

# INTRODUÇÃO AOS SISTEMAS DINÂMICOS

---

YURI LIMA

*Universidade Federal do Ceará*

[yurilima@gmail.com](mailto:yurilima@gmail.com)

---

## Sumário

<b>I</b>	<b>Equações diferenciais ordinárias</b>	<b>1</b>
1	Definição e exemplos	1
2	EDO de ordem superior	7
3	Teorema de existência e unicidade de EDO's	9
3.1	Preparação para a prova do Teorema 14	12
3.2	Prova do Teorema 14	14
3.3	Exemplos	16
3.4	Métricas adaptadas	17
3.5	Melhorando o domínio da solução	19
4	Teorema de Peano	21
5	Soluções maximais	23
6	Continuidade de $\varphi$	24
7	Desigualdade de Gronwall e estimativa Lipschitz de $\varphi(t, t_0, \cdot)$	26
8	Diferenciabilidade de $\varphi$	28
9	Teorema de comparação de Sturm	30
9.1	Estimativas para a equação de Jacobi	31
9.2	Equação de Riccati	32
10	Exercícios	33
<b>II</b>	<b>Mecânica celeste</b>	<b>37</b>
11	Problema de Kepler	37
12	Problema dos dois corpos	43
13	Problema dos três corpos	46

14 Exercícios	48
III Sistemas dinâmicos: definições básicas e homeomorfismos do intervalo	50
15 Transformações e fluxos	50
16 Homeomorfismos do intervalo	56
17 Exercícios	60
IV Endomorfismos do intervalo, teoremas de Li-Yorke e Sharkovsky	62
18 Exemplo introdutório: família quadrática	62
19 Grafos de Markov e o teorema de Li-Yorke	64
20 Teorema de Sharkovsky	69
20.1 Fatos gerais	69
20.2 Órbitas de período $2^k$	71
20.3 Órbitas de período ímpar	73
20.4 Conclusão da prova do teorema de Sharkovsky	78
21 Exercícios	79
V Transformações no círculo	81
22 Projeção de transformações	81
23 Levantamento de transformações	82
24 Rotações irracionais	83
25 Número de rotação	86
26 Exercícios	90
VI Shifts topológicos de Markov	91

27 Espaços simbólicos	91
28 Palavras, cilindros e topologia produto	93
29 Shifts topológicos de Markov	94
30 Exercícios	96
VII Exemplos	98
31 Mapas unidimensionais markovianos	98
32 Conjuntos de Cantor dinamicamente definidos	105
33 Ferradura de Smale	109
34 Automorfismos lineares hiperbólicos do toro	113
35 Exercícios	120
VIII Apêndice	121
36 Convergência uniforme	121
37 O círculo $\mathbb{S}^1$	123
38 Teorema de Kronecker	125
39 Dimensão de Hausdorff em $\mathbb{R}$	126

---

# Equações diferenciais ordinárias

## SEÇÃO 1

### Definição e exemplos

**Definição 1** (Equação diferencial ordinária – EDO) Uma *equação diferencial ordinária* (EDO) é uma relação

$$x^{(k)} = F(t, x, x', x'', \dots, x^{(k-1)}) \quad (1.1)$$

onde  $F : U \rightarrow \mathbb{R}^d$  é uma função contínua e  $U$  é um aberto de  $\mathbb{R}^{1+kd}$ .

Acima,  $t \in \mathbb{R}$  e cada  $x, x', x'', \dots, x^{(k)} \in \mathbb{R}^d$ . A *ordem* ou *grau* da EDO é  $k$ , e a dimensão da EDO é  $d$ . A ideia da equação (1.1) é que as derivadas de ordem inferior e a evolução no tempo descrevem o comportamento da  $k$ -ésima derivada. Assim, uma EDO de ordem  $k$  é uma equação *implícita* em que  $k$  é a quantidade de derivadas envolvidas. Nosso objetivo é, dada a função  $F$ , encontrar/caracterizar/descrever as soluções  $x$  da EDO. Antes, definimos formalmente o significado de solução.

**Definição 2** (Solução de EDO) Uma *solução* de (1.1) é uma função  $\gamma : I \rightarrow \mathbb{R}^d$  que é  $k$  vezes diferenciável e satisfaz as seguintes condições:

- (1)  $I$  é um intervalo aberto de  $\mathbb{R}$ .
- (2)  $(t, \gamma(t), \gamma'(t), \dots, \gamma^{(k-1)}(t)) \in U$  para todo  $t \in I$ .
- (3)  $\gamma^{(k)}(t) = F(t, \gamma(t), \gamma'(t), \dots, \gamma^{(k-1)}(t)) \in U$  para todo  $t \in I$ .

Noutras palavras, apenas substituímos  $x$  por  $\gamma(t)$  em (1.1). Essa noção também nos permite definir uma EDO com condições iniciais prescritas.

**Definição 3** (EDO com condições iniciais) Uma *EDO com condições iniciais* é um sistema da forma

$$\begin{cases} x^{(k)} = F(t, x, x', x'', \dots, x^{(k-1)}) \\ x(t_0) = x_0, x'(t_0) = x'_0, \dots, x^{(k-1)}(t_0) = x_0^{(k-1)}, \end{cases} \quad (1.2)$$

onde  $F : U \rightarrow \mathbb{R}^d$  é uma função contínua em um aberto  $U$  de  $\mathbb{R}^{1+kd}$  e  $x_0, x'_0, \dots, x_0^{(k-1)} \in \mathbb{R}^d$  são dados.

**Definição 4** (Solução de EDO com condições iniciais) Uma *solução de (1.1) com condições iniciais*

$$\begin{cases} x^{(k)} = F(t, x, x', x'', \dots, x^{(k-1)}) \\ x(t_0) = x_0, x'(t_0) = x'_0, \dots, x^{(k-1)}(t_0) = x_0^{(k-1)}, \end{cases} \quad (1.3)$$

é uma função  $\gamma : I \rightarrow \mathbb{R}^d$  que satisfaz a Definição 2 e adicionalmente:

$$(4) \quad \gamma(t_0) = x_0, \gamma'(t_0) = x'_0, \dots, \gamma^{(k-1)}(t_0) = x_0^{(k-1)}.$$

Vamos analisar alguns exemplos.

**Exemplo 5** Considere  $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  dada por  $F(t, x) = 0$ . A EDO é

$$x' = 0,$$

com  $k = d = 1$ . Fixado  $x_0 \in \mathbb{R}$ , seja  $\gamma : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $\gamma(t) = x_0$  para todo  $t \in \mathbb{R}$ . Então  $\gamma$  é solução da EDO, pois

$$\frac{d\gamma}{dt}(t) = \gamma'(t) = 0 = F(t, \gamma(t)).$$

Note que a EDO tem infinitas soluções, uma para cada  $x_0 \in \mathbb{R}$ . Podemos também considerar a versão multidimensional desse exemplo, tomando  $F : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $F(t, x) = 0$ . Nesse caso,  $k = 1$  e a dimensão é  $d$ . As soluções são obtidas da mesma maneira.

**Exemplo 6** Considere  $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $F(t, x) = x$ . A EDO é

$$x' = x,$$

novamente com  $k = d = 1$ . As soluções dessa EDO são  $\gamma(t) = ce^t$ , onde  $c$  é uma constante. De fato:

$$\gamma'(t) = \gamma(t) = F(t, \gamma(t)).$$

**Exemplo 7** Seja  $F : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  dada por  $F(t, x, x') = -x$ . A EDO é

$$x'' = -x \iff x'' + x = 0,$$

com  $k = 2$  e  $d = 1$ . Dadas constantes  $a, b \in \mathbb{R}$ , seja  $\gamma(t) = a \cos t + b \sin t$ . Temos

$$\gamma'(t) = -a \sin t + b \cos t \Rightarrow \gamma''(t) = -(a \cos t + b \sin t) = -\gamma(t) = F(t, \gamma(t), \gamma'(t)).$$

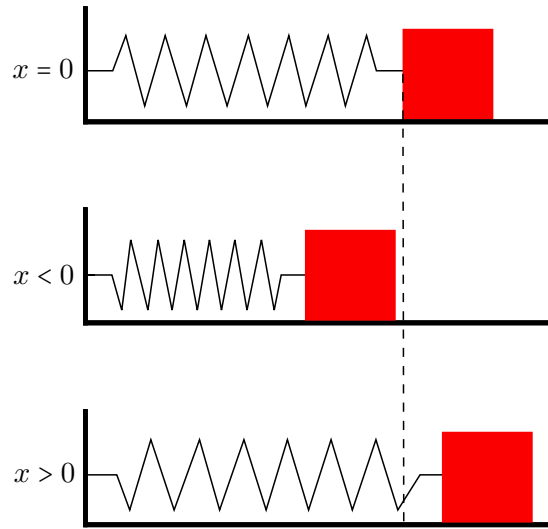
Logo,  $\gamma(t)$  é solução da EDO.

Os próximos exemplos são oriundos da mecânica clássica (leis de Newton).

**Exemplo 8** (Lei de Hooke) Considere uma mola de comprimento  $\ell$  em equilíbrio (sem estar expandida ou comprimida). Fixe uma de suas pontas a uma parede e a outra ponta a um bloco  $B$ . A *lei de Hooke* descreve a força necessária a ser aplicada em  $B$  para que a mola seja expandida/comprimida a partir de sua posição de repouso:

$$F = kx,$$

onde  $k$  é uma constante que depende apenas da mola e  $x$  é a distância da expansão/compressão medida a partir da posição de equilíbrio. A



Pela terceira lei de Newton, podemos enunciar a lei de Hooke afirmando que a *força de restauração da mola* é

$$F = -kx.$$

O sinal negativo é o que indica que a força é de restauração:

- Quando a mola está expandida, a força aponta para a esquerda, ou seja, a mola quer se desexpandir.
- Quando a mola está comprimida, a força aponta para a direita, ou seja, a mola quer se descomprimir.

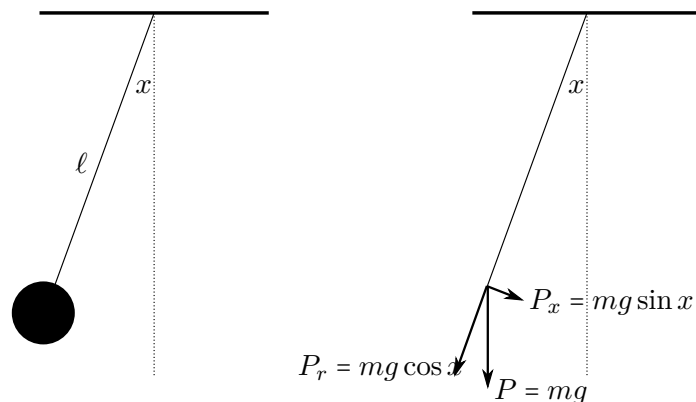
Em outras palavras, a mola sempre busca retornar à sua posição de repouso. Queremos entender o que acontece quando retiramos o bloco: a partir do conhecimento da força que a mola sofre inicialmente e das leis de Newton, queremos descrever o movimento  $x$  que a extremidade da direita da mola sofre. A força é  $F = -kx$ . Pela segunda lei de Newton,  $F = ma = mx''$ . Igualando as duas expressões, temos

$$mx'' = -kx \implies x'' = -\frac{k}{m}x.$$

Essa EDO é uma variação do exemplo anterior. São soluções dessa EDO as funções da forma

$$\gamma(t) = a \cos \left[ \sqrt{\frac{k}{m}} t \right] + b \sin \left[ \sqrt{\frac{k}{m}} t \right].$$

**Exemplo 9** (Pêndulo harmônico) Considere um haste rígida presa ao teto em uma extremidade e a um corpo de massa  $m$  na outra extremidade. A haste pode girar na extremidade da parede. O movimento que o corpo realiza é chamado de *pêndulo harmônico*. Afirmamos que ele é descrito por uma EDO.



De fato, há duas forças atuando no corpo: seu peso e a força de tensão da haste. Após decompor o peso em duas forças, uma paralela e outra perpendicular à haste, obtemos da segunda lei de Newton que

$$mx'' = mg \sin x \implies x'' = g \sin x,$$

que é uma EDO com  $k = 2$  e  $d = 1$ . Infelizmente (ou felizmente), é impossível fornecer uma fórmula para as soluções dessa EDO utilizando as funções que conhecemos.

Observe que no segundo, terceiro e quarto exemplos, as soluções da EDO utilizam funções bem mais complicadas do que a função  $F$  que define a EDO, e que no último exemplo, com a função  $F(t, x, x') = g \sin x$  relativamente simples, já não é possível descrever as soluções utilizando funções elementares. Essa maior complexidade das soluções em comparação à EDO de fato ocorre com muita frequência, e é muito improvável que as soluções possam ser explicitamente descritas. Dessa forma, é necessário encontrar outras maneiras de descrever as soluções.

A ideia fundamental de Poincaré no final do século XIX foi trocar a descrição quantitativa usual da época por uma descrição mais qualitativa, que explique as soluções de outro modo, por exemplo a partir de suas propriedades topológicas. Noutras palavras: em muitos casos não precisamos (e nem é possível) encontrar a equação explícita das soluções, mas podemos descrever suas propriedades. Vamos entender essa mudança de perspectiva a partir do próximo exemplo.

**Exemplo 10** Considere  $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  dada por  $F(t, x) = x(x - 1)$ . A EDO é

$$x' = x(x - 1),$$

com  $k = d = 1$ . Para encontrar as soluções, escrevemos  $x' = \frac{dx}{dt}$  e resolvemos a



equação:

$$\begin{aligned} \frac{dx}{dt} = x(x-1) &\implies \frac{dx}{x(x-1)} = dt \implies \left(-\frac{1}{x} + \frac{1}{x-1}\right) dx = dt \\ \implies \int \left(-\frac{1}{x} + \frac{1}{x-1}\right) dx &= \int dt \implies -\ln|x| + \ln|x-1| = t + c_0 \\ \implies \ln\left|\frac{x-1}{x}\right| = t + c_0 &\implies \left|\frac{x-1}{x}\right| = e^{t+c_0} = ce^t, \end{aligned}$$

onde  $c_0$  e  $c = e^{c_0}$  são constantes. A depender do comportamento de  $x$  entre os valores 0 e 1, podemos resolver a última equação:

◦  $x > 1$ : temos

$$ce^t = \left|\frac{x-1}{x}\right| = \frac{x-1}{x} \implies x(t) = \frac{1}{1-ce^t}.$$

◦  $0 < x < 1$ : temos

$$ce^t = \left|\frac{x-1}{x}\right| = \frac{1-x}{x} \implies x(t) = \frac{1}{1+ce^t}.$$

◦  $x < 0$ : similarmente ao primeiro caso, obtemos  $x(t) = \frac{1}{1-ce^t}$ .

Devemos ter cuidado quando  $x = 0$  ou  $x = 1$ , pois nesse caso os cálculos anteriores não são formalmente justificados (por exemplo, dividimos por 0 ou tomamos o logaritmo de 0). Ademais, notamos também que as soluções não podem ser definidas para todo  $t$ , pois  $1 - ce^t$  pode ser zero. Para realizar uma análise matematicamente correta, passamos para uma análise qualitativa das soluções, seguindo os passos abaixo.

PASSO 1: Análise de  $x = 0$  e  $x = 1$  separadamente.

Notamos que  $\gamma(t) \equiv 0$  e  $\gamma(t) \equiv 1$  são soluções:

$$\begin{aligned} \gamma(t) \equiv 0 &\implies \begin{cases} \gamma'(t) = 0 \\ F(t, \gamma(t)) = 0(0-1) = 0. \end{cases} \\ \gamma(t) \equiv 1 &\implies \begin{cases} \gamma'(t) = 0 \\ F(t, \gamma(t)) = 1(1-1) = 0. \end{cases} \end{aligned}$$

PASSO 2: Análise local do espaço de soluções.

Desenhamos o plano  $(t, x)$ , e representamos as soluções  $(t, \gamma(t))$  como gráficos nesse plano. As duas soluções do Passo 1 são as retas horizontais  $x = 0$  e  $x = 1$ . Dado um ponto  $(t_0, x_0)$ , o valor  $F(t_0, x_0)$  é o coeficiente angular da reta tangente ao gráfico da solução nesse ponto:

- Se  $x_0 > 1$ , então a solução  $\gamma$  com  $\gamma(t_0) = x_0$  satisfaz

$$\gamma'(t_0) = F(t_0, \gamma(t_0)) = F(t_0, x_0) = x_0(x_0 - 1) > 0$$

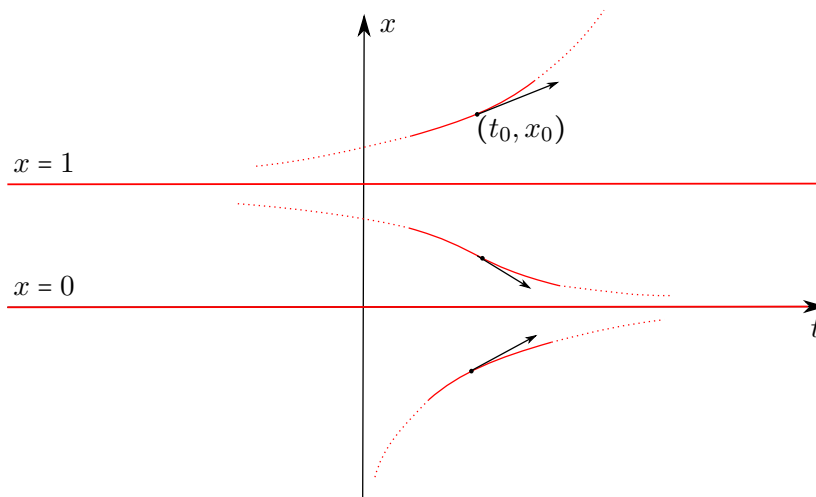
e portanto  $\gamma$  é crescente numa vizinhança de  $t = t_0$ .

- Se  $0 < x_0 < 1$ , então

$$\gamma'(t_0) = x_0(x_0 - 1) < 0$$

e portanto  $\gamma$  é decrescente numa vizinhança de  $t = t_0$ .

- Se  $x_0 < 0$ , então novamente  $\gamma$  é crescente numa vizinhança de  $t = t_0$ .



PASSO 3: As soluções independem do tempo.

Essa propriedade de fato ocorre para toda  $F$  que não depende de  $t$ . Tais EDO's são conhecidas como *autônomas*.

**Definição 11** (EDO autônoma) Dizemos que a EDO (1.1) é *autônoma* se  $F$  não depende de  $t$ , isto é, se existe  $G: V \rightarrow \mathbb{R}^d$  onde  $V$  é um aberto de  $\mathbb{R}^{dk}$  tal que

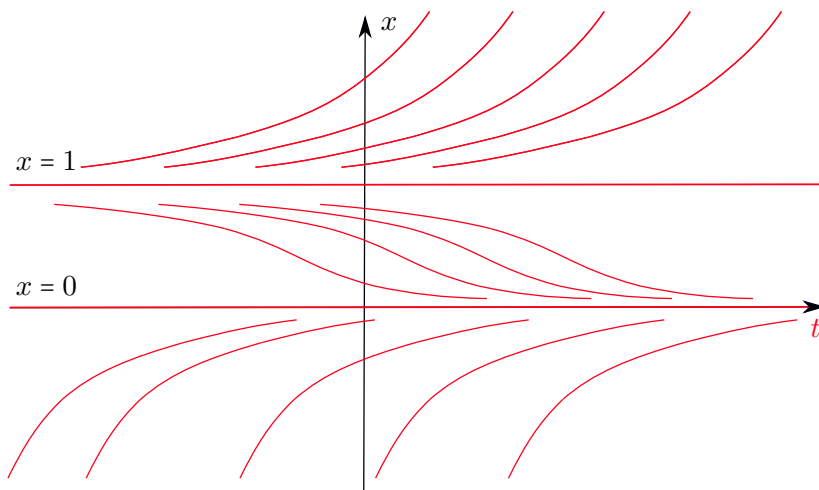
$$F(t, x_0, x_1, \dots, x_{k-1}) = G(x_0, x_1, \dots, x_{k-1})$$

para todo  $(t, x_0, x_1, \dots, x_{k-1}) \in U$ .

**Lema 12** Seja  $x' = F(t, x)$  uma EDO autônoma. Se  $\gamma(t)$  é solução da EDO, então  $\beta(t) = \gamma(t + c)$  também é solução da EDO, para todo  $c \in \mathbb{R}$ .

PROVA Temos  $\beta'(t) = \gamma'(t + c) = F(t + c, \gamma(t + c)) = F(t, \gamma(t + c)) = F(t, \beta(t))$ , onde na terceira igualdade utilizamos que  $F$  é autônoma.  $\square$

Pelo lema acima, soluções particulares podem ser horizontalmente transladadas no plano  $(t, x)$  para gerar novas soluções. Pelas soluções explícitas encontradas acima, o espaço de soluções fica da forma abaixo:



Observe que, na figura acima, as curvas cobrem todo o plano e quaisquer duas delas não se intersectam. O resultado principal desse capítulo, que representa o teorema básico na teoria qualitativa de EDO's, mostrará que esse é de fato o caso para uma grande classe de funções  $F$ .

SEÇÃO 2

**EDO de ordem superior**

---

Na seção anterior, vimos exemplos de EDO's de ordem superior ( $k > 1$ ). Agora, veremos que toda EDO de ordem superior é equivalente a uma EDO de ordem 1. Isso se tornará útil quando provarmos resultados gerais na teoria, podendo nos restringir ao caso  $k = 1$ . Para entender essa equivalência, vejamos como fazê-la quando  $k = 2$ . Considere a EDO de ordem 2 e dimensão  $d$

$$x'' = F(t, x, x')$$

onde  $F : U \subset \mathbb{R}^{1+2d} \rightarrow \mathbb{R}^d$ . A ideia é introduzir a variável  $y = x'$  e considerar o par  $X = (x, y)$ . Temos

$$\begin{cases} x' = y & , \text{ por definição} \\ y' = x'' = F(t, x, x') = F(t, x, y) & , \text{ pela EDO de ordem 2.} \end{cases}$$

Definindo  $G : U \subset \mathbb{R}^{1+2d} \rightarrow \mathbb{R}^{2d}$  por  $G(t, X) = G(t, (x, y)) = (y, F(t, x, y))$ , podemos escrever as igualdades acima como

$$X' = G(t, X),$$

que é uma EDO de ordem  $k = 1$  e dimensão  $2d$ . O próximo resultado realiza essa equivalência para qualquer  $k$ .

**Proposição 13** | Toda EDO de ordem  $k$  é equivalente a uma EDO de ordem 1. Mais especifi-

camente, se

$$x^{(k)} = F(t, x, x', \dots, x^{(k-1)})$$

é uma EDO de ordem  $k$ , onde  $F : U \subset \mathbb{R}^{1+kd} \rightarrow \mathbb{R}^d$  com  $U$  aberto, então existe  $G : U \rightarrow \mathbb{R}^{kd}$  tal que as EDO's com condições iniciais

$$\begin{cases} x^{(k)} = F(t, x, x', \dots, x^{(k-1)}) \\ x(t_0) = x_0, x'(t_0) = x'_0, \dots, x^{(k-1)}(t_0) = x_0^{(k-1)} \end{cases} \quad (2.1)$$

e

$$\begin{cases} X' = G(t, X) \\ X(t_0) = (x_0, x'_0, \dots, x_0^{(k-1)}) \end{cases} \quad (2.2)$$

possuem soluções em relação biunívoca:  $\gamma : I \rightarrow \mathbb{R}^d$  é solução de (2.1) se e somente se  $\Gamma : I \rightarrow \mathbb{R}^{kd}$ ,  $\Gamma(t) = (\gamma(t), \gamma'(t), \dots, \gamma^{(k-1)}(t))$ , é solução de (2.2).

PROVA Introduza as variáveis  $y_1, y_2, \dots, y_{k-1}$  pondo

$$y_1 = x', y_2 = x'', \dots, y_{k-1} = x^{(k-1)},$$

e seja  $X = (x, y_1, y_2, \dots, y_{k-1}) \in \mathbb{R}^{kd}$ . Então  $X' = (x', y'_1, y'_2, \dots, y'_{k-1})$  satisfaz

$$\begin{cases} x' = y_1 \\ y'_1 = x'' = y_2 \\ \vdots \\ y'_{k-2} = x^{(k-1)} = y_{k-1} \\ y'_{k-1} = x^{(k)} = F(t, x, x', \dots, x^{(k-1)}) = F(t, x, y_1, \dots, y_{k-1}) = F(t, X), \end{cases}$$

que pode ser escrita vetorialmente como

$$X' = \begin{bmatrix} x' \\ y'_1 \\ \vdots \\ y'_{k-2} \\ y'_{k-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{k-1} \\ F(t, X) \end{bmatrix}.$$

Defina  $G : U \rightarrow \mathbb{R}^{kd}$  por

$$G(t, X) = G(t, x, y_1, \dots, y_{k-1}) = (y_1, y_2, \dots, y_{k-1}, F(t, x, y_1, \dots, y_{k-1})).$$

FATO 1. Se  $\gamma : I \rightarrow \mathbb{R}^d$  é solução de (2.1) então  $\Gamma : I \rightarrow \mathbb{R}^{kd}$ ,  $\Gamma(t) = (\gamma(t), \gamma'(t), \dots, \gamma^{(k-1)}(t))$ , é solução de (2.2).

Para provar isso, note que

$$\Gamma'(t) = \begin{bmatrix} \gamma'(t) \\ \gamma''(t) \\ \vdots \\ \gamma^{(k-1)}(t) \\ \gamma^{(k)}(t) \end{bmatrix} = \begin{bmatrix} \gamma'(t) \\ \gamma''(t) \\ \vdots \\ \gamma^{(k-1)}(t) \\ F(t, \gamma(t), \dots, \gamma^{(k-1)}(t)) \end{bmatrix}$$

e

$$G(t, \Gamma(t)) = G(t, \gamma(t), \gamma'(t), \dots, \gamma^{(k-1)}(t)) = \begin{bmatrix} \gamma'(t) \\ \gamma''(t) \\ \vdots \\ \gamma^{(k-1)}(t) \\ F(t, \gamma(t), \dots, \gamma^{(k-1)}(t)) \end{bmatrix},$$

portanto  $\Gamma'(t) = G(t, \Gamma(t))$ . Ademais,  $\Gamma(t_0) = (\gamma(t_0), \gamma'(t_0), \dots, \gamma^{(k-1)}(t_0)) = (x_0, x'_0, \dots, x_0^{(k-1)})$ .

FATO 2. Se  $\Gamma : I \rightarrow \mathbb{R}^{kd}$ ,  $\Gamma(t) = (\Gamma_0(t), \dots, \Gamma_{k-1}(t))$ , é solução de (2.2) com  $\Gamma(t_0) = (x_0, x'_0, \dots, x_0^{(k-1)})$ , então  $\gamma : I \rightarrow \mathbb{R}^d$ ,  $\gamma(t) = \Gamma_0(t)$ , é solução de (2.1) com  $\gamma(t_0) = x_0, \gamma'(t_0) = x'_0, \dots, \gamma^{(k-1)}(t_0) = x_0^{(k-1)}$ .

Para provar esse fato, note que  $\Gamma'(t) = G(t, \Gamma(t))$  se escreve como

$$\begin{bmatrix} \Gamma'_0(t) \\ \Gamma'_1(t) \\ \vdots \\ \Gamma'_{k-2}(t) \\ \Gamma'_{k-1}(t) \end{bmatrix} = \begin{bmatrix} \Gamma_1(t) \\ \Gamma_2(t) \\ \vdots \\ \Gamma_{k-1}(t) \\ F(t, \Gamma_0(t), \dots, \Gamma_{k-1}(t)) \end{bmatrix}$$

Assim, se  $\gamma = \Gamma_0$ , então  $\Gamma_i(t) = \gamma^{(i)}(t)$  para  $i = 0, \dots, k-1$ , e portanto

$$\gamma^{(k)}(t) = \Gamma'_{k-1}(t) = F(t, \Gamma_0(t), \dots, \Gamma_{k-1}(t)) = F(t, \gamma(t), \dots, \gamma^{(k-1)}(t)).$$

As condições iniciais são

$$\begin{cases} \gamma(t_0) = \Gamma_0(t_0) = x_0 \\ \gamma'(t_0) = \Gamma_1(t_0) = x'_0 \\ \vdots \\ \gamma^{(k-1)}(t_0) = \Gamma_{k-1}(t_0) = x_0^{(k-1)}. \end{cases}$$

□

### SEÇÃO 3

## Teorema de existência e unicidade de EDO's

Como vimos na seção anterior, os resultados sobre EDO's de ordem superior podem ser obtidos a partir de suas versão com  $k = 1$ . É nesse último contexto que enunciamos o próximo teorema, que corresponde ao principal resultado desse capítulo.

**Teorema 14** (Existência e unicidade) Considere a EDO

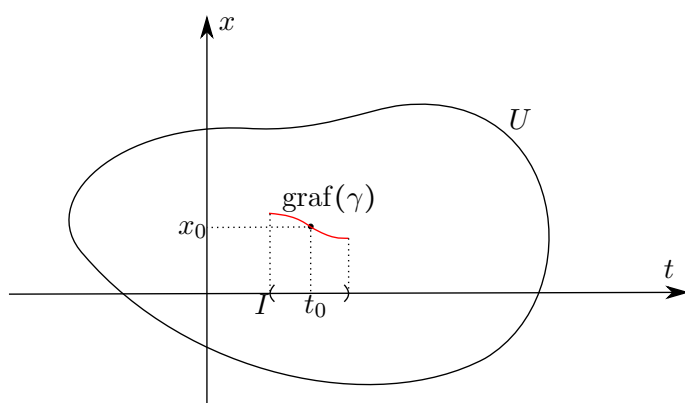
$$x' = F(t, x),$$

onde  $F : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  é uma função contínua e localmente Lipschitz na

segunda coordenada e  $U$  é aberto. Então:

- (1) **EXISTÊNCIA:** para todo  $(t_0, x_0) \in U$ , existe  $\gamma : I \rightarrow \mathbb{R}^d$  solução da EDO com condição inicial  $\gamma(t_0) = x_0$ .
- (2) **UNICIDADE:** Se  $\gamma_1 : I \rightarrow \mathbb{R}^d$  e  $\gamma_2 : J \rightarrow \mathbb{R}^d$  são ambas soluções da EDO e se existe  $t \in I \cap J$  tal que  $\gamma_1(t) = \gamma_2(t)$ , então  $\gamma_1(t) = \gamma_2(t)$  para todo  $t \in I \cap J$ .

A primeira parte diz que para toda condição inicial  $(t_0, x_0) \in U$ , existe uma solução da EDO com essa condição inicial. Esperamos que se  $(t_0, x_0)$  está próximo de  $\partial U$ , então o domínio  $I$  de definição da solução é pequeno. A segunda parte diz que os gráficos das soluções não podem se cruzar



Resta explicar as condições requeridas na função  $F$ . Usualmente, a norma considerada em  $\mathbb{R}^{1+d}$  é a euclidiana,

$$\|(t, x)\| = \sqrt{t^2 + x^2}.$$

Como  $\mathbb{R}^{1+d}$  tem dimensão finita, todas as normas são equivalentes e, para o que estamos interessados, é mais conveniente considerar a norma da soma com respeito à decomposição  $\mathbb{R}^{1+d} = \mathbb{R} \times \mathbb{R}^d$ ,

$$\|(t, x)\| = |t| + \|x\|,$$

onde  $\|x\|$  é a norma euclidiana em  $\mathbb{R}^d$ .

#### Definição 15

(Função Lipschitz) Dizemos que  $G : V \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$  é *Lipschitz* se existe  $C > 0$  tal que

$$\|G(z_1) - G(z_2)\| \leq C\|z_1 - z_2\|, \quad \forall z_1, z_2 \in V.$$

Dizemos que  $G$  é *localmente Lipschitz* se, para todo  $z \in V$ , existem  $C, \varepsilon > 0$  tais que  $B(z, \varepsilon) \subset V$  e

$$\|G(z_1) - G(z_2)\| \leq C\|z_1 - z_2\|, \quad \forall z_1, z_2 \in B(z, \varepsilon).$$

Para  $n = m = 1$ , podemos interpretar o quociente  $\frac{|G(z_1) - G(z_2)|}{|z_1 - z_2|}$  como a in-

clinação da reta secante aos pontos  $(z_1, G(z_1))$  e  $(z_2, G(z_2))$  do gráfico de  $G$ . Logo, ser Lipschitz significa que o gráfico de  $G$  está contido nos cones horizontais definidos pelos pontos de  $V$  e abertura que está relacionada com o valor de  $C$ .

**Exemplo 16** Se  $G$  é  $C^1$ , então  $G$  é localmente Lipschitz. De fato, pela desigualdade do valor médio temos em  $B(z, \varepsilon) \subset U$  que

$$\|G(z_1) - G(z_2)\| \leq \sup_{w \in B(z, \varepsilon)} \|G'(w)\| \cdot \|z_1 - z_2\|.$$

Basta tomar  $C = \sup_{w \in B(z, \varepsilon)} \|G'(w)\| < \infty$ .

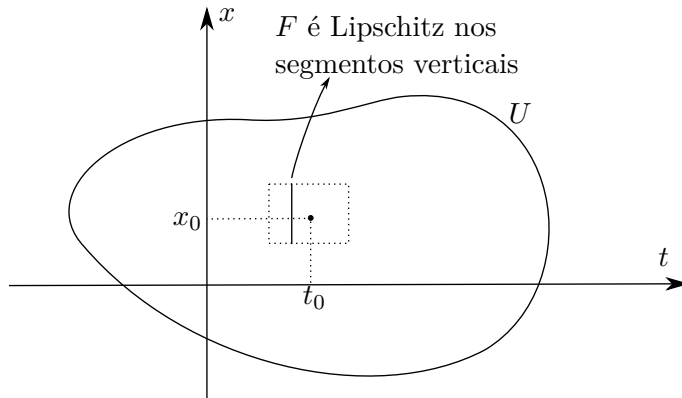
**Exemplo 17** A função  $G : \mathbb{R} \rightarrow \mathbb{R}$ ,  $G(z) = \sqrt{|z|}$ , não é localmente Lipschitz em  $z = 0$ . Geometricamente, isso é fácil de provar notando que o gráfico de  $G$  escapa de qualquer cone horizontal centrado em  $(0, 0)$ . Analiticamente, basta notar que  $\lim_{z \rightarrow 0} |G'(z)| = \infty$ . O mesmo vale para  $G(z) = |z|^\alpha$ ,  $\alpha \in (0, 1)$ .

Agora procedemos à definição que aparece no enunciado do Teorema 14.

**Definição 18** (Localmente Lipschitz na segunda coordenada) Dizemos que  $F : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  é *localmente Lipschitz na segunda coordenada* se, para todo  $(t_0, x_0) \in U$ , existem  $C, \varepsilon > 0$  tais que:

- (1)  $B(t_0, \varepsilon) \times B(x_0, \varepsilon) \subset U$ , onde  $B(t_0, \varepsilon) = (t_0 - \varepsilon, t_0 + \varepsilon) \subset \mathbb{R}$  e  $B(x_0, \varepsilon) \subset \mathbb{R}^d$ .
- (2) Para  $t \in B(t_0, \varepsilon)$  e  $x_1, x_2 \in B(x_0, \varepsilon)$  vale que

$$\|F(t, x_1) - F(t, x_2)\| \leq C \|x_1 - x_2\|.$$



**Exemplo 19** Se  $F \in C^1$ , então  $F$  é localmente Lipschitz na segunda coordenada. Por outro lado, se  $F$  possui  $\sqrt{|x|}$  em sua expressão, por exemplo  $F(t, x) = t + 2\sqrt{|x|}$ , então  $F$  não é localmente Lipschitz na segunda coordenada. O mesmo ocorre se  $F$  possui  $|x|^\alpha$ ,  $\alpha \in (0, 1)$ , em sua expressão.

Antes de proceder ao desenvolvimento dos pré-requisitos e posterior prova do Teorema 14, vamos mostrar que a condição de  $F$  ser localmente Lipschitz na

segunda coordenada é necessária.

**Exemplo 20** Seja  $F : \mathbb{R} \rightarrow \mathbb{R}$ ,  $F(x) = x^{2/3}$ . Então as soluções da EDO não são únicas. Mais especificamente, existem duas soluções para a EDO com algumas condições iniciais. Para ver isso, comece notando que  $\gamma(t) \equiv 0$  é solução. Vamos encontrar outras soluções com  $\gamma(t_0) = 0$ . A EDO é

$$\begin{aligned} x' = x^{2/3} &\implies \frac{dx}{dt} = x^{2/3} \implies x^{-2/3} dx = dt \implies \int x^{-2/3} dx = \int dt \\ \implies \frac{x^{1/3}}{1/3} = t + c &\implies x(t) = \left(\frac{t+c}{3}\right)^3. \end{aligned}$$

Note que  $x(-c) = 0$ . Geometricamente, os gráficos dessas funções cruzam o eixo  $x = 0$ .

### SUBSEÇÃO 3.1

#### Preparação para a prova do Teorema 14

**Definição 21** (Espaço métrico) Um *espaço métrico* é um par  $(X, d)$ , onde  $X$  é um conjunto e  $d : X \times X \rightarrow \mathbb{R}$  é uma distância, isto é, satisfaz as seguintes propriedades:

- (1) POSITIVIDADE:  $d(x, y) \geq 0$ , com igualdade se e somente se  $x = y$ .
- (2) SIMETRIA:  $d(x, y) = d(y, x)$ .
- (3) DESIGUALDADE TRIANGULAR:  $d(x, z) \leq d(x, y) + d(y, z)$ ,  $\forall x, y, z \in X$ .

**Exemplo 22** Em  $X = \mathbb{R}^n$ , temos várias distâncias úteis. A distância euclidiana é

$$d((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}.$$

A distância da soma é

$$d_{\text{soma}}((x_1, \dots, x_n), (y_1, \dots, y_n)) = |x_1 - y_1| + \dots + |x_n - y_n|.$$

A distância do máximo é

$$d_{\infty}((x_1, \dots, x_n), (y_1, \dots, y_n)) = \max_{1 \leq i \leq n} |x_i - y_i|.$$

**Exemplo 23** (Distância em espaços de funções) Tome

$$X = C^0[0, 1] = \{f : [0, 1] \rightarrow \mathbb{R} : f \text{ é contínua}\}.$$

Uma possível distância em  $X$  é

$$d_{\infty}(f, g) = \sup_{t \in [0, 1]} |f(x) - g(x)| = \max_{t \in [0, 1]} |f(x) - g(x)|.$$



Outras distâncias são as normas  $L^p$ : dado  $p \geq 1$ , defina

$$d_{L^p}(f, g) = \left( \int_0^1 |f(x) - g(x)|^p dx \right)^{1/p}.$$

**Definição 24** (Espaço métrico completo) Dizemos que o espaço métrico  $(X, d)$  é *completo* se toda sequência de Cauchy  $(x_n)_{n \geq 0}$  em  $X$  possui um limite  $x \in X$ .

**Exemplo 25** O espaço métrico  $([0, 1], d)$  com  $d(x, y) = |x - y|$  é completo, pois  $[0, 1]$  é compacto. Já  $((0, 1], d)$  não é completo, pois  $(x_n) = (\frac{1}{n})_{n \geq 1}$  é de Cauchy mas  $\lim x_n = 0 \notin (0, 1]$ . O espaço métrico  $(C^0[a, b], d_\infty)$  é completo (Exercício 1).

**Definição 26** (Contração) Dado um espaço métrico  $(X, d)$ , dizemos que  $f : X \rightarrow X$  é uma *contração* se existir  $\lambda \in (0, 1)$  tal que

$$d(f(x), f(y)) \leq \lambda d(x, y), \quad \forall x, y \in X.$$

Em outras palavras,  $f$  contrai distâncias. É fácil notar que toda contração é contínua (Exercício 2).

**Exemplo 27** Tome  $f : [0, 1] \rightarrow [0, 1]$  diferenciável com  $|f'| \leq \lambda < 1$ . Pelo teorema do valor médio, temos que

$$|f(x) - f(y)| \leq |f'(\xi)| \cdot |x - y| \leq \lambda |x - y|.$$

**Exemplo 28** Tome  $X = C^0[a, b]$  e  $\mathcal{F} : X \rightarrow X$  dada por  $(\mathcal{F}f)(x) = \frac{1}{2}f(x)$ . Temos

$$|(\mathcal{F}f)(x) - (\mathcal{F}g)(x)| = \frac{1}{2}|f(x) - g(x)|,$$

logo  $d(\mathcal{F}f, \mathcal{F}g) = \frac{1}{2}d(f, g)$ , provando que  $\mathcal{F}$  é uma contração.

Faremos uso essencial do seguinte resultado.

**Teorema 29** (Ponto fixo para contrações) Sejam  $(X, d)$  um espaço métrico completo e  $\mathcal{F} : X \rightarrow X$  uma contração. Então  $\mathcal{F}$  possui um único ponto fixo  $\bar{x} \in X$ . Ademais,

$$\bar{x} = \lim_{n \rightarrow \infty} \mathcal{F}^n x, \quad \forall x \in X,$$

e a convergência é exponencial.

**PROVA** É fácil ver que  $\mathcal{F}$  possui no máximo um ponto fixo : se  $x_0, x_1 \in X$  são pontos fixos com  $x_0 \neq x_1$ , então

$$d(x_0, x_1) = d(\mathcal{F}x_0, \mathcal{F}x_1) \leq \lambda d(x_0, x_1),$$

o que contradiz o fato de  $\lambda < 1$ . Logo, resta provar que para qualquer  $x \in X$  a sequência  $(x_n)_{n \geq 0}$  definida por  $x_n = \mathcal{F}^n x$  converge para um ponto fixo de  $\mathcal{F}$ .

**FATO 1.** A sequência  $(x_n)_{n \geq 0}$  é de Cauchy.

PROVA (Fato 1) Temos

$$d(x_n, x_{n+1}) \leq \lambda d(x_{n-1}, x_n) \leq \dots \leq \lambda^n d(x_0, x_1).$$

Logo, para  $n < m$  temos

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{m-1}, x_m) \\ &\leq (\lambda^n + \lambda^{n+1} + \dots + \lambda^{m-1})d(x_0, x_1) < \frac{\lambda^n}{1-\lambda}d(x_0, x_1). \end{aligned}$$

Dado  $\varepsilon > 0$ , tome  $n_0$  tal que  $\frac{\lambda^n}{1-\lambda}d(x_0, x_1) < \varepsilon$ . Assim,  $d(x_n, x_m) < \varepsilon$  para todos  $m, n \geq n_0$ .  $\square$

FATO 2. O limite  $\bar{x} = \lim_{n \rightarrow \infty} x_n$  é ponto fixo de  $\mathcal{F}$ .

PROVA (Fato 2) Temos

$$\mathcal{F}(\bar{x}) = \mathcal{F}(\lim x_n) = \lim \mathcal{F}x_n = \lim x_{n+1} = \bar{x},$$

onde na segunda passagem usamos que  $\mathcal{F}$  é contínua.  $\square$

Resta notar que  $\bar{x} = \lim_{n \rightarrow \infty} x_n$  exponencialmente rápido. De fato,

$$d(\bar{x}, x_n) = d(\mathcal{F}\bar{x}, \mathcal{F}x_{n-1}) \leq \lambda d(\bar{x}, x_{n-1}) \leq \dots \leq \lambda^n d(\bar{x}, x_0).$$

$\square$

### SUBSEÇÃO 3.2

#### Prova do Teorema 14

A ideia é trabalhar com a equação integral equivalente à EDO. Vamos explicar o que isso significa. No contexto do Teorema 14, seja  $\gamma : I \rightarrow \mathbb{R}^d$  uma solução da EDO com condições iniciais

$$\begin{cases} x' = F(t, x) \\ x(t_0) = x_0. \end{cases} \quad (3.1)$$

Então  $\gamma'(t) = F(t, \gamma(t))$  para todo  $t \in I$ . Fixando  $t_0$  e integrando de  $t_0$  a  $t$ , o Teorema Fundamental do Cálculo dá que

$$\begin{aligned} \gamma(t) - \gamma(t_0) &= \int_{t_0}^t \gamma'(s) ds = \int_{t_0}^t F(s, \gamma(s)) ds \\ \implies \gamma(t) &= x_0 + \int_{t_0}^t F(s, \gamma(s)) ds. \end{aligned}$$

Dizemos que

$$x(t) = x_0 + \int_{t_0}^t F(s, \gamma(s)) ds$$

é a equação integral associada à EDO com condições iniciais (3.1). Tal equação é de fato equivalente a (3.1): como vimos, uma solução de (3.1) é solução da

equação integral acima; reciprocamente, se  $\gamma(t)$  satisfaz a equação integral, então

$$\begin{cases} \gamma'(t) = F(t, \gamma(t)), \quad \forall t \in I \\ \gamma(t_0) = x_0 + \int_{t_0}^t F(s, \gamma(s)) ds = x_0, \end{cases}$$

onde na primeira igualdade utilizamos mais uma vez o Teorema Fundamental do Cálculo. Assim:

EDO com condições iniciais

Equação integral

$$\begin{cases} x' = F(t, x) \\ x(t_0) = x_0 \end{cases} \longleftrightarrow x(t) = x_0 + \int_{t_0}^t F(s, \gamma(s)) ds$$

Agora procedemos à prova do Teorema 14. Fixe  $(t_0, x_0) \in U$ . Tome  $\delta > 0$  pequeno tal que

$$V = [t_0 - \delta, t_0 + \delta] \times \overline{B(x_0, \delta)} \subset U.$$

Como  $F$  é contínua,  $M := \max_V \|F(t, x)\|$  é finito. Seja  $C > 0$  uma constante de Lipschitz na segunda coordenada em  $V$ . Tome  $\varepsilon < \min\{\delta, \frac{\delta}{M}, \frac{1}{C}\}$ , e defina

$$X = \left\{ \gamma : [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \mathbb{R}^d : \begin{array}{l} \gamma \text{ é contínua e } \gamma(t) \in \overline{B(x_0, \delta)} \\ \text{para todo } 0 \leq |t - t_0| \leq \varepsilon \end{array} \right\}.$$

com a norma  $d$  do máximo (que coincide com a norma do supremo). Então  $(X, d)$  é espaço métrico completo. Vamos definir uma contração em  $(X, d)$ , de modo que seu ponto fixo é uma solução da equação integral. Tome  $\mathcal{F} : X \rightarrow X$  pondo

$$(\mathcal{F}\gamma)(t) = x_0 + \int_{t_0}^t F(s, \gamma(s)) ds, \quad t \in [t_0 - \varepsilon, t_0 + \varepsilon].$$

Pela equivalência entre a EDO e a equação integral,  $\mathcal{F}\gamma = \gamma$  se e somente se  $\gamma$  é solução da EDO com condições iniciais (3.1).

FATO 1.  $\mathcal{F}$  está bem-definida.

PROVA (Fato 1) A integral acima está bem-definida porque  $\gamma$  é contínua. Ademais,  $\mathcal{F}\gamma$  é diferenciável, logo contínua. Resta checar que  $(\mathcal{F}\gamma)(t) \in \overline{B(x_0, \delta)}$ . Ora,

$$\begin{aligned} \|(\mathcal{F}\gamma)(t) - x_0\| &= \left\| \int_{t_0}^t F(s, \gamma(s)) ds \right\| \leq \int_{t_0}^t \|F(s, \gamma(s))\| ds \\ &\leq M|t - t_0| \leq \varepsilon M \leq \delta. \end{aligned}$$

□

FATO 2.  $\mathcal{F}$  é contração.

PROVA (Fato 2) Tome  $\gamma_1, \gamma_2 \in X$ . Temos

$$(\mathcal{F}\gamma_1)(t) - (\mathcal{F}\gamma_2)(t) = \int_{t_0}^t [F(s, \gamma_1(s)) - F(s, \gamma_2(s))] ds$$

e portanto

$$\begin{aligned} \|(\mathcal{F}\gamma_1)(t) - (\mathcal{F}\gamma_2)(t)\| &\leq \int_{t_0}^t \|F(s, \gamma_1(s)) - F(s, \gamma_2(s))\| ds \\ &\leq \int_{t_0}^t C \|\gamma_1(s) - \gamma_2(s)\| ds = C|t - t_0|d(\gamma_1, \gamma_2) \leq C\varepsilon d(\gamma_1, \gamma_2). \end{aligned}$$

Fazendo  $\lambda := C\varepsilon < 1$ , concluímos que  $\|(\mathcal{F}\gamma_1)(t) - (\mathcal{F}\gamma_2)(t)\| \leq \lambda d(\gamma_1, \gamma_2)$  para todo  $|t - t_0| \leq \varepsilon$  e assim  $d(\mathcal{F}\gamma_1, \mathcal{F}\gamma_2) \leq \lambda d(\gamma_1, \gamma_2)$ .  $\square$

Pelo Teorema 29 segue que  $\mathcal{F}$  possui um (único) ponto fixo. Isso prova a parte (1) do Teorema 14.

Provemos agora a unicidade. Assuma que  $\gamma_1 : I \rightarrow \mathbb{R}^d$  e  $\gamma_2 : J \rightarrow \mathbb{R}^d$  são ambas soluções de (3.1) e que exista  $\bar{t} \in I \cap J$  tal que  $\gamma_1(\bar{t}) = \gamma_2(\bar{t})$ . Vamos utilizar um argumento de *conexidade* para mostrar que  $\gamma_1(t) = \gamma_2(t)$  para todo  $t \in I \cap J$ . Defina

$$A = \{t \in I \cap J : \gamma_1(t) = \gamma_2(t)\}.$$

Por hipótese,  $\bar{t} \in A$  e portanto  $A \neq \emptyset$ . É claro que  $A$  é fechado, pois  $\gamma_1, \gamma_2$  são contínuas. Resta mostrar que  $A$  é aberto. Fixe  $t_0 \in A$ , seja  $x_0 := \gamma_1(t_0) = \gamma_2(t_0)$  e tome  $\varepsilon > 0$  como na prova da parte (1) e de modo que  $[t_0 - \varepsilon, t_0 + \varepsilon] \subset I \cap J$ . Defina  $\mathcal{F}$  como acima. Então as restrições de  $\gamma_1, \gamma_2$  ao intervalo  $[t_0 - \varepsilon, t_0 + \varepsilon]$  são ambas pontos fixos de  $\mathcal{F}$ . Pela unicidade do ponto fixo, concluímos que  $\gamma_1(t) = \gamma_2(t)$  para todo  $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$  e assim  $(t_0 - \varepsilon, t_0 + \varepsilon) \subset A$ . Isso conclui a prova da parte (2), e também do teorema.

### SUBSEÇÃO 3.3

#### Exemplos

Como vimos no Teorema 29, a partir de qualquer condição inicial  $x \in X$  podemos encontrar o único ponto fixo  $\bar{x}$  por meio de iterações sucessivas:

$$\bar{x} = \lim_{n \rightarrow \infty} \mathcal{F}^n(x).$$

Em particular, podemos obter uma sequência de funções que converge uniformemente e exponencialmente rápido para a única solução da EDO com condições iniciais, fornecida pelo Teorema 14. Vamos exemplificar isso em duas situações (EDO autônoma e EDO não-autônoma).

#### Exemplo 30

Seja  $F(t, x) = x$ , que define a EDO autônoma  $x' = x$ . Vamos achar a solução da EDO com condições iniciais  $x(0) = 1$  por meio de iterações sucessivas. Sabemos que ela é  $\gamma(t) = e^t$ . Considere a curva inicial constante com condição inicial requerida, i.e.  $\gamma_0(t) = 1$  para todo  $t \in \mathbb{R}$ . A contração  $\mathcal{F} : X \rightarrow X$  é dada por

$$(\mathcal{F}\gamma)(t) = x_0 + \int_{t_0}^t F(s, \gamma(s)) ds = 1 + \int_0^t \gamma(s) ds.$$

Sendo  $\gamma_n = \mathcal{F}^n(\gamma_0)$ , temos:

- $\gamma_1(t) = 1 + \int_0^t 1 ds = 1 + t$ .
- $\gamma_2(t) = 1 + \int_0^t (1 + s) ds = 1 + t + \frac{t^2}{2!}$ .

$$\circ \gamma_3(t) = 1 + \int_0^t \left(1 + s + \frac{s^2}{2}\right) ds = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!}.$$

$$\circ \text{Por indução, } \gamma_n(t) = \sum_{k=0}^n \frac{t^k}{k!}.$$

Assim, a solução é  $\gamma(t) = \lim_{n \rightarrow \infty} \gamma_n(t) = e^t$ .

**Exemplo 31** Seja  $F(t, x) = 2tx$ , que define a EDO não-autônoma  $x' = 2tx$ . Considere a condição inicial  $x(0) = 1$ . Nesse caso,

$$(\mathcal{F}\gamma)(t) = 1 + \int_0^t 2s\gamma(s) ds.$$

Considerando novamente  $\gamma_0(t) \equiv 1$ , temos:

$$\circ \gamma_1(t) = 1 + \int_0^t 2s ds = 1 + t^2.$$

$$\circ \gamma_2(t) = 1 + \int_0^t 2s(1 + s^2) ds = 1 + t^2 + \frac{t^4}{2} = 1 + \frac{t^2}{1!} + \frac{t^4}{2!}.$$

$$\circ \gamma_3(t) = 1 + \int_0^t 2s \left(1 + \frac{s^2}{1!} + \frac{s^4}{2!}\right) ds = 1 + \frac{t^2}{1!} + \frac{t^4}{2!} + \frac{t^6}{3!}.$$

$$\circ \text{Por indução, } \gamma_n(t) = \sum_{k=0}^n \frac{t^{2k}}{k!}.$$

Portanto,  $\gamma(t) = \lim_{n \rightarrow \infty} \gamma_n(t) = e^{t^2}$ .

### SUBSEÇÃO 3.4

#### Métricas adaptadas

Vimos na prova do Teorema 14 que a EDO com condições iniciais possui uma única solução  $\gamma$  definida no intervalo  $(t_0 - \varepsilon, t_0 + \varepsilon)$ , onde  $\varepsilon > 0$  satisfaz as três desigualdades abaixo:

$$\circ \varepsilon \leq \delta,$$

$$\circ \varepsilon \leq \frac{\delta}{M}, \text{ onde } M = \max_{(t,x) \in V} \|F(t, x)\| \text{ para } V = [t_0 - \delta, t_0 + \delta] \times \overline{B(x_0, \delta)},$$

$$\circ \varepsilon < \frac{1}{C}, \text{ onde } C \text{ é a constante de Lipschitz na segunda coordenada da restrição de } F \text{ a } V.$$

As duas primeiras desigualdades são intrínsecas ao problema e necessárias para definir formalmente  $\mathcal{F} : X \rightarrow X$ , mas a terceira é um artefato da prova, utilizado para garantir que  $\mathcal{F}$  é uma contração e portanto possui um único ponto fixo. Abaixo, vamos utilizar uma versão refinada do teorema do ponto fixo, que não necessita da terceira desigualdade. A conclusão é que a única solução da EDO com condições iniciais está definida no intervalo  $(t_0 - \varepsilon, t_0 + \varepsilon)$  com  $\varepsilon = \min\{\delta, \frac{\delta}{M}\}$ . A nova versão do teorema do ponto fixo faz uso da noção de métricas equivalentes. Seja  $X$  um conjunto.

**Definição 32** (Métricas equivalentes) Dizemos que duas distâncias  $d, D$  em  $X$  são *equivalen-*

tes se existe  $A > 0$  tal que

$$\frac{1}{A}d(x, y) \leq D(x, y) \leq Ad(x, y), \quad \forall x, y \in X.$$

Se  $d, D$  são equivalentes, então os espaços métricos  $(X, d)$  e  $(X, D)$  são essencialmente iguais. Por exemplo, eles são iguais como espaços topológicos. Outra propriedade preservada é a convergência:

$$\lim x_n = x \text{ em } (X, d) \iff \lim x_n = x \text{ em } (X, D).$$

Uma propriedade que *não* é preservada é a de um mapa ser contração: uma contração em  $(X, d)$  é Lipschitz (com constante de Lipschitz  $A$ ) em  $(X, D)$ , mas não é necessariamente uma contração. O próximo resultado mostra, por outro lado, que se alguma potência de  $\mathcal{F}$  é uma contração em  $(X, d)$ , então podemos definir uma métrica equivalente  $D$  para a qual a própria  $\mathcal{F}$  é uma contração em  $(X, D)$ . Isso é útil porque, em muitos casos, o mapa que definimos não é uma contração mas apenas uma potência alta.

**Proposição 33** (Métrica adaptada) Seja  $(X, d)$  um espaço métrico, e seja  $F : (X, d) \rightarrow (X, d)$  Lipschitz. Se existe  $k \geq 1$  tal que  $F^k : (X, d) \rightarrow (X, d)$  é uma contração, então existe uma métrica  $D$  equivalente a  $d$  para a qual  $F : (X, D) \rightarrow (X, D)$  é uma contração. Em particular, se  $(X, d)$  é completo então  $F$  possui um único ponto fixo.

PROVA Por hipótese, existe  $\lambda_0 < 1$  tal que

$$d(F^k x, F^k y) \leq \lambda_0 d(x, y), \quad \forall x, y \in X.$$

Tome  $\lambda = \lambda_0^{1/k} \in (\lambda_0, 1)$ , e defina  $D : X \times X \rightarrow [0, \infty)$  por

$$\begin{aligned} D(x, y) &= \sum_{i=0}^{k-1} \lambda^{-i} d(F^i x, F^i y) \\ &= d(x, y) + \lambda^{-1} d(Fx, Fy) + \dots + \lambda^{-k+1} d(F^{k-1}x, F^{k-1}y). \end{aligned}$$

É fácil notar que  $D$  é uma distância (exercício). Vamos mostrar que

$$D(Fx, Fy) \leq \lambda D(x, y), \quad \forall x, y \in X. \quad (3.2)$$

Temos

$$\begin{aligned} D(Fx, Fy) &= d(Fx, Fy) + \lambda^{-1} D(F^2x, F^2y) + \dots + \lambda^{-k+1} d(F^k x, F^k y) \\ &\leq d(Fx, Fy) + \lambda^{-1} D(F^2x, F^2y) + \dots + \lambda^{-k+1} \lambda_0 d(x, y) \\ &= d(Fx, Fy) + \lambda^{-1} D(F^2x, F^2y) + \dots + \lambda d(x, y) \\ &= \lambda [d(x, y) + \lambda^{-1} d(Fx, Fy) + \dots + \lambda^{-k+1} d(F^{k-1}x, F^{k-1}y)] = \lambda D(x, y). \end{aligned}$$

Resta mostrar que  $d, D$  são equivalentes:

- $d(x, y) \leq D(x, y)$  claramente.

- o Para o outro lado, utilizamos que  $F$  é Lipschitz em  $(X, d)$ . Assumindo que  $d(Fx, Fy) \leq Cd(x, y)$ , temos

$$\begin{aligned} D(x, y) &= d(x, y) + \lambda^{-1}d(Fx, Fy) + \dots + \lambda^{-k+1}d(F^{k-1}x, F^{k-1}y) \\ &\leq d(x, y) + \left(\frac{C}{\lambda}\right)d(Fx, Fy) + \dots + \left(\frac{C}{\lambda}\right)^{k-1}d(x, y) \\ &= \underbrace{\left[1 + \frac{C}{\lambda} + \dots + \left(\frac{C}{\lambda}\right)^{k-1}\right]}_{=:A} d(x, y), \end{aligned}$$

donde  $D(x, y) \leq Ad(x, y)$ .

□

A métrica  $D$  construída na proposição acima é chamada *métrica adaptada*, e tem diversas aplicações em dinâmica, por exemplo em dinâmica hiperbólica.

### SUBSEÇÃO 3.5

#### Melhorando o domínio da solução

Utilizando a Proposição 33, vamos remover a hipótese  $\varepsilon < \frac{1}{C}$  na prova do Teorema 14 e obter uma solução definida no intervalo  $(t_0 - \varepsilon, t_0 + \varepsilon)$  com

$$\varepsilon = \min\left\{\delta, \frac{\delta}{M}\right\}.$$

Fixe  $\varepsilon$  como acima e defina  $\mathcal{F} : (X, d) \rightarrow (X, d)$  como na Seção 3.2. Vamos mostrar que  $\mathcal{F}^k$  é uma contração para  $k$  grande.

**Lema 34** Para todo  $k \geq 0$ , temos

$$d(\mathcal{F}^k(\gamma), \mathcal{F}^k(\beta)) \leq \frac{(C\varepsilon)^k}{k!} d(\gamma, \beta), \quad \forall \gamma, \beta \in X.$$

Como  $\lim_{k \rightarrow \infty} \frac{(C\varepsilon)^k}{k!} = 0$ , concluímos que  $\mathcal{F}^k$  é contração para  $k$  grande e portanto  $\mathcal{F}$  possui um único ponto fixo.

**PROVA** Vamos estimar  $\|(\mathcal{F}^k\gamma)(t) - (\mathcal{F}^k\beta)(t)\|$ . Sem perda de generalidade, assuma  $t > t_0$ . Escreva  $\gamma_k = \mathcal{F}^k\gamma$ ,  $\beta_k = \mathcal{F}^k\beta$ , e lembre que  $C > 0$  é a constante de Lipschitz na segunda coordenada de  $F$  restrito a  $V$ . Vamos mostrar que  $d(\gamma_k, \beta_k) \leq C^k d(\gamma, \beta) I_k(t)$ , onde  $I_k(t)$  é uma integral múltipla em  $k$  variáveis. Para  $k = 1$ , temos

$$\begin{aligned} \|\gamma_1(t) - \beta_1(t)\| &= \left\| \int_{t_0}^t [F(s_1, \gamma(s_1)) - F(s_1, \beta(s_1))] ds_1 \right\| \\ &\leq \int_{t_0}^t \|F(s_1, \gamma(s_1)) - F(s_1, \beta(s_1))\| ds_1 \leq \int_{t_0}^t C \|\gamma(s_1) - \beta(s_1)\| ds_1 \\ &\leq Cd(\gamma, \beta) \underbrace{\int_{t_0}^t ds_1}_{=:I_1(t)} = Cd(\gamma, \beta) I_1(t). \end{aligned}$$

Para  $k = 2$ , temos

$$\begin{aligned} \|\gamma_2(t) - \beta_2(t)\| &= \left\| \int_{t_0}^t [F(s_1, \gamma_1(s_1)) - F(s_1, \beta_1(s_1))] ds_1 \right\| \\ &\leq C \int_{t_0}^t \|\gamma_1(s_1) - \beta_1(s_1)\| ds_1 \leq C \int_{t_0}^t [Cd(\gamma, \beta)I_1(s_1)] ds_1 \\ &= C^2 d(\gamma, \beta) \underbrace{\int_{t_0}^t \int_{t_0}^{s_1} ds_2 ds_1}_{=: I_2(t)} = C^2 d(\gamma, \beta) I_2(t), \end{aligned}$$

onde na terceira passagem utilizamos o caso  $k = 1$ . Por indução, obtemos que

$$\|\gamma_k(t) - \beta_k(t)\| \leq C^k d(\gamma, \beta) I_k(t),$$

onde  $I_k(t) = \int_{t_0}^t \int_{t_0}^{s_1} \dots \int_{t_0}^{s_{k-1}} ds_k \dots ds_2 ds_1$ . Portanto, basta estimar  $I_k(t)$ .

FATO.  $I_k(t) = \frac{(t-t_0)^k}{k!}$  para todo  $k \geq 1$ .

PROVA (Fato) O caso  $k = 1$  é claro. Por indução, se  $I_k(t) = \frac{(t-t_0)^k}{k!}$  então

$$I_{k+1}(t) = \int_{t_0}^t I_k(s) ds = \int_{t_0}^t \frac{(s-t_0)^k}{k!} ds = \frac{(t-t_0)^{k+1}}{(k+1)!},$$

o que prova o Fato. □

Pelo Fato, segue que  $\|\gamma_k(t) - \beta_k(t)\| \leq C^k d(\gamma, \beta) \frac{(t-t_0)^k}{k!} \leq \frac{(C\varepsilon)^k}{k!} d(\gamma, \beta)$ , o que prova o Lema. □

Desse modo, obtemos a seguinte versão quantitativa do Teorema 14.

**Teorema 35** (Existência - versão quantitativa) Considere a EDO

$$x' = F(t, x),$$

onde  $F : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  é uma função contínua e localmente Lipschitz na segunda coordenada e  $U$  é aberto. Considere  $(t_0, x_0) \in U$ , e tome  $\delta, M > 0$  tais que

$$V := [t_0 - \delta, t_0 + \delta] \times \overline{B(x_0, \delta)} \subset U \text{ e } M := \max_{(t,x) \in V} \|F(t, x)\|.$$

Então para  $\varepsilon = \min\{\delta, \frac{\delta}{M}\}$  existe uma solução  $\gamma : (t_0 - \varepsilon, t_0 + \varepsilon) \rightarrow \mathbb{R}^d$  da EDO com condições iniciais  $x(t_0) = x_0$ .

**Corolário 36** Se, nas condições acima,  $U = \mathbb{R}^{1+d}$  e

$$M = \sup_{(t,x) \in \mathbb{R}^{1+d}} \|F(t, x)\| < \infty,$$

então para todo  $(t_0, x_0) \in \mathbb{R}^{1+d}$  a única solução da EDO com  $\gamma(t_0) = x_0$  é globalmente definida, i.e.  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^d$ .



PROVA | Podemos tomar  $\delta = \infty$ , e portanto o domínio de definição de  $\gamma$  é  $\mathbb{R}$ .  $\square$

## SEÇÃO 4

### Teorema de Peano

O próximo teorema, conhecido como Teorema de Peano, relaxa uma das hipóteses do Teorema 14. Assumindo apenas que  $F$  é contínua, ele garante a existência de solução para a EDO com condições iniciais. Já sabemos, pelo Exemplo 20, que unicidade não é garantida se  $F$  não for localmente Lipschitz na segunda coordenada.

**Teorema 37** (Peano) Considere a EDO

$$x' = F(t, x),$$

onde  $F : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  é uma função contínua. Para todo  $(t_0, x_0) \in U$ , existe  $\gamma : I \rightarrow \mathbb{R}^d$  solução da EDO com condição inicial  $\gamma(t_0) = x_0$ .

PROVA | Dado  $(t_0, x_0) \in U$ , seja  $\delta > 0$  tal que

$$V := [t_0 - \delta, t_0 + \delta] \times \overline{B(x_0, \delta)} \subset U,$$

e seja  $M_0 = \max_V \|F(t, x)\| < \infty$ . Pelo Teorema de Weierstrass (Teorema 157), podemos tomar uma sequência  $(F_n)_{n \geq 1}$  de polinômios  $F_n(t, x)$  tais que suas restrições a  $V$  convergem uniformemente a  $F \upharpoonright_V$ . Em particular, existe  $n_0 > 0$  tal que  $\|F_n \upharpoonright_V - F \upharpoonright_V\| < 1$  para todo  $n \geq n_0$ . Definindo

$$M = \max\{\|F_1 \upharpoonright_V\|, \dots, \|F_{n_0} \upharpoonright_V\|, M_0 + 1\},$$

temos que  $\|F_n(t, x)\|, \|F(t, x)\| \leq M$  para todo  $(t, x) \in V$  e todo  $n \geq 1$ . Defina  $\varepsilon = \min\{\delta, \frac{\delta}{M}\}$ . Cada  $F_n$  é  $C^\infty$  e portanto satisfaz as hipóteses do Teorema 35, logo a EDO com condições iniciais

$$\begin{cases} x' = F_n(t, x) \\ x(t_0) = x_0 \end{cases}$$

possui uma solução  $\gamma_n : (t_0 - \varepsilon, t_0 + \varepsilon) \rightarrow \mathbb{R}^d$ . Vamos obter  $\gamma : I \rightarrow \mathbb{R}^d$  como limite uniforme de uma subsequência de  $(\gamma_n)$ .

FATO. Se  $(\gamma_{n_k})$  converge uniformemente para  $\gamma$ , então  $\gamma$  é solução de  $x' = F(t, x)$  com  $x(t_0) = x_0$ .

PROVA (Fato) Por simplicidade de notação, assuma que  $(\gamma_n)$  converge uniformemente para  $\gamma$ . Então  $F_n(\cdot, \gamma_n(\cdot))$  converge uniformemente para  $F(\cdot, \gamma(\cdot))$ . Por hipótese, temos

$$\gamma_n(t) = x_0 + \int_{t_0}^t F_n(s, \gamma_n(s)) ds.$$

Passando o limite, obtemos que

$$\gamma(t) = x_0 + \int_{t_0}^t F(s, \gamma(s)) ds.$$

Por fim,  $\gamma(t_0) = \lim \gamma_n(t_0) = \lim x_0 = x_0$ .  $\square$

Resta mostrar que  $(\gamma_n)$  possui uma subsequência que converge uniformemente. Para isso, aplicamos o Teorema de Arzelà-Ascoli (Teorema 159), notando que:

◦  $(\gamma_n)$  é equilimitada: para quaisquer  $t$  e  $n$ , temos

$$\|\gamma_n(t)\| \leq \|x_0\| + \int_{t_0}^t \|F_n(s, \gamma_n(s))\| ds \leq \|x_0\| + \varepsilon M.$$

◦  $(\gamma_n)$  é equicontínua: dados  $t_1, t_2 \in (t_0 - \varepsilon, t_0 + \varepsilon)$ , temos

$$\begin{aligned} \|\gamma_n(t_1) - \gamma_n(t_2)\| &= \left\| \int_{t_0}^{t_1} F_n(s, \gamma_n(s)) ds - \int_{t_0}^{t_2} F_n(s, \gamma_n(s)) ds \right\| \\ &\leq |t_1 - t_2| \|F_n \upharpoonright_V\| \leq M|t_1 - t_2|, \end{aligned}$$

logo as  $\gamma_n$  possuem uma constante de Lipschitz comum. Isso garante a equicontinuidade.  $\square$

Temos um resultado interessante que decorre da prova do teorema acima.

**Proposição 38** Sejam  $F, F_n : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  tais que  $F_n$  converge uniformemente para  $F$  em compactos. Assuma que para todo  $(t_0, x_0) \in U$  as EDO's com condições iniciais

$$\begin{cases} x' = F(t, x) \\ x(t_0) = x_0 \end{cases} \quad (4.1)$$

e

$$\begin{cases} x' = F_n(t, x) \\ x(t_0) = x_0 \end{cases} \quad (4.2)$$

possuem uma única solução local. Se  $\gamma : I \rightarrow \mathbb{R}^d$  é solução de (4.1),  $\gamma_n : I \rightarrow \mathbb{R}^d$  é solução de (4.2) com  $x(t_n) = x_n$  e  $(t_n, x_n) \rightarrow (t_0, x_0)$ , então  $\gamma_n$  converge uniformemente para  $\gamma$  em compactos.

PROVA Temos

$$\gamma_n(t) = x_n + \int_{t_n}^t F_n(s, \gamma_n(s)) ds.$$

Se  $\gamma_{n_k}$  converge uniformemente para  $\tilde{\gamma}$  em um intervalo compacto, então pas-

sando o limite na igualdade acima obtemos que

$$\tilde{\gamma}(t) = x_0 + \int_{t_0}^t F(s, \gamma(s)) ds.$$

Pela unicidade de solução de (4.1), concluímos que  $\tilde{\gamma} = \gamma$ . Assim, todo limite uniforme de subsequências de  $\gamma_n$  é igual a  $\gamma$ . Isso prova a proposição.  $\square$

## SEÇÃO 5

### Soluções maximais

Voltamos a assumir que  $F$  é contínua e localmente Lipschitz na segunda coordenada. Nosso objetivo agora é passar de soluções locais para soluções globais, definidas nos maiores intervalos possíveis. Isso induz a definição de *solução maximal*, como veremos abaixo. Nosso interesse é definir, quando possível, as soluções em toda a reta  $\mathbb{R}$  e, quando não possível, entender o porquê.

#### Definição 39

(Solução maximal) Para cada  $(t_0, x_0) \in U$ , existe um intervalo aberto maximal  $(\omega_-, \omega_+) = (\omega_-(t_0, x_0), \omega_+(t_0, x_0))$  e uma solução  $\gamma : (\omega_-, \omega_+) \rightarrow \mathbb{R}^d$  da EDO com condição inicial

$$\begin{cases} x' = F(t, x) \\ x(t_0) = x_0 \end{cases}$$

tal que qualquer outra solução  $\beta : I \rightarrow \mathbb{R}^d$  da EDO com condição inicial acima satisfaz  $I \subset (\omega_-, \omega_+)$ . Chamamos  $\gamma$  de *solução maximal*.

Para definir  $\gamma : (\omega_-, \omega_+) \rightarrow \mathbb{R}^d$  formalmente, tome

$$\mathcal{A} = \{\gamma_J : J \rightarrow \mathbb{R}^d \text{ solução de } x' = F(t, x) \text{ com } \gamma(t_0) = x_0\}.$$

Defina  $(\omega_-, \omega_+) = \bigcup_{\gamma \in \mathcal{A}} J$  e para  $t \in J$  defina  $\gamma(t) = \gamma_J(t)$ . A unicidade local garante que  $\gamma$  está bem-definida e é solução de  $x' = F(t, x)$  com  $x(t_0) = x_0$ .

#### Proposição 40

Nas condições acima, se  $\omega_+ < +\infty$  então para todo compacto  $K \subset U$  existe  $t(K)$  tal que

$$t > t(K) \implies (t, \gamma(t)) \notin K.$$

#### PROVA

A prova é por contradição: se a solução não escapa de  $K$ , então ela se acumula em  $K$  e pode ser estendida além de  $\omega_+$ . Vejamos a justificativa formal. Suponha que exista uma sequência  $t_n \rightarrow \omega_+$  tal que  $(t_n, \gamma(t_n)) \in K$ . Passando a uma subsequência, podemos assumir que  $(t_n, \gamma(t_n)) \rightarrow (\omega_+, x_1) \in K$ . Tome  $\delta > 0$  tal que

$$V = [\omega_+ - 2\delta, \omega_+ + 2\delta] \times \overline{B(x_1, 2\delta)} \subset U$$

e  $M = \max_{(t,x) \in V} \|F(t, x)\|$ . Sejam também

$$W = [\omega_+ - \delta, \omega_+ + \delta] \times \overline{B(x_1, \delta)}$$

e  $\varepsilon = \min\{\delta, \frac{\delta}{M}\}$ . Pelo Teorema 35, para todo  $(t, x) \in W$  existe  $\beta :$

$(t - \varepsilon, t + \varepsilon) \rightarrow \mathbb{R}^d$  solução da EDO com  $\beta(t) = x$ .  
Tome  $n$  grande tal que  $|t_n - \omega_+| < \frac{\varepsilon}{2}$  e  $(t_n, \gamma(t_n)) \in W$ . Então existe  $\beta : (t_n - \varepsilon, t_n + \varepsilon) \rightarrow \mathbb{R}^d$  solução da EDO com  $\beta(t_n) = \gamma(t_n)$ , e portanto a solução maximal se define pelo menos até  $t_n + \frac{\varepsilon}{2} > \omega_+$ , o que contradiz a maximalidade de  $\omega_+$ .  $\square$

Terminamos essa seção acoplando todas as soluções maximais em uma única função.

**Definição 41** (Função  $\varphi$ ) Defina

$$D = \{(t, t_0, x_0) : (t_0, x_0) \in U \text{ e } t \in (\omega_-(t_0, x_0), \omega_+(t_0, x_0))\} \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d$$

e  $\varphi : D \rightarrow \mathbb{R}^d$  por  $\varphi(t, t_0, x_0) = \gamma_{(t_0, x_0)}(t)$ , onde  $\gamma_{(t_0, x_0)}$  é a solução maximal de  $x' = F(t, x)$  com  $x(t_0) = x_0$ .

Entender o domínio  $D$  e as propriedades de  $\varphi$  corresponde a entender a dependência de soluções da EDO com respeito às condições iniciais. Abaixo listamos duas propriedades básicas de  $\varphi$ , e nas próximas seções faremos um estudo mais detalhado:

- $\varphi(t_0, t_0, x_0) = x_0$ .
- $\frac{\partial \varphi}{\partial t}(t, t_0, x_0) = F(t, \varphi(t, t_0, x_0))$ .

## SEÇÃO 6

### Continuidade de $\varphi$

Nessa seção, mostraremos que  $\varphi : D \rightarrow \mathbb{R}^d$  é contínua. Para isso, precisamos de um melhoramento da Proposição 38.

**Proposição 42** Sejam  $F, F_n : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  contínuas e localmente Lipschitz na segunda coordenada tais que  $F_n$  converge uniformemente para  $F$  em compactos. Sejam  $(t_0, x_0), (t_n, x_n) \in U$  tais que  $(t_n, x_n) \rightarrow (t_0, x_0)$ , e sejam  $\gamma : (\omega_-, \omega_+) \rightarrow \mathbb{R}^d$  e  $\gamma_n : (\omega_-^{(n)}, \omega_+^{(n)}) \rightarrow \mathbb{R}^d$  as soluções maximais de

$$\begin{cases} x' = F(t, x) \\ x(t_0) = x_0 \end{cases} \quad \text{e} \quad \begin{cases} x' = F_n(t, x) \\ x(t_n) = x_n \end{cases}$$

respectivamente. Se  $[a, b] \subset (\omega_-, \omega_+)$  então  $[a, b] \subset (\omega_-^{(n)}, \omega_+^{(n)})$  para todo  $n$  suficientemente grande e as restrições de  $\gamma_n$  a  $[a, b]$  convergem uniformemente para a restrição de  $\gamma$  a  $[a, b]$ .

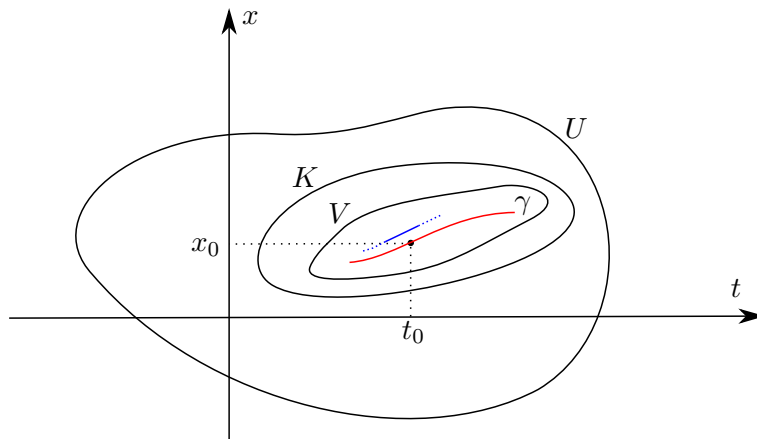
Em outras palavras: se  $\gamma$  está bem-definida em um intervalo compacto, então o mesmo ocorre para todas as condições iniciais suficientemente próximas, e as soluções convergem uniformemente.

**PROVA** A diferença para a Proposição 38 é que não sabemos, a priori, que  $\gamma_n$  está definida em  $[a, b]$ . Faremos isso aumentando localmente, a partir de  $t_0$ , as soluções pelo menos uma quantidade fixa. Para isso, tome  $V$  aberto e  $K$

compacto tais que

$$\text{graf}(\gamma \upharpoonright_{[a,b]}) \subset V \subset K,$$

e tome  $\delta > 0$  tal que  $[t - \delta, t + \delta] \times \overline{B(x, \delta)} \subset K$  para todo  $(t, x) \in V$ .



Se  $M = \max\{\|F(t, x)\|, \|F_n(t, x)\| : (t, x) \in K \text{ e } n \geq 1\}$  e  $2\varepsilon = \min\{\delta, \frac{\delta}{M}\}$ , então a EDO's  $x' = F(t, x)$  e  $x' = F_n(t, x)$  com condição inicial  $(t, x) \in V$  estão bem-definidas em  $(t - 2\varepsilon, t + 2\varepsilon)$ . Logo, se  $n$  é grande então  $\gamma_n$  está definida em  $(t_n - 2\varepsilon, t_n + 2\varepsilon) \supset [t_0 - \varepsilon, t_0 + \varepsilon]$ . Pela Proposição 38, as restrições de  $\gamma_n$  a  $[t_0 - \varepsilon, t_0 + \varepsilon]$  convergem para a restrição de  $\gamma$  a  $[t_0 - \varepsilon, t_0 + \varepsilon]$ . Em particular,

$$(t_0 \pm \varepsilon, \gamma_n(t_0 \pm \varepsilon)) \rightarrow (t_0 \pm \varepsilon, \gamma(t_0 \pm \varepsilon)).$$

Repita o mesmo argumento para essa nova sequência convergente de pontos, concluindo que para  $n$  grande  $\gamma_n$  está definida em  $[t_0 - 2\varepsilon, t_0 + 2\varepsilon]$ . Novamente pela Proposição 38, as restrições de  $\gamma_n$  a  $[t_0 - 2\varepsilon, t_0 + 2\varepsilon]$  convergem para a restrição de  $\gamma$  a  $[t_0 - 2\varepsilon, t_0 + 2\varepsilon]$ . Continue repetindo esse argumento. Em cada passo, o raio do domínio de definição de  $\gamma_n$  aumenta  $\varepsilon$  e portanto, após uma quantidade finita de passos, concluímos que  $[a, b] \subset (\omega_-^{(n)}, \omega_+^{(n)})$  para  $n$  grande. A convergência uniforme em  $[a, b]$  decorre novamente da Proposição 38.  $\square$

Obtemos, assim, o seguinte resultado.

**Teorema 43**

Dada  $F : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  contínua e localmente Lipschitz na segunda coordenada, seja  $\varphi : D \subset \mathbb{R}^{2+d} \rightarrow \mathbb{R}^d$  a função construída acima. Então:

- (1)  $D$  é aberto.
- (2)  $\omega_+ : U \rightarrow \mathbb{R} \cup \{+\infty\}$  é semicontínua inferiormente, e  $\omega_- : U \rightarrow \mathbb{R} \cup \{-\infty\}$  é semicontínua superiormente.
- (3)  $\varphi$  é contínua.

**PROVA**

(1) Dado  $(t, t_0, x_0) \in D$ , seja  $\omega_{\pm} = \omega_{\pm}(t_0, x_0)$ . Tome  $a, b$  tais que  $\omega_- < a < t < b < \omega_+$ . Pela proposição anterior, existe  $V$  vizinhança de  $(t_0, x_0)$  tal que  $[a, b] \subset (\omega_-(t_1, x_1), \omega_+(t_1, x_1))$  para todo  $(t_1, x_1) \in V$ . Assim,  $(a, b) \times V \subset D$ .

(2) Diretamente da proposição anterior, temos:

$$\liminf_{(t_n, x_n) \rightarrow (t_0, x_0)} \omega_+(t_n, x_n) \geq \omega_+(t_0, x_0)$$

$$\limsup_{(t_n, x_n) \rightarrow (t_0, x_0)} \omega_-(t_n, x_n) \leq \omega_-(t_0, x_0).$$

(3) Fixe  $(t, t_0, x_0) \in D$ . Dado  $(r, t_1, x_1) \in D$ , temos

$$\|\varphi(r, t_1, x_1) - \varphi(t, t_0, x_0)\| \leq \|\varphi(r, t_1, x_1) - \varphi(r, t_0, x_0)\| + \|\varphi(r, t_0, x_0) - \varphi(t, t_0, x_0)\|.$$

O primeiro termo no lado direito converge para zero quando  $(t_1, x_1) \rightarrow (t_0, x_0)$ , pela proposição anterior. Se  $\gamma$  é a solução da EDO com  $\gamma(t_0) = x_0$ , então o segundo termo é  $\|\int_t^r F(s, \gamma(s)) ds\|$ , que converge para zero quando  $r \rightarrow t$ .  $\square$

## SEÇÃO 7

### Desigualdade de Gronwall e estimativa Lipschitz de $\varphi(t, t_0, \cdot)$

Nessa seção, provamos algumas estimativas, chamadas desigualdades de Gronwall, que são clássicas na teoria de EDO. Utilizaremos uma das desigualdade de Gronwall para obter uma estimativa Lipschitz de  $\varphi(t, t_0, x_0)$  com respeito à entrada  $x_0$ .

Existem diversas versões da desigualdade de Gronwall. Em todas elas, temos uma estimativa inicial de uma função  $u(t)$  em termos de um produto  $u(t)\beta(t)$ . A desigualdade de Gronwall fornece uma estimativa de  $u(t)$  em termos de  $\beta$  apenas. No que segue, apresentaremos duas versões dessa desigualdade, uma diferencial e outra integral.

**Proposição 44** (Desigualdade de Gronwall - versão diferencial) Sejam  $u, \beta : [a, b] \rightarrow \mathbb{R}$  funções tais que  $u$  é diferenciável.

(1) Se  $u'(t) \geq u(t)\beta(t)$  para todo  $t \in [a, b]$ , então

$$u(t) \geq u(a) \exp \left[ \int_a^t \beta(s) ds \right].$$

(2) Se  $u'(t) \leq u(t)\beta(t)$  para todo  $t \in [a, b]$ , então

$$u(t) \leq u(a) \exp \left[ \int_a^t \beta(s) ds \right].$$

Em outras palavras, podemos comparar  $u$  com a solução da EDO  $x' = \beta x$ .

**PROVA** (1) Defina  $w(t) = \exp \left[ \int_a^t \beta(s) ds \right]$ , que é positiva com  $w(a) = 1$  e diferenciável com

$$w'(t) = \exp \left[ \int_a^t \beta(s) ds \right] \beta(t) = w(t)\beta(t).$$

Defina  $f = u/w$ . Temos  $f(a) = u(a)$ . Afirmamos que  $f$  é crescente. De fato,

$f' = \frac{u'w - uw'}{w^2} \geq 0$  pois como  $w > 0$  temos  $u'w \geq (u\beta)w = u(\beta w) = uw'$ . Assim, segue que

$$\frac{u(t)}{w(t)} = f(t) \geq f(a) = u(a) \Rightarrow u(t) \geq u(0)w(t)$$

para todo  $t \in [a, b]$ .

(2) A função  $v = -u$  satisfaz  $v' \geq vw$ . Pela parte (1), temos

$$v(t) \geq v(0) \exp \left[ \int_a^t \beta(s) ds \right] \implies u(t) \leq u(0) \exp \left[ \int_a^t \beta(s) ds \right].$$

□

**Proposição 45**

(Desigualdade de Gronwall - versão integral) Sejam  $\alpha \in \mathbb{R}$  e  $u, \beta : [a, b] \rightarrow \mathbb{R}$  funções tais que  $\beta \geq 0$  e

$$u(t) \leq \alpha + \int_a^t u(s)\beta(s)ds.$$

Então

$$u(t) \leq \alpha \exp \left[ \int_a^t \beta(s) ds \right].$$

## PROVA

Introduza a função  $w$  pelo lado direito da desigualdade:

$$w(t) = \alpha + \int_a^t u(s)\beta(s)ds.$$

Então  $u \leq w$  e  $w' = u\beta$ . Como  $\beta \geq 0$ , segue que  $w' = u\beta \leq w\beta$ . Pela versão diferencial da desigualdade de Gronwall, temos

$$w(t) \leq w(0) \exp \left[ \int_a^t \beta(s) ds \right] = \alpha \exp \left[ \int_a^t \beta(s) ds \right]$$

e portanto

$$u(t) \leq w(t) \leq \alpha \exp \left[ \int_a^t \beta(s) ds \right].$$

□

Agora procedemos para uma aplicação da última proposição.

**Proposição 46**

Seja  $F : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  contínua e localmente Lipschitz na segunda coordenada. Então  $\varphi$  é localmente Lipschitz na coordenada  $x_0$ . Mais especificamente, para todo  $(s, s_0, y_0) \in D$  existe uma vizinhança  $V \subset D$  de  $(s, s_0, y_0)$  e  $C > 0$  tais que

$$\|\varphi(t, t_0, x_0) - \varphi(t, t_0, x_1)\| \leq \exp [C|t - t_0|] \|x_0 - x_1\|$$

para todos  $(t, t_0, x_0), (t, t_0, x_1) \in V$ .

## PROVA

Seja  $V$  vizinhança de  $(s, s_0, y_0)$  tal que  $F \upharpoonright_V$  é localmente Lipschitz na segunda coordenada, e seja  $C > 0$  a constante de Lipschitz. Tome  $(t, t_0, x_0), (t, t_0, x_1) \in V$ , e assumamos que  $t \geq t_0$  (o caso  $t \leq t_0$  é tratado no Exercício 17). Pela equação

integral, obtemos que

$$\begin{cases} \varphi(t, t_0, x_0) = x_0 + \int_{t_0}^t F(s, \varphi(s, t_0, x_0)) ds \\ \varphi(t, t_0, x_1) = x_1 + \int_{t_0}^t F(s, \varphi(s, t_0, x_1)) ds \end{cases}$$

Subtraindo e aplicando a desigualdade triangular, temos

$$\begin{aligned} & \|\varphi(t, t_0, x_0) - \varphi(t, t_0, x_1)\| \\ & \leq \|x_0 - x_1\| + \int_{t_0}^t \|F(s, \varphi(s, t_0, x_0)) - F(s, \varphi(s, t_0, x_1))\| ds \\ & \leq \|x_0 - x_1\| + \int_{t_0}^t C \|\varphi(s, t_0, x_0) - \varphi(s, t_0, x_1)\| ds. \end{aligned}$$

Chamando  $\alpha = \|x_0 - x_1\|$ ,  $u(t) = \|\varphi(t, t_0, x_0) - \varphi(t, t_0, x_1)\|$  e  $\beta(t) \equiv C$ , a última desigualdade diz que  $u(t) \leq \alpha + \int_{t_0}^t u(s)\beta(s)ds$ . Pela versão integral da desigualdade de Gronwall (Proposição 45), segue que

$$u(t) \leq \|x_0 - x_1\| \exp \left[ \int_{t_0}^t C ds \right] = \exp [C|t - t_0|] \|x_0 - x_1\|.$$

Isso conclui a prova.  $\square$

## SEÇÃO 8

### Diferenciabilidade de $\varphi$

Agora vamos estudar a dependência diferenciável de  $\varphi$  com respeito à variável  $x_0$ . Naturalmente, devemos fazer hipóteses mais fortes sobre a função  $F$ . No que segue,  $\mathcal{L}(\mathbb{R}^d)$  denota o espaço vetorial das transformações lineares em  $\mathbb{R}^d$ . No próximo enunciado, precisamos considerar a versão matricial de EDO.

**Definição 47** (EDO matricial) Dada a EDO  $x' = F(t, x)$ , a *EDO matricial* associada é  $X' = F(t, X)$ , onde  $X$  é uma matriz  $d \times d$ . Dizemos que  $\Gamma : (t_0 - \varepsilon, t_0 + \varepsilon) \rightarrow \mathcal{L}(\mathbb{R}^d)$  é *solução* da EDO matricial com condições iniciais

$$\begin{cases} X' = F(t, X) \\ X(t_0) = X_0 \end{cases}$$

se, para todo  $v \in \mathbb{R}^d$ , a curva  $\gamma : (t_0 - \varepsilon, t_0 + \varepsilon) \rightarrow \mathbb{R}^d$  definida por  $\gamma(t) = \Gamma(t)v$  é a solução local da EDO com condições iniciais

$$\begin{cases} x' = F(t, x) \\ x(t_0) = X_0 v. \end{cases}$$

Em geral, podemos tomar  $X_0 = \text{Id}$ .



**Teorema 48**

(Diferenciabilidade de  $\varphi$ ) Seja  $F : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$  contínua e diferenciável na segunda coordenada com  $\frac{\partial F}{\partial x} : U \rightarrow \mathcal{L}(\mathbb{R}^d)$  contínua. Então a derivada parcial  $\frac{\partial \varphi}{\partial x} : D \rightarrow \mathcal{L}(\mathbb{R}^d)$  existe, é contínua, e é solução da EDO matricial

$$\begin{cases} X' = \frac{\partial F}{\partial x}(t, \varphi(t, t_0, x_0))X \\ X(t_0) = \text{Id.} \end{cases} \quad (8.1)$$

**PROVA**

Sem justificativas formais, podemos derivar  $\frac{\partial \varphi}{\partial t}(t, t_0, x_0) = F(t, \varphi(t, t_0, x_0))$  e obter que, caso exista,  $X = \frac{\partial \varphi}{\partial x}(t, t_0, x_0)$  satisfaz a EDO matricial abaixo:

$$\begin{aligned} \frac{\partial}{\partial x} \left[ \frac{\partial \varphi}{\partial t}(t, t_0, x_0) \right] &= \frac{\partial}{\partial x} [F(t, \varphi(t, t_0, x_0))] \\ \frac{\partial}{\partial t} \left[ \frac{\partial \varphi}{\partial x}(t, t_0, x_0) \right] &= \frac{\partial F}{\partial x}(t, \varphi(t, t_0, x_0)) \cdot \frac{\partial \varphi}{\partial x}(t, t_0, x_0) \\ X' &= \frac{\partial F}{\partial x}(t, \varphi(t, t_0, x_0)) \cdot X. \end{aligned}$$

A ideia é justamente ver  $\frac{\partial \varphi}{\partial x}$  como solução dessa EDO.

Como nem sabemos, a priori, se  $\frac{\partial \varphi}{\partial x}$  existe, vamos considerar EDO's que aproximam (8.1) e usar a dependência contínua para concluir que as soluções convergem uniformemente. O limite uniforme é, por definição, a derivada parcial  $\frac{\partial \varphi}{\partial x}$ . O fato abaixo permite identificar as funções que definem as EDO's que aproximam (8.1).

**FATO.** Seja  $f : (a, b) \times A \rightarrow \mathbb{R}^d$  uma função contínua, onde  $A \subset \mathbb{R}^d$  é convexo, tal que  $\frac{\partial f}{\partial x} : (a, b) \times A \rightarrow \mathcal{L}(\mathbb{R}^d)$  existe e é contínua. Então existe uma função contínua  $g : (a, b) \times A \times A \rightarrow \mathcal{L}(\mathbb{R}^d)$  tal que

$$f(t, x_2) - f(t, x_1) = g(t, x_1, x_2) \cdot (x_2 - x_1), \quad \forall t \in (a, b), \forall x_1, x_2 \in A.$$

Em particular,  $\frac{\partial f}{\partial x}(t, x, x) = g(t, x, x)$  para todo  $(t, x) \in (a, b) \times A$ .

**PROVA**

(Fato) Pela condição requerida, devemos ter

$$\begin{aligned} g(t, x_1, x_2) \cdot (x_2 - x_1) &= f(t, x_1) - f(t, x_2) \\ &= \int_0^1 \frac{\partial f}{\partial x}(t, x_1 + s(x_2 - x_1)) \cdot (x_2 - x_1) ds \end{aligned}$$

e portanto  $g$  se define unicamente por

$$g(t, x_1, x_2) := \int_0^1 \frac{\partial f}{\partial x}(t, x_1 + s(x_2 - x_1)) ds.$$

A hipótese sobre  $f$  garante que  $g$  é contínua. Por fim,

$$g(t, x, x) = \int_0^1 \frac{\partial f}{\partial x}(t, x) ds = \frac{\partial f}{\partial x}(t, x).$$

□

Agora procedemos para finalizar a prova do teorema. Fixe  $v \in \mathbb{R}^d$ . Para cada

$h \in \mathbb{R}$ , vamos identificar a EDO que o quociente de Newton

$$z_h = \frac{\varphi(t, t_0, x_0 + hv) - \varphi(t, t_0, x_0)}{h}$$

satisfaz. Note que  $z_h(t_0) = v$  para todo  $h$ . Temos

$$\begin{aligned} z'_h(t) &= \frac{1}{h} [F(t, \varphi(t, t_0, x_0 + hv)) - F(t, \varphi(t, t_0, x_0))] \\ &= \frac{1}{h} g(t, \varphi(t, t_0, x_0), \varphi(t, t_0, x_0 + hv)) \cdot (\varphi(t, t_0, x_0 + hv) - \varphi(t, t_0, x_0)) \\ &= G_h(t, t_0, x_0) \cdot z_h(t), \end{aligned}$$

onde  $G_h(t, t_0, x_0) := g(t, \varphi(t, t_0, x_0), \varphi(t, t_0, x_0 + hv))$  e na segunda passagem utilizamos o Fato. Também pelo Fato, temos  $\lim_{h \rightarrow 0} G_h = \frac{\partial F}{\partial x}$ . Pelo teorema da dependência contínua (Teorema 43), os quocientes de Newton  $z_h$  convergem uniformemente para a solução  $z$  da EDO  $x' = \frac{\partial F}{\partial x}(t, \varphi(t, t_0, x_0)) \cdot x$  com  $x(t_0) = v$ . Assim,  $\frac{\partial \varphi}{\partial x} \cdot v$  existe e é igual a  $z$ . Em particular,  $\frac{\partial \varphi}{\partial x} \cdot v$  é contínua e solução de  $x' = \frac{\partial F}{\partial x}(t, \varphi(t, t_0, x_0)) \cdot v$  com  $x(t_0) = v$ . Como  $v \in \mathbb{R}^d$  é arbitrário, concluímos a prova do teorema.  $\square$

## SEÇÃO 9

### Teorema de comparação de Sturm

Assim como a desigualdade de Gronwal, o teorema de comparação de Sturm também possui diversas versões, cada uma enunciada em um contexto específico de interesse. Abaixo, enunciaremos de um modo aplicável em geometria riemanniana (de fato, a versão abaixo pode ser vista como um “baby case” do teorema de comparação de Rauch).

#### Definição 49

(Equação de Jacobi) Dada  $f : [a, b] \rightarrow \mathbb{R}$  contínua, a *equação de Jacobi* de  $f$  é a EDO de ordem dois

$$u'' + fu = 0.$$

É fácil ver que a EDO acima tem única solução local. De fato, ela é uma equação linear de ordem dois definida por  $F(t, x, x') = -f(t)x$ . Introduzindo  $y = x'$ , na variável  $X = (x, y)$  temos a EDO de ordem 1 equivalente  $X' = G(t, X)$ , onde  $G(t, X) = G(t, (x, y)) = (y, -f(t)x)$  é contínua e localmente Lipschitz na segunda variável  $X$ . Na versão abaixo, o teorema de comparação de Sturm compara soluções de equações de Jacobi com mesmas condições iniciais. Relembre que, por terem ordem dois, as condições iniciais de equações de Jacobi são  $(u(a), u'(a)) = (u_0, u'_0)$ .

#### Teorema 50

(Comparação de Sturm) Sejam  $f, g : [a, b] \rightarrow \mathbb{R}$  funções com  $f \leq g$ , e sejam  $u, v : [a, b] \rightarrow \mathbb{R}$  soluções das equações de Jacobi

$$\begin{cases} u'' + fu = 0 \\ v'' + gv = 0 \end{cases}$$

tais que  $(u(a), u'(a)) = (v(a), v'(a))$  e  $v > 0$ . Então:

- (1)  $u > 0$ .
- (2)  $u \geq v$ .

A parte (1) diz que  $u$  não se anula antes de  $v$ .

PROVA | A prova do teorema seguirá da seguinte igualdade. Seja  $w = u'v - uv'$ . Temos  $w(a) = u'(a)v(a) - u(a)v'(a) = u'(a)u(a) - u(a)u'(a) = 0$  e  $w' = u''v - uv'' = (-fu)v - u(-gv) = (g-f)uv$ , e daí pelo TFC obtemos

$$w(t) = \int_a^t (g-f)uv. \quad (9.1)$$

(1) Por contradição, assuma que  $u$  se anule em  $[a, b]$ , e seja  $t_0 = \min\{t \in [a, b] : u(t) = 0\}$ . Pela minimalidade, temos  $u'(t_0) \leq 0$ . Mais ainda:  $u'(t_0) < 0$ , pois se  $u'(t_0) = 0$  então  $u$  é solução com  $u(t_0) = u'(t_0) = 0$  e daí, pela unicidade de solução,  $u \equiv 0$  (o que contradiz  $u(a) > 0$ ). Por um lado,  $w(t_0) = u'(t_0)v(t_0) - u(t_0)v'(t_0) = u'(t_0)v(t_0) < 0$ . Por outro lado,  $\int_a^{t_0} (g-f)uv \geq 0$  pois o integrando é não-negativo em  $[a, t_0]$ . Isso contradiz a igualdade (9.1) para  $t = t_0$ .

(2) Pela parte (1) e a igualdade (9.1), obtemos que  $w \geq 0$ , que se reescreve como

$$u'v \geq uv' \implies \frac{u'}{u} \geq \frac{v'}{v} \implies (\log u)' \geq (\log v)'$$

As divisões podem ser feitas e preservam as desigualdades porque  $u, v > 0$ . Integrando e usando que  $u(a) = v(a)$ , segue que  $\log u \geq \log v$  e assim  $u \geq v$ .  $\square$

## SUBSEÇÃO 9.1

### Estimativas para a equação de Jacobi

Pelo teorema de comparação de Sturm, podemos obter estimativas das soluções de uma equação de Jacobi comparando-a com a solução com mesmas condições iniciais de uma equação de Jacobi com função constante, a qual podemos descrever explicitamente as soluções. No que segue, obteremos tal representação.

No contexto do Exercício 5, parte (3), se  $f : [a, b] \rightarrow \mathbb{R}$  é contínua e para todo  $x_0 \in \mathbb{R}$  a equação de Jacobi  $u'' + fu = 0$  com condições iniciais  $(a, x_0)$  possui solução em  $[a, b]$ , então o mapa

$$\begin{aligned} \Phi_t : \mathcal{S} &\rightarrow \mathbb{R}^2 \\ u &\mapsto (u(t), u'(t)) \end{aligned}$$

é um isomorfismo linear, onde  $\mathcal{S} = \mathcal{S}_{[a,b]}$ . Logo, se  $u_1, u_2 \in \mathcal{S}$  são as soluções com  $\Phi(u_1) = (1, 0)$  e  $\Phi(u_2) = (0, 1)$ , então a única  $u \in \mathcal{S}$  com  $u(a) = A$  e  $u'(a) = B$  é

$$u = Au_1 + Bu_2.$$

Vejam quem são  $u_1, u_2$  quando  $f$  é constante, digamos  $f \equiv K$ . Por simplicidade, tomamos  $a = 0$ . Temos três casos.

CASO 1.  $K = 0$ .

A equação de Jacobi é  $u'' = 0$ , cuja solução com  $u(0) = A$  e  $u'(0) = B$  é

$$u(t) = Bt + A.$$

CASO 2.  $K = -a^2$ .

A equação de Jacobi é  $u'' = a^2u$ . As funções  $e^{at}$  e  $e^{-at}$  são soluções, mas não correspondem à base canônica de  $\mathbb{R}^2$  via  $\Phi$ . Tais soluções são  $u_1(t) = \cosh(at)$  e  $u_2(t) = \sinh(at)$ . Portanto, a única solução de  $u'' = a^2u$  com  $u(0) = A$  e  $u'(0) = B$  é

$$u(t) = A \cosh(at) + B \sinh(at).$$

CASO 3.  $K = a^2$ .

A equação de Jacobi é  $u'' = -a^2u$ . As funções  $\cos(at)$  e  $\sin(at)$  são soluções com  $\Phi(u_1) = (1, 0)$  e  $\Phi(u_2) = (0, a)$ . Normalizando, definimos  $u_1(t) = \cos(at)$  e  $u_2(t) = \frac{1}{a} \sin(at)$ , de modo que a única solução de  $u'' = -a^2u$  com  $u(0) = A$  e  $u'(0) = B$  é

$$u(t) = A \cos(at) + \frac{B}{a} \sin(at).$$

## SUBSEÇÃO 9.2

### Equação de Riccati

Outra maneira de estimar o crescimento da solução de uma equação de Jacobi é considerar sua derivada logarítmica. Isso só pode ser feito para soluções com sinal. Assuma que  $u'' + fu = 0$  e que  $u$  é positiva em  $[a, b]$ . A *derivada logarítmica* de  $u$  é a função

$$v = \frac{u'}{u} = (\log u)'$$

Qualquer estimativa em  $v$  se traduz em uma estimativa exponencial em  $u$ , pois pelo TFC temos

$$u(t) - u(a) = \exp[\log u(t) - \log u(a)] = \exp\left[\int_a^t v(s) ds\right].$$

A função  $v$  também satisfaz uma EDO. De fato,

$$u' = uv \implies u'' = u'v + uv' \implies -fu = u'v + uv' \implies v' + v^2 + f = 0.$$

**Definição 51** (Equação de Riccati) Dada  $f : [a, b] \rightarrow \mathbb{R}$ , a *equação de Riccati* de  $f$  é a EDO de ordem 1

$$v' + v^2 + f = 0.$$

Como vimos acima, toda solução positiva da equação de Jacobi define uma solução da equação de Riccati. Logo, reduzimos o estudo de soluções positivas de uma EDO de ordem dois para o estudo de soluções de uma EDO de ordem 1. O preço a pagar é que a nova EDO não é mais linear. A relação entre as

equações de Jacobi e Riccati é muito explorada no estudo da dinâmica de fluxos geodésicos.

## SEÇÃO 10

**Exercícios**

**Exercício 1** | Prove que o espaço métrico  $(C^0[a, b], d_\infty)$  definido no Exemplo 25 é completo.

**Exercício 2** | Mostre que toda contração é contínua.

**Exercício 3** | Usando a Proposição 13, enuncie e prove o Teorema 14 para EDO's de ordem superior.

**Exercício 4** | Considere a EDO com condições iniciais

$$x' = 3x + 5, x(0) = 1.$$

Pelo teorema de existência e unicidade, a solução da EDO é igual ao limite de funções  $\{\gamma_n\}_{n \geq 1}$  calculadas iterativamente.

(a) Tomando  $\gamma_0(t) \equiv 1$ , calcule  $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ .

(b) Mostre  $\gamma_n$  é da forma

$$\gamma_n(t) = 1 + \sum_{k=1}^n a_k t^k,$$

onde  $\{a_k\}_{k \geq 1}$  satisfaz  $a_k = \frac{3a_{k-1}}{k}, \forall k \geq 2$ .

(c) Ache a solução  $\lim_{n \rightarrow \infty} \gamma_n(t)$ .

**Exercício 5** | (Equações lineares) Dado um intervalo aberto  $I$ , seja  $F: I \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  dada por  $F(t, x) = A(t)x + B(t)$ , onde  $A, B: I \rightarrow \text{GL}(d)$  são contínuas.

(1) Mostre que para todo  $(t_0, x_0) \in I \times \mathbb{R}^d$  a EDO com condições iniciais

$$\begin{cases} x' = F(t, x) \\ x(t_0) = x_0 \end{cases}$$

possui única solução local.

(2) Assuma que  $B \equiv 0$ . Dado  $[a, b] \subset \mathbb{R}$ , assumo também que a EDO com condições iniciais  $(a, x_0)$  tem solução  $\gamma_{x_0}: [a, b] \rightarrow \mathbb{R}^d$  para todo  $x_0 \in \mathbb{R}^d$ . Mostre que o espaço de soluções

$$\mathcal{S}_{[a,b]} = \{\gamma_{x_0} : x_0 \in \mathbb{R}^d\}$$

é um espaço vetorial. Mostre também que para todo  $t \in [a, b]$  o mapa  $\gamma \in \mathcal{S}_{[a,b]} \mapsto \gamma(t) \in \mathbb{R}^d$  é um isomorfismo linear.

(3) Enuncie e prove a versão de (2) para EDO's de ordem superior.

**Exercício 6** (Equação linear com coeficientes constantes) Seja  $A \in \text{GL}(d)$ . Dado  $x_0 \in \mathbb{R}^d$ , mostre que a única solução de

$$\begin{cases} x' = Ax \\ x(0) = x_0 \end{cases}$$

é  $\gamma(t) = e^{tA}x_0$ , onde

$$e^B = \sum_{n \geq 0} \frac{B^n}{n!}$$

para toda  $B \in \text{GL}(d)$ .

**Exercício 7** (Equação linear matricial com coeficientes constantes) Seja  $A \in \text{GL}(d)$ . Mostre que a única solução  $X : \mathbb{R} \rightarrow \text{GL}(d)$  de

$$\begin{cases} X' = AX \\ X(0) = \text{Id} \end{cases}$$

é  $X(t) = e^{tA}$ .

**Exercício 8** Seja  $F : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  uma função Lipschitz. Mostre que para todo  $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^d$  a EDO com condições iniciais

$$\begin{cases} x' = F(t, x) \\ x(t_0) = x_0 \end{cases}$$

possui uma única solução  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^d$ .

**Exercício 9** Sejam  $F, G : \mathbb{R} \rightarrow \mathbb{R}$  funções contínuas tais que  $F$  é Lipschitz. Mostre que para todo  $(t_0, x_0) \in \mathbb{R}^2$  o sistema

$$\begin{cases} x' = F(x), x(t_0) = x_0 \\ y' = G(x)y, y(t_0) = y_0 \end{cases}$$

possui única solução local. O resultado permanece válido se  $F$  não for Lipschitz?

**Exercício 10** (Teorema de Kneser) Na notação do Teorema de Peano (Teorema 37), assumamos que  $d = 1$ . Dado  $c \in [t_0, t_0 + \varepsilon)$ , mostre que o conjunto

$$S_c = \{\gamma(c) : \gamma : I \rightarrow \mathbb{R} \text{ é solução de } x' = F(t, x), x(t_0) = x_0, \text{ com } c \in I\}$$

é um intervalo fechado.

**Exercício 11** Sejam  $F, G : \mathbb{R}^2 \rightarrow \mathbb{R}$  contínuas tais que as EDO's com condições iniciais possuem única solução local. Suponha que  $F(t, x) > G(t, x)$  para todo  $(t, x) \in \mathbb{R}^2$ . Dado  $(t_0, x_0)$ , sejam  $\gamma : I \rightarrow \mathbb{R}$  e  $\beta : J \rightarrow \mathbb{R}$  soluções de

$$\begin{cases} x' = F(t, x) \\ x(t_0) = x_0 \end{cases} \quad \text{e} \quad \begin{cases} x' = G(t, x) \\ x(t_0) = x_0 \end{cases}$$

respectivamente. Mostre que  $\gamma(t) \geq \beta(t)$  para todo  $t > t_0$  em  $I \cap J$ . Mostre que a mesma conclusão vale se supusermos apenas que  $F(t, x) \geq G(t, x)$  para

todo  $(t, x) \in \mathbb{R}^2$ .

**Exercício 12** (Fluxo de campos de vetores) Sejam  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$  um campo de vetores contínuo, localmente Lipschitz na segunda coordenada e tal que  $\sup_{x \in \mathbb{R}^d} \|F(x)\| < \infty$ .

Mostre os seguintes itens:

(1) Para todo  $x_0 \in \mathbb{R}^d$ , o problema

$$\begin{cases} x' = F(x) \\ x(0) = x_0 \end{cases}$$

possui uma única solução  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^d$ .

(2) Para cada  $t \in \mathbb{R}$ , seja  $\varphi_t : \mathbb{R}^d \rightarrow \mathbb{R}^d$  a função  $\varphi_t(x_0) = \gamma_{x_0}(t)$ , onde  $\gamma_{x_0}$  é a solução de (1). Então  $\varphi_t$  é um homeomorfismo para todo  $t \in \mathbb{R}$  e  $\varphi_{t+s} = \varphi_t \circ \varphi_s$  para todos  $t, s \in \mathbb{R}$ . A família  $\varphi = \{\varphi_t\}_{t \in \mathbb{R}}$  é chamada de *fluxo* associado ao campo de vetores.

**Exercício 13** (Campo gradiente) Dada  $F : U \subset \mathbb{R}^d \rightarrow \mathbb{R}$ , o *campo gradiente* de  $F$  é definido por  $X = \nabla F : U \rightarrow \mathbb{R}^d$ . Seja  $\gamma$  solução de

$$\begin{cases} x' = F(x) \\ x(0) = x_0. \end{cases}$$

Mostre que ou  $\gamma(t) = x_0$  para todo  $t \in \mathbb{R}$  ou a composição  $F \circ \gamma$  é estritamente crescente. Em particular, as únicas soluções periódicas são constantes.

**Exercício 14** Sejam  $X = (X_1, \dots, X_d) : \mathbb{R}^d \rightarrow \mathbb{R}^d$  um campo vetorial de classe  $C^1$  e  $V : \mathbb{R}^d \rightarrow \mathbb{R}$  diferenciável tais que  $\sum_{i=1}^d \frac{\partial V}{\partial x_i} X_i \leq 0$  e  $V(x) \geq \|x\|^2$  para todo  $x \in \mathbb{R}^d$ . Mostre que para todo  $x_0 \in \mathbb{R}^d$  a solução de

$$\begin{cases} x' = X(x) \\ x(0) = x_0 \end{cases}$$

está definida em  $(0, +\infty)$ .

**Exercício 15** Prove que a função  $\varphi$  definida na Seção 5 é localmente Lipschitz na primeira coordenada.

**Exercício 16** Assumindo que  $u > 0$ , prove a desigualdade de Gronwall como abaixo:

(1) Introduza a *derivada logarítmica*  $f = u'/u = (\log u)'$ . Se  $u' \geq u\beta$ , conclua que  $f \geq \beta$ .

(2) Aplique o teorema fundamental do cálculo e conclua que

$$\log \left[ \frac{u(t)}{u(a)} \right] \geq \int_a^t \beta(s) ds.$$

(3) Prove a desigualdade de Gronwall.

**Exercício 17** (Reversão temporal) Dada  $F : U \subset \mathbb{R}^{1+d} \rightarrow \mathbb{R}^d$ , seja  $V = \{(-t, x) : (t, x) \in U\}$  e defina  $G : V \rightarrow \mathbb{R}^d$  por  $G(-t, x) = -F(t, x)$ . Mostre que  $\gamma : I \rightarrow \mathbb{R}^d$  é solução de  $x' = F(t, x)$  se e somente se  $\beta : -I \rightarrow \mathbb{R}^d$  dada por  $\beta(-t) = \gamma(t)$  é solução de  $x' = G(t, x)$ . Use isto para provar o caso  $t \leq t_0$  na Proposição 46.

**Exercício 18** Seja  $F : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  uma função par, isto é,  $F(t, x) = F(-t, -x)$  para todo  $(t, x) \in \mathbb{R} \times \mathbb{R}^d$ . Sabendo que  $F$  é contínua e localmente Lipschitz na variável  $x$ , mostre que se  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^d$  é solução da EDO com condições iniciais

$$x' = F(t, x), x(0) = 0,$$

então  $\gamma$  é ímpar, isto é,  $\gamma(t) = -\gamma(-t)$  para todo  $t \in \mathbb{R}$ .

**Exercício 19** Seja  $a > 0$ , e seja  $K : \mathbb{R} \rightarrow \mathbb{R}$  tal que  $K(t) \leq -a^2, \forall t \in \mathbb{R}$ . Seja  $u : \mathbb{R} \rightarrow \mathbb{R}$  uma solução da EDO

$$\begin{cases} u'' + Ku = 0 \\ u(0) = 1, u'(0) = 0. \end{cases}$$

Mostre que  $u(t) \geq \cosh(at), \forall t \geq 0$ .

**Exercício 20** Seja  $f : [a, b] \rightarrow \mathbb{R}$  contínua com  $f \leq 0$ . Mostre que se  $u$  é solução não-identicamente nula da equação de Jacobi  $u'' + fu = 0$ , então  $u$  possui no máximo um zero. (Dica: caso haja dois, existe  $c \in [a, b]$  tal que  $u'(c) = 0$ . Trocando  $u$  por  $-u$  se necessário, podemos assumir que  $u(c) > 0$ . Agora aplique o teorema de comparação de Sturm).

**Exercício 21** Seja  $f : [a, b] \rightarrow \mathbb{R}$  contínua com  $f \geq c^2$  para uma constante  $c > 0$ . Mostre que para qualquer intervalo  $I \subset [a, b]$  de comprimento  $> \pi/c$  existe  $t \in I$  tal que  $u(t) = 0$ .



# Mecânica celeste

## SEÇÃO 11

### Problema de Kepler

O *problema de Kepler*, também conhecido como *problema da força central*, é um caso particular do problema dos dois corpos, em que assumimos que um dos corpos celestes tem massa bem maior do que a do outro a tal ponto de o considerarmos parado. A discussão desse problema tem três papéis:

- sendo uma simplificação do problema dos dois corpos, o problema de Kepler serve como um exemplo introdutório à mecânica celeste.
- de fato, o problema dos dois corpos pode ser reduzido ao problema de Kepler, como veremos na próxima seção.
- introduziremos elementos básicos da *dinâmica hamiltoniana*, que é extensivamente utilizado em sistemas dinâmicos e constitui uma área de pesquisa independente.

Considere dois corpos celestes de massas  $m_1, m_2$ . Pelas leis de gravitação de Newton, o corpo de massa  $m_1$  exerce sobre o corpo de massa  $m_2$  uma força de intensidade  $F = \frac{Gm_1m_2}{r^2}$ , onde  $r$  é a distância entre os corpos. Pela segunda lei de Newton, a aceleração  $a_2$  do corpo de massa  $m_2$  satisfaz

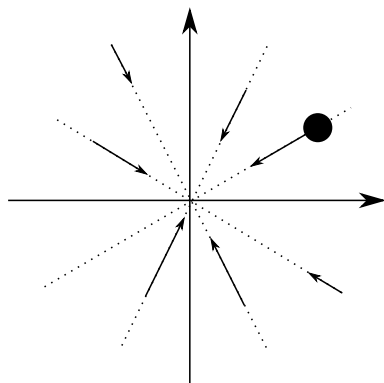
$$m_2a_2 = \frac{Gm_1m_2}{r^2} \implies a_2 = \frac{Gm_1}{r^2}.$$

Similarmente, a aceleração do corpo de massa  $m_1$  é  $a_1 = \frac{Gm_2}{r^2}$ . Se  $m_1 \gg m_2$  então  $a_1 \ll a_2$ , de modo que uma *simplificação* do problema consiste em assumir que o corpo de massa  $m_1$  é fixo, i.e. não se move. Essa aproximação, embora matematicamente útil para o que faremos na sequência, não representa a priori a situação real do problema dos dois corpos: mesmo que um dos corpos tenha aceleração negligível em comparação ao outro, ele ainda se move e portanto pode ter um comportamento extremamente complicado ao passar do tempo.

Feitas as devidas advertências, vamos proceder à formulação matemática do problema de Kepler. Coloque o corpo de massa  $m_1$  na origem do plano cartesiano. Como ele permanece parado, podemos desconsiderá-lo e apenas assumir que o outro corpo celeste, a partir de agora com massa denotada por  $m$ , está sujeito à ação de uma força apontando para a origem de intensidade

$$\|F\| = \frac{k}{r^2},$$

onde  $r$  é a distância para a origem. Definimos, assim, um campo gravitacional  $F : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2$  dado por  $F(x) = -k \frac{x}{\|x\|^3}$ . O sinal negativo significa que a força é central, apontando para a origem.



Como feito acima, a aceleração do corpo é igual a  $x'' = -\frac{k}{m} \frac{x}{\|x\|^3}$ , que define uma EDO autônoma de ordem dois. As condições iniciais do problema são a posição inicial  $x_0 \in \mathbb{R}^2$  e a velocidade inicial  $v_0 \in \mathbb{R}^2$ . Como  $(x_0, v_0) \in \mathbb{R}^4$ , dizemos que o grau de liberdade da EDO é quatro. Chegamos assim, à formulação abaixo.

**PROBLEMA DE KEPLER.** Descrever as leis de movimento de um corpo de massa  $m$  sujeito ao campo gravitacional central  $F(x) = -\frac{k}{m} \frac{x}{\|x\|^3}$  com posição e velocidade iniciais  $x_0, v_0 \in \mathbb{R}^2$ .

A função que define a EDO é  $f : \mathbb{R} \times (\mathbb{R}^2 \setminus \{0\}) \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$  dada por  $f(t, x, x') = -\frac{k}{m} \frac{x}{\|x\|^3}$ . Como  $F$  é contínua e localmente Lipschitz na segunda coordenada (exercício), cada condição inicial do problema define uma única solução. Assim, as soluções existem. Abaixo, vamos identificá-las qualitativamente. O primeiro passo é reduzir a EDO para uma EDO de ordem 1. Introduzindo a variável  $v = x'$ , obtemos a EDO equivalente

$$\begin{cases} x' = v \\ v' = -\frac{k}{m} \frac{x}{\|x\|^3}. \end{cases}$$

Na formulação acima, a EDO é um caso particular de uma classe importante de equações da dinâmica hamiltoniana. Essa perspectiva permite entender, do ponto de vista matemático, várias propriedades físicas intrínsecas ao problema, como a lei de conservação de energia. Antes de definir o formalismo básico da dinâmica hamiltoniana, vamos observar que o campo gravitacional  $F$  é um campo gradiente.

**Lema 52** | O campo  $F$  é gradiente. Mais especificamente,  $F = \nabla P$  onde  $P : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$  é definido por  $P(x) = \frac{k}{\|x\|}$ .

PROVA | Se  $x = (x_1, x_2)$ , então  $P(x) = k(x_1^2 + x_2^2)^{-\frac{1}{2}}$  e daí

$$\frac{\partial P}{\partial x_i} = k \left(-\frac{1}{2}\right) (x_1^2 + x_2^2)^{-\frac{3}{2}} 2x_i = -k \frac{x_i}{\|x\|^3}.$$

Assim,  $\nabla P(x) = F(x)$ . □

**Definição 53** (Hamiltoniano) A *função hamiltoniano* ou simplesmente *hamiltoniano* do problema de Kepler é a função  $H : (\mathbb{R}^2 \setminus \{0\}) \times \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por

$$H(x, v) = \frac{\|v\|^2}{2} - \frac{1}{m}P(x) = \frac{\|v\|^2}{2} - \frac{k}{m} \frac{1}{\|x\|}.$$

A priori, o hamiltoniano parece ser algo que não fornece muita informação. O ponto de partida de sua utilidade reside no cálculo de suas derivadas parciais

$$\begin{cases} \frac{\partial H}{\partial v}(x, v) = v \\ \frac{\partial H}{\partial x}(x, v) = -\frac{1}{m}\nabla P(x) = \frac{k}{m} \frac{x}{\|x\|^3}, \end{cases}$$

que são exatamente as funções que descrevem a EDO de ordem 1 do problema de Kepler. Esse formalismo permite reescrever tal EDO por meio das equações de Hamilton-Jacobi.

**Definição 54** (Equações de Hamilton-Jacobi) As *equações de Hamilton-Jacobi* definidas pelo hamiltoniano  $H$  são

$$\begin{cases} x' = \frac{\partial H}{\partial v} \\ v' = -\frac{\partial H}{\partial x}. \end{cases}$$

Agora é fácil provar que  $H$  é constante ao longo das soluções.

**Definição 55** (Energia) A *energia* do problema de Kepler é definida por

$$E(t) = H(x(t), v(t)) = \frac{\|v(t)\|^2}{2} - \frac{1}{m}P(x(t)),$$

onde  $\frac{\|v(t)\|^2}{2}$  representa a energia cinética e  $-\frac{1}{m}P(x(t))$  a energia potencial.

**Proposição 56** (Lei de conservação de energia) Se  $(x(t), v(t))$  satisfazem as equações de Hamilton-Jacobi, então a energia  $E(t)$  é constante.

PROVA Pela regra da cadeia e equações de Hamilton-Jacobi, temos

$$E'(t) = \frac{d}{dt}H(x(t), v(t)) = \left\langle \frac{\partial H}{\partial x}, x' \right\rangle + \left\langle \frac{\partial H}{\partial v}, v' \right\rangle = \left\langle \frac{\partial H}{\partial x}, \frac{\partial H}{\partial v} \right\rangle - \left\langle \frac{\partial H}{\partial v}, \frac{\partial H}{\partial x} \right\rangle = 0.$$

□

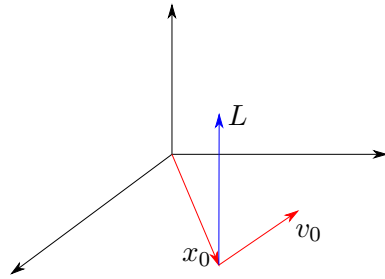
A função  $H$  também é chamada de *primeira integral do movimento*. Ela descreve uma relação implícita entre  $x$  e  $v$ , e assim reduz o grau de liberdade da EDO de quatro para três. Se acharmos mais integrais “linearmente independentes”, podemos reduzir ainda mais o grau de liberdade. Caso ele seja reduzido a um, obtemos uma descrição geométrica para a curva solução como sendo a interseção de subvariedades de codimensão 1. É isso que faremos, introduzindo mais dois invariantes ao problema de Kepler. Tais invariantes são propriedades físicas. A próxima é a preservação do momento angular.

**Definição 57** (Momento angular) O *momento angular* é definido por

$$L = x \times p = m(x \times v),$$

onde  $p = mv$  é o momento.

Note que  $L$  é um vetor em  $\mathbb{R}^3$ , perpendicular ao plano cartesiano em que o corpo de massa  $m$  se move.



**Lema 58** (Lei de conservação do momento angular) No problema de Kepler,  $L$  é preservado ao longo de trajetórias.

Em outras palavras, o módulo de  $L$  é constante.

PROVA | A derivada de  $L$  é

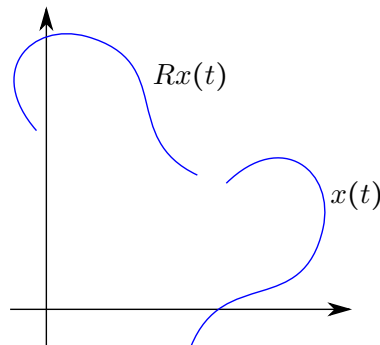
$$L' = m(x' \times v + x \times v') = m \left[ v \times v + x \times \left( -\frac{k}{m} \frac{x}{\|x\|^3} \right) \right] = 0.$$

□

Portanto,  $L$  é uma segunda integral de movimento. A última invariância do problema é a invariância por rotações: como o campo gravitacional é central, ele é simétrico com respeito à origem e portanto as soluções são invariantes por rotações. Vamos formalizar esse fato. Dado  $\beta \in \mathbb{R}$ , seja  $R = R_\beta$  a rotação de ângulo  $\beta$ :

$$R : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad R = \begin{bmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{bmatrix}.$$

**Lema 59** (Invariância por rotações) Se  $x(t)$  é a solução com condições iniciais  $(t_0, x_0, v_0)$ , então  $Rx(t)$  é a solução com condições iniciais  $(t_0, Rx_0, Rv_0)$ .



PROVA Pelo teorema de existência e unicidade (Teorema 14), basta checar que  $\tilde{x}(t) = Rx(t)$  satisfaz a EDO com condições iniciais

$$\begin{cases} \tilde{x}'' = -\frac{k}{m} \frac{\tilde{x}}{\|\tilde{x}\|^3} \\ \tilde{x}(t_0) = Rx_0, \tilde{x}'(t_0) = Rv_0. \end{cases}$$

É claro que as condições iniciais são satisfeitas. Resta checar a EDO. Como  $R$  preserva norma, temos

$$\tilde{x}'' = Rx'' = R \left[ -\frac{k}{m} \frac{x}{\|x\|^3} \right] = -\frac{k}{m} \frac{Rx}{\|Rx\|^3} = -\frac{k}{m} \frac{\tilde{x}}{\|\tilde{x}\|^3}.$$

□

Agora vamos descrever qualitativamente as soluções do problema de Kepler. Para isso, vamos usar os Lemas 58 e 59. Inicialmente, observe que se  $x_0, v_0$  são colineares ( $L$  é o vetor nulo), então o movimento ocorre ao longo da reta  $\{tx_0 : t \in \mathbb{R}\}$  e portanto a EDO é unidimensional. A partir de agora, assumiremos que  $x_0, v_0$  não são colineares, i.e.  $L \neq 0$ . Vamos provar que a trajetória  $x(t)$  está contida em uma cônica. Por não sabermos a priori qual cônica obteremos, trabalharemos com coordenadas polares  $(r, \theta)$ , pois nessas coordenadas todas as cônicas têm a representação universal

$$r(1 + e \cos \theta) = \text{const.} \tag{11.1}$$

O tipo de cônica depende da excentricidade  $e$ :

- $e = 0$ : circunferência.
- $0 < |e| < 1$ : elipse.
- $|e| = 1$ : parábola.
- $|e| > 1$ : hipérbole.

Sejam  $r(t), \theta(t)$  as coordenadas polares de  $x(t)$ , i.e.  $x(t) = r(t)(\cos \theta(t), \sin \theta(t))$ . Quando possível, omitiremos a variável  $t$ . No que segue,  $'$  denota a derivação com respeito a  $t$ . Defina os vetores

$$\begin{cases} C = (\cos \theta, \sin \theta) \\ N = (-\sin \theta, \cos \theta), \end{cases}$$

que formam uma base de  $\mathbb{R}^2$ . Para obter (11.1), vamos calcular  $L$  usando a base  $\{C, N\}$  e a base canônica. Começamos com a base  $\{C, N\}$ . Temos  $x = r(\cos \theta, \sin \theta) = rC$  e

$$v = x' = r'(\cos \theta, \sin \theta) + r(-\sin \theta, \cos \theta)\theta' = r'C + r\theta'N,$$

assim

$$L = m(x \times v) = m(rC) \times (r'C + r\theta'N) = mr^2\theta'(C \times N) = mr^2\theta' \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (11.2)$$

Como  $L$  é constante não-nulo, segue que  $mr^2\theta' = L_0$  é constante e não-nulo. Em particular,  $\theta' \neq 0$  e portanto podemos escrever  $t = t(\theta)$  e representar  $x, v$  na variável  $\theta$ . Vamos identificar a EDO que  $v$  satisfaz com respeito a essa variável. Pela EDO, temos

$$v' = -\frac{k}{m} \frac{x}{\|x\|^3} = -\frac{k}{m} \frac{rC}{r^3} = -\frac{k}{mr^2} C.$$

Pela regra da cadeia,

$$v_\theta = \frac{dv}{d\theta} = \frac{\frac{dv}{dt}}{\frac{d\theta}{dt}} = \frac{v'}{\theta'} = -\frac{k}{mr^2 L_0} C = -\frac{K}{L_0} (\cos \theta, \sin \theta).$$

Assim, pelo Teorema Fundamental do Cálculo

$$v = -\frac{k}{L_0} [(\sin \theta, -\cos \theta) + (\alpha_1, \alpha_2)],$$

onde  $\alpha_1, \alpha_2$  são constantes. Pelo Lema 59, podemos aplicar uma rotação e assumir que  $\alpha_1 = 0$ . Escreva  $(\alpha_1, \alpha_2) = (0, -e)$ . Então  $v = -\frac{k}{L_0} (\sin \theta, -\cos \theta - e)$ . Relembrando que  $x = r(\cos \theta, \sin \theta)$ , obtemos que

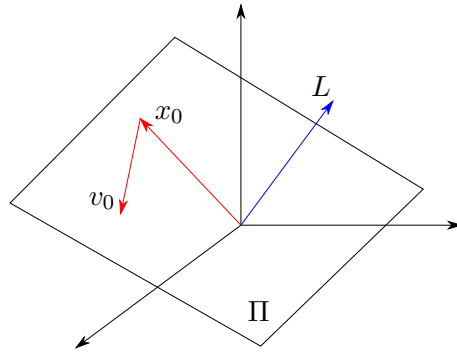
$$L = m(x \times v) = m \left( -\frac{kr}{L_0} \right) \begin{vmatrix} e_1 & e_2 & e_3 \\ \cos \theta & \sin \theta & 0 \\ \sin \theta & -\cos \theta - e & 0 \end{vmatrix} = \frac{mkr}{L_0} (1 + e \cos \theta) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Igualando com (11.2), concluímos que

$$\frac{mkr}{L_0} (1 + e \cos \theta) = L_0 \implies r(1 + e \cos \theta) = \frac{L_0^2}{mk},$$

que é a equação de uma cônica.

Para finalizar essa seção, vamos considerar o problema de Kepler em  $\mathbb{R}^3$  e mostrar que ele se reduz ao problema de Kepler em  $\mathbb{R}^2$ . Para isso, basta notar que a prova do Lema 58 se aplica *ipsis literis* e portanto o momento angular  $L = m(x \times v)$  é constante. Isso diz que os vetores  $x, v$  estão sobre o plano  $\Pi$  que passa pela origem e contém  $x_0, v_0$ , veja a figura.



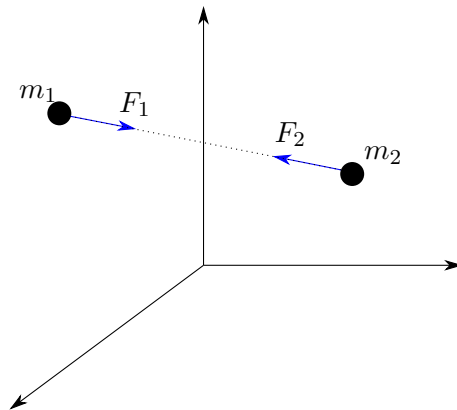
Assim, o corpo está sujeito ao problema de Kepler em  $\Pi \equiv \mathbb{R}^2$ .

SEÇÃO 12

**Problema dos dois corpos**

---

Agora estudaremos o problema dos dois corpos em  $\mathbb{R}^3$ . Considere dois corpos de massas  $m_1, m_2$  em  $\mathbb{R}^3$ , sujeitos apenas às forças gravitacionais que um exerce no outro. Sejam  $x_i =$  posição do corpo  $i$  e  $v_i =$  velocidade do corpo  $i$ .



As forças resultantes no corpos 1 e 2 são

$$\begin{cases} F_1 = k \frac{x_2 - x_1}{\|x_2 - x_1\|^3} = -k \frac{x_1 - x_2}{\|x_1 - x_2\|^3} \\ F_2 = -k \frac{x_2 - x_1}{\|x_2 - x_1\|^3}. \end{cases}$$

Pela segunda lei de Newton aplicada ao corpo 1, temos

$$m_1 v_1' = F_1 = -k \frac{x_1 - x_2}{\|x_1 - x_2\|^3} \implies v_1' = -\frac{k}{m_1} \frac{x_1 - x_2}{\|x_1 - x_2\|^3}.$$

Fazendo o mesmo para o corpo 2, obtemos a EDO de ordem 1 do problema dos

dois corpos, dada pelas quatro equações abaixo:

$$\begin{cases} x'_1 = v_1 \\ v'_1 = -\frac{k}{m_1} \frac{x_1 - x_2}{\|x_1 - x_2\|^3} \end{cases} \quad \text{e} \quad \begin{cases} x'_2 = v_2 \\ v'_2 = -\frac{k}{m_2} \frac{x_2 - x_1}{\|x_2 - x_1\|^3} \end{cases}.$$

A função que define essa EDO autônoma é  $f : \mathbb{R} \times V \rightarrow \mathbb{R}^{12}$  definida por

$$f(t, x_1, v_1, x_2, v_2) = \left( v_1, -\frac{k}{m_1} \frac{x_1 - x_2}{\|x_1 - x_2\|^3}, v_2, -\frac{k}{m_2} \frac{x_2 - x_1}{\|x_2 - x_1\|^3} \right),$$

onde  $V = \{(x_1, v_1, x_2, v_2) \in \mathbb{R}^{12} : x_1 \neq x_2\}$  é aberto. Como  $f$  é contínua e localmente Lipschitz na segunda coordenada (exercício), o teorema de existência e unicidade (Teorema 14) garante que cada condição inicial  $(t_0, x_1(t_0), v_1(t_0), x_2(t_0), v_2(t_0)) \in \mathbb{R} \times V$  define uma única solução local. Assim como no problema de Kepler, podemos descrever explicitamente tal solução. Uma das abordagens é introduzir integrais de movimento a partir das propriedades físicas do sistema. Ao invés disso, vamos aproveitar o que já fizemos na seção anterior, mostrando que a EDO acima pode ser simplificada e interpretada como dois problemas de Kepler separados. De todo modo, é interessante observar que a EDO acima também possui um hamiltoniano que descreve a energia total, definido por

$$H(x_1, v_1, x_2, v_2) = \frac{m_1 \|v_1\|^2}{2} + \frac{m_2 \|v_2\|^2}{2} - k \frac{1}{\|x_1 - x_2\|},$$

veja o Exercício 23.

Para reduzir o problema dos dois corpos a dois problemas de Kepler, vamos utilizar o centro de massa.

**Definição 60** (Centro de massa) O *centro de massa* do problema dos dois corpos é

$$\bar{x} = \frac{m_1}{m_1 + m_2} x_1 + \frac{m_2}{m_1 + m_2} x_2 = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2}.$$

Intuitivamente, o centro de massa é a posição onde podemos concentrar toda a massa do sistema. Ele costuma possuir um comportamento mais simples do que o corpos que o definem. Isso de fato ocorre no problema dos dois corpos.

**Lema 61** | No problema dos dois corpos, temos  $\bar{x}'' = 0$ .

Em outras palavras,  $\bar{x}$  possui um comportamento linear.

PROVA | Temos

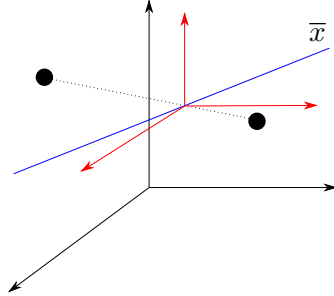
$$\begin{aligned} (m_1 + m_2)\bar{x}'' &= m_1 x_1'' + m_2 x_2'' \\ &= m_1 \left[ -\frac{k}{m_1} \frac{x_1 - x_2}{\|x_1 - x_2\|^3} \right] + m_2 \left[ -\frac{k}{m_2} \frac{x_2 - x_1}{\|x_2 - x_1\|^3} \right] = 0. \end{aligned}$$

□

Portanto, existem  $A, B \in \mathbb{R}^3$  tais que  $\bar{x} = At + B$ . Quando isso ocorre,



os físicos consideram um referencial se movendo como o centro de massa, de modo que o centro de massa pode ser considerado fixo e localizado na origem do referencial, e entendem como os corpos se movem com respeito a esse novo referencial.



Do ponto de vista matemática, isso consiste em introduzir novas coordenadas ao problema:

$$\begin{cases} \tilde{x}_1 = x_1 - At - B \\ \tilde{v}_1 = v_1 - A \end{cases} \quad \text{e} \quad \begin{cases} \tilde{x}_2 = x_2 - At - B \\ \tilde{v}_2 = v_2 - A. \end{cases}$$

Para entender como os corpos se movem nesse novo referencial, usamos a EDO satisfeita por  $(x_1, v_1, x_2, v_2)$  para encontrar a EDO que  $(\tilde{x}_1, \tilde{v}_1, \tilde{x}_2, \tilde{v}_2)$  satisfaz. Vejamos:

- $\tilde{x}'_1 = x'_1 - A = v_1 - A = \tilde{v}_1$ .
- $\tilde{v}'_1 = v'_1 = -\frac{k}{m_1} \frac{x_1 - x_2}{\|x_1 - x_2\|^3} = -\frac{k}{m_1} \frac{\tilde{x}_1 - \tilde{x}_2}{\|\tilde{x}_1 - \tilde{x}_2\|^3}$ .

O mesmo vale para  $\tilde{x}'_2, \tilde{v}_2$ . Note que agora o centro de massa é

$$\frac{m_1(x_1 - At - B) + m_2(x_2 - At - B)}{m_1 + m_2} = \bar{x} - (At + B) = 0.$$

Assim, no novo referencial temos o centro de massa na origem. A partir de agora, toda a discussão será feita com respeito a esse referencial. Por simplicidade, vamos denotar as coordenadas do novo referencial por  $x_1, v_1, x_2, v_2$ . Logo  $m_1 x_1 + m_2 x_2 = 0$ . Essa identidade é a chave para reduzir o problema dos dois corpos a dois problemas de Kepler. Temos

$$x_2 = -\frac{m_1}{m_2} x_1 \implies x_1 - x_2 = \left(\frac{m_2 - m_1}{m_2}\right) x_1$$

e portanto

$$v'_1 = -\frac{k}{m_1} \frac{\left(\frac{m_2 - m_1}{m_2}\right) x_1}{\left(\frac{m_2 - m_1}{m_2}\right)^3 \|x_1\|^3} = -k_1 \frac{x_1}{\|x_1\|^3},$$

onde  $k_1 = \frac{km_2^2}{m_1(m_2 - m_1)^2}$  é constante. Assim,  $(x_1, v_1)$  satisfazem

$$\begin{cases} x'_1 = v_1 \\ v'_1 = -k_1 \frac{x_1}{\|x_1\|^3}, \end{cases}$$

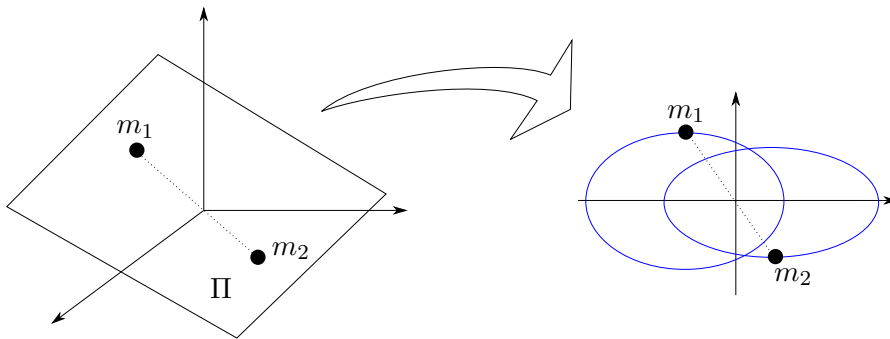
que é um problema de Kepler em  $\mathbb{R}^3$ . Similarmente,  $(x_2, v_2)$  satisfazem um problema de Kepler em  $\mathbb{R}^3$ :

$$\begin{cases} x_2' = v_2 \\ v_2' = -k_2 \frac{x_2}{\|x_2\|^3}. \end{cases}$$

Assim, com respeito ao referencial do centro de massa, cada um dos corpos realiza um movimento cônico em um plano. De fato, os planos dos movimentos coincidem. Para provar isso, calculamos o momento angular de cada corpo separadamente. Derivando  $x_2 = -\frac{m_1}{m_2}x_1$ , temos  $v_2 = -\frac{m_1}{m_2}v_1$ . Assim,

$$L_2 = m_2(x_2 \times v_2) = m_2\left(-\frac{m_1}{m_2}x_1\right) \times \left(-\frac{m_1}{m_2}v_1\right) = \frac{m_1}{m_2} [m_1(x_1 \times v_1)] = \frac{m_1}{m_2}L_1,$$

ou seja, os momentos angulares são proporcionais. Assim, os planos  $\Pi_1, \Pi_2$  que contêm os movimentos dos corpos 1 e 2 coincidem. Nesse plano, cada corpo se move ao longo de uma cônica que tem a origem como um dos focos.



Finalizamos essa seção notando que, de fato, o grau de liberdade da EDO do problema dos dois corpos pode ser reduzido a um por meio de integrais de movimento. Relembre que as condições iniciais são  $(x_1, v_1, x_2, v_2) \in \mathbb{R}^{12}$ , logo o grau de liberdade inicial é 12. Temos os seguintes invariantes:

- Hamiltoniano  $H$  (veja o Exercício 23): grau de liberdade = 1.
- Centro de massa  $\bar{x} = At + B$ , com  $(A, B) \in \mathbb{R}^6$ : grau de liberdade = 6.
- Momento angular  $L$ : grau de liberdade = 1.
- Invariância pelo grupo de rotação  $\text{SO}(3)$ : grau de liberdade =  $\dim(\text{SO}(3)) = 3$ .

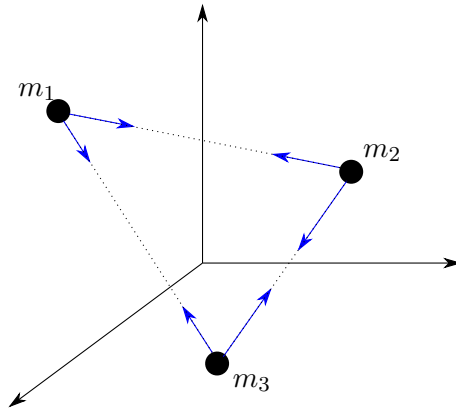
Juntos, os invariantes reduzem  $1 + 6 + 1 + 3 = 11$  graus de liberdade.

### SEÇÃO 13

## Problema dos três corpos

Finalizamos esse capítulo discutindo o problema dos três corpos em  $\mathbb{R}^3$ . Considere três corpos de massas  $m_1, m_2, m_3$  em  $\mathbb{R}^3$ , sujeitos apenas às forças gra-

vitacionais que exercem uns nos outros. Sejam  $x_i =$  posição do corpo  $i$  e  $v_i =$  velocidade do corpo  $i$ .



A força resultante no corpo  $i$  é

$$F_i = - \sum_{j \neq i} G m_i m_j \frac{x_i - x_j}{\|x_i - x_j\|^3} = -G m_i \sum_{j \neq i} m_j \frac{x_i - x_j}{\|x_i - x_j\|^3}.$$

Pela segunda lei de Newton, temos

$$m_i v_i' = F_i = -G m_i \sum_{j \neq i} m_j \frac{x_i - x_j}{\|x_i - x_j\|^3} \implies v_i' = -G \sum_{j \neq i} m_j \frac{x_i - x_j}{\|x_i - x_j\|^3}.$$

Assim, o problema dos três corpos satisfaz uma EDO de ordem 1 dada pelas seis equações abaixo:

$$\begin{cases} x_1' = v_1 \\ v_1' = -G \sum_{j \neq 1} m_j \frac{x_1 - x_j}{\|x_1 - x_j\|^3} \\ x_2' = v_2 \\ v_2' = -G \sum_{j \neq 2} m_j \frac{x_2 - x_j}{\|x_2 - x_j\|^3} \\ x_3' = v_3 \\ v_3' = -G \sum_{j \neq 3} m_j \frac{x_3 - x_j}{\|x_3 - x_j\|^3}. \end{cases}$$

A função que define essa EDO autônoma é  $f : \mathbb{R} \times V \rightarrow \mathbb{R}^{18}$  definida por

$$f(t, x_1, v_1, x_2, v_2, x_3, v_3) = \left( v_1, -G \sum_{j \neq 1} m_j \frac{x_1 - x_j}{\|x_1 - x_j\|^3}, v_2, -G \sum_{j \neq 2} m_j \frac{x_2 - x_j}{\|x_2 - x_j\|^3}, v_3, -G \sum_{j \neq 3} m_j \frac{x_3 - x_j}{\|x_3 - x_j\|^3} \right),$$

onde  $V = \{(x_1, v_1, x_2, v_2, x_3, v_3) \in \mathbb{R}^{12} : x_1 \neq x_2 \neq x_3 \neq x_1\}$  é aberto. Como  $f$  é

contínua e localmente Lipschitz na segunda coordenada (exercício), o teorema de existência e unicidade (Teorema 14) garante que cada condição inicial

$$(t_0, x_1(t_0), v_1(t_0), x_2(t_0), v_2(t_0), x_3(t_0), v_3(t_0)) \in \mathbb{R} \times V$$

define uma única solução local. Porém, nesse caso não temos integrais de movimento suficientes para baixarem o grau de liberdade a um. É por isso que, dentre outras razões, o problema dos três corpos permanece sem solução.

## SEÇÃO 14

**Exercícios**

**Exercício 22** (Prova alternativa do problema de Kepler) Vamos fornecer outra maneira de provar que as trajetórias no problema de Kepler estão em uma cônica. Utilizando a mesma notação da Seção 11, assuma que já temos  $mr^2\theta' = L_0 \neq 0$ .

- (1) Introduza a variável  $u = r\theta'$ . Nessa variável, a equação da cônica é equivalente a  $u = c(1 + e \cos \theta)$  para constantes  $c, e$ .
- (2) Usando a EDO do problema de Kepler, mostre que

$$v' = -\frac{k\theta'}{L_0}C.$$

Usando que  $v = r'C + r\theta'N$ , mostre que

$$v' = (r'' - u\theta')C + (r'\theta' + u')N.$$

Conclua que

$$\begin{cases} r'' - u\theta' = -\frac{k}{L_0}\theta' & \text{(I)} \\ r'\theta' + u' = 0. & \text{(II)} \end{cases}$$

- (3) Use (II) e a regra da cadeia para mostrar que  $u_\theta = -r'$ . Derive mais uma vez para obter que  $r'' = -u_{\theta\theta}\theta'$ . Substitua essa igualdade em (I) para concluir que

$$u_{\theta\theta} + u = \frac{k}{L_0}.$$

- (4) Obtenha que  $u = \frac{k}{L_0}(1 + e \cos(\theta - \theta_0))$ . Pelo Lema 59, assuma que  $\theta_0 = 0$  e conclua que

$$u = \frac{k}{L_0}(1 + e \cos \theta).$$

**Exercício 23** Utilizando a notação da Seção 12, defina  $H : \mathbb{R} \times V \rightarrow \mathbb{R}$  por

$$H(x_1, v_1, x_2, v_2) = \frac{m_1\|v_1\|^2}{2} + \frac{m_2\|v_2\|^2}{2} - k \frac{1}{\|x_1 - x_2\|}.$$

(1) Mostre que

$$\begin{cases} \frac{\partial H}{\partial x_1} = -k \frac{x_1 - x_2}{\|x_1 - x_2\|^3} \\ \frac{\partial H}{\partial v_1} = m_1 v_1 \end{cases} \quad \text{e} \quad \begin{cases} \frac{\partial H}{\partial x_2} = -k \frac{x_2 - x_1}{\|x_2 - x_1\|^3} \\ \frac{\partial H}{\partial v_2} = m_2 v_2. \end{cases}$$

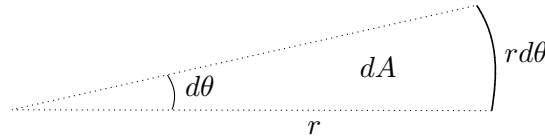
(2) Conclua que se  $(x_1(t), v_1(t), x_2(t), v_2(t))$  é solução do problema dos dois corpos, então  $E(t) = H(x_1(t), v_1(t), x_2(t), v_2(t))$  é constante.

**Exercício 24** Considere o problema de Kepler, descrito pela EDO

$$\begin{cases} x' = v \\ v' = -\frac{k}{m} \frac{x}{|x|^3} \end{cases},$$

onde  $(x, v) \in \mathbb{R}^2 \times \mathbb{R}^2$  representam a posição e velocidade,  $x \neq 0$ . O objetivo desse problema é provar a segunda lei de Kepler: uma solução do problema de Kepler varre áreas iguais em tempos iguais. Mais precisamente, vamos mostrar que se  $A(t)$  = área varrida pela trajetória até o tempo  $t$ , então  $\frac{dA}{dt}$  é constante. Sejam  $(r, \theta)$  as coordenadas polares, e seja  $L = m(x \times v)$  o momento angular. Pela igualdade (11.2),  $L$  é uma integral do movimento com comprimento igual a  $L_0 = mr^2\theta'$ .

(a) Considere a trajetória entre ângulos  $\theta$  e  $\theta + d\theta$ , e seja  $dA$  a área varrida entre esses tempos, como na figura abaixo. Mostre que  $dA \approx \frac{r^2 d\theta}{2}$ .

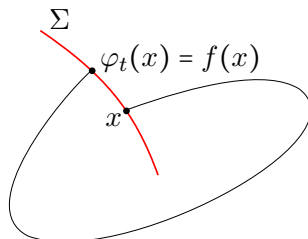


(b) Prove que  $\frac{dA}{dt} = \frac{L_0}{2m}$ , uma constante.

**Exercício 25** Enuncie o problema dos  $n$  corpos em  $\mathbb{R}^3$ , e descreva a EDO de ordem 1 que o problema satisfaz. Qual sua dimensão (grau de liberdade)?

# *Sistemas dinâmicos: definições básicas e homeomorfismos do intervalo*

Os exemplos iniciais aqui tratados foram motivados pela física, e naturalmente são definidos para todo tempo  $t \in \mathbb{R}$  (tempo contínuo). Um dos modos de estudar tais problemas foi desenvolvido por Poincaré, ainda no final do século XIX: considerando uma seção  $\Sigma$  transversal a uma trajetória recorrente, podemos definir uma transformação de retorno  $f : \Sigma \rightarrow \Sigma$ , onde  $f(x)$  é a primeira posição em que a trajetória positiva de  $x$  intersecta  $\Sigma$ .



A transformação  $f$  é o que chamamos de um sistema dinâmico com tempo discreto. Entender  $f$  permite entender algumas propriedades do problema inicial. No que segue, vamos definir formalmente essas duas possibilidades para dinâmicas, com tempo discreto e com tempo contínuo. Introduziremos conjuntos que são utilizados para analisar o comportamento de sistemas dinâmicos. A razão de introduzi-los é, além de apresentar ao leitor conceitos importantes, possibilitar que eles trabalhem com os conceitos e obtenham um melhor entendimento de dinâmicas básicas. Finalizamos o capítulo estudando a dinâmica dos homeomorfismos do intervalo. Tal classe é interessante pois podemos caracterizar todo o comportamento, no futuro e passado.

## SEÇÃO 15

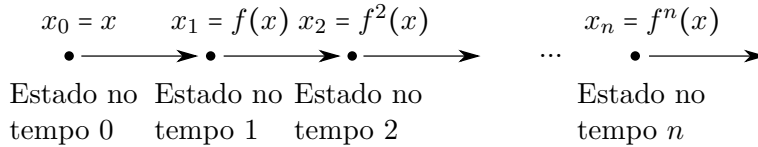
### Transformações e fluxos

Fixe um conjunto  $X$ , que representará nosso espaço de fase, igual a todas as possíveis configurações do sistema. Em geral, assumimos que  $X$  possui estruturas adicionais, por exemplo que ele é um espaço topológico/métrico compacto ou uma variedade Riemanniana compacta. Definida no conjunto  $X$ , consideraremos uma *dinâmica*, que representa uma regra que descreve como as configurações do sistema variam com o tempo. O objeto principal de estudo é o par (espaço de fase, dinâmica). Começamos definindo dinâmicas com tempo discreto.

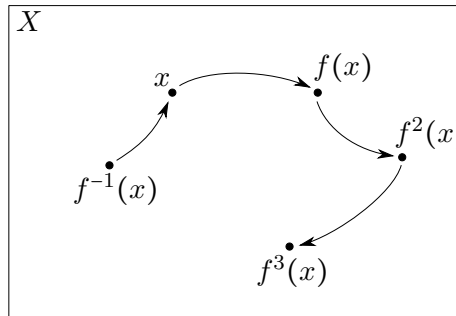
**Definição 62**

(Transformação) Uma *transformação* em  $X$  é uma função  $f : X \rightarrow X$ . O par  $(X, f)$  é chamado de *dinâmica com tempo discreto*.

Intuitivamente,  $f$  é a maneira que permite entender a configuração do sistema em toda unidade discreta e positiva de tempo, de modo que se  $x$  representa a configuração no tempo 0, então  $f(x)$  representa a configuração no tempo 1,  $f^2(x)$  representa a configuração no tempo 2 e assim sucessivamente.



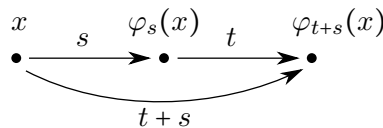
Se, adicionalmente,  $f$  é invertível, então também podemos entender a configuração do sistema em unidades discretas e negativas de tempo:  $f^{-1}(x)$  representa a configuração no tempo  $-1$ ,  $f^{-2}(x)$  representa a configuração no tempo  $-2$  e assim sucessivamente.



De agora em diante, para simplificar o entendimento dos conceitos, focaremos nas definições para  $f$  invertível. Antes, vejamos a definição de dinâmica com tempo contínuo.

**Definição 63**

(Fluxo) Um *fluxo* em  $X$  é uma família  $\varphi = \{\varphi_t\}_{t \in \mathbb{R}}$  de transformações  $\varphi_t : X \rightarrow X$  tais que  $\varphi_0 = \text{Id}$  e  $\varphi_t \circ \varphi_s = \varphi_{t+s}$  para todos  $t, s \in \mathbb{R}$ . O par  $(X, \varphi)$  é chamado *dinâmica com fluxo contínuo*.



Cada  $\varphi_t$  é invertível, com inversa  $\varphi_{-t}$ . Em outras palavras, se  $f$  é invertível então  $f$  define uma  $\mathbb{Z}$ -ação em  $X$ , enquanto que um fluxo define uma  $\mathbb{R}$ -ação (família a 1 parâmetro). Em termos bem abrangentes, podemos dizer que o principal objetivo dos sistemas dinâmicos é o seguinte.

**OBJETIVO DOS SISTEMAS DINÂMICOS.** Entender o comportamento dos sistemas ao longo do tempo, isto é, entender  $f^n(x)$  e  $\varphi_t(x)$  quando  $n, t \rightarrow +\infty$ , para cada ou a maioria das condições iniciais  $x \in X$ .

Para isso, introduzimos a noção de órbita.

**Definição 64** (Órbita ou trajetória) A *órbita* ou *trajetória* de  $x$  para  $(X, f)$  é

$$\mathcal{O}_f(x) = \{f^n(x) : n \in \mathbb{Z}\}.$$

A *órbita* ou *trajetória* de  $x$  para  $(X, \varphi)$  é

$$\mathcal{O}_\varphi(x) = \{\varphi_t(x) : t \in \mathbb{R}\}.$$

Desse modo, nosso objetivo é entender os conjuntos  $\mathcal{O}_f(x)$  e  $\mathcal{O}_\varphi(x)$ . Certamente, as órbitas mais simples são as periódicas.

**Definição 65** (Ponto periódico) Dizemos que  $x \in X$  é *periódico* para  $f$  se existir  $n \neq 0$  tal que  $f^n(x) = x$ . O menor  $n$  tal que  $f^n(x) = x$  é chamado de *período* de  $x$ . Similarmente, dizemos que  $x \in X$  é *periódico* para o fluxo  $\varphi$  se existir  $t \neq 0$  tal que  $\varphi_t(x) = x$ . O menor  $t > 0$  tal que  $\varphi_t(x) = x$  é chamado de *período* de  $x$ .

Denotamos por  $\text{Per}_f(x) = \text{Per}(x)$  o período de  $x$  com respeito a  $f$ , e similarmente por  $\text{Per}_\varphi(x)$  o período de  $x$  com respeito a  $\varphi$ . As órbitas periódicas dependem apenas de um intervalo finito de tempos. Para analisar as órbitas não periódicas, as seguintes noções tornam-se úteis.

**Definição 66** ( $\omega$ -limite e  $\alpha$ -limite) Dados  $(X, f)$ , para cada  $x \in X$  definimos o conjunto  $\omega$ -limite de  $x$  por

$$\omega_f(x) = \omega(x) = \{y \in X : \exists n_k \rightarrow +\infty \text{ tal que } f^{n_k}(x) \rightarrow y\}$$

e o conjunto  $\alpha$ -limite de  $x$  por

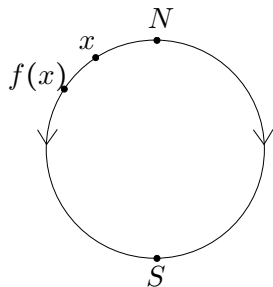
$$\omega_f(x) = \omega(x) = \{y \in X : \exists n_k \rightarrow -\infty \text{ tal que } f^{n_k}(x) \rightarrow y\}.$$

Definições análogas valem para fluxos. A nomenclatura vem do grego:  $\alpha$  é a primeira letra do alfabeto, e está associada a “nascimento”, i.e. onde a órbita de  $x$  começa;  $\omega$  é a última letra e está associada a “morte”, i.e. onde a órbita de  $x$  termina. Claramente, se  $x$  é periódico então  $\omega(x) = \alpha(x) = \mathcal{O}(x)$ .

**Exemplo 67** Considere a dinâmica pólo norte - pólo sul  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  tal que:

1.  $f(N) = N$  e  $f(S) = S$ .
2.  $f$  afasta pontos de  $N$  e aproxima de  $S$ , i.e. se  $x \neq N, S$  então  $f^n(x) \rightarrow S$  quando  $n \rightarrow \infty$  e  $f^n(x) \rightarrow N$  quando  $n \rightarrow -\infty$ .

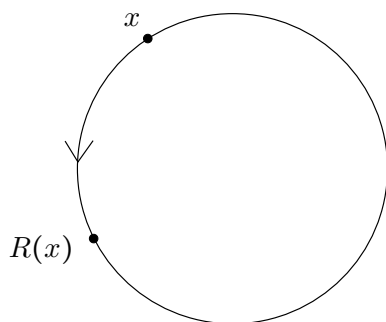




Temos  $\text{Per}(f) = \text{Fix}(f) = \{N, S\}$ . Para  $x \neq N, S$ , temos  $\alpha(x) = \{N\}$  e  $\omega(x) = \{S\}$ .

**Exemplo 68** (Rotações em  $\mathbb{S}^1$ ) Dado  $\beta \in \mathbb{R}$ , defina  $R = R_\beta : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  por

$$R(x) = x + \beta \pmod{1}.$$



O comportamento de  $R$  depende das propriedades diofantinas de  $\beta$ . Se  $\beta = p/q \in \mathbb{Q}$  com  $\text{mdc}(p, q) = 1$  e  $q > 0$ , então  $R^q(x) = x$  para todo  $x \in \mathbb{S}^1$ , e  $q$  é o período. Mais ainda,  $p$  é igual ao número de voltar que a órbita de  $x$  faz ao redor de  $\mathbb{S}^1$  antes de retornar para  $x$ . Quando  $\beta \notin \mathbb{Q}$ , obtemos um comportamento totalmente diferente:  $\mathcal{O}(x)$  é densa em  $\mathbb{S}^1$ . Isso decorre diretamente do Lema de Kronecker, veja o Apêndice (Seção 38). Mais ainda,  $\alpha(x) = \omega(x) = \mathbb{S}^1$  (isso decorre de uma versão mais refinada do Lema de Kronecker, que afirma que se  $\beta \in \mathbb{R} \setminus \mathbb{Q}$  então o conjunto  $\{m + n\beta : m \in \mathbb{Z}, n > 0\}$  também é denso em  $\mathbb{R}$ ).

Para fluxos contínuos em espaços topológicos conexos e compactos, temos propriedades interessantes para  $\alpha(x)$  e  $\omega(x)$ . Dizemos que  $\varphi = \{\varphi_t\}_{t \in \mathbb{R}}$  é um fluxo contínuo se a transformação

$$\begin{aligned} \mathbb{R} \times X &\rightarrow X \\ (t, x) &\mapsto \varphi_t(x) \end{aligned}$$

é contínua. Em particular, cada  $\varphi_t$  é contínua.

**Proposição 69** Sejam  $X$  um espaço topológico conexo compacto e  $\varphi$  um fluxo contínuo em  $X$ . Para todo  $x \in X$ , valem as seguintes afirmações:

- (1)  $\omega(x) \neq \emptyset$ .

- (2)  $\omega(x)$  é fechado.
- (3)  $\omega(x)$  é invariante, isto é,  $\varphi_t(\omega(x)) = \omega(x)$  para todo  $t \in \mathbb{R}$ .
- (4)  $\omega(x)$  é conexo.

PROVA (1) Como  $X$  é compacto,  $\{\varphi_n(x)\}_{n \geq 0}$  possui uma subsequência convergente  $\varphi_{n_k}(x) \rightarrow y \in \omega(x)$ .

(2) Se  $y \notin \omega(x)$ , então existem  $V \ni y$  aberto e  $t_0 > 0$  tais que  $\varphi_t(x) \notin V$  para todo  $t > t_0$ . Então nenhum ponto de  $V$  está em  $\omega(x)$ , logo  $V \subset X \setminus \omega(x)$ . Isso prova que  $\omega(x)$  é fechado.

(3) Basta mostrar que  $\varphi_t(\omega(x)) \subset \omega(x)$  para todo  $t \in \mathbb{R}$ , pois então aplicando essa inclusão para  $-t$  obtemos

$$\varphi_{-t}(\omega(x)) \subset \omega(x) \implies \omega(x) \subset \varphi_t(\omega(x)),$$

estabelecendo a inclusão reversa. Fixe  $t \in \mathbb{R}$ . Se  $\varphi_{t_k}(x) \rightarrow y \in \omega(x)$  com  $t_k \rightarrow +\infty$ , então

$$\varphi_t(y) = \varphi_t(\lim \varphi_{t_k}(x)) = \lim \varphi_{t+t_k}(x),$$

com  $t+t_k \rightarrow +\infty$ , e portanto  $\varphi_t(y) \in \omega(x)$ .

(4) Por absurdo, suponha que existam  $V_1, V_2 \subset X$  abertos disjuntos tais que  $\omega(x) \subset V_1 \cap V_2$  e  $V_i \cap \omega(x) \neq \emptyset$  para  $i = 1, 2$ . Então existem  $t_k, s_k \rightarrow +\infty$  tais que  $\varphi_{t_k}(x) \in V_1$  e  $\varphi_{s_k}(x) \in V_2$ . Pelo teorema da alfândega, existe uma sequência  $r_k \rightarrow +\infty$  tal que  $\varphi_{r_k}(x) \notin V_1 \cap V_2$ . Como  $X \setminus (V_1 \cap V_2)$  é compacto, existe uma subsequência convergente  $\varphi_{r_{k_j}}(x) \rightarrow y \in X \setminus (V_1 \cap V_2)$ . Isso é uma contradição, pois  $y \in \omega(x) \subset V_1 \cup V_2$ .  $\square$

Continuando nosso estudo, vamos introduzir mais conceitos.

**Definição 70** (Conjunto limite) O *conjunto limite* de  $f$  é definido por  $L(f) = L_+(f) \cup L_-(f)$  onde

$$L_+(f) = \overline{\bigcup_{x \in X} \omega(x)} \text{ e } L_-(f) = \overline{\bigcup_{x \in X} \alpha(x)}.$$

Quando  $f$  não é invertível, definimos

$$L(f) = \overline{\bigcup_{x \in X} \omega(x)}.$$

Definições análogas valem para fluxos. Para a próxima definição, enunciamos separadamente as versões para transformações e para fluxos.

**Definição 71** (Conjunto não-errante para transformações) Dado  $(X, f)$ , dizemos que  $x \in X$  é *errante* para  $f$  se existir  $U \ni x$  aberto tal que  $f^n(U) \cap U = \emptyset$  para todo  $n \geq 1$ . Quando  $x$  não é errante, dizemos que  $x$  é *não-errante*. O *conjunto não-errante* de  $f$  é

$$\Omega(f) = \{x \in X : x \text{ é não-errante para } f\}.$$

Note que  $\Omega(f)$  é fechado e  $L(f) \subset \Omega(f)$  (Exercício 30). Para fluxos, devemos ter atenção ao seguinte fato: se  $V$  é aberto e  $\varphi$  é contínuo, então  $\varphi_t(X) \cap X \neq \emptyset$  para  $t$  pequeno.

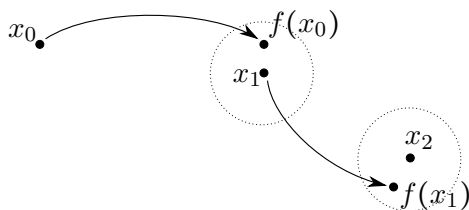
**Definição 72** (Conjunto não-errante para fluxos) Dado  $(X, \varphi)$ , dizemos que  $x \in X$  é *errante* para  $\varphi$  se existir  $U \ni x$  aberto e  $t_0 > 0$  tais que  $\varphi_t(U) \cap U = \emptyset$  para todo  $t > t_0$ . Quando  $x$  não é errante, dizemos que  $x$  é *não-errante*. O *conjunto não-errante* de  $\varphi$  é

$$\Omega(\varphi) = \{x \in X : x \text{ é não-errante para } \varphi\}.$$

Alternativamente,  $x \in X$  é não-errante se, para todos  $U \ni x$  aberto e  $t_0 > 0$ , existe  $t > t_0$  tal que  $\varphi_t(U) \cap U \neq \emptyset$ . Para o último conjunto que definiremos, necessitamos introduzir um importante conceito. De agora em diante, nos restringiremos a transformações e assumiremos que  $X$  é espaço métrico, com distância  $d$ .

**Definição 73** (Pseudo-órbita) Dado  $\varepsilon > 0$ , dizemos que uma sequência  $x_0, x_1, x_n$  em  $X$  é uma  $\varepsilon$ -*pseudo-órbita* se

$$d(f(x_i), x_{i+1}) < \varepsilon, \quad i = 0, \dots, n-1.$$



Esse conceito é um enfraquecimento da noção de órbita, onde  $f(x_i)$  coincide com  $x_{i+1}$ . Definições análogas podem ser aplicadas para sequências infinitas  $x_0, x_1, \dots$ . Sua grande importância intrínseca reside no fato de identificar possíveis recorrências não identificadas pelos conjuntos  $\text{Per}(f), L(f), \Omega(f)$ . Sua importância prática surge quando pedimos para um computador calcular os iterados de um ponto: por conta de erros de aproximação, o computador não retorna uma órbita real mas uma pseudo-órbita: em cada iteração, ocorre algum arredondamento que introduz um pequeno erro.

Para analisar as propriedades de um sistema dinâmico, focamos nas pseudo-órbitas periódicas.

**Definição 74** (Conjunto recorrente por cadeias) Dizemos que  $x \in X$  é *recorrente por cadeias* se, para todo  $\varepsilon > 0$ , existir uma  $\varepsilon$ -pseudo-órbita periódica  $x_0 = x, x_1, \dots, x_n = x$ . O *conjunto recorrente por cadeias* de  $f$  é definido por

$$R(f) = \{x \in X : x \text{ é recorrente por cadeias}\}.$$

O conjunto  $R(f)$  é fechado e contém  $\Omega(f)$ . Assim, temos a sequência de inclusões

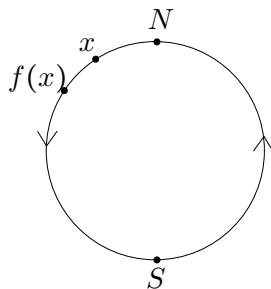
$$\text{Fix}(f) \subset \text{Per}(f) \subset L(f) \subset \Omega(f) \subset R(f),$$

veja o Exercício 30.

**Exemplo 75** | Considere a transformação do Exemplo 67. Temos

$$\text{Fix}(f) = \text{Per}(f) = L(f) = \Omega(f) = R(f) = \{N, S\}.$$

**Exemplo 76** | Considere a transformação  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  dada por:



Temos  $\text{Fix}(f) = \text{Per}(f) = L(f) = \Omega(f) = R(f) = \{N, S\}$  e  $R(f) = \mathbb{S}^1$ .

## SEÇÃO 16

### Homeomorfismos do intervalo

Nessa seção, consideramos e caracterizamos a dinâmica de homeomorfismos de intervalos compactos. Sem perda de generalidade, tomamos  $X = [0, 1]$ . Relembre que um homeomorfismo  $f : X \rightarrow X$  é uma transformação contínua invertível e com inversa contínua. Por exemplo,  $f(x) = x$  e  $f(x) = 1 - x$  são homeomorfismos, mas transformações da família quadrática não são. Em todo o estudo de dinâmicas unidimensionais, faremos uso recorrente do teorema do valor intermediário (TVI). Tal teorema faz uso crucial da unidimensionalidade de  $\mathbb{R}$ , e de fato não é válido em dimensão maior.

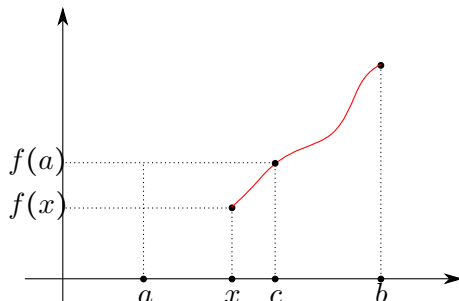
**Lema 77** | (Monotonicidade) Se  $f : [a, b] \rightarrow \mathbb{R}$  é contínua e injetiva, então  $f$  é estritamente monótona, isto é, ou  $f$  é estritamente crescente ou estritamente decrescente.

No que segue, escreveremos apenas crescente/decrescente.

**PROVA** | Temos dois casos:  $f(a) < f(b)$  ou  $f(a) > f(b)$ . Assumimos inicialmente que  $f(a) < f(b)$ . Vamos mostrar que  $f$  é crescente.

**FATO 1.** Se  $x \in (a, b)$ , então  $f(x) \in (f(a), f(b))$ .

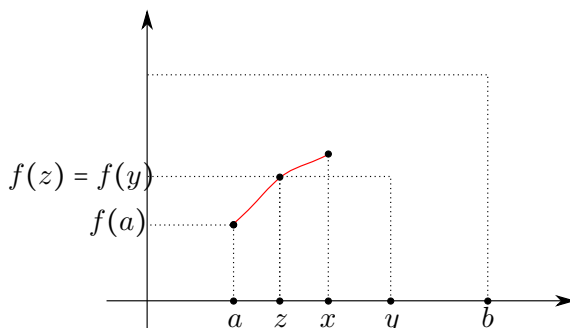
PROVA (Fato 1) Vamos mostrar que  $f(x) > f(a)$  (a prova de que  $f(x) < f(b)$  é análoga). Por absurdo, suponha que  $f(x) < f(a)$ . Pelo TVI aplicado ao intervalo  $(x, b)$ , existe  $c \in (x, b)$  tal que  $f(c) = f(a)$ , o que contraria a injetividade de  $f$ .



□

FATO 2. Se  $x < y$ , então  $f(x) < f(y)$ .

PROVA (Fato 2.) Por contradição, assuma que  $f(x) > f(y)$ . Pelo TVI aplicado ao intervalo  $(a, x)$ , existe  $z \in (a, x)$  tal que  $f(z) = f(y)$ , contrariando novamente a injetividade de  $f$ .



□

Pelos Fatos 1 e 2,  $f$  é crescente. No caso  $f(a) > f(b)$ , aplique o caso anterior à função  $-f$ . □

Em particular, a imagem de  $f$  é igual ao intervalo de extremos  $f(a)$  e  $f(b)$ . Assim, se  $f : [0, 1] \rightarrow [0, 1]$  é contínua e injetiva, então:

- Se  $f$  é crescente, então  $f(0) = 0$  e  $f(1) = 1$ ; portanto  $0, 1 \in \text{Fix}(f)$ .
- Se  $f$  é decrescente, então  $f(0) = 1$  e  $f(1) = 0$ ; nesse caso  $0, 1 \in \text{Per}_2(f)$ .

No que segue, assuma que  $f$  é crescente. O próximo resultado caracteriza o comportamento de  $f$  entre dois pontos fixos.

**Proposição 78** Seja  $f : [0, 1] \rightarrow [0, 1]$  contínua e crescente, e sejam  $x_0, x_1 \in \text{Fix}(f)$  tais que  $(x_0, x_1) \cap \text{Fix}(f) = \emptyset$ . Então uma das situações abaixo ocorre:

- (1)  $\{f^n(x)\}_{n \geq 0}$  é crescente e converge para  $x_1$ , para todo  $x \in (x_0, x_1)$ .
- (2)  $\{f^n(x)\}_{n \geq 0}$  é decrescente e converge para  $x_0$ , para todo  $x \in (x_0, x_1)$ .

PROVA Temos dois casos:  $f(x) > x$  para todo  $x \in (x_0, x_1)$  ou  $f(x) < x$  para todo  $x \in (x_0, x_1)$ . No primeiro caso, a sequência  $x < f(x) < f^2(x) < \dots$  é crescente. Como é limitada, existe  $\bar{x} = \lim f^n(x) \in (x, x_1]$ , que é ponto fixo pois

$$f(\bar{x}) = f(\lim f^n(x)) = \lim f^{n+1}(x) = \bar{x}.$$

Assim,  $\bar{x} = x_1$ . O outro caso é tratado similarmente. □

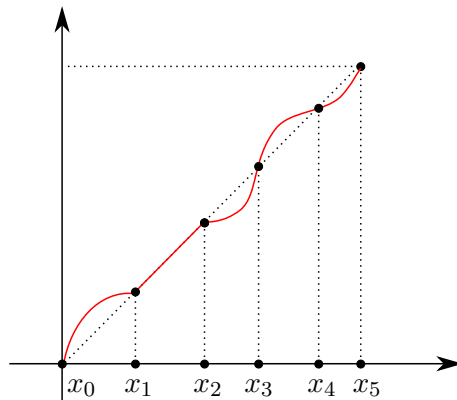
Quando  $f$  é homeomorfismo, um comportamento similar ocorre para iterados negativos:

- se  $f(x) > x$  para todo  $x \in (x_0, x_1)$ , então  $\{f^n(x)\}_{n \leq 0}$  é decrescente e converge para  $x_0$ , para todo  $x \in (x_0, x_1)$ .
- se  $f(x) < x$  para todo  $x \in (x_0, x_1)$ , então  $\{f^n(x)\}_{n \leq 0}$  é crescente e converge para  $x_1$ , para todo  $x \in (x_0, x_1)$ .

Desse modo, caracterizamos o comportamento de homeomorfismos crescentes do intervalo.

**Corolário 79** Se  $f$  é um homeomorfismo crescente de  $[0, 1]$ , então para todo  $x \notin \text{Fix}(f)$  a sequência  $\{f^n(x)\}_{n \in \mathbb{Z}}$  é monótona e  $\lim_{n \rightarrow +\infty} f^n(x)$  e  $\lim_{n \rightarrow -\infty} f^n(x)$  são pontos fixos de  $f$ . Em particular,  $\text{Per}_n(f) = \emptyset$  para todo  $n > 1$ .

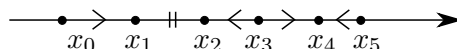
**Exemplo 80** Tome  $f : [0, 1] \rightarrow [0, 1]$  como na figura abaixo.



Identificamos  $\text{Fix}(f) = \{x_0 = 0\} \cap [x_1, x_2] \cup \{x_3, x_4, x_5 = 1\}$ . Temos:

- $x \in (x_0, x_1)$ :  $\lim_{n \rightarrow +\infty} f^n(x) = x_1$  e  $\lim_{n \rightarrow -\infty} f^n(x) = x_0$ .
- $x \in (x_2, x_3)$ :  $\lim_{n \rightarrow +\infty} f^n(x) = x_2$  e  $\lim_{n \rightarrow -\infty} f^n(x) = x_3$ .
- $x \in (x_3, x_4)$ :  $\lim_{n \rightarrow +\infty} f^n(x) = x_4$  e  $\lim_{n \rightarrow -\infty} f^n(x) = x_3$ .
- $x \in (x_4, x_5)$ :  $\lim_{n \rightarrow +\infty} f^n(x) = x_4$  e  $\lim_{n \rightarrow -\infty} f^n(x) = x_5$ .

Obtemos assim a seguinte descrição da dinâmica:



Passamos agora à caracterização de homeomorfismos decrescentes. Nesse caso, já sabemos que  $f(0) = 1$ ,  $f(1) = 0$  e que  $f$  possui um único ponto fixo (pelo TVI). Para obter informações adicionais, consideramos a composição  $g = f \circ f$ , que é um homeomorfismo crescente. Assim,  $g$  satisfaz o Corolário 79. Note que  $\text{Fix}(g) = \text{Fix}(f) \cup \text{Per}_2(f)$ . Ademais,  $\text{Per}_n(f) = \emptyset$  para todo  $n > 2$  (Exercício 35). Se  $x \notin \text{Fix}(g)$ , então existem  $x_0, x_1 \in \text{Fix}(g)$  tais que  $\lim_{n \rightarrow +\infty} g^n(x) = x_0$  e  $\lim_{n \rightarrow -\infty} g^n(x) = x_1$ . Temos dois casos para os iterados positivos de  $x$ :

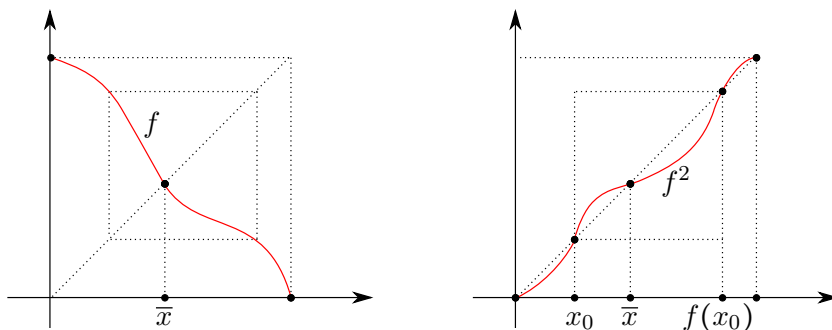
◦ Se  $x_0 \in \text{Fix}(f)$ , então  $\lim_{n \rightarrow +\infty} f^n(x) = x_0$ .

◦ Se  $x_0 \in \text{Per}_2(f)$ , então

$$\lim_{n \rightarrow +\infty} f^{2n}(x) = x_0 \quad \text{e} \quad \lim_{n \rightarrow +\infty} f^{2n+1}(x) = f(x_0).$$

O mesmo vale para os iterados negativos. Isso caracteriza os possíveis comportamentos de  $f$ .

**Exemplo 81** Considere  $f$  tal que  $f^2$  tem o gráfico como abaixo.



Identificamos  $\text{Fix}(f) = \{x_1\}$  e  $\text{Per}_2(f) = \{0, x_0, f(x_0), 1\}$ . Temos:

◦  $x \in (0, x_0)$ :

$$\begin{cases} \lim_{n \rightarrow +\infty} f^{2n}(x) = 0 \\ \lim_{n \rightarrow +\infty} f^{2n+1}(x) = 1 \end{cases} \quad \text{e} \quad \begin{cases} \lim_{n \rightarrow -\infty} f^{2n}(x) = x_0 \\ \lim_{n \rightarrow -\infty} f^{2n+1}(x) = f(x_0). \end{cases}$$

◦  $x \in (x_0, x_1)$ :

$$\lim_{n \rightarrow +\infty} f^n(x) = x_1 \quad \text{e} \quad \begin{cases} \lim_{n \rightarrow -\infty} f^{2n}(x) = x_0 \\ \lim_{n \rightarrow -\infty} f^{2n+1}(x) = f(x_0). \end{cases}$$

◦  $x \in (x_1, f(x_0))$ :

$$\lim_{n \rightarrow +\infty} f^n(x) = x_1 \quad \text{e} \quad \begin{cases} \lim_{n \rightarrow -\infty} f^{2n}(x) = f(x_0) \\ \lim_{n \rightarrow -\infty} f^{2n+1}(x) = x_0. \end{cases}$$

◦  $x \in (f(x_0), 1)$ :

$$\left\{ \begin{array}{l} \lim_{n \rightarrow +\infty} f^{2n}(x) = 1 \\ \lim_{n \rightarrow +\infty} f^{2n+1}(x) = 0 \end{array} \right. \quad \text{e} \quad \left\{ \begin{array}{l} \lim_{n \rightarrow -\infty} f^{2n}(x) = f(x_0) \\ \lim_{n \rightarrow -\infty} f^{2n+1}(x) = x_0. \end{array} \right.$$

## SEÇÃO 17

**Exercícios**

**Exercício 26** | Sejam  $(X, f)$  e  $(X, \varphi)$  dinâmicas.

(1) Se  $n = \text{Per}_f(x)$ , então  $f^m(x) = x$  se e somente se  $n$  divide  $m$ .

(2) Se  $t = \text{Per}_\varphi(x)$ , então  $\varphi_s(x) = x$  se e somente se  $s/t$  é inteiro.

**Exercício 27** | Considere  $(X, f)$ , onde  $f$  não é invertível. Nesse contexto, temos uma definição mais fraca de periodicidade. Dizemos que  $x \in X$  é *pré-periódico* se existirem  $m, n > 0$  distintos tais que  $f^m(x) = f^n(x)$ . Todo ponto periódico é pré-periódico, mas a recíproca não é necessariamente válida.

**Exercício 28** | Se  $X$  é finito e  $f : X \rightarrow X$ , então todo  $x \in X$  é pré-periódico. Se  $f$  é invertível, então todo  $x \in X$  é periódico.

**Exercício 29** | Mostre que se  $X$  é espaço topológico conexo compacto e  $f : X \rightarrow X$  é contínua, então os itens (1)–(3) da Proposição 69 são válidas, mas não necessariamente (4).

**Exercício 30** | Mostre que para  $f$  invertível e não-invertível, valem as inclusões

$$\text{Fix}(f) \subset \text{Per}(f) \subset L(f) \subset \Omega(f) \subset R(f)$$

e que todos esses conjuntos são fechados. Mostre que o mesmo vale para fluxos.

**Exercício 31** | Prove as afirmações dos Exemplos 75 e 76.

**Exercício 32** | Identifique os conjuntos  $\text{Fix}(f)$ ,  $\text{Per}(f)$ ,  $L(f)$ ,  $\Omega(f)$ ,  $R(f)$  para as rotações de  $\mathbb{S}^1$ .

**Exercício 33** | Considere o disco unitário  $\mathbb{D} = \{(r, \theta) : 0 \leq r \leq 1, \theta \in \mathbb{S}^1\}$  em coordenadas polares, e considere a transformação  $f : \mathbb{D} \rightarrow \mathbb{D}$  definida por

$$f(r, \theta) = (r, \theta + r).$$

(1) Descreva a dinâmica de  $f$  restrita a cada círculo  $r = \text{constante}$ .

(2) Quem são  $\text{Per}(f)$ ,  $L(f)$ ,  $\Omega(f)$ ,  $R(f)$ ?

**Exercício 34** | (1) Mostre que se  $f : [0, 1] \rightarrow [0, 1]$  é homeomorfismo, então  $\text{Fix}(f)$  é fechado.

(2) Mostre que se  $A \subset [0, 1]$  é fechado, então existe  $f : [0, 1] \rightarrow [0, 1]$  homeomorfismo crescente tal que  $\text{Fix}(f) = A$ .



**Exercício 35** | Mostre que se  $f : [0, 1] \rightarrow [0, 1]$  é homeomorfismo decrescente, então  $\text{Per}_n(f) = \emptyset$  para todo  $n > 2$ . (Dica: considere  $n$  par e  $n$  ímpar).

**Exercício 36** | Para cada função abaixo, descreva o comportamento de  $f$ , como feito no Exemplo 81.

(1)

$$f(x) = \begin{cases} 1 - \frac{x}{2}, & \text{se } 0 \leq x \leq \frac{1}{2} \\ \frac{3}{2}(1 - x) & \text{se } \frac{1}{2} \leq x \leq 1. \end{cases}$$

(2)

$$f(x) = \begin{cases} 1 - \frac{5}{4}x, & \text{se } 0 \leq x \leq \frac{1}{2} \\ \frac{3}{4}(1 - x) & \text{se } \frac{1}{2} \leq x \leq 1. \end{cases}$$

# *Endomorfismos do intervalo, teoremas de Li-Yorke e Sharkovsky*

Um endomorfismo do intervalo é uma transformação contínua  $f : [0, 1] \rightarrow [0, 1]$ . Ao contrário do último capítulo,  $f$  não é necessariamente injetiva nem sobrejetiva. As possíveis dinâmicas de tais transformações são muito mais variadas do que as de homeomorfismos. Em geral, não há uma descrição tão precisa da dinâmica quanto no caso de homeomorfismos. Por outro lado, podem aparecer pontos periódicos de diversos períodos (o que não ocorre para homeomorfismos). Nesse capítulo, discutiremos dois resultados clássicos sobre pontos periódicos de endomorfismos do intervalo: o teorema de Li-Yorke (1975) e o teorema de Sharkovsky (1964). Embora tenha sido provado onze anos depois, o teorema de Li-Yorke é um caso particular do teorema de Sharkovsky. À época, o planeta encontrava-se à luz da Guerra Fria, e muitos dos trabalhos feitos no oriente não eram conhecidos pelo ocidente e vice-versa. Soma-se a isso o fato de Sharkovsky ter publicado seu resultado em russo, o que dificultou à comunidade do ocidente tomar conhecimento de seus argumentos. No que segue, provaremos ambos os resultados. Para isso, introduziremos uma ferramenta muito útil para gerar pontos periódicos: os grafos de Markov de pontos periódicos.

## SEÇÃO 18

### Exemplo introdutório: família quadrática

Para cada  $0 \leq a \leq 4$ , considere o endomorfismo  $f_a : [0, 1] \rightarrow [0, 1]$  definido por  $f(x) = ax(1 - x)$ .

#### Definição 82

(Família quadrática) A *família quadrática* ou *logística* é a família de endomorfismos  $\{f_a\}_{0 \leq a \leq 4}$ . Cada  $f_a$  é chamado de *transformação quadrática* ou *logística*.

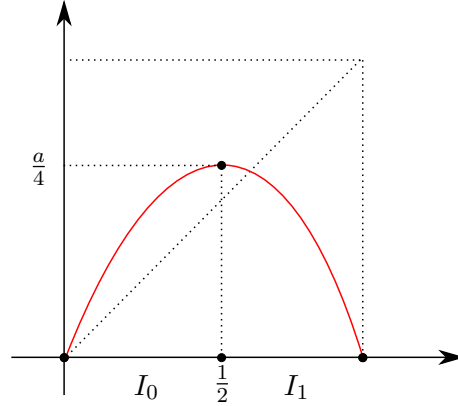
Transformações logísticas aparecem em diversas áreas do conhecimento, por exemplo em dinâmica de populações. Imagine uma colônia de bactérias que necessita de um certo substrato para sobreviver. Se há muito substrato, a colônia cresce; quando ela cresce, a quantidade de substrato se torna insuficiente para a colônia, e assim ela decresce. O crescimento da colônia depende dessa relação entre seu tamanho e o total de substrato, e pode ser modelado pela iteração  $x \mapsto ax(1 - x)$  para certos valores de  $a$ .

Abaixo listamos algumas propriedades gerais de  $f_a$ :

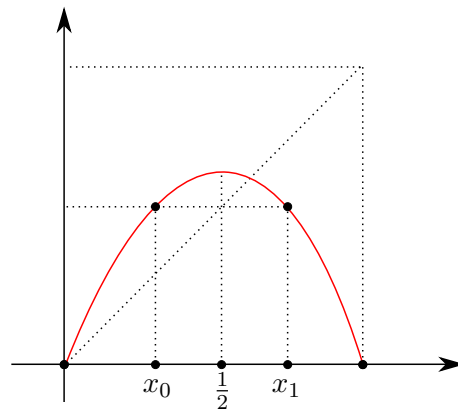
- $f(0) = f(1) = 0$ .
- $f\left(\frac{1}{2}\right) = \frac{a}{4}$ .

- $f$  é *unimodal*: existe o único valor  $x = \frac{1}{2}$  onde  $f$  é crescente em  $[0, x]$  e decrescente em  $[x, 1]$ .

Escreva  $I_0 = [0, \frac{1}{2}]$  e  $I_1 = [\frac{1}{2}, 1]$  para os intervalos de monotonicidade de  $f$ .



É essa mudança de monotonicidade que permite uma maior riqueza dinâmica da família quadrática, em comparação com homeomorfismos do intervalo. Por exemplo, se  $a > 2$  então  $f^2$  é uma transformação *multimodal*, isto é, muda de monotonicidade mais de uma vez. Para ver isso, comece notando que  $\frac{1}{2} < \frac{a}{4}$  e portanto  $\frac{1}{2} \in \text{Im}(f)$ , logo existem  $x_0, x_1 \in [0, 1]$  tais que  $f(x_0) = f(x_1) = \frac{1}{2}$ .



Defina os intervalos

$$I_{00} = [0, x_0], \quad I_{01} = [x_0, \frac{1}{2}], \quad I_{11} = [\frac{1}{2}, x_1], \quad I_{10} = [x_1, 1].$$

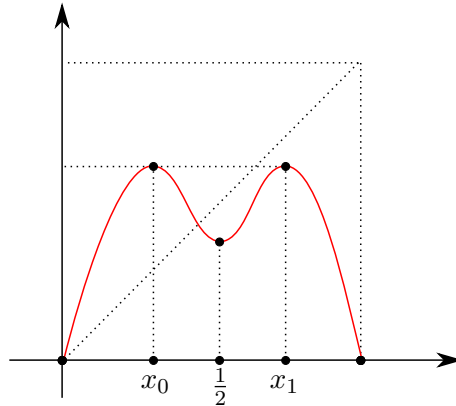
A escolha dos índices como acima reflete em qual dos intervalos  $I_0, I_1$  os pontos  $x, f(x)$  pertencem:

- $x \in I_{00} \implies x \in I_0$  e  $f(x) \in I_0$ .
- $x \in I_{01} \implies x \in I_0$  e  $f(x) \in I_1$ .
- $x \in I_{11} \implies x \in I_1$  e  $f(x) \in I_1$ .
- $x \in I_{10} \implies x \in I_1$  e  $f(x) \in I_0$ .

Desse modo, obtemos as seguintes informações sobre a monotonicidade de  $f^2$ :

- Em  $I_{00}$ ,  $f^2$  é a composição de funções crescentes, logo  $f^2 \upharpoonright_{I_{00}}$  é crescente.
- Em  $I_{01}$ ,  $f^2$  é a composição de uma função crescente e outra decrescente, logo  $f^2 \upharpoonright_{I_{01}}$  é decrescente.
- Similarmente,  $f^2 \upharpoonright_{I_{11}}$  é crescente e  $f^2 \upharpoonright_{I_{10}}$  é decrescente.

Veja a figura abaixo.



Quando  $3 < a \leq 4$ , o gráfico acima intersecta a diagonal  $y = x$  em quatro pontos  $y_1 = 0, y_2, y_3, y_4$  (Exercício 37). Dois desses pontos,  $y_1$  e  $y_2$ , são pontos fixos de  $f$ . Os outros dois são pontos de período 2. É possível continuar esse estudo, mostrando que existe  $a_k$  tal que se  $a_k < a \leq 4$  então  $f$  possui pontos de período  $2^k$ . Desse modo, ao aumentar o parâmetro novos períodos aparecem. É também possível mostrar que existe um parâmetro  $a_\infty$  tal que se  $a_\infty < a \leq 4$  então  $f$  possui pontos de todos os períodos.

## SEÇÃO 19

### Grafos de Markov e o teorema de Li-Yorke

Nessa seção, introduzimos grafos de Markov e os utilizamos para provar o seguinte teorema.

**Teorema 83** (Li-Yorke) Seja  $f : [0, 1] \rightarrow [0, 1]$  contínua. Se  $f$  possui um ponto periódico de período três, então  $f$  possui pontos periódicos de todos os períodos  $n \geq 1$ .

Esse teorema é também conhecido como “período 3 implica período de qualquer ordem”. A principal ferramenta utilizada será o lema abaixo.

**Lema 84** Seja  $f : I \rightarrow I$  contínua, onde  $I \subset \mathbb{R}$  é intervalo compacto.

- (1) Se  $J \subset I$  é um intervalo fechado tal que  $f(J) \supset J$ , então existe  $x \in J$  tal que  $f(x) = x$ .
- (2) Se  $\{I_n\}_{n \geq 0}$  é uma sequência de intervalos fechados de  $I$  tais que  $f(I_n) \supset I_{n+1}$  para todo  $n \geq 0$ , então existe uma cadeia descendente  $\{J_n\}_{n \geq 0}$  tal

que  $f^n(J_n) = I_n$  para todo  $n \geq 0$ . Em particular, existe  $x \in I_0$  tal que  $f^n(x) \in I_n$  para todo  $n \geq 0$ .

PROVA (1) Seja  $J = [a, b]$ . Então existem  $z, w \in J$  tais que  $f(z) = a$  e  $f(w) = b$ . Defina  $g : J \rightarrow \mathbb{R}$  por  $g(x) = f(x) - x$ . Temos:

- $g(z) = f(z) - z = a - z \leq 0$ .
- $g(w) = f(w) - w = b - w \geq 0$ .

Pelo TVI, existe  $x$  entre  $z$  e  $w$  tal que  $g(x) = 0$ , isto é,  $f(x) = x$ .

(2) Procedemos por indução. Começamos com o caso inicial: assumindo  $g : I \rightarrow I$  contínua com  $g(I_0) \supset I_1$ , queremos encontrar  $J_1 \subset I_0$  tal que  $g(J_1) = I_1$ . A ideia é ver a última realização do extremo esquerdo de  $I_1$  e a primeira realização do extremo direito de  $I_1$  logo após. Vejamos os detalhes. Escreva  $I_1 = [c, d]$ . Existem  $a', b' \in I_0$  tais que  $g(a') = c$  e  $g(b') = d$ . Se  $I'$  é o intervalo de extremos  $a', b'$ , então  $g(I') \supset I_1$ . Logo, podemos trocar, se necessário,  $I_0$  por  $I'$  e assumir que  $I_0 = [a, b]$  com  $\{g(a), g(b)\} = \{c, d\}$ . Defina

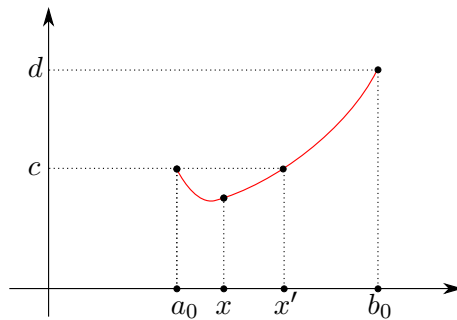
$$a_0 = \max\{x \in [a, b] : g(x) = g(a)\}$$

$$b_0 = \min\{x \in [a_0, b] : g(x) = g(b)\}.$$

Isso define o intervalo  $[a_0, b_0]$ , que afirmarmos satisfazer  $g([a_0, b_0]) = I_1 = [c, d]$ . Para ver isso, note que pelo TVI já sabemos que  $g([a_0, b_0]) \supset I_1$ . Para provar a inclusão reversa, consideramos dois casos.

CASO 1.  $g(a) = c$  e  $g(b) = d$ .

Tome  $x \in [a_0, b_0]$ . Se  $g(x) < c$ , então pelo TVI existe  $x' \in [x, b_0]$  tal que  $g(x') = c$ . Mas  $x' \geq x > a_0$ , o que contraria a maximalidade de  $a_0$ . Veja a figura abaixo.



Similarmente, se  $g(x) > d$  então existe  $x' \in [a_0, x]$  tal que  $g(x') = d$ . Nesse caso,  $x' \leq x < b_0$ , o que contraria a minimalidade de  $b_0$ .

CASO 2.  $g(a) = d$  e  $g(b) = c$ .

Tome  $x \in [a_0, b_0]$ . Se  $g(x) < c$ , então pelo TVI existe  $x' \in [a_0, x]$  tal que  $g(x') = c$ . Mas  $x' \leq x < b_0$ , o que contraria a minimalidade de  $a_0$ . Similar-

mente, se  $g(x) > d$  então existe  $x' \in [x, b_0]$  tal que  $g(x') = d$ . Nesse caso,  $x' \geq x > a_0$ , o que contraria a maximalidade de  $a_0$ .

Em qualquer dos casos,  $J_1 = [a_0, b_0]$  satisfaz  $g(J_1) = I_1$ . Para o caso geral, assumamos por indução que já construímos  $J_n \subset J_{n-1} \subset \dots \subset J_1 \subset I_0 =: J_0$  tais que  $f^k(J_k) = I_k$  para  $k = 1, \dots, n$ . Temos  $f^{n+1}(J_n) = f(f^n(J_n)) = f(I_n) \supset I_{n+1}$ . Aplicando o caso inicial para  $g = f^{n+1}$ , existe  $J_{n+1} \subset J_n$  tal que  $g(J_{n+1}) = I_{n+1}$ , isto é,  $f^{n+1}(J_{n+1}) = I_{n+1}$ . Para a parte final, tome  $x \in \bigcap_{n \geq 0} J_n$ , onde a interseção é não-vazia pelo teorema de interseção de Cantor. Claramente,  $x \in J_n$  e portanto  $f^n(x) \in f^n(J_n) = I_n$  para todo  $n \geq 0$ .  $\square$

**Corolário 85** | Seja  $f : I \rightarrow I$  contínua, e sejam  $I_0, I_1, \dots, I_{n-1}$  intervalos fechados de  $I$  tais que  $f(I_0) \supset I_1, f(I_1) \supset I_2, \dots, f(I_{n-2}) \supset I_{n-1}$  e  $f(I_{n-1}) \supset I_0$ . Então existe  $x \in I_0$  tal que  $f^n(x) = x$  e  $f^k(x) \in I_k$  para  $k = 0, \dots, n-1$ .

PROVA | Existem  $I_0 \supset J_1 \supset \dots \supset J_{n-1} \supset J_n$  tais que  $f^k(J_k) = I_k$  para  $k = 0, 1, \dots, n$  (aqui, escrevemos  $I_n = I_0$ ). Como  $f^n(J_n) = I_n = I_0 \supset J_n$ , a primeira parte do Lema 84 implica que existe  $x \in J_n$  tal que  $f^n(x) = x$ . Nesse caso,  $f^k(x) \in f^k(J_k) = I_k$  para  $k = 0, \dots, n-1$ .  $\square$

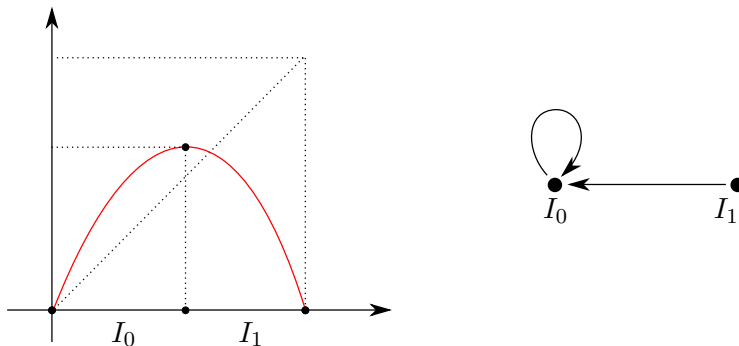
Notamos, assim, que uma sequência de intervalos encaixados por iterações de  $f$  dão origem a pontos fixos de  $f^n$ . Para melhor entender esse comportamento, introduzimos a seguinte estrutura combinatória.

**Definição 86** | (Grafo de Markov) Dados  $f : I \rightarrow I$  contínua e  $\mathcal{I} = \{I_n\}_{n \geq 0}$  uma família de intervalos fechados de  $I$ , o *grafo de Markov* associado é o grafo  $G = (V, E)$  onde  $V = \mathcal{I}$  e

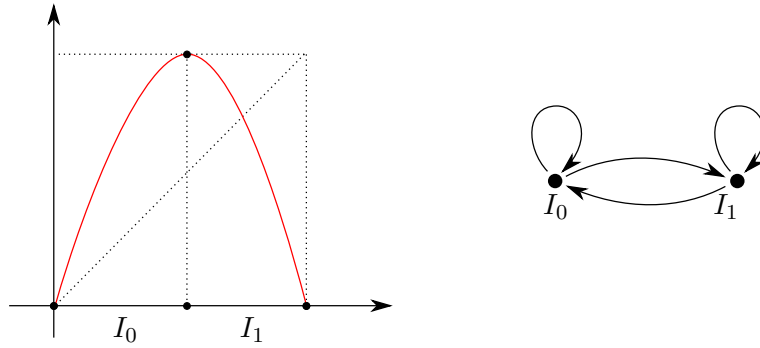
$$E = \{I_j \rightarrow I_k : f(I_j) \supset I_k\}.$$

Vamos entender esse conceito por meio de exemplos.

**Exemplo 87** | Considere  $f = f_a$  com  $2 \leq a < 4$  e  $\mathcal{I} = \{I_0 = [0, \frac{1}{2}], I_1 = [\frac{1}{2}, 1]\}$ . Então  $f(I_0) = f(I_1) = [0, \frac{a}{4}]$  contém  $I_0$  mas não contém  $I_1$ , de modo que o grafo de Markov é:



**Exemplo 88** | Considere  $f = f_4$ , e seja  $\mathcal{I}$  como acima. Então  $f(I_0) = f(I_1) = I_0 \cup I_1$ , logo o grafo de Markov é:



Nesse caso, podemos aplicar o Corolário 85 para mostrar que  $\text{Per}_2(f) \neq \emptyset$ . De fato, o ciclo  $I_0 \rightarrow I_1 \rightarrow I_0$  de tamanho 2 garante que existe  $x \in I_0$  tal que  $f^2(x) = x$  e  $f(x) \in I_1$ . Se  $x$  não tivesse período 2, então  $f(x) = x$ , e daí  $x \in I_0$  e  $x = f(x) \in I_1$ , implicando que  $x = \frac{1}{2}$ , que claramente não é fixo por  $f^2$ .

O último exemplo mostra como gerar pontos periódicos a partir de ciclos com vértices não repetidos. Focaremos nos grafos de Markov associados a pontos periódicos. Tome  $x \in \text{Per}_n(f)$ , com

$$\mathcal{O}(x) = \{x, f(x), \dots, f^{n-1}(x)\} = \{x_0 < x_1 < \dots < x_{n-1}\}.$$

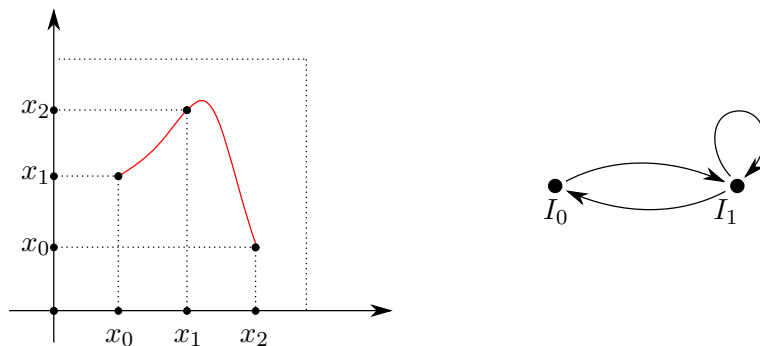
**Definição 89** (Grafo de Markov de um ponto periódico) Na notação acima, o *grafo de Markov* de  $x$  é o grafo de Markov associado à família de intervalos  $\mathcal{I} = \{[x_0, x_1], [x_1, x_2], \dots, [x_{n-2}, x_{n-1}]\}$ .

Em geral, escreveremos  $\mathcal{I} = \{I_0, \dots, I_{n-2}\}$ .

**Exemplo 90** Vamos identificar os possíveis grafos de Markov de um ponto periódico de período 3. Temos  $\mathcal{O}(x) = \{x_0 < x_1 < x_2\}$ . Sejam  $I_0 = [x_0, x_1]$  e  $I_1 = [x_1, x_2]$ . Temos dois casos.

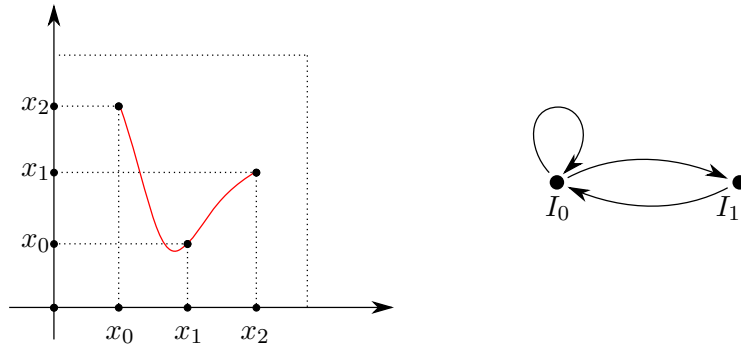
CASO 1.  $f(x_0) = x_1$ .

A combinatória em  $\mathcal{O}(x)$  por  $f$  é  $x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow x_0$ , isto é,  $f(x_1) = x_2$  e  $f(x_2) = x_0$ . Então  $f(I_0) \supset I_1$  e  $f(I_1) \supset I_0 \cup I_1$ , e o grafo de Markov é:



CASO 2.  $f(x_0) = x_2$ .

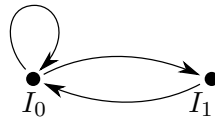
A combinatoria em  $\mathcal{O}(x)$  por  $f$  é  $x_0 \rightarrow x_2 \rightarrow x_1 \rightarrow x_0$ , isto é,  $f(x_2) = x_1$  e  $f(x_1) = x_0$ . Então  $f(I_0) \supset I_0 \cup I_1$  e  $f(I_1) \supset I_0$ , e o grafo de Markov é:



Note que os grafos acima representam a estrutura mínima de cada um dos grafos de Markov: é possível existirem mais arestas. De todo modo, a estrutura mínima já permite provar o teorema de Li-Yorke. A ideia é que eles possuem um loop e portanto ciclos de qualquer tamanho.

**Teorema 91** (Li-Yorke) Seja  $f : [0, 1] \rightarrow [0, 1]$  contínua. Se  $f$  possui um ponto periódico de período três, então  $f$  possui pontos periódicos de todos os períodos  $n \geq 1$ .

**PROVA** Podemos assumir, sem perda de generalidade, que o grafo de Markov de  $x \in \text{Per}_3(f)$  é dado pelo Caso 2 ( $f(x_0) = x_2$ ):



Para  $n = 2$ , considere o ciclo  $I_0 \rightarrow I_1 \rightarrow I_0$ . Pelo Corolário 85, existem  $x \in I_0$  tal que  $f^2(x) = x$  e  $f(x) \in I_1$ . Se  $x \in \text{Per}_2(f)$ , então  $f(x) = x$  e daí  $x \in I_0 \cap I_1 = \{x_1\}$ , que tem período 3, uma contradição. Assim,  $x \in \text{Per}_2(f)$ .

Tome agora  $n \geq 4$ . Considere o seguinte ciclo de tamanho  $n$ :

$$\underbrace{I_0 \rightarrow I_0 \rightarrow \dots \rightarrow I_0}_{n-2 \text{ arestas}} \rightarrow I_1 \rightarrow I_0.$$

Pelo Corolário 85, existe  $x \in I_0$  tal que  $f^n(x) = x$  e  $x, f(x), \dots, f^{n-2}(x) \in I_0$ ,  $f^{n-1}(x) \in I_1$ . Por contradição, assumamos que  $x \notin \text{Per}_n(f)$ . Então  $x \in \text{Per}_d(f)$  para algum  $d$  divisor próprio de  $n$ . Em particular,  $d - 1 \leq \frac{n}{2} - 1 < n - 2$ . Mas então  $y = f^{d-1}(x) = f^{n-1}(x)$  satisfaz:

- $y = f^{d-1}(x) \in I_0$ , pois  $d - 1 < n - 2$ .
- $y = f^{n-1}(x) \in I_1$ .

Logo,  $y \in I_0 \cap I_1 = \{x_1\} \implies y = x_1 \implies x = f(x_1) = x_0 \implies f(x) = x_2 \notin I_0$ , contradição. Isso conclui a prova.  $\square$



## SEÇÃO 20

**Teorema de Sharkovsky**

Nessa seção, vamos discutir o teorema de Sharkovsky. Provado em 1964, onze anos antes do teorema de Li-Yorke, esse teorema representa uma compreensão completa sobre pontos periódicos de endomorfismos do intervalo. Ele é bem mais geral do que o teorema de Li-Yorke e, em particular, também prova que período três implica período de qualquer ordem. Para enunciá-lo, introduzimos a ordem de Sharkovsky.

**Definição 92**

(Ordem de Sharkovsky) A *ordem de Sharkovsky* é a ordem total em  $\mathbb{N} = \{1, 2, \dots\}$  definida por:

$$\begin{array}{ccccccccccc} 3 & \triangleright & 5 & \triangleright & 7 & \triangleright & 9 & \triangleright & \dots & \triangleright & \\ 2 \cdot 3 & \triangleright & 2 \cdot 5 & \triangleright & 2 \cdot 7 & \triangleright & 2 \cdot 9 & \triangleright & \dots & \triangleright & \\ 4 \cdot 3 & \triangleright & 4 \cdot 5 & \triangleright & 4 \cdot 7 & \triangleright & 4 \cdot 9 & \triangleright & \dots & \triangleright & \\ & & & & \vdots & & & & & & \\ 2^n \cdot 3 & \triangleright & 2^n \cdot 5 & \triangleright & 2^n \cdot 7 & \triangleright & 2^n \cdot 9 & \triangleright & \dots & \triangleright & \\ & & & & \vdots & & & & & & \\ \dots & \triangleright & 2^n & \triangleright & 2^{n-1} & \triangleright & \dots & \triangleright & 2 & \triangleright & 1 \end{array}$$

Nessa ordem, temos por exemplo que  $5 \triangleright 62 \triangleright 404 \triangleright 1024$ , pois  $5 \triangleright 2 \cdot 31 \triangleright 4 \cdot 401 \triangleright 2^{10}$ .

**Teorema 93**

(Sharkovsky) Seja  $f : I \rightarrow I$  uma função contínua, onde  $I \subset \mathbb{R}$  é um intervalo compacto. Se  $f$  possui ponto de período  $n$ , então  $f$  possui ponto de período  $m$  para todo  $m \triangleleft n$ .

Vamos tomar  $x \in \text{Per}_n(x)$  e estudar seu grafo de Markov. No que segue, assumimos  $n > 1$ . Escreva  $\mathcal{O}(x) = \{x, f(x), \dots, f^{n-1}(x)\} = \{x_0 < x_1 < \dots < x_{n-1}\}$ , seja  $\mathcal{I} = \{[x_0, x_1], [x_1, x_2], \dots, [x_{n-2}, x_{n-1}]\}$  a família de  $n-1$  intervalos definidos por  $\mathcal{O}(x)$ , e seja  $G = (V, E)$  o grafo de Markov associado ( $V = \mathcal{I}$ ). A prova será dividida em subseções.

## SUBSEÇÃO 20.1

**Fatos gerais**

Vamos começar mostrando que  $G$  possui um loop.

**Lema 94** |  $G$  possui um loop  $I_1 = [x_i, x_{i+1}]$  com  $f(x_i) \geq x_{i+1}$  e  $f(x_{i+1}) \leq x_i$ .

## PROVA

Vamos analisar a monotonicidade de  $f$  restrita a  $\mathcal{O}(x)$ . Temos  $f(x_0) > x_0$  e  $f(x_{n-1}) < x_{n-1}$ , logo podemos tomar

$$i = \max\{0 \leq j < n-1 : f(x_j) > x_j\}.$$

Por definição,  $f(x_i) > x_i$  e daí  $f(x_i) \geq x_{i+1}$ . Por maximalidade,  $f(x_{i+1}) \leq x_{i+1}$  e, como  $n > 1$ , temos  $f(x_{i+1}) < x_{i+1} \implies f(x_{i+1}) \leq x_i$ . O lema está provado.  $\square$

Agora, vamos mostrar que desse vértice especial podemos chegar a qualquer outro vértice de  $G$ .

**Lema 95** | Seja  $I_1$  um vértice satisfazendo o Lema 94. Então, partindo de  $I_1$ , podemos chegar a qualquer outro vértice de  $G$ .

PROVA | Para cada  $k \geq 1$ , seja  $V_k = \{I \in \mathcal{I} : \exists \text{ caminho } I_1 \rightarrow \dots \rightarrow I \text{ de tamanho } k\}$ . Se  $\mathcal{C} : I_1 \rightarrow \dots \rightarrow I$  tem tamanho  $k$ , então

$$\mathcal{C}' : I_1 \rightarrow \underbrace{I_1 \rightarrow \dots \rightarrow I}_c$$

é caminho de tamanho  $k+1$ , logo  $V_k \subset V_{k+1}$ , e portanto

$$\emptyset \neq V_1 \subset V_2 \subset \dots \subset \mathcal{I}.$$

Como  $\mathcal{I}$  é finito, a cadeia ascendente acima se estabiliza: existe  $k \geq 1$  tal que  $A = V_k = V_{k+1} = \dots$ . Vamos mostrar que, nesse momento, necessariamente  $A = \mathcal{I}$ . Isso concluirá a prova do lema. A ideia é olhar para o conjunto  $U$  igual à união dos intervalos de  $A$ , e mostrar que  $U = [x_0, x_{n-1}]$ . Temos  $U \supset I_1$ . Seja  $C$  a componente conexa de  $U$  que contém  $I_1$ , digamos  $C = [x_j, x_k]$  com  $j \leq i < k$ . Vamos usar  $x_i$  como âncora para mostrar dois fatos.

FATO 1.  $f(x_i) \in C$ .

Caso contrário,  $f(I_1)$  contém o intervalo  $J = [x_{j-1}, x_j]$  ou  $[x_k, x_{k+1}]$ , de modo que  $J \in V_{k+1} = A$ , o que contraria a maximalidade de  $C$  como componente conexa.

FATO 2.  $f(\{x_j, \dots, x_k\}) \subset C$ .

Assuma que não. Temos dois casos:

- Se existe  $i < \ell \leq k$  tal  $f(x_\ell) \notin C$ , tome o menor tal  $\ell$ . Então  $f(x_{\ell-1}, x_\ell)$  é um intervalos que intersecta  $C$  e seu complemento, e portanto intersecta  $[x_{j-1}, x_j]$  ou  $[x_k, x_{k+1}]$ . Isso contraria a maximalidade de  $C$ .
- Se existe  $j \leq \ell < i$  tal  $f(x_\ell) \notin C$ , tome o maior tal  $\ell$ , e proceda como no caso anterior.

Portanto,  $\{x_j, \dots, x_k\} \subset \mathcal{O}(x)$  é  $f$ -invariante. O único subconjunto não-vazio e invariante de  $\mathcal{O}(x)$  é o próprio  $\mathcal{O}(x)$ , assim  $\{x_j, \dots, x_k\} = \mathcal{O}(x)$ , o que prova que  $A = \mathcal{I}$ .  $\square$

Alguns livros fornecem uma prova incompleta do lema acima. Eles provam o fato enunciado no Exercício 41, de que o conjunto dos extremos dos intervalos de  $A$  é igual a  $\mathcal{O}(x)$ . Isso não necessariamente implica que  $A = \mathcal{I}$  (tome, por exemplo,  $\mathcal{I} \setminus \{[x_1, x_2]\}$ ).

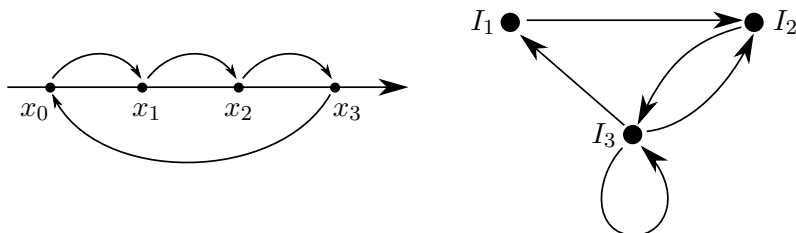
## SUBSEÇÃO 20.2

### Órbitas de período $2^k$

Para adquirir intuição, vamos começar o estudo analisando todas as possíveis combinatórias das órbitas de período 4. No que segue, vamos escrever  $x \rightarrow$  quando  $f(x) = y$ . Sejam  $I_1 = [x_0, x_1]$ ,  $I_2 = [x_1, x_2]$  e  $I_3 = [x_2, x_3]$ .

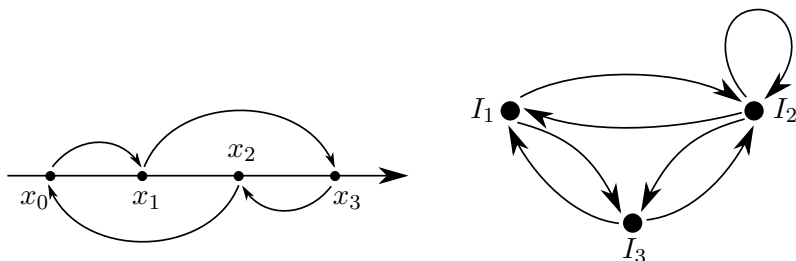
CASO 1.  $x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_0$ .

A combinatória da permutação é  $\begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 0 \end{pmatrix}$ . O grafo de Markov é:



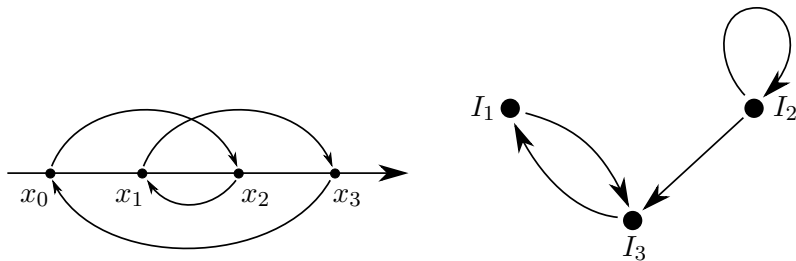
CASO 2.  $x_0 \rightarrow x_1 \rightarrow x_3 \rightarrow x_2 \rightarrow x_0$ .

A combinatória da permutação é  $\begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 3 & 0 & 2 \end{pmatrix}$ . O grafo de Markov é:



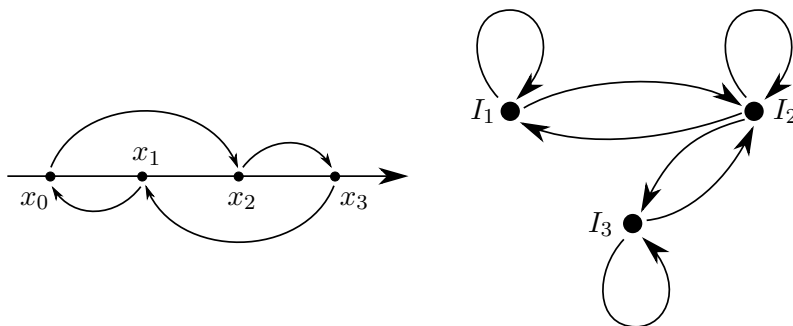
CASO 3.  $x_0 \rightarrow x_2 \rightarrow x_1 \rightarrow x_3 \rightarrow x_0$ .

A combinatória da permutação é  $\begin{pmatrix} 0 & 1 & 2 & 3 \\ 2 & 3 & 1 & 0 \end{pmatrix}$ . O grafo de Markov é:



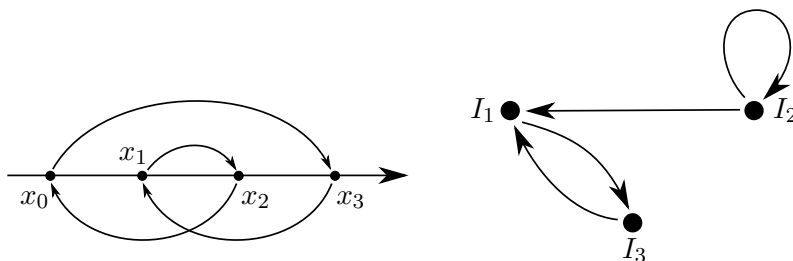
CASO 4.  $x_0 \rightarrow x_2 \rightarrow x_3 \rightarrow x_1 \rightarrow x_0$ .

A combinatória da permutação é  $\begin{pmatrix} 0 & 1 & 2 & 3 \\ 2 & 0 & 3 & 1 \end{pmatrix}$ . O grafo de Markov é:



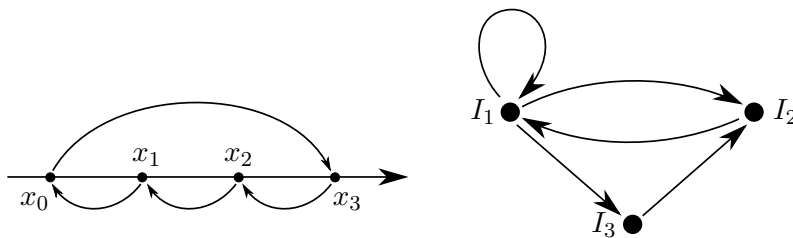
CASO 5.  $x_0 \rightarrow x_3 \rightarrow x_1 \rightarrow x_2 \rightarrow x_0$ .

A combinatória da permutação é  $\begin{pmatrix} 0 & 1 & 2 & 3 \\ 3 & 2 & 0 & 1 \end{pmatrix}$ . O grafo de Markov é:



CASO 6.  $x_0 \rightarrow x_3 \rightarrow x_2 \rightarrow x_1 \rightarrow x_0$ .

A combinatória da permutação é  $\begin{pmatrix} 0 & 1 & 2 & 3 \\ 3 & 0 & 1 & 2 \end{pmatrix}$ . O grafo de Markov é:



Note que os casos 1 e 6 são simétricos por meio de uma reflexão, e o mesmo ocorre com os casos 3 e 5. Mas os casos 2 e 4 *não* são simétricos. Desse modo, temos na prática 4 possibilidades distintas para a combinatória. Note que, em qualquer dos casos, há um ciclo de tamanho 2, e portanto existe  $y$  tal que  $f^2(y) = y$ . Necessariamente  $y \in \text{Per}_2(f)$ , pois caso contrário  $y \in I_j \cap I_k$  e daí  $y = x_i$  tem período 4. Em suma,  $\text{Per}_4(f) \neq \emptyset \implies \text{Per}_2(f) \neq \emptyset$ . O lema abaixo generaliza essa implicação.

**Lema 96** (Períodos da forma  $2^k$ )

- (1) Se  $\text{Per}_4(f) \neq \emptyset$ , então  $\text{Per}_2(f) \neq \emptyset$ .
- (2) Se  $\text{Per}_{2^{k+1}}(f) \neq \emptyset$ , então  $\text{Per}_{2^k}(f) \neq \emptyset$ .

PROVA | A primeira parte foi provada acima. Vamos provar a parte (2). Tome  $x \in \text{Per}_{2^{k+1}}(f)$ . Escrevendo  $g = f^{2^{k-1}}$ , temos  $x \in \text{Per}_{2^{k+1}}(f) = \text{Per}_4(g)$ , onde na última igualdade usamos o Exercício 42(2). Pela parte (1) e novamente pelo Exercício 42(2), existe  $y \in \text{Per}_2(g) = \text{Per}_{2^k}(f)$ , e portanto  $\text{Per}_{2^k}(f) \neq \emptyset$ .  $\square$

O lema acima prova o teorema de Sharkovsky quando estamos na última linha da ordem de Sharkovsky. Quando esse não for o caso, procedemos a uma análise mais refinada, conforme a próxima subseção.

### SUBSEÇÃO 20.3

#### Órbitas de período ímpar

Relembramos que pontos periódicos são garantidos pela existência de ciclos no grafo de Markov. Pelo Lema 95, temos vários candidatos a ciclos. Escrevendo novamente  $I_1 = [x_i, x_{i+1}]$  para o vértice do Lema 94, resta provar a existência de arestas  $I_k \rightarrow I_1$  para algum  $k \neq 1$ . Quando  $n$  é ímpar, tais arestas são garantidas pelo próximo resultado.

**Lema 97** | Na notação acima, se não existe nenhuma aresta  $I_k \rightarrow I_1$  com  $k \neq 1$ , então  $f$  é uma bijeção entre  $\{x_0, \dots, x_i\}$  e  $\{x_{i+1}, \dots, x_{n-1}\}$ . Em particular,  $n$  é par e  $\text{Per}_2(f) \neq \emptyset$ .

PROVA | Escreva  $E = \{x_0, \dots, x_i\}$  para os vértices à esquerda e  $D = \{x_{i+1}, \dots, x_{n-1}\}$  para os vértices à direita. Por absurdo, assuma que  $f$  leva algum elemento de  $E$  em  $E$ . Seja  $j = \max\{1 \leq \ell \leq i : f(x_j) \leq x_i\}$  o maior índice com essa propriedade. Como  $f(x_i) \geq x_{i+1}$ , temos  $j < i$ . Pela maximalidade,  $f(x_{j+1}) \geq x_{i+1}$ . Então  $f[x_j, x_{j+1}] \supset I_1 \implies [x_j, x_{j+1}] \rightarrow I_1$ , absurdo. Assim,  $f(x_j) \geq x_{i+1}$  para todo  $j = 1, \dots, i$ . Similarmente,  $f(x_j) \leq x_i$  para todo  $j = i+1, \dots, n-1$ . Isso prova a primeira parte do lema, que por sua vez implica que  $n$  é par. Por fim, note que

$$\begin{cases} f[x_0, x_i] \supset [x_{i+1}, x_{n-1}] \\ f[x_{i+1}, x_{n-1}] \supset [x_0, x_i]. \end{cases}$$

Logo, existe  $x \in [x_0, x_i]$  tal que  $f^2(x) = x$  e  $f(x) \in [x_{i+1}, x_{n-1}]$ . É claro que  $x \in \text{Per}_2(f)$ .  $\square$

Ao longo do restante dessa subseção, assumimos que  $n$  é ímpar. Dos lemas 94, 95 e 97, temos as seguintes propriedades do grafo de Markov:

- Existe  $I_1 \rightarrow I_1$ , com  $I_1 = [x_i, x_{i+1}]$  tal que  $f(x_i) \geq x_{i+1}$  e  $f(x_{i+1}) \leq x_i$ .
- Para todo  $I_j$ , existe um caminho de  $I_1$  a  $I_j$ .
- Existe  $I_j \rightarrow I_1$  para algum  $j \neq 1$ .

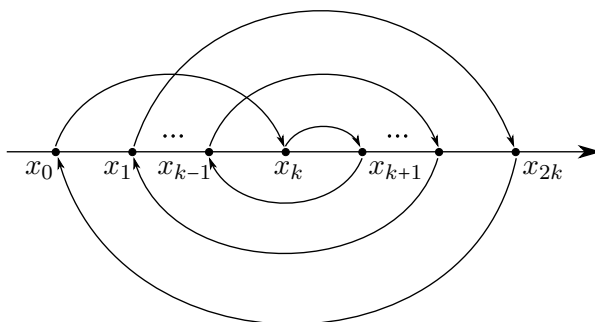
Vamos agora introduzir uma combinatória particular. Escreva  $n = 2k + 1$  e  $\mathcal{O}(x) = \{x_0, \dots, x_{2k}\}$ .

**Definição 98** | (Configuração do tipo Štefan) Dizemos que  $x \in \text{Per}_n(f)$  tem *configuração do tipo Štefan* se a combinatória de  $f$  ao longo de  $\mathcal{O}(x)$  for uma espiral que parte

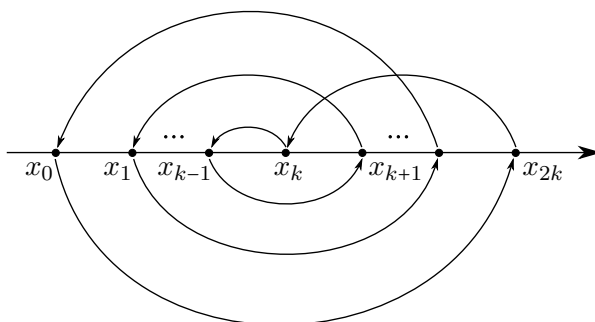
do ponto central  $x_k$  para fora, alternando entre os lados esquerdo e direito de  $x_k$ .

Noutras palavras, há duas possibilidades para a configuração do tipo Štefan:

◦  $x_k \rightarrow x_{k+1} \rightarrow x_{k-1} \rightarrow \dots \rightarrow x_1 \rightarrow x_{2k} \rightarrow x_0 \rightarrow x_k$ :



◦  $x_k \rightarrow x_{k-1} \rightarrow x_{k+1} \rightarrow \dots \rightarrow x_{2k-1} \rightarrow x_0 \rightarrow x_{2k} \rightarrow x_k$ :



É claro que as duas configurações são simétricas por uma reflexão.

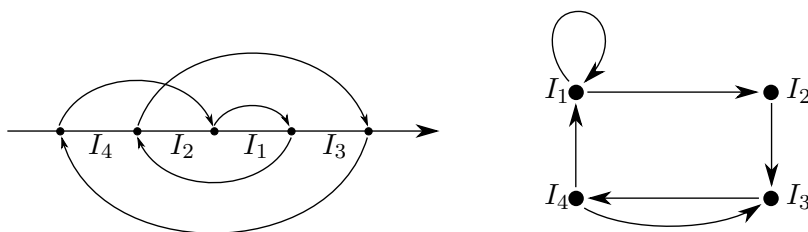
**Exemplo 99**

Vamos exibir os grafos de Markov de configurações do tipo Štefan para  $n = 5$ .

Se  $x_2 \rightarrow x_3 \rightarrow x_1 \rightarrow x_4 \rightarrow x_0 \rightarrow x_2$ , escreva

$$I_1 = [x_2, x_3], I_2 = [x_1, x_2], I_3 = [x_3, x_4], I_4 = [x_0, x_1].$$

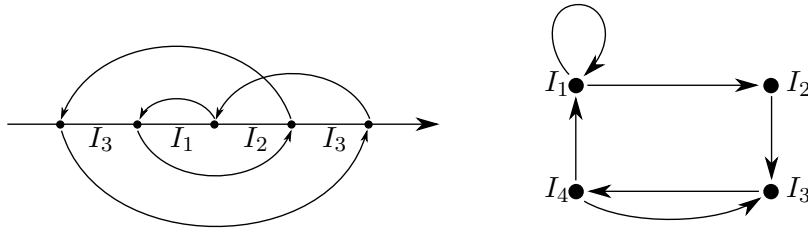
A combinatória é  $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 & 0 \end{pmatrix}$  e o grafo de Markov é:



Se  $x_2 \rightarrow x_1 \rightarrow x_3 \rightarrow x_0 \rightarrow x_4 \rightarrow x_2$ , escreva

$$I_1 = [x_1, x_2], I_2 = [x_2, x_3], I_3 = [x_0, x_1], I_4 = [x_3, x_4].$$

A combinatória é  $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 0 & 2 \end{pmatrix}$  e o grafo de Markov é:



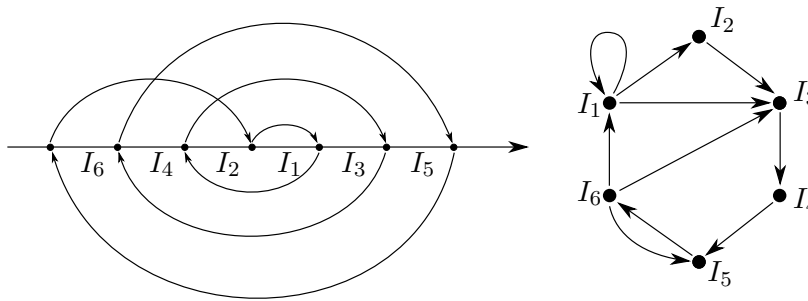
Note que, com nossa notação de enumerar os intervalos de acordo com a espiral, os grafos de Markov coincidem.

**Exemplo 100**

Vamos exibir os grafos de Markov de configurações do tipo Štefan para  $n = 7$ . Se  $x_3 \rightarrow x_4 \rightarrow x_2 \rightarrow x_5 \rightarrow x_1 \rightarrow x_6 \rightarrow x_0 \rightarrow x_3$ , escreva

$$I_1 = [x_3, x_4], I_2 = [x_2, x_3], I_3 = [x_4, x_5], I_4 = [x_1, x_2], I_5 = [x_5, x_6], I_6 = [x_0, x_1].$$

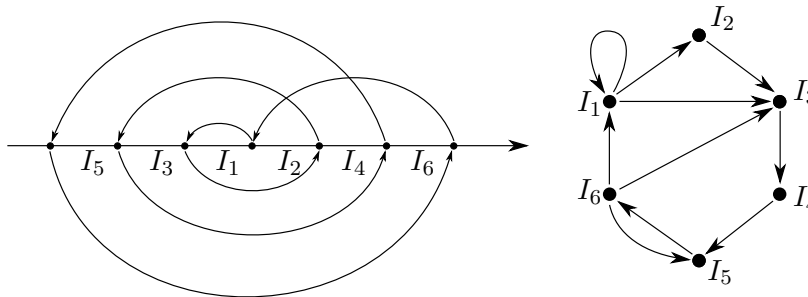
A combinatória é  $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 5 & 4 & 2 & 1 & 0 \end{pmatrix}$  e o grafo de Markov é:



Se  $x_3 \rightarrow x_2 \rightarrow x_4 \rightarrow x_1 \rightarrow x_5 \rightarrow x_0 \rightarrow x_6 \rightarrow x_3$ , escreva

$$I_1 = [x_2, x_3], I_2 = [x_3, x_4], I_3 = [x_1, x_2], I_4 = [x_4, x_5], I_5 = [x_0, x_1], I_6 = [x_5, x_6].$$

A combinatória é  $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 2 & 1 & 0 & 3 \end{pmatrix}$  e o grafo de Markov é:



Novamente, os grafos de Markov coincidem.

Mais geralmente, se  $x \in \text{Per}_n(f)$  tem configuração do tipo Štefan, então o grafo de Markov possui as seguintes arestas:

- Loop  $I_1 \rightarrow I_1$ .
- Ciclo hamiltoniano  $I_1 \rightarrow I_2 \rightarrow \dots \rightarrow I_{2k} \rightarrow I_1$ .
- Arestas  $I_{2k} \rightarrow I_1, I_3, \dots, I_{2k-1}$ .

Tal grafo origina diversos pontos periódicos, conforme o próximo lema.

**Lema 101** Se  $x \in \text{Per}_n(f)$  tem configuração do tipo Štefan, então  $f$  possui ponto periódico de qualquer período  $\geq n$  e qualquer  $m < n$  par (i.e.  $m = 2, 4, \dots, n-1$ ).

PROVA Se  $m \geq n$ , considere o ciclo

$$\mathcal{C} : \underbrace{I_1 \rightarrow I_1 \rightarrow \dots \rightarrow I_1}_{m-n+1 \text{ arestas}} \rightarrow \underbrace{I_2 \rightarrow \dots \rightarrow I_{2k} \rightarrow I_1}_{n-1 \text{ arestas}}.$$

Existe  $y \in I_1$  tal que  $f^m(y) = y$  que segue o itinerário de  $\mathcal{C}$ . É de fácil verificação mostrar que  $y \in \text{Per}_m(f)$  (Exercício 43).

Se  $m < n$  é par, considere o ciclo de tamanho  $m$

$$\mathcal{C}' : I_{2k} \rightarrow I_{2k+1-m} \rightarrow I_{2k+1-m+1} \rightarrow \dots \rightarrow I_{2k-1} \rightarrow I_{2k}.$$

Novamente, existe  $y \in I_{2k}$  tal que  $f^m(y) = y$  que segue o itinerário de  $\mathcal{C}'$ , e novamente  $y \in \text{Per}_m(f)$  (Exercício 43).  $\square$

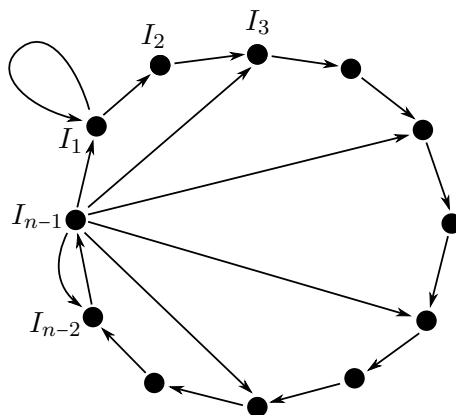
Até agora, introduzimos uma configuração especial, sem saber que ela existe. O próximo resultado mostra que se  $f$  possui ponto periódico de período ímpar, então o menor de tais períodos define uma configuração do tipo Štefan.

**Lema 102** (Menor período ímpar é Štefan) Se  $x \in \text{Per}_n(f)$ , onde  $n$  é o menor período ímpar maior que 1 de  $f$ , então  $x$  tem configuração do tipo Štefan. Logo, existem:

- Loop  $I_1 \rightarrow I_1$ .
- Ciclo hamiltoniano  $I_1 \rightarrow I_2 \rightarrow \dots \rightarrow I_{n-1} \rightarrow I_1$ .
- Arestas  $I_{n-1} \rightarrow I_1, I_3, \dots, I_{n-2}$ .

Adicionalmente, não existe nenhuma aresta  $I_j \rightarrow I_\ell$  com  $\ell - j > 1$  nem  $I_j \rightarrow I_1$  com  $j \neq 1, n-1$ .





PROVA Relembramos que  $I_1 = [x_i, x_{i+1}]$  é o intervalo dado pelo Lema 94. Pelos lemas 95 e 97, existe um ciclo não-trivial partindo de  $I_1$ . Seja  $\mathcal{C}$  tal ciclo de tamanho mínimo, digamos  $\mathcal{C} : I_1 \rightarrow I_2 \rightarrow \dots \rightarrow I_k \rightarrow I_1$ , com  $k \neq 1$ . Como  $k$  é mínimo, os vértices de  $\mathcal{C}$  são distintos. Afirmamos que  $\mathcal{C}$  é hamiltoniano, i.e.  $k = n - 1$ . Mostraremos isso por contradição. Assuma que  $k < n - 1$ . Temos dois casos:

- $k$  é ímpar:  $\mathcal{C}$  gera um ponto  $y \in \text{Per}_k(f)$ , que contraria a minimalidade de  $n$ .
- $k$  é par: então  $k \leq n - 3 \implies k + 1 \leq n - 2$ . O ciclo  $\mathcal{C} \rightarrow I_1$  tem tamanho  $k + 1$  e portanto gera um ponto  $y$  tal que  $f^{k+1}(y) = y$ . Logo,  $y \in \text{Per}_d(f)$  para algum divisor próprio  $d$  de  $k + 1$ . Mas então  $d$  é ímpar e  $1 < d \leq k + 1 < n$ , novamente contrariando a minimalidade de  $n$ .

Assim,  $k = n - 1$  com  $\mathcal{C} : I_1 \rightarrow I_2 \rightarrow \dots \rightarrow I_{n-1} \rightarrow I_1$ . Ainda pela minimalidade, não existem arestas  $I_j \rightarrow I_\ell$  com  $\ell - j > 1$  nem  $I_j \rightarrow I_1$  com  $j \neq 1, n - 1$ . Ademais, o grau de saída de  $I_1$  é 2 e de  $I_2, \dots, I_{n-1}$  é 1 (Exercício 44). Resta provar que  $x$  é do tipo Štefan. Temos  $f(x_i) \geq x_{i+1}$  e  $f(x_{i+1}) \leq x_i$ . Como  $x$  não possui período 2, alguma das desigualdades é estrita. Assumindo que  $f(x_{i+1}) < x_i$ , vamos mostrar que a espiral se inicia com  $f(x_i) = x_{i+1}$ .

FATO 1.  $f(x_i) = x_{i+1}$  e  $f(x_{i+1}) = x_{i-1}$ .

PROVA (Fato 1) Temos  $f(x_{i+1}) \leq x_{i-1}$ . Se  $f(x_i) \geq x_{i+2}$ , então  $f(I_1) \supset [x_{i-1}, x_{i+2}]$  e portanto o grau de saída de  $I_1$  é pelo menos 3, absurdo. Assim,  $f(x_i) = x_{i+1}$ . Similarmente, se  $f(x_{i+1}) \leq x_{i-2}$ , então  $f(I_1) \supset [x_{i-2}, x_{i+1}]$ , novamente um absurdo. Assim,  $f(x_{i+1}) = x_{i-1}$ . Em particular,  $I_2 = [x_{i-1}, x_i]$ .  $\square$

FATO 2.  $f(x_{i-1}) = x_{i+2}$ .

PROVA (Fato 2) O grau de saída de  $I_2$  é 1. Se  $f(x_{i-1}) \leq x_{i-1}$  ou  $f(x_{i-1}) > x_{i+2}$ , então  $f(I_2)$  contém pelo menos dois intervalos, absurdo. Assim,  $f(x_{i-1}) = x_{i+2}$ .  $\square$

Indutivamente, utilizando que os graus de saída de  $I_2, \dots, I_{n-1}$  são 1, podemos concluir que  $x$  tem configuração do tipo Štefan.  $\square$

**Corolário 103** Se  $f$  possui ponto de período ímpar  $n > 1$ , então  $f$  possui pontos periódicos de qualquer período  $m \geq n$  e  $m < n$  par (i.e.  $m = 2, 4, \dots, n - 1$ ).

Esse resultado estende o Lema 101, sem requerer que  $n$  seja o período de uma configuração do tipo Štefan.

PROVA Seja  $n_0$  o menor período ímpar. Notando que

$$\{m : m \geq n \text{ ou } m < n \text{ par}\} \subset \{m : m \geq n \text{ ou } m < n_0 \text{ par}\},$$

o resultado segue pelos Lemas 101 e 102.  $\square$

## SUBSEÇÃO 20.4

### Conclusão da prova do teorema de Sharkovsky

Relembre a ordem de Sharkovsky:

$$\begin{array}{cccccccc} 3 & \triangleright & 5 & \triangleright & 7 & \triangleright & 9 & \triangleright \dots \triangleright \\ 2 \cdot 3 & \triangleright & 2 \cdot 5 & \triangleright & 2 \cdot 7 & \triangleright & 2 \cdot 9 & \triangleright \dots \triangleright \\ 4 \cdot 3 & \triangleright & 4 \cdot 5 & \triangleright & 4 \cdot 7 & \triangleright & 4 \cdot 9 & \triangleright \dots \triangleright \\ & & & & \vdots & & & \\ 2^n \cdot 3 & \triangleright & 2^n \cdot 5 & \triangleright & 2^n \cdot 7 & \triangleright & 2^n \cdot 9 & \triangleright \dots \triangleright \\ & & & & \vdots & & & \\ \dots & \triangleright & 2^n & \triangleright & 2^{n-1} & \triangleright & \dots & \triangleright 2 \triangleright 1 \end{array}$$

**Teorema 104** (Sharkovsky) Seja  $f : I \rightarrow I$  uma função contínua, onde  $I \subset \mathbb{R}$  é um intervalo compacto. Se  $f$  possui ponto de período  $n$ , então  $f$  possui ponto de período  $m$  para todo  $m \triangleleft n$ .

PROVA Se  $n = 2^b$ , o teorema segue do Lema 96. Assuma que  $n = a \cdot 2^b$ , onde  $a > 1$  é ímpar. Tome  $m \triangleleft n$ . Temos dois casos.

CASO 1.  $m$  não está na mesma linha de  $n$ .

Temos  $m = c \cdot 2^b$ , onde  $c > 1$  é par. Retraduzindo as informações para a função  $f^{2^b}$ , temos

$$\emptyset \neq \text{Per}_n(f) = \text{Per}_{a \cdot 2^b}(f) \subset \text{Per}_a(f^{2^b})$$

e portanto  $\text{Per}_a(f^{2^b}) \neq \emptyset$ . Pelo Corolário 103, obtemos  $\text{Per}_c(f^{2^b}) \neq \emptyset$ . Pelo Exercício 42(2),

$$\emptyset \neq \text{Per}_c(f^{2^b}) = \text{Per}_{c \cdot 2^b}(f)$$

e assim  $\text{Per}_m(f) \neq \emptyset$ .

CASO 2.  $m$  pertence à mesma linha de  $n$ .

Temos  $m = c \cdot 2^b$ , onde  $c > a$  é ímpar. Temos

$$\emptyset \neq \text{Per}_n(f) = \text{Per}_{a \cdot 2^b}(f) \subset \text{Per}_a(f^{2^b}).$$

Como  $c > a$ , o Corolário 103 implica que  $\text{Per}_c(f^{2^b}) \neq \emptyset$ . Pelo Exercício 42(3),

$$\emptyset \neq \text{Per}_c(f^{2^b}) \subset \text{Per}_c(f) \cup \text{Per}_{c \cdot 2}(f) \cup \dots \cup \text{Per}_{c \cdot 2^b}(f),$$

logo existe  $\ell \in \{0, 1, \dots, b\}$  tal que  $\text{Per}_{c \cdot 2^\ell}(f) \neq \emptyset$ . Se  $\ell = b$ , acabou. Se  $\ell < b$ , então  $n' = c \cdot 2^\ell$  é um número em uma linha anterior à linha de  $m$  e  $n$ . Como  $\text{Per}_{n'}(f) \neq \emptyset$ , o Caso 1 garante que  $\text{Per}_m(f) \neq \emptyset$ .  $\square$

## SEÇÃO 21

**Exercícios**

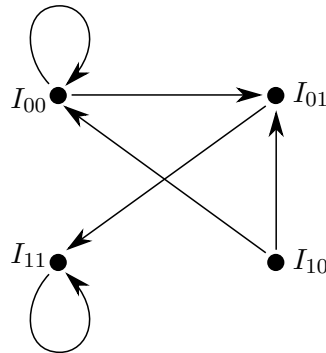
**Exercício 37** | Mostre que se  $3 < a \leq 4$  então a transformação quadrática  $f_a(x) = ax(1-x)$  possui pontos periódicos de período 2.

**Exercício 38** | Seja  $f : [0, 1] \rightarrow [0, 1]$  dada por

$$f(x) = \begin{cases} 2x & , x \in [0, \frac{1}{2}] \\ 2 - 2x & , x \in [\frac{1}{2}, 1]. \end{cases}$$

Descreva o grafo de Markov associado aos intervalos  $[0, \frac{1}{2}]$  e  $[\frac{1}{2}, 1]$ .

**Exercício 39** | Tome  $f = f_a$  com  $1 + \sqrt{5} \leq a \leq 4$ , e sejam  $I_{00}, I_{01}, I_{11}, I_{10}$  como na página 63. Mostre que o grafo de Markov associado é:



**Exercício 40** | Seja  $f : [0, 1] \rightarrow [0, 1]$  dada por  $f(x) = 4x(1-x)$ . Notando que  $\text{Per}_2(f) = \{x_0 < x_1\}$ , faça os seguintes itens.

- (1) Descreva o grafo de Markov associado aos intervalos  $[0, x_0], [x_0, \frac{1}{2}], [\frac{1}{2}, x_1]$  e  $[x_1, 1]$ .
- (2) Identifique um ciclo de tamanho 5 no grafo que passa pelo vértice  $[x_0, \frac{1}{2}]$ , e mostre que existe  $x \in [x_0, \frac{1}{2}]$  de período 5.

**Exercício 41** | Na notação do Lema 95, seja  $B$  o conjunto dos extremos dos intervalos em  $A = V_k = V_{k+1} = \dots$ . Siga os itens abaixo para mostrar que  $B = \mathcal{O}(x)$ .

- (1) É suficiente provar que  $B$  é  $f$ -invariante. Melhor ainda: como  $f$  é uma bijeção em  $\mathcal{O}(x)$ , basta mostrar que  $f(B) \subset B$ .
- (2) Para provar (1), fixe  $I = [x_j, x_{j+1}] \in A$ . Escrevendo  $f(x_j) = x_\ell$ , algum

dos casos abaixo ocorre

$$\begin{cases} f(I) \supset [x_{\ell-1}, x_{\ell}] \implies I \rightarrow [x_{\ell-1}, x_{\ell}] \\ f(I) \supset [x_{\ell}, x_{\ell+1}] \implies I \rightarrow [x_{\ell}, x_{\ell+1}]. \end{cases}$$

- (3) Assumindo a primeira inclusão, segue que  $[x_{\ell-1}, x_{\ell}] \in V_{k+1} = A \implies x_{\ell} \in B$ , i.e.  $f(x_j) \in B$ . Similarmente,  $f(x_{j+1}) \in B$ .

Relembramos que a invariância de  $B$  não implica diretamente que  $A = \mathcal{I}$ .

**Exercício 42** Prove os seguintes itens.

- (1) Para  $m, n \geq 1$ , temos  $\text{Per}_{mn}(f) \subset \text{Per}_m(f^n)$ .
- (2) Se  $n$  é par, então  $\text{Per}_{2n}(f) = \text{Per}_n(f^2)$ . Mais geralmente,  $\text{Per}_{2^k n}(f) = \text{Per}_n(f^{2^k})$ .
- (3) Se  $n$  é ímpar, então  $\text{Per}_{2n}(f) \subset \text{Per}_n(f^2) \subset \text{Per}_n(f) \cup \text{Per}_{2n}(f)$ . Mais geralmente,

$$\text{Per}_{2^k n}(f) \subset \text{Per}_n(f^{2^k}) \subset \text{Per}_n(f) \cup \text{Per}_{2n}(f) \cup \dots \cup \text{Per}_{2^k n}(f).$$

- (4) Em geral, se  $p$  é primo, então valem os seguintes fatos:
  - (a) Se  $p$  divide  $n$ , então  $\text{Per}_{p^k n}(f) = \text{Per}_n(f^{p^k})$ .
  - (b) Se  $p$  não divide  $n$ , então

$$\text{Per}_{p^k n}(f) \subset \text{Per}_n(f^{p^k}) \subset \text{Per}_n(f) \cup \text{Per}_{pn}(f) \cup \dots \cup \text{Per}_{p^k n}(f).$$

**Exercício 43** Na notação da prova do Lema 101, mostre que  $y \in \text{Per}_m(f)$  em ambos os casos.

**Exercício 44** Na notação da prova do Lema 102, prove que o grau de saída de  $I_1$  é 2 e de  $I_2, \dots, I_{n-1}$  é 1.

# Transformações no círculo

Nesse curto capítulo, vamos fazer um breve estudo da vasta área de transformações no círculo  $\mathbb{S}^1$ . Já vimos exemplos de tais transformações, dadas pelas rotações. Agora vamos estudar transformações mais gerais. Para isso, utilizaremos a interpretação de  $\mathbb{S}^1$  dada no Apêndice (Seção 37), e relacionaremos as transformações em  $\mathbb{S}^1$  com seus levantamentos em  $\mathbb{R}$ .

## SEÇÃO 22

### Projeção de transformações

Relembramos que  $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z} = \mathbb{R}/\sim$ , onde  $x \sim y$  se e somente se  $x - y \in \mathbb{Z}$ . Seja  $\pi : \mathbb{R} \rightarrow \mathbb{S}^1$  a projeção canônica da relação de equivalência,  $\pi(x) = [x]$ . Vamos relacionar transformações em  $\mathbb{S}^1$  com transformações em  $\mathbb{R}$ . Trataremos a projeção de transformações em  $\mathbb{R}$  e o levantamento de transformações em  $\mathbb{S}^1$ . Começemos com a projeção.

Dada  $F : \mathbb{R} \rightarrow \mathbb{R}$ , queremos saber quando  $F$  induz uma transformação  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  de modo canônico, i.e. definida pela igualdade

$$f([x]) = [F(x)].$$

Necessariamente, se  $x \sim y$  então  $F(x) \sim F(y)$ . Essa é de fato condição suficiente.

#### Definição 105

(Projeção de  $F$ ) Se  $F : \mathbb{R} \rightarrow \mathbb{R}$  satisfaz a implicação  $x \sim y \implies F(x) \sim F(y)$ , então  $F$  induz uma transformação  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  definida por  $f([x]) = [F(x)]$ ,  $\forall [x] \in \mathbb{S}^1$ .

Nesse caso, o diagrama abaixo comuta:

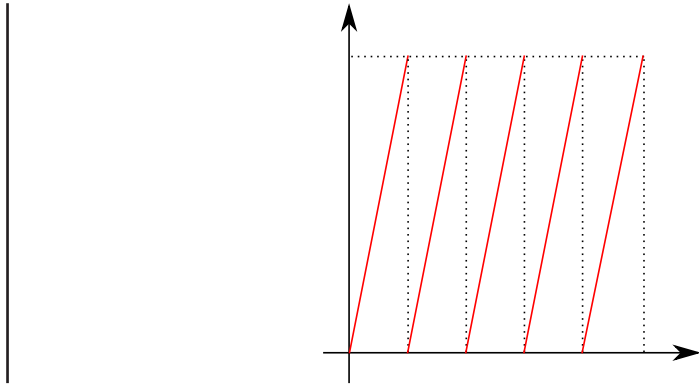
$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{F} & \mathbb{R} \\ \pi \downarrow & & \downarrow \pi \\ \mathbb{S}^1 & \xrightarrow{f} & \mathbb{S}^1 \end{array}$$

#### Exemplo 106

Dado  $d \in \mathbb{Z}$ , seja  $F : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $F(x) = dx$ . Se  $x = y + n$  com  $n \in \mathbb{Z}$ , então

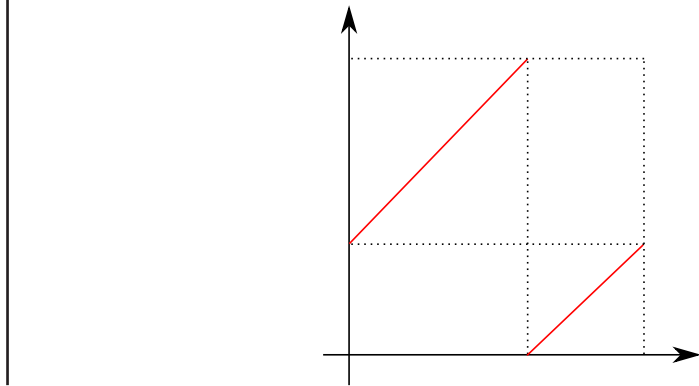
$$F(x) = dx = d(y + n) = dy + dn = F(y) + \underbrace{dn}_{\in \mathbb{Z}}$$

e portanto  $F(x) \sim F(y)$ . Assim,  $F$  induz  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ , conhecida como  $dx \pmod{1}$ . A representação de  $f$  no domínio fundamental  $[0, 1)$  é:



**Exemplo 107** | Tome agora  $\beta \notin \mathbb{Z}$ , e considere  $F(x) = \beta x$ . Temos  $0 \sim 1$  mas  $F(1) - F(0) = \beta \notin \mathbb{Z}$ . Logo,  $F$  não se projeta em  $\mathbb{S}^1$ .

**Exemplo 108** | Revisitamos o Exemplo 68. Tome  $\beta \in \mathbb{R}$ , e considere  $F : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $F(x) = x + \beta$ . Se  $x = y + n$ , então  $F(x) = x + \beta = F(y) + n$ , logo  $F$  induz  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ , conhecida como a rotação em  $\mathbb{S}^1$  de ângulo  $\beta$ . A representação de  $f$  no domínio fundamental  $[0, 1)$  é:



## SEÇÃO 23

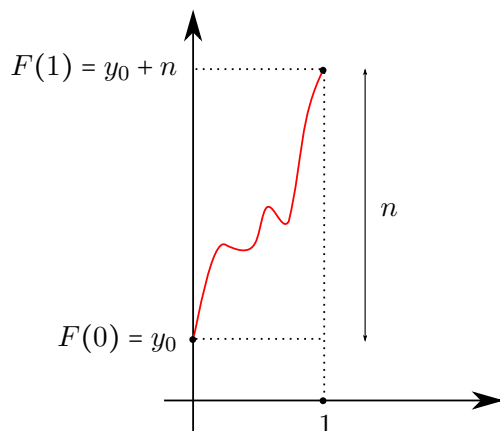
### Levantamento de transformações

Agora, dada  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  queremos encontrar  $F : \mathbb{R} \rightarrow \mathbb{R}$  tal que  $\pi \circ F = f \circ \pi$ . Assumindo que  $f$  é contínua, vamos descrever como levantar  $f$ .

- Tome  $y_0 \in \mathbb{R}$  tal que  $f([0]) = [y_0]$ , e defina  $F(0) = y_0$ .
- Se  $\varepsilon > 0$  é pequeno, então para  $|x| < \varepsilon$  existe único  $y \in (y_0 - \frac{1}{4}, y_0 + \frac{1}{4})$  tal que  $f([x]) = [y]$ . Defina  $F(x) = y$ .
- De fato, pela continuidade uniforme de  $f$ , podemos tomar  $\varepsilon > 0$  independente do ponto base: se definimos  $F(x_0) = z_0$ , então para  $|x - x_0| < \varepsilon$  existe único  $y \in (z_0 - \frac{1}{4}, z_0 + \frac{1}{4})$  tal que  $f([x]) = [y]$ . Isso define  $F$  em  $(x_0 - \varepsilon, x_0 + \varepsilon)$ .
- Indutivamente, definimos  $F$  em toda a reta  $\mathbb{R}$ .

De agora em diante, consideramos levantamentos da maneira acima. Pela construção,  $F$  é contínua. Há liberdade na escolha de  $y_0$ . De fato, se  $F$  é levantamento e  $n \in \mathbb{Z}$ , então  $F + n$  também é.

Como  $0 \sim 1$ , temos  $F(0) \sim F(1)$  e portanto  $F(1) - F(0) = n$  para algum  $n \in \mathbb{Z}$ .



Mais geralmente:  $x \sim x + 1$  e portanto  $F(x + 1) - F(x) \in \mathbb{Z}$ . Isso define  $G : \mathbb{R} \rightarrow \mathbb{Z}$  contínua por  $G(x) = F(x + 1) - F(x)$ . Logo,  $G$  é constante, i.e.  $F(x + 1) = F(x) + n$  para todo  $x \in \mathbb{R}$ . O valor  $n$  representa, para o levantamento  $F$ , o *número de voltas* que  $f$  dá em  $\mathbb{S}^1$  a partir de qualquer ponto.

**Definição 109** (Grau de  $f$ ) O grau de  $f$  é definido por  $F(x + 1) - F(x) \in \mathbb{Z}$ , que não depende de  $x \in \mathbb{R}$  nem do levantamento contínuo  $F$ .

A definição acima requer justificativa. Já mostramos que de fato  $F(x + 1) - F(x)$  não depende de  $x$ . Resta mostrar que essa diferença não depende de  $F$ . Sejam  $F_1, F_2$  levantamentos (contínuos) de  $F$ . Como  $F_1(x) \sim F_2(x)$ , temos a função contínua  $F_1 - F_2 : \mathbb{R} \rightarrow \mathbb{Z}$ , que deve ser constante, digamos  $F_1 - F_2 \equiv n$  para algum  $n \in \mathbb{Z}$ . Daí,  $F_1(x + 1) - F_1(x) = F_2(x + 1) - F_2(x)$ .

**Exemplo 110** O grau de  $dx \pmod{1}$  é  $d$ , pois  $F(1) - F(0) = d - 0 = d$ . Cada rotação tem grau 1, pois  $F(x) = x + \beta$  satisfaz  $F(1) - F(0) = (1 + \beta) - \beta = 1$ .

É fácil provar que se  $f$  é homeomorfismo, então o grau de  $f$  é igual a 1 ou  $-1$  (Exercício 45). Quando for 1, dizemos que  $f$  preserva orientação; quando for  $-1$ ,  $f$  reverte orientação.

## SEÇÃO 24

### rotações irracionais

Tome  $\beta \in \mathbb{R} \setminus \mathbb{Q}$  e seja  $f = R_\beta : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  a rotação de ângulo  $\beta$ . Já sabemos que  $f$  é um homeomorfismo que preserva orientação (grau 1). Vimos também no Exemplo 68, como consequência do Lema de Kronecker, que  $f$  é *minimal*, i.e. a órbita de todo ponto é densa em  $\mathbb{S}^1$ . O Exercício 47 fornece outra prova desse fato, utilizando o princípio da casa dos pombos.

Nessa seção, vamos mostrar uma propriedade mais forte: toda órbita de  $f$  é equidistribuída em  $\mathbb{S}^1$ . A equidistribuição diz que a frequência de visita de toda órbita a todo intervalo é igualitária (nenhum ponto nem intervalo é preferido). Começamos introduzindo os conceitos necessários.

**Definição 111** (Frequência) Dado um intervalo  $I \subset \mathbb{S}^1$ , para  $x \in \mathbb{S}^1$  e  $n > 0$  seja

$$\tau_n(I, x) := \frac{\#\{0 \leq i < n : f^i(x) \in I\}}{n}$$

a fração de visitas a  $I$  dos iterados de  $x$  até a ordem  $n$ . A *frequência* de visitas de  $x$  a  $I$  por  $f$  é

$$\tau(I, x) := \lim_{n \rightarrow +\infty} \tau_n(I, x),$$

quando o limite existir.

Em geral, a existência do limite é garantida por teoremas ergódicos (teoremas de Birkhoff e de von Neumann). Por exemplo, o teorema de Birkhoff estabelece que se  $f$  preserva uma medida de probabilidade, então  $\tau(I, x)$  existe para *quase todo ponto* com respeito à medida. A rotação irracional tem uma propriedade especial:  $\tau(I, x)$  existe para *todo ponto*. Assumimos isso no que segue, e calcularemos  $\tau(I, x)$ .

**Definição 112** (Equidistribuição) Dizemos que  $f$  é *equidistribuída* se para todo intervalo  $I \subset \mathbb{S}^1$  e todo  $x \in \mathbb{S}^1$  vale que

$$\tau(I, x) = |I|.$$

Para rotações, temos  $\tau(I, x) = \tau(I - x, 0)$  onde  $I - x = \{a - x : a \in I\}$ , logo para provar a equidistribuição da rotação irracional é suficiente tomar  $x = 0$ . Escreva  $\tau(I) = \tau(I, 0)$ . O próximo lema fornece as propriedades básicas de  $\tau(I)$ .

**Lema 113** As seguintes propriedades valem.

(1)  $\tau(\emptyset) = 0$  e  $\tau(\mathbb{S}^1) = 1$ .

(2) ADITIVIDADE: Se  $I, J \subset \mathbb{S}^1$  são disjuntos então

$$\tau(I \cup J) = \tau(I) + \tau(J).$$

(3) INVARIÂNCIA:  $\tau(I + \beta) = \tau(I)$ . Mais geralmente,  $\tau(I + n\beta) = \tau(I)$  para todo  $n \in \mathbb{Z}$ .

A prova do lema é deixada como exercício (Exercício 48). Vamos agora provar a equidistribuição.

**Teorema 114** (Rotação irracional é equidistribuída) A rotação irracional  $f$  é equidistribuída, i.e. para todo intervalo  $I \subset \mathbb{S}^1$  vale que

$$\tau(I) = |I|.$$



PROVA | Afirmamos que é suficiente provar a desigualdade  $\tau(I) \leq |I|$ . De fato, se isso ocorrer então

$$\begin{cases} \tau(I) \leq |I| \\ \tau(I^c) \leq |I^c| \end{cases} \implies 1 = \tau(\mathbb{S}^1) = \tau(I) + \tau(I^c) \leq |I| + |I^c| = 1,$$

de modo que  $\tau(I) = |I|$ . Vamos provar a desigualdade  $\tau(I) \leq |I|$  em três passos. Sem perda de generalidade, assumimos que  $I$  é fechado, digamos  $I = [a, b]$ .

FATO 1. Se  $|I| < \frac{1}{n}$  então  $\tau(I) \leq \frac{1}{n}$ .

PROVA | (Fato 1) Se  $J$  é intervalo de comprimento  $> |I|$ , então existe  $k > 0$  tal que  $f^k(I) \subset J$ . De fato, se  $J = [c, d]$ , basta tomar  $k > 0$  tal que  $f^k(a) \in \left[ c, c + \frac{|J|-|I|}{2} \right]$ , que existe porque a órbita de  $a$  é densa em  $\mathbb{S}^1$ . Divida  $\mathbb{S}^1$  em  $n$  intervalos  $J_1, \dots, J_n$  de comprimento  $\frac{1}{n}$ . Pelo exposto acima, para cada  $1 \leq i \leq n$  existe  $k_i$  tal que  $f^{k_i}(I) \subset J_i$ . Em particular,  $\{f^{k_i}(J_i) : 1 \leq i \leq n\}$  são disjuntos, e daí

$$\sum_{i=1}^n \tau(f^{k_i}(I)) \leq 1 \implies n\tau(I) \leq 1 \implies \tau(I) \leq \frac{1}{n},$$

onde na segunda passagem utilizamos o Lema 113(3). □

FATO 2. Se  $|I| < \frac{m}{n}$  então  $\tau(I) \leq \frac{m}{n}$ .

PROVA | (Fato 2) Para provar isso, divida  $I = I_1 \cup \dots \cup I_m$  com  $|I_j| < \frac{1}{n}$  e aplique o Fato 1:

$$\tau(I) = \tau(I_1) + \dots + \tau(I_m) \leq \frac{1}{n} + \dots + \frac{1}{n} = \frac{m}{n}.$$

□

FATO 3.  $\tau(I) \leq |I|$ .

PROVA | (Fato 3) Tome racionais  $\frac{m_k}{n_k}$  que convergem para  $|I|$  pela direita. Temos

$$|I| < \frac{m_k}{n_k} \implies \tau(I) \leq \frac{m_k}{n_k} \implies \tau(I) \leq \lim_{k \rightarrow +\infty} \frac{m_k}{n_k} = |I|,$$

onde na primeira passagem utilizamos o Fato 2. □

□

Há um modo analítico de provar o Teorema 114, utilizando o *critério de Weyl* (veja o Exercício 49). Note que a órbita de 0 pela rotação irracional é igual a  $\{p(n) \pmod{1} : n \geq 0\}$  para o polinômio  $p(X) = \beta X \in \mathbb{R}[X]$ . Nessa roupagem, o Teorema 114 possui a seguinte generalização.

**Teorema 115** (Weyl) Se  $p(X) = a_d X^d + \dots + a_1 X \in \mathbb{R}[X]$  é polinômio de grau  $d \geq 1$  tal que algum  $a_i \notin \mathbb{Q}$ , então o conjunto

$$\{p(n) \pmod{1} : n \geq 0\}$$

é equidistribuído em  $\mathbb{S}^1$ .

Como exemplos, obtemos que se  $\beta \notin \mathbb{Q}$  então os conjuntos  $\{n^2\beta : n \geq 0\}$  e  $\{n^{2021}\pi : n \geq 0\}$  são equidistribuídos em  $\mathbb{S}^1$ .

SEÇÃO 25

**Número de rotação**

Seja  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  homeomorfismo. De agora em diante, assumimos que  $f$  preserva orientação (caso contrário, tomamos  $f^2$ ). Vamos introduzir uma “rotação média” de  $f$ .

**Definição 116** (Número de rotação) O *número de rotação* de  $f$  é igual a

$$\rho(f) = \lim_{n \rightarrow +\infty} \frac{F^n(x) - x}{n} \pmod{1}$$

para algum (qualquer)  $x \in \mathbb{R}$ , onde  $F : \mathbb{R} \rightarrow \mathbb{R}$  é algum (qualquer) levantamento contínuo de  $f$ .

A definição acima requer explicação de que o limite não depende de  $F$  nem de  $x \in \mathbb{R}$ . Antes de proceder à boa definição, vamos introduzir exemplos. Relembre que se  $F$  é levantamento de  $f$ , então  $F(x+1) = F(x) + 1$  para todo  $x \in \mathbb{R}$ . Usando que  $F^q$  é levantamento de  $f^q$ , obtemos que  $F^q(x+1) = F^q(x) + 1$  para todo  $q \in \mathbb{Z}$  e todo  $x \in \mathbb{R}$ .

**Exemplo 117** Se  $f$  é a rotação de ângulo  $\beta$ , então  $\rho(f) = \beta$ . De fato,  $F(x) = x + \beta$  é levantamento e daí

$$\rho(f) = \lim_{n \rightarrow +\infty} \frac{F^n(x) - x}{n} = \lim_{n \rightarrow +\infty} \frac{(x + n\beta) - x}{n} = \beta.$$

**Exemplo 118** Suponha que  $f$  possua um ponto período de período  $q$ , digamos  $f^q(x_0) = x_0$ . Se  $[x] = x_0$ , existe  $p > 0$  inteiro tal que

$$F^q(x) = x + p,$$

pois  $[F^q(x)] = f^q(x_0) = x_0 = [x]$ . Então

$$F^{2q}(x) = F^q(x + p) = F^q(x) + p = x + 2p,$$

e por indução obtemos que  $F^{kq}(x) = x + kp$ . Assim, o número de rotação de  $f$  é igual a

$$\rho(f) = \lim_{k \rightarrow +\infty} \frac{F^{kq}(x) - x}{kq} = \lim_{k \rightarrow +\infty} \frac{kp}{kq} = \frac{p}{q}.$$

Observe que o denominador é o período e  $p$  é o número de voltas que a órbita de  $x_0$  faz ao redor de  $\mathbb{S}^1$  antes de fechar o período.

Provamos acima que se  $\text{Per}(f) \neq \emptyset$  então  $\rho(f) \in \mathbb{Q}$ . A recíproca é verdadeira, conforme provaremos mais adiante. Antes, vamos provar que o número de

rotação está bem-definido.

**Teorema 119** (Poincaré) O número de rotação  $\rho(f)$  está bem-definido.

PROVA Mostraremos o teorema em três partes.

FATO 1.  $\rho(f)$  não depende de  $F$ .

PROVA (Fato 1) Sejam  $F, G : \mathbb{R} \rightarrow \mathbb{R}$  levantamentos de  $f$ . Então existe  $p \in \mathbb{Z}$  tal que  $G(x) = F(x) + p$ . Por outro lado, temos

$$\begin{cases} F(x+1) = F(x) + 1 \\ G(x+1) = G(x) + 1 \end{cases} \implies \begin{cases} F(x+p) = F(x) + p \\ G(x+p) = G(x) + p. \end{cases}$$

Assim,

$$G^2(x) = G(G(x)) = G(F(x) + p) = G(F(x)) + p = F(F(x)) + 2p = F^2(x) + 2p.$$

Por indução, temos  $G^n(x) = F^n(x) + np$  para todo  $n \in \mathbb{Z}$ , e portanto

$$\frac{G^n(x) - x}{n} = \frac{F^n(x) - x}{n} + p,$$

provando que  $\rho(x) \in \mathbb{S}^1$  não depende do levantamento.  $\square$

FATO 2.  $\rho(f)$  não depende de  $x \in \mathbb{S}^1$ .

PROVA (Fato 2) Como  $f$  tem grau 1, se  $x, y \in \mathbb{R}$  são tais que  $|x - y| \leq 1$  então  $|F(x) - F(y)| \leq 1$ . Por indução, segue que  $|F^n(x) - F^n(y)| \leq 1$ . Assim,

$$\left| \frac{F^n(x) - x}{n} - \frac{G^n(x) - x}{n} \right| \leq \frac{|F^n(x) - F^n(y)|}{n} + \frac{|x - y|}{n} \leq \frac{2}{n}.$$

Assim, os limites calculados em  $x$  e  $y$  coincidem. Note também que

$$\frac{F^n(x+1) - (x+1)}{n} = \frac{F^n(x) - x}{n} + 1 - \frac{1}{n}$$

e portanto, módulo 1, os limites calculados em  $x+1$  e  $x$  também coincidem.

Isso prova que  $\rho(f)$  não depende de  $x \in \mathbb{R}$ .  $\square$

Procedemos para a última parte, de que o limite existe. Para isso, fixamos um levantamento  $F$  e  $x = 0$ . Vamos mostrar que a sequência  $\left\{ \frac{F^n(0)}{n} \right\}_{n \geq 1}$  é de Cauchy. Defina  $G_n(x) = F^n(x) - x$ .

FATO 3.  $|G_n(x) - G_n(y)| \leq 2$ , para todos  $x, y \in \mathbb{R}$ .

PROVA (Fato 3) Como já usado anteriormente,  $F^n(x+1) = F^n(x) + 1$  (pois  $F^n$  é levantamento de  $f^n$ , que tem grau 1). Assim,  $G_n(x+1) = G_n(x)$ , i.e.  $G_n$  é 1-periódica. Logo, podemos assumir que  $x, y \in \mathbb{R}$ . Daí,

$$|G_n(x) - G_n(y)| \leq \underbrace{|F^n(x) - F^n(y)|}_{\leq 1} + \underbrace{|x - y|}_{\leq 1} \leq 2.$$

□

Pelo Fato 3, se  $x, y \in \mathbb{R}$  então

$$G_n(x) - 2 \leq G_n(y) \leq G_n(x) + 2 \implies F^n(x) - x - 2 \leq F^n(y) - y \leq F^n(x) - x + 2.$$

Tomando  $x = 0$  e  $y = F^{(k-1)n}(0)$ , temos

$$F^n(0) - 2 \leq F^{kn}(0) - F^{(k-1)n}(0) \leq F^n(0) + 2$$

para todo  $k \geq 1$ . Somando em  $k = 1, \dots, m$ , temos

$$m[F^n(0) - 2] \leq F^{mn}(0) - \underbrace{F^0(0)}_{=0} \leq m[F^n(0) + 2]$$

e portanto

$$\frac{F^n(0)}{n} - \frac{2}{n} \leq \frac{F^{mn}(0)}{mn} \leq \frac{F^n(0)}{0} + \frac{2}{n} \implies \frac{F^{mn}(0)}{mn} \in \left[ \frac{F^n(0)}{n} - \frac{2}{n}, \frac{F^n(0)}{n} + \frac{2}{n} \right].$$

Trocando  $m, n$  de lugar, temos também que  $\frac{F^{mn}(0)}{mn} \in \left[ \frac{F^n(0)}{n} - \frac{2}{n}, \frac{F^n(0)}{n} + \frac{2}{n} \right]$ .

Mas então

$$\begin{aligned} & \left| \frac{F^n(0)}{n} - \frac{F^m(0)}{m} \right| \\ & \leq \left| \left[ \frac{F^n(0)}{n} - \frac{2}{n} \right] - \frac{F^{mn}(0)}{mn} \right| + \left| \left[ \frac{F^m(0)}{m} - \frac{2}{m} \right] - \frac{F^{mn}(0)}{mn} \right| + \frac{2}{n} + \frac{2}{m} \\ & \leq \frac{4}{n} + \frac{4}{m}. \end{aligned}$$

Portanto,  $\left\{ \frac{F^n(0)}{n} \right\}_{n \geq 1}$  é de Cauchy. □

Dentre as importâncias do número de rotação, temos a invariância por conjugações.

**Definição 120**

(Conjugação de homeomorfismos em  $\mathbb{S}^1$ ) Dizemos que dois homeomorfismos  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  e  $g : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  são *conjugados* se existe  $h : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  homeomorfismo que comuta o diagrama abaixo, i.e.  $h \circ g = f \circ h$ .

$$\begin{array}{ccc} \mathbb{S}^1 & \xrightarrow{g} & \mathbb{S}^1 \\ h \downarrow & & \downarrow h \\ \mathbb{S}^1 & \xrightarrow{f} & \mathbb{S}^1 \end{array}$$

A transformação  $h$  é chamada de conjugação. Nesse caso, temos  $g = h^{-1} \circ f \circ h$  e portanto  $g^n = h^{-1} \circ f^n \circ h$  para todo  $n \in \mathbb{Z}$ .

**Lema 121** | Na notação acima, se  $f, g$  são conjugados então  $\rho(f) = \rho(g)$ .

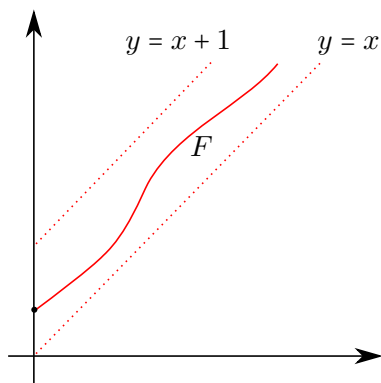
Deixamos a prova e outras consequências da conjugação nos Exercícios 50 e 51. Uma consequência do Lema 121 é que rotações distintas não são conjugadas.

Finalizamos essa seção concluindo a discussão sobre números de rotação racionais.

**Proposição 122** | O número de rotação de  $f$  é racional se e somente se  $\text{Per}(f) \neq \emptyset$ .

PROVA | A implicação ( $\Leftarrow$ ) foi provada no Exemplo 118. Provaremos a implicação ( $\Rightarrow$ ). Mostraremos que  $\text{Fix}(f) = \emptyset$  se e somente se  $\rho(f) \neq 0$ . Uma vez provado isso, segue que  $\text{Fix}(f^q) = \emptyset$  se e somente se  $\rho(f) \neq \frac{p}{q}$ . De fato, como  $\rho(f^q) = q\rho(f)$  (Exercício 52), temos que  $\rho(f) \neq \frac{p}{q} \iff \rho(f^q) \neq 0 \iff \text{Fix}(f^q) = \emptyset$ . Isso completará a prova: se  $\text{Fix}(f^q) = \emptyset$  para todo  $q \in \mathbb{Z}$  então  $\rho(f) \neq \frac{p}{q}$  para todo  $q \in \mathbb{Z}$ , ou seja,  $\rho(f)$  é irracional.

Tome  $F$  levantamento de  $f$  com  $F(0) \in (0, 1)$ . Como  $\text{Fix}(f) = \emptyset$ , o gráfico de  $F$  não intersecta as retas  $y = x$  e  $y = x + 1$ , veja a figura abaixo.



Logo,  $0 < F(x) - x < 1$  para todo  $x \in \mathbb{R}$ . Como  $F(x) - x$  é 1-periódica, existe  $\delta > 0$  tal que  $\delta < F(x) - x < 1 - \delta$  para todo  $x \in \mathbb{R}$ . Assim,

$$\begin{cases} \delta < F(x) - x < 1 - \delta \\ \delta < F^2(x) - F(x) < 1 - \delta \\ \vdots \\ \delta < F^n(x) - F^{n-1}(x) < 1 - \delta \end{cases}$$

Somando, obtemos que  $n\delta < F^n(x) - x < n(1 - \delta)$  e portanto

$$\delta < \frac{F^n(x) - x}{n} < 1 - \delta \implies \delta < \rho(f) < 1 - \delta,$$

o que completa a prova. □

## SEÇÃO 26

### Exercícios

**Exercício 45** | Mostre que se  $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  é um homeomorfismo, então o grau de  $f$  é 1 ou  $-1$ .

**Exercício 46** | Mostre que se  $f, g : \mathbb{S}^1 \rightarrow \mathbb{S}^1$  são contínuas, então o grau de  $f \circ g$  é igual ao produto dos graus de  $f$  e  $g$ .

**Exercício 47** | Seja  $f$  a rotação irracional de ângulo  $\beta$ . Vamos mostrar que a órbita positiva  $\mathcal{O}_+(x) = \{f^n(x) : n \geq 0\}$  de todo  $x \in \mathbb{S}^1$  é densa em  $\mathbb{S}^1$ .

(1) Mostre que é suficiente provar o afirmado para  $x = 0$ , i.e. que  $\{n\beta : n \geq 0\}$  é denso em  $\mathbb{S}^1$ .

(2) Mostre que se para todo  $\varepsilon > 0$  existe  $n_0 \geq 0$  tal que  $f^{n_0}(0) \in (0, \varepsilon) \subset \mathbb{S}^1$  então vale (1).

(3) Dado  $\varepsilon > 0$ , vamos mostrar que existe  $n_0$  satisfazendo (2). Tome  $N = \lceil \frac{1}{\varepsilon} \rceil$  e considere o conjunto  $\{0, f(0), \dots, f^N(0)\}$ . Pelo princípio da casa dos pombos, existem  $0 \leq i < j \leq N$  tais que  $0 < |f^i(0) - f^j(0)| < \varepsilon$ . Conclua o resultado.

**Exercício 48** | Prove o Lema 113.

**Exercício 49** | O *critério de Weyl* afirma que a sequência  $\{x_n\}_{n \geq 1} \subset \mathbb{S}^1$  é equidistribuída se e somente se

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i m x_n} = 0$$

para todo  $m \in \mathbb{Z} \setminus \{0\}$ . Utilizando o critério de Weyl, mostre que a rotação irracional é equidistribuída.

**Exercício 50** | Prove o Lema 121.

**Exercício 51** | Mostre que se  $f, g$  são conjugação  $h$  tal que  $h \circ f = f \circ h$ , então  $h(\text{Per}_n(g)) = \text{Per}_n(f)$  para todo  $n \geq 1$ .

**Exercício 52** | Prove o afirmado na Proposição 122:  $\rho(f^q) = q\rho(f)$ .

# Shifts topológicos de Markov

Nesse capítulo, introduziremos um sistema dinâmico de extrema importância na dinâmica e na teoria ergódica. Ele constitui um exemplo intrínseco que, surpreendentemente, modela diversos sistemas dinâmicos que aparecem em várias áreas do conhecimento e em aplicações. São os *shifts topológicos de Markov*.

## SEÇÃO 27

### Espaços simbólicos

Tome um conjunto finito  $\mathcal{A}$  com  $|\mathcal{A}| = N$ . Chamamos  $\mathcal{A}$  de alfabeto. No que segue, consideramos  $\mathcal{A}$  fixo.

#### Definição 123

(Espaços simbólicos) O *espaço simbólico unilateral* com alfabeto  $\mathcal{A}$  é definido por

$$\Sigma^+ = \Sigma_{\mathcal{A}}^+ = \mathcal{A}^{\{0,1,\dots\}} = \{\underline{x} = (x_0, x_1, x_2, \dots) : x_i \in \mathcal{A}, \forall i \geq 0\}.$$

O *espaço simbólico bilateral* com alfabeto  $\mathcal{A}$  é definido por

$$\Sigma = \Sigma_{\mathcal{A}} = \mathcal{A}^{\mathbb{Z}} = \{\underline{x} = (\dots, x_{-1}; x_0, x_1, \dots) : x_i \in \mathcal{A}, \forall i \in \mathbb{Z}\}.$$

Note que, no caso bilateral, separamos as posições  $-1$  e  $0$  por um ponto e vírgula. Um exemplo clássico na literatura é quando  $\mathcal{A} = \{0, 1, \dots, N-1\}$ .

Existem diversas maneiras de introduzir métricas em  $\Sigma^+$  e  $\Sigma$ . Consideramos uma classe delas, que compara os maiores intervalos centrados na origem em que duas sequências coincidem. Dados  $\underline{x}, \underline{y} \in \Sigma^+$ , defina  $N(\underline{x}, \underline{y}) \in \mathbb{N} \cup \infty$  por

$$N(\underline{x}, \underline{y}) = \max\{n \geq 0 : x_i = y_i, 0 \leq i < n\} = \min\{n \geq 0 : x_n \neq y_n\}.$$

Similarmente, dados  $\underline{x}, \underline{y} \in \Sigma$ , defina

$$N(\underline{x}, \underline{y}) = \max\{n \geq 0 : x_i = y_i, -n < i < n\} = \min\{|n| : x_n \neq y_n\}.$$

Tome  $\rho > 1$ .

#### Definição 124

(Métricas  $d_{\rho}^+$  e  $d_{\rho}$ ) Defina a métrica  $d_{\rho}^+ : \Sigma^+ \times \Sigma^+ \rightarrow [0, \infty)$  por

$$d_{\rho}^+(\underline{x}, \underline{y}) = \rho^{-N(\underline{x}, \underline{y})}.$$

Similarmente, defina a métrica  $d_{\rho} : \Sigma \times \Sigma \rightarrow [0, \infty)$  por

$$d_{\rho}(\underline{x}, \underline{y}) = \rho^{-N(\underline{x}, \underline{y})}.$$

**Lema 125** | Para todo  $\rho > 1$ ,  $(\Sigma^+, d_{\rho}^+)$  e  $(\Sigma, d_{\rho})$  são espaços métricos.

PROVA | Por simplicidade, omitimos a dependência em  $\rho$  das métricas. Façamos a prova para  $d^+$  (a prova para  $d$  é similar). Começamos notando que se  $\underline{x}, \underline{y}, \underline{z} \in \Sigma^+$  então

$$N(\underline{x}, \underline{y}) \geq \min\{N(\underline{x}, \underline{z}), N(\underline{z}, \underline{y})\}. \quad (27.1)$$

Para checar essa desigualdade, temos dois casos a considerar:

- Se  $N(\underline{x}, \underline{z}) > N(\underline{z}, \underline{y})$  então  $N(\underline{x}, \underline{y}) = N(\underline{z}, \underline{y})$  e obtemos igualdade em (27.1). O mesmo vale se ocorre a desigualdade reversa.
- Se  $N(\underline{x}, \underline{z}) = N(\underline{z}, \underline{y}) = N$  então  $x_i = y_i$  para  $0 \leq i < N$  e portanto  $N(\underline{x}, \underline{y}) \geq N$ .

Agora procedemos para a prova de que  $d^+$  é métrica:

- POSITIVIDADE:  $d^+(\underline{x}, \underline{y}) \geq 0$  claramente, e  $d^+(\underline{x}, \underline{y}) = 0$  se e somente se  $N(\underline{x}, \underline{y}) = \infty$  se e somente se  $\underline{x} = \underline{y}$ .
- SIMETRIA:  $d^+(\underline{x}, \underline{y}) = d^+(\underline{y}, \underline{x})$  é claro.
- DESIGUALDADE TRIANGULAR: pela estimativa (27.1), temos

$$d^+(\underline{x}, \underline{y}) = \rho^{-N(\underline{x}, \underline{y})} \leq \rho^{-\min\{N(\underline{x}, \underline{z}), N(\underline{z}, \underline{y})\}} = \max\{d^+(\underline{x}, \underline{z}), d^+(\underline{z}, \underline{y})\},$$

que implica a desigualdade triangular.

Isso conclui a prova.  $\square$

A mesma versão forte da desigualdade triangular também vale para  $(\Sigma, d_\rho)$ . Isso mostra que  $(\Sigma^+, d_\rho^+)$  e  $(\Sigma, d_\rho)$  são *espaços ultramétricos*. Agora mostremos a compacidade desses espaços.

**Lema 126** | Os espaços métricos  $(\Sigma^+, d_\rho^+)$  e  $(\Sigma, d_\rho)$  são sequencialmente compactos.

PROVA | Provamos o resultado para  $(\Sigma^+, d_\rho^+)$ , sendo a prova para  $(\Sigma, d_\rho)$  análoga. Tome  $\{\underline{x}^{(n)}\}_{n \geq 1} \subset \Sigma^+$ , com  $\underline{x}^{(n)} = (x_0^{(n)}, x_1^{(n)}, \dots)$ . Como  $x_0^{(n)} \in \mathcal{A}$  e  $\mathcal{A}$  é finito, existe uma subsequência, que também denotaremos por  $\{\underline{x}^{(n)}\}_{n \geq 1}$ , tal que  $x_0^{(n)} = x_0$  é constante. Procedendo por indução coordenada a coordenada, construímos uma subsequência, ainda denotada por  $\{\underline{x}^{(n)}\}_{n \geq 1}$ , e letras  $x_0, x_1, \dots \in \mathcal{A}$  tais que  $x_k^{(n)} = x_k$  para todo  $n$  suficientemente grande. Tomando  $\underline{x} = (x_0, x_1, \dots) \in \Sigma^+$ , é claro que  $d_\rho^+(\underline{x}^{(n)}, \underline{x}) \rightarrow 0$  e portanto  $\underline{x}^{(n)} \rightarrow \underline{x}$ .  $\square$

Finalizamos essa seção introduzindo outros espaços simbólicos, definidos por grafos orientados. Seja  $G = (V, E)$  um grafo orientado, onde  $V = \mathcal{A}$ .

**Definição 127** (Espaços simbólicos associados a  $G$ ) O *espaço simbólico unilateral* com alfabeto  $\mathcal{A}$  é definido por

$$\Sigma^+ = \Sigma_{\mathcal{A}}^+ = \mathcal{A}^{\{0,1,\dots\}} = \{\underline{x} = (x_0, x_1, x_2, \dots) : x_i \in \mathcal{A}, \forall i \geq 0\}.$$



O espaço simbólico bilateral com alfabeto  $\mathcal{A}$  é definido por

$$\Sigma = \Sigma_{\mathcal{A}} = \mathcal{A}^{\mathbb{Z}} = \{\underline{x} = (\dots, x_{-1}; x_0, x_1, \dots) : x_i \in \mathcal{A}, \forall i \in \mathbb{Z}\}.$$

Observamos que as definições e resultados para  $\Sigma^+$  e  $\Sigma$  também valem para os espaços simbólicos associados a grafos  $\Sigma^+(G)$  e  $\Sigma(G)$ , com provas análogas.

Podemos representar a combinatoria de  $G$  por meio de uma matriz de adjacência  $A = [a_{ij}]_{i,j \in \mathcal{A}}$ , onde

$$a_{ij} = \begin{cases} 1, & \text{se } (i, j) \in E \\ 0, & \text{se } (i, j) \notin E. \end{cases}$$

**Exemplo 128** Se  $G = (\mathcal{A}, \mathcal{A} \times \mathcal{A})$  é o grafo completo, cuja matriz de adjacência é  $A = [1]_{i,j \in \mathcal{A}}$ , então  $\Sigma^+(G) = \Sigma^+$  e  $\Sigma(G) = \Sigma$ . Por isso, costumamos chamar  $(\Sigma^+, \sigma_+)$  e  $(\Sigma, \sigma)$  de TMS completos.

**Exemplo 129** Seja  $\mathcal{A} = \{0, 1\}$  e  $G$  o grafo abaixo:



A matriz de adjacência é  $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ . Então  $\Sigma^+(G)$  é o conjunto das sequências de  $\Sigma^+$  sem dois 1's consecutivos, e o mesmo ocorre para  $\Sigma(G)$ .

## SEÇÃO 28

### Palavras, cilindros e topologia produto

**Definição 130** (Palavras e linguagem) As *palavras* de tamanho  $n$  formadas pelo alfabeto  $\mathcal{A}$  são

$$\mathcal{A}^n = \{(x_0, \dots, x_{n-1}) : x_i \in \mathcal{A}\}.$$

A *linguagem* do alfabeto  $\mathcal{A}$  é definida por  $\cup_{n \geq 1} \mathcal{A}^n$ .

**Definição 131** (Cilindros) Dada uma palavra  $(a_m, \dots, a_{m+n-1})$  da linguagem,  $m \geq 0, n \geq 1$ , definimos o *cilindro* em  $\Sigma^+$  por

$$\begin{aligned} m[a_m, \dots, a_{m+n-1}] &= C_{m, m+n-1}(a_m, \dots, a_{m+n-1}) \\ &= \{\underline{x} \in \Sigma^+ : x_m = a_m, \dots, x_{m+n-1} = a_{m+n-1}\}. \end{aligned}$$

Cilindros em  $\Sigma$  são definidos do mesmo modo, permitindo  $m \in \mathbb{Z}$ .

Por exemplo, se  $\mathcal{A} = \{0, 1, 2\}$  então em  $\Sigma^+$  temos o cilindro

$$2021[1] = \{(x_0, \dots, x_{2020}, 1, x_{2022}, \dots) : x_i \in \{0, 1, 2\}\}.$$

Em  $\{0, 1\}^{\mathbb{Z}}$ , temos

$$C_{-1,1}(1, 0, 0) = \{(\dots, x_{-2}, 1; 0, 0, x_2, \dots) : x_i \in \{0, 1\}\}.$$

Vamos entender de que modo as definições acima se relacionam com a topologia produto. Relembre a definição da topologia produto: se  $(X_\lambda, \tau_\lambda)$  são espaços topológicos, com  $\lambda \in \Lambda$ , definimos uma topologia  $\tau$  em  $X = \prod_{\lambda \in \Lambda} X_\lambda$  de modo que cada projeção  $p_\lambda : X \rightarrow X_\lambda$  é contínua, i.e.  $\tau$  é gerada por conjuntos da forma  $\prod_{\lambda \in \Lambda} U_\lambda$ , onde  $U_\lambda \in \tau_\lambda$  e  $U_\lambda \neq X_\lambda$  para uma quantidade finita de  $\lambda$ .

No caso em estudo, considere a topologia discreta em  $\mathcal{A}$  (i.e. todo subconjunto  $U \subset \mathcal{A}$  é aberto e fechado), e tome as topologias produto em  $\Sigma^+$  e  $\Sigma$ . Essas topologias são geradas pelos seguintes tipos de conjuntos:

- $\Sigma^+$ : conjuntos  $U_0 \times U_1 \times \dots$  com  $U_k \neq \mathcal{A}$  para uma quantidade finita de  $k$ 's. Em particular, os cilindros

$${}_m[a_m, \dots, a_{m+n-1}] = \mathcal{A} \times \dots \times \{a_m\} \times \dots \times \{a_{m+n-1}\} \times \dots$$

são abertos. De fato, os cilindros da forma acima formam uma base da topologia (Exercício 54).

- $\Sigma$ : conjuntos  $\dots \times U_{-1} \times U_0 \times U_1 \times \dots$  com  $U_k \neq \mathcal{A}$  para uma quantidade finita de  $k$ 's. Similarmente, os cilindros

$${}_m[a_m, \dots, a_{m+n-1}] = \dots \times \{a_m\} \times \dots \times \{a_{m+n-1}\} \times \dots$$

são abertos.

Por outro lado, os cilindros também são fechados. De fato, seus complementares

$${}_m[a_m, \dots, a_{m+n-1}]^c = \bigcup_{\substack{(b_m, \dots, b_{m+n-1}) \neq \\ (a_m, \dots, a_{m+n-1})}} {}_m[b_m, \dots, b_{m+n-1}]$$

são a união de cilindros, logo são abertos. Em inglês, dizemos que cada cilindro é *clopen*. Isso implica, dentre outras propriedades, que  $\Sigma^+$  e  $\Sigma$  são totalmente desconexos.

Pelo teorema de Tychonoff, com respeito à topologia produto os espaços  $\Sigma^+, \Sigma$  são compactos. De fato, já sabíamos disso, pois a topologia produto é a mesma da topologia induzida pelas métricas. Mostraremos esse fato na próxima seção.

Finalizamos a presente seção notando que as definições e resultados provados acima também valem para os espaços simbólicos associados a grafos  $\Sigma^+(G)$  e  $\Sigma(G)$ , com provas análogas. Por exemplo, podemos também definir cilindros em  $\Sigma^+(G)$  e  $\Sigma$ , restringindo às sequências desses conjuntos.

## SEÇÃO 29

### Shifts topológicos de Markov

---

Para definir os shifts topológicos de Markov, devemos antes definir dinâmicas em  $\Sigma^+$  e  $\Sigma$ .

**Definição 132** (Transformações shift) O *shift unilateral* é a transformação  $\sigma_+ : \Sigma^+ \rightarrow \Sigma^+$  definida por  $\sigma_+(x_0, x_1, \dots) = (x_1, x_2, \dots)$ . O *shift bilateral* é a transformação  $\sigma : \Sigma \rightarrow \Sigma$  definida por  $\sigma(\dots, x_{-1}; x_0, x_1, \dots) = (\dots, x_0; x_1, x_2, \dots)$ .

Em outras palavras,  $\sigma_+$  e  $\sigma$  deslocam a sequência uma unidade à esquerda. No caso unilateral, apagamos a primeira entrada e consideramos a sequência remanescente. É fácil notar que  $\sigma^+$  não é invertível, ao passo que  $\sigma$  é.

**Lema 133** Com respeito às métricas  $d_\rho^+$  e  $d_\rho$ , as transformações  $\sigma_+$  e  $\sigma$  são  $\rho$ -Lipschitz. Mais especificamente:

$$d_\rho^+(\sigma_+(\underline{x}), \sigma_+(\underline{y})) \leq \rho d_\rho^+(\underline{x}, \underline{y}), \quad \forall \underline{x}, \underline{y} \in \Sigma^+$$

$$\rho^{-1} d_\rho(\underline{x}, \underline{y}) \leq d_\rho(\sigma(\underline{x}), \sigma(\underline{y})) \leq \rho d_\rho(\underline{x}, \underline{y}), \quad \forall \underline{x}, \underline{y} \in \Sigma.$$

**PROVA** Vamos provar a estimativa para  $\sigma_+$ . Escreva  $\underline{x} = (x_0, x_1, \dots)$  e  $\underline{y} = (y_0, y_1, \dots)$ . Temos dois casos:

- $x_0 \neq y_0$ : temos  $d_\rho^+(\underline{x}, \underline{y}) = 1$  e  $d_\rho^+(\sigma_+(\underline{x}), \sigma_+(\underline{y})) \leq 1$ , logo a desigualdade é clara.
- $x_0 = y_0$ : nesse caso,  $N(\sigma_+(\underline{x}), \sigma_+(\underline{y})) = N(\underline{x}, \underline{y}) - 1$  e daí

$$d_\rho^+(\sigma_+(\underline{x}), \sigma_+(\underline{y})) = \rho d_\rho^+(\underline{x}, \underline{y}).$$

□

**Definição 134** (Shift topológico de Markov com alfabeto  $\mathcal{A}$ ) O *shift topológico de Markov unilateral* com alfabeto  $\mathcal{A}$  é o par  $(\Sigma^+, \sigma^+)$ . O *shift topológico de Markov bilateral* com alfabeto  $\mathcal{A}$  é o par  $(\Sigma, \sigma)$ .

Em inglês, eles são conhecidos como topological Markov shifts, com a sigla TMS. No que segue, utilizaremos essa sigla.

Seja agora  $G = (V, E)$  um grafo orientado, onde  $V = \mathcal{A}$ , e relembre a definição de  $\Sigma^+(G)$  e  $\Sigma(G)$  ao final da Seção 27. É fácil notar que  $\sigma^+[\Sigma^+(G)] \subset \Sigma^+(G)$  e  $\sigma[\Sigma(G)] = \Sigma(G)$ . Assim, podemos restringir as transformações shift a esses conjuntos e considerar as seguintes definições.

**Definição 135** (Shift topológico de Markov associado a  $G$ ) O *shift topológico de Markov unilateral* associado a  $G$  é o par  $(\Sigma^+(G), \sigma^+)$ . O *shift topológico de Markov bilateral* associado a  $G$  é o par  $(\Sigma(G), \sigma)$ .

Agora provamos que as duas topologias introduzidas nos espaços simbólicos coincidem.

**Teorema 136** As topologias produto de  $\Sigma^+$  e  $\Sigma$  coincidem com as topologias induzidas por  $d_\rho^+$  e  $d_\rho$ .

PROVA | Fixe  $\rho > 1$ . Focamos no caso unilateral. Sejam  $\tau$  a topologia produto em  $\Sigma^+$  e  $\tau_{d^+}$  a topologia induzida por  $d^+$ . A prova seguirá se mostrarmos que uma base da topologia de  $\tau$  está contida em  $\tau_{d^+}$  e vice-versa. Tome um cilindro  $m[a_m, \dots, a_{m+n-1}]$ . Temos que:

- $C = {}_0[a_m, \dots, a_{m+n-1}] = B(\underline{x}, \rho^{-n})$  para qualquer  $\underline{x} \in {}_0[a_m, \dots, a_{m+n-1}]$ , logo  $C$  é aberto de  $\tau_{d^+}$ .
- ${}_m[a_m, \dots, a_{m+n-1}] = \sigma_+^{-n}(C)$  é aberto de  $\tau_{d^+}$ , pois  $\sigma_+$  é contínua.

Reciprocamente, tome uma bola  $B(\underline{x}, r)$  da topologia  $\tau_{d^+}$ . Se  $n$  é tal que  $\rho^{-n} \leq r < \rho^{-n+1}$ , então  $B(\underline{x}, r) = B(\underline{x}, \rho^{-n}) = {}_0[x_0, \dots, x_{n-1}] \in \tau$ .  $\square$

Finalizamos esse capítulo provando que os TMS  $(\Sigma^+, \sigma^+)$  e  $(\Sigma, \sigma)$  são transitivos, i.e. possuem pontos com órbita densa.

**Lema 137** |  $(\Sigma^+, \sigma^+)$  e  $(\Sigma, \sigma)$  possuem pontos com órbita densa.

PROVA | Podemos fixar  $\rho > 1$  e tomar as distâncias  $d^+$  e  $d$ . Relembre que para  $\underline{x} \in \Sigma^+$  temos  $\mathcal{O}_{\sigma^+}(\underline{x}) = \{\sigma_+^n(\underline{x}) : n \geq 0\}$ ; para  $\underline{x} \in \Sigma$  temos  $\mathcal{O}_\sigma(\underline{x}) = \{\sigma^n(\underline{x}) : n \in \mathbb{Z}\}$ . Note que  $d^+(\sigma_+^m(\underline{x}), \underline{y}) < \rho^{-n}$  se e somente se  $(x_m, \dots, x_{m+n-1}) = (y_0, \dots, y_{m-1})$ , e uma equivalência similar vale para  $\sigma$ . Assim,  $\underline{x} \in \Sigma^+$  possui órbita densa se e somente se toda palavra da linguagem aparece em  $\underline{x}$ . Isso é fácil de construir, bastando definir  $\underline{x}$  por uma concatenação lado a lado de todas as palavras da linguagem:

$$\underline{x} = (\mathcal{A}^1, \mathcal{A}^2, \dots),$$

onde  $\mathcal{A}^n$  significa que concatenamos lado a lado as  $|\mathcal{A}|^n$  palavras de tamanho  $n$ . O mesmo pode ser feito no caso bilateral.  $\square$

Por exemplo, em  $\{0, 1\}^{\{0, 1, \dots\}}$  temos o seguinte ponto com órbita densa:

$$\underline{x} = (\underbrace{0, 1}_{\mathcal{A}^1}, \underbrace{0, 0, 0, 1, 1, 0, 1, 1}_{\mathcal{A}^2}, \underbrace{0, 0, 0, 0, 1, 0, 1, 0, \dots, 1, 1, 1, \dots}_{\mathcal{A}^3}).$$

## SEÇÃO 30

### Exercícios

**Exercício 53** | Forneça detalhes para todas as provas não explicitadas do caso bilateral.

**Exercício 54** | Mostre que os cilindros  $m[a_m, \dots, a_{m+n-1}]$  formam uma base para a topologia produto de  $\Sigma^+$  e  $\Sigma$ .

**Exercício 55** | Seja  $G = (\{1, 2, \dots, n\}, E)$  um grafo orientado, e seja  $A = [a_{ij}]$  sua matriz de adjacência. Seja  $A^k = [a_{ij}^{(k)}]$ . Mostre os seguintes itens.

- (1) O total de caminhos de tamanho  $k$  que começam em  $i$  e terminam em  $j$  é  $a_{ij}^{(k)}$ .
- (2) O total de ciclos de tamanho  $k$  é  $\text{tr}(A^k) = \sum a_{ii}^{(k)}$ .

- (3) Conclua que o total de pontos fixos de  $\sigma_+^n, \sigma_-^n$  em  $\Sigma^+(G), \Sigma(G)$  é  $\text{tr}(A^k) = \sum a_{ii}^{(k)}$ .

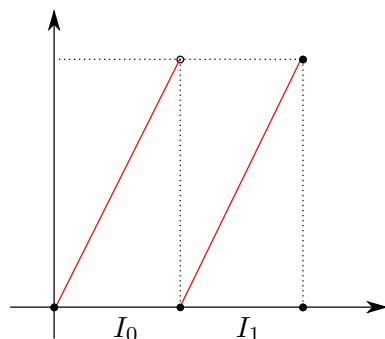
# Exemplos

SEÇÃO 31

## Mapas unidimensionais markovianos

Vamos começar com um exemplo introdutório. Seja  $f : [0, 1] \rightarrow [0, 1]$  a transformação  $f(x) = 2x \pmod{1}$ , i.e.

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



A transformação  $f$  é intimamente relacionada com a representação binária de números reais. Dado  $x \in [0, 1] \setminus \mathbb{Q}$ , há um único modo de escrever  $x$  como a soma de potências de dois:

$$x = 0.a_1a_2\cdots = \sum_{n \geq 1} \frac{a_n}{2^n} = \frac{a_1}{2^1} + \frac{a_2}{2^2} + \cdots,$$

onde cada  $a_n \in \{0, 1\}$ . A transformação  $f$  nos permite identificar  $a_1, a_2, \dots$ , do seguinte modo. Sejam  $I_0 = [0, \frac{1}{2}]$  e  $I_1 = [\frac{1}{2}, 1]$ . Note que:

- $a_1 = 0 \iff 0 < x < \frac{1}{2} \iff x \in I_0$ .
- $a_1 = 1 \iff \frac{1}{2} < x < 1 \iff x \in I_1$ .

Noutras palavras, identificamos  $x$  pela inclusão  $x \in I_{a_1}$ . Para achar  $a_2$ , aplicamos  $f$ :

$$x = 0.a_1a_2\cdots = \sum_{n \geq 1} \frac{a_n}{2^n} \implies f(x) = 0.a_2a_3\cdots = \sum_{n \geq 1} \frac{a_{n+1}}{2^n}.$$

Portanto,  $f(x) \in I_{a_2}$ . Indutivamente, temos  $f^{n-1}(x) = 0.a_n a_{n+1} \cdots$  e portanto  $f^{n-1}(x) \in I_{a_n}$ . Desse modo, geramos de modo dinâmico por meio de  $f$  os dígitos de  $x$  em base dois. O mesmo vale em base  $d$ , bastando tomar  $f(x) = dx \pmod{1}$ .

Uma consequência é a seguinte associação (que, por enquanto, é quase uma tautologia): cada  $x \in [0, 1] \setminus \mathbb{Q}$  se associa a uma sequência unitaleral

$$\theta(x) = (a_1, a_2, \dots) \in \{0, 1\}^{\{1, \dots\}}$$

onde  $a_n$  é definido por  $f^{n-1}(x) \in I_{a_n}$ . Definimos assim uma codificação  $\theta : [0, 1] \setminus \mathbb{Q} \rightarrow \{0, 1\}^{\{1, \dots\}}$ . Para entender melhor a codificação  $\theta$ , veja que a relação entre  $x$  e  $\theta(x)$  é a seguinte:

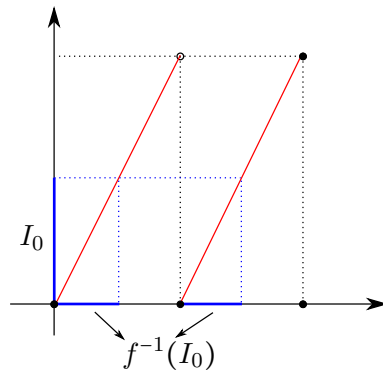
$$\begin{cases} x \in I_{a_1} \\ f(x) \in I_{a_2} \\ f^2(x) \in I_{a_3} \\ \dots \end{cases} \iff \begin{cases} x \in I_{a_1} \\ x \in f^{-1}(I_{a_2}) \\ x \in f^{-2}(I_{a_3}) \\ \dots \end{cases} \implies x \in \bigcap_{n \geq 0} f^{-n}(I_{a_{n+1}}).$$

Afirmamos que, pelas propriedades de  $f$ , a interseção acima consiste de um único ponto. No que segue, assumimos que  $f$  deixa de ser uma função em  $x = \frac{1}{2}$ , e escrevemos que  $f(\frac{1}{2}) = 1$  pelo ramo da esquerda e  $f(\frac{1}{2}) = 0$  pelo ramo da direita. Como  $x \notin \mathbb{Q}$ , nossa análise não é afetada por essa hipótese. Por outro lado, essa condição garante que as pré-imagens de intervalos fechados pelos ramos  $f \upharpoonright_{I_0}$  e  $f \upharpoonright_{I_1}$  permanecem fechados. Vejamos alguns exemplos.

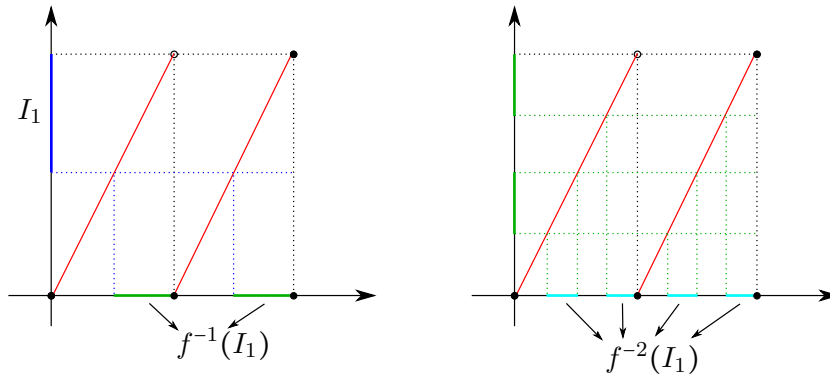
**Exemplo 138**

Vamos identificar  $A = I_1 \cap f^{-1}(I_0) \cap f^{-2}(I_1) = \{x \in [0, 1] : a_1 = 1, a_2 = 0, a_3 = 1\}$ . Para isso, identificamos sucessivamente  $I_1$ ,  $I_1 \cap f^{-1}(I_0)$  e finalmente  $A$ . Temos  $I_1 = [\frac{1}{2}, 1]$ . Para identificar  $I_1 \cap f^{-1}(I_0)$ , observe que  $f^{-1}(I_0) = J_0 \cup J_1 = [0, \frac{1}{4}] \cup [\frac{2}{4}, \frac{3}{4}]$ , com  $J_i \subset I_i$ . Assim,

$$I_1 \cap f^{-1}(I_0) = [\frac{2}{4}, \frac{3}{4}].$$



Por fim, para identificar  $A$ , note que  $f^{-2}(I_1)$  consiste de quatro intervalos fechados, cada um de comprimento  $\frac{1}{8}$ .



Apenas um deles intersecta  $I_1 \cap f^{-1}(I_0)$ , e portanto  $A = [\frac{5}{8}, \frac{6}{8}]$ .  
 Observe que se não tivéssemos definido  $f(\frac{1}{2})$  como dois valores teríamos  $f^{-1}(I_1) = [\frac{1}{4}, \frac{2}{4}) \cup [\frac{3}{4}, 1]$ , e portanto  $f^{-2}(I_1)$  não consistiria de quatro intervalos fechados.

Retornando ao nosso objetivo, escrevemos

$$\bigcap_{n \geq 0} f^{-n}(I_{a_{n+1}}) = \bigcap_{n \geq 0} \underbrace{I_{a_1} \cap \dots \cap f^{-(n-1)} I_{a_n}}_{=F_n} = \bigcap_{n \geq 0} F_n,$$

onde  $\{F_n\}_{n \geq 0}$  é uma cadeia descendente de intervalos fechados com  $|F_n| = \frac{1}{2^{n+1}}$  (isso é facilmente provado por indução em  $n$ ). Pelo teorema de interseção de Cantor,  $\bigcap_{n \geq 0} f^{-n}(I_{a_{n+1}})$  consiste de um único ponto. Desse modo, definimos uma inversa para  $\theta$  por  $\pi : \{0, 1\}^{\{1, \dots\}} \rightarrow [0, 1]$  dada por

$$\{\pi(\underline{x})\} = \bigcap_{n \geq 0} f^{-n}(I_{a_{n+1}}).$$

Mais ainda,  $\pi$  comuta o diagrama abaixo:

$$\begin{array}{ccc} \{0, 1\}^{\{1, \dots\}} & \xrightarrow{\sigma_+} & \{0, 1\}^{\{1, \dots\}} \\ \pi \downarrow & & \downarrow \pi \\ [0, 1] & \xrightarrow{f} & [0, 1] \end{array}$$

De fato, se  $\underline{x} = (x_0, x_1, \dots)$  e  $\pi(\underline{x}) = y$ , então

$$\begin{cases} y \in I_{x_1} \\ f(y) \in I_{x_2} \\ f^2(y) \in I_{x_3} \\ \vdots \end{cases} \iff \begin{cases} f(y) \in I_{x_2} \\ f(f(y)) \in I_{x_3} \\ \vdots \end{cases}$$

e portanto  $f(\pi(\underline{x})) = f(y) = \pi(x_2, x_3, \dots) = \pi(\sigma_+(\underline{x}))$ .

As propriedades fundamentais de  $f$  que permitiram provar que a interseção é uma cadeia descendente de intervalos de comprimento convergindo para zero são duas:  $f$  é expansora e satisfaz a propriedade de Markov, conforme definiremos abaixo. No que segue, vamos definir uma classe maior de transformações



para as quais podemos proceder exatamente como fizemos com  $2x \pmod{1}$ , e forneceremos todos os detalhes formais.

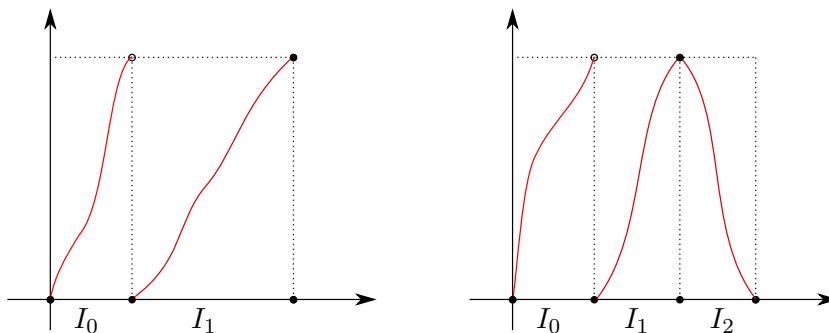
**Definição 139** (Mapa unidimensional markoviano) Dizemos que  $f : [0, 1] \rightarrow [0, 1]$  é um *mapa unidimensional markoviano* se existirem  $z_0 = 0 < z_1 < \dots < z_{n-1} < z_n = 1$  e  $\lambda > 1$  tais que:

- (1) *f* É EXPANSORA: *f* é diferenciável com  $|f'| \geq \lambda$  em cada  $(z_k, z_{k+1})$ , para  $k = 0, \dots, n - 1$ .
- (2) PROPRIEDADE DE MARKOV:  $\lim_{x \rightarrow z_k^-} f(x)$  e  $\lim_{x \rightarrow z_k^+} f(x)$  pertencem ao conjunto  $\{z_0, \dots, z_n\}$ , para  $k = 0, \dots, n$ .

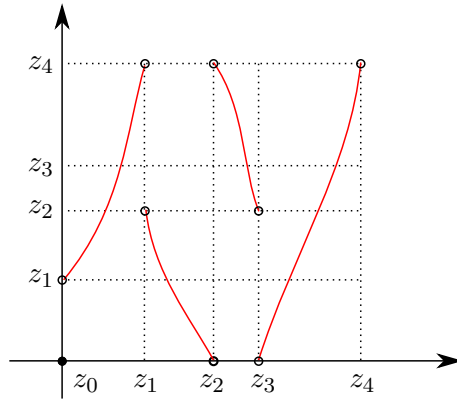
Nesse caso, chamamos a família  $\{[z_0, z_1], \dots, [z_{n-1}, z_n]\}$  de *partição de Markov*.

Em particular, a restrição de *f* a cada  $(z_k, z_{k+1})$  é estritamente monótona, para  $k = 0, \dots, n - 1$ . Observe que nem descrevemos o valor de  $f(z_k)$ . Isso não é importante e, de fato, *f* nem precisa estar definida nesses pontos. Mas, para que as pré-imagens de intervalos fechados permaneçam fechadas em cada ramo de *f* (como fizemos para  $f(x) = 2x \pmod{1}$ ), assumimos que  $f(z_k) = \lim_{x \rightarrow z_k^-} f(x)$  no ramo  $f \upharpoonright_{(z_{k-1}, z_k)}$  e  $f(z_k) = \lim_{x \rightarrow z_k^+} f(x)$  no ramo  $f \upharpoonright_{(z_k, z_{k+1})}$ .

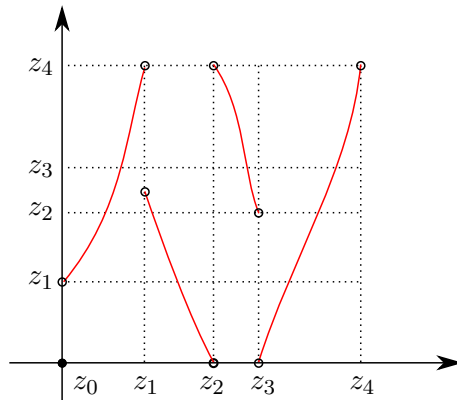
**Exemplo 140** O mapa  $f(x) = 2x \pmod{1}$  visto acima é unidimensional markoviano. Mais geralmente, se  $d \in \mathbb{Z}$  então  $f(x) = dx \pmod{1}$  é unidimensional markoviano. Melhor ainda, *f* nem precisa ter derivada constante: se *f* é uma transformação expansora com ramos completos, i.e.  $f(z_k, z_{k+1}) = [0, 1]$  para  $k = 0, \dots, n - 1$  então *f* é unidimensional markoviana. Veja exemplos na figura abaixo.



**Exemplo 141** O exemplo abaixo é unidimensional markoviano. Note que *f* possui ramos não completos.



**Exemplo 142** O exemplo abaixo *não* é unidimensional markoviano. De fato,  $\lim_{x \rightarrow z_1^+} f(x)$  não pertence a  $\{z_0, z_1, z_2, z_3, z_4\}$ .



A construção feita para  $f(x) = 2x \pmod{1}$  se aplica *ipsis literis* para todo mapa unidimensional markoviano. Vamos recapitular a discussão, dividindo-a em passos.

PASSO 1. Tome  $G = (V, E)$  onde  $V = \{I_0 = [z_0, z_1], \dots, I_{n-1} = [z_{n-1}, z_n]\}$  e  $E = \{I_i \rightarrow I_j : f(I_i) \supset I_j\}$ .<sup>1</sup>

PASSO 2. Considere o TMS unilateral  $(\Sigma^+, \sigma_+)$  associado a  $G$ .

PASSO 3. Defina  $\pi : \Sigma^+ \rightarrow [0, 1]$  por

$$\{\pi(\underline{x})\} = \bigcap_{n \geq 0} f^{-n}(I_{x_n}).$$

Note que, uma vez que justificamos a boa definição de  $\pi$ , obtemos que  $\pi(\underline{x}) =$

<sup>1</sup>Relembre que  $f(z_i) = \lim_{x \rightarrow z_i^+} f(x)$  e  $f(z_{i+1}) = \lim_{x \rightarrow z_{i+1}^-} f(x)$  em  $I_i$ , de modo que as arestas de

$G$  estão definidas, com respeito à  $f$  original, pela inclusão  $\overline{f(I_i)} \supset I_j$ . Esse é um benefício de estender  $f$  nos pontos  $\{z_0, \dots, z_n\}$ , pois assim não precisamos tomar fechos. Por outro lado, a discussão perde a precisão (injetividade) no conjunto  $\{z_0, \dots, z_n\}$  e em suas pré-imagens mas, como o conjunto de pré-imagens é enumerável, ele é “irrelevante” do ponto de vista de medida de Lebesgue.

$y$  é o único ponto tal que

$$\begin{cases} y \in I_{x_0} \\ f(y) \in I_{x_1} \\ f^2(y) \in I_{x_2} \\ \vdots \end{cases}$$

A boa definição e propriedades que  $\pi$  satisfaz estão enunciadas no próximo resultado.

**Teorema 143**

Na notação acima,  $\pi : \Sigma^+ \rightarrow [0, 1]$  está bem-definida, é contínua e comuta o diagrama abaixo:

$$\begin{array}{ccc} \Sigma^+ & \xrightarrow{\sigma^+} & \Sigma^+ \\ \pi \downarrow & & \downarrow \pi \\ [0, 1] & \xrightarrow{f} & [0, 1] \end{array}$$

PROVA

Para ver que  $\pi$  está bem-definida, devemos provar que a interseção  $\bigcap_{n \geq 0} f^{-n}(I_{x_n})$  consiste de um único elemento. Temos

$$\bigcap_{n \geq 0} f^{-n}(I_{x_n}) = \bigcap_{n \geq 0} \underbrace{I_{x_0} \cap \dots \cap f^{-n}I_{x_n}}_{=F_n} = \bigcap_{n \geq 0} F_n,$$

onde  $\{F_n\}_{n \geq 0}$  é uma cadeia descendente de conjuntos fechados.

FATO 1. Se  $I_{a_0} \rightarrow I_{a_1} \rightarrow \dots \rightarrow I_{a_n}$  é um caminho em  $G$ , então

$$\bigcap_{k=0}^n f^{-k}(I_{a_k})$$

é um intervalo não-vazio.

PROVA

(Fato 1) Esse resultado é consequência direta da propriedade de Markov. Vamos prová-la por indução. O caso  $n = 1$  decorre da definição de aresta: se  $f(I_{a_0}) \supset I_{a_1}$   $I_{a_0}$  intersecta  $f^{-1}(I_{a_1})$ ; pela monotonicidade de  $f \upharpoonright_{I_{a_0}}$ , a interseção  $I_{a_0} \cap f^{-1}(I_{a_1})$  é um intervalo. Assuma o resultado válido para  $n - 1$ , e considere  $I_{a_0} \rightarrow I_{a_1} \rightarrow \dots \rightarrow I_{a_n}$  caminho de tamanho  $n$  em  $G$ . Por indução,

$$X = I_{a_1} \cap \dots \cap f^{-(n-1)}(I_{a_n})$$

é um intervalo não-vazio. Temos  $X \subset I_{a_1} \subset f(I_{a_0})$ , e portanto  $f^{-1}(X)$  intersecta  $I_{a_0}$ , com a interseção  $I_{a_0} \cap f^{-1}(X)$  igual a um intervalo. Assim,

$$I_{a_0} \cap f^{-1}(I_{a_1}) \dots \cap f^{-n}(I_{a_n}) = I_{a_0} \cap f^{-1}(X)$$

é um intervalo não-vazio. □

FATO 2. Se  $I_{a_0} \rightarrow I_{a_1} \rightarrow \dots \rightarrow I_{a_n}$  é um caminho em  $G$ , então o intervalo

$$\bigcap_{k=0}^n f^{-k}(I_{a_k})$$

tem comprimento  $\leq \lambda^{-n}$ .

PROVA (Fato 2) A prova é novamente por indução. Considere  $n = 1$ . O mapa  $f$  leva o intervalo  $I_{a_0} \cap f^{-1}(I_{a_1})$  em um subintervalo de  $[0, 1]$ . Como  $f$  é expansora com derivada  $\geq \lambda$ , segue que  $I_{a_0} \cap f^{-1}(I_{a_1})$  tem comprimento  $\leq \lambda^{-1}$ . Assuma o resultado válido para  $n - 1$ . Utilizando a mesma notação da prova do Fato 1, temos que  $X$  é um intervalo de comprimento  $\leq \lambda^{-(n-1)}$ . Pela mesma argumentação do caso  $n = 1$ , concluímos que  $I_{a_0} \cap f^{-1}(I_{a_1}) \cdots \cap f^{-n}(I_{a_n}) = I_{a_0} \cap f^{-1}(X)$  tem comprimento  $\leq \lambda^{-n}$ .  $\square$

Pelos Fatos 1 e 2, segue que do teorema de interseção de Cantor que  $\pi$  está bem-definida. Passamos agora para provar a continuidade. Fixe  $\rho > 1$ , e seja  $d^+ = d_\rho^+$  em  $\Sigma^+$ .

FATO 3. Se  $\underline{x}, \underline{y} \in \Sigma^+$  com  $d^+(\underline{x}, \underline{y}) < \rho^{-n}$ , então  $d^+(\pi(\underline{x}), \pi(\underline{y})) \leq \lambda^{-n}$ . Em particular,  $\pi$  é contínua.

PROVA (Fato 3) Escreva  $\underline{x} = (x_0, x_1, \dots)$  e  $\underline{y} = (y_0, y_1, \dots)$ . Então  $(x_0, \dots, x_n) = (y_0, \dots, y_n)$ . Escreva

$$\begin{cases} \pi(\underline{x}) = \bigcap_{k \geq 0} \underbrace{I_{x_0} \cap \dots \cap f^{-k}(I_{x_k})}_{=F_k} = \bigcap_{k \geq 0} F_k \\ \pi(\underline{y}) = \bigcap_{k \geq 0} \underbrace{I_{y_0} \cap \dots \cap f^{-k}(I_{y_k})}_{=G_k} = \bigcap_{k \geq 0} G_k. \end{cases}$$

Temos  $F_k = G_k$  para  $k = 0, 1, \dots, n$ . Como  $\pi(\underline{x}), \pi(\underline{y}) \in F_n = G_n$ , concluímos que  $d^+(\pi(\underline{x}), \pi(\underline{y})) \leq \text{diam}(F_n) \leq \lambda^{-n}$ , onde na última passagem usamos o Fato 2. Isso prova o Fato 3.  $\square$

Resta mostrar que  $\pi \circ \sigma_+ = f \circ \pi$ . Dado  $\underline{x} \in \Sigma^+$ , temos que  $\pi(\underline{x})$  é o único ponto tal que

$$\begin{cases} \pi(\underline{x}) \in I_{x_0} \\ f(\pi(\underline{x})) \in I_{x_1} \\ f^2(\pi(\underline{x})) \in I_{x_2} \\ \vdots \end{cases} \implies \begin{cases} f(\pi(\underline{x})) \in I_{x_1} \\ f(f(\pi(\underline{x}))) \in I_{x_2} \\ \vdots \end{cases}$$

e portanto  $f(\pi(\underline{x})) = \pi(x_1, x_2, \dots) = \pi(\sigma_+(\underline{x}))$ .  $\square$

É importante entender também a imagem de  $\pi$ . Afirmamos que  $\pi$  é “quase” bijetiva: se  $y \in [0, 1]$  é tal que  $f^k(y) \notin \{z_0, \dots, z_n\}$  para todo  $k \geq 0$  então existe único  $\underline{x} \in \Sigma^+$  tal que  $\pi(\underline{x}) = y$ . De fato, basta notar que existe único  $x_k \in \{0, 1, \dots, n-1\}$  tal que  $f^k(y) \in I_{x_k}$ , o que define uma sequência  $(x_0, x_1, \dots)$ . Afirmamos que essa sequência está em  $\Sigma^+$ . Como  $f^k(y) \in \text{int}(I_{x_k})$  e  $f^{k+1}(y) \in \text{int}(I_{x_{k+1}})$ , temos que  $f(\text{int}(I_{x_k}))$  intersecta  $\text{int}(I_{x_{k+1}})$ . Pelo TVI e pela propriedade de Markov (propriedade (2) na Definição 139), segue que  $f(I_{x_k}) \supset I_{x_{k+1}}$ .

A observação do último parágrafo permitiu passar de uma interseção não-vazia (de interiores) para uma inclusão ( $f(\text{int}(I_i)) \cap I_j \neq \emptyset \implies f(I_i) \supset I_j$ ). Podemos, portanto, reinterpretar a propriedade de Markov do seguinte modo: “se intersecta, então contém”. Explicado desse modo, podemos dar uma definição alternativa para o conjunto de arestas  $E$ , sem precisar estender  $f$  nos pontos

$\{z_0, \dots, z_n\}$ :

$$E = \{I_i \rightarrow I_j : f(\text{int}(I_i)) \cap \text{int}(I_j) \neq \emptyset\}.$$

Note que a definição não considera  $f(z_k)$ . Em suma, a propriedade de Markov implica inclusões que, por sua vez, implicam concatenação de arestas como fizemos na prova do Teorema 143.

SEÇÃO 32

## Conjuntos de Cantor dinamicamente definidos

---

Nessa seção vamos definir uma classe de conjuntos de  $\mathbb{R}$ , chamados conjuntos de Cantor dinamicamente definidos. Sugerimos ao leitor estudar a Seção 39 sobre dimensão de Hausdorff em  $\mathbb{R}$ . Essa é uma maneira de medir conjuntos que possuem medida de Lebesgue nula. Chamamos um subconjunto de  $\mathbb{R}$  de *fractal* se sua dimensão de Hausdorff estiver em  $(0, 1)$ . Os conjuntos de Cantor dinamicamente definidos são fractais, definidos a partir de funções. Iniciaremos a discussão com a introdução do conjunto de Cantor clássico  $K \subset \mathbb{R}$ . Antes de defini-lo de modo dinâmico, vamos considerar a maneira tradicional de definição, por meio de um processo iterativo, conforme descreveremos abaixo.

PASSO 0. Tome  $K_0 = [0, 1]$ .

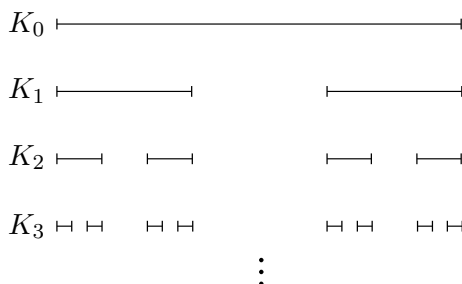
PASSO 1. Divida  $K_0$  em três intervalos de mesmo comprimento e exclua o terço médio aberto (i.e. os extremos são mantidos). O conjunto resultante é

$$K_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

PASSO 2. Para cada um dos dois intervalos de  $K_1$ , repita o argumento, i.e. exclua seu terço médio aberto. O conjunto resultante é

$$K_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{3}{9}\right] \cup \left[\frac{6}{9}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

O passo  $n$  é realizado de modo análogo, para todo  $n \geq 3$ . Desse modo, geramos uma cadeia descendente de compactos não-vazios  $K_0 \supset K_1 \supset K_2 \supset \dots$  em que  $K_n$  é a união de  $2^n$  intervalos compactos de comprimento  $\frac{1}{3^n}$  cada.



**Definição 144**

(Conjunto de Cantor clássico) O *conjunto de Cantor clássico* é definido por

$$K = \bigcap_{n \geq 0} K_n.$$

As seguintes propriedades valem:

- $K$  é um compacto não-vazio, pelo teorema de interseção de Cantor.
- $K$  possui medida de Lebesgue igual a zero. De fato,  $K_n$  tem medida de Lebesgue igual a  $2^n \times \frac{1}{3^n} = \left(\frac{2}{3}\right)^n$ , que converge para zero quando  $n \rightarrow +\infty$ .
- $K$  é igual ao conjunto dos números de  $[0, 1]$  com apenas 0 e 2 em base 3 (Exercício 57):

$$K = \left\{ \sum_{n \geq 0} \frac{a_n}{3^n} : a_n \in \{0, 2\} \right\}.$$

Embora  $K$  seja negligível para a medida de Lebesgue, ele é um conjunto grande. De fato, afirmamos que  $K$  possui dimensão de Hausdorff igual a  $\frac{\log 2}{\log 3}$ . Para ver isso, escreva  $K_n = \bigcup_{1 \leq i \leq 2^n} I_n^i$  onde  $|I_n^i| = \frac{1}{3^n}$ . Dado  $d \in [0, 1]$ , temos

$$\begin{aligned} m_d(K) &= \lim_{n \rightarrow +\infty} \sum_{1 \leq i \leq 2^n} |I_n^i|^d = \lim_{n \rightarrow +\infty} 2^n \left(\frac{1}{3^n}\right)^d \\ &= \lim_{n \rightarrow +\infty} \left(\frac{2}{3^d}\right)^n = \begin{cases} +\infty & , \text{ se } d < \frac{\log 2}{\log 3} \\ 0 & , \text{ se } d > \frac{\log 2}{\log 3}. \end{cases} \end{aligned}$$

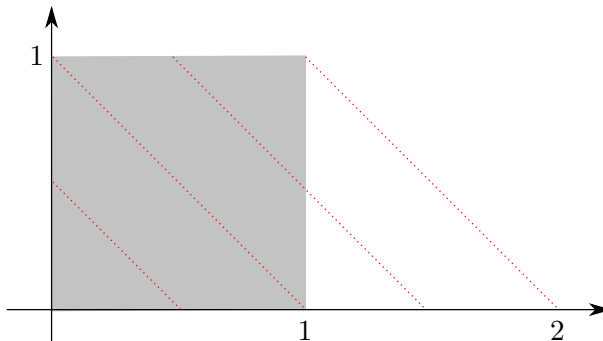
Logo,  $K$  é um conjunto fractal.

Outra propriedade que mostrar  $K$  ser grande é que  $K + K = [0, 2]$ . Forneceremos uma prova geométrica dessa igualdade. Começamos observando que se  $A, B \subset \mathbb{R}$  então a soma aritmética  $A + B = \{a + b : a \in A, b \in B\}$  é igual à projeção de  $45^\circ$  do conjunto  $A \times B \subset \mathbb{R}^2$  em  $\mathbb{R}$ . Isso é verdade porque tal projeção  $\pi : \mathbb{R}^2 \rightarrow \mathbb{R}$  é dada por  $\pi(x, y) = x + y$ , e portanto  $\pi(A \times B) = A + B$ . Em particular,  $K + K = \pi(K \times K)$ . Como

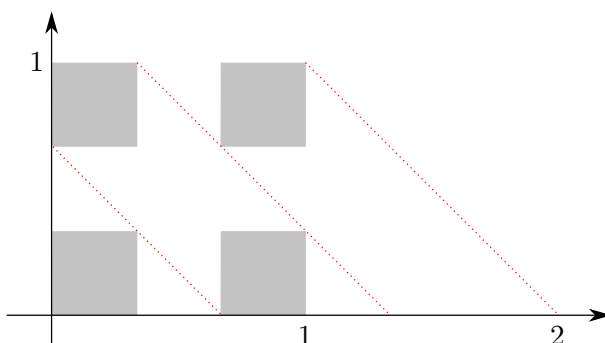
$$K \times K = \bigcap_{n \geq 0} K_n \times K_n,$$

basta provarmos que  $\pi(K_n \times K_n) = [0, 2]$ . Isso será provado por indução, em cada um dos passos da construção.

PASSO 0.  $K_0 \times K_0 = [0, 1]^2$  e portanto  $\pi(K_0 \times K_0) = [0, 2]$ .



PASSO 1.  $K_1 \times K_1$  é a união de quatro quadrados, cada um deles com lado  $\frac{1}{3}$ . A união da quatro projeções coincide com a projeção do quadrado  $[0, 1]^2$ , veja na figura abaixo. Assim,  $\pi(K_1 \times K_1) = \pi([0, 1]^2) = [0, 2]$ .



PASSO 2. Cada quadrado  $Q$  de  $K_1 \times K_1$  vira quatro quadrados em  $K_2 \times K_2$  de modo que a união das quatro projeções coincide com a projeção de  $Q$ . Assim,  $\pi(K_2 \times K_2) = \pi(K_1 \times K_1) = [0, 2]$ .

Por indução, segue que  $\pi(K_n \times K_n) = [0, 2]$  para todo  $n \geq 0$ , conforme afirmamos. Agora, introduzimos uma classe bem maior de conjuntos de Cantor.

**Definição 145** (Conjunto de Cantor dinamicamente definido) Dizemos que um conjunto  $K \subset [0, 1]$  é um *conjunto de Cantor dinamicamente definido* de classe  $C^k$ ,  $k \geq 1$ , se:

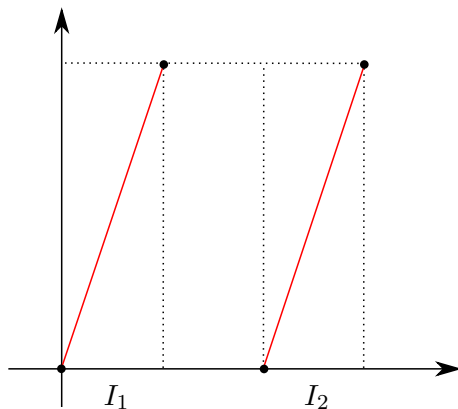
- (1) Existem intervalos compactos disjuntos  $I_1, I_2, \dots, I_r \subset [0, 1]$  tais que  $K \subset I_1 \cup I_2 \cup \dots \cup I_r$  e as extremidades de cada  $I_j$  pertencem a  $K$ .
- (2) Existe uma função  $f$  de classe  $C^k$  e expansora, definida em uma vizinhança da união  $I_1 \cup I_2 \cup \dots \cup I_r$  tal que  $f(I_j)$  é o fecho convexo de intervalos dentre  $I_1, \dots, I_r$ .
- (3)  $K = \bigcap_{n \geq 0} f^{-n}(I_1 \cup I_2 \cup \dots \cup I_r)$ .

Nesse caso, chamamos a família  $\{I_1, I_2, \dots, I_r\}$  de *partição de Markov*.

A condição (1) diz que os intervalos são definidos de modo maximal. A condição (2) sobre a imagem  $f(I_j)$  é a *propriedade de Markov* (“se intersecta, então contém”). O conjunto  $K$  é o conjunto de pontos que podem ser iterados por  $f$  para todo  $n \geq 0$ . Note que, ao contrário dos mapas unidimensionais markovianos,  $f$  é definida nos extremos de  $I_1, \dots, I_r$ .

**Exemplo 146** O conjunto de Cantor clássico é dinamicamente definido. Para ver isso, tome  $I_1 = [0, \frac{1}{3}]$  e  $I_2 = [\frac{2}{3}, 1]$  e considere  $f : I_1 \cup I_2 \rightarrow [0, 1]$  definida por

$$f(x) = \begin{cases} 3x & , \text{ se } x \in I_1 \\ 3x - 2 & , \text{ se } x \in I_2 \end{cases}$$



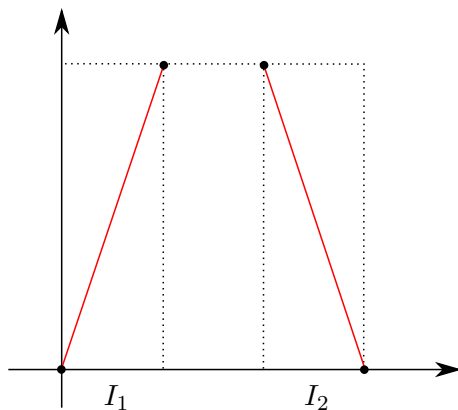
As condições (1) e (2) da Definição 145 são automáticas. Para checar (3), basta notar que:

- $f^{-1}([0, 1]) = I_1 \cup I_2 = K_1$ .
- $f^{-2}([0, 1]) = f^{-1}(K_1) = K_2$ .
- Por indução  $f^{-n}([0, 1]) = K_n$ .

Assim,  $K = \bigcap_{n \geq 0} f^{-n}(I_1 \cup \dots \cup I_r)$ .

**Exemplo 147** | Existem outras maneiras de provar que o conjunto de Cantor clássico é dinamicamente definido. Podemos também tomar  $g : I_1 \cup I_2 \rightarrow [0, 1]$  por

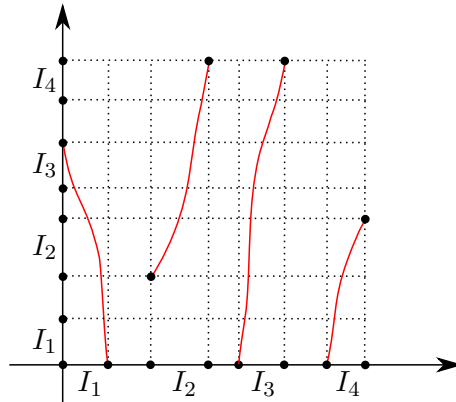
$$f(x) = \begin{cases} 3x & , \text{ se } x \in I_1 \\ 3 - 3x & , \text{ se } x \in I_2 \end{cases}$$



O conjunto de Cantor gerado por  $g$  coincide com  $K$ .

**Exemplo 148** | Considere  $r = 4$  e  $f$  dada pela figura abaixo.





É bem mais difícil identificar quem é o conjunto de Cantor dinamicamente definido.

Como vimos no último exemplo, para  $f$  geral é muito difícil identificar o respectivo conjunto de Cantor. Entender as propriedades geométricas desses conjuntos de fato constitui uma área de pesquisa atual e profunda. Por outro lado, do ponto de vista combinatório, podemos realizar a mesma construção da Seção 31:

- Tome  $G = (V, E)$  onde  $V = \{I_1, \dots, I_r\}$  é a partição de Markov e  $E = \{I_i \rightarrow I_j : f(I_i) \supset I_j\}$ .
- Considere o TMS unilateral  $(\Sigma^+, \sigma_+)$  associado a  $G$ .
- Defina  $\pi : \Sigma^+ \rightarrow [0, 1]$  por

$$\{\pi(\underline{x})\} = \bigcap_{n \geq 0} f^{-n}(I_{x_n}).$$

Assim como na seção anterior,  $\pi$  comuta o diagrama abaixo:

$$\begin{array}{ccc} \Sigma^+ & \xrightarrow{\sigma_+} & \Sigma^+ \\ \pi \downarrow & & \downarrow \pi \\ K & \xrightarrow{f} & K \end{array}$$

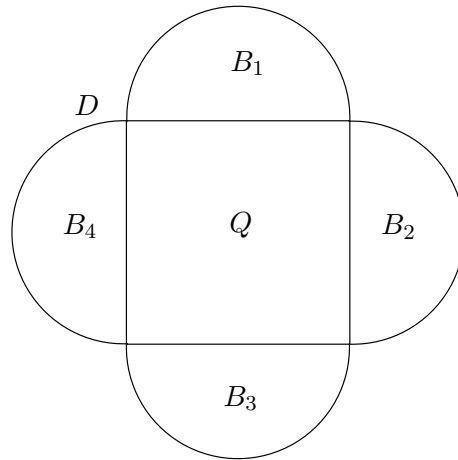
SEÇÃO 33

**Ferradura de Smale**

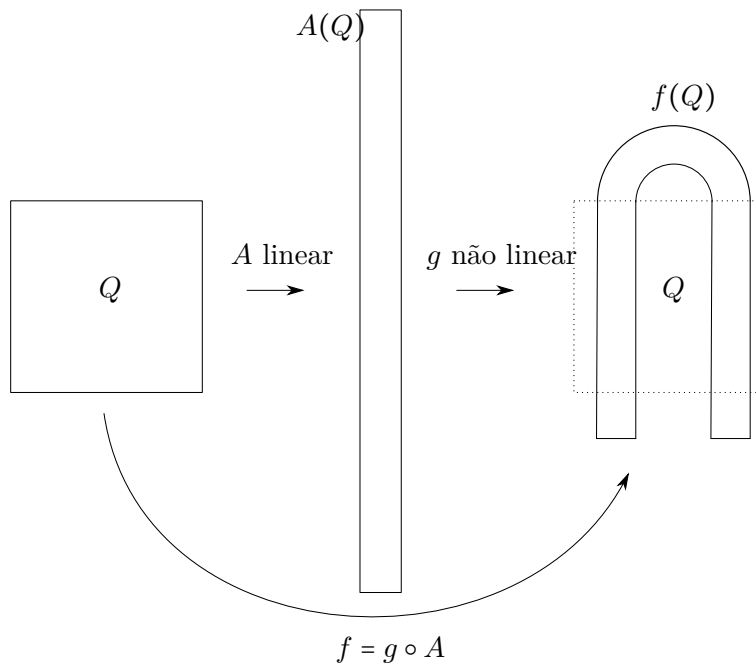
A ferradura de Smale revolucionou a área de sistemas dinâmicos. Antes de sua descoberta, feita por Steve Smale em 1962, parte da comunidade (e o próprio Smale) acreditava que a maior parte dos sistemas dinâmicos era bem comportada. Em particular, eles deveriam ser do tipo gradiente, possuindo uma quantidade finita de pontos periódicos. O mecanismo descoberto por Smale mostrou que não: existem dinâmicas robustas (permanecem sendo as mesmas após pequenas perturbações) que possuem uma infinidade de pontos periódicos, dentre

outras propriedades. Esse mecanismo se mostrou bem presente na teoria ao ponto de, em alguns contextos (por exemplo em dimensão baixa), ser o único responsável por caoticidade. Hoje em dia, conhecemos a dinâmica descoberta por Smale de *ferradura de Smale* (em inglês, Smale's horseshoe). Vamos descrever como construí-la.

Começamos considerando o domínio  $D$  abaixo, formado pela união do quadrado  $Q$  e de quatro semicírculos  $B_1, B_2, B_3, B_4$ :

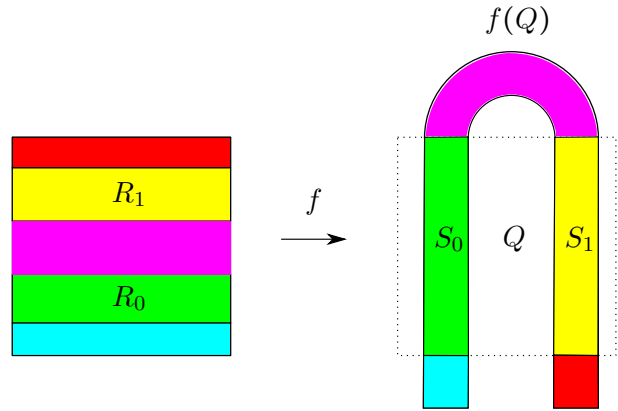


Vamos definir uma transformação  $f : D \rightarrow D$ . Definimos  $f$  na união  $B_1 \cup Q \cup B_3$  de modo que a imagem desse domínio seja uma ferradura, definida pela composição de uma transformação linear e outra não linear no terço médio do retângulo obtido:

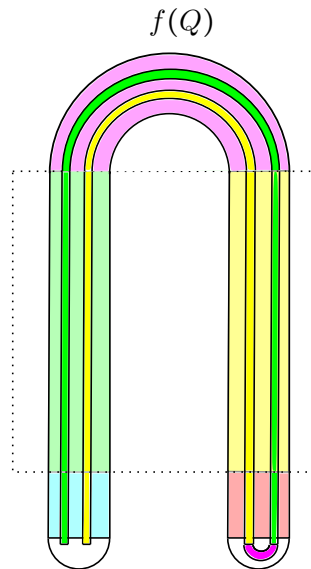


No que segue, olharemos apenas para os pontos que permanecem no quadrado por todo o tempo. Em particular, não nos preocupamos como  $f$  é definida

em  $B_2 \cup B_4$  (os pontos nessa união não pertencem a  $Q$  no tempo  $n = 0$ ). A dinâmica de  $f$  em  $Q$  pode ser melhor descrita com a divisão abaixo:



Acima,  $f(Q)$  intersecta  $Q$  em dois retângulos verticais  $S_0 = f(R_0)$ , que chamamos de perna esquerda da ferradura, e  $S_1 = f(R_1)$ , que chamamos de perna direita. Desse modo, os pontos que estão em  $Q$  para  $n = 0, 1$  são a interseção  $Q \cap f^{-1}(Q) = R_0 \cup R_1$ . Para entender os pontos que permanecem em  $Q$  para  $n = 2$ , notamos que  $f^2(Q)$  tem o formato de uma ferradura “dupla”, conforme a figura abaixo:



**Figura 1.** A região verde mais escura é igual a  $f(S_0)$ , e a região amarela mais escura é igual a  $f(S_1)$ .

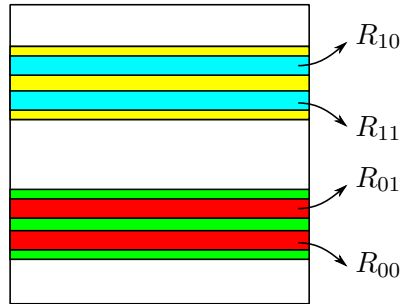
Portanto,  $f^2(Q)$  intersecta  $Q$  em quatro retângulos verticais, dois deles contidos em  $S_0$  e dois deles em  $S_1$ . Assim,

$$Q \cap f^{-1}(Q) \cap f^{-2}(Q) = \underbrace{R_{00} \cup R_{01}}_{\subset R_1} \cup \underbrace{R_{10} \cup R_{11}}_{\subset R_2},$$

onde cada  $R_{ij}$  é um retângulo horizontal da forma  $[0, 1] \times I_{ij}$ . Note que:

- $R_{00}$  = pontos  $x \in Q$  tais que  $f(x), f^2(x) \in S_0$ .
- $R_{01}$  = pontos  $x \in Q$  tais que  $f(x) \in S_0$  e  $f^2(x) \in S_1$ .
- $R_{10}$  = pontos  $x \in Q$  tais que  $f(x) \in S_1$  e  $f^2(x) \in S_0$ .
- $R_{11}$  = pontos  $x \in Q$  tais que  $f(x), f^2(x) \in S_1$ .

Noutras palavras, os índices descrevem o itinerário com respeito aos retângulos  $S_0, S_1$ .



Por indução, temos

$$Q \cap f^{-1}(Q) \cap \dots \cap f^{-n}(Q) = \bigcup_{i,j=0,1} R_{i_1, \dots, i_n},$$

onde

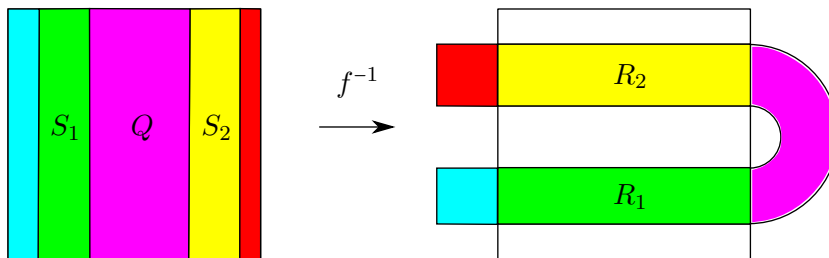
$$R_{i_1, \dots, i_n} = [0, 1] \times I_{i_1, \dots, i_n} = \{x \in Q : f^j(x) \in S_{i_j}\}.$$

Desse modo,

$$\bigcap_{n \geq 0} f^{-n}(Q) = [0, 1] \times K_2,$$

onde  $K_2$  é um conjunto de Cantor. O conjunto de Cantor  $K_2$  é dinamicamente definido por uma função com dois ramos completos. Em suma: um ponto permanece em  $Q$  para toda iteração positiva se e somente se sua segunda coordenada está no conjunto de Cantor  $K_2$ .

Similarmente, definimos  $f^{-1}$  em  $B_2 \cup Q \cup B_4$  em forma de uma ferradura (horizontal) de modo que  $f^{-1}(S_i) = R_i$  para  $i = 0, 1$ , conforme a figura abaixo:



Assim, procedendo analogamente, obtemos que

$$\bigcap_{n \leq 0} f^{-n}(Q) = K_1 \times [0, 1],$$

onde  $K_1$  é um conjunto de Cantor. Portanto, um ponto permanece em  $Q$  para toda iteração negativa se e somente se sua primeira coordenada está no conjunto de Cantor  $K_1$ .

Finalmente, obtemos que

$$\bigcap_{n \in \mathbb{Z}} f^n(Q) = K_1 \times K_2$$

é o produto cartesiano de dois conjuntos de Cantor. Do ponto de vista combinatório, todo  $x \in K_1 \times K_2$  está associado a um itinerário  $(\dots, i_{-1}; i_0, \dots)$  e, reciprocamente, todo itinerário  $(\dots, i_{-1}; i_0, \dots)$  está associado a um ponto em  $K_1 \times K_2$ . Assim, podemos definir uma bijeção  $\pi : \Sigma = \{0, 1\}^{\mathbb{Z}} \rightarrow K_1 \times K_2$  por

$$\pi(\dots, i_{-1}; i_0, \dots) = \left[ \bigcap_{n \geq 0} f^{-n}(S_{i_n}) \right] \cap \left[ \bigcap_{n < 0} f^{-n}(R_{i_n}) \right],$$

ou seja,  $x$  é o único ponto tal que  $f^n(x) \in S_{i_n}$  para  $n \geq 0$  e  $f^n(x) \in R_{i_n}$  para  $n < 0$ . Note que a primeira interseção  $\bigcap_{n \geq 0} f^{-n}(S_{i_n})$  corresponde a um segmento horizontal, que define a órbita futura de  $x$ , enquanto  $\bigcap_{n \leq 0} f^{-n}(R_{i_n})$  corresponde a um segmento vertical, que define a órbita passada de  $x$ . Visto desse modo, é claro que  $\pi$  está bem definida. Considerando o TMS bilateral  $(\Sigma, \sigma)$ , temos também que  $\pi$  comuta o diagrama abaixo:

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma} & \Sigma \\ \pi \downarrow & & \downarrow \pi \\ K_1 \times K_2 & \xrightarrow{f} & K_1 \times K_2 \end{array}$$

Logo, a dinâmica de  $f$  restrita a  $K_1 \times K_2$  é a dinâmica de um shift completo de dois símbolos. Em particular, obtemos que  $f^n$  tem exatamente  $2^n$  pontos fixos. Essa conclusão é praticamente impossível de obter se olharmos apenas para a ferradura geometricamente. Outras propriedades, por exemplo sobre medidas invariantes, também podem ser melhor entendidas do ponto de vista simbólico.

## SEÇÃO 34

### Automorfismos lineares hiperbólicos do toro

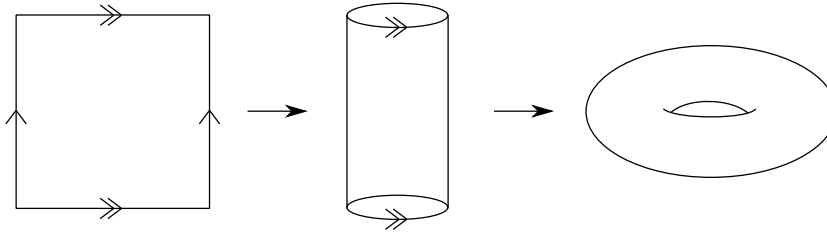
Relembre a definição do círculo  $\mathbb{S}^1 = \mathbb{R}/\sim = \mathbb{R}/\mathbb{Z}$ , onde  $x \sim y \iff x - y \in \mathbb{Z}$ . Podemos considerar a versão multidimensional dessa definição, conforme agora explicamos. Em  $\mathbb{R}^n$ , considere a relação de equivalência  $\sim$  dada por  $x \sim y \iff x - y \in \mathbb{Z}^n$  (Exercício 59).

**Definição 149**

(Toro  $\mathbb{T}^n$ ) O *toro  $n$ -dimensional*  $\mathbb{T}^n$  é definido pelo quociente

$$\mathbb{T}^n = \mathbb{R}^n / \sim = \mathbb{R}^n / \mathbb{Z}^n.$$

Denotamos a projeção canônica definida por  $\sim$  por  $\pi : \mathbb{R}^n \rightarrow \mathbb{T}^n$ . Quando  $n = 1$ , temos que  $\mathbb{T} = \mathbb{S}^1$  conforme definido na Seção 37. Nessa seção, focamos no caso  $n = 2$ . Chamamos  $\mathbb{T}^2$  de toro bidimensional plano. Vejamos as justificativas dessa nomenclatura. O primeiro passo é visualizar  $\mathbb{T}^2$  topologicamente. O produto Note que a restrição de  $\pi$  ao quadrado  $[0, 1]^2$  é sobrejetiva em  $\mathbb{T}^2$  e só deixa de ser injetiva nos lados:  $(1, y) \sim (0, y)$  e  $(x, 1) \sim (x, 0)$  para  $0 \leq x, y, \leq 1$ . Desse modo,  $\mathbb{T}^2$  é obtido tomando o quadrado  $[0, 1]^2$  com a identificação dos dois lados verticais e dos dois lados horizontais. A figura abaixo mostra que tais identificações definem um toro do ponto de vista topológico:



Geometricamente, a curvatura de  $\mathbb{T}^2$  é igual à de  $\mathbb{R}^2$ , i.e. é nula. Por fim, observamos que  $\mathbb{T}^2$  é compacto.

Nosso objetivo é introduzir e estudar uma classe de transformações em  $\mathbb{T}^2$ . Para isso, tomamos transformações em  $\mathbb{R}^2$  e projetamos. Como no caso do círculo (veja a Seção 22),  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  se projeta em  $f : \mathbb{T}^2 \rightarrow \mathbb{T}^2$  via  $f([x]) = [F(x)]$  se e somente se  $F$  satisfaz a seguinte condição:

$$x \sim y \implies F(x) \sim F(y). \quad (34.1)$$

Considere a matriz  $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$ , e seja  $F = A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  induzido por  $A$ , i.e.

$F(x) = Ax$ . É fácil ver que  $F$  satisfaz (34.1). De fato, se  $x - y = \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{Z}^2$  então

$$F(x) - F(y) = Ax - Ay = A(x - y) = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2a + b \\ a + b \end{bmatrix} \in \mathbb{Z}^2.$$

Essa propriedade vale para qualquer matriz de coordenadas inteiras.

Afirmamos que  $F$  é invertível. De fato, como  $A$  possui determinante 1, sua inversa  $A^{-1} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$  também possui entradas inteiras, e portanto induz  $g : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ . Como  $AA^{-1} = A^{-1}A = \text{Id}_{\mathbb{R}^2}$ , segue que  $f \circ g = g \circ f = \text{Id}_{\mathbb{T}^2}$ . Essas conclusões valem para qualquer  $A \in \text{GL}(2, \mathbb{Z})$  com determinante  $\pm 1$  (Exercício 60). Chegamos à classe de transformações de interesse.

**Definição 150**

(Automorfismo linear de  $\mathbb{T}^2$ ) Um *automorfismo linear do toro*  $\mathbb{T}^2$  é uma transformação  $f : \mathbb{T}^2 \rightarrow \mathbb{T}^2$  induzida por uma matriz  $A \in \text{GL}(2, \mathbb{Z})$  com  $\det(A) = \pm 1$ .

O automorfismo linear induzido por  $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$  acima é chamado *mapa de*

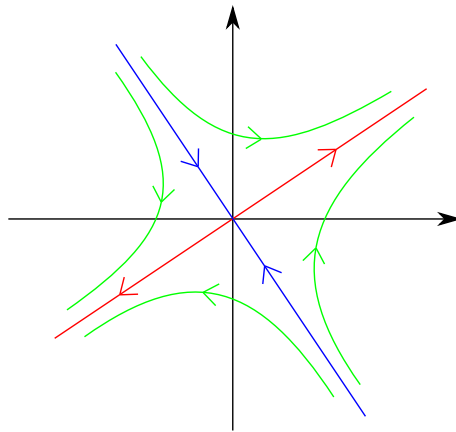
*Arnold* ou *cat map*. No restante dessa seção, focamos nossa atenção no cat map, buscando entender como representá-lo de modo mais simples de modo a entender suas propriedades dinâmicas, e.g. a quantidade de pontos periódicos. O cat map constitui um exemplo clássico de automorfismo linear *hiperbólico*. Vejamos o que isso significa.

O polinômio característico de  $A$  é  $p(\lambda) = \lambda^2 - \text{tr}(A)\lambda + \det(A) = \lambda^2 - 3\lambda + 1$ , logo  $A$  possui autovalores  $\lambda_1 = \frac{3+\sqrt{5}}{2}$  e  $\lambda_2 = \frac{3-\sqrt{5}}{2}$ . Como  $0 < |\lambda_2| < 1 < |\lambda_1|$ ,  $A$  é uma matriz hiperbólica. Assim,  $f$  é um automorfismo linear hiperbólico.

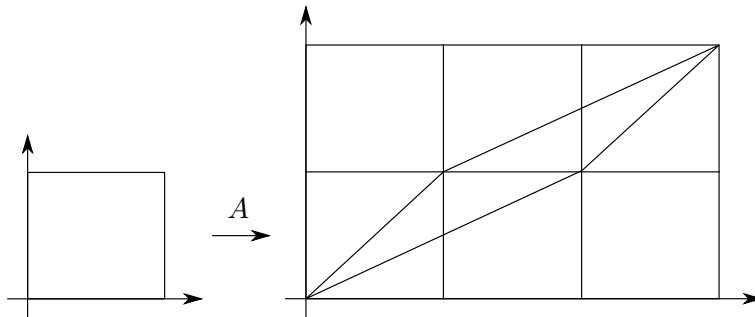
Para compreender  $f$ , compreendemos antes  $F$ . É fácil encontrar autodireções  $v_1 = \begin{bmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{bmatrix}$  e  $v_2 = \begin{bmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{bmatrix}$ . Note que  $v_1, v_2$  são ortogonais (de fato,  $A$  é simétrica). É claro que:

- $v \in \langle v_1 \rangle \implies F^n(v) = \lambda_1^n v$  para todo  $n \in \mathbb{Z}$ .
- $v \in \langle v_2 \rangle \implies F^n(v) = \lambda_2^n v$  para todo  $n \in \mathbb{Z}$ .

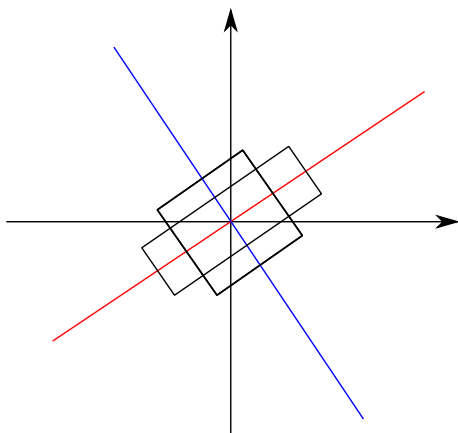
Mais geralmente, se  $v = Av_1 + Bv_2$  então  $F^n(v) = A\lambda_1^n v_1 + B\lambda_2^n v_2$  para todo  $n \in \mathbb{Z}$ . Isso caracteriza completamente o comportamento dinâmico de  $F$ , descrito pela figura abaixo:



Esse comportamento é chamado de *hiperbólico* ou *do tipo sela*. O que as equações acima dizem é que as autodireções, por serem direções invariantes, permitem mais facilmente compreender  $F$ . De fato, se queremos compreender  $F$  por meio dos eixos coordenados  $x$  e  $y$ , então retângulos são levados em paralelogramos inclinados:

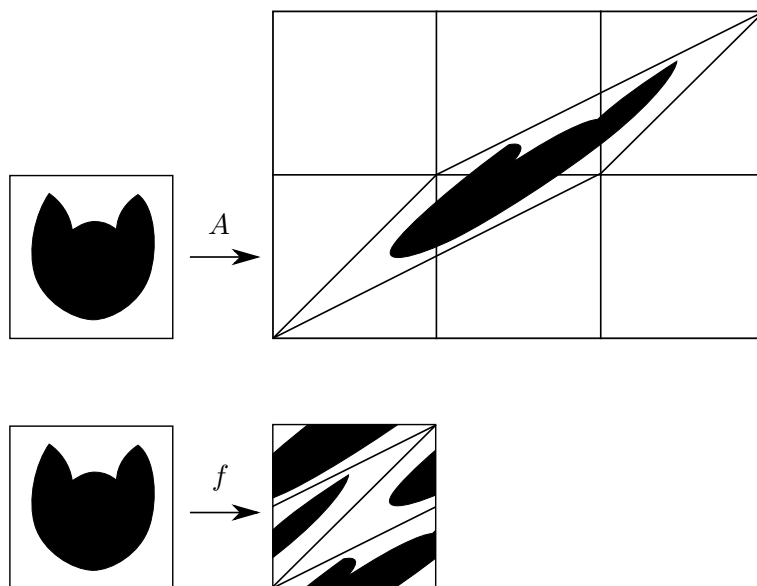


Por outro lado, retângulos com lados paralelos às autodireções preservam essa propriedade por iteração de  $F$ :



Do ponto de vista dinâmico, as autodireções de  $A$  têm a seguinte nomenclatura:  $\langle v_1 \rangle$  é chamado de direção *expansora* ou *instável*, e  $\langle v_2 \rangle$  é chamado de direção *contratora* ou *estável*.

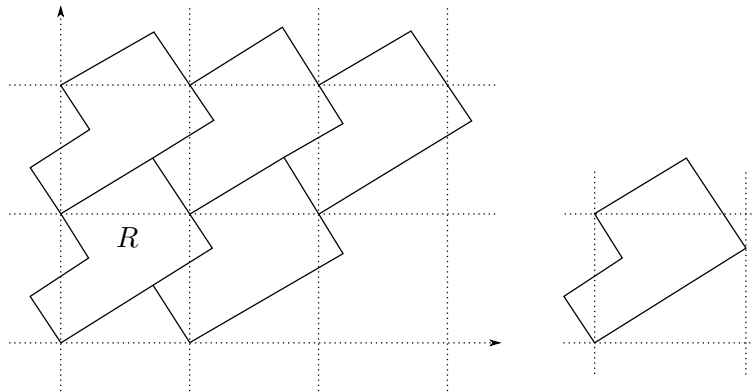
Agora, vamos entender  $f$ . Conforme explicado acima, os eixos coordenados não fornecem um bom sistema de coordenadas para isso. De fato, se considerarmos o domínio fundamental  $[0, 1)^2$ , então a ação de  $f$  é descrita pela figura abaixo, onde aplicamos a matriz  $A$  ao domínio fundamental e depois aplicamos translações inteiras às partes exteriores ao domínio fundamental para que retornem para ele.



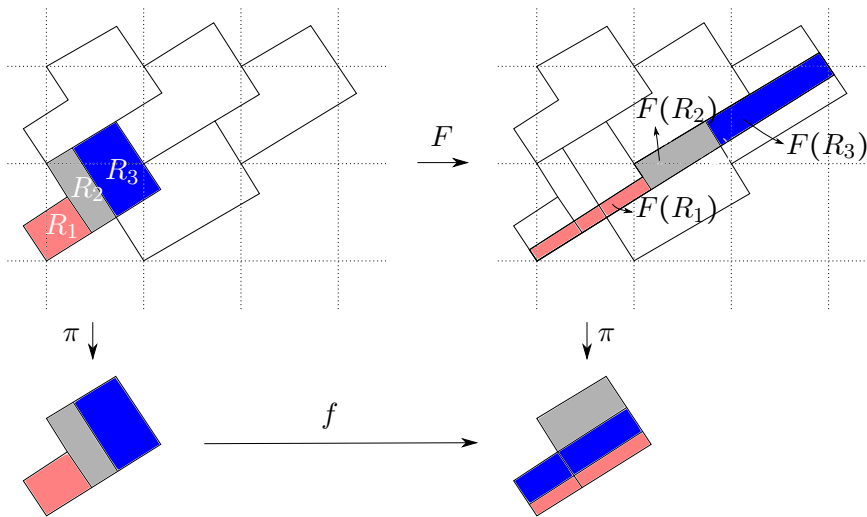
Veja como o gato é deformado no toro. Entender a segunda iteração em diante se torna algo difícil se continuarmos utilizando o domínio fundamental  $[0, 1)^2$ . Adler & Weiss propuseram, em 1967, um modo diferente de representar  $f$ . A ideia é buscar um domínio fundamental formado por lados paralelos



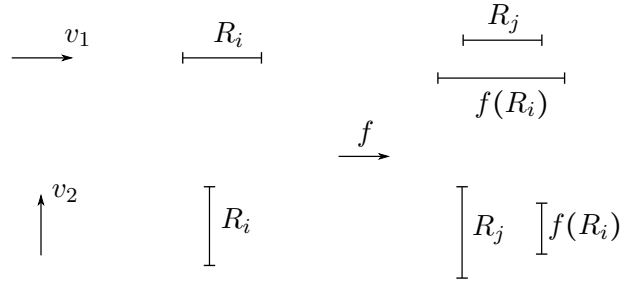
às autodireções. Existem várias possibilidades. Abaixo, mostramos uma, que utilizaremos na sequência.



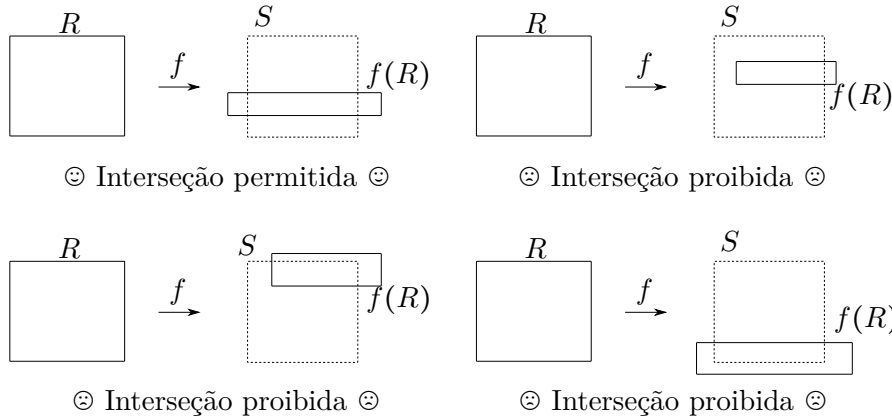
É fácil notar que a figura à direita de fato é um domínio fundamental do toro. Seja  $R$  a cópia que intersecta a origem  $(0,0)$  como na figura. Vamos entender a ação de  $f$  por meio da ação de  $F$  sobre  $R$ . A imagem  $F(R) = AR$  é descrita na figura abaixo. Temos  $[F(R)] = f([R]) = f(\mathbb{T}^2) = \mathbb{T}^2$ . De fato,  $F(R)$  também é domínio fundamental de  $\mathbb{T}^2$  (Exercício 61). Para entender  $f$ , notamos que  $F(R)$  intersecta três cópias de  $R$ . A ação de  $f$  consiste em transladar essas cópias todas para um mesmo domínio fundamental. Para isso, divida  $R$  em três retângulos  $R_1, R_2, R_3$ , como na figura abaixo, de modo que a imagem de cada retângulo intersecta exatamente uma única cópia de  $R$ . Logo, a ação de  $f$  nos retângulos é dada pela parte inferior da figura.



Esses retângulos satisfazem duas propriedades de Markov, uma para cada autodireção. Na direção instável, os retângulos satisfazem a propriedade usual “se intersecta, então contém” com respeito a  $f$ . Na direção estável, os retângulos satisfazem essa propriedade com respeito à inversa  $f^{-1}$ ; com respeito a  $f$ , a propriedade vira “se intersecta, então está contido” (de fato, a direção estável para  $f$  é instável para  $f^{-1}$ ). Veja a figura:

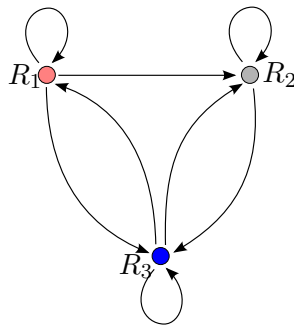


Juntas, essas duas propriedades de Markov unidimensionais definem a propriedade de Markov bidimensional. Tal propriedade pode ser informalmente descrita como “se intersecta, então cruza de um lado ao outro”. Mostramos abaixo a maneira que tal propriedade de Markov é satisfeita, e três exemplos em ela não é.



Uma vez que  $R_1, R_2, R_3$  satisfazem a propriedade de Markov acima, podemos concatenar arestas como fizemos nos exemplos unidimensionais: escrevendo  $R_i \rightarrow R_j$  quando  $f(R_i) \cap R_j \neq \emptyset$ , se  $R_i \rightarrow R_j$  e  $R_j \rightarrow R_k$ , então existe  $x \in R_i$  tal que  $f(x) \in R_j$  e  $f^2(x) \in R_k$ . Essa propriedade de concatenação de arestas permite-nos repetir a construção simbólica feita nas seções anteriores, tomando:

- o  $G = (V, E)$  onde  $V = \{R_1, R_2, R_3\}$  e  $E = \{R_i \rightarrow R_j : f(R_i) \cap R_j \neq \emptyset\}$ . O grafo é



e a matriz de adjacência de  $G$  é

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

- $(\Sigma, \sigma)$  é o TMS bilateral definido por  $G$ .
- $\pi : \Sigma \rightarrow \mathbb{T}^2$  é definido por

$$\begin{aligned} \{\pi(\dots, R_{i_{-1}}; R_{i_0}, R_{i_1}, \dots)\} &= \bigcap_{n \in \mathbb{Z}} f^{-n}(R_{i_n}) \\ &= \bigcap_{n \geq 0} \underbrace{f^n(R_{i_{-n}}) \cap \dots \cap R_{i_0} \cap \dots \cap f^{-n}(R_{i_n})}_{=F_n}, \end{aligned}$$

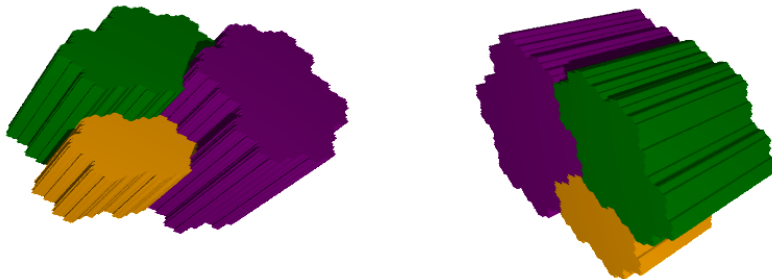
onde  $\{F_n\}_{n \geq 0}$  é uma família descendente de fechados com diâmetros convergindo a zero (Exercício 62).

- $\pi$  comuta o diagrama abaixo:

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma} & \Sigma \\ \pi \downarrow & & \downarrow \pi \\ \mathbb{T}^2 & \xrightarrow{f} & \mathbb{T}^2 \end{array}$$

É importante ressaltar que a situação acima, em que os conjuntos que formam a partição de Markov são retângulos geométricos, é bem particular da dimensão dois (em que as autodireções têm codimensão um). De fato, Bowen mostrou em 1978 que para alguns automorfismos lineares hiperbólicos em dimensão três, os bordos dos conjuntos não podem ser suaves. Nas figuras abaixo<sup>2</sup>, mostramos uma partição de Markov para o automorfismo linear induzido pela

matriz  $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ , vista sob duas perspectivas:



Note que os bordos dos conjuntos são fractais.

<sup>2</sup>Figuras elaboradas por Timo Jolivet utilizando o programa Sage.

## SEÇÃO 35

## Exercícios

**Exercício 56** | Com respeito a  $f(x) = 2x \pmod{1}$ , identifique os seguintes conjuntos:

(1)  $\{x \in [0, 1] : a_1 = 0, a_2 = 1\}$ .

(2)  $\{x \in [0, 1] : a_2 = 1\}$ .

(3)  $\{x \in [0, 1] : a_1 = 1, a_3 = 0\}$ .

**Exercício 57** | Prove que o conjunto de Cantor clássico é igual a

$$K = \left\{ \sum_{n \geq 0} \frac{a_n}{3^n} : a_n \in \{0, 2\} \right\}.$$

Use essa igualdade para provar que  $K + K = [0, 2]$ .

**Exercício 58** | Para  $N \in \mathbb{N}$ , seja  $\mathcal{A} = \{0, 1, \dots, N-1\}$  e considere o espaço simbólico  $\mathcal{A}^{\{0,1,\dots\}}$ . A ordem lexicográfica em  $\mathcal{A}^{\{0,1,\dots\}}$  é definida da seguinte maneira: se  $\underline{x}, \underline{y}$  são distintos e  $n = N(\underline{x}, \underline{y})$ , dizemos que  $\underline{x} < \underline{y}$  se  $x_n < y_n$ .

(1) Mostre que  $<$  é uma ordem total em  $\mathcal{A}^{\{0,1,\dots\}}$ .

(2) Mostre que para  $g$  no Exemplo 147,  $\pi : \Sigma^+ \rightarrow K$  definida na Seção 32 preserva  $<$ , i.e. se  $\underline{x} < \underline{y}$  então  $\pi(\underline{x}) < \pi(\underline{y})$ .

**Exercício 59** | Mostre que a relação  $x \sim y \iff x - y \in \mathbb{Z}^n$  em  $\mathbb{R}^n$  é de equivalência.

**Exercício 60** | Se  $A \in \text{GL}(2, \mathbb{Z})$  tem determinante  $\pm 1$ , então  $A$  induz um homeomorfismo em  $\mathbb{T}^2$ .

**Exercício 61** | Utilizando a notação da Seção 34, mostre que  $F(R)$  é domínio fundamental do toro.

**Exercício 62** | Utilizando a notação da Seção 34, vamos provar que  $F_n$  tem diâmetro convergindo a zero. Sejam  $\lambda_1 > 1 < \lambda_2$  os autovalores de  $A$ . Mostre que na direção estável  $R_{i_0} \cap \dots \cap f^{-n}(R_{i_n})$  tem diâmetro menor que  $\lambda_2^n$ , e que na direção instável  $f^n(R_{i_{-n}}) \cap \dots \cap R_{i_0}$  tem diâmetro menor que  $\lambda_1^{-n}$ . Conclua que  $F_n$  é um retângulo com lados menores que  $\lambda_2^n$ .

# Apêndice

## SEÇÃO 36

## Convergência uniforme

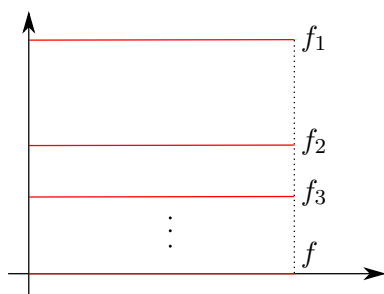
Vamos relembra alguns fatos elementares sobre convergência de funções. Dado  $A \subset \mathbb{R}^m$ , seja  $X = \{f : A \rightarrow \mathbb{R}^d\}$ . Uma sequência em  $X$  é uma sequência de funções  $(f_n)_{n \geq 1}$ .

**Definição 151** (Convergência pontual) Dizemos que  $(f_n)_{n \geq 1} \subset X$  converge pontualmente para  $f : A \rightarrow \mathbb{R}^d$  se  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  para todo  $x \in A$ .

**Exemplo 152** Tome  $f_n : [0, 1] \rightarrow \mathbb{R}^d$ ,  $f_n(t) \equiv \frac{1}{n}$ . Então  $(f_n)_{n \geq 1}$  converge pontualmente para  $f(x) \equiv 0$ .

**Exemplo 153** Tome  $f_n : [0, 1] \rightarrow \mathbb{R}^d$ ,  $f_n(t) = t^n$ . Temos:

- $t = 1$ :  $\lim f_n(1) = \lim 1 = 1$ .
- $t < 1$ :  $\lim f_n(t) = \lim t^n = 0$ .



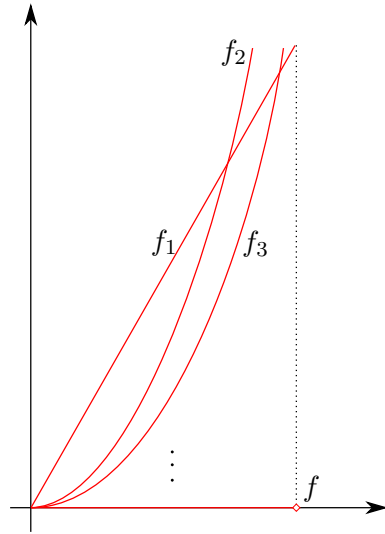
Assim,  $(f_n)_{n \geq 1}$  converge pontualmente para

$$f(x) = \begin{cases} 0, & 0 \leq t < 1 \\ 1, & t = 1. \end{cases}$$

Embora cada  $f_n$  seja contínua, o limite pontual não o é.

**Exemplo 154** Tome  $f_n(t) = (n+1)t^n$ ,  $\forall n \geq 1$ . Temos:

- $t = 1$ :  $\lim f_n(1) = \lim(n+1) = \infty$ .
- $t < 1$ :  $\lim f_n(t) = \lim(n+1)t^n = 0$ , pois toda exponencial cresce mais rápido do que polinômios.



Assim,  $(f_n)_{n \geq 1}$  converge pontualmente para

$$f(x) = \begin{cases} 0, & 0 \leq t < 1 \\ \infty, & t = 1. \end{cases}$$

Em cada exemplo acima, calcule  $\lim \int_0^1 f_n(t) dt$  e compare com  $\int_0^1 f(t) dt$ . Nos dois primeiros casos, temos igualdade e no último temos

$$\lim \int_0^1 f_n(t) dt = \lim 1 = 1 \neq 0 = \int f(t) dt.$$

Portanto, convergência pontual não preserva integral. Para garantir isso, devemos requerer alguma propriedade mais forte. A convergência uniforme é uma de tais propriedades. Considere a norma do supremo:  $d(f, g) = \sup\{\|f(x) - g(x)\| : x \in A\}$ .

**Definição 155**

(Convergência uniforme) Dizemos que  $(f_n)_{n \geq 1} \subset X$  converge uniformemente para  $f : A \rightarrow \mathbb{R}^d$  se, para todo  $\varepsilon > 0$ , existe  $n_0 > 0$  tal que  $d(f_n, f) < \varepsilon$  para todo  $n \geq n_0$ , i.e.

$$|f_n(x) - f(x)| < \varepsilon, \quad \forall n \geq n_0, \forall x \in A.$$

Essa hipótese é mais forte do que convergência pontual (exercício). Algumas propriedades básicas de convergência uniforme (exercício):

- Se  $f_n$  converge uniformemente para  $f$  e  $g_n$  converge uniformemente para  $g$ , então  $f_n + g_n$  converge uniformemente para  $f + g$ .
- Se  $f_n : A \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$  converge uniformemente para  $f$  e  $g_n : B \subset \mathbb{R}^n \rightarrow \mathbb{R}^d$  converge uniformemente para  $g$ , então  $g_n \circ f_n$  converge uniformemente para  $g \circ f$ .

**Proposição 156**

Se  $A$  é compacto e  $(f_n)_{n \geq 1} \in X$  converge uniformemente para  $f : A \rightarrow \mathbb{R}^d$ , então

$$\int f(x) dx = \lim_{n \rightarrow \infty} \int f_n(x) dx.$$

Veja, por exemplo, o livro de Análise do Elon. Em particular, a convergência do último exemplo *não* é uniforme.

**Teorema 157** (Aproximação de Weierstrass) Os polinômios são densos em espaços de funções contínuas. Mais especificamente: dada  $f : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^d$  contínua e  $K \subset U$  compacto, existe uma sequência de polinômios  $(f_n)_{n \geq 1}$  tal que suas restrições a  $K$  convergem uniformemente para  $f \upharpoonright_K$ .

**Exemplo 158** Para uma função igual à sua série de Taylor, esse resultado é simples. Por exemplo, se  $f : [0, 1] \rightarrow \mathbb{R}$ ,  $f(t) = e^t$ , podemos tomar

$$f_n(t) = 1 + t + \frac{t^2}{2!} + \dots + \frac{t^n}{n!}.$$

Outro exemplo: se  $f(t) = \cos t$ , tome

$$f_n(t) = 1 - \frac{t^2}{2!} + \dots + (-1)^n \frac{t^{2n}}{(2n)!}.$$

O último resultado importante sobre convergência uniforme é o seguinte.

**Teorema 159** (Arzelà-Ascoli) Sejam  $K \subset \mathbb{R}^m$  um conjunto compacto e  $(f_n)_{n \geq 1}$  uma sequência de funções  $f_n : K \rightarrow \mathbb{R}^d$  que é:

(i) *Equicontínua*: para todo  $\varepsilon > 0$  existe  $\delta > 0$  tal que

$$x, y \in K, \|x - y\| < \delta \implies \|f_n(x) - f_n(y)\| < \varepsilon, \forall n \geq 1.$$

(ii) *Equilimitada*: existe  $M > 0$  tal que

$$\|f_n(x)\| < M, \forall x \in K, \forall n \geq 1.$$

Então  $(f_n)_{n \geq 1}$  possui uma subsequência que converge uniformemente.

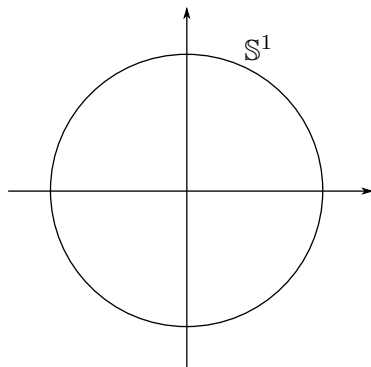
Como cada  $f_n$  é contínua, o limite uniforme da subsequência é uma função contínua. Em termos práticos, equicontinuidade é garantida pela limitação nas derivadas: se existe  $L > 0$  tal que  $\|f'_n(x)\| < L$  para todos  $x \in K$  e  $n \geq 1$ , então pela desigualdade do valor médio temos equicontinuidade: dado  $\varepsilon > 0$ , tome  $\delta = \frac{\varepsilon}{L}$ .

## SEÇÃO 37

### O círculo $\mathbb{S}^1$

Em geometria, definimos a circunferência unitária  $\mathbb{S}^1$  por

$$\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} = \{z \in \mathbb{C} : \|z\| = 1\}.$$



Tal conjunto é unidimensional e possui várias propriedades geométricas e métricas importantes, por exemplo:  $\mathbb{S}^1$  é compacto, possui curvatura constante igual a 1 e *não* é simplesmente conexo. Aqui, estamos considerando a métrica induzida pela métrica euclidiana usual de  $\mathbb{R}^2$ . Para o que estamos interessados, algumas das propriedades citadas acima não são relevantes. Por exemplo, não nos preocupamos com a curvatura de  $\mathbb{S}^1$ , e tampouco com seu comprimento euclidiano (igual a  $2\pi$ ). Para o que estamos interessados, podemos definir  $\mathbb{S}^1$  de modo diferente, em que as propriedades de interesse são preservadas e as não essenciais não são levadas em conta. Defina a seguinte relação de equivalência em  $\mathbb{R}$ :

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

É claro que  $\sim$  é relação de equivalência:

- REFLEXIVIDADE:  $x \sim x$ , pois  $x - x = 0 \in \mathbb{Z}$ .
- SIMETRIA: temos  $x \sim y \iff y \sim x$ , pois  $x - y \in \mathbb{Z} \iff y - x \in \mathbb{Z}$ .
- TRANSITIVIDADE: se  $x \sim y$  e  $y \sim z$ , então

$$x - z = \underbrace{(x - y)}_{\in \mathbb{Z}} + \underbrace{(y - z)}_{\in \mathbb{Z}} \in \mathbb{Z}.$$

Dado  $x \in \mathbb{R}$ , seja

$$[x] = x + \mathbb{Z} = \{x + n : n \in \mathbb{Z}\}$$

sua classe de equivalência. Cada  $[x]$  é uma translação de  $\mathbb{Z}$  por  $x$ .

**Definição 160**

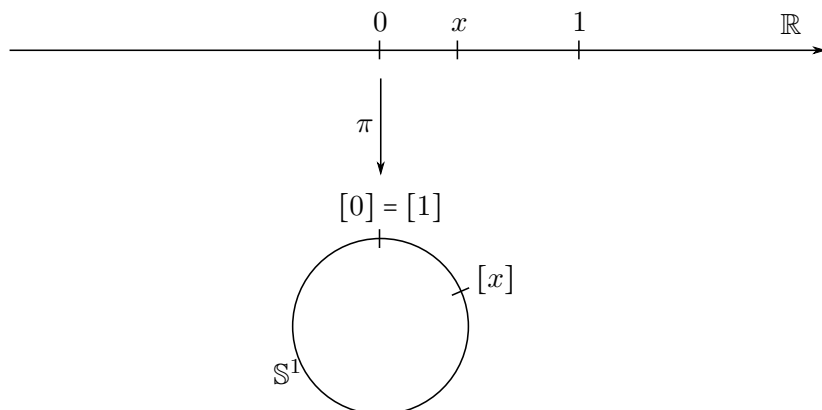
(Círculo  $\mathbb{S}^1$ ) Chamamos de *círculo*  $\mathbb{S}^1$  o espaço quociente de  $\mathbb{R}$  pela relação de equivalência  $\sim$ :

$$\mathbb{S}^1 = \mathbb{R} / \sim = \mathbb{R} / \mathbb{Z}.$$

Note que  $\mathbb{S}^1$  é, topologicamente, uma circunferência: se consideramos as classes  $[x]$  para  $0 \leq x < 1$ , então todas as classes são distintas duas a duas exceto as extremas  $[0] = [1]$ . O mesmo vale para qualquer intervalo  $[t, t + 1]$ . Em particular, qualquer intervalo  $[t, t + 1)$  é um domínio fundamental de  $\mathbb{S}^1$ , de modo que  $\mathbb{S}^1$  é obtido “enrolando”  $\mathbb{R}$  ao redor de um círculo.

A projeção canônica da relação de equivalência é  $\pi : \mathbb{R} \rightarrow \mathbb{S}^1$ ,  $\pi(x) = [x]$ .





A métrica usual  $\bar{d}(x, y) = |x - y|$  de  $\mathbb{R}$  se projeta via  $\pi$  em uma métrica  $d$  em  $\mathbb{S}^1$ :

$$d([x], [y]) = \bar{d}(\pi^{-1}[x], \pi^{-1}[y]) = \min\{|x - y + n| : n \in \mathbb{Z}\}.$$

Com respeito a essa métrica,  $\mathbb{S}^1$  possui tamanho 1. Ademais,  $\mathbb{S}^1$  é um espaço métrico compacto e completo e não é simplesmente conexo (de fato,  $\pi_1(\mathbb{S}^1) = \mathbb{Z}$ ). Essa última propriedade distingue  $\mathbb{S}^1$  do intervalo  $[0, 1]$ , e tem sérias consequências. Um delas é que o teorema do valor intermediário não é válido em  $\mathbb{S}^1$ . A razão é a seguinte: enquanto que ao retirarmos um elemento  $c$  interior a um intervalo  $[a, b]$  desconectamos o intervalo, ao fazer isso com um intervalo do círculo permanecemos com um conjunto conexo (de fato, no círculo temos duas maneiras de ir de um ponto  $[x]$  a um ponto  $[y]$ ). Desse modo, muitas provas e resultados válidos para intervalos não são válidos em  $\mathbb{S}^1$ . Por outro lado, isso permite novos fenômenos dinâmicos, como a presença de transformações sem pontos fixos.

## SEÇÃO 38

### Teorema de Kronecker

Enunciamos e provamos abaixo um resultado muito importante sobre números irracionais. Dado  $\alpha \in \mathbb{R}$ , defina

$$G_\alpha = \{m + n\alpha : m, n \in \mathbb{Z}\}.$$

**Teorema 161** (Kronecker) Se  $\alpha$  é irracional, então  $G_\alpha$  é denso em  $\mathbb{R}$ .

PROVA Note que  $G_\alpha$  é um subgrupo aditivo de  $\mathbb{R}$ , isto é,

$$g_1, g_2 \in G_\alpha \implies g_1 + g_2, g_1 - g_2 \in G_\alpha.$$

Os subgrupos aditivos de  $\mathbb{R}$  são caracterizadas pelo fato abaixo.

FATO. Se  $G$  é um subgrupo aditivo de  $\mathbb{R}$ , então uma das situações ocorre:

- (1)  $G$  é denso em  $\mathbb{R}$ .
- (2) Existe  $g_0 \in G$  tal que  $G = \{pg_0 : p \in \mathbb{Z}\}$ .

PROVA (Fato) O conjunto  $G$  é denso em  $\mathbb{R}$  se e somente se  $0$  é ponto de acumulação de  $G$ . Tome

$$g_0 = \inf(G \cap (0, +\infty)).$$

Se  $g_0 = 0$ , então  $0$  é ponto de acumulação de  $G$  e portanto ocorre (1). Suponha  $g_0 > 0$ . Afirmamos que  $g_0 \in G$ . Por absurdo, se  $g_0 \notin G$ , existem  $g_1, g_2 \in G$  tais que

$$g_0 < g_1 < g_2 < 2g_0$$

e daí  $g_2 - g_1$  é um elemento de  $G \cap (0, +\infty)$  menor que  $g_0$ , absurdo. Assim,  $g_0 \in G$ , o que estabelece a inclusão

$$G \supset \{pg_0 : p \in \mathbb{Z}\}.$$

Reciprocamente, se existe  $g \in G$  tal que  $pg_0 < g < (p+1)g_0$  para algum  $p \in \mathbb{Z}$ , então

$$g - pg_0 \in G \cap (0, +\infty) \cap (0, g_0),$$

novamente contrariando a minimalidade de  $g_0$ . □

Para concluir a prova do Teorema 161, basta observar que  $G_\alpha$  não satisfaz (2). De fato, suponha que  $g_0 = m_0 + n_0\alpha$  satisfaça

$$G = \{pg_0 : p \in \mathbb{Z}\}.$$

Como  $1, \alpha \in G$ , temos  $m_0n_0 \neq 0$ . Por outro lado, existe  $p_0 \in \mathbb{Z}$  tal que

$$-m_0 + n_0\alpha = p_0(m_0 + n_0\alpha) \implies \begin{cases} -m_0 = p_0m_0 \\ n_0 = p_0n_0 \end{cases} \implies \begin{cases} p_0 = -1 \\ p_0 = 1 \end{cases},$$

absurdo. □

### SEÇÃO 39

## Dimensão de Hausdorff em $\mathbb{R}$

Em  $\mathbb{R}$ , a medida de Lebesgue da bola de raio  $r$  é  $2r \approx r^1$ . Em  $\mathbb{R}^2$ , a medida de Lebesgue da bola de raio  $r$  é  $\pi r^2 \approx r^2$ . Os expoentes 1 e 2 nessas expressões determinam dimensões em  $\mathbb{R}$  em  $\mathbb{R}^2$  com respeito à medida de Lebesgue. Dado um conjunto  $X \subset \mathbb{R}$ , podemos tentar medir a dimensão de  $X$  de modo similar, calculando a medida de  $X \cap [x-r, x+r]$  para  $x \in X$  e  $r > 0$  e tentando achar o valor de  $d \in [0, 1]$  para o qual essa medida é aproximadamente  $r^d$ . Essa é a ideia intuitiva da dimensão de Hausdorff. Vamos proceder aos detalhes de como defini-la. Fixe  $d \in [0, 1]$ .

**Definição 162** ( $d$ -medida de Hausdorff) A  $d$ -medida de Hausdorff de  $X \subset \mathbb{R}$  é

$$m_d(X) = \liminf_{\varepsilon \rightarrow 0} \left\{ \sum_{U \in \mathcal{U}} |U|^d : \mathcal{U} \text{ é cobertura aberta de } X \text{ com } \|\mathcal{U}\| < \varepsilon \right\}.$$

Acima,  $\|\mathcal{U}\|$  denota o diâmetro de  $\mathcal{U}$ . Em particular,  $m_1$  é a medida de Lebesgue em  $\mathbb{R}$ . Note que

$$m_0(X) = \liminf_{\varepsilon \rightarrow 0} \{\#\mathcal{U} : \mathcal{U} \text{ é cobertura aberta de } X \text{ com } \|\mathcal{U}\| < \varepsilon\}$$

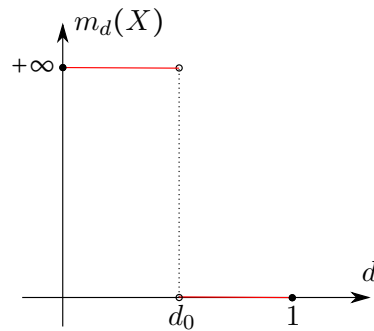
e portanto  $m_0(X) = +\infty$  sempre que  $X$  é infinito. De agora em diante, assuma isso. Outras propriedades sobre  $m_d$  são as seguintes:

- Se  $d < d'$  e  $m_{d'}(X) = +\infty$  então  $m_d(X) = +\infty$ .
- Se  $d < d'$  e  $m_d(X) = 0$  então  $m_{d'}(X) = 0$ .

**Definição 163**

(Dimensão de Hausdorff) A *dimensão de Hausdorff* de  $X$  é definida por

$$\text{HD}(X) = \sup\{d \geq 0 : m_d(X) = +\infty\} = \inf\{d \geq 0 : m_d(X) = 0\}.$$



Ou seja, a dimensão de Hausdorff é o expoente  $d_0$  onde há uma transição nos valores de  $m_d(X)$ . Em geral, porém, não sabemos de que modo a transição ocorre: pode ocorrer de  $m_{d_0}(X)$  ser infinito, finito ou zero.