{2}{3},1\right] \\ A_2 &= \left[0,\frac{1}{9}\right] \cup \left[\frac{2}{9},\frac{1}{3}\right] \cup \left[\frac{2}{3},\frac{7}{9}\right] \cup \left[\frac{8}{9},1\right] \\ &\vdots \\ A_n &= \cdots \end{aligned}$$

onde A_n é obtido dividindo cada um dos intervalos que constituem A_{n-1} em 3 subintervalos com o mesmo diâmetro e retirando o subintervalo intermediário. Deste modo, $A_n \subset A_{n-1}$ para todo $n \in \mathbb{N}$. Denominamos

$$C = \bigcap_{n=0}^{\infty} A_n \text{ o conjunto de Cantor.}$$

Consideremos a transformação $f : [0,1] \to [0,1]$ definida por $f(x) = 3x(\text{mod}1)$.
Afirmação: O conjunto C é invariante por f e $f : C \to C$ é uma transformação expansora.
De fato: Para concluir que C é invariante por f, basta verificar que os pontos de C são os pontos do intervalo $[0,1]$ que tem representação na base 3 apenas com $0's$ e $2's$ como dígitos e que $f^{-1}(x) = \left\{\frac{x}{3}, \frac{x+1}{3}, \frac{x+2}{3}\right\}$. Assim, basta ver que se x tem uma representação na base 3 apenas com $0's$ e $2's$, o mesmo acontece com qualquer elemento de $f^{-1}(x)$. O fato de f ser expansora resulta simplesmente do reconhecimento de que $\| Df^n(x)v\| = 3^n \| v\|$ sempre que a derivada está definida.

Teorema 2.6. Seja $f : M \longrightarrow M$ uma transformação expansora numa variedade compacta conexa M. Assuma que o logarítmo

$$M \ni x \longmapsto \log|\det Df(x)|$$

do determinante jacobiano de f é ν-Hölder para algum $\nu > 0$. Então, f admite uma única probabilidade invariante μ absolutamente contínua com relação à medida de Lebesgue m. Além disso,

1. μ é ergódica.
2. O suporte de μ coincide com M.
3. A bacia de μ tem medida de Lebesgue total na variedade M.

A estratégia da demonstração é a seguinte. Note que a pré-imagem por f de um conjunto com medida Lebesgue m nula também tem medida de

Lebesgue nula. Isto significa que a imagem $f_*\nu$ por f de qualquer medida ν absolutamente contínua com relação a m também é absolutamente contínua com relação a m. Em particular, a n-ésima imagem $f_*^n m$ é absolutamente contínua com relação à medida de Lebesgue m para todo $n \geq 1$.

Provaremos na Proposição 2.9 que a derivada de Radon-Nikodym de cada $f_*^n m$ com relação a m (também chamada **densidade**) é limitada por alguma constante independente de $n \geq 1$. Deduziremos que todo ponto de acumulação, na topologia fraca*, de seqüência

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

é uma probabilidade invariante absolutamente contínua com relação à medida de Lebesgue, com densidade limitada pela mesma constante.

Um argumento adicional provará que tal ponto de acumulação é único e satisfaz todas as propriedades no enunciado do teorema.

2.1 Lema de distorção

Lema 2.7. Existe $k \geq 1$ tal que todo ponto $y \in \mathrm{M}$ tem exatamente k pré-imagens por f. Além disso, existe $\rho_0 > 0$ tal que, para qualquer pré-imagem x de um ponto $y \in \mathrm{M}$, existe uma aplicação $h : \mathrm{B}(y, \rho_0) \longrightarrow \mathrm{M}$ de classe C^1 tal que

$$f \circ h = id,\ h(y) = x \ \text{ e } \ d(h(y_1), h(y_2)) \leq \sigma^{-1} d(y_1, y_2)$$

para todo $y_1, y_2 \in \mathrm{B}(y, \rho_0)$.

Demonstração: A condição $\|\mathrm{D}f(x)v\| \geq \sigma \| v\|$ implica que a derivada $\mathrm{D}f$ é um isomorfismo em todo ponto. Logo, dado qualquer $x \in \mathrm{M}$, existe $\rho_0 > 0$ tal que f envia alguma vizinhança V_x de x difeomorficamente sobre a bola de raio ρ_0 em torno de $y = f(x)$. Por compacidade, podemos tomar ρ_0 independente de x. Então, o número de pré-imagens de qualquer $y \in \mathrm{M}$ deve ser limitado. Além disso, o conjunto dos pontos que têm exatamente n pré-imagens é aberto, para todo $n \geq 0$. Portanto, por conexidade, o número de pré-imagens deve ser o mesmo para todo $y \in \mathrm{M}$. Finalmente, seja $h = (f|_{V_x})^{-1}$. Note que

$$\begin{aligned} \|\mathrm{D}f(x)v\| \ \geq \ \sigma \| v\| &\Longrightarrow \frac{1}{\|\mathrm{D}f(x)v\|} \leq \frac{1}{\sigma \| v\|} \Longrightarrow \frac{\| v\|}{\|\mathrm{D}f(x)v\|} \leq \frac{1}{\sigma} \\ &\Longrightarrow \ \|\mathrm{D}f(x)^{-1}\| \leq \sigma^{-1}. \end{aligned} \tag{2.1.2}$$

Derivando ambos os membros da igualdade $f \circ h = id$, temos

$$\mathrm{D}f(h(z)).\mathrm{D}h(z) = 1 \Longrightarrow \mathrm{D}h(z) = \mathrm{D}f(h(z))^{-1}.$$

Combinando $h(y) = x$ com (2.1.2), obtemos

$$\| \mathrm{D}h(z)\| = \| \mathrm{D}f(h(z))^{-1}\| = \| \mathrm{D}f(x)^{-1}\| \leq \sigma^{-1}$$

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

para todo z no domínio de h. Portanto, h contrai distâncias à taxa σ^{-1} como enunciamos. ∎

As transformações h enunciadas no Lema 2.7 são chamadas **ramos inversos** de f. Mais geralmente, podemos definir **ramo inverso** h^n de f^n, para todo $n \geq 1$, da seguinte maneira. Dados $y \in \mathrm{M}$ e $x \in f^{-n}(y)$, sejam $h_1, \ldots, h_n$ ramos inversos de f com

$$h_j\left(f^{n-j+1}(x)\right) = f^{n-j}(x) \text{ para todo } 1 \leq j \leq n.$$

Como cada h_j é uma **contração**, a sua imagem está contida numa bola de raio menor que ρ_0 em torno de $f^{n-j}(x)$. Então, $h^n = h_n \circ \cdots \circ h_1$, está bem definida na bola de raio ρ_0 em torno de y. Note que $f^n \circ h^n = id$ e $h^n(y) = x$.

O próximo resultado fornece um bom controle da distorção de iterados de f e seus ramos inversos, que é fundamental para a demonstração do Teorema 2.6. Este é o único lugar onde se usa a hipótese de que o determinante jacobiano é Hölder.

A hipótese de que a transformação $f : \mathrm{M} \longrightarrow \mathrm{M}$ é expansora serve para obter a seguinte proposição.

Proposição 2.8 (Lema de Distorção). Seja $f : \mathrm{M} \longrightarrow \mathrm{M}$ expansora com $\det \mathrm{D}f$ ν-Hölder. Então, existe uma constante $\mathrm{C}_1 > 0$ tal que, dados qualquer $n \geq 1$, qualquer $y \in \mathrm{M}$ e qualquer ramo inverso $h^n : \mathrm{B}(y, \rho_0) \longrightarrow \mathrm{M}$ de f^n, tem-se

$$\frac{|\det \mathrm{D}h^n(y_1)|}{|\det \mathrm{D}h^n(y_2)|} \leq \exp(\mathrm{C}_1 d((y_1, y_2)^\nu) \leq \exp(\mathrm{C}_1 (2\rho_0)^\nu)$$

para todo $y_1, y_2 \in \mathrm{B}(y, \rho_0)$.

Demonstração: Escrevemos h^n como composição $h^n = h_n \circ \cdots \circ h_1$ de ramos inversos de f. Também escrevemos $h^j = h_j \circ \cdots \circ h_1$ para $1 \leq j \leq n$, bem como $h^0 = id$. Então,

$$\begin{aligned} \log \frac{|\det \mathrm{D}h^n(y_1)|}{|\det \mathrm{D}h^n(y_2)|} &= \log|\det \mathrm{D}h^n(y_1)| - \log|\det \mathrm{D}h^n(y_2)| \\ &= \sum_{j=1}^{n} \log|\det \mathrm{D}h_j(h^{j-1}(y_1))| - \sum_{j=1}^{n} \log|\det \mathrm{D}h_j(h^{j-1}(y_2))|. \end{aligned}$$

Note que $\log|\det \mathrm{D}h_j| = -\log|\det \mathrm{D}f| \circ h_j$ e, por hipótese, $\log|\det \mathrm{D}f|$ é (C_0, ν)-Hölder para algum $\mathrm{C}_0 > 0$. Além disso, pelo Lema 2.7, cada h_j é uma σ^{-1}-contração. Logo,

$$\log \frac{|\det \mathrm{D}h^n(y_1|}{|\det \mathrm{D}h^n(y_2|} \leq \sum_{j=0}^{n} \mathrm{C}_0 \; d(h^j(y_1), h^j(y_2))^\nu \leq \sum_{j=0}^{n} \mathrm{C}_0 \sigma^{-j\nu} d(y_1, y_2)^\nu.$$

Portanto, para prova o lema basta tomar $\mathrm{C}_1 = \sum_{j=0}^{n} \mathrm{C}_0 \sigma^{-j\nu}$. ∎

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

2.2 Medidas absolutamente contínuas

O próximo resultado , que é uma conseqüência do Lema de Distorção, mostra que as imagens $f_*^n m$ têm densidades uniformemente limitadas.

Proposição 2.9. Existe uma constante C_2 tal que $f_*^n m(B) \leq C_2 m(B)$ para todo conjunto mensurável $B \subset M$ e todo $n \geq 1$.

Demonstração: Podemos, sem restrição, supor que B está contido em alguma bola $B_0 = B(z, \rho_0)$ de raio ρ_0 centrada em algum ponto $z \in M$. A Proposição 2.8 implica que

$$\frac{m(h^n(B))}{m(h^n(B_0))} = \frac{\int_B |\det Dh^n| \, dm}{\int_{B_0} |\det Dh^n| \, dm} \leq \exp(C_1(2\rho_0)^\nu) \frac{m(B)}{m(B_0)}$$

para todo ramo inverso h^n e f^n no ponto z. Além disso, $f_*^n m(B) = m(f^{-n}(B))$ é a soma de $m(h^n(B))$ sobre todos os ramos inversos, e analogamente para B_0. Deste modo, obtemos

$$\frac{f_*^n m(B)}{f_*^n m(B_0)} \leq \exp(C_1(2\rho_0)^\nu) \frac{m(B)}{m(B_0)}.$$

Claro que $f_*^n m(B_0) \leq f_*^n m(M) = 1$. Além disso, a medida de Lebesgue das bolas com um raio fixado ρ_0 está limitada de zero por alguma constante $\alpha_0 > 0$ que só depende de ρ_0. Então, para obter a conclusão da proposição basta tomar $C_2 = \dfrac{\exp(C_1(2\rho_0)^\nu)}{\alpha_0}$. ∎

Lema 2.10. Sejam ν uma probabilidade num espaço métrico compacto N e $\varphi : N \longrightarrow [0, +\infty)$ uma função integrável com relação a ν. Seja μ_i, $i \geq 1$, uma seqüência de probabilidades em N convergindo para uma probabilidade μ na topologia fraca*. Se $\mu_i \leq \varphi\nu$ para todo $i \geq 1$ então $\mu \leq \varphi\nu$.

Demonstração: Seja B um conjunto mensurável qualquer. Para cada $\varepsilon > 0$, seja K_ε um conjunto compacto de B tal que $\mu(B \setminus K_\varepsilon)$ e $(\varphi\nu)(B \setminus K_\varepsilon)$ são ambos menores que ε. Então, seja A_ε uma vizinhança aberta de K_ε da forma $A_\varepsilon = \{z;\ d(z, K_\varepsilon) < r\}$ com $r > 0$ suficientemente pequeno para que a medida de $A_\varepsilon \setminus K_\varepsilon$ seja menor que ε, tanto para μ como para $\varphi\nu$. Mudando r ligeiramente, caso necessário, podemos supor que o bordo de A_ε tem μ-medida zero: há no máximo uma quantidade enumerável de valores de r para os quais isso não acontece. Então,

$$\mu = \lim \mu_i \Longrightarrow \mu(A_\varepsilon) = \lim \mu_i(A_\varepsilon) \leq (\varphi\nu)(A_\varepsilon).$$

Fazendo $\varepsilon \to 0$, obtemos $\mu(B) \leq (\varphi\nu)(B)$. ∎

Aplicando este lema na nossa situação, obtemos

Corolário 2.11. Todo ponto de acumulação μ da seqüência $\dfrac{1}{n}\sum_{j=0}^{n-1} f_*^j m$ é uma probabilidade invariante por f absolutamente contínua com relação à medida

de Lebesgue.

Demonstração: Tomemos no Lema 2.10, φ constante igual a C_2 e $\nu = m$, bem como $\mu_i = \frac{1}{n_i} \sum_{j=0}^{n_i-1} f_*^j m$, para qualquer subseqüência $(n_i)_i$ tal que $(\mu_i)_i$ converge para uma medida μ. A Proposição 2.9 garante que $\nu_i \leq \varphi\nu$. Então, também temos $\mu \leq \varphi\nu = C_2 m$, pelo Lema 2.10. Isto implica que $\mu \ll m$, com densidade limitada por C_2. ∎

2.3 Existência de medidas ergódicas

Agora vamos mostrar que a medida μ que acabamos de construir é a única probabilidade invariante absolutamente contínua com relação à medida de Lebesgue e, além disso, é ergódica por f.

Começamos por fixar uma partição $\mathcal{P}_0 = \{U_1, \ldots, U_s\}$ de M em regiões com interior não vazio e diâmetro menor que ρ_0. Então, para cada $n \geq 1$, definimos $\mathcal{P}_n$ como sendo a partição de M que consiste das imagens de cada um dos U_i, $1 \leq i \leq s$, pelos respectivos ramos inversos de f^n. O diâmetro da partição $\mathcal{P}_n$, ou seja, o supremo dos diâmetros dos seus elementos, é menor que $\rho_0 \sigma^{-n}$.

Lema 2.12. Seja $\mathcal{P}_n$, $n \geq 1$, uma seqüência de partições num espaço métrico compacto com diâmetro convergindo para zero quando $n \to +\infty$. Sejam ν uma probabilidade nesse espaço e B qualquer conjunto mensurável com $\nu(B) > 0$. Então existem $V_n \in \mathcal{P}_n$, $\nu(V_n) > 0$ para $n \geq 1$, tais que

$$\frac{\nu(B \cap V_n)}{\nu(V_n)} \to 1 \quad \text{quando } n \to +\infty.$$

Demonstração: Dado qualquer $0 < \varepsilon < \nu(B)$, seja K_ε um subconjunto compacto de B com $\nu(B \setminus K_\varepsilon) < \varepsilon$. Como o diâmetro das partições converge para zero, a medida da união $A_{\varepsilon,n}$ de todos os elementos de $\mathcal{P}_n$ que intersectam K_ε satisfaz $\nu(A_{\varepsilon,n} \setminus K_\varepsilon) < \varepsilon$ para todo n suficientemente grande. Se tivéssemos

$$\nu(K_\varepsilon \cap V_n) \leq \frac{\nu(B) - \varepsilon}{\nu(V_n) + \varepsilon} \nu(V_n)$$

para todo $V_n \in \mathcal{P}_n$ que intersecta K_ε, deduziríamos que

$$\nu(K_\varepsilon) \leq \frac{\nu(B) - \varepsilon}{\nu(V_n) + \varepsilon} \nu(A_{\varepsilon,n}) \leq \frac{\nu(B) - \varepsilon}{\nu(V_n) + \varepsilon} \ (\nu(K_\varepsilon) + \varepsilon) \leq \nu(B) - \varepsilon,$$

o que é uma contradição. Portanto, deve existir algum $V_n \in \mathcal{P}_n$ tal que

$$\nu(B \cap V_n) \geq \nu(K_\varepsilon \cap V_n) > \frac{\nu(B) - \varepsilon}{\nu(V_n) + \varepsilon} \nu(V_n)$$

e isto implica $\nu(V_n) > 0$. Fazendo $\varepsilon \to 0$ obtemos a conclusão do lema. ∎

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Recorde que $A \subset M$ é um conjunto invariante por $f : M \longrightarrow M$ se $f^{-1}(A) = A$. Mas geralmente, dizemos que A é **positivamente invariante** por f se $f(A) = A$. Note que conjuntos invariantes são sempre positivamente invariantes, pois f é sobrejetora.

Lema 2.13. Se $A \subset M$ é um conjunto positivamente invariante por f com medida de Lebesgue $m(A) > 0$, então A tem medida de Lebesgue total em algum $U_i \in \mathcal{P}_0$, ou seja, existe $1 \leq i \leq s$ tal que $m(U_i \setminus A) = 0$.
Demonstração: Pelo Lema 2.12, existe $V_n \in \mathcal{P}_n$ tal que

$$\frac{m(V_n \setminus A)}{m(V_n)}$$

converge para zero quando $n \to \infty$. Seja $U_{i(n)} = f^n(V_n)$. Aplicando a Proposição 2.8 ao ramo inverso de f^n que envia $U_{i(n)}$ em V_n, concluímos que

$$\frac{m(U_{i(n)} \setminus A)}{m(U_{i(n)})} \leq \frac{m(f^n(V_n \setminus A))}{m(f^n(V_n))} \leq \exp(C_1(2\rho_0)^\nu) \frac{m(V_n \setminus A)}{m(V_n)}$$

também converge para zero. Como $\mathcal{P}_0$ é finita, deve existir $1 \leq i \leq s$ tal que $i(n) = i$ para infinitos valores de n. Logo, $m(U_{i(n)} \setminus A) = 0$. ∎

Corolário 2.14. A transformação $f : M \longrightarrow M$ admite alguma probabilidade invariante ergódica e absolutamente contínua com relação à medida de Lebesgue.
Demonstração: Como conseqüência do Lema 2.13, existem no máximo $\#\mathcal{P}_0$ conjuntos invariantes com medida de Lebesgue positiva disjuntos dois-a-dois. Portanto, M pode ser particionado num número finito de conjuntos invariantes $A_1, \ldots, A_s$, $s \leq \#\mathcal{P}_0$ com medida de Lebesgue positiva e que são minimais no sentido de que não existam subconjuntos invariantes $B_i \subset A_i$ com $0 < m(B_i) < m(A_i)$. Dada qualquer medida invariante absolutamente contínua μ, existe algum i tal que $\mu(A_i) > 0$. Então, a restrição normalizada μ_i de μ a A_i,

$$\mu_i(B) = \frac{\mu(B \cap A_i)}{\mu(A_i)}$$

é invariante e absolutamente contínua. Além disso, μ_i é ergódica, pois A_i é minimal. ∎

2.4 Unicidade e conclusão da prova

O argumento anterior também mostra que existe apenas um número finito de probabilidades ergódicas e absolutamente contínuas. O último passo da demonstração é mostrar que, de fato, tal probabilidade é única. Para isso usamos o fato de que f é **topologicamente misturadora**, isto é, para quaisquer dois abertos não-vazios $U, V \subset M$, existe $k \in \mathbb{N}$ tal que $f^{-n}(U) \cap V \neq \emptyset$ para todo $n > k$.

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Lema 2.15. Dado qualquer aberto não vazio $U \subset M$, existe $N \geq 1$ tal que $f^N(U) = M$.

Demonstração: Seja $x \in U$ e $r > 0$ tais que a bola de raio r centrada em x está contida em U. Dado qualquer $n \geq 1$, suponha que $f^n(U)$ não cobre toda a variedade. Então existe alguma curva γ ligando $f^n(x)$ a um ponto $y \in (M \setminus f^n(U))$ e essa curva pode ser tomada com comprimento menor que $\text{diam}M + 1$. Levantando γ pelo difeomorfismo local f^n, obtemos uma curva γ_n ligando x a um ponto $y_n \in (M \setminus U)$. Então, $r \leq \text{comp}(\gamma_n) \leq \sigma^{-1}(\text{diam}M + 1)$. Isto dá uma cota superior para o valor possível de n. Portanto, $f^n(U) = M$ para todo n suficientemente grande, como afirmamos. ∎

Corolário 2.16. Se $A \subset M$ é um conjunto positivamente invariante com medida de Lebesgue positiva, então A tem medida de Lebesgue total na variedade M.

Demonstração: Sejam U o interior de um conjunto U_i como no Lema 2.13 e $N \geq 1$ tal que $f^N(U) = M$. Então, $m(U \setminus A) = 0$, e portanto

$$M \setminus A = (f^N(U) \setminus f^N(A)) \subset f^N(U \setminus A)$$

também tem medida de Lebesgue zero. ∎

O próximo enunciado completa a demonstração do Teorema 2.6.

Corolário 2.17. Seja μ qualquer probabilidade invariante absolutamente contínua. Então, μ é ergódica e a sua bacia $B(\mu)$ tem medida de Lebesgue total em M. Além disso, o suporte de μ é toda a variedade M.

Demonstração: Se A é um subconjunto invariante qualquer, então pelo Corolário 2.16, A tem medida de Lebesgue zero ou A^c tem medida de Lebesgue zero. Uma vez que μ é absolutamente contínua, segue-se que $\mu(A) = 0$ ou $\mu(A^c) = 0$. Isto prova que μ é ergódica. Então, $B(\mu)$ é um conjunto invariante com medida de Lebesgue positiva e, conseqüentemente, deve ter medida de Lebesgue total. Analogamente, como o suporte de μ é um conjunto compacto positivamente invariante, ele tem que coincidir com M. ∎

Observação 2.18. A conclusão do Teorema 2.6 é falsa, em geral, se omitirmos a hipótese de que o jacobiano é Hölder.



Capítulo 3

Transformações expansoras por partes

A noção de variação de uma função, ferramenta essencial no estudo da existência de medidas invariantes absolutamente contínuas, é muito sensível à geometria do seu domínio. Este fato é causador de grandes dificuldades quando se pretendem estudar tais medidas em dimensão superior a um, e para as ultrapassar, é necessário ter em conta os domínios e diferenciabilidade das transformações. O nosso objetivo é expor resultados da existência de medidas invariantes ergódicas absolutamente contínuas com relação à medida de Lebesgue total para transformações expansoras por partes em dimensão um. Neste capítulo vamos considerar $M = S^1$ ou $M = [0, 1]$.

Definição 3.1. Dizemos que $\mathcal{P} = \{\eta_i;\ i \in \mathbb{N}\}$ é uma **partição** por intervalos abertos para M se:

- η_i é um intervalo aberto.
- $\eta_i \cap \eta_j = \emptyset$ para $i \neq j$.
- $m\left(M \setminus \bigcup_{i \in \mathbb{N}} \eta_i\right) = 0$, onde m é a medida de Lebesgue.

Segue destas propriedades que $\mathcal{P}$ é uma partição **finita** se $\mathcal{P} = \{\eta_1, \ldots, \eta_k\}$ ou **enumerável** se $\mathcal{P} = \{\eta_1, \ldots, \eta_k, \ldots\}$. Os elementos de $\mathcal{P}$ são denominados **átomos**.

Se $\mathcal{P}$ e $\mathcal{Q}$ são partições, dizemos $\mathcal{Q}$ é mais fina que $\mathcal{P}$, e denotamos $\mathcal{P} \leq \mathcal{Q}$, se todo átomo de $\mathcal{Q}$ está contido $\mod(0)$ em algum átomo de $\mathcal{P}$. Isto implica que todo átomo de $\mathcal{P}$ pode-se escrever como união de átomos de $\mathcal{Q}$ $\mod(0)$.

Chamamos $\mathcal{P}^1$ uma **partição por intervalos de regularidade** de f. Dizemos que f tem um número finito de intervalos de regularidades se a partição $\mathcal{P}^1$ a ser escolhida for finita. Os pontos fronteiras dos intervalos de regularidades $\eta \in \mathcal{P}^1$ que não estão na fronteira de M são chamados **pontos singulares** de f.

unesp

3.1 Funções de variação limitada

Nesta seção, apresentamos resultados simples sobre **funções de variação limitada em dimensão um**, às quais estão intimamente relacionadas com as funções monótonas.

Definição 3.2. Sejam $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$ uma função e $\eta = [a, b]$ um intervalo compacto em M. A **variação** de φ em η é

$$\operatorname*{var}_{\eta}\varphi = \sup\left\{\sum_{i=1}^{n}|f(t_i) - f(t_{i-1})|;\ \ n \in \mathbb{N} \ \text{ e } \ a = t_o < t_1 < \ldots < t_n = b\right\}.$$

A variação $\operatorname*{var}_{\eta}\varphi$ de φ num subconjunto conexo η de M arbitrário (incluindo $\eta = \mathrm{M} = \mathrm{S}^1$) é o supremo destas variações sobre todos intervalos compactos contidos em η. Representamos a **variação** sobre toda variedade M simplesmente como:

$$\operatorname{var}\varphi = \sup\left\{\operatorname*{var}_{\eta}\varphi;\ \eta \subset \mathrm{M}\right\}.$$

Definição 3.3. Uma função $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$ tem **variação limitada** se varφ é finita. Dado qualquer subconjunto conexo η de M, φ tem variação limitada em η se $\operatorname*{var}_{\eta}\varphi < \infty$.

Exemplo 3.4. A função $\varphi : [a, b] \longrightarrow \mathbb{R}$ definida por $\varphi(x) = \operatorname{sen}(x)$ é de variação limilitada em $[a, b]$.
De fato: Sejam $x, y \in [a, b]$ com $0 < b - a < \pi$, como a função seno é contínua em $[a, b]$, pelo Teorema do Valor Médio, existe $z \in (a, b)$ tal que $\operatorname{sen}(x) - \operatorname{sen}(y) = \cos(z)(x - y)$. Como $|\cos(z)| \leq 1$, obtemos

$$|\operatorname{sen}(x) - \operatorname{sen}(y)| = |\cos(z)(x - y)| = |\cos(z)||x - y| \leq |x - y|.$$

Dada $\{x_o, x_1, \cdots, x_n\}$ uma partição de $[a, b]$, temos

$$\begin{aligned}\sum_{i=1}^{n}|\operatorname{sen}(x_i) - \operatorname{sen}(x_{i-1})| &\leq \sum_{i=1}^{n}|x_i - x_{i-1}| = \sum_{i=1}^{n}(x_i - x_{i-1}) \\ &= (x_1 - x_0) + (x_2 - x_1) + \cdots + (x_n - x_{n-1}) \\ &= x_n - x_0 = b - a.\end{aligned}$$

Pela Definição 3.2, esta desigualdade implica que $\operatorname*{var}_{[a,b]}\varphi \leq b - a$. ∎

Proposição 3.5. Seja $\varphi : [a, b] \longrightarrow \mathbb{R}$ uma função de variação limitada. Então, φ é limitada em $[a, b]$.
Demonstração: Seja $\{x_o, x_1, \ldots, x_n\}$ uma partição de $[a, b]$. Note que

$$\begin{aligned}\sum_{i=1}^{n}|\varphi(x_i) - \varphi(x_{i-1})| &= |\varphi(x_1) - \varphi(x_0)| + \cdots + |\varphi(x_n) - \varphi(x_{n-1})| \\ &\geq |\varphi(x_1)| - |\varphi(x_0)| + \cdots + |\varphi(x_n)| - |\varphi(x_{n-1})| \\ &= |\varphi(x_n)| - |\varphi(x_0)| = |\varphi(b)| - |\varphi(a)|.\end{aligned}$$

unesp

Isto mostra que $|\varphi(b)| - |\varphi(a)| \leq \sum_{i=1}^{n} |\varphi(x_i) - \varphi(x_{i-1})|$. Tendo em vista que $\sum_{i=1}^{n} |\varphi(x_i) - \varphi(x_{i-1})| \leq \operatorname{var}_{[a,b]} \varphi$, temos que $|\varphi(b)| - |\varphi(a)| \leq \operatorname{var}_{[a,b]} \varphi$. Note que, dado $x \in [a,b]$ com $a < x < b$, obtemos

$$\begin{aligned} |\varphi(b)| - |\varphi(a)| &\leq |\varphi(b) - \varphi(a)| = |\varphi(b) - \varphi(x) + \varphi(x) - \varphi(a)| \\ &\leq |\varphi(b) - \varphi(x)| + |\varphi(x) - \varphi(a)| > |\varphi(x) - \varphi(a)| \\ &\geq |\varphi(x)| - |\varphi(a)|. \end{aligned}$$

Portanto, $|\varphi(x)| - |\varphi(a)| \leq |\varphi(b)| - |\varphi(a)| \leq \operatorname{var}_{[a,b]} \varphi$ o que implica $|\varphi(x)| \leq |\varphi(a)| + \operatorname{var}_{[a,b]} \varphi$, para todo $x \in [a,b]$. Isto mostra que φ é limitada em $[a,b]$. ∎

Observação 3.6. Uma função contínua e/ou limitada num intervalo não tem necessariamente variação limitada nesse intervalo.

Exemplo 3.7. A função

$$\varphi(x) = \begin{cases} x \operatorname{sen} \frac{1}{x}, & \text{se } 0 < x < 2\pi \\ 0, & \text{se } x = 0 \end{cases}$$

é limitada em $[0, 2\pi]$ mas não tem variação limitada.

De fato: Tome a seqüência $x_k = \dfrac{1}{k\pi + \frac{\pi}{2}}$. Então, $\varphi(x_k) = (-1)^k \dfrac{2}{2k\pi + \pi}$ e $\varphi(x_{k-1}) = (-1)^{k-1} \dfrac{2}{2k\pi - \pi}$. Note que

$$\begin{aligned} \sum_{k=1}^{n} |\varphi(x_k) - \varphi(x_{k-1})| &= \sum_{k=1}^{n} \left| \frac{2}{2k\pi + \pi} + \frac{2}{2k\pi - \pi} \right| \\ &= \sum_{k=1}^{n} \left(\frac{2}{2k\pi - \pi} + \frac{2}{2k\pi + \pi} \right) \\ &= \left(\frac{2}{\pi} + \frac{2}{3\pi} \right) + \left(\frac{2}{3\pi} + \frac{2}{5\pi} \right) + \left(\frac{2}{5\pi} + \frac{2}{7\pi} \right) + \cdots \\ &\cdots + \left(\frac{2}{2n\pi - 3\pi} + \frac{2}{2n\pi - \pi} \right) + \left(\frac{2}{2n\pi - \pi} + \frac{2}{2n\pi + \pi} \right) \\ &> \frac{2}{\pi} + \frac{2}{2\pi} + \frac{2}{3\pi} + \cdots + \frac{2}{(n-1)\pi} + \frac{2}{n\pi} = \sum_{k=1}^{n} \frac{2}{k\pi} \\ &= \frac{2}{\pi} \sum_{k=1}^{n} \frac{1}{k} \longrightarrow +\infty. \end{aligned}$$

unesp

Como $\operatorname{var}_{[0,2\pi]} \varphi > \frac{2}{\pi} \sum_{k=1}^{n} \frac{1}{k} \longrightarrow +\infty$, então a função φ não tem variação limitada em $[0, 2\pi]$. ∎

Lema 3.8. Sejam $\varphi_1,\ \varphi_2 : M \longrightarrow \mathbb{R}$ funções de variação limitada e η um subconjunto conexo de M. Então,

1. $\operatorname{var}_{\eta}(\varphi_1 \pm \varphi_2) \leq \operatorname{var}_{\eta} \varphi_1 \pm \operatorname{var}_{\eta} \varphi_2$.
2. $\operatorname{var}_{\eta}(\varphi_1\, \varphi_2) \leq \operatorname{var}_{\eta} \varphi_1 \sup_{\eta} |\varphi_2| + \sup_{\eta} |\varphi_1| \operatorname{var}_{\eta} \varphi_2$.
3. $\sup_{\eta} \varphi_1 \leq \operatorname{var}_{\eta} \varphi_1 + \inf_{\eta} \varphi_1 \leq \operatorname{var}_{\eta} \varphi_1 + \frac{1}{m(\eta)} \int_{\eta} \varphi_1\, dm$.
4. $\sup_{\eta} (\varphi_1 \circ \mathrm{h}) = \sup_{\mathrm{h}(\eta)} \varphi_1$ se $h : \eta \longrightarrow \mathrm{h}(\eta)$ é um homeomorfismo.

Demonstração: Seja $\{x_o, x_1, \ldots x_n\}$ uma partição de η.
1. Note que

$$\begin{aligned}\sum_{i=1}^{n} |(\varphi_1 + \varphi_2)(x_i) - (\varphi_1 + \varphi_2)(x_{i-1})| &= \sum_{i=1}^{n} |\varphi_1(x_i - x_{i-1}) + \varphi_2(x_i - x_{i-1})| \\ &\leq \sum_{i=1}^{n} |\ \varphi_1(x_i) - \varphi_1(x_{i-1})| \\ &\quad + \sum_{i=1}^{n} |\ \varphi_2(x_i) - \varphi_2(x_{i-1})| \\ &\leq \operatorname{var}_{\eta} \varphi_1 + \operatorname{var}_{\eta} \varphi_2.\end{aligned}$$

Pela Definição 3.2, esta desigualdade implica que

$$\operatorname{var}_{\eta}(\varphi_1 + \varphi_2) \leq \operatorname{var}_{\eta} \varphi_1 + \operatorname{var}_{\eta} \varphi_2.$$

De modo análogo, mostra-se que $\operatorname{var}_{\eta}(\varphi_1 - \varphi_2) \leq \operatorname{var}_{\eta} \varphi_1 - \operatorname{var}_{\eta} \varphi_2$.

2. Considere $|\varphi_1(x_i)| = \sup_{\eta}|\varphi_1|$ e $|\varphi_2(x_i)| = \sup_{\eta}|\varphi_2|$. Note que

$$\begin{aligned}
\sum_{i=1}^{n}|\,(\varphi_1\varphi_2)(x_i)-(\varphi_1\varphi_2)(x_{i-1})| &\leq \sum_{i=1}^{n}|\,\varphi_1(x_i)\varphi_2(x_i)-\varphi_1(x_{i-1})\varphi_2(x_i)| \\
&\quad+\sum_{i=1}^{n}|\,\varphi_1(x_{i-1})\varphi_2(x_i)-\varphi_1(x_{i-1})\varphi_2(x_{i-1})| \\
&\leq \sum_{i=1}^{n}|\,\varphi_1(x_i)-\varphi_1(x_{i-1})||\varphi_2(x_i)| \\
&\quad+\sum_{i=1}^{n}|\,\varphi_1(x_{i-1})||\,\varphi_2(x_i)-\varphi_2(x_{i-1})| \\
&\leq \sum_{i=1}^{n}|\,\varphi_1(x_i)-\varphi_1(x_{i-1})|\,\sup_{\eta}|\varphi_2| \\
&\quad+\sup_{\eta}\,|\varphi_1|\sum_{i=1}^{n}|\,\varphi_2(x_i)-\varphi_2(x_{i-1})| \\
&\leq \operatorname*{var}_{\eta}\,\varphi_1\sup_{\eta}|\varphi_2| \;+\; \sup_{\eta}|\varphi_1|\operatorname*{var}_{\eta}\varphi_2.
\end{aligned}$$

Pela Definição 3.2, esta desigualdade implica que

$$\operatorname*{var}_{\eta}(\varphi_1\,\varphi_2)\leq\operatorname*{var}_{\eta}\,\varphi_1\sup_{\eta}|\varphi_2| \;+\; \sup_{\eta}|\varphi_1|\operatorname*{var}_{\eta}\varphi_2.$$

Observação 3.9. Tendo em vista que a **função identicamente nula** e a **função constante** têm variação limitada, pelos ítens 1 e 2 do Lema 3.8, resulta que o **conjunto das funções de variação limitada** tem uma **estrutura de espaço vetorial** sobre o corpo $\mathbb{R}$.

Proposição 3.10. Seja $\varphi : [a,b] \longrightarrow \mathbb{R}$ uma função monótona. Então, φ tem variação limitada em $[a,b]$.

Demonstração: Suponha que φ é não-decrescente. Então, $a < b$ implica $\varphi(a) \leq \varphi(b)$. Note que $|\,\varphi(a)| \leq |\,\varphi(b)| \Longrightarrow 0 \leq |\,\varphi(b)| - |\,\varphi(a)|$. Como $\operatorname*{var}_{[a,b]}\varphi \geq |\,\varphi(b)| - |\,\varphi(a)|$ e $|\,\varphi(b)| - |\,\varphi(a)| \leq |\,\varphi(b) - \varphi(a)|$, temos que $\operatorname*{var}_{[a,b]}\varphi \leq |\,\varphi(b) - \varphi(a)|$. Isto prova que φ tem variação limitada em $[a,b]$. ■

De modo análogo, prova-se os demais casos.

Proposição 3.11. Seja $\varphi : [a,b] \longrightarrow \mathbb{R}$ uma função lipschitziana. Então, φ tem variação limitada em $[a,b]$.

Demonstração: Como φ é lipschitziana em $[a,b]$, então existe $K > 0$ tal que

$$|\,\varphi(b) - \varphi(a)| \leq \mathrm{K}|\,b-a|.$$

unesp

Usando a Proposição 3.10, obtemos

$$\operatorname*{var}_{[a,b]} \varphi \leq \mathrm{K}|\, b-a|.$$

Isto prova que φ tem variação limitada em $[a,b]$. ∎

Observação 3.12. A inversa de uma função de variação limitada não é necessariamente de variação limitada, conseqüentemente o quociente de funções de variação limitada pode não ser de variação limitada.

Exemplo 3.13. Seja $\varphi : [a,b] \longrightarrow \mathbb{R}$ uma função de variação limitada. Se $x_0 \in [a,b]$ e $\varphi(x) \longrightarrow 0$ quando $x \longrightarrow 0$, então a função $\dfrac{1}{\varphi}$ não é limitada em $[a,b]$, conseqüentemente não tem variação limitada em $[a,b]$.

Excluindo quocientes em que a função do denominador tome valores muito próximos de zero, o Lema 3.8 pode ser estendido a quocientes de funções de variação limitada.

Proposição 3.14. Sejam $\varphi : [a,b] \longrightarrow \mathbb{R}$ uma função de variação limitada e uma constante $\mathrm{C} > 0$ tal que $|\,\varphi(x)| \geq \mathrm{C}$, para todo $x \in [a,b]$. Então, a função $\dfrac{1}{\varphi}$ tem variação limitada em $[a,b]$ e

$$\operatorname*{var}_{[a,b]} \frac{1}{\varphi} \leq \frac{1}{\mathrm{C}^2} \operatorname*{var}_{[a,b]} \varphi.$$

Demonstração: Seja $\{x_o, x_1, \ldots x_n\}$ uma partição de $[a,b]$. Note que

$$\begin{aligned}
\sum_{i=1}^{n} \left| \frac{1}{\varphi(x_i)} - \frac{1}{\varphi(x_{i-1})} \right| &= \sum_{i=1}^{n} \left| \frac{\varphi(x_{i-1}) - \varphi(x_i)}{\varphi(x_i)\,\varphi(x_{i-1})} \right| = \sum_{i=1}^{n} \frac{|\,\varphi(x_{i-1}) - \varphi(x_i)|}{|\,\varphi(x_i)||\,\varphi(x_{i-1})|} \\
&\geq \sum_{i=1}^{n} \frac{|\,\varphi(x_{i-1}) - \varphi(x_i)|}{\mathrm{C}^2} = \frac{1}{\mathrm{C}^2} \sum_{i=1}^{n} |\,\varphi(x_{i-1}) - \varphi(x_i)| \\
&\geq \frac{1}{\mathrm{C}^2} \left(|\,\varphi(a)| - |\,\varphi(b)| \right).
\end{aligned}$$

Como $\operatorname*{var}_{[a,b]} \dfrac{1}{\varphi} \geq \displaystyle\sum_{i=1}^{n} \left| \frac{1}{\varphi(x_i)} - \frac{1}{\varphi(x_{i-1})} \right| \geq \frac{1}{\mathrm{C}^2} \left(|\,\varphi(a)| - |\,\varphi(b)| \right)$ e $|\,\varphi(a)| - |\,\varphi(b)| \leq |\,\varphi(a) - \varphi(b)|$, então $\operatorname*{var}_{[a,b]} \dfrac{1}{\varphi} \leq |\,\varphi(a) - \varphi(b)| \leq \operatorname*{var}_{[a,b]} \varphi$. ∎

O teorema seguinte dá uma importante caracterização das funções de variação limitada na reta.

Teorema 3.15. Se $\varphi \in \mathrm{C}^1([a,b])$ então φ tem variação limitada em $[a,b]$ e

$$\operatorname*{var}_{[a,b]} \varphi = \int_a^b |\,\varphi'(x)|\, dx.$$

unesp

Demonstração: Como $\varphi \in C^1([a,b])$ implica que φ' é contínua em $[a,b]$. Pelo Teorema do Valor Médio, para $a < x < y < b$, existe $x < z < y$ tal que $\varphi(y) - \varphi(x) = \varphi'(z)(y-x)$. Seja $|\varphi'(z)| \leq K$ com $K > 0$. Note que

$$|\varphi(y) - \varphi(x)| = |\varphi'(z)(y-x)| = |\varphi'(z)||y-x| \leq K|y-x|$$

para todo $x, y \in [a,b]$. Portanto, φ é lipschitziana em $[a,b]$ e pela Proposição 3.11 tem variação limitada em $[a,b]$. Dado $\varepsilon > 0$ qualquer, existe uma partição $\{x_0, x_1, \ldots, x_n\}$ em $[a,b]$ tal que

$$\operatorname*{var}_{[a,b]} \varphi - \varepsilon \leq \sum_{i=1}^{n} |\varphi(x_i) - \varphi(x_{i-1})| \leq \operatorname*{var}_{[a,b]} \varphi.$$

Usando o Teorema Fundamental do Cálculo, note que

$$\sum_{i=1}^{n} |\varphi(x_i) - \varphi(x_{i-1})| = \sum_{i=1}^{n} \left| \int_{x_{i-1}}^{x_i} \varphi'(x)\, dx \right| \leq \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} |\varphi'(x)|\, dx$$
$$= \int_a^b |\varphi'(x)|\, dx.$$

Deste modo,

$$\operatorname*{var}_{[a,b]} \varphi - \varepsilon \leq \sum_{i=1}^{n} |\varphi(x_i) - \varphi(x_{i-1})| \leq \int_a^b |\varphi'(x)|\, dx.$$

Portanto, como ε é arbitrário, obtemos

$$\operatorname*{var}_{[a,b]} \varphi \leq \int_a^b |\varphi'(x)|\, dx.$$

Por outro lado, como φ' é contínua em $[a,b]$, φ' é uniformemente contínua em $[a,b]$. Dado $\varepsilon > 0$ arbitrário, como φ' é uniformemente contínua em $[a,b]$, existe $\delta > 0$ tal que

$$|\varphi'(y) - \varphi'(x)| < \varepsilon \text{ sempre que } |y-x| < \delta, \quad \text{para todo } x, y \in [a,b].$$

Seja $\{x_0, x_1, \cdots, x_n\}$ uma partição em $[a,b]$ cuja amplitude máxima é menor do que δ. Se $x_{i-1} \leq x \leq x_i$, temos $|\varphi'(x)| \leq |\varphi'(x_i)| + \varepsilon$. Integrando membro a membro, obtemos

$$\begin{aligned}
\int_{x_{i-1}}^{x_i} |\varphi'(x)|\, dx - \varepsilon \Delta x_i &\leq \int_{x_{i-1}}^{x_i} |\varphi'(x_i)|\, dx = \int_{x_{i-1}}^{x_i} |\varphi'(x) + \varphi'(x_i) - \varphi'(x)|\, dx \\
&\leq \left| \int_{x_{i-1}}^{x_i} \varphi'(x)\, dx \right| + \left| \int_{x_{i-1}}^{x_i} (\varphi'(x_i) - \varphi'(x))\, dx \right| \\
&\leq \int_{x_{i-1}}^{x_i} |\varphi'(x)|\, dx + |\varphi'(x_i) - \varphi'(x)| \int_{x_{i-1}}^{x_i} dx \\
&< |\varphi'(x_i) - \varphi'(x_{i-1})| + \varepsilon\, \Delta x_i.
\end{aligned}$$

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Aplicando o somatório no primeiro e no último membros da desigualdade, com $i \in \{1, \cdots, n\}$, obtemos

$$\int_a^b |\ \varphi'(x)|\ dx < \sum_{i=1}^{n} |\ \varphi(x_i) - \varphi(x_{i-1})| + 2\varepsilon(b-a) < \operatorname*{var}_{[a,b]}\ \varphi + 2\varepsilon(b-a).$$

Como ε é arbitrário, segue-se que $\int_a^b |\ \varphi'(x)|\ dx < \operatorname*{var}_{[a,b]}\ \varphi$. O que completa a demonstração do teorema. ∎

Observação 3.16. Existem funções descontínuas que são de variação limitada. Por exemplo, a função

$$\varphi(x) = \begin{cases} 1, & \text{se} \quad 1 \leq x \leq 2 \\ 0, & \text{se} \quad 0 \leq x \leq 1 \end{cases}$$

é de variação limitada.

Lema 3.17. Suponha que $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$ tem variação limitada em algum intervalo $\eta = [a, b] \subset \mathrm{M}$. Então,

1. Para todo $x, y \in \eta$ com $x > y$ implica que

$$|\ \varphi(x) - \varphi(y)| \leq \operatorname*{var}_{[a,x]}\ \varphi - \operatorname*{var}_{[a,y]}\ \varphi.$$

2. Para todo $p \in \eta$, os limites laterais $\lim_{x \to p\ \pm} \varphi(x)$ existem (para pontos na fronteira considere apenas o limite no interior de η).

3. O conjunto dos pontos de descontinuidade é no máximo enumerável.

4. A restrição φ em η pode ser escrita como diferença de duas funções não-decrescentes (também de variação limitada).

Demonstração:
1. Seja $\{y_0, y_1, \cdots, y_n\}$ uma partição de $[a, y]$, então

$$\begin{aligned} \operatorname*{var}_{[a,x]}\ \varphi &\geq \sum_{i=1}^{n} |\ \varphi(y_i) - \varphi(y_{i-1})| + |\ \varphi(x) - \varphi(y)| \Longrightarrow \operatorname*{var}_{[a,x]}\ \varphi \\ &\geq \operatorname*{var}_{[a,y]}\ \varphi + |\ \varphi(x) - \varphi(y)| \Longrightarrow \operatorname*{var}_{[a,x]}\ \varphi - \operatorname*{var}_{[a,y]}\ \varphi \geq |\ \varphi(x) - \varphi(y)|. \end{aligned}$$

2. Como $\mathrm{var}\varphi < \infty$, então $\mathrm{var}\varphi$ é uma função não-decrescente assumindo valores finitos. Em particular, dado $p \in \eta$, existem os limites laterais

$$\lim_{x \to p\ \pm} \operatorname*{var}_{[a,x]}\ \varphi.$$

unesp

Tomemos o caso em que $x \to p^+$. Assim, pela propriedade 1, tomando $p^+ < y < x$, temos

$$|\varphi(x) - \varphi(y)| \leq \operatorname*{var}_{[a,x]} \varphi - \operatorname*{var}_{[a,y]} \varphi \longrightarrow 0.$$

Se $x, y \longrightarrow p$, então como afirmamos, existem os limites laterais para $\mathrm{var}\varphi$. Portanto, existem também os limites laterais para φ.

De modo análogo, prova-se o caso $x \to p^-$.

3. Pela propriedade 1 e por raciocínio análogo ao da propriedade 2, vale que $\mathrm{var}\varphi$ é contínua num ponto $p \in \eta$ o que implica φ contínua em p. Mas $\mathrm{var}\varphi$ é monótona não-decrescente, e sabemos de análise na reta que as funções monótonas têm seu conjunto de descontinuidades enumerável.

4. Considere as funções ϕ e ψ tais que $\varphi = \phi - \psi$. Tomemos $\phi = \frac{1}{2}(\mathrm{var}\varphi - \varphi)$ e $\psi = \frac{1}{2}(\mathrm{var}\varphi - \varphi)$. Pela propriedade 1, se $x \geq y$, temos

- $\varphi(x) - \varphi(y) \leq |\varphi(x) - \varphi(y)| \leq \operatorname*{var}_{[a,x]} \varphi - \operatorname*{var}_{[a,y]} \varphi \Longrightarrow \operatorname*{var}_{[a,x]} \varphi - \varphi(x) \geq \operatorname*{var}_{[a,y]} \varphi - \varphi(y)$.
- $\varphi(y) - \varphi(x) \leq |\varphi(x) - \varphi(y)| \leq \operatorname*{var}_{[a,x]} \varphi - \operatorname*{var}_{[a,y]} \varphi \Longrightarrow \operatorname*{var}_{[a,x]} \varphi + \varphi(x) \geq \operatorname*{var}_{[a,y]} \varphi + \varphi(y)$.

Portanto, ϕ e ψ são não-decrescentes. Como φ e $\mathrm{var}\varphi$ são de variação limitada, seque-se que ϕ e ψ também são de variação limitada. ∎

Agora vamos provar o **Teorema de Helly**: conjuntos de funções uniformemente limitada que tem uniformemente variação limitada são relativamente compactos em $\mathrm{L}^1(\mu)$.

Lema 3.18. Sejam K_1, K_2 e $\psi_n : \mathrm{M} \longrightarrow \mathbb{R}$, $n \geq 1$, uma seqüência de funções em M tal que $\sup \psi_n \leq \mathrm{K}_1$ e $\mathrm{var}\ \psi_n \leq \mathrm{K}_2$ para todo $n \geq 1$. Então, existem uma subseqüência $(\psi_{n_k})_k$ e uma função $\psi_0 : \mathrm{M} \longrightarrow \mathbb{R}$ com $\sup|\psi_0| \leq \mathrm{K}_1$ e $\mathrm{var}\ \psi_0 \leq \mathrm{K}_2$ tais que ψ_{n_k} converge para ψ_0 quando $k \longrightarrow \infty$, μ-quase toda parte em $\mathrm{L}^1(\mu)$.

Demonstração: Consideremos $\mathrm{M} = [0,1]$. Escreva

$$\psi_n^+(x) = \operatorname*{var}_{[0,x]} \psi_n \quad \text{e} \quad \psi_n^- = \psi_n^+ - \psi_n.$$

Então, $(\psi_n^+)_n$ e $(\psi_n^-)_n$ são seqüências de funções não-decrescentes uniformemente limitadas. Escolha uma seqüência de índices $(n_k)_k$ tal que, para todo racional $p \in [0,1]$, quando $k \longrightarrow \infty$,

- $\psi_{n_k}^+(p)$ converge para algum número real $\phi_0^+(p)$; e
- $\psi_{n_k}^-(p)$ converge para algum número real $\phi_0^-(p)$.



Note que $\phi_0^+(p_1) \leq \phi_0^+(p_2)$ e $\phi_0^-(p_1) \leq \phi_0^-(p_2)$ sempre que $p_1 \leq p_2$. Então, estendemos ϕ_0^+ e ϕ_0^- para funções não-decrescentes em todo $[0,1]$, tomando

$$\phi_0^+(x) = \inf\{\phi_0^+(p);\ p \in [x,1] \cap \mathbb{Q}\} \quad \text{e} \quad \phi_0^-(x) = \inf\{\phi_0^-(p);\ p \in [x,1] \cap \mathbb{Q}\}.$$

Afirmação: $\psi_{n_k}^+(x)$ converge para $\phi_0^+(x)$ e $\psi_{n_k}^-(x)$ converge para $\phi_0^-(x)$ quando $k \longrightarrow \infty$, para todo ponto x de continuidade de ϕ_0^+ e ϕ_0^- (um conjunto co-enumerável).

Prova: Dados algum tal x e qualquer $\delta > 0$, devemos fixar números racionais $p_1 \leq x \leq p_2$ tais que

- $\phi_0^+(x) - \delta \leq \phi_0^+(p_1) \leq \phi_0^+(x) \leq \phi_0^+(p_2) \leq \phi_0^+(x) + \delta$.
- $\phi_0^-(x) - \delta \leq \phi_0^-(p_1) \leq \phi_0^-(x) \leq \phi_0^-(p_2) \leq \phi_0^-(x) + \delta$.

Então, para todo k suficientemente grande, obtemos

- $\phi_0^+(x) - 2\delta \leq \phi_0^+(p_1) - \delta \leq \psi_{n_k}^+(p_1) \leq \psi_{n_k}^+(x)$.
- $\phi_0^-(x) - 2\delta \leq \phi_0^-(p_1) - \delta \leq \psi_{n_k}^-(p_1) \leq \psi_{n_k}^-(x)$.

Analogamente, obtemos

- $\psi_{n_k}^+(x) \leq \psi_{n_k}^+(p_2 \leq \phi_0^+(p_2) + \delta \leq \phi_0^+(x) + 2\delta$.
- $\psi_{n_k}^-(x) \leq \psi_{n_k}^-(p_2 \leq \phi_0^-(p_2) + \delta \leq \phi_0^-(x) + 2\delta$.

Isto prova a afirmação.

Sejam ψ_0^+ e ψ_0^- funções contínuas à direita coincidindo, respectivamente, com ϕ_0^+ e ϕ_0^- em todo ponto de continuidade de ϕ_0^+ e ϕ_0^-. Definindo $\psi_0 = \psi_0^+ - \psi_0^-$, segue-se que ψ_{n_k} converge para ψ_0, exceto, possivelmente, num conjunto enumerável de pontos A. Em particular $\psi_{n_k} \longrightarrow \psi_0$ μ-quase toda parte em $L^1(\mu)$. Finalmente, $|\psi_0(x)| = \lim\limits_{k\to\infty} |\psi_{n_k}(x)| \leq \sup\limits_k \sup \psi_{n_k}$ e

$$\sum_{i=1}^{s} |\psi_0(x_i) - \psi_0(x_{i-1})| = \lim_{k\to\infty} \sum_{i=1}^{s} |\psi_{n_k}(x_i) - \psi_{n_k}(x_{i-1})| \leq \sup_k \operatorname*{var}_\eta \psi_{n_k},$$

para todo x e $x_0 \leq x_1 \leq \cdots \leq x_s$ em $[0,1] \setminus A$. Como ψ_0 é contínua à direita, implica que $\sup|\psi_0| \leq K_1$ e $\operatorname{var}|\psi_0| \leq K_2$. ∎

Observação 3.19. A demonstração do Lema 3.18 no caso $M = S^1$ é análoga ao caso $M = [0,1]$.

O Teorema de Helly é de grande utilidade para estender a noção de variação limitada para elementos do espaço $L^1(\mu)$, conforme a seguir.

Definição 3.20. Sejam $\varphi : M \longrightarrow \mathbb{R}$ uma função, η um intervalo em M e $[\varphi]$ um elemento de $L^1(\mu)$. A **variação** de $[\varphi]$ em η é

$$\operatorname*{var}_\eta [\varphi] = \inf \operatorname*{var}_\eta \varphi$$

tomado sobre todas representantes de $[\varphi]$. Dizemos que $[\varphi]$ tem **variação limitada** em η se $\operatorname*{var}_\eta [\varphi] < \infty$.

unesp

Observação 3.21. Uma representante φ torna o inf na Definição 3.20 se, e somente se, φ é contínua nos pontos fronteira de η e $\varphi(z)$ está entre $\lim_{k\to z^-} \varphi(x)$ e $\lim_{k\to z^+} \varphi(x)$ para todo ponto z no interior de η. Além disso, a variação de uma tal função φ sobre qualquer $[a,b] \subset \eta$ coincide com o

$$\sup\left\{\sum_{i=1}^{n} |\ \varphi(x_i) - \varphi(x_{i-1})|;\ n \in \mathbb{N}\ \text{ e }\ a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b\right\}.$$

Definição 3.22. Dizemos que $f : \mathrm{M} \longrightarrow \mathrm{M}$ é uma **transformação expansora por partes** se existe uma partição $\mathcal{P}^1$ cobrindo a medida de Lebesgue total de subconjuntos de M tal que

1. **regularidade:** a restrição de f para cada $\eta_i = (a_{i-1}, a_i) \in \mathcal{P}^1$ é uma transformação monótona de classe C^1, com $|\ \mathrm{D}f(x)| > 0$ para todo $x \in \eta_i$ e $i = 1, \cdots, k$. Além disso, a função

$$g_{\eta_i} = \frac{1}{\left|\ \mathrm{D}f_{|\eta_i}\right|}$$

tem variação limitada para $i = 1, \cdots, k$.

2. **expansividade:** existem constantes $\mathrm{C} > 0$ e $\sigma > 1$ tais que

$$|\ \mathrm{D}f^n(x)| \geq \mathrm{C}\ \sigma^n$$

para todo $n \geq 1$ e todo $x \in \mathrm{M}$, para os quais a derivada é definida.

Observação 3.23. $f_{|\eta_i}$ e g_{η_i} admitem extensões contínuas em $\overline{\eta}_i = [a_{i-1}, a_i]$, para cada $i = 1, \cdots, k$. Assim, modificando os valores de uma transformação sobre um conjunto finito de pontos não mudam as propriedades estatísticas (ergódicas). Devemos assumir que f é contínua à direita, ou contínua à esquerda, ou ambas, em a_i, para cada $i = 1, \cdots, k$. Então, seja $\mathcal{P}^1$ qualquer partição de M por intervalos η tal que $\eta_i \subset \eta \subset \overline{\eta}_i$ para qualquer i e $f_{|\eta_i}$ é contínua. Além disso, para $n \geq 1$, seja $\mathcal{P}^n$ uma **partição por intervalos de regularidade** de f^n, definida por

$$\mathcal{P}^n(x) = \mathcal{P}^n(y) \Longleftrightarrow \mathcal{P}^1\left(f^j(x)\right) = \mathcal{P}^1\left(f^j(y)\right)$$

para todo $0 \leq j < n$, onde $\mathcal{P}^n(x)$ denota o átomo de $\mathcal{P}^n$ que contém x.

Par cada $\xi \in \mathcal{P}^1$, definimos a função $\widehat{g}_\xi : \mathrm{M} \longrightarrow \mathbb{R}$ por

$$\widehat{g}_\xi(x) = \begin{cases} \dfrac{1}{|\ \mathrm{D}f(x)|}, & \text{se } \ x \in \xi \\ 0, & \text{se } \ x \notin \xi \end{cases} \tag{3.1.1}$$

unesp

Observação 3.24. Usando as condições 1 e 2 da Definição 3.22 na definição acima implica que cada $\widehat{g}_\xi$ tem variação limitada.

Para cada $n \geq 1$ e $\eta \in \mathcal{P}^n$, definimos a função $g_\eta^n : \mathrm{M} \longrightarrow \mathbb{R}$ por

$$g_\eta^n(y) = \begin{cases} \dfrac{1}{|\,\mathrm{D}f^n|} \circ \left(f^n|_\eta\right)^{-1}(y), & \text{se} \quad y \in f^n(\eta) \\ 0, & \text{se} \quad y \notin f^n(\eta) \end{cases} \tag{3.1.2}$$

Por simplicidade, escrevemos $g_\eta = g_\eta^1$ para $\eta \in \mathcal{P}^1$.

3.2 Medidas invariantes absolutamente contínuas

Sejam $f : \mathrm{M} \longrightarrow \mathrm{M}$ uma transformação expansora por partes e $\mathcal{P}^1$ uma partição por intervalos de regularidade. Nesta seção vamos assumir que $\mathcal{P}^1$ é finita. Porém, enumerável na Proposição 3.29.

Observação 3.25. Para cada $\eta \in \mathcal{P}^1$, temos

1. $\sup g_\eta = \sup \widehat{g}_\eta$.
2. $\mathrm{var} g_\eta \leq \mathrm{var} \widehat{g}_\eta$.

De fato: Note que $f|_\eta : \eta \longrightarrow f(\eta) \Longrightarrow (f|_\eta)^{-1} : f(\eta) \longrightarrow \eta$. Tomemos $g_\eta = \dfrac{1}{|\,\mathrm{D}f|} \circ (f|_\eta)^{-1} \chi_{f(\eta)}$. Como $\mathrm{var}\chi_{f(\eta)} = 2$, usando as propriedades 2 e 4 do Lema 3.8, obtemos

1. $$\sup g_\eta = \sup_{f(\eta)} \frac{1}{|\,\mathrm{D}f|} \circ (f|_\eta)^{-1} \sup \chi_{f(\eta)} = \sup_\eta \frac{1}{|\,\mathrm{D}f|} \sup \chi_\eta = \sup \widehat{g}_\eta.$$

2. $$\mathrm{var} g_\eta \leq \operatorname*{var}_{f(\eta)} \frac{1}{|\,\mathrm{D}f|} \circ (f|_\eta)^{-1} \sup \chi_{f(\eta)} + \mathrm{var}\chi_{f(\eta)} \sup_{f(\eta)} \frac{1}{|\,\mathrm{D}f|} \circ (f|_\eta)^{-1}$$
$$\leq \operatorname*{var}_\eta \frac{1}{|\,\mathrm{D}f|} + 2 \sup_\eta \frac{1}{|\,\mathrm{D}f|} = \mathrm{var}\widehat{g}_\eta. \qquad \blacksquare$$

Na verdade, as duas variações coincidem, exceto quando $f(\eta) = \mathrm{M}$.

Em geral, dada qualquer função $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$, consideramos $\varphi \circ (f|_\eta)^{-1}$ como um função em M, **identicamente nula**, no complemento de $f^n(\eta)$.

Lema 3.26. Para toda função integrável $\varphi : \mathrm{M} \longrightarrow \mathbb{R}$ e todo $n \geq 1$, o iterado f^n com a medida de Lebesgue pode ser escrito como $f_*^n(\varphi\mu) = \varphi_n \mu$, com

$$\varphi_n = \sum_\eta g_\eta^n \cdot (\varphi \circ (f|_\eta)^{-1})$$

unesp

onde o somatório é tomado sobre todo $\eta \in \mathcal{P}^n$ tal que $\mu(\eta) > 0$.

Demonstração: Seja $A \subset M$ um conjunto mensurável qualquer. Por definição, temos

$$f_*^n(\varphi\mu|_\eta)(A) = \varphi\mu|_\eta(f^{-n}(A)) = \varphi\mu(f^{-n}(A) \cap \eta).$$

Pelo Teorema de Radon-Nikodyn,

$$\varphi\mu(f^{-n}(A) \cap \eta) = \int_{f^{-n}(A)\cap\eta} \varphi(x)\, d\mu(x).$$

Portanto,

$$f_*^n(\varphi\mu|_\eta)(A) = \int_{f^{-n}(A)\cap\eta} \varphi(x)\, d\mu(x). \tag{3.2.3}$$

Fazendo a mudança de variável, $y = (f^n|_\eta)(x)$, obtemos

$$\begin{cases} x = (f^n|_\eta)^{-1}(y) \\ f^{-n}(A) \cap \eta = A \cap f^n(\eta) \\ \dfrac{d\mu(y)}{d\mu(x)} = |\, Df^n| \Longrightarrow d\mu(x) = \dfrac{d\mu(y)}{|\, Df^n|}. \end{cases} \tag{3.2.4}$$

Substituindo (3.2.4) em (3.2.3), obtemos

$$\begin{aligned} f_*^n(\varphi\mu|_\eta)(A) &= \int_{(A)\cap f^n(\eta)} \frac{1}{|\, Df|} \cdot (\varphi \circ (f|_\eta)^{-1}(y))\, d\mu(y) \\ &= \int_A g_\eta^n \cdot (\varphi \circ (f|_\eta)^{-1})\, d\mu. \end{aligned}$$

Como $\varphi \circ (f|_\eta)^{-1}) \equiv 0$ no complemento de $f^n(\eta)$, então

$$\int_A g_\eta^n \cdot (\varphi \circ (f|_\eta)^{-1})\, d\mu \equiv 0 \quad \text{no complemento de} \quad f^n(\eta).$$

Isto prova que a **densidade** para cada $f_*^n(\varphi\mu|_\eta)$ é dada por $g_\eta^n \cdot (\varphi \circ (f|_\eta)^{-1})$, para todo $\eta \in \mathcal{P}^n$. Por outro lado,

$$f_*^n(\varphi\mu) = \sum_{\eta \in \mathcal{P}^n} f_*^n(\varphi\mu|_\eta)$$

para todos os intervalos com $\mu(\eta) > 0$, pois os intervalos com $\mu(\eta) = 0$ têm os correspondentes termos $f_*^n(\varphi\mu|_\eta) \equiv 0$. A conclusão do lema segue pelo somatório de todos os $\eta \in \mathcal{P}^n$ com $\mu(\eta) > 0$. ∎

unesp

Observação 3.27. Por definição, para todo $n \geq 1$, temos

$$\int \varphi_n \, d\mu = \int 1 \, d(f_*^n(\varphi\mu)) = \int 1 \, d(\varphi\mu) = \int \varphi \, d\mu.$$

Seja $|\varphi|_n$ uma seqüência como a obtida no Lema 3.26, onde φ é substituída por $|\varphi|$. Então,

$$-|\varphi|_n \leq \varphi_n \leq |\varphi|_n, \quad \text{visto que} \quad -|\varphi| \leq \varphi \leq |\varphi|.$$

Em particular, para todo $n \geq 1$, temos

$$\int |\varphi_n| \, d\mu \leq \int |\varphi|_n \, d\mu = \int |\varphi| \, d\mu. \tag{3.2.5}$$

Lema 3.28. Para todo $n \geq 1$ e todo $\eta \in \mathcal{P}^n$, existem constantes C_1 e $0 < \lambda_1 < 1$ tais que

1. $\sup g_\eta^n \leq C_1 \lambda_1^n$.
2. $\text{var} g_\eta^n \leq C_1 \lambda_1^n$.

Demonstração: Combinando a propriedade 4 do Lema 3.8 com a condição de expansividade da Definição 3.22 , temos

$$\sup g_\eta^n = \sup_{f(\eta)} \frac{1}{|Df^n|} \circ (f|\eta)^{-1} = \sup_\eta \frac{1}{|Df^n|} \geq \sup_\eta \frac{1}{C_1 \sigma^n}.$$

Tomando $C_1 > \dfrac{1}{C}$ e $\lambda_1 > \dfrac{1}{\sigma}$, obtemos $\sup g_\eta^n \geq \sup_\eta C_1 \lambda_1^n \leq C_1 \lambda_1^n$. Isto prova a propriedade 1.

Dados $\eta \in \mathcal{P}^n$ e $0 \leq j < n$, sejam $\xi_j \in \mathcal{P}^j$, $\eta_j \in \mathcal{P}^1$ e $\zeta_j \in \mathcal{P}^{n-j-1}$ definidos por

$$\eta \subset \xi_j, \quad f^j(\eta) \subset \eta_j \quad \text{e} \quad f^{j+1}(\eta) \subset \zeta_j.$$

Afirmação: Pela definição acima, existe alguma constante $C_2 > 0$ tal que, para todo $\xi \in \mathcal{P}^1$, $\text{var} g_\xi \leq C_2$.

Prova: Pelo ítem 2 da Observação 3.25, temos

$$\text{var} g_\xi \leq \text{var}_\xi \frac{1}{|Df|} + 2 \sup_\xi \frac{1}{|Df|}. \tag{3.2.6}$$

Como o termo sup limita a variação de $\text{var} g_\xi$ nos pontos fronteira de $f(\xi)$, então $\text{var} g_\xi \leq C_2$. Sendo $g_\eta^n = g_{\xi_j}^j \, g_{\eta_j} \, g_{\zeta_j}^{n-j-1}$ para todo $0 \leq j < n$ e usando a propriedade 2 do Lema 3.8, temos

$$\begin{aligned} \text{var} g_\eta^n &\leq \sum_{j=0}^{n-1} \sup g_{\xi_j}^j \, \text{var} g_{\eta_j} \sup g_{\zeta_j}^{n-j-1} \leq \sum_{j=0}^{n-1} C\lambda^j . C_2 . C\lambda^{n-j-1} \\ &= \sum_{j=0}^{n-1} C^2 C_2 \lambda^{n-1} = C^2 C_2 n \lambda^{n-1}. \end{aligned}$$

unesp

Fixando $\lambda_1 > \frac{1}{\lambda}$ e escolhendo $C_1 > \sup\left\{\frac{C^2C_2}{\lambda}\, n\left(\frac{\lambda}{\lambda_1}\right)^n ;\ n \geq 1\right\}$, obtemos que $\mathrm{var} g_\eta^n \leq \lambda_1 C_2$. Isto prova a propriedade 2. ∎

A proposição seguinte é de fundamental importância na parte dos resultados desta seção, a qual afirma que a variação da densidade de φ_n tende a decrescer quando n cresce.

Proposição 3.29. Seja $\varphi : M \longrightarrow \mathbb{R}$ qualquer função de variação limitada. Então, existem constantes C_0 e $0 < \lambda_0 < 1$ tais que

$$\mathrm{var}\varphi_n \leq C_0\lambda_0^n \mathrm{var}\varphi + C_0 \int |\varphi|\, d\mu$$

para todo $n \geq 1$, onde tomamos $\varphi_n = \sum_\eta g_\eta^n \cdot (\varphi \circ (f|\eta)^{-1})$ sobre todo $\eta \in \mathcal{P}^n$ tal que $\mu(\eta) > 0$.

Demonstração: Tome $\varphi_n = \sum_\eta g_\eta^n \cdot (\varphi \circ (f|\eta)^{-1}) \cdot \chi_{f(\eta)}$. Como $\mathrm{var}\chi_{f(\eta)} = 2$ e usando as propriedades 2 e 4 do Lema 3.8, temos

$$\begin{aligned}
\mathrm{var}\varphi_n &\leq \sum_\eta \mathrm{var} g_\eta^n \cdot (\varphi \circ (f|\eta)^{-1}) \sup \chi_{f(\eta)} \\
&+ \sum_\eta \mathrm{var}\chi_{f(\eta)} \sup_{f(\eta)} \left| g_\eta^n \cdot (\varphi \circ (f|\eta)^{-1})\right| \\
&\leq \sum_\eta \mathrm{var} g_\eta^n \cdot \sup_{f(\eta)} |\varphi \circ (f|\eta)^{-1}| + \sum_\eta \mathop{\mathrm{var}}_{f(\eta)} \varphi \circ (f|\eta)^{-1} \sup |g_\eta^n| \\
&+ \sum_\eta 2\sup|g_\eta^n| \sup_{f(\eta)} |\varphi \circ (f|\eta)^{-1}| \\
&\leq \sum_\eta \left\{ \mathrm{var} g_\eta^n \sup_\eta |\varphi| + \mathop{\mathrm{var}}_\eta \varphi \sup g_\eta^n + 2\sup g_\eta^n \sup_\eta |\varphi| \right\} \\
&= \sum_\eta \left\{ \mathrm{var} g_\eta^n \sup_\eta |\varphi| + \sup g_\eta^n (\mathop{\mathrm{var}}_\eta \varphi + 2\sup_\eta |\varphi|) \right\}. \qquad (3.2.7)
\end{aligned}$$

Agora, usando o Lema 3.28, obtemos

$$\begin{aligned}
\mathrm{var}\varphi_n &\leq \sum_\eta \left\{ C_1\lambda_1^n \sup_\eta |\varphi| + C_1\lambda_1^n (\mathop{\mathrm{var}}_\eta \varphi + 2\sup_\eta |\varphi|) \right\} \\
&= \sum_\eta \left\{ 3C_1\lambda_1^n \sup_\eta |\varphi| + C_1\lambda_1^n \mathop{\mathrm{var}}_\eta \varphi \right\}. \qquad (3.2.8)
\end{aligned}$$

Note que, levando em conta a distância na fronteira de $f^n(\eta)$, a expressão $\mathop{\mathrm{var}}_\eta \varphi + 2\sup_\eta |\varphi|)$ é um limite superior para a variação de $\varphi \circ (f|\eta)^{-1}$. Usando

unesp

a propriedade 3 do lema 3.8, temos

$$\begin{aligned}
\operatorname{var}\varphi_n &\leq \sum_\eta \left\{ 3C_1\lambda_1^n \left(\operatorname*{var}_\eta \varphi + \frac{1}{\mu(\eta)} \int_\eta |\varphi| \, d\mu \right) + C_1\lambda_1^n \operatorname*{var}_\eta \varphi \right\} \\
&= \sum_\eta \left\{ 4C_1\lambda_1^n \operatorname*{var}_\eta \varphi + 3C_1\lambda_1^n \frac{1}{\mu(\eta)} \int_\eta |\varphi| \, d\mu \right\} \\
&= 4C_1\lambda_1^n \sum_\eta \operatorname*{var}_\eta \varphi + 3C_1\lambda_1^n \sum_\eta \frac{1}{\mu(\eta)} \int_\eta |\varphi| \, d\mu \\
&= 4C_1\lambda_1^n \operatorname*{var}_\eta \varphi + K(n) \int_\eta |\varphi| \, d\mu, \qquad (3.2.9)
\end{aligned}$$

onde $K(n) = 3C_1\lambda_1^n \left\{ \dfrac{1}{\mu(\eta)};\ \eta \in \mathcal{P}^n \right\}$. Recorde que o somatário é tomado apenas sobre os intervalos $\eta \in \mathcal{P}^n$ com medida de Lebesgue positiva. Para completar a demonstração, devemos eliminar a dependência de K em n. Para isso, fixemos $N \geq 1$ tal que $4C_1\lambda_1^N \leq \dfrac{1}{2}$ e denotemos $K_0 = \{K(n);\ 1 \leq n \leq N\}$. Então, dado qualquer $n \geq 1$, escrevemos $n = qN + r$ com $q \geq 0$ e $1 \leq r \leq N$. Desta forma,

$$\operatorname{var}\varphi_n \leq K_0 \int |\varphi| \, d\mu + \frac{1}{2}\operatorname{var}\varphi.$$

Substituindo φ por $\varphi_{n-N}, \ldots, \varphi_{n-qN}$ e φ, respectivamente, e recordando (3.2.5), obtemos

$$\begin{aligned}
\varphi_n &\leq K_0 \textstyle\int |\varphi_{n-N}| \, d\mu + \dfrac{1}{2}\operatorname{var}\varphi_{n-N} \\
&\leq \left(\frac{1}{2^0} + \frac{1}{2^1} + \frac{1}{2^2} + \cdots + \frac{1}{2^{q-1}} \right) K_0 \textstyle\int |\varphi| \, d\mu + \dfrac{1}{2} \cdot \dfrac{1}{2^{q-1}} \operatorname{var}\varphi \\
&\leq \left(1 + \cdots + \frac{1}{2^q} \right) K_0 \textstyle\int |\varphi| \, d\mu + \dfrac{1}{2^q} 4C_1\lambda_1^r \operatorname{var}\varphi \\
&\leq 2K_0 \textstyle\int |\varphi| \, d\mu + \dfrac{1}{2^q} 4C_1\lambda_1^r \operatorname{var}\varphi.
\end{aligned}$$

Escolhendo $C_0 \geq \max\{ 2K_0, 4C_1 \}$ e $\lambda_0 \geq \max\left\{ \dfrac{1}{2^N}, \lambda_1 \right\}$, concluímos que $\operatorname{var}\varphi_n \leq C_0\lambda_1^n \operatorname{var}\varphi + C_0 \int |\varphi| \, d\mu$. ∎

Observação 3.30. Sejam ν e μ_n, $n \geq 1$, medidas finitas num espaço métrico compacto tais que μ_n é absolutamente contínua com relação a ν para todo $n \geq 1$. Se as derivadas de Radon-Nikodyn $\dfrac{d\mu_n}{d\nu}$ convergem em $L^1(\nu)$ para alguma função ψ, então μ_n converge para $\mu = \psi\nu$ na topologia fraca*.

De fato: Note que $d\mu = \psi d\nu$ e $d\mu_n - d\mu = d\mu_n - \psi d\nu$. Portanto, $d\mu_n - d\mu = \left(\dfrac{d\mu_n}{d\nu} - \psi \right) d\nu$. Dada qualquer função contínua $\varphi : M \longrightarrow \mathbb{R}$,

unesp

obtemos

$$\begin{aligned}\left|\int \varphi\, d\mu_n - \int \varphi\, d\mu\right| &= \left|\int \varphi\,(d\mu_n - d\mu)\right| = \left|\int \varphi \left(\frac{d\mu_n}{d\nu} - \psi\right) d\nu\right| \\ &\leq \int |\,\varphi| \left|\frac{d\mu_n}{d\nu} - \psi\right| d\nu \leq \sup|\,\varphi| \left\|\frac{d\mu_n}{d\nu} - \psi\right\|_1\end{aligned}$$

onde o último termo converge para zero quando $n \longrightarrow \infty$.

Em particular, as densidades ϕ_n das imagens $f_*^n\mu$ da medida de Lebesgue são uniformemente limitadas.

De fato: Pela Proposição 3.29, existe uma constante $C_0 > 0$ tal que $\mathrm{var}\phi_n \leq C_0 \int |\,\phi|\, d\mu$ para todo $n \geq 1$. Usando a propriedade 3 do Lema 3.8 e a Observação 3.27, como μ é uma probabilidade, obtemos

$$\begin{aligned}\sup \phi_n &\leq \mathrm{var}\phi_n + \int \phi_n\, d\mu \leq C_0 \int |\,\phi_n|\, d\mu + \int \phi_n\, d\mu \\ &= C_0 \int |\,1| + \int 1\, d\mu = (C_0 + 1)\int 1\, d\mu = (C_0 + 1) \qquad (3.2.10)\end{aligned}$$

para todo $n \geq 1$.

Corolário 3.31. A transformação $f : M \longrightarrow M$ admite alguma probabilidade invariante ν absolutamente contínua com relação à medida de Lebesgue cuja densidade $\dfrac{d\nu}{d\mu}$ tem variação limitada.

Demonstração: Sejam ϕ_n as densidades das imagens $f_*^n\mu$ da medida de Lebesgue e definimos $\psi_n = \dfrac{1}{n}\displaystyle\sum_{j=0}^{n-1} \phi_j$ para cada $n \geq 1$. Então, $\nu_n = \dfrac{1}{n}\displaystyle\sum_{j=0}^{n-1} f_*^n\mu$ pode ser escrita como $\nu_n = \psi_n\mu$. Pela Proposição 3.29 implica que $(\psi_n)_n$ tem uniformemente variação limitada:

$$\mathrm{var}\psi_n \leq \frac{1}{n}\sum_{j=0}^{n-1} \mathrm{var}\phi_j \leq \frac{1}{n}\sum_{j=0}^{n-1} C_0 \int 1\, d\mu = C_0$$

para todo $n \geq 1$. Note que, usando a Observação 3.27, obtemos

$$\int \psi_n\, d\mu = \frac{1}{n}\sum_{j=0}^{n-1} \int \phi_j\, d\mu = \frac{1}{n}\sum_{j=0}^{n-1} \int 1\, d\mu = 1$$

para todo $n \geq 1$. Logo, por (3.2.10), $(\psi_n)_n$ também tem uniformemente variação limitada:

$$\sup \psi_n \leq \mathrm{var}\psi_n + \int \psi_n\, d\mu \leq C_0 + 1$$

para todo $n \geq 1$. Portanto, pelo Teorema de Helly, existe uma subseqüência $(\psi_{n_k})_k$ convergindo em $L^1(\mu)$ para alguma φ_0 com variação limitada ($\mathrm{var}\varphi_0 \leq C_0$). Em particular, conforme Observação 3.30, $(\psi_{n_k})_k$ converge para $\nu = \varphi_0\mu$ na topologia fraca*. Isto garante que ν é invariante por f. ■



Apêndice A

Elementos de teoria da medida e integração

Neste apêndice recordamos algumas noções e resultados básicos que serão necessários no desenvolvimento deste trabalho. Em particular, faremos uma breve introdução sobre elementos de teoria da medida e integração. Tratando-se, na maioria, de resultados conhecidos, tais demonstrações foram omitidas, uma vez que estas poder ser facilmente encontradas nas referências bibliográficas.

A.1 Espaços mensuráveis

Sejam M um conjunto e $\mathcal{A}$ um subconjunto das partes de M. Dizemos que $\mathcal{A}$ é uma **álgebra** se:

1. $M \in \mathcal{A}$;
2. $A \in \mathcal{A}$ implica $A^c = M \setminus A \in \mathcal{A}$;
3. $A, B \in \mathcal{A}$ implica $A \cup B \in \mathcal{A}$.

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

Uma álgebra é dita uma **σ-álgebra** de subconjuntos de M se, além disso, for fechada para uniões enumeráveis:

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

Então, $A_j \in \mathcal{A}$ para $j = 1, 2, \ldots, n, \ldots$ implica $\bigcap_{j=1}^{\infty} A_j = \left(\bigcup_{j=1}^{\infty} A_j^c \right)^c \in \mathcal{A}$.

Chama-se **σ-álgebra gerada** por uma família $\mathcal{D}$ de subconjuntos de M, a **menor** σ-álgebra que contém $\mathcal{D}$, ou equivalentemente, a **interseção** de todas

as σ-álgebras que contém $\mathcal{D}$. Se M é um **espaço topológico** e $\mathcal{B}$ a σ-álgebra gerada pela família de todos os subconjuntos abertos de M, dizemos que $\mathcal{B}$ é uma **σ-álgebra de Borel** e aos seus elementos **borelianos** de M.

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

A.2 Espaços de medida

Uma **medida** num espaço mensurável $(M, \mathcal{A})$ é uma função $\mu : \mathcal{A} \longrightarrow [0, +\infty]$ tal que

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

Um **espaço de medida** é uma terna $(M, \mathcal{A}, \mu)$, onde $(M, \mathcal{A})$ é um espaço mensurável e μ é uma medida definida em $\mathcal{A}$. O espaço de medida diz-se **finito** quando $\mu(M) < \infty$. Dizemos que μ é uma **probabilidade** se $\mu(M) = 1$ e, nesse caso, o espaço de medida diz-se **normalizado** ou **espaço de probabilidade**.

Seja $(M, \mathcal{A}, \mu)$ um espaço de medida. Dizemos que a medida μ é **σ-finita** se existe uma seqüência crescente $(A_n)_{n \in \mathbf{N}}$ de elementos de $\mathcal{A}$ tais que

$$M = \bigcup_{n=1}^{\infty} A_n \quad \text{e} \quad \mu(A_n) < \infty \quad \text{para todo } n.$$

Neste caso, diremos que um ponto x **vai para infinito** se, para qualquer n, existe apenas **um número finito** de iterados de x que estão em A_n.

A propriedade 2 na definição de medida é chamada **σ-aditiva**. Dizemos que uma função $\mu : \mathcal{A} \longrightarrow [0, +\infty]$ é **finitamente aditiva** se

- $\mu\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} \mu(A_j)$ para qualquer família finita $A_1, \ldots, A_n \in \mathcal{A}$ de subconjuntos disjuntos dois-a-dois.

Observação A.1. Toda função σ-aditiva é finitamente aditiva.

Sejam $(M, \mathcal{A}, \mu)$ um espaço de medida e A um subconjunto de M. Dizemos que A tem **medida nula** se existe $B \in \mathcal{A}$ tal que $B \subset A$ e $\mu(B) = 0$. Uma propriedade sobre os elementos de M vale em **quase todo ponto** se, o conjunto dos pontos onde ele não vale tem medida nula.

Proposição A.2. Sejam $\mathcal{A}$ uma álgebra e $\mu : \mathcal{A} \longrightarrow [0, +\infty]$ uma função σ-aditiva.

1. Se $A, B \in \mathcal{A}$ tais que $A \subset B$, então $\mu(A) \leq \mu(B)$.



2. Se $(A_n)_n$ é uma seqüência de elementos de $\mathcal{A}$ tais que $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, então

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Demonstração: Ver referência [3], p.23.

Proposição A.3. Sejam $(M, \mathcal{A}, \mu)$ um espaço de medida e $(A_n)_n$ uma seqüência de elementos de $\mathcal{A}$.

1. Se $A_1 \subset A_2 \subset \ldots$ então $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$.

2. Se $A_1 \supset A_2 \supset \ldots$ e $\mu(A_1) < \infty$ então $\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$.

Demonstração: Ver referência [1], p.21.

A.3 Extensão

O resultado a seguir é muito útil para construir medidas.

Proposição A.4 (Teorema de Extensão). Se $\mathcal{A}_0$ é uma álgebra de subconjuntos e $\mu_0 : \mathcal{A}_0 \longrightarrow [0, +\infty]$ é finitamente aditiva, então existe uma única função finitamente aditiva $\mu : \mathcal{A} \longrightarrow [0, +\infty]$ que é uma extensão de μ_0 (isto é: μ restrita a $\mathcal{A}_0$ coincide com μ_0) à σ-álgebra $\mathcal{A}$ gerada por $\mathcal{A}_0$. Se μ_0 é σ-aditiva então μ também será. Se, além disso, $\mu_0(M) = 1$ então μ é uma probabilidade.
Demonstração: Ver referência [3], p.29.

Em geral, ao aplicar este resultado, o mais difícil é verificar a propriedade σ-aditiva. O critério mais usado para esse efeito é a proposição seguinte.

Proposição A.5 (Teorema de σ-aditividade). Sejam $\mathcal{A}_0$ uma álgebra e $\mu_0 : \mathcal{A}_0 \longrightarrow [0, +\infty]$ uma função finitamente aditiva com $\mu_0(M) = 1$. Suponha que $\mu\left(\bigcap_{j=1}^{\infty} A_j\right) = 0$ para toda seqüência $A_1 \supset A_2 \supset \ldots A_j \supset \ldots$ de conjuntos mensuráveis tal que $\bigcap_{j=1}^{\infty} A_j = \emptyset$. Então, μ_0 é σ-aditiva.
Demonstração: Ver referência [3].

Também temos que todo elemento de A da σ-álgebra gerada por uma álgebra é aproximado por algum elemento A_0 da álgebra, no sentido de que a medida da **diferença simétrica** $A \triangle A_0 = (A \setminus A_0) \cup (A_0 \setminus A)$ tem medida pequena.

unesp

Proposição A.6 (Teorema de Aproximação). Sejam $(M, \mathcal{A}, \mu)$ um espaço de probabilidade e $\mathcal{A}_0$ uma álgebra que gera a σ-álgebra de $\mathcal{A}$. Então, para todo $\varepsilon > 0$ e $A \in \mathcal{A}$, existe $A_0 \in \mathcal{A}$ tal que $\mu(A \triangle A_0) < \varepsilon$.
Demonstração: Ver referência [3], p.34.

A.4 Medida de Lebesgue

Consideremos $M = [0, 1]$ e seja $\mathcal{A}_0$ a família de todos os subconjuntos da forma $A = I_1 \cup \cdots \cup I_n$, onde $I_1, \ldots, I_n$ são intervalos disjuntos dois-a-dois. Note que $\mathcal{A}_0$ é uma álgebra de subconjuntos de M. Além disso, temos uma função $\mu_0 : \mathcal{A}_0 \longrightarrow [0, 1]$ definida nesta álgebra por

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

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

Note que $\mu_0(M) = 1$. Além disso, a σ-álgebra gerada por $\mathcal{A}_0$ coincide com a σ-álgebra de Borel de M, já que todo aberto pode ser escrito como união enumerável de intervalos disjuntos dois-a-dois. Pela proposição A.4, existe uma única probabilidade μ definida na σ-álgebra de $[0, 1]$ que é uma extensão de μ_0 à σ-álgebra $\mathcal{A}$ gerada por $\mathcal{A}_0$. 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 I_d são intervalos, e definimos

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

Em seguida, consideremos a álgebra $\mathcal{A}_0$ dos subconjuntos 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 $\mathcal{A}_0$, que coincide com a σ-álgebra boreliana de M. Finalmente, definimos a medida de Lebesgue num espaço euclidiano $\mathbb{R}^d$ decompondo o espaço em cubos de lado unitário

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

e definindo, para cada subconjunto mensurável E,

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

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

A.5 Funções integráveis

Sejam $(M, \mathcal{A})$ um espaço mensurável e $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. Dizemos que uma função $f : M \longrightarrow \overline{\mathbb{R}}$ é **mensurável**, se para todo boreliano A de $\mathbb{R}$ tivermos $f^{-1}(A) \in \mathcal{A}$. São exemplos imediatos de funções mensuráveis, as funções constantes e as funções características dos elementos de $\mathcal{A}$. Se M for um espaço topológico e $\mathcal{A}$ for a σ-álgebra de Borel, então as funções contínuas são mensuráveis.

Observação A.7. Para as operações com os símbolos $+\infty$ e $-\infty$, além das convenções usuais, faremos as seguintes convenções: $(\pm\infty) \cdot 0 = 0$ e $0 \cdot (\pm\infty) = 0$. Não atribuiremos significado a $\infty - \infty$.

Proposição A.8. Se $c \in \mathbb{R}$ e $f, g : M \longrightarrow \overline{\mathbb{R}}$ são funções mensuráveis, então são mensuráveis: $f + c$, cf, $f + g$ (sempre que façam sentido) e fg.
Demonstração: Ver referência [1], p.33.

Proposição A.9. Se $(f_n)_n$ é uma seqüência de funções mensuráveis, então são mensuráveis:

1. $\sup_{n\geq 1} f_n$ e $\inf_{n\geq 1} f_n$;
2. $\lim_{n\to+\infty} \sup f_n$ e $\lim_{n\to+\infty} \inf f_n$.

Demonstração: Ver referência [3], p.55.

Seja $(M, \mathcal{A}, \mu)$ um espaço de medida. Dizemos que uma seqüência de funções mensuráveis $(f_n)_n$ **converge em medida** para a função mensurável f se:

$$\lim_{n\to\infty} \mu(\{x \in M;\ |\ f_n(x) - f(x)| \geq \varepsilon\}) = 0 \quad \text{para todo } \varepsilon > 0.$$

As convergências em medida e em quase todo ponto relacionam-se pelo resultado a seguir.

Teorema A.10. Toda seqüência que converge em medida possui uma subseqüência que converge em quase todo ponto. Se o espaço de medida for finito, a convergência em quase todo ponto implica a convergência em medida.
Demonstração: Ver referência [1].

Sejam $(M, \mathcal{A}, \mu)$ um espaço de medida e $A \subset M$. Denotamos por χ_A a **função característica** de A. Dizemos que uma função mensurável não negativa φ é uma **função simples**, se pudermos escrever

$$\varphi = \sum_{i=1}^{n} a_i \chi_{A_i},$$

unesp

onde $a_i \in \mathbb{R}^+$ e $A_i \in \mathcal{A}$ são disjuntos dois-a-dois. Definimos a **integral** de uma função simples $\varphi = \sum_{i=1}^{n} a_i \chi_{A_i}$ como

$$\int \varphi \, d\mu = \sum_{i=1}^{n} a_i \, \mu(A_i).$$

Este valor não depende da representação de φ como combinação linear das funções características. Se f é uma função mensurável não negativa, definimos a **integral** de f como

$$\int f \, d\mu = \sup \left\{ \int \varphi \, d\mu; \ \varphi \text{ é uma função simples e } \varphi \leq f \right\}.$$

Segue-se imediatamente da definição apresentada que se f e g são funções mensuráveis não negativas, então

$$f \leq g \ \text{ implica que } \ \int f \, d\mu \leq \int g \, d\mu.$$

Lema A.11 (Fatou). Seja $(f_n)_n$ uma seqüência de funções reais mensuráveis, limitadas inferiormente por uma função integrável. Se $\liminf_{n\to\infty} \int f_n \, d\mu < \infty$ então $\liminf_{n\to\infty} f_n$ é integrável e

$$\int \liminf_{n\to\infty} f_n \, d\mu \leq \liminf_{n\to\infty} \int f_n \, d\mu.$$

Demonstração: Ver referência [1], p.33.

Corolário A.12. Seja f uma função mensurável não negativa. Então, $\int f \, d\mu = 0$ se, e somente se, $f = 0$ quase todo ponto.
Demonstração: Ver referência [1], p.34.

O resultado a seguir dá uma maneira alternativa de definir a integral de uma função mensurável.

Corolário A.13. Se f é uma função mensurável não negativa, então existe uma seqüência crescente $(\varphi_n)_n$ de funções simples convergindo pontualmente para f com

$$\lim_{n\to\infty} \int \varphi_n \, d\mu = \int f \, d\mu.$$

Demonstração: Ver referência [1], p.35.

O resultado a seguir dá uma condição suficiente para que a integral do limite de uma seqüência de funções coincide com o limite das integrais dessas funções.



Teorema A.14 (Convergência Monótona). Se $(f_n)_n$ é uma seqüência de funções mensuráveis não negativas tais que $f_1 \leq f_2 \leq \ldots$, então

$$\int \lim_{n\to\infty} f_n \, d\mu = \lim_{n\to\infty} \int f_n \, d\mu.$$

Demonstração: Ver referência [3], p.70.

Teorema A.15. Sejam $f_1, f_2, \cdots : \mathrm{M} \longrightarrow [0, +\infty]$ funções mensuráveis. Então

$$\int \sum_{n=1}^{\infty} f_n \, d\mu = \sum_{n=1}^{\infty} \int f_n \, d\mu.$$

Demonstração: Ver referência [19], p.22.

Agora, vamos generalizar a noção de integral para funções mensuráveis que podem tomar valores negativos e introduzir o conceito de função integrável. Dada uma função mensurável f, consideremos respectivamente, as suas **parte positiva** e **parte negativa**,

$$f^+(x) = \max\{f(x), 0\} \quad \text{e} \quad f^-(x) = \max\{-f(x), 0\}.$$

Desta forma, f pode ser escrita como diferença de duas funções não negativas, mais precisamente, $f = f^+ - f^-$. Dizemos que uma função mensurável f é **integrável** se

$$\int f^+ \, d\mu < \infty \quad \text{e} \quad \int f^- \, d\mu < \infty.$$

No caso de alguma destas duas integrais ser finita, definimos a **integral** de f como

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu.$$

Note que $\|f| = f^+ + f^-$. Portanto, f é integrável se, se somente se, $\|f\|$ é integrável. Dados um conjunto mensurável A e uma função mensurável f, dizemos que f é integrável em A se $f\chi_{\mathrm{A}}$ for integrável. Nesse caso, definimos o integral de f em A por

$$\int_{\mathrm{A}} f \, d\mu = \int f\chi_{\mathrm{A}} \, d\mu.$$

Proposição A.16. Sejam $c \in \mathbb{R}$ e f, g funções integráveis.

1. cf é integrável e $\int cf \, d\mu = c \int f \, d\mu$.
2. $f + g$ é integrável e $\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu$.
3. $f \leq g$ implica $\int f \, d\mu \leq \int g \, d\mu$.
4. Se A e B são conjuntos mensuráveis disjuntos, então

$$\int_{\mathrm{A}\cup\mathrm{B}} f \, d\mu = \int_{\mathrm{A}} f \, d\mu + \int_{\mathrm{B}} g \, d\mu.$$



Demonstração: Ver referência [3].

Proposição A.17. Se f é uma função integrável, então $\left|\int f\ d\mu\right| \leq \int |f|\ d\mu$, e temos a igualdade se, e somente se, $f \geq 0$ quase todo ponto ou $f \leq 0$ quase todo ponto.
Demonstração: Ver referência [1], p.43.

Teorema A.18 (Convergência Dominada). Seja $(f_n)_n$ uma seqüência de funções integráveis convergindo quase todo ponto para f tais que $|f_n| \leq g$, onde g é integrável. Então, f é integrável e

$$\lim_{n\to\infty} \int f_n\ d\mu = \int f\ d\mu.$$

Demonstração: Ver referência [3], p.75.

O teorema seguinte caracteriza a relação entre a noção de integral de uma função real de variável real segundo Lebesgue (em relação à medida de Lebesgue nos borelianos de $\mathbb{R}$) com a noção de função integrável segundo Riemann.

Teorema A.19. Se f é integrável segundo Riemann, então f é integrável segundo Lebesgue e as integrais coincidem.
Demonstração: Ver referência [3].

A.6 Teorema de Radon-Nikodym

Sejam ν e μ duas medidas num espaço mensurável $(\mathrm{M}, \mathcal{A})$. Dizemos que ν é **absolutamente contínua** com relação a μ se, qualquer que seja o conjunto mensurável A, $\mu(\mathrm{A}) = 0$ implica $\nu(\mathrm{A}) = 0$. Nesse caso, escrevemos $\nu \ll \mu$. As medidas dizem-se **equivalente** se $\mu \ll \nu$ e $\nu \ll \mu$. Se $(\mathrm{M}, \mathcal{A}, \mu)$ é um espaço de medida e f é uma função não negativa integrável, então

$$\nu : \mathrm{A} \longmapsto \int_{\mathrm{A}} f\ d\mu.$$

define uma medida finita em $(\mathrm{M}, \mathcal{A})$ que é absolutamente contínua com relação a μ.

Teorema A.20. Sejam ν e μ medidas finitas num espaço mensurável $(\mathrm{M}, \mathcal{A})$ tais que $\nu \ll \mu$, então existe uma função mensurável $\rho : \mathrm{M} \longrightarrow [0, +\infty]$ tal que $\nu = \rho\mu$, ou seja, tal que

$$\nu(\mathrm{A}) = \int_{\mathrm{A}} \rho\ d\mu \quad \text{para todo conjunto mensurável } \mathrm{A} \subset \mathrm{M}.$$

Além disso, tal função é essencialmente única: duas funções quaisquer coincidem num conjunto com medida total para μ.
Demonstração: Ver referência [3], p.133.

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

Observação A.21. A função ρ dada no teorema acima chama-se **derivada de Radon-Nikodym** ou **densidade** da medida ν com relação a μ e denotada-se por $\frac{d\nu}{d\mu}$.

A.7 Teorema de decomposição ergódico

Na seqüência dos resultados da subseção 1.7, é natural perguntar se toda medida invariante é uma combinação linear de medidas ergódicas. O teorema que enunciaremos nesta seção afirma que a resposta é afirmativa, exceto quando o número de **parcelas** nesta combinação não é necessariamente finito, nem mesmo enumerável em geral.

Teorema A.22. Seja $f : M \longrightarrow M$ uma transformação contínua num espaço compacto. Então existem um conjunto mensurável $M_0 \subset M$, uma partição $\mathcal{P}$ de M_0 e uma família de probabilidade $\{\nu_P;\ P \in \mathcal{P}\}$ satisfazendo

1. $\nu_P(P) = 1$ para todo elemento P de $\mathcal{P}$;
2. A aplicação $P \longmapsto \nu_P$ é mensurável;
3. toda ν_P é invariante e ergódica para f;

tais que, dada qualquer probabilidade μ invariante por f, o conjunto M_0 tem $\mu(M_0) = 1$. Além disso,

$$\mu(A) = \int \nu_P(A)\, d\widehat{\mu}(P) \quad \text{para todo conjunto mensurável } A \subset M, \qquad (A.7.1)$$

onde $\widehat{\mu}$ é a medida projeção de μ em $\mathcal{P}$.

A relação (A.7.1) significa que μ é uma combinação convexa das várias probabilidades ergódicas ν_P, em que cada ν_P entra com "coeficiente"igual a $\widehat{\mu}(P)$. De momento, só faremos alguns comentários explicativos do enunciado. Dada qualquer partição $\mathcal{P}$ de M fica definida a projeção natural $\pi : M \longrightarrow \mathcal{P}$ que associa a cada ponto $x \in M$ o elemento $P(x)$ da partição que o contém. Isto permite definir o que é um subconjunto mensurável da partição: $\Omega \subset \mathcal{P}$ é mensurável, se e somente se,

$$\pi^{-1}(\Omega) = \text{união dos } P \in \Omega$$

é um subconjunto mensurável de M. É fácil ver que esta definição está correta: a família dos subconjuntos mensuráveis é uma σ-álgebra em $\mathcal{P}$. A medida projeção de μ está definida nesta σ-álgebra por

$$\widehat{\mu}(\Omega) = \mu(\pi^{-1}(\Omega)).$$

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

A.8 Teorema de derivação de Lebesgue

Consideremos M $= \mathbb{R}^d$ com $d \geq 1$ e $m =$ medida de Lebesgue (volume usual). Seja A um subconjunto mensurável de um espaço euclidiano M. Um ponto p$\in$ M diz-se **ponto de densidade** de A se,

$$\liminf_{\varepsilon \to 0} \left\{ \frac{m(\mathrm{B} \cap \mathrm{A})}{m(\mathrm{B})};\ \mathrm{B} \text{ bola contida em } \mathrm{B}(\mathrm{p}, \varepsilon) \text{ e contendo p} \right\} = 1. \quad \text{(A.8.2)}$$

Note que $\mathrm{B}(\mathrm{p}, \varepsilon)$ representa a bola de centro p e raio ε. É usual escrever

$$\lim_{\epsilon \to 0} \frac{m(\mathrm{B}(\mathrm{p}, \varepsilon) \cap \mathrm{A})}{m(\mathrm{B}(\mathrm{p}, \varepsilon))} = 1 \quad \text{(A.8.3)}$$

no lugar da condição (A.8.2).

Observação A.23.

- Um ponto p não precisa pertencer a A para ser ponto de densidade de A.
- Se p $\in$ M é ponto de densidade de A, então o conjunto A preenche a maior parte do volume de bolas na vizinhança de p que tenha um raio suficientemente pequeno.

Teorema A.24. Seja A um subconjunto de um espaço euclidiano M com medida de Lebesgue $m(\mathrm{A}) > 0$. Então, m-quase todo ponto p $\in$ A é ponto de densidade de A.

Demonstração: Dado qualquer subconjunto B de M, defina $\mu(\mathrm{B}) = m(\mathrm{B} \cap \mathrm{A})$. Daí, μ é uma medida de Lebesgue positiva absolutamente contínua com relação a m, o que implica que em m-quase todo ponto $p \in \mathrm{M}$, existem os limites

$$\lim_{\epsilon \to 0} \frac{\mu(\mathrm{B}(\mathrm{p}, \varepsilon)}{m(\mathrm{B}(\mathrm{p}, \varepsilon))} = \lim_{\epsilon \to 0} \frac{m(\mathrm{B}(\mathrm{p}, \varepsilon) \cap \mathrm{A})}{m(\mathrm{B}(\mathrm{p}, \varepsilon))} = D_\mu(p) = \frac{d\mu}{dm}(p),$$

essa última igualdade (válida em m-quase todo ponto $p \in \mathrm{M}$), devido ao fato de que $\mu \ll m$. Note que para todo $f \in \mathrm{L}^1(m)$, vale

$$\int_{\mathrm{M}} f\, dm = \int_{\mathrm{A}} f\, dm = \int f \chi_{\mathrm{A}}\, dm.$$

Logo, pela unicidade da derivada de Radon-Nikodym, vale $\dfrac{d\mu}{dm}(p) = \chi_{\mathrm{A}}(p)$. ■

A.9 Espaços L^p

Seja $(\mathrm{M}, \mathcal{A}, \mu)$ uma espaço de medida. Dado $1 \leq p \leq \infty$, define-se o **espaço de funções** $\mathrm{L}^p(\mu)$ como a classe das funções mensuráveis f tais que

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18

$|f|^p$ é integrável, com a identificação de duas funções que coincidem em μ-quase todo ponto. Definimos para $f \in \mathrm{L}^p(\mu)$

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

Definimos o **espaço de funções** $\mathrm{L}^\infty(\mu)$ como a classe das funções mensuráveis f tais que existe algum $\mathrm{K} > 0$ tal que $|f(x)| \leq \mathrm{K}$ μ-quase todo ponto, mais uma vez com a identificação de duas funções que coincidem em μ-quase todo ponto. Denotamos por $\|f\|_\infty$ o ínfimo dos valores K com esta propriedade. Mais precisamente, definimos

$$\|f\|_\infty = \inf \{\mathrm{K} > 0; \ |f(x)| \leq \mathrm{K}\}.$$

Resulta do modo como definimos estes espaços que, para $1 \leq p \leq \infty$, qualquer $f \in \mathrm{L}^p$ está identificada como uma função mensurável que nunca toma os valores $\pm\infty$.

Referências Bibliográficas

[1] Bartle, Robert G. *The elements of intregation and Lebesgue measure.* John Wiley & Sons Inc., New York, 1995.

[2] D'Ambrosio, Ubiratan. *Métodos da topologia: uma introdução e aplicações.* Rio de Janeiro, LTC, 1977.

[3] Fernandez, P. J. *Medida e integração.* Projeto Euclides. IMPA, Rio de Janeiro, 1976.

[4] Halmos, Paul R. *Lectures on ergodic thorey.* Chelsea Publising Company, New York, 1956.

[5] Katznelson, Y e Weiss, B. *A simple proof of some ergodic theorems.* Israel Journal of Mathematics 42: 291-296, 1982.

[6] Lima, Elon Lages. *Curso de análise I.* Projeto Euclides. IMPA/CNPq, Rio de Janeiro, 1982.

[7] Lima, Elon Lages. *Elementos de topologia geral.* Ao Livro Técnico, S.A., Rio de Janeiro, 1970.

[8] Lima, Elon Lages. *Espaços métricos.* Projeto Euclides. IMPA/CNPq, Rio de Janeiro, 1983.

[9] Lequain, Y. A. E. *Aproximação de um número real por números racionais.* XIX-CBM, IMPA, Rio de Janeiro, 1993.

[10] Machado, A. *Introdução à análise funcional.* Escola Editora, Lisboa, 1991.

[11] Mañé, R. *Teoria ergódica.* Projeto Euclides. IMPA, Rio de Janeiro, 1983.

[12] Munckres, James Raymond. *Topology: a first couse.* Springer-Verlag, New York, 1974.

[13] M, Rychlik. *Bounded variation and invariante measures.* Studia Mathematics 76: 68-80, 1983.

[14] Viana, M. *Stochastic dynamics of deterministic systems.* Lecture Notes XXI-CBM, IMPA, Rio de Janeiro, 1997.

unesp

[15] Viana, M. *Introdução à teoria ergódica.* Escola de verão 2003, Departamento de Matemátaica-UFPe, Recife - PE.

[16] Viana, M. *Dynamics: a probabilistic and geometric perspective.* In Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998), volume 1998, pages 557-578.

[17] Viana, M. *Towards a theory of chaotic dynamics.* Preprint 2000.

[18] Viana, M. *Mathematics Unlimited - 2001 and Beyond.* Springer Verlag, 2000.

[19] Rudin, W. *Real and complex analysis.* McGrawo-Hill, third edition, 1987.

[20] Walters, Peter. *An introduction to ergodic theory.* Springer-Verlag, New York, 1982.

cm 1 2 3 4 5 6 7 unesp 11 12 13 14 15 16 17 18