

Relatório Final de Iniciação Científica - FAPESP

Dinâmica Simbólica: uma Introdução via Exemplos Hiperbólicos

Processo: #2015/16911-8.

Aluno: Pedro Griguol de Mattos.

Orientador: Prof. Dr. José Régis Azevedo Varão Filho.

Período do relatório: 10/03/2016 a 30/09/2016.

Vigência da bolsa: 01/10/2015 a 30/09/2016.

Sumário

1	Relato das Atividades do Relatório	3
2	Resumo do Estudo	4
3	Introdução	6
4	Rotação no Círculo	6
5	Cojugação de Sitemas Dinâmicos	10
6	Shift de Bernoulli	13
7	Expansão no Círculo	18
8	A Ferradura de Smale	22
9	Automorfismo Hiperbólico no Toro	26
9.1	Entropia do sistema (\mathbb{T}^2, L_A)	35
10	Cadeias de Markov	40
11	Conclusão	42

1 Relato das Atividades do Relatório

O período vigente deste projeto de iniciação científica é de 01 de outubro de 2015 a 30 de setembro de 2016. Nesse período de trabalho, realizei encontros semanais com o professor responsável. Esses encontros tinham como objetivo a apresentação dos tópicos estudados e dos resultados obtidos na semana, bem como a indicação por parte do professor de como proceder aos estudos. No segundo semestre de 2015, cursei quatro disciplinas pelo curso de Matemática do IMECC — a saber, Matemática IV, Análise I, Teoria Aritmética dos Números e Álgebra Linear Avançada, sendo esta última disciplina ministrada em conjunto para alunos de Graduação e de Mestrado. No primeiro semestre de 2016, cursei duas disciplinas pelo curso de Matemática do IMECC — Análise II e Grupos e Representações, sendo a segunda disciplina ministrada em conjunto para alunos de Graduação e de Mestrado.

2 Resumo do Estudo

O objetivo da primeira parte deste projeto de Iniciação Científica foi analisar alguns exemplos clássicos de Sistemas Dinâmicos e conhecer os principais conceitos da área. A principal referência utilizada para este trabalho foi o livro [1]. As seções §3 a §9 foram baseadas nele. Para seção §9, no entanto, foi principalmente utilizado o livro [2].

Na primeira seção, procura-se introduzir a ideia de o que é um sistema dinâmico, discorrendo sobre sua definição formal e qual a ideia por trás dessa definição. Ainda na Introdução, definem-se alguns dos principais conceitos associados ao estudo dos sistemas dinâmicos, como a noção de iterados de um elemento do sistema dinâmico e a noção de órbita de um elemento do sistema dinâmico, bem como de periodicidade da órbita.

Na seção §4, discute-se o primeiro sistema dinâmico estudado, a Rotação no Círculo. Nessa seção, são definidos o conjunto círculo, de modo propício ao estudo de uma dinâmica de rotação, e a função rotação em si, que é em seguida classificada em dois tipos: rotação racional e rotação irracional. Aparecem, ainda, a definição de métrica e de conjunto denso, e se apresenta uma métrica no círculo. Três importantes teoremas são demonstrados. Os dois primeiros relacionam rotações iracionais e racionais com a periodicidade das órbitas dos pontos do círculo. O outro afirma que a órbita de pontos sob rotações irracionais é densa no círculo.

A seção §5 apresenta a definição de conjugação entre sistemas dinâmicos e algumas proposições imediatas. Como o tema de conjugação é motivo da segunda parte deste projeto — principalmente a conjugação dos sistemas estudados na primeira parte com o shift de Bernoulli — o assunto não é muito prolongado. Na seção §6 trata-se do Shift de Bernoulli. Para isso, são definidos o conjunto das sequências e a função shift, e um teorema imediato sobre a periodicidade de sequência sob o shift é demonstrado. A seguir, exibe-se uma métrica para o conjunto das sequências, utilizada para provar mais alguns teoremas envolvendo o conceito de densidade. A seção seguinte trata da expansão no círculo. Nessa seção, define-se a expansão no conjunto círculo discutido anteriormente. As proposições provadas nessa seção se resumem a exibir uma conjugação entre a expansão no círculo e o shift de Bernoulli.

A seção §9 é a maior e mais importante seção deste trabalho. Nela é definido o objeto toro e o automorfismo hiperbólico no toro. Para isso, uma relação de equivalência é usada. São provadas algumas propriedades relacionadas à matriz descrita na definição do automorfismo. Em seguida, é definida uma métrica no toro, necessária para proposições envolvendo o conceito de densidade. Por fim, definem-se as variedades estáveis e instáveis do de um ponto do toro e prova-se que elas são densas. O conceito de uma dinâmica topologicamente transitiva é introduzido e mostra-se que o automorfismo hiperbólico o é. Na subseção, é definida a entropia de um sis-

tema dinâmico. Alguns calculos com potências da matriz do automorfismo são feitos para permitir calcular a quantidade de pontos periódicos do automorfismo e, então, a entropia do sistema do toro com o automorfismo hiperbólico.

As cadeias de Markov são brevemente tratadas na seção final, e são calculados os pontos periódicos. Na seção de conclusão do relatório, são discutidos o contato do aluno com a pesquisa e os objetivos atingidos neste relatório. A bibliografia principal adotada como base de estudo e pesquisa está referenciada no final do relatório.

3 Introdução

A noção de sistema dinâmico adotada neste trabalho é a de um *conjunto* não-vazio munido de um *endomorfismo*. Trata-se, portanto, do estudo de um conjunto X e uma função $f : X \rightarrow X$. Por simplicidade, porém, nos referimos muitas vezes à função em si como sendo a *dinâmica*, e o conjunto em que ela atua fica subentendido. Definições mais específicas de sistemas dinâmicos costumam impor certas restrições tanto a esse conjunto como à função a ele associada, mas essa abordagem foge ao escopo deste estudo de iniciação científica. O objetivo do estudo é aproximar-se da teoria dos sistemas dinâmicos através de alguns exemplos e estudá-los à luz de métodos clássicos.

Ao definirmos um sistema dinâmico e estudarmos algumas propriedades recorrentes entre eles, surge a necessidade de se definirem alguns outros conceitos. Chamaremos de X o conjunto estudado e de f o endomorfismo a ele associado, de modo que nos refiramos ao sistema dinâmico como (X, f) . Se a função f for invertível, denotaremos sua inversa por f^{-1} . Denotaremos a n -ésima composição de f consigo mesma por $f^n := f \circ f \circ \dots \circ f$ e a n -ésima composição de f^{-1} por $f^{-n} := f^{-1} \circ f^{-1} \circ \dots \circ f^{-1}$. Os valores que um elemento x do conjunto X assume quando aplicamos a ele a função f repetidamente — $f(x), f^2(x), \dots$ — é uma característica importante desse elemento e, conseqüentemente, do sistema dinâmico estudado. A tais elementos damos o nome de *iterados* de x — mais especificamente, $f^n(x)$ é o n -ésimo iterado de x — e o conjunto de todos os iterados de x é chamado de *órbita* de x e denotado por $\mathcal{O}_f(x)$. Quando ocorre $f^n(x) = x$ para algum natural n , dizemos que a órbita de x é *periódica* e que n é o *período* da órbita; podemos também nos referir a um ponto cuja órbita é periódica como um ponto periódico e a n como seu período. Denotaremos o conjunto de todos os elementos cujas órbitas são periódicas em X sob ação de f por $Per(f)$. Quando x não tem órbita periódica, mas existem naturais m e n tais que $f^m(x) = f^n(x)$, dizemos que a órbita de x é *eventualmente periódica* e que seu período é a menor diferença positiva entre tais m e n , como sugere a intuição. A seguir, estudaremos um exemplo simples, mas importante, de sistema dinâmico.

4 Rotação no Círculo

O primeiro sistema dinâmico a ser estudado neste trabalho é composto por um conjunto chamado de círculo munido de uma função que rotaciona seus pontos por um ângulo fixo. Esse sistema é simples e funciona como um exemplo que consolida a ideia de sistemas dinâmicos e seu funcionamento, evidenciando como podemos trabalhar com suas propriedades e definições iniciais para concluir outras propriedades do sistema. Para isso, devemos definir o conjunto círculo e a função rotação.

Definição 4.1. O conjunto *círculo* é definido como o intervalo $[0, 1]$ com o 0 e o 1 identificados como sendo o mesmo elemento. Ele é denotado por S^1 e chamamos seus elementos de pontos.

A definição sugere que podemos entender o intervalo $[0, 1]$ como um círculo se identificarmos seus extremos. Percebemos que o círculo como objeto geométrico que conhecemos usualmente compartilha com esse intervalo não a característica de equidistância de todos pontos a um centro, mas sim uma característica topológica de objeto unidimensional em que todos os pontos do conjuntos têm pontos vizinhos. Essa definição nos permite trabalhar com o conjunto círculo algebricamente em vez de geometricamente, mas exige que usemos a noção de aritmética modular, como a definição a seguir de rotação evidencia.

Definição 4.2. A *função rotação* é definida a partir de uma constante real α como

$$R_\alpha : S^1 \rightarrow S^1 \\ x \mapsto x + \alpha \pmod{1}.$$

Usamos a aritmética modular para estabelecer uma definição de rotação que dispensa a noção geométrica de ângulo. Para isso, somamos a constante real α a um ponto do círculo e tomamos a parte fracionária, pois o resultado deve estar no intervalo $[0, 1]$. Essa dinâmica é chamada de rotação no círculo, pois representa uma rotação se imaginarmos que cada elemento se desloca α ao longo do perímetro do círculo. Ela equivale a uma rotação de $2\pi\alpha$ radianos. Classificamos as rotações no círculo de acordo com a natureza da constante α . Essa classificação será importante nos teoremas seguintes e, portanto, evidenciamo-la abaixo.

Definição 4.3. Se α for um número racional, diremos que R_α é uma *rotação racional* e, caso contrário, que é uma *rotação irracional*.

A classificação das rotações como racional ou irracional nos permite classificar, também, as órbitas dos pontos do círculo sob ação de tais rotações. Os teoremas a seguir relacionam a periodicidade das rotações com a racionalidade dos ângulos que as determinam.

Teorema 4.1. *Todo ponto do círculo tem órbita periódica sob rotação racional.*

Demonstração. Sejam x_0 um ponto do círculo e α um número racional associado à rotação R_α definida anteriormente. Para mostrar que a órbita de x_0 é periódica, devemos achar um natural n tal que $R_\alpha^n(x_0) = x_0$. A constante α é um número racional e, por isso, pode ser escrita em forma de fração irredutível como $\alpha = \frac{j}{k}$, sendo j e k inteiros e k não nulo. Tomando $n = k$, temos $R_\alpha^k(x_0) = x_0 + k\alpha \pmod{1}$ e, finalmente, como $k\alpha$ é um inteiro, $R_\alpha^k(x_0) = x_0$. Como x_0 é um ponto genérico do círculo, todo ponto do círculo tem órbita periódica sob rotação racional. ■

Teorema 4.2. *Nenhum ponto do círculo tem órbita periódica sob rotação irracional.*

Demonstração. Sejam x_0 um ponto do círculo e α um número irracional associado à rotação R_α definida anteriormente. Podemos provar a afirmação acima por uma redução ao absurdo, notando que, se houvesse algum inteiro n tal que $R_\alpha^n(x_0) = x_0$, como $R_\alpha^n(x_0) = x_0 + n\alpha \pmod{1}$, teríamos $n\alpha$ inteiro. Mas como n é inteiro e α irracional, isso é impossível. Por esse absurdo, nenhum ponto do círculo pode ter órbita periódica sob rotação irracional. ■

Os teoremas acima parecem relativamente simples, mas nos garantem maior entendimento da natureza das rotações. A partir deles, concluímos que um ponto do círculo, quando sob rotação, tem órbita periódica se, e somente se, essa rotação é racional. As rotações irracionais não permitem periodicidade de nenhum ponto do círculo. A consequência desse resultado é importante e constitui a base para o último teorema dessa seção (4.4), que é menos imediato que os anteriores. Antes de enunciá-lo, porém, é importante definir a noção de um conjunto denso — o que, por sua vez, exige que definamos a noção de uma métrica.

Definição 4.4. Uma *métrica* é definida em um conjunto não-vazio X como uma função $d : X \times X \rightarrow \mathbb{R}$ que satisfaz, para todo $x, y, z \in X$, as seguintes propriedades:

$$\text{Não-negatividade : } d(x, y) \geq 0;$$

$$\text{Coincidência : } d(x, y) = 0 \Leftrightarrow x = y;$$

$$\text{Simetria : } d(x, y) = d(y, x);$$

$$\text{Desigualdade triangular : } d(x, z) \leq d(x, y) + d(y, z).$$

A partir dessa definição, vamos estabelecer uma distância no círculo que será usada ao longo de todo o estudo. A noção modular de círculo nos sugere que definamos uma distância que indique, por exemplo, a proximidade dos pontos 0,1 e 0,9 como sendo maior que a dos pontos 0,1 e 0,8; no entanto, a ideia de distância num intervalo também deve ser levada em consideração. Para isso, devemos medir a distância entre os pontos como pontos do intervalo $[0, 1]$ e compará-la com a soma da distância entre 0 e o ponto mais próximo de 0 com a distância entre 1 e o ponto mais próximo de 1.

Definição 4.5. A função distância no círculo é definida por

$$d_{S^1} : S^1 \times S^1 \rightarrow \mathbb{R}$$

$$(x, y) \mapsto \min\{|x - y|, 1 - |x - y|\}.$$

Para provar que a definição acima de fato procede, vamos analisar cada uma das quatro propriedades de distância.

Proposição 4.3. A função $d_{S^1} : S^1 \times S^1 \rightarrow \mathbb{R}$ é uma métrica.

Demonstração. Sejam x, y e z pontos de S^1 .

Não-negatividade. Claramente, como x e y estão no intervalo $[0, 1]$, vale $0 \leq |x - y| < 1$ e $0 < 1 - |x - y| \leq 1$. Desse modo, $d_{S^1}(x, y) = \min\{|x - y|, 1 - |x - y|\} \geq 0$.

Coincidência. Notamos que, se $x = y$, $d_{S^1}(x, y) = \min\{0, 1\} = 0$. Por outro lado, se $d_{S^1}(x, y) = \min\{|x - y|, 1 - |x - y|\} = 0$, devemos ter $|x - y| = 0$ ou $|x - y| = 1$. No primeiro caso temos $x = y$; o segundo é impossível, pois 0 e 1 são identificados no círculo. Porém, mesmo se fôssemos considerar usar os valores 0 e 1 separados, para termos $|x - y| = 1$, deveríamos ter $x = 1$ e $y = 0$ ou $x = 0$ e $y = 1$. Em ambos os casos, $x = y$, pois $0 = 1$.

Simetria. Como $|x - y| = |y - x|$, temos $\min\{|x - y|, 1 - |x - y|\} = \min\{|y - x|, 1 - |y - x|\}$; portanto $d_{S^1}(x, y) = d_{S^1}(y, x)$.

Desigualdade triangular. Por fim, devemos provar que vale a desigualdade triangular. Essa é a propriedade mais trabalhosa entre as quatro. Vamos separar a demonstração em vários casos e subcasos, conforme necessário.

No primeiro caso, $d_{S^1}(x, y) = |x - y|$ e $d_{S^1}(y, z) = |y - z|$. Assim, temos que $d_{S^1}(x, y) + d_{S^1}(y, z) = |x - y| + |y - z|$. Mas $|x - y| + |y - z| \geq |(x - y) + (y - z)| = |x - z|$. Se $d_{S^1}(x, z) = |x - z|$, a propriedade segue direto da desigualdade anterior. Caso contrário, $d_{S^1}(x, z) = 1 - |x - z|$, e vale $|x - z| \geq 1 - |x - z|$, e também segue desta e daquela desigualdades a propriedade.

O segundo caso é $d_{S^1}(x, y) = |x - y|$ e $d_{S^1}(y, z) = 1 - |y - z|$. Disso temos que $d_{S^1}(x, y) + d_{S^1}(y, z) = |x - y| + 1 - |y - z|$. Agora notemos que $|x - y| - |z - y| \geq |(x - y) - (z - y)| = |x - z|$. Como $|y - z| = |z - y|$, vale $|x - y| + 1 - |y - z| \geq |x - z|$. Como no caso anterior, seja $d_{S^1}(x, z) = |x - z|$ ou $d_{S^1}(x, z) = 1 - |x - z|$, temos a desigualdade triangular.

No terceiro caso, $d_{S^1}(x, y) = 1 - |x - y|$ e $d_{S^1}(y, z) = |y - z|$. Semelhantemente ao caso passado, temos $d_{S^1}(x, y) + d_{S^1}(y, z) = 1 - |x - y| + |y - z|$ e $1 - |x - y| + |y - z| \geq 1 - |(z - y) - (x - y)| = 1 - |x - z|$. Assim, se tivermos $d_{S^1}(x, z) = 1 - |x - z|$, segue a propriedade; quando $d_{S^1}(x, z) = |x - z|$, vale $1 - |x - z| \geq |x - z|$, e temos, novamente, a desigualdade triangular.

No último caso, $d_{S^1}(x, y) = 1 - |x - y|$ e $d_{S^1}(y, z) = 1 - |y - z|$. Daí segue que $d_{S^1}(x, y) + d_{S^1}(y, z) = 1 - |x - y| + 1 - |y - z| = 2 - |x - y| - |y - z|$. Notemos, agora, que para quaisquer pontos x e $y \in S^1$, temos que $d_{S^1}(x, y) \leq \frac{1}{2}$. Isso ocorre pois, se $d_{S^1}(x, y) = |x - y|$, devemos ter $|x - y| \leq 1 - |x - y|$ e, rearrumando a desigualdade, vem que $|x - y| \leq \frac{1}{2}$; se $d_{S^1}(x, y) = 1 - |x - y|$, devemos ter $1 - |x - y| \leq |x - y|$ e, rearrumando a desigualdade, $|x - y| \geq \frac{1}{2}$ e segue que $1 - |x - y| \leq \frac{1}{2}$. Assim, notamos que $|x - y| + |y - z| + |x - z| \leq \frac{1}{2} + \frac{1}{2} + \frac{1}{2} \leq 2$, e temos $2 - |x - y| - |y - z| \geq |x - z|$. Como anteriormente, caso $d_{S^1}(x, z) = |x - z|$, a propriedade está provada e, caso contrário, $|x - z| \geq 1 - |x - z|$, e também vale a desigualdade triangular.

Como se pode notar, em todos os casos, teremos

$$d_{S^1}(x, y) + d_{S^1}(y, z) \geq d_{S^1}(x, z),$$

concluindo que vale a desigualdade triangular para quaisquer três pontos de S^1 e que, portanto, a função d_{S^1} é uma métrica. ■

Podemos analisar as órbitas de cada ponto do círculo sob rotação irracional e notar que elas são, de fato, densas em S^1 . Para isso, definimos abaixo o conceito de densidade.

Definição 4.6. Sejam X e Y conjuntos que satisfazem $X \subseteq Y$ e munidos da métrica $d : Y \times Y \rightarrow \mathbb{R}$. Dizemos que o conjunto X é denso no conjunto Y quando, para todo $\varepsilon > 0$ e $y \in Y$, existe $x \in X$ tal que $d(y, x) < \varepsilon$.

Teorema 4.4. *A órbita de um ponto do círculo é densa sob rotação irracional.*

Demonstração. Seja a um número irracional e x_0 um ponto do círculo. Devemos mostrar que, dados $\varepsilon > 0$ e $y \in S^1$, existe um inteiro n_0 tal que $d_{S^1}(y, R_a^{n_0}(x_0)) < \varepsilon$. Para isso, acharemos dois iterado de x_0 a distância menor que ε entre si e, a partir destes, acharemos o iterado procurado.

Podemos pensar no círculo como o intervalo $[0, 1]$ e dividi-lo em intervalos menores iguais entre si; para um inteiro m e para todo $k \in \{0, \dots, m-1\}$, consideremos os intervalos $I_k := [\frac{k}{m}, \frac{k+1}{m})$. O comprimento de cada um dos k intervalos é $\frac{1}{m}$. Escolhendo m suficientemente grande, podemos ter $\frac{1}{m} < \varepsilon$. Como pelo teorema 4.2 nenhum ponto do círculo tem órbita periódica sob rotação irracional, deverão existir, pelo princípio das gavetas, dois iterados distintos de x_0 que estejam no mesmo intervalo. Sejam esses iterados $R_a^i(x_0)$ e $R_a^j(x_0)$ e suponhamos que $i > j$. Consideremos, então, a função $R_a^{(i-j)}$, e notemos que $(i-j)$ é um natural estritamente positivo. Sabemos que $d_{S^1}(R_a^i(x_0), R_a^j(x_0)) < \varepsilon$ e, portanto, que iterados seguidos de x_0 sob $R_a^{(i-j)}$ têm distância menor que ε entre si. Isso significa que deve haver um iterado de x_0 sob $R_a^{(i-j)}$ que estará no intervalo em que y está. Seja esse iterado o n -ésimo iterado de $R_a^{(i-j)}$. Tomamos $n_0 = (i-j)n$ e temos $d_{S^1}(y, R_a^{n_0}(x_0)) < \varepsilon$. ■

5 Cojugação de Sistemas Dinâmicos

Um dos objetivos principais deste estudo é procurar classificar diferentes sistemas dinâmicos usando conjugações. Conjugações entre sistemas dinâmicos são uma noção de equivalência, um modo de associar um sistema dinâmico a outro para poder estudar o primeiro através do segundo. Faz-se isso estabelecendo uma função que

leva elementos do conjunto de um sistema no do outro. Naturalmente, deve-se esperar que essa função satisfaça algumas propriedades para que se preserve a estrutura do sistema inicial; por exemplo, seria proveitoso que pontos fixos e períodos fossem preservados. Uma das exigências para essa função é que ela seja contínua. Portanto dou as definições formais a seguir.

Definição 5.1. Sejam X e Y conjuntos munidos das métricas $d_X : X \times X \rightarrow \mathbb{R}$ e $d_Y : Y \times Y \rightarrow \mathbb{R}$, respectivamente, e seja $f : X \rightarrow Y$ uma função. Dizemos que a função f é *contínua em um ponto* x de seu domínio se, dado $\varepsilon > 0$, existe $\delta > 0$ tal que, para todo $y \in X$, vale

$$d_X(x, y) < \delta \Rightarrow d_Y(f(x), f(y)) < \varepsilon.$$

Se uma função é contínua em todos pontos de seu domínio, dizemos que ela é uma *função contínua*.

Definição 5.2. Sejam (X, f) e (Y, g) sistemas dinâmicos e X e Y espaços métricos. Dizemos que uma função $\pi : X \rightarrow Y$ é uma *semiconjugação* do sistema (X, f) para o sistema (Y, g) se é uma função sobrejetiva, contínua e satisfaz a relação comutativa $\pi \circ f = g \circ \pi$ ilustrada pelo diagrama abaixo. Caso π seja também invertível e sua inversa seja contínua, dizemos que é uma *conjugação*.

$$\begin{array}{ccc} X & \xrightarrow{f} & X \\ \pi \downarrow & & \downarrow \pi \\ Y & \xrightarrow{g} & Y \end{array}$$

A partir das propriedades da conjugação, vamos deduzir alguns resultados imediatos. Podemos notar que as funções compostas f^n e g^n ainda satisfazem a propriedade comutativa que f e g satisfazem. Segue a proposição.

Proposição 5.1. *Sejam (X, f) e (Y, g) sistemas dinâmicos e $\pi : X \rightarrow Y$ uma função satisfazendo $\pi \circ f = g \circ \pi$. Então, para todo natural $n \geq 1$, vale a relação $\pi \circ f^n = g^n \circ \pi$.*

Demonstração. Vamos provar a proposição acima usando indução matemática. Primeiro, notamos que a propriedade claramente vale para $n = 1$. Agora, supondo que

valha $\pi \circ f^{n-1} = g^{n-1} \circ \pi$, temos que

$$\begin{aligned}\pi \circ f^n &= \pi \circ (f^{n-1} \circ f) \\ &= (\pi \circ f^{n-1}) \circ f \\ &= (g^{n-1} \circ \pi) \circ f \\ &= g^{n-1} \circ (\pi \circ f) \\ &= g^{n-1} \circ (g \circ \pi) \\ &= (g^{n-1} \circ g) \circ \pi \\ &= g^n \circ \pi,\end{aligned}$$

o que conclui a indução. ■

Ainda, podemos ver que a conjugação preserva órbitas periódicas, no sentido em que, se algum elemento do conjunto inicial tem órbita periódica, então sua imagem pela conjugação terá, também, órbita periódica. Segue a proposição.

Proposição 5.2. *Sejam (X, f) e (Y, g) sistemas dinâmicos e $\pi : X \rightarrow Y$ uma conjugação entre eles. Considere $x \in X$. Se x tem órbita de período n em X , então $\pi(x)$ também tem órbita de período n em Y .*

Demonstração. Para notar isso, seja $x \in X$ tal que $x = f^n(x)$. Aplicando a conjugação nessa igualdade, temos

$$\begin{aligned}\pi(x) &= \pi(f^n(x)) \\ &= (\pi \circ f^n)(x) \\ &= (g^n \circ \pi)(x) \\ &= g^n(\pi(x)).\end{aligned}$$
■

A última propriedade que será provada nessa seção é a de que a conjugação preserva a densidade das órbitas; ou seja, se a órbita de um ponto é densa no domínio da conjugação, a imagem desse ponto também terá órbita densa no contradomínio da conjugação. Essa propriedade é uma consequência da continuidade de da conjugação, e a demonstração dessa afirmação vem a seguir.

Proposição 5.3. *Sejam (X, f) e (Y, g) sistemas dinâmicos, X e Y munidos das métricas $d_X : X \times X \rightarrow \mathbb{R}$ e $d_Y : Y \times Y \rightarrow \mathbb{R}$, respectivamente, e seja $\pi : X \rightarrow Y$ uma conjugação entre eles. Então, se um ponto $x_0 \in X$ tem órbita densa em X , o ponto $\pi(x_0)$ também tem órbita densa em Y .*

Demonstração. Para mostrar que $\pi(x_0)$ tem órbita densa em Y , sejam $\varepsilon > 0$ e $y \in Y$. Como π é sobrejetiva, existe $x \in X$ tal que $\pi(x) = y$. Considerando o fato de que x_0 tem órbita densa em X , sabemos que, para tal x e qualquer $\varepsilon' > 0$, existe $n \in \mathbb{N}$ tal que $d_X(x, f^n(x_0)) \leq \varepsilon'$. Agora, como π é contínua, para o ε acima, existe $\delta > 0$ tal que, para $f^n(x_0)$, vale

$$d_X(x, f^n(x_0)) < \delta \Rightarrow d_Y(\pi(x), \pi(f^n(x_0))) < \varepsilon.$$

Por fim, pela proposição 5.1, sabemos que vale $\pi \circ f^n = g^n \circ \pi$, e, ainda, temos $\pi(x) = y$. Tomando $\delta = \varepsilon'$, concluímos que, dados $\varepsilon > 0$ e $y \in Y$, existe $n \in \mathbb{N}$ tal que $d_Y(y, g^n(\pi(x_0))) < \varepsilon$. ■

6 Shift de Bernoulli

Nesta seção, estudaremos o sistema mais importante deste projeto, o sistema do Shift de Bernoulli. Esse sistema é composto por um conjunto de sequências numéricas cujas entradas são números naturais com uma cota superior especificada. Associada a esse conjunto está uma função chamada shift de Bernoulli, que desloca cada entrada de uma sequência para a posição anterior a ela. A importância do sistema é devida à facilidade de lidar com a função shift e por haver tantos outros sistemas que podem ser conjugados com ele. Devemos, portanto, definir mais formalmente os conceitos acima explicados.

Definição 6.1. O conjunto Σ_m , definido a partir de um natural $m > 1$, é o conjunto de todas sequências da forma $(x_n)_{n \in \mathbb{N}} := (x_0, x_1, x_2, \dots)$, em que as entradas x_0, x_1, x_2, \dots pertencem a $\{0, \dots, m-1\}$. Denotamos isso por

$$\Sigma_m := \{0, \dots, m-1\}^{\mathbb{N}}.$$

Costumamos dizer que as sequências infinitas descritas acima são unilaterais, pois suas entradas têm como conjunto de índices os naturais \mathbb{N} . Se tratamos de sequências da forma $(\dots, x_{-1}; x_0, x_1, \dots)$, dizemos que elas são bilaterais e o conjunto de todas tais sequências é denotado por $\{0, \dots, m-1\}^{\mathbb{Z}}$.

A função shift pode ser definida tanto para sequências unilaterais como para sequências bilaterais. Geralmente usamos as sequências unilaterais para semiconjugações e as bilaterais para conjugações. O motivo disso é que quando o endomorfismo de um sistema não é bijetivo, 'informação' se perde, pois não há inversa, e o endomorfismo shift reflete essa perda. A definição a seguir evidenciará esse fenômeno.

Definição 6.2. A função *shift* é definida como

$$\begin{aligned} \sigma : \Sigma_m &\rightarrow \Sigma_m \\ (x_0, x_1, \dots) &\mapsto (x_1, x_2, \dots) \end{aligned}$$

Para sequências unilaterais, a entrada x_0 se perde ao aplicarmos o shift. Também não existe inversa do shift. Quando se trata de uma sequência bilateral, o shift leva a sequência $(\dots, x_{-1}; x_0, x_1, \dots)$ na sequência $(\dots, x_0; x_1, x_2, \dots)$ e informação não é perdida.

Podemos prosseguir agora para estudar algumas propriedades do shift de Bernoulli. Uma propriedade inicial importante do sistema (Σ_m, σ) fala sobre sequências de órbita periódica. Ela afirma que um elemento do conjunto das sequências tem órbita periódica se e somente se ele é uma repetição de sequência de dígitos. Definimos e provamos formalmente tal afirmação no próximo teorema.

Teorema 6.1. *Seja $(x_n)_{n \in \mathbb{N}} := \{x_0, x_1, \dots\}$ uma sequência do conjunto Σ_m . Então a órbita de (x_n) é periódica de período $k > 0$ sob σ se, e somente se, para todo natural i , vale $x_i = x_{i+k}$.*

Demonstração. Se vale $x_i = x_{i+k}$ para todo natural i , a órbita da sequência $(x_n)_{n \in \mathbb{N}}$ claramente será periódica, pois teremos $\sigma^k((x_n)_{n \in \mathbb{N}}) = (x_n)_{n \in \mathbb{N}}$. Por outro lado, se $(x_n)_{n \in \mathbb{N}}$ tiver órbita periódica, valerá $\sigma^k((x_n)_{n \in \mathbb{N}}) = (x_n)_{n \in \mathbb{N}}$ para algum k . Nesse caso, teremos que $x_i = x_{i+k}$ para todo natural i . ■

Esse teorema fala sobre a natureza dos elementos periódicos de (Σ_m, σ) , que são simplesmente sequências que se repetem infinitamente. Analogamente, provamos que um elemento tem órbita eventualmente periódica se ocorre $x_i = x_{i+k}$ para algum natural $i > 0$. A noção de densidade será necessária no próximo teorema e, portanto, devemos definir uma métrica no conjunto Σ_m . Essa definição virá seguida da demonstração de que ela é, de fato, uma métrica.

Definição 6.3. Para sequências $x = (x_n)_{n \in \mathbb{N}}$ e $y = (y_n)_{n \in \mathbb{N}}$ de Σ_m , definimos a função distância $d_{\Sigma_m} : \Sigma_m \times \Sigma_m \rightarrow \mathbb{R}$ como

$$d_{\Sigma_m}(x, y) = \begin{cases} 0 & \text{se } x = y \\ \left(\frac{1}{2}\right)^{n(x,y)} & \text{se } x \neq y, \end{cases}$$

em que $n(x, y) := \max\{i \in \mathbb{N} : x_j = y_j, \forall j < i\}$ mede quantas estradas iniciais de duas sequências x e y são iguais.

Corolário 6.3.1. Como consequência dessa definição, podemos notar que, quando $x \neq y$, vale $n(x, y) = k_0 - 1$, sendo k_0 o menor natural tal que $x_{k_0} \neq y_{k_0}$.

Demonstração. Como $x \neq y$, existe algum natural k tal que $x_k \neq y_k$. O conjunto de todos esses k é subconjunto dos naturais; portanto tem um menor elemento k_0 . Por definição, $n(x, y) = \max\{i \in \mathbb{N} : x_j = y_j, \forall j < i\}$. Como todo natural i menor que k_0 cumpre $x_j = y_j, \forall j < i$, o maior deles é $k_0 - 1$ e temos $n(x, y) = k_0 - 1$. ■

Para esclarecer a definição acima, calculamos alguns exemplos, considerando $m = 10$. Para as sequências

$$x = (3, 1, 4, 1, 5, 9, 2, 6, 5, 3, 5, \dots) \text{ e } y = (3, 1, 4, 1, 5, 9, 2, 6, 5, 3, 3, \dots),$$

temos $n(x, y) = 10$; para $x = (3, 1, 5, 1, 5, 9, \dots)$ e $y = (3, 1, 4, 1, 5, 9, \dots)$, temos $n(x, y) = 2$; para $x = (1, 1, 4, \dots)$ e $y = (3, 1, 4, \dots)$, temos $n(x, y) = 0$. Prossequimos para a próxima porposição.

Proposição 6.2. *A função $d_{\Sigma_m} : \Sigma_m \times \Sigma_m \rightarrow \mathbb{R}$ é uma métrica.*

Demonstração. Sejam $x = (x_n)_{n \in \mathbb{N}}$, $y = (y_n)_{n \in \mathbb{N}}$ e $z = (z_n)_{n \in \mathbb{N}}$ sequências de Σ_m .

Não-negatividade. Consideramos dois casos. Se $x = y$, temos $d_{\Sigma_m}(x, y) = 0$. Se $x \neq y$, seja $n(x, y) = k$. Então $d_{\Sigma_m}(x, y) = (\frac{1}{2})^k$ e, como $(\frac{1}{2})^{k_0} \geq 0$ para todo natural k_0 , concluímos que $d_{\Sigma_m}(x, y) \geq 0$.

Coincidência. Se $x = y$, sabemos que $d_{\Sigma_m}(x, y) = 0$ por definição. Por outro lado, caso $d_{\Sigma_m}(x, y) = 0$, como não existe natural k tal que $(\frac{1}{2})^{k_0} = 0$, devemos ter $x = y$.

Simetria. Novamente, consideramos dois casos. Se $x = y$, claramente temos $d_{\Sigma_m}(x, y) = d_{\Sigma_m}(y, x)$. Caso $x \neq y$, notemos que $n(x, y)$ é simétrica por definição, pois $\max\{i \in \mathbb{N} : x_j = y_j, \forall j < i\} = \max\{i \in \mathbb{N} : y_j = x_j, \forall j < i\}$. Disso segue $d_{\Sigma_m}(x, y) = d_{\Sigma_m}(y, x)$.

Desigualdade triangular. Para provar a desigualdade triangular, devemos antes notar uma propriedade da função expoente n envolvendo três sequências quaisquer de Σ_m . Quando temos três sequências $x, y, z \in \Sigma_m$ e calculamos os valores da função expoente n entre cada uma delas, dois dos três valores serão sempre iguais entre si e o terceiro será sempre maior ou igual aos outros dois. Em outras palavras, uma das três sequências terá sempre o mesmo número de casas iniciais coincidentes com as outras duas, e as outras duas coincidirão em mais casas iniciais entre si. Assim, entre os valores $n(x, y)$, $n(y, z)$ e $n(x, z)$, dois serão iguais e o terceiro maior ou igual aos outros dois. Suponhamos que $n(x, y) = a$, $n(y, z) = b$ e, sem perda de generalidade, que $a \leq b$. O que devemos perceber é que $n(x, z)$ não poderá ter um valor arbitrário, independentemente de quais sejam as sequências x, y e z . Esse valor está determinado pelos valores a e b . De fato, se $a = b$, $n(x, z) \geq a$ e, se $a < b$, $n(x, z) = a$. Isso ocorre pois, como $n(x, y) = a$, $x_j = y_j$ para todo $j < a$ e, da mesma forma, como $n(y, z) = b$, $y_j = z_j$ para todo $j < b$, o que implica que $x_j = z_j$ para todo $j < a$, pois $a \leq b$. Assim concluímos que $n(x, z) \geq a$. Desse modo, o caso em que $a = b$ já está provado e basta restringir o resultado para o caso em que $a < b$. Nesse caso, suponhamos, por absurdo, que $n(x, z) > a$. Sendo assim, teríamos $x_{a+1} = z_{a+1}$ e, como $a < b = n(y, z)$, também valeria $y_{a+1} = z_{a+1}$. Mas isso implica $x_{a+1} = y_{a+1}$, o que contradiz $n(x, y) = a$. Portanto fica provado que, para três sequências quaisquer de Σ_m , dois dos valores da

função expoente devem ser iguais entre si e o terceiro deve ser maior ou igual aos outros dois. Esse resultado é essencial para a demonstração a seguir.

Para provar a desigualdade triangular, devemos provar que vale

$$d_{\Sigma_m}(x, y) + d_{\Sigma_m}(y, z) \geq d_{\Sigma_m}(x, z).$$

Caso $x = y$, $x = z$ ou $y = z$, a igualdade é imediata, pois ao menos um dos três valores será nulo. Caso contrário, se $x \neq y$, $x \neq z$ e $y \neq z$, devemos mostrar que

$$\left(\frac{1}{2}\right)^{n(x,y)} + \left(\frac{1}{2}\right)^{n(y,z)} \geq \left(\frac{1}{2}\right)^{n(x,z)}.$$

Como mostrado no parágrafo anterior, no mínimo dois desses três expoentes são iguais. Dividimos, portanto, a demonstração em dois casos, considerando a o menor valor entre os expoentes $n(x, y)$, $n(y, z)$ e $n(x, z)$ e b o outro valor, e sabendo que vale $a \leq b$. No primeiro caso, devemos mostrar que

$$\left(\frac{1}{2}\right)^a + \left(\frac{1}{2}\right)^a \geq \left(\frac{1}{2}\right)^b.$$

De fato, $\left(\frac{1}{2}\right)^a + \left(\frac{1}{2}\right)^a = \left(\frac{1}{2}\right)^{a-1}$; como $a - 1 < a \leq b$, temos $\left(\frac{1}{2}\right)^{a-1} \geq \left(\frac{1}{2}\right)^b$, e segue a desigualdade. No segundo caso, devemos mostrar que

$$\left(\frac{1}{2}\right)^a + \left(\frac{1}{2}\right)^b \geq \left(\frac{1}{2}\right)^a.$$

Como, para todo $b \in \mathbb{N}$, $\left(\frac{1}{2}\right)^b \geq 0$, segue também a desigualdade e, assim, fica demonstrada a desigualdade triangular por completo e, conseqüentemente, que a função d_{Σ_m} é uma métrica. ■

A métrica acima definida é uma métrica que leva em conta para definir proximidade de duas seqüências de Σ_m somente a quantidade de entradas iniciais que elas compartilham, seguindo a ideia de que quanto mais casas iniciais coincidem, menor é a distância entre as seqüências. Essa métrica não leva em consideração as entradas posteriores às coincidentes nas duas seqüências. De fato, definimos a métrica em dois casos, $x = y$ e $x \neq y$, mas podemos pensar que o primeiro caso está contido no segundo se consideramos que duas seqüências iguais coincidem em todas as suas infinitas casas e, no limite, quando o expoente de $\frac{1}{2}$ tende ao infinito, a distância entre as seqüência tende a zero.

As seqüências de Σ_m cujas órbitas são periódicas estão, de certa forma, bem distribuídas pelo conjunto. Levando em conta a noção de proximidade a partir da métrica definida anteriormente, para qualquer seqüência do conjunto Σ_m , existe uma seqüência de órbita periódica tão próxima quanto se queira daquela seqüência. Essa afirmação é formalizada pelo seguinte teorema.

Teorema 6.3. *O conjunto $Per(\sigma)$ das sequências cujas órbitas são periódicas é denso em Σ_m .*

Demonstração. Para provar isso, basta mostrar que, dados $\varepsilon > 0$ e $y \in \Sigma_m$, existe ponto periódico p tal que $d_{\Sigma_m}(y, p) < \varepsilon$. De fato, existe natural n tal que $(\frac{1}{2})^n < \varepsilon$ — qualquer natural maior que o logaritmo de ε na base $(\frac{1}{2})$. Assim, construímos p tal que suas n primeiras entradas sejam iguais às n primeiras entradas de y e teremos o ponto periódico procurado. Explicitamente, se y é dado por $y := (y_n)_{n \in \mathbb{N}}$ e p por $p := (p_n)_{n \in \mathbb{N}}$, fazemos $p_k = y_k$ para todo $k \in \{0, \dots, n-1\}$ e $p_{i+n} = p_i$ para todo natural i e obtemos $\max\{i \in \mathbb{N} : y_j = p_j, \forall j < i\} \geq n$, o que resulta em $d_{\Sigma_m}(y, p) = (\frac{1}{2})^n < \varepsilon$. ■

Definição 6.4. Uma *sequência finita* é uma sequência do tipo (x_1, \dots, x_n) , definida a partir de um natural n . Nesse caso, dizemos que essa sequência tem n entradas e a denotamos por (x_n) .

Essa definição é usada para enunciar e provar o lema a seguir, que será usado na demonstração do próximo teorema.

Lema 6.4. *O conjunto das sequências finitas com entradas em $\{0, \dots, m-1\}$ é enumerável.*

Demonstração. Definimos o conjunto A_i , para um natural i estritamente positivo, como o conjunto de todas as sequências finitas com i entradas. Notemos que todos esses conjuntos são finitos. Há m sequências finitas com uma entrada — a saber,

$$(0), (1), \dots, (m-1);$$

da mesma forma, há m^2 sequências finitas com duas entradas e, de modo geral, m^k sequências finitas com k entradas. Por serem finitos, os conjuntos A_i são todos enumeráveis. Há uma quantidade natural desses conjuntos e, portanto, a união

$$A := \bigcup_{i \in \mathbb{N}} A_i$$

de todos eles é enumerável. Mas o conjunto A é o conjunto de todas as sequências finitas e o lema está provado. ■

Teorema 6.5. *Existe sequência cuja órbita é densa em Σ_m .*

Demonstração. Pelo lema anterior, o conjunto das sequências finitas com entradas em $\{0, \dots, m-1\}$ é enumerável e podemos, portanto, nomeá-las:

$$p_0 = (0), p_1 = (1), \dots, p_m = (0, 0), \dots$$

Agora, vejamos que, para provar o teorema, devemos mostrar que existe $x \in \Sigma_m$ tal que, dado $\varepsilon > 0$ e $y \in \Sigma_m$, há um natural n satisfazendo $d_{\Sigma_m}(y, \sigma^n(x)) < \varepsilon$. Nesse caso, tomamos $x := (p_0, p_1, p_2, \dots)$, a sequência de todas as possíveis sequências finitas com dígitos em $\{0, \dots, m-1\}$ — para ilustrar, x terá a forma

$$(0, 1, \dots, (m-1), 0, 0, 0, 1, \dots, 0, (m-1), 1, 0, \dots, (m-1), (m-1), 0, 0, 0, 0, 0, 1, \dots).$$

Para simplificar, se tivéssemos $m = 2$, teríamos

$$x = (0; 1; 0; 0; 0; 1; 1; 0; 1; 1; 0; 0; 0; 0; 0; 1; 0; 1; 0; 0; 1; 1; 1; 0; 0; \dots),$$

em que os pontos-e-vírgulas foram colocados somente para evidenciar cada bloco de sequência finita. Sendo assim, para todo ε dado, existirá um natural k tal que $(\frac{1}{2})^k < \varepsilon$. As k primeiras entradas de y são, de fato, uma sequência de dígitos em $\{0, \dots, m-1\}$ e, portanto, aparecem em alguma posição na sequência x . Escolhe-se n , por fim, para que $\sigma^n(x)$ coincida nas k primeiras entradas com y e, daí, vem que $d_{\Sigma_m}(y, \sigma^n(x)) = (\frac{1}{2})^k < \varepsilon$. ■

7 Expansão no Círculo

O próximo exemplo a ser analisado é uma outra função no círculo. A dinâmica, nesse caso, é chamada de expansão no círculo.

Definição 7.1. A *expansão* do círculo é definida a partir de uma constante natural m maior que 1 como

$$E_m : S^1 \rightarrow S^1 \\ x \mapsto mx \pmod{1}.$$

Determinar as propriedades principais do sistema dinâmico (S^1, E_m) pode não ser tão fácil; achar os elementos com órbita periódica, por exemplo, não é trivial. Para facilitar esse processo, vamos exibir uma semiconjugação π do sistema da expansão (S^1, E_m) para o shift de Bernoulli (Σ_m, σ) . Esse exemplo será, portanto, a primeira conjugação deste trabalho. Para isso, devemos definir a função π e provar que ela é, de fato, uma conjugação.

Definição 7.2. Primeiramente, definimos, para todo $k \in \{0, \dots, m-1\}$, os intervalos

$$I_k := \left[\frac{k}{m}, \frac{k+1}{m} \right]$$

e, a partir desses, para qualquer sequência (x_n) , finita ou infinita unilateral, com entradas em $\{0, \dots, m-1\}$, o intervalo $I_{(x_n)} = I_{(x_0, x_1, x_2, \dots)}$ é definido como

$$I_{(x_n)} := I_{x_0} \cap E_m^{-1}(I_{x_1}) \cap E_m^{-2}(I_{x_2}) \cap \dots$$

A função inversa de E_m não está bem definida para alguns elementos de S^1 . Todos os extremos dos intervalos I_k têm a mesma imagem, o elemento $0 \sim 1$, tal que este não tem pré-imagem única. Da mesma forma, nenhum número número que seja somatório de racionais cujo denominador é uma potência de m tem pré-imagem única e, por isso, devemos excluir todos esses números de qualquer consideração posterior. Os gráficos abaixo ilustram essa definição.

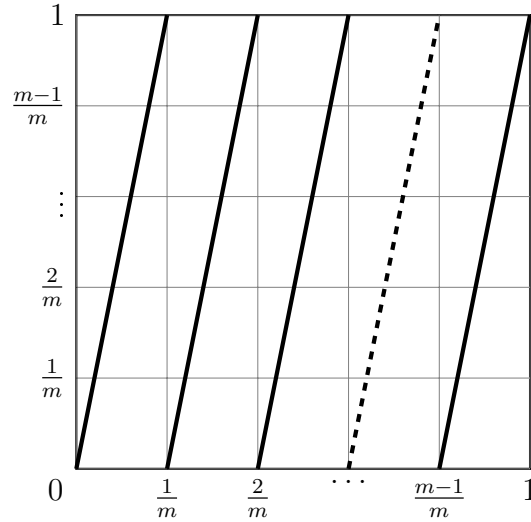


Figura 1: Gráfico da função E_m no quadrado $[0, 1] \times [0, 1]$ e intervalos I_k .

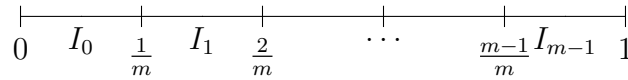


Figura 2: Representação dos intervalos I_k em S^1 .

Vale notar que, dada uma sequência finita (x_n) e o intervalo $I_{(x_n)}$, podemos saber o valor de seus extremos a partir das entradas da sequência (x_n) . Segue a proposição.

Proposição 7.1. *Sejam (x_n) uma sequência finita com entradas em $\{0, \dots, m-1\}$ e $I_{(x_n)}$ o intervalo por ela definido. Então os extremos do intervalo são dados por*

$$I_{(x_n)} := \left[\sum_{i=0}^{n-1} \frac{x_i}{m^{i+1}}, \frac{1}{m^n} + \sum_{i=0}^{n-1} \frac{x_i}{m^{i+1}} \right].$$

Demonstração. Para mostrarmos que o inferior do intervalo é dado pelo somatório do enunciado, notemos que seu primeiro termo é $\frac{x_0}{m}$, que corresponde à distância da

origem ao extremo inferior do intervalo I_{x_0} em que $I_{(x_n)}$ está contido. O segundo termo do somatório é $\frac{x_1}{m^2}$, que corresponde à distância do extremo inferior do intervalo I_{x_0} ao extremo inferior do intervalo $I_{(x_0, x_1)}$, contido em I_{x_0} . Assim por diante, conclui-se que o somatório representa o valor procurado. ■

Definição 7.3. A semiconjugação π do sistema da expansão (S^1, E_m) para o shift de Bernoulli (Σ_m, σ) é definida como

$$\begin{aligned}\pi : S^1 &\rightarrow \Sigma_m \\ x &\mapsto (x_n)_{n \in \mathbb{N}}\end{aligned}$$

tal que $x_n = k$ se $(E_m)^n(x) \in I_k$.

A semiconjugação associa, portanto, a cada elemento de S^1 , uma sequência cuja n -ésima entrada é o índice do intervalo I_k em que o n -ésimo iterado do elemento está.

Notemos que, ao aplicarmos E_m em somatórios de números racionais cujo denominador é uma potência de m , seus iterados eventualmente chegam nos extremos dos intervalos I_k e, então, no ponto $0 \sim 1$. À parte destes, a função está bem definida em todos os valores de $x \in S^1$ pois sempre existe um único intervalo I_k que contém $(E_m)^n(x)$. Assim, definimos a semiconjugação π somente para esses últimos valores.

Uma vez definida a função π , devemos provar que ela de fato é uma semiconjugação de (S^1, E_m) para (Σ_m, σ) . A próxima proposição demonstra que isso é verdade.

Teorema 7.2. *A função π é uma semiconjugação do sistema (S^1, E_m) para o sistema (Σ_m, σ) .*

Demonstração. Primeiro mostraremos que π é uma função sobrejetiva. Para isso, devemos mostrar que para toda sequência de Σ_m existe um respectivo elemento de S^1 que é sua pré-imagem. Seja $(x_n)_{n \in \mathbb{N}}$ uma sequência se Σ_m . Consideremos os intervalos $I_{x_0}, I_{(x_0, x_1)}, I_{(x_0, x_1, x_2)}, \dots$. Nota-se que esses intervalos são intervalos fechados e que satisfazem $I_{x_0} \supset I_{(x_0, x_1)} \supset I_{(x_0, x_1, x_2)} \supset \dots$. Portanto, pelo teorema dos intervalos encaixados, concluímos que a interseção desses intervalos é não vazia: $I_{x_0} \cap I_{(x_0, x_1)} \cap I_{(x_0, x_1, x_2)} \cap \dots \neq \emptyset$. Ainda, podemos concluir que existe um único elemento pertencente a essa interseção, pois o comprimento dos intervalos diminui geometricamente por um fator de $\frac{1}{m}$, tal que o comprimento final da interseção é 0. Seja esse elemento x . Por construção, vale que $(E_m)^n(x) \in I_{x_n}$ e, assim, concluímos que $\pi(x) = (x_n)_{n \in \mathbb{N}}$.

Mostremos, agora, que vale $\pi \circ E_m = \sigma \circ \pi$. Seja x um elemento de S^1 que não é soma de racionais com denominadores que são potência de m e $\pi(x) = (x_0, x_1, x_2, \dots)$. Nesse caso, sabemos que $x \in I_{(x_0, x_1, x_2, \dots)}$. Mas $I_{(x_0, x_1, x_2, \dots)} = I_{x_0} \cap (E_m)^{-1}(I_{x_1}) \cap$

$(E_m)^{-2}(I_{x_2}) \cap \dots$ e, aplicando E_m a $I_{(x_0, x_1, x_2, \dots)}$, temos

$$\begin{aligned} E_m(I_{(x_0, x_1, x_2, \dots)}) &= E_m(I_{x_0}) \cap I_{x_1} \cap (E_m)^{-1}(I_{x_2}) \cap \dots \\ &= S^1 \cap I_{x_1} \cap (E_m)^{-1}(I_{x_2}) \cap \dots \\ &= I_{x_1} \cap (E_m)^{-1}(I_{x_2}) \cap \dots \\ &= I_{(x_1, x_2, \dots)}. \end{aligned}$$

Assim, devemos ter $E_m(x) \in I_{(x_1, x_2, \dots)}$ e, portanto, $\pi(E_m(x)) = (x_1, x_2, \dots)$. Por outro lado, como $\pi(x) = (x_0, x_1, x_2, \dots)$, temos $\sigma(\pi(x)) = (x_1, x_2, \dots)$. Logo $\pi \circ E_m = \sigma \circ \pi$.

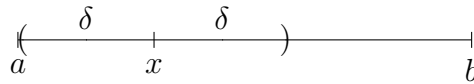
A seguir, mostramos que a função π é uma função contínua. Seja $x \in S^1$ e $\varepsilon > 0$. Devemos achar $\delta > 0$ tal que, para todo $y \in S^1$, valha

$$d_{S^1}(x, y) < \delta \Rightarrow d_{\Sigma_m}(\pi(x), \pi(y)) < \varepsilon.$$

Se $y = x$, $d_{\Sigma_m}(\pi(x), \pi(y)) = 0 < \varepsilon$ e qualquer δ satisfaz a condição procurada. Consideremos, portanto, quando $y \neq x$. Nesse caso, $d_{\Sigma_m}(\pi(x), \pi(y)) = \left(\frac{1}{2}\right)^{n(\pi(x), \pi(y))}$. Agora, seja k o menor natural maior que $\log_{\left(\frac{1}{2}\right)} \varepsilon$. Vamos determinar δ para que as seqüências $\pi(x)$ e $\pi(y)$ tenham ao menos k entradas iguais. Para isso, devemos fazer com que x e y estejam num intervalo em que seqüências com k primeiras entradas estão. Seja $\pi(x) := (x_n)_{n \in \mathbb{N}}$ e $\pi(y) := (y_n)_{n \in \mathbb{N}}$, e consideremos o intervalo $I := I_{(x_0, x_1, \dots, x_{k-1})}$. Pela proposição 7.1, sabemos determinar os extremos de I . Porém podemos defini-los, por conveniência, como $I := [a, b]$. Sabemos que $x \in I$. Tomando $\delta = \min\{x - a, b - x\}$, para todo $y \in S^1$ tal que $d_{S^1}(x, y) < \delta$, garantimos que $y \in I$ e que, conseqüentemente, $\pi(x)$ e $\pi(y)$ terão ao menos k entradas iguais; ou seja, $n(\pi(x), \pi(y)) \geq k$, pois $x_i = y_i$ para todo $i \in \{0, \dots, k-1\}$. Considerando a definição de k , isso implica que

$$d_{\Sigma_m}(\pi(x), \pi(y)) = \left(\frac{1}{2}\right)^{n(\pi(x), \pi(y))} \leq \left(\frac{1}{2}\right)^k < \varepsilon,$$

o que garante a continuidade. O esquema abaixo ilustra o intervalo I , seus extremos, um possível ponto x e a escolha de δ .



■

8 A Ferradura de Smale

Nesta seção, trataremos de um sistema dinâmico conhecido por Ferradura de Smale. O conjunto usado nesta seção será criado a partir de um conjunto $D \in \mathbb{R}^2$. Esse conjunto D é dividido em cinco conjuntos disjuntos; ou seja, $D = D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5$. Os conjuntos D_2, D_3, D_4 formam um quadrado unitário $R := D_2 \cup D_3 \cup D_4$ e os conjuntos D_1, D_5 são semicírculos de diâmetro unitário, como na figura abaixo.

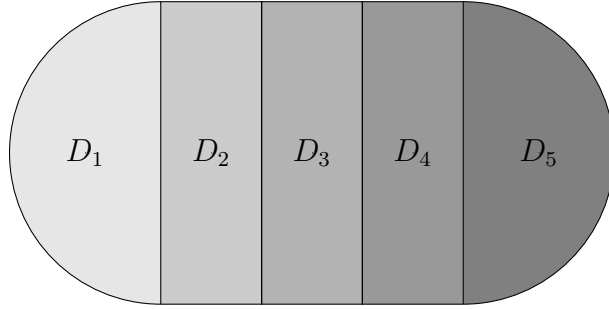


Figura 3: O conjunto $D = D_1 \cup R \cup D_5$.

Associamos ao conjunto D o endomorfismo diferenciável $f : D \rightarrow D$ que estica e entorta o conjunto D nele mesmo, como na figura a seguir.

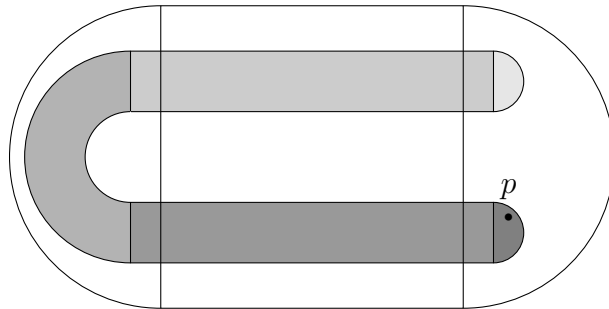


Figura 4: A imagem de D sob o endomorfismo f .

Assuma que f estica $D_2 \cup D_4$ uniformemente na direção horizontal por um fator $\mu > 2$ e contrai uniformemente $D_2 \cup D_4$ na direção vertical por um fator $\lambda < \frac{1}{2}$. Ainda, note que, pelo teorema do ponto fixo de Brower, existe um ponto fixo $p \in f(D_5)$, pois $f(D_5) \in D_5$.

Queremos estudar esse sistema usando uma conjugação com o shift de Bernoulli. Para isso, vamos definir alguns conjuntos de modo análogo a como fizemos para criar a conjugação da expansão no círculo. Vamos definir os seguintes conjuntos:

$$R_0 := f(D_2) \cap R \quad R_1 := f(D_4) \cap R.$$

Note que $f(R) \cap R = R_0 \cup R_1$. A partir desses conjuntos, vamos definir uma sequência de conjuntos da mesma forma. Assim como $f(R) \cap R$ consiste em dois retângulos, que nomeamos R_0 e R_1 , o conjunto $f^2(R) \cap f(R) \cap R = f^2(R) \cap R$ consiste em quatro retângulos que nomearemos R_{00}, R_{01}, R_{10} e R_{11} . Vale notar que a altura dos retângulos R_0 e R_1 é λ e a dos retângulos R_{00}, R_{01}, R_{10} e R_{11} é λ^2 .

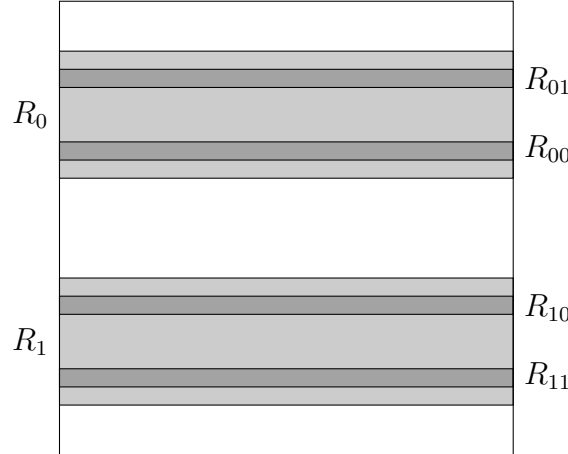


Figura 5: O quadrado R e conjuntos R_ω horizontais usados na conjugação.

De modo geral, para qualquer sequência finita $\omega = (\omega_0, \omega_1, \dots, \omega_n)$ de zeros e uns, definimos

$$R_\omega := R_{\omega_0} \cap f(R_{\omega_1}) \cap f^2(R_{\omega_2}) \cap \dots \cap f^n(R_{\omega_n}),$$

que será um retângulo horizontal de altura λ^n e $f^n(R) \cap R$ será a união dos 2^n possíveis retângulos dessa forma. Para uma sequência infinita $\omega = (\omega_n)_{n \in \mathbb{N}} \in \Sigma_2$, o retângulo R_ω é definido como

$$R_\omega := \bigcap_{i \in \mathbb{N}} f^i(R_{\omega_i}).$$

Construiremos, de modo semelhante, retângulos verticais a partir da função inversa f^{-1} . Note que $f^{-1}(R_0) = f^{-1}(R) \cap D_2$ e $f^{-1}(R_1) = f^{-1}(R) \cap D_4$ são retângulos verticais de largura μ^{-1} . Assim, para uma sequência finita $\omega = (\omega_1, \dots, \omega_n)$ de zeros e uns, definimos

$$R'_\omega := f^{-1}(R_{\omega_1}) \cap f^{-2}(R_{\omega_2}) \cap \dots \cap f^{-n}(R_{\omega_n}),$$

que será um retângulo vertical de largura μ^{-n} e $f^{-n}(R) \cap R$ será a união dos 2^n possíveis retângulos dessa forma. Para uma sequência infinita $\omega = (\omega)_{n \in \mathbb{N} \setminus \{0\}} \in \Sigma_2$, definimos

$$R'_\omega := \bigcap_{i \in \mathbb{N} \setminus \{0\}} f^{-i}(R_{\omega_i}).$$

Após essa discussão, podemos finalmente definir o conjunto ferradura que será estudado nesta seção e uma conjugação do sistema associado a ele. Para isso, é importante fazer alguns comentários. Construiremos uma conjugação da ferradura com um tipo de shift de Bernoulli um pouco diferente do estudado na seção 6. No entanto, as diferenças não são grandes, de modo que será desnecessário dar um tratamento tão minucioso às definições desse sistema de shift como foi feito na seção 6. O conjunto considerado para esse shift de Bernoulli será o conjunto $\Sigma := \{0, 1\}^{\mathbb{Z}}$, o espaço das sequências infinitas bilaterais com entradas 0 ou 1. Diferente de Σ_2 definido nas seções anteriores, esse espaço tem sequências indexadas pelos inteiros, de modo que o shift associado a ela será uma bijeção. A necessidade de usarmos sequências infinitas bilaterais e um shift bijetivo se dá pela natureza da ferradura, que permite uma conjugação completa, não somente uma semiconjugação, como no caso da expansão da seção 7. A definição do mapa shift também se diferencia do shift da seção passada. A função σ é definida como

$$\begin{aligned}\sigma : \Sigma &\rightarrow \Sigma \\ (x_n)_{n \in \mathbb{Z}} &\mapsto (x_{n-1})_{n \in \mathbb{Z}};\end{aligned}$$

ou seja, leva a sequência $(\dots, x_{-1}; x_0, x_1, \dots)$ na sequência $(\dots; x_{-1}, x_0, x_1, \dots)$. Esse shift desloca as sequências bilaterais para a direita, não para a esquerda, como acontecia nas seções anteriores. Nessa seção, todos os usos do σ consideram essa definição. Por fim, vamos definir uma função distância no espaço Σ para poder trabalhar com a noção de continuidade; não demonstraremos que a função é de fato uma métrica. Costumamos nos referir à função distância no espaço Σ como a distância determinada por um cilindro. Na prática, isso quer dizer que, dada duas sequências em Σ , a distância entre elas é o número de casas que elas têm iguais no lado positivo e negativos das sequências. A definição formal é a seguinte função. Para sequências $x = (x_n)_{n \in \mathbb{Z}}$ e $y = (y_n)_{n \in \mathbb{Z}}$, a função $d_\Sigma : \Sigma \times \Sigma \rightarrow \mathbb{R}$ é definida como

$$d_\Sigma(x, y) = \begin{cases} 0 & \text{se } x = y \\ (\frac{1}{2})^{N(x, y)} & \text{se } x \neq y, \end{cases}$$

em que $N(x, y) := \max\{i \in \mathbb{N} : \forall j \in \mathbb{Z} \ |j| < i, x_j = y_j\}$. Sendo assim, prosseguimos para as definições necessárias.

Definição 8.1. A *ferradura* é o conjunto H definido como

$$H := \bigcap_{i \in \mathbb{Z}} f^i(R).$$

Lembrando das definições dos retângulos no começo da seção, podemos perceber que H é a união de todos os conjuntos $R_\omega \cap R'_\omega$, com ω variando em Σ .

Definição 8.2. A conjugação ϕ do sistema da ferradura (H, f) para o shift de Bernoulli (Σ, σ) é definida como

$$\begin{aligned}\phi : H &\rightarrow \Sigma \\ x &\mapsto (x_n)_{n \in \mathbb{Z}}\end{aligned}$$

tal que $x_n = k$ se $f^n(x) \in R_k$.

Perceba que, assim como na semiconjugação da expansão da seção 7, definimos conjuntos propícios para relacionar um ponto do espaço a uma sequência numérica. Diferentemente do caso da semiconjugação da expansão, no entanto, esse caso nos permite definir uma conjugação, ou seja, uma função bijetiva entre H e Σ . A demonstração de que ϕ é de fato uma conjugação será dada na proposição a seguir.

Proposição 8.1. *A função ϕ é uma conjugação do sistema (H, f) para o sistema (Σ, σ) .*

Demonstração. Primeiramente, vamos demonstrar que a função ϕ é bijetiva. Para mostrar a sobrejetividade, consideremos $(x_n)_{n \in \mathbb{Z}} \in \Sigma$. Definamos as sequências positiva e negativa de (x_n) como as sequências infinitas unilaterais

$$(x_n)^+ := (x_0, x_1, \dots) \quad (x_n)^- := (x_{-1}, x_{-2}, \dots).$$

Consideremos os retângulos horizontais $R_{x_0}, R_{(x_0, x_1)}, R_{(x_0, x_1, x_2)}, \dots$. Esses retângulos estão cada um contido no anterior, $R_{x_0} \supset R_{(x_0, x_1)} \supset R_{(x_0, x_1, x_2)} \supset \dots$. Se consideramos suas interseções com alguma reta vertical em R , temos intervalos encaixados $I_{x_0}, I_{(x_0, x_1)}, I_{(x_0, x_1, x_2)}, \dots$, respectivamente, que são intervalos fechados e satisfazem $I_{x_0} \supset I_{(x_0, x_1)} \supset I_{(x_0, x_1, x_2)} \supset \dots$. Portanto, pelo teorema dos intervalos encaixados, concluímos que a interseção desses intervalos é não vazia: $I_{x_0} \cap I_{(x_0, x_1)} \cap I_{(x_0, x_1, x_2)} \cap \dots \neq \emptyset$. Ainda, podemos concluir que existe um único elemento pertencente a essa interseção, pois o comprimento de cada intervalo é a altura de seu respectivo retângulo e, portanto, diminui geometricamente por um fator de $\lambda < 1$, tal que o comprimento final da interseção é 0. Como esse argumento vale para qualquer reta vertical em R , concluímos que $R_{(x_n)^+} = \bigcap_{i \in \mathbb{N}} f^i(R_{x_i})$, que é o conjunto de todos esses pontos, é uma reta horizontal em R . Da mesma forma, consideramos os retângulos verticais $R'_{x_{-1}}, R'_{(x_{-1}, x_{-2})}, R'_{(x_{-1}, x_{-2}, x_{-3})}, \dots$. Suas interseções com uma reta horizontal em R são intervalos encaixados $I'_{x_{-1}}, I'_{(x_{-1}, x_{-2})}, I'_{(x_{-1}, x_{-2}, x_{-3})}, \dots$, respectivamente, que são intervalos fechados e satisfazem $I'_{x_{-1}} \supset I'_{(x_{-1}, x_{-2})} \supset I'_{(x_{-1}, x_{-2}, x_{-3})} \supset \dots$. Pelo teorema dos intervalos encaixados, concluímos que a interseção desses intervalos é não vazia e, como anteriormente, como o comprimento de cada intervalo é a largura de seu respectivo retângulo e, portanto, diminui geometricamente por um fator de $\mu^{-1} < 1$, a interseção desses intervalos terá um único elemento. Novamente, esse argumento vale

para qualquer reta horizontal em R , e concluimos que $R'_{(x_n)-} = \bigcap_{i \in \mathbb{N} \setminus \{0\}} f^{-i}(R_{x_{-i}})$, o conjunto de todos esses pontos, é uma reta vertical em R . Por fim, notamos que as retas horizontal e vertical definidas acima se tocam em um único ponto x . Isso que dizer que $R_{(x_n)+} \cap R'_{(x_n)-} = \{x\}$. Portanto, por construção, vale que $f^n(x) \in R_{x_n}$ para todo inteiro n e, assim, concluimos que $\phi(x) = (x_n)_{n \in \mathbb{Z}}$. Isso prova a sobrejetividade.

Para mostrar a injetividade, consideremos dois elementos $a, b \in H$ e sejam suas imagens $\phi(a) = (a_n)_{n \in \mathbb{Z}}$ e $\phi(b) = (b_n)_{n \in \mathbb{Z}}$. Suponhamos que $(a_n)_{n \in \mathbb{Z}} = (b_n)_{n \in \mathbb{Z}}$. Nesse caso, sabemos que $R_{(a_n)_{n \in \mathbb{Z}}} = R_{(b_n)_{n \in \mathbb{Z}}}$ e que, portanto, a e b pertencem a $R_{(a_n)_{n \in \mathbb{Z}}}$. Mas foi demonstrado acima que, dado $\omega \in \Sigma$, o conjunto R_ω tem somente um elemento, o que implica que $a = b$.

Vamos agora mostrar que vale a relação comutativa $\phi \circ f = \sigma \circ \phi$. Seja $x \in H$ e $\phi(x) = (x_n)_{n \in \mathbb{Z}}$. Nesse caso, sabemos que $x \in R_{(x_n)_{n \in \mathbb{Z}}}$. Aplicando f a $R_{(x_n)}$, temos

$$\begin{aligned} f(R_{(x_n)_{n \in \mathbb{Z}}}) &= f\left(\bigcap_{i \in \mathbb{Z}} f^i(R_{x_i})\right) \\ &= \bigcap_{i \in \mathbb{Z}} f^{i+1}(R_{x_i}) \\ &= \bigcap_{i \in \mathbb{Z}} f^i(R_{x_{i-1}}) \\ &= R_{(x_{n-1})_{n \in \mathbb{Z}}}. \end{aligned}$$

Mas isso implica que $f(x) \in R_{(x_{n-1})_{n \in \mathbb{Z}}}$ e, portanto, pela bijetividade de ϕ , temos $\phi(f(x)) = (x_{n-1})_{n \in \mathbb{Z}}$. Por outro lado, como $\phi(x) = (x_n)_{n \in \mathbb{Z}}$, temos $\sigma(\phi(x)) = (x_{n-1})_{n \in \mathbb{Z}}$. Logo $\phi \circ f = \sigma \circ \phi$.

Por fim, devemos mostrar a continuidade de ϕ . Seja $x \in H$ e $\varepsilon > 0$. Devemos achar $\delta > 0$ tal que, para todo $y \in H$, valha

$$d_{\mathbb{R}^2}(x, y) < \delta \Rightarrow d_\Sigma(\phi(x), \phi(y)) < \varepsilon.$$

Para isso, consideremos um natural N tal que $(\frac{1}{2})^N < \varepsilon$ e seja $\phi(x) = (x_n)_{n \in \mathbb{Z}}$. Claramente, todas sequências de Σ tais que as primeiras N entradas à direita e à esquerda são iguais às de $(x_n)_{n \in \mathbb{Z}}$ estão a uma distância menor que ε de $\phi(x)$. Mas essas sequências são todas sequências tais que sua imagem inversa y satisfaz $y \in R_{(x_0, x_1, \dots, x_n)} \cap R'_{(x_{-1}, \dots, x_{-n})}$. Tomando δ tal que $y \in R_{(x_0, x_1, \dots, x_n)} \cap R'_{(x_{-1}, \dots, x_{-n})}$, temos o resultado. ■

9 Automorfismo Hiperbólico no Toro

Após o estudo de duas dinâmicas envolvendo o círculo, a dinâmica de rotação na seção 4 e a de expansão na seção 7, voltamos nossa atenção nesta seção a um novo

objeto geométrico: o toro. O toro é um objeto semelhante a uma rosquinha ou a um pneu, e pode ser definido usando-se coordenadas cartesianas em um espaço \mathbb{R}^3 . Sua maior importância, porém, está na área da Topologia e uma descrição precisa do toro por coordenadas cartesianas não é relevante nessa área, mas sim a ideia de que o toro é um objeto com um único furo. Neste trabalho, introduziremos a ideia do toro como o conjunto quociente de \mathbb{R}^2 por uma relação de equivalência que identifica todos os pontos do plano cuja diferença é um ponto \mathbb{Z}^2 . Essa definição preserva a mesma ideia da definição do círculo S^1 : a identificação de pontos inteiros. De fato, poderíamos ter definido o círculo na seção 4 usando uma relação de equivalência semelhante na reta. O toro pode ser visto também como o produto cartesiano de dois círculos. Quando fazemos esse produto cartesiano estamos, na realidade, fazendo o produto cartesiano de dois intervalos $[0, 1]$ e identificando, para qualquer $x \in S^1$, todos os elementos da forma $(0, x)$ com elementos da forma $(1, x)$ e todos os elementos da forma $(x, 0)$ com elementos da forma $(x, 1)$. Antes de definirmos o toro, definimos o que é relação de equivalência, classe de equivalência e conjunto quociente.

Definição 9.1. Uma *relação de equivalência* \sim em um conjunto X é uma relação binária que satisfaz, para todo $x, y, z \in X$, as seguintes propriedades

Reflexividade: $x \sim x$;

Simetria: $x \sim y \Leftrightarrow y \sim x$;

Transitividade: $x \sim y$ e $y \sim z \Rightarrow x \sim z$.

Uma *classe de equivalência* de X por \sim é um subconjunto de X , definido a partir de um elemento $a \in X$ como

$$[a] := \{x \in X : x \sim a\}.$$

Ainda, o *conjunto quociente* de X por \sim é o conjunto de todas classes de equivalência de X , definido como

$$X/\sim := \{[x] : x \in X\}.$$

Definição 9.2. O *toro* é definido como

$$\mathbb{T}^2 := \mathbb{R}^2 / \sim$$

em que \sim é a relação de equivalência definida para todo $x, y \in \mathbb{R}^2$ por

$$x \sim y \Leftrightarrow x - y \in \mathbb{Z}^2,$$

que claramente satisfaz as três propriedades da definição de relação de equivalência. Assim, os elementos de \mathbb{T}^2 são as classes de equivalência

$$[x] := \{x + z : z \in \mathbb{Z}^2\}.$$

Por simplicidade, consideraremos frequentemente que pontos do toro são pontos de $S^1 \times S^1$ e esses pontos serão utilizados como representantes das classes de equivalência. O sistema dinâmico a ser estudado nesta seção tem como conjunto o toro, descrito acima, e a dinâmica que age nele é chamada de automorfismo hiperbólico. Essa função é dada por uma matriz de dimensão 2×2 com algumas características específicas. Como a próxima definição depende muito da utilização de matrizes, convém definir algumas coisas sobre elas antes de abordarmos a dinâmica no toro. Denotaremos uma matriz sempre com uma letra maiúscula e seus elementos com a respectiva letra minúscula acompanhados dos índices; escrevemos $A = [a_{ij}]$. A dimensão de uma matriz será explicitada somente quando necessário e será simplificada para matrizes quadradas, como em $A_{2 \times 2} = A_2$. Denotaremos o determinante de uma matriz A por $|A|$ ou $\det(A)$ e sua matriz inversa por A^{-1} . Convenções esclarecidas, definimos o automorfismo formalmente.

Definição 9.3. O *automorfismo hiperbólico* no toro é definido a partir de uma matriz quadrada A , de dimensão 2, como

$$L_A : \mathbb{T}^2 \rightarrow \mathbb{T}^2 \\ [x] \mapsto [Ax].$$

A matriz A deve ser uma matriz com entradas inteiras, determinante unitário e autovalores distintos.

Para justificar a nomenclatura dessa classe de funções como automorfismos hiperbólicos, devemos demonstrar algumas propriedades tanto da matriz A como da função L_A . A matriz é dada por

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

As entradas devem ser inteiras para que a operação $L_A(x)$ esteja bem definida no toro. Isso ocorre pois, para todo $y \in [x]$, temos $y = x + z$, com $z \in \mathbb{Z}^2$, e segue que

$$[Ay] = [A(x + z)] = [Ax + Az] = [Ax],$$

pois $Az \in \mathbb{Z}^2$.

Devemos justificar por que essa função é um automorfismo. Claramente, a função L_A é um endomorfismo por ter como domínio e contradomínio o mesmo conjunto. Para ser um automorfismo, deve ser também um isomorfismo. Isso ocorre porque o determinante da matriz A é igual a 1, o que faz com que sua matriz inversa

$$A^{-1} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

seja também uma matriz com entradas inteiras, determinante unitário e autovalores distintos, e a função inversa $L_{A^{-1}} := [A^{-1}x]$ esteja bem definida no toro. O automorfismo L_A recebe o nome de hiperbólico porque matrizes cujos autovalores têm valor absoluto diferente de 1 recebem esse nome. Sabemos que o determinante de uma matriz de dimensão 2 é o produto de seus autovalores. Como os autovalores da matriz A são distintos e seu determinante é 1, ela deve ter um autovalor maior e um menor que 1 e é, portanto, uma matriz hiperbólica. Vale notar, por fim, que para todo inteiro k , vale

$$(L_A)^k = L_{A^k}.$$

A partir da definição da matriz A , podemos deduzir algumas características imediatas dos autovalores, autovetores e determinante da matriz.

Proposição 9.1. *A matriz A da definição 9.3 tem autovalores*

$$\lambda_+ = \frac{1}{2}(a_{11} + a_{22} + \sqrt{(a_{11} + a_{22})^2 - 4}) \quad e \quad \lambda_- = \frac{1}{2}(a_{11} + a_{22} - \sqrt{(a_{11} + a_{22})^2 - 4})$$

e as entradas satisfazem $a_{11} + a_{22} \neq \pm 2$, $a_{12} \neq 0$ e $a_{21} \neq 0$.

Demonstração. Primeiro, notemos que o polinômio característico de A será

$$\chi_A(\lambda) = \lambda^2 - (a_{11} + a_{22})\lambda + 1.$$

Assim, isolando λ pela fórmula quadrática, obtemos λ_+ e λ_- como no enunciado. Para termos $\lambda_+ \neq \lambda_-$, devemos ter $\sqrt{(a_{11} + a_{22})^2 - 4} \neq 0$, o que implica $a_{11} + a_{22} \neq \pm 2$. Por fim, se $a_{12} = 0$ ou $a_{21} = 0$, temos $a_{12}a_{21} = 0$, o que implica $\det A = a_{11}a_{22} = 1$. Teríamos, portanto, $a_{22} = \frac{1}{a_{11}}$, mas como todas as entradas são inteiras, isso implicaria $a_{11} = a_{22} = \pm 1$, e teríamos a contradição $a_{11} + a_{22} = \pm 2$. ■

Lema 9.2. *Os autovalores de A são irracionais.*

Demonstração. Vamos demonstrar o lema por absurdo. Primeiramente notamos que os autovalores de A devem ser reais pois, caso contrário, seriam complexos conjugados. Então seu produto seria $\lambda_+\lambda_- = |\lambda_+|^2 = |\lambda_-|^2 = \pm 1$. Mas isso é absurdo, pois os módulos dos autovalores são diferentes de 1. Consideremos o autovalor λ_+ e suponhamos que ele é um número racional. Usando o resultado da proposição 9.1, ao isolar a raiz na expressão de λ_+ obtemos

$$\sqrt{(a_{11} + a_{22})^2 - 4} = 2\lambda_+ - (a_{11} + a_{22}).$$

Mas $(a_{11} + a_{22})^2 - 4$ é inteiro e sabemos que a raiz de um número inteiro é inteira ou irracional. Devemos ter $2\lambda_+ - (a_{11} + a_{22})$ inteiro, portanto. Fazendo $x = a_{11} + a_{22}$ e $y = 2\lambda_+ - (a_{11} + a_{22})$, devemos ter x e y inteiros e, rearrumando a equação,

$$x^2 - y^2 = 4.$$

De fato, para achar uma solução inteira da equação acima, notamos que $x^2 - y^2 = (x + y)(x - y)$ e temos três opções, já que $4 = 2 \cdot 2 = 4 \cdot 1$. A primeira é $x + y = 2$ e $x - y = 2$. Isso implica $x + y = x - y$, que implica $y = 0$ e, por fim, $a_{11} + a_{22} = x = 2$, o que contradiz a proposição 9.1. As outras duas opções são $x + y = 4$ e $x - y = 1$ ou $x + y = 1$ e $x - y = 4$. Ambas, no entanto, implicam $x = \frac{5}{2}$, o que é um absurdo porque $x = a_{11} + a_{22}$ é um número inteiro. Assim, concluímos que o autovalor λ_+ é irracional. Como $\lambda_- = \frac{1}{\lambda_+}$, ambos autovalores são ambos irracionais. ■

Lema 9.3. *Os autovetores de A são dados por*

$$v_{\lambda_+} = \left(1, \frac{\lambda_+ - a_{11}}{a_{12}}\right) \quad e \quad v_{\lambda_-} = \left(1, \frac{\lambda_- - a_{11}}{a_{12}}\right).$$

Demonstração. Para calcular os autovalores de A , devemos resolver o sistema de equações $(A - \lambda I)v = 0$, em que λ representa qualquer um dos dois autovalores. De fato,

$$(A - \lambda I) = \begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}.$$

Notemos que $a_{11} \neq \lambda$, pois a_{11} é um inteiro e, como provado no lema 9.2, os autovalores de A são irracionais. Assim, concluímos que $a_{11} \neq \lambda$, seja $\lambda = \lambda_+$ ou $\lambda = \lambda_-$. Podemos, portanto, resolver o sistema inicial $(A - \lambda I)v = 0$ usando a matriz M a seguir, obtida a partir da matriz $(A - \lambda I)$ ao se dividir a primeira linha por $a_{11} - \lambda$ e subtraí-la, multiplicada por a_{21} , da segunda linha. Nota-se que, ao se dar o trabalho de calcular o descrito anteriormente, a expressão que aparece na segunda entrada da linha debaixo da matriz M vale 0, pois o numerador é o determinante de $(A - \lambda I)$.

$$M = \begin{bmatrix} 1 & \frac{a_{12}}{a_{11} - \lambda} \\ 0 & 0 \end{bmatrix}.$$

Ao resolvermos $Mv = 0$ para $v = (x, y)$, obtemos a expressão $x + \frac{a_{12}}{a_{11} - \lambda}y = 0$. Por fim, notemos que, pela proposição 9.1, $a_{12} \neq 0$, e temos $y = \frac{\lambda - a_{11}}{a_{12}}x$. Tomando $x = 1$, resultam os autovetores como no enunciado. ■

Definição 9.4. A função distância no toro é definida como

$$d_{\mathbb{T}^2} : \mathbb{T}^2 \times \mathbb{T}^2 \rightarrow \mathbb{R} \\ ([x], [y]) \mapsto \min\{d_{\mathbb{R}^2}(x + m, y + n) : m, n \in \mathbb{Z}^2\}.$$

Essa definição de distância no toro segue a mesma ideia da distância definida no círculo. Consideram-se as menores distâncias entre representantes das classes de equivalência de x e y usando-se a distância padrão definida em \mathbb{R}^2 . Claramente, a função

está bem definida, pois são considerados todos pontos das classes de equivalência de x e y na definição. A função $d_{\mathbb{T}^2}$ é uma distância pois herda as propriedades de distância de $d_{\mathbb{R}^2}$. Essas propriedades não serão demonstradas neste estudo.

Definição 9.5. Um *ponto racional* de \mathbb{T}^2 é um ponto da forma $[r]$ em que $r \in \mathbb{Q}^2$.

De fato, esses pontos estão bem definidos pois, para todo $s \in [r]$, temos $s = r + m$, com $m \in \mathbb{Z}^2$. Logo temos $y \in \mathbb{Q}^2$, já que a soma de um racional com um inteiro ainda é racional.

Lema 9.4. O conjunto dos pontos racionais de \mathbb{T}^2 é denso em \mathbb{T}^2 .

Demonstração. Sejam $\varepsilon > 0$ e $p \in \mathbb{T}^2$. Para mostrar que os pontos racionais são densos em \mathbb{T}^2 , devemos exibir ponto racional r tal que $d_{\mathbb{T}^2}(p, r) \leq \varepsilon$. Seja r um ponto racional de \mathbb{T}^2 . Podemos achar um denominador comum para suas entradas e escrever $r := [(\frac{a}{k}, \frac{b}{k})]$, em que a, b e k são naturais, k não nulo. Ao escolhermos k como o menor natural tal que $\frac{1}{\varepsilon} < k$, podemos dividir o toro como na figura a seguir, em que os intervalos $[0, 1]$ foram divididos em k partes iguais, cada uma com comprimento $\frac{1}{k}$.

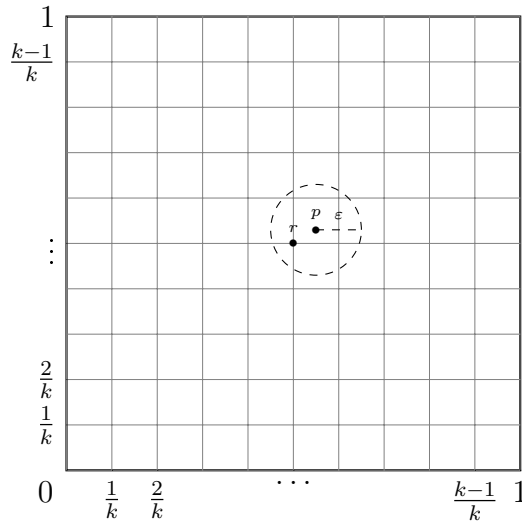


Figura 6: Gráfico do toro \mathbb{T}^2 quadriculado em quadrados de lado $\frac{1}{k}$.

O ponto p estará sempre em algum quadrado da divisão e, como $\frac{1}{k} < \varepsilon$, sempre haverá um ponto racional que esteja próximo o suficiente de p ; basta-nos determinar a e b para termos $d_{\mathbb{T}^2}(p, r) \leq \varepsilon$. ■

Proposição 9.5. O conjunto $Per(L_A)$ dos pontos do toro cujas órbitas são periódicas sob L_A é denso em \mathbb{T}^2 .

Demonstração. No lema anterior, provamos que os pontos racionais de \mathbb{T}^2 são densos em \mathbb{T}^2 . Nessa demonstração, mostraremos que esses pontos são periódicos. Isso significa que os pontos periódicos são, de fato, densos em \mathbb{T}^2 . Seja r um ponto racional de \mathbb{T}^2 da forma $r := \left[\left(\frac{a}{k}, \frac{b}{k}\right)\right]$, em que a , b e k são naturais, k não nulo. Notemos que, ao aplicarmos L_A em r , temos

$$L_A(r) = \left[\left(\frac{(a_{11}a + a_{12}b)}{k}, \frac{(a_{21}a + a_{22}b)}{k} \right) \right],$$

que é da mesma forma que r , pois $(a_{11}a + a_{12}b)$ e $(a_{21}a + a_{22}b)$ são inteiros. Assim concluímos que pontos da forma $\left[\left(\frac{a}{k}, \frac{b}{k}\right)\right]$ continuam dessa forma após aplicarmos L_A e, de fato, L_A simplesmente permuta pontos dessa forma. Como, para termos classes de equivalência diferentes, há k escolhas de valores tanto para a como para b , existem k^2 pontos do toro dessa forma. Pelo princípio das gavetas, deve haver inteiros i e j tais que $L_A^i(r) = L_A^j(r)$ e $|i - j| \leq k^2$. Sem perda de generalidade, supomos $j \geq i$ e, aplicando L_A^{-i} ao dois lados dessa igualdade, temos $r = L_A^{j-i}(r)$; ou seja, r tem órbita periódica de período menor ou igual a k^2 . Concluimos, assim, que o conjunto dos pontos cujas órbitas são periódicas sob L_A é denso em \mathbb{T}^2 . ■

Após essa discussão inicial sobre o sistema (\mathbb{T}^2, L_A) , seguimos para abordar a ideia de variedades estável e instável de um ponto do toro. Para um ponto p do toro, definimos a variedade estável de p como a projeção no toro da reta em \mathbb{R}^2 que passa por p e é paralela ao autovetor v_{λ_-} de A associado ao autovalor $\lambda_- < 1$. Analogamente, definimos a variedade instável de p como a projeção no toro da reta em \mathbb{R}^2 que passa por p e é paralela ao autovetor v_{λ_+} de A associado ao autovalor $\lambda_+ > 1$. Segue a definição formal e sucinta.

Definição 9.6. Seja $[p] \in \mathbb{T}^2$. Definimos a *variedade estável* de $[p]$ como

$$W^s([p]) := \{[p + \alpha v_{\lambda_-}] : \alpha \in \mathbb{R}\},$$

Analogamente, definimos a *variedade instável* de $[p]$ por

$$W^u([p]) := \{[p + \alpha v_{\lambda_+}] : \alpha \in \mathbb{R}\}.$$

Note que as variedades estável e instável estão bem definidas pois, para todo $q \in [p]$, temos $q = p + m$, com $m \in \mathbb{Z}^2$. Assim, $[q + \alpha v_{\lambda_-}] = [p + m + \alpha v_{\lambda_-}] = [p + \alpha v_{\lambda_-}]$ e $[q + \alpha v_{\lambda_+}] = [p + m + \alpha v_{\lambda_+}] = [p + \alpha v_{\lambda_+}]$, e segue que $W^s([q]) = W^s([p])$ e $W^u([q]) = W^u([p])$. A próxima proposição justifica a nomeação dos conjuntos $W^s([p])$ e $W^u([p])$.

Teorema 9.6. *Seja $[p]$ um ponto do toro. Então, para um ponto $[p'] \in W^s([p])$,*

$$\lim_{n \rightarrow \infty} d_{\mathbb{T}^2}(L_A^n([p']), L_A^n([p])) = 0.$$

Da mesma forma, para um ponto $[p'] \in W^u([p])$,

$$\lim_{n \rightarrow \infty} d_{\mathbb{T}^2}(L_A^{-n}([p']), L_A^{-n}([p])) = 0.$$

Demonstração. Vamos demonstrar a proposição para a variedade instável. A outra demonstração é análoga. Se $[p'] \in W^s([p])$, ele é da forma $[p'] = [p + \alpha v_{\lambda_-}]$, para $\alpha \in \mathbb{R}$. Assim,

$$L_A([p']) = L_A([p + \alpha v_{\lambda_-}]) = [A^n(p + \alpha v_{\lambda_-})] = [A^n p + \alpha(\lambda_-)^n v_{\lambda_-}]$$

e $L_A^n([p]) = [A^n p]$. Notemos que, como $|\lambda_-| \leq 1$, para n suficientemente grande a distância no toro é igual à distância no plano, pois $|\lambda_-|^n \rightarrow 0$. Assim, vale

$$d_{\mathbb{T}^2}([A^n p + \alpha(\lambda_-)^n v_{\lambda_-}], [A^n p]) = d_{\mathbb{R}^2}(A^n p + \alpha(\lambda_-)^n v_{\lambda_-}, A^n p) = d_{\mathbb{R}^2}(\alpha(\lambda_-)^n v_{\lambda_-}, 0)$$

e segue o enunciado. ■

Proposição 9.7. *As variedades estável e instável de um ponto do toro são densas em \mathbb{T}^2 .*

Demonstração. Seja $[p] = [(p_1, p_2)]$ um ponto de \mathbb{T}^2 . Na demonstração, usaremos $W([p])$ para representar qualquer uma das duas variedades de $[p]$. Queremos mostrar que, dado $\varepsilon > 0$ e $[t] = [(t_1, t_2)] \in \mathbb{T}^2$, existe $[w] = [w_1, w_2] \in W([p])$ tal que $d_{\mathbb{T}^2}([t], [w]) \leq \varepsilon$. Sem perda de generalidade, consideraremos que $p, t, w \in S^1 \times S^1$ para facilitar a demonstração.

Consideremos o conjunto $C := \{[(t_1, r)] : r \in \mathbb{R}\}$. Esse conjunto é um subconjunto de \mathbb{T}^2 e auxiliará na demonstração. Acharemos o ponto $[w]$ nesse conjunto, considerando o conjunto $W([p]) \cap C$. Esse conjunto interseção é interessante pois a distância entre seus pontos pode ser calculada usando a distância do círculo. Tomemos $[c] \in C$, com $c = (t_1, r) \in S^1 \times S^1$ sem perda de generalidade. Notamos que

$$\begin{aligned} d_{\mathbb{T}^2}([t], [c]) &= \min\{d_{\mathbb{R}^2}(t + m, c + n) : m, n \in \mathbb{Z}\} \\ &= \min\{\sqrt{(t_1 + m_1 - t_1 - n_1)^2 + (t_2 + m_2 - r - n_2)^2} : m_1, m_2, n_1, n_2 \in \mathbb{Z}\} \\ &= \min\{\sqrt{(t_2 - r - z)^2} : z \in \mathbb{Z}\} \\ &= \min\{|t_2 - r - z| : z \in \mathbb{Z}\} \\ &= d_{S^1}(t_2, r), \end{aligned}$$

pois $t_2, r \in S^1$. Isso mostra que a distância de um ponto de C ao ponto $[t]$ é igual à distância de suas segundas coordenadas no círculo.

Agora, definamos $l := \frac{\lambda - a_{11}}{a_{12}}$, sendo λ o autovalor λ_+ ou λ_- , de acordo com qual variedade estamos considerando, e seja $\alpha \in \mathbb{R}$. Os pontos de $W([p]) \cap C$ são da forma

$$[p + \alpha v_\lambda] = [(p_1 + \alpha, p_2 + \alpha l)],$$

de modo que $p_1 + \alpha = t_1 + k$, com $k \in \mathbb{Z}$; ou seja, definindo $m := p_2 - (p_1 - t_1)l \bmod 1$, temos pontos da forma

$$[(t_1, m + kl)].$$

Isso significa que os pontos de $W([p]) \cap C$ podem ser obtidos através de uma rotação irracional R_l (definida na seção 4) aplicada em $m \in S^1$. Como sabemos pelo teorema 4.4 que as rotações irracionais são densas em S^1 para qualquer ponto do círculo, dados $\varepsilon > 0$ e $t_2 \in S^1$, existe inteiro a tal que

$$d_{S^1}(t_2, m + al) = d_{S^1}(t_2, R_l^a(m)) \leq \varepsilon.$$

Por fim, tomando $[w] = [t_1, m + al]$, temos $[w] \in W([p]) \cap C$; ou seja, $[w] \in W([p])$, vale

$$d_{\mathbb{T}^2}([t], [w]) = d_{S^1}(t_2, m + al) \leq \varepsilon$$

e está provada a proposição. ■

Proposição 9.8. *As variedades estável e instável de um ponto $[p]$ do toro são invariantes sob ação do automorfismo L_A no sentido que valem $L_A(W^s([p])) = W^s(L_A([p]))$ e $L_A(W^u([p])) = W^u(L_A([p]))$.*

Demonstração. Na demonstração, usaremos W para representar qualquer uma das duas variedades estável e instável. Da mesma forma, v para o autovetor e λ para o autovalor associado à respectiva variedade.

Primeiro demonstraremos que $L_A(W([p])) \subset W(L_A([p]))$. Seja $[x] \in L_A(W([p]))$. Então existe $\alpha \in \mathbb{R}$ tal que $[x] = L_A([p + \alpha v])$. Logo

$$[x] = [A(p + \alpha v)] = [Ap + \alpha Av] = [Ap + \alpha \lambda v].$$

Fazendo $\beta := \alpha \lambda$, temos $\beta \in \mathbb{R}$ e $[x] = [Ap + \beta v] = [L_A([p]) + \beta v]$. Portanto $[x] \in W(L_A([p]))$.

Agora demonstramos que $W(L_A([p])) \subset L_A(W([p]))$. Tomamos $[x] \in W(L_A([p]))$. Então existe $\alpha \in \mathbb{R}$ tal que $[x] = [L_A([p]) + \alpha v]$. Como $\lambda \neq 0$, fazemos $\beta := \frac{\alpha}{\lambda}$. Então temos $\beta \in \mathbb{R}$ e

$$[x] = [Ap + \beta \lambda v] = [Ap + \beta Av] = [A(p + \beta v)] = L_A([p + \beta v])$$

e, portanto, $x \in L_A(W^s(p))$. ■

A seguir, vamos introduzir a ideia de uma dinâmica topologicamente transitiva.

Definição 9.7. Seja (X, f) um sistema dinâmico. Dizemos que a função f é *topologicamente transitiva* se, dados dois conjuntos não-vazios e abertos $U, V \in X$, existe um natural n tal que $f^{-n}(U) \cap V \neq \emptyset$.

Teorema 9.9. A dinâmica do sistema (\mathbb{T}^2, L_A) é topologicamente transitiva.

Demonstração. Para demonstrar isso, tomemos dois conjuntos não-vazios abertos quaisquer no toro, U e V . Agora, como a variedade estável é densa no toro, pegamos um ponto $u \in U$ e tomamos um segmento de reta I_U passando por u que esteja contido de U e seja paralelo a v_{λ_-} . Da mesma forma, pegamos um ponto $v \in V$ e tomamos um segmento de reta I_V passando por v que esteja contido em V e seja paralelo a v_{λ_+} . Como as variedades estável e instável são densas em \mathbb{T}^2 e o comprimento dos intervalos em U e V ambos aumentam ao aplicarmos neles, respectivamente, L_A^{-1} e L_A , existem $p, q \in \mathbb{N}$ tais que $L_A^{-p}(I_U) \cap L_A^q(I_V) \neq \emptyset$. Assim, tomamos um ponto w' nesse conjunto. Segue que $L_A^{-q}(w') \in V$ e $L_A^p(w') \in U$. Definindo $w := L_A^{-q}(w')$ e tomando $n = p + q$, temos que $w \in L_A^{-n}(V) \cap U$ e está provado o teorema. ■

9.1 Entropia do sistema (\mathbb{T}^2, L_A)

Nesta subseção, provaremos algumas propriedades que a matriz A satisfaz para nos permitir calcular a entropia do sistema (\mathbb{T}^2, L_A) . Os cálculos são extensos e podem ser pulados, se assim desejado, pois não contribuem com nenhuma informação importante para o entendimento das proposições. Vamos definir a entropia de um sistema dinâmico e calcular a quantidade de pontos periódicos para um período dado. Após isso, o cálculo em si da entropia do sistema será bem simples.

Proposição 9.10. A n -ésima potência da matriz A é dada por

$$A^n = \begin{bmatrix} a_{11} \cdot \alpha_n - \alpha_{n-1} & a_{12} \cdot \alpha_n \\ a_{21} \cdot \alpha_n & a_{22} \cdot \alpha_n - \alpha_{n-1} \end{bmatrix}$$

em que λ_+ e λ_- são os autovalores de A e

$$\alpha_n := \frac{(\lambda_+)^n - (\lambda_-)^n}{(\lambda_+) - (\lambda_-)}.$$

Demonstração. Primeiramente, notamos que α_n satisfaz a relação de recorrência

$$\alpha_{n+1} = (a_{11} + a_{22})\alpha_n - \alpha_{n-1}.$$

De fato, como os autovalores de A são raízes do polinômio $\lambda^2 - (a_{11} + a_{22})\lambda + 1$, temos

$$\begin{aligned}\alpha_{n+1} &= \frac{(\lambda_+)^{n+1} - (\lambda_-)^{n+1}}{(\lambda_+) - (\lambda_-)} = \frac{(\lambda_+)^{n-1}(\lambda_+)^2 - (\lambda_-)^{n-1}(\lambda_-)^2}{(\lambda_+) - (\lambda_-)} \\ &= \frac{(\lambda_+)^{n-1}((a_{11} + a_{22})\lambda_+ - 1) - (\lambda_-)^{n-1}((a_{11} + a_{22})\lambda_- - 1)}{(\lambda_+) - (\lambda_-)} \\ &= (a_{11} + a_{22})\frac{(\lambda_+)^n - (\lambda_-)^n}{(\lambda_+) - (\lambda_-)} - \frac{(\lambda_+)^{n-1} - (\lambda_-)^{n-1}}{(\lambda_+) - (\lambda_-)} \\ &= (a_{11} + a_{22})\alpha_n - \alpha_{n-1}.\end{aligned}$$

Os autovetores de A são dados por

$$v_{\lambda_+} = \left(1, \frac{\lambda_+ - a_{11}}{a_{12}}\right) \quad \text{e} \quad v_{\lambda_-} = \left(1, \frac{\lambda_- - a_{11}}{a_{12}}\right).$$

Como $\lambda_+ \neq \lambda_-$, eles são linearmente independentes. Isso garante que a matriz A pode ser escrita como conjugada de uma matriz diagonal, como na fórmula $A = PDP^{-1}$. As matrizes D e P são dadas por

$$D = \begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix} \quad \text{e} \quad P = \begin{bmatrix} 1 & 1 \\ \frac{\lambda_+ - a_{11}}{a_{12}} & \frac{\lambda_- - a_{11}}{a_{12}} \end{bmatrix}.$$

Invertendo P , temos

$$P^{-1} = \begin{bmatrix} \frac{\lambda_- - a_{11}}{\lambda_- - \lambda_+} & -a_{12} \\ -\frac{\lambda_+ - a_{11}}{\lambda_- - \lambda_+} & a_{12} \end{bmatrix}.$$

Assim, calculando A^n , temos que

$$A^n = (PDP^{-1})^n = PD^nP^{-1},$$

pois os fatores P e P^{-1} se cancelam no produto. Logo, temos

$$\begin{aligned}A^n &= \begin{bmatrix} 1 & 1 \\ \frac{\lambda_+ - a_{11}}{a_{12}} & \frac{\lambda_- - a_{11}}{a_{12}} \end{bmatrix} \begin{bmatrix} (\lambda_+)^n & 0 \\ 0 & (\lambda_-)^n \end{bmatrix} \begin{bmatrix} \frac{\lambda_- - a_{11}}{\lambda_- - \lambda_+} & -a_{12} \\ -\frac{\lambda_+ - a_{11}}{\lambda_- - \lambda_+} & a_{12} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ \frac{\lambda_+ - a_{11}}{a_{12}} & \frac{\lambda_- - a_{11}}{a_{12}} \end{bmatrix} \begin{bmatrix} (\lambda_+)^n \frac{\lambda_- - a_{11}}{\lambda_- - \lambda_+} & -a_{12}(\lambda_+)^n \\ -(\lambda_-)^n \frac{\lambda_+ - a_{11}}{\lambda_- - \lambda_+} & a_{12}(\lambda_-)^n \end{bmatrix} \\ &= \begin{bmatrix} \frac{(\lambda_+)^n(\lambda_- - a_{11}) - (\lambda_-)^n(\lambda_+ - a_{11})}{\lambda_- - \lambda_+} & -a_{12}((\lambda_+)^n - (\lambda_-)^n) \\ \frac{(\lambda_+ - a_{11})(\lambda_- - a_{11})((\lambda_+)^n - (\lambda_-)^n)}{a_{12}(\lambda_- - \lambda_+)} & -\frac{(\lambda_+)^n(\lambda_+ - a_{11}) - (\lambda_-)^n(\lambda_- - a_{11})}{\lambda_- - \lambda_+} \end{bmatrix}.\end{aligned}$$

Notando que

$$\begin{aligned}
(\lambda_+ - a_{11})(\lambda_- - a_{11}) &= \lambda_+ \lambda_- - 2a_{11}(\lambda_+ + \lambda_-) + a_{11}^2 \\
&= 1 - a_{11}(a_{11} + a_{22}) + a_{11}^2 \\
&= 1 - a_{11}(a_{11} + a_{22}) + a_{11}^2 \\
&= -a_{12}a_{21}
\end{aligned}$$

e que

$$\alpha_{n+1} - a_{11}\alpha_n = a_{22}\alpha_n - \alpha_{n-1},$$

concluimos que

$$A^n = \begin{bmatrix} a_{11} \cdot \alpha_n - \alpha_{n-1} & a_{12} \cdot \alpha_n \\ a_{21} \cdot \alpha_n & a_{22} \cdot \alpha_n - \alpha_{n-1} \end{bmatrix}.$$

■

Proposição 9.11. *O determinante de $A^n - I$ é $\det(A^n - I) = 2 - (\lambda_+)^n - (\lambda_-)^n$.*

Demonstração. Vamos calcular $\det(A^n - I)$.

$$\begin{aligned}
\det(A^n - I) &= \begin{vmatrix} a_{11}\alpha_n - \alpha_{n-1} - 1 & a_{12}\alpha_n \\ a_{21}\alpha_n & a_{22}\alpha_n - \alpha_{n-1} - 1 \end{vmatrix} \\
&= (a_{11}\alpha_n - (\alpha_{n-1} + 1))(a_{22}\alpha_n - (\alpha_{n-1} + 1)) - a_{12}a_{21}(\alpha_n)^2 \\
&= (a_{11}a_{22} - a_{12}a_{21})(\alpha_n)^2 - (a_{11} + a_{22})\alpha_n(\alpha_{n-1} + 1) + (\alpha_{n-1} + 1)^2 \\
&= 1 + [2\alpha_{n-1} - (a_{11} + a_{22})\alpha_n] + [(\alpha_n)^2 - (a_{11} + a_{22})\alpha_n\alpha_{n-1} + (\alpha_{n-1})^2].
\end{aligned}$$

Como, pela relação de recorrência de α_n ,

$$\begin{aligned}
2\alpha_{n-1} - (a_{11} + a_{22})\alpha_n &= \alpha_{n-1} - \alpha_{n+1} \\
(a_{11} + a_{22})\alpha_n\alpha_{n-1} - (\alpha_{n-1})^2 &= \alpha_{n+1}\alpha_{n-1},
\end{aligned}$$

segue que

$$\det(A^n - I) = 1 - \alpha_{n+1} + \alpha_{n-1} + (\alpha_n)^2 - \alpha_{n+1}\alpha_{n-1}.$$

Vamos provar algumas relações úteis para prosseguir.

$$\begin{aligned}
(\alpha_n)^2 - \alpha_{n+1}\alpha_{n-1} &= \frac{[(\lambda_+)^n - (\lambda_-)^n]^2 - [(\lambda_+)^{n+1} - (\lambda_-)^{n+1}][(\lambda_+)^{n-1} - (\lambda_-)^{n-1}]}{(\lambda_+ - \lambda_-)^2} \\
&= \frac{(\lambda_+)^{2n} - 2 + (\lambda_-)^{2n} - (\lambda_+)^{2n} + (\lambda_+)^2 + (\lambda_-)^2 - (\lambda_-)^{2n}}{(\lambda_+ - \lambda_-)^2} \\
&= \frac{(\lambda_+)^2 - 2 + (\lambda_-)^2}{(\lambda_+ - \lambda_-)^2} = 1.
\end{aligned}$$

$$\begin{aligned}
\alpha_{n-1} - \alpha_{n+1} &= \frac{(\lambda_+)^{n-1} - (\lambda_-)^{n-1} - (\lambda_+)^{n+1} + (\lambda_-)^{n+1}}{\lambda_+ - \lambda_-} \\
&= \frac{(\lambda_+)^n(\lambda_+^{-1} - \lambda_+) - (\lambda_-)^n(\lambda_-^{-1} - \lambda_-)}{\lambda_+ - \lambda_-} \\
&= -(\lambda_+)^n - (\lambda_-)^n
\end{aligned}$$

Daí, segue facilmente que $\det(A^n - I) = 2 - (\lambda_+)^n - (\lambda_-)^n$. ■

Definição 9.8. A *entropia* de um sistema dinâmico (X, f) é definida como

$$h_f := \lim_{n \rightarrow \infty} \frac{|Per_n(f)|}{n}$$

em que $Per_n(f) := \{x \in X : f^n(x) = x\}$.

Teorema 9.12. $|Per_n(L_A)| = |\det(A^n - I)|$.

Demonstração. Para achar a quantidade de pontos de período n sob L_A , temos que achar as soluções de

$$A^n x = x + m, \tag{1}$$

sendo x in $S^1 \times S^1$ e $m \in \mathbb{Z}^2$. De fato, se observamos a transformação $T := A^n - I$ aplicada ao quadrado $S^1 \times S^1$, subconjunto de \mathbb{R}^2 , vemos que o quadrado é levado em um paralelogramo $T(S^1 \times S^1)$ cujos vértices são pontos inteiros de \mathbb{R}^2 . Notemos que os pontos inteiros que pertencem ao paralelogramo são as soluções de (1). Pelo Teorema de Pick, sabemos que, se a é a área de $T(S^1 \times S^1)$, i a quantidade de pontos inteiros internos a $T(S^1 \times S^1)$ e b a quantidade de pontos inteiros na fronteira de $T(S^1 \times S^1)$, então

$$a = i + \frac{b}{2} - 1.$$

É possível mostrar que os únicos pontos na fronteira do paralelogramo $T(S^1 \times S^1)$ são seus vértices e assim concluir que $a = i + 1$. A partir disso, como os vértices do paralelogramo representam todos o mesmo ponto de \mathbb{T}^2 , concluímos que os pontos que satisfazem (1) são somente o ponto $(0, 0)$ e os pontos internos do paralelogramo, e a quantidade deles é, portanto, igual à área do paralelogramo. Basta-nos, portanto, calcular a área do paralelogramo $T(S^1 \times S^1)$. Mas isso é simples, pois ela é igual ao módulo do determinante da transformação, já que a área de $S^1 \times S^1$ é 1. Assim, concluímos que $|Per_n(L_A)| = |\det(A^n - I)|$. ■

Teorema 9.13. A *entropia* de sistema (\mathbb{T}^2, L_A) associado à matriz A é

$$h_{L_A} = \log |\lambda_+|.$$

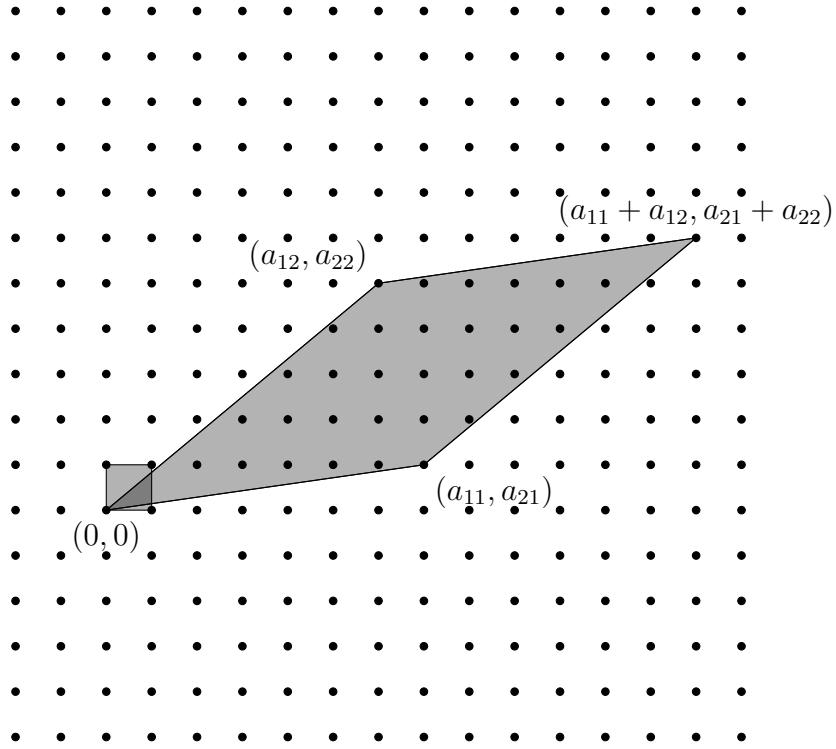


Figura 7: Exemplo de um possível paralelogramo $T(S^1 \times S^1)$

Demonstração. De fato, dos resultados anteriores vem que

$$\begin{aligned}
 h_{L_A} &= \frac{\log |\det(A^n - I)|}{n} \\
 &= \frac{\log |2 - (\lambda_+)^n - (\lambda_-)^n|}{n} \\
 &= \frac{\log |(\lambda_+)^n [1 - 2(\lambda_-)^n + (\lambda_-)^{2n}]|}{n} \\
 &= \log |\lambda_+| + \frac{\log |1 - 2(\lambda_-)^n + (\lambda_-)^{2n}|}{n}
 \end{aligned}$$

Daí, segue claramente o resultado procurado, pois $|\lambda_-| < 1$. ■

10 Cadeias de Markov

Nesta seção, estudaremos um tipo de sistema dinâmico que está relacionado ao Shift de Bernoulli. Consideraremos um subespaço do espaço dos símbolos Σ_n definido por uma matriz chamada matriz de transição da cadeia de Markov, cujas entradas assumem somente os valores 0 ou 1. Como na definição de Σ_n , podemos considerar sequências unilaterais ou bilaterais. Nesse capítulo, consideraremos somente sequências bilaterais; ou seja, consideraremos $\Sigma_n := \{0, \dots, n-1\}^{\mathbb{Z}}$. Ainda é importante ressaltar que, visando a facilitar a notação, consideraremos que os índices de uma matriz $n \times n$ variam em $\{0, \dots, n-1\}$; isto é, as primeiras linha e coluna da matriz são as 0-ésimas e as últimas são as $(n-1)$ -ésimas.

Definição 10.1. Seja $M = (m_{ij})$ uma matriz $n \times n$ com entradas $m_{ij} \in \{0, 1\}$ para todo $i, j \in \{0, \dots, n-1\}$. O espaço dos símbolos restrito pela matriz M é definido por

$$\Sigma_M := \{(s_k)_{k \in \mathbb{N}} \in \Sigma_n : m_{s_k s_{k+1}} = 1 \quad \forall k \in \mathbb{N}\}.$$

Definição 10.2. Uma cadeia de Markov com respeito à matriz de transição M , também chamada de *subshift finito*, é a restrição da função shift σ , definida na seção 6, ao espaço Σ_M . Essa função é denotada por σ_M e dada por

$$\begin{aligned} \sigma_M : \Sigma_M &\rightarrow \Sigma_M \\ x &\mapsto \sigma(x). \end{aligned}$$

Na próxima proposição, vamos calcular a quantidade de pontos periódicos de um subshift σ_M associado a uma matriz de transição M qualquer.

Teorema 10.1. *Seja M uma matriz $n \times n$, Σ_M o espaço dos símbolos restrito por M e σ_M o respectivo subshift. Então a quantidade de pontos periódicos de período $p > 0$ é igual a $|Per_p(\sigma_M)| = tr(A^p)$.*

Demonstração. Seja $(s_k)_{k \in \mathbb{N}} \in \Sigma_M$. Sabemos, pelo teorema 6.1, que (s_k) é periódico de período p se, e somente se, para todo inteiro i , vale $s_i = s_{i+p}$. Sendo assim, os pontos procurados serão da forma

$$(s_k) = (\dots, s_{i_1}, s_{i_2}, \dots, s_{i_{p-1}}, s_{i_p}, \dots)$$

e, portanto, devemos achar todos os vetores $(i_1, \dots, i_p) \in \{0, \dots, n-1\}^p$ tais que a sequência

$$i_1 \rightarrow i_2 \rightarrow i_3 \rightarrow \dots \rightarrow i_{p-1} \rightarrow i_p \rightarrow i_1$$

seja uma sequência permitida pela matriz M . Essa condição equivale a

$$m_{i_1, i_2} = m_{i_2, i_3} = \dots = m_{i_{p-1}, i_p} = m_{i_p, i_1} = 1,$$

e, como o produto de entradas da matriz M vale 1 se todas entradas valem 1 e zero caso contrário, isso significa que a quantidade de pontos periódicos procurada é

$$\begin{aligned}
 |Per_p(\sigma_M)| &= \sum_{\substack{(i_1, \dots, i_p) \\ \in \\ \{0, \dots, n-1\}^p}} m_{i_1, i_2} m_{i_2, i_3} \cdots m_{i_{p-1}, i_p} m_{i_p, i_1} \\
 &= \sum_{i_1=1}^n \sum_{\substack{(i_2, \dots, i_p) \\ \in \\ \{0, \dots, n-1\}^{p-1}}} m_{i_1, i_2} m_{i_2, i_3} \cdots m_{i_{p-1}, i_p} m_{i_p, i_1} \\
 &= \sum_{i_1=1}^n (A^p)_{i_1 i_1} = tr(A^p).
 \end{aligned}$$

■

11 Conclusão

Este projeto foi muito produtivo para a compreensão dos conceitos básicos da teoria dos sistemas dinâmicos. Neste relatório, foi possível explorar os conceitos de o que são sistemas dinâmicos e entender as motivações matemáticas relacionadas a eles. O contato com a abordagem acadêmica e o ritmo de pesquisa adotado distoam do paradigma de estudos da graduação e são parte importante da formação do aluno. Os objetivos listados no projeto de pesquisa para o primeiro e segundo relatórios foram cumpridos. O estudo da rotação e expansão no círculo foram uma boa introdução à linguagem de sistemas dinâmicos. A ferradura de Smale foi importante para a compreensão de um modelo clássico da área. O estudo mais aprofundado dos automorfismos hiperbólicos foi bastante instigante e criou a possibilidade de uma possível continuação dos estudos de automorfismos no toro. De modo geral, o projeto foi decisivo na vida acadêmica do aluno, possibilitando acesso a muitas áreas diferentes da matemática e demonstrando como é o exercício de uma atividade acadêmica.

Referências

- [1] Brin, Michael, and Stuck, Garret. *Introduction to Dynamical Systems*. Cambridge University Press, 2003.
- [2] Devaney, Robert L. *An Introduction to Chaotic Dynamical Systems* 2nd ed. Addison-Wesley, 1989
- [3] Pollicott, M. and Yuri, M., *Dynamical Systems and Ergodic Theory*, London Math. Soc., 1998.
- [4] Barreira, L. and Valls, C., *Dynamical Systems - An Introduction*, IST Press 2012, Instituto Superior Técnico.