



Universidade Estadual de Campinas

Instituto de Matemática, Estatística e Computação Científica

RENATA POSSOBON

Desintegração de medidas: propriedades e aplicações

Campinas
2020

RENATA POSSOBON

Desintegração de medidas: propriedades e aplicações

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em Matemática Aplicada.

ORIENTADOR: CHRISTIAN DA SILVA RODRIGUES

Este trabalho corresponde à versão final da dissertação defendida pela aluna Renata Possobon e orientada pelo Professor Doutor Christian da Silva Rodrigues.

Campinas
2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

P845d Possobon, Renata, 1995-
Desintegração de medidas : propriedades e aplicações / Renata Possobon.
– Campinas, SP : [s.n.], 2020.

Orientador: Christian da Silva Rodrigues.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de
Matemática, Estatística e Computação Científica.

1. Desintegração de medidas. 2. Teorema de Rokhlin. 3. Transporte ótimo.
4. Teoria ergódica. 5. Sistemas dinâmicos. I. Rodrigues, Christian da Silva,
1978-. II. Universidade Estadual de Campinas. Instituto de Matemática,
Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Disintegration of measures : properties and applications

Palavras-chave em inglês:

Disintegration of measures

Rokhlin's theorem

Optimal transportation

Ergodic theory

Dynamical systems

Área de concentração: Matemática Aplicada

Titulação: Mestra em Matemática Aplicada

Banca examinadora:

Christian da Silva Rodrigues [Orientador]

Pedro Jose Catuogno

Ali Tahzibi

Data de defesa: 28-09-2020

Programa de Pós-Graduação: Matemática Aplicada

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0002-9922-763X>

- Currículo Lattes do autor: <http://lattes.cnpq.br/0215263356109906>

**Dissertação de Mestrado defendida em 28 de setembro de 2020 e aprovada
pela banca examinadora composta pelos Profs. Drs.**

Prof(a). Dr(a). CHRISTIAN DA SILVA RODRIGUES

Prof(a). Dr(a). PEDRO JOSE CATUOGNO

Prof(a). Dr(a). ALI TAHZIBI

A Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação do Instituto de Matemática, Estatística e Computação Científica.

Agradecimentos

Aos meus pais, meu irmão e meu namorado, por todo apoio e amor.

Ao Professor Doutor Christian da Silva Rodrigues, por sua dedicação, paciência e por me proporcionar tantas oportunidades de crescimento acadêmico e pessoal.

Ao Professor Doutor Stefano Galatolo, pela solicitude.

Aos professores da Universidade Estadual de Campinas, da Universidade de Pisa e da *Scuola Normale Superiore di Pisa* que contribuíram em minha formação.

Ao Instituto de Matemática, Estatística e Computação Científica, pela oportunidade e suporte.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro nos processos 2018/05309-3 e 2019/14724-7.

Aos grandes amigos que fizeram parte desta jornada, por todo apoio e pelos ótimos momentos juntos. Em especial, agradeço à Cristiane Brunhara, Rita Fiengo e às grandes mulheres do *Istituto Suore Figlie di Nazareth*, que me acolheram tão bem.

Resumo

Neste trabalho estuda-se o conceito de desintegração de medidas e o Teorema de Rokhlin em diferentes contextos. Em particular, discute-se a relação de desintegração de medidas com o problema central da Teoria de Transporte Ótimo (o problema dos planos de transporte) e aborda-se uma conexão desta perspectiva de desintegração com o caso em espaços particionados. Além disso, estuda-se algumas propriedades geométricas e estatísticas de sistemas dinâmicos através do operador de transferência associado e, neste contexto, aplica-se a desintegração de medidas para definir espaços apropriados ao sistema.

Palavras-chave: desintegração de medidas, Teorema de Rokhlin, Transporte Ótimo, Teoria Ergódica, sistemas dinâmicos.

Abstract

In this work, we study the disintegration of measures and Rokhlin's Theorem in different contexts. In particular, we argue the relationship between the disintegration of measures and the central problem of the Optimal Transport Theory (optimal transport problem). Also, we show a connection of this perspective of disintegration with the case of partitioned spaces. In addition, we study some geometric and statistical properties of dynamical systems by the associated transfer operator, and, within this context, we apply the disintegration of measures to define suitable spaces to the system.

Keywords: disintegration of measures, Rokhlin's Theorem, Optimal Transport, Ergodic Theory, dynamical systems.

Lista de Símbolos

$\mathcal{B}(X)$	σ -álgebra de Borel em X .
$\#A$	Cardinalidade do conjunto A .
$\mathcal{C}(X)$	Classe de funções reais contínuas definidas em X .
C_c^+	Conjunto das funções contínuas de X a $\mathbb{R}^+$ com suporte compacto.
$Lip_1(X)$	Conjunto das funções Lipschitz em X com constante Lipschitz ≤ 1 .
$BM(X)$	Conjunto das funções reais mensuráveis limitadas definidas em X .
$\mathcal{SB}(X)$	Conjunto das medidas borelianas com sinal em X .
$\Pi(\mu, \nu)$	Conjunto dos planos de transporte de μ para ν .
$\coprod_{n \in \mathbb{N}} A_n$	Co-produto dos conjuntos A_n .
$W_p(\mu, \nu)$	Distância de Wasserstein de ordem p entre μ e ν .
$C_b^0(X)$	Espaço das funções reais contínuas e limitadas definidas em X .
$\{0, 1\}^{\mathbb{N}}$	Espaço de Cantor.
$\mathcal{P}_p(X)$	Espaço de Wasserstein de ordem p .
$T_p M$	Espaço tangente à M no ponto p .
$\mathcal{L}^1(\mu)$	Família das funções reais μ -integráveis.
$\mathcal{P}(X)$	Família das probabilidades borelianas em X .
$\mathbb{1}_A$	Função indicadora do conjunto A .
δ_x	Medida de Dirac no ponto x .
$\mu \times \nu$	Medida produto.
$\ \cdot\ _p$	Norma no espaço L^p .
$f _A$	Restrição da função f ao conjunto A .
$\overline{\mathbb{R}}$	Reta estendida; $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$.
$\text{supp}(\mu)$	Suporte da medida μ .

Sumário

Introdução	10
1 Preliminares	12
1.1 Teoria da Medida	12
1.2 Transporte Ótimo	19
1.3 Variedades	21
2 Desintegração de medidas	26
2.1 Uma definição formal de desintegração	26
2.2 Teorema de Rokhlin	29
3 Desintegração em diferentes contextos	51
3.1 Desintegração e Transporte Ótimo	52
3.2 Desintegração e variedades: do Teorema 2.3 ao Corolário 3.2	58
3.2.1 Sobre desintegração de medidas absolutamente contínuas	59
3.2.2 Restrição e medidas com sinal	61
4 Algumas propriedades estatísticas de sistemas dinâmicos	65
4.1 Explorando exemplos	68
4.1.1 Transformações expansoras	68
4.1.2 Mapa solenoidal	76
Referências	81

Introdução

No final do século XIX, algumas questões fundamentais relacionadas à evolução e estabilidade do sistema solar levaram ao desenvolvimento da Teoria de Sistemas Dinâmicos que, de modo geral, compreende o estudo do comportamento de sistemas ao longo do tempo, e hoje é um vasto e rico campo que possui aplicações em física, biologia, meteorologia, astronomia, economia e outras áreas [8].

Heuristicamente, um sistema dinâmico consiste de um conjunto, que abrange todos os possíveis estados para o sistema em estudo, e uma função que descreve a dinâmica. Um pouco mais especificamente, um *sistema dinâmico de tempo discreto* consiste em um espaço X e uma aplicação $f : X \rightarrow X$ tal que, para cada $n \in \mathbb{N}$, a n -ésima iteração de f é $f^n = f \circ f \circ \dots \circ f$ (n vezes) e $f^0 = id$, e um *sistema dinâmico de tempo contínuo* consiste em um espaço X e uma família de transformações $\{f^t : X \rightarrow X\}_{t \in \mathbb{R}}$, denominada fluxo, satisfazendo $f^0 = id$ e $f^{t+s} = f^t \circ f^s$, $\forall t, s \in \mathbb{R}$.

Existem modelos de sistemas dinâmicos que são determinísticos, no sentido que, dado um estado x do sistema, tomado em qualquer momento, é possível determinar toda a trajetória futura (e, quando o sistema é invertível, também a trajetória passada). Ao mesmo tempo, um sistema pode exibir dependência sensível das condições iniciais, de modo que pequenas alterações no estado inicial levam a comportamentos de longo prazo dramaticamente diferentes, o que é considerado um comportamento caótico [8]. Em sistemas caóticos, embora a trajetória possa ser imprevisível pontualmente, as propriedades estatísticas podem exibir certa regularidade, fornecendo informações relevantes sobre o comportamento do sistema. O estudo das propriedades estatísticas de sistemas dinâmicos é uma abordagem da dinâmica que foi implementada na chamada Teoria Ergódica.

A desintegração de medidas é um conceito que exerce um importante papel no estudo de Sistemas Dinâmicos e Teoria Ergódica. Em termos de medidas condicionais, dado um espaço de probabilidade $(X, \mathcal{F}, \mu)$ e uma partição $\mathcal{P}$ de X , uma desintegração de μ relativamente a $\mathcal{P}$ é uma família de probabilidades $\{\mu_P : P \in \mathcal{P}\}$ tal que cada μ_P tem suporte no respectivo elemento indexado e é possível reobter μ a partir da integração desta família. Neste trabalho, estuda-se e aplica-se este conceito em diferentes contextos, juntamente com resultados relevantes, como o Teorema de Rokhlin. Mais detalhadamente:

- No Capítulo 1 apresenta-se uma síntese de conceitos e resultados que embasam o estudo.

- No Capítulo 2 aborda-se uma definição formal de desintegração de medidas e estuda-se, sob diferentes perspectivas, o Teorema de Rokhlin, resultado que afirma a existência da desintegração de medidas sob certas condições. Especificamente, são apresentadas três versões deste teorema: uma formulação para medidas de probabilidade definidas em um espaço métrico completo separável particionado, uma formulação um pouco mais moderna, que enfatiza alguns aspectos geométricos dos espaços envolvidos, e, por fim, uma generalização para uma categoria de espaços σ -finitos com morfismos absolutamente contínuos.
- No Capítulo 3 aborda-se a desintegração de medidas considerando espaços produto. Na Seção 3.1, estuda-se como a desintegração de medidas nesta perspectiva está relacionada com o problema central da Teoria de Transporte Ótimo. Na Seção 3.2, considerando variedades riemannianas compactas, estuda-se uma conexão desta perspectiva de desintegração com o caso em espaços particionados.
- Por fim, no Capítulo 4, estuda-se algumas propriedades geométricas e estatísticas de sistemas dinâmicos, como a existência de medidas invariantes, mistura, convergência para o equilíbrio e estabilidade. Mais especificamente, estuda-se o comportamento do operador de transferência em dois exemplos de sistemas dinâmicos, sendo que um deles é um mapa solenoidal, onde o conceito de desintegração de medidas é aplicado para definir um espaço adequado para o sistema.

CAPÍTULO 1

Preliminares

Este capítulo traz algumas definições e resultados que são empregados ao longo do trabalho. A intenção é criar uma fonte de referência sobre os principais conceitos utilizados, a fim de favorecer uma boa leitura. Em geral, a demonstração dos resultados apresentados neste capítulo são omitidas, mas, ao longo do texto, são indicadas as referências para a consulta das respectivas demonstrações.

1.1 Teoria da Medida

A seguir são apresentados alguns fundamentos e resultados de Teoria da Medida, que são abordados nas referências [5], [10], [13], [21], [23] e [30]. Considere um conjunto X não vazio. Uma álgebra de X é uma família $\mathcal{A}$ de subconjuntos de X que contém o conjunto vazio e tal que, para cada $A \in \mathcal{A}$, A^c , o complementar do conjunto A em relação a X , também pertence à $\mathcal{A}$, e, dados dois conjuntos $A, B \in \mathcal{A}$, a união $A \cup B$ pertence à $\mathcal{A}$. Uma σ -álgebra é uma álgebra que é fechada para as uniões enumeráveis, isto é:

Definição 1.1. *Seja X um conjunto não vazio. Uma σ -álgebra em X é uma família $\mathcal{F}$ de subconjuntos de X tal que:*

1. $\emptyset, X \in \mathcal{F}$.
2. Se $A \in \mathcal{F}$, então $A^c \in \mathcal{F}$.
3. Se $A_n \in \mathcal{F}$ para todo $n \in \mathbb{N}$, então $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

O par $(X, \mathcal{F})$ é denominado espaço mensurável, e um conjunto qualquer $A \in \mathcal{F}$ é denominado conjunto mensurável. Dada uma família $\mathcal{E}$ de subconjuntos de X , denomina-se σ -álgebra gerada por $\mathcal{E}$, $\sigma(\mathcal{E})$, a menor σ -álgebra que contém esta família. Além disso, se X é um espaço topológico, a σ -álgebra gerada pelos abertos de X é denominada σ -álgebra de Borel, e denotada por $\mathcal{B}(X)$. Neste caso, os conjuntos $A \in \mathcal{B}(X)$ são denominados borelianos.

A partir do conceito de espaço mensurável, é possível estabelecer a ideia de função mensurável, que é uma aplicação que preserva a família de conjuntos mensuráveis:

Definição 1.2. Considere $(X, \mathcal{F})$ um espaço mensurável. Uma função $f : X \rightarrow \overline{\mathbb{R}}$ é mensurável se $f^{-1}(A) \in \mathcal{F}$, para todo $A \in \mathcal{B}(\overline{\mathbb{R}})$. Mais geralmente, se $(X, \mathcal{F})$ e $(Y, \mathcal{A})$ são espaços mensuráveis, uma função $f : X \rightarrow Y$ é mensurável se, para todo $A \in \mathcal{A}$, $f^{-1}(A) = \{x \in X : f(x) \in A\} \in \mathcal{F}$.

Se $f, g : X \rightarrow \overline{\mathbb{R}}$ são funções mensuráveis e $\alpha \in \mathbb{R}$ é uma constante, é possível demonstrar que αf , $f + g$, fg e $|f|$ são, também, funções mensuráveis. Além disso, se $(f_n)_{n \in \mathbb{N}}$ é uma sequência de funções mensuráveis tal que $f_n : X \rightarrow \overline{\mathbb{R}}$, então $\inf f_n(x)$, $\sup f_n(x)$, $\liminf f_n(x)$ e $\limsup f_n(x)$ são mensuráveis e, conseqüentemente, se o limite da sequência existe, ele é mensurável. Dada uma função $f : X \rightarrow \mathbb{R}$, é conveniente definir

$$f^+(x) := \max\{f(x), 0\}$$

$$f^-(x) := \max\{-f(x), 0\}$$

e escrever $f(x) = f^+(x) - f^-(x)$, pois isto permite generalizar resultados válidos, inicialmente, apenas para funções mensuráveis não negativas. Note que f é mensurável se, e somente se, f^+ e f^- são mensuráveis. Uma função $\varphi : X \rightarrow \overline{\mathbb{R}}$ mensurável que assume apenas um conjunto finito de valores é dita função simples. Ou seja, se φ é uma função simples, é possível escrevê-la da forma

$$\varphi(x) = \sum_{i=1}^n a_i \mathbb{1}_{A_i}(x)$$

onde $a_1, \dots, a_n$ são os valores de φ e $A_i = \{x \in X : \varphi(x) = a_i\}$, para $i = 1, 2, \dots, n$. Toda função simples não negativa possui uma representação única, denominada representação *standard*, caracterizada por valores a_i distintos e conjuntos A_i disjuntos. Um resultado importante associado à esta definição é o seguinte:

Lema 1.3. Se $f : X \rightarrow [0, \infty]$ é mensurável, então existe uma sequência (φ_n) de funções simples $\varphi_n : X \rightarrow [0, \infty]$ tais que:

1. $0 \leq \varphi_n(x) \leq \varphi_{n+1}(x)$ para cada $x \in X$ e para todo $n \in \mathbb{N}$.
2. $f(x) = \lim_{n \rightarrow \infty} \varphi_n(x)$ para cada $x \in X$.

Demonstração. Veja [5, Lema 2.11].

□

Outro conceito fundamental relacionado à espaços mensuráveis é o de medida:

Definição 1.4. *Uma medida em um espaço mensurável $(X, \mathcal{F})$ é uma função $\mu : \mathcal{F} \rightarrow [0, \infty]$ tal que $\mu(\emptyset) = 0$ e μ é σ -aditiva, isto é, se $\{A_n : n \in \mathbb{N}\}$ é uma família de conjuntos mensuráveis tal que $A_n \cap A_m = \emptyset$ para $n \neq m$, então*

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n).$$

A tripla $(X, \mathcal{F}, \mu)$ é denominada espaço de medida. Dependendo das características de uma medida μ definida em $(X, \mathcal{F})$, utiliza-se certas denominações especiais. Por exemplo, μ é dita: finita, se $\mu(X) < \infty$; probabilidade, se $\mu(X) = 1$; σ -finita, se existem conjuntos $A_n \in \mathcal{F}$ tais que $\mu(A_n) < \infty$ para todo n e $X = \bigcup A_n$; boreliana, se $\mathcal{F} = \mathcal{B}(X)$; se ν é outra medida em $(X, \mathcal{F})$, diz-se que μ absolutamente contínua com respeito a ν , e denota-se $\mu \ll \nu$, se para todo conjunto $A \in \mathcal{F}$ tal que $\nu(A) = 0$ tem-se que $\mu(A) = 0$; se μ é boreliana e para todo $B \in \mathcal{B}(X)$ e todo $\epsilon > 0$ existe um conjunto fechado F e um aberto A tais que $F \subset B \subset A$ e $\mu(A - F) \leq \epsilon$, diz-se que μ é regular.

Observação. *Por definição, uma medida tem valor positivo. Porém, em alguns casos é útil generalizar essa definição, no seguinte sentido: em um espaço mensurável $(X, \mathcal{F})$, uma função $\mu : \mathcal{F} \rightarrow (-\infty, \infty]$ ou $\mu : \mathcal{F} \rightarrow [-\infty, \infty)$ que é σ -aditiva e tal que $\mu(\emptyset) = 0$ é denominada medida com sinal. Se μ assume valores no intervalo aberto $(-\infty, \infty)$, é denominada medida com sinal finita. O Teorema da Decomposição de Jordan é um resultado no qual, dada uma medida com sinal μ , tem-se que existem únicas medidas μ^+ e μ^- tais que $\mu = \mu^+ - \mu^-$ e $\mu^+ \perp \mu^-$, onde $\mu^+ \perp \mu^-$ representa que as medidas μ^+ e μ^- são mutuamente singulares, isto é: existem dois conjuntos mensuráveis, P e N , tais que $X = P \cup N$, $P \cap N = \emptyset$, $\mu^+(N) = 0$ e $\mu^-(P) = 0$. Com este teorema é possível generalizar os resultados obtidos para medidas. Além disso, tem-se que a soma $|\mu| = \mu^+ + \mu^-$ é uma medida, denominada variação total, e é possível demonstrar que a função $\|\mu\| = |\mu|(X)$ é uma norma completa no espaço vetorial real constituído pelo conjunto das medidas com sinal finitas. Ainda mais, se X é um espaço métrico compacto, tem-se que este espaço é isomorfo ao dual do espaço $\mathcal{C}(X)$ das funções reais contínuas em X .*

Dado um espaço de medida $(X, \mathcal{F}, \mu)$, um subconjunto $A \subset X$ é um conjunto de medida nula se existir $B \in \mathcal{F}$ tal que $A \subset B$ e $\mu(B) = 0$. Diz-se que uma propriedade é válida em μ -quase todo ponto (ou μ -quase sempre) se esta é válida em todo X exceto, possivelmente, em um conjunto de medida nula. Pode ser conveniente que todos os conjuntos de medida nula sejam, também, mensuráveis. Um espaço de medida com esta propriedade é denominado

completo (ou diz-se que a medida μ é completa). É possível realizar o completamento de um espaço, i.e., tornar $(X, \mathcal{F}, \mu)$ qualquer em um espaço completo. Para isso, considere a família de conjuntos $\tilde{\mathcal{F}} = \{A \subset X : \exists F_1, F_2 \in \mathcal{F} \text{ com } F_1 \subset A \subset F_2 \text{ e } \mu(F_2 - F_1) = 0\}$. Note que $\tilde{\mathcal{F}}$ é uma σ -álgebra que contém $\mathcal{F}$. Com isso, a função $\tilde{\mu} : \tilde{\mathcal{F}} \rightarrow [0, \infty]$ tal que $\tilde{\mu}(A) = \mu(F_1) = \mu(F_2)$ é uma medida em $\tilde{\mathcal{F}}$, cuja restrição a $\mathcal{F}$ coincide com μ . Por construção, o espaço de medida $(X, \tilde{\mathcal{F}}, \tilde{\mu})$ é completo.

Estabelecidos estes conceitos, é possível definir a integral de uma função com respeito a uma medida: No espaço de medida $(X, \mathcal{F}, \mu)$, a integral de $\varphi : X \rightarrow [0, \infty]$, uma função simples com representação *standard*, com respeito a μ é definida por

$$\int_X \varphi \, d\mu = \sum_{i=1}^n a_i \mu(A_i).$$

Para uma função mensurável $f : X \rightarrow [0, \infty]$, define-se a integral de f com respeito a μ por

$$\int_X f \, d\mu = \sup \left\{ \int_X \varphi \, d\mu : 0 \leq \varphi \leq f \text{ e } \varphi \text{ simples} \right\}.$$

Além disso, diz-se que uma função $f : X \rightarrow \overline{\mathbb{R}}$ é integrável se é mensurável e f^+ e f^- possuem integrais finitas. Neste caso, define-se:

$$\int_X f \, d\mu = \int_X f^+ \, d\mu - \int_X f^- \, d\mu.$$

Observação. O conjunto $\mathcal{L}^1(\mu)$ das funções reais integráveis é um espaço vetorial real e a aplicação $L : \mathcal{L}^1(\mu) \rightarrow \mathbb{R}$ que associa cada $f \in \mathcal{L}^1(\mu)$ à integral $\int f \, d\mu$ é um funcional linear positivo.

A seguir, são apresentados alguns resultados importantes sobre o comportamento da integral relativamente à passagem ao limite de sequências.

Teorema 1.5 (Teorema da Convergência Monótona). *Considere um espaço de medida $(X, \mathcal{F}, \mu)$ e uma sequência de funções mensuráveis $(f_n)_{n=1}^\infty$ tal que $f_n : X \rightarrow [0, \infty]$ para todo $n \in \mathbb{N}$, $f_1 \leq f_2 \leq \dots \leq f_n \leq f_{n+1} \leq \dots$ e $\lim f_n = f$. Então, f é mensurável e*

$$\int_X f \, d\mu = \lim_{n \rightarrow \infty} \int_X f_n \, d\mu.$$

Demonstração. Veja [5, Teorema 4.6].

□

Lema 1.6 (Lema de Fatou). *Considere um espaço de medida $(X, \mathcal{F}, \mu)$ e uma sequência de funções mensuráveis $(f_n)_{n=1}^{\infty}$ tal que $f_n : X \rightarrow [0, \infty]$ para todo $n \in \mathbb{N}$. Então*

$$\int_X \liminf f_n d\mu \leq \liminf \int_X f_n d\mu.$$

Demonstração. Veja [5, Lema 4.8]. □

Teorema 1.7 (Teorema da Convergência Dominada). *Considere um espaço de medida $(X, \mathcal{F}, \mu)$ e uma sequência de funções integráveis $(f_n)_{n=1}^{\infty}$ que converge μ -quase sempre para a função mensurável $f : X \rightarrow \mathbb{R}$. Se existe uma função integrável g tal que $|f_n| \leq g$ para todo $n \in \mathbb{N}$, então f é integrável e*

$$\int_X f d\mu = \lim_{n \rightarrow \infty} \int_X f_n d\mu.$$

Demonstração. Veja [5, Teorema 5.6]. □

Outro resultado importante que será utilizado neste trabalho é o Teorema de Radon-Nikodym, que estabelece uma relação entre duas medidas σ -finitas absolutamente contínuas definidas em um mesmo espaço mensurável:

Teorema 1.8 (Teorema de Radon-Nikodym). *Sejam ν e μ duas medidas σ -finitas em $(X, \mathcal{F})$ e suponha que $\nu \ll \mu$. Então, existe uma função $f : X \rightarrow [0, \infty]$ tal que*

$$\nu(E) = \int_E f d\mu$$

para $E \in \mathcal{F}$. A função f é denominada densidade ou derivada de Radon-Nikodym de ν com respeito a μ e é denotada por $\frac{d\nu}{d\mu}$. Ainda mais, a função f é μ -quase sempre unicamente determinada.

Demonstração. Veja [5, Teorema 8.9]. □

Agora que os conceitos fundamentais de Teoria da Medida foram apresentados, pode-se explorar alguns conceitos relacionados à medidas de transporte, também denominadas medidas *push forward*. Esse tipo de medida é obtido transferindo (“*pushing forward*”) uma medida de um espaço mensurável para outro usando uma função mensurável. Especificamente, considere X_1 e X_2 espaços métricos separáveis e $\mathcal{P}(X_1)$ e $\mathcal{P}(X_2)$ as famílias de probabilidades

borelianas em X_1 e X_2 , respectivamente. Se $\mu \in \mathcal{P}(X_1)$ e $T : X_1 \rightarrow X_2$ é uma função mensurável, denota-se por $T_*\mu \in \mathcal{P}(X_2)$ o *push forward* de μ por T , definido por

$$T_*\mu(B) := \mu(T^{-1}(B))$$

para todo $B \in \mathcal{B}(X_2)$. Mais geralmente, tem-se

$$\int_{X_1} f(T(x)) d\mu(x) = \int_{X_2} f(y) dT_*\mu(y)$$

para toda função $f : X_2 \rightarrow \mathbb{R}$ boreliana limitada (ou $T_*\mu$ -integrável). Note que se $\nu \ll \mu$, então $T_*\nu \ll T_*\mu$, para todo $\nu, \mu \in \mathcal{P}(X_1)$. Ainda mais, $(T \circ S)_*\mu = T_*(S_*\mu)$ para $S : X_1 \rightarrow X_2$, $T : X_2 \rightarrow X_3$ e $\mu \in \mathcal{P}(X_1)$.

A seguir são apresentadas algumas definições correlatas, cuja principal referência é [30], que serão úteis ao longo do trabalho.

Definição 1.9. *Seja X um espaço métrico separável, $\mathcal{B}(X)$ a σ -álgebra de Borel em X e $\mathcal{P}(X)$ a família de todas as medidas de probabilidade borelianas em X . O suporte de $\mu \in \mathcal{P}(X)$ é o conjunto $\text{supp } \mu = \{x \in X : \mu(\mathcal{U}) > 0 \text{ para cada vizinhança } \mathcal{U} \text{ de } x\}$.*

Definição 1.10. *Um espaço X é denominado Radon se é um espaço métrico separável tal que para toda $\mu \in \mathcal{P}(X)$, para todo $B \in \mathcal{B}(X)$, existe um conjunto compacto $K_\epsilon \subset B$ tal que $\mu(B - K_\epsilon) \leq \epsilon$, para todo $\epsilon > 0$.*

Definição 1.11. *Considere $(X, \mathcal{F}, \mu)$ um espaço de probabilidade e $(E, \mathcal{B}(E))$ um espaço mensurável. Toda função mensurável $\phi : (X, \mathcal{F}, \mu) \rightarrow (E, \mathcal{B}(E))$ induz uma probabilidade em E dada por*

$$\mu_\phi(B) := \phi_*\mu(B) = \mu(\phi^{-1}(B)) = \mu(\{x \in X : \phi(x) \in B\}),$$

para $B \in \mathcal{B}(E)$. Neste caso, μ_ϕ é denominada lei (ou distribuição) de ϕ e, também pode ser denotada por $\text{lei}(\phi)$.

Definição 1.12. *Considere (X, μ) e (Y, ν) espaços de probabilidade. Acoplar μ e ν consiste em construir duas funções mensuráveis (que também são denominadas variáveis aleatórias) $\mathcal{X}$ e $\mathcal{Y}$ em um espaço de probabilidade $(\Omega, \mathbb{P})$ de tal forma que $\text{lei}(\mathcal{X}) = \mu$ e $\text{lei}(\mathcal{Y}) = \nu$. O par $(\mathcal{X}, \mathcal{Y})$ é denominado coupling de (μ, ν) .*

Se μ e ν são as únicas leis do problema, então, sem perda de generalidade, pode-se escolher $\Omega = X \times Y$. Em uma formulação mais teórica, acoplar μ e ν significa construir uma medida γ em $X \times Y$ tal que γ admita μ e ν como marginais em X e Y , respectivamente. As seguintes afirmações são formas equivalentes desta condição marginal:

1. $(\text{proj}_X)_*\gamma = \mu$, $(\text{proj}_Y)_*\gamma = \nu$, onde (proj_X) e (proj_Y) representam, respectivamente, as projeções $(x, y) \mapsto x$ e $(x, y) \mapsto y$;
2. $\gamma[A \times Y] = \mu[A]$, $\gamma[X \times B] = \nu[B]$, $\forall A \subset X$, $B \subset Y$ mensuráveis;
3. Para todas funções integráveis φ , ψ definidas em X e Y , respectivamente, tem-se

$$\int_{X \times Y} (\varphi(x) + \psi(y)) d\gamma(x, y) = \int_X \varphi d\mu + \int_Y \psi d\nu.$$

Definição 1.13. Considerando a notação utilizada na Definição 1.12, um coupling $(\mathcal{X}, \mathcal{Y})$ é dito determinístico se existe uma função mensurável $T : X \rightarrow Y$ tal que $\mathcal{Y} = T(\mathcal{X})$. Além disso, para dizer que $(\mathcal{X}, \mathcal{Y})$ é um coupling determinístico de μ e ν , as seguintes afirmações são estritamente equivalentes:

1. $(\mathcal{X}, \mathcal{Y})$ é um coupling determinístico de μ e ν cuja lei γ está concentrada no gráfico de uma função mensurável $T : X \rightarrow Y$;
2. $\gamma = (\text{Id}, T)_*\mu$;
3. $\mathcal{X}$ tem lei μ e $\mathcal{Y} = T(\mathcal{X})$, onde $T_*\mu = \nu$;
4. $\mathcal{X}$ tem lei μ e $\mathcal{Y} = T(\mathcal{X})$, onde T é a mudança de variável de μ para ν , ou seja, para toda função ν -integrável φ tem-se

$$\int_Y \varphi(y) d\nu(y) = \int_X \varphi(T(x)) d\mu(x).$$

A função T nas afirmações acima é a mesma e é definida unicamente μ -quase sempre (quando a lei conjunta de $(\mathcal{X}, \mathcal{Y})$ está fixa). A implicação inversa é verdadeira: se T e $\tilde{T}$ coincidem μ -quase sempre, então $T_*\mu = \tilde{T}_*\mu$. É comum denominar T como mapa de transporte. Com isso, tem-se o ponto de partida para o estudo em Transporte Ótimo, que é o assunto da próxima seção.

1.2 Transporte Ótimo

Os primeiros estudos relacionados à Teoria de Transporte Ótimo remontam o final do século XVIII, quando Gaspard Monge, matemático francês, propôs a modelagem de um problema logístico. Informalmente, este problema consiste em transportar certa quantidade de areia de um ponto a outro já pré-estabelecido, de forma que o custo de transporte seja mínimo [29]. Formalmente:

Problema de Monge. *Considere X e Y espaços Radon, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$ e uma função boreliana $c : X \times Y \rightarrow [0, \infty]$. O problema de transporte ótimo de Monge é encontrar*

$$\inf_T \left\{ \int_X c(x, T(x)) d\mu : T_*\mu = \nu \right\} \quad (1.1)$$

onde T é um mapa de transporte.

Nos termos anteriormente estabelecidos, Monge buscava estabelecer um *coupling* determinístico ótimo. Ainda mais, note que, nesta formulação formal, as medidas μ e ν modelam a quantidade de areia em cada espaço, no sentido que, dados $A \subset X$ e $B \subset Y$ conjuntos mensuráveis, $\mu(A)$ representa a quantidade areia que está em A e $\nu(B)$ representa a quantidade de areia que deve ser transportada para B . Além disso, $c(x, y)$ representa o custo de transportar uma unidade de massa do ponto $x \in X$ para o ponto $y \in Y$, e o mapa de transporte $T : X \rightarrow Y$ associa, unicamente, cada x à localização y , preservando volume.

Monge levou em consideração que o custo de transporte depende da massa de areia a ser transportada e da distância percorrida no transporte. Estudou o problema em três dimensões para uma distribuição contínua de massa, obtendo avanços, mas não uma conclusão. O problema foi indiretamente retomado pelo matemático russo Leonid Vitaliyevich Kantorovich. Em 1948, Kantorovich reapresentou o problema, incorporando novos resultados desenvolvidos por ele:

Problema de Monge-Kantorovich. *Considere X e Y espaços Radon, medidas $\mu \in \mathcal{P}(X)$ e $\nu \in \mathcal{P}(Y)$ e uma função boreliana $c : X \times Y \rightarrow [0, \infty]$. O problema de Monge-Kantorovich consiste em encontrar*

$$\min_{\gamma \in \Pi(\mu, \nu)} \left\{ \int_{X \times Y} c(x, y) d\gamma(x, y) \right\} \quad (1.2)$$

onde $\Pi(\mu, \nu)$ é o conjunto de todas as medidas $\gamma \in \mathcal{P}(X \times Y)$ com marginais μ e ν .

Em geral, as medidas $\gamma \in \mathcal{P}(X \times Y)$ com marginais μ e ν são denominadas planos de transporte de μ para ν e as medidas que satisfazem o Problema de Monge-Kantorovich são denominadas planos de transporte ótimo. Informalmente, nesta nova formulação tem-se que a quantidade de areia em dado ponto x pode ser dividida e transportada para pontos distintos, e $d\gamma(x, y)$ representa quantidade de areia que é transferida de x para y .

Note que podem ocorrer casos em que não existe uma função $T : X \rightarrow Y$ mensurável tal que $T_*\mu = \nu$, de modo que o Problema de Monge é insolúvel, independentemente da função custo. Isso não ocorre para o Problema de Monge-Kantorovich, uma vez que $\mu \times \nu \in \Pi(\mu, \nu)$. Além disso, quando a função custo é semicontínua inferiormente, é possível provar a existência de um plano de transporte ótimo [3] e, ainda mais, tem-se o seguinte resultado:

Teorema 1.14 (Dualidade). *Se $c : X \times Y \rightarrow [0, \infty]$ é uma função própria semicontínua inferiormente, então o mínimo da equação (1.2) é igual a*

$$\sup \left\{ \int_X \varphi(x) d\mu + \int_Y \psi(y) d\nu \right\}$$

onde o supremo é tomado sob todos $(\varphi, \psi) \in C_b^0(X) \times C_b^0(Y)$ tais que $\varphi(x) + \psi(y) \leq c(x, y)$.

Demonstração. Veja [3, Teorema 6.1.1]. □

A partir desses problemas floresceu a Teoria de Transporte Ótimo, que hoje provê ferramentas para o tratamento de problemas em diversas áreas como Probabilidade e Sistemas Dinâmicos. Algumas referências relevantes sobre Transporte Ótimo são [2], [3], [29] e [30].

A seguir define-se um conceito, muito importante na Teoria de Transporte Ótimo, que será utilizado neste trabalho. No estudo de problemas de transporte pode ser interessante considerar que o custo depende das medidas transportadas, e pensar no custo ótimo como um tipo de distância entre μ e ν :

Definição 1.15 (Distância de Wasserstein). *Sejam (X, d) um espaço métrico completo separável e $p \in [1, \infty)$. Para duas medidas de probabilidade μ, ν em X , a distância de Wasserstein de ordem p entre μ e ν é dada por*

$$W_p(\mu, \nu) := \left(\inf_{\gamma \in \Pi(\mu, \nu)} \int d(x_1, x_2)^p d\gamma(x_1, x_2) \right)^{\frac{1}{p}}.$$

Com essa definição, W_p não é uma distância no sentido estrito, pois pode assumir o valor $+\infty$. Ademais, satisfaz os axiomas de distância. Assim, para concluir a construção, é natural restringir W_p à um subconjunto no qual esta toma valores finitos.

Definição 1.16. *O espaço de Wasserstein de ordem p é definido como*

$$\mathcal{P}_p(X) := \left\{ \mu \in \mathcal{P}(X) : \int d(x, \tilde{x})^p \mu(dx) < +\infty \right\}$$

para algum $\tilde{x} \in X$ arbitrário.

Deste modo, W_p define uma distância (finita) em $\mathcal{P}_p(X)$. É possível demonstrar que, se (X, d) é um espaço métrico completo separável e d é uma métrica limitada, então a distância p -Wasserstein metriza a topologia fraca sobre $\mathcal{P}(X)$. Ainda mais, se X é compacto, então $(\mathcal{P}(X), W_p)$ é um espaço métrico compacto. Para detalhes e mais informações, consulte [30, Capítulo 6].

1.3 Variedades

Nesta seção são apresentadas definições importantes relacionadas ao conceito de variedade. A principal referência utilizada é [9]. Em termos gerais, uma variedade de dimensão n é um espaço topológico que se assemelha localmente ao espaço euclidiano n -dimensional em uma vizinhança de cada ponto. Ainda mais, uma variedade compacta é uma variedade que é compacta como espaço topológico.

Considere M um espaço topológico. Uma carta local é um par (U, φ) onde U é um aberto de M e $\varphi : U \rightarrow \mathbb{R}^m$ é um homeomorfismo sobre o aberto $\varphi(U) \subset \mathbb{R}^m$. Denomina-se atlas de dimensão m e classe C^r sobre M uma coleção $\mathcal{A}$ de cartas locais tais que:

1. $\bigcup_{(U, \varphi) \in \mathcal{A}} U = M$;
2. Se $(U, \varphi), (\tilde{U}, \tilde{\varphi}) \in \mathcal{A}$ e $U \cap \tilde{U} \neq \emptyset$, então a aplicação

$$\tilde{\varphi} \circ \varphi^{-1} : \varphi(U \cap \tilde{U}) \rightarrow \tilde{\varphi}(U \cap \tilde{U})$$

é um difeomorfismo C^r entre abertos de $\mathbb{R}^m$.

Os difeomorfismos acima são denominados mudanças de coordenadas. Além disso, o atlas $\mathcal{A}$ é denominado atlas maximal quando contém todas as cartas locais $(\tilde{U}, \tilde{\varphi})$ cujas mudanças de coordenadas com elementos $(U, \varphi) \in \mathcal{A}$ são difeomorfismos C^r . Um espaço topológico de Hausdorff M com base enumerável, munido de um atlas maximal é denominado variedade diferenciável.

Considerando M uma m -variedade com atlas $\mathcal{A} = \{\varphi_a : U_a \rightarrow X_a\}$, pode-se considerar o conceito de subvariedade. Dado $N \subset M$, N é dita subvariedade de dimensão $n < m$ se existir algum atlas $\mathcal{B} = \{\psi_b : V_b \rightarrow Y_b\}$ de M tal que $\mathcal{A}$ e $\mathcal{B}$ são compatíveis, isto é, as mudanças de coordenadas $\psi_b \circ \varphi_a^{-1}$ e $\varphi_a \circ \psi_b^{-1}$ são diferenciáveis em seus domínios para

todo a, b , e a carta local ψ_b envia $V'_b = V_b \cap N$ sobre um aberto $Y'_b \subset \mathbb{R}^n \times 0^{m-n}$ para todo b . Note que, em particular, toda subvariedade é uma variedade.

Outro conceito importante é o de espaço tangente à uma variedade M . Para cada $p \in M$, considere o conjunto $\mathcal{C}(p)$ das curvas $c : I \rightarrow M$, onde I é um intervalo aberto contendo zero, tais que $c(0) = p$ e c é diferenciável em 0. Duas curvas $c_1, c_2 \in \mathcal{C}(p)$ são equivalentes se $(\varphi \circ c_1)'(0) = (\varphi \circ c_2)'(0)$ para toda carta local $\varphi : U \rightarrow X$, com $p \in U$. Seja $[c]$ a classe de equivalência de qualquer curva $c \in \mathcal{C}(p)$. O espaço tangente à M no ponto p , denotado por $T_p M$, é o conjunto de tais classes de equivalência.

Dada uma carta local $\varphi_a : U_a \rightarrow X_a$ qualquer, com $p \in U_a$, a aplicação

$$\begin{aligned} D\varphi_a(p) : T_p M &\rightarrow \mathbb{R}^m \\ [c] &\mapsto (\varphi_a \circ c)'(0) \end{aligned}$$

está bem definida e é uma bijeção, de modo que é possível identificar $T_p M$ com $\mathbb{R}^d$ e o espaço tangente adquire uma estrutura de espaço vetorial que não depende da escolha da carta local. De fato, dada qualquer outra carta local $\psi_b : V_b \rightarrow Y_b$ com $p \in V_b$, tem-se

$$D\psi_b(p) = D(\psi_b \circ \varphi_a^{-1})(\varphi_a(p)) \circ D\varphi_a(p).$$

Como $D(\psi_b \circ \varphi_a^{-1})(\varphi_a(p))$ é um isomorfismo linear, segue que as estruturas de espaço vetorial transportadas por $D\varphi_a(p)$ e $D\psi_b(p)$ do espaço euclidiano para $T_p M$ coincidem.

Com a noção de espaço tangente é possível introduzir a ideia de derivada de uma função entre duas variedades: a aplicação $f : M \rightarrow N$, onde M e N são variedades, é diferenciável se $\psi_b \circ f \circ \varphi_a^{-1} : \varphi_a(U_a \cap f^{-1}(V_b)) \rightarrow \psi_b(V_b \cap f(U_a))$ é uma aplicação diferenciável para toda carta local $\varphi_a : U_a \rightarrow X_a$ de M e toda carta local $\psi_b : V_b \rightarrow Y_b$ com $f(U_a) \cap V_b \neq \emptyset$. Além disso, f é de classe C^r se M e N são variedades C^r e toda aplicação $\psi_b \circ f \circ \varphi_a^{-1}$ é de classe C^r . A derivada de uma aplicação diferenciável entre duas variedades $f : M \rightarrow N$ no ponto $p \in M$ é a aplicação linear $Df(p) : T_p M \rightarrow T_{f(p)} N$ definida por

$$Df(p) = D\psi_b(f(p))^{-1} \circ D(\psi_b \circ f \circ \varphi_a^{-1})(\varphi_a(p)) \circ D\varphi_a(p)$$

onde $\varphi_a : U_a \rightarrow X_a$ é uma carta local de M com $p \in U_a$ e $\psi_b : V_b \rightarrow Y_b$ é uma carta local de N com $f(p) \in V_b$. Ainda mais, pode-se escrever a derivada de uma transformação diferenciável entre duas variedades pensando no conceito de fibrado tangente. O fibrado tangente à variedade M é a união disjunta

$$TM = \bigcup_{p \in M} T_p M$$

de todos os espaços tangentes a M . Para cada carta local $\varphi : U_a \rightarrow X_a$, considere

$$T_{U_a}M = \bigcup_{p \in U_a} T_pM$$

e a aplicação $D\varphi_a : T_{U_a}M \rightarrow X_a \times \mathbb{R}^m$ dada por $[c] \mapsto ((\varphi_a \circ c)(0), (\varphi_a \circ c)'(0))$. Considere em TM a topologia que torna cada $D\varphi$ um homeomorfismo. Supondo que o atlas $\{\varphi_a : U \rightarrow X_a\}$ de M é de classe C^r então, a mudança de coordenadas

$$D\varphi_b \circ \varphi_a^{-1} : \varphi_a(U_a \cap U_b) \times \mathbb{R}^m \rightarrow \varphi_b(U_a \cap U_b) \times \mathbb{R}^m$$

é uma aplicação de classe C^{r-1} , para a, b quaisquer tais que $U_a \cap U_b \neq \emptyset$. Deste modo, TM está munido com a estrutura de variedade de classe C^{r-1} e dimensão $2m$. Assim, a derivada de uma transformação diferenciável $f : M \rightarrow N$ é a aplicação $Df : TM \rightarrow TN$ cuja restrição a cada espaço tangente T_pM está dada por $Df(p)$. Se f é de classe C^r , então Df é de classe C^{r-1} , relativamente à estrutura de variedade nos fibrados tangentes TM e TN introduzida anteriormente.

Por fim, um campo de vetores na variedade M é uma aplicação $X : M \rightarrow TM$ tal que $\pi \circ X = id$, onde $\pi : TM \rightarrow M$ associa a cada $v \in TM$ o único ponto $p \in M$ tal que $v \in T_pM$. Ou seja, a aplicação X associa p a um elemento do espaço tangente. O campo de vetores é de classe C^k , com $k \leq r - 1$, se X for de classe C^k . Suponha que $k \geq 1$. Então o teorema de existência e unicidade de soluções de equações diferenciais garante que, para todo $p \in M$, existe uma única curva $c_p : I_p \rightarrow M$ tal que $c_p(0) = p$ e $c_p'(t) = X(c_p(t))$ para todo $t \in I_p$, onde I_p é intervalo aberto maximal com esta propriedade. Se M é compacta, tem-se $I_p = \mathbb{R}$ para qualquer $p \in M$. Além disso, as transformações $f^t : M \rightarrow M$ definidas por $f^t(p) = c_p(t)$ são difeomorfismos de classe C^k , com $f^0 = id$ e $f^s \circ f^t = f^{s+t}$ para quaisquer $s, t \in \mathbb{R}$. A família $f^t : t \in \mathbb{R}$ é chamada fluxo do campo de vetores X .

Variedades riemannianas Uma variedade riemanniana é uma variedade M munida de uma métrica riemanniana, que consiste de uma aplicação que associa a cada $p \in M$ um produto interno no espaço tangente T_pM , ou seja, uma aplicação bilinear simétrica

$$\cdot_p : T_pM \times T_pM \rightarrow \mathbb{R}$$

tal que $v \cdot_p v > 0$ para todo vetor não nulo $v \in T_pM$. Ainda, é necessário que este produto interno varie diferenciavelmente com o ponto p , no seguinte sentido: dada uma carta local $\varphi : U \rightarrow X$ de M qualquer, para cada $p \in U$ é possível identificar T_pM com $\mathbb{R}^m$, via $D\varphi(p)$, e então considerar $\cdot_p$ como um produto interno no espaço euclidiano. Com isso, considerando $e_1, e_2, \dots, e_m$ uma base de $\mathbb{R}^m$, toma-se as funções $g_{i,j}(p) = e_i \cdot_p e_j$ diferenciáveis para todo par (i, j) e para qualquer escolha da carta local e da base.

Toda subvariedade N de uma variedade riemanniana M herda uma estrutura de variedade riemanniana, dada pela restrição do produto interno $\cdot_p$ de M ao subespaço tangente $T_p N$ de cada $p \in N$. Toda variedade compacta admite (infinitas) métricas riemannianas. Isso segue do Teorema de Whitney, o qual afirma que toda variedade compacta pode ser realizada como subvariedade de algum espaço euclidiano.

Definição 1.17. Considere M uma variedade riemanniana e $\gamma : [a, b] \rightarrow M$ uma curva diferenciável. O comprimento de γ é dado por

$$\ell(\gamma) = \int_a^b \|\gamma'(t)\|_{\gamma(t)} dt$$

onde $\|v\|_p = \sqrt{(v \cdot_p v)}$.

Definição 1.18. Considere uma variedade riemanniana M . A distância d associada à métrica riemanniana entre dois pontos $p, q \in M$ é

$$d(p, q) = \inf_{\gamma \in \Lambda_{p,q}} \ell(\gamma)$$

onde $\Lambda_{p,q}$ denota o conjunto das curvas diferenciáveis ligando p e q .

A curva diferenciável $\gamma : [a, b] \rightarrow M$ é dita minimizante se $\ell = d(\gamma(a), \gamma(b))$. Uma curva diferenciável $\gamma : I \rightarrow M$, definida em um intervalo aberto I , é chamada geodésica se é localmente minimizante: para todo $a \in I$ existe $\delta > 0$ tal que a restrição de γ ao intervalo $[a - \delta, a + \delta]$ é minimizante. Note que toda curva minimizante é geodésica, mas a recíproca não é verdadeira. Além disso, se γ é geodésica, então $\|\gamma'(t)\|_{\gamma(t)}$ é constante em I , e, para todo $p \in M$ e todo $v \in T_p M$ existe uma única geodésica $\gamma_{p,v} : I_{p,v} \rightarrow M$ tal que $\gamma_{p,v}(0) = p$, $\gamma'_{p,v}(0) = v$ e $I_{p,v}$ é intervalo maximal tal que $\gamma_{p,v}$ é localmente minimizante. O fluxo geodésico no fibrado tangente é

$$\begin{aligned} f^t : TM &\rightarrow TM \\ (p, v) &\mapsto (\gamma_{p,v}(t), \gamma'_{p,v}(t)). \end{aligned}$$

Folheação Grosso modo, uma folheação de dimensão n de uma m -variedade diferenciável M é um tipo de decomposição de M em subvariedades conexas de dimensão $n < m$, denominadas folhas, que se empilham localmente. São como subconjuntos de $\mathbb{R}^n \times \mathbb{R}^{m-n}$ com a segunda coordenada constante. Formalmente: Seja M uma variedade de dimensão m e classe C^∞ . Uma folheação de dimensão $n < m$ é um atlas maximal $\mathcal{F}$ de M com as seguintes propriedades:

1. Se $(U, \varphi) \in \mathcal{F}$, então $\varphi(U) = U_1 \times U_2 \subset \mathbb{R}^n \times \mathbb{R}^{m-n}$ onde U_1 e U_2 são discos abertos.
2. Se $(U, \varphi) \in \mathcal{F}$ e $(V, \psi) \in \mathcal{F}$ são tais que $U \cap V \neq \emptyset$, então o mapa mudança de coordenadas $\psi \circ \varphi^{-1} : \varphi(U \cap V) \rightarrow \psi(U \cap V)$ é da forma $\psi \circ \varphi^{-1} = (h_1(x, y), h_2(y))$.

Dada uma folheação $\mathcal{F}$ de classe C^r e dimensão n , com $0 < n < m$, de uma m -variedade M , considere $(U, \varphi) \in \mathcal{F}$ com $\varphi(U) = U_1 \times U_2 \subset \mathbb{R}^n \times \mathbb{R}^{m-n}$. Os conjuntos da forma $\varphi^{-1}(U_1 \times \{c\})$, onde $c \in U_2$ são denominados placas de $\mathcal{F}$. Além disso, denomina-se por caminho de placas de $\mathcal{F}$ a sequência $\alpha_1, \dots, \alpha_k$ de placas tais que $\alpha_i \cap \alpha_{i+1} \neq \emptyset$ para $i = 1, \dots, k-1$. Considere a relação de equivalência $p \sim q \iff \exists$ um caminho de placas $\alpha_1, \dots, \alpha_k$ tal que $p \in \alpha_1, q \in \alpha_k$. As classes de equivalência desta relação são denominadas folhas de $\mathcal{F}$. É imediato da definição que cada folha é um conjunto conexo por caminhos.

CAPÍTULO 2

Desintegração de medidas

Von Neumann, em 1932, introduziu uma definição rigorosa de desintegração de medidas de probabilidade [31]. Desde então, este conceito tem sido uma ferramenta muito útil no estudo de diversas propriedades de sistemas dinâmicos. Como um exemplo, pode-se citar o Teorema da Decomposição Ergódica, que estabelece, sob certas condições do espaço, que toda medida de probabilidade invariante pode ser escrita como uma combinação convexa de medidas ergódicas (para mais detalhes, veja Teorema 5.1.3 em [21]). Este teorema está relacionado com um importante resultado da Teoria da Medida, o Teorema de Rokhlin, que afirma que, asseguradas certas condições do espaço, é possível realizar a desintegração de uma probabilidade em termos de medidas condicionais.

Neste capítulo, é apresentada uma definição formal do conceito de desintegração de medidas, em termos de medidas condicionais. Ademais, o Teorema de Rokhlin é apresentado em diferentes perspectivas: na primeira, baseada em [21], considera-se um espaço métrico completo e separável e uma partição mensurável deste; em outra perspectiva, considera-se uma formulação mais moderna, apresentada originalmente em [25], que vai mais além, trazendo, também, uma generalização deste teorema para uma categoria de espaços σ -finitos com morfismos absolutamente contínuos.

2.1 Uma definição formal de desintegração

Considere $(X, \mathcal{B}, \mu)$ um espaço de probabilidade e $\mathcal{P}$ uma partição de X em subconjuntos mensuráveis. Denote por π a projeção natural que associa cada ponto $x \in X$ ao elemento da partição $P \in \mathcal{P}$ que o contém:

$$\begin{aligned}\pi : X &\rightarrow \mathcal{P} \\ x &\mapsto P(x).\end{aligned}$$

Note que a aplicação π fornece uma estrutura de espaço de probabilidade à partição $\mathcal{P}$. De fato, definindo que $A \subset \mathcal{P}$ é mensurável se, e somente se, $\pi^{-1}(A)$ é um subconjunto mensurável de X , tem-se que $\hat{\mathcal{B}} = \{A \subset \mathcal{P} : A \text{ é mensurável}\}$ é uma σ -álgebra em $\mathcal{P}$. Assim,

considerando $\hat{\mu}$ o *push forward* de μ por π , isto é:

$$\hat{\mu}(A) := \pi_*\mu(A) = \mu(\pi^{-1}(A))$$

para todo $A \subset \mathcal{P}$ mensurável, define-se uma probabilidade em $\mathcal{P}$. A medida $\hat{\mu}$ é denominada medida quociente. Neste contexto, define-se uma desintegração de μ da seguinte maneira:

Definição 2.1. *Uma desintegração de μ relativamente à uma partição $\mathcal{P}$ de X é uma família $\{\mu_P : P \in \mathcal{P}\}$ de probabilidades em X tal que, para $E \subset X$ mensurável:*

1. $\mu_P(P) = 1$ para $\hat{\mu}$ -quase todo $P \in \mathcal{P}$.
2. $P \mapsto \mu_P(E)$ é mensurável.
3. $\mu(E) = \int \mu_P(E) d\hat{\mu}(P)$.

Neste caso, as probabilidades μ_P são denominadas probabilidades condicionais de μ com relação à $\mathcal{P}$.

Proposição 2.2. *Se a σ -álgebra $\mathcal{B}$ admite um gerador enumerável e se $\{\mu_P : P \in \mathcal{P}\}$ e $\{\mu'_P : P \in \mathcal{P}\}$ são desintegrações de μ relativamente a $\mathcal{P}$, então $\mu_P = \mu'_P$ para $\hat{\mu}$ -quase todo $P \in \mathcal{P}$.*

Demonstração. Considere Γ um gerador enumerável de $\mathcal{B}$ e $\mathcal{A}$ a álgebra gerada por Γ , de modo que $\mathcal{A}$ seja enumerável. Para cada $A \in \mathcal{A}$, considere:

$$\mathcal{D}_A = \{P \in \mathcal{P} : \mu_P(A) > \mu'_P(A)\}$$

$$\mathcal{E}_A = \{P \in \mathcal{P} : \mu_P(A) < \mu'_P(A)\}.$$

Se $P \in \mathcal{D}_A$, então $P \subset \pi^{-1}(\mathcal{D}_A)$ e, pela definição de desintegração, tem-se

$$\mu_P(A \cap \pi^{-1}(\mathcal{D}_A)) = \mu_P(A).$$

Se $P \notin \mathcal{D}_A$, então $P \cap \pi^{-1}(\mathcal{D}_A) = \emptyset$, de modo que

$$\mu_P(A \cap \pi^{-1}(\mathcal{D}_A)) = 0.$$

Ademais, o mesmo é válido ao considerar μ'_P . Logo, pela propriedade 3 da Definição 2.1, segue que:

$$\mu(A \cap \pi^{-1}(\mathcal{D}_A)) = \begin{cases} \int \mu_P(A \cap \pi^{-1}(\mathcal{D}_A)) d\hat{\mu}(P) = \int_{\mathcal{D}_A} \mu_P(A) d\hat{\mu}(P) \\ \int \mu'_P(A \cap \pi^{-1}(\mathcal{D}_A)) d\hat{\mu}(P) = \int_{\mathcal{D}_A} \mu'_P(A) d\hat{\mu}(P). \end{cases}$$

Uma vez que $\mu_P(A) > \mu'_P(A)$, segue que $\hat{\mu}(\mathcal{D}_A) = 0$ para todo $A \in \mathcal{A}$. Com o mesmo argumento, segue que $\hat{\mu}(\mathcal{E}_A) = 0$ para todo $A \in \mathcal{A}$. Assim,

$$\hat{\mu}\left(\bigcup_{A \in \mathcal{A}} (\mathcal{D}_A \cup \mathcal{E}_A)\right) = 0.$$

Para todo P em $\left(\bigcup_{A \in \mathcal{A}} (\mathcal{D}_A \cup \mathcal{E}_A)\right)^c$, as medidas μ_P e μ'_P coincidem em $\mathcal{A}$ e, portanto, coincidem em $\mathcal{B}$, ou seja, $\mu_P = \mu'_P$ para $\hat{\mu}$ -quase todo $P \in \mathcal{P}$. $\square$

A Proposição 2.2 afirma a unicidade da desintegração para um caso bem geral, pois vale sempre que X é um espaço topológico com base enumerável de abertos e $\mathcal{B}$ é uma σ -álgebra de Borel. Apesar disso, ainda não há garantia da existência de uma desintegração. Por exemplo, considere a relação de equivalência em $\mathbb{R}$ dada por

$$x \sim y \iff x - y \in \mathbb{Z}.$$

O espaço S^1 de todas essas classes de equivalência é canonicamente identificado com o círculo unitário no plano complexo. Tomando $\pi : \mathbb{R} \rightarrow S^1$ que associa cada $x \in \mathbb{R}$ à sua classe de equivalência, defina a medida de Lebesgue λ em S^1 por

$$\lambda(E) := m(\pi^{-1}(E) \cap [k, k+1))$$

para qualquer $k \in \mathbb{Z}$, onde m é a medida de Lebesgue em $\mathbb{R}$. Observe que λ é uma medida de probabilidade e é invariante por rotações, isto é, se $R : S^1 \rightarrow S^1$ é uma rotação, então $\lambda(A) = \lambda(R^{-1}(A))$ para todo $A \subset S^1$ mensurável. Considere $f : S^1 \rightarrow S^1$ uma rotação irracional e $\mathcal{P}$ uma partição de S^1 cujos elementos são as órbitas $\{f^n(x) : n \in \mathbb{Z}\}$ de f . Suponha que existe uma desintegração $\{\lambda_P : P \in \mathcal{P}\}$ de λ relativamente à $\mathcal{P}$. Considerando a família $\{f_*\lambda_P : P \in \mathcal{P}\}$, tem-se que: como todos os elementos da partição são conjuntos invariantes, $f_*\lambda_P(P) = \lambda_P(P) = 1$ para $\hat{\lambda}$ -quase todo $P \in \mathcal{P}$; Para $E \subset S^1$ mensurável, $P \mapsto f_*\lambda_P(E)$ é mensurável, uma vez que $\{\lambda_P\}$ é desintegração; Como λ é invariante por f ,

$$\lambda(E) = \lambda(f^{-1}(E)) = \int f_*\lambda_P(E) d\hat{\lambda}(P).$$

Deste modo, $\{f_*\lambda_P : P \in \mathcal{P}\}$ é uma desintegração de λ relativamente à $\mathcal{P}$. Entretanto, pela unicidade, $f_*\lambda_P = \lambda_P$ para $\hat{\lambda}$ -quase todo $P \in \mathcal{P}$, de modo que λ_P é f -invariante, o que é uma contradição, pois $\mathcal{P}$ é conjunto infinito enumerável e portanto não pode existir nenhuma probabilidade invariante que dê peso positivo a P . Então, não é possível desintegrar λ relativamente à $\mathcal{P}$ [21, Exemplo 5.1.8].

O Teorema de Rokhlin afirma a existência da desintegração de uma medida sob certas condições do espaço X e da partição $\mathcal{P}$ de X : é necessário que X seja espaço métrico

completo e separável e que $\mathcal{P}$ seja mensurável. A partição $\mathcal{P}$ é mensurável se existe algum conjunto mensurável $X_0 \subset X$ com medida total tal que $\mathcal{P} = \bigvee_{n=1}^{\infty} \mathcal{P}_n$ restrito a X_0 , para uma sequência crescente de partições enumeráveis $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots \prec \mathcal{P}_n \prec \cdots$. Note que $\mathcal{P}_i \prec \mathcal{P}_{i+1}$ significa que todo elemento de $\mathcal{P}_{i+1}$ está contido em algum elemento de $\mathcal{P}_i$.

2.2 Teorema de Rokhlin

Considerando a mesma notação da seção anterior, o Teorema de Rokhlin pode ser enunciado da seguinte forma:

Teorema 2.3. *Seja X um espaço métrico completo e separável, seja $\mathcal{P}$ uma partição mensurável de X e seja μ uma medida de probabilidade. Então μ admite uma desintegração relativamente à $\mathcal{P}$, ou seja, existe $\{\mu_P : P \in \mathcal{P}\}$ tal que, para todo $E \subset X$ mensurável:*

1. $\mu_P(P) = 1$ para $\hat{\mu}$ -quase todo $P \in \mathcal{P}$.
2. $P \mapsto \mu_P(E)$ é mensurável.
3. $\mu(E) = \int \mu_P(E) d\hat{\mu}(P)$.

Para demonstrar o Teorema 2.3, são necessários alguns resultados. Considere $P_n(x)$ o elemento da partição $\mathcal{P}_n$ que contém o ponto $x \in X$. Dada uma função mensurável limitada $\varphi : X \rightarrow \mathbb{R}$, defina $e_n(\varphi) : X \rightarrow \mathbb{R}$, a esperança condicional de φ relativamente à $\mathcal{P}_n$, por:

$$e_n(\varphi, x) = \begin{cases} \frac{1}{\mu(P_n(x))} \int_{P_n(x)} \varphi d\mu, & \text{se } \mu(P_n(x)) > 0 \\ 0, & \text{se } \mu(P_n(x)) = 0. \end{cases}$$

Observe que $e_n(\varphi)$ é constante em cada $P_n \in \mathcal{P}_n$ e que é possível escrever:

$$\begin{aligned} \int e_n(\varphi) d\mu &= \sum_{P_n} \int_{P_n} e_n(\varphi) d\mu \\ &= \sum_{P_n} \int_{P_n} \left(\frac{1}{\mu(P_n)} \int_{P_n} \varphi d\mu \right) d\mu \\ &= \sum_{P_n} \left(\frac{1}{\mu(P_n)} \int_{P_n} \varphi d\mu \right) \int_{P_n} 1 d\mu \\ &= \sum_{P_n} \int_{P_n} \varphi d\mu \\ &= \int \varphi d\mu. \end{aligned}$$

Com isso, considere o seguinte lema:

Lema 2.4. *Dada uma função mensurável limitada $\varphi : X \rightarrow \mathbb{R}$, existe $X_\varphi \subset X$, com $\mu(X_\varphi) = 1$, tal que:*

1. $e(\varphi, x) = \lim_{n \rightarrow \infty} e_n(\varphi, x)$ existe para todo $x \in X_\varphi$.
2. $e(\varphi) : X_\varphi \rightarrow \mathbb{R}$ é mensurável e constante em cada $P \in \mathcal{P}$.
3. $\int \varphi d\mu = \int e(\varphi) d\mu$.

Demonstração. Suponha, primeiramente, que $\varphi \geq 0$. Dados α, β , com $\alpha < \beta$, tome

$$S(\alpha, \beta) = \{x \in X : \liminf_n e_n(\varphi, x) < \alpha < \beta < \limsup_n e_n(\varphi, x)\}.$$

Note que $e_n(\varphi, x)$ diverge se, e somente se, $x \in S(\alpha, \beta)$, com α e β números racionais, isto é, $e(\varphi, x)$ existe se, e somente se, $x \in X_\varphi$, onde

$$X_\varphi = \bigcap_{\substack{\alpha, \beta \in \mathbb{Q} \\ \alpha < \beta}} S(\alpha, \beta)^c.$$

Afirmção. $\mu(X_\varphi) = 1$.

Demonstração. Considere α e β fixos e, dado $x \in S(\alpha, \beta)$, fixe uma sequência de inteiros $1 \leq a_1^x < b_1^x < \dots < a_i^x < b_i^x < \dots$ tal que $e_{a_i^x}(\varphi, x) \leq \alpha$ e $e_{b_i^x}(\varphi, x) \geq \beta$, $\forall i \geq 1$. Considere os conjuntos

$$A_i = \bigcup_{x \in S(\alpha, \beta)} \mathcal{P}_{a_i^x}(x)$$

$$B_i = \bigcup_{x \in S(\alpha, \beta)} \mathcal{P}_{b_i^x}(x)$$

de modo que $S(\alpha, \beta) \subset A_{i+1} \subset B_i \subset A_i$, uma vez que $\mathcal{P}_n \prec \mathcal{P}_{n+1}$ para todo $n \in \mathbb{N}$. Deste modo, tem-se que

$$S(\alpha, \beta) \subset \tilde{S}(\alpha, \beta) = \bigcap_{i=1}^{\infty} B_i = \bigcap_{i=1}^{\infty} A_i.$$

Ainda, note que, dados dois conjuntos quaisquer que formam A_i , por exemplo, $\mathcal{P}_{a_i^{x_1}}(x_1)$ e $\mathcal{P}_{a_i^{x_2}}(x_2)$, ou $\mathcal{P}_{a_i^{x_1}}(x_1) \cap \mathcal{P}_{a_i^{x_2}}(x_2) = \emptyset$ ou um está contido no outro. Deste modo, os conjuntos $\mathcal{P}_{a_i^x}(x)$ maximais são disjuntos dois-a-dois, formando uma partição de A_i . Somando sobre estes conjuntos maximais com medida positiva, e considerando a definição de $e_{a_i^x}(\varphi, x)$, tem-se:

$$\begin{aligned}
\int_{A_i} \varphi \, d\mu &= \sum_{\mathcal{P}_{a_i^x}(x)} \int_{\mathcal{P}_{a_i^x}(x)} \varphi \, d\mu \\
&= \sum_{\mathcal{P}_{a_i^x}(x)} e_{a_i^x}(\varphi, x) \, \mu(\mathcal{P}_{a_i^x}(x)) \\
&\leq \sum_{\mathcal{P}_{a_i^x}(x)} \alpha \, \mu(\mathcal{P}_{a_i^x}(x)) \\
&= \alpha \, \mu(A_i).
\end{aligned}$$

Analogamente,

$$\begin{aligned}
\int_{B_i} \varphi \, d\mu &= \sum_{\mathcal{P}_{b_i^x}(x)} \int_{\mathcal{P}_{b_i^x}(x)} \varphi \, d\mu \\
&= \sum_{\mathcal{P}_{b_i^x}(x)} e_{b_i^x}(\varphi, x) \, \mu(\mathcal{P}_{b_i^x}(x)) \\
&\geq \sum_{\mathcal{P}_{b_i^x}(x)} \beta \, \mu(\mathcal{P}_{b_i^x}(x)) \\
&= \beta \, \mu(B_i).
\end{aligned}$$

Como $A_i \supset B_i$, e como $\varphi \geq 0$ por hipótese, tem-se:

$$\alpha \, \mu(A_i) \geq \int_{A_i} \varphi \, d\mu \geq \int_{B_i} \varphi \, d\mu \geq \beta \, \mu(B_i). \quad (2.1)$$

Por convergência monótona,

$$\lim_{i \rightarrow \infty} \int_{A_i} \varphi \, d\mu = \int_{\tilde{S}(\alpha, \beta)} \varphi \, d\mu = \lim_{i \rightarrow \infty} \int_{B_i} \varphi \, d\mu.$$

Com isso, ao tomar o limite da equação (2.1) quando $i \rightarrow \infty$, tem-se que

$$\alpha \, \mu(\tilde{S}(\alpha, \beta)) \geq \beta \, \mu(\tilde{S}(\alpha, \beta))$$

o que implica que $\mu(\tilde{S}(\alpha, \beta)) = 0$, uma vez que $\alpha < \beta$. Agora, como $S(\alpha, \beta) \subset \tilde{S}(\alpha, \beta)$, tem-se que $\mu(S(\alpha, \beta)) = 0$. Como X_φ é uma interseção enumerável, segue que $\mu(X_\varphi) = 0$.

◁

Para o caso geral, em que φ não é necessariamente uma função não negativa, pode-se escrever $\varphi = \varphi^+ - \varphi^-$ onde φ^+ e φ^- são mensuráveis, não negativas e limitadas.

Com isso, existem conjuntos $X_{\varphi^+}, X_{\varphi^-} \subset X$, com $\mu(X_{\varphi^+}) = \mu(X_{\varphi^-}) = 1$, tais que

$$e(\varphi^+, x) = \lim_{n \rightarrow \infty} e_n(\varphi^+, x)$$

$$e(\varphi^-, x) = \lim_{n \rightarrow \infty} e_n(\varphi^-, x).$$

Uma vez que $e_n(\varphi, x) = e_n(\varphi^+, x) - e_n(\varphi^-, x)$, o limite

$$\lim_{n \rightarrow \infty} e_n(\varphi, x) = \lim_{n \rightarrow \infty} e_n(\varphi^+, x) - \lim_{n \rightarrow \infty} e_n(\varphi^-, x)$$

existe para todo $x \in X_\varphi$, onde o conjunto $X_\varphi = X_{\varphi^+} \cap X_{\varphi^-}$ é tal que $\mu(X_\varphi) = 1$. Com isso, fica demonstrada a primeira parte do lema. Para a segunda parte, observe que $e(\varphi)$ é mensurável como limite de funções mensuráveis. Ainda mais, dado que $\mathcal{P}_n \prec \mathcal{P}$, tem-se que $e_n(\varphi)$ é constante em cada $P \in \mathcal{P}$, restrito a um subconjunto de X com medida total. Logo o mesmo vale para $e(\varphi)$. Por fim, para a terceira parte, note que $|e_n(\varphi)| \leq \sup |\varphi|$, para todo $n \geq 1$. Sabe-se que

$$\int \varphi \, d\mu = \int e_n(\varphi) \, d\mu$$

então, pelo Teorema 1.7, tem-se

$$\int e(\varphi) \, d\mu = \lim_{n \rightarrow \infty} \int e_n(\varphi) \, d\mu = \int \varphi \, d\mu.$$

□

Demonstração do Teorema 2.3. Considere um subconjunto $A \subset X$ mensurável e a função $\varphi = \mathbb{1}_A$. Pelo Lema 2.4, tem-se

$$e(\varphi, x) = e(\mathbb{1}_A, x) = \lim_{n \rightarrow \infty} \frac{\mu(\mathcal{P}_n(x) \cap A)}{\mu(\mathcal{P}_n(x))}.$$

Além disso, considere o conjunto

$$\mathcal{P}_A := \{P \in \mathcal{P} : P \cap X_{\mathbb{1}_A} \neq \emptyset\}$$

de modo que $\hat{\mu}(\mathcal{P}_A) = 1$. Defina a função $E(A) : \mathcal{P}_A \rightarrow \mathbb{R}$ que associa cada $P \in \mathcal{P}_A$ ao valor $E(A, P) = e(\mathbb{1}_A, x)$ para qualquer $x \in P \cap X_{\mathbb{1}_A}$. Observe que $E(A)$ está bem definida e é mensurável, pois $e(\mathbb{1}_A) = E(A) \circ \pi$ (lembrando que $\pi : X \rightarrow \mathcal{P}$ é a aplicação que associa cada elemento $x \in X$ com a partição que o contém), de modo que

$$\int \varphi \, d\mu = \int e(\varphi) \, d\mu = \int e(\mathbb{1}_A) \, d\mu = \int E(A) \circ \pi \, d\mu = \int E(A) \, d\hat{\mu}.$$

Como X é separável, possui base enumerável de abertos. Seja $\mathcal{U} = \{U_k : k \in \mathbb{N}\}$ uma base de abertos de X e seja $\mathcal{A}$ a álgebra gerada por $\mathcal{U}$. $\mathcal{A}$ gera a σ -álgebra de Borel de X e, além disso $\mathcal{A}$ é enumerável, pois coincide com a união das álgebras geradas pelos subconjuntos $\{U_k : 1 \leq k \leq n\}$, para cada $n \geq 1$, que são finitas. Defina

$$\mathcal{P}_* = \bigcap_{A \in \mathcal{A}} \mathcal{P}_A.$$

Então

$$\hat{\mu}(\mathcal{P}_*) = \hat{\mu}\left(\bigcap_{A \in \mathcal{A}} \mathcal{P}_A\right) = \mu\left(\pi^{-1}\left(\bigcap_{A \in \mathcal{A}} \mathcal{P}_A\right)\right).$$

Observe que

$$X - \bigcap_{A \in \mathcal{A}} \mathcal{P}_A = X \cap \left(\bigcap_{A \in \mathcal{A}} \mathcal{P}_A\right)^c = X \cap \left(\bigcup_{A \in \mathcal{A}} \mathcal{P}_A^c\right) = \bigcup_{A \in \mathcal{A}} (X \cap \mathcal{P}_A^c) = \bigcup_{A \in \mathcal{A}} \mathcal{P}_A^c$$

e, deste modo,

$$\hat{\mu}\left(X - \bigcap_{A \in \mathcal{A}} \mathcal{P}_A\right) = \hat{\mu}\left(\bigcup_{A \in \mathcal{A}} \mathcal{P}_A^c\right) = 0.$$

Portanto, segue que $\hat{\mu}(\mathcal{P}_*) = 1$. Agora, para cada $P \in \mathcal{P}_*$, defina

$$\begin{aligned} \mu_P : \mathcal{A} &\rightarrow [0, 1] \\ A &\mapsto E(A, P). \end{aligned}$$

Note que:

$\rightarrow \mu_P(X) = E(X, P) = e(\mathbb{1}_X, x)$ para qualquer $x \in X_{\mathbb{1}_X}$. Portanto,

$$\begin{aligned} \mu_P(X) &= \lim_{n \rightarrow \infty} e_n(\mathbb{1}_X, x) \\ &= \lim_{n \rightarrow \infty} \frac{1}{\mu(P_n(x))} \int_{P_n(x)} \mathbb{1}_X d\mu \\ &= \lim_{n \rightarrow \infty} \frac{\mu(P_n(x) \cap X)}{\mu(P_n(x))} \\ &= 1. \end{aligned}$$

→ μ_P é uma função aditiva. De fato, tome A e B disjuntos. Então, para todo $P \in \mathcal{P}_*$:

$$\begin{aligned}
 E(A \cup B, P) &= e(\mathbb{1}_{A \cup B}, x) = \lim_{n \rightarrow \infty} \frac{1}{\mu(P_n(x))} \int_{P_n(x)} \mathbb{1}_{A \cup B} d\mu \\
 &= \lim_{n \rightarrow \infty} \frac{\mu(P_n(x) \cap (A \cup B))}{\mu(P_n(x))} \\
 &= \lim_{n \rightarrow \infty} \frac{\mu((P_n(x) \cap A) \cup (P_n(x) \cap B))}{\mu(P_n(x))} \\
 &= \lim_{n \rightarrow \infty} \frac{\mu(P_n(x) \cap A) + \mu(P_n(x) \cap B)}{\mu(P_n(x))} \\
 &= \lim_{n \rightarrow \infty} \frac{\mu(P_n(x) \cap A)}{\mu(P_n(x))} + \lim_{n \rightarrow \infty} \frac{\mu(P_n(x) \cap B)}{\mu(P_n(x))} \\
 &= E(A, P) + E(B, P).
 \end{aligned}$$

Por [21, Proposição 5.2.2], esta função se estende a uma medida de probabilidade (que ainda será denotada por μ_P) definida na σ -álgebra de Borel de X . Assim, obtém-se uma família de medidas de probabilidade $\{\mu_P : P \in \mathcal{P}_*\}$. Agora, basta mostrar que esta família é, de fato, uma desintegração de μ relativamente à $\mathcal{P}_*$:

Condição 1. $\mu_P(P) = 1$ para $\hat{\mu}$ -quase todo $P \in \mathcal{P}_*$.

Seja $P \in \mathcal{P}_*$. Para cada $n \in \mathbb{N}$, considere P_n o elemento de $\mathcal{P}_n$ que contém P . Note que, se $A \in \mathcal{A}$ é tal que $A \cap P_n = \emptyset$ para dado n , tem-se que $A \cap P_m = \emptyset$ para todo $m \geq n$, já que $P_m \subset P_n$ para todo $m > n$. Então,

$$\mu_P(A) = E(A, P) = \lim_{m \rightarrow \infty} \frac{\mu(A \cap P_m)}{\mu(P_m)} = 0.$$

Considere n fixo. Para cada $s \geq 1$, seja P_n^s a união de todos os conjuntos da forma $U_1 \cap \dots \cap U_s$ que intersectam P_n . Note que, pela observação anterior, $\mu_P((P_n^s)^c) = 0$ e então $\mu_P(P_n^s) = 1$, $\forall s \geq 1$. Assim, tomando o limite quando $s \rightarrow \infty$, tem-se que $\mu_P(U) = 1$ para todo conjunto aberto U que contém P_n . Como toda medida de probabilidade em um espaço métrico é regular, μ_P é regular e, portanto, $\mu_P(P_n) = 1$. Assim, no limite $n \rightarrow \infty$, tem-se que $\mu_P(P) = 1$ para todo $P \in \mathcal{P}_*$.

Condições 2 e 3. $P \mapsto \mu_P(E)$ é mensurável $\forall E \subset X$ mensurável e $\mu(E) = \int \mu_P(E) d\hat{\mu}(P)$.

Por construção, tem-se que, dado qualquer $A \in \mathcal{A}$, a função $P \mapsto \mu_P(A)$, onde

$\mu_P(A) = E(A, P)$, é mensurável e satisfaz

$$\begin{aligned}\mu(A) &= \int E(A, P) \, d\hat{\mu}(P) \\ &= \int \mu_P(A) \, d\hat{\mu}(P).\end{aligned}$$

Agora, note que família $\mathfrak{C} = \{B \subset X : B \text{ satisfaz as condições 2 e 3}\}$ é uma classe monótona, isto é, dados $B_1 \subset B_2 \subset \dots$ em $\mathfrak{C}$, tem-se $\bigcup_{n \in \mathbb{N}} B_n \in \mathfrak{C}$ e, dados subconjuntos $B_1 \supset B_2 \supset \dots$ em $\mathfrak{C}$, tem-se $\bigcap_{n \in \mathbb{N}} B_n \in \mathfrak{C}$. De fato, considere $(B_n)_{n \in \mathbb{N}}$ sequência crescente de conjuntos em $\mathfrak{C}$ e $B = \bigcup_{n \in \mathbb{N}} B_n$. Então

$$P \mapsto \mu_P(B) = \sup_n \mu_P(B_n)$$

é uma função mensurável e, pelo Teorema da Convergência Monótona

$$\begin{aligned}\mu(B) &= \lim_{n \rightarrow \infty} \mu(B_n) \\ &= \lim_{n \rightarrow \infty} \int \mu_P(B_n) \, d\hat{\mu} \\ &= \int \lim_{n \rightarrow \infty} \mu_P(B_n) \, d\hat{\mu} \\ &= \int \mu_P(B) \, d\hat{\mu}.\end{aligned}$$

Logo, $B \in \mathfrak{C}$. Analogamente, se B é interseção de uma sequência decrescente $(B_n)_{n \in \mathbb{N}}$ em $\mathfrak{C}$, então $P \mapsto \mu_P(B)$ é mensurável e $\mu(B) = \int \mu_P(B) \, d\hat{\mu}$. Assim, $\mathfrak{C}$ é uma classe monótona que contém $\mathcal{A}$. Uma vez que a menor classe monótona que contém uma álgebra $\mathcal{A}$ coincide com a σ -álgebra $\sigma(\mathcal{A})$ gerada por $\mathcal{A}$, segue que as propriedades 2 e 3 são válidas em toda a σ -álgebra de Borel de X . $\square$

Observação. Nas hipóteses do Teorema de Rokhlin, X é espaço métrico completo e separável. Assim, segue pela Proposição 2.2 a unicidade da desintegração obtida.

O Teorema de Rokhlin pode ser abordado sob diversas perspectivas. Por exemplo, o trabalho de Simmons em [25] fornece uma formulação (cheia de sutilezas) distinta do estudo original de Rokhlin [22] e uma generalização deste resultado para a categoria de espaços σ -finitos com morfismos absolutamente contínuos. Para abordar esse novo panorama, é necessário estabelecer algumas notações e definições.

Definição 2.5. Considere um espaço de probabilidade (X, μ) , um espaço mensurável Y , uma função mensurável $\pi : X \rightarrow Y$ e defina a medida $\hat{\mu} := \pi_*\mu$ em Y . Denomina-se sistema de medidas condicionais de μ com respeito a (X, π, Y) a coleção de medidas $(\mu_y)_{y \in Y}$ tal que

1. Para cada $y \in Y$, μ_y é uma medida em $\pi^{-1}(y)$.
2. μ_y é medida de probabilidade, para $\hat{\mu}$ -quase todo $y \in Y$.
3. As medidas $(\mu_y)_{y \in Y}$ satisfazem, $\forall A \subset X$ mensurável, a lei da probabilidade total:

$$\mu(A) = \int \mu_y(A) d\hat{\mu}(y). \quad (2.2)$$

Notação 2.6. Dado o espaço de probabilidade (X, μ) , considere $A \subset X$ mensurável, com $\mu(A) > 0$, e uma função $\psi : X \rightarrow \mathbb{R}$ integrável. Define-se:

$$P_\mu(x \in F \mid x \in A) := \mu_A(F) := \frac{\mu(F \cap A)}{\mu(A)}$$

$$E_\mu(\psi(x) \mid x \in A) := \int \psi(x) d\mu_A(x).$$

Definição 2.7. Considere (X, μ) um espaço de probabilidade topológico, Y um espaço métrico, $\pi : X \rightarrow Y$ uma função mensurável e $\hat{\mu} := \pi_*\mu$. Dado $y \in Y$ e $B(y, \epsilon)$ a bola aberta centrada em y e com raio ϵ , denomina-se medida condicional topológica de μ com respeito a (X, π, y, Y) o limite fraco*

$$\mu_y := \lim_{\epsilon \rightarrow 0} \mu_{\pi^{-1}(B(y, \epsilon))} = \lim_{\epsilon \rightarrow 0} \frac{\mu|_{\pi^{-1}(B(y, \epsilon))}}{\hat{\mu}(B(y, \epsilon))} \quad (2.3)$$

se este existe e tem suporte em $\pi^{-1}(y)$.

Definição 2.8. Um espaço mensurável $(X, \mathcal{B})$ é denominado Borel standard se existe uma métrica em X que o torna um espaço métrico completo separável, de modo que $\mathcal{B}$ seja a σ -álgebra de Borel. Neste trabalho, considera-se como espaço Borel standard o espaço de Cantor $\{0, 1\}^{\mathbb{N}}$ com a σ -álgebra de Borel associada à topologia produto (da topologia discreta em $\{0, 1\}$).

Definição 2.9. Um subconjunto de um espaço métrico completo separável X é denominado universalmente mensurável se é mensurável em relação à toda probabilidade boreliana completa

de X . Ainda, X é denominado *universalmente mensurável* se é um subconjunto mensurável de seu complemento. Como neste trabalho considera-se o espaço de Cantor, define-se que um espaço universalmente mensurável é um espaço mensurável X para o qual existe um mergulho isomórfico $i_X : X \rightarrow \{0, 1\}^{\mathbb{N}}$ tal que, para toda medida boreliana μ em $\{0, 1\}^{\mathbb{N}}$, $i_X(X)$ está no complemento de μ .

Definição 2.10. Um espaço métrico X é denominado *espaço ultra-métrico* se satisfaz a desigualdade triangular ultra-métrica:

$$d(x, z) \leq \max(d(x, y), d(y, z)), \forall x, y, z \in X.$$

Com essas definições, o enunciado do Teorema de Rokhlin em [25] é dado da seguinte forma:

Teorema 2.11 (Teorema de Rokhlin). Considere X um espaço universalmente mensurável, Y um espaço mensurável tal que existe uma aplicação mensurável injetiva de Y no espaço Borel standard. Ainda, tome $\pi : X \rightarrow Y$ uma função mensurável e μ uma probabilidade boreliana em X . Então, existe um sistema de medidas condicionais $(\mu_y)_{y \in Y}$ de μ com respeito a (X, π, Y) e esse sistema é único.

Para a demonstração do Teorema 2.11, é utilizado o seguinte teorema auxiliar:

Teorema 2.12. Considere (X, μ) um espaço de probabilidade métrico compacto, Y um espaço ultra-métrico separável localmente compacto ou uma variedade riemanniana separável. Seja $\pi : X \rightarrow Y$ mensurável. Então, para $\hat{\mu}$ -quase todo $y \in Y$, a medida condicional topológica de μ com respeito a (X, π, y, Y) existe como na Definição 2.7. Além disso, a família $(\mu_y)_{y \in Y}$ é um sistema de medidas condicionais como na Definição 2.5 (se μ_y não existe, considera-se $\mu_y = 0$).

Agora, considerando o Teorema 2.12, prova-se o Teorema 2.11:

Demonstração do Teorema 2.11. Considere $X' = \{0, 1\}^{\mathbb{N}}$. Note que $\{0, 1\}$ é compacto, pois é finito, e, deste modo, $\{0, 1\}^{\mathbb{N}}$ é compacto pelo Teorema de Tychonoff. Além disso, considere $i_X : X \rightarrow X'$ uma função inclusão, cuja existência é garantida pela mensurabilidade universal de X . Seja $\mu' = i_{X*}\mu$ uma medida de probabilidade em X' . Deste modo, (X', μ')

é um espaço de probabilidade métrico compacto. Considere, também, $i_Y : Y \rightarrow Y'$ uma aplicação injetiva mensurável, onde $Y' = \{0, 1\}^{\mathbb{N}}$ equipado com a σ -álgebra de Borel. Note que Y' é um espaço ultra-métrico (pois foi tomada a métrica discreta), separável (uma vez que $\{0, 1\}^{\mathbb{N}}$ é polonês) e localmente compacto (compacidade implica compacidade local). Como $i_Y : Y \rightarrow Y'$ é injetiva mensurável, pode-se considerar $\pi : X \rightarrow Y'$. Agora, considere o seguinte lema.

Lema 2.13. *Considere um espaço metrizável $(X, \mathcal{T})$, um subconjunto A de X , um espaço polonês Y e uma aplicação boreliana $f : A \rightarrow Y$. Então, existe uma topologia metrizável mais fina $\mathcal{T}'$ em X que gera a mesma σ -álgebra de Borel, tal que f é contínua com respeito a $\mathcal{T}'$. Ainda mais, $f : A \rightarrow Y$ admite uma extensão boreliana $g : X \rightarrow Y$.*

Demonstração. Veja [26, 3.2.3, página 92]. ◁

Com isso, tem-se que $\pi : X \rightarrow Y'$ admite uma extensão $\pi' : X' \rightarrow Y'$ Borel mensurável. Assim, são satisfeitas as hipóteses do Teorema 2.12 para (X', μ', π', Y') , ou seja: para $\widehat{\mu}'$ -quase todo $y' \in Y'$ a medida condicional topológica de μ' com respeito a (X', π', y', Y') existe e tem-se que

$$\mu'_{y'} := \lim_{\epsilon \rightarrow 0} \mu'_{(\pi')^{-1}(B(y', \epsilon))}$$

e $(\mu'_{y'})_{y' \in Y'}$ é um sistema de medidas condicionais de μ' com respeito a (X', π', Y') tal que:

1. Para cada $y' \in Y'$, $\mu'_{y'}$ é uma medida em $(\pi')^{-1}(y')$.
2. $\mu'_{y'}$ é uma medida de probabilidade para $\widehat{\mu}'$ -quase todo $y' \in Y'$.
3. As medidas $(\mu'_{y'})_{y' \in Y'}$ satisfazem, para todo $B' \subset X'$ mensurável,

$$\mu'(B') = \int \mu'_{y'}(B') d\widehat{\mu}'(y'). \quad (2.4)$$

Para cada $y \in Y$, seja

$$\mu_y = \begin{cases} (\mu'_{i_Y(y)} \mid i_X(X)) \circ (i_X^{-1})^{-1}, & \text{se } \text{supp}(\mu'_{i_Y(y)}) \subset i_X(X) \\ 0, & \text{se } \text{supp}(\mu'_{i_Y(y)}) \not\subset i_X(X). \end{cases}$$

O objetivo é mostrar que $(\mu_y)_{y \in Y}$, onde cada μ_y é definida de acordo com a equação acima, é um sistema de medidas condicionais de μ com respeito a (X, π, Y) . Para

isso, note que como π' é extensão de π , tem-se $i_Y \circ \pi = \pi' \circ i_X$ e

$$\begin{aligned} \hat{\mu} \circ i_Y^{-1} &= \mu \circ \pi^{-1} \circ i_Y^{-1} \\ &= \mu \circ (i_Y \circ \pi)^{-1} \\ &= \mu \circ (\pi' \circ i_X)^{-1} \\ &= \mu \circ i_X^{-1} \circ (\pi')^{-1} \\ &= \mu' \circ (\pi')^{-1} \\ &= \hat{\mu}'. \end{aligned}$$

Note que $\mu'_{i_Y(y)}$ é uma medida em $(\pi')^{-1}(i_Y(y))$, para cada $y \in Y$. Ainda, se $\mu'_{i_Y(y)}(X' - i_X(X)) > 0$, então $\mu_y = 0$ é uma medida em $\pi^{-1}(y)$. Se $\mu'_{i_Y(y)}(X' - i_X(X)) = 0$, tem-se que $\mu_y = (\mu'_{i_Y(y)} \mid i_X(X)) \circ (i_X^{-1})^{-1}$ tem suporte em $i_X^{-1}((\pi')^{-1}(i_Y(y)) \cap i_X(X))$ que, pela injetividade de i_Y , é igual a $\pi^{-1}(y)$. Assim, $\mu'_{i_Y(y)} = \mu_y \circ i_X^{-1}$, de modo que, se $\mu'_{i_Y(y)}$ é uma medida de probabilidade, então μ_y é uma medida de probabilidade.

Como $\mu' = \mu \circ i_X^{-1}$, tem-se $\mu'(X' - i_X(X)) = \mu \circ i_X^{-1}(X' - i_X(X)) = \mu(\emptyset) = 0$. Pela equação (2.4), segue que $\mu'_{y'}(X' - i_X(X)) = 0$. Então, tem-se que, para $\hat{\mu}$ -quase todo $y \in Y$, $\mu'_{i_Y(y)}$ é uma medida de probabilidade e $\mu'_{i_Y(y)}(X' - i_X(X)) = 0$. Assim, μ_y é uma medida em $\pi^{-1}(y)$ para todo $y \in Y$, sendo medida de probabilidade para $\hat{\mu}$ -quase todo $y \in Y$, e $\mu'_{i_Y(y)} = \mu_y \circ i_X^{-1}$, e as duas primeiras condições da Definição 2.5 são satisfeitas.

Agora, para demonstrar a condição 3: Fixe $B \subset X$ mensurável. Como i_X é um mergulho, existe um conjunto boreliano $B' \subset X'$ tal que $B = i_X^{-1}(B')$. Como, para $\hat{\mu}$ -quase todo $y \in Y$, $\mu'_{i_Y(y)} = \mu_y \circ i_X^{-1}$, pode-se escrever

$$\mu'_{i_Y(y)}(B') = \mu_y \circ i_X^{-1}(B') = \mu_y(B).$$

Assim, a função $y \mapsto \mu_y(B)$ é igual, $\hat{\mu}$ -quase sempre, à composição de i_Y com a função $y' \mapsto \mu'_{y'}(B')$, e, sendo assim, é $\hat{\mu}$ -mensurável. Ainda, note que $\mu'(B') = \mu \circ i_X^{-1}(B') = \mu(B)$, de modo que

$$\begin{aligned} \mu(B) &= \mu'(B') = \int \mu'_{y'}(B') \, d\hat{\mu}'(y') \\ &= \int \mu'_{i_Y(y)}(B') \, d\hat{\mu}(y) \\ &= \int \mu_y \circ i_X^{-1}(B') \, d\hat{\mu}(y) \\ &= \int \mu_y(B) \, d\hat{\mu}(y) \end{aligned}$$

concluindo, assim, a demonstração da condição 3. Portanto, $(\mu_y)_{y \in Y}$ é um sistema de medidas condicionais de μ com respeito a (X, π, Y) .

Ainda resta demonstrar a unicidade. Esta segue de suposições mais fracas: ao invés de assumir as propriedades de X e Y definidas no enunciado, pode-se assumir apenas que X e Y são espaços mensuráveis e que a σ -álgebra de subconjuntos mensuráveis de X é gerada de maneira enumerável (note que isto decorre da mensurabilidade universal de X , uma vez que $\{0, 1\}^{\mathbb{N}}$ é separável e qualquer subconjunto de um espaço mensurável separável é separável). Primeiramente, será provada a unicidade de cada evento individualmente e, posteriormente, a separabilidade de X permitirá a extensão à unicidade enunciada.

Lema 2.14. *Se $(\mu_y)_{y \in Y}$ é um sistema de medidas condicionais de μ relativamente a (X, π, Y) , então para todos $S \subset Y$ e $B \subset X$ mensuráveis, tem-se*

$$\mu(\pi^{-1}(S) \cap B) = \int_S \mu_y(B) d\hat{\mu}(y). \quad (2.5)$$

Demonstração. Como μ_y possui suporte em $\pi^{-1}(y)$, tem-se que

$$\mu_y(\pi^{-1}(S) \cap B) = \mathbb{1}_S(y) \mu_y(B)$$

e, portanto, a equação (2.5) segue diretamente da equação (2.2). $\triangleleft$

Corolário 2.15. *Um sistema de medidas condicionais é único no sentido que, se $(\mu_y)_{y \in Y}$ e $(\nu_y)_{y \in Y}$ são dois sistemas de medidas condicionais para uma mesma medida μ , então, para todo $B \subset X$ mensurável, $\mu_y(B) = \nu_y(B)$ para $\hat{\mu}$ -quase todo $y \in Y$.*

Demonstração. Pelo Lema 2.14 segue que

$$\int_S \mu_y(B) d\hat{\mu}(y) = \int_S \nu_y(B) d\hat{\mu}(y)$$

para todo $S \subset Y$ mensurável. Considere os conjuntos

$$S_1 = \{y \in Y : \mu_y(B) < \nu_y(B)\}$$

$$S_2 = \{y \in Y : \mu_y(B) > \nu_y(B)\}.$$

Se $\mu(S_1) > 0$, tem-se que $\int_S \mu_y(B) d\hat{\mu}(y) > \int_S \nu_y(B) d\hat{\mu}(y)$, o que contradiz a igualdade acima. Então, $\mu(S_1) = 0$. Do mesmo modo, $\mu(S_2) = 0$. Portanto $\mu_y(B) = \nu_y(B)$ para $\hat{\mu}$ -quase todo $y \in Y$. $\triangleleft$

O Corolário 2.15 garante a unicidade para eventos individuais. Com esse resultado pode-se apenas garantir que $\mu_y(B) = \nu_y(B)$ para uma coleção enumerável de subconjuntos

mensuráveis $B \subset X$. Mas, assim como feito anteriormente, é possível usar o fato de que a σ -álgebra de X é gerada de forma enumerável: Seja $(B_n)_{n \in \mathbb{N}}$ uma sequência geradora, de modo que a coleção de interseções finitas $(\cap_{n \in F} B_n)_{F \subset \mathbb{N}; \#F < \infty}$ é, também, enumerável. Assim, para $\hat{\mu}$ -quase todo $y \in Y$, $\mu_y(B) = \nu_y(B)$ para cada B nesta coleção. Por [10, página 45, 1.6.2], isso implica que para $\hat{\mu}$ -quase todo $y \in Y$, $\mu_y(B) = \nu_y(B)$ para cada $B \subset X$ mensurável, isto é, $\mu_y = \nu_y$ para $\hat{\mu}$ -quase todo $y \in Y$. $\square$

Demonstração do Teorema 2.12. Para esta demonstração, considere o lema:

Lema 2.16. *Seja $\psi : X \rightarrow \mathbb{R}$ é integrável. Então*

$$y \mapsto E_\mu(\psi(x) \mid \pi(x) = y) := \lim_{\epsilon \rightarrow 0} E_\mu(\psi(x) \mid \pi(x) \in B(y, \epsilon)) \quad (2.6)$$

onde $B(y, \epsilon)$ é a bola centrada em y e com raio ϵ , está bem definida para $\hat{\mu}$ -quase todo $y \in Y$, e é uma derivada de Radon-Nikodym de $\pi_*(\psi\mu)$ com respeito a $\hat{\mu}$.

Demonstração. Note que:

$$\begin{aligned} E_\mu(\psi(x) \mid \pi(x) \in B(y, \epsilon)) &= \int \psi(x) d\mu_{\pi^{-1}(B(y, \epsilon))}(x) \\ &= \int \psi(x) d\left(\frac{\mu(\pi^{-1}(B(y, \epsilon)) \cap x)}{\mu(\pi^{-1}(B(y, \epsilon)))}\right) \\ &= \frac{(\psi\mu) \circ \pi^{-1}(B(y, \epsilon))}{\hat{\mu}(B(y, \epsilon))}. \end{aligned}$$

Ainda, tem-se que este lema é, na verdade, o Teorema da Diferenciação de Lebesgue aplicado ao espaço Y e às medidas $\pi_*(\psi\mu)$ e $\hat{\mu}$. $\triangleleft$

Agora, partindo do Lema 2.16, é possível checar as propriedades de μ_y :

$\rightarrow$ Para $\hat{\mu}$ -quase todo $y \in Y$, $\mu_y := \lim_{\epsilon \rightarrow 0} \mu_{\pi^{-1}(B(y, \epsilon))}$.

Considere a classe de funções contínuas, $\mathcal{C}(X)$. Sabe-se que $\mathcal{C}(X)$ é separável, isto é, existe uma coleção enumerável de funções contínuas que é densa na topologia uniforme. Assim, seja Ψ um subconjunto enumerável denso de $\mathcal{C}(X)$, de modo que $N = \bigcup_{\psi \in \Psi} N_\psi$ é um conjunto nulo. Agora, note que para todo $\psi \in \mathcal{C}(X) = \overline{\Psi}$, $N_\psi \subset N$, onde $\overline{\Psi}$ é o fecho uniforme de Ψ . De fato, tome a sequência $\psi_n \in \Psi$ que converge uniformemente para ψ .

Então, para todo $y \in (Y - N)$, $E_\mu(\psi_n(x) \mid \pi(x) = y)$ existe para todo $n \in \mathbb{N}$. Então

$$\begin{aligned} E_\mu(\psi(x) \mid \pi(x) = y) &= \lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} E_\mu(\psi_n(x) \mid \pi(x) \in B(y, \epsilon)) \\ &= \lim_{n \rightarrow \infty} \lim_{\epsilon \rightarrow 0} E_\mu(\psi_n(x) \mid \pi(x) \in B(y, \epsilon)) \\ &= \lim_{n \rightarrow \infty} E_\mu(\psi_n(x) \mid \pi(x) = y) \end{aligned}$$

onde a troca de limites é justificada porque a convergência é uniforme em relação à n . Em particular, note que se $E_\mu(\psi(x) \mid \pi(x) = y)$ existe, então $y \in Y - N_\psi$. Assim, por contrapositiva, segue que $N_\psi \subset N$.

Pelo resultado anterior, fixando $y \in (Y - N)$ tem-se que $\forall \psi \in \mathcal{C}(X)$, o limite (2.6) existe. Note que

$$L_n(\psi) = E_\mu\left(\psi(x) \mid \pi(x) \in B\left(y, \frac{1}{n}\right)\right), \quad \psi \in \Psi$$

define um funcional linear não negativo e limitado em Ψ . Se

$$L(\psi) = \lim_{n \rightarrow \infty} L_n(\psi), \quad \psi \in \Psi, \quad |L_n(\psi)| < \infty$$

existe, é claro que L é também um funcional linear não negativo e limitado em Ψ . Note que este limite existe, pois é o mesmo limite de (2.6). Assim, de acordo com [6, página 77], L determina unicamente uma medida em $\pi^{-1}(Y)$, o que implica que as medidas $\mu_{\pi^{-1}(B(y, \epsilon))}$ convergem na topologia fraca* para uma medida limite

$$\mu_y := \lim_{\epsilon \rightarrow 0} \mu_{\pi^{-1}(B(y, \epsilon))}.$$

Note que μ_y é uma medida de probabilidade, uma vez que X é compacto, entretanto, ainda não foi provado que μ_y é suportada inteiramente em $\pi^{-1}(y)$.

→ A coleção de medidas $(\mu_y)_{y \in Y - N}$ satisfaz a lei da probabilidade total.

Note que, pelo Lema 2.16, se ψ é qualquer função contínua, integrando em ambos os lados da equação (2.3), obtém-se

$$\begin{aligned} \int \psi \, d\mu_y &= \int \psi \, d\left(\lim_{\epsilon \rightarrow 0} \mu_{\pi^{-1}(B(y, \epsilon))}\right) \\ &= \lim_{\epsilon \rightarrow 0} \int \psi \, d\mu_{\pi^{-1}(B(y, \epsilon))} \\ &= \lim_{\epsilon \rightarrow 0} E_\mu(\psi(x) \mid \pi(x) \in B(y, \epsilon)) \\ &= E_\mu(\psi(x) \mid \pi(x) = y) \end{aligned}$$

de modo que a função $y \mapsto \int \psi d\mu_y$ está bem definida para $\hat{\mu}$ -quase todo $y \in Y$, é $\hat{\mu}$ -mensurável e é uma derivada de Radon-Nikodym de $\pi_*(\psi\mu)$ com respeito a $\hat{\mu}$. De acordo com a definição da derivada Radon-Nikodym, isso significa que, para qualquer conjunto mensurável $S \subset Y$,

$$\int_{\pi^{-1}(S)} \psi d\mu = (\pi_*(\psi\mu))(S) = \int_S \left[\int \psi d\mu_y \right] d\hat{\mu}(y) \quad (2.7)$$

onde o lado direito é tomado para incluir a suposição de que a função $y \mapsto \int \psi d\mu_y$ é mensurável. Agora, o objetivo é generalizar (2.7) para todas as ψ mensuráveis limitadas.

Seja $BM(X)$ o conjunto de funções mensuráveis limitadas de X a $\mathbb{R}$. Sabe-se que $\mathcal{C}(X) \subseteq BM(X)$ é um subconjunto denso, se $BM(X)$ está munido com a topologia da convergência monótona [26, Teorema de Lebesgue-Hausdorff, página 91]. Então, basta mostrar que o conjunto de todas as $\psi \in BM(X)$ que satisfazem a equação (2.7) é fechado na topologia da convergência monótona. Denote por A o conjunto de todas as $\psi \in BM(X)$ que satisfazem a equação (2.7). Considere $(\psi_n)_{n \in \mathbb{N}}$, com $\psi_n \in A$, tal que $\psi_n \rightarrow \psi$. Como resultado de Topologia Geral, sabe-se que $\psi \in \overline{A}$ (e a recíproca é verdadeira, uma vez que o conjunto é metrizável). Note que se $\psi_n \rightarrow \psi$ monotonicamente, então $\int \psi_n d\mu_y \rightarrow \int \psi d\mu_y$ monotonicamente, pelo Teorema da Convergência Monótona. Assim, se $y \mapsto \int \psi_n d\mu_y$ é $\hat{\mu}$ -mensurável para todo $n \in \mathbb{N}$, então $y \mapsto \int \psi d\mu_y$ é, também, $\hat{\mu}$ -mensurável. Portanto, a equação (2.7) é válida para todas as funções em $\overline{A} \implies A = \overline{A}$, ou seja, A é fechado. Assim, a equação (2.7) é verdadeira para todas as ψ mensuráveis limitadas. Em particular, se $\psi = \mathbb{1}_B$, então a equação (2.7) é simplificada para a equação (2.5). Se, além disso, $S = Y$, então a equação (2.5) é simplificada para a equação (2.2).

→ Para $\hat{\mu}$ -quase todo $y \in Y$, a medida μ_y está inteiramente suportada em $\pi^{-1}(y)$.

Note que dizer que μ_y está suportada inteiramente em $\pi^{-1}(y)$ é o mesmo que dizer que $\pi_*\mu_y = \delta_y$, onde δ_y é a medida de Dirac em y . Tomando a equação (2.5) considerando $B = \pi^{-1}(C)$, onde C é um subconjunto enumerável de Y , tem-se:

$$\mu(\pi^{-1}(S) \cap \pi^{-1}(C)) = \int_S \mu_y \circ \pi^{-1}(C) d\hat{\mu}(y).$$

Mas,

$$\mu(\pi^{-1}(S) \cap \pi^{-1}(C)) = \int \mathbb{1}_{(\pi^{-1}(S) \cap \pi^{-1}(C))} d\mu = \int_S \mathbb{1}_C(y) d\hat{\mu}(y)$$

logo

$$\int_S \mathbb{1}_C(y) d\hat{\mu}(y) = \int_S \mu_y \circ \pi^{-1}(C) d\hat{\mu}(y)$$

para todos os conjuntos mensuráveis $C, S \subset Y$. Por um argumento semelhante ao utilizado na demonstração do Corolário 2.15, tem-se que para cada conjunto mensurável $C \subset Y$ e para

$\hat{\mu}$ -quase todo $y \in Y$

$$\delta_y(C) = \mathbb{1}_C(y) = \mu_y \circ \pi^{-1}(C). \quad (2.8)$$

Tomando uma coleção densa enumerável de subconjuntos mensuráveis de Y , pode-se inverter a ordem dos quantificadores, notando que a coleção de todos os conjuntos mensuráveis para os quais (2.8) vale é fechada sob convergência monótona. Agora, tome $C = \{y\}$. Assim, para $\hat{\mu}$ -quase todo $y \in Y$, $\mu_y(\pi^{-1}(y)) = 1$. $\square$

Até aqui, atribuiu-se uma interpretação probabilística ao Teorema de Rokhlin. Em seu trabalho, Simmons apresenta uma generalização desse resultado para uma categoria de espaços de medida σ -finitos com morfismos absolutamente contínuos, tendo em vista que existem casos de medidas σ -finitas das quais as medidas de probabilidade surgem naturalmente. Essa perspectiva será abordada agora. Primeiramente, considere uma versão generalizada da Definição 2.5 e da Definição 2.7:

Definição 2.17. *Sejam (X, μ) e (Y, ν) espaços de medidas, e $\pi : X \rightarrow Y$ uma função mensurável. Um sistema de medidas concionais de μ com respeito a (X, π, Y, ν) é uma coleção de medidas $(\nu_y)_{y \in Y}$ tal que*

1. *Para cada $y \in Y$, ν_y é uma medida em $\pi^{-1}(y)$.*
2. *Para cada $B \subset X$ mensurável, as medidas $(\nu_y)_{y \in Y}$ satisfazem a lei da probabilidade total generalizada:*

$$\mu(B) = \int \nu_y(B) d\nu(y). \quad (2.9)$$

Definição 2.18. *Considere (X, μ) um espaço de medida topológico, (Y, ν) um espaço de medida métrico e $\pi : X \rightarrow Y$ uma função mensurável. Dado $y \in Y$ e $B(y, \epsilon)$ a bola aberta centrada em y e com raio ϵ , a medida condicional topológica de μ com respeito a (X, π, y, Y, ν) é o limite fraco**

$$\nu_y := \lim_{\epsilon \rightarrow 0} \frac{\mu|_{\pi^{-1}(B(y, \epsilon))}}{\nu(B(y, \epsilon))} \quad (2.10)$$

se este existe, é localmente finito, e possui suporte $\pi^{-1}(y)$.

É válido ressaltar que, em ambas as definições, tem-se ν_y no lugar de μ_y , de modo que se $\nu = \hat{\mu}$ e μ é uma medida de probabilidade, essas definições são equivalentes às originais. Com isso, é possível enunciar os seguintes teoremas:

Teorema 2.19 (Teorema de Rokhlin generalizado). *Considere (X, μ) um espaço de medida σ -finito universalmente mensurável, (Y, ν) um espaço de medida σ -finito tal que existe uma função mensurável injetiva de Y em $\{0, 1\}^{\mathbb{N}}$. Ainda, seja $\pi : X \rightarrow Y$ uma aplicação mensurável e considere a medida $\hat{\mu} = \pi_*\mu$. Suponha que $\hat{\mu} \ll \nu$. Então, existe um sistema de medidas condicionais $(\nu_y)_{y \in Y}$ de μ com respeito a (X, π, Y, ν) . Para ν -quase todo $y \in Y$, ν_y é uma medida σ -finita. Além disso, as medidas condicionais são únicas, no sentido que se $(\gamma_y)_{y \in Y}$ é qualquer outro sistema de medidas condicionais de μ com respeito a (X, π, Y, ν) , tem-se $\nu_y = \gamma_y$ para ν -quase todo $y \in Y$.*

Teorema 2.20. *Sejam (X, μ) e (Y, ν) espaços de medida métricos, separáveis, localmente finitos e localmente compactos. Assuma que Y é um espaço ultra-métrico ou uma variedade riemanniana. Considere a aplicação mensurável $\pi : X \rightarrow Y$ e a medida $\hat{\mu} = \pi_*\mu$. Suponha que $\hat{\mu} \ll \nu$. Então, para ν -quase todo $y \in Y$, a medida condicional topológica de μ com respeito a (X, π, y, Y, ν) existe, de acordo com a Definição 2.18. Além disso, a coleção de medidas $(\nu_y)_{y \in Y}$ de μ é um sistema de medidas condicionais, como na Definição 2.17.*

Demonstração do Teorema 2.19. Uma vez que (X, μ) é um espaço de medida σ -finito universalmente mensurável, considere $X = \coprod_{n \in \mathbb{N}} A_n$, onde cada A_n possui medida finita e é universalmente mensurável. Para cada $n \in \mathbb{N}$, considere $i_n : A_n \rightarrow \{0, 1\}^{\mathbb{N}}$ um mergulho isomórfico, de modo que $i_n(A_n)$ é universalmente mensurável. “Colando” adequadamente as i_n , obtém-se um mergulho isomórfico $i_X : X \rightarrow X' := \mathbb{N} \times \{0, 1\}^{\mathbb{N}}$ tal que $i_X(X)$ é universalmente mensurável. Tomando $\mu' = i_{X*}\mu$, tem-se que (X', μ') é um espaço de medida ultra-métrico, separável, localmente compacto e localmente finito. Da mesma forma, é possível construir uma função mensurável injetiva $i_Y : Y \rightarrow Y' := \mathbb{N} \times \{0, 1\}^{\mathbb{N}}$ tal que $\nu' := i_{Y*}\nu$ é localmente finita. Assim, tem-se que (Y', ν') é um espaço de medida ultra-métrico, separável, localmente finito e localmente compacto. Ainda tem-se, por [26, 3.2.3, página 92], que a aplicação π pode ser estendida a uma função Borel mensurável $\pi' : X' \rightarrow Y'$. Com isso, tem-se as hipóteses do Teorema 2.20 para $(X', \mu', \pi', Y', \nu')$. Então, considere $(\nu'_{y'})_{y' \in Y'}$ um sistema de medidas condicionais de μ' com respeito a (X', π', Y', ν') .

Note que a mesma demonstração da existência do sistema de medidas condicionais realizada para o Teorema 2.11 é válida para este teorema, substituindo $\hat{\mu}$ por ν , $\hat{\mu}'$ por ν' , μ_y por ν_y e “probabilidade” por “ σ -finito”, lembrando que uma medida localmente finita em um espaço métrico separável localmente compacto é σ -finita, de modo que as medidas oriundas da Definição 2.18 são, necessariamente, σ -finitas.

Para provar a unicidade, assumamos que μ é finita. Neste caso,

$$\int \nu_y(X) d\nu(y) = \mu(X) < \infty$$

então $\nu_y(X) < \infty$ para ν -quase todo $y \in Y$. Deste modo, a prova de unicidade dada para o Teorema 2.11 é válida tomando as mesmas substituições citadas acima. Agora, se μ é uma medida σ -finita, como $X = \coprod_{n \in \mathbb{N}} A_n$, para todo sistema de medidas condicionais $(\nu_y)_{y \in Y}$ de μ relativamente a (X, π, ν) , para cada $n \in \mathbb{N}$, tem-se que $(\nu_y|_{A_n})_{y \in Y}$ é um sistema de medidas condicionais de $\mu|_{A_n}$ relativamente a (A_n, π, ν) . Assim, se $(\nu_y)_{y \in Y}$ e $(\gamma_y)_{y \in Y}$ são dois sistemas de medidas condicionais de μ com respeito a (X, π, ν) , então para cada $n \in \mathbb{N}$, $(\nu_y|_{A_n})_{y \in Y}$ e $(\gamma_y|_{A_n})_{y \in Y}$ são sistemas de medidas condicionais de $\mu|_{A_n}$ com respeito a (A_n, π, ν) . A unicidade para o caso finito implica que $\nu_y|_{A_n} = \gamma_y|_{A_n}$, para ν -quase todo $y \in Y$. Tomando $y \in Y$ fixo e variando $n \in \mathbb{N}$ tem-se que, para ν -quase todo $y \in Y$, $\nu_y|_{A_n} = \gamma_y|_{A_n}$ para todo $n \in \mathbb{N}$ (de fato, a enumerabilidade de $\mathbb{N}$ justifica esta inversão de quantificadores). E, para cada $y \in Y$, tem-se

$$\nu_y = \sum_{n \in \mathbb{N}} \nu_y|_{A_n} = \sum_{n \in \mathbb{N}} \gamma_y|_{A_n} = \gamma_y.$$

Assim, segue que $\nu_y = \gamma_y$ para ν -quase todo $y \in Y$. □

Agora, falta demonstrar o Teorema 2.20. Note que existem duas generalizações feitas a partir do Teorema 2.12: $\nu = \hat{\mu}$ para $\hat{\mu} \ll \nu$ e espaços compactos para espaços localmente compactos. Em vista disso, a demonstração será realizada considerando dois casos: X como espaço compacto e o caso geral (X como espaço localmente compacto).

Demonstração do Teorema 2.20.

Caso 1: X é compacto. Neste caso, como μ é localmente finita por hipótese, tem-se que μ é finita. Com isso, as hipóteses do Teorema 2.12 são satisfeitas para a medida normalizada

$$\mu_X := \frac{\mu}{\mu(X)}.$$

A primeira afirmação do Teorema 2.12 é que para $\hat{\mu}$ -quase todo $y \in Y$, o limite fraco* dado pela equação (2.3) existe e está suportado em $\pi^{-1}(y)$. Ao comparar as equações (2.3) e (2.10), nota-se que estas se diferem de um fator

$$f(y) = \lim_{\epsilon \rightarrow 0} \frac{\hat{\mu}(B(y, \epsilon))}{\nu(B(y, \epsilon))}. \quad (2.11)$$

Pelo Teorema da Diferenciação de Lebesgue, este limite existe e é finito para ν -quase todo $y \in Y$. Ainda mais, a função f assim definida é uma derivada de Radon-Nikodym de $\hat{\mu}$ relativamente a ν . Deste modo, segue que o limite (2.10) existe e é igual a $f(y)\mu_y$, para $\hat{\mu}$ -quase todo $y \in Y$. Agora, é necessário mostrar que este existe para ν -quase todo $y \in Y$: Seja F o conjunto de todos $y \in Y$ tal que a medida condicional μ_y exista de acordo com a Definição 2.8. Então $N := (X - F)$ é um conjunto $\hat{\mu}$ -nulo, e

$$\int_N f(y) d\nu(y) = \int_N \lim_{\epsilon \rightarrow 0} \frac{\hat{\mu}(B(y, \epsilon))}{\nu(B(y, \epsilon))} d\nu(y) = \hat{\mu}(N) = 0.$$

Assim $f(y) = 0$ para ν -quase todo $y \in N$. Como a equação (2.3) é limitada e $f(y) = 0$ tem-se que o limite (2.10) é zero para ν -quase todo $y \in N$. Em conclusão, tem-se que para ν -quase todo $y \in Y$, o limite (2.10) existe e é dado por:

$$\nu_y = \begin{cases} f(y)\mu_y, & \text{se } y \notin N \\ 0, & \text{se } y \in N \end{cases} \quad (2.12)$$

onde μ_y é uma medida condicional.

Para completar, falta mostrar que para ν -quase todo $y \in Y$, ν_y é finita e com suporte em $\pi^{-1}(y)$ e que a equação (2.9) é válida para todo conjunto mensurável $B \subset X$. Considere $y \in Y$ e note que, se $y \in N$, então $\nu_y = 0$ é trivialmente finito e suportado em $\pi^{-1}(y)$. Por outro lado, se $y \in F$, então μ_y existe e satisfaz a Definição 2.8, ou seja, é uma medida de probabilidade e está inteiramente suportada em $\pi^{-1}(y)$. Agora, se $f(y) < \infty$, então ν_y é, também, finita e inteiramente suportada em $\pi^{-1}(y)$. Como $f(y) < \infty$ para ν -quase todo $y \in Y$, não há mais nada o que mostrar. Note que o Teorema 2.12 implica a equação (2.2), de modo que basta mostrar que o lado direito das equações (2.2) e (2.9) são iguais. De fato:

$$\begin{aligned} \int \nu_y(B) d\nu(y) &= \int_F f(y)\mu_y(B) d\nu(y) \\ &= \int_F \left(\lim_{\epsilon \rightarrow 0} \frac{\hat{\mu}(B(y, \epsilon))}{\nu(B(y, \epsilon))} \right) \mu_y(B) d\nu(y) \\ &= \int \mu_y(B) d\hat{\mu}(y). \end{aligned}$$

◁

Caso 2: geral. Considere o seguinte lema:

Lema 2.21. *Se X é um espaço métrico separável localmente compacto, então existe uma sequência crescente de conjuntos compactos $K_n \subset X$ tal que $X = \bigcup_{n \in \mathbb{N}} K_n$, e tal que para qualquer conjunto compacto $K \subset X$, existe N tal que $K \subset K_N$.*

Demonstração. Uma vez que X é localmente compacto, existe uma cobertura aberta de conjuntos abertos relativamente compactos. Como X é métrico separável, segue que X é Lindelöf, e assim existe uma subcobertura enumerável $(U_j)_{j \in \mathbb{N}}$. Seja $K_n := \overline{\bigcup_{j < n} U_j}$, de modo que K_n é compacto. Então $X \subset \bigcup_{j \in \mathbb{N}} U_j \subset \bigcup_{n \in \mathbb{N}} K_n \subset X$. Finalmente, suponha que $K \subset X$ é compacto. Então a cobertura $(U_j)_{j \in \mathbb{N}}$ possui subcobertura finita $(U_j)_{j < N}$. Assim, $K \subset \bigcup_{j < N} U_j \subset \overline{\bigcup_{j < N} U_j} = K_N$. ◀

Seja $(K_n)_{n \in \mathbb{N}}$ como no Lema 2.21. Para cada $n \in \mathbb{N}$, a compacidade de K_n implica que $(K_n, \mu|_{K_n}, \pi, Y, \nu)$ satisfaz as condições do Teorema 2.20, no caso já demonstrado. Denote a medida condicional topológica de $\mu|_{K_n}$ com respeito a (K_n, π, y, Y, ν) por $\nu_{n,y}$, se ela existe e satisfaz as condições da Definição 2.18.

Afirmção 1. *Dado $y \in Y$ fixo, suponha que, para cada $n \in \mathbb{N}$, a medida condicional $\nu_{n,y}$ existe e satisfaz as condições da Definição 2.18. Então, a sequência $(\nu_{n,y})_{n \in \mathbb{N}}$ é monótona crescente, e a medida $\nu_y := \lim_{n \rightarrow \infty} \nu_{n,y}$ é a medida condicional topológica de μ com respeito à (X, π, y, Y, ν) .*

Demonstração. O fato de $(\nu_{n,y})_{n \in \mathbb{N}}$ ser uma sequência monótona crescente segue diretamente do fato de $(K_n)_{n \in \mathbb{N}}$ ser uma sequência de conjuntos monótona crescente. Antes de prosseguir, considere o seguinte resultado:

Lema 2.22. *Seja X um espaço de medida e (μ_n) uma sequência monotonicamente crescente de medidas em X . Então, $\mu(A) := \lim_{n \rightarrow \infty} \mu_n(A)$ é uma medida. Ainda mais, para toda função mensurável não negativa $\psi : X \rightarrow \mathbb{R}$, tem-se*

$$\int \psi \, d\mu = \lim_{n \rightarrow \infty} \int \psi \, d\mu_n. \quad (2.13)$$

Demonstração. μ é medida, pois:

1. $\mu(\emptyset) = \lim_{n \rightarrow \infty} \mu_n(\emptyset) = 0$;
2. Seja $A = \coprod_{m \in \mathbb{N}} A_m$. Então

$$\begin{aligned} \mu(A) &= \lim_{n \rightarrow \infty} \mu_n(A) \\ &= \lim_{n \rightarrow \infty} \sum_{m \in \mathbb{N}} \mu_n(A_m) \\ &= \sum_{m \in \mathbb{N}} \lim_{n \rightarrow \infty} \mu_n(A_m) \\ &= \sum_{m \in \mathbb{N}} \mu(A_m). \end{aligned}$$

Note que, se ψ é uma função indicadora, então a equação (2.13) segue da definição. Se ψ é uma função simples, a equação (2.13) também é claramente estabelecida. Para uma ψ qualquer, considere uma sequência monotonicamente crescente de funções simples, (ψ_m) , que converge para ψ . Então:

$$\begin{aligned}
\int \psi d\mu &= \lim_{m \rightarrow \infty} \int \psi_m d\mu \\
&= \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \int \psi_m d\mu_n \\
&\leq \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \int \psi d\mu_n \\
&= \lim_{n \rightarrow \infty} \int \psi d\mu_n.
\end{aligned}$$

Ainda, como $\mu_n \leq \mu$, tem-se $\int \psi d\mu_n \leq \int \psi d\mu$ e, tomando o limite quando $n \rightarrow \infty$, obtém-se a desigualdade oposta. Assim, segue (2.13).

◁

Voltando a demonstração da Afirmação 1, note que $\nu_y(A) := \lim_{n \rightarrow \infty} \nu_{n,y}(A)$ é uma medida, mas é necessário mostrar que é localmente finita e que é igual ao limite fraco* (2.10). Para cada $x \in X$, considere uma vizinhança U de x que seja relativamente compacta. Por [10, 7.1.8, página 199], existe uma função $\psi \in C_c^+(X)$ com $\mathbb{1}_U \leq \psi$. Então,

$$\nu_y(U) \leq \int \psi d\nu_y = \lim_{\epsilon \rightarrow 0} \frac{\int \psi d\mu|_{\pi^{-1}(B(y,\epsilon))}}{\nu(B(y,\epsilon))} < \infty. \quad (2.14)$$

De fato, tome N grande o suficiente, de modo que $\text{supp}(\psi) \subseteq K_N$. Então para todo $n \geq N$, o lado direito desta equação pode ser calculado inteiramente em K_n e, por definição,

$$\lim_{\epsilon \rightarrow 0} \frac{\int \psi d\mu|_{\pi^{-1}(B(y,\epsilon))}}{\nu(B(y,\epsilon))} = \int \psi d\nu_{n,y} < \infty.$$

Então, para todo $n \geq N$, tem-se que $\int \psi d\nu_{n,y}$ é independente de n . Tomando o limite quando $n \rightarrow \infty$, a equação (2.14) fica provada devido a (2.13), e, assim, tem-se que ν_y está de acordo com (2.10) e é localmente finita. Agora, considere $y \in Y$ fixo e $\psi := \mathbb{1}_{X - \pi^{-1}(y)}$, de modo que, considerando o resultado do Lema 2.22, tem-se:

$$\begin{aligned}
\nu_y(X - \pi^{-1}(y)) &= \int \psi d\nu_y \\
&= \lim_{n \rightarrow \infty} \int \psi d\nu_{n,y} \\
&= \lim_{n \rightarrow \infty} \nu_{n,y}(X - \pi^{-1}(y)) \\
&= 0.
\end{aligned}$$

Assim, segue que ν_y está inteiramente suportada em $\pi^{-1}(y)$, o que conclui a demonstração da Afirmação 1. ◀

As hipóteses da Afirmação 1 são satisfeitas para ν -quase todo $y \in Y$, de modo que a primeira afirmação do Teorema 2.20 fica demonstrada. Agora, falta demonstrar que a equação (2.9) é satisfeita para todo $B \subset X$ mensurável. Para isso, dado $y \in Y$, considere $\psi := \mathbb{1}_B$, de modo que $\nu_y(B) = \lim_{n \rightarrow \infty} \nu_{n,y}(B)$. Integrando com respeito a ν , tem-se:

$$\begin{aligned} \int \nu_{n,y}(B) \, d\nu(y) &= \int \lim_{n \rightarrow \infty} \nu_{n,y}(B) \, d\nu(y) \\ &= \lim_{n \rightarrow \infty} \int \nu_{n,y}(B) \, d\nu(y) \\ &= \lim_{n \rightarrow \infty} \mu(B \cap K_n) \\ &= \mu(B) \end{aligned}$$

concluindo, assim, a demonstração do Teorema 2.20.

◁

□

CAPÍTULO 3

Desintegração em diferentes contextos

As diferentes formulações do Teorema de Rokhlin evidenciam que o conceito de desintegração de medidas pode ser abordado em diferentes contextos. Por exemplo, considere o seguinte teorema de desintegração [3, Teorema 5.3.1]:

Teorema 3.1. *Sejam X e Z espaços métricos separáveis Radon, $\gamma \in \mathcal{P}(Z)$, $\pi : Z \rightarrow X$ uma aplicação boreliana e $\mu \in \mathcal{P}(X)$ dada por $\mu = \pi_*\gamma$. Então existe uma família de probabilidades borelianas $\{\gamma_x : x \in X\} \subset \mathcal{P}(Z)$, μ -quase sempre determinada, tal que, para μ -quase todo $x \in X$,*

$$\gamma_x(Z - \pi^{-1}(x)) = 0 \quad (3.1)$$

e, para qualquer função boreliana $\varphi : Z \rightarrow [0, \infty]$,

$$\int_Z \varphi(z) d\gamma(z) = \int_X \left(\int_{\pi^{-1}(x)} \varphi(z) d\gamma_x(z) \right) d\mu(x). \quad (3.2)$$

Corolário 3.2. *Em particular, se $Z := X \times Y$, tem-se $\gamma \in \mathcal{P}(X \times Y)$ e $\mu = \pi_*\gamma$ obtido via projeção no primeiro componente. Com isso, pode-se identificar canonicamente cada fibra $\pi^{-1}(x)$ com Y e encontrar uma família de probabilidades borelianas $\{\gamma_x : x \in X\} \subset \mathcal{P}(Y)$ que é μ -quase sempre unicamente determinada, e tal que $\gamma := \int_X \gamma_x d\mu(x)$.*

Esta interessante perspectiva da desintegração de medidas pode ser abordada no contexto da Teoria de Transporte Ótimo: é possível considerar um plano γ , cujo primeiro marginal seja $\mu \in \mathcal{P}(X)$, como um tipo de mapa de transporte “estocástico” $x \mapsto \gamma_x$, sendo aplicável ao Problema de Monge-Kantorovich, e correspondendo a um tipo de mapa de transporte “determinístico” quando γ_x for uma medida de Dirac. Além disso, a desintegração neste contexto tem sido aplicada em exemplos relacionados à dinâmica hiperbólica, para a construção de espaços funcionais que permitem o estudo de diversas propriedades de determinados sistemas (veja, por exemplo, [15] e [16]). Neste capítulo, a desintegração de medidas é explorada dentro deste cenário.

3.1 Desintegração e Transporte Ótimo

A desintegração de planos de transporte pode ser utilizada para introduzir uma perspectiva diferente para os problemas de otimização propostos por Monge e Kantorovich, como abordado em [4] e [18]. Sejam $X, Y \subset \mathbb{R}^N$ conjuntos compactos. Considere os espaços X e $X \times Y$, uma medida $\gamma \in \mathcal{P}(X \times Y)$ cujo primeiro marginal seja μ , e o Corolário 3.2. Neste contexto, denota-se γ por $\mu \otimes \gamma_x$ (com essa notação procura-se evidenciar o primeiro marginal de γ e a desintegração considerada). Assim, define-se um *mapa de desintegração* pela função

$$\begin{aligned} f : X &\rightarrow (\mathcal{P}(Y), W_1) \\ x &\mapsto \gamma_x \end{aligned}$$

tal que $\gamma = \mu \otimes f_{(x)}$ e onde W_1 é a distância 1-Wasserstein. A partir desta definição, tem-se:

Lema 3.3. *Uma aplicação $f : X \rightarrow (\mathcal{P}(Y), W_1)$ é um mapa de desintegração se, e somente se, é Borel.*

Demonstração. ($\Rightarrow$) Pelo Teorema de Ascoli-Arzelá, o espaço $Lip_1(Y)$ é compacto relativamente à convergência uniforme [3, Proposição 3.3.1]. Considere $D \subset Lip_1(Y)$ denso enumerável e $\varphi \in Lip_1(Y)$. No caso particular em que medidas $\nu_1, \nu_2 \in \mathcal{P}(Y)$ possuem suporte limitado, é possível usar a fórmula da dualidade (Teorema 1.14) para obter

$$W_1(\nu_1, \nu_2) = \sup_{\varphi \in Lip_1(Y)} \left\{ \int_Y \varphi d(\nu_1 - \nu_2) \right\} = \sup_{\varphi \in D} \int_Y \varphi d(\nu_1 - \nu_2).$$

Para detalhes, consulte [30, Capítulo 6] ou [3, Seção 7.1]. Como $x \mapsto f_{(x)}$ é boreliana, para todo $\varphi \in Lip_1(Y)$, tem-se que

$$\begin{aligned} \psi_\varphi : X &\rightarrow \mathbb{R} \\ x &\mapsto \int_Y \varphi d(\nu - f_{(x)}) \end{aligned}$$

é, também, boreliana. Com essas condições, para que f seja Borel basta que

$$f^{-1}(B(\nu, r)) = \bigcap_{\varphi \in D} \psi_\varphi^{-1}((-r, r)) := A$$

onde $B(\nu, r)$ é a bola aberta em $(\mathcal{P}(Y), W_1)$ centrada em ν e com raio r . De fato, se $x \in A$, $|\psi_\varphi(x)| < r$ para todo $\varphi \in D$ e, pela definição de ψ_φ , tem-se que $W_1(\nu, f_{(x)}) < r$, de modo que $f_{(x)} \in B(\nu, r)$. Do mesmo modo, se $f_{(x)} \in B(\nu, r)$, então $W_1(\nu, f_{(x)}) < r$ e, pela fórmula da dualidade, $|\psi_\varphi(x)| := \left| \int_Y \varphi d(\nu - f_{(x)}) \right| < r$ para todo $\varphi \in D$, de modo que $x \in A$.

($\Leftarrow$) Considere um conjunto $A \subset Y$ aberto. Note que

$$f_{(x)}(A) = \int_Y \mathbb{1}_A(y) df_{(x)}.$$

Para uma função φ semicontínua inferiormente sobre Y , considere I_φ dada por:

$$\begin{aligned} I_\varphi : (\mathcal{P}(Y), W_1) &\rightarrow \mathbb{R} \\ \lambda &\mapsto \int_Y \varphi(y) d\lambda. \end{aligned} \quad (3.3)$$

Uma vez que W_1 metriza a topologia fraca* de $\mathcal{P}(Y)$ [3, Capítulo 7], tem-se que a função I_φ é semicontínua inferiormente. Para todo $x \in X$, pode-se escrever $\int_Y \varphi(y) df_{(x)} = I_\varphi(f_{(x)})$. Por hipótese, f é boreliana, e então $f_{(\cdot)}(A) : X \rightarrow \mathbb{R}$ é composição de uma aplicação semicontínua inferiormente e uma aplicação boreliana, portanto Borel. Logo, f é mapa de desintegração. $\square$

Pensando no problema de otimização formulado por Monge, dado um mapa de transporte $T : X \rightarrow Y$, uma medida $\mu \in \mathcal{P}(X)$ e a medida $\nu = T_*\mu \in \mathcal{P}(Y)$, o mapa de desintegração é da forma $x \mapsto \delta_{T(x)}$, onde $\delta_{T(x)}$ é uma medida de Dirac. É possível mostrar que existe uma relação entre medidas sobre o espaço $(\mathcal{P}(Y), W_1)$, dadas pelo *push forward* de μ por mapas de desintegração, e o segundo marginal dos planos de transporte induzidos por mapas de transportes distintos:

Lema 3.4. *Considere $T : X \rightarrow Y$ e $S : X \rightarrow Y$ mapas de transporte, uma medida $\mu \in \mathcal{P}(X)$ e as funções f e g definidas por:*

$$\begin{aligned} f : X &\rightarrow (\mathcal{P}(Y), W_1) \\ x &\mapsto \delta_{T(x)} \end{aligned}$$

$$\begin{aligned} g : X &\rightarrow (\mathcal{P}(Y), W_1) \\ x &\mapsto \delta_{S(x)}. \end{aligned}$$

Tem-se que $T_\mu = S_*\mu$ se, e somente se, $f_*\mu = g_*\mu$.*

Demonstração. ($\Rightarrow$) Considere uma função $\varphi(y) = \psi(\delta_y)$ para $\psi \in \mathcal{C}((\mathcal{P}(Y), W_1))$ qualquer. Então,

$$\int_Y \psi d(f_*\mu) = \int_X \psi(f_{(x)}) d\mu = \int_X \psi(\delta_{T(x)}) d\mu = \int_X \varphi(T(x)) d\mu = \int_X \varphi(S(x)) d\mu = \int_Y \psi d(g_*\mu).$$

Pela arbitrariedade de ψ , segue que $f_*\mu = g_*\mu$.

($\Leftarrow$) Seja $\varphi \in \mathcal{C}(Y)$. Considere a função I_φ definida pela equação (3.3). Tem-se que

$$\begin{aligned} \int_{\mathcal{P}(Y)} I_\varphi(\lambda) d(f_*\mu) &= \int_{\mathcal{P}(Y)} I_\varphi(\lambda) d(g_*\mu) \iff \int_X I_\varphi(f(x)) d\mu = \int_X I_\varphi(g(x)) d\mu \\ &\iff \int_X \left(\int_Y \varphi(y) df_{(x)} \right) d\mu = \int_X \left(\int_Y \varphi(y) dg_{(x)} \right) d\mu. \end{aligned}$$

Uma vez que

$$\int_Y \varphi(y) df_{(x)} = \varphi(T(x))$$

e

$$\int_Y \varphi(y) dg_{(x)} = \varphi(S(x))$$

a última equação pode ser escrita como

$$\int_X \varphi(T(x)) d\mu = \int_X \varphi(S(x)) d\mu.$$

Pela arbitrariedade de φ , infere-se que $T_*\mu = S_*\mu$. □

Corolário 3.5. *Considere $T : X \rightarrow Y$ e $S : X \rightarrow Y$ mapas de transporte, uma medida $\mu \in \mathcal{P}(X)$, os mapas de desintegração f e g definidos como no Lema 3.4 e as medidas $\gamma = \mu \otimes f_{(x)}$ e $\eta = \mu \otimes g_{(x)}$. Então $f_*\mu = g_*\mu$ se, e somente se, $\pi_{y*}\gamma = \pi_{y*}\eta$, onde π_y é a projeção na segunda componente, $(x, y) \mapsto y$.*

Dados $\gamma, \eta \in \Pi(\mu, \nu)$, com $\gamma = \mu \otimes f_{(x)}$ e $\eta = \mu \otimes g_{(x)}$, diz-se que γ é equivalente a η por desintegração, e denota-se $\gamma \approx \eta$, se $f_*\mu = g_*\mu$. Com esta equivalência, é possível definir uma espécie de classe de transporte, da seguinte maneira:

Definição 3.6. *Dado $\gamma \in \Pi(\mu, \nu)$ com $\gamma = \mu \otimes f_{(x)}$, a classe de transporte de γ é definida como a classe de equivalência $[\gamma] = \{\eta = \mu \otimes g_{(x)} : g_*\mu = f_*\mu\}$.*

Assim, tem-se que todos os planos de transporte induzidos por mapas de transporte pertencem à mesma classe de transporte. Ainda mais, considere a seguinte proposição:

Proposição 3.7. *Considere um mapa de transporte $T : X \rightarrow Y$ tal que, para uma medida $\mu \in \mathcal{P}(X)$ não atômica, $T_*\mu = \nu$ e $\gamma = \mu \otimes \delta_{T(x)}$. Se $\eta \in [\gamma]$, então existe um mapa de transporte $S : X \rightarrow Y$ tal que $\eta = \mu \otimes \delta_{S(x)}$.*

Demonstração. Por [1, Teorema 9.3], é possível encontrar uma sequência de funções borelianas $S_n : X \rightarrow Y$ tal que $\eta = \lim_{n \rightarrow \infty} \mu \otimes \delta_{S_n(x)}$ e $(S_n)_* \mu = \nu$ para todo $n \in \mathbb{N}$, de modo que $\mu \otimes \delta_{S_n(x)} \in [\gamma] \forall n \in \mathbb{N}$. Considere $\varphi(y) = |y|^2$ e note que

$$\int_X |S_n(x)|^2 d\mu = \int_X \varphi(S_n(x)) d\mu = \int_Y \varphi(y) d\nu = \int_Y |y|^2 d\nu < \infty.$$

Considere $\psi \in \mathcal{C}((\mathcal{P}(Y), W_1))$, definida por $\psi(\delta_y) = |y|^2$, e $\Delta \subset \mathcal{P}(Y)$ o conjunto das medidas de Dirac. Note que ψ é Lipschitz sobre Δ , com respeito à distância 1-Wasserstein. Então, considere qualquer extensão Lipschitz de ψ sobre todo $\mathcal{P}(Y)$. Como $(\delta_{S_n})_* \mu = (\delta_T)_* \mu$, tem-se que

$$\int_X |S_n(x)|^2 d\mu = \int_X \psi(\delta_{S_n(x)}) d\mu = \int_X \psi(\delta_{T(x)}) d\mu = \int_X |T(x)|^2 d\mu$$

para todo $n \in \mathbb{N}$. Portanto, passando à uma subsequência, pode-se supor que S_n é fracamente convergente para um limite fraco S . Por [1, Lema 9.1] o resultado segue. $\square$

Dentro deste contexto, o problema de transporte de Monge pode ser interpretado como: minimizar $\int_{X \times Y} c(x, y) d\gamma$ em uma classe de transporte fixa de $\Pi(\mu, \nu)$, isto é, encontrar

$$\min \left\{ \int_{X \times Y} c(x, y) d\gamma : \gamma \in [\mu \otimes \delta_T] \right\}$$

para um dado mapa de transporte T . Agora, em relação ao problema de Monge-Kantorovich, observe que a segunda parte da demonstração do Lema 3.4 vale para planos de transportes gerais, de modo que o segundo marginal pode ser fixado pelos mapas de desintegração, da seguinte maneira:

Lema 3.8. *Considere uma medida $\mu \in \mathcal{P}(X)$, os mapas de desintegração $f : X \rightarrow \mathcal{P}(Y)$ e $g : X \rightarrow \mathcal{P}(Y)$ e os planos $\gamma = \mu \otimes f_{(x)}$ e $\eta = \mu \otimes g_{(x)}$. Tem-se que $f_* \mu = g_* \mu \implies \pi_{y*} \gamma = \pi_{y*} \eta$, onde π_y é a projeção na segunda componente, $(x, y) \mapsto y$.*

Porém, a recíproca do Lema 3.8 não é verdadeira. Por exemplo, considere o seguinte problema de transporte: Três fábricas, x_1, x_2 e x_3 , com produção idêntica, fornecem 100% de seus produtos para duas lojas, y_1 e y_2 , sendo que a loja y_2 tem uma demanda cinco vezes maior que a loja y_1 . Considere μ a medida que fornece a quantidade de produtos produzidos nas fábricas e ν a medida relacionada à quantidade de produtos entregue nas lojas. Dentro deste cenário, estas medidas são dadas por:

$$\mu = \frac{1}{3}\delta_{x_1} + \frac{1}{3}\delta_{x_2} + \frac{1}{3}\delta_{x_3}$$

$$\nu = \frac{1}{6}\delta_{y_1} + \frac{5}{6}\delta_{y_2}.$$

Considere o plano de transporte ilustrado na Figura 3.1, que divide os produtos produzidos na fábrica x_1 entre as duas lojas, de modo que o correspondente mapa de desintegração, representado por f , é dado por $f(x_1) = \frac{1}{2}\delta_{y_1} + \frac{1}{2}\delta_{y_2}$, $f(x_2) = \delta_{y_2}$ e $f(x_3) = \delta_{y_2}$.

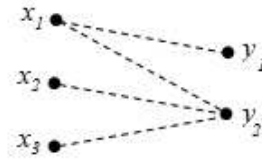


Figura 3.1: Exemplo: Plano de transporte 1.

Note que, como μ é da forma $\sum_i \alpha_i \delta_{x_i}$, tem-se que $f_*\mu = \sum_i \alpha_i \delta_{f(x_i)}$. Agora, suponha que ocorra uma alteração na logística e, no novo plano de transporte, ilustrado na Figura 3.2, a fábrica x_1 entregue sua produção apenas na loja y_2 e a fábrica x_2 divida a produção entre as duas lojas. Nesta nova situação, o mapa de desintegração, representado por g , é dado por $g(x_1) = \delta_{y_2}$, $g(x_2) = \frac{1}{2}\delta_{y_1} + \frac{1}{2}\delta_{y_2}$ e $g(x_3) = \delta_{y_2}$. Observe que $f_*\mu = g_*\mu$, ou seja, a classe de transporte não muda. Por outro lado, se neste novo plano de transporte as fábricas x_1 e x_2 entreguem em ambas as lojas, como ilustrado na Figura 3.3, de modo que x_1 envie 30% de sua produção para y_1 e 70% para y_2 e x_2 envie 20% de sua produção para y_1 e 80% para y_2 , o mapa de desintegração, agora representado por h , seria $h(x_1) = \frac{3}{10}\delta_{y_1} + \frac{7}{10}\delta_{y_2}$, $h(x_2) = \frac{2}{10}\delta_{y_1} + \frac{8}{10}\delta_{y_2}$ e $h(x_3) = \delta_{y_2}$. Neste caso, $f_*\mu \neq h_*\mu$ e, deste modo, os planos de transporte 1 e 3 não pertencem à mesma classe de transporte.

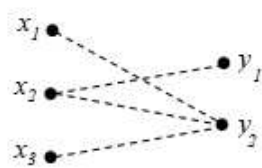


Figura 3.2: Exemplo: Plano de transporte 2.

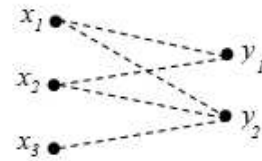


Figura 3.3: Exemplo: Plano de transporte 3.

Ainda mais, se ocorre uma pequena mudança no plano de transporte 3, de modo que x_1 passe a entregar 10% de sua produção em y_1 e 90% em y_2 e x_2 entregue 40% de sua produção em y_1 e 60% em y_2 , o novo mapa de desintegração k é da seguinte forma: $k(x_1) = \frac{1}{10}\delta_{y_1} + \frac{9}{10}\delta_{y_2}$, $k(x_2) = \frac{4}{10}\delta_{y_1} + \frac{6}{10}\delta_{y_2}$ e $k(x_3) = \delta_{y_2}$. Deste modo $h_*\mu \neq k_*\mu$. Assim, com a definição adotada para classe de transporte, ao alterar o número de fábricas que entregam produtos em mais de uma loja, ou alterar o modo como a divisão da entrega é feita em cada fábrica, altera-se a classe de transporte.

Por essas considerações, é necessária outra definição de classe de transporte para o contexto do problema de Monge-Kantorovich. Considere $\mu \otimes f \in \Pi(\mu, \nu)$ e $\Lambda = f_*\mu$. Uma vez que $(\pi_y)_*(\mu \otimes f) = \nu$, para todo $\varphi \in \mathcal{C}(Y)$ tem-se que

$$\begin{aligned} \int_Y \varphi(y) d\nu &= \int_X \left(\int_Y \varphi(y) df_{(x)} \right) d\mu \\ &= \int_X I_\varphi(f_{(x)}) d\mu \\ &= \int_{\mathcal{P}(Y)} I_\varphi(\lambda) d\Lambda(\lambda) \\ &= \int_{\mathcal{P}(Y)} \left(\int_Y \varphi(y) d\lambda \right) d\Lambda \end{aligned}$$

onde I_φ é a função definida pela equação (3.3) e $\lambda \in \mathcal{P}(Y)$. Deste modo, toda medida de probabilidade Λ sobre $(\mathcal{P}(Y), W_1)$ que satisfaz

$$\int_{\mathcal{P}(Y)} \lambda d\Lambda = \nu \quad (3.4)$$

define uma classe de transporte $[\gamma] = \{\mu \otimes f : f_*\mu = \Lambda\}$. Por exemplo, a classe de transporte $[\mu \times \nu]$ corresponde à medida $\Lambda = \delta_\nu$. Deste ponto de vista, o problema de Monge-Kantorovich na classe Λ pode ser interpretado como:

$$MK_\Lambda(c, \mu, \nu) = \inf_{\gamma} \left\{ \int_{X \times Y} c(x, y) d\gamma : \gamma = \mu \otimes f, f_*\mu = \Lambda \right\}.$$

Esta noção de classe de transporte permite construir o problema de Monge-Kantorovich como um problema de Monge abstrato entre os espaços X e $\mathcal{P}(Y)$, considerando uma função custo

$$\tilde{c}(x, \lambda) = \int_Y c(x, y) d\lambda.$$

De fato, note que, para qualquer mapa de desintegração $f : X \rightarrow \mathcal{P}(Y)$ tal que $f_*\mu = \Lambda$,

$$\begin{aligned} \int_X \tilde{c}(x, f_{(x)}) d\mu &= \int_X \left(\int_Y c(x, y) df_{(x)} \right) d\mu \\ &= \int_{X \times Y} c(x, y) d(\mu \otimes f) \end{aligned}$$

e, deste modo, o problema $MK_\Lambda(c, \mu, \nu)$ é equivalente ao problema de Monge $M(\tilde{c}, \mu, \Lambda)$.

3.2 Desintegração e variedades: do Teorema 2.3 ao Corolário 3.2

Em [15, Seção 10] é apresentado um exemplo no qual considera-se o seguinte sistema: um mapa solenoidal $F : \Sigma \rightarrow \Sigma$, onde $\Sigma = S^1 \times D^2$, e tal que F é o *skew product* $F(x, y) = (T(x), G(x, y))$, de modo que a aplicação $T : S^1 \rightarrow S^1$ é expansora e $G : \Sigma \rightarrow D^2$ é uma aplicação diferenciável com $\|\frac{\partial G}{\partial x}\|_\infty < \infty$. Além disso, considera-se uma folheação $\mathcal{F}^s = \{\{x\} \times D^2\}_{x \in S^1}$, de tal modo que $\mathcal{F}^s$ esteja contraindo, ou seja, $\exists \alpha \in \mathbb{R}$, $0 < \alpha < 1$, tal que, para todo $x \in S^1$, tem-se $|G(x, y_1) - G(x, y_2)| \leq \alpha|y_1 - y_2|$, para todo $y_1, y_2 \in D^2$. Então, considera-se o Teorema de Rokhlin e uma medida de probabilidade μ definida em Σ , descreve-se uma desintegração de μ ao longo das folhas de $\mathcal{F}^s$ e, com isso, estuda-se propriedades estatísticas do sistema em questão. Este exemplo será apresentado com mais detalhes no próximo capítulo. Nesta seção, explora-se esta abordagem de aplicação do Teorema de Rokhlin considerando uma folheação (de forma um pouco mais abrangente do que em [15]) e, a partir dela, uma conexão entre o Teorema 2.3 e o Corolário 3.2.

Considere $\Sigma = M_1 \times M_2$, onde M_1 e M_2 são variedades riemannianas compactas finitas e a folheação de Σ dada por $\mathcal{F}^s = \{\{x\} \times M_2\}_{x \in M_1}$. Denote por $\mathcal{SB}(\Sigma)$ o conjunto das medidas borelianas com sinal em Σ e por $\mathcal{AB}(\Sigma)$ o conjunto $\{\mu \in \mathcal{SB}(\Sigma) : \pi_{x*}\mu^+ \ll m_1 \text{ e } \pi_{x*}\mu^- \ll m_1\}$, onde $\pi_x : \Sigma \rightarrow M_1$ é a projeção dada por $(x, y) \mapsto x$, m_1 é a medida de Lebesgue em M_1 e μ^+ e μ^- representam as partes positiva e negativa, respectivamente, da decomposição de Jordan de μ . Considere uma medida de probabilidade $\mu \in \mathcal{AB}(\Sigma)$, a função $\pi : \Sigma \rightarrow \mathcal{F}^s$, que associa cada ponto $(x, y) \in \Sigma$ o elemento ζ de $\mathcal{F}^s$ que o contém, e a medida obtida pelo *push forward* de μ por π , $\hat{\mu} = \pi_*\mu$. Observe que é possível construir uma sequência crescente de partições enumeráveis considerando a ideia de partição da unidade subordinada à cobertura, de modo que $\mathcal{F}^s$ é uma partição mensurável de Σ . Além disso, observe que Σ é espaço métrico completo e separável, uma vez que é produto de variedades riemannianas compactas. Assim, o Teorema 2.3 descreve uma desintegração de μ relativamente a $\mathcal{F}^s$ por uma família $\{\mu_\zeta : \zeta \in \mathcal{F}^s\}$ de probabilidades em Σ , tal que, para $E \subset \Sigma$ mensurável, tem-se:

1. $\mu_\zeta(\zeta) = 1$, para $\hat{\mu}$ -quase todo $\zeta \in \mathcal{F}^s$.
2. $\zeta \mapsto \mu_\zeta(E)$ é mensurável.
3. $\mu(E) = \int \mu_\zeta(E) d\hat{\mu}(\zeta)$.

Indo em direção ao Corolário 3.2, ao invés de associar as probabilidades condicionais obtidas via desintegração à elementos de $\mathcal{F}^s$, é necessário associá-las à elementos de M_1 . Para isso, reescreva $\mathcal{F}^s$ por $\{\zeta_x\}_{x \in M_1}$, onde $\zeta_x = \{x\} \times M_2$ para cada $x \in M_1$, e considere a medida $\mu_x = \pi_{x*}\mu$. Para cada $x \in M_1$, considere μ_{ζ_x} a restrição da probabilidade condicional μ_ζ para

a qual $\zeta = \{x\} \times M_2$ à ζ . Observe que $\pi^{-1}(\zeta) = \{x\} \times M_2 = \zeta_x = \pi_x^{-1}(x)$, de modo que

$$\begin{aligned} \int \mu_\zeta(E) d\hat{\mu}(\zeta) &= \int \mu_\zeta(E) \mu(\pi^{-1}(d\zeta)) \\ &= \int_{M_1} \mu_{\zeta_x}(E \cap \zeta_x) \mu(\pi_x^{-1}(dx)) \\ &= \int_{M_1} \mu_{\zeta_x}(E \cap \zeta_x) d\mu_x \end{aligned}$$

ou seja, pode-se escrever

$$\mu(E) = \int_{M_1} \mu_{\zeta_x}(E \cap \zeta_x) d\mu_x.$$

Note que, deste modo, μ_{ζ_x} está suportada em $\pi_x^{-1}(x) = \zeta_x$. Então, é possível considerar a desintegração $\{\mu_{\zeta_x} : \zeta_x \in \mathcal{F}^s\}$ ao longo das folhas de Σ , associada à medida μ_x . Ainda mais, considere $\pi_y : \Sigma \rightarrow M_2$ a projeção definida por $\pi_y(x, y) = y$ e $\pi_y|_{\zeta_x} : \zeta_x \rightarrow M_2$ a restrição de π_y à folha ζ_x . Observe que com a restrição $\pi_y|_{\zeta_x}$ é possível identificar cada fibra $(\pi_x)^{-1}(x)$ com M_2 , munindo cada folha com a estrutura da variedade M_2 . Assim, pode-se definir uma medida em M_2 por $\mu'_{\zeta_x}(A) = \pi_y|_{\zeta_x*} \mu_{\zeta_x}(A)$, para $A \subset M_2$ mensurável. Ou seja, ao considerar uma probabilidade boreliana μ em Σ e a medida $\mu_x = \pi_{x*} \mu$, é possível obter uma família $\{\mu'_{\zeta_x} : x \in M_1\} \subset \mathcal{P}(M_2)$, μ_x -quase sempre unicamente determinada, tal que $\mu = \int_{M_1} \mu'_{\zeta_x} d\mu_x$, isto é, a desintegração descrita no Corolário 3.2.

3.2.1 Sobre desintegração de medidas absolutamente contínuas

A continuidade absoluta de medidas é uma importante ferramenta para o estudo de diversas propriedades de sistemas dinâmicos. Dentro do contexto de desintegração trabalhado nesta seção, e considerando a abordagem realizada em [16], é possível mostrar que, se uma medida ν é absolutamente contínua em relação a μ , então, $\nu_{\zeta_x} \ll \mu_{\zeta_x}$. Para isso, considere, como anteriormente, $\Sigma = M_1 \times M_2$, a folheação $\mathcal{F}^s$ e as aplicações π_x e $\pi_y|_{\zeta_x}$. Tome uma medida de probabilidade $\mu \in \mathcal{AB}(\Sigma)$ e considere uma desintegração $\{\mu_{\zeta_x} : x \in M_1\}$ associada à medida $\mu_x = \pi_{x*} \mu$. Note que, como $\mu \in \mathcal{AB}(\Sigma)$, μ_x pode ser identificada com uma densidade marginal não negativa $\phi_x : M_1 \rightarrow \mathbb{R}$, e pode-se definir a restrição de μ em ζ_x , $\mu|_{\zeta_x}(A) = \pi_y|_{\zeta_x*}(\phi_x(x) \mu_{\zeta_x})(A)$, para todo $A \subset M_2$ mensurável. Ainda, considere $f : M_1 \times M_2 \rightarrow \mathbb{R}$ uma função μ -integrável e denote por $\nu = f\mu$ a medida definida por $\nu(E) := \int_E f d\mu$. Denote $f|_{\zeta_x}$ a restrição de f ao conjunto $\zeta_x = \{x\} \times M_2$ e $f_x : M_2 \rightarrow \mathbb{R}$ a função definida por $f_x(y) = f(x, y)$. Observe que $f_x = f|_{\zeta_x} \circ \pi_y|_{\zeta_x}^{-1}$, e $f|_{\zeta_x} = f_x \circ \pi_y|_{\zeta_x}$.

Proposição 3.9. *Se $\{\nu_{\zeta_x}\}$ é uma desintegração de ν associada à $\nu_x := \pi_{x*} \nu$, então tem-se que $\nu_x \ll \mu_x$ e $\nu_{\zeta_x} \ll \mu_{\zeta_x}$. Ainda mais,*

$$\tilde{f}(x) = \frac{d\nu_x}{d\mu_x}(x) = \int_{M_2} f_x d\mu|_{\zeta_x} \quad (3.5)$$

e para ν_x -quase todo $x \in M_1$

$$\frac{d\nu_{\zeta_x}}{d\mu_{\zeta_x}}(y) = \begin{cases} \frac{f|_{\zeta_x}(y)}{\int f|_{\zeta_x}(y)d\mu_{\zeta_x}}, & \text{se } x \in B^c \\ 0, & \text{se } x \in B. \end{cases} \quad (3.6)$$

para $y \in M_2$, onde $B = \tilde{f}^{-1}(0)$.

Demonstração. Para cada conjunto mensurável $A \subset M_1$, note que:

$$\begin{aligned} \int_A \frac{d\nu_x}{dm_1} dm_1 &= \int_{M_1 \times M_2} \mathbb{1}_A \circ \pi_x d(f\mu) \\ &= \int_{M_1 \times M_2} \mathbb{1}_{\pi_x^{-1}(A)} f d\mu \\ &= \int_{M_1} \left(\int_{M_2} \mathbb{1}_{\pi_x^{-1}(A)} |f|_{\zeta_x} d\mu_{\zeta_x} \right) d\mu_x \\ &= \int_{M_1} \left(\phi_x(x) \int_{M_2} \mathbb{1}_{\pi_x^{-1}(A)} |f|_{\zeta_x} d\mu_{\zeta_x} \right) dm_1 \\ &= \int_A \left(\int_{M_2} f_x d\mu|_{\zeta_x} \right) dm_1 \end{aligned}$$

e, deste modo, tem-se que

$$\frac{d\nu_x}{dm_1} = \int_{M_2} f_x d\mu|_{\zeta_x}$$

para m_1 -quase todo $x \in M_1$. Ainda, note que

$$\int f_x d\mu|_{\zeta_x} = \phi_x(x) \int f|_{\zeta_x} d\mu_{\zeta_x} \quad (3.7)$$

de modo que $\tilde{f}(x) = 0$ se, e somente se, $\phi_x(x) = 0$ ou $\int f|_{\zeta_x} d\mu_{\zeta_x} = 0$. Agora, considere a função $h_x : M_2 \rightarrow \mathbb{R}$ dada por

$$h_x(y) = \begin{cases} \frac{f|_{\zeta_x}(y)}{\int f|_{\zeta_x}(y)d\mu_{\zeta_x}}, & \text{se } x \in B^c \\ 0, & \text{se } x \in B. \end{cases}$$

para $y \in M_2$, onde $B = \tilde{f}^{-1}(0)$. Note que, para todo $E \subset M_1 \times M_2$, tem-se:

$$\begin{aligned}
\nu(E) &= \int_E f d\mu \\
&= \int_{M_1} \int_{E \cap \zeta_x} f|_{\zeta_x} d\mu_{\zeta_x} d\mu_x \\
&= \int_{B^c} \int_{E \cap \zeta_x} f|_{\zeta_x} d\mu_{\zeta_x} d\mu_x \\
&= \int_{B^c} \int_{E \cap \zeta_x} \frac{\int f|_{\zeta_x} d\mu_{\zeta_x}}{\int f|_{\zeta_x} d\mu_{\zeta_x}} f|_{\zeta_x} d\mu_{\zeta_x} d\mu_x \\
&= \int_{B^c} \left(\phi_x(x) \int f|_{\zeta_x} d\mu_{\zeta_x} \right) \left(\frac{1}{\int f|_{\zeta_x} d\mu_{\zeta_x}} \int_{E \cap \zeta_x} f|_{\zeta_x} d\mu_{\zeta_x} \right) dm_1 \\
&= \int_{B^c} \tilde{f}(x) \left(\frac{1}{\int f|_{\zeta_x} d\mu_{\zeta_x}} \int_{E \cap \zeta_x} f|_{\zeta_x} d\mu_{\zeta_x} \right) dm_1 \\
&= \int_{B^c} \left(\frac{1}{\int f|_{\zeta_x} d\mu_{\zeta_x}} \int_{E \cap \zeta_x} f|_{\zeta_x} d\mu_{\zeta_x} \right) d\tilde{f} m_1 \\
&= \int_{B^c} \int_{E \cap \zeta_x} h_x d\mu_{\zeta_x} d\nu_x \\
&= \int_{M_1} \int_{E \cap \zeta_x} h_x d\mu_{\zeta_x} d\nu_x
\end{aligned}$$

e, portanto, segue a equação (3.6). □

3.2.2 Restrição e medidas com sinal

Até o momento considerou-se apenas medidas positivas $\mu \in \mathcal{AB}(\Sigma)$, mas é possível generalizar os resultados para medidas com sinal, considerando a decomposição de Jordan.

Para uma medida com sinal $\mu \in \mathcal{AB}(\Sigma)$ e a decomposição $\mu = \mu^+ - \mu^-$, seguindo a abordagem de [16], define-se a restrição de μ em ζ_x por $\mu|_{\zeta_x} = \mu^+|_{\zeta_x} - \mu^-|_{\zeta_x}$. Pode-se demonstrar que esta restrição $\mu|_{\zeta_x}$ não depende da decomposição, ou seja, se $\mu = \mu_1 - \mu_2$, onde μ_1 e μ_2 são medidas positivas quaisquer, então $\mu|_{\zeta_x} = \mu_1|_{\zeta_x} - \mu_2|_{\zeta_x}$ para m_1 -quase todo $x \in M_1$. Para isso, são necessários alguns resultados:

Proposição 3.10. *Sejam $\mu_1, \mu_2 \in \mathcal{AB}(\Sigma)$ medidas positivas cujas densidades marginais são, respectivamente, $d(\mu_{1x}) = \phi_x dm_1$ e $d(\mu_{2x}) = \psi_x dm_1$, onde $\phi_x, \psi_x \in L^1(m_1)$. Então $(\mu_1 + \mu_2)|_{\zeta_x} = \mu_1|_{\zeta_x} + \mu_2|_{\zeta_x}$ para m_1 -quase todo $x \in M_1$.*

Demonstração. Note que $d(\mu_1 + \mu_2) = (\phi_x + \psi_x) dm_1$. Ainda, considere a desintegração de $\mu_1 + \mu_2$ dada por $\{(\mu_1 + \mu_2)_{\zeta_x}\}_{\zeta_x}$, onde

$$(\mu_1 + \mu_2)_{\zeta_x} = \begin{cases} \frac{\phi_x(x)}{\phi_x(x) + \psi_x(x)} \mu_{1x} + \frac{\psi_x(x)}{\phi_x(x) + \psi_x(x)} \mu_{2x}, & \text{se } \phi_x(x) + \psi_x(x) \neq 0 \\ 0, & \text{se } \phi_x(x) + \psi_x(x) = 0 \end{cases}$$

Pela unicidade da desintegração, tem-se que, para m_1 -quase todo $x \in M_1$,

$$(\phi_x + \psi_x)(x)(\mu_1 + \mu_2)_{\zeta_x} = (\phi_x + \psi_x)(x) \left(\frac{\phi_x(x)}{\phi_x(x) + \psi_x(x)} \mu_{1x} + \frac{\psi_x(x)}{\phi_x(x) + \psi_x(x)} \mu_{2x} \right)$$

então

$$(\phi_x + \psi_x)(x)(\mu_1 + \mu_2)_{\zeta_x} = \phi_x(x)\mu_{1x} + \psi_x(x)\mu_{2x}.$$

Portanto, $(\mu_1 + \mu_2)|_{\zeta_x} = \mu_1|_{\zeta_x} + \mu_2|_{\zeta_x}$ para m_1 -quase todo $x \in M_1$. $\square$

Definição 3.11. Diz-se que uma medida positiva λ_1 é disjunta de uma medida positiva λ_2 se $(\lambda_1 - \lambda_2)^+ = \lambda_1$ e $(\lambda_1 - \lambda_2)^- = \lambda_2$.

Lema 3.12. Considere μ e ν medidas com sinal e suas respectivas decomposições de Jordan $\mu = \mu^+ - \mu^-$ e $\nu = \nu^+ - \nu^-$. Então, existem medidas positivas $\mu_1, \mu_2, \mu^{++}, \mu^{--}, \nu^{++}$ e ν^{--} tais que $\mu^+ = \mu^{++} + \mu_1$, $\mu^- = \mu^{--} + \mu_2$, $\nu^+ = \nu^{++} + \mu_2$ e $\nu^- = \nu^{--} + \mu_1$.

Demonstração. Suponha que $\mu = \nu_1 - \nu_2$ e considere $\mu' = \nu_1 - \mu^+$, de modo que $\nu_1 = \mu^+ + \mu'$. Como $\mu = \mu^+ - \mu^-$, tem-se que $\mu^+ - \mu^- = \nu_1 - \nu_2$, de modo que $\nu_2 = \nu_1 - \mu^+ + \mu^-$ e então $\nu_2 = \mu^- + \mu'$. Agora, considere a medida $\mu^+ - \nu^-$, com sua decomposição de Jordan $(\mu^+ - \nu^-)^+ - (\mu^+ - \nu^-)^-$. Seguindo o mesmo raciocínio, tem-se

$$\mu^+ = (\mu^+ - \nu^-)^+ + \mu_1$$

$$\nu^- = (\mu^+ - \nu^-)^- + \mu_1$$

para alguma medida positiva μ_1 . Do mesmo modo, considerando $(\nu^+ - \mu^-)$ e sua decomposição de Jordan $(\nu^+ - \mu^-)^+ - (\nu^+ - \mu^-)^-$, pode-se escrever

$$\nu^+ = (\nu^+ - \mu^-)^+ + \mu_2$$

$$\mu^- = (\nu^+ - \mu^-)^- + \mu_2$$

para alguma medida positiva μ_2 .

É claro que $(\mu^+ - \nu^-)^+$ é disjunta de $(\mu^+ - \nu^-)^-$ e que $(\nu^+ - \mu^-)^+$ é disjunta de $(\nu^+ - \mu^-)^-$. Ainda, como μ^+ e μ^- são disjuntas, $(\mu^+ - \nu^-)^+$ e $(\nu^+ - \mu^-)^-$ são também disjuntas. Além disso, como ν^+ e ν^- são disjuntas, segue que $(\nu^+ - \mu^-)^+$ e $(\mu^+ - \nu^-)^-$ são disjuntas. Assim, segue que

$$((\mu^+ - \nu^-)^+ + (\nu^+ - \mu^-)^+) - ((\nu^+ - \mu^-)^- + (\mu^+ - \nu^-)^-)$$

é uma decomposição de Jordan de $\mu + \nu$. Assim, tem-se o resultado desejado tomando

$$\mu^{++} = (\mu^+ - \nu^-)^+$$

$$\mu^{--} = (\nu^+ - \mu^-)^-$$

$$\nu^{++} = (\nu^+ - \mu^-)^+$$

$$\nu^{--} = (\mu^+ - \nu^-)^-.$$

□

Proposição 3.13. *Considere $\mu, \nu \in \mathcal{AB}(\Sigma)$ medidas com sinal. Então, $(\mu + \nu)|_{\zeta_x} = \mu|_{\zeta_x} + \nu|_{\zeta_x}$ para m_1 -quase todo $x \in M_1$.*

Demonstração. Dadas as decomposições de Jordan de μ e ν , por definição, tem-se

$$\mu|_{\zeta_x} = \mu^+|_{\zeta_x} - \mu^-|_{\zeta_x}$$

$$\nu|_{\zeta_x} = \nu^+|_{\zeta_x} - \nu^-|_{\zeta_x}.$$

De acordo com o Lema 3.12, considere $\mu^+ = \mu^{++} + \mu_1$, $\mu^- = \mu^{--} + \mu_2$, $\nu^+ = \nu^{++} + \mu_2$ e $\nu^- = \nu^{--} + \mu_1$. Assim,

$$(\mu + \nu)^+ = (((\mu^{++} + \mu_1) - (\mu^{--} + \mu_2)) + ((\nu^{++} + \mu_2) - (\nu^{--} + \mu_1)))^+ = \mu^{++} + \nu^{++}$$

$$(\mu + \nu)^- = (((\mu^{++} + \mu_1) - (\mu^{--} + \mu_2)) + ((\nu^{++} + \mu_2) - (\nu^{--} + \mu_1)))^- = \mu^{--} + \nu^{--}.$$

Pela Proposição 3.10, tem-se que

$$\mu^+|_{\zeta_x} = \mu^{++}|_{\zeta_x} + \mu_1|_{\zeta_x}$$

$$\mu^-|_{\zeta_x} = \mu^{--}|_{\zeta_x} + \mu_2|_{\zeta_x}$$

$$\nu^+|_{\zeta_x} = \nu^{++}|_{\zeta_x} + \mu_2|_{\zeta_x}$$

$$\nu^-|_{\zeta_x} = \nu^{--}|_{\zeta_x} + \mu_1|_{\zeta_x}$$

e, ainda,

$$(\mu + \nu)^+|_{\zeta_x} = \mu^{++}|_{\zeta_x} + \nu^{++}|_{\zeta_x}$$

$$(\mu + \nu)^-|_{\zeta_x} = \mu^{--}|_{\zeta_x} + \nu^{--}|_{\zeta_x}$$

de modo que

$$\begin{aligned}
 (\mu + \nu)|_{\zeta_x} &= (\mu + \nu)^+|_{\zeta_x} - (\mu + \nu)^-|_{\zeta_x} \\
 &= (\mu^{++}|_{\zeta_x} + \nu^{++}|_{\zeta_x}) - (\mu^{--}|_{\zeta_x} + \nu^{--}|_{\zeta_x}) \\
 &= \mu^+|_{\zeta_x} - \mu_1|_{\zeta_x} + \nu^+|_{\zeta_x} - \mu_2|_{\zeta_x} - \mu^-|_{\zeta_x} + \mu_2|_{\zeta_x} - \nu^-|_{\zeta_x} + \mu_1|_{\zeta_x} \\
 &= \mu^+|_{\zeta_x} - \mu^-|_{\zeta_x} + \nu^+|_{\zeta_x} - \nu^-|_{\zeta_x} \\
 &= \mu|_{\zeta_x} + \nu|_{\zeta_x}.
 \end{aligned}$$

□

Corolário 3.14. *Considere $\mu \in \mathcal{AB}(\Sigma)$. Se μ_1 e μ_2 são medidas positivas quaisquer tais que $\mu = \mu_1 - \mu_2$, então $\mu|_{\zeta_x} = \mu_1|_{\zeta_x} - \mu_2|_{\zeta_x}$, i.e., a restrição não depende da decomposição de μ .*

CAPÍTULO 4

Algumas propriedades estatísticas de sistemas dinâmicos

A Teoria Ergódica estuda propriedades estatísticas dos sistemas dinâmicos. Em termos gerais, lida com sistemas dinâmicos munidos com uma medida invariante. Considerando um espaço de medida $(X, \mathcal{B}, \mu)$, μ é invariante pela transformação mensurável $f : X \rightarrow X$ se $\mu(A) = \mu(f^{-1}(A))$ para todo $A \in \mathcal{B}$, ou seja, f preserva μ se, e somente se, $\int \varphi d\mu = \int \varphi \circ f d\mu$, para toda função $\varphi : X \rightarrow \mathbb{R}$ μ -integrável. Existem medidas invariantes que são “especiais”, pois decodificam propriedades estatísticas do sistema. Tais medidas são denominadas *físicas* e são definidas do seguinte modo:

Definição 4.1. *Considere uma variedade compacta M , uma aplicação $f : M \rightarrow M$ de classe C^1 e uma medida de probabilidade f -invariante μ . A medida μ é denominada medida física se o conjunto*

$$B(\mu) = \left\{ x \in M : \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \int \varphi d\mu, \forall \varphi : M \rightarrow \mathbb{R} \text{ contínua} \right\}$$

tem medida de Lebesgue positiva. O conjunto $B(\mu)$ é chamado de bacia de μ .

Além disso, diversas propriedades estão associadas à medidas invariantes, como, por exemplo, mistura e ergodicidade. Dada uma transformação mensurável $f : X \rightarrow X$ e uma probabilidade f -invariante μ , o sistema (f, μ) é dito (*fortemente*) *misturador* se

$$\lim_{n \rightarrow \infty} \mu(f^{-n}(A) \cap B) = \mu(A)\mu(B)$$

e *fracamente* *misturador* se

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} |\mu(f^{-k}(A) \cap B) - \mu(A)\mu(B)| = 0$$

para quaisquer conjuntos A e B mensuráveis. Dado qualquer conjunto mensurável $E \subset X$, o sistema (f, μ) é dito *ergódico* se $\tau(E, x) = \mu(E)$ para μ -quase todo $x \in X$, onde

$$\tau(E, x) = \lim_{n \rightarrow \infty} \frac{1}{n} \# \{0 \leq j < n : f^j \in E\} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathbb{1}_E(f^j(x))$$

é o *tempo médio de visita* de x a E . Outras formas equivalentes de definir a ergodicidade de (f, μ) são [21, Proposição 4.1.3]:

- Para todo conjunto mensurável $E \subset X$, a função $\tau(E, \cdot)$ é constante μ -quase sempre.
- Para toda função integrável $\varphi : X \rightarrow \mathbb{R}$ tem-se que $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \int \varphi d\mu$ para μ -quase todo ponto, isto é, médias temporais coincidem com médias espaciais.
- Para toda função integrável $\varphi : X \rightarrow \mathbb{R}$ a média temporal $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x))$ é constante em μ -quase todo ponto.
- Para todo subconjunto invariante B tem-se $\mu(B) = 0$ ou $\mu(B) = 1$, ou seja, o sistema é dinamicamente indivisível.

Observação. Uma função mensurável $\psi : X \rightarrow \mathbb{R}$ é dita *invariante* se $\psi \circ f = \psi$ em μ -quase todo ponto. Um conjunto $B \subset X$ é dito *invariante* se sua função característica $\mathbb{1}_B$ é uma função invariante.

Observação. Da definição, todo sistema misturador é fracamente misturador. Ainda, a mistura é uma propriedade mais forte que a ergodicidade, no seguinte sentido: Suponha que existe algum conjunto invariante $A \subset X$ com $0 < \mu(A) < 1$ e tome $B = A^c$. Então, para todo $n \in \mathbb{N}$, $f^{-n}(A) \cap B = \emptyset$. Logo, $\mu(f^{-n}(A) \cap B) = 0$ para todo n , enquanto que $\mu(A)\mu(B) \neq 0$ e o sistema (f, μ) não é misturador. Com isso, segue que todo sistema misturador é ergódico e, pelo mesmo argumento, todo sistema fracamente misturador é ergódico.

O estudo de medidas invariantes pode conduzir à informações importantes sobre o comportamento dinâmico do sistema, dificilmente obtidas de outro modo. São conhecidos resultados que garantem a existência de medidas invariantes sob certas condições. Por exemplo, se X é um espaço métrico compacto e f é contínua, sabe-se que existe pelo menos uma medida de probabilidade f -invariante em X [21, Teorema 2.1]. O estudo do comportamento do operador de transferência associado ao sistema é uma maneira de encontrar tais medidas, além de outros resultados relevantes [15].

Dado um espaço métrico X e uma aplicação boreliana $T : X \rightarrow X$, o operador de transferência associado à T é a função linear $L : \mathcal{SB}(X) \rightarrow \mathcal{SB}(X)$ definida de modo que, se $\nu \in \mathcal{SB}(X)$, então $L[\nu] \in \mathcal{SB}(X)$ é tal que $L[\nu](A) = \nu(T^{-1}(A))$ para todo $A \subset X$ mensurável. Se a medida considerada é absolutamente contínua com respeito à medida de Lebesgue m , de modo que $d\nu = f dm$ com $f \in L^1(X, \nu)$, e se T é não singular com respeito a m (ou seja, $m(T^{-1}(A)) = 0$ se, e somente se, $m(A) = 0$), este operador induz outro operador $\tilde{L} : L^1(m) \rightarrow L^1(m)$ que age nas medidas densidade, é positivo e preserva a integral $\int \tilde{L}f dm = \int f dm$. Assim, segue diretamente que $\tilde{L}$ é uma contração fraca para a norma L^1 , isto é, se f é uma densidade em L^1 , então $\|\tilde{L}f\|_1 \leq \|f\|_1$. Além disso:

Proposição 4.2. Se $f \in L^1(m)$ e $g \in L^\infty(m)$, então

$$\int g \tilde{L}(f) dm = \int g \circ T f dm.$$

Demonstração. Se g uma função característica, i. e., $g = \mathbb{1}_A$, tem-se que:

$$\int g \circ T f dm = \int \mathbb{1}_A \circ T f dm = \int \mathbb{1}_{T^{-1}(A)} f dm = \int_A \tilde{L}f dm = \int g \tilde{L}(f) dm.$$

Uma vez que $g \in L^\infty$ por hipótese, é possível aproximá-la por funções simples. Deste modo, considere $\hat{g} = \sum_i a_i \mathbb{1}_{A_i}$, com $\|g - \hat{g}\|_\infty \leq \epsilon$ e

$$\int \hat{g} L(f) dm = \int \hat{g} \circ T f dm.$$

Assim

$$\int g \tilde{L}(f) dm = \int [g - \hat{g}] \tilde{L}(f) dm + \int \hat{g} \tilde{L}(f) dm$$

e

$$\int g \circ T f dm = \int [g - \hat{g}] \circ T f dm + \int \hat{g} \circ T f dm.$$

Note que:

$$\left| \int [g - \hat{g}] \circ T f dm \right| \leq \| [g - \hat{g}] \circ T \|_\infty \|f\|_1 \leq \| [g - \hat{g}] \|_\infty \|f\|_1 \leq \epsilon$$

e

$$\left| \int [g - \hat{g}] \tilde{L}(f) dm \right| \leq \| [g - \hat{g}] \|_\infty \| \tilde{L}f \|_1 \leq \epsilon$$

e então segue o resultado. $\square$

Neste capítulo estuda-se a existência de medidas invariantes e algumas propriedades em dois exemplos de sistemas dinâmicos, utilizando o operador de transferência associado. Primeiramente, considera-se sistemas cuja dinâmica é uma transformação expansora. Em seguida, considera-se o caso do sistema com mapa solenoidal uniformemente hiperbólico, no qual a desintegração de medidas é utilizada como ferramenta na definição de espaços apropriados ao sistema.

4.1 Explorando exemplos

4.1.1 Transformações expansoras

As transformações expansoras $f : M \rightarrow M$ têm como propriedade característica expandirem distâncias entre pontos próximos, ou seja, são tais que existe uma constante $\sigma > 1$ e $d(f(x), f(y)) \geq \sigma d(x, y)$, para $d(x, y)$ suficientemente pequena. Se M é uma variedade compacta e f é uma transformação de classe C^1 , diz-se que f é *expansora* se existe $\sigma > 1$ e alguma métrica riemanniana em M tal que

$$\|Df(x)v\| \geq \sigma \|v\| \quad (4.1)$$

para todo $x \in M$ e todo $v \in T_x M$. Em particular, f é um difeomorfismo local, pois a equação (4.1) implica que $Df(x)$ é um isomorfismo para todo $x \in M$. As transformações expansoras exibem um comportamento dinâmico rico, tanto do ponto de vista métrico e topológico, quanto do ponto de vista ergódico, permitindo estabelecer paradigmas que são úteis para compreender muitos sistemas complexos.

Nesta seção, é considerado o mapa de expansão $T : S^1 \rightarrow S^1$ tal que $T \in C^2$ e $|T'(x)| > 1$ para todo x , e são exploradas algumas propriedades, tendo como embasamento o artigo [15]. Considerando o operador de transferência associado à T , o espaço “forte” $W^{1,1}$ (que é o espaço de Banach das funções absolutamente contínuas, para as quais $f(x) = f(0) + \int_0^x f'(t) dt$ para algum $f' \in L^1$, com a norma $\|f\| = \|f\|_1 + \|f'\|_1$) e o espaço “fraco” L^1 , demonstra-se que os iterados de L satisfazem uma desigualdade de Lasota-Yorke; Com isso, demonstra-se a existência de medidas invariantes e propriedades de mistura, convergência exponencial para o equilíbrio, gap espectral e estabilidade.

Desigualdade de Lasota-Yorke Considere L o operador de transferência associado à T . Para mostrar que os iterados de L satisfazem uma desigualdade de Lasota-Yorke, considere a fórmula explícita para o operador de transferência associado à T :

$$Lf(x) = \sum_{y \in T^{-1}(x)} \frac{f(y)}{|T'(y)|} \quad (4.2)$$

cujas demonstrações são apresentadas em [7, página 85]. Tomando a derivada da equação (4.2), nota-se que

$$(Lf)' = L\left[\frac{1}{T'} f'\right] - L\left[\frac{T''}{(T')^2} f\right]$$

de modo que

$$\begin{aligned} \|(Lf)'\|_1 &\leq \left\|\frac{f'}{T'}\right\|_1 + \left\|\frac{T''f}{(T')^2}\right\|_1 \\ &\leq \alpha\|f'\|_1 + \left\|\frac{T''}{(T')^2}\right\|_\infty \|f\|_1 \end{aligned}$$

onde $\alpha = \max(\frac{1}{T'})$. Então, é possível escrever

$$\|(Lf)'\|_1 + \|Lf\|_1 \leq \alpha\|f'\|_1 + \alpha\|f\|_1 + \left(\left\|\frac{T''}{(T')^2}\right\|_\infty + 1\right)\|f\|_1$$

e assim

$$\|Lf\| \leq \alpha\|f\| + \left(\left\|\frac{T''}{(T')^2}\right\|_\infty + 1\right)\|f\|_1.$$

Iterando esta desigualdade, tem-se

$$\|(L^n f)\| \leq \alpha^n \|f\| + \frac{\left(\left\|\frac{T''}{(T')^2}\right\|_\infty + 1\right)}{1 - \alpha} \|f\|_1 \quad (4.3)$$

que é uma desigualdade de Lasota-Yorke. Além disso, começando com $f \in W^{1,1}$, todos os elementos da sequência $L^n f$ de iterados de f estão em $W^{1,1}$, e suas normas são uniformemente limitadas. Pela equação (4.3), tem-se que $\|(L^n f)\|$ é uniformemente limitada, e também $\|(L^n f)\|_\infty$ o é. Ainda, uma vez que o operador de transferência é positivo, tem-se que

$$\sup_{n, \|f\|_\infty=1} \|(L^n f)\|_\infty = \sup_n \|(L^n 1)\|_\infty.$$

Defina $M := \sup_n \|(L^n 1)\|_\infty$. Deste modo, existe n_1 tal que $\alpha^{n_1} M < 1$. Seja $T_2 = T^{n_1}$ uma nova transformação, que ainda possui as mesmas propriedades de regularidade, e expande uniformemente em S^1 . Seja L_2 o operador de transferência associado à T_2 . Da equação (4.2), tem-se que

$$\begin{aligned} \|(L_2 f)'\|_\infty &\leq \left\|L_2\left(\frac{1}{T_2'} f'\right)\right\|_\infty + \left\|L_2\left(\frac{T_2''}{(T_2')^2} f\right)\right\|_\infty \\ &\leq \alpha^{n_1} M \|f'\|_\infty + M \left\|\frac{T_2''}{(T_2')^2}\right\|_\infty \|f\|_\infty. \end{aligned} \quad (4.4)$$

Considere a norma $\|\cdot\|_\ell$ definida por $\|f\|_\ell = \|f'\|_\infty + \|f\|_\infty$ e tome $\lambda = \alpha^{n_1} M < 1$. Com isso, segue que L_2 satisfaz uma desigualdade de Lasota-Yorke: $\|L_2^n f\|_\ell \leq \lambda^n \|f\|_\ell + B\|f\|_\infty$.

Daí, tem-se que

$$\begin{aligned} \|L^{n_{n_1+q}}f\|_\ell &\leq \lambda^n \|L^q f\|_\ell + B \|L^q f\|_\infty \\ &\leq \lambda^n M \|f\|_\ell + BM \|f\|_\infty \end{aligned} \quad (4.5)$$

e então L também satisfaz uma desigualdade de Lasota-Yorke com essas normas.

Existência de medida invariante Com a desigualdade de Lasota-Yorke é possível provar a existência de uma medida invariante absolutamente contínua em L^1 (com densidade em $W^{1,1}$) para T . O seguinte teorema fornece um argumento de compacidade necessário para tal.

Teorema 4.3. ($W^{1,1}$ é compactamente imerso em L^1) *Se $B \subset W^{1,1}$ é um conjunto fortemente limitado, então para cada $\epsilon > 0$, B tem uma ϵ -rede finita para a topologia L^1 . Em particular, qualquer sequência limitada $f_n \in W^{1,1}$ possui uma subsequência f_{n_k} convergindo fracamente a $f \in L^1$.*

Demonstração. Veja [15, Proposição 5]. □

Considere uma sequência

$$g_n = \frac{1}{n} \sum_{i=1}^{n-1} L^i 1$$

onde 1 é a densidade da medida de Lebesgue normalizada m . Pela desigualdade de Lasota-Yorke, a sequência tem uma norma $W^{1,1}$ uniformemente limitada e pelo Teorema 4.3, existe uma subsequência g_{n_k} convergindo em L^1 para um limite h . Como L é contínuo na norma L^1 , tem-se

$$Lh = L\left(\lim_{k \rightarrow \infty} g_{n_k}\right) = \lim_{k \rightarrow \infty} Lg_{n_k} = h$$

de modo que h é uma densidade invariante. Ainda falta demonstrar que $h \in W^{1,1}$. Para isso, note que $L^m(g_n) \in W^{1,1}$ e as normas são uniformemente limitadas. Pela desigualdade de Lasota-Yorke, segue que

$$\|L^{a+m}(g_{n_1}) - L^{b+m}(g_{n_2})\| \leq \alpha^m \|L^a(g_{n_1}) - L^b(g_{n_2})\| + B \|L^a(g_{n_1}) - L^b(g_{n_2})\|_1.$$

Note, também, que se $\|L^0(g_{n_k}) - h\|_1 \leq \epsilon$, então $\|L^j(g_{n_k}) - h\|_1 \leq \epsilon$ para todo $j \geq 0$. Assim a sequência $L^k(g_{n_k})$ converge para h em L^1 e, pela equação acima, é uma sequência de Cauchy na norma $W^{1,1}$. De fato, suponha $k_1 < k_2$. Então

$$\|L^{k_1}(g_{n_{k_1}}) - L^{k_2}(g_{n_{k_2}})\| \leq \alpha^{k_1} \|L^0(g_{n_{k_1}}) - L^{k_2-k_1}(g_{n_{k_2}})\| + B \|L^0(g_{n_{k_1}}) - L^{k_2-k_1}(g_{n_{k_2}})\|_1.$$

Uma vez que $W^{1,1}$ é completo, a convergência é em $W^{1,1}$ para algum limite, que é forçado a ser h . Portanto, $h \in W^{1,1}$.

Mistura Considere os espaços forte e fraco de densidade média zero:

$$V_s = \left\{ g \in W^{1,1} : \int g \, dm = 0 \right\}$$

$$V_w = \left\{ g \in L^1 : \int g \, dm = 0 \right\}$$

onde m é medida de Lebesgue normalizada e considere as seguintes proposições:

Proposição 4.4. *Se para algum $f \geq 0$ tal que $\int f \, dm = 1$, $L^n f$ converge fracamente em L^1 para h , então T possui uma medida de probabilidade invariante absolutamente contínua com densidade h .*

Demonstração. Suponha, sem perda de generalidade, que $\int h \, dm = 1$. Por hipótese tem-se que $L^n f$ converge fracamente para $h \int f \, dm$, mas, isto ocorre se, e somente se, para todo $\phi \in L^\infty$ tem-se

$$\int \phi h \, dm = \lim_{n \rightarrow \infty} \int \phi L^{n+1} f \, dm = \lim_{n \rightarrow \infty} \int (\phi \circ T) L^n f \, dm = \int (\phi \circ T) h \, dm = \int \phi L h \, dm.$$

□

Proposição 4.5. *Se o sistema possui uma medida invariante absolutamente contínua com densidade $h \in L^1$ e, para todo $f \in L^1$ tal que $\int f \, dm = 0$, tem-se que $L^n f$ converge fracamente para 0, então T é misturador.*

Demonstração. Pelas hipóteses, dado um $f \in L^1$, tem-se que $L^n f - h \int f \, dm$ converge fracamente para 0, de modo que $L^n f$ converge fracamente para $h \int f \, dm$. Considere dois conjuntos mensuráveis A e B e denote por hm a medida com densidade h com respeito à medida de Lebesgue m . Tem-se

$$[hm](A \cap T^{-n}(B)) = \int \mathbb{1}_A(\mathbb{1}_B \circ T^n) h \, dm = \int \mathbb{1}_B L^n(h \mathbb{1}_A) \, dm$$

que, quando $n \rightarrow \infty$, converge para

$$\int \mathbb{1}_B \left(h \int h \mathbb{1}_A \, dm \right) dm = \int (h \mathbb{1}_B) \, dm \cdot \int (h \mathbb{1}_A) \, dm = [hm](E)[hm](F).$$

□

Proposição 4.6. *Para cada $g \in V_s$ vale que $\lim_{n \rightarrow \infty} \|L^n g\|_1 = 0$.*

Demonstração. Suponha que $\|g\|_\ell < \infty$. Pela equação (4.5) tem-se que $\|L^n g\|_\ell \leq \bar{M}$. Escreva $g = g^+ - g^-$, de modo que $\|g\|_1 = 2 \int g^+ dm$. Note que existe um ponto $\bar{x}$ tal que $g^+(\bar{x}) \geq \frac{1}{2} \|g\|_1$. Considere a vizinhança $N = B(\bar{x}, \frac{1}{4} \|g\|_1 \bar{M}^{-1})$ de $\bar{x}$ e observe que, para cada ponto $x \in N$, tem-se que $g^+(x) \geq \frac{1}{4} \|g\|_1$. Considere também $d = \min |T'|$ e $D = \max |T'|$. Se n_1 é o menor inteiro tal que $d^{n_1} \frac{1}{2} \|g\|_1 \bar{M}^{-1} > 1$, então $T^{n_1}(N) = S^1$ e $L^{n_1} g^+$ tem densidade ao menos

$$\frac{\|g\|_1}{4D^{n_1}} \geq \frac{\|g\|_1}{4De^{\log(2\|g\|_1^{-1}\bar{M}) \frac{\log D}{\log d}}} = \frac{\|g\|_1}{4D} (2\|g\|_1^{-1}\bar{M})^{\frac{\log D}{\log d}}$$

em S^1 . O mesmo vale para g^- . Deste modo, depois de n_1 iterações, a parte constante positiva se anula com a parte constante negativa correspondente e tem-se

$$\|L^{n_1} g\|_1 \leq \|g\|_1 - C \|g\|_1^{\frac{\log D}{\log d} + 1}$$

onde $C = \frac{(2\bar{M}^{-1})^{\frac{\log D}{\log d}}}{4D}$. Denote $g_1 = L^{n_1} g$ e repita a construção, de modo que

$$\|g_2\|_1 = \|L^{n_2} g_1\|_1 \leq \|g_1\|_1 - \|g_1\|_1^{\frac{\log D}{\log d} + 1}.$$

Assim, tem-se uma sequência tal que

$$\|g_{n+1}\|_1 \leq \|g_n\|_1 - \|g_n\|_1^{\frac{\log D}{\log d} + 1}$$

e então $\|g_n\|_1 \rightarrow 0$. Se $g \in W^{1,1}$, é possível aproximar g por $\tilde{g}$ tal que $\|\tilde{g}\|_\ell < \infty$, no sentido que $\|g - \tilde{g}\|_1 \leq \epsilon$. Uma vez que $\|L\|_{L^1 \rightarrow L^1} \leq 1$ tem-se que $\lim_{n \rightarrow \infty} \|L^n(g - \tilde{g})\|_1 \leq \epsilon$ e o resultado segue. $\square$

Pelas três proposições acima, segue:

Corolário 4.7. *Mapas expansores do círculo com medida $W^{1,1}$ invariante são misturadores.*

Ainda mais, note que se h_1 e h_2 são duas densidades de probabilidade invariantes, então $h_1 - h_2 \in V_s$ e é invariante, o que é impossível. Deste modo, existe uma única medida invariante em $W^{1,1}$ para mapas expansores em S^1 .

Convergência para o equilíbrio Para um tratamento um pouco mais geral, denote o espaço “forte” por $(B_s, \|\cdot\|_s)$ e o espaço “fraco” por $(B_w, \|\cdot\|_w)$. Diz-se que uma transformação $\mathfrak{T}$ em X , com operador de transferência associado $\mathfrak{L}$, tem convergência para o equilíbrio com velocidade ϕ , com respeito à essas normas, se para qualquer $f \in V_s$, onde V_s é o espaço forte de densidade média zero, tem-se

$$\|\mathfrak{L}^n f\|_w \leq \phi(n) \|f\|_s. \quad (4.6)$$

Se $\mathfrak{L}$ é um operador agindo no espaço de Banach complexo $(B, \|\cdot\|)$, o espectro de $\mathfrak{L}$ é definido por $\text{spec}(\mathfrak{L}) = \{\lambda : (\lambda \text{id} - \mathfrak{L}) \text{ não possui inversa limitada}\}$ e o raio espectral de $\mathfrak{L}$ por $\rho(\mathfrak{L}) = \sup\{|\lambda| : \lambda \in \text{spec}(\mathfrak{L})\}$. Outra forma de expressar o raio espectral é

$$\rho(\mathfrak{L}) = \lim_{n \rightarrow \infty} \sqrt[n]{\|\mathfrak{L}^n\|}. \quad (4.7)$$

Deste modo, é possível notar que existe uma conexão entre as propriedades espectrais do operador e o comportamento assintótico dos iterados. Diz-se que $\mathfrak{L}$ possui um *gap espectral* se $\mathfrak{L} = \lambda P + N$, onde P é uma projeção e $\dim(\text{Im}(P)) = 1$, $PN = NP = 0$ e $\rho(N) < |\lambda|$.

Agora, supondo que $\mathfrak{L}$ age nos espaços $(B_s, \|\cdot\|_s)$, $(B_w, \|\cdot\|_w)$ e que:

1. Para cada $g \in B_s$ tem-se $\|\mathfrak{L}^n g\| \leq A\lambda_1^n \|g\|_s + B\|g\|_w$;
2. Para cada $g \in V_s$ tem-se que $\lim_{n \rightarrow \infty} \|\mathfrak{L}^n g\|_w = 0$;
3. V_s é compactamente imerso em V_w ;
4. A norma fraca do operador restrita a V_s satisfaz $\sup_n \|\mathfrak{L}^n|_{V_s}\|_w < \infty$.

é possível demonstrar (veja [15, páginas 12 e 13]) que existem $C_2 > 0$ e $\rho_2 < 1$ tais que para todo $g \in V_s$

$$\|L^n g\|_s \leq C_2 \rho_2^n \|g\|_s. \quad (4.8)$$

Pela suposição 1 e pela equação (4.7), segue que o raio espectral de $\mathfrak{L}$ em B_s não pode ser maior que 1. Pela segunda hipótese, existe apenas um ponto fixo de $\mathfrak{L}$ em B_s e apenas uma medida de probabilidade invariante, que será denotada por h . Como existe uma medida invariante em B_s , então o raio espectral é igual a um. Para todo $g \in B_s$ pode-se escrever $g = [g - hg(X)] + [hg(X)]$. A função $P : B_s \rightarrow B_s$ definida como $P(g) = hg(X)$ é uma projeção e a função $N : B_s \rightarrow B_s$ definida como $N(g) = \mathfrak{L}[g - hg(X)]$ é tal que $N(B_s) \subset V_s$, $N|_{V_s} = \mathfrak{L}|_{V_s}$ e, pela equação (4.8) satisfaz $\rho(N) < |\lambda|$. Observe que vale que $\mathfrak{L} = P + N$ e $PN = NP = 0$. Deste modo, sob as hipóteses de 1 a 4, $\mathfrak{L}$ possui gap espectral.

Agora, note que pelo Teorema 4.3, Proposição 4.6 e a desigualdade de Lasota-Yorke, as hipóteses de 1 a 4 são verificadas para o aplicação expansora $T : S^1 \rightarrow S^1$, para a norma $W^{1,1}$ (com a norma L^1 para a norma fraca). Além disso, note que a equação (4.8)

implica convergência exponencial para o equilíbrio. Deste modo tem-se que T tem convergência para o equilíbrio e L possui gap espectral.

Estabilidade A estabilidade de um sistema está relacionada com sua resposta quando sofre perturbações. As perturbações podem ser determinísticas ou estocásticas, sendo que, no primeiro caso, o operador de transferência é perturbado por pequenas mudanças na dinâmica adjacente. Já as perturbações estocásticas possuem diversas formas, sendo que a mais simples consiste em adicionar algum ruído ao resultado da dinâmica determinística a cada iteração.

Sabe-se que, se uma medida varia continuamente, muitas propriedades estatísticas do sistema são estáveis sob perturbações. Porém, as medidas físicas podem mudar descontinuamente. É possível demonstrar que, sob certas condições relacionadas à convergência para o equilíbrio e o tipo de perturbação, a medida física pode ter uma variação contínua. Para isso, considere, novamente, os espaços $(B_s, \|\cdot\|_s)$, $(B_w, \|\cdot\|_w)$, V_s e V_w , juntamente com a transformação $\mathfrak{T}$, e considere uma família de operadores $\mathfrak{L}_\delta$, $\delta \in [0, 1)$, tal que $\mathfrak{L}_\delta(B_s) \subset B_s$ e $\mathfrak{L}_\delta(B_w) \subset B_w$. Suponha que:

UF1. Existem constantes $A, B, \lambda_1 \in \mathbb{R}$, com $\lambda_1 < 1$, tais que para todo $f \in B_s$, todo $n \geq 1$ e todo $\delta \in [0, 1)$, cada operador satisfaz uma desigualdade de Lasota-Yorke

$$\|\mathfrak{L}_\delta^n f\|_s \leq A\lambda_1^n \|f\|_s + B\|f\|_w. \quad (4.9)$$

UF2. $\mathfrak{L}_\delta$ aproxima $\mathfrak{L}_0$ quando δ é pequeno, no seguinte sentido: existe $C \in \mathbb{R}$ tal que, para todo $g \in B_s$,

$$\|(\mathfrak{L}_\delta - \mathfrak{L}_0)g\|_w \leq \delta C \|g\|_s. \quad (4.10)$$

UF3. $\mathfrak{L}_0$ possui convergência exponencial para o equilíbrio, relativamente à $\|\cdot\|_w$ e $\|\cdot\|_s$.

UF4. Existe M tal que, para todo $\delta \in [0, 1)$, $n \in \mathbb{N}$ e $g \in B_s$, tem-se $\|\mathfrak{L}_\delta^n g\|_w \leq M\|g\|_w$.

Sob essas hipóteses, pode-se garantir que a medida invariante do sistema varia continuamente na norma fraca quando $\mathfrak{L}_0$ é perturbado para $\mathfrak{L}_\delta$, para pequenos valores de δ . Observe que as condições UF1 e UF3 juntas implicam que $\mathfrak{L}_0$ eventualmente contrai V_s . De fato, seja $f \in V_s$, então

$$\begin{aligned} \|\mathfrak{L}_0^{n+m} f\|_s &\leq A\lambda_1^n \|\mathfrak{L}_0^m f\|_s + B\|\mathfrak{L}_0^m f\|_w \\ &\leq A\lambda_1^n \|\mathfrak{L}_0^m f\|_s + BE\lambda_2^m \|f\|_s \\ &\leq A\lambda_1^n (B + A)\|f\|_s + BE\lambda_2^m \|f\|_s \end{aligned}$$

para os quais existem m, n suficientemente grande tais que $\|\mathfrak{L}_0^{n+m} f\|_s \leq \frac{1}{2}\|f\|_s$.

Antes de prosseguir, considere que os operadores $\mathfrak{L}_0$ e $\mathfrak{L}_\delta$ preservam um espaço normado de medidas $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ contido no conjunto das medidas borelianas em X . Suponha que $f_0, f_\delta \in \mathcal{B}$ são medidas de probabilidade fixas, de $\mathfrak{L}_0$ e $\mathfrak{L}_\delta$, respectivamente.

Lema 4.8. *Suponha que*

$$a. \|\mathfrak{L}_\delta f_\delta - \mathfrak{L}_0 f_\delta\|_{\mathcal{B}} < \infty.$$

$$b. \mathfrak{L}_0^i \text{ é contínua em } \mathcal{B}, \text{ de modo que existe } C_i \text{ tal que para todo } g \in \mathcal{B}, \|\mathfrak{L}_0^i g\|_{\mathcal{B}} \leq C_i \|g\|_{\mathcal{B}}.$$

Então, para cada N , tem-se

$$\|f_\delta - f_0\|_{\mathcal{B}} \leq \|\mathfrak{L}_0^N(f_\delta - f_0)\|_{\mathcal{B}} + \|\mathfrak{L}_\delta f_\delta - \mathfrak{L}_0 f_\delta\|_{\mathcal{B}} \sum_{i \in [0, N-1]} C_i \quad (4.11)$$

A demonstração deste lema provém de cálculo direto (para detalhes, consulte [15, página 16]). Agora, supondo $B_w = \mathcal{B}$, segue:

Proposição 4.9. *Suponha que $\mathfrak{L}_\delta$ é uma família de operadores satisfazendo as condições UF1 - UF4, f_0 é a única medida de probabilidade invariante de $\mathfrak{L}_0$ e f_δ é uma medida de probabilidade invariante de $\mathfrak{L}_\delta$. Então $\|f_\delta - f_0\|_w = \mathcal{O}(\delta \log \delta)$.*

Demonstração. Pela desigualdade de Lasota-Yorke uniforme, tem-se que $\|f_\delta\|_s \leq M$ são uniformemente limitados, e então, $\|\mathfrak{L}_\delta f_\delta - \mathfrak{L}_0 f_\delta\|_w \leq \delta CM$. Por UF4, $C_i \leq M_2$, de modo que

$$\|f_\delta - f_0\|_w \leq \delta CMM_2N + \|\mathfrak{L}_0^N(f_0 - f_\delta)\|_w.$$

Pela convergência exponencial para o equilíbrio de $\mathfrak{L}_0$, tem-se

$$\|\mathfrak{L}_0^N(f_0 - f_\delta)\|_w \leq C_2 \rho_2^N \|(f_\delta - f_0)\|_s \leq C_2 \rho_2^N M$$

e então

$$\|(f_\delta - f_0)\|_{\mathcal{B}} \leq \delta CMM_2N + C_2 \rho_2^N M.$$

Tomando $N = \lfloor \frac{\log \delta}{\log \rho_2} \rfloor$, segue que

$$\|(f_\delta - f_0)\|_{\mathcal{B}} \leq \delta \log \delta C M_2 M \frac{1}{\log \rho_2} + C_2 \delta M.$$

□

É possível aplicar esse tratamento geral para a aplicação expansora $T : S^1 \rightarrow S^1$, considerando uma família de transformações expansoras T_δ , $\delta \in [0, 1)$ que satisfaz

$$\|T_\delta - T_0\| \leq K\delta \quad (4.12)$$

para algum $K \in \mathbb{R}$. Para cada um desses mapas, associe um operador de transferência L_δ agindo em $W^{1,1}$. Note que esta família satisfaz as condições *UF1*, *UF3* e *UF4*. Realizando alguns cálculos [15, Proposição 26], mostra-se que a equação (4.12) implica a condição *UF2*, ou seja, se L_0 e L_δ são operadores de transferência de mapas de expansão T_0 e T_δ satisfazendo a equação (4.12), então existe $C \in \mathbb{R}$ tal que para todo $g \in W^{1,1}$

$$\|(L_\delta - L_0)f\|_1 \leq \delta C \|f\|_{W^{1,1}}. \quad (4.13)$$

Com isso, pode-se aplicar a Proposição 4.9 para provar a estabilidade de uma medida invariante para essa família de aplicações e ter uma estimativa explícita para o módulo de continuidade: Considerando h_δ uma família de medidas invariante em L^1 para os mapas T_δ , tem-se que $\|h_0 - h_\delta\|_1 = \mathcal{O}(\delta \log \delta)$.

4.1.2 Mapa solenoidal

Nesta seção, é considerado um sistema cuja dinâmica é modelada por um mapa solenoidal F , tendo como base o exemplo apresentado em [15, Seção 10]. A desintegração de medidas abordada na Seção 3.2 é aplicada neste exemplo para a definição de espaços adaptados para o sistema e, considerando o operador de transferência associado à F , demonstra-se a existência de uma medida invariante.

Considere mapa solenoidal $F : \Sigma \rightarrow \Sigma$, onde $\Sigma = S^1 \times D^2$, tal que F é o *skew product* $F(x, y) = (T(x), G(x, y))$ onde $T : S^1 \rightarrow S^1$ é transformação expansora e $G : \Sigma \rightarrow D^2$ é uma aplicação diferenciável com $\|\frac{\partial G}{\partial x}\|_\infty < \infty$. Adicionalmente, suponha que T possui grau q , e, com abuso de notação, considere $T : [0, 1] \rightarrow [0, 1]$, cujas ramificações são T_i , $i = 1, 2, \dots, q$. Considere a folheação $\mathcal{F}^s = \{\{x\} \times D^2\}_{x \in S^1}$, de tal modo que $\mathcal{F}^s$ seja contração, ou seja, que exista um $\alpha \in \mathbb{R}$, $0 < \alpha < 1$, tal que para todo $x \in S^1$ tem-se $|G(x, y_1) - G(x, y_2)| \leq \alpha|y_1 - y_2|$, para todo $y_1, y_2 \in D^2$.

Denote por $\mathcal{AB}(\Sigma)$ o conjunto $\{\mu \in \mathcal{SB}(\Sigma) : \pi_{x*}\mu^+ \ll m \text{ e } \pi_{x*}\mu^- \ll m\}$, onde $\pi_x : \Sigma \rightarrow S^1$ é a projeção dada por $(x, y) \mapsto x$ e m é a medida de Lebesgue em S^1 . Considere uma medida de probabilidade $\mu \in \mathcal{AB}(\Sigma)$ e a medida $\mu_x = \pi_{x*}\mu$, que pode ser identificada com uma densidade marginal não negativa $\phi_x : S^1 \rightarrow \mathbb{R}$. Além disso, denote a folheação $\mathcal{F}^s$ por $\{\zeta_x\}_{x \in S^1}$, com $\zeta_x = \{x\} \times D^2$ para cada $x \in S^1$. Note que a abordagem da Seção 3.2 é válida para este sistema. Deste modo, obtêm-se uma família $\{\mu_{\zeta_x} : x \in S^1\} \subset \mathcal{P}(D^2)$, μ_x -quase sempre unicamente determinada, tal que $\mu = \int_{S^1} \mu_{\zeta_x} d\mu_x$.

Com essa desintegração é possível construir um espaço normado de medidas com sinal para estudar propriedades estatísticas relacionadas à este sistema:

Definição 4.10. *Seja $\mathfrak{L}^1 \subset \mathcal{AB}(\Sigma)$ definido por*

$$\mathfrak{L}^1 = \left\{ \mu \in \mathcal{AB}(\Sigma) : \int_{S^1} W_1^0(\mu^+|_{\zeta_x}, \mu^-|_{\zeta_x}) dm < \infty \right\} \quad (4.14)$$

onde $\mu^+|_{\zeta_x}, \mu^-|_{\zeta_x}$ representam as partes positiva e negativa, respectivamente, da decomposição de Jordan da restrição de μ em ζ_x , no sentido das seções 3.2.1 e 3.2.2. Defina uma norma em $\mathfrak{L}^1$, $\|\cdot\|_{\mathfrak{L}^1} : \mathfrak{L}^1 \rightarrow \mathbb{R}$ por

$$\|\mu\|_{\mathfrak{L}^1} = \int_{S^1} W_1^0(\mu^+|_{\zeta_x}, \mu^-|_{\zeta_x}) dm. \quad (4.15)$$

Observação. *Para um espaço métrico compacto (Γ, d) , $g : \Gamma \rightarrow \mathbb{R}$ uma função Lipschitz e*

$$L(g) = \sup_{x,y \in \Gamma} \left\{ \frac{|g(x) - g(y)|}{d(x,y)} \right\},$$

dadas μ e ν duas medidas com sinal em Γ , a distância 1-Wasserstein entre μ e ν pode ser escrita como

$$W_1^0(\mu, \nu) = \sup_{L(g) \leq 1, \|g\|_\infty \leq 1} \left| \int g d\mu - \int g d\nu \right|. \quad (4.16)$$

Denote $W_1^0(0, \mu)$ por $\|\mu\|_W$. Tem-se que $\|\cdot\|_W$ define uma norma no espaço vetorial das medidas com sinal definidas em um espaço métrico compacto.

Ainda, é válido ressaltar uma propriedade da norma $\|\cdot\|_W$ relacionada com o operador de transferência: dado um espaço métrico Γ qualquer, se a aplicação $H : \Gamma \rightarrow \Gamma$ é uma contração fraca e L_H é o operador de transferência associado à H , tem-se que

$$\|L_H \mu\|_W \leq \|\mu\|_W \quad (4.17)$$

para toda $\mu \in \mathcal{SM}(\Gamma)$. De fato, uma vez que H é contração, se g é tal que $\|g\|_\infty \leq 1$ e $Lip(g) \leq 1$, o mesmo vale para $g \circ H$ e então

$$\left| \int g d(L_H(\mu)) \right| = \left| \int g \circ H d\mu \right| \leq \|\mu\|_W.$$

Tomando o supremo em $\|g\|_\infty \leq 1$ e $Lip(g) \leq 1$, obtém-se a desigualdade (4.17).

Seja L_F o operador de transferência associado a F , i. e., $[L_F\mu](E) = \mu(F^{-1}(E))$ para $\mu \in \mathcal{AB}(\Sigma)$ e $E \subset \Sigma$ conjunto mensurável. Em [16] é abordada uma caracterização do operador de transferência que o faz funcionar como um operador unidimensional. Apesar de ser tratado no artigo de modo mais geral, esta caracterização se aplica a este caso, fornecendo:

Proposição 4.11. *Para uma folha $\zeta_x \in \mathcal{F}^s$, defina o mapa $F_{\zeta_x} : D^2 \rightarrow D^2$ por*

$$F_{\zeta_x} = \pi_y \circ F|_{\zeta_x} \circ \pi_y|_{\zeta_x}^{-1}$$

onde $\pi_y|_{\zeta_x}$ é a projeção π_y definida por $(x, y) \mapsto y$, restrita à folha ζ_x . Para todo $\mu \in \mathfrak{L}^1$ e para quase todo $x \in S^1$, tem-se que

$$(L_F\mu)|_{\zeta_x} = \sum_{i=1}^q \frac{F_{T_i^{-1}(x)}_* \mu|_{T_i^{-1}(x)}}{|T'_i \circ T_i^{-1}(x)|}. \quad (4.18)$$

Agora, são exploradas algumas propriedades associadas à L_F , a fim de mostrar a existência de uma medida invariante em $\mathfrak{L}^1$ para o sistema solenoidal.

Proposição 4.12. *Se $\mu \in \mathfrak{L}^1$, então $\|L_F\mu\|_{\text{"1"}} \leq \|\mu\|_{\text{"1"}}$.*

Demonstração. Considere, para todo i , a mudança de variável $\gamma = T_i(\alpha)$. Então

$$\begin{aligned} \|L_F\mu\|_{\text{"1"}} &= \int_{S^1} \|(L_F\mu)|_{\zeta_x}\|_W dm \\ &\leq \sum_{i=1}^q \int_{T(x_i)} \left\| \frac{F_{T_i^{-1}(\gamma)}_* \mu|_{T_i^{-1}(\gamma)}}{|T'_i \circ T_i^{-1}(\gamma)|} \right\|_W dm \\ &= \sum_{i=1}^q \int_{x_i} \|F_{\alpha*} \mu|_{\alpha}\|_W dm(\alpha) \\ &= \sum_{i=1}^q \int_{x_i} \|\mu|_{\alpha}\|_W dm(\alpha) \\ &= \|\mu\|_{\text{"1"}} \end{aligned}$$

□

Considere o seguinte resultado:

Lema 4.13. *Seja $F : \Gamma \rightarrow \Gamma$ uma contração, onde Γ é um espaço métrico compacto. Para toda $\mu \in \mathcal{SM}(\Gamma)$, vale $\|F_*\mu\|_W \leq \alpha\|\mu\|_W + \mu(\Gamma)$, onde α é a taxa de contração de F .*

Demonstração. Se g é tal que $Lip(g) \leq 1$ e $\|g\|_{\infty} \leq 1$, então tem-se que $g \circ F$ é α -Lipschitz

e $\|g \circ F_{\zeta_x} - \theta\|_\infty \leq \alpha$ para algum $\theta \leq 1$. Assim,

$$\begin{aligned} \left| \int g \, d(F_*\mu) \right| &= \left| \int g \circ F \, d\mu \right| \\ &= \left| \int g \circ F - \theta \, d\mu \right| + \left| \int \theta \, d\mu \right| \\ &= \alpha \left| \int \frac{1}{\alpha} g \circ F - \theta \, d\mu \right| + \theta \mu(\Gamma) \\ &\leq \alpha \|\mu\|_W + \mu(\Gamma). \end{aligned}$$

Tomando o supremo em $Lip(g) \leq 1$ e $\|g\|_\infty \leq 1$, segue o resultado. $\square$

Com esse resultado, pode-se demonstrar a seguinte proposição.

Proposição 4.14. *Para toda medida com sinal $\mu \in \mathfrak{L}^1$ tem-se*

$$\|F_*\mu\|_{\mathfrak{L}^1} \leq \alpha \|\mu\|_{\mathfrak{L}^1} + (\alpha + 1) \|\phi_x\|_1. \quad (4.19)$$

Demonstração. Veja [15, Proposição 57]. $\square$

Note que, iterando a equação (4.19) obtém-se $\|L_F^n \mu\|_{\mathfrak{L}^1} \leq \alpha^n \|\mu\|_{\mathfrak{L}^1} + \bar{\alpha} \|\phi_x\|_1$, onde $\bar{\alpha} = \frac{1+\alpha}{1-\alpha}$. Considere o conjunto das medidas média zero dado por

$$\mathcal{V} = \{\mu \in \mathfrak{L}^1 : \mu(\Sigma) = 0\}.$$

Uma vez que $\pi_{x*}\mu = \mu_x = \phi_x m$, tem-se que $\int \phi_x \, dm = 0$ se $\mu \in \mathcal{V}$. Deste modo, pela equação de iteração acima e pela convergência para o equilíbrio para aplicações expansoras com respeito às normas de L^1 e $W^{1,1}$, discutidas na seção anterior, segue:

Proposição 4.15 (Convergência exponencial para o equilíbrio). *Existem $D \in \mathbb{R}$ e $0 < \beta_1 < 1$ tais que, para toda $\mu \in \mathcal{V}$ e todo $n \geq 1$, tem-se*

$$\|L_F^n \mu\|_{\mathfrak{L}^1} \leq D_2 \beta_1^n (\|\mu\|_1 + \|\phi_x\|_{W^{1,1}}).$$

Agora pode-se demonstrar a existência de uma medida invariante no conjunto $\mathfrak{L}^1$ para o sistema solenoidal.

Proposição 4.16. *Existe uma única $\mu \in \mathfrak{L}^1$ tal que $L_F\mu = \mu$.*

Demonstração. O mapa T é expansor e possui uma medida invariante absolutamente contínua, que será denominada φ_x . Considere a medida $\nu = \varphi_x \times m$ e a sequência $\nu_n = L_F^n \nu$. Pela Proposição 4.15 tem-se que $\|\nu_n - \nu_p\|_{1''} \leq D_2 \beta_1^n$, de modo que ν_n é uma sequência de Cauchy. Considere uma subsequência ν_{n_k} tal que, para cada folha ζ_x , $\nu_{n_k}|_{\zeta_x}$ é uma sequência de Cauchy para a norma $\|\cdot\|_W$. Como o conjunto das medidas de Borel positivas em Σ é completo com respeito à $\|\cdot\|_W$, essa sequência possui um limite, que é uma medida positiva, e isso define uma medida limite μ , que é invariante. A integrabilidade de $\|\nu|_{\zeta_x}\|_W$ segue do fato de que $\|\nu_{n_k}|_{\gamma}\|_W \leq \sup \varphi_x$ (e então é uma sequência limitada) e $\|\nu_{n_k}|_{\zeta_x}\|_W$ converge pontualmente para $\|\nu|_{\zeta_x}\|_W$, e a unicidade segue diretamente da Proposição 4.15. $\square$

No artigo [15] ainda é apresentada a construção de um espaço normado “forte” para este sistema, no qual existe uma desigualdade de Lasota-Yorke para o operador de transferência aplicado à medidas não negativas e, ainda, é demonstrado que a iteração L_F^n mantém uma regularidade limitada nessa norma “forte”.

Nestes exemplos observa-se como este tipo de abordagem, considerando o operador de transferência, permite obter inúmeras informações relevantes relacionadas às propriedades estatísticas de sistemas dinâmicos. Além disso, este exemplo do sistema solenoidal ilustra como a desintegração de medidas é um conceito bastante versátil, no sentido que pode ser aplicado em diversos cenários.

Referências

- [1] Ambrosio, L. (2000). Lecture Notes on Optimal Transport Problems. Lectures notes, Euro Summer School “Mathematical aspects of evolving interfaces”.
- [2] Ambrosio, L. & Gigli, N. (2013). A user’s guide to Optimal Transport. In L. Ambrosio et. al., Modelling and Optimisation of Flows on Networks. Lecture Notes in Mathematics, 2062. Springer Verlag Berlin.
- [3] Ambrosio, L. & Gigli, N. & Savaré, G. (2005). Gradient Flows in Metric Spaces and in the Space of Probability Measures. Birkhäuser Verlag.
- [4] Ambrosio L. & Pratelli A. (2003). Existence and stability results in the L1 theory of optimal transportation. In: Optimal Transportation and Applications. Lecture Notes in Mathematics (Fondazione C.I.M.E., Firenze), 1813. Springer, Berlin.
- [5] Bartle, R. G. (1966). The Elements of Integration. John Wiley & Sons.
- [6] Bergstrom, H. (1982). Weak convergence of measures, Probability and Mathematical Statistics. Academic Press.
- [7] Boyarsky, A. & Góra, P. F. (1997). Laws of Chaos : Invariant Measures and Dynamical Systems in One Dimension.
- [8] Brin M. & Stuck, G. (2003). Introduction to Dynamical Systems. Cambridge University Press.
- [9] Camacho, C. & Lins Neto, A. (1985). Geometric Theory of Foliations. Birkhäuser.
- [10] Cohn, D. (1993). Measure theory. Reprint of the 1980 original. Birkhauser Boston.
- [11] Dellacherie, C. & Meyer, P. A. (1978). Probabilities and potential. North-Holland Mathematics Studies, 29. North-Holland Publishing Co. Amsterdam.
- [12] Do Carmo, M. P. (1976). Differential geometry of curves and surfaces. Prentice-Hall.
- [13] Folland, G. B. (1984). Real Analysis: Modern techniques and their applications. Wiley-Interscience.

- [14] Galatolo, S. (2018). Quantitative statistical stability, speed of convergence to equilibrium and partially hyperbolic skew products. *Journal de l'École polytechnique Mathématiques*.
- [15] Galatolo, S. (2017). Statistical properties of Dynamics: Introduction to the functional analytic approach. *arXiv:1510.02615v2*.
- [16] Galatolo, S. & Lucena, R. (2018). Spectral gap and quantitative statistical stability for systems with contracting fibers and lorenz-like maps. *arXiv:1507.08191v8*.
- [17] Glendinning, P. (1994). *Stability, instability and chaos*. Cambridge University Press.
- [18] Granieri L. & Maddalena, F. (2013). Transport problems and disintegration maps. *ESAIM: Control, Optimisation and Calculus of Variations*, 19 n. 3, pp. 888-905.
- [19] Layek, G. C. (2015). *An introduction to dynamical systems and chaos*. Springer.
- [20] Milnor, J. W. (1965). *Topology from the differential point of view*. The University Press of Virginia.
- [21] Oliveira, K. & Viana, M. (2014). *Fundamentos da Teoria Ergódica*. IMPA.
- [22] Rokhlin, V. A. (1962). On the fundamental ideas of measure theory. *American Mathematical Society Translations*, 10. 1-52.
- [23] Rudin, W. (1987). *Real and Complex Analysis*. McGraw-Hill.
- [24] Sarig, O. (2020). Introduction to the transfer operator method. *Winter School on Dynamics*. Hausdorff Research Institute for Mathematics.
- [25] Simmons, D. (2012). Conditional measures and conditional expectation; Rohlin's disintegration theorem. *Discrete & Continuous Dynamical Systems*, 32, 2565-2582.
- [26] Srivastava, S. M. (1998). *A course on Borel sets*. Graduate Texts in Mathematics, 180. Springer-Verlag.
- [27] Sternberg, S. (2014). *Dynamical Systems*. Dover Publications.
- [28] Viana, M. (2008). Disintegration into conditional measures: Roklin's theorem. *Página pessoal*. IMPA.
- [29] Villani, C. (2003). *Topics in Optimal Transportation*. Graduate Studies in Mathematics, 58. American Mathematical Society.
- [30] Villani, C. (2009). *Optimal Transport: old and new*. Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag Berlin Heidelberg.
- [31] Von Neumann, J. (1932). Zur Operatorenmethode in der klassischen Mechanik. *Annals of Mathematics*, 33, 587-642.