UNIVERSIDADE FEDERAL DE ALAGOAS

INSTITUTO DE MATEMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

DISSERTAÇÃO DE MESTRADO

TEOREMA CENTRAL DO LIMITE PARA SISTEMAS DINÂMICOS DE APLICAÇÕES EXPANSORAS

MATHEUS BARBOSA MARTINS

MACEIÓ - AL

MAIO DE 2018

MATHEUS BARBOSA MARTINS

TEOREMA CENTRAL DO LIMITE PARA SISTEMAS DINÂMICOS DE APLICAÇÕES EXPANSORAS

Dissertação de Mestrado, na área de concentração de Sistemas Dinâmicos submetida a banca examinadora, designada pelo Programa de Mestrado 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 de Oliveira

MACEIÓ - AL

MAIO DE 2018

Catalogação na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecária Responsável: Helena Cristina Pimentel do Vale – CRB4 - 661

M386t Martins, Matheus Barbosa.
Teorema central do limite para sistemas dinâmicos de aplicações expansoras / Matheus Barbosa Martins. – 2018.
48 f.

Orientador: Krerley Irraciel Martins de Oliveira.
Dissertação (mestrado em Matemática) – Universidade Federal de Alagoas. Instituto de Matemática. Maceió, 2018.

Bibliografia: f. 42.
Apêndices: f. 43-48.

1. Aplicações expansoras. 3. Teorema central do limite. 3. Medida invariante. 4. Exponencialmente misturadora. 5. Lebesgue, medida de. I. Título.

CDU: 517.518.11

Matheus Barbosa Martins

Teorema Central do Limite Para Sistemas Dinâmicos de Aplicações Expansoras

Dissertação de Mestrado na área de Sistemas Dinâmicos, submetida em 10/05/2018 à banca examinadora, designada pelo Programa de Mestrado em Matemática da Universidade Federal de Alagoas, como parte dos requisitos necessários à obtenção do grau de Mestre em Matemática.

Banca Examinadora

Krerley Oliveira

Prof. Dr. Krerley Oliveira - Orientador - UFAL

Prof. Dr. Giovane Ferreira - UFMA

Wagner Ranter

Prof. Dr. Wagner Ranter - UFAL

Agradecimentos

Agradeço à Deus, por me conceder as bençãos necessárias para superar cada desafio que se apresentou ao longo desses anos.

À minha família, por todo o apoio longo desse mestrado. Principalmente nos momentos de desânimo e cansaço. Agradeço à minha namorada pelo amor, carinho e apoio que foram importantes nessa reta final.

À todos os meus amigos do Intituto de Matemática, pela amizade que tornou melhor cada fase vivida.

Ao meu orientador, Prof. Dr. Krerley Irraciel Martins de Oliveira, pelo seu apoio, comprometimento e orientação para o desenvolvimento deste trabalho. Agradeço também aos professores Dr. Rafael Nóbrega de Oliveira Lucena e Dr. Wagner Ranter Gouveia da Silva pela ajuda nas correções e esclarecimento de dúvidas.

Por fim, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pelo suporte financeiro durante todo o mestrado através da bolsa PICME.

Resumo

Nesse trabalho iremos mostrar que se $f : M \to M$ é uma aplicação expansora $C^{1+\nu_0}$ em uma variedade conexa compacta M. Então:

I) f admite uma única medida invariante μ_0 que é absolutamente contínua em relação a medida de Lebesgue. Além disso, μ_0 é exata (então ergódica);

II) (f, μ_0) é exponencialmente misturadora e satisfaz o teorema central do limite, no espaço das funções ν-Hölder contínuas, para qualquer $\nu \in (0, \nu_0]$.

Palavras-chaves: Aplicações expansoras, medida invariante, medida de Lebesgue, exponencialmente misturadora, teorema central do limite.

Abstract

In this paper we will show that if $f : M \to M$ is a $C^{1+\nu_0}$ expanding map on a compact connected manifold M. Then:

I) f admits a unique invariant measure μ_0 which is absolutely continuous with relation to the Lebesgue measure. Moreover, μ_0 is exact (thus ergodic);

II) (f, μ_0) is exponentially mixing and satisfies the central limit theorem, in the space of ν-Hölder continuous functions, for any $\nu \in (0, \nu_0]$.

Keywords: Expanding maps, invariant measure, Lebesgue measure, exponentially mixing, central limit theorem.

Sumário

INTRODUÇÃO

Nosso objetivo é mostrar que, dada uma aplicação expansora $f : M \to M$ numa variedade compacta conexa M, podemos obter uma medida μ_0 que é f-invariante e absolutamente contínua em relação a medida de Lebesgue, e depois demonstrar que as funções Hölder contínuas satisfazem o teorema central do limite para o sistema (f, μ_0).

Para a primeira parte, definiremos o operador de transferência $\mathcal{L} : E \to E$ dado por

$$(\mathcal{L}\varphi)(y) = \sum_{f(x)=y} \frac{\varphi(x)}{|\det Df(x)|},$$

(mostraremos que essa soma tem um número finito de termos), onde $E = C^0(M, \mathbb{R})$. Uma propriedade importante do operador de transferência é que vale

$$\int (\mathcal{L}\varphi)\psi dm = \int \varphi(\psi \circ f) dm. \tag{1}$$

(Quando as integrais fazem sentido).

A relação (1) implica, em particular, que pontos fixos de $\mathcal{L}$ estão relacionados com medidas f-invariantes absolutamente contínuas, pois se $\varphi_0 \in E$ é um ponto fixo de $\mathcal{L}$ então $\mu_0 = \varphi_0 m$ é uma medida f-invariante absolutamente contínua.

Para provar que tal ponto fixo existe, usaremos a noção de métrica projetiva em um cone convexo de um espaço vetorial, introduzida por G. Birkhoff. Essa noção e algumas propriedades serão dadas na seção 2.2.

Então, na seção 2.3 construiremos um cone $C(a, \nu)$ no espaço das funções Hölder contínuas tal que $\mathcal{L} : C(a, \nu) \to C(a, \nu)$ será uma contração com respeito a métrica projetiva θ associada a $C(a, \nu)$.

Na seção 2.4 mostraremos que (C_+, θ_+) é um espaço métrico completo, onde C_+ é o cone das funções contínuas não-negativas. Através de algumas relações entre as métricas θ e θ_+, junto com a continuidade de $\mathcal{L}$ em relação a métrica $||\cdot||_0$, mostraremos a existência de um ponto fixo único para $\mathcal{L}$ em $C(a, \nu)$ e, assim, obteremos a medida μ_0 mencionada anteriormente.

Na seção 2.5, mostraremos que o sistema (f, μ_0) tem decaimento exponencial de correlações no espaço das funções Hölder contínuas.

Finalmente, na seção 2.6 usaremos estimativas obtidas na seção 2.5, junto com um teorema central do limite abstrato para sistemas dinâmicos, Teorema 2.6.2, para obter que as funções Hölder contínuas satisfazem o teorema central do limite para (f, μ_0).

1 CONCEITOS DE TEORIA ERGÓDICA

1.1 Medidas Invariantes e Ergodicidade

Definição 1.1.1: Sejam (M, Σ, μ) um espaço de medida e $f : M \to M$ uma transformação mensurável. Dizemos que que a medida μ é **invariante** por f se

$$\mu(E) = \mu(f^{-1}(E)) \qquad para \ \ todo \ \ E \in \Sigma. \tag{1.2}$$

Nesse caso também dizemos que f preserva μ, e que μ é f-invariante.

Note que a Definição 1.1.1 faz sentido, pois a pré-imagem de um conjunto mensurável continua sendo mensurável. De um ponto de vista mais heurístico, essa definição significa que a probabilidade de um ponto estar num dado conjunto mensurável é igual a probabilidade de que a sua imagem esteja nesse mesmo conjunto.

Agora, apresentaremos uma proposição que nos dá uma condição equivalente a Definição 1.1.1. Essa condição é por muitas vezes de verificação mais imediata que a própria definição.

Proposição 1.1.1: Sejam $f : M \to M$ uma função mensurável e μ uma medida em M. Então f preserva μ se, e somente se,

$$\int \varphi d\mu = \int \varphi \circ f d\mu, \tag{1.3}$$

para toda função μ-integrável $\varphi : M \to \mathbb{R}$.

Demonstração: Suponhamos que a medida μ é invariante. Vamos mostrar que a relação (1.3) é válida para classes de funções cada vez mais amplas. Inicialmente, mostraremos que vale para as funções características. Para isso basta observar que

$$\int \chi_B d\mu = \mu(B) \quad e \quad \mu(f^{-1}(B)) = \int (\chi_B \circ f) d\mu,$$

e, por hipótese, temos $\mu(B) = \mu(f^{-1}(B))$ para todo $B \in \Sigma$. Agora, usando a linearidade da integral, podemos estender a validade de (1.3) para funções simples. Em seguida, vamos usar um argumento de aproximação para concluir que (1.3) vale para toda função integrável. Dada qualquer função integrável $\varphi : M \to \mathbb{R}$, consideremos sua parte positiva $\varphi^+(x) = \max\{\varphi(x), 0\}$. Pela Proposição A.1, podemos considerar uma sequência crescente

$(\varphi_n)_n$ de funções simples convergindo para φ^+. Então, usando o Teorema da Convergência Monótona (Teorema A.2) duas vezes:

$$\int \varphi^+ d\mu = \lim_n \int \varphi_n d\mu = \lim_n \int (\varphi_n \circ f) d\mu = \int (\varphi^+ \circ f) d\mu.$$

De uma maneira análoga, podemos mostrar que essa relação também vale para a parte negativa $\varphi^-(x) = \max\{-\varphi(x), 0\}$. Logo, pela linearidade da integral e por $\varphi = \varphi^+ - \varphi^-$, temos que a relação vale para φ.

Para mostrar a recíproca, basta notar que a relação (1.2) na definição de medida invariante se torna um caso particular de (1.3), onde $\varphi = \chi_E$ para cada $E \in \Sigma$. $\square$

O próximo resultado afirma que, dada qualquer medida invariante finita, quase todo ponto de qualquer conjunto mensurável E regressa a E um número infinito de vezes:

Teorema 1.1.2 (Recorrência de Poincaré): Seja $f : M \to M$ uma transformação mensurável e seja μ uma medida invariante por f. Seja $E \subset M$ qualquer conjunto mensurável com $\mu(E) > 0$. Então, para μ-quase todo ponto $x \in E$ existem infinitos valores de n para os quais $f^n(x)$ também está em E.

Consideremos um conjunto mensurável $E \subset M$ com medida positiva e um ponto $x \in M$ qualquer. Queremos analisar o conjunto dos iterados de x que visitam E, isto é,

$$\{j \geq 0 : f^j(x) \in E\}.$$

Por exemplo, o teorema de recorrência de Poincaré afirma que, para quase todo $x \in E$, este conjunto é infinito. Gostaríamos de ter informação mais precisa, de natureza quantitativa.

Definição 1.1.2: Chamamos **tempo médio de visita** de x a E o valor de

$$\tau(E, x) = \lim_{n\to\infty} \frac{1}{n} \#\{0 \leq j < n : f^j(x) \in E\}. \tag{1.4}$$

Seria interessante saber, por exemplo, em que condições esse tempo médio de visita é positivo. Antes de abordar este problema, é necessário responder a uma questão ainda mais básica: o limite (1.3) existe?

Representando por φ a função característica do conjunto E, podemos reescrever

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

Isso sugere uma generalização natural da nossa pergunta inicial: o limite acima existe para funções φ muito gerais, por exemplo, para todas as funções integráveis? O próximo teorema responde a essa pergunta.

Teorema 1.1.3 (Ergódico de Birkhoff): Seja $f : M \to M$ uma transformação mensurável e μ uma probabilidade invariante por f. Dada qualquer função integrável $\varphi : M \to \mathbb{R}$, o limite

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

existe em μ-quase todo ponto $x \in M$. Além disso, a função $\tilde{\varphi}$ definida dessa forma é integrável e satisfaz

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

O limite $\tilde{\varphi}$ é chamado **média temporal**, ou **média orbital**, de φ.

Definição 1.1.3: Dizemos que um sistema (f, μ) é ergódico se, dado qualquer conjunto mensurável E, temos $\tau(E, x) = \mu(E)$ para μ-quase todo ponto $x \in M$.

Dizemos que uma função mensurável $\varphi : M \to \mathbb{R}$ é **invariante** se $\varphi = \varphi \circ f$ em μ-quase todo ponto. Ou seja, a menos de um conjunto de medida nula, a função é constante em toda a trajetória de f. Além disso, dizemos que um conjunto mensurável $B \subset M$ é **invariante** se a sua função característica χ_B é uma função invariante. Em outras palavras, B é invariante se ele difere da sua pré-imagem $f^{-1}(B)$ por um conjunto de medida nula:

$$\mu(B \triangle f^{-1}(B)) = 0.$$

A seguinte proposição coleta diversas maneiras diferentes de definir ergodicidade.

Proposição 1.1.4: Seja μ uma probabilidade invariante de uma transformação mensurável $f : M \to M$. As seguintes condições são equivalentes:

(a) Para todo conjunto mensurável $B \subset M$ a função $\tau(B, \cdot)$ é constante em μ-quase todo ponto.

(b) Para toda função integrável $\varphi : M \to \mathbb{R}$ a média temporal $\tilde{\varphi} : M \to \mathbb{R}$ é constante em μ-quase todo ponto.

(c) Toda função integrável invariante $\psi : M \to \mathbb{R}$ é constante em μ-quase todo ponto.

(d) Para todo subconjunto invariante A tem-se $\mu(A) = 0$ ou $\mu(A) = 1$.

Definição 1.1.4: Uma medida ν diz-se **absolutamente contínua** com relação a outra medida μ se $\mu(E) = 0$ implica $\nu(E) = 0$. Nesse caso escrevemos $\nu \ll \mu$. Esta relação é transitiva: se $\nu \ll \mu$ e $\mu \ll \lambda$ então $\nu \ll \lambda$.

A próxima proposição afirma que probabilidades ergódicas são minimais para esta relação de ordem.

Proposição 1.1.5: Se μ e ν são probabilidades invariantes tais que μ é ergódica e ν é absolutamente contínua em relação a μ, então $\mu = \nu$.

Todas as demonstrações omitidas nessa seção podem ser encontradas em [2] nos capítulos 1,3 e 4.

1.2 Correlações

Em Teoria da Probabilidade, chamamos **correlação** de duas variáveis aleatórias, X e Y, ao número

$$C(X,Y) = \mathbb{E}\left[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\right] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Dada uma probabilidade invariante μ de um sistema dinâmico $f : M \to M$ e dadas funções mensuráveis $\varphi, \psi : M \to \mathbb{R}$, queremos analisar a evolução da sequência de correlações

$$C_n(\varphi,\psi) = C(\varphi \circ f^n, \psi) = \int (\varphi \circ f^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu, \quad n \in \mathbb{N},$$

quando o tempo n vai para infinito. Podemos pensar em φ e ψ como grandezas que medimos no sistema, tais como a temperatura, o pH, a energia cinética, etc. Sendo f a dinâmica do sistema, então $C_n(\varphi,\psi)$ mede como o valor da grandeza φ na n-ésima iterada da dinâmica se correlaciona com o valor da grandeza ψ medida no estado inicial, ou seja, até que ponto o valor inicial de ψ "influencia no futuro"do valor de φ.

Definição 1.2.1: Dizemos que o sistema (f,μ) é **misturador** se

$$\lim_n C_n(\chi_A,\chi_B) = \lim_n \mu(f^{-n}(A) \cap B) - \mu(A)\mu(B) = 0,$$

para quaisquer conjuntos mensuráveis $A, B \subset M$.

Em muitos casos importantes, se restringirmos φ, ψ a subconjuntos adequados de funções, é possível estimar a velocidade de decaimento das sequências de correlações $C_N(\varphi,\psi)$, num sistema misturador. Como mostraremos a seguir, as correlações $(\varphi,\psi) \mapsto C_n(\varphi,\psi)$ são funções bilineares. Logo, é natural considerar subconjuntos que são subespaços vetoriais. De fato, temos que para todo $\alpha \in \mathbb{R}$ e toda função mensurável $\gamma : M \to \mathbb{R}$, velem:

$I)$ $C_n(\alpha\varphi,\psi) \;= \int(\alpha\varphi \circ f^n)\psi d\mu - \int \alpha\varphi d\mu \int \psi d\mu = \;\alpha\left(\int (\varphi \circ f^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu\right)$
$= \alpha C_n(\varphi,\psi).$

$$\begin{aligned} II)\ \ C_n(\varphi+\gamma,\psi) &= \int((\varphi+\gamma)\circ f^n)\psi d\mu - \int(\varphi+\gamma)d\mu\int\psi d\mu = \int(\varphi\circ f^n)\psi d\mu \\ +\int(\gamma\circ f^n)\psi d\mu &- \int\varphi d\mu\int\psi d\mu - \int\gamma d\mu\int\psi d\mu = C_n(\varphi,\psi)+C_n(\gamma,\psi). \end{aligned}$$

A linearidade em relação a segunda entrada pode ser demonstrada de maneira análoga.

Definição 1.2.2: Dizemos que (f,μ) tem **decaimento exponencial de correlações** num dado espaço vetorial $\mathcal{V}$ se existe $\lambda<1$ e para todo $\varphi,\psi\in\mathcal{V}$ existe $A(\varphi,\psi)>0$ tal que

$$|C_n(\varphi,\psi)|\leq A(\varphi,\psi)\lambda^n,$$

para todo $n\geq 1$.

Agora, vamos desenvolver uma noção de teorema central do limite para (f,μ), que descreve a oscilação das médias em tempo finito

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

em torno do seu valor esperado $\int\varphi d\mu$. Em Teoria da Probabilidade, temos o seguinte teorema:

Teorema 1.2.1 (Teorema Central do Limite para V.A.I.I.D.): Sejam $X_0,...,X_n,...$ variáveis aleatórias independentes e identicamente distribuídas tomando valores em $\mathbb{R}$, com média $\bar{X}=E(X_n)<\infty$ e variância $\sigma^2=E((X_n-\bar{X})^2)\in(0,+\infty)$. Então, dado qualquer intervalo aberto $A\subset\mathbb{R}$, a probabilidade de

$$\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(X_j-\bar{X})\in A \quad \textit{converge para} \quad \frac{1}{\sqrt{2\pi}\sigma}\int_A e^{\frac{-t^2}{2\sigma^2}}dt,$$

quando $n\to+\infty$.

Trazendo esse conceito para o contexto de sistemas dinâmicos, dizemos que $\varphi: M\to\mathbb{R}$ satisfaz o **teorema central do limite para** (f,μ) se existe $\sigma>0$ tal que, para todo intervalo $A\subset\mathbb{R}$,

$$\mu(\{x\in M:\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}(\varphi(f^j(x))-\int\varphi d\mu)\in A\})\to\frac{1}{\sqrt{2\pi}\sigma}\int_A e^{\frac{-t^2}{2\sigma^2}}dt,$$

quando $n\to+\infty$.

As noções básicas de teoria da probabilidade podem ser encontradas em [4].

2 TRANSFORMAÇÕES EXPANSORAS

Definição 2.1.1: Dadas uma variedade Riemanniana compacta M e $f : M \to M$ uma aplicação de classe C^1. Dizemos que f é expansora se existe $\sigma > 1$ tal que

$$\|Df(x) \cdot v\| \geq \sigma \|v\| \text{ para quaisquer } x \in M \text{ e } v \in T_x M \ .$$

No que se segue, chamaremos **medida de Lebesgue** em M à medida de volume m induzida por tal métrica Riemanniana. Observe que a escolha particular da métrica não tem muita importância, uma vez que as medidas de volume associadas a distintas métricas Riemannianas são todas equivalentes.

A definição de função expansora implica, em particular, que a derivada Df é um isomorfismo em todo ponto e, pelo Teorema da Função Inversa, f é um difeomorfismo local. A próxima proposição nos mostrará duas propriedades importantes de f.

Proposição 2.1.1: Seja $f : M \to M$ uma aplicação expansora em uma Variedade Riemanniana compacta e conexa M, então:

(i) Existe $k \geq 1$ tal que todo ponto $y \in M$ tem exatamente k pré-imagens sob f.

(ii) Existe $\rho_0 > 0$ tal que, para todo $x \in f^{-1}(y)$, existe uma aplicação $h : B(y, \rho_0) \to M$ de classe C^1 tal que $f \circ h = \text{id}$, $h(y) = x$ e

$$d(h(y_1), h(y_2)) \leq \sigma^{-1} d(y_1, y_2) \text{ para todo } y_1, y_2 \in B(y, \rho_0).$$

Demonstração:

(i) Pelo teorema da função inversa, para todo $x \in M$, existe uma vizinhança $V(x)$ de x e $r(x) > 0$ tal que $f|_{V(x)}$ é um difeomorfismo sobre a bola de centro $y = f(x)$ e raio $r(x)$.

Afirmação 1: Dado $y \in M$, temos que $\# f^{-1}(y)$ é finita.

Suponhamos, por absurdo, que $f^{-1}(y) = \{y_i : i \in \mathbb{N}\}$. Por compacidade, temos que $(y_n)_n$ se acumula, a menos de uma subsequência, em um ponto $x \in M$. Logo, a partir de um certo $n_0 \in \mathbb{N}$, todos os y_n pertencem a vizinhança $V(x)$. Isso é um absurdo, pois f é injetiva em $V(x)$.

Como f é um difeomorfismo local, a imagem $f(M)$ é um conjunto aberto em M. Por outro lado, como f é contínua e M é compacto, temos que $f(M)$ é compacto e, em particular, é fechado em M. Portanto, por conexidade, $f(M) = M$. Assim, podemos tomar a cobertura de M formada por $\{B(f(x), r(x)) : x \in M\}$ e, por compacidade, essa cobertura admite uma subcobertura finita $\{B(f(x_k), r(x_k) : k = 1, ..., m\}$. Ainda por compacidade, essa cobertura finita tem um número de Lebesgue r, isto é, dado qualquer subconjunto $A \in M$ com diâmetro menor ou igual a r, temos $A \subset B(f(x_k), r(x_k))$ para algum $k \in \{1, ..., m\}$.

Afirmação 2: Se x é uma pré-imagem qualquer de $y \in M$, então existe uma vizinhança $W(x)$ de x tal que $f|_{W(x)}$ é um difeomorfismo sobre $B(y, \frac{r}{2})$.

Como diam$(B(y, \frac{r}{2})) = r$, existe $k \in \{1, ..., m\}$ tal que $B(y, \frac{r}{2}) \subset B(f(x_k), r(x_k))$. Sendo f um difeomorfismo entre $V(x_k)$ e $B(f(x_k), r(x_k))$, existe uma aplicação $g : B(f(x_k), r(x_k)) \to M$ de classe C^1 tal que $f \circ g = \text{id}$ e $g(y) = x$. Então, basta tomarmos $W(x) = g(B(y, \frac{r}{2}))$.

Agora, consideremos os conjuntos $D_n = \{x \in M : \#f^{-1}(x) = n\}$. Queremos mostrar que D_n é aberto e fechado para todo $n \geq 1$. Seja $x \in D_n$ e $f^{-1}(x) = \{x_1, ..., x_n\}$, considere $\rho = \min\{r(x_1), ..., r(x_n), \frac{r}{2}\}$. Logo, para cada x_i existe uma vizinhança $W(x_i)$ tal que $f|_{W(x_i)}$ é um difeomorfismo sobre $B(x, \rho)$. Portanto, qualquer $z \in B(x, \rho)$ tem pelo menos n pré-imagens (uma em cada $W(x_i)$). Suponha que existe $z \in B(x, \rho)$ que tem alguma outra pré-imagem, digamos z_{n+1}, então pela Afirmação 2 existe uma vizinhança $W(z_{n+1})$ de z_{n+1} tal que $f|_{W(z_{n+1})}$ é um difeomorfismo sobre $B(x, \frac{r}{2})$, como $\rho \leq \frac{r}{2}$, segue que x tem uma pré-imagem em $W(z_{n+1})$, contradizendo a suposição de que x tem exatamente n pré-imagens. Com isso, concluímos que $\#f^{-1}(y) = n$ para todo $y \in B(x, \rho)$. Segue diretamente que D_n e D_n^C são conjuntos abertos para todo $n \geq 1$ e, por conexidade, temos $D_n = \emptyset$ ou $D_n = M$ para cada $n \geq 1$. Finalmente, como a quantidade de pré-imagens de cada ponto é limitada, existe $k > 1$ tal que $D_k = M$.

(ii) Vamos considerar $\rho_0 = \frac{r}{2}$, pela Afirmação 2 temos que dado qualquer $x \in \{f^{-1}(y)\}$, existe uma vizinhança $W(x)$ de x tal que $f|_{W(x)}$ é um difeomorfismo sobre $B(y, \rho_0)$. Logo, basta tomarmos $h = (f|W(x))^{-1}$. Temos que $\|Dh(z)\| = \|Df(h(z))^{-1}\| \leq \sigma^{-1}$ para todo z no domínio de h. Pela Desigualdade do Valor Médio, obtemos

$$d(h(y_1), h(y_2)) \leq \sigma^{-1} d(y_1, y_2) \text{ para todo } y_1, y_2 \in B(y, \rho_0). \qquad \square$$

Transformações h como neste enunciado são chamadas **ramos inversos** de f. O número $k = \#f^{-1}(y)$ de pré-imagens de um ponto qualquer de M é chamado **grau** da aplicação.

Ao longo do texto, consideraremos E como sendo o espaço normado $(C^0(M,\mathbb{R}), ||\cdot||_0)$, onde $||\cdot||_0$ é a norma do supremo dada por $\|\varphi\|_0 = \sup_{x\in M}|\varphi(x)|$.

Observação 2.1.1: Podemos obter $R_0 > 0$ tal que qualquer bola aberta de raio R_0 esteja no domínio de cada ramo inverso h_i de f, $i = 1,...,k$. Vamos construir R_0 da seguinte maneira:
Dado $x \in M$, temos $f^{-1}(x) = \{x_1,...,x_k\}$. Para cada x_j existem uma vizinhança $V(x_j)$ e $r_j > 0$ tais que $f|_{V(x_j)}$ é um difeomorfismo sobre $B(x,r_j)$, ou seja, $B(x,r_j)$ está no domínio de um ramo inverso h_k. Seja $R(x) = \min\{r_1,...,r_k\}$, temos $B(x,R(x))$ contido nos domínios de todos os k ramos inversos. Basta tomar R_0 como sendo o número de Lebesgue da cobertura aberta $\{B(x,R(x)) : x \in M\}$.

Definição 2.1.2: Definimos o operador $\mathcal{L} : E \to E$ por

$$(\mathcal{L}\varphi)(y) = \sum_{f(x)=y} \varphi(x)|\det Df(x)|^{-1} = \sum_{i=1}^{k} \varphi(y_i)|\det Df(y_i)|^{-1},$$

onde $f^{-1}(y) = \{y_1,...,y_k\}$. O operador $\mathcal{L}$ é conhecido como **Operador de Transferência** ou **Operador de Perron-Frobenius**.

De fato, temos que $\mathcal{L}g \in E$ para todo $g \in E$. Pois, dado $y \in M$ e $h_i : B(y,R_0) \to M$ os ramos inversos de y por f, então

$$\mathcal{L}g = \sum_{i=1}^{k}(g \circ h_i)|\det Df|^{-1}$$

em $B(y,R_0)$. Logo, $\mathcal{L}g$ é contínua em y.

Note que $\mathcal{L}$ é linear. Logo, para que $\mathcal{L}$ seja contínua, basta ser limitada. De fato, temos

$$\|\mathcal{L}\varphi\|_0 \le k \sup_{x\in M} |\det Df(x)|^{-1}\|\varphi\|_0,$$

onde $k = \#f^{-1}(y)$ para todo $y \in M$. Por f ser de classe C^1, temos que $|\det Df(x)|^{-1}$ atinge valor de máximo no compacto M. Usamos a norma $\|\varphi\|_0 = \sup_{x\in M}|\varphi(x)|$ em E.

A próxima proposição mostrará uma propriedade importante do Operdor de Transferência. Ele está diretamente relacionado com o **Operador de Koopman** $U : E \to E$ dado por $U(\varphi) = \varphi \circ f$.

Proposição 2.1.2: Dadas φ,$\psi \in E$, temos

$$\int (\mathcal{L}\varphi)\psi dm = \int \varphi(\psi \circ f)dm.$$

Demonstração:

Sabemos que, se $f : U \to V$ é um difeomorfismo entre os abertos $U, V \in M$, então vale o Teorema da Mudança de Variáveis em qualquer compacto $K \subset U$.

Seja $R_0 > 0$ o número obtido na Observação 2.1.1. Condideremos a cobertura aberta $\{B(x, \frac{R_0}{3}) : x \in M\}$ e, por compacidade, tomemos uma subcobertura finita $\{B(x_j, \frac{R_0}{3}) : j = 1, ..., m\}$. A vantagem dessa nova cobertura é que ela foi construída de modo que o fecho de cada $B_j = B(x_j, \frac{R_0}{3})$ está contido no domínio de cada ramo inverso h_i de f , $i = 1, ..., k$.

Agora, considere o conjunto β cujos elementos são todas as interseções finitas entre $B_1, B_1^C, B_2, B_2^C, ..., B_m, B_m^C$. Note que β é uma cobertura aberta (a menos de um conjunto de medida nula contido na união dos bordos dos elementos de β), finita e disjunta dois-a-dois. Sendo β uma cobertura finita, podemos enumerar seu elementos, digamos $\beta = \{\beta_1, ..., \beta_n\}$. Agora, considere um novo conjunto $\beta' = \{\bar{\beta_j} : j = 1, ..., n\}$. Note que β' é uma cobertura finita de M formada por conjuntos compactos cujas intersecções tem medida nula, pois eles se intersectam apenas nos bordos. Como o fecho de cada β_j está contido no domínio de cada ramo inverso h_i de f, vale o Teorema da Mudança de Variáveis para $h_{j,i}$ em cada $\bar{\beta_j}$. Logo,

$$\begin{aligned}
\int_{\bar{\beta_j}} (\mathcal{L}\varphi)\psi dm &= \int_{\bar{\beta_j}} \sum_{i=1}^{k} \varphi \circ h_{j,i}\psi |\det Df(h_{j,i})|^{-1} dm \\
&= \sum_{i=1}^{k} \int_{h_{j,i}(\bar{\beta_j})} \varphi(\psi \circ f) |\det Df|^{-1} |\det Df| dm \\
&= \int_{f^{-1}(\bar{\beta_j})} \varphi(\psi \circ f) dm,
\end{aligned}$$

pois $f^{-1}(\bar{\beta_j}) = \bigcup_{i=1}^{k} h_{j,i}(\bar{\beta_j})$ (união disjunta dois-a-dois). Segue que

$$\begin{aligned}
\int_M (\mathcal{L}\varphi)\psi dm &= \sum_{j=1}^{n} \int_{\bar{\beta_j}} (\mathcal{L}\varphi)\psi dm \\
&= \sum_{j=1}^{n} \int_{f^{-1}(\bar{\beta_j})} \varphi(\psi \circ f) dm \\
&= \int_M \varphi(\psi \circ f) dm.
\end{aligned}$$

Onde na última igualdade usamos os fatos de que:

I) $\bigcup_{j=1}^{n} f^{-1}(\bar{\beta}_j) = f^{-1}(\bigcup_{j=1}^{n} \bar{\beta}_j) = f^{-1}(M) = M$;

II) As pré-imagens dos $\bar{\beta}_j{}'$s por f são disjuntas duas-a-duas a menos de um conjunto de medida nula. De fato, para quaisquer $j, m \in \{1, ..., n\}$, temos

$$\begin{aligned} f^{-1}(\bar{\beta}_j) \cap f^{-1}(\bar{\beta_m}) &= \left(\bigcup_{i=1}^{k} h_{j,i}(\bar{\beta}_j)\right) \bigcap \left(\bigcup_{p=1}^{k} h_{m,p}(\bar{\beta_m})\right) \\ &= \bigcup_{i,p=1}^{k} \left(h_{j,i}(\bar{\beta}_j) \cap (h_{m,p}(\bar{\beta_m})\right), \end{aligned}$$

e como os β_j's se intersectam apenas nos bordos, as intersecções acima deixam de ser vazias apenas em conjuntos contidos nas pré-imagens desses bordos pelos ramos inversos. Uma vez que esses ramos inversos são difeomorfismos em seus domínios, eles preservam a dimensão de tais bordos, as quais são menores que a dimensão da variedade e, portanto, tem medida de Lebesgue zero. $\square$

2.2 Cones e Métricas Projetivas.

Apresentaremos nessa seção alguns resultados sobre a teoria de cones e métricas projetivas associada a operadores lineares. Entre os resultados está o teorema de Birkhoff, que garante a contração, na métrica projetiva, de operadores lineares restritos a cones invariantes.

Seja E um espaço vetorial. Chamaremos **cone** a qualquer subconjunto $C \subset E \backslash \{0\}$ satisfazendo:

$$v \in C \text{ e } t > 0 \Rightarrow tv \in C.$$

O cone é dito ser **convexo** se

$$v_1, v_2 \subset C \quad \text{e} \quad t_1, t_2 > 0 \quad \Rightarrow \quad t_1 v_1 + t_2 v_2 \in C.$$

Nós definimos o **fecho** $\overline{C}$ de C por

$$w \in \overline{C} \quad \Leftrightarrow \quad \text{existem } v \in C \text{ e } (t_n)_n \searrow 0 \text{ tais que } (w + t_n v) \in C \text{ para todo } n \geq 1.$$

Estamos interessados em eliminar cones degenerados, como por exemplo, semi-planos. Para isso, exigiremos que

$$\overline{C} \cap (-\overline{C}) = \{0\}. \qquad (2.2.1)$$

Denominaremos os cones convexos com a propriedade acima de **cones projetivos**. Passaremos agora a definir a **métrica projetiva** associada ao cone projetivo C. Dados v_1, v_2 em C, defina

$$\alpha(v_1, v_2) = \sup\{t > 0 : v_2 - tv_1 \in C\} \quad \text{e} \quad \beta(v_1, v_2) = \inf\{s > 0 : sv_1 - v_2 \in C\},$$

com a convenção de que $\sup \emptyset = 0 \;\; e \;\; \inf \emptyset = +\infty$.

A partir das definições acima, podemos demonstrar as seguintes propriedades:

1) $\alpha(v_1, v_2) \leq \beta(v_1, v_2)$ para todo $v_1, v_2 \in C$:

$v_2 - tv_1 \in C$, $\;sv_1 - v_2 \in C \Rightarrow (s-t)v_1 \in C$ (por convexidade) $\Rightarrow (s-t) \geqslant 0$ (caso contrário $-v_1 \in C$, contradizendo (2.2.1))

2) $\alpha(v_1, v_2) < +\infty$:

$\alpha(v_1, v_2) = +\infty \Rightarrow$ existe uma sequência $(t_n)_n \to +\infty$ tal que $v_2 - t_n v_1 \in C$ para todo $n \geqslant 1 \Rightarrow s_n = 1/t_n$ seria uma sequência de números positivos convergindo para zero, tais que $s_n v_2 - v_1 \in C$ para todo $n \geqslant 1 \Rightarrow -v_1 \in \overline{C}$, contradizendo (2.2.1).

3) $\beta(v_1, v_2) > 0$:

Podemos mostrar de maneira análoga a anterior que $\beta(v_1, v_2) = 0 \Rightarrow -v_2 \in \overline{C} \quad \square$

Dado qualquer cone $C \subset E$ e quaisquer $v_1, v_2 \in C \backslash \{0\}$, definimos

$$\theta(v_1, v_2) = \log \frac{\beta(v_1, v_2)}{\alpha(v_1, v_2)} \quad (\text{com } \theta = +\infty \text{ se } \alpha = 0 \text{ ou } \beta = +\infty\ ).$$

As propriedades 1, 2 e 3 nos garantem que $\theta(v_1, v_2)$ está bem definido e toma valores em $[0, +\infty]$.

A proposição seguinte afirma que θ define uma métrica no quociente projetivo de $C \backslash \{0\}$, ou seja, no conjunto das classes de equivalência da relação $\sim$ definida por $v_1 \sim v_2 \Leftrightarrow v_1 = tv_2$ para algum $t > 0$.

Proposição 2.2.1: $\theta(\cdot, \cdot)$ é uma métrica no quociente projetivo de C, isto é,

a) $\theta(v_1, v_2) = \theta(v_2, v_1)$ para quaisquer $v_1, v_2 \in C$,

b) $\theta(v_1, v_2) + \theta(v_2, v_3) \geq \theta(v_1, v_3)$ para quaisquer $v_1, v_2, v_3 \in C$,

c) $\theta(v_1, v_2) = 0 \Leftrightarrow$ existe $t > 0$ tal que $v_1 = tv_2$,

d) $\theta(t_1v_1, t_2v_2) = \theta(v_1, v_2)$ para quaisquer $v_1, v_2 \in C$ e $t_1, t_2 > 0$.

Demonstração:

a) Se $\alpha(v_2, v_1) > 0$ então

$\alpha(v_2, v_1) = \sup\{t > 0 : v_1 - tv_2 \in C\} = \sup\{t > 0 : \frac{1}{t}v_1 - v_2 \in C\} = (\inf\{s > 0 : sv_1 - v_2 \in C\})^{-1} = \beta(v_1, v_2)^{-1}$.

Por outro lado,

$\alpha(v_2, v_1) = 0 \Leftrightarrow v_1 - tv_2 \notin C$ para todo $t > 0 \Leftrightarrow sv_1 - v_2 \notin C$ para todo $s > 0 \Leftrightarrow \beta(v_1, v_2) = +\infty$.

Portanto, $\alpha(v_2, v_1) = \beta(v_1, v_2)^{-1}$ em todos os casos. Analogamente, temos $\beta(v_2, v_1) = \alpha(v_1, v_2)^{-1}$.

b) Afirmamos que $\alpha(v_1, v_2)\alpha(v_2, v_3) \leq \alpha(v_1, v_3)$ pra quaisquer $v_1, v_2, v_3 \in C$. Isto é óbvio se $\alpha(v_1, v_2) = 0$ ou $\alpha(v_2, v_3) = 0$; portanto, suporemos que $\alpha(v_1, v_2) > 0$ e $\alpha(v_2, v_3) > 0$. Então existem sequências de números positivos $(r_n)_n \nearrow \alpha(v_1, v_2)$ e $(s_n)_n \nearrow \alpha(v_2, v_3)$ tais que:

$v_2 - r_nv_1 \in C$ e $v_3 - s_nv_2 \in C \Rightarrow (v_3 - s_nv_2) + s_n(v_2 - r_nv_1) = v_3 - s_nr_nv_1 \in C$ (por convexidade) $\Rightarrow s_nr_n \leq \alpha(v_1, v_2)$ para todo $n \geq 1$. Passando o limite quando $n \to +\infty$, obtemos a afirmação. Analogamente, temos $\beta(v_1, v_2)\beta(v_2, v_3) \geq \beta(v_1, v_3)$. A desigualdade em b) segue como consequência imediata.

c) $\theta(v_1, v_2) = 0 \Leftrightarrow \alpha(v_1, v_2) = \beta(v_1, v_2) = \gamma \in (0, +\infty)$. Logo, existem sequências $(t_n)_n \nearrow \gamma$ e $(s_n)_n \searrow \gamma$ tais que

$v_2 - t_nv_1 \in C$ para todo $n \geq 1 \Rightarrow v_2 - \gamma v_1 \in \overline{C}$.

$s_nv_1 - v_2 \in C$ para todo $n \geq 1 \Rightarrow \gamma v_1 - v_2 \in \overline{C}$.

Por (2.2.1), concluímos que $v_2 - \gamma v_1 = 0$.

d) Considere quaisquer $t_1, t_2 > 0$ e $v_1, v_2 \in C$. Por definição,

$$\alpha(t_1v_1, t_2v_2) = \tfrac{t_2}{t_1}\alpha(v_1, v_2) \text{ e } \beta(t_1v_1, t_2v_2) = \tfrac{t_2}{t_1}\beta(v_1, v_2).$$

Portanto, $\theta(t_1v_1, t_2v_2) = \theta(v_1, v_2)$. $\square$

Chamamos $\theta(\cdot, \cdot)$ de **métrica projetiva** associada ao cone convexo C. Note que a métrica projetiva depende do cone de uma maneira monótona. De fato, sejam $C_1 \subset C_2$ dois cones convexos em E e $\alpha_i(\cdot,\cdot), \beta_i(\cdot,\cdot), \theta_i(\cdot,\cdot)$ as suas correspondentes métricas projetivas, $i = 1, 2$. Por definição, temos claramente que

$$\alpha_1(v_1, v_2) \leq \alpha_2(v_1, v_2) \text{ e } \beta_1(v_1, v_2) \geq \beta_2(v_1, v_2)$$

e assim $\theta_1(v_1, v_2) \geq \theta_2(v_1, v_2)$ para todo $v_1, v_2 \in C_1 \subset C_2$.

Generalizando, sejam E_1, E_2 dois espaços vetoriais e $C_i \subset E_i$, $i = 1, 2$, cones convexos. Seja $L : E_1 \to E_2$ um operador linear e assuma que $L(C_1) \subset C_2$. Então

$$\begin{aligned} \alpha_1(v_1, v_2) &= \sup\{t > 0 : v_2 - tv_1 \in C_1\} \\ &\leq \sup\{t > 0 : L(v_2 - tv_1) \in C_2\} \quad (\textit{pois} \;\; L(C_1) \subset C_2) \\ &= \sup\{t > 0 : L(v_2) - tL(v_1) \in C_2\} = \alpha_2(L(v_1), L(v_2)) \end{aligned}$$

e, de maneira análoga, podemos mostrar que $\beta_1(v_1, v_2) \geq \beta_2(L(v_1), L(v_2))$. Portanto,

$$\theta_1(v_1, v_2) \geq \theta_2(L(v_1), L(v_2)) \text{ para quaisquer } v_1, v_2 \in C_1.$$

Em geral, L pode não ser uma contração estrita, com respeito a θ_1 e θ_2. A próxima proposição afirma que este caso ocorre quando $L(C_1)$ tem θ_2-diâmetro finito.

Teorema 2.2.1 (Teorema de Birkhoff): Sejam E_1 e E_2 espaços vetoriais e sejam $C_1 \subset E_1$ e $C_2 \subset E_2$ cones projetivos. Se $L : E_1 \to E_2$ é um operador linear tal que $L(C_1) \subset C_2$ e $D = \sup\{\theta_2(L(v_1), L(v_2)) : v_1, v_2 \in C_1\}$ é finito, então

$$\theta_2(L(v_1), L(v_2)) \leq (1 - e^{-D})\theta_1(v_1, v_2) \text{ para quaisquer } v_1, v_2 \in C_1.$$

Demonstração: Nos casos em que $\alpha_1(v_1, v_2) = 0$ ou $\beta_1(v_1, v_2) = +\infty$ o resultado segue imediatamente. Suponhamos que $\alpha_1(v_1, v_2) > 0$ e $\beta_1(v_1, v_2) < +\infty$, então existem sequências de números positivos $(t_n)_n \nearrow \alpha_1(v_1, v_2)$ e $(s_n)_n \searrow \beta_1(v_1, v_2)$ tais que $v_2 - t_n v_1 \in C_1$ e $s_n v_1 - v_2 \in C_1$. Logo, $\theta_2(L(v_2 - t_n v_1), L(s_n v_1 - v_2)) \leq D$ para todo $n \geq 1$.

Considere as sequências $T_n = \alpha_2(L(v_2 - t_n v_1), L(s_n v_1 - v_2))$ e $S_n = \beta_2(L(v_2 - t_n v_1), L(s_n v_1 - v_2))$. Observe que $\lim(\log \frac{S_n}{T_n}) = \lim \theta_2(L(v_2 - t_n v_1), L(s_n v_1 - v_2)) \leq D$. Além disso,

$$L(s_n v_1 - v_2) - T_n L(v_2 - t_n v_1) \in C_2 \Leftrightarrow (s_n + t_n T_n)L(v_1) - (1 + T_n)L(v_2) \in C_2 \Rightarrow$$
$$\beta_2(L(v_1), L(v_2)) \leq \frac{s_n + t_n T_n}{1 + T_n}.$$

$$S_nL(v_2 - t_nv_1) - L(s_nv_1 - v_2) \in C_2 \Rightarrow \alpha_2(L(v_1), L(v_2)) \geq \frac{s_n + t_nS_n}{1 + S_n}.$$

Usando essas desigualdades, obtemos

$$\begin{aligned}
\theta_2(L(v_1), L(v_2)) &\leq \log(\frac{s_n + t_nT_n}{1 + T_n} \cdot \frac{1 + S_n}{s_n + t_nS_n}) \\
&= \log(\frac{s_n}{t_n} + T_n) - \log(1 + T_n) - \log(\frac{s_n}{t_n} + S_n) + \log(1 + S_n) \\
&= \int_0^{\log(s_n/t_n)} (\frac{e^x dx}{e^x + T_n} - \frac{e^x dx}{e^x + S_n}) \\
&\leq \log(\frac{s_n}{t_n}) \cdot \sup_{x>0} \frac{e^x(S_n - T_n)}{(e^x + T_n)(e^x + S_n)} \\
&\leq \log(\frac{s_n}{t_n}) \cdot (1 - \frac{T_n}{S_n}),
\end{aligned}$$

na última desigualdade usamos $(e^x + T_n)(e^x + S_n) > e^x S_n$.

Passando o limite quando $n \to +\infty$, concluímos que

$$\theta_2(L(v_1), L(v_2)) \leq \theta(v_1, v_2) \cdot (1 - e^{-D}). \qquad \square$$

Exemplo 2.2.1: Sejam $E = \mathbb{R}^2$ e $C = \{(x, y) : y > |x|\}$. O quociente projetivo de C pode ser identificado de uma maneira natural com $(-1, 1) \times \{1\}$, e assim, com o intervalo $(-1, 1)$. Dados $-1 < x_1 < x_2 < 1$, temos que

$$\begin{aligned}
\alpha(x_1, x_2) &= \sup\{t > 0 : (x_2, 1) - t(x_1, 1) \in C\} \\
&= \sup\{t > 0 : 1 - t > x_2 - tx_1\} = \frac{1 - x_2}{1 - x_1}. \\
\beta(x_1, x_2) &= \frac{x_2 + 1}{x_1 + 1}.
\end{aligned}$$

e então $\theta(x_1, x_2) = \log(\frac{x_2 + 1}{x_1 + 1} \cdot \frac{1 - x_1}{1 - x_2})$.

Exemplo 2.2.2: Sejam X um espaço métrico compacto e $E = C^0(X)$ o espaço das funções reais contínuas definidas em X. Tome o cone convexo

$$C = C_+ = \{\varphi \in E : \varphi(x) > 0 \text{ para todo } x \in X\}.$$

Então, para quaisquer $\varphi_1, \varphi_2 \in C$

$$\begin{aligned}
\alpha(\varphi_1, \varphi_2) &= \sup\{t > 0 : (\varphi_2 - t\varphi_1)(x) > 0 \ \ para \ \ todo \ \ x \in X\} \\
\sup & \ \ \{t > 0 : t < \frac{\varphi_2}{\varphi_1}(x) \ \ para \ \ todo \ \ x \in X\} \\
&= \inf\{\frac{\varphi_2}{\varphi_1}(x) : x \in X\} \ \ e \\
\beta(\varphi_1, \varphi_2) &= \sup\{\frac{\varphi_2}{\varphi_1}(x) : x \in X\}.
\end{aligned}$$

Portanto,

$$\theta(\varphi_1,\varphi_2) = \log \frac{\sup(\varphi_2/\varphi_1)}{\inf(\varphi_2/\varphi_1)} = \log \sup \left\{ \frac{\varphi_2(x)\varphi_1(y)}{\varphi_1(x)\varphi_2(y)} : x, y \in X \right\}.$$

Exemplo 2.2.3: Sejam X e E como no exemplo anterior e tome $C = C(a,\nu)$, o conjunto de todas as $\varphi \in E$ tais que $\varphi > 0$ para todo $x \in X$ e $\log \varphi$ é (a,ν)-Hölder contínua. A segunda condição equivale a dizer que

$$\begin{aligned} &\sup \left\{ \frac{|\log\varphi(x) - \log\varphi(y)|}{d(x,y)^\nu} : x, y \in X, x \neq y \right\} \leq a \\ \Leftrightarrow \quad & \exp(-ad(x,y)^\nu) \leq \frac{\varphi(x)}{\varphi(y)} \leq \exp(ad(x,y)^\nu) \ \ \textit{para} \ \ \textit{quaisquer} \ \ x, y \in X. \end{aligned}$$

Dados $\varphi_1, \varphi_2 \in C$ e $t_1, t_2 > 0$, temos que

$$\exp(-ad(x,y)^\nu) \leq \frac{t_1\varphi_1(x) + t_2\varphi_2(x)}{t_1\varphi_1(y) + t_2\varphi_2(y)} \leq \exp(ad(x,y)^\nu) \text{ para quaisquer } x, y \in X.$$

Logo, C é um cone convexo. Pela definição, $\alpha(\varphi_1,\varphi_2)$ será o supremo de todos os $t > 0$ satisfazendo, para quaisquer $x, y \in X$, a duas seguintes condições

i) $(\varphi_2 - t\varphi_1)(x) > 0 \Leftrightarrow t < \dfrac{\varphi_2}{\varphi_1}(x)$.

ii) $\dfrac{(\varphi_2 - t\varphi_1)(x)}{(\varphi_2 - t\varphi_1)(y)} \leq \exp(ad(x,y)^\nu)$
$\Leftrightarrow t(\exp(ad(x,y)^\nu)\varphi_1(y) - \varphi_1(x)) \leq \exp(ad(x,y)^\nu)\varphi_2(y) - \varphi_2(x)$
$\Leftrightarrow t \leq \dfrac{\exp(ad(x,y)^\nu)\varphi_2(y) - \varphi_2(x)}{\exp(ad(x,y)^\nu)\varphi_1(y) - \varphi_1(x)}$.

Por outro lado,

$$\frac{(\varphi_2 - t\varphi_1)(x)}{(\varphi_2 - t\varphi_1)(y)} \geq \exp(ad(x,y)^\nu) \Leftrightarrow t \leq \frac{\exp(ad(x,y)^\nu)\varphi_2(x) - \varphi_2(y)}{\exp(ad(x,y)^\nu)\varphi_1(x) - \varphi_1(y)}.$$

Assim, concluímos que $\alpha(\varphi_1,\varphi_2)$ é igual ao

$$\inf \left\{ \frac{\varphi_2(x)}{\varphi_1(x)}, \frac{\exp(ad(x,y)^\nu)\varphi_2(x) - \varphi_2(y)}{\exp(ad(x,y)^\nu)\varphi_1(x) - \varphi_1(y)} : x, y \in X, x \neq y \right\},$$

e $\beta(\varphi_1,\varphi_2)$ é dada por uma expressão semelhante, mudando apenas o supremo no lugar do ínfimo.

2.3 Operadores de Transferência e Cones Invariantes

Seja $f : M \to M$ uma aplicação expansora $C^{1+\nu_0}$ sobre uma variedade conexa compacta M. A menos de um redimensionamento, podemos supor que o diâmetro de M é menor que 1, faremos isso de agora em diante. Consideraremos m como sendo a medida de Lebesgue, normalizada de tal maneira que $m(M) = 1$.

Por tudo que foi discutido no início deste capítulo, em resumo, temos as duas seguintes propriedades:

(i) Existe $k \geq 1$ tal que todo ponto $y \in M$ tem exatamente k pré-imagens por f;

(ii) existe $\rho_0 > 0$ tal que, dados $y_1, y_2 \in M$ com $d(y_1, y_2) \leq \rho_0$, podemos escrever $f^{-1}(y_j) = \{x_{j1}, ..., x_{jk}\}, j = 1, 2$, com

$$d(x_{1i}, x_{2i}) \leq \sigma^{-1} d(y_1, y_2) \text{ para cada } i = 1, ..., k.$$

Consideraremos os dois operadores lineares $\mathcal{L}$ e U sobre $E = C^0(M, \mathbb{R})$, definidos por

$$(\mathcal{L}\varphi)(y) = \sum_{i=1}^{k} \varphi(x_i)|\det Df(x_i)|^{-1} \quad e \quad (U\varphi)(x) = \varphi \circ f(x).$$

Então, podemos escrever

$$\int (\mathcal{L}\varphi)\psi dm = \int \varphi(U\psi) dm.$$

Fixando $\rho_0 > 0$ como em (ii) e, para $a > 0$ e $0 < \nu \leq \nu_0$, definimos $C(a, \nu)$ como sendo o cone convexo das funções $\varphi \in E$ satisfazendo

(1) $\varphi(x) > 0$ para todo $x \in M$;

(2) $\log \varphi$ é (a, ν)-Hölder contínua em ρ_0-vizinhanças, isto é,

$$d(y_1, y_2) \leq \rho_0 \Rightarrow \varphi(y_1) \leq \exp(a d(y_1, y_2)^\nu)\varphi(y_2).$$

Proposição 2.3.1 (Invariância): Existe $\lambda_1 < 1$ tal que $\mathcal{L}(C(a, \nu)) \subset C(\lambda_1 a, \nu)$ para todo $a > 0$ suficientemente grande.

Demonstração: Dado $\varphi \in C(a, \nu)$, devemos mostrar que

a) $\mathcal{L}\varphi > 0$,

b) $\log \mathcal{L}\varphi$ é $C(\lambda_1 a, \nu)$-Hölder contínua em ρ_0-vizinhanças.

A parte a) segue diretamente da expressão de $(\mathcal{L}\varphi)$ ser uma soma de termos positivos e da sobrejetividade de f, que garante a existência de pelo menos uma pré-imagem para qualquer $x \in M$.

Agora, demonstraremos a condição b). Sejam $y_1, y_2 \in M$ com $d(y_1, y_2) \leq \rho_0$ e denote $f^{-1}(y_j) = \{x_{j1}, ..., x_{jk}\}$, $j = 1, 2$, assim como fizemos em (ii). Usaremos o fato de que $\log|\det Df(x)|$ é (a_0,ν_0)-Hölder para algum $a_0>0$, isto é, para quaisquer $x,y\in M$ vale

$$\begin{aligned} |\det Df(y)| &\leq \exp(a_0 d(x,y)^{\nu_0})|\det Df(x)| \\ \Rightarrow |\det Df(x)|^{-1} &\leq |\det Df(y)|^{-1}\exp(a_0 d(x,y)^{\nu_0}). \end{aligned}$$

Logo, temos que para todo $\varphi\in C(a,\nu)$, vale

$$\begin{aligned} (\mathcal{L}\varphi)(y_1) &= \sum_{i=1}^{k}\varphi(x_{1i})|\det Df(x_{1i})|^{-1} \\ &\leq \sum_{i=1}^{k}\varphi(x_{2i})\exp(ad(x_{1i},x_{2i})^{\nu})\cdot|\det Df(x_{2i})|^{-1}\exp(a_0 d(x_{1i},x_{2i})^{\nu_0}) \\ &\leq \exp((a\sigma^{-\nu}+a_0)d(y_1,y_2)^{\nu})\sum_{i=1}^{k}\varphi(x_{2i})|\det Df(x_{2i})|^{-1}. \end{aligned}$$

Na última desigualdade, usamos o fato de que $d(x_{1i},x_{2i})^{\nu_0}\leq d(y_1,y_2)^{\nu_0}\leq d(y_1,y_2)^{\nu}$ (pois $\nu\leq\nu_0$ e o diâmetro de M é menor que 1). Usamos também a desigualdade $d(x_{1i},x_{2i})\leq \sigma^{-1}d(y_1,y_2)$.

Agora, considere a função contínua $g(x)=a\sigma^{-\nu}-ax+a_0$, observe que $g(\sigma^{-\nu})=a_0>0$ e $g(1)<0\Leftrightarrow a>\dfrac{a_0}{(1-\sigma^{-\nu})}$. Pelo Teorema do Valor Intermediário, se $a>\dfrac{a_0}{(1-\sigma^{-\nu})}$ então existe $\lambda_1\in(\sigma^{-\nu},1)$ tal que $g(\lambda_1)=0\Rightarrow a\sigma^{-\nu}+a_0=\lambda_1 a$. Segue que

$$\begin{aligned} (\mathcal{L}\varphi)(y_1) &\leq \exp((a\sigma^{-\nu}+a_0)d(y_1,y_2)^{\nu})\sum_{i=1}^{k}\varphi(x_{2i})|\det Df(x_{2i})|^{-1} \\ &= \exp(\lambda_1 a d(y_1,y_2)^{\nu})\cdot(\mathcal{L}\varphi)(y_2). \quad \square \end{aligned}$$

Denotamos por $\theta_{a,\nu}$ a métrica projetiva associada ao cone convexo $C(a,\nu)$. Então, pelo Exemplo 1.3, $\theta(\varphi_1,\varphi_2)=\log(\beta(\varphi_1,\varphi_2)/\alpha(\varphi_1,\varphi_2))$ onde $\alpha(\varphi_1,\varphi_2)$ é dado pelo

$$\inf\left\{\frac{\varphi_2(x)}{\varphi_1(x)},\frac{\exp(ad(x,y)^{\nu})\varphi_2(x)-\varphi_2(y)}{\exp(ad(x,y)^{\nu})\varphi_1(x)-\varphi_1(y)}: x,y\in X, x\neq y \;\; e \;\; d(x,y)\leq\rho_0\right\},$$

e $\beta(\varphi_1,\varphi_2)$ tem uma expressão similar, com o sup no lugar do inf. No que se segue, também usaremos o cone convexo

$$C_+=\{\varphi\in E:\varphi(x)>0 \text{ para todo } x\in M\}.$$

Pelo Exemplo 2.2.3, temos que a métrica projetiva θ_+ associada a C_+ é dada por $\theta_+(\varphi_1,\varphi_2)=\log(\beta_+(\varphi_1,\varphi_2)/\alpha_+(\varphi_1,\varphi_2))$ com

$$\alpha_+(\varphi_1,\varphi_2)=\inf\{\frac{\varphi_2}{\varphi_1}(x):x\in M\}\quad \text{e}\quad \beta_+(\varphi_1,\varphi_2)=\sup\{\frac{\varphi_2}{\varphi_1}(x):x\in M\}.$$

Proposição 2.3.2 (Diâmetro Finito): $D_1=\sup\{\theta(\varphi_1,\varphi_2):\varphi_1,\varphi_2\in C(\lambda_1 a,\nu)\}$ é finito, para quaisquer $a>0,\nu>0,\lambda_1<1$.

Demonstração: A demonstração será feita em dois passos: Primeiro mostraremos que

$$\theta\text{-diâmetro}(C(\lambda_1 a,\nu))\leq\theta_+\text{-diâmetro}(C(\lambda_1 a,\nu))+K'(\lambda_1),$$

E em seguida, obteremos

$$\theta_+\text{-diâmetro}(C(a,\nu))\leq K''(a),$$

Onde $K'(\cdot),K''(\cdot)<+\infty$.

Passo 1: Dados quaisquer $\varphi_1,\varphi_2\in C(\lambda_1 a,\nu)$,

$$\begin{aligned}
&\frac{\exp(ad(x,y)^\nu)\varphi_2(x)-\varphi_2(y)}{\exp(ad(x,y)^\nu)\varphi_1(x)-\varphi_1(y)}\\
\geq\ &\frac{\exp(ad(x,y)^\nu)\varphi_2(x)-\varphi_2(x)\exp(a\lambda_1 d(x,y)^\nu)}{\exp(ad(x,y)^\nu)\varphi_1(x)-\varphi_1(x)\exp(-a\lambda_1 d(x,y)^\nu)}\\
=\ &\frac{\varphi_1}{\varphi_2}(x)\cdot\frac{\exp(ad(x,y)^\nu)-\exp(a\lambda_1 d(x,y)^\nu)}{\exp(ad(x,y)^\nu)-\exp(-a\lambda_1 d(x,y)^\nu)}\geq K_1\frac{\varphi_1}{\varphi_2}(x),
\end{aligned}$$

onde

$$K_1=\inf\left\{\frac{k-k^{\lambda_1}}{k-k^{-\lambda_1}}:k>1\right\}.$$

Note que $K_1\in(0,1)$, pois

$$\lim_{k\to+\infty}\frac{k-k^{\lambda_1}}{k-k^{-\lambda_1}}=1,$$

pela hipótese de $\lambda_1<1$, e

$$\lim_{k\to 1^+}\frac{k-k_1^{\lambda}}{k-k^{-\lambda_1}}=\frac{1-\lambda_1}{1+\lambda_1}<1.$$

Segue, das expressões para α, α_+ e a desigualdade obtida, que $\alpha(\varphi_1,\varphi_2)\geq K_1\alpha_+(\varphi_1,\varphi_2)$. Analogamente, temos $\beta(\varphi_1,\varphi_2)\leq K_2\beta_+(\varphi_1,\varphi_2)$, com

$$K_2=\sup\left\{\frac{k-k^{-\lambda_1}}{k-k^{\lambda_1}}:k>1\right\}\in(1,+\infty).$$

Dessa forma, obtemos $\theta(\varphi_1, \varphi_2) \leq \theta_+(\varphi_1, \varphi_2) + \log K_2 - \log K_1$ e isto conclui o primeiro passo da demonstração.

Passo 2: Começaremos mostrando que se $\varphi \in C(a, \nu)$, então $\log \varphi$ é (b, ν)-Hölder contínua sobre toda a variedade M (não apenas em ρ_0-vizinhanças), para algum $b > 0$. Para isso, mostraremos que existe $N \geq 1$ (dependendo apenas de M e ρ_0) tal que

(*) Dados quaisquer $x, y \in M$ existem $z_0 = x, z_1, ..., z_N = y$ com $d(z_{i-1}, z_i) \leq \rho_0$ para todo $i = 1, ..., N$.

Considere a cobertura aberta dada por $\{B(x, \frac{\rho_0}{2}) : x \in M\}$. Por compacidade, podemos tomar uma subcobertura aberta finita, digamos $\{B_i = B(x_i, \frac{\rho_0}{2}) : i = 1, ..., k\}$. Seja β a cobertura aberta formada por todas as interseções finitas entre $B_1, B_1^C, B_2, B_2^C, ..., B_k, B_k^C$. Note que β tem uma quantidade finita de elementos e $N = \#\beta$ satisfaz a condição (*).

Se $d(x, y) < \rho_0$, então a Hölder continuidade de $\log \varphi$ é garantida pelo fato de $\varphi \in C(a, \nu)$. Logo, suporemos que $d(x, y) \geq \rho_0$. Segue que

$$
\begin{aligned}
\frac{\varphi(y)}{\varphi(x)} &= \prod_{i=1}^{N} \frac{\varphi(z_i)}{\varphi(z_{i-1})} \leq \prod_{i=1}^{N} \exp(ad(z_{i-1}, z_i)^\nu) \\
&= \exp(\sum_{i=1}^{N} ad(z_{i-1}, z_i)^\nu) \leq \exp(\sum_{i=1}^{N} a \cdot (\rho_0)^\nu) \\
&\leq \exp(\sum_{i=1}^{N} ad(x, y)^\nu) = \exp(Nad(x, y)^\nu),
\end{aligned}
$$

e basta tomarmos $b = Na$.

Agora, observe que

$$
\theta_+(\varphi_1, \varphi_2) = \log \frac{\beta_+(\varphi_1, \varphi_2)}{\alpha_+(\varphi_1, \varphi_2)} = \log \sup \left\{ \frac{\varphi_2(x)}{\varphi_1(x)} \cdot \frac{\varphi_1(y)}{\varphi_2(y)} : x, y \in M \right\}.
$$

A propriedade de $\log \varphi$ ser (b, ν)-Hölder contínua implica que

$$
\log \frac{\varphi_2(x)}{\varphi_2(y)} \leq bd(x, y)^\nu \leq b \quad \text{e} \quad \log \frac{\varphi_1(y)}{\varphi_1(x)} \leq b \quad \text{para quaisquer } x, y \in M.
$$

Portanto, $\theta_+(\varphi_1, \varphi_2) \leq 2 \log b$. $\square$

Como consequência do Teorema 2.2.1(Birkhoff) e das proposições 2.3.1 e 2.3.2, temos que o operador $\mathcal{L} : C(a, \nu) \to C(a, \nu)$ é uma Λ_1-contração com respeito a métrica $\theta = \theta_{a,\nu}$, onde $\Lambda_1 = 1 - e^{-D_1}$.

2.4 Medida Invariante Absolutamente Contínua.

Nessa seção, mostraremos que f admite uma medida invariante μ_0 que é absolutamente contínua com respeito a Lebesgue. Para a construção dessa medida, será necessário que $\mathcal{L}$ tenha algum ponto fixo φ_0. Começaremos mostrando que o cone convexo C_+ é completo em relação a correspondente métrica projetiva θ_+.

Observação 2.4.1: Dado um cone convexo C, a sua respectiva métrica projetiva θ define uma distância no quociente projetivo de $C\backslash\{0\}$, ou seja, no conjunto das classes de equivalência da relação $\sim$ definida por $v_1 \sim v_2 \Leftrightarrow v_1 = tv_2$ para algum $t > 0$. Assim, dado algum φ em $C/\sim$, podemos considerar a normalização $\int \varphi dm = 1$. De fato, basta tomarmos $t = 1/\int \varphi dm > 0$ e teremos que o representante $\varphi_1 = t\varphi \in \overline{\varphi}$ cumpre a normalização.

Proposição 2.4.1 (Completude): Toda sequência θ_+-Cauchy $(\varphi_n)_n$ em C_+ é θ_+-convergente em C_+. Além disso, se normalizarmos $\int \varphi_n dm = 1$ para todo $n \geq 1$, então $(\varphi_n)_n$ é também uniformemente convergente.

Demonstração: Seja $(\varphi_n)_n$ uma sequência θ_+-Cauchy, normalizada por $\int \varphi_n dm = 1$ para todo $n \geq 1$. Em particular, temos que $(\varphi_n)_n$ é θ_+-limitada e então

$$\begin{aligned}\theta_+(\varphi_1, \varphi_n) \quad &= \quad \log\sup\left\{\frac{\varphi_n(x)\varphi_1(y)}{\varphi_1(x)\varphi_n(y)} : x, y \in X\right\} \leq R \\ &\Longrightarrow \quad \exp(-R) \leq \frac{\varphi_n(x)\varphi_1(y)}{\varphi_1(x)\varphi_n(y)} \leq \exp(R).\end{aligned}$$

para todo $x, y \in M$ e todo $n \geq 1$. Em particular,

$$\frac{1}{R_1} \leq \frac{\varphi_n(x)}{\varphi_n(y)} \leq R_1,$$

para todo $x, y \in M$ e todo $n \geq 1$, onde $R_1 = \exp(R) \cdot \sup\{\varphi_1(s)/\varphi_1(t) : s, t \in M\}$. Por outro lado, temos que $\inf \varphi_n = \inf \varphi_n m(M) \leq \int \varphi_n dm = 1$ e analogamente $\sup \varphi_n \geq 1$. Segue que

$$\frac{1}{R_1} \leq \frac{\varphi_n(x)}{\sup \varphi_n} \leq \varphi_n(x) \leq \frac{\varphi_n(x)}{\inf \varphi_n} \leq R_1 \text{ para todo } x \in M \text{ e } n \geq 1.$$

A condição de Cauchy é equivalente a afirmar que, dado qualquer $\varepsilon > 0$, existe $N \geq 1$ tal que para quaisquer $k, l \geq N$

$$\frac{\sup(\varphi_k/\varphi_l)}{\inf(\varphi_k/\varphi_l)} \leq e^{\varepsilon} \Rightarrow e^{-\varepsilon} \leq \inf \frac{\varphi_k}{\varphi_l} \leq 1 \leq \sup \frac{\varphi_k}{\varphi_l} \leq e^{\varepsilon}, \tag{2.5}$$

pois $\int \varphi_k dm = \int \varphi_l dm$ implica que existem $x, y \in M$ tais que $\varphi_k(x) \geq \varphi_l(x)$ e $\varphi_k(y) \leq \varphi_l(y)$ e assim, temos $\inf(\varphi_k/\varphi_l) \leq 1 \leq \sup(\varphi_k/\varphi_l)$. Segue que

$$\sup|\varphi_k - \varphi_l| \leq \sup|\varphi_l| \cdot \sup\left|\frac{\varphi_k}{\varphi_l} - 1\right| \leq R_2(e^{\varepsilon} - 1) \quad com \quad R_2 = \sup \varphi_l.$$

Isto significa que $(\varphi_n)_n$ é uma sequência de Cauchy com respeito a métrica uniforme (supremo), e então é uniformemente convergente. Seja φ_0 o limite uniforme, observe que $\varphi_0 \geq R_1^{-1}$ e assim $\varphi_0 \in C_+$. Tomando o limite na desigualdade (2.5), quando $l \to +\infty$, obtemos

$$e^{-\varepsilon} \leq \inf \frac{\varphi_k}{\varphi_0} \leq 1 \leq \sup \frac{\varphi_k}{\varphi_0} \leq e^{\varepsilon},$$

para todo $k \geq N$. Isto prova que ambos $\sup(\varphi_k/\varphi_0)$ e $\inf(\varphi_k/\varphi_0)$ convergem para 1, e dessa maneira $\theta_+(\varphi_n, \varphi_0)$ converge para zero, quando $n \longrightarrow +\infty$. $\square$

Aplicaremos o resultado anterior na sequência $\varphi_n = \mathcal{L}^n 1$. Observe que as expressões de θ_+ e $\theta = \theta_{a,\nu}$ junto ao fato de que $C(a,\nu) \subset C_+$, implicam que $\theta_+ \leq \theta$. Então, como $\mathcal{L}$ é uma θ-contração, segue que $(\varphi_n)_n$ é θ-Cauchy e assim é também θ_+-Cauchy. Por outro lado,

$$\int \varphi_n dm = \int (\mathcal{L}^n 1) \cdot 1 dm = \int 1 \cdot (1 \circ f^n) = \int 1 dm = 1.$$

Logo, φ_n converge uniformemente para algum $\varphi_0 \in C_+$. Note que, como a sequência $(\varphi_n)_n$ está em $C(a,\nu)$ e a condição de Hölder continuidade na definição desses cones é fechada para limites uniformes, temos $\varphi_0 \in C(a,\nu)$.

Como o operador $\mathcal{L}$ é contínuo em E em relação a métrica $||\cdot||_0$, então

$$\varphi_{n+1} = \mathcal{L}\varphi_n \Rightarrow \lim_n \varphi_{n+1} = \lim_n \mathcal{L}\varphi_n \Rightarrow \varphi_0 = \mathcal{L}\varphi_0.$$

Segue que φ_0 é um ponto fixo de $\mathcal{L}$. Como uma consequência, $\mu_0 = \varphi_0 m$ é uma medida de probabilidade f-invariante:

$$\int \varphi d\mu_0 = \int \varphi\varphi_0 dm = \int \varphi(\mathcal{L}\varphi_0) dm = \int (\varphi \circ f)\varphi_0 dm = \int (\varphi \circ f) d\mu_0$$

para toda $\varphi \in C^0(M, \mathbb{R})$. Finalmente, $d\mu_0/dm = \varphi_0 \geq R_1^{-1} > 0$ e então μ_0 é equivalente a medida de Lebesgue m.

2.5 Mistura Exponencial

Seja $\varphi \in C(\lambda_1 a, \nu)$. Usando o Teorema 2.2.1(Birkhoff) e o fato de φ_0 ser um ponto fixo de $\mathcal{L}$, temos que

$$\theta_+(\mathcal{L}^n\varphi, \varphi_0) \leq \theta(\mathcal{L}^n\varphi, \varphi_0) = \theta(\mathcal{L}^n\varphi, \mathcal{L}^n\varphi_0) \leq \theta(\varphi, \varphi_0)\Lambda_1^n \leq D_1\Lambda_1^n,$$

onde $D_1 = \sup\{\theta(\varphi_1, \varphi_2) : \varphi_1, \varphi_2 \in C(a, \nu)\}$ e $\Lambda_1 = 1 - e^{-D}$. Então

$$\frac{\sup(\mathcal{L}^n\varphi/\varphi_0)}{\inf(\mathcal{L}^n\varphi/\varphi_0)} \leq e^{D_1\Lambda_1^n} \Rightarrow e^{-D_1\Lambda_1^n} \leq \inf\frac{\mathcal{L}^n\varphi}{\varphi_0} \leq 1 \leq \sup\frac{\mathcal{L}^n\varphi}{\varphi_0} \leq e^{D_1\Lambda_1^n},$$

pois $\int \mathcal{L}^n\varphi dm = \int \varphi dm = 1 = \int \varphi_0 dm$ implica que $\inf(\mathcal{L}^n\varphi/\varphi_0) \leq 1 \leq \sup(\mathcal{L}^n\varphi/\varphi_0)$, analogamente ao que foi feito na seção anterior. Segue que

$$\sup|\mathcal{L}^n\varphi - \varphi_0| \leq \sup|\varphi_0| \cdot \sup\left|\frac{\mathcal{L}^n\varphi}{\varphi_0} - 1\right| \leq R_2(e^{D_1\Lambda_1^n} - 1) \leq R_3\Lambda_1^n \text{ (com } R_2 = \sup\varphi_0),$$

para alguma constante $R_3 > 0$ e todo $n \geq 1$. Isso pode ser visto como uma demonstração da perda exponencial de memória no sistema: os iterados $(\mathcal{L}^n\varphi)m$ de uma distribuição de massa inicial φm converge exponencialmente rápido para a distribuição equilíbrio $\varphi_0 m$. Mostraremos agora, que também existe um limite exponencial para as funções de correlação.

Proposição 2.5.1: Dadas φ uma função ν-Hölder e $\psi \in L^1(m)$ sobre M, existe $K_0 = K_0(\varphi, \psi) > 0$ tal que

$$\left|\int (\psi \circ f^n)\varphi dm - \int \psi d\mu_0 \int \varphi dm\right| \leq K_0\Lambda_1^n$$

para todo $n \geq 0$.

Demonstração: Suponha primeiramente que $\varphi \in C(\lambda_1 a, \nu)$. Sem perda de generalidade, assuma que $\int \varphi dm = 1$. Então, denotando $||\psi||_1 = \int |\psi| d\mu_0$, temos

$$\begin{aligned}
\left|\int (\psi \circ f^n)\varphi dm - \int \psi d\mu_0\right| &= \left|\int \psi(\frac{\mathcal{L}^n\varphi}{\varphi_0} - 1)d\mu_0\right| \leq \left\|\frac{\mathcal{L}^n\varphi}{\varphi_0} - 1\right\|_0 \cdot ||\psi||_1 \\
&\leq (e^{D_1\Lambda_1^n} - 1)||\psi||_1 \leq R_3||\psi||_1\Lambda_1^n.
\end{aligned}$$

Agora, seja φ uma função ν-Hölder qualquer e $A > 0$ tal que φ é (A, ν)-Hölder. Para $B > 0$ qualquer, escrevemos

$$\varphi = \varphi_B^+ - \varphi_B^- \quad \text{onde} \quad \varphi_B^\pm = \frac{1}{2}(|\varphi| \pm \varphi) + B.$$

Claramente, $\varphi_B^{\pm}$ são (A,ν)-Hölder contínuas e $\varphi_B^{\pm} \geq B$. Se tomarmos $B = (A/\lambda_1 a)$ teremos $\varphi_B^{\pm} \in C(\lambda_1 a, \nu)$ e então a proposição é válida para $\varphi_B^{\pm}$. Por linearidade a proposição se dá para φ. $\square$

Observação 2.5.1: Note que a constante $K_0 = K_0(\varphi,\psi)$ construída na demonstração da proposição 2.5.1 tem a forma

$$K_0(\varphi,\psi) \leq K_0' ||\psi||_1(||\varphi||_1 + H_\nu(\varphi)),$$

onde $H_\nu(\varphi)$ denota algum número A tal que φ é (A,ν)-Hölder, e $K_0' > 0$ é independente de φ e ψ. De fato, pela primeira parte da demonstração temos

$$K_0(\varphi_B^{\pm},\psi) \leq R_3||\psi||_1 \int \varphi_B^{\pm} dm \leq R_3||\psi||_1(\int |\varphi| dm + B) \leq R_3||\psi||_1(R_2||\varphi||_1 + \frac{A}{a\lambda_1}).$$

e concluímos a afirmação notando que $K_0(\varphi,\psi) \leq K_0(\varphi_B^{+},\psi) + K_0(\varphi_B^{-},\psi)$.

Corolário 2.5.2 (Decaimento exponencial de Correlações): Dadas as funções ν-Hölder contínuas φ_1 e φ_2, existe $K = K(\varphi_1,\varphi_2) > 0$ tal que

$$\left| \int (\varphi_1 \circ f^n)\varphi_2 d\mu_0 - \int \varphi_1 d\mu_0 \int \varphi_2 d\mu_0 \right| \leq K\Lambda_1^n,$$

para todo $n \geq 0$.

Demonstração: Basta tomar $\psi = \varphi_1$ e $\varphi = \varphi_2\varphi_0$ na Proposição 2.5.1. $\square$

Agora vamos definir o que significa uma medida ser exata:

Sejam $\mathcal{B}_n = f^{-n}(\mathcal{B})$, para $n \geq 0$, onde $\mathcal{B}$ é a σ-álgebra de Borel de M. Cada $\mathcal{B}_n$ é claramente uma σ-álgebra em M. Uma função $\xi : M \to \mathbb{R}$ é mensurável com respeito a $\mathcal{B}_n$ se e somente se $\xi = \xi_n \circ f^n$ para alguma função ξ_n mensurável (em relação a $\mathcal{B}$). Além disso, $\mathcal{B}_0 = \mathcal{B} \supset \mathcal{B}_1 \supset \cdots \supset \mathcal{B}_n \supset \cdots$. Uma medida μ f-invariante é chamada **exata** se a σ-álgebra

$$\mathcal{B}_\infty = \bigcap_{n\geq 0} \mathcal{B}_n$$

é μ-trivial, no sentido de que todas as funções $\mathcal{B}_\infty$-mensurável são constantes em μ-quase todo ponto. Note que medidas exatas são ergódicas: Se $A \subset M$ é um conjunto f-invariante, então $\chi_A \in \mathcal{B}_\infty$.

Corolário 2.5.3 (Unicidade e Exatidão): A medida μ_0 é exata e é a única medida f-invariante que é absolutamente contínua com respeito a m.

Demonstração: Já mostramos que μ_0 e m são medidas equivalentes. Seja $\psi \in L^1(\mu_0)$ uma função $\mathcal{B}_\infty$-mensurável. Logo, para todo $n \geq 0$ existe uma função mensurável ψ_n tal que $\psi = \psi_n \circ f^n$. Note que $||\psi||_1 = \int |\psi| d\mu_0 = \int |\psi_n \circ f^n| d\mu_0 = \int |\psi_n| d\mu_0 = ||\psi_n||_1$. Pela Proposição 2.5.1 e a Observação 2.5.1, dada qualquer função ν-Hölder contínua φ, existe $K_0''(\varphi) > 0$ tal que

$$\begin{aligned}\left|\int (\psi - \int \psi d\mu_0)\varphi dm\right| &= \left|\int (\psi_n \circ f^n)\varphi dm - \int \psi_n d\mu_0 \int \varphi dm\right| \\ &\leq K_0' ||\psi_n||_1 (||\varphi||_1 + H_\nu(\varphi))\Lambda_1^n = K_0''(\varphi)||\psi_n||_1 \Lambda_1^n \\ &= K_0''(\varphi)||\psi||_1 \Lambda_1^n \to 0,\end{aligned}$$

e então

$$\int (\psi - \int \psi d\mu_0)\varphi dm = 0.$$

Segue que $\psi = \int \psi d\mu_0$ em quase todo ponto (com respeito a m e μ_0), isso equivale a dizer que ψ é constante em μ_0-quase todo ponto. Portanto, μ_0 é uma medida exata. Finalmente, já que μ_0 ser exata implica em μ_0 ser ergódica, temos que se μ é outra medida f-invariante com $\mu \ll m$ então $\mu \ll \mu_0$ e assim, por ergodicidade, $\mu = \mu_0$. $\square$

Seja $L^2(\mathcal{B}_n) = \{\xi \in L^2(\mu_0) : \xi \text{ é } \mathcal{B}_n\text{-mensurável}\}$, para cada valor de $n \geq 0$. Observe que $L^2(\mu_0) = L^2(\mathcal{B}_0) \supset L^2(\mathcal{B}_1) \supset ... \supset L^2(\mathcal{B}_n) \supset ...$ e

$$\mu_0 \quad \text{é} \quad \text{exata} \quad \Leftrightarrow \quad \bigcap_{n \geq 0} L^2(\mathcal{B}_n) = \{\text{constantes}\}.$$

Dado $\varphi \in L^2(\mu_0)$ e $n \geq 0$, denotaremos por $E(\varphi \mid \mathcal{B}_n)$ a projeção ortogonal de φ sobre $L^2(\mathcal{B}_n)$. Analogamente, denotaremos por $E(\varphi \mid \mathcal{B}_n^\perp)$, a projeção ortogonal de φ sobre $L^2(\mathcal{B}_n)^\perp$ (onde o colplemento ortogonal é tomado no espaço $L^2(\mu_0)$).

Corolário 2.5.4: Para toda função φ ν-Hölder contínua com $\int \varphi d\mu_0 = 0$, existe $R_0 = R_0(\varphi)$ tal que $||E(\varphi \mid \mathcal{B}_n)||_2 \leq R_0 \Lambda_1^n$ para todo $n \geq 0$.

Demonstração: Basta observar que

$$\begin{aligned}&\sup \quad \{\int \xi \varphi d\mu_0 : \xi \in L^2(\mathcal{B}_n) \quad e \quad ||\xi||_2 = 1\} = \\ &\sup \quad \{\int \xi E(\varphi \mid \mathcal{B}_n) d\mu_0 : \xi \in L^2(\mathcal{B}_n) \quad e \quad ||\xi||_2 = 1\} = \\ &\sup \quad \{||E(\varphi \mid \mathcal{B}_n)||_2 \cdot ||\xi||_2 \cos\theta(E(\varphi \mid \mathcal{B}_n), \xi) : \xi \in L^2(\mathcal{B}_n) \quad e \quad ||\xi||_2 = 1\},\end{aligned}$$

onde $\theta(E(\varphi \mid \mathcal{B}_n), \xi)$ é o ângulo entre $E(\varphi \mid \mathcal{B}_n)$ e ξ. Esse supremo será atingido em

$$\xi = \frac{1}{||(E(\varphi \mid \mathcal{B}_n)||_2} \cdot E(\varphi \mid \mathcal{B}_n).$$

Portanto, temos

$$\begin{aligned}||E(\varphi \mid \mathcal{B}_n)||_2 &= \sup\{\int \xi\varphi d\mu_0 : \xi \in L^2(\mathcal{B}_n) \ \ e \ \ ||\xi||_2 = 1\} \\ &= \sup\{\int (\psi \circ f^n)\varphi\varphi_0 dm : \psi \in L^2(\mu_0) \ \ e \ \ ||\psi||_2 = 1\} \\ &= K_0''(\varphi, \varphi_0)\Lambda_1^n,\end{aligned}$$

pois $||\psi||_1 \leq ||\psi||_2 = 1$ e supomos que $\int \varphi d\mu_0 = \int \varphi\varphi_0 dm = 0$. $\square$

2.6 Teorema Central do Limite.

Baseado nas análises prévias, nessa seção mostraremos que a oscilação das somas de Birkhoff de uma função ν-Hölder em torno de seu valor esperado converge, em distribuição, para o processo gaussiano. Primeiro, enunciaremos o Teorema Central do Limite para martingales, que será muito útil para nossas pretensões.

Começamos relembrando algumas noções de teoria da probabilidade. Dada uma sequência X_n, $n \geq 0$, de variáveis aleatórias reais em um espaço de probabilidade (M, Σ, μ), considere a aplicação shift

$$\tau : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}, \qquad (y_n)_{n\geq 0} \mapsto (y_{n+1})_{n\geq 0},$$

e a medida de probabilidade ν em $\mathbb{R}^{\mathbb{N}}$ dada por

$$\nu([k; A_0, ..., A_m]) = \mu(X_k^{-1}(A_0)) \cdots \mu(X_{k+m}^{-1}(A_m)),$$

(onde $[k; A_0, ..., A_m] = \{(y_i)_{i\geq 0} : y_k \in A_0, ..., y_{k+m} \in A_m\}$), para cada $k \geq 0$, $m \geq 0$ e intervalos $A_0, ..., A_m \subset \mathbb{R}$. Então, a sequência $(X_n)_{n\geq 0}$ é dita ser **estacionária** se a medida ν é τ-invariante, e **ergódica** se ν é τ-ergódica. Um caso particular relevante ocorre quando temos $X_n = X_0 \circ f^n$ para todo $n \geq 0$, onde X_0 é alguma função mensurável e $f : M \to M$ é uma aplicação mensurável. Para este caso, temos que a sequência é estacionária se μ é f-invariante, e ela é ergódica se μ é f-ergódica.

Definição 2.6.1: Seja $(\mathcal{F}_n)_n$, $n \geq 0$, uma sequência não-crescente de σ-álgebras em M. Então $(X_n)_{n\geq 0}$ é uma **diferença martingale reversa** para $(\mathcal{F}_n)_{n\geq 0}$ se

(a) X_n é $\mathcal{F}_n$-mensurável para todo $n \geq 0$;

(b) $\int_A X_n d\mu = 0$ para todo $A \in \mathcal{F}_{n+1}$.

Definimos todos os conceitos necessários para enunciar o seguinte Teorema.

Teorema 2.6.1 (Central do Limite Para Martingales): Seja $(X_n)_{n\geq 0}$ uma sequência estacionária, ergódica, diferença martingale reversa, tal que $\sigma^2 = E(X_0^2)$ é estritamente positiva e finita. Então, para todo $z \in \mathbb{R}$,

$$\mu\left(\{x \in M : \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} X_j(x) < z\}\right) \to \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{z} e^{-t^2/2\sigma^2}dt, \quad quando \quad n \to \infty.$$

Demonstraremos uma versão do Teorema Central do Limite para transformações mensuráveis. Antes disso, é útil fazer uma observação que será utilizada nessa demonstração.

Observação 2.6.1: Sejam V um espaço vetorial normado e $A, B \subset V$ subespaços vetoriais. Vamos mostrar que se $B \subset A$ então para todo $x \in V$, vale que $E(x \mid A) - E(x \mid B)$ é a projeção ortogonal de $E(x \mid A)$ em $B^\perp$ (onde $E(x \mid A)$ denota a projeção ortogonal de x sobre A). De fato, x pode ser escrito de maneira única na forma

$$x = y + y', \quad onde \quad y = E(x \mid B) \quad e \quad y' = E(x \mid B^\perp).$$

Logo,

$$E(x \mid A) = E(y \mid A) + E(y' \mid A) = y + E(y' \mid A), \tag{2.6}$$

pois $y \in B \subset A$. Analogamente, y' pode ser escrito de forma única como

$$y' = E(y' \mid A) + E(y' \mid A^\perp) \Rightarrow E(y' \mid A) = y' - E(y' \mid A^\perp).$$

Como $B \subset A$ implica $A^\perp \subset B^\perp$, temos $E(y' \mid A) = y' - E(y' \mid A^\perp) \in B^\perp$, por ser a soma de dois vetores em $B^\perp$. Então a igualdade (2.6) exprime $E(x \mid A)$ como uma soma de um vetor em B com um vetor em $B^\perp$ e, por unicidade, $E(x \mid A) - y = E(x \mid A) - E(x \mid B)$ é a projeção ortogonal de $E(x \mid A)$ sobre $B^\perp$.

Teorema 2.6.2: Sejam $(M, \mathcal{B}, \mu)$ um espaço de probabilidade, $f : M \to M$ uma aplicação mensurável tal que μ é f-invariante e f-ergódica, e $\phi \in L^2(\mu)$ tal que $\int \phi d\mu = 0$. Seja $\mathcal{B}_n$ a sequência não decrescente das σ-álgebras $\mathcal{B}_n = f^{-n}(\mathcal{B}), n \geq 0$. Assuma que

$$\sum_{n=0}^{\infty} ||E(\phi \mid \mathcal{B}_n)||_2 < \infty.$$

Então $\sigma \geq 0$ dado por

$$\sigma^2 = \int \phi^2 d\mu + 2\sum_{n=1}^{\infty}\int \phi(\phi \circ f^n)d\mu \tag{2.7}$$

é finito e $\sigma = 0$ se e somente se $\phi = u \circ f - u$ para algum $u \in L^2(\mu)$. Por outro lado, se $\sigma > 0$ então, dado qualquer intervalo $A \subset \mathbb{R}$,

$$\mu\left\{x \in M : \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \phi(f^j(x)) \in A\right\} \to \frac{1}{\sigma\sqrt{2\pi}} \int_A e^{-\frac{t^2}{2\sigma^2}} dt,$$

quando $n \to +\infty$.

Como uma consequência imediata desse Teorema 2.6.2 e do Corolário 2.5.4, obtemos o seguinte teorema central do limite para aplicações expansoras:

Proposição 2.6.3 (Teorema Central do Limite): Sejam φ uma função ν-Hölder contínua e

$$\sigma^2 = \int \phi^2 d\mu_0 + 2\sum_{j=1}^{\infty} \int \phi(\phi \circ f^j) d\mu_0,$$

onde $\phi = \varphi - \int \varphi d\mu_0$. Então σ está bem definida e $\sigma = 0$ se e somente se $\varphi = u \circ f - u$ para algum $u \in L^2(\mu_0)$. Se $\sigma > 0$ então, para todo intervalo $A \subset \mathbb{R}$,

$$\mu_0\left\{x \in M : \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} (\varphi(f^j(x)) - \int \varphi d\mu_0) \in A\right\} \to \frac{1}{\sigma\sqrt{2\pi}} \int_A e^{-\frac{t^2}{2\sigma^2}} dt,$$

quando $n \to +\infty$.

No restante dessa seção iremos demonstrar o Teorema 2.6.2.

Demonstração: A ideia geral é escrever $\phi = \eta + \varsigma \circ f - \varsigma$, com funções $\eta, \varsigma \in L^2(\mu)$ tais que

i) $\dfrac{1}{\sqrt{n}}(\varsigma \circ f^n - \varsigma)$ converge para zero em medida, quando $n \to +\infty$ (provaremos a convergência em $L^2(\mu)$, que é um fato mais forte);

ii) as variáveis aleatórias $\eta \circ f^n$ satisfazem a seguinte condição de independência

$$E(\eta \circ f^n \mid \mathcal{B}_{n+1}) = 0 \text{ para todo } n \geq 0.$$

A primeira propriedade significa que, dado qualquer $\varepsilon > 0$,

$$\lim_n \mu\left\{x \in M : \frac{1}{\sqrt{n}} |\zeta(f^n(x)) - \zeta(x)| > \varepsilon\right\} \to 0,$$

e implica que limite da distribuição de

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} (\phi \circ f^j) = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} (\eta \circ f^j) + \frac{1}{\sqrt{n}} (\zeta \circ f^n - \zeta)$$

é o mesmo limite da distribuição de $\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \eta \circ f^j$.

A propriedade (ii) significa, na liguagem da teoria da probabilidade, que $(\eta \circ f^n)_n$ é uma **diferença martingale reversa** para $(\mathcal{B}_n)_{n\geq 0}$. Com efeito, dado qualquer conjunto $A \in \mathcal{B}_{n+1}$, temos $A = f^{-n-1}(A_0)$, onde A_0 pertence a σ-álgebra de Borel de M. Logo,

$$\begin{aligned}\int_A \eta \circ f^n d\mu &= \int \chi_A(\eta \circ f^n) d\mu = \int \chi_{f^{-n-1}(A_0)}(\eta \circ f^n) d\mu \\ &= \int (\chi_{A_0} \circ f^{n+1})(\eta \circ f^n) d\mu = 0,\end{aligned}$$

pois $\chi_{A_0} \circ f^{n+1} \in L^2(\mathcal{B}_{n+1})$ e $E(\eta \circ f^n \mid \mathcal{B}_{n+1}) = 0$ para todo $n \geq 0$. A propriedade (ii) nos permitirá usar o Teorema Central do Limite pra martingales.

Apresentado o esquema geral que vamos seguir, agora começaremos, de fato, a demonstração do Teorema 2.6.2.

Iniciamos introduzindo o operador $U : L^2(\mu) \to L^2(\mu)$ dado por $U\varphi = \varphi \circ f$, e seja $\mathcal{P}$ o seu operador adjunto $\int (\mathcal{P}\psi)\varphi d\mu = \int \psi(U\varphi) d\mu$. Note que

$$\int |U\varphi|^2 d\mu = \int |\varphi|^2 \circ f d\mu = \int |\varphi| d\mu.$$

Segue que U é uma isometria sobre $L^2(\mathcal{B}_1) \subset L^2(\mathcal{B}_0) = L^2(\mu)$, e então $\mathcal{P} : L^2(\mathcal{B}_1) \to L^2(\mathcal{B}_0)$ é também uma isometria. Além disso,

$$U(L^2(\mathcal{B}_n)) = L^2(\mathcal{B}_{n+1}) \;\; e \;\; \mathcal{P}(L^2(\mathcal{B}_{n+1})) = L^2(\mathcal{B}_n),$$

para cada $n \geq 1$. Definamos

$$\zeta = -\sum_{j=1}^{\infty} \mathcal{P}^j(E(\phi \mid \mathcal{B}_j)) \;\; e \;\; \eta = \sum_{j=0}^{\infty} \mathcal{P}^j(E(\phi \mid \mathcal{B}_j) - E(\phi \mid \mathcal{B}_{j+1})).$$

Como $||\mathcal{P}^j(E(\phi \mid \mathcal{B}_j))||_2 = ||E(\phi \mid \mathcal{B}_j)||_2$, a hipótese do teorema garante que a série definindo ζ converge em $L^2(\mu)$. Sendo assim, temos que

$$\left\| \frac{1}{\sqrt{n}}(\zeta \circ f^n - \zeta) \right\|_2 \leq \frac{1}{\sqrt{n}}(||\zeta \circ f^n||_2 + ||\zeta||_2) = \frac{2}{\sqrt{n}}||\zeta||_2 \to 0 \;\; quando \;\; n \to +\infty.$$

Como uma consequência, $\frac{1}{\sqrt{n}}(\zeta \circ f^n - \zeta)$ converge para zero em medida.

Pela Observação 2.6.1 (com $A = L^2(\mathcal{B}_j)$ e $B = L^2(\mathcal{B}_{j+1})$), temos que $E(\phi \mid \mathcal{B}_j) - E(\phi \mid \mathcal{B}_{j+1})$ coincide com a projeção ortogonal de $E(\phi \mid \mathcal{B}_j)$ em $L^2(\mathcal{B}_{j+1})^{\perp}$. Logo,

$$\left\| \mathcal{P}^j(E(\phi \mid \mathcal{B}_j) - E(\phi \mid \mathcal{B}_{j+1})) \right\|_2 = \|E(\phi \mid \mathcal{B}_j) - E(\phi \mid \mathcal{B}_{j+1})\|_2 \leq ||E(\phi \mid \mathcal{B}_j)||_2,$$

e então a série definindo η também converge em $L^2(\mu)$. Observe que η pode ser reescrito da forma

$$\begin{aligned}\eta \;\; = \;\; & E(\phi \mid \mathcal{B}_0) + \mathcal{P}E(\phi \mid \mathcal{B}_1) + \mathcal{P}^2E(\phi \mid \mathcal{B}_2) + ... \\ & - E(\phi \mid \mathcal{B}_1) - \mathcal{P}E(\phi \mid \mathcal{B}_2) - \mathcal{P}^2E(\phi \mid \mathcal{B}_3) - ...\end{aligned}$$

Claramente, $E(\phi \mid \mathcal{B}_0) = \phi$. Além disso, para todo $j \geq 1$,

$$\mathcal{P}^{j-1}(E(\phi \mid \mathcal{B}_j)) = U\mathcal{P}^j(E(\phi \mid \mathcal{B}_j)) = \mathcal{P}^j(E(\phi \mid \mathcal{B}_j)) \circ f.$$

De fato, $E(\phi \mid \mathcal{B}_j) \in L^2(\mathcal{B}_j)$ implica que $\mathcal{P}^{j-1}(E(\phi \mid \mathcal{B}_j)) \in L^2(E(\mathcal{B}_1))$, e $U\mathcal{P} = \mathrm{id} \mid L^2(\mathcal{B}_1)$, pois os operadores adjuntos U e $\mathcal{P}$ são isometrias sobre suas imagens. Portanto, podemos reescrever

$$\begin{aligned} \eta &= \phi + \sum_{j=1}^{\infty} \mathcal{P}^j(E(\phi \mid \mathcal{B}_j)) - \sum_{j=1}^{\infty} \mathcal{P}^j(E(\phi \mid \mathcal{B}_j)) \circ f \\ &= \phi - \zeta + \zeta \circ f. \end{aligned}$$

Dado qualquer $\psi \in L^2(\mathcal{B}_1)$, pelo fato de que $E(\phi \mid \mathcal{B}_j) - E(\phi \mid \mathcal{B}_{j+1}) \in L^2(\mathcal{B}_{j+1})^{\perp}$, temos

$$\int \mathcal{P}^j(E(\phi \mid \mathcal{B}_j) - E(\phi \mid \mathcal{B}_{j+1}))\psi d\mu = \int (E(\phi \mid \mathcal{B}_j) - E(\phi \mid \mathcal{B}_{j+1}))(\psi \circ f^j) d\mu = 0,$$

pois $\psi \in L^2(\mathcal{B}_1) \Rightarrow \psi \circ f^j \in L^2(\mathcal{B}_{j+1})$, para todo $j \geq 0$. Segue que $\mathcal{P}^j(E(\phi \mid \mathcal{B}_j) - E(\phi \mid \mathcal{B}_{j+1})) \in L^2(\mathcal{B}_1)^{\perp}$, para todo $j \geq 0$, e então

$$\begin{aligned} \eta_k &= \sum_{j=0}^{k} \mathcal{P}^j(E(\phi \mid \mathcal{B}_j) - E(\phi \mid \mathcal{B}_{j+1})) \in L^2(\mathcal{B}_1)^{\perp}, \quad \forall k \geq 0 \\ &\Rightarrow \quad \eta = \lim_k \eta_k \in L^2(\mathcal{B}_1)^{\perp} \\ &\Leftrightarrow \quad E(\eta \mid \mathcal{B}_1) = 0. \end{aligned}$$

Agora, vamos obter a propriedade martingale que mencionamos em (ii).

Começamos observando que

$$\begin{aligned} \eta &= \xi + \xi', \quad \textit{onde} \;\; \xi = E(\eta \mid \mathcal{B}_1) \;\; \textit{e} \;\; \xi' = E(\eta \mid \mathcal{B}_1^{\perp}) \\ &\Rightarrow \quad \eta \circ f^n = \xi \circ f^n + \xi' \circ f^n, \quad \forall n \geq 0. \end{aligned}$$

Claramente $\xi \circ f^n \in L^2(\mathcal{B}_{n+1})$. Dado $z \in L^2(\mathcal{B}_{n+1})$, temos

$$\int (\xi' \circ f^n) z d\mu = \int \xi'(\mathcal{P}^n z) d\mu = 0, \quad \textit{pois} \;\; \xi' \in L^2(\mathcal{B}_1)^{\perp} \;\; \textit{e} \;\; \mathcal{P}^n z \in L^2(\mathcal{B}_1), \quad \forall n \geq 0.$$

Assim, concluimos que $\xi' \circ f^n \in L^2(\mathcal{B}_{n+1})^{\perp}$. Portanto,

$$E(\eta \circ f^n \mid \mathcal{B}_{n+1}) = \xi \circ f^n = E(\eta \mid \mathcal{B}_1) \circ f^n = 0, \quad \text{para todo } n \geq 0.$$

Em particular, as variáveis $\eta \circ f^n$ são duas-a-duas ortogonais:

$$\int (\eta \circ f^k)(\eta \circ f^n) d\mu = 0,$$

para todo $k > n \geq 0$, pois $\eta \circ f^k \in L^2(\mathcal{B}_{n+1})$. Isto nos dá

$$||\eta||_2^2 = \frac{1}{n}\sum_{j=0}^{n-1} ||\eta \circ f^j||_2^2 = \frac{1}{n}||\sum_{j=0}^{n-1} \eta \circ f^j||_2^2 = ||\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \eta \circ f^j||_2^2.$$

Logo,

$$||\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \phi \circ f^j - \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \eta \circ f^j||_2 = ||\frac{1}{\sqrt{n}}(\zeta \circ f^n - \zeta)||_2 \to 0.$$

A partir disso, concluímos que

$$\begin{aligned}
||\eta||_2^2 &= \lim_n ||\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} \phi \circ f^j||_2^2 \\
&= \lim_n \frac{1}{n}\left(\sum_{k=0}^{n-1}\int (\phi \circ f^k)^2 d\mu + 2\sum_{0\leq k<l\leq n-1}\int (\phi \circ f^k)(\phi \circ f^l) d\mu\right) \\
&= \lim_n \left(\int \phi^2 d\mu + 2\sum_{j=1}^{n-1}\left(1 - \frac{j}{n}\right)\int \phi(\phi \circ f^j) d\mu\right) \\
&= \int \phi^2 d\mu + 2\sum_{j=1}^{\infty}\int \phi(\phi \circ f^j) d\mu - 2\lim_n \sum_{j=1}^{n-1}\frac{j}{n}\int \phi(\phi \circ f^j) d\mu.
\end{aligned}$$

A última igualdade é válida somente se as séries forem convergentes, vamos demonstrar que isso realmente ocorre. Observe que

$$\begin{aligned}
\int \phi(\phi \circ f^j) d\mu &= \int (E(\phi \mid \mathcal{B}_j) + E(\phi \mid \mathcal{B}_j^{\perp})) \cdot (\phi \circ f^j) d\mu \\
&= \int E(\phi \mid \mathcal{B}_j)(\phi \circ f^j) d\mu \quad (\textit{pois} \quad \phi \circ f^j \in L^2(\mathcal{B}_j)) \\
&\leq ||E(\phi \mid \mathcal{B}_j)||_2 ||\phi \circ f^j||_2 = ||E(\phi \mid \mathcal{B}_j)||_2 ||\phi||_2.
\end{aligned}$$

Segue imediatamente que

$$||\sum_{j=1}^{\infty}\int \phi(\phi \circ f^j) d\mu||_2 \leq ||\phi||_2 \sum_{j=1}^{\infty} ||E(\phi \mid \mathcal{B}_j)||_2 < +\infty.$$

Agora mostraremos que

$$\lim_n \sum_{j=1}^{n-1}\frac{j}{n}\int \phi(\phi \circ f^j) d\mu = 0.$$

Para isso, começaremos observando a limitação

$$||\sum_{j=1}^{n-1}\frac{j}{n}\int \phi(\phi \circ f^j) d\mu||_2 \leq ||\phi||_2 \sum_{j=1}^{n-1}\frac{j}{n}||E(\phi \mid \mathcal{B}_j)||_2.$$

Logo, basta mostrar que esse último somatório tende para zero quando n vai para o infinito.

Dado $\varepsilon > 0$, existe $k \in \mathbb{N}$ tal que $\sum_{j=k+1}^{\infty} ||E(\phi \mid \mathcal{B}_j)||_2 < \frac{\varepsilon}{2}$. Tome $n_0 \in \mathbb{N}$ de modo que $n_0 > \frac{2k}{\varepsilon} \sum_{j=1}^{k} ||E(\phi \mid \mathcal{B}_j)||_2$. Então

$$\begin{aligned} n > n_0 \Rightarrow \sum_{j=1}^{n-1} \frac{j}{n} ||E(\phi \mid \mathcal{B}_j)||_2 &= \sum_{j=1}^{k} \frac{j}{n} ||E(\phi \mid \mathcal{B}_j)||_2 + \sum_{j=k+1}^{n-1} \frac{j}{n} ||E(\phi \mid \mathcal{B}_j)||_2 \\ &\leq \frac{k}{n} \sum_{j=1}^{k} ||E(\phi \mid \mathcal{B}_j)||_2 + \sum_{j=k+1}^{\infty} ||E(\phi \mid \mathcal{B}_j)||_2 \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

concluindo o que queríamos mostrar. Como consequência, temos

$$||\eta||_2^2 = \int \phi^2 d\mu + 2 \sum_{j=1}^{\infty} \int \phi(\phi \circ f^j) d\mu = \sigma^2,$$

relembrando que σ está definido por (2.7). Em particular, temos que $\sigma < +\infty$. Além disso, $\sigma = 0$ implica em $\eta = 0$ e então $\phi = \zeta \circ f - \zeta$. Reciprocamente, se ϕ tem a forma $\phi = u \circ f - u$ então podemos tomar $\zeta = u$ e $\eta = 0$ e assim, pelos argumentos anteriores, $\sigma = ||\eta||_2 = 0$. Isto prova a primeira parte do teorema. A partir de agora suporemos $\sigma > 0$. Sejam $a < b$ fixos. Dado $\delta > 0$, tomemos $\varepsilon > 0$ tal que

$$\begin{aligned} \Phi(a-\varepsilon, b+\varepsilon) &= \Phi(a,b) + \Phi(a-\varepsilon, a) + \Phi(b, b+\varepsilon) \\ &\leq \Phi(a,b) + \delta, \quad \textit{onde} \quad \Phi(r,s) = \frac{1}{\sigma\sqrt{2\pi}} \int_r^s e^{-\frac{t^2}{2\sigma^2}} dt. \end{aligned}$$

Então tomamos $n_1 \geq 1$ tal que

$$\mu \left\{ x \in M : \left| \frac{1}{\sqrt{n}} (\zeta \circ f^n - \zeta)(x) \right| > \varepsilon \right\} \leq \delta \quad \textit{para} \quad \textit{todo} \quad n \geq n_1.$$

Pelo Teorema Central do Limite para martingales (Teorema 2.6.1), existe $n_2 \geq n_1$ tal que

$$\mu \left\{ x \in M : \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \eta(f^j(x)) \in (a-\varepsilon, b+\varepsilon) \right\} \leq \Phi(a-\varepsilon, b+\varepsilon) + \delta,$$

para todo $n \geq n_2$. Observe que se tivermos ao mesmo tempo

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \eta \circ f^j \notin (a-\varepsilon, b+\varepsilon) \quad e \quad \left| \frac{1}{\sqrt{n}} (\zeta \circ f^n - \zeta) \right| < \varepsilon,$$

então $\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \phi \circ f^j = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \eta \circ f^j + \frac{1}{\sqrt{n}} (\zeta \circ f^n - \zeta) \notin (a,b)$. Portanto, $\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \phi \circ f^j \in (a,b)$, implica em

$$\frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \eta \circ f^j \in (a-\varepsilon, b+\varepsilon) \quad ou \quad \left| \frac{1}{\sqrt{n}} (\zeta \circ f^n - \zeta) \right| > \varepsilon,$$

e concluímos que

$$\mu\left\{x\in M:\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\phi(f^j(x))\in(a,b)\right\}\leq\Phi(a-\varepsilon,b+\varepsilon)+2\delta\leq\Phi(a,b)+3\delta,$$

para todo $n\geq n_2$. Isto mostra que

$$\limsup_n\mu\left\{x\in M:\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\phi(f^j(x))\in(a,b)\right\}\leq\Phi(a,b),$$

e um argumento similar nos dá $\liminf_n\geq\Phi(a,b)$. $\square$

APÊNDICE - INTEGRAÇÃO

Nessa seção,introduziremos a teoria de integração em espaços mensuráveis abstratos, definiremos primeiramente a integral para funções mensuráveis simples e depois para funções mensuráveis quaisquer. Os principais resultados serão os Teoremas da Convergência Monótona e da Convergência Dominada, que serão ferramentas básicas para o que se seguirá.

Ao longo desta seção consideraremos um espaço de medida (X, Σ, μ) fixado. Denotaremos a coleção de todas as funções Σ-mensuráveis em X por $M = M(X, \Sigma)$ e a coleção de todas as funções Σ-mensuráveis não negativas por $M^+ = M^+(X, \Sigma)$. Começaremos introduzindo a noção de uma função simples.

Definição A.1: Dizemos que uma função $\varphi : X \to \mathbb{R}$ é **simples** se existem constantes $\alpha_1, ..., \alpha_k \in \mathbb{R}$ e conjuntos mensuráveis $E_1, ..., E_n \in \Sigma$ disjuntos dois-a-dois tais que

$$\varphi = \sum_{j=1}^{n} \alpha_j \chi_{E_j}, \tag{3.8}$$

onde χ_{E_j} é a função característica do conjunto E_j.

Em outras palavras, funções simples são funções mensuráveis que tomam apenas um número finito de valores em $\mathbb{R}$. Observe que cada $E_j = \varphi^{-1}(\alpha_j)$, logo são conjuntos mensuráveis.

Proposição A.1: Se $f \in M^+(X, \Sigma)$, então existe uma sequência crescente de funções simples $(\varphi_n)_n$ em $M^+(X, \Sigma)$ convergindo para f.

Demonstração: Seja n um número natural fixo. Se $k = 0, 1, ..., n2^n - 1$, tome E_{kn} como sendo o conjunto

$$E_{kn} = \{x \in X : k2^{-n} \leq f(x) < (k+1)2^{-n}\},$$

e se $k = n2^n$, tome $E_{kn} = \{x \in X : f(x) \geq n\}$. Observe que os conjuntos $\{E_{kn} : k = 0, 1, ..., n2^n\}$ são disjuntos, pertencem a Σ e tem união igual a X. Se definirmos $\varphi_n = k2^{-n}$ em E_{kn}, então $\varphi_n \in M(X, \Sigma)$ e satisfaz:

a) $0 \leq \varphi_n(x) \leq \varphi_{n+1}(x)$ para todo $x \in X$, $n \in \mathbb{N}$.

b) $f(x) = \lim_n \varphi_n(x)$ para cada $x \in X$.

c) Cada φ_n toma apenas um número finito de valores em $\mathbb{R}$.

Isso conclui a demonstração. $\square$

Definição A.2: Se φ é uma função simples em $M^+(X, \Sigma)$ com a representação padrão (3.8), definimos a **integral** de φ com respeito a μ por

$$\int \varphi d\mu = \sum_{j=1}^{n} \alpha_j \mu(E_j).$$

Estamos usando a convenção de que $0 \cdot (+\infty) = 0$, logo a integral de uma função identicamente nula é igual a 0 independentemente do espaço ter medida finita ou infinita. Observe que o valor da integral de uma função simples em M^+ está bem definido (podendo ser $+\infty$), pois todos os α_j são não-negativos, e assim não ocorrem indeterminações do tipo $(+\infty) - (+\infty)$.

Agora estamos preparados para introduzir a definição de integral para uma função arbitrária em M^+.

Definição A.3: Se $f \in M^+(X, \Sigma)$, definimos a **integral de f com respeito a μ** por

$$\int f d\mu = \sup \int \varphi d\mu,$$

onde o supremo é tomado sobre todas as funções simples $\varphi \in M^+(X, \Sigma)$ satisfazendo $0 \leq \varphi(x) \leq f(x)$ para todo $x \in X$.

Se $f \in M^+$ e $E \in \Sigma$, então $f\chi_E \in M^+$ e definimos a **integral de f sobre E com respeito a μ** por

$$\int_E f d\mu = \int f \chi_E d\mu.$$

Definimos a integral para uma função mensurável positiva. Agora, estenderemos essa definição para uma função mensurável qualquer. Para isso, basta observar que dada uma função $f : X \to [-\infty, +\infty]$ sempre podemos escrever $f = f^+ - f^-$ onde

$$f^+(x) = \max\{f(x), 0\} \text{ e } f^-(x) = \max\{-f(x), 0\}.$$

Note que f^+ e f^- são não negativas e também serão mensuráveis se f for mensurável.

Definição A.4: Dizemos que uma função mensurável $f : X \to [-\infty, +\infty]$ é **integrável com respeito a μ** se ambas as partes positiva e negativa de f^+, f^-, de f tem integral finita em relação a μ. Nesse caso, definimos a **integral de f com respeito a μ** por

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu.$$

Se $E \in \Sigma$, definimos

$$\int_E f d\mu = \int_E f^+ d\mu - \int_E f^- d\mu.$$

Denotaremos por $L = L(X, \Sigma, \mu)$ a coleção de todas as funções integráveis.

Teorema A.2 (Convergência Monótona): Se $(f_n)_n$ é uma sequência monótona crescente de funções em $M^+(X, \Sigma)$ que converge para f, então

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

Demonstração: A função f é mensurável, pois é o limite de uma sequência de funções mensuráveis. Como $f_n \leq f_{n+1} \leq f$, temos

$$\int f_n d\mu \leq \int f_{n+1} d\mu \leq \int f d\mu,$$

para todo $n \in \mathbb{N}$. Portanto,

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

Para estabelecer a desigualdade oposta, seja α um número real satisfazendo $0 < \alpha < 1$ e seja φ uma função mensurável simples satisfazendo $0 \leq \varphi \leq f$. Seja

$$A_n = \{x \in X : f_n(x) \geq \alpha\varphi(x)\},$$

de modo que $A_n \in X$, $A_n \subseteq A_{n+1}$, e $X = \bigcup A_n$. Logo,

$$\int_{A_n} \alpha\varphi d\mu \leq \int_{A_n} f_n d\mu \leq \int f_n d\mu.$$

Como a sequência $(A_n)_n$ é monótona crescente e tem união X, segue que

$$\int \varphi d\mu = \lim_n \int_{A_n} \varphi d\mu.$$

Portanto, tomando o limite na desigualdade anterior, obtemos

$$\alpha \int \varphi d\mu \leq \lim_n \int f_n d\mu.$$

Já que isso ocorre para todo α com $0 < \alpha < 1$, então

$$\int \varphi d\mu \leq \lim_n \int f_n d\mu$$

e como φ é uma função simples arbitrária em M^+ satisfazendo $0 \leq \varphi \leq f$, finalmente concluímos que

$$\int f d\mu = \sup_{\varphi} \int \varphi d\mu \leq \lim_n \int f_n d\mu.$$

Assim, obtemos o resultado. $\square$

Proposição A.3 (Lema de Fatou): Seja $(f_n)_n$ uma sequência de funções em $M^+(X, \Sigma)$, então a função $f : X \to [0, +\infty]$ definida por $f(x) = \liminf_n f_n(x)$ é integrável e vale

$$\int (\liminf_n f_n) d\mu \leq \liminf_n \int f_n d\mu.$$

Demonstração: Sejam $g_m = \inf\{f_m, f_{m+1}, ...\}$ tais que $g_m \leq f_n$ sempre que $m \leq n$. Portanto,

$$\int g_m d\mu \leq \int f_n d\mu, \quad m \leq n,$$

$$\Rightarrow \int g_m d\mu \leq \liminf_n \int f_n d\mu.$$

Como a sequência $(g_m)_m$ é crescente e converge para $\liminf_n f_n$, o Teorema da Convergência Monótona implica que

$$\int (\liminf_n f_n) d\mu = \lim_m \int g_m d\mu \leq \liminf_n \int f_n d\mu. \qquad \square$$

Corolário A.4: Se $f \in M^+$ e λ é definida em X por

$$\lambda(E) = \int_E f d\mu,$$

então λ é uma medida.

Demonstração: Como $f \geq 0$, segue que $\lambda(E) \geq 0$. Se $E = \emptyset$, então $f\chi_E$ é identicamente nula e assim $\lambda(\emptyset) = 0$. Para ver que λ é σ-aditiva, seja $(E_n)_n$ uma sequência de conjuntos disjuntos em Σ cuja união resulta em E, e seja f_n definida por

$$f_n = \sum_{k=1}^{n} \int f\chi_{E_k} d\mu = \sum_{k=1}^{n} \lambda(E_k).$$

Como $(f_n)_n$ é uma sequência crescente em M^+ convergindo para $f\chi_E$, o Teorema da Convergência Monótona implica que

$$\lambda(E) = \int f\chi_E d\mu = \lim_n \int f_n d\mu = \sum_{k=1}^{\infty} \lambda(E_k). \qquad \square$$

Proposição A.5: Uma função mensurável f pertence a L se e somente se $|f|$ pertence a L. Nesse caso,

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

Demonstração: Por definição, $f \in L$ se e somente se $f^+, f^- \in M^+$ e tem integrais finitas. Como $|f|^+ = f^+ + f^-$ e $|f|^- = 0$, a primeira afirmação é imediata. Além disso,

$$\left|\int f d\mu\right| = \left|\int f^+ d\mu - \int f^- d\mu\right|$$

$$\leq \int f^+ d\mu + \int f^- d\mu = \int |f| d\mu. \qquad \square$$

Teorema A.6 (Convergência Dominada): Seja $f_n : X \to \mathbb{R}$ uma sequência de funções mensuráveis que converge em μ-quase todo ponto para uma função f. Se existe uma função integrável g tal que $|f_n(x)| \leq g(x)$ para μ-q.t.p. e todo $n \geq 1$, então f é integrável e

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

Demonstração: Já que $|f_n(x)| \leq g(x)$ para μ-q.t.p. e todo $n \geq 1$, temos também $|f(x)| \leq g(x)$ em μ-q.t.p., segue diretamente do Teorema 2.5 (pois g é integrável) que as funções f_n e f são integráveis para todo $n \geq 1$. Como $g + f_n \geq 0$, podemos aplicar o Lema de Fatou para obter

$$\int g d\mu + \int f d\mu = \int (g + f) d\mu = \int \liminf_n (g + f_n) d\mu$$

$$\leq \liminf_n \int (g + f_n) d\mu = \liminf_n (\int g d\mu + \int f_n d\mu)$$

$$= \int g d\mu + \liminf_n \int f_n d\mu.$$

Portanto, segue que

$$\int f d\mu \leq \liminf_n \int f_n d\mu.$$

Como $g - f_n \geq 0$, outra aplicação do Lema de Fatou implica que

$$\int g d\mu - \int f d\mu = \int (g - f) d\mu = \int \liminf_n (g - f_n) d\mu$$

$$\leq \liminf_n \int (g - f_n) d\mu = \liminf_n (\int g d\mu - \int f_n d\mu)$$

$$= \int g d\mu - \limsup_n \int f_n d\mu,$$

do qual segue que

$$\limsup_n \int f_n d\mu \leq \int f d\mu.$$

Combinando as duas desigualdades obtidas, concluímos que

$$\int f d\mu = \lim_n \int f_n d\mu. \qquad \square$$

Referências

1 Bartle, Robert G., *Elements of Integration and Lebesgue Measure*, John Wiley & Sons, 1995. Nenhuma citação no texto.

2 K. Oliveira, M. Viana. *Fudamentos da Teoria Ergódica*, Colecao Fronteiras da Matematica, SBM, (2014). Citado na página 14.

3 M. Viana. *Stochastic dynamics of deterministic systems*. Brazillian Math. Colloquium 1997, IMPA, <http://w3.impa.br/~viana/out/sdds.pdf>. Nenhuma citação no texto.

4 J. Jacod, P. Protter. *Probability Essentials*, Springer, 2004. Citado na página 15.