Sistemas Dinâmicos Discretos Lineares
notas de aula — Análise Linear

Isabel S. Labouriau

Novembro de 1991

Índice

Capítulo 1

Introdução, definições e exemplos

A teoria de *sistemas dinâmicos discretos* ou *relações de recorrência* estuda as imagens sucessivas de uma aplicação $F : A \longrightarrow A$. A aplicação F é chamada *relação de recorrência* ou *processo iterativo* ou *sistema dinâmico discreto*. Tratamos aquí apenas os casos em que A é um subconjunto de $\mathbf{R}^n$. Inicialmente estudaremos aplicações F com inversa F^{-1}.

Denotamos por F^n a composta de F n vezes, isto é $F^n = F \circ F^{n-1}$, com a convenção de que F^0 é a aplicação identidade de A e $F^{-n} = (F^{-1})^n$. A *órbita* $\mathcal{O}_F(x)$ de um ponto $x \in A$ é a sucessão $(F^m(x))$ com $m \in \mathbf{Z}$, isto é

$$\mathcal{O}_F(x) = \left(\ldots, F^{-n}(x), \ldots, F^{-2}(x), F^{-1}(x), x, F(x), F^2(x) \ldots, F^n(x), \ldots\right).$$

Analogamente definimos a *semi-órbita positiva* $\mathcal{O}_F^+(x)$ e a *semi-órbita negativa*, $\mathcal{O}_F^-(x)$ de um ponto $x \in A$ como as sucessões

$$\mathcal{O}_F^+(x) = (F^m(x)) \qquad \mathcal{O}_F^-(x) = \left(F^{-m}(x)\right) \qquad m \in \mathbf{N}.$$

Um ponto $x \in A$ é um *ponto fixo* da relação de recorrência F quando $F(x) = x$, isto é, $\mathcal{O}_F(x) = \{x\}$. Dizemos que um ponto $x \in A$ é um *ponto periódico* da relação de recorrência F quando $F^k(x) = x$ para algum $k \in \mathbf{Z}$, isto é, quando $\{F^n(x)\}$ é um conjunto finito. O período de x é o menor $k > 0$ para o qual $F^k(x) = x$. Observe-se que neste caso x é um ponto fixo da relação de recorrência F^k.

Dada uma relação de recorrência, estamos interessados em descrever os seus pontos periódicos (incluindo os pontos fixos). Para os demais pontos queremos descrever a geometria das órbitas e o comportamento de $F^n(x)$ quando n tende para $\pm\infty$.

1.1 exemplo: transformação linear em R

Para cada $\alpha \in \mathbf{R} - \{0\}$, a função $F_\alpha : \mathbf{R} \longrightarrow \mathbf{R}$, $F_\alpha(x) = \alpha x$ define uma relação de recorrência em $\mathbf{R}$. Como $\mathcal{O}_{F_\alpha}(x) = (\alpha^n x)$ para $n \in \mathbf{Z}$, os pontos fixos de F_α são soluções da equação $F_\alpha^n(x) = \alpha^n x = x$. Se $|\alpha| \neq 1$ então o único ponto fixo de F_α é $x = 0$ e não há pontos periódicos de nenhum período diferente de 1.

Se $|\alpha| < 1$ então para qualquer ponto $x \in \mathbf{R}$, $x \neq 0$, a sucessão $|\alpha|^n|x| = |F_\alpha^n(x)|$ satisfaz

$$\lim_{n\to+\infty} |F_\alpha^n(x)| = 0 \qquad e \qquad \lim_{n\to-\infty} |F_\alpha^n(x)| = \infty.$$

Neste caso 0 é o único ponto de acumulação de $\mathcal{O}^+_{F_\alpha}(x)$ e da órbita $\mathcal{O}_{F_\alpha}(x)$ para qualquer valor de $x \in \mathbf{R}$. Para $x \neq 0$, a semi-órbita $\mathcal{O}^-_{F_\alpha}(x)$ não tem pontos de acumulação.

Para $|\alpha| > 1$, temos:

$$\lim_{n\to-\infty} |F_\alpha^n(x)| = 0 \qquad e \qquad \lim_{n\to-\infty} |F_\alpha^n(x)| = \infty,$$

e a origem é o único ponto de acumulação de $\mathcal{O}_{F_\alpha}(x)$ e de $\mathcal{O}^-_{F_\alpha}(x)$

Se $\alpha = 1$ então todos os pontos de $\mathbf{R}$ são pontos fixos (F_α é a aplicação identidade). Para $\alpha = -1$ só a origem é um ponto fixo de F_α, todos os demais pontos de $\mathbf{R}$ são periódicos com período 2 (pontos fixos da recorrência F_α^2 que é a aplicação identidade).

1.2 exemplo: teorema da contração

Se existir $K \in \mathbf{R}$, $0 < K < 1$ tal que $\forall x, y \in \mathbf{R}^n \ |F(x) - F(y)| < K|x - y|$ então dizemos que $F : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ é uma *contração*.

Teorema 1.2.1 *Se F é uma contração então F tem um único ponto fixo, que é o único ponto de acumulação de qualquer semi-órbita positiva de F.*

Demonstração: Primeiro demonstraremos que todas as semi-órbitas positivas têm um único ponto de acumulação, verificando que para qualquer $x \in \mathbf{R}^n$ a sucessão $(F^m(x))$ é uma sucessão de Cauchy. A seguir verificaremos que o limite desta sucessão é um ponto fixo de F e por fim que o ponto fixo é único.

Para verificar que $(F^m(x))$ é uma sucessão de Cauchy, é necessário estimar $|F^{m+j}(x) - F^m(x)|$ dados $m, j \in \mathbf{N}$ e $x \in \mathbf{R}^n$. Como $F^{m+j}(x) = F^m(F^j(x))$, podemos calcular:

$$|F^{m+j}(x) - F^m(x)| = |F^m(F^j(x)) - F^m(x)| \leq K^m|F^j(x) - x|.$$

Como $K < 1$ sabemos que $\lim_{m\to\infty} K^m = 0$ e por isso se obtivermos uma estimativa de $|F^j(x) - x|$ que não dependa de j poderemos concluir que $(F^m(x))$ é uma sucessão de Cauchy. Para isto observamos que

$$\begin{aligned} |F^j(x) - x| &\leq |F^j(x) - F^{j-1}(x)| + |F^{j-1}(x) - x| \\ &\leq K^{j-1}|F(x) - x| + |F^{j-1}(x) - x| \\ &\leq K^{j-1}|F(x) - x| + K^{j-2}|F(x) - x| + |F^{j-2}(x) - x| \\ &\leq \left(K^{j-1} + K^{j-2} + \cdots + K^2 + K + 1\right)|F(x) - x|. \end{aligned}$$

Podemos estimar o coeficiente de $|F(x)-x|$ usando a série geométrica: como $\sum_{n=0}^{\infty} K^n = 1/(1-K)$ porque $0 < K < 1$ e então $(K^{j-1} + \cdots + K^2 + K + 1) < 1/(1-K)$. Usando a primeira desigualdade, obtemos:

$$|F^{m+j}(x) - F^m(x)| \leq \frac{K^m}{1-K}|F(x) - x|.$$

Sabendo que $K^m/(1-K)$ tende para zero quando m tende para infinito e que $|F(x)-x|$ não depende de m, concluímos que para qualquer x dado $(F^m(x))$ é uma sucessão de Cauchy e por isso é convergente.

Se L é o limite da sucessão $(F^m(x))$ então é fácil calcular $F(L)$. A condição de ser contração garante que F é contínua e por isso a imagem do limite de uma sucessão convergente é o limite da imagem da sucessão, isto é, $F(L)$ é o limite da sucessão

$$F(x), F^2(x), F^3(x), F^4(x), F^5(x), \ldots \longrightarrow F(L)$$

que converge para o mesmo limite que a sucessão

$$x, F(x), F^2(x), F^3(x), F^4(x), F^5(x), \ldots \longrightarrow L$$

Assim vemos que $F(L) = L$, isto é, L é um ponto fixo.

Resta mostrar que o ponto fixo é único. Se existissem dois pontos fixos, x_1 e x_2 então ele satisfariam $|x_1-x_2| = |F(x_1)-F(x_2)| < K|x_1-x_2| < |x_1-x_2|$, o que é um absurdo. □

O uso deste teorema em análise numérica é conhecido como *método iterativo simples*.

É fácil ver que se G é a inversa de uma contração então existe $D \in \mathbf{R}$, $D > 1$ tal que $\forall x, y \in \mathbf{R}^n$ $|G(x) - G(y)| > D|x - y|$. Neste caso dizemos que G é uma *dilatação*.

1.3 exemplo: método de Newton

Dada uma função $f : \mathbf{R} \longrightarrow \mathbf{R}$, diferenciável, uma maneira de calcular os zeros de f é aplicar a relação de recorrência

$$F(x) = x - \frac{f(x)}{f'(x)}.$$

Os pontos fixos de F serão os zeros de f e o método converge sempre que F está bem definida e é uma contração.

1.4 exemplo: discretização de uma equação diferencial

Se $v : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ é um campo de vetores com fluxo associado $\varphi_v(t,p)$ podemos definir uma relação de recorrência estreitamente ligada ao fluxo (a *aplicação tempo 1* de v) por $F(p) = \varphi_v(1,p)$ desde que o fluxo φ_v esteja sempre definido para $t = 1$. Se φ_v estiver definido para todo $t \in \mathbf{R}$, então segue-se das propriedades dos fluxos que $F^n(p) = \varphi_v(n,p)$ para todo ponto p.

Os pontos de equilíbrio da equação diferencial são sempre pontos fixos de F, já que se $\varphi_v(t,p) = p$ para todo t então em particular $F(p) = \varphi_v(1,p) = p$. Outros pontos fixos são os valores iniciais em $\mathbf{R}^n$ que correspondem a soluções periódicas da equação diferencial cujo período é $1/n$ para algum $n \in \mathbf{N}$.

A órbita de p pela relação de recorrência F é sempre um subconjunto da trajetória de p pela equação diferencial $\dot{x} = v(p)$, já que $F^n(p) = \varphi_v(n,p)$, e as imagens sucessivas de p por L estão ordenadas sobre a curva $x(t) = \varphi_v(t,p)$ segundo a orientação da trajetória.

Capítulo 2

Sistemas dinâmicos lineares em $\mathbf{R}^2$

Uma transformação linear com inversa $L : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ é chamada um *automorfismo linear* de $\mathbf{R}^n$. Para começar, estudaremos os automorfismos lineares do plano:

2.1 automorfismo diagonalizável

Seja $L : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ dado por uma matriz diagonal:

$$D = \begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix} \qquad D^n = \begin{pmatrix} {\gamma_1}^n & 0 \\ 0 & {\gamma_2}^n \end{pmatrix} \tag{2.1}$$

Para garantir que L tem inversa supomos que $\det(L) = \gamma_1 \cdot \gamma_2 \neq 0$.

Os dois eixos coordenados de $\mathbf{R}^2$ (os autoespaços V_{γ_1} e V_{γ_2}) são invariantes por L . Em cada um destes espaços L é a multiplicação pelo número real γ_j e o comportamento de L em V_{γ_j} é o descrito na seção 1.1. A órbita de um ponto qualquer, $(p_1, p_2) \in \mathbf{R}^2$ é dada por $\mathcal{O}_L(p_1, p_2) = ((\gamma_1^n p_1, \gamma_2^n p_2))$.

Suponhamos, por exemplo, que $0 < \gamma_1 < \gamma_2 < 1$. Neste caso para qualquer ponto (p_1, p_2) do plano

$$|L(p_1, p_2)| = |(\gamma_1 p_1, \gamma_2 p_2)| \leq |\gamma_2| \, |(p_1, p_2)| \tag{2.2}$$

e L é uma contração. Usando o teorema 1.2.1 concluímos que L tem um único ponto fixo, que, como L é linear, é a origem de $\mathbf{R}^2$. Se $p \neq 0$ é um ponto de $\mathbf{R}^2$ então a semi-órbita positiva de p aproxima-se da origem e os pontos da semi-órbita negativa afastam-se de $(0, 0)$.

Como γ_1 e γ_2 são positivos, podemos encontrar uma matriz diagonal $\tilde{D}$ tal que L é a aplicação tempo 1 (ver exemplo (1.4)) da equação $\dot{x} = \tilde{D}x$. De fato, se tomarmos

$$\tilde{D} = \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix} \qquad \text{com} \qquad \begin{array}{l} \delta_1 = \ln \gamma_1 \\ \delta_2 = \ln \gamma_2. \end{array}$$

então o fluxo associado à equação $\dot{x} = \tilde{D}x$ é dado por

$$\varphi_{\tilde{D}}(t, (p_1, p_2)) = (\mathrm{e}^{t\delta_1} p_1, \mathrm{e}^{t\delta_2} p_2)$$

e a aplicação tempo 1 é dada por

$$\varphi_{\tilde{D}}(1, (p_1, p_2)) = (\mathrm{e}^{\delta_1} p_1, \mathrm{e}^{\delta_2} p_2) = (\mathrm{e}^{\ln \gamma_1} p_1, \mathrm{e}^{\ln \gamma_2} p_2) = (\gamma_1 p_1, \gamma_2 p_2) = L(p_1, p_2).$$

Para qualquer ponto p do plano, $\mathcal{O}_L(p)$ está contido na trajetória da equação diferencial que passa por p, como na figura a seguir. Para cada valor de p o automorfismo L transforma pontos de uma das curvas $x(t) = \varphi_{\tilde{D}}(t, p)$ em pontos da mesma curva, já que $Lx(t) = \varphi_{\tilde{D}}(t+1, p)$. Quando uma curva tem esta propriedade dizemos que a curva é *invariante por* L.

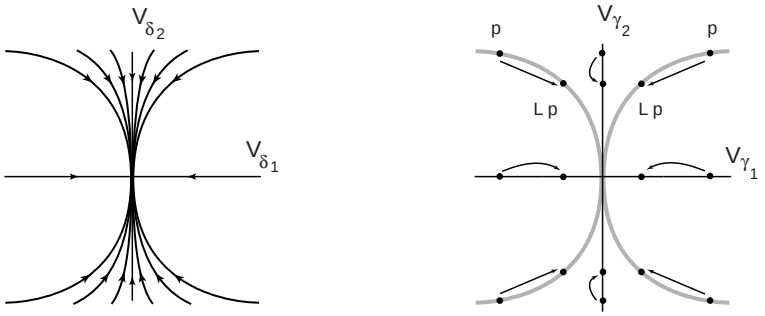


Diagrama de fase de $\dot{x} = \tilde{D}x$ e órbitas de L.

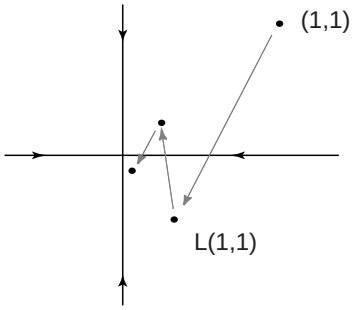
2.2 exemplo: autovalores negativos

Sempre que $|\gamma_1| < |\gamma_2| < 1$ podemos usar a expressão (2.2) para mostrar que L é uma contração, mas nem sempre é possível encontrar uma equação diferencial linear como acima para a qual L é a aplicação tempo 1.

Suponha que L seja dado pela matriz:

$$C = \begin{pmatrix} 1/3 & 0 \\ 0 & -1/2 \end{pmatrix}$$

É fácil ver que se L fosse a aplicação tempo 1 de $\dot{x} = \tilde{D}x$ então os subespaços invariantes por L e por $\tilde{D}$ seriam os mesmos. Então $\tilde{D}$ seria uma matriz diagonal como acima, cada um dos semi-eixos horizontais seria uma trajetória de $\dot{x} = \tilde{D}x$ e a origem seria um ponto de equilíbrio. As imagens sucessivas por L de cada ponto estariam sobre uma trajetória da equação diferencial. Como duas trajetórias não se cruzam, então a órbita de um ponto do semi-plano superior estaria contida neste semi-plano. Calculando imagens sucessivas do ponto $(1,1)$ obtemos $L(1,1) = (1/3, -1/2)$, $L^2(1,1) = (1/9, 1/4)$, $L^3(1,1) = (1/27, -1/8)$ que oscilam entre o semi-plano supeiror e o semi-plano inferior de $\mathbf{R}^2$. A trajetória de $(1,1)$ pela equação diferencial teria que passar por todos estes pontos, o que é impossível se a trajetória não sair do semi-plano superior.



Voltando ao caso geral em que L é diagonalizável, na base em que L se escreve como a matriz diagonal (2.1), a órbita (x_n, y_n) de um ponto (p_1, p_2) de $\mathbf{R}^2$ com $p_1 \neq 0$ e $p_2 \neq 0$ satisfaz

$$\frac{y_n}{p_2} = \left(\frac{\gamma_2}{\gamma_1}\right)^n \frac{x_n}{p_1}$$

e se $|\gamma_1| < |\gamma_2|$ então quando n tende para $+\infty$ a órbita se aproxima do eixo y. Em geral numa base qualquer $\mathcal{O}_L^+(p)$ se aproxima do autoespaço associado ao autovalor de maior módulo, neste caso γ_2, desde que p não pertença ao outro autoespaço que aquí é o associado a γ_1.

Analogamente se começarmos com um ponto p fora do autoespaço associado a γ_2 então quando n tende para infinito os pontos $L^{-n}(p)$ se aproximam do autoespaço associado a γ_1 que é o autovalor de menor módulo.

Como consequência obtemos:

Teorema 2.2.1 *Suponha que $L : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ é um automorfismo linear com autovalores reais distintos e considere a aplicação $F : \mathbf{R}^2 - \{0\} \longrightarrow \mathbf{R}^2 - \{0\}$ dada por*

$$F(p) = \frac{L^2 p}{|L^2 p|}.$$

Para qualquer ponto de $\mathbf{R}^2 - \{0\}$ *a sucessão* $(F^n(p))$ *converge para um autovetor de* L *com norma igual a 1.*

Este resultado é conhecido em análise numérica como *método da potência iterada.*

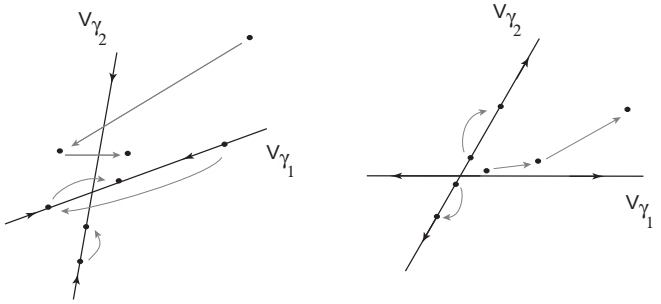
Demonstração: Consideremos a aplicação

$$G(p) = \frac{Lp}{|Lp|}.$$

Já vimos que as imagens sucessivas de um ponto $p \neq 0$ qualquer, $L^n(p)$ aproximam-se sempre de um dos autoespaços de L — se p estiver em algum dos autoespaços a sua órbita está contida neste autoespaço, caso contrário a órbita se aproxima do autoespaço associado ao autovalor de maior módulo. Cada um dos vetores $G^n(p)$ é paralelo a $L^n(p)$ e tem norma 1, de modo que se a sucessão $(G^n(p))$ convergir o seu limite será um autovetor de L com norma 1.

No entanto esta sucessão nem sempre converge: se o autoespaço de que a sucessão se aproxima corresponder a um autovalor negativo e se v for um autovetor com norma 1 então a sucessão se aproxima alternadamente de v e de $-v$. Este problema é evitado se tomarmos apenas iterações pares de $G(p)$, ou, o que é equivalente, se iterarmos a aplicação F. □

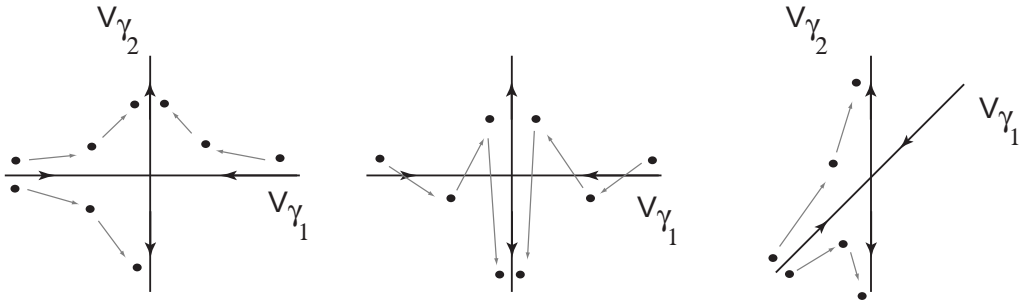
Em geral, se L é um automorfismo diagonalizável do plano com autovalores γ_1 e γ_2, então L é uma contração se e só se $|\gamma_1| < 1$ e $|\gamma_2| < 1$, e L é uma dilatação se e só se $|\gamma_1| > 1$ e $|\gamma_2| > 1$.



exemplos: à esquerda, contração com $-1 < \gamma_1 < 0 < \gamma_2 < 1$; á direita dilatação com $\gamma_1 > \gamma_2 > 1$.

Quando $|\gamma_1| < 1 < |\gamma_2|$, em um dos autoespaços L é uma contração e no outro uma dilatação. Na base dos autovetores de L um ponto $p = (p_1, p_2) \in \mathbf{R}^2$ tem órbita dada por $\mathcal{O}_L(p_1, p_2) = ((\gamma_1^n p_1, \gamma_2^n p_2))$. Se p não está

sobre os eixos, quando n cresce a primeira coordenada de $L^n(p)$ tende para zero e a segunda para infinito — a semi-órbita positiva de p aproxima-se do autoespaço associado a γ_2, o autoespaço em que L é uma dilatação e a forma das órbitas é a indicada na figura abaixo.



Um automorfismo linear L cujos autovalores têm sempre módulo diferente de 1 é chamado um *automorfismo hiperbólico*. Vimos que neste caso, se L é diagonalizável, então existe um único ponto fixo e não há pontos periódicos.

Um automorfismo diagonalizável que não é hiperbólico pode ter mais pontos fixos ou periódicos além da origem. Todo ponto no autoespaço associado ao autovalor 1 é um ponto fixo. Todos os pontos do autoespaço associado a -1 são periódicos de período 2, exceto a origem que é um ponto fixo.

2.3 representação da multiplicação complexa

Podemos identificar o plano $\mathbf{R}^2$ com o conjunto $\mathbf{C}$ dos números complexos através da aplicação:

$$\begin{array}{rcl} \mathcal{J} : \mathbf{R}^2 & \longrightarrow & \mathbf{C} \\ (x,y) & \mapsto & x+iy = \mathcal{J}(x,y). \end{array}$$

Pensando em $\mathbf{C}$ como um espaço vetorial de dimensão 2 sobre o o corpo $\mathbf{R}$, é fácil ver que $\mathcal{J}$ é um isomorfismo, com $\mathcal{J}^{-1}(z) = (Re(z), Im(z))$.

A multiplicação por um número complexo $\omega \neq 0$ define uma transformação linear $L_\omega : \mathbf{C} \longrightarrow \mathbf{C}$, $L_\omega(z) = \omega z$, com inversa $(L_\omega)^{-1} = L_{1/\omega}$. Usando a identificação acima, obtemos um automorfismo $M_\omega : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$

dado por $M_\omega(x,y) = \mathcal{J}^{-1} \circ L_\omega \circ \mathcal{J}$ isto é, para $\omega = a + ib$ com a e $b \in \mathbf{R}$:

$$\begin{array}{ccc} \mathbf{R}^2 & \xrightarrow{\mathcal{J}} & \mathbf{C} \\ (x,y) & & x+iy = \mathcal{J}(x,y) \\ M_{a+ib} \quad \downarrow & & \downarrow \quad L_{a+ib} \\ \mathbf{R}^2 & \xrightarrow{\mathcal{J}} & \mathbf{C} \\ (ax-by, bx+ay) & & ax-by+i(bx+ay) \end{array}$$

O automorfismo M_{a+ib} com a e $b \in \mathbf{R}$ é representado na base usual de $\mathbf{R}^2$ pela matriz:

$$M_{a+ib} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{2.3}$$

Escrevendo $a+ib$ em coordenadas polares, $a+ib = \rho(\cos\theta + i\,\mathrm{sen}\,\theta)$, a matriz (2.3) toma a forma:

$$M_{\rho(\cos\theta+i\,\mathrm{sen}\,\theta)} = \begin{pmatrix} \rho\cos\theta & -\rho\,\mathrm{sen}\,\theta \\ \rho\,\mathrm{sen}\,\theta & \rho\cos\theta \end{pmatrix} = \begin{pmatrix} \rho & 0 \\ 0 & \rho \end{pmatrix} \cdot \begin{pmatrix} \cos\theta & -\,\mathrm{sen}\,\theta \\ \mathrm{sen}\,\theta & \cos\theta \end{pmatrix} \tag{2.4}$$

Assim vemos que a multiplicação pelo número complexo $\rho(\cos\theta + i\,\mathrm{sen}\,\theta)$ representa no plano a rotação de um ângulo θ (multiplicação por $\cos\theta + i\,\mathrm{sen}\,\theta$), seguida da homotetia de razão ρ que é uma contração ou uma dilatação do plano, conforme ρ seja maior ou menor que 1. Esta última aplicação é um múltiplo escalar da identidade de $\mathbf{R}^2$ e por isso comuta com qualquer automorfismo de $\mathbf{R}^2$. Em particular, $M^2_{\rho(\cos\theta+i\,\mathrm{sen}\,\theta)}(x,y) = \rho^2 M_{(\cos 2\theta+i\,\mathrm{sen}\,2\theta)}(x,y)$, já que $M^2_{(\cos\theta+i\,\mathrm{sen}\,\theta)}(x,y) = M_{(\cos 2\theta+i\,\mathrm{sen}\,2\theta)}(x,y)$ (rodar duas vezes de um ângulo θ é o mesmo que rodar uma só vez de um ângulo 2θ) e por isso

$$M^n_{\rho(\cos\theta+i\,\mathrm{sen}\,\theta)}(x,y) = \rho^n M_{(\cos n\theta+i\,\mathrm{sen}\,n\theta)}(x,y)$$

para qualquer n inteiro. Assim, para qualquer ponto (x,y) de $\mathbf{R}^2$, temos:

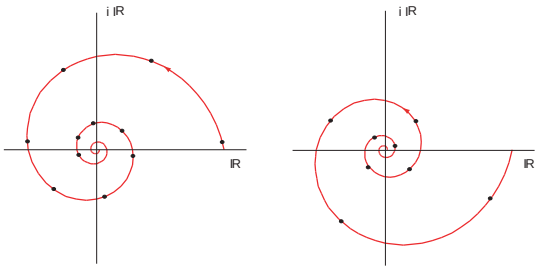
$$|M_{\cos\theta+i\,\mathrm{sen}\,\theta}(x,y)| = |(x,y)| \tag{2.5}$$

$$\left|M^n_{\rho(\cos\theta+i\,\mathrm{sen}\,\theta)}(x,y)\right| = \rho^n\,|(x,y)|\,. \tag{2.6}$$

Se $\rho > 1$ então $M_{\rho(\cos\theta+i\,\mathrm{sen}\,\theta)}$ é uma dilatação. Os pontos $M^n_{\rho(\cos\theta+i\,\mathrm{sen}\,\theta)}(x,y)$ afastam-se da origem exponencialmente quando $n \longrightarrow +\infty$ e tendem para 0 quando $n \longrightarrow -\infty$. O único ponto fixo é a origem e a órbita de qualquer outro ponto está contida numa espiral passando por este ponto.

Para $\rho < 1$ o automorfismo $M_{\rho(\cos\theta+i\,\mathrm{sen}\,\theta)}$ é uma contração. A origem ainda é o único ponto fixo, mas a semi-órbita positiva de $(x,y) \neq 0$ por

$M_{\rho(\cos\theta+i\,\mathrm{sen}\,\theta)}$ espirala em direção à origem e quando $n \longrightarrow -\infty$ os pontos $M^n_{\rho(\cos\theta+i\,\mathrm{sen}\,\theta)}(x,y)$ afastam-se da origem exponencialmente.



Se $\rho = 1$ então todos os pontos de uma mesma órbita terão normas iguais, por (2.6). Isto quer dizer que a órbita de $(x,y) \in \mathbf{R}^2$ está contida na circunferência de centro na origem que passa por (x,y). Cada uma dessas circunferências é rodada uniformemente pela aplicação $M_{\cos\theta+i\,\mathrm{sen}\,\theta}$, que define assim um automorfismo em cada uma daquelas circunferências. O estudo detalhado deste automorfismo será feito na seção 4.

Finalmente se $L : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ é um automorfismo cujos autovalores são $a \pm ib$ com a e b reais então sabemos que é possível obter uma base de $\mathbf{R}^2$ na qual L é representado pela matriz (2.3) e temos uma descrição completa das órbitas de L. Se escolhermos uma outra base para $\mathbf{R}^2$, as espirais serão deformadas mas o seu comportamento assintótico (a norma tender para 0 ou ∞ quando $n \longrightarrow \pm\infty$) se mantém. No caso de $\rho = |a+ib| = 1$, uma mudança linear de coordenadas pode transformar as circunferências em elipses. O comportamento das órbitas sobre cada elipse é o mesmo que sobre a circunferência original.

Capítulo 3

Sistemas dinâmicos lineares em $\mathbf{R}^n$

Nos dois exemplos acima quando a norma dos autovalores é menor que 1, obtemos um único ponto fixo que é ponto de acumulação das órbitas de todos os pontos do plano. O mesmo ocorre em $\mathbf{R}^n$:

Teorema 3.0.1 *Se todos os autovalores de um automorfismo linear* $L : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ *tiverem norma menor que 1, então* $\lim_{n\to\infty} L^n(p) = 0 \quad \forall p \in \mathbf{R}^n$.

Para demonstrar o Teorema 3.0.1 utilizaremos o seguinte resultado:

Lema 3.0.1 *Seja* $L : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ *um automorfismo linear. Se existirem* γ *e* $\delta \in \mathbf{R}^+$ *tais que todos as raízes* λ *do polinômio característico de* L *satizfazem* $\gamma < |\lambda| < \delta$, *então existe uma norma em* $\mathbf{R}^n$ *tal que:*

$$\forall p \neq 0 \in \mathbf{R}^n \qquad \gamma|p| < |Lp| < \delta|p|.$$

Demonstração do Lema: Basta demonstrar o lema para vetores p cuja norma seja 1. Além disso se o lema for verdadeiro em dois subespaços vetoriais L-invariantes V_1 e V_2 com $V_1 \cap V_2 = \{0\}$, então ele é verdadeiro em $V_1 \oplus V_2$, com a norma $|x+y|^2 = |x|^2_{V_1} + |y|^2_{V_2}$ para $x \in V_1$, $y \in V_2$. Basta então mostrar que o lema é vedadeiro em cada um dos autoespaços generalizados de L.

A base da forma canônica real de L pode ser modificada de maneira que cada bloco elementar tome uma das formas abaixo, com $\varepsilon > 0$ arbitrário:

$$\begin{pmatrix} \lambda_j & 0 & & \\ \varepsilon & \lambda_j & & \\ & & \ddots & \\ & & \varepsilon & \lambda_j \end{pmatrix} \qquad \begin{pmatrix} R_j & 0 & & \\ \varepsilon \mathrm{I} & R_j & & \\ & & \ddots & \\ & & \varepsilon \mathrm{I} & R_j \end{pmatrix} \qquad R_j = \begin{pmatrix} a_j & -b_j \\ b_j & a_j \end{pmatrix}.$$

No primeiro caso, tem-se para $|p| = 1$

$$\begin{aligned} |Lp|^2 &= \lambda_j^2|p|^2 + \varepsilon^2 \sum_{k\geq 2} p_k^2 + 2\lambda_j \varepsilon \sum_{k\geq 2} p_{k-1}p_k \\ &\leq (\lambda_j^2 + \varepsilon^2)|p|^2 + 2|\lambda_j||\varepsilon| \sum_{k\geq 2} |p_{k-1}p_k| \\ &= (\lambda_j^2 + \varepsilon^2) + 2|\lambda_j||\varepsilon| \sum_{k\geq 2} |p_{k-1}p_k|. \end{aligned}$$

Tomando ε suficientemente pequeno para que

$$\varepsilon^2 + 2|\lambda_j||\varepsilon| \max_{|p|=1} \sum_{k\geq 2} |p_{k-1}p_k| < \delta^2 - \lambda_j^2,$$

fica demonstrado que $|Lp|^2 < \delta^2|p|^2$. A demonstração dos outros casos é inteiramente análoga. □

Demonstração do Teorema 3.0.1: Como todos os autovalores de L têm norma menor que 1, podemos escolher $\delta \in \mathbf{R}$ com $0 < \delta < 1$ de maneira que δ seja maior que todos os autovalores. Usando o Lema 3.0.1 e a linearidade de L verifica-se que L é uma contração. Usando o Teorema da contração (Teorema 1.2.1) obtemos o resultado. □

Como $\lambda \in \mathbf{C}$ é autovalor de L se e só se $1/\lambda$ é autovalor de L^{-1} segue-se diretamente do Teorema 3.0.1:

Corolário 3.0.1 *Se todos os autovalores de um automorfismo linear* $L : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ *tiverem norma maior que 1, então* $\lim_{n\to-\infty} L^n(p) = 0 \quad \forall p \in \mathbf{R}^n$.

Um automorfismo linear L cujos autovalores têm sempre módulo diferente de 1 é chamado um *automorfismo hiperbólico*.

Teorema 3.0.2 *Se o automorfismo* $L : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ *for hiperbólico, então existem dois subespaços* L *invariantes de* $\mathbf{R}^n$, W_c *e* W_d *com* $\mathbf{R}^n = W_c \oplus W_d$, *e tais que* $L\big|_{W_c}$ *é uma contração e* $L\big|_{W_d}$ *é uma dilatação, isto é,* $\left(L\big|_{W_d}\right)^{-1}$ *é uma contração, para alguma norma em* $\mathbf{R}^n$.

Demonstração: Como L é hiperbólico, tomamos W_c como a soma dos autoespaços generalizados associados aos autovalores de L cujo módulo é menor que 1 e W_d como a soma dos autoespaços generalizados associados

aos autovalores de L cujo módulo é maior que 1. O resultado segue-se aplicando o Lema 3.0.1 a $L\big|_{W_c}$, $L\big|_{W_d}$ e a $L^{-1}\big|_{W_d}$. □

Os espaços L-invariantes W_c e W_d são chamados *espaço de contração e espaço de dilatação de* L, respectivamente.

As órbitas de um automorfismo de $\mathbf{R}^n$ podem sempre ser descritas comparando-o a uma aplicação tempo 1, como nos exemplos 1.4 e indiretamente em 2.1 e 2.3.

Teorema 3.0.3 *Dado um automorfismo linear* $L : \mathbf{R}^n \longrightarrow \mathbf{R}^n$*, existem sempre um automorfismo* $S : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ *e um operador linear* $T : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ *com as propriedades:*

1. $L = S\mathrm{e}^T$.
2. $L \cdot S = S \cdot L$ *e* $S \cdot T = T \cdot S$.
3. $S^2 = \mathrm{I}$.
4. *Se todos os autovalores reais de* L *forem positivos então* $S = \mathrm{I}$.

Demonstração: Para encontrar um operador linear cuja exponencial é o automorfismo dado, trabalhamos em cada um dos autoespaços generalizados e observamos que a exponencial de um bloco de Jordan é uma matriz triangular com cujas as entradas abaixo da diagonal são idênticas, como por exemplo:

$$J = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix} \qquad \mathrm{e}^J = \begin{pmatrix} \mathrm{e}^\lambda & 0 & 0 \\ \mathrm{e}^\lambda & \mathrm{e}^\lambda & 0 \\ \mathrm{e}^\lambda & \mathrm{e}^\lambda & \mathrm{e}^\lambda \end{pmatrix}.$$

Alterando convenientemente a base de Jordan podemos escrever qualquer automorfismo L como e^T sempre que os autovalores reais de L são positivos. Se o automorfismo tiver autovalores negativos, o bloco de Jordan, multiplicado por -1 será a exponencial de uma matriz. □

Capítulo 4

Sistemas dinâmicos no círculo S^1

O automorfismo linear do plano que representa a multiplicação por um número complexo com norma 1 corresponde a uma rotação do plano, como vimos em 2.3. Para estudar este exemplo em mais detalhe, bem como os exemplos em dimensão mais alta, precisamos conhecer melhor a geometria do círculo S^1.

4.1 recobrimento de S^1

Usaremos as seguintes descrições alternativas (e equivalentes) da circunferência S^1:

1. $S^1 \subset \mathbf{R}^2$, $S^1 = \{x \in \mathbf{R}^2 : |x| = 1\}$

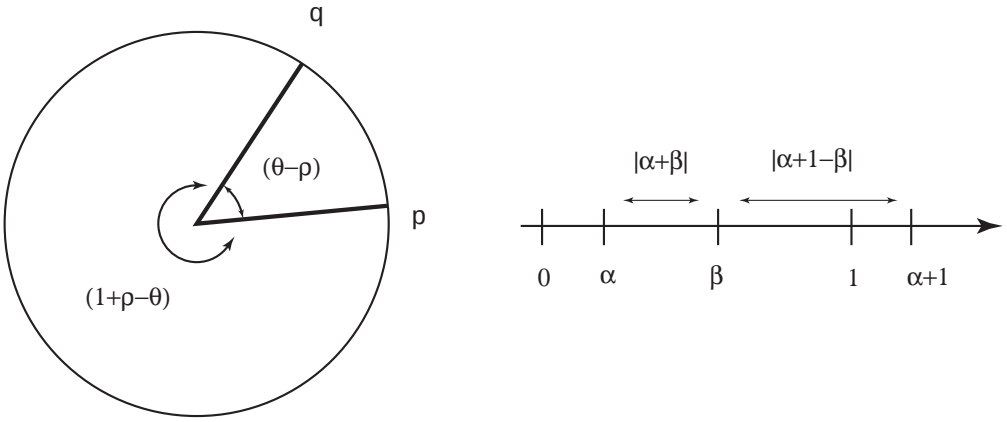
2. Carta (ou parametrização) de S^1 no intervalo $[0, 1[$, dada por: $c : [0, 1[\longrightarrow \mathbf{R}^2$, $c(\rho) = (\cos 2\pi\rho, \operatorname{sen} 2\pi\rho)$, isto é, $c(\rho)$ é a intersecção de S^1 com a semi-reta pela origem que faz um ângulo $2\pi\rho$ com o semi-eixo positivo dos xx. A aplicação c é injetiva, diferenciável e sua imagem é a circunferência S^1.

3. Recobrimento de S^1 por $\mathbf{R}$ dado pela aplicação $\mathcal{C} : \mathbf{R} \longrightarrow \mathbf{R}^2$, $\mathcal{C}(\rho) = (\cos 2\pi\rho, \operatorname{sen} 2\pi\rho)$, que é periódica de período 1, diferenciável e cuja imagem é a circunferência S^1. A restrição de $\mathcal{C}$ a qualquer intervalo de comprimento menor ou igual a 1 é injetiva.

4. O recobrimento $\mathcal{C}$ pode ser definido como a composta da carta c com a aplicação m que a cada número associa a sua mantissa: $m : \mathbf{R} :\longrightarrow$

$[0,1[$, $m(\rho) = \rho - [\rho]$, onde $[\rho]$ é o maior inteiro menor ou igual a ρ. Temos que $\forall \rho \in \mathbf{R}$, $\mathcal{C}(m(\rho)) = \mathcal{C}(\rho)$.

5. A circunferência pode também ser identificada com o espaço quociente $\mathbf{R}/\mathbf{Z}$ usando a relação de equivalência definida por $\rho \sim \theta \iff \rho - \theta \in \mathbf{Z}$. Isto faz sentido porque $\rho \sim \theta \iff m(\rho) = m(\theta) \iff \mathcal{C}(\rho) = \mathcal{C}(\theta)$.

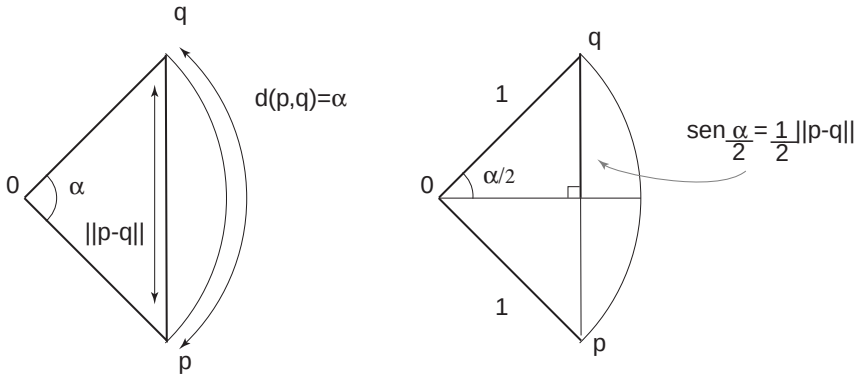


Usando a carta de S^1 definida em 2, podemos definir uma nova noção de distância $d(p,q)$ entre dois pontos p e q de S^1 como menor o ângulo (medido em radianos) entre as semi-retas pela origem por p e q. Se $p = \mathcal{C}(\rho)$ e $q = \mathcal{C}(\theta)$ então

$$d(p,q) = 2\pi \min\{m(|\rho - \theta|), m(|\rho + 1 - \theta|)\}$$

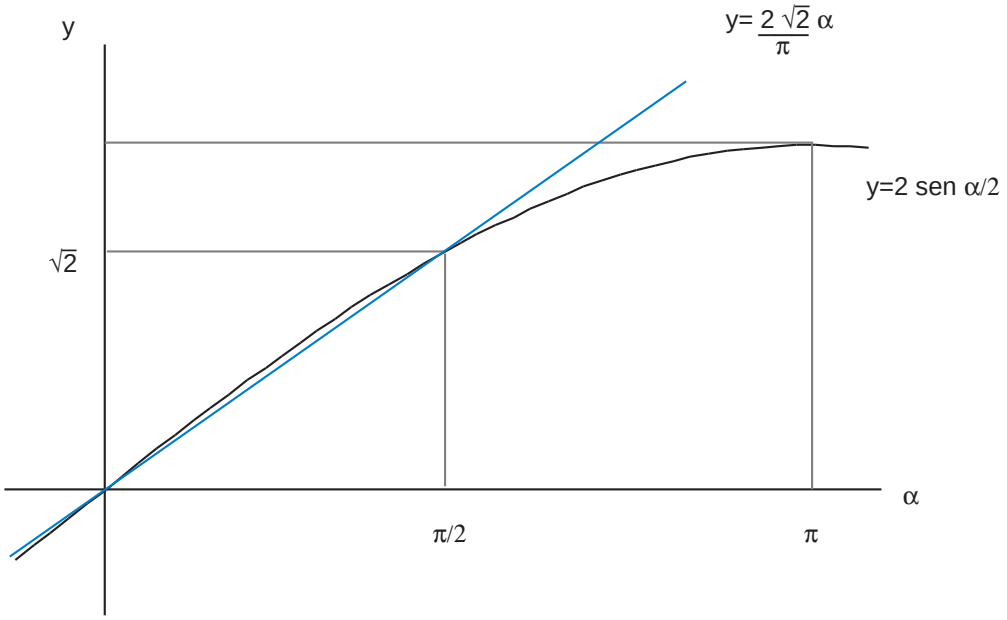
isto é,

$$d(p,q) = 2\pi \min_{\mathcal{C}(\alpha)=p, \mathcal{C}(\beta)=q} \{m(|\alpha - \beta|)\}.$$



A nova noção corresponde a medir a distância ao longo de S^1, já que $d(p,q)$ é exatamente o comprimento do menor arco de circunferência entre p e q. Por isto é sempre verdade que $||p - q|| \leq d(p,q)$ onde $||.||$ é a norma

usual em $\mathbf{R}^2$ — o comprimento da secante é sempre menor que o do arco (ver figura). Se $d(p,q) = \alpha$, então $||p-q|| = 2\,\mathrm{sen}\,\alpha/2 \leq \alpha$. Por outro lado, é fácil ver (mas dá trabalho) que $d(p,q) \leq \frac{\pi}{2\sqrt{2}}||p-q||$ já que $||p-q|| = 2\,\mathrm{sen}\,\alpha/2 \geq \frac{2\sqrt{2}}{\pi}\alpha$ para todo $\alpha \in [0, \pi/2]$ (ver figura).



A relação $||p-q|| \leq d(p,q) \leq \frac{\pi}{2\sqrt{2}}||p-q||$ implica que a nova distância d e a norma euclideana $||.||$ definem a mesma noção de proximidade em S^1: uma sucessão p_n de pontos em S^1 converge para um ponto $p_0 \in S^1$ se e só se $\lim_{n\to\infty} d(p_n, p_0) = 0$. Em outras palavras, d e $||.||$ definem a mesma topologia em S^1.

4.2 sistema dinâmico em S^1: rotação

Podemos definir um sistema dinâmico em S^1 através do recobrimento $\mathcal{C}$ — se $F : \mathbf{R} \longrightarrow \mathbf{R}$ é uma aplicação tal que $\forall \rho \in \mathbf{R} \quad \forall M \in \mathbf{Z} \quad \exists N \in \mathbf{Z}$: $F(\rho + M) = F(\rho) + N$, então $m(F(\rho))$ só depende de $m(\rho)$ e por isso $\mathcal{C} \circ F \circ c^{-1}$ define uma relação de recorrência em S^1. Neste caso temos que $(m(F(\rho)))^j = m(F^j(\rho))$ para todos os valores de $j \in \mathbf{Z}$ e de $\rho \in \mathbf{R}$ e por isso $(\mathcal{C}(F(\rho)))^j = \mathcal{C}(F^j(\rho))$.

Como exemplo estudaremos a aplicação induzida em S^1 pela translação $F_\alpha(\rho) = \alpha + \rho$ com $\alpha \in \mathbf{R}$. Esta recorrência $L_\alpha : S^1 \longrightarrow S^1$ é a rotação de um ângulo $2\alpha\pi$ e $F^n_\alpha(\rho) = \rho + n\alpha$. Consideraremos separadamente os casos em que α é racional e irracional.

Teorema 4.2.1 *A aplicação $L_\alpha : S^1 \longrightarrow S^1$ dada pela rotação de um ângulo $2\alpha\pi$ tem órbitas periódicas de período q se e só se $\alpha = p/q$ com p e $q \in \mathbf{Z}$ e* mdc $(p, q) = 1$. *Neste caso todas as órbitas são periódicas de período q.*

Demonstração: Como $\alpha = p/q$ com p e $q \in \mathbf{Z}$, $q \neq 0$ então $F^q_{p/q}(\rho) = \rho + p \sim \rho$ para qualquer ponto ρ de $\mathbf{R}$. Isto quer dizer que $L^q_{p/q}$ é a identidade, isto é, que todos os pontos de S^1 são periódicos, com período q.

Reciprocamente, suponhamos que existe algum $\rho \in [0, 1[$ tal que $F^q_\alpha(\rho) \sim \rho$, o que acontece se e só se $q\alpha + \rho - \rho = q\alpha \in \mathbf{Z}$. Tomando $p = q\alpha \in \mathbf{Z}$ podemos escrever $\alpha = p/q \in \mathbf{Q}$. □

Teorema 4.2.2 *Toda órbita da aplicação $L_\alpha : S^1 \longrightarrow S^1$ dada pela rotação de um ângulo $2\alpha\pi$ com $\alpha \notin \mathbf{Q}$ é densa em S^1.*

Demonstração: Se $\alpha \notin \mathbf{Q}$ então, pelo teorema 4.2.1, L_α não tem órbitas periódicas de nenhum período: quaisquer que sejam $\rho \in [0, 1[$ e $q \in \mathbf{Z}$, $q \neq 0$ temos $m(F^q_\alpha(\rho)) \neq \rho$. Queremos mostrar que dado $\rho \in [0, 1[$ todo ponto em S^1 pode ser aproximado por pontos da órbita de $c(\rho)$ por Ł$_\alpha$. Para isto usaremos a distância d que definimos na secção 4.1.

Se fixarmos $\rho \in [0, 1[$ e $q \in \mathbf{Z}$, então os $q + 1$ pontos no intervalo $[0, 1[$:

$$\rho, m(F_\alpha(\rho)), m(F^2_\alpha(\rho)), \ldots, m(F^q_\alpha(\rho)) \tag{4.1}$$

são todos distintos. Estes pontos não aparecem necessariamente dispostos pela ordem de iteração de F.

Como os pontos (4.1) são todos distintos, então pelo menos dois dentre eles estão a uma distância menor que $1/q$ — caso contrário a soma de suas distâncias seria maior que o comprimento do intervalo $[0, 1[$. Denotamos por $m(F^j_\alpha(\rho))$ e $m(F^l_\alpha(\rho))$ dois destes pontos tais que

$$|m(F^j_\alpha(\rho)) - m(F^l_\alpha(\rho))| < 1/q. \tag{4.2}$$

Tomando $\theta = m(F^l_\alpha(\rho))$ e $s = j - l$, a expressão (4.2) passa a ser

$$|m(F^s_\alpha(\theta)) - m(\theta)| = |m(\theta + s\alpha) - m(\theta)| < \frac{1}{q}.$$

Isto quer dizer que dois termos sucessivos da sequência

$$m(\theta), m(F^s_\alpha(\theta)), m(F^{2s}_\alpha(\theta)), m(F^{3s}_\alpha(\theta)), \ldots \tag{4.3}$$

estão sempre a uma distância menor que $1/q$ entre sí. É claro que as imagens por $\mathcal{C}$ dos pontos da sequência (4.3) pertencem todas à órbita $\mathcal{O}_{L_\alpha}(\mathcal{C}(\rho))$.

Assim um ponto qualquer de S^1 está sempre entre dois pontos sucessivos da imagem por $\mathcal{C}$ da sequência (4.3) e por isso a uma distância menor que $1/q$ de algum ponto de $\mathcal{O}_{L_\alpha}(\mathcal{C}(\rho))$. □

Capítulo 5

Sistemas dinâmicos no toro $\mathbf{T}^2$

Um automorfismo linear não hoperbólico com dois pares de autovalores complexos com norma 1 define uma transformação no toro $\mathbf{T}^2 = S^1 \times S^1$. Estudaremos o toro em mais detalhe, começando pela sua geometria, e passando a sistemas dinâmicos no toro definidos por automorfismos lineares.

5.1 geometria do toro $\mathbf{T}^2$

Da mesma maneira que fizemos para S^1, usaremos várias descrições equivalentes do toro $\mathbf{T}^2$ para estudar relações de recorrência:

1. $\mathbf{T}^2 \subset \mathbf{R}^4$,

$$\begin{aligned} \mathbf{T}^2 &= \{(x_1, x_2, x_3, x_4) \in \mathbf{R}^4 : |(x_1, x_2)| = |(x_3, x_4)| = 1\} \\ &= \{(x_1, x_2, x_3, x_4) \in \mathbf{R}^4 : x_1^2 + x_2^2 = x_3^2 + x_4^2 = 1\}. \end{aligned}$$

2. Carta (ou parametrização) de $\mathbf{T}^2$ no quadrado $[0, 1[^2$, dada por:

$$\begin{aligned} &\gamma : [0, 1[^2 \longrightarrow \mathbf{R}^4 \\ &\gamma(\rho, \theta) = (\cos 2\pi\rho, \operatorname{sen} 2\pi\rho, \cos 2\pi\theta, \operatorname{sen} 2\pi\theta). \end{aligned}$$

   Esta aplicação é análoga à carta c de S^1, definida em 2) na seção 4.1. A carta γ é injetiva, diferenciável e sua imagem é o toro $\mathbf{T}^2$ definido em 1) acima.

3. Recobrimento de $\mathbf{T}^2$ por $\mathbf{R}^2$ dado pela aplicação

$$\begin{aligned} &\mathcal{P} : \mathbf{R}^2 \longrightarrow \mathbf{R}^4 \\ &\mathcal{P}(\rho, \theta) = (\cos 2\pi\rho, \operatorname{sen} 2\pi\rho, \cos 2\pi\theta, \operatorname{sen} 2\pi\theta) \end{aligned}$$

que é periódica de período 1, diferenciável e cuja imagem é o toro $\mathbf{T}^2$. A restrição de $\mathcal{P}$ a qualquer retângulo cujos lados tenham comprimento menor ou igual a 1 é injetiva.

4. O recobrimento $\mathcal{P}$ pode ser definido como a composta da carta γ com a aplicação:
$$\Pi : \mathbf{R}^2 :\longrightarrow [0,1[^2$$
$$\Pi(\rho,\theta) = (m(\rho), m(\theta))$$
onde m é a aplicação que a cada número associa a sua mantissa, definida na seção 4.1. Temos que $\forall(\rho,\theta) \in \mathbf{R}^2$, $\mathcal{P}(\Pi(\rho,\theta)) = \mathcal{P}(\rho,\theta)$.

5. Analogamente ao que foi feito em 5) na seção 4.1, definimos em $\mathbf{R}^2$ uma relação de equivalência módulo $\mathbf{Z}^2$: dois pontos $q = (\rho,\theta)$ e $\tilde{q} = (\tilde{\rho},\tilde{\theta})$ são equivalentes (o que denotamos por $q \approx \tilde{q}$) se e só se $q - \tilde{q} \in \mathbf{Z}^2$, isto é, se e só se $\rho \sim \tilde{\rho}$ e $\theta \sim \tilde{\theta}$, onde $\sim$ é a relação de equivalência que definimos na seção 4.1.

   Do mesmo modo que em S^1, temos que $q \approx \tilde{q} \iff \Pi(q) = \Pi(\tilde{q}) \iff \mathcal{P}(q) = \mathcal{P}(\tilde{q})$. Assim, o toro pode também ser identificado com o espaço quociente $\mathbf{R}^2/\mathbf{Z}^2$.

6. De 1) é fácil ver que $\mathbf{T}^2 = S^1 \times S^1 \subset \mathbf{R}^4$.

7. O toro pode também ser representado em $\mathbf{R}^3$ como a superfície de revolução gerada pela rotação em torno do eixo dos zz de uma circunferência no semi-plano (x,z), $x > 0$.

Podemos definir uma nova noção de distância $\delta(q,\tilde{q})$ entre dois pontos q e $\tilde{q}$ de $\mathbf{T}^2$ usando o recobrimento γ de $\mathbf{T}^2$ e uma norma, $||.||$, em $\mathbf{R}^2$, como fizemos na seção 4.1 para a distância d em S^1. A nova distância será:

$$\delta(q,\tilde{q}) = 2\pi \cdot \min_{\mathcal{P}(p)=q, \mathcal{P}(\tilde{p})=\tilde{q}} \{||p - \tilde{p}||\}\,.$$

Se $\mathcal{P}(p_0) = q$ e $\mathcal{P}(\tilde{p}_0) = \tilde{q}$, então

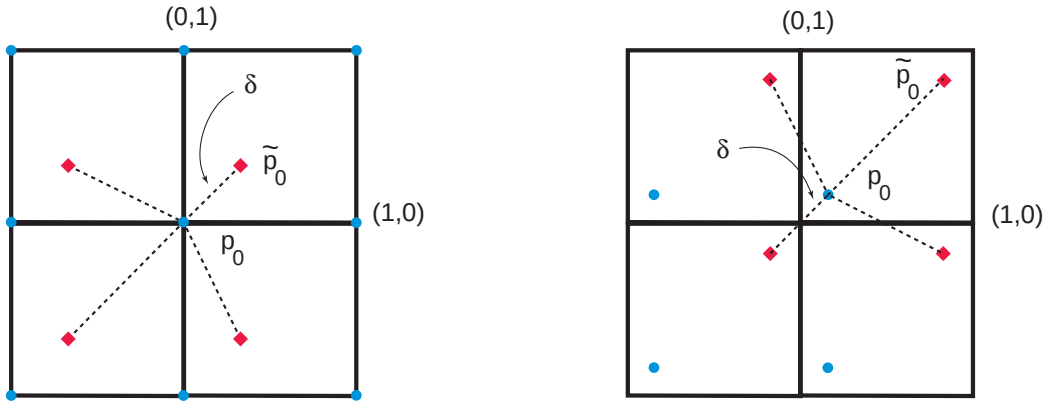
$$\delta(q,\tilde{q}) = 2\pi \cdot \min_{M \in \mathbf{Z}^2} \{||\Pi(p_0 - \tilde{p}_0 + M)||\}\,.$$

A nova distância δ corresponde a medir o comprimento da menor curva sobre o toro que liga q a $\tilde{q}$. Como no caso de S^1, a distância δ define a mesma noção de convergência em $\mathbf{T}^2$ que a restrição a $\mathbf{T}^2$ da norma euclideana de $\mathbf{R}^4$.

5.2 exemplos usando a norma euclideana em $\mathbf{R}^2$

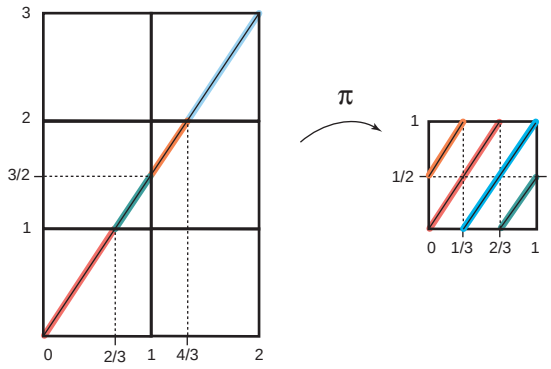
1. $p_0 = (0,0)$; $\tilde{p}_0 = (\frac{1}{3}, \frac{1}{3})$; (ver figura) $\delta(\mathcal{P}(p_0)), \mathcal{P}(\tilde{p}_0)) = \left\|\left(\frac{1}{3}, \frac{1}{3}\right)\right\| = \frac{\sqrt{2}}{3}$.

2. $p_0 = (\frac{1}{6}, \frac{1}{6})$; $\tilde{p}_0 = (\frac{5}{6}, \frac{5}{6})$; (ver figura)
   $\delta(\mathcal{P}(p_0)), \mathcal{P}(\tilde{p}_0)) = \|p_0 - (\tilde{p}_0 - (1,1))\| = \left\|(\frac{-2}{3}, \frac{-2}{3}) + (1,1)\right\| = \frac{\sqrt{2}}{3}$.



Nas figuras os pontos equivalentes a p_0 estão representados por um disco azul e os equivalentes a $\tilde{p_0}$ por um quadrado vermelho.

As retas horizontais e verticais do plano $\rho \times \theta$ dadas por $\theta =$ constante e $\rho =$ constante são transformadas por $\mathcal{P}$ em circunferências sobre o toro que chamaremos paralelos e meridianos respectivamente. Podemos estudar as imagens por $\mathcal{P}$ das demais retas do plano, como a reta de inclinação 3/2 do desenho.



Teorema 5.2.1 *A imagem por $\mathcal{P}$ de uma reta de inclinação racional é uma curva fechada.*

Demonstração: Seja $y = (p/q)x$ a equação da reta, com p e $q \in \mathbf{Z}$ e $q > 0$. Como a reta passa nos pontos $(0,0)$ e (q,p) e como (q,p) pertence a $\mathbf{Z}^2$, então $\Pi(q,p) = (0,0)$ e $\mathcal{P}$ transforma o segmento de reta $y = (p/q)x$ com $x \in [0,q]$ numa curva fechada em $\mathbf{T}^2$. Os segmentos de reta $y = (p/q)x$ com $x \in [kq,(k+1)q]$, para $k \in \mathbf{Z}$ são transformados na mesma curva fechada, já que $(x,(p/q)x) \approx (m(x),(p/q)m(x)) \approx (x - kq,(p/q)(x - kq))$. Logo, $\mathcal{P}$ transforma a reta $y = (p/q)x$ numa curva fechada em $\mathbf{T}^2$. A demostração para o caso de uma reta que não passa pela origem é inteiramente análoga.□

Teorema 5.2.2 *A imagem por $\mathcal{P}$ de uma reta de inclinação irracional é uma curva densa em $\mathbf{T}^2$.*

Demonstração: Seja $y = \alpha x$, $\alpha \notin \mathbf{Q}$ a equação da reta (o caso $y = \alpha x + \beta$ com $\beta \neq 0$ é análogo). A reta $y = \alpha x$ intersecta o feixe de retas verticais $x = \rho + N$ com $N \in \mathbf{Z}$ nos pontos $(\rho+N, \alpha\rho+\alpha N) \approx (m(\rho), m(\alpha\rho))$. Pelo Teorema 4.2.2 a imagem destes pontos por $\mathcal{P}$ é densa no meridiano. Como a imagem da reta intersecta cada meridiano de $\mathbf{T}^2$ num conjunto denso e como todo ponto do toro pertence a algum meridiano, todo ponto de $\mathbf{T}^2$ pode ser aproximado por pontos da forma $\mathcal{P}(x,\alpha x)$, $x \in \mathbf{R}$ sobre o meridiano e portanto a imagem da reta é densa no toro. □

5.3 automorfismos de Anosov em $\mathbf{T}^2$

Do mesmo modo que fizemos para S^1, podemos definir recorrências em $\mathbf{T}^2$ através do recobrimento $\mathcal{P}$ — se $F : \mathbf{R}^2 \longrightarrow \mathbf{R}^2$ é uma aplicação que sempre transforma p e $\tilde{p} \in \mathbf{R}^2$ com $p \approx \tilde{p}$ em pontos equivalentes $F(p) \approx F(\tilde{p})$, então $\Pi(F(p))$ só depende do valor de $\Pi(p)$. Neste caso, $\mathcal{P} \circ F \circ \gamma^{-1}$ define uma relação de recorrência no toro.

No caso de F ser uma transformação linear, exigir que $F(p) \approx F(\tilde{p})$ sempre que $p \approx \tilde{p}$ é o mesmo que pedir que se $p - \tilde{p} \in \mathbf{Z}^2$ então $F(p) - F(\tilde{p}) = F(p - \tilde{p}) \in \mathbf{Z}^2$, isto é, que F transforme $\mathbf{Z}^2 \subset \mathbf{R}^2$ em si mesmo. Isto acontece se e só se a matriz de F na base usual de $\mathbf{R}^2$ tiver entradas inteiras. Estudaremos nesta seção um caso especial de relação de recorrência induzida em $\mathbf{T}^2$ por uma transformação linear do plano:

Definição 5.3.1 *Seja A uma matriz 2×2 tal que:*

1. *Todas as entradas de A são inteiras*

2. $\det A = \pm 1$

3. *A define um automorfismo hiperbólico de* $\mathbf{R}^2$*, i.e., A não tem autovalores com módulo igual a 1*

A recorrência $L_A : \mathbf{T}^2 \longrightarrow \mathbf{T}^2$ *induzida por A é chamada um* automorfismo de Anosov no toro.

Proposição 5.3.1 *Se a matriz A satisfaz as condições 1–3 da definição 5.3.1, então:*

1. *A imagem de um paralelogramo por A é um paralelogramo de mesma área.*

2. *Os autovalores de A são reais e distintos, um deles com módulo maior que 1 e outro com módulo menor que 1.*

3. *A tem inversa que satisfaz as condições 1–3 da definição 5.3.1.*

Demonstração:

1) A multiplicação por uma matriz transforma retas paralelas em retas paralelas, logo a imagem de um paralelogramo será um paralelogramo. A área de um paralelogramo com vértices em $(0,0)$, $\mathbf{u} = (u_1, u_2)$ $\mathbf{v} = (v_1, v_2)$ e $\mathbf{u} + \mathbf{v} \in \mathbf{R}^2$ é dada por $|\det(B)|$, onde B é a matriz

$$B = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} = (\mathbf{u}, \mathbf{v}) \ .$$

O paralelogramo de lados $A\mathbf{u}$ e $A\mathbf{v}$ terá então área dada por $|\det(C)|$ onde $C = (A\mathbf{u}, A\mathbf{v}) = AB$ e a afirmação segue-se da propriedade 2).

2) Suponhamos por absurdo que A tem um par de autovalores complexos $a \pm ib$ com $b \neq 0$, isto é, que A é semelhante à matriz

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

cujo determinante é $a^2 + b^2 = |a \pm ib|^2$. Como matrizes semelhantes têm sempre o mesmo determinante, pela hipótese 2 concluimos que $|a \pm ib|^2 = 1$ o que contradiz a hipótese 3. Portanto os autovalores de A são sempre reais.

Se A tivesse um autovalor duplo λ então seria semelhante a uma das matrizes:

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \qquad \text{ou} \qquad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Em qualquer dos casos $\det(A) = \lambda^2 = 1$, o que novamente contradiz a hipótese de hiperbolicidade.

Concluimos então que A tem dois autovalores reais distintos, λ_1 e λ_2, com $|\lambda_j| \neq 1$, para $j = 1, 2$. Assim A é semelhante à matriz $\mathrm{diag}(\lambda_1, \lambda_2)$, e por isso $\det(A) = \lambda_1\lambda_2 = \pm 1$, isto é, $\lambda_1 = \pm 1/\lambda_2$.

3) A tem inversa porque $\det A \neq 0$, e usando a regra de Cramer e o fato de que $\det A = \pm 1$ segue-se imediatamente que as entradas de A^{-1} são inteiras. A hiperbolicidade de A^{-1} é garantida por os autovalores de A^{-1} serem o inverso multiplicativo dos de A. □

O resultado a seguir mostra que a propriedade 2) é essencial para garantir a existência da inversa de L_A:

Proposição 5.3.2 *Se A é uma matriz com entradas inteiras e $\det(A) = \pm 1$ então A induz uma aplicação $L_A : \mathbf{T}^2 \longrightarrow \mathbf{T}^2$ que é inversível.*

Demonstração: Para mostrar que L_A é injetiva basta verificar que se $Ap \approx (0,0)$ com $p \in [0,1[^2$ então $p = (0,0)$. Suponhamos então que $p = (p_1, p_2) \in [0,1[^2$ e que $Ap = (M_1, M_2) \in \mathbf{Z}^2$. O paralelogramo com vértices $(0,0)$, p, $(1,0)$ e $p + (1,0)$ tem área $|p_2| \in [0,1[$. Por outro lado, se $A = (a_{ij})$ então a imagem do paralelogramo tem área $|M_1 a_{21} - M_2 a_{11}| \in \mathbf{Z}$. Logo $|M_1 a_{21} - M_2 a_{11}| = |p_2|$ e portanto p_2 será inteiro, o que implica que $p_2 = 0$. Para mostrar que $p_1 = 0$ consideramos o paralelogramo com vértices $(0,0)$, p, $(0,1)$ e $p + (0,1)$ e usamos o mesmo argumento.

Como consequência, a restrição de Π à imagem $A[0,1[^2$ é inversível, isto é, não há pontos equivalentes entre sí em $A[0,1[^2$. Por outro lado, $A[0,1[^2$ é um paralelogramo com área 1 e portanto $\Pi(A[0,1[^2) \subset [0,1[^2$ tem área 1. Logo $\Pi(A[0,1[^2) = [0,1[^2$ e por isso L_A é sobrejetiva. □

Denotamos os autovalores de A por λ_d e λ_c, onde $|\lambda_d| > 1$ e $|\lambda_c| < 1$ e os autoespaços de A por W_c e W_d.

Teorema 5.3.1 *Se a matriz A induz um automorfismo de Anosov em $\mathbf{T}^2$ então as imagens $\mathcal{P}(W_c)$ e $\mathcal{P}(W_d)$ são densas em $\mathbf{T}^2$.*

Demonstração: Os autoespaços W_c e W_d são retas A-invariantes. Se $(M, N) \in \mathbf{Z}^2$ for um ponto de coordenadas inteiras numa destas retas com $(M, N) \neq (0,0)$, então para qualquer $n \in \mathbf{Z}$ os pontos $A^n(M, N)$ estarão sobre a mesma reta e terão coordenadas inteiras. Como isto contradiz o fato de $\lim_{n\to\infty} A^n(M, N) = 0$ em W_c e $\lim_{n\to-\infty} A^n(M, N) = 0$ em W_d, conclui-se

que estas retas nunca passam por pontos de coordenadas inteiras que não a origem, e por isso que são retas de inclinação irracional. □

5.4 exemplo de automorfismo de Anosov

A matriz: $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ satisfaz as condições 1–3 da definição 5.3.1:

1. Todas as entradas de A são inteiras.
2. $\det A = 1$.
3. Os autovalores de A são:

$$-1 < \lambda_c = \frac{-3+\sqrt{5}}{2} < 0 \qquad \lambda_d = \frac{-3-\sqrt{5}}{2} < -1.$$

Por isso A define um automorfismo hiperbólico de $\mathbf{R}^2$, i.e., A não tem autovalores com módulo igual a 1.

Os autoespaços de A são:

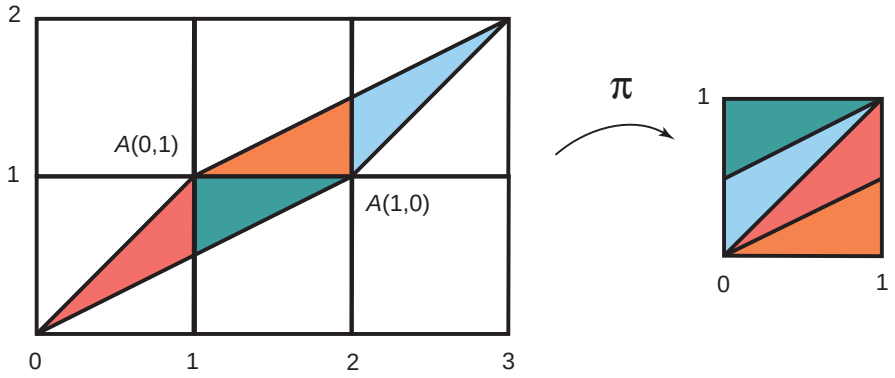
$$W_c = \left\{(x,y) : y = \frac{-5+\sqrt{5}}{2}\right\} \qquad W_d = \left\{(x,y) : y = \frac{-5-\sqrt{5}}{2}\right\}.$$

O único ponto fixo de L_A é a imagem, $C(0,0)$, da origem de $\mathbf{R}^2$, já que $A(x,y)^T \sim (x,y)^T$ se e só se existirem M e $N \in \mathbf{Z}$ tais que:

$$\left\{\begin{array}{rcl} 2x+y &=& x+M \\ x+y &=& y+N \end{array}\right. \iff \left\{\begin{array}{rcll} y &=& M-N & \in \mathbf{Z} \\ x &=& N & \in \mathbf{Z} \end{array}\right.$$

ou seja, se e só se $(x,y) \sim (0,0)$.

A imagem de $[0,1[^2$ por A e a sua projeção $\Pi\left(A\left([0,1[^2\right)\right)$ estão representadas na figura abaixo.



Em geral se L_A é um automorfismo de Anosov no toro, a imagem da origem por $\mathcal{P}$ é sempre um ponto fixo. Para encontrar pontos periódicos consideramos os pontos de $\mathbf{R}^2$ da forma $(\frac{\alpha}{\gamma}, \frac{\beta}{\gamma})$, com α, β e $\gamma \in \mathbf{Z}$, $\gamma \neq 0$. Como as entradas de A são inteiras, o ponto $A(\frac{\alpha}{\gamma}, \frac{\beta}{\gamma})$ também tem esta forma, isto é, existem α' e $\beta' \in \mathbf{Z}$, tais que $A(\frac{\alpha}{\gamma}, \frac{\beta}{\gamma}) = (\frac{\alpha'}{\gamma}, \frac{\beta'}{\gamma})$. Como $\Pi(\frac{\alpha'}{\gamma}, \frac{\beta'}{\gamma}) = (\frac{\alpha''}{\gamma}, \frac{\beta''}{\gamma})$, com $0 \leq \alpha'', \beta'' < \gamma$, então para cada valor de γ fixo, há exatamente γ^2 pontos da forma $\Pi(\frac{\alpha'}{\gamma}, \frac{\beta'}{\gamma})$ em $[0,1[^2$. Isto quer dizer que se $q = \mathcal{P}(\frac{\alpha}{\gamma}, \frac{\beta}{\gamma})$, então no conjunto $\left\{q, L_A q, L_A^2 q, \cdots, L_A^{\gamma^2} q\right\}$ há pelo menos uma repetição e portanto que q é um ponto periódico, com período menor ou igual a γ^2.

Teorema 5.4.1 *O conjunto dos pontos do toro cuja órbita por um automorfismo de Anosov é periódica é denso no toro.*

Demonstração: Basta observar que o conjunto

$$\left\{\left(\frac{\alpha}{\gamma}, \frac{\beta}{\gamma}\right), \alpha, \beta, \gamma \in \mathbf{Z}, \gamma \neq 0\right\} \cap [0,1[^2$$

é denso em $[0,1[^2$ e que portanto sua imagem pela carta do toro é densa no toro. □

Definição 5.4.1 *Se $q \in \mathbf{T}^2$ é um ponto periódico do automorfismo de Anosov L_A e se $p \in \mathbf{R}^2$ pertence à imagem inversa de q, isto é, se $\mathcal{P}(p) = q$, então definimos a curva de contração de q, $C_c(q)$, como a imagem por $\mathcal{P}$ da reta por p paralela a W_c. Analogamente definimos a curva de dilatação de q , $C_d(q)$, como a imagem por $\mathcal{P}$ da reta por p paralela a W_d.*

É fácil ver que a definição acima não depende da escolha do ponto p, isto é, que se $p \approx \tilde{p}$, então $\mathcal{P}$ transforma na mesma curva as duas retas paralelas a W_c por p e $\tilde{p}$. Além disso temos:

Teorema 5.4.2 *Se $q \in \mathbf{T}^2$ é um ponto periódico com período γ para o automorfismo de Anosov L_A, então $C_c(q)$ e $C_d(q)$ são invariantes por L_A^γ.*

Demonstração: Dado $p \in \mathbf{R}^2$ com $\mathcal{P}(p) = q$, podemos escrever:

$$C_c(q) = \mathcal{P}\left(\{p + u : u \in W_c\}\right).$$

Como W_c é invariante por A, então $u \in W_c \Longrightarrow A^\gamma u \in W_c$ donde

$$A^\gamma(p + u) \approx p + A^\gamma u \in \{p + u : u \in W_c\}.$$

Da mesma maneira podemos demonstrar a invariância de $C_d(q)$. □

Teorema 5.4.3 *Suponhamos que L_A é um automorfismo de Anosov e que $q \in \mathbf{T}^2$ é um ponto periódico. Se $\tilde{q} \in C_c(q)$ e $\hat{q} \in C_d(q)$ então*

$$\lim_{n\to\infty} \delta\left(L_A^n(q), L_A^n(\tilde{q})\right) = 0 \qquad \lim_{n\to-\infty} \delta\left(L_A^n(q), L_A^n(\hat{q})\right) = 0$$

Demonstração: Para calcular os limites acima observamos primeiro que as imagens de segmentos de reta paralelos a W_c e W_d por A e A^{-1} são também segmentos de reta paralelos a W_c e W_d. Para terminar a demonstração usaremos o seguinte lema:

Lema 5.4.1 *Se A é uma matriz que induz um automorfismo de Anosov em $\mathbf{T}^2$ e se l e $\hat{l}$ são segmentos de reta paralelos a W_c e W_d respectivamente, então $|A(l)| = |\lambda_c| \cdot |l|$ e $|A^{-1}(\hat{l})| = |\lambda_c| \cdot |\hat{l}|$, onde $|s|$ denota o comprimento do segmento de reta s.*

O teorema fica então demonstrado se aplicarmos o lema aos segmentos de reta $l = \overline{p\tilde{p}}$ e $\hat{l} = \overline{p\hat{p}}$ onde $\mathcal{P}p = q$, $\mathcal{P}\tilde{p} = \tilde{q}$ e $\mathcal{P}\hat{p} = \hat{q}$. □

Demonstração do Lema 5.4.1: Os segmentos l e $\hat{l}$ podem ser descritos como:

$$l = \{p + tu : u \in W_c, t \in [0,1]\} \qquad \hat{l} = \{\hat{p} + t\hat{u} : \hat{u} \in W_d, t \in [0,1]\}$$

e com esta notação temos $|l| = \|u\|$, $|\hat{l}| = \|\hat{u}\|$ e

$$\begin{aligned} A(l) &= \{Ap + tAu : u \in W_c, t \in [0,1]\} \\ A^{-1}(\hat{l}) &= \left\{A^{-1}\hat{p} + tA^{-1}\hat{u} : \hat{u} \in W_d, t \in [0,1]\right\}. \end{aligned}$$

Portanto $|A(l)| = \|Au\| = |\lambda_c| \cdot \|u\|$ e $|A^{-1}(\hat{l})| = \|A^{-1}\hat{u}\| = |\lambda_c| \cdot \|\hat{u}\|$. □

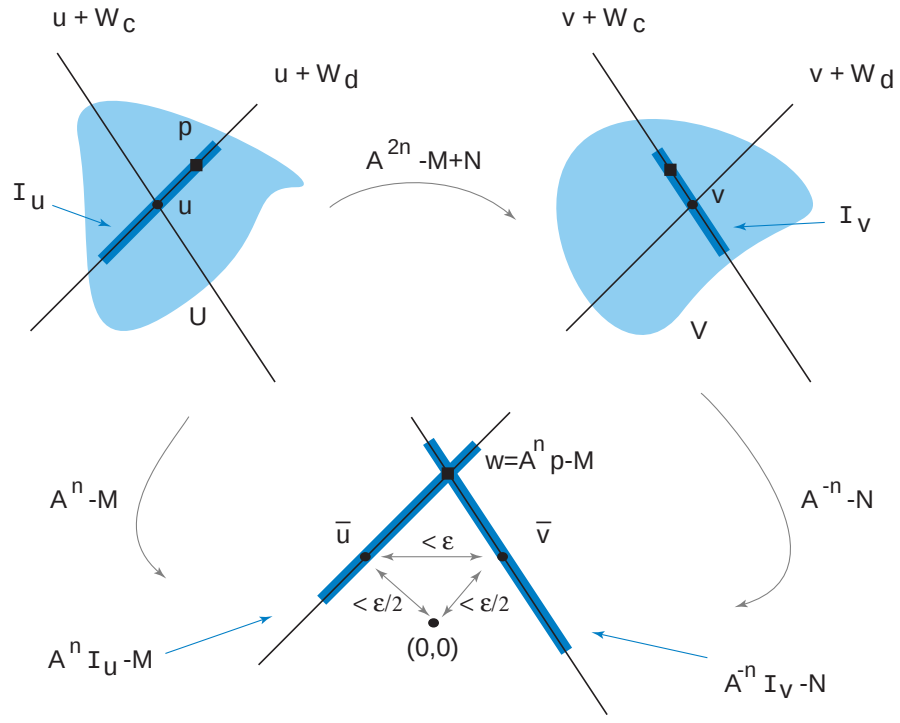
Como os conjuntos $C_c(q)$ e $C_d(q)$ são imagens de retas transversais com inclinação irracional, para cada $q \in \mathbf{T}^2$ estas curvas se intersectarão em um conjunto infinito de pontos. Um tal ponto de intersecção $\tilde{q}$ é chamado um *ponto homoclínico* a q e tem a propriedade de que $\lim_{n\to\pm\infty} \delta\left(L_A(q), L_A(\tilde{q})\right) = 0$. Como os pontos periódicos de um automorfismo de Anosov são densos em $\mathbf{T}^2$ então o conjunto de todos os pontos homoclínicos também é denso em $\mathbf{T}^2$. Isto quer dizer que em qualquer subconjunto aberto do toro existe um ponto homoclínico. (Nota: os subconjuntos abertos de $\mathbf{T}^2$ são exatamente as imagens por $\mathcal{P}$ de subconjuntos abertos de $\mathbf{R}^2$.)

Teorema 5.4.4 *Se L_A é um automorfismo de Anosov em $\mathbf{T}^2$, então dados dois subconjuntos abertos de $\mathbf{T}^2$, $\tilde{U}$ e $\tilde{V}$ existe um ponto $q \in \tilde{U}$ tal que $L_A(q) \in \tilde{V}$.*

Uma relação de recorrência com esta propriedade diz-se *topologicamente transitiva*.

Demonstração: Sejam U e $V \subset \mathbf{R}^2$ dois subconjuntos abertos tais que $\mathcal{P}(U) = \tilde{U}$ e $\mathcal{P}(V) = \tilde{V}$. Queremos mostrar que existem $p \in U$ e $\hat{p} \in V$ tais que $A^n p \approx \hat{p}$ para algum $m \in \mathbf{Z}$.

Começamos por tomar dois pontos $u \in U$ e $v \in V$ tais que $\mathcal{P}(u)$ e $\mathcal{P}(v)$ sejam pontos homoclínicos a $\mathcal{P}(0,0)$. Tomando $n \in \mathbf{N}$ suficientemente grande podemos garantir, para qualquer valor de $\varepsilon > 0$ dado, que $A^n u \approx \bar{u}$ com $|\bar{u} - (0,0)| < \frac{\varepsilon}{2}$ e $A^{-n} v \approx \bar{v}$ com $|\bar{v} - (0,0)| < \frac{\varepsilon}{2}$, (donde $|\bar{u} - \bar{v}| < \varepsilon$) isto é, podemos garnntir que existem $M, N \in \mathbf{Z}^2$ tais que $|A^n u - M| = |\bar{u} - M| < \frac{\varepsilon}{2}$ e $|A^{-n} v - N| = |\bar{v} - N| < \frac{\varepsilon}{2}$.



Queremos agora encontrar um ponto w perto de $\bar{u}$ e de $\bar{v}$ tal que $A^{-n}w + M$ está suficientemente próximo de u e portanto em U e que $A^n w + N$ está suficientemente próximo de v e portanto em V. Se encontrarmos um tal ponto, então basta tomar $p = A^{-n}w + M$ para garantir que $A^{2n} p \approx A^n w + N \in V$.

Para encontrar o ponto w escolhemos dois segmentos de reta I_u e I_v de comprimento $2\eta > 0$, tendo u e v respectivamente como pontos médios e com I_u paralelo a W_d, I_v paralelo a W_c (ver figura). Tomando η suficientemente pequeno podemos garantir que os intervalos estão contidos em U e V respectivamente.

Pelo lema 5.4.1, $|A^n(I_u)| = |\lambda_d|^n|I_u| = 2|\lambda_d|^n\eta$ e $|A^{-n}(I_v)| = |\lambda_d|^n|I_v| = 2|\lambda_d|^n\eta$. Além disso $\bar{u} + M \in A^n(I_u)$ e $\bar{v} + N \in A^{-n}(I_v)$. Como $A^n(I_u)$ e $A^{-n}(I_v)$ são sempre paralelos a W_c e a W_d respectivamente, tomando n suficientemente grande podemos garantir que que os segmentos $A^n(I_u) - M$ e $A^{-n}(I_v) - N$ se intersectam, já que podemos tomar $\bar{u}$ e $\bar{v}$ suficientemente próximos e que o comprimento das imagens dos dois segmentos tende para ∞ quando n cresce. Qualquer ponto $w \in A^n(I_u) - M \cap A^{-n}(I_v) - N$ será então equivalente à imagem por A^n de um ponto de U e à imagem por A^{-n} de um ponto de V.

A demonstração fica completa tomando $q = \mathcal{P}(A^{-n}w) \in \tilde{U}$ e $m = 2n$, donde $L_A^m(q) \in \tilde{V}$. □

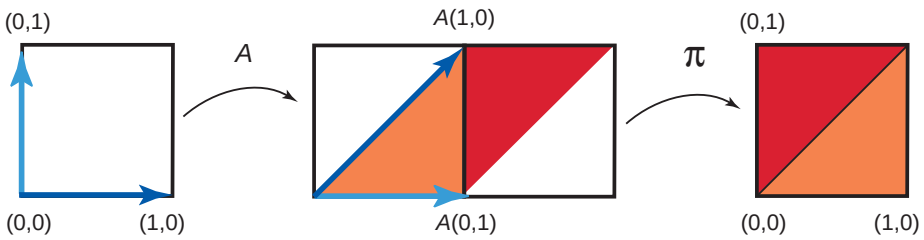
5.5 dinâmica simbólica

Para descrever em detalhe a dinâmica de um automorfismo de Anosov no toro, representaremos cada ponto p de $\mathbf{T}^2 \approx [0,1[^2$ como uma sucessão de números. Esta sucessão é escolhida de maneira a descrever a posição de cada ponto da órbita de p. Embora a construção possa ser feita para um automorfismo de Anosov qualquer, trabalharemos aqui com um exemplo.

A matriz

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

induz um automorfismo de Anosov em $\mathbf{T}^2$, já que tem entradas inteiras, com $\det(A) = -1$ e que A é hiperbólica.



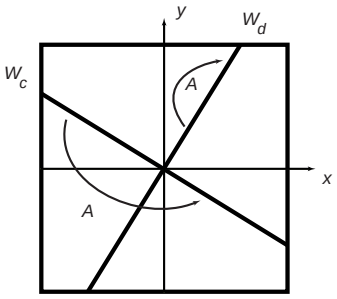
Os autovalores de A são:

$$\lambda_d = \frac{1+\sqrt{5}}{2} > 1 \qquad -1 < \lambda_c = \frac{1-\sqrt{5}}{2} < 0$$

e os autoespaços correspondentes são:

$$W_d = \{(x,y) : y = -\lambda_c x\} \qquad W_c = \{(x,y) : y = -\lambda_d x\}\,.$$

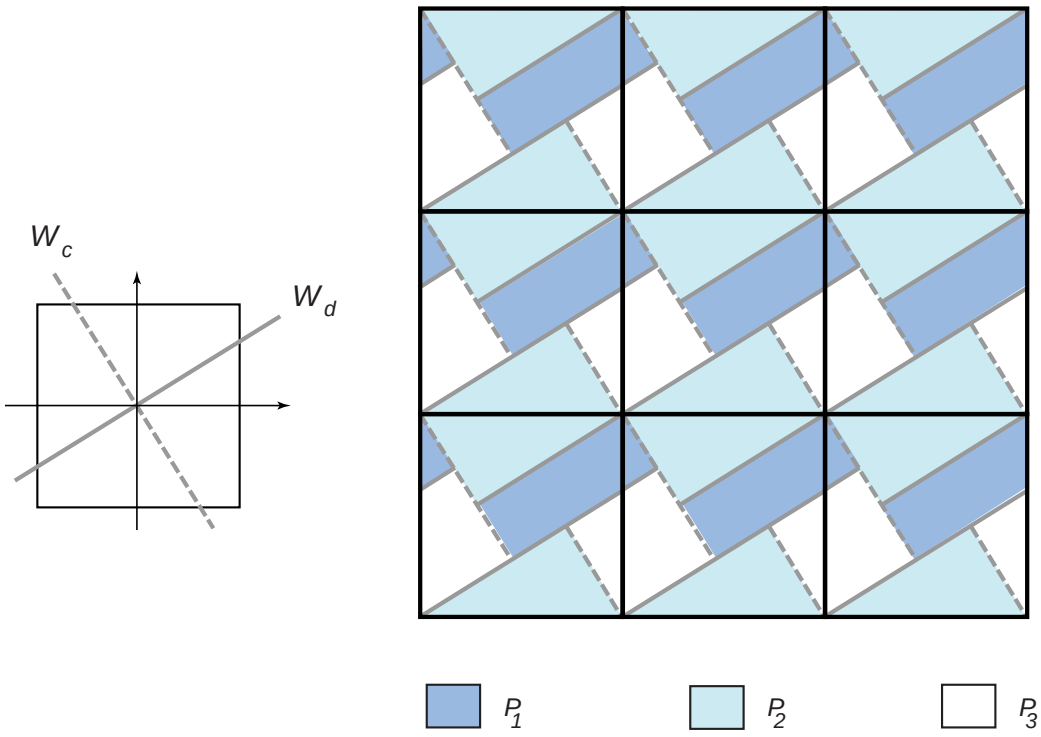
Observe-se que A inverte a orientação do plano.



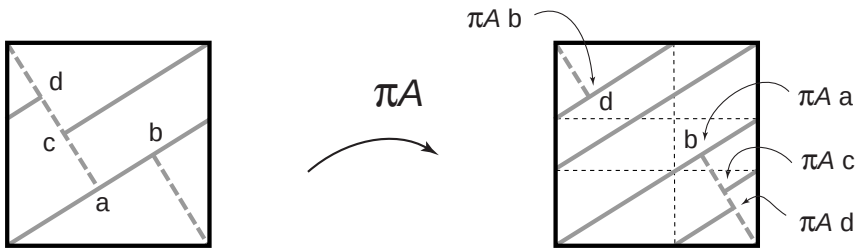
Podemos usar paralelas a W_c e W_d por pontos com coordenadas inteiras para dividir o plano em paralelogramos com as seguintes características:

1. Todos os paralelogramos têm lados paralelos a W_c ou a W_d.
2. Todos os paralelogramos são equivalentes (por $\approx$) a um de 3 paralelogramos P_1, P_2 e P_3 (ver figura abaixo).
3. Cada paralelogramo tem ao menos um vértice num ponto de coordenadas inteiras, isto é, num ponto equivalente a $(0,0)$.
4. $[0,1[^2 = \Pi(P_1) \cup \Pi(P_2) \cup \Pi(P_3)$.
5. Dois paralelogramos só se intersectam ao longo dos seus lados.

A decomposição de $[0,1[^2$ nas imagens dos paralelogramos P_j é chamada uma *partição de Markov de* $[0,1[^2$.

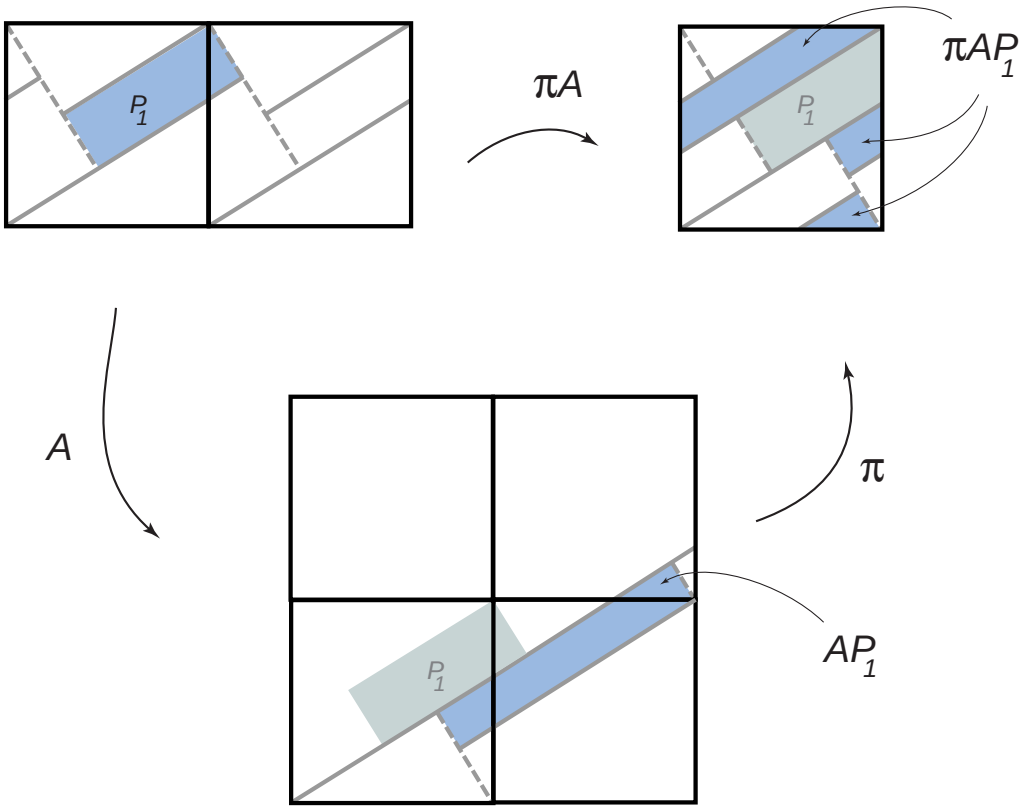


Podemos calcular as imagens por $\Pi \circ A$ de P_1, P_2 e P_3. Os cálculos são simples e indicamos aquí os seus resultados. Começamos por observar os vértices de $\Pi(P_1)$, $\Pi(P_2)$ e $\Pi(P_3)$ — a origem e os pontos a, b, c e d da figura abaixo. A origem é um ponto fixo e como os outros vértices são pontos homoclínicos suas imagens por $\Pi \circ A$ também o são (ver figura).



Denotando o interior de P_j por $\overset{o}{P}_j$, temos:

$$\Pi \circ A\left(\overset{o}{P}_1\right) \subset P_2 \cup P_3 \qquad \Pi \circ A\left(\overset{o}{P}_2\right) \subset P_1 \cup P_3 \qquad \Pi \circ A\left(\overset{o}{P}_3\right) \subset P_2.$$



Dado um ponto $p \in [0,1[^2$, sabemos que algum dos pontos equivalentes a p pertence a um dos paralelogramos P_1, P_2 ou P_3. Associamos a p o número $s_0(p)$ que é o índice deste paralelogramo, isto é:

$$p \in \Pi P_j \Longrightarrow s_0(p) = j.$$

É claro que esta definição é ambígua se p estiver na projeção por Π de um dos lados dos paralelogramos. Para evitar esta dificuldade consideramos inicialmente apenas pontos em $\bigcup \Pi \left(\overset{o}{P}_j\right)$.

Podemos associar a p números $s_n(p)$ que indicam a qual paralelogramo pertence cada ponto $\Pi \circ A^n(p)$, desde que todos os pontos da órbita de p estejam no interior de algum paralelogramo, isto é, associamos a p o número $s_n(p) \in \{1,2,3\}$ de modo que:

$$(\Pi \circ A)^n(p) \in \Pi \left(\overset{o}{P}_j\right) \Longrightarrow s_n(p) = j = s_0\left((\Pi \circ A)^n(p),\right).$$

Para cada ponto $p \in [0,1[^2$ obtivemos assim uma sequência "bi-infinita" de números, $(s_n(p))_{n=-\infty}^{\infty}$, com ambiguidade apenas no caso de alguma imagem de p pertencer a um dos lados de $\Pi(P_j)$. Da informação da página 33

concluimos que nem todas as sucessões $(s_n)_{n=-\infty}^{\infty}$ com $s_n \in \{1,2,3\}$ podem ocorrer. Por exemplo, como $\Pi \circ A\left(\overset{o}{P}_1\right) \cap \overset{o}{P}_1 = \emptyset$, nunca ocorrem dois 1 seguidos em $(s_n(p))_{n=-\infty}^{\infty}$.

A seqência $s_n(p)$ codifica informações sobre a órbita de p. Por exemplo, se p for um ponto periódico de período γ então $s_n(p) = s_{n+\gamma}(p)$. Assim a um ponto $p \in P_1$ de período 3 tal que $\Pi \circ Ap \in P_3$ e $(\Pi \circ A)^2 p \in P_2$ associamos a sequência

$$(s_n(p)) = \ldots 213213213\underline{1}3213213\ldots$$

em que $s_0(p)$ está sublinhado. Denotaremos uma sequência periódica como a deste exemplo por $\overline{213}$. É fácil ver que $s_n(\Pi \circ A(p)) = s_{n+1}(p)$.

Dizemos que uma sequência $(s_n)_{n=-\infty}^{\infty}$ com $s_n \in \{1,2,3\}$ é uma *sequência permitida* se satisfizer as seguintes condições:

1. $s_n \neq s_{n+1}$
2. Se $s_n = 3$ então $s_{n+1} \neq 1$.

Usando os dados da página 33 concluimos que toda sucessão obtida sem ambiguidade a partir de um ponto $p \in [0,1[^2$ é permitida — a imagem do interior de cada paralelogramo não intersecta o mesmo paralelogramo e a imagem de $\overset{o}{P}_3$ está contida em P_2.

Se denotarmos por $\mathcal{S}$ o conjunto das sequências permitidas podemos definir a *aplicação translação* $\sigma : \mathcal{S} \longrightarrow \mathcal{S}$, como $\sigma(s_n) = (s_{n+1})$. Temos então que $(s_n(\Pi \circ A(p))) = (s_{n+1}(p)) = \sigma(s_n(\Pi \circ A(p)))$ e assim σ é uma descrição de L_A em têrmos de sequências permitidas.

Resta estender a construção da sequência $(s_n(p))$ a pontos cuja órbita passa pelo lado de algum paralelogramo P_j.

Começamos pelos lados de paralelogramos que são segmentos de reta paralelos a W_d. Usando a notação da figura da página 33, com $o = (0,0)$, estes lados são todos equivalentes a partes do segmento l que passa em $\tilde{c}$, o, a, b, e $\tilde{d}$ (ver figura abaixo), em que $\tilde{c} \approx c$ e $\tilde{d} \approx d$. Como $o \in l \subset W_d$ então $A^{-1}(l) \subset l$, já que $A^{-1}\big|_{W_d}$ é uma contração. Isto quer dizer que se um ponto pertence a l todas as suas imagens por A^{-n} com $n > 0$ também estarão em l.

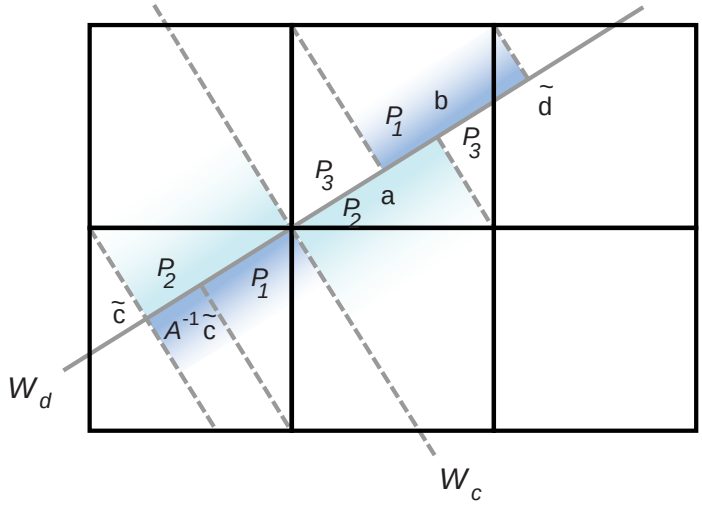
Usando a informação da figura abaixo e a das páginas 32 e 33, obtemos as imagens recíprocas dos lados paralelos a W_d. Em particular usamos: $A^{-1}\tilde{d} =$

b, $A^{-1}b = a$, e $A^{-1}a \in \overline{oa}$, $A^{-1}\tilde{c} \in \overline{o\tilde{c}}$. Os resultados são:

$$p \in P_1 \cap P_2 \;=\; \begin{cases} \overline{ab} \implies \Pi \circ A^{-1}p \in P_2 \cap P_3 & (A^{-1}(\overline{ab}) \subset \overline{oa}) \\ \cup \\ \overline{\tilde{c}o} \implies \Pi \circ A^{-1}p \in P_1 \cap P_2 & (A^{-1}(\overline{\tilde{c}o}) \subset \overline{\tilde{c}o}) \end{cases}$$

$$p \in P_1 \cap P_3 \;=\; \overline{b\tilde{d}} \implies \Pi \circ A^{-1}p \in P_1 \cap P_2 \qquad (A^{-1}(\overline{b\tilde{d}}) = \overline{ab})$$

$$p \in P_2 \cap P_3 \;=\; \overline{oa} \implies \Pi \circ A^{-1}p \in P_2 \cap P_3 \qquad (A^{-1}(\overline{oa}) \subset \overline{oa})$$



Com os dados da página 36 podemos então descrever as ambiguidades que podem aparecer na sequência $(s_n(p))$ quando p pertence a $P_i \cap P_j$. Há três casos a considerar:

1. $p \in P_2 \cap P_3$, logo há duas escolhas para $s_0(p)$:
   - $s_0(p) = 2$, e como $\Pi \circ A^{-1}p \in P_2 \cap P_3$, então $s_{-1}(0) = 2$ ou 3. Como 22 não pode aparecer numa sequência permitida, então temos que escolher $s_{-1}(0) = 3$.
   - $s_0(p) = 3$, e como $\Pi \circ A^{-1}p \in P_2 \cap P_3$, então $s_{-1}(0) = 2$ ou 3 e temos que escolher $s_{-1}(0) = 2$.

   Aplicando recursivamente o mesmo raciocínio a $s_{-n}(p)$ concluimos que $(s_n(p)) = \ldots 2323232\underline{2}s_1s2\ldots$ ou $(s_n(p)) = \ldots 3232323\underline{3}s_1s2\ldots$. Denotaremos estas duas sucessões respectivamente por $\overleftarrow{23}\underline{2}s_1s2\ldots$ e $\overleftarrow{32}\underline{3}s_1s2\ldots$, isto é, $\overleftarrow{a_1\ldots a_n}$ denota a repetição do bloco $a_1\ldots a_n$ indefinidamente

para a esquerda. No caso que estamos considerando temos que identificar quaisquer seqências que contenham os inícios $\overleftarrow{32}$ e $\overleftarrow{23}$. Indicaremos esta identificação por $\overleftarrow{32} \equiv \overleftarrow{23}$.

2. Se $p \in P_1 \cap P_2$, logo $s_0(p) = 1$ ou 2. Aquí há dois casos a considerar:
   (a) $p \in \overline{ab} \Longrightarrow \Pi \circ A^{-1}p \in P_2 \cap P_3$, logo $s_{-n}(p) = \overleftarrow{23}1 \equiv \overleftarrow{32}\underline{1} \equiv \overleftarrow{23}\underline{2}$ para $n = 0, 1, 2, \ldots$.
   (b) $p \in \overline{\tilde{c}o} \Longrightarrow \Pi \circ A^{-1}p \in P_1 \cap P_2$, logo $s_{-n}(p) = \overleftarrow{12} \equiv \overleftarrow{21}$ para $n = 0, 1, 2, \ldots$.
3. $p \in P_1 \cap P_3 = \overline{bd}$, logo $\Pi \circ A^{-1}p \in \overline{ab} \subset P_1 \cap P_2$ e $\Pi \circ A^{-2}p \in \overline{oa} \subset P_2 \cap P_3$. Portanto as possibilidades para $s_{-n}(p)$, $n = 0, 1, 2, \ldots$ são $\overleftarrow{23}2\underline{1}$, $\overleftarrow{23}2\underline{3}$ e $\overleftarrow{32}1\underline{3}$, que consideraremos equivalentes.

Em resumo, basta identificar todas as sequências que contenham os termos $\overleftarrow{23} \equiv \overleftarrow{23}$ bem como as subsucessões $\overleftarrow{12} \equiv \overleftarrow{21}$. A identificação $\equiv$ entre subsucessões que definimos acima é análoga à que se faz na representação decimal de números reais, onde as duas dízimas $0,439999999\ldots = 0,43\overline{9}$ e $0,44000000\ldots = 0,44\overline{0} = 0,44$ são identificadas como representando o mesmo número.

O mesmo estudo pode ser feito para os lados dos paralelogramos na direção W_c. Os pontos cuja órbita passa por algum destes lados serão representados por "dízimas periódicas à direita" do tipo $\ldots s_{-2}s_{-1}\underline{2}12121212\ldots$ que denotamos por $\ldots s_{-2}s_{-1}\underline{2}\overrightarrow{12}$. A única identificação que ocorre neste caso é $\overrightarrow{21} \equiv \overrightarrow{32}$. Assim obtemos para cada ponto de $[0,1[^2$ uma sequência permitida que o representa.

Teorema 5.5.1 *Toda sequência permitida representa um ponto de* $[0,1[^2$.

Demonstração: Dada uma sequência permitida $(a_n)_{n=-\infty}^{\infty}$ com $a_n \in \{1,2,3\}$ queremos encontrar $p \in [0,1[^2$ tal que $s_n(p) = a_n$, isto é que $\Pi \circ A^n(p) \in \Pi(P_{a_n})$, ou seja, que $p \in \Pi \circ A^{-n}(P_{a_n})$, para cada $n \in \mathbf{Z}$. O teorema ficará demonstrado se verificarmos que $\bigcap_{n=-\infty}^{\infty} \Pi \circ A^{-n}(P_{a_n})$ contém um só ponto.

Como (a_n) é uma sequência permitida, então

$$\Pi \circ A^{-n}(P_{a_n}) \cap \Pi \circ A^{-(n+1)}(P_{a_{n+1}}) \neq \emptyset.$$

Um dos lados paralelos a W_d de cada P_j está contido numa reta equivalente a W_d e por isto os lados paralelos a W_d de $A^{-n}(P_{a_n})$ terão comprimento menor

que $|\lambda_c|^n = 1/|\lambda_d|^n$. Assim $\bigcap_{n=0}^{k} \Pi \circ A^{-n}(P_{a_n})$ está sempre contido num paralelogramo tal que o comprimento dos lados paralelos a W_d tende para zero quando k tende para infinito. Pelo teorema dos intervalos encaixantes, a intersecção para todo $k \geq 0$ destes paralelogramos será um segmento de reta paralelo a W_c (ou seja, o produto cartesiano de um ponto por um segmento de reta).

Analogamente, $\left(\bigcap_{n=-k}^{0} \Pi \circ A^{-n}(P_{a_n})\right) \cap (\bigcap_{n=0}^{\infty} \Pi \circ A^{-n}(P_{a_n}))$ está sempre contido num segmento paralelo a W_c cujo comprimento tende para zero quando k tende para infinito e por isto $\bigcap_{n=-\infty}^{\infty} \Pi \circ A^{-n}(P_{a_n})$ contém um só ponto. □

Como consequência do teorema 5.5.1, para conhecer as órbitas de L_A em $\mathbf{T}^2$ basta estudar a aplicação translação σ no conjunto das sequências permitidas com as identificações acima. Por exemplo, se p é representado pela seqência $\overleftarrow{213}\underline{2}\overrightarrow{132}$, então p é um ponto periódico de período 3. Mais ainda, p, $\Pi \circ A(p)$ e $\Pi \circ A^2(p)$ são os únicos pontos com este período, já que não há outra sequência permitida com 3 algarismos. É fácil construir pontos que não são periódicos — basta encontrar uma sequência permitida que não o seja, por exemplo, $\overleftarrow{213}\underline{2}12123121212312121212312121212123\ldots$.