Universidade Federal de Alagoas
Instituto de Matemática
Programa de Pós-Graduação em Matemática
Dissertação de Mestrado

Desigualdade de Ruelle

Davi dos Santos Lima

Maceió, Brasil
Setembro de 2012

Davi dos Santos Lima

Desigualdade de Ruelle

Dissertação de Mestrado na área de concentração em Sistemas Dinâmicos submetida em 27 de Novembro de 2012 à banca examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Matemática da Universidade Federal de Alagoas, como parte dos requisitos necessários à obtenção do grau de mestre em Matemática.

Orientador: Prof. Dr. Krerley Irraciel Martins Oliveira.

Maceió
2012

Banca Examinadora:

Prof. Dr. Krerley Irraciel Martins Oliveira (Orientador)

Prof. Dr. Walter Huaraca Vargas

Prof. Dr. Carlos Bocker

Agradecimentos

Agradeço a DEUS pelas oportunidades e a Maria por sua interseção.

Aos meus pais e familiares, por sempre me acolher e me dá forças.

Ao meu orientador, Krerley Oliveira, que acreditou que eu seria capaz no momento mais difícil da minha graduação, pela amizade e pelo companheirismo.

A todos os professores e funcionários do IM que direta ou indiretamente contribuiram para o meu avanço matemático.

A todos os meus amigos, matemáticos e não matemáticos, os primeiros, certamente me ajudaram a crescer profissionalmente com conversas sobre matemática e os não matemáticos pelos momentos descontraídos, divertidos e pelas palavras de apoio.

Ao CNPq pelo apoio financeiro.

Enfim, a todos que estiveram comigo em muitos momentos de incertezas, sou muito grato!

Resumo

Neste trabalho, revisaremos fatos básicos de Teoria Ergódica, um importante invariante métrico que é a Entropia, apresentaremos uma ferramenta útil para o estudo do comportamento dinâmico de uma aplicação, os Expoentes de Lyapunov sobre uma variedade compacta e relacionamos a entropia com os expoentes via a Desigualdade de Ruelle, fato principal do trabalho.

Palavras-chave: Teoria Ergódica; Entropia; Expoentes de Lyapunov; Desigualdade de Ruelle.

Abstract

In this work, we review basic facts of Ergodic Theory, an important metric invariant, Entropy, we present a useful tool for studying the dynamic behavior of an application, the Lyapunov exponents, on a compact manifold and relate entropy with exponents via the Ruelle's inequality, because the main work.

Sumário

Capítulo 1

Introdução

Nesta dissertação vamos abordar como se relaciona duas formas de estudar o comportamento de um sistema (f, μ). Um deles, a entropia, analisa o comportamento do sistema com a informação entre as órbitas de pontos do espaço considerado; a outra, os expoentes de Lyapunov, analisa o comportamento assintótico da derivada dos iterados dos pontos, i.e., um estudo geométrico do sistema. Naturalmente, gostaríamos de relacionar os dois conceitos. Para estudar o comportamento do sistemas com propriedades geométricas, fixaremos o nosso ambiente, uma variedade compacta. Para definirmos o que vem a ser entropia, precisamos de uma partição do ambiente em que estamos, pois vamos entender dois pontos como iguais se eles estão no mesmo elemento da partição. Isso se mostrará bastante conveniente, pois iterando os pontos podemos distinguí-los a partir do momento que seus iterados distinguem. Com essa abordagem, temos um perspectiva puramente dinâmica e probabilística do sistema a ser estudado. Em 1978, no Boletim da Sociedade Brasileira de Matemática, o matemático David Ruelle, publicou um artigo no qual o principal resultado é uma relação muito bonita entre os dois conceitos. Para isso ele supunha um sistema invariante, (f, μ) de classe C^1 definido sobre sobre uma variedade compacta C^∞. Aqui, reproduziremos uma das demonstrações da Desigualdade de Ruelle, tal como a relação ficou conhecida, e para tanto definiremos e veremos algumas propriedades de entropia e expoentes de Lyapunov.

Capítulo 2

Um Pouco de Teoria Ergódica

2.1 Medidas Invariantes

Fixemos M um conjunto, $\mathcal{S}$ σ-álgebra de subconuntos de M. Um par (f, μ) será dito invariante se $\mu(f^{-1}(A)) = \mu(A)$, onde $f : M \to M$ é uma aplicação mensurável e A é um conjunto qualquer mensurável. Diremos ainda que μ é f-invariante ou que f é μ-invariante quando a condição acima ocorrer. A uma dupla $(M, \mathcal{S})$ como acima dá-se o nome de *espaço mensurável*, e quando a um espaço mensurável acrescenta-se uma medida μ a tripla $(M, \mathcal{S}, \mu)$ é chamado *espaço de medida*.

Teorema 2.1.1 (Recorrência de Poincaré). *Seja $(M, \mathcal{S}, \mu)$ um espaço de medida, $f : M \to M$ uma tranformação mensurável e $A \in \mathcal{S}$ um conjunto com medida positiva. Quase todo ponto $x \in A$ retorna a A.*

Corolário 2.1.1.1. *Nas condições do Teorema de Recorrência de Poincaré, quase todo ponto $x \in A$ retorna infinitas vezes a A.*

Capítulo 3

Mais um pouco de Teoria Ergódica

3.1 Entropia

Em Teoria da Medida, é de costume desprezar conjuntos cuja medida é nula. Escreve-se $A = B(mod.0)$ quando $\mu(A\Delta B) = 0$.

Definição 3.1.1. *Dado um espaço de medida $(M, \mathcal{B}, \mu)$, uma partição do mesmo, é uma coleção de conjuntos mensuráveis $\xi = \{\xi_\alpha\}_{\alpha\in\Lambda}$ indexados por algum conjunto não vazio de índices, Λ, satisfazendo*

$$\bigcup_{\alpha\in\Lambda} \xi_\alpha = M(mod.0) \text{ e } \xi_\alpha \cap \xi_\beta = \emptyset,\ \alpha \neq \beta.$$

Observação 3.1.1. *Veremos logo mais, que sempre poderemos para os nossos objetivos, usar Λ finito.*

Dadas duas partições ξ e η, denotaremos $\xi \prec \eta$ para significar que todo elemento de η está contido em algum elemento de ξ. Com duas partições, ξ e η, podemos definir a soma $\xi \vee \eta := \{\xi_i \cap \eta_j, \xi_i \in \xi, \eta_j \in \eta\}$ que satisfaz portanto, $\xi \prec \xi \vee \eta$ e $\eta \prec \xi \vee \eta$.

Seja $f : M \to M$ uma transformação mensurável. Dada uma partição ξ de M, temos outra partição determinada por f: $f^{-1}(\xi) := \{f^{-1}(\xi_i); \xi_i \in \xi\}$. Dada uma família enumerável de partições $\mathcal{P}_n$ definimos

$$\bigvee_{n\in\mathbb{N}} \mathcal{P}_n = \{\cap_n P_n; P_n \in \mathcal{P}_n\}.$$

Denotamos

$$\xi^n = \bigvee_{j=0}^{n-1} f^{-j}(\xi), \forall n \geq 1.$$

Para μ-q.t.p $x \in M$ existe um elemento de ξ que o contém, tal elemento será denotado por $\xi(x)$. Para cada $n \geq 1$ o elemento $\xi^n(x)$ que contém $x \in M$ está dado por:

$$\xi^n(x) = \xi(x) \cap f^{-1}(\xi(f(x))) \cap ... \cap f^{-(n-1)}(\xi(f^{-(n-1)}(x))).$$

Note que a sequência ξ^n é não decrescente, ou seja, $\xi^n \prec \xi^{n+1}$ para todo n.

Para uma dada partição ξ, definimos $I_\xi : M \to \mathbb{R} \cup \{\infty\}$ por $I_\xi(x) = -\log \mu(\xi(x))$. Claramente I_ξ é mensurável.

A continuidade da função $\varphi : [0,1] \to [0,1]$, $\varphi(x) = -x \log x$ nos permite convencionar que $0 \log 0 = 0 = \lim_{x\to 0} \varphi(x)$. Com essa convenção definimos

Definição 3.1.2. *A entropia de ξ é*

$$H_\mu(\xi) = \int I_\xi(x) d\mu(x) = -\sum_{C\in\xi} \mu(C) \log \mu(C)$$

.

Uma noção importante de entropia é o de entropia condicional conforme a definição abaixo.

Definição 3.1.3. *Sejam ξ e η partições de M, a entropia de ξ dado η é*

$$\begin{aligned} H_\mu(\xi/\eta) &= -\sum_{D\in\eta} \mu(D) \sum_{C\in\xi} \frac{\mu(C\cap D)}{\mu(D)} \log \frac{\mu(C\cap D)}{\mu(D)} \\ &= -\sum_{C\in\xi, D\in\eta} \mu(C\cap D) \log \frac{\mu(C\cap D)}{\mu(D)}. \end{aligned} \tag{3.1}$$

Tomando $\eta = \{\emptyset, M\}$ obtemos que $H_\mu(\xi/\eta) = H_\mu(\xi)$

Lema 3.1.1. *Sejam ξ, η, ζ partições com entropia finita. Valem,*

1. $H_\mu(\xi \vee \eta/\zeta) = H_\mu(\xi/\zeta) + H_\mu(\eta/\xi \vee \zeta)$.
2. $\xi \prec \eta$ *então* $H_\mu(\xi/\zeta) \leq H_\mu(\eta/\zeta)$ *e* $H_\mu(\zeta/\xi) \geq H_\mu(\zeta/\eta)$.
3. *Se* $\xi \prec \eta$ *então* $H_\mu(\xi/\eta) = 0$
4. *Se* μ *é* f*-invariante então* $H_\mu(f^{-1}(\xi)/f^{-1}(\eta)) = H_\mu(\xi/\eta)$

Demonstração. Vamos começar provando *1*:

$$\begin{aligned} H_\mu(\xi \vee \eta/\zeta) &= \sum_{C\in\xi, D\in\eta, E\in\zeta} -\mu(C\cap D\cap E) \log \frac{\mu(C\cap D\cap E)}{\mu(E)} \\ &= \sum_{C\in\xi, D\in\eta E\in\zeta} -\mu(C\cap D\cap E) \log \frac{\mu(C\cap D\cap E)}{\mu(C\cap E)} + \\ &+ \sum_{C\in\xi, D\in\eta, E\in\zeta} -\mu(C\cap D\cap E) \log \frac{\mu(C\cap E)}{\mu(E)} \end{aligned} \tag{3.2}$$

Usamos a propriedade do logarítmo do produto. A soma do lado direito pode ser reescrita como

$$\sum_{F\in\xi\vee\zeta, D\in\eta} -\mu(F\cap D)\log\frac{\mu(F\cap D)}{\mu(F)} + \sum_{C\in\xi, E\in\zeta} -\mu(C\cap E)\log\frac{\mu(C\cap E)}{\mu(E)},$$

mas isto é por definição, $H_\mu(\eta/\xi\vee\zeta)+H_\mu(\xi/\zeta)$. Isto termina o item *1*.

Para demonstração de *2*, veja que se $\xi\prec\eta$ então

$$\begin{aligned} H_\mu(\xi/\zeta) &= \sum_{C\in\xi, E\in\zeta}\sum_{D\subset C, D\in\eta} -\mu(D\cap E)\log\frac{\mu(C\cap E)}{\mu(E)} \\ &\leq \sum_{C\in\xi, E\in\zeta}\sum_{D\subset C, D\in\eta} -\mu(D\cap E)\log\frac{\mu(D\cap E)}{\mu(E)} = H_\mu(\eta/\zeta). \end{aligned} \tag{3.3}$$

Para a demonstração de 3 veja que se $\xi\prec\eta$ então cada elemento de η está em algum elemento de ξ, portanto $\mu(C\cap D)=\mu(D)$ ou $\mu(C\cap D)=0$, em qualquer um desses casos temos,

$$\mu(C\cap D)\log\frac{\mu(C\cap D)}{\mu(D)}=0,$$

donde, $H_\mu(\xi/\eta)=0$.

O ítem 4 decorre da definição, pois

$$\mu(f^{-1}(C\cap D)\log\frac{\mu(f^{-1}(C\cap D))}{\mu(f^{-1}(D))} = \mu(C\cap D)\log\frac{\mu(C\cap D)}{\mu(D)}$$

. □

A partir da observação feita depois da definição de $H_\mu(\xi/\eta)$ quando $\eta=\{\emptyset,M\}$ vemos que o ítem 1 do lema acima de reduz a

$$H_\mu(\xi\vee\eta)=H_\mu(\xi)+H_\mu(\eta/\xi), \tag{3.4}$$

isso é obtido tomando $\zeta=\{\emptyset,M\}$.

Além disso, do ítem 2 e da observação anterior, temos

$$H_\mu(\xi\vee\eta)\leq H_\mu(\xi)+H_\mu(\eta) \tag{3.5}$$

Lema 3.1.2. *Dado $k\geq 1$ e $\epsilon>0$ existe $\delta>0$ tal que para quaisquer partições finitas $\xi=\{\xi_1,...,\xi_k\}$ e $\eta=\{\eta_1,...,\eta_k\}$,*

$$\mu(\xi_i\Delta\eta_i)<\delta\ \forall i=1,...,k\Rightarrow H_\mu(\eta/\xi)<\epsilon.$$

Demonstração. Fixe $\epsilon>0$ e $k\geq 1$. Pela continuidade da função $\phi:[0,1]\to\mathbb{R}$ dada por $\phi(x)=-x\log x$ existe $\rho>0$ tal que $\phi(x)<\epsilon/k^2$ para todo $x\in[0,\rho)\cup(1-\rho,1]$. Tome $\delta=\rho/k$. Dadas as partições ξ e η como no enunciado, seja ζ a partição de M cujos elementos são os cojuntos da forma $\xi_i\cap\eta_j$ quando $i\neq j$ e o conjunto $\cup_{i=1}^k\xi_i\cap\eta_i$. Note que $\xi_i\cap\eta_j\subset\bigcup_{i=1}^k\xi_i\Delta\eta_i$ e dai, $\mu(\xi_i\cap\eta_j)\leq\sum_{i=1}^k\mu(\xi_i\Delta\eta_i)<k\delta=\rho$ sempre que $i\neq j$ e desde que $A\cap B=A\cup B-A\Delta B$ temos

$$\mu(\bigcup_{i=1}^{k}\xi_i \cap \eta_i) \geq \sum_{i=1}^{k}\mu(\xi_i) - \mu(\xi_i \Delta \eta_i) > \sum_{i=1}^{k}\mu(\xi_i) - \rho = 1-\rho.$$

Portanto,

$$H_\mu(\zeta) = \sum_{\zeta' \in \zeta}\phi(\mu(\zeta')) < \mathrm{card}(\zeta)\epsilon/k^2 \leq \epsilon.$$

Da definição temos $\xi \vee \eta = \xi \vee \zeta$, e daí

$$H_\mu(\eta/\xi) = H_\mu(\xi \vee \eta) - H_\mu(\xi) = H_\mu(\xi \vee \zeta) - H_\mu(\xi),$$

a expressão acima é $H_\mu(\zeta/\xi) \leq H_\mu(\zeta) < \epsilon$. □

A função $\phi(x) = -x\log x$ é côncava em $[0,1]$. Podemos usar essa propriedade para mostrar que se ξ é uma partição finita, então $H_\mu(\xi) \leq \log \mathrm{card}(\xi)$.

Com efeito, suponha $\mathrm{card}(\xi) = n$. Temos

$$\begin{aligned}
\frac{1}{n}H_\mu(\xi) &= \frac{1}{n}\sum_{C\in\xi}\phi(\mu(C)) \\
&\leq \phi(\frac{1}{n}\sum_{C\in\xi}\mu(C)) \\
&= \phi(\frac{1}{n}) \\
&= -\frac{1}{n}\log\frac{1}{n} \qquad (3.6)
\end{aligned}$$

donde, $H_\mu(\xi) \leq \log n$.

Lema 3.1.3. *Se $\{a_n\}_{n\geq 0}$ é uma sequência de números reais satisfazendo $a_{n+p} \leq a_n + a_p$ para todo $n, p \geq 1$ então $\lim\limits_{n\to\infty}\dfrac{a_n}{n}$ existe e é igual a $\inf_n a_n/n$. O limite pode ser $-\infty$, mas se $\dfrac{a_n}{n}$ for limitada inferiormente, o limite será $c > -\infty$.*

Demonstração. Fixe $p > 0$. Cada $n > 0$ pode escrito como $n = kp + i$ com $0 \leq i < p$ (aqui k depende de n). Observe que $a_{kp} \leq a_k + a_{k(p-1)} \leq 2a_k + a_{k(p-2)} \leq ... \leq ka_k$. Segue-se que

$$\frac{a_n}{n} = \frac{a_{kp+i}}{kp+i} \leq \frac{a_i}{kp} + \frac{a_{kp}}{kp} \leq \frac{a_i}{kp} + \frac{ka_p}{kp} = \frac{a_i}{kp} + \frac{a_p}{p}$$

Agora veja que quando $n \to \infty$ temos que $k \to \infty$. Dessa forma, como a_i permanece limitado temos

$$\limsup_{n\to\infty}\frac{a_n}{n} \leq \frac{a_p}{p},$$

logo,

$$\limsup_{n\to\infty}\frac{a_n}{n} \leq \inf\frac{a_p}{p}.$$

Por outro lado, $\inf_p \frac{a_p}{p} \leq \frac{a_n}{n}$ qualquer que seja n, e portanto,

$$\inf \frac{a_p}{p} \leq \liminf_{n\to\infty} \frac{a_n}{n}.$$

Isto mostra que $\lim_{n\to\infty} \frac{a_n}{n}$ existe e vale $\inf_n \frac{a_n}{n}$. □

Lema 3.1.4. *$H_\mu(\xi^{m+n}) \leq H_\mu(\xi^m) + H_\mu(\xi^n)$ para todo $m, n \geq 1$*

Demonstração. Temos $\xi^{m+n} = \bigvee_{j=0}^{m+n-1} f^{-j}(\xi) = \xi^m \vee f^{-m}(\xi^n)$. Pelo Lema 1.1.1 temos

$$H_\mu(\xi^{m+n}) \leq H_\mu(\xi^m) + H_\mu(f^{-m}(\xi^n)).$$

Como μ é f-invariante temos a desigualdade desejada. □

A partir do Lema anterior e do Lema 3.1.3 temos que o limite

$$\lim_{n\to\infty} \frac{1}{n} H_\mu(\xi^n)$$

existe. Este será denotado por $h_\mu(f,\xi)$, e chamado de *entropia de f com respeito a ξ.*

Exemplo 3.1.1. *Seja $f : [0,1] \to [0,1]$ a função dada por $f(x) = 10x(mod.1) = 10x - [10x]$, onde $[10x]$ representa o maior inteiro menor ou igual a $10x$. Para cada $i = 0, ...9$ faça $\xi_i = [i/10, i+1/10]$. Seja $\xi = \{\xi_0, ..., \xi_9\}$. Como se $x \in f^{-1}(\xi_i)$ então $f(x) \in \xi_i$ temos que $\xi^2 = \xi \vee f^{-1}(\xi) = \{\xi_j, j = 0, ..., 99\}$, onde $\xi_j = [j/10^2, (j+1)/10^2]$. Para o caso geral, não é difícil notar que $\xi^n = \{\xi_j; j = 0, ..., 10^n - 1$, onde $\xi_j = [j/10^n, (j+1)/10^n]\}$. Seja m a medida de Lebesgue do intervalo $[0,1]$. Veja que $H_m(\xi^n) = \sum_{j=0}^{10^n-1} m(\xi_j) \log \mu(\xi_j)$. Mas, $m(\xi_j) = 10^{-n}$, disto*

$$\begin{aligned} H_m(\xi^n) &= \sum_{j=0}^{10^n-1} 10^{-n} \log 10^{-n} \\ &= n 10^{-n} \log 10 \sum_{j=0}^{10^n-1} 1 \\ &= n \log 10 \end{aligned} \tag{3.7}$$

Segue-se de 3.7 que $h_\mu(f,\xi) = \log 10$.

Teorema 3.1.1. *Se $f : M \to M$ preserva μ e ξ é uma partição de M então $\frac{1}{n} H_\mu(\xi^n)$ decresce para $h_\mu(f,\xi)$.*

Demonstração. Vamos mostrar inicialmente que,

$$H_\mu(\xi^n) = H_\mu(\xi) + \sum_{j=1}^{n-1} H_\mu(\xi / f^{-1}(\xi^j)).$$

De fato, para $n = 2$, temos

$$H_\mu(\xi^2) = H_\mu(\xi \vee f^{-1}(\xi)) = H_\mu(\xi) + H_\mu(\xi/f^{-1}(\xi)),$$

conforme a observaçao (1.4). Assuma validade para $n = p \geq 2$. Para $n = p + 1$ temos,

$$\begin{aligned} H_\mu(\xi^{p+1}) &= H_\mu(f^{-1}(\xi^p) \vee \xi) \\ &= H_\mu(f^{-1}(\xi^p)) + H_\mu(\xi/f^{-1}(\xi^p)) \\ &= H_\mu(\xi^p) + H_\mu(\xi/f^{-1}(\xi^p)) \\ &= H_\mu(\xi) + \sum_{j=1}^{p} H_\mu(\xi/f^{-1}(\xi^j)). \end{aligned} \tag{3.8}$$

A última passagem vem da hipótese de indução. Desta fórmula e do ítem 2 do lema 3.1.1 temos $H_\mu(\xi^n) \geq nH_\mu(\xi/f^{-1}(\xi^n))$ e daí,

$$\begin{aligned} H_\mu(\xi^{n+1}) &= n[H_\mu(\xi^n) + H_\mu(\xi/f^{-1}(\xi^n))] \\ &\leq (n+1)H_\mu(\xi^n). \end{aligned} \tag{3.9}$$

Ou seja,

$$\frac{1}{n+1}H_\mu(\xi^{n+1}) \leq \frac{1}{n}H_\mu(\xi^n)$$

□

Definição 3.1.4. *O supremo de $h_\mu(f, \xi)$ sobre todas as partições com entropia finita é a entropia do sistema (f, μ), e é denotado por $h_\mu(f)$.*

Exemplo 3.1.2. *Sejam $f : M \to M$ e $x \in M$ ponto periódico de f, i.e., $f^p(x) = x$ para algum $p \in \mathbb{N}$. Considere*

$$\mu = \frac{1}{p}(\delta_x + \delta_{f(x)} + ... + \delta_{f^{p-1}(x)}).$$

Neste caso, a medida toma somente um número finito de valores. Consequentemente, a entropia $H_\mu(\xi)$ toma somente um número finito de valores quando consideramos as partições enumeráveis de M. Em particular, $\lim_{n\to\infty} n^{-1}H_\mu(\xi^n) = 0$ para toda ξ. Disto, $h_\mu(f) = 0$.

Lema 3.1.5. *$h_\mu(f, \eta) \leq h_\mu(f, \xi) + H_\mu(\eta/\xi)$ para quaisquer partições ξ e η de M com entropia finita.*

Demonstração. Sabemos que $H_\mu(\xi \vee \eta/\zeta) = H_\mu(\xi/\zeta) + H_\mu(\eta/\xi \vee \zeta)$. Segue-se que para quaisquer $n \geq 1$ temos

$$\begin{aligned} H_\mu(\eta^{n+1}/\xi^{n+1}) &= H_\mu(\eta^n \vee f^{-n}(\eta)/\xi^n \vee f^{-n}(\xi)) \\ &= H_\mu(\eta^n/\xi^n \vee f^{-n}(\xi)) + H_\mu(f^{-n}(\eta)/\eta^n \vee \xi^n \vee f^{-n}(\xi)) \\ &\leq H_\mu(\eta^n/\xi^n) + H_\mu(f^{-n}(\eta)/f^{-n}(\xi)) \end{aligned} \tag{3.10}$$

aqui estamos usando o Ítem 2 do Lema 3.1.1. Como μ é f-invariante, temos do ítem 4 do Lema 3.1.1 que $H_\mu(f^{-n}(\eta)/f^{-n}(\xi)) = H_\mu(\eta/\xi)$. Portanto, se $a_k = H_\mu(\eta^k/\xi^k)$ temos $a_{k+1} - a_k \leq a_1$ e portanto $a_n \leq na_1$, i.e.,

$$H_\mu(\eta^n/\xi^n) \leq nH_\mu(\eta/\xi).$$

Assim, $H_\mu(\eta^n) \leq H_\mu(\xi^n \vee \eta^n) = H_\mu(\xi^n) + H_\mu(\eta^n/\xi^n) \leq H_\mu(\xi^n) + nH_\mu(\eta/\xi)$. Dividindo por n e tomando o limite obtemos a conclusão.

$\square$

Com o Lema acima, vamos justificar a observação 3.1.1, para mostrar que a definição de $h_\mu(f)$ pode ser feita considerando o supremo sobre todas partições finitas.

Lema 3.1.6. $h_\mu(f) = \sup\{h_\mu(f,\xi); \xi \text{ finita}\}$.

Demonstração. Seja $\widetilde{h}_\mu(f) = \sup\{h_\mu(f,\xi); \xi \text{ finita}\}$. Claramente, $\widetilde{h}_\mu(f) \leq h_\mu(f)$. Para mostrar a desigualdade contrária, dado $\epsilon > 0$, existe $\xi = \{C_1, C_2, ...\}$ uma partição com entropia finita, tal que $h_\mu(f) < h_\mu(f,\xi) + \epsilon/2$. Desde que ξ tem entropia finita, existe $N \in \mathbb{N}$ tal que $\sum_{i>N} -\mu(C_i)\log\mu(C_i) < \epsilon/2$.

Tome $\eta = \{C_1, C_2, ..., C_N, \cup_{j>N} C_j\}$. Veja que $H_\mu(\eta/\xi) = \sum_{i>N} -\mu(C_i)\log\mu(C_i) + \mu(\cup_{j>N}C_j)\log\mu(\cup_{j>N}C_j) < \epsilon/2$, e do Lema 3.1.5 temos

$$h_\mu(f) < h_\mu(f,\xi) + \epsilon/2 \leq h_\mu(f,\eta) + H_\mu(\eta/\xi) + \epsilon/2 < h_\mu(f,\eta) + \epsilon.$$

Como η é finita, temos que $h_\mu(f) \leq \widetilde{h}_\mu(f)$. $\square$

Observação 3.1.2. *Ainda do Lema 3.1.5 temos que se $\xi \prec \eta$ então $h_\mu(f,\xi) \leq h_\mu(f,\eta)$, uma vez que pelo Lema 3.1.1 sabemos que $H_\mu(\xi/\eta) = 0$.*

Lema 3.1.7. $h_\mu(f,\xi) = \lim_{n\to\infty} H_\mu(\xi/\bigvee_{j=1}^n f^{-j}(\xi))$ *para qualquer partição ξ com entropia finita.*

Demonstração. Sabemos do Teorema 3.1.1 que

$$H_\mu(\xi^n) = H_\mu(\xi) + \sum_{k=1}^{n-1} H_\mu(\xi/f^{-1}(\xi^k)).$$

Logo, $h_\mu(f,\xi) = \lim\limits_n \dfrac{1}{n} H_\mu(\xi^n) = \lim\limits_n \dfrac{1}{n} \displaystyle\sum_{k=0}^{n-1} H_\mu(\xi/f^{-1}(\xi^k))$. Por lado, $\{H_\mu(\xi/f^{-1}(\xi^n)\}_n$ é uma sequência decrescente e então seu limite coincide com o limite à Cesaro. $\square$

Quando f é inversível chamamos $\xi^{\pm n} = \bigvee_{j=-n}^{n-1} f^j(\xi)$.

Lema 3.1.8. *Se ξ é uma partição com entropia finita então $h_\mu(f,\xi) = h_\mu(f,\xi^k)$ para todo $k \geq 1$. Se f é inversível então $h_\mu(f,\xi) = h_\mu(f,\xi^{\pm k})$ para todo $k \geq 1$*

Demonstração. Veja que para qualquer $n \geq 1$ temos

$$\bigvee_{j=0}^{n-1} f^{-j}(\xi^k) = \bigvee_{j=0}^{n-1} f^{-j}(\bigvee_{i=0}^{k-1} f^{-i}(\xi)) = \bigvee_{j=0}^{n+k-1} f^{-j}(\xi) = \xi^{n+k}.$$

Assim, $h_\mu(f, \xi^k) = \lim_n \frac{1}{n} H_\mu(\xi^{n+k}) = \lim_n \frac{n+k}{n} \frac{1}{n+k} H_\mu(\xi^{n+k}) = h_\mu(f, \xi)$

No caso inversível,

$$\bigvee_{j=0}^{n-1} f^{-j}(\xi^{\pm k}) = \bigvee_{j=0}^{n-1} f^{-j}(\bigvee_{i=-k}^{k-1} f^{-i}(\xi)) = \bigvee_{l=-k}^{n+k-1} f^{-l}(\xi) = f^{-k}(\xi^{n+2k}),$$

para todo n e para todo k. Portanto,

$$h_\mu(f, \xi^{\pm k}) = \lim_n \frac{1}{n} H_\mu(f^{-k}(\xi^{n+2k})) = \lim_n \frac{n+2k}{n} \frac{1}{n+2k} H_\mu(\xi^{n+2k}) = h_\mu(f, \xi).$$

□

Proposição 3.1.1 (Fórmula de Abramov). *$h_\mu(f^k) = kh_\mu(f)$ para todo $k \in \mathbb{N}$. Se f é inversível $h_\mu(f^k) = |k|h_\mu(f)$ para todo $k \in \mathbb{Z}$.*

Demonstração. Para uma melhor compreensão na demonstração, denotaremos

$$\xi_f^k = \xi \vee f^{-1}(\xi) \vee ... \vee f^{-(k-1)}(\xi).$$

Observe que

$$\begin{aligned}
(\xi_f^k)_{f^k}^m &= (\bigvee_{j=0}^{k-1} f^{-j}(\xi))_{f^k}^m \\
&= \bigvee_{i=0}^{k-1} f^{-i}(\xi) \vee \bigvee_{i=k}^{2k-1} f^{-i}(\xi) \vee ... \vee \bigvee_{i=mk-m}^{mk-1} f^{-i}(\xi) \\
&= \bigvee_{i=0}^{mk-1} f^{-i}(\xi) \\
&= \xi_f^{km}
\end{aligned} \tag{3.11}$$

Considere $g = f^k$ e tome ξ uma partição de M com entropia finita. Portanto, ξ_f^k também tem entropia finita, ja que $H_\mu(\xi_f^k) \leq kH_\mu(\xi)$.

Assim,

$$kh_\mu(f, \xi) = \lim_m \frac{1}{m} H_\mu(\xi_f^{km}) = \lim_m \frac{1}{m} H_\mu((\xi_f^k)_g^m) = h_\mu(g, \xi_f^k).$$

Tomando o supremo sobre todas as partições ξ obtemos, $h_\mu(f^k) \geq kh_\mu(f)$. Por outro lado, pela observação 3.1.2, desde que $\xi \prec \xi_f^k$ temos $h_\mu(g, \xi) \leq h_\mu(g, \xi_f^k) = kh_\mu(f, \xi)$

tomando o supremo sobre ξ com entropia finita obtemos $h_\mu(f^k) \le k h_\mu(f)$. Com isto concluímos o caso não inversivel.

Para o caso inversível, tome ξ com entropia finita e observe que para todo $n \ge 1$

$$H_\mu(\vee_{j=0}^{n-1} f^{-j}(\xi)) = H_\mu(f^{-n+1}(\vee_{i=0}^{n-1} f^i(\xi))) = H_\mu(\vee_{j=0}^{n-1} f^j(\xi)),$$

pois μ é f-invariante. Segue-se que dividindo por n e tomando o limite $h_\mu(f,\xi) = h_\mu(f^{-1},\xi)$. Tomando o supremo sobre todas as partiçoes ξ obtemos $h_\mu(f) = h_\mu(f^{-1})$. Trocando f por f^k temos, $h_\mu(f^{-k}) = h_\mu(f^k) = k h_\mu(f)$ para todo k natural. □

Teorema 3.1.2. *Seja $\xi_1 \prec \xi_2 \prec ... \prec \xi_n \prec ...$ uma sequência não decrescente de partições com entropia finita tais que $\cup_{n\in\mathbb{N}}\xi_n$ gera a σ-álgebra dos conjuntos mensuráveis. Então*

$$h_\mu(f) = \lim_{n\to\infty} h_\mu(f,\xi_n).$$

Lema 3.1.9. $\lim H_\mu(\eta/\xi_n) = 0$ *para qualquer partição finita η.*

Demonstração. Seja $\eta = \{\eta_1, ..., \eta_k\}$. Dado $\epsilon > 0$, fixe δ como no Lema 3.1.2. Seja $\mathcal{A}$ a álgebra formada pelas uniões $\cup_n\xi_n$. Por hipótese, $\mathcal{A}$ gera a σ-álgebra dos mensuráveis de M. Portanto, para cada $i = 1, ..., k$ existe A_i tal que

$$\mu(\eta_i \Delta A_i) < \delta/(4k) \tag{3.12}$$

Disto, para cada $i = 1, ..., k$ temos,

$$\mu(A_i \cap \cup_{j\neq i} A_j) \le \mu(\cup_{j=1}^k A_j \Delta \eta_j) < \delta/4, \tag{3.13}$$

e

$$\mu(M \setminus \cup_{i=1}^k A_i) = \mu(\cup_{i=1}^k \eta_i \setminus A_i) < \delta/4. \tag{3.14}$$

Agora defina

$$\eta_i' = \begin{cases} A_i & \text{se} \quad i = 1 \\ A_i \setminus \cup_{j=1}^{i-1} A_j & \text{se} \quad 1 < i < k \\ M \setminus \cup_{j=1}^{k-1} A_j & \text{se} \quad i = k \end{cases}$$

Temos que $\eta' = \{\eta_1', ..., \eta_k'\}$ é uma partição de M. Afirmamos que

$$\mu(A_i \Delta \eta_i') < \delta/2, \ \forall i = 1, 2, ..., k. \tag{3.15}$$

De fato, vale para $i = 1$ trivialmente da definição de η_1'. Para $1 < i < k$ temos que $A_i \setminus \eta_i' = A_i \cap \cup_{j<i} A_j$ e usando 3.13 temos que $\mu(A_i \Delta \eta_i') < \delta/4$, já que $\eta_i' \setminus A_i = \emptyset$. Finalmente, para $i = k$ temos que $\eta_k' \setminus A_k$ está contido no complementar de $\cup_{i=1}^k A_i$ e usando 3.14 temos $\mu(\eta_k' \setminus A_k) < \delta/4$. Segue-se que $\mu(\eta_k' \Delta A_k) < \delta/2$. Logo,

$$\mu(\eta_i \Delta \eta_i') \le \mu(\eta_i \Delta A_i) + \mu(A_i \Delta \eta_i') < \delta \tag{3.16}$$

para todo $i = 1, 2, ..., k$. É óbvio que $\eta_i' \in \mathcal{A}$ para todo i. Então, como $\mathcal{A}$ é gerada por $\cup_n \xi_n$, podemos encontrar m tal que todo η_i' é uma união finita de elementos de ξ_m. Em outras palavras, $\eta' \prec \xi_m$. Disto, usando os Lemas 3.1.1, a equação 3.16 e 3.1.2, temos

$$H_\mu(\eta/\xi_n) \leq H_\mu(\eta/\xi_m) \leq H_\mu(\eta/\eta') < \epsilon, \ \forall m \geq n.$$

Isto demonstra o lema. □

Vamos demostrar o teorema:

Demonstração. Pelo Lema 3.1.5 temos

$$h_\mu(f,\eta) \leq h_\mu(f,\xi_n) + H_\mu(\eta/\xi_n) \ \forall n.$$

Fazendo $n \to \infty$ e tomando o supremo sobre todas as partições η temos acabado o teorema. □

O seguinte corolário é o importante teorema de Kolmogorov-Sinai.

Corolário 3.1.2.1 (Kolmogorov-Sinai)**.** *Seja ξ uma partição cujos iterados ξ^n geram a σ-álgebra dos mensuráveis. Então $h_\mu(f) = h_\mu(f,\xi)$.*

Demonstração. Basta aplicar o teorema acima para a sequência ξ^n, usando que $h_\mu(f,\xi) = h_\mu(f,\xi^n)$. □

Analogamente o temos quando $f : M \to M$ é inversível. Como o

Corolário 3.1.2.2. *Suponha que $f : M \to M$ é inversível. Seja ξ uma partição com entropia finita tal que a $\cup_n \xi^{\pm n}$ gera a σ-álgebra dos mensuráveis de M. Então, $h_\mu(f) = h_\mu(f,\xi)$.*

Demonstração. A prova é análoga ao corolário anterior, mas usando que $h_\mu(f,\xi^{\pm n}) = h_\mu(f,\xi)$ para todo $n \geq 1$. □

Corolário 3.1.2.3. *Suponha $f : M \to M$ inversível e que existe ξ com entropia finita tal que $\cup_{n\in\mathbb{N}}\xi^n$ gera a σ-álgebra dos mensuráveis de M. Então, $h_\mu(f) = 0$*

Demonstração. De fato, sabemos que $h_\mu(f) = h_\mu(f,\xi) = \lim_n H_\mu(\xi/f^{-1}(\xi^n))$, usando o Corolário 3.1.2.1 e o Lema 3.1.7. Como $\cup_n\xi^n$ gera a σ-álgebra $\mathcal{B}$ dos mensuráveis de M, $\cup_n f^{-1}(\xi^n)$ gera a σ-ágebra $f^{-1}(\mathcal{B})$, e por f ser inversível $f^{-1}(\mathcal{B}) = \mathcal{B}$. Usando o Lema 3.1.9 temos que $\lim_n H_\mu(\xi/f^{-1}(\xi^n)) = 0$. Portanto, $h_\mu(f) = 0$. □

Suponha que M é um espaço métrico e esteja munido da σ-álgebra de Borel.

Corolário 3.1.2.4. *Seja $\xi_1 \prec \xi_2 \prec ... \prec \xi_n$... uma sequência não-decrescente de partições com entropia finita e $lim_n \mathrm{diam}\xi(x) = 0$ para μ-quase todo $x \in M$. Então*

$$h_\mu(f) = \lim_n h_\mu(f,\xi_n).$$

Demonstração. Seja U um aberto qualquer de M. A hipótese garate que para cada x existe $n(x)$ para o qual $\xi_x = \xi_{n(x)}(x)$ está contido em U. Claramente ξ_x pertence a álgebra $\mathcal{A}$ gerada por $\cup_n\xi_n$. Observe também que esta álgebra é enumerável, já que ela está formada pelas uniões finitas de elementos de ξ_n. Em particular, existe uma quantidade enumerável de ξ_x's, donde $U = \cup_x\xi_x$ está em $\mathcal{A}$. Isto mostra que a σ-álgebra gerada por $\cup_n\xi_n$ contém os borelianos. A conclusão, portanto, segue-se aplicando o Teorema 3.1.2. □

Exemplo 3.1.3. *Seja* $f : [0,1] \to [0,1]$ *a função dada por* $f(x) = 10x(mod.1)$ *e* ξ *como no exemplo 3.1.1. Vemos que* $\lim_{n\to\infty} \operatorname{diam}(\xi^n) = 0$*, donde* ξ *é geradora. Logo, de 3.1.2.4 e de 3.1.2.1 temos que* $h_m(f) = \log 10$.

Exemplo 3.1.4. *Seja* $f : S^1 \to S^1$ *um homeomorfismo e* μ *uma probabilidade invariante por* f*. Dada qualquer partição finita* ξ *de* S^1 *em subintervalos, denotemos por* $x_1, ..., x_m$ *os seus pontos extremos. Para qualquer* $j \geq 1$*, a partição* $f^{-j}(\xi)$ *é formada pelos subintervalos cujos extremos são os pontos* $f^{-j}(x_i)$*. Isto implica que para cada* $n \geq 1$*, os elementos de* ξ^n *têm os seus pontos extremos no conjunto*

$$\{f^{-j}(x_i); j = 0, 1, ..., n-1 \text{ e } i = 1, ..., m\}.$$

Em particular, $card(\xi) \leq mn$*. Segue-se que*

$$h\mu(f,\xi) = \lim \frac{1}{n} H_\mu(\xi^n) \leq \frac{1}{n} \log card(\xi^n) \leq \frac{1}{n} \log mn = 0.$$

Disto, $h_\mu(f) = 0$*, pois qualquer partição finita está na condições do Corolário 3.1.2.4.*

Dizemos que $f : M \to M$ é expansiva se existe $\epsilon > 0$(*constante de expansividade*) tal que

$$d(f^j(x), f^j(y)) \leq \epsilon, \ \forall j \geq 0 \Rightarrow x = y.$$

Equivalentemente, $f : M \to M$ é expansiva se existe $\epsilon > 0$ tal que dados $x \neq y$ existe $n \in \mathbb{N}$ tal que $d(f^n(x), f^n(y)) \geq \epsilon$.

Quando a transformação é inversivel existe uma versão análoga, trocando $\mathbb{N}$ por $\mathbb{Z}$ na definição anterior.

Proposição 3.1.2. *Seja* $f : M \to M$ *expansiva num espaço métrico compacto e* $\epsilon > 0$ *uma contante de expansividade. Então tem-se* $lim_n \operatorname{diam}(\xi^n) = 0$ *para toda partição finita com* $\operatorname{diam}(\xi) < \epsilon$

Demonstração. A sequência $\operatorname{diam}(\xi^n)$ é não crescente. Seja δ o seu ínfimo e suponha que $\delta > 0$. Então, para todo $n \geq 1$ existem pontos x_n e y_n tais que $d(x_n, y_n) > \delta/2$ mas x_n e y_n pertencem ao mesmo elemento de ξ^n, e portanto satisfazem,

$$d(f^j(x_n), f^j(y_n)) \leq \operatorname{diam}(\xi) < \epsilon, \ \forall j = 0, ..., n-1.$$

Por compacidade existe $\mathbb{N}_1 \subset \mathbb{N}$ tal que existe $\lim_{n\in\mathbb{N}_1} x_n = x$ e $\mathbb{N}_2 \subset \mathbb{N}_1$ para o qual existe $\lim_{n\in\mathbb{N}_2} y_n = y$. Temos, $x \neq y$ mas $d(f^j(x), f^j(y)) < \epsilon$ para todo $j \geq 0$, contrariando a hipótese de expansividade. $\square$

A proposição anterior mostra que toda transformação expansiva admite partição geradora.

3.2 Decomposição Ergódica

Um sistema (f,μ) é ergódico se todo conjunto f-invariante E, i.e. $f^{-1}(E) = E$, satistaz $\mu(E)\mu(E^c) = 0$.

Faça $\mathcal{M}_1(f)$ o conjunto das probabilidades invariantes por f e $\mathcal{M}_e(f)$ o conjunto das probabilidades para as quais (f,μ) é ergódico.

Uma medida ν é absolutamente contínua com respeito a μ se $\mu(E) = 0$ implica $\nu(E) = 0$. Nesse caso, escrevemos $\nu \ll \mu$. Observe que $\nu \ll \mu$ e $\mu \ll \rho$ então $\nu \ll \rho$.

ν e μ são mutuamente singulares quando ν e μ estão suportadas em conjuntos disjuntos.

Proposição 3.2.1. *Se μ e ν são probabilidades invariantes por f, com μ ergódica e $\nu \ll \mu$ então $\nu = \mu$.*

Se ν e μ são ergódicas então ou elas coincidem ou são mutuamente singulares.

Demonstração. Seja φ uma função mensurável limitada qualquer. Como μ é invariante e ergódica com respeito a f temos pelo teorema de Birkhoff que

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

é constante em μ quase todo x e $\tilde{\varphi}(x) = \int \varphi d\mu$. Como $\nu \ll \mu$ temos que a igualdade anterior vale em ν quase todo x. Em particular,

$$\int \tilde{\varphi} d\nu = \int \varphi d\nu = \int \varphi d\mu$$

(a primeira igualdade decorre do teorema de Birkhoff). Como as intergrais com respeito a μ e ν são iguais quando φ é uma função mensurável qualquer, temos que $\mu = \nu$.

Agora suponha que μ e ν são ergódicas para f e que ν não é absolutamente contínua com respeito a μ. Logo, existe um mensurável B com $\mu(B) = 0$ e $\nu(B) > 0$. O conjunto $A = \cup_{n\geq 0} f^{-n}(B)$ é invariante módulo zero (i.e. $A = f^{-1}(A)(mod.0)$). É facil ver que, $\mu(A) = 0$ e $\nu(A) > 0$, por σ-aditividade e porque $B \subset A$, portanto a ergodicidade de (f,ν) nos diz que $\nu(A) = 1$, donde μ e ν sao mutuamente singulares. $\square$

Veja que se μ_1 e μ_2 são medidas invariantes por f então para qualquer $t \in [0,1]$ temos que $(1-t)\mu_1 + t\mu_2$ é invariante por f. Isto significa que o conjunto $\mathcal{M}_1(f)$ é *convexo.* A proposição seguinte caracteriza as medidas ergódicas desse convexo; elas são os elementos extremais do mesmo.

Proposição 3.2.2. *Uma probabilidade μ é ergódica para f se, e somente se, não é possível escrever a mesma como combinação convexa $(1-t)\mu_1 + t\mu_2$ com $t \in (0,1)$ e μ_1 e μ_2 distintas.*

Demonstração. Se μ não é ergódica, então existe A mensurável f-invariante tal que $0 < \mu(A) < 1$. Defina μ_1 e μ_2 as restrições normalizadas de μ a A e a A^c, respectivamente,

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

Como A e A^c são invariantes por f, tanto μ_1 quanto μ_2 sao invariantes por f e se t=$\mu(A)$ então $1-t=\mu(A^c)$ e

$$\mu(E)=(1-t)\mu_1(E)+t\mu_2(E),$$

donde se μ não é ergódica então μ não é extremal. Para a recíproca, suponha que μ é ergódica e que $\mu=(1-t)\mu_1+t\mu_2$ com $t\in(0,1)$. É claro que $\mu(E)=0$ implica $\mu_1(E)=0=\mu_2(E)$, ou seja, $\mu_1\ll\mu$ e $\mu_2\ll\mu$. Pela Proposição 3.2.1 temos $\mu=\mu_1=\mu_2$ e assim μ é extremal.

O lema a seguir mostra que as probabilidades ergódicas estão suportadas em conjuntos disjuntos do espaço M.

□

Lema 3.2.1. *Seja $\mathcal{I}$ um conjunto enumerável(podendo ser finito), e seja $\{\mu_i; i\in\mathcal{I}\}$ uma família de probabilidades com índices em $\mathcal{I}$ todas distintas. Então existem subconjuntos mensuráveis $\{P_i; i\in\mathcal{I}\}$ dois a dois disjuntos, tais que*

$$f^{-1}(P_i)=P_i \text{ e } \mu_j(P_k)=\begin{cases}1 & se \quad i=k\\ 0 & se \quad i\neq k.\end{cases}$$

Demonstração. Fixe j,k naturais distintos. Pela Proposição 3.2.1, a probabilidade μ_j não pode ser absolutamente contínua com respeito a μ_k. Logo, existe um subconjunto mensurável $A_{j,k}$ com $\mu_j(A_j,k)>0$ e $\mu_k(A_{j,k})=0$. Denote por $B_{j,k}=\cup_{n\geq 0}f^{-n}(A_{j,k})$. Disto, $\mu_j(B_{j,k})>0$ pois $A_{j,k}\subset B_{j,k}$ e $\mu_k(B_{j,k})=0$ pois μ_k é f-invariante. Além disso, temos $f^{-1}(B_{j,k})\subset B_{j,k}$. Denote $C_{j,k}=\cap_{n\geq 0}f^{-n}(B_{j,k})$. De imediato, $C_{j,k}\subset f^{-1}(C_{j,k})$. Por outro lado, de $f^{-1}(B_{j,k})\subset B_{j,k}$ temos $f^{-(n+1)}(B_{j,k})\subset f^{-n}(B_{j,k})$ e portanto

$$f^{-1}(C_{j,k})=\bigcap_{n\geq 0}f^{-(n+1)}(B_{j,k})\subset\bigcap_{n\geq 0}f^{-n}(B_{j,k})=C_{j,k} \tag{3.17}$$

Isso mostra que $f^{-1}(C_{j,k})=C_{j,k}$. Desde que a sequência $\{f^{-n}(B_{j,k})\}_n$ é decrescente e μ_j é uma propabilidade temos que

$$\mu_j(C_{j,k})=\lim_n\mu_j(f^{-n}(B_{j,k}))=\mu_j(B_{j,k})>0.$$

Logo, por ergodicidade obtemos $\mu(C_{j,k})=1$. Além disso, $\mu_k(C_{j,k})=0$ pois $C_{j,k}\subset B_{j,k}$. Agora defina,

$$D_j=\bigcap_{k\neq j}C_{j,k} \text{ e } P_j=D_j\setminus\bigcup_{k\neq j}D_k.$$

Como $f^{-1}(C_{j,k})=C_{j,k}$ e $\mu_j(C^c_{j,k})=0$ temos $f^{-1}(D_j)=D_j$ e $\mu_j(D^c_j)=\mu_j(\cup_{j\neq k}C^c_{j,k})\leq\sum_{j\neq k}\mu(C^c_{j,k})=0$, ou seja, $\mu_j(D_j)=1$, ainda $\mu_k(D_j)=0$ para $k\neq j$. Daí, $f^{-1}(P_j)=P_j$, $\mu_j(P_j)=1$ e $\mu_k(P_j)=0$ para $j\neq k$. Além disso, os P'_js sao dois a dois disjuntos. □

No caso de subconjuntos convexos de espaços vetoriais de dimensão finita todo elemento pode ser escrito como combinação convexa dos elementos extremais. Por exemplo, num triângulo ABC todo ponto pode ser escrito como combinação convexa do vértices, sabemos em particular, que o baricentro $G = (A + B + C)/3$.

Naturalmente surge a pergunta: Pode toda probabilidade invariante ser escrita como combinação convexa de probabilidades ergódicas?

O principal teorema desta seção dar uma resposta a essa pergunta, e exceto pelo número de parcelas a mesma é afirmativa.

Exemplo 3.2.1. *Seja $f : [0, 1] \to [0, 1]$ dada por $f(x) = x^2$. As medidas de Dirac no 0 e no 1, δ_0 e δ_1, são ergódicas e invariantes. Veja que todo ponto em $[0, 1)$ vai para o 0. Assim, qualquer probabilidade invariante de dar peso total a $\{0, 1\}$. Em particular, $\mu = \mu(\{0\})\delta_0 + \mu_1\{1\}\delta_1$. Ou seja, μ é uma combinação convexa (nesse caso finita) de medidas ergódicas.*

Exemplo 3.2.2. *Consedere $f : \mathbb{T}^2 \to \mathbb{T}^2$ dada por $f(x, y) = (x + y, y)$. Como o determinante de f é identicamente 1, a medida de Lebesgue m é preservada por f. Todo círculo horizontal $S^1_y = S^1 \times \{y\}$ é invariante por f. Além disso, a restrição $f : S^1_y :\to S^1_y$ é a rotação R_y. Denote por m_y a medida de Lebesgue em S^1_y. Observe que m_y tambem é invariante por f e é ergódica para y irracional. Por outro lado, o Teorema de Fubini nos diz que*

$$m(E) = \int m_y(E)dy$$

para todo mensurável E. Como os racionais tem medida de Lebesgue nula, a igualdade anterior não é afetada restringindo a integral ao conjunto dos números irracionais. Assim, a igualdade anterior nos dá m como combinação convexa (nesse caso não-enumerável) de medidas ergódicas.

Fixemos $(M, \mathcal{B}, \mu)$ um espaço de probabilidade e ξ uma partição mensurável de M em conjuntos mensuráveis. Denotaremos por $\pi : M \to \xi$ a projeção natural que associa a cada elemento $x \in M$ o elemento $\xi(x)$ da partição ξ que o contém. Essa projeção permite munir ξ de uma estrutura de espaço de probabilidade da seguinte forma. Primeiramente, diremos que um conjunto C de ξ é mensurável se, e somente, $\pi^{-1}(C) \in \mathcal{B}$. A coleção $\hat{\mathcal{B}}$ de tais C de fato é uma σ-álgebra como se pode ver facilmente. Em seguida definimos a *medida quociente* por

$$\hat{\mu}(C) = \mu(\pi^{-1}(C))$$

para cada $C \in \hat{\mathcal{B}}$.

Teorema 3.2.1 (Decomposição Ergódica)**.** *Considere M um espaço métrico completo e separável, $f : M \to M$ uma transformação mensurável e μ uma probabilidade invariante. Então, existe um conjunto mensurável $M_0 \subset M$ com $\mu(M_0) = 1$, uma partição ξ de M_0 em subconjuntos mensuráveis e uma família de probabilidades $\{\mu_P; P \in \xi\}$ em M, satisfazendo*

(a) $\mu_P(P) = 1$ para $\hat{\mu}$-quase todo $P \in \xi$;

(b) A aplicação $P \mapsto \mu_P(E)$ é mensurável para todo $E \in \mathcal{B}$;

(c) μ_P é invariante e ergódica para $\hat{\mu}$-quase todo $P \in \xi$;

(d) $\mu(E) = \int \mu_P(E)d\hat{\mu}(P)$ para todo $E \in \mathcal{B}$.

3.3 Semicontinuidade da Entropia

Vamos analisar a função $h_f : \mathcal{M}_1(f) \to \mathbb{R}^*$ definida por $h_f(\mu) = h_\mu(f)$. Nosso primeiro exemplo mostrará que essa função nem sempre é contínua. Nesta seção mostraremos, no entanto, que em um amplo contexto, a mesma é *semicontínua superiormente*: Dado $\epsilon > 0$ $h_f(\nu) < h_f(\mu) + \epsilon$, i.e., $h_\nu(f) < h_\mu(f) + \epsilon$ para toda ν suficientemente próxima de μ na topologia fraca*.

O exemplo seguinte vem nos mostrar que nem sempre a aplicação h_f é contínua.

Exemplo 3.3.1. *Seja $f : [0,1] \to [0,1]$ dada por $f(x) = 10x(mod.1)$. A entropia de f com respeito a medida de de Lebesgue m é $h_m(f) = \log 10$. Para cada $k \geq 1$ seja F_k o conjunto dos pontos fixos do iterado f^k. Observe que F_k é um conjunto invariante com $card(F_k) = 10^k$, e que estes pontos estão equidistribuídos no seguinte sentido: cada intervalo $A_{i,k} = [(i-1)/10^k, i/10^k]$, $1 \leq i \leq 10^k - 1$ contém exatamente um ponto de F_k. Considere a sequência de medidas*

$$\mu_k = \frac{1}{10^k} \sum_{x \in F_k} \delta_x.$$

A coleção dos conjuntos $A_{i,k}$ geram os boreliano de $[0,1]$ e são conjuntos de continuidade para a medida de Lebesgue, i.e., $m(\partial A_{i,k}) = 0$. Segue-se que para mostrar que $\mu_k \to m$ é suficiente mostrar que

$$\lim_{n \to \infty} \mu_n(A_{i,k}) = m(A_{i,k}).$$

Para isso, note que $\mu_n(A_{i,k}) = \dfrac{1}{10^n} \displaystyle\sum_{x \in F_n} \delta_x(A_{i,k})$, e como o número de pontos periódicos de período $n > k$ em $A_{i,k}$ é 10^{n-k} (de fato, escolher um número em $A_{i,k}$ de período n significa escolher $n-k$ dígitos de 0 a 9, uma vez que os k primeiros ja estão determinados) temos

$$\mu_n(A_{i,k}) = \frac{1}{10^n} 10^{n-k} = 10^{-k} = m(A_{i,k}).$$

Isto mostra que $\mu_n \to m$. Porém, $h_{\mu_n}(f) = 0$ e $h_m(f) = \log 10$.

O exemplo acima mostra que a função h_f não varia continuamente.

Considere qualquer partição finita ξ de M cujo bordo

$$\partial \xi = \bigcup_{C \in \xi} \partial C$$

satisfaz $\mu(\partial \xi) = 0$. A função $\nu \mapsto \nu(C)$ é contínua no ponto μ uma vez que C é conjunto de continuidade para μ, para todo $C \in \xi$. Portanto, a função

$$\nu \mapsto H_\nu(\xi) = \sum_{C \in \xi} -\nu(C) \log \nu(C)$$

também é contínua em μ. A hipótese sobre ξ também implica que $\mu(\xi^n) = 0$ para todo $n \geq 1$ uma vez que

$$\partial\xi^n \subset \bigcup_{j=0}^{n-1} f^{-j}(\partial\xi).$$

Segue-se que a função $\nu \mapsto H_\nu(\xi^n)$ é contínua para todo n.

Proposição 3.3.1. *Seja ξ uma partição finita tal que $\mu(\partial\xi) = 0$, A função $\nu \mapsto h_\nu(f, \xi)$ é semicontinua superiormente em μ.*

Demonstração. De fato, por definição

$$h_\nu(f, \xi) = \inf_n H_\nu(\xi^n)$$

e o infimo de uma família de funções contínuas é semicontínua superiormente. □

Corolário 3.3.0.1. *Suponha que existe uma partição ξ tal que $\mu(\partial\xi) = 0$ e $\cup_n \xi^n$ gera a σ-álgebra dos mensuráveis de M. A função $\mu \mapsto h_\mu(f)$ é semicontínua superiormente.*

Demonstração. Pelo Corolário 3.3.0.1, dado $\epsilon > 0$ existe uma vizinhança U de μ na topologia fraca* tal que $h_\nu(f, \xi) < h_\mu(f, \xi) + \epsilon$ para toda ν em U. Temos da definiçao que $h_\mu(f, \xi) \leq h_\mu(f)$. Por ser geradora 3.1.2.1 nos diz que $h_\nu(f, \xi) = h_\nu(f)$, qualquer que seja ν. Portanto, $h_\nu(f) \leq h_\mu(f) + \epsilon$. Isso termina a prova do corolário. □

Suponhamos agora que M é um espaço métrico compacto e μ uma probabilidade boreliana em M. Nesse caso temos versão mais especializada do corolário anterior

Corolário 3.3.0.2. *Suponha que existe $\epsilon > 0$ tal que toda partição finita ξ com $\operatorname{diam}\xi < \epsilon$ satisfaz $\lim_n \operatorname{diam}\xi^n = 0$. Então, a função $\nu \mapsto h_\nu(f)$ semicontínua superiormente. Consequentemente, essa função é limitada e o seu supremo é atingido por alguma medida μ.*

Demonstração. Sabemos que $\lim_n \operatorname{diam}\xi^n = 0$ implica que $\cup_n \xi^n$ gera a σ-álgebra dos mensuráveis de M. Dado qualquer medida μ, escolha $r_x \in (0, \epsilon)$ tal que $\mu(\partial B(x, r_x)) = 0$. Seja $\mathcal{U}$ uma cobertura finita de M por tais bolas; tome para ξ a partição associada a $\mathcal{U}$, i.e., a partição cujos elementos são os conjuntos maximais para $\mathcal{U}$ que, para cada $U \in \mathcal{U}$ estao contidos em U ou U^c. Segue do corolário anterior que a função entropia é semicontínua superiormente em μ. Como μ é arbitraria a prova está concluída. As demais afirmações decorrem do fato de $\mathcal{M}_1(M)$, das probabilidades invariantes, ser compacto e a função entropia ser semicontínua superiormente. □

Proposição 3.3.2. *Seja $f : M \to M$ uma transformação expansiva de um espaço métrico compacto e seja $\epsilon > 0$ uma constante de expansividade de f. Então,*

$$\lim_n \operatorname{diam}(\xi^n) = 0$$

para toda partição finita ξ com $\operatorname{diam}(\xi) < \epsilon$.

Demonstração. A sequência $\text{diam}(\xi^n)$ é não-crescente. Seja δ o ínfimo de da mesma e suponha que $\delta > 0$. Então, para todo $n \geq 1$ existem pontos x_n e y_n com $d(x_n, y_n) > \delta/2$ mas x_n e y_n estão no mesmo elemento de ξ^n e portanto satisfazem

$$d(f^j(x_n), f^j(y_n)) \leq \text{diam}(\xi) < \epsilon$$

para todo $0 \leq j < n$. Por compacidade, existe $(n_j)_j$ sequência de naturais com $n_j \to \infty$ tal que $x = \lim_j x_{n_j}$ e $y = \lim_j y_{n_j}$. Então, $x \neq y$ mas $d(f^j(x), f^j(y)) < \epsilon$, contradizendo o fato de ϵ ser constante de expansividade de f. Isto mostra que $\delta = 0$ e portanto, $\lim_n \text{diam}(\xi^n) = 0$ □

Para encerrar essa seção definiremos, a *entropia topológica* de f como sendo o número $h(f) = \sup\{h_\mu(f); \mu \in \mathcal{M}_1(f)\}$, o qual está bem definido sempre que f é continua e f está definida num compacto.

Capítulo 4

Desigualdade de Ruelle

Vamos neste capítulo abordar uma importante estimativa superior para entropia que ajuda, por exemplo, no estudo de medidas maximais para uma dada f de classe C^1 definida sobre uma variedade M compacta C^∞.

Teorema 4.0.1. *Sejam M e f nas condições acima. Para qualquer probabilidade invariante boreliana μ temos*

$$h_\mu(f) \leq \int \sum_{i:\lambda_i(x)>0} \lambda_i(x) m_i(x) d\mu(x)$$

Se o sistema (f, μ) é ergódico, então os expoentes de Lyapunov são constantes μ-q.t.p. e assima desigualdade de Ruelle se reduz a

$$h_\mu(f) \leq \sum_{i;\lambda_>0} \lambda_i m_i \tag{4.1}$$

4.1 Teorema de Oseledets

O Teorema Ergódico de Birkhoff diz que se φ é uma função mensurável e μ é uma probabilidade invariante por uma transformação mensurável $f : M \to M$ de tal modo que $\varphi \in L^1(\mu)$ então o limite

$$\tilde{\varphi}(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)), \tag{4.2}$$

existe para quase todo $x \in M$.

Exemplo 4.1.1. *Considere $f : \mathbb{R} \to \mathbb{R}$ de classe C^1, $\phi : \mathbb{R} \to \mathbb{R}$ uma função mensurável, μ uma medida f-invariante, sabemos pelo teorema ergódico de Birkhoff que*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi(f^i(x))$$

existe num conjunto de medida total. Tomando em particular, $\phi = \log|f'|$ *temos*

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1}\phi(f^i(x)) &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\log|f'(f^i(x))| \\ &= \lim_{n\to\infty}\frac{1}{n}\log(\prod_{i=0}^{n-1}|f'(f^i(x))|) \end{aligned} \tag{4.3}$$

e pela regra da cadeia, a última expressão vale

$$\lim \frac{1}{n}\log|(f^n)'(x)|.$$

Isto nos dá uma parte do Teorema de Oseledets em dimensão 1, uma vez que o limite anterior existe μ*-q.t.p.*

O exemplo acima nos permite estudar o comportamente assintótico de $(f^n)'$, e portanto de f^n; quando n é suficientemente grande temos $|(f^n)'(x)| \approx \exp(n\tilde{\varphi}(x))$.

Definição 4.1.1. *Uma sequência de funções* $\{\phi_n\}_n$*, com* $\phi_n : M \to \mathbb{R}$ *é subaditiva para uma transformação mensurável* $f : M \to M$ *quando*

$$\phi_{m+n} \leq \phi_m + \phi_n \circ f^m, \ \forall n, m \geq 1. \tag{4.4}$$

Exemplo 4.1.2. *Seja* $A : M \to GL(d)$ *uma função mensurável, com valores no conjunto* $GL(d)$ *das matrizes quadradas inversíveis com entradas reais e dimensao d. Defina* $\phi(x) = A(x)$ *e* $\phi^n(x) = A(f^{n-1}(x))...A(f(x))A(x)$ *para todo* $n \geq 1$ *e* $x \in M$*. Então a sequência* $\{\varphi_n\}$*, onde* $\varphi_n(x) = \log \parallel \phi^n(x) \parallel$ *é subaditiva. De fato,*

$$\phi^{m+n}(x) = \phi^n(f^m(x)) + \phi^n(x).$$

Logo,

$$\begin{aligned} \varphi_{m+n}(x) &= \log\|\phi^n(f^m(x))\phi_n(x)\| \\ &\leq \log\|\phi^n(f^m(x))\| + \log\|\phi_n(x)\| \\ &= \varphi_n(f^m(x)) + \varphi_m(x), \end{aligned} \tag{4.5}$$

para quaisquer $m, n \in \mathbb{N}$ *e* $x \in M$*.*

Dada uma função $\varphi : M \to \mathbb{R}$ denotamos $\varphi^+(x) = \max\{0, \varphi(x)\}$. O teorema a seguir é equivalente a Birkhoff e para uma demonstração o leitor pode consultar [O-V].

Teorema 4.1.1 (Kingman)**.** *Sejam* μ *uma probabilidade invariante para uma transformação* $f : M \to M$ *e* $\varphi_n : M \to \mathbb{R}$ *uma sequência subaditiva de funções mensuráveis, tal que* $\varphi_1^+ \in L^1(\mu)$*. Então, a sequência* $\{\varphi_n/n\}_n$ *converge em* μ*-q.t.p.* $x \in M$ *para uma função mensurável* $\varphi : M \to [-\infty, \infty)$*. Além disso,* $\varphi^+ \in L^1(\mu)$ *e*

$$\int \varphi d\mu = \lim_n \frac{1}{n}\int \varphi_n d\mu = \inf_n \frac{1}{n}\int \varphi_n d\mu \in [-\infty, \infty).$$

Exemplo 4.1.3 (Furstenberg-Kesten). *Seja M um espaço mensurável, $f : M \to M$ transformação mensurável e $T : M \to M_{m\times m}(\mathbb{R})$ mensurável tal que $\log^+ \|T(.)\| \in L^1(\mu)$. Escreva $T_x^n = T(f^{n-1}(x))...T(f(x))T(x)$. Então existe uma função mensurável $\lambda : M \to \mathbb{R}\cup\{-\infty\}$ tal que*

$$\lambda(x) = \lim_{n\to\infty} \frac{1}{n} \log \|T_x^n\|$$

existe em μ quase todo ponto. Além disso, $\lambda \in L^1(\mu)$ e

$$\int \lambda d\mu = \lim_{n\to\infty} \frac{1}{n} \int \log \|T_x^n\| d\mu = \inf_n \int \log \|T_x^n\| d\mu.$$

A demonstração desses fatos se gue-se diretamente do Teorema de Kingman, e do exemplo 4.1.2

Proposição 4.1.1. *Seja $\{T_n\}_{n\in\mathbb{N}}$ uma sequência de matrizes $m \times m$ com entradas em $\mathbb{R}$ tais que*

$$\limsup_{n\to\infty} \frac{1}{n} \log \|T_n\| \leq 0, \tag{4.6}$$

escrevemos

$$T^n = T_n...T_2T_1$$

e assumimos que os limites

$$\lim_{n\to\infty} \frac{1}{n} \log \|(T^n)^{\wedge_k}\| \tag{4.7}$$

existem para $k = 1, ..., m$. Então

(a) $\lim_{n\to\infty}(T^{n*}T^n)^{1/2n} = \Lambda$ *existe, onde T^* denota a transposta de T.*

(b) Sejam $\exp(\lambda^{(1)}) < ... < \exp(\lambda^{(s)})$ os autovalores de Λ ($\lambda^{(r)}$ são reais, e possivelmente $\lambda^{(1)} = -\infty$), e $U^{(1)}$,..., $U^{(s)}$ os autoespaços correspondentes. Escrevendo $V^{(0)} = \{0\}$ e $V^{(r)} = U^{(1)} + ... + U^{(r)}$, temos:

$$\lim_{n\to\infty} \frac{1}{n} \log \|T^n u\| = \lambda^{(r)}, \text{ quando } u \in V^{(r)} \backslash V^{(r-1)}$$

para $r = 1, 2, ..., s$.

Uma demonstração da Proposição acima pode ser encontrada em [R2].

Teorema 4.1.2 (Teorema Ergódico Multiplicativo de Oseledets). *Considere $f : M \to M$ e μ uma probabilidade f-invariante. Seja $T : M \to M_{m\times m}(\mathbb{R})$ um função mensurável tal que*

$$\log^+ \|T(\cdot)\| \in L^1(\mu) \tag{4.8}$$

Escrevemos $T_x^n = T(f^{n-1}(x))...T(f(x))T(x)$. Usamos T^* para significar a transposta de T.

Existe $\Gamma \subset M$ tal que $f(\Gamma) \subset \Gamma$, $\mu(\Gamma) = 1$ e as seguintes propriedades ocorrem se $x \in \Gamma$:

(a) $\lim(T_x^{n*}T_x^n)^{1/2n} = \Lambda_x$ *existe.*

(b) Sejam $\exp(\lambda_x^{(1)}) < ... < \exp(\lambda_x^{(s)})$ *os autovalores de* Λ_x *(onde* $s = s(x)$*, os* $\lambda_x^{(s)}$ *são reais, e possivelmente* $\lambda_x^{(1)} = -\infty$*), e* $U_x^{(1)}$*,...,*$U_x^{(s)}$ *os autoespaços correspondentes. Faça* $m_x^{(r)} = \dim U_x^{(r)}$*. As funções* $x \mapsto \lambda_x^{(r)}$*,* $x \mapsto m_x^{(r)}$ *são* f*-invariantes. Escrevendo* $V_x^{(0)} = \{0\}$ *e* $V_x^{(r)} = U_x^{(1)} + ... + U_x^{(r)}$*, temos:*

$$\lim_{n\to\infty} \frac{1}{n} \log \|T_x^n u\| = \lambda_x^{(r)}$$

sempre que $u \in V_x^{(r)} \backslash V_x^{(r-1)}$*, para* $r = 1, ..., s$.

Vamos mostrar que

$$\lim_n \frac{1}{n} \log^+ \|T(f^{n-1}(x))\| = 0,$$

para quase todo $x \in M$.

Lema 4.1.1. *Se* φ *é uma função integrável então* $\lim_n n^{-1}\varphi(f^{n-1}(x)) = 0$ *para* μ *quase todo* $x \in M$.

Demonstração. Pelo Teorema Ergódico de Birkhoff, existe um conjunto $B = B(\varphi)$, invariante por f, de medida um para qualquer medida invariante por f tal que

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

converge para todo $x \in B$. Temos que

$$\frac{1}{n}\varphi(f^{n-1}(x)) = S_n(x) - \frac{n-1}{n} S_{n-1}(x).$$

Segue-se que $\lim_{n\to\infty} \varphi(f^{n-1}(x)) = 0$. □

Do Lema 4.11, para $\varphi(\cdot) = \log \|T(\cdot)\|$, que existe Γ_1 tal que $f(\Gamma_1) \subset \Gamma_1$, $\mu(\Gamma_1) = 1$, e

$$\lim_{n\to\infty} \frac{1}{n} \log^+ \|T(f^{n-1}(x))\| = 0 \textit{ se } x \in \Gamma_1.$$

Pelo exemplo 4.1.3, existe ainda Γ_2 tal que $f(\Gamma_2) \subset \Gamma_2$, $\mu(\Gamma_2) = 1$, e, para $q = 1, 2, ..., m$

$$\lim_{n\to\infty} \frac{1}{n} \log \|(T_x^n)^{\wedge q}\|$$

existe, e é uma função f invariante de x.

Tome $\Gamma = \Gamma_1 \cap \Gamma_2$. O teorema segue-se da Proposição 4.1.1 aplicada à sequência $T_n = T(f^{n-1}(x))$ para $x \in \Gamma$.

Corolário 4.1.2.1. *Seja $x \in T_xM$, $u \in T_xM$; temos*

$$\lim_{n\to\infty} \frac{1}{n} \log \|T_x^n u\| = \chi(x,u) \tag{4.9}$$

existe, finito ou $-\infty$. Se $\lambda \in \mathbb{R}$, os espaços lineares

$$V_x^\lambda = \{u \in T_xM; \chi(x,u) \leq \lambda\}$$

é uma função mensurável de $x \in \Gamma$.

Isto é uma consequência imediata do Teorema anterior. Nós temos $\chi(x,u) = \lambda_x^{(r)}$ se $u \in V_x^{(r)} \backslash V_x^{(r-1)}$, e $V_x^\lambda = \cup\{V_x^{(r)}; \lambda_x^{(r)} \leq \lambda\}$.

Observação 4.1.1. *4.9 implica que*

$$\chi(f(x), T(x)u) = \chi(x,u).$$

De fato, observe que $T_{f(x)}^n = T(f^n(x))...T(f^2(x))T(f(x))$, donde $T_{f(x)}^n T(x)u = T_x^{n+1}u$. Em particular, temos $T(x)V_x^\lambda \subset V_{f(x)}^\lambda$, $T(x)V_x^{(r)} \subset V_{f(x)}^{(r)}$. Se $\lambda_x^{(1)} \neq -\infty$, temos $T(x)$ inversível e portanto, $T(x)V_x^\lambda = V_{f(x)}^\lambda$, $T(x)V_x^{(r)} = V_{f(x)}^{(r)}$.

4.1.1 Espectro

Fixado $(M.\mathcal{B}, \mu)$ um espaço de medida e $f : M \to M$ uma aplicação mensurável μ preservando μ. Seja $T : M \to M_m(\mathbb{R})$ uma aplicação mensurável tal que

$$\log^+ T(\cdot) \in L^1(M, \mu).$$

Como na seção anterior, escrevemos $T_x^n = T(f^{n-1}(x))...T(f(x))T(x)$. Pelas observações anteriores podemos definir Λ_x; $s = s(x)$; $\lambda_x^{(1)} < ... < \lambda_x^{(s)} = \chi(x)$; $U_x^{(1)}$,...,$U_x^{(s)}$; $\{0\} = V_x^{(0)} \subset V_x^{(1)} \subset ... \subset V_x^{(s)} = T_xM$, e as funções $u \mapsto \chi(x,u)$ e $\lambda \mapsto V_x^\lambda$.

Faça $m_x^{(r)} = \dim U_x^{(r)} = \dim V_x^{(r)} - \dim V_x^{(r-1)}$. Os números $\lambda_x^{(r)}$ são os *expoentes de Lyapunov* com multiplicadade $m_x^{(r)}$, eles constituem o espectro de (f, T). Diremos que $V_x^{(1)} \subset V_x^{(2)} \subset ... \subset V_x^{(s)}$ é a *filtração associada* de T_xM. Note que o espectro é f-invariante e portante se (f, μ) é ergódico o espectro constante em μ quase todo ponto.

Espectro de $(f, T^\wedge)$

Considere $T^{\wedge p} : M \to M_{\binom{m}{p}}(\mathbb{R})$ a p-ésima potência exterior de T. Temos:

$$\begin{aligned}
(T^{\wedge p})_x^n(u_1 \wedge ... \wedge u_p) &= T^{\wedge p}(f^{n-1}(x))...T^{\wedge p}(x)(u_1 \wedge ... \wedge u_p) \\
&= T^{\wedge p}(f^{n-1}(x))...T^{\wedge p}(f(x))(T(x)u_1 \wedge ... \wedge T(x)u_p) \\
&= T^{\wedge p}(f^{n-1}(x))...(T(f(x)T(x)u_1 \wedge ... \wedge T(f(x))T(x)u_p) \\
&= ... \\
&= T(f^{n-1}(x))...T(x)u_1 \wedge ... \wedge T(f^{n-1}(x))...T(x)u_p \\
&= (T_x^n)^{\wedge p}(u_1 \wedge ... \wedge u_p)
\end{aligned} \tag{4.10}$$

Segue-se que
$$(T^{\wedge p})^n_x = (T^n_x)^{\wedge p}.$$
Veja que se A_n é uma sequência de matrizes com $A_n \to A$ então $A_n^{\wedge p} \to A^{\wedge p}$. É fácil ver que $(T^n_x)^{\wedge p *}(T^n_x)^{\wedge p} = ((T^n_x)^* T^n_x)^{\wedge p}$. Destas considerações temos que:
$$\lim_{n\to\infty}((T^n_x)^{\wedge p*}(T^n_x)^{\wedge p})^{1/2n} = \Lambda_x^{\wedge p}.$$

Isto determina o espectro de $T^{\wedge p}$ e a filtração associada de $(T_xM)^{\wedge p}$. Escrevendo $T^{\wedge} = \oplus_{p=0}^m T^{\wedge p}$, obtemos em particular que
$$\lim_{n\to\infty}\frac{1}{n}\log\|(T^n_x)^{\wedge}\| = \sum_{r:\lambda_x^{(r)}>0} m_x^{(r)}\lambda_x^{(r)}. \tag{4.11}$$
e por convergência domimada
$$\lim_{n\to\infty}\frac{1}{n}\int\log\|(T^n_x)^{\wedge}\| = \int\sum_{r:\lambda_x^{(r)}>0} m_x^{(r)}\lambda_x^{(r)}d\mu(x). \tag{4.12}$$

Espectro de (f^{-1}, T^*)

Suponha agora f inversível, com inversa mensurável, vejamos quem é o espectro de (f^{-1}, T^*). Faça $\hat{\Lambda}_x = \lim_{n\to\infty}(\hat{T}^{n*}_x\hat{T}^n_x)^{1/2n}$, onde $\hat{T}^n_x = T^*(f^{-n+1}(x))...T^*(f^{-1}(x))T^*(x)$. Desde que o espectro de $\hat{\Lambda}_x$ é f-invariante ele também é o espectro de $\lim_{n\to\infty}(\check{T}^{n*}_x\check{T}^n_x)^{1/2n}$ para quase todo $x \in M$, onde
$$\check{T}^n_x = T^*(x)T^*(f(x))...T^*(f^{n-1}(x)) = \hat{T}^n_{f^{n-1}(x)}.$$

Por outro lado, o espectro de $\check{T}^{n*}_x\check{T}^n_x$ é o mesmo de $\check{T}^n_x\check{T}^{n*}_x = T^{n*}_x T^n_x$. Portanto o espectro de $\hat{\Lambda}_x$ é o mesmo de Λ_x.

Espectro de $(f, (T^*)^{-1})$

Suponha que T é inversível em μ-quase todo ponto e que
$$\log^+\|T^{-1}(\cdot)\| \in L^1(M, \mu).$$
Defina $\widetilde{\Lambda}_x = \lim_{n\to\infty}(\widetilde{T}^{n*}_x\widetilde{T}^n_x)^{1/2n}$, onde
$$\widetilde{T}^n_x = (T^*)^{-1}(f^{n-1}(x))...(T^*)^{-1}(f(x))(T^*)^{-1}(x).$$
Note que $\widetilde{\Lambda}_x^{-1} = \Lambda_x$, já que $(\widetilde{T}^{n*}_x\widetilde{T}^n_x)^{-1} = T^{n*}_x T^n_x$. Segue-se que o espectro de $(f, (T^*)^{-1})$ é obtido mudando o sinal do espectro de (f, T): $\widetilde{\lambda}_x^{(r)} = -\lambda_x^{(s-r+1)}$. A filtração de T_xM associada com $(f, (T^*)^{-1})$ é a ortogonal da filtração associada a (f, T): $\widetilde{V}_x^{(r)} = V_x^{(s-r)\perp}$.

O caso inversível

Agora suponhamos $f : M \to M$ inversível com inversa mensurável preservando μ.

Teorema 4.1.3. *Seja $T : M \to GL_m(\mathbb{R})$ uma função mensurável no conjunto das matrizes inversiveis $m \times m.$, tal que*

$$\log^+ \|T(\cdot)\|,\ \log^+ \|T^{-1}(\cdot)\| \in L^1(M, \mu).$$

Escreva:

$$\begin{aligned} T_x^n &= T(f^{n-1}(x))...T(f(x))T(x) \\ T_x^{-n} &= T^{-1}(f^{-n}(x))...T^{-1}(f^{-2}(x))T^{-1}(x). \end{aligned}$$

Então, existe $\Delta \subset M$ tal que $f(\Delta) = \Delta$, $\mu(\Delta) = 1$, e uma decomposição mensurável $x \mapsto W_x^{(1)} \oplus ... \oplus W_x^{(s)}$ de T_xM sobre Δ (com $s = s(x)$), tal que

$$\lim_{n\to\pm\infty} \frac{1}{k} \log \|T_x^k u\| = \lambda_x^{(r)},\ \text{se } u \in W_x^{(r)} \backslash \{0\}.$$

Novamente os números $\lambda_x^{(1)} < ... < \lambda_x^{(s)}$ com multiplicidades $m_x^{(1)}$,...,$m_x^{(s)}$ constituem o espectro de (f, T) em x. Seja $V_x^{(1)} \subset ... \subset V_x^{(s)}$ a filtração associada de T_xM. Dos comentários anteriores, sabemos que o espectro de $(f^{-1}, T^{-1} \circ f^{-1})$ em x consiste dos números $-\lambda_x^s < ... < -\lambda_x^{(1)}$ com multiplicidades $m_x^{(s)}, ..., m_x^{(1)}$. Faça:

$$V_x^{-s} \subset ... \subset V_x^{-1}$$

a filtração associada. Suponha que

$$V_x^{(r-1)} \cap V^{(-r)} = \{0\} \tag{4.13}$$

e

$$V_x^{(r-1)} + V^{(-r)} = T_xM \tag{4.14}$$

par $r = 2, 3, ..., s$. Então, pondo

$$W_x^{(r)} = V^{(r)} \cap V_x^{(-r)}$$

obtemos:

$$\begin{aligned} T_xM &= V_x^{(-1)} \cap (V_x^{(1)} + V_x^{(-2)}) \cap (V_x^{(2)} + V_x^{(-3)}) \cap ... \cap V_x^{(s)} \\ &= W_x^{(1)} \oplus W_x^{(2)} \oplus ... \oplus W_x^{(s)} \end{aligned}$$

e o teorema ocorre. Resta-nos mostrar que valem 4.13 e 4.14. Defina S como o conjunto dos pontos x tal que 4.13 não ocorre. Dado $\delta > 0$ e $2 \leq r \leq s$, seja S_n o subconjutnto de S tal que se $x \in S_n$,

$$\|T_x^n u\| \leq \|u\| \exp n(\lambda_x^{(r-1)} + \delta) \tag{4.15}$$

e

$$\|T_x^{-n}u\| \leq \|u\| \exp n(-\lambda_x^{(-r)} + \delta) \tag{4.16}$$

para todo $u \in V_x^{(r-1)} \cap V_x^{(-r)}$. De 4.16, se $x \in f^{-n}(S_n)$,

$$\|T_x^n u\| \geq \|u\| \exp n(\lambda_x^{(r)} - \delta) \tag{4.17}$$

para todo $u \in V_x^{(r-1)} \cap V_x^{(-r)}$. Para $x \in S_n \cap f^{-n}(S_n)$, 4.15 e 4.17 nos dá $\lambda_x^{(r)} - \lambda_x^{(r-1)} \leq 2\delta$. Como $\mu(S_n \cap f^{-n}(S_n)) \to \mu(S)$ nós temos que $\lambda_x^{(r)} - \lambda_x^{(r-1)} \leq 2\delta$ para quase todo $x \in S$ e, desde que δ é arbitrário, nós obtemos que $\mu(S) = 0$. Assim, 4.13 está provado. 4.14 segue-se porque

$$\dim V_x^{(r-1)} + \dim V_x^{(-r)} = m.$$

De agora em diante escreveremos $\lambda_r(x)$ e $m_r(x)$ para $\lambda_x^{(r)}$ e $m_x^{(r)}$ respectivamente.

4.2 Demonstração do Teorema 4.0.1

Sejam $0 \leq \chi_1(A)^2 \leq \chi_2(A) \leq ... \leq \chi_m(A)^2$ os autovalores de uma matriz A^*A de ordem m. Um fato de álgebra linear que será usado é o

Lema 4.2.1. *Suponha que X e Y são dois espaços vetoriais de dimensão m e $A : X \to Y$ uma aplicação linear. Seja b uma constante positiva. Existe uma constante $C = C(m, b)$, tal que para qualquer $r > 0$,*

$$\mathrm{Vol}(V_{br}(A(B(0,r)))) \leq Cr^m \prod_{i=1}^{m} \max\{\chi_i(A), 1\}$$

Além disso, se g é uma aplicação C^1 a continuidade da aplicação $x \mapsto Dg(x)^* Dg(x)$ garante a existência de um $\epsilon_1 > 0$ tal que se $d(x,y) < \epsilon_1$ então

$$|\chi_i(D(g(x))) - \chi(Dg(y))| \leq \frac{1}{2} \tag{4.18}$$

onde estamos usando a dependência contínua dos autovalores.

De 4.18 temos

$$\frac{2}{3} \leq \frac{\max\{1, \chi_i(Dg(x))\}}{\max\{1, \chi_i(Dg(y))\}} \leq \frac{3}{2},\ 1 \leq i \leq m. \tag{4.19}$$

De fato, sejam $\{a_i\}_{i=1}^m$ e $\{b_i\}_{i=1}^m$ sequências de números reais tais que

$$|a_i - b_i| \leq \frac{1}{2}\ \forall i = 1, ..., m \tag{4.20}$$

Veja inicialmente que $\max\{1,a_i\} \geq 1$ e $\max\{1,b_i\} \geq 1$, logo

$$\begin{aligned}
0 &\geq a_i - b_i + |a_i - b_i| - 1 \\
&\geq a_i - b_i + |1-a_i| - |1-b_i| - 1 \\
&= 1 + a_i + |1-a_i| - (1 + b_i + |1-b_i|) - 1 \\
&= 2\max\{1,a_i\} - 2\max\{1,b_i\} - 1 \\
&\geq 2\max\{1,a_i\} - 2\max\{1,b_i\} - \max\{1,b_i\} \\
&\geq 2\max\{1,a_i\} - 3\max\{1,b_i\},
\end{aligned} \tag{4.21}$$

isto mostra 4.19 completamente devido à simetria ao substituirmos $a_i = \chi_i(Dg(x))$ e $b_i = \chi_i(Dg(y))$.

Fixe $n \geq 1$. Da compacidade de M e por f^n ser C^1 que existe ϵ_2 tal que $f^n(B(x,\epsilon_2)) \subset B(f^n(x), \rho_0/2)$ para todo $x \in M$. Tome ϵ_0 como em 5.3, para $g = f^n$, onde ρ_0 é o raio de injetividade da aplicação exponencial. Considere $\epsilon > 0$ menor que os números ϵ_i, $i = 0,1,2$ e menor que $\rho_0/4$.

Para cada $k \in \mathbb{N}$, seja E_k um conjunto maximal ϵ/k separado de M. Definimos uma partição finita $\xi_k = \{\xi_k(x); x \in E_k\}$ de M tal que $\xi_k(x) \subset \overline{\mathrm{Int}(\xi_k(x))}$ e

$$\mathrm{Int}(\xi_k(x)) = \{y \in M; d(y,x) < d(y,x') \text{ } se \text{ } x \neq x' \in E_k\}$$

para todo $x \in E_k$. Quando $y \in \xi_k(x)$ então $d(y,x) < \epsilon/k$, pois se $d(y,x) > \epsilon/k$ então para $x' \neq x$ $d(y,x') \geq d(y,x) > \epsilon/k$, donde $\{y\} \cup E_k$ é ϵ/k separado, donde E_k não é maximal. Isto mostra que para todo $x \in E_k$, $\xi_k(x) \subset B(x;\epsilon/k)$ e portanto, $\mathrm{diam}(\xi_k) \leq 2\epsilon/k$. Pela Fórmula de Abramov e por 3.1.2.4 temos

$$nh_\mu(f) = h_\mu(f^n) = \lim_{k\to\infty} h_\mu(f^n, \xi_k). \tag{4.22}$$

Veja que

$$\begin{aligned}
h_\mu(f,\xi) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(f^{-n+1}(\xi) \vee \ldots \vee f^{-1}(\xi) \vee \xi) \\
&= \lim_{n\to\infty} \frac{1}{n} [H_\mu(f^{-n+1}(\xi) \vee \ldots \vee f^{-1}(\xi)|\xi) + H_\mu(\xi)] \\
&= \lim_{n\to\infty} \frac{1}{n} [H_\mu(f^{-n+1}(\xi) \vee \ldots \vee f^{-2}(\xi)|f^{-1}(\xi) \vee \xi) + H_\mu(f^{-1}(\xi)|\xi)] \\
&\leq \lim_{n\to\infty} \frac{1}{n} [H_\mu(f^{-n+1}(\xi) \vee \ldots \vee f^{-2}(\xi)|f^{-1}(\xi)) + H_\mu(f^{-1}(\xi)|\xi)] \\
&\leq \lim_{n\to\infty} \frac{1}{n} [\sum_{i=1}^{n-1} H_\mu(f^{-i}(\xi)|f^{-i+1}(\xi))] \\
&= \lim_{n\to\infty} \frac{n-1}{n} H_\mu(f^{-1}(\xi)|\xi),
\end{aligned} \tag{4.23}$$

ou seja,

$$h_\mu(f,\xi) \leq H_\mu(f^{-1}(\xi)|\xi). \tag{4.24}$$

Aplicando 4.24 à f^n e à ξ_k temos

$$\begin{aligned} h_\mu(f^n,\xi_k) &\leq H_\mu(f^{-n}(\xi_k)|\xi_k) \\ &= \sum_{x\in E_k} \mu(\xi_k)H_\mu(f^{-n}(\xi_k)|\xi_k(x)) \\ &\leq \sum_{x\in E_k} \mu(\xi_k(x))\log K_n(x), \end{aligned} \tag{4.25}$$

onde

$$\begin{aligned} K_n(x) &= \#\{x' \in E_k; f^{-n}(\xi_k(x')) \cap \xi_k(x) \neq \emptyset\} \\ &= \#\{x' \in E_k; \xi_k(x') \cap f^n(\xi_k(x)) \neq \emptyset\}. \end{aligned} \tag{4.26}$$

Disto, vemos ser importante estimarmos o número de elementos $x' \in E_k$ para os quais $\xi_k(x')$ intersectam um dado conjunto da forma $f^n(\xi_k(x))$. Portanto, estimaremos $K_n(x)$.

Seja $b = b(\rho_0) \geq 1$ como no Lema 5.1.2. Denotemos por $V_\delta(F) = \{y : d(y,F) < \delta\}$ a vizinhança de raio δ em torno de F.

Lema 4.2.2. *Denotando $B(0,\epsilon/k)$ a bola centrada na origem de T_xM e raio ϵ/k e $\delta = \epsilon/k$ temos:*

$$f^n(\xi_k(x)) \subset f^n(\exp_x(\overline{B(0,\delta)})) \subset \exp_{f^n(x)} V_{b\delta}(Df^n(x)B(0,\delta))$$

Demonstração. Denotando $B_M(x,\delta)$ a bola com centro x e raio δ. Tem-se $\exp_x(B(0,\delta)) = B_M(x,\delta)$. Segue-se disto a primeira inclusão, já que $\overline{\xi_k(x) \subset B_M(x,\delta)}$.

Para a outra inclusão, tome $v \in \exp_{f^n(x)}^{-1} f^n(\exp(\overline{B(0,\delta)}))$. Queremos mostrar que

$$|v - Df^n(x)w| \leq b\delta$$

para algum $w \in B(0,\delta)$. Existe $y = f^n(z)$ com $z \in \exp_x(\overline{B(0,\delta)})$ tal que $v = \exp_{f^n(x)}^{-1}(y)$. Usando 5.3 para $g = f^n$ e chamando $z = \exp_x(w)$, temos $d(x,z) = |w| \leq \epsilon/k$, donde $f^n(z) \in B(f^n(x),\rho/2)$. Além disso, (lembre-se que $y = f^n(z)$)

$$\begin{aligned} d(f^n(x), \exp_{f^n(x)} Df^n(x)\exp_x^{-1}(z)) &\leq d(f^n(x),y) + d(y, \exp_{f^n(x)} Df^n(x)\exp_x^{-1}(z)) \\ &\leq \rho/2 + d(x,z) \\ &\leq \rho_0/2 + |w| \\ &\leq \rho_0/2 + \epsilon/k \\ &< \rho_0 \end{aligned} \tag{4.27}$$

Segue-se que podemos usar 5.1.2. Note que

$$\begin{aligned} |v - Df^n(x)w| &= |\exp^{-1}_{f^n(x)}(y) - \exp^{-1}_{f(x)}(\exp_{f^n(x)} Df^n(x)\exp^{-1}_x(z))| \\ &\leq bd(y, \exp_{f^n(x)} Df^n(x)\exp^{-1}_x(z)) \\ &\leq bd(x,z) \\ &\leq b\delta \end{aligned} \tag{4.28}$$

como queríamos demonstrar. $\square$

Se

$$\xi_k(x') \bigcap f^n(\xi_k(x)) \neq \emptyset$$

para algum $x' \in E_k$ então

$$B\left(x', \frac{\epsilon}{2k}\right) \bigcap \exp_{f^n(x)} V_{\widetilde{\delta}}\left(Df^n(x)B\left(0, \frac{\epsilon}{k}\right)\right) \neq \emptyset,$$

onde $\widetilde{\delta} = ((2b+1)/2)\epsilon/k$. Com efeito, Se $y \in \xi_k(x') \bigcap f^n(\xi_k(x))$ então $d(y, B(x', \epsilon/(2k))) < \epsilon/(2k)$. Assim, aumentando a vizinhança de raio δ em $\epsilon/(2k)$, temos o desejado.

Chamemos $\widetilde{b} = (2b+1)/2 > b$. Temos que

$$B\left(\exp^{-1}_{f^n(x)}(x'), \frac{\widetilde{b}^{-1}\epsilon}{2k}\right) \subset \exp^{-1}_{f^n(x)} B\left(x', \frac{\epsilon}{2k}\right) \subset V_{2\widetilde{\delta}}\left(Df^n(x)B\left(0, \frac{\epsilon}{k}\right)\right).$$

De fato, se $v \in B\left(\exp^{-1}_{f^n(x)}(x'), \frac{\widetilde{b}^{-1}\epsilon}{2k}\right)$, então existe $z \in M$ tal que $v = \exp^{-1}_{f^n(x)}(z)$. Observe que $d(x', f^n(x)) \leq d(y, x') + d(y, f^n(x))$, para algum $y \in \xi_k(x') \cap f^n(\xi_k(x))$. Logo, $y = f^n(y')$ com $y' \in \xi_k(x)$, i.e., $d(f^n(y'), f^n(x)) \leq \rho_0/2$ e então $d(x', f^n(x)) \leq \epsilon/k + \rho_0/2 < \rho_0$. Além disso,

$$\begin{aligned} d(z, f^n(x)) &= |v| \\ &\leq |v - \exp^{-1}_{f^n(x)} x'| + |\exp^{-1}_{f^n(x)} x'| \\ &< \widetilde{b}^{-1}\epsilon/(2k) + d(x', f^n(x)) \\ &< \widetilde{b}^{-1}\epsilon/(2k) + \epsilon/k + \rho_0/2 \\ &< \rho_0. \end{aligned} \tag{4.29}$$

Portanto, podemos usar 5.1.2

$$\widetilde{b}^{-1}d(z,x') \leq b^{-1}d(z,x') \leq |\exp^{-1}_{f^n(x)}(z) - \exp^{-1}_{f^n(x)}(x')| \leq b^{-1}\frac{\epsilon}{2k} \leq \widetilde{b}\frac{\epsilon}{2k}.$$

ou seja, $z \in B(x', \frac{\epsilon}{2k})$ e portanto $v \in \exp^{-1}_{f^n(x)} B(x', \frac{\epsilon}{2k})$

Para a outra continência, tome $y \in B\left(x', \frac{\epsilon}{2k}\right) \cap \exp_{f^n(x)} V_{\widetilde{\delta}}(Df^n(x)B(0, \epsilon/k))$. Note que se $v \in \exp^{-1}_{f^n(x)}(B(x', \epsilon/(2k)))$ então $\exp_{f^n(x)}(v) \in B(x', \epsilon/(2k))$. Se

$$y \in B\left(x', \frac{\epsilon}{2k}\right) \bigcap \exp_{f^n(x)} V_{\widetilde{\delta}}(Df^n(x)B\left(0, \frac{\epsilon}{k}\right))$$

então $d(\exp_{f^n(x)}(v), y) \leq \epsilon/k$ e assim $|\exp^{-1}_{f^n(x)}(y) - v| \leq b\epsilon/k$. Como

$$d(y, \exp_{f^n(x)}(Df^n(x)B(0,\epsilon/k))) \leq \widetilde{b}\epsilon/k$$

temos que

$$d(v, V_{\widetilde{\delta}}(Df^n(x)B(0,\epsilon/k))) \leq 2\widetilde{b}\epsilon/k.$$

Aqui usamos que $\exp_p(V_\delta(A)) = V_\delta(\exp_p(A))$, onde $A \subset T_pM$.

Desde que $B(x', \epsilon/(2k))$, $x' \in E_k$ sao disjuntas (pois E_k é ϵ/k separado), temos que $B(\exp^{-1}_{f^n(x)}(x'), \widetilde{b}^{-1}\epsilon/k)$, $x' \in E_k$ também são disjuntas. Segue-se que

$$\begin{aligned} K_n(x) &\leq \#\{B(\exp^{-1}_{f^n(x)}(x'), \widetilde{b}^{-1}\epsilon/(2k)); x \in E_k \text{ e } \xi_k(x') \cap f^n(\xi_k(x)) \neq \emptyset\} \\ &\leq \operatorname{Vol}(V_{2\widetilde{\delta}}(Df^n(x)B(0,\epsilon/k)))/\min\{\operatorname{Vol}(B(\exp^{-1}_{f^n(x)}(x'), \epsilon/(2\widetilde{b}k))); x' \in E_k\}. \end{aligned}$$

Sabemos do Lema 4.2.1 que

$$\operatorname{Vol}(V_{2\widetilde{\delta}}(Df^n(x)B(0,\epsilon/k))) \leq C_1 \left(\frac{\epsilon}{k}\right)^m \prod_{i=1}^m \max\{1, \chi_i(Df^n(x))\}.$$

O volume de uma bola $B(\exp^{-1}_{f^n(x)}(x'), \epsilon/(2\widetilde{b}k))$ é da forma $C_2\left(\dfrac{\epsilon}{k}\right)^m$, onde C_2 depende somente de m e de $2\widetilde{b}$. Portanto, para $C = C(m, 2\widetilde{b}) = C_1/C_2$ temos,

$$K_n(x) \leq C \prod_{i=1}^m \max\{\chi_i(Df^n(x)), 1\}$$

Observe que $\log^+ \chi_i(Df^n(x)) = \max\{0, \log \chi_i(Df^n(x))\} = \log \max\{1, \chi_i(Df^n(x))\}$. Por 4.18, para $y \in \xi_k(x)$ temos que

$$\log^+ \chi_i(Df^n(x)) \leq \log 2 + \log^+ \chi_i(Df^n(y))$$

logo,

$$\log K_n(x) \leq \log C + m \log 2 + \sum_{i=1}^m \log^+ \chi_i(Df^n(y)).$$

Segue-se que

$$\begin{aligned} H_\mu(f^{-n}(\xi_k)|\xi_k) &\leq \sum_{x \in E_k} \int_{\xi_k(x)} \log K_n(x) d\mu(y) \\ &\leq \log C + m \log 2 + \int_M \sum_{i=1}^m \log^+ \chi_i(Df^n(y)) d\mu(y). \end{aligned} \tag{4.30}$$

Isto, fazendo $k \to \infty$, juntamente com 4.22 temos

$$nh_\mu(f) \leq \log C + m\log 2 + \int_M \sum_{i=i}^{m} \log^+ \chi_i(Df^n(x))d\mu(x)$$

Sabemos que $\|Df^n(x)^{\wedge p}\| = \prod_{i=m-p+1}^{m} \chi_i(Df^n(x)),\ 1 \leq p \leq m$, temos

$$\|Df^n(x)^\wedge\| \geq \prod_{i=1}^{m} \max\{1, \chi_i(Df^n(x))\}.$$

Daí, $nh_\mu(f) \leq \log C + m\log 2 + \int_M \log|Df^n(x)^\wedge|d\mu(x)$.
Da expressão anterior

$$\begin{aligned} h_\mu(f) &\leq \frac{\log C + m\log 2}{n} + \frac{1}{n}\int_M \log|Df^n(x)^\wedge|d\mu(x) \\ &\leq \lim_{n\to\infty} \frac{1}{n}\int_M \log|Df^n(x)^\wedge|d\mu(x) \\ &= \int \sum_{i;\lambda_i(x)>0} \lambda_i(x)m_i(x)d\mu(x). \end{aligned} \tag{4.31}$$

Isto termina a demonstração.

Corolário 4.2.0.1. *Seja $f : M \to M$ uma transformação C^1 e μ uma probabilidade boreliana f-invariante ergódica. Se $h_\mu(f) > 0$ então f tem pelo menos um expoente de Lyapunov positivo.*

Demonstração. De fato, a condição $h_\mu(f) > 0$ implica $\sum_{i;\lambda_i>0} \lambda_i m_i > 0$, assim para o qual existe i com $\lambda_i > 0$. □

Corolário 4.2.0.2. *Se $f : M \to M$ é um difeomorfismo C^1, μ ergódica e $h_\mu(f) > 0$ então f tem um expoente positivo e um negativo.*

Demonstração. Pelo Corolário anterior, desde que $h_\mu(f) = h_\mu(f)$ e é claro (f^{-1}, μ) tambem é ergódico, logo temos que f^{-1} tem um expoente positivo. Mas, o espectro de f^{-1} é simétrico ao de f. Logo, f tem um expoente negativo. □

Capítulo 5

Apêndice

Neste apêndice faremos uma breve menção de alguns conceitos de Geometria Riemanniana, a fim de explicitar as propriedades básicas da Aplicação exponencial que foi usada na prova da Desigualdade de Ruelle.

5.1 Geometria Riemanniana

5.1.1 Conexões Afins

Um campo de vetores X numa variedade diferenciável M é uma correspondência que associa a cada ponto $p \in M$ um vetor $X(p) \in T_pM$.

Definição 5.1.1. *Um campo vetorial ao longo de uma curva $c : I \to M$ é uma aplicação que a cada $t \in I$ associa um vetor tangente $V(t) \in T_{c(t)M}$, aplicaçao esta diferenciável em M no seguinte sentido: se f é uma função diferenciável em M, então a função $t \mapsto V(t)f$ é uma função diferenciável em I.*

O campo $dc(\frac{d}{dt})$ indicado por $\frac{dc}{dt}$ é chamado *campo velocidade*(ou tangente) de c.

A restrição de uma curva c a um intervalo fechado $[a, b] \subset I$ chama-se um *segmento*. Se M é Riemannian, definimos o comprimento de um segmento por

$$l_a^b = \int |\frac{dc}{dt}|^{1/2} dt.$$

Proposição 5.1.1. *Toda variedade diferenciável (de Hausdorff e com base enumerável) possui uma métrica Riemanniana.*

Indicaremos por $\mathcal{C}(M)$ o conjunto dos campos de vetores de classe C^∞ em M e por $\mathcal{D}(M)$ o anel das funções reais de classe C^∞ definidas em M.

Definição 5.1.2. *Uma conexão afim ∇ numa variedade diferenciável M é uma aplicação*

$$\nabla : \mathcal{C}(M) \times \mathcal{C}(M) \to \mathcal{C}(M)$$

definida por $\nabla(X, Y) = \nabla_X Y$ e que satisfaz:

1. $\nabla_{fX+gY} = f\nabla_X Z + g\nabla_Y Z$

2. $\nabla_X(Y+Z) = \nabla_X Y + \nabla_X Z$

3. $\nabla_X(fY) = f\nabla_X Y + X(f)Y$

onde $X, Y, Z \in \mathcal{C}(M)$ *e* $f, g \in \mathcal{D}(M)$.

Proposição 5.1.2. *Seja M uma variedade diferenciável com uma conexão afim ∇ . Então existe uma única lei associa a um campo vetorial V ao longo da curva diferenciável $c : I \to M$ um outro campo vetorial $\frac{DV}{dt}$ ao longo de c, denominado* ***derivada covariante*** *de V ao longo de c, tal que*

(a) $\dfrac{D(V+W)}{dt} = \dfrac{DV}{dt} + \dfrac{DW}{dt}$,

(b) $\dfrac{D(fV)}{dt} = \dfrac{df}{dt}V + f\dfrac{DV}{dt}$,
onde W é um campo de vetores ap longo de c e $f \in \mathcal{D}(\mathcal{M})$

(c) Se V é induzido por um campo de vetores $Y \in \mathcal{C}(M)$, i.e., $V(t) = Y(c(t))$, então de c e $\dfrac{DV}{dt} = \nabla_{dc/dt}Y$

A proposição anterior mostra que a escolha de uma conexão afim em M dá origem a uma "derivada" de campos de vetores ao longo de curvas.

5.1.2 Conexão Riemanniana

Definição 5.1.3. *Seja M uma variedade diferenciável com uma conexão afim ∇ e uma métrica Riemanniana $\langle\ ,\ \rangle$. A conexão é dita ser compatível com a métrica quando para toda curva diferenciável c e quaisquer pares de campos de vetores paralelos P e P' ao longo de c tivermos $\langle P, P' \rangle$ =constante*

Proposição 5.1.3. *Suponha que uma variedade Riemanniana M tem uma conexão ∇ compatível com a métrica. Sejam V e W campos de vetores ao longo da curva diferenciável $c : I \to M$. Então*

$$\frac{d}{dt}\langle V, W\rangle = \langle \frac{DV}{dt}, W\rangle + \langle V, \frac{Dw}{dt}\rangle,\ t \in I. \tag{5.1}$$

Definição 5.1.4. *Uma conexão afim ∇ em uma variedade difenrenciável M é dita simétrica quando*

$$\nabla_X Y - \nabla_Y X = [X, Y]$$

Teorema 5.1.1 (Levi-Civita). *Dada uma variedade Riemanniana M existe uma única conexão afim ∇ em M satisfazendo*

(a) ∇ é simétrica

(b) ∇ é compativel com a métrica Riemanniana.

5.1.3 Aplicação Exponencial

No que se segue M será uma variedade Riemanniana munida de sua conexão Riemanniana.

Definição 5.1.5. *Uma curva parametrizada $\gamma : I \to M$ é uma geodesica em $t_0 \in I$ se $\frac{D}{dt}\frac{d\gamma}{dt} = 0$. Se γ é uma geodésica para todo $t \in I$, dizemos que γ é uma geodésica. Se $[a,b] \subset I$ e $\gamma : I \to M$ é uma geodésica, a restrição de γ a $[a,b]$ é chamada (segmento de) geodésica ligando $\gamma(a)$ e $\gamma(b)$.*

Se γ é uma geodésica, então

$$\frac{d}{dt}\langle\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\rangle = 2\langle\frac{D}{dt}\frac{d\gamma}{dt}, \frac{d\gamma}{dt}\rangle = 0,$$

usando (3.1) da proposição 5.1.3. Isto mostra que o comprimento do vetor tangente $\frac{d\gamma}{dt}$ é contante. Suporemos de agora em diante, que $|\frac{d\gamma}{dt}| = c \neq 0$. i.e., excuiremos geodésicas que se reduzem a pontos. O comprimento de arco s de γ, a partir de uma origem fica, digamos $t = t_0$, e então dado por

$$s(t) = \int_{t_0}^{t} |\frac{d\gamma}{dt}| dt = c(t - t_0).$$

Quando $c = 1$ temos uma geodesica normalizada.

Proposição 5.1.4. *Dado $p \in M$, existem um aberto $V \subset M$, $p \in V$, números $\delta > 0$, $\epsilon_1 > 0$ e uma aplicação C^∞*

$$\gamma : (-\delta, \delta) \times \mathcal{U},\ \mathcal{U} = \{(q, v), q \in V, v \in T_qM, |v| < \epsilon_1\},$$

tais que a curva $t \mapsto \gamma(t, v, q)$, $t \in (-\delta, \delta)$, é a única geodésica de M que no instante $t = 0$ passa por q com velocidade v, para cada $q \in V$ e cada $v \in T_qM$ com $|v| < \epsilon_1$.

A Proposição 5.1.4 afirma que se $|v| < \epsilon$, a geodésica $\gamma(t, q, v)$ existe no intervalo $(-\delta, \delta)$ e é única. Em verdade, é possível aumentar a velocidade de uma geodésica diminuindo seu intervalo de definição, ou vice-versa. Isto decorre so seguinte lema

Lema 5.1.1 (Homogeneidade de uma geodésica)**.** *Se a geodésica $\gamma(t, q, v)$ está definida no intervalo $(-\delta, \delta)$, então a geodésica $\gamma(t, q, av)$, $a \in \mathbb{R}^+$, está definida no intervalo $(-\delta/a, \delta/a)$ e*

$$\gamma(at, q, v) = \gamma(t, q, av).$$

A Proposição 5.1.4 juntamente com o Lemma de Homogeneidade , permite tornar o intervalo de definição de uma geodésica uniformemente grande em uma vizinhança de p. Mais precisamente, temos o seguinte resultado

Proposição 5.1.5. *Dado $p \in M$, existem uma vizinhança V de p em M, um número $\epsilon > 0$ e uma aplicação C^∞, $\gamma : (-2, 2) \times \mathcal{U} \to M$, $\mathcal{U} = \{(q, w) \in TM; q \in V, w \in T_qM, |w| < \epsilon\}$, tal que $t \mapsto \gamma(t, q, w)$, $t \in (-2, 2)$, é a única geodésica de M que no instante $t = 0$ passa por q com velocidade w, para cada $q \in V$ e cada $w \in T_qM$, com $|w| < \epsilon$.*

Demonstração. A geodésica $\gamma(t,q,v)$ da Proposição 5.1.4 está definida para todo $|t|<\delta$ e para $|v|>\epsilon_1$. Pelo lema de homogeneidade, $\gamma(t,q,\delta v/2)$ está definida para $|t|<2$. Tomando $\epsilon<\delta\epsilon_1/2$, a geodésica $\gamma(t,q,w)$ está definida para $|t|<2$ e $|w|<\epsilon$. □

Seja $p\in M$ e $\mathcal{U}\in TM$ um aberto dado pela proposição 5.1.4. A aplicação $\exp:\mathcal{U}\to M$ dada por

$$\exp(q,v)=\gamma(1,q,v)=\gamma\left(|v|,q,\frac{v}{|v|}\right),\ (q,v)\in\mathcal{U},$$

é chamada a *aplicação exponencial* em $\mathcal{U}$.

É claro que exp é diferenciálvel. Na maior parte das aplicações utiliza-se a restrição de exp ao espaço tangente T_qM, isto é, definiremos

$$\exp_q : B(0,\epsilon)\subset T_qM\to M$$

por $\exp_q(v)=\exp(q,v)$.

Geometricamente, $\exp_q(v)$ é o ponto de M obtido percorrendo um comprimento igual a $|v|$, a partir de q sobre a geodésica que passa por q com velocidade $v/|v|$. Em particular, $d(\exp_q(v),q)=|v|$.

Proposição 5.1.6. *Dado $q\in M$, existe $\epsilon>0$ tal que $\exp_q : B(0,\epsilon)\subset T_qM\to M$ é um difeomorfismo C^∞ de $B(0,\epsilon)$ sobre um aberto de M.*

Demonstração. Calculemos $d(\exp_q)_0$:

$$\begin{aligned} d(\exp_q)_0(v) &= \frac{d}{dt}(\exp_q(tv))\Big|_{t=0} \\ &= \frac{d}{dt}(\gamma(1,q,tv))\Big|_{t=0} \\ &= \frac{d}{dt}(\gamma(t,q,v))\Big|_{t=0}=v. \end{aligned} \tag{5.2}$$

Logo, $d(\exp_q)_0$ é a identidade de T_qM, donde pelo Teorema da Função Inversa, $\exp_q$ é um difeomorfismo C^∞ numa vizinhança de 0. □

Exemplo 5.1.1. *Seja $M=\mathbb{R}^n$. Como a derivação covariante coincide com a usual, as geodésicas são retas parametrizadas proporcionalmente ao comprimento de arco. A exponencial é a identidade (com a identificação usual do espaço tangente em p com $\mathbb{R}^n$).*

Vamos enunciar algumas propriedades da aplicação exponencial que nos serão úteis na demonstração da desiguldade de Ruelle.

A compacidade de M garante a existência de um número universal $\rho_0>0$, chamado de raio de injetividade tal que

$$\exp_x : B(0,\rho_0)\subset T_xM\to B(x,\rho_0)\subset M$$

é um difeomorfismo C^∞ para todo $x\in M$. O lema seguinte é um fato conhecido de geometria riemanniana

Lema 5.1.2. *Para qualquer $0 < r \leq \rho_0$ existe um número $b = b(r) \geq 1$ tal que para qualquer $x \in M$, y e z estão em $B(x,r) \subset M$, então*

$$b^{-1}d(y,z) \leq |\exp_x^{-1}(y) - \exp_x^{-1}(z)| \leq bd(y,z).$$

Além disso, se $g : M \to M$ é uma função C^1 existe ϵ_0 tal que

$$d(x,y) \leq \epsilon_0 \Rightarrow d(g(y), \exp_{g(x)} \circ Dg(x) \circ \exp_x^{-1}(y)) \leq d(x,y). \tag{5.3}$$

5.2 Álgebra Exterior

Seja E um espaço vetorial de dimensão m e $\mathcal{B} = \{e_1, ..., e_m\}$ uma base de E munido de produto interno. O espaço $E^{\wedge p}$, $p \leq m$, é o conjunto das combinações lineares da forma $v_1 \wedge v_2 \wedge ... \wedge v_p$ com $v_i \in E, i = 1, ..., p$, e $\wedge$ satisfaz

1. Se α e β são reais, então

$$v_1 \wedge ... \wedge (\alpha v_i + \beta v_i') \wedge ... \wedge v_p = \alpha v_1 \wedge ... \wedge v_i \wedge ... \wedge v_p + \beta v_1 \wedge ... \wedge v_i' \wedge ... \wedge v_p$$

2. Para quaisquer pares de vetores v_i e v_j vale:

$$v_1 \wedge ... \wedge v_i ... \wedge v_j \wedge ... v_p = -v_1 \wedge ... \wedge v_j \wedge ... \wedge v_i \wedge ... v_p$$

O conjunto $\mathcal{B}_1 = \{e_{i_1} \wedge ... \wedge e_{i_p}; 1 \leq i_1 < ... < i_p \leq m\}$ é uma base de $E^{\wedge p}$. Assim, $\dim E^{\wedge} = \binom{m}{p}$. Uma transformação linear $T : E \to E$ induz de modo natural uma transformação k-linear alternada $T^{\wedge p} : E^{\wedge p} \to E^{\wedge p}$, dada por

$$T^{\wedge p}(v_1 \wedge ... \wedge v_p) = Tv_1 \wedge ... \wedge Tv_p$$

Observe que $T^{\wedge p} + S^{\wedge} = (T + P)^{\wedge P}$ e $\alpha T^{\wedge p} = (\alpha T)^{\wedge p}$ Vamos munir $E^{\wedge p}$ de um produto interno declarando que

$$\langle e_{i_1} \wedge ... \wedge e_{i_p}, e_{j_1} \wedge ... \wedge e_{j_p} \rangle = \begin{cases} 1, & \text{se} \quad (i_1, ..., i_p) = (j_1, ..., j_p), \\ 0, & \text{se} \quad (i_1, ..., i_p) \neq (j_1, ..., j_p), \end{cases}$$

Disto $\langle v_1 \wedge ... \wedge v_p, w_1 \wedge ... \wedge w_p \rangle = \det(\langle v_i, w_j \rangle)$. O conjunto $E^{\wedge p}$ é a p-ésima potência exterior de E e a transformação $T^{\wedge p}$ a p-ésima potência exterior de T.

Proposição 5.2.1. *Seja $T : E \to E$ uma tranformação linear. Então*

$$\|T^{\wedge p}\| = \prod_{i=m-p+1}^{m} \chi_i(T), \tag{5.4}$$

onde $0 \leq \chi_1^2(T) \leq ... \leq \chi_i^2(T) \leq ... \chi_m^2(T)$ são os autovalores de T^*T.

Segue-se disso que $\|T^{\wedge p}\| \leq \|T\|^p$. Portanto, se temos uma sequência de matrizes $T_n \to T$ então $T_n^{\wedge p} \to T^{\wedge p}$. Com efeito, dado $\epsilon > 0$ tome $n_0 \in \mathbb{N}$ tal que $n > n_0$ implique $\|T_n - T\| < (\epsilon)^{1/p}$, temos

$$\begin{aligned} \|T_n^{\wedge p} - T^{\wedge p}\| &= \|(T_n - T)^{\wedge p}\| \\ &\leq \|T_n - T\|^p \\ &\leq \epsilon. \end{aligned} \tag{5.5}$$

Bibliografia

[R] D. Ruelle. *An inequality for the entropy of differential maps*. Bol. Soc. Bras. Mat., 9 (1978), 83-87.

[R2] D. Ruelle. *Ergodic Theory of Dynamical Systems*. Publ. Math. IHES 50, 275-305 (1979)

[Man] R. Mañe. *Ergodic Theory and Differentiable Dynamics*. Springer-Verlag, 1987.

[QXZ] M. Qian, J. Xie, S. Zhu. *Smooth Ergodic Theory for Endomorphisms*. Springer Verlag, 2009.

[Wal] P. Walters. *An Introduction to Ergodic Theory (Graduate Texts in Mathematics 79)*. Springer, New York, 1982.

[YP] Y. Pesin. *Dimension theory in dynamical systems: Contemporary views and applications*. The University of Chicago, Chicago-London, 1997.

[1] K. Oliveira, M. Viana. *Fundamentos de Teoria Ergódica*. A ser publicado.