

# Sumário

<b>1</b>	<b>Resumo do projeto</b>	<b>2</b>
<b>2</b>	<b>Atividades realizadas no período</b>	<b>3</b>
	Formas diferenciais e variedades . . . . .	3
	Variedades diferenciáveis, integração e o Teorema de Stokes . . . . .	5
	Dinâmicas expansoras . . . . .	6
	Cones e métricas projetivas . . . . .	7
	Operadores de transferência e cones invariantes . . . . .	8
	Estabilidade estocástica . . . . .	9
	Entropia em sistemas dinâmicos . . . . .	9
	Teoria da Informação . . . . .	9
	Entropia topológica . . . . .	12
<b>3</b>	<b>Participação em evento científico</b>	<b>14</b>
<b>4</b>	<b>Lista dos trabalhos preparados ou submetidos</b>	<b>15</b>

# Resumo do projeto

O projeto teve como objetivo expor o aluno de graduação a uma introdução à Teoria Ergódica, e em particular, ao estudo de diversas propriedades ergódicas de sistemas dinâmicos. A área de Sistemas Dinâmicos, na matemática, tem como um de seus objetivos mais básicos entender o comportamento qualitativo das aplicações iteradas  $f \circ f \circ \dots \circ f$  de uma mesma função  $f: X \rightarrow X$ . Em muitos casos, é mais produtivo entender tal comportamento de forma estatística, colocando uma medida de probabilidade no conjunto e buscando resultados que são válidos para quase todo ponto. Nesse sentido, a Teoria Ergódica se preocupa com essas dinâmicas quando são munidas de medidas invariantes, e dispõe em seus resultados mais fundamentais algumas consequências estatísticas da existência dessas medidas no comportamento do sistema.

Uma classe de sistemas com comportamento interessante do ponto de vista ergódico são os difeomorfismos expansores em variedades compactas. Com o intuito de dar ao aluno mais familiaridade com a presença desses objetos, e também a fim de conhecimento geral, iniciamos o período com tópicos introdutórios sobre variedades e formas diferenciais, seguindo os livros *Smooth Manifolds*[2] e *Formas Diferenciais e Aplicações*[4]. Em seguida, estudamos as propriedades ergódicas dos difeomorfismos expansores a partir do livro *Stochastic Dynamics of Deterministic Systems*[5], vendo também como esses sistemas são estocasticamente estáveis.

Por fim, se a Teoria Ergódica nos diz como caracterizar um sistema a partir de uma medida, também é interessante entender quando dois sistemas possuem comportamento equivalente por esse ponto de vista. Nisso, são importantes as invariantes por equivalência ergódica, propriedades que são compartilhadas por quaisquer dois sistemas ergodicamente equivalentes. As entropias topológica e de Kolmogorov-Sinai são duas dessas invariantes, sendo originalmente desenvolvidas como uma forma de discernir sistemas que não eram distinguíveis por outras invariantes conhecidas. No projeto, também estudamos as definições e alguns resultados básicos sobre ambas entropias.

# Atividades realizadas no período

## Formas diferenciais e variedades

Para cada ponto  $p \in \mathbb{R}^n$ , o *espaço tangente em  $p$*  pode ser pensado como o conjunto de vetores que partem de  $p$ . Mais concretamente, representaremos esse espaço como o conjunto  $\mathbb{R}_p^n = \{v_p : v \in \mathbb{R}^n\}$  que herda uma estrutura de espaço vetorial, e onde cada vetor carrega na notação uma “etiqueta” que o associa ao ponto  $p$ . Sendo um espaço vetorial, podemos falar no espaço dual  $(\mathbb{R}_p^n)^* = \text{Hom}_{\mathbb{R}}(\mathbb{R}_p^n, \mathbb{R})$  e no espaço das  $k$ -formas lineares  $\text{Hom}_{\mathbb{R}}^k(\mathbb{R}_p^n, \mathbb{R})$ , ou seja, das funções  $\phi: \underbrace{\mathbb{R}_p^n \times \cdots \times \mathbb{R}_p^n}_{k \text{ vezes}} \rightarrow \mathbb{R}$  que são lineares quando se fixam todas menos uma

das entradas. Denotaremos por  $\bigwedge^k(\mathbb{R}_p^n)^* \subseteq \text{Hom}_{\mathbb{R}}^k(\mathbb{R}_p^n, \mathbb{R})$  o subespaço formado pelas formas *alternadas*, ou seja, tais que

$$\varphi(v_1, \dots, v_i, v_{i+1}, \dots, v_k) = -\varphi(v_1, \dots, v_{i+1}, v_i, \dots, v_k)$$

para todo  $1 \leq i < k$ . Dadas  $\varphi_1, \dots, \varphi_k \in (\mathbb{R}_p^n)^*$ , podemos construir um elemento de  $\bigwedge^k(\mathbb{R}_p^n)^*$  dado por

$$(\varphi_1 \wedge \cdots \wedge \varphi_k)(v_1, \dots, v_k) = \det([\varphi_i(v_j)]_{i,j}),$$

onde  $[\varphi_i(v_j)]_{i,j}$  é uma matriz  $k \times k$ .

Seja  $\mathcal{C}^\infty(\mathbb{R}^n)$  o conjunto das funções  $\mathbb{R}^n \rightarrow \mathbb{R}$  infinitamente diferenciáveis. Notamos que para toda  $f \in \mathcal{C}^\infty(\mathbb{R}^n)$  e  $p \in \mathbb{R}^n$ , a *diferencial*  $(df)_p$  no ponto  $p$  é um elemento de  $(\mathbb{R}_p^n)^*$  que leva um vetor  $v \in \mathbb{R}_p^n$  ao valor da derivada direcional de  $f$  ao longo de  $v$  no ponto  $p$ . Assim, sendo  $x_i: \mathbb{R}^n \rightarrow \mathbb{R}$  a função que retorna a  $i$ -ésima coordenada de um vetor, as derivadas  $\{(dx_i)_p : 1 \leq i \leq n\}$  formam a base dual da base canônica de  $\mathbb{R}_p^n$ , e o conjunto

$$\beta = \{(dx_{i_1} \wedge \cdots \wedge dx_{i_k})_p : i_j \in \{1, \dots, n\} \text{ para cada } j \text{ e } i_1 < \cdots < i_k\}$$

forma uma base de  $\bigwedge^k(\mathbb{R}_p^n)^*$ . Notamos que cada escolha  $i_1, \dots, i_k$  de índices na direita determina um subconjunto de  $k$  elementos de  $\{1, \dots, n\}$ , e reciprocamente, todo subconjunto de  $k$  elementos pode ser ordenado em ordem crescente, e assim corresponde a uma escolha de  $i_1, \dots, i_k$ . Como existem  $\binom{n}{k}$  desses subconjuntos, o espaço  $\bigwedge^k(\mathbb{R}_p^n)^*$  possui dimensão  $\binom{n}{k}$ .

Uma  *$k$ -forma exterior* de  $\mathbb{R}^n$  é uma função  $\omega: \mathbb{R}^n \rightarrow \bigsqcup_{p \in \mathbb{R}^n} \bigwedge^k(\mathbb{R}_p^n)^*$  tal que  $\omega(p) \in \bigwedge^k(\mathbb{R}_p^n)^*$ . Como o conjunto  $\beta$  é uma base do espaço das formas alternadas em  $p$ , podemos escrever

$$\omega(p) = \sum_{i_1 < \cdots < i_k} a_{i_1, \dots, i_k}(p)(dx_{i_1} \wedge \cdots \wedge dx_{i_k})_p,$$

onde  $a_{i_1, \dots, i_k}$  são funções  $\mathbb{R}^n \rightarrow \mathbb{R}$ . Por convenção, uma 0-forma é uma função  $\omega: \mathbb{R}^n \rightarrow \mathbb{R}$ . Se todas as funções  $a_{i_1, \dots, i_k}$  forem diferenciáveis,  $\omega$  é dita uma  *$k$ -forma diferencial*. Denotaremos

por  $\Omega^k(\mathbb{R}^n)$  o espaço vetorial formado pelas  $k$ -formas diferenciais de  $\mathbb{R}^n$ , onde as operações são herdadas de  $\bigwedge^k(\mathbb{R}_p^n)^*$  em cada ponto  $p$ .

Podemos estender a operação  $\wedge$  para qualquer par de formas do mesmo espaço, da seguinte maneira: sendo  $\omega = \sum_I a_I dx_I$  uma  $k$ -forma e  $\theta = \sum_J b_J dx_J$  uma  $l$ -forma, definimos o seu produto cunha como a  $k+l$ -forma

$$(\omega \wedge \theta)(p) = \sum_{I,J} a_I(p)b_J(p)(dx_I \wedge dx_J)_p,$$

onde  $I = (i_1, \dots, i_k)$  e  $J = (j_1, \dots, j_l)$  são abreviações para tuplas de índices em ordem crescente, como no parágrafo anterior. Se uma das formas for uma 0-forma, o produto cunha simplesmente multiplica 0-forma pela outra forma em cada ponto. Para formas  $\omega$ ,  $\theta$  e  $\varphi$ , não é difícil mostrar que vale  $(\omega + \theta) \wedge \varphi = \omega \wedge \varphi + \theta \wedge \varphi$ ,  $\varphi \wedge \omega + \varphi \wedge \theta = \varphi \wedge (\omega + \theta)$  e  $(\lambda\omega) \wedge (\mu\theta) = (\lambda\mu)(\omega \wedge \theta)$  para  $\lambda, \mu \in \mathbb{R}$ , fazendo do espaço  $\Omega(\mathbb{R}^n) = \bigoplus_{k=0}^{\infty} \Omega^k(\mathbb{R}^n) = \bigoplus_{k=0}^n \Omega^k(\mathbb{R}^n)$  das formas diferenciais em  $\mathbb{R}^n$  uma álgebra sobre o corpo  $\mathbb{R}$ .

Fixada  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ , cada  $k$ -forma diferencial  $\omega$  de  $\mathbb{R}^m$  se associa de forma natural a uma  $k$ -forma  $f^*\omega$  de  $\mathbb{R}^n$ , seu *pullback* por  $f$ :

$$(f^*\omega)(p)(v_1, \dots, v_k) = (\omega \circ f)(p)(df_p(v_1), \dots, df_p(v_k)).$$

O mapa  $f^*: \Omega(\mathbb{R}^m) \rightarrow \Omega(\mathbb{R}^n)$  é linear e também é um homomorfismo de álgebras, no sentido que  $f^*(\omega \wedge \theta) = f^*(\omega) \wedge f^*(\theta)$ .

É simples ver que  $\bigwedge^1(\mathbb{R}_p^n)^*$  é isomorfo a  $(\mathbb{R}_p^n)^*$  como espaço vetorial. Assim, para toda  $g \in \mathcal{C}^\infty(\mathbb{R}^n)$ , a aplicação

$$p \mapsto (dg)_p = \sum_{i=1}^n \frac{\partial g}{\partial x_i} \Big|_p (dx_i)_p$$

é uma 1-forma diferencial, denotada por  $dg$ . Estenderemos esse operador  $d$ , que será chamado de *derivada exterior*, para todas as formas da seguinte maneira: se  $\omega = \sum_I a_I dx_I$  é uma  $k$ -forma,

$$d\omega = \sum_I da_I \wedge dx_I.$$

Essa extensão possui duas propriedades que também a determinam unicamente como operador linear:

1.  $d(df) = 0$  para toda  $f \in \mathcal{C}^\infty(\mathbb{R}^n)$ ;
2.  $d(\omega \wedge \theta) = d\omega \wedge \theta + (-1)^k \omega \wedge d\theta$ , onde  $\omega$  é uma  $k$ -forma.

É bastante simples integrar  $n$ -formas diferenciais em  $\mathbb{R}^n$ : sabemos que toda  $n$ -forma pode ser escrita como  $\omega(x) = a(x)dx_1 \wedge \dots \wedge dx_n$ . Assim, sendo o conjunto  $K = \{x : a(x) \neq 0\}$  compacto, definimos:

$$\int \omega = \int_K a(x_1, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

## Variedades diferenciáveis, integração e o Teorema de Stokes

No primeiro período da iniciação científica, vimos algumas propriedades básicas sobre variedades diferenciáveis. Já nesse segundo momento, passamos a estudar a maneira pela qual as formas diferenciais em  $\mathbb{R}^n$  se generalizam a esses objetos.

Seja  $\mathbb{H}^n = \{x \in \mathbb{R}^n : x_1 \geq 0\}$  um semi-espaço fechado de  $\mathbb{R}^n$ . Uma função entre abertos desses semi-espaços será dita *diferenciável* se for possível estender seu domínio a um aberto de  $\mathbb{R}^n$  tal que a extensão seja  $C^\infty$ . Relembramos que uma *variedade diferenciável de dimensão  $n$  com (ou sem) bordo* é um espaço topológico  $M$  satisfazendo o axioma de Hausdorff e que possui uma base enumerável de vizinhanças, juntamente a uma família de *cartas*  $\{(f_i, U_i)\}_{i \in \mathcal{I}}$  (chamada de *atlas diferenciável*) satisfazendo:

1. cada  $U_i$  é aberto e cada  $f_i: U_i \subseteq M \rightarrow \mathbb{H}^n$  é um homeomorfismo sobre sua imagem;
2.  $\bigcup_i U_i = M$ ;
3. para cada par de cartas  $(f_i, U_i)$  e  $(f_j, U_j)$  no atlas, se  $U_i \cap U_j = W \neq \emptyset$  então a aplicação  $f_j \circ f_i^{-1}: f_i(W) \rightarrow f_j(W)$  é infinitamente diferenciável;
4. se um homeomorfismo  $f: U \subseteq M \rightarrow f(U) \subseteq \mathbb{H}^n$  é tal que a condição 3 é satisfeita pelo par  $(f, U)$  para todo  $i \in \mathcal{I}$ , então  $(f, U) \in \{(f_i, U_i)\}_{i \in \mathcal{I}}$  (dizemos que essa família é *maximal*).

Diversos objetos geométricos são exemplos de variedades diferenciáveis:  $\mathbb{R}^n$ , o espaço projetivo, esferas, toros e a garrafa de Klein são alguns exemplos bem conhecidos. A faixa de Möbius é um exemplo de variedade com bordo.

Um mapa  $\varphi: M \rightarrow N$  entre variedades é dito diferenciável se para todo  $p \in M$  e para algum (logo todo) par de cartas  $(f, U)$  contendo  $p$  e  $(g, V)$  contendo  $\varphi(p)$ , o mapa  $g \circ \varphi \circ f^{-1}$  é diferenciável. Denotaremos por  $C^\infty(M)$  o conjunto dos mapas diferenciáveis  $M \rightarrow \mathbb{R}$ .

Sejam  $M$  uma variedade diferenciável e  $p \in M$ . Para cada curva diferenciável  $\gamma: I \rightarrow M$ , onde  $I \subseteq \mathbb{R}$  é um intervalo contendo 0 e  $\gamma$  é tal que  $\gamma(0) = p$ , podemos definir um funcional linear  $\gamma'(0): C^\infty(M) \rightarrow \mathbb{R}$  chamado de *vetor tangente à  $\gamma$  em  $p$*  da seguinte maneira:

$$\gamma'(0)(\varphi) = \left. \frac{d}{dt}(\varphi \circ \gamma) \right|_{t=0}$$

O *espaço tangente em  $p$*   $T_p M$  é o espaço vetorial formado por todos os vetores tangentes às curvas em  $p$ , e possui dimensão igual à dimensão de  $M$ . Notamos que todo mapa  $f: M \rightarrow N$  entre variedades induz em cada  $p \in M$  um mapa linear  $df_p: T_p M \rightarrow T_{f(p)} N$ , que envia cada  $\gamma'(0) \in T_p M$  ao vetor tangente  $(f \circ \gamma)'(0) \in T_{f(p)} N$ .

Agora, definimos uma  $k$ -forma diferencial da seguinte maneira: sendo  $T_p M$  um espaço vetorial, existem as mesmas noções de espaço dual  $(T_p M)^*$  e subespaço das  $k$ -formas lineares alternadas  $\bigwedge^k (T_p M)^* \subseteq \text{Hom}_{\mathbb{R}}^k(T_p M, \mathbb{R})$ , e a mesma discussão sobre base e dimensão se aplica visto que  $T_p M \cong \mathbb{R}^n$ . Uma  $k$ -forma exterior em  $M$  é uma aplicação  $\omega: M \rightarrow \bigsqcup_p \bigwedge^k (T_p M)^*$  que leva  $p \in M$  a uma forma em  $\bigwedge^k (T_p M)^*$ , que será *diferencial* se para toda carta  $(f, U)$  de

$M$ ,  $f^*(\omega) \in \Omega^k(f(U))$  é uma forma diferencial no aberto  $f(U) \subseteq \mathbb{R}^n$ . As operações de cunha  $\wedge$  e derivada exterior  $d$  são também definidas localmente, pois se mostra que não dependem da escolha de cartas.

Em seguida, vimos como a integração de  $n$ -formas diferenciais pode ser estendida para uma variedade  $M$  orientável e compacta de dimensão  $n$ . Isso é possível porque para toda cobertura  $\{U_i\}_{i \in \mathcal{I}}$  de  $M$  por abertos, existem difeomorfismos  $\varphi_1, \dots, \varphi_m: M \rightarrow \mathbb{R}$  satisfazendo  $\sum_{i=1}^m \varphi_i = 1$ ,  $0 \leq \varphi_i \leq 1$  e  $\text{supp } \varphi_i \subseteq U_{j_i}$ ,  $j_i \in \mathcal{I}$ , onde  $\text{supp } \varphi_i$  é o suporte de  $\varphi_i$ : formando o que chamamos de *partição da unidade*. Assim, tomando a cobertura formada por um atlas  $\{(f_i, U_i)\}_i$  compatível com uma orientação da variedade, definimos a integral da  $n$ -forma  $\omega$  por:

$$\int_M \omega = \sum_{i=1}^m \int_{f(U_{j_i})} f_{j_i}^*(\varphi_i \cdot \omega),$$

e de fato, é possível mostrar que a expressão na direita não depende da partição escolhida, graças à hipótese de orientabilidade.

Recordamos que como nossas variedades são localmente homeomorfas a  $\mathbb{H}^n$ , definimos o bordo de uma variedade  $M$  como  $\partial M = \bigcup_{i \in \mathcal{I}} f^{-1}(\partial \mathbb{H}^n)$ , onde  $\{f_i, U_i\}_{i \in \mathcal{I}}$  é um atlas de  $M$  e  $\partial \mathbb{H}^n = \{x \in \mathbb{H} : x_1 = 0\}$ . É possível mostrar que o bordo de uma  $n$ -variedade é uma variedade de dimensão  $n - 1$ , que também é orientável se  $M$  for orientável. Por fim, vimos a demonstração do

**Teorema 1** (Teorema de Stokes). *Seja  $M$  uma variedade diferencial de dimensão  $n$ , compacta, orientada e com bordo não vazio. Seja  $\omega$  uma  $(n - 1)$ -forma diferencial em  $M$  e  $i: \partial M \rightarrow M$  a inclusão. Então*

$$\int_{\partial M} i^* \omega = \int_M d\omega.$$

## Dinâmicas expansoras

Durante esta etapa do projeto, estudamos dinâmicas expansoras em variedades riemannianas compactas. Essas dinâmicas são caracterizadas por afastar pontos próximos e um exemplo específico foi estudado no começo do projeto: a transformação  $f(x) = nx \pmod{1}$  no intervalo  $[0, 1]$ , onde  $n$  é um inteiro maior ou igual a dois. Baseamos o estudo no primeiro capítulo do livro *Stochastic Dynamics of Deterministic Systems*[5].

Sendo  $M$  uma variedade riemanniana, um difeomorfismo  $f: M \rightarrow M$  é dito *expansor* se existe uma constante  $\lambda > 1$  tal que  $\|df_x(v)\| \geq \lambda \|v\|$  para todos  $x \in M$ ,  $v \in T_x M$ . Como nesse caso  $df_x(v) = 0 \iff \|v\| = 0$ , mapas expansores são difeomorfismos locais. Uma propriedade importante de mapas expansores é que quando  $M$  é uma variedade compacta, tais mapas possuem um número fixo de pré-imagens:

**Proposição 2.** *Seja  $M$  uma variedade compacta,  $N$  uma variedade conexa, e  $f: M \rightarrow N$  um difeomorfismo local. Então existe um inteiro  $k \geq 1$  tal que  $\#f^{-1}(y) = k$  para todo  $x \in N$  (tal inteiro é chamado de grau do mapa  $f$ ).*

Para demonstrar o teorema, precisaremos de alguns lemas:

**Lema 2.1.** *Sejam  $X, Y$  espaços métricos,  $f : X \rightarrow Y$  função contínua e  $K \subset X$  um compacto. Então  $f(K)$  é um compacto.*

**Lema 2.2.** *Agora, supomos  $X, Y$  e  $f$  como no lema anterior e que o espaço  $X$  seja compacto. Então  $f$  é um mapa fechado.*

**Lema 2.3.** *Seja  $M, N$ , e  $f : M \rightarrow N$  como na Proposição 2. Então para todo  $y \in N$ ,  $\#f^{-1}(y)$  é finito.*

*Demonstração.* De fato, caso não fosse finito,  $f^{-1}(y)$  teria um ponto limite no espaço compacto  $M$ . Toda vizinhança desse ponto contém infinitas pré-imagens de  $y$ , portanto a restrição de  $f$  a nenhuma delas poderia ser bijetiva. Isso contradiz  $f$  ser difeomorfismo local.  $\square$

*Prova da Proposição 2.* Pelo Lema 2.2,  $f$  é um mapa fechado. Seja  $y \in N$ . Usando o último lema, podemos escrever  $\{x_i\}_{i=1}^K = f^{-1}(y)$ . Então existem vizinhanças abertas  $U_i \ni x_i$  tais que  $f|_{U_i}$  é difeomorfismo para cada  $i$ , e podemos diminuir essas vizinhanças de forma que sejam disjuntas. Seja  $F = M \setminus \bigcup_{i=1}^K U_i$ . Então  $f(F)$  é fechado porque  $F$  e  $f$  são fechados, logo  $V = \bigcap_{i=1}^K f(U_i) \setminus f(F)$  é um aberto contendo  $y$  visto que nenhuma das  $K$  pré-imagens de  $y$  está em  $F$ .

Agora, por definição  $f^{-1}(V) \subset F^c = \bigcup_{i=1}^K U_i$ , e já sabemos que  $f(U_i \cap f^{-1}(V)) = V$  para cada  $i$ . Mas  $U_i \cap f^{-1}(V)$  são vizinhanças disjuntas, portanto todo ponto do aberto  $V$  tem exatamente  $K$  pré-imagens, uma para cada  $i$ .

O que acabamos de mostrar foi que as classes da relação de equivalência  $a \sim b \iff \#f^{-1}(a) = \#f^{-1}(b)$  são abertas em  $N$ . Como  $N$  é conexa, então só há uma classe de equivalência. Sendo  $M$  não vazia, existe um ponto com pelo menos uma pré-imagem, então concluímos que existe  $k$  como no enunciado.  $\square$

## Cones e métricas projetivas

Seja  $V$  um espaço vetorial. Um *cone convexo* é definido como um subconjunto  $C \subset E \setminus \{0\}$  satisfazendo as seguintes propriedades: para todos  $v \in C$  e  $t > 0$ ,  $tv \in C$ , e  $C$  é convexo. Em particular, essas condições garantem que  $C$  é fechado por combinações lineares com coeficientes positivos. O *fecho* de um cone  $C$  é o conjunto  $\bar{C}$  formado pelos pontos  $w \in V$  tais que existe  $v \in C$  e uma sequência  $(t_n)_n$  de reais positivos convergindo para zero satisfazendo  $(w + t_n v) \in C$  para todo  $n \geq 1$ .

Fixemos um cone  $C$  tal que  $\bar{C} \cap (-\bar{C}) = \{0\}$ . Nesse caso, conseguimos definir uma métrica  $\theta$  no quociente projetivo de  $C$ , chamada de *métrica projetiva* associada ao cone  $C$ , da seguinte forma:

$$\theta(v, w) = \log \frac{\sup\{t > 0 \mid w - tv \in C\}}{\inf\{s > 0 \mid sv - w \in C\}}. \quad (2.1)$$

Vimos que a função  $\theta$  possui as seguintes propriedades, que justificam chamá-la

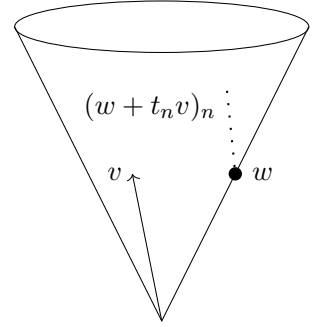
de métrica projetiva: para todos  $u, v, w \in C$ ,

1.  $\theta(u, v) = \theta(v, u)$
2.  $\theta(u, w) \leq \theta(u, v) + \theta(v, w)$
3.  $\theta(u, v) = 0 \iff \exists t > 0, u = tv$ .

Essas métricas projetivas se aplicam em alguns cones relevantes aos mapas expansores: o cone  $C_+$  de funções contínuas positivas e o cone  $C(a, \nu)$  das funções  $(a, \nu)$ -Hölder contínuas positivas, ou seja, das funções  $\varphi > 0$  tais que

$$\exp(-a \cdot d(x, y)^\nu) \leq \frac{\varphi(x)}{\varphi(y)} \leq \exp(a \cdot d(x, y)^\nu) \text{ para todos } x, y \in M,$$

onde em ambos os cones as funções tomam valores reais em um espaço métrico  $M$ .



### Operadores de transferência e cones invariantes

Como vimos na Proposição 2, se temos uma transformação expansora  $f$  com constante  $\sigma$  em uma variedade riemanniana compacta  $M$ , então existe um número  $k$  tal que todo ponto tem  $k$  pré-imagens. Assim, os operadores *de transferência*  $\mathcal{L}$  e *de Koopman*  $U_f$  em  $C^1(M, \mathbb{R})$  são:

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

onde  $\{y_1, \dots, y_k\} = f^{-1}(y)$ . Há uma relação especial entre os dois operadores, expressa na seguinte igualdade:

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

a primeira proposição importante diz que os cones  $C(a, \nu)$  são invariantes por  $\mathcal{L}$ :

**Proposição 3** (invariância). *Existe  $\lambda < 1$  tal que  $\mathcal{L}(C(a, \nu)) \subseteq C(\lambda a, \nu)$  para todo  $a > 0$  suficientemente grande.*

Fixados quaisquer  $a, \lambda$  e  $\nu$  dados pela proposição acima, a métrica projetiva  $\theta$  no cone  $C(a, \nu)$  também tem a propriedade importante de ter diâmetro finito, ou seja,  $\sup\{\theta(\varphi_1, \varphi_2) : \varphi_1, \varphi_2 \in C(\lambda a, \nu)\} < \infty$ . Juntos, esses resultados implicam em  $\mathcal{L}$  ser uma contração no cone  $C(a, \nu)$ . Aplicamos esse fato notando que a métrica projetiva  $\theta_+$  em  $C_+$  tem a propriedade de toda sequência  $(\varphi_n)_n$  Cauchy ser convergente, e se  $\int \varphi_i \, dm = 1$  para todo  $i$ , a convergência ser uniforme. Isso é importante porque a sequência  $(\mathcal{L}^n 1)_n$  é tal que

$$\int (\mathcal{L}^n 1) \cdot 1 \, dm = \int 1 \cdot 1 \, dm = 1,$$

e também é  $\theta_+$ -Cauchy, pois é uma contração para a métrica  $\theta$  e de  $C_+ \subseteq C(a, \mu)$  segue que  $\theta_+ \leq \theta$ . Como o limite em  $C_+$  é uniforme, obtemos um ponto fixo  $\varphi_0 \in C(a, \nu)$  pois a propriedade Hölder é preservada por limites uniformes. Esse ponto fixo define uma medida invariante  $\mu_0 = \varphi_0 m$ , e vimos que essa medida tem a propriedade de ser exata, e portanto ergódica.



## Estabilidade estocástica

Também é possível mostrar que mapas expansores  $f: M \rightarrow M$  são *estocasticamente estáveis*, como será definido abaixo. Sendo  $C^{1+\nu_0}$  o espaço das funções  $M \rightarrow \mathbb{R}$  diferenciáveis com derivada  $\nu_0$ -Hölder contínuas, consideramos uma vizinhança  $T$  de  $f$  nesse espaço. Também precisaremos de uma família  $\{\theta_\varepsilon\}_{\varepsilon>0}$  de medidas de probabilidade regulares de  $T$  tal que  $\text{supp } \theta_\varepsilon \rightarrow \{f\}$  quando  $\varepsilon \rightarrow 0$ , no sentido que o diâmetro do suporte em torno de  $f$  converge para zero. Só para fins de notação, definiremos  $f_t = t$  para  $t \in T$ . Com isso, podemos definir versões perturbadas dos operadores  $U$  e  $\mathcal{L}$  que foram vistos acima:

$$(U_t)(x) = \varphi(f_t(x)) \quad (\mathcal{L}_t\varphi)(y) = \sum_{x \in f_t^{-1}(y)} \varphi(x) |Df_t(x)|^{-1},$$

$$(\widehat{U}_\varepsilon\varphi)(x) = \int (U_t\varphi)(x) d\theta_\varepsilon(t) \quad (\widehat{\mathcal{L}}_\varepsilon\varphi)(x) = \int (\mathcal{L}_t\varphi)(x) d\theta_\varepsilon(t).$$

É possível mostrar que  $\widehat{U}_\varepsilon$  e  $\widehat{\mathcal{L}}_\varepsilon$  compartilham muitas das propriedades dos operadores originais, como dualidade e a invariância dos cones de funções. Isso nos permite, como no caso original, encontrar um ponto fixo  $\varphi_\varepsilon$  de  $\widehat{\mathcal{L}}_\varepsilon$  para cada  $\varepsilon$ , associado a uma medida  $\mu_\varepsilon = \varphi_\varepsilon m$ . Essa medida também governa o comportamento estatístico de trajetórias aleatórias, no sentido que tomando  $t_1, t_2, \dots \in T$  distribuídos de acordo com a probabilidade  $\theta_\varepsilon$ , vale

$$\frac{1}{n} \sum_{j=1}^n \varphi(f_{t_j} \circ \dots \circ f_{t_1}(x)) \xrightarrow{n \rightarrow \infty} \int \varphi d\mu_\varepsilon,$$

para  $\varphi$  Hölder contínua e  $m$ -quase todo  $x \in M$ .

Por fim, vimos que para  $\varphi_0$  sendo o ponto fixo de  $\mathcal{L}$  e  $\mu_0 = \varphi_0 m$ ,  $\varphi_\varepsilon \rightarrow \mu_0$  uniformemente e  $\mu_\varepsilon \rightarrow \mu_0$  na topologia fraca\* quando  $\varepsilon \rightarrow 0$ . Isso significa que a dinâmica é *estocasticamente estável*, no sentido que nos aproximamos do comportamento determinístico quando consideramos perturbações menores para a versão estocástica.

## Entropia em sistemas dinâmicos

A palavra entropia teve sua origem durante o desenvolvimento da Termodinâmica no século XIX, expressando uma medida de “desordem” de um sistema em equilíbrio. Em meados do século XX, o conceito inspirou novas definições de entropia em Teoria da Informação, principalmente pelos trabalhos de Claude Shannon, e também na Teoria Ergódica, pelos matemáticos Andrey Kolmogorov e Yakov Sinai. Nesta última, a entropia foi introduzida como uma invariante de equivalência ergódica que pudesse distinguir sistemas não equivalentes (por exemplo, os diferentes deslocamentos de Bernoulli) que não eram distinguíveis por outras invariantes como o espectro do operador de Koopman.

### Teoria da Informação

Na Teoria da Informação, a entropia representa uma medida da quantidade informação transmitida através de um canal de comunicação. Consideremos um alfabeto  $\mathcal{A}$  com um número

finito de símbolos, e um canal de comunicação pelo qual duas instâncias se comunicam enviando caracteres desse alfabeto. Cada símbolo transmitido carrega um “fator surpresa” para o ouvinte, mas esse fator também depende da sequência de símbolos transmitida anteriormente: por exemplo, ao se transmitir a sequência “M E L A N C I” em um contexto em que os interlocutores falam português, transmitir ou não uma última letra “A” praticamente não adicionaria informação à mensagem — não há outras possibilidades no idioma que façam sentido. Por isso, a nossa definição de informação irá envolver não apenas caracteres, mas listas de caracteres.

Tendo em vista a interpretação da informação como um “fator surpresa”, não é de se surpreender que a definição formal do conceito envolva probabilidades: quanto mais provável consideramos a ocorrência uma sequência de caracteres, menos informação ela nos traz, e vice-versa. Mas como podemos “compor” as entropias para estimar a quantidade de informação de múltiplos eventos distintos e independentes? Por exemplo, em um jogo perfeito de cara ou coroa, sabemos uma sequência de  $n$  caras consecutivas acontece com probabilidade  $1/2^n$ . Mas seria razoável dizer que cada um dos  $n$  resultados obtidos *adiciona* a mesma quantidade de informação para quem assiste — de certa forma, queremos que a multiplicação de probabilidades independentes vire a *soma* de informações. Por isso, nada mais adequado que a função logaritmo para definir informação. De fato, dada uma palavra  $a_1 a_2 \dots a_n$ , que ocorre no meio de comunicação com probabilidade  $p_{a_1 \dots a_n}$ , onde cada  $a_i$  é um elemento do alfabeto, definimos sua *quantidade de informação* como  $I(a_1 \dots a_n) = -\log p_{a_1 \dots a_n}$ . Agora, sendo  $\mathcal{A}^n$  o conjunto de todas as palavras de tamanho  $n$ , definimos a informação desse conjunto como o valor esperado

$$I(\mathcal{A}^n) = \sum_{a_1 \dots a_n \in \mathcal{A}^n} p_{a_1 \dots a_n} I(a_1 \dots a_n). \quad (2.2)$$

Finalmente, a *entropia* do canal de comunicação é o limite do valor esperado acima quando consideramos palavras cada vez maiores e dividimos o valor pelo tamanho das mesmas, obtendo assim o quanto de informação por símbolo é transmitido, em média, na interação pelo canal:

$$I = \lim_n \frac{1}{n} I(\mathcal{A}^n). \quad (2.3)$$

A construção do conceito de entropia para sistemas dinâmicos é similar a de teoria da informação, mas já percebemos de início que ambas tratam sobre domínios distintos. Fixemos um espaço de probabilidade  $(M, \mathcal{B}, \mu)$  e uma função mensurável  $f : M \rightarrow M$  até o fim da seção. Como o espaço  $M$  não precisa ser finito, o papel de um alfabeto para a dinâmica será desempenhado por uma partição enumerável do espaço  $M$ , definida como uma família enumerável de subconjuntos de  $M$  disjuntos dois a dois e cuja união é o espaço inteiro. Notamos que cada partição  $\mathcal{P}$  define uma função  $\mathcal{P} : X \rightarrow I$ , onde  $\mathcal{P}(x)$  é o elemento de  $\mathcal{P}$  contendo  $x$  e  $I$  é a coleção dos subconjuntos de  $X$ .

Podemos definir a *soma de duas partições*  $\mathcal{P}$  e  $\mathcal{Q}$  como  $\mathcal{P} \vee \mathcal{Q} = \{P \cap Q : P \in \mathcal{P}, Q \in \mathcal{Q}\}$ , e também para famílias enumeráveis de partições  $(\mathcal{P}_i)_{i \in \mathcal{I}}$  como

$$\bigvee_{i \in \mathcal{I}} \mathcal{P}_i = \left\{ \bigcap_{i \in \mathcal{I}} P_i : P_i \in \mathcal{P}_i \text{ para cada } i \right\}.$$

Cada partição  $\mathcal{P}$  está associada a uma função informação  $I_{\mathcal{P}} : X \rightarrow \mathbb{R}$  definida por  $I_{\mathcal{P}}(x) = -\log \mu(\mathcal{P}(x))$ . A entropia  $H_{\mu}(\mathcal{P})$  da partição  $\mathcal{P}$  é simplesmente o valor esperado dessa função:

$$H_{\mu}(\mathcal{P}) = \int I_{\mathcal{P}} d\mu = \sum_{P \in \mathcal{P}} -\mu(P) \log \mu(P).$$

Também podemos definir a entropia condicional de  $\mathcal{P}$  relativamente a outra partição  $\mathcal{Q}$  da seguinte forma:

$$H_{\mu}(\mathcal{P}/\mathcal{Q}) = \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} -\mu(P \cap Q) \log \frac{\mu(P \cap Q)}{\mu(Q)}.$$

Intuitivamente, a entropia de uma partição é a quantidade de informação associada a saber em qual elemento da partição o sistema está. Nesse sentido, a entropia condicional de  $\mathcal{P}$  relativa a  $\mathcal{Q}$  mede a informação adicional dada por  $\mathcal{P}$  em um contexto em que já se conhece a informação dada por  $\mathcal{Q}$ .

Podemos definir uma ordem parcial nas partições de  $M$  dizendo que  $\mathcal{P}$  é *menos fina* que  $\mathcal{Q}$  se todo elemento de  $\mathcal{Q}$  de medida não-nula está contido em um elemento de  $\mathcal{P}$ . Nesse caso, escrevemos  $\mathcal{P} \prec \mathcal{Q}$ .

Munidos das definições acima, prosseguimos vendo propriedades importantes da função informação. Enumeramos aqui algumas delas:

1.  $H_{\mu}(\mathcal{P} \vee \mathcal{Q}/\mathcal{R}) = H_{\mu}(\mathcal{P}/\mathcal{R}) + H_{\mu}(\mathcal{Q}/(\mathcal{P} \vee \mathcal{R}))$
2. se  $\mathcal{P} \prec \mathcal{Q}$  então  $H_{\mu}(\mathcal{P}/\mathcal{R}) \leq H_{\mu}(\mathcal{Q}/\mathcal{R})$  e  $H_{\mu}(\mathcal{R}/\mathcal{P}) \geq H_{\mu}(\mathcal{R}/\mathcal{Q})$
3.  $\mathcal{P} \prec \mathcal{Q}$  se, e somente se,  $H_{\mu}(\mathcal{P}/\mathcal{Q}) = 0$ .

Podemos interpretar um pouco as propriedades: a primeira diz que “a informação conjunta de  $\mathcal{P}$  e  $\mathcal{Q}$  se já sabemos  $\mathcal{R}$  é a informação de  $\mathcal{P}$  se já sabemos  $\mathcal{R}$  mais a informação de  $\mathcal{Q}$  se já sabemos  $\mathcal{P}$  e  $\mathcal{R}$ ”. Também podemos interpretar  $\mathcal{P} \prec \mathcal{Q}$  como “a informação de  $\mathcal{P}$  está contida na informação de  $\mathcal{Q}$ ” — de fato, o item 3 diz que isso é equivalente à  $\mathcal{P}$  ter informação nula dado que já sabemos a informação de  $\mathcal{Q}$ .

Seja  $f: M \rightarrow M$  uma transformação mensurável para uma medida  $\mu$ . Em analogia às palavras de tamanho  $n$ , definimos para cada  $n \geq 1$

$$\mathcal{P}^n = \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{P}).$$

Vimos que a sequência  $(H_{\mu}(\mathcal{P}^n))_{n=1}^{\infty}$  é subaditiva, ou seja,  $H_{\mu}(\mathcal{P}^{n+m}) \leq H_{\mu}(\mathcal{P}^n) + H_{\mu}(\mathcal{P}^m)$ . Isso implica no limite

$$h_{\mu}(f, \mathcal{P}) = \lim_n \frac{1}{n} H_{\mu}(\mathcal{P}^n) = \inf_n \frac{1}{n} H_{\mu}(\mathcal{P}^n)$$

existir, sendo chamado de *entropia de  $f$  com respeito à medida  $\mu$  e à partição  $\mathcal{P}$* . Finalmente, a *entropia do sistema dinâmico  $(f, \mu)$*  é o supremo  $h_{\mu}(f) = \sup_{\mathcal{P}} h_{\mu}(f, \mathcal{P})$ , onde  $\mathcal{P}$  percorre todas as partições com entropia finita.

## Entropia topológica

Também estudamos outra definição de entropia, dessa vez baseada apenas na topologia do espaço  $M$ . Se  $\alpha$  é uma cobertura aberta de um espaço compacto  $M$ , por definição  $\alpha$  possui subcobertura finita; assim, a menor cardinalidade de uma subcobertura, denotada por  $N(\alpha)$ , é um número natural. Nessa seção, todas as coberturas serão abertas; definimos a entropia de uma cobertura simplesmente como  $H(\alpha) = \log N(\alpha)$ . Também podemos somar coberturas, definindo a soma  $\alpha_1 \vee \dots \vee \alpha_n$  como a cobertura formada pelas interseções  $A_1 \cap \dots \cap A_n$ , em que  $A_i \in \alpha_i$ .

Fixemos uma função contínua  $f: M \rightarrow M$ ; então a pré-imagem de uma cobertura  $f^{-1}(\alpha)$  também é cobertura, o que nos motiva a definir  $\alpha^n = \alpha \vee f^{-1}(\alpha) \vee \dots \vee f^{-n+1}(\alpha)$ , de forma bastante parecida à versão para partições. Também é verdade nesse caso que  $(H(\alpha^n))_{n=1}^\infty$  é uma sequência subaditiva, e assim existe

$$h(f, \alpha) = \lim_n \frac{1}{n} H(\alpha^n),$$

chamada de *entropia de  $f$  com respeito à cobertura  $\alpha$* . Por fim, definimos a *entropia topológica de  $f$*  como o número real estendido  $h(f) = \sup\{h(f, \alpha) : \alpha \text{ é cobertura de } M\}$ .

Outra definição de entropia topológica, que dessa vez requer que  $M$  seja um espaço métrico não necessariamente compacto, é feita através de *bolas dinâmicas* e *conjuntos geradores e separados*. Definimos a bola dinâmica de centro  $a \in M$ , raio  $r$  e comprimento  $n$  como  $B(a, n, r) = \{x : d(f^i(x), f^i(a)) < r, \forall i \in \{0, \dots, n-1\}\}$ . Dizemos que um subconjunto  $E \subseteq M$  é  $(n, \varepsilon)$ -gerador um compacto  $K \subseteq M$  se

$$K \subseteq \bigcup_{x \in E} B(x, n, \varepsilon).$$

Notamos que as bolas dinâmicas são abertas, logo por compacidade, para cada  $K$  sempre existem  $(n, \varepsilon)$ -geradores finitos. Assim, a menor cardinalidade de um conjunto  $(n, \varepsilon)$ -gerador  $g_n(f, \varepsilon, K)$  é um número natural. Definindo

$$g(f, \varepsilon, K) = \limsup_n \frac{1}{n} \log g_n(f, \varepsilon, K),$$

vemos que  $\varepsilon \mapsto g(f, \varepsilon, K)$  é uma função positiva não crescente. Isso garante que existe o limite  $g(f, K) = \lim_{\varepsilon \rightarrow 0} g(f, \varepsilon, K)$ . Como último passo, para que  $g$  dependa apenas de  $f$ , definimos  $g(f) = \sup\{g(f, K) : K \subseteq M \text{ é compacto}\}$ .

A função  $g$  também é chamada de entropia topológica, e de fato podemos mostrar que ela é igual à função  $h$  quando  $M$  é um espaço métrico compacto. Já a relação entre a entropia topológica e a entropia de Kolmogorov-Sinai é mais sutil: essa relação fica expressa no princípio variacional, que não foi tema de estudo do projeto.

## Referências Bibliográficas

- [1] Oliveira, K. e Viana, M. *Fundamentos da Teoria Ergódica*, 2<sup>a</sup> ed., SBM, Rio de Janeiro, 2019.
- [2] Lee, J. M., *Introduction to Smooth Manifolds*, 2<sup>a</sup> ed., Graduate Texts in Mathematics, Springer-Verlag, Berlim, 2013.
- [3] Lee, J. M., *Introduction to Topological Manifolds*, 2<sup>a</sup> ed., Graduate Texts in Mathematics, Springer-Verlag, Berlim, 2011.
- [4] do Carmo, M. P., *Formas Diferenciais e Aplicações*, 1<sup>a</sup> ed., SBM, Rio de Janeiro, 2015.
- [5] Viana, M., *Stochastic Dynamics of Deterministic Systems*, IMPA, 1997.

## Participação em evento científico

No período compreendido entre os dias 30 de novembro e 3 de dezembro de 2020, o aluno participou como expositor no 28º Congresso de Iniciação Científica da Unicamp, que aconteceu de forma virtual. O trabalho apresentado consistiu de um pequeno artigo, enviado em anexo ([congresso.pdf](#)), e um vídeo-pôster, disponível em <https://youtu.be/04uEX8S1W0w>, que foram disponibilizados aos visitantes. Durante o evento, os visitantes puderam escrever perguntas sobre o projeto para serem respondidas pelo aluno. A página do congresso pode ser acessada pelo link:

<https://www.prp.unicamp.br/pibic/congressos/xxviiiicongresso/index.php>

## Lista dos trabalhos preparados ou submetidos

O trabalho “Propriedades Ergódicas de Sistemas Dinâmicos” (arquivo `congresso.pdf`), que foi apresentado no congresso acima, será posteriormente publicado na Revista dos Trabalhos de Iniciação Científica da Unicamp.