Estudo de Sistemas Dinâmicos

Pedro G. Mattos

13 de março de 2018

Sumário

1 Sistemas Dinâmicos

1.1 Sistemas Dinâmicos, Periodicidade e Invariância

Definição 1.1. Um *sistema dinâmico* é uma dupla $\mathcal{S} = (X, f)$ em que

1. X é um conjunto, o *espaço de fase* do sistema;

2. $f : \boldsymbol{T} \curvearrowright X$ é uma ação de monoides em X, a *dinâmica* do sistema, e $\boldsymbol{T}$ é um monoide, o *espaço temporal* do sistema, que pode ser $(\mathbb{I}_+, +)$, $(\mathbb{R}_+, +)$, $(\mathbb{I}, +)$ ou $(\mathbb{R}, +)$.

   Um sistema dinâmico é

   (a) *revertível* se $\boldsymbol{T}$ é um grupo, ou seja, se $\boldsymbol{T} = (\mathbb{I}, +)$ ou $\boldsymbol{T} = (\mathbb{R}, +)$;
   irrevetível caso contrário, ou seja, se $\boldsymbol{T} = (\mathbb{I}_+, +)$ ou $\boldsymbol{T} = (\mathbb{R}_+, +)$.

   (b) *discreto* se $\boldsymbol{T} = (\mathbb{I}_+, +)$ ou $\boldsymbol{T} = (\mathbb{I}, +)$;
   contínuo se $\boldsymbol{T} = (\mathbb{R}_+, +)$ ou $\boldsymbol{T} = (\mathbb{R}, +)$.

A seguir, estudaremos somente **sistemas dinâmicos discretos**, ou seja, sistemas cujo espaço temporal é $\mathbb{I}_+$ ou $\mathbb{I}$. Nesse caso, identificamos a dinâmica $f : \boldsymbol{T} \curvearrowright X$ com a função $f^1 : X \longrightarrow X$ por simplicidade, já que todas funções da ação podem ser representadas como composições de f^1, e também de $(f^1)^{-1}$ caso $\boldsymbol{T}$ seja um grupo.

Definição 1.2 (Iterados, Órbitas e Periodicidade)**.** Sejam $\mathcal{S} = (X, f)$ um sistema dinâmico, $x \in X$ e $n \in T$.

1. O *n-ésimo iterado* de x em $\mathcal{S}$ é o ponto $f^n(x)$;

2. A *órbita* de x em $\mathcal{S}$ é o conjunto

$$\mathcal{O}(x) := \{f^n(x) \mid n \in T\},$$

   a órbita de x sob a ação de f; a *semiórbita positiva* de x é o conjunto $\mathcal{O}_+(x) := \{f^n(x) \mid n \in T_+\}$;

3. Um ponto *n-periódico* é um ponto $x \in X$ que satisfaz $f^n(x) = x$ e n é um *período* de x; sua órbita é uma *órbita n-periódica.* O conjunto dos elementos periódicos de $\mathcal{S}$ é denotado $\mathrm{Per}(\mathcal{S})$ (ou $\mathrm{Per}(f)$ se for necessário ressaltar a dinâmica).

Proposição 1.1. *Sejam $\mathcal{S} = (X, f)$ um sistema dinâmico e $C \subseteq X$.*

1. *O conjunto*

$$\bigcup_{t \in T_+} f^t(C) = \bigcup_{x \in C} \mathcal{O}_+(x)$$

   é o conjunto das semiórbitas positivas dos pontos de C;

2. *O conjunto*

$$\bigcup_{t \in T_+} f^{-t}(C) = \{x \in X \mid \mathcal{O}_+(x) \cap C \neq \varnothing\}.$$

   é o conjunto dos pontos cujas semiórbitas positivas passam por C;

3. Se $\mathcal{S}$ é um discreto, o conjunto

$$\limsup_{t\in\mathbb{I}_+} f^{-t}(C) = \bigcap_{m=0}^{\infty}\bigcup_{n=m}^{\infty} f^{-t}(C) = \{x \in X \mid |\mathcal{O}_+(x)\cap C| = |\mathbb{N}|\}$$

é o conjunto dos pontos cujas seminórbitas positivas passam infinitamente por C.

Demonstração. 1. Para a primeira igualdade, basta notar que

$$\bigcup_{t\in T_+} f^t(C) = \bigcup_{t\in T_+}\bigcup_{x\in C}\{f^t(x)\} = \bigcup_{x\in C}\bigcup_{t\in T_+}\{f^t(x)\} = \bigcup_{x\in C}\mathcal{O}_+(x).$$

2. Para a segunda igualdade, basta notar que

$$\begin{aligned} x \in \bigcup_{t\in T_+} f^{-t}(C) &\Leftrightarrow \exists t \in T_+ \; x \in f^{-t}(C) \\ &\Leftrightarrow \exists t \in T_+ \; f^t(x) \in C \\ &\Leftrightarrow \mathcal{O}_+(x)\cap C \neq \varnothing. \end{aligned}$$

3. O conjunto $\limsup_{t\in\mathbb{I}_+} f^{-t}(C)$ é o conjunto dos pontos que pertencem a infinitos dos conjuntos $f^{-t}(C)$, ou seja, os pontos cujas órbitas intersecionam C infinitas vezes. ∎

Definição 1.3. Seja $\mathcal{S} = (X, f)$ um sistema dinâmico. Um conjunto *invariante* de $\mathcal{S}$ é um conjunto $C \subseteq X$ tal que $f^t(C) \subseteq C$ para todo $t \in T$. Um conjunto *positivamente invariante* de $\mathcal{S}$ é um conjunto $C \subseteq X$ tal que $f^t(C) \subseteq C$ para todo $t \in T_+$. Um conjunto *negativamente invariante* de $\mathcal{S}$ é um conjunto $C \subseteq X$ tal que $f^{-t}(C) \subseteq C$ para todo $t \in T_+$.

Proposição 1.2. *Sejam $\mathcal{S} = (X, f)$ um sistema dinâmico e $C \subseteq X$. Então*

1. C é positivamente invariante se, e somente se, as semiórbitas positivas de todo ponto de C está em C;

2. C é negativamente invariante se, e somente se, os pontos cujas semiórbitas positivas passam por C estão em C;

3. Se C é negativamente invariante, então $C^{\complement}$ é positivamente invariante.

Demonstração.

Evidente das igualdades de conjuntos da proposição anterior.

Evidente das igualdades de conjuntos da proposição anterior.

Se C é negativamente invariante, então $f^{-1}(C) \subseteq C$, logo $C^{\complement} \subseteq (f^{-1}(C))^{\complement} = f^{-1}(C^{\complement})$, portanto $f(C^{\complement}) \subseteq f(f^{-1}(C^{\complement})) \subseteq C^{\complement}$. ∎

No caso de sistemas dinâmicos discretos, pode-se mostrar por indução que essa definição é equivalente a exigir que $f(C) \subseteq C$ e $f^{-1}(C) \subseteq C$.

Proposição 1.3. *Sejam $\mathcal{S} = (X, f)$ um sistema dinâmico discreto e $C \subseteq X$. Então*

1. *Se* $f(C) \subseteq C$*, então* C *é positivamente invariante;*

2. *Se* $f^{-1}(C) \subseteq C$*, então* C *é negativamente invariante.*

Demonstração. 1. Mostraremos por indução que, para todo $t \in \mathbb{I}_+$, $f^t(C) \subseteq C$. Para $t = 0$, temos que $f^0(C) = C \subseteq C$. Agora seja $t \in \mathbb{I}_+{}^*$ e suponha que $f^{t-1}(C) \subseteq C$. Então, como $f(C) \subseteq C$,

$$f^t(C) = f^{t-1}(f(C)) \subseteq f^{t-1}(C) \subseteq C.$$

2. A mesma demonstração vale ao substituir f^t por f^{-t}.

■

Proposição 1.4. *Seja* $\mathcal{S} = (X, f)$ *um sistema dinâmico discreto e* $C \subseteq X$*. Então*

1. C *é invariante se, e somente se,* $f^{-1}(C) = C$*;*

2. *Se* f *é sobrejetiva e* C *é invariante, então* $f(C) = C$*;*

3. *Se* f *é injetiva e* $f(C) = C$*, então* C *é invariante.*

Demonstração. 1. ($\Rightarrow$) Como C é positivamente invariante, vale $f(C) \subseteq C$, logo $f^{-1}(f(C)) \subseteq f^{-1}(C)$. Mas para quaisquer função vale que $C \subseteq f^{-1}(f(C))$, portanto $C \subseteq f^{-1}(f(C)) \subseteq f^{-1}(C)$. Como C é negativamente invariante, $f^{-1}(C) \subseteq C$, o que mostra que $f^{-1}(C) = C$.

($\Leftarrow$) Como $f^{-1}(C) = C$, segue da proposição 1.3 que C é negativamente invariante. Também dessa relação segue que $f(C) = f(f^{-1}(C))$ e, como $f(f^{-1}(C) \subseteq C$ para qualquer função, segue que $f(C) \subseteq C$, portanto C é positivamente invariante, o que mostra que ele é invariante.

2. Como f é sobrejetiva, $f(f^{-1}(C)) = C$. Da invariância negativa de C, segue que $f^{-1}(C) \subseteq C$, portanto $f(f^{-1}(C)) \subseteq f(C)$. Mas então $C \subseteq f(C)$. Da invariância positiva de C, segue que $f(C) \subseteq C$ e, portanto, $f(C) = C$.

3. Como $f(C) = C$,segue que $f^{-1}(C) = f^{-1}(f(C))$ e, como $f^{-1}(f(C)) = C$ para função injetiva, segue que $f^{-1}(C) = C$, portanto C é invariante pela proposição acima.

■

1.2 Conjugação de Sistemas Dinâmicos

Definição 1.4. Sejam $\mathcal{S}_1 = (X_1, f_1)$ e $\mathcal{S}_2 = (X_2, f_2)$ sistemas dinâmicos. Uma *semi-conjugação* (morfismo de sistemas dinâmicos) de $\mathcal{S}_1$ para $\mathcal{S}_2$ é uma função sobrejetiva $\phi : X_1 \longrightarrow X_2$ que satisfaz $\phi \circ f_1 = f_2 \circ \phi$ (o diagrama comuta).

$$\begin{tikzcd}
X_1 \arrow[r, "f_1"] \arrow[d, "\phi"'] & X_1 \arrow[d, "\phi"] \\
X_2 \arrow[r, "f_2"'] & X_2
\end{tikzcd}$$

Denota-se $\phi : \mathcal{S}_1 \longrightarrow \mathcal{S}_2$. Uma *conjugação* (isomorfismo de sistemas dinâmicos) de $\mathcal{S}_1$ para $\mathcal{S}_2$ é uma semiconjugação $\phi : \mathcal{S}_1 \longrightarrow \mathcal{S}_2$ que é invertível.

Sempre que temos uma função sobrejetiva $\phi : X_1 \longrightarrow X_2$, para todo $y \in X_2$ existe $x \in X_1$ tal que $y = \phi(x)$. Assim podemos induzir a dinâmica de (X_1, f_1) em X_2, definindo $f_2(y) := \phi(f_1(x))$, de modo que a dinâmica f_2 é uma função de X_2 para X_2 e vale $f_2 \circ \phi = \phi \circ f_1$. Isso faz com que os sistemas (X_1, f_1) e (X_2, f_2) sejam semiconjugados.

A partir das propriedade da conjugação, vamos deduzir alguns resultados imediatos. Podemos notar que as funções compostas ${f_1}^n$ e ${f_2}^n$ ainda satisfazem a propriedade comutativa que f_1 e f_2 satisfazem. Isso basicamente diz que em sistemas discretos, preservar a função f é preservar a ação da dinâmica. Segue a proposição.

Proposição 1.5. *Sejam (X_1, f_1) e (X_2, f_2) sistemas dinâmicos e $\phi : X_1 \longrightarrow X_2$ uma semiconjugação. Então, para todo natural $n \geq 1$, vale $\phi \circ {f_1}^n = {f_2}^n \circ \phi$ (o diagrama comuta).*

$$\begin{tikzcd}
X_1 \arrow[r, "{f_1}^n"] \arrow[d, "\phi"'] & X_1 \arrow[d, "\phi"] \\
X_2 \arrow[r, "{f_2}^n"'] & X_2
\end{tikzcd}$$

Demonstração. Vamos provar a proposição acima usando indução. Primeiro, notamos que a propriedade vale para $n = 1$ por definição. Agora, supondo que valha $\phi \circ {f_1}^{n-1} = {f_2}^{n-1} \circ \phi$, temos que

$$\begin{aligned}
\phi \circ {f_1}^n &= \phi \circ ({f_1}^{n-1} \circ f_1) \\
&= (\phi \circ {f_1}^{n-1}) \circ f_1 \\
&= ({f_2}^{n-1} \circ \phi) \circ f_1 \\
&= {f_2}^{n-1} \circ (\phi \circ f_1) \\
&= {f_2}^{n-1} \circ (f_2 \circ \phi) \\
&= ({f_2}^{n-1} \circ f_2) \circ \phi \\
&= {f_2}^n \circ \phi,
\end{aligned}$$

o que conclui a indução. ∎

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

Proposição 1.6 (Semiconjugação preserva peridiocidade da órbita.)**.** *Sejam* (X_1, f_1) *e* (X_2, f_2) *sistemas dinâmicos e* $\phi : X_1 \longrightarrow X_2$ *uma semiconjugação. Para todo* $x \in X$*, se* x *é um ponto* n*-periódico, então* $\phi(x)$ *é um ponto* n*-periódico. Ainda, o período mínimo de* $\phi(x)$ *divide o período mínimo de* x*.*

Demonstração. Seja $x \in X$ tal que $x = {f_1}^n(x)$. Pela propriedade da semiconjugação,

$$\begin{aligned}\phi(x) &= \phi({f_1}^n(x)) \\ &= (\phi \circ {f_1}^n)(x) \\ &= ({f_2}^n \circ \phi)(x) \\ &= {f_2}^n(\phi(x)).\end{aligned}$$

Como $\phi(x)$ é n-periódico, segue que seu período mínimo divide n. ■

1.3 Exemplos

1.3.1 Rotação em $\mathbb{T}^1$

Definição 1.5. Seja $[\alpha] \in \mathbb{T}$. A *rotação* por $[\alpha]$ em $\mathbb{T}^1$ é a função

$$\begin{aligned} R_\alpha \colon \mathbb{T}^1 &\longrightarrow \mathbb{T}^1 \\ [x] &\longmapsto [x + \alpha]. \end{aligned}$$

Uma *rotação racional* é uma rotação cuja constante α é racional e uma *rotação irracional* é uma rotação cuja constante α é irracional.

Proposição 1.7. *A dupla* $(\mathbb{T}^1, R_\alpha)$ *é um sistema dinâmico discreto revertível para todo* $\alpha \in \mathbb{R}$ *e*

1. *se* α *é racional, todos os pontos de* $\mathbb{T}^1$ *são periódicos;*

2. *se* α *é irracional, nenhum ponto de* $\mathbb{T}^1$ *é periódico.*

Demonstração. Por definição a dupla é um sistema dinâmico discreto e, como $R_{-\alpha} = {R_\alpha}^{-1}$, segue que o sistema é invertível.

1. Seja $[x] \in \mathbb{T}^1$. Para mostrar que $[x]$ é periódico, devemos achar um natural n tal que ${R_\alpha}^n([x]) = [x]$. A constante α é um número racional e, por isso, pode ser escrita em forma de fração como $\alpha = \frac{p}{q}$, sendo p e q inteiros e q não nulo. Tomando $n = q$, e notando que $q\alpha = p$ é um inteiro, segue que

$$ {R_\alpha}^q([x]) = [x + q\alpha] = [x] $$

o que mostra que $[x]$ é q-periódico.

2. Seja $[x] \in \mathbb{T}^1$. Podemos provar a afirmação acima por contradição, notando que, se houvesse algum inteiro n tal que ${R_\alpha}^n([x]) = [x]$, como ${R_\alpha}^n([x]) = [x + n\alpha]$, teríamos $n\alpha$ inteiro. Mas como n é inteiro e α irracional, isso é impossível. ■

1.3.2 Translação no Toro $\mathbb{T}^d$

Um caso mais geral que a rotação no círculo $\mathbb{T}^1$ é o caso da tranlação no toro $\mathbb{T}^d$. Ela é basicamente uma rotação em cada entrada e é chamada de translação pois é a transformação induzida no espaço quociente $\mathbb{T}^d$ de uma translação em $\mathbb{R}^d$. Ela é também um caso específico de translações em grupos topológicos.

Definição 1.6. Seja $[\alpha] \in \mathbb{T}^d$. A *translação* por $[\alpha]$ em $\mathbb{T}^d$ é a função

$$\begin{aligned} T_\alpha \colon \mathbb{T}^d &\longrightarrow \mathbb{T}^d \\ [x] &\longmapsto [x+\alpha]. \end{aligned}$$

1.3.3 Expansão em $\mathbb{T}^1$

Definição 1.7. Seja $N \in \mathbb{I}_+$. A *expansão por* N em $\mathbb{T}^1$ é a função

$$\begin{aligned} E_N \colon \mathbb{T}^1 &\longrightarrow \mathbb{T}^1 \\ [x] &\longmapsto [Nx]. \end{aligned}$$

Note que a função está bem definida porque N é um número inteiro.

1.4 Deslocamentos Simbólicos

Adotaremos a convenção de que o natural $N \in \mathbb{N}$ é o conjunto dos naturais menores que ele, seguindo a construção de ordinais do Von Neumann. Isso simplificará a notação. Assim temos que $N = \{0, \cdots, N-1\}$. Nesta seção consideraremos duas dinâmicas semelhantes, pois ambas angem em sequências, mas diferentes, pois uma considera sequências unilaterais e outra, bilaterais.

1.4.1 Deslocamento Unilateral

Consideremos o conjunto

$$N^{\mathbb{I}_+} = \left\{ (x_i)_{i\in\mathbb{I}_+} \,\middle|\, \forall i \in \mathbb{I}_+ \quad x_i \in N \right\}$$

das *sequências unilaterais* em N símbolos, as funções do tipo

$$\begin{aligned} x \colon \mathbb{I}_+ &\longrightarrow N \\ i &\longmapsto x_i. \end{aligned}$$

Definição 1.8. O *deslocamento unilateral* em $N^{\mathbb{I}_+}$ é a função

$$\begin{aligned} \sigma \colon N^{\mathbb{I}_+} &\longrightarrow N^{\mathbb{I}_+} \\ (x_i)_{i\in\mathbb{I}_+} &\longmapsto (x_{i+1})_{i\in\mathbb{I}_+}. \end{aligned}$$

Essa dinâmica é conhecida como *deslocamento de Bernoulli*. Basicamente, o deslocamento desloca cada entrada de uma sequência $(x_0, x_1, x_2, \ldots)$ à esquerda, levando-a para sequência $(x_1, x_2, x_3, \ldots)$. Claramente essa função não é invertível, logo o sistema dinâmico formado por ela não é invertível. Rassaltamos isso na proposição a seguir.

Proposição 1.8. *A dupla* $(N^{\mathbb{I}_+}, \sigma)$ *é um sistema dinâmico discreto irrevertível.*

Proposição 1.9. *Seja* $k \in \mathbb{I}_+$. *Uma sequência unilateral* $(x_i)_{i\in\mathbb{I}_+} \in N^{\mathbb{I}_+}$ *é um ponto* k*-periódico de* $(N^{\mathbb{I}_+}, \sigma)$ *se, e somente se, para todo* $i \in \mathbb{I}_+$, *vale* $x_i = x_{i+k}$.

Demonstração. Se vale $x_i = x_{i+k}$ para todo $i \in \mathbb{I}_+$, a sequência $(x_i)_{i\in\mathbb{I}_+}$ claramente será k-periódica, pois

$$\sigma^k((x_i)_{i\in\mathbb{I}_+}) = (x_{i+k})_{i\in\mathbb{I}_+} = (x_i)_{i\in\mathbb{I}_+}.$$

Por outro lado, se $(x_i)_{i\in\mathbb{I}_+}$ tiver órbita periódica, valerá $\sigma^k((x_i)_{i\in\mathbb{I}_+}) = (x_i)_{i\in\mathbb{I}_+}$ para algum k. Nesse caso, teremos que $x_i = x_{i+k}$ para todo $i \in \mathbb{I}_+$. ■

Proposição 1.10. *A função*

$$\begin{aligned}\phi\colon N^{\mathbb{I}_+} &\longrightarrow \mathbb{T}^1\\ (x_i)_{i\in\mathbb{I}_+} &\longmapsto \left[\sum_{i\in\mathbb{I}_+} \frac{x_i}{N^{i+1}}\right]\end{aligned}$$

é uma semiconjugação de $(N^{\mathbb{I}_+}, \sigma)$ *para* $(\mathbb{T}^1, E_N)$.

Demonstração. Seja $(x_i)_{i\in\mathbb{I}_+} \in N^{\mathbb{I}_+}$. Então

$$\begin{aligned}\phi \circ \sigma((x_i)_{i\in\mathbb{I}_+}) &= \phi((x_{i+1})_{i\in\mathbb{I}_+})\\ &= \left[\sum_{i\in\mathbb{I}_+} \frac{x_{i+1}}{N^{i+1}}\right]\\ &= \left[N \sum_{i\in\mathbb{I}_+} \frac{x_{i+1}}{N^{i+2}}\right]\\ &= \left[N \left(\frac{x_0}{N} + \sum_{i\in\mathbb{I}_+^*} \frac{x_i}{N^{i+1}}\right)\right]\\ &= \left[N \sum_{i\in\mathbb{I}_+} \frac{x_i}{N^{i+1}}\right]\\ &= E_N \circ \phi((x_i)_{i\in\mathbb{I}_+}).\end{aligned}$$

Somamos x_0 na quarta lilnha pois é um interio, não altera a classe de equivalência. ■

Notemos que a função ϕ não é uma conjugação porque não é invertível – existe mais de uma representação N-ária para alguns números reais em $[0, 1]$, portanto mais de uma sequência em $N^{\mathbb{I}_+}$ que o representa. No entanto, veremos mais à frente que isso não será problema quando considerarmos sistemas dinâmicos de medida, pois o conjunto desses pontos problemáticos terá medida nula e, portanto, a conjugação poderá ser definida. Isso será importante pois significa que os dois sistemas são o mesmo sob a ótica de sistemas dinâmicos de medida.

1.4.2 Deslocamento Bilateral

Consideremos o conjunto

$$N^{\mathbb{I}} = \{(x_i)_{i\in\mathbb{I}} \mid \forall i \in \mathbb{I} \quad x_i \in N\}$$

das *sequências bilaterais* em N símbolos, as funções do tipo

$$\begin{aligned} x\colon \mathbb{I} &\longrightarrow N \\ i &\longmapsto x_i. \end{aligned}$$

Definição 1.9. O *deslocamento bilateral (à esquerda)* em $N^{\mathbb{I}}$ é a função

$$\begin{aligned} \sigma\colon N^{\mathbb{I}} &\longrightarrow N^{\mathbb{I}} \\ (x_i)_{i\in\mathbb{I}} &\longmapsto (x_{i+1})_{i\in\mathbb{I}}. \end{aligned}$$

O *deslocamento bilateral à direita* é σ^{-1}.

Proposição 1.11. *A dupla* $(N^{\mathbb{I}}, \sigma)$ *é um sistema dinâmico discreto revertível.*

1.4.3 Subdeslocamentos

Os deslocamentos definidos anteriormente são deslocamentos completos, no sentido de que todas as sequências são consideradas no sistema. É possível, porém, selecionar subconjuntos dessas sequências tal que o deslocamento ainda está bem definido nelas. Definimos que cada entrada de uma sequência $(x_i)_{i\in\mathbb{I}} \in N^{\mathbb{I}}$ tem entradas em $\{0, \ldots, N-1\}$. Esses subconjuntos são escolhidos a partir de uma regra que indica qual entrada de uma sequência pode ser seguida de qual outra entra – por exemplo, para $N = 9$, sequências em que uma entrada 4 só pode ser segida de uma entrada 2 ou uma entrada 8. Para ter toda a informação necessária, essas sequências são definidas a partir de uma *matriz de adjacência*, uma matriz que indica qual entrada pode ser adjacente a qual outra.

Definição 1.10. Uma *matriz de adjacência* de ordem N é uma matriz $A \in \mathbb{M}_N(2) = \mathbb{M}_N(\mathbb{I}_2)$ de entradas 0 ou 1 tal que $a_{i,j} = 0$ se uma entrada com valor i não pode ser seguida de uma entrada com valor j e $a_{i,j} = 1$ se uma entrada de valor i pode ser seguuida de uma entrada de valor j.

Definição 1.11. Sejam $N \in \mathbb{I}_+^*$ e A uma matriz de adjacência de ordem N. O conjunto das sequências de $N^{\mathbb{I}}$ restritas por A é o conjunto

$$N^{\mathbb{I}}|_A := \left\{(x_i)_{i\in\mathbb{I}} \,\middle|\, \forall i \in \mathbb{I}, a_{x_i,x_{i+1}} = 1\right\}.$$

O conjunto das sequências de $N^{\mathbb{I}_+}$ restritas por A é o conjunto

$$N^{\mathbb{I}_+}|_A := \left\{(x_i)_{i\in\mathbb{I}_+} \,\middle|\, \forall i \in \mathbb{I}_+, a_{x_i,x_{i+1}} = 1\right\}.$$

O *subdeslocamento bilateral* com *matriz de transição* A é a função $\sigma|_A := \sigma|_{N^{\mathbb{I}}|_A}$ e o *subdeslocamento unilateral* com *matriz de transição* A é a função $\sigma|_A := \sigma|_{N^{\mathbb{I}_+}|_A}$. Os sistemas $(N^{\mathbb{I}}|_A, \sigma|_A)$ e $(N^{\mathbb{I}_+}|_A, \sigma|_A)$ são *sistemas de deslocamento parcial* e a matriz A é a *matriz de transição* dos sistemas.

Esses sistemas também são conhecidos como *cadeias de Markov*[1].

Proposição 1.12. *Sejam* $N \in \mathbb{I}_+^*$ *e* A *uma matriz de adjacência de ordem* N*. A inclusão* $\iota : N^{\mathbb{I}}|_A \longrightarrow N^{\mathbb{I}}$ *é uma semiconjugação do sistema* $(N^{\mathbb{I}}|_A, \sigma|_A)$ *para o sistema* $(N^{\mathbb{I}}|_A, \sigma|_A)$.

Demonstração. Queremos mostrar que $\iota \circ \sigma|_A = \sigma \circ \iota$. Seja $(x_i)_{i\in\mathbb{I}} \in N^{\mathbb{I}}|_A$. Então

$$\begin{aligned}
\iota \circ \sigma|_A((x_i)_{i\in\mathbb{I}}) &= \iota(\sigma|_A((x_i)_{i\in\mathbb{I}})) \\
&= \iota((x_{i+1})_{i\in\mathbb{I}}) \\
&= (x_{i+1})_{i\in\mathbb{I}} \\
&= \sigma((x_i)_{i\in\mathbb{I}}) \\
&= \sigma(\iota((x_i)_{i\in\mathbb{I}})) \\
&= \sigma \circ \iota((x_i)_{i\in\mathbb{I}}).
\end{aligned}$$

■

O mesmo vale para os sietmas unilaterais. A partir da matriz da matriz de transição podemos determinar quantos pontos periódicos de um determinado período existem.

Proposição 1.13. *Sejam* $p \in \mathbb{I}_+^*$ *e* $(N^{\mathbb{I}}|_A, \sigma|_A)$ *um sistema de subdeslocamento com matriz de transição* A*. A quantidade de pontos* p*-periódicos de* $(N^{\mathbb{I}}|_A, \sigma|_A)$ *é* $\mathrm{tr}(A^p)$.

Demonstração. Sabemos que $(x_i)_{i\in\mathbb{I}}$ é p-periódico se, e somente se, para todo $i \in \mathbb{I}$ vale $x_i = x_{i+p}$. Sendo assim, procuramos os pontos da forma

$$(\ldots, x_{i_0}, x_{i_1}, \ldots, x_{i_{p-2}}, x_{p-1}, \ldots)$$

e pportanto devemos achar os vetores $(i_0, \ldots, i_{p-1}) \in N^p$ tais que a sequência de entradas

$$i_0 \to i_1 \to \cdots i_{p-2} \cdots i_{p-1}$$

é uma sequência permitida pela matriz de transição. Isso é equivalente a termos

$$\bigtimes_{k=0}^{p-2} a_{i_k,i_{k+1}} = a_{i_0,i_1} a_{i_1,i_2} \cdots a_{i_{p-3},i_{p-2}} a_{i_{p-2},i_{p-1}} = 1,$$

pois esse produto é 1 se, e somente se todos os fatores são 1. Assim, a quantidade de pontos periódicos é a soma de todos esses produtos

$$\begin{aligned}
|\mathrm{Per}_p(\sigma_A)| &= \sum_{(i_0,\ldots,i_{p-1})\in N^p} \bigtimes_{k=0}^{p-2} a_{i_k,i_{k+1}} \\
&= \sum_{i_0=0}^{N-1} \sum_{(i_1,\ldots,i_{p-1})\in N^{p-1}} \bigtimes_{k=0}^{p-2} a_{i_k,i_{k+1}} \\
&= \sum_{i_0=0}^{N-1} (A^p)_{i_0,i_0} = \mathrm{tr}(A^p).
\end{aligned}$$

■

[1] *Andrei Andreyevich Markov* (1856 - 1922), matemático russo.

1.4.4 Itinerários

Definição 1.12. Sejam (X, f) um sistema dinâmico discreto, $\mathcal{P} = (P_i)_{i\in N}$ uma partição finita de X ($N \in \mathbb{N}$) e $x \in X$. O *itinerário* de x conforme $\mathcal{P}$ é a sequência $(x_t)_{t\in T} \in N^T$ tal que $x_t = k$ se, e somente se, $f^t(x) \in P_k$.

Proposição 1.14. *Sejam (X, f) um sistema dinâmico discreto, $\mathcal{P} = (P_i)_{i\in N}$ uma partição finita de X e $\phi : X \longrightarrow N^T$ a função que leva $x \in X$ em seu itinerário. Então*

$$\phi \circ f = \sigma \circ \phi.$$

Demonstração. Sejam $x \in X$ e $\phi(x) = (x_t)_{t\in T}$. Então $\phi(f(x)) = (x_{t+1})_{t\in T}$, pois $f^t(f(x)) = f^{t+1}(x)$. Portanto

$$\phi \circ f(x) = \phi(f(x)) = (x_{t+1})_{t\in T} = \sigma((x_t)_{t\in T}) = \sigma(\phi(x)) = \sigma \circ \phi(x),$$

o mostra que $\phi \circ f = \sigma \circ \phi$. ■

A imagem $\phi(X)$ é invariante por σ, pois

$$\sigma(\phi(X)) = \phi(f(X)) = \phi(X),$$

de modo que ϕ é uma semiconjugação de X para $\phi(X)$, pois satisfaz a propriedade comutativa e é sobrejetiva por definição. Nem sempre tem-se ϕ injetiva, isso depende da partição e do sistema que se estuda. Ainda, nem sempre ϕ preserva as estruturas adicionais de X. Nas seções posteriores serão apresentadas partições que preservam tais estruturas.

Sejam (X, f) um sistema dinâmico discreto, $\mathcal{P} = (P_i)_{i\in N}$ uma partição finita de X e $(x_t)_{t\in T} \in N^T$. O conjunto

$$\bigcap_{t\in T} f^{-t}(P_{x_t}).$$

é o conjunto dos pontos $x \in X$ tal que $f^t(x) \in P_{x_t}$ para todo $t \in T$, pois $f^t(x) \in P_{x_t}$ se, e somente se, $x \in f^{-t}(P_{x_t})$; ou seja, o conjunto de todos os pontos tais que $\phi(x) = (x_t)_{t\in T}$.

1.4.5 Ferradura

Nesta seção, trataremos de um sistema dinâmico conhecido por Ferradura de Smale. O conjunto usado nesta seção será criado a partir de um conjunto $D \subseteq \mathbb{R}^2$. Esse conjunto D é particionado em cinco conjuntos: $D = D_1 \cup D_2 \cup D_3 \cup D_4 \cup D_5$. 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.

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

Assuma que f estica $D_2 \cup D_4$ uniformemente na direção horizontal por um fator $m > 2$ 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 deslocamento simbólico. Vamos definir os seguintes conjuntos:

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

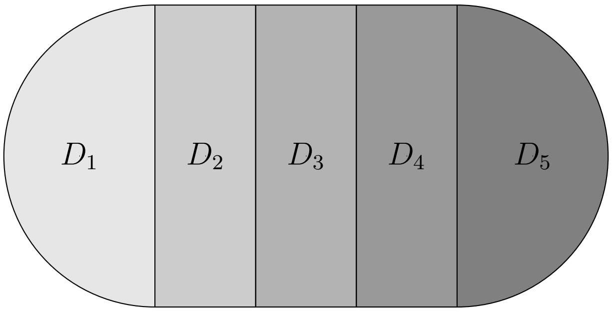

Figura 1: O conjunto $D = D_1 \cup R \cup D_5$

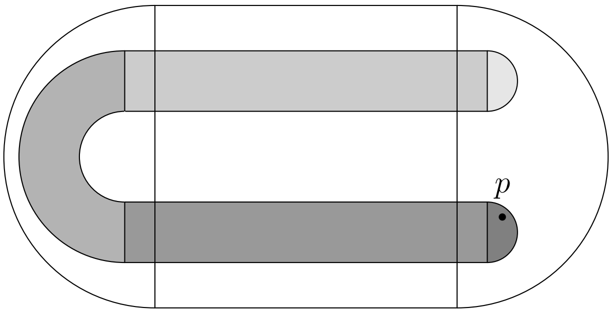

Figura 2: A imagem de D sob o endomorfismo f

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.

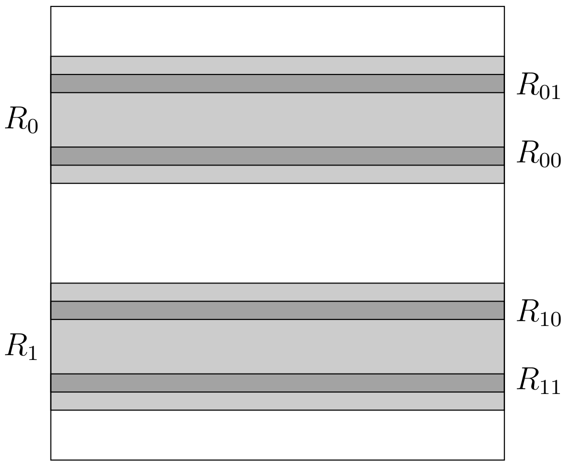

Figura 3: O quadrado R e conjuntos R_ω horizontais usados na conjugação

De modo geral, para qualquer sequência finita $\omega = (\omega_0, \omega_1, \ldots, \omega_n) \in 2^{n+1}$ de zeros e uns, definimos

$$R_\omega := \bigcap_{i=0}^{n} f^i(R_{\omega_i}) = R_{\omega_0} \cap f(R_{\omega_1}) \cap f^2(R_{\omega_2}) \cap \cdots \cap f^n(R_{\omega_n}),$$

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 unilateral $\omega = (\omega_i)_{i \in \mathbb{I}_+} \in 2^{\mathbb{I}_+}$, o retângulo R_ω é definido como

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

Construiremos, de modo semelhante, retângulos verticais a parir 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 m^{-1}. Assim, para uma sequência finita $\omega = (\omega_1, \ldots, \omega_n) \in 2^n$ de zeros e uns, definimos

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

que será um retângulo vertical de largura m^{-n} e $f^{-n}(R) \cap R$ será a união dos 2^n possíveis retângulos dessa forma. Para uma sequência unilateral $\omega = (\omega)_{i \in \mathbb{I}_+^*} \in 2^{\mathbb{I}_+^*}$, definimos

$$R'_\omega := \bigcap_{i \in \mathbb{I}_+^*} f^{-i}(R_{\omega_i}).$$

Definição 1.13. A *ferradura* é o conjunto F definido como

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

Lembrando das definições dos retângulos no começo da seção, podemos perceber que F é a união de todos os conjuntos $R_\omega \cap R'_\omega$, com ω variando em $2^{\mathbb{I}}$.

Definição 1.14. A conjugação ϕ do sistema da ferradura (H, f) para o deslocamento bilateral $(2^{\mathbb{I}}, \sigma^{-1})$ é definida como

$$\begin{aligned} \phi \colon F &\longrightarrow 2^{\mathbb{I}} \\ x &\longmapsto (x_i)_{i \in \mathbb{I}} \end{aligned}$$

tal que $x_i = k$ se $f^i(x) \in R_k$.

Proposição 1.15. *A função ϕ é uma conjugação do sistema (F, f) para o sistema $(2^{\mathbb{I}}, \sigma^{-1})$.*

Demonstração. Primeiramente, vamos demonstrar que a função ϕ é bijetiva. Para mostrar a sobrejetividade, consideremos $(x_i)_{i \in \mathbb{I}} \in 2^{\mathbb{I}}$. Definamos as sequências positiva e negativa de (x_i) como as sequências unilaterais

$$(x_i)^+ := (x_0, x_1, \ldots) \text{ e } (x_i)^- := (x_{-1}, x_{-2}, ...).$$

Consideremos os retângulos horizontais $R_{x_0}, R_{(x_0,x_1)}, R_{(x_0,x_1,x_2)}, \ldots$. Esses retâgulos estão cada um contido no anterior, $R_{x_0} \supset R_{(x_0,x_1)} \supset R_{(x_0,x_1,x_2)} \supset \ldots$. Se consideramos suas interseções com alguma reta vertical em R, temos intervalos encaixados $I_{x_0}, I_{(x_0,x_1)}, I_{(x_0,x_1,x_2)}, \ldots$, respectivamente, que são intervalos fechados e satisfazem $I_{x_0} \supset I_{(x_0,x_1)} \supset I_{(x_0,x_1,x_2)} \supset \ldots$. Portanto, pelo teorema dos intervalos encaixados, concluímos que a interseção desses intervalos é não vazia: $I_{x_0} \cap I_{(x_0,x_1)} \cap I_{(x_0,x_1,x_2)} \cap \ldots \neq \varnothing$. 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 $R'_{x_{-1}}, R'_{(x_{-1},x_{-2})}, R'_{(x_{-1},x_{-2},x_{-3})}, \ldots$. Suas interseções com uma reta horizontal em R são intervalos encaixados $I'_{x_{-1}}, I'_{(x_{-1},x_{-2})}, I'_{(x_{-1},x_{-2},x_{-3})}, \ldots$, respectivamente, que são intervalos fechados e e satisfazem $I'_{x_{-1}} \supset I'_{(x_{-1},x_{-2})} \supset I'_{(x_{-1},x_{-2},x_{-3})} \supset \ldots$. 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 $m^{-1} < 1$, a interseção desses intervalos terá um único elemento. Novamente, esse argumento vale para qualquer reta horizontal em R, e concluímos 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, concluímos que $\phi(x) = (x_i)_{i\in\mathbb{I}}$. Isso prova a sobrejetividade.

Para mostrar a injetividade, consideremos dois elementos $a, b \in F$ e sejam suas imagens $\phi(a) = (a_n)_{n\in\mathbb{I}}$ e $\phi(b) = (b_n)_{n\in\mathbb{I}}$. Suponhamos que $(a_n)_{n\in\mathbb{I}} = (b_n)_{n\in\mathbb{I}}$. Nesse caso, sabemos que $R_{(a_n)_{n\in\mathbb{I}}} = R_{(b_n)_{n\in\mathbb{I}}}$ e que, portanto, a e b pentencem a $R_{(a_n)_{n\in\mathbb{I}}}$. Mas foi demonstrado acima que, dado $\omega \in 2^{\mathbb{I}}$, 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 $x \in F$ e $\phi(x) = (x_n)_{n\in\mathbb{I}}$. Nesse caso, sabemos que $x \in R_{(x_n)_{n\in\mathbb{I}}}$. Aplicando f a $R_{(x_n)}$, temos

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

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

1.4.6 Automorfismos Hiperbólicos em $\mathbb{T}^2$

Após o estudo de duas dinâmicas envolvendo o círculo, a dinâmica de rotação na e a de expansão, voltamos nossa atenção nesta seção a um novo objeto geométrico: o toro. O sistema dinâmico a ser estudado nesta seção tem como conjunto o toro $\mathbb{T}^2$, 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 dependente 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 $A = [a_{i,j}]$. A dimensão de uma matriz será explicitada somente quando necessário e será simplificada para matrizes quadradas, como em $A_{2x2} = A_2$. Denotaremos o determinante de uma matriz A por $|A|$ ou $\det(A)$ e sua matriz inversa por A^{-1}. Convenções esclarecidas, definimos o automorfismo.

Definição 1.15. Seja $A \in \mathrm{SL}_2(\mathbb{I})$ uma matriz cujos autovalors têm valor absoluto não unitário. O *automorfismo hiperbólico* no toro é definido a partir de A como

$$
\begin{aligned}
[A] \colon \mathbb{T}^2 &\longrightarrow \mathbb{T}^2 \\
[x] &\longmapsto [Ax].
\end{aligned}
$$

Para justificar a nomenclatura dessa classe de funções como automorfismos hiperbólicos, devemos demonstrar algumas propriedades tanto da matriz A como da função $[A]$. 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 $[A](x)$ esteja bem definida no toro. Isso ocorre pois, para todo $y \in [x]$, temos $y = x + k$, com $k \in \mathbb{I}^2$, e segue que

$$
[Ay] = [A(x + k)] = [Ax + Ak] = [Ak],
$$

pois $Ak \in \mathbb{I}^2$.

Devemos justificar por que essa função é um automorfismo. Claramente, a função $[A]$ 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 $[A]$ recebe o nome de hiperbólico porque matrizes cujos autovalores têm valor absoluto diferente de 1 recebem esse nome. Vale notar, por fim, que para todo inteiro k, vale

$$
[A]^k = [A^k].
$$

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

Proposição 1.16. *Seja $A \in \mathrm{SL}_2(\mathbb{I})$ uma matriz cujos autovalors têm valor absoluto não unitário. Então*

1. *Os autovalores de A são*

$$
\begin{aligned}
\lambda_+ &:= \frac{1}{2}\left(a_{11} + a_{22} + \sqrt{(a_{11} + a_{22})^2 - 4}\right) \\
\lambda_- &:= \frac{1}{2}\left(a_{11} + a_{22} - \sqrt{(a_{11} + a_{22})^2 - 4}\right)
\end{aligned}
$$

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

2. Os autovalores de λ_+ e λ_- são irracionais;

3. Os autovetores de A são

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

Demonstração. 1. 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 entradas são inteiras, isso implicaria $a_{11} = a_{22} = \pm 1$, e teríamos a contradição $a_{11} + a_{22} = \pm 2$.

2. 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. 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 **??**. 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.

3. 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 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 $a_{12} \neq 0$, e temos $y = \frac{\lambda - a_{11}}{a_{12}}x$. Tomando $x = 1$, resultam os autoveotres como no enunciado.

■

Definição 1.16. Um *ponto racional* de $\mathbb{T}^2$ é um ponto $[r] \in \mathbb{T}^2$ tal que $r \in \mathbb{Q}^2$.

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

Proposição 1.17. *O conjunto $Per([A])$ é o conjunto dos pontos racionais de $\mathbb{T}^2$.*

Demonstração. Seja r um ponto racional de $\mathbb{T}^2$ da forma $r := [(\frac{a}{k}, \frac{b}{k})]$, em que a, b e k são naturais, k não nulo. Notemos que, ao aplicarmos $[A]$ em r, temos

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

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 $[(\frac{a}{k}, \frac{b}{k})]$ continuam dessa forma após aplicarmos $[A]$ e, de fato, $[A]$ 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 $[A]^i(r) = [A]^j(r)$ e $|i - j| \leq k^2$. Sem perda de generalidade, supomos $j \geq i$ e, aplicando $[A]^{-i}$ ao dois lados dessa igualdade, temos $r = [A]^{j-i}(r)$; ou seja, r tem órbita periódica de período menor ou igual a k^2. ■

Proposição 1.18. *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

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

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 calcelam 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

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

concluímos 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 1.19. *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,

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

seque 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 $\det(A^n - I) = 2 - (\lambda_+)^n - (\lambda_-)^n$. ■

Proposição 1.20. *Seja $A \in \mathrm{SL}_2(\mathbb{I})$ uma matriz cujos autovalores não são raízes da unidade. Então*

$$|\mathrm{Per}_n([A])| = |\det(A^n - I)|.$$

Demonstração. Vamos calcular a quantidade de pontos fixos de A e mostrar que $|\mathrm{Per}_1([A])| = |\det(A-I)|$. Como A^n é uma matriz da forma de A, segue que $|\mathrm{Per}_n([A])| = |\det(A^n - I)|$. Para achar a quantidade fixos sob $[A]$, temos que achar as soluções de $Ax = x + m$, com $x \in [0,1]^2$ e $m \in \mathbb{I}^2$; ou seja, as soluções de

$$(A - I)x = m, \tag{1}$$

sendo $x \in [0,1]^2$ e $m \in \mathbb{I}^2$. De fato, se observamos a transformação $T := A - I$ aplicada ao quadrado $[0,1]^2$, subconjunto de $\mathbb{R}^2$, vemos que o quadrado é levado em um paralelogramo $T([0,1]^2)$ cujos vértices são pontos inteiros de $\mathbb{R}^2$. Os pontos inteiros que pertencem ao paralelogramo são as soluções m de (1). Pelo Teorema de Pick, sabemos que, se a é a área de $T([0,1]^2)$, i a quatidade de pontos inteiros internos a $T([0,1]^2)$ e f a quatidade de pontos inteiros na fronteira de $T([0,1]^2)$, então

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

É possível mostrar que os únicos pontos na fronteira do paralelogramo $T([0,1]^2)$ 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 quatidade deles é, portanto, igual à área do paralelogramo.

Para mostrar que os únicos pontos da fronteira do paralelogramo são seus vértices, mostraremos que cada um de seus lados tem 2 pontos inteiros. Note que basta mostrar a afirmação para dois lados não paralelos do paralelogramo. O lado de extremidades $(0,0)$ e $(1,0)$ do quadrado é levado no lado de extremidades $(0,0)$ e (a_{11}, a_{21}). A inclinação desse lado é $m := \frac{a_{21}}{a_{11}}$ e, portanto, qualquer ponto inteiro nesse lado deve ter essa inclinação. A quantidade de pontos inteiros $(x,y) \neq (0,0)$ que satisfazem $\frac{y}{x} = m$, tal que $x \leq a_{11}$ é $\mathrm{mdc}(a_{11}, a_{21})$. Mas, como $|A| = 1$, temos que $\mathrm{mdc}(a_{11}, a_{21}) = 1$, pois a_{11} e a_{21} estão na mesma coluna de A, o que implica que $|A|$ é múltiplo de $\mathrm{mdc}(a_{11}, a_{21})$. Assim, concluímos que há dois pontos inteiros nesse lado, $(0,0)$ e (a_{11}, a_{21}). Da mesma forma, como $\mathrm{mdc}(a_{12}, a_{22}) = 1$, conluímos o mesmo para o lado de extremidades $(0,0)$ e (a_{12}, a_{22}) do paralelogramo, a imagem do lado de extremidades $(0,0)$ e $(0,1)$ do quadrado. Assim, concluímos que todos lados têm somente os vértices como pontos inteiros.

Basta-nos, portanto, calcular a área do paralelogramo $T([0,1]^2)$. Mas isso é simples, pois ela é igual ao módulo do determinante da tranformação, já que a área de $[0,1]^2$ é 1. Assim, concluímos que $|\mathrm{Per}_1([A])| = |\det(A - I)|$. ∎

2 Sistemas Dinâmicos Topológicos

Definição 2.1. Um *sistema dinâmico topológico* é uma dupla $\mathcal{S} = (\boldsymbol{X}, f)$ em que $\boldsymbol{X}$ é um espaço topológico, (X, f) é um sistema dinâmico e f é uma ação contínua em $\boldsymbol{X}$.

2.1 Conjugação Topológica

Definição 2.2. Sejam $\boldsymbol{\mathcal{S}_1} = (\boldsymbol{X_1}, f_1)$ e $\boldsymbol{\mathcal{S}_2} = (X_2, f_2)$ sistemas dinâmicos topológicos. Uma *semiconjugação topológica* (morfismo de sistemas dinâmicos topológico) de $\boldsymbol{\mathcal{S}_1}$ para

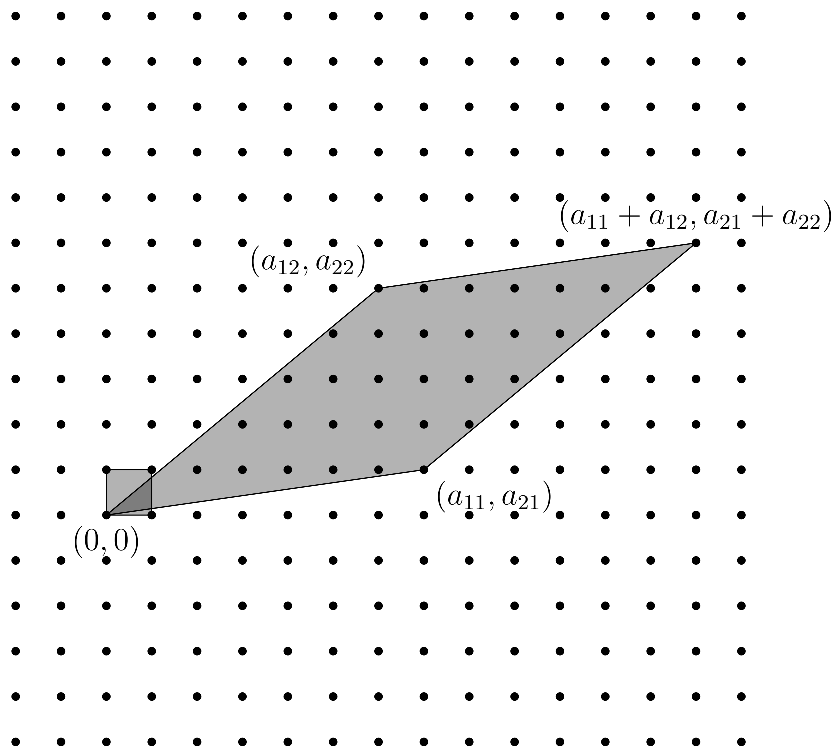

Figura 4: O quadrado $[0,1]^2$ e o paralelogramo $T([0,1]^2)$

$\boldsymbol{\mathcal{S}_2}$ é uma semiconjugação $\phi : \mathcal{S}_1 \longrightarrow \mathcal{S}_2$ que é uma função contínua de $\boldsymbol{X_1}$ para $\boldsymbol{X_2}$. Denota-se $\phi : \boldsymbol{\mathcal{S}_1} \longrightarrow \boldsymbol{\mathcal{S}_2}$. Uma *conjugação topológica* (isomorfismo de sistemas dinâmicos topológicos) de $\boldsymbol{\mathcal{S}_1}$ para $\boldsymbol{\mathcal{S}_2}$ é uma semiconjugação topológica $\phi : \boldsymbol{\mathcal{S}_1} \longrightarrow \boldsymbol{\mathcal{S}_2}$ que é um homeomorfismo de $\boldsymbol{X_1}$ para $\boldsymbol{X_2}$.

2.2 Transitividade Topológica e Minimalidade

Definição 2.3. Um sistema dinâmico *topologicamente transitivo* é um sistema dinâmico topológico $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ em que existe ponto $x \in X$ cuja órbita $\mathcal{O}_f(x)$ é densa em X.

Uma definição equivalente é: para todos abertos não vazios $A, B \subseteq X$ existe $n \in \mathbb{I}_+$ tal que $f^{-n}(A) \cap B \neq \varnothing$.

Definição 2.4. Um sistema dinâmico *mínimo* é um sistema dinâmico topológico $\boldsymbol{\mathcal{S}} = (\boldsymbol{X}, f)$ cujos pontos têm órbita densa em X.

Proposição 2.1. *O sistema de rotação $(\mathbb{T}^1, R_\alpha)$ é mínimo se α é irracional.*

Proposição 2.2. *Seja $\boldsymbol{G}$ um grupo topológico, $\alpha \in G$ e $E_\alpha : G \longrightarrow G$ a translação à esquerda por α. Se o sistema $(\boldsymbol{G}, E_\alpha)$ é topologicamente transitivo, então ele é mínimo.*

Demonstração. Seja ■

2.3 Exemplos

2.3.1 Automorfismos no Toro

Após essa discussão inicial sobre o sistema $(\mathbb{T}^2, [A])$, 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 rigorosa.

Definição 2.5. Seja $[p] \in \mathbb{T}^2$. A *variedade estável* de $[p]$ é o conjunto

$$W^-([p]) := \left\{ [p + \alpha v_{\lambda_-}] \;\middle|\; \alpha \in \mathbb{R} \right\},$$

Analogamente, a *variedade instável* de $[p]$ é o conjunto

$$W^+([p]) := \left\{ [p + \alpha v_{\lambda_+}] \;\middle|\; \alpha \in \mathbb{R} \right\}.$$

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

Teorema 2.3. *Seja $[p]$ um ponto do toro. Então, para um ponto $[p'] \in W^s([p])$,*

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

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

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

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,

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

e $[A]^n([p]) = [A^n p]$. Notemos que, como $|\lambda_-| \leq 1$, para n suficientemente grande a distância no toro é igual à distância no plano, pois $|\lambda_-|^n \to 0$. 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 2.4. *As variedades estável e instável de um ponto $[p]$ do toro são invariavéis sob ação do automorfismo $[A]$ no sentido que valem $[A](W^s([p])) = W^s([A]([p]))$ e $[A](W^u([p])) = W^u([A]([p]))$.*

Demonstração. Na demosntração, 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 $[A](W([p])) \subset W([A]([p]))$. Seja $[x] \in [A](W([p]))$. Então existe $\alpha \in \mathbb{R}$ tal que $[x] = [A]([p + \alpha v])$. Logo

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

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

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

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

e, portanto, $x \in [A](W^s(p))$. ∎

Proposição 2.5. *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 demosntração, 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 potos 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

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

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 $k \in \mathbb{I}$; 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 obtido atráves de uma rotação irracional R_l (definida na seção **??**) aplicada em $m \in S^1$. Como sabemos pelo teorema **??**

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. ■

Teorema 2.6. *A dinâmica do sistema* $(\mathbb{T}^2, [A])$ *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 intervalo em U e V ambos aumentam ao aplicarmos neles, respectivamente, $[A]^-1$ e $[A]$, existem $p, q \in \mathbb{N}$ tais que $[A]^{-p}(I_U) \cap [A]^q(I_V) \neq \varnothing$. Assim, tomamos um ponto w' nesse conjunto. Segue que $[A]^{-q}(w') \in V$ e $[A]^p(w') \in U$. Definindo $w := [A]-q(w')$ e tomando $n = p + q$, temos que $w \in [A]^{-n}(V) \cap U$ e está provado o teorema. ■

Entropia

Definição 2.6. A *entropia* de um sistema dinâmico (X, f) é definida como

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

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

Teorema 2.7. *A entropia de sistema* $(\mathbb{T}^2, [A])$ *associado à matriz* A *é*

$$h_{[A]} = \log |\lambda_+|.$$

Demonstração. De fato, dos resultados anteriores vem que

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

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

3 Sistemas Dinâmicos Diferenciais

Definição 3.1. Um *sistema dinâmico diferencial* é uma dupla $\mathcal{S} = (\boldsymbol{X}, f)$ em que $\boldsymbol{X}$ é uma variedade diferencial, (X, f) é um sistema dinâmico e f é uma ação diferencial em $\boldsymbol{X}$.

4 Sistemas Dinâmicos de Medida

Definição 4.1. Um *sistema dinâmico de medida* (ou *sistema que preserva medida*) é uma dupla $\mathcal{S} = (\boldsymbol{X}, f)$ em que $\boldsymbol{X}$ é um espaço de medida, (X, f) é um sistema dinâmico e f é uma ação que preserva medida em $\boldsymbol{X}$.

Como estudaremos somente o caso **discreto**, a função f é no caso um morfismo de medida, uma função que preserva a medida do espaço. Ainda, estudaremos somente **espaços de probabilidade**. Introduzimos agora noções de morfismo de sistemas dinâmicos de probabilidade. Em geral, quando tratamos da teoria da medida, podemos considerar funções que são iguais a menos de um conjunto de medida nula.

4.1 Conjugação que Preserva Medida

Definição 4.2. Sejam $\mathcal{S}_\mathbf{1} = (\boldsymbol{X}_\mathbf{1}, f_1)$ e $\mathcal{S}_\mathbf{2} = (\boldsymbol{X}_\mathbf{2}, f_2)$ sistemas que preservam medida. Uma *semiconjugação de medida* de $\mathcal{S}_\mathbf{1}$ para $\mathcal{S}_\mathbf{2}$ é uma semiconjugação $\phi : C_1 \longrightarrow C_2$ que preserva medida, definida em conjuntos mensuráveis $C_1 \in \mathcal{M}_1$ e $C_2 \in \mathcal{M}_2$ que satisfazem $f_1(C_1) \subseteq C_1$, $f_2(C_2) \subseteq C_2$ e $m_1(C_1) = m_2(C_2) = 1$. Nesse caso, o sistema $\mathcal{S}_\mathbf{2}$ é um *fator* do sistema $\mathcal{S}_\mathbf{1}$. Uma *conjugação de medida* de $\mathcal{S}_\mathbf{1}$ para $\mathcal{S}_\mathbf{2}$ é uma semiconjugação de medida que é uma conjugação.

$$\begin{tikzcd}
X_1 \arrow[r, "f_1"] \arrow[d, "\phi"'] & X_1 \arrow[d, "\phi"] \\
X_2 \arrow[r, "f_2"'] & X_2
\end{tikzcd}$$

Como comentado anteriormente, podemos considerar ϕ como definida de X_1 para X_2 sem problemas, pois só não está definida em um conjunto de medida nula.

4.2 Recorrência

Lema 4.1. *Sejam $(X, \mathcal{M}, m)$ um espaço de medida finito e $(A_n)_{n\in\mathbb{N}}$ uma sequência decrescente de conjuntos mensuráveis. Então*

$$m\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\longrightarrow+\infty} m(A_n).$$

Teorema 4.2 (Recorrência de Poincaré)**.** *Sejam* $(\boldsymbol{X}, f)$ *um sistema dinâmico de medida e* $E \in \mathcal{M}$ *um conjunto mensurável tal que* $E \not\doteq \varnothing$*. Então quase todo ponto de* E *retorna infinitamente a* E*:*

$$\{x \in E \mid |\mathcal{O}_+(x) \cap E| = |\mathbb{N}|\} \doteq E.$$

Demonstração. Notemos que

$$\{x \in E \mid |\mathcal{O}_+(x) \cap E| = |\mathbb{N}|\} = E \cap \limsup_{t \in \mathbb{I}_+} f^{-t}(E).$$

Esse é o conjunto dos pontos de E que retornam infinitamente para E. Queremos provar que esse conjunto tem a mesma medida de E. Consideremos os conjuntos

$$E_N := \bigcup_{n=N}^{\infty} f^{-n}(E).$$

Temos que $f^{-1}(E_N) = E_{N+1}$, e, como f preserva m, segue que $m(E_N) = m(E_{N+1})$. Isso implica que $m(E_0) = m(E_N)$ para todo $N \in \mathbb{N}$. Observe que $E_0 \supset E_1 \supset E_2 \ldots$, então segue, pelo lema 4.1, que

$$m\left(\bigcap_{N=0}^{\infty} E_N\right) = \lim_{N \longrightarrow \infty} m(E_N) = m(E_0).$$

Notando que $\bigcap_{N=0}^{\infty} E_N = \limsup_{t \in \mathbb{I}_+} f^{-t}(E)$, consequentemente segue que

$$m(F) = m\left(E \cap \limsup_{t \in \mathbb{I}_+} f^{-t}(E)\right) = m(E \cap E_0) = m(E).$$

A segunda igualdade segue de $\bigcap_{N=0}^{\infty} E_N = E_0$ quase sempre, e a terceira vem de $E \subseteq E_0$. ■

Como mencionado antes, a hipótese de que o sistema tem medida finita é fundamental para usar o lema 4.1.

4.3 Ergodicidade

Nesta seção, introduziremos o conceito de ergodicidade. A definição de transformação em sistemas ergódicos virá seguida de uma proposição que mostra várias equivalências de propriedades que caracterizam a ergodicidade. Essas equivalências são importantes não só como resultados teóricos interessantes e importantes ferramentas da teoria, mas também como embasamento intuitivo da ideia de ergodicidade. Relembramos que um conjunto é quase vazio quando tem medida 0 e quase total quando seu complementar tem medida 0.

Definição 4.3. Um *sistema dinâmico ergódico* é um sistema dinâmico de medida $\mathcal{S} = (\boldsymbol{X}, f)$ tal que todo conjunto mensurável invariante de $\mathcal{S}$ é quase vazio ou quase total: para todo $M \in \mathcal{M}$ tal que $f^{-1}(M) = M$, $M \doteq \varnothing$ ou $M \doteq X$. A dinâmica f é uma *dinâmica ergódica*.

Proposição 4.3. *Seja* $\mathcal{S} = (\boldsymbol{X}, f)$ *um sistema dinâmico de medida. São equivalentes*

1. *f é ergódica;*
2. *Todo conjunto quase invariante de $\mathbf{S}$ é quase vazio ou quase total (para todo $M \in \mathcal{M}$ tal que $f^{-1}(M) \doteq M$, $M \doteq \varnothing$ ou $M \doteq X$);*
3. *A semiórbita positiva de quase todo ponto passa por qualquer conjunto de medida positiva: para todo $M \in \mathcal{M}$ tal que $M \not\doteq \varnothing$,*
$$\bigcup_{n \in \mathbb{I}_+} f^{-n}(M) = \{x \in X \mid \mathcal{O}_+(x) \cap M \neq \varnothing\} \doteq X;$$
4. *Se a semiórbita positiva de quase nenhum ponto de um conjunto mensurável passa por um outro conjunto mensurável, algum dos conjuntos é quase vazio: para todos $N, M \in \mathcal{M}$ tais que*
$$N \cap \bigcup_{n \in \mathbb{I}_+} f^{-n}(M) \doteq \varnothing,$$
$N \doteq \varnothing$ ou $M \doteq \varnothing$.

Demonstração. (1) $\Rightarrow$ (2) Pela proposição B.8, existe um conjunto invariante $N \in \mathcal{M}$ tal que $M \doteq N$. Pela ergodicidade, segue que $N \doteq \varnothing$ ou $N \doteq X$, portanto segue que $M \doteq \varnothing$ ou $M \doteq X$.

(2) $\Rightarrow$ (3) Seja $M \in \mathcal{M}$ com $m(M) > 0$ e definamos

$$N := \bigcup_{n \in \mathbb{I}_+} f^{-n}(M).$$

Segue que $f^{-1}(N) \subseteq N$. Isso implica que $f^{-1}(N) \cup N = N$ e que $f^{-1}(N) \cap N = f^{-1}(N)$. Portanto, como $m(f^{-1}(N)) = m(N)$, pois f preserva medida, segue que

$$\begin{aligned} m(N \triangle f^{-1}(N)) &= m(f^{-1}(N) \cup N) - m(f^{-1}(N) \cap N) \\ &= m(N) - m(f^{-1}(N)) = 0. \end{aligned}$$

Assim, temos que $f^{-1}(N) \doteq N$, e de (2) segue que $N \doteq \varnothing$ ou $N \doteq X$. Mas $N \not\doteq \varnothing$, pois $f^{-1}(M) \subseteq N$ e $m(f^{-1}(M)) = m(M) > 0$, logo temos que $N \doteq X$.

(3) $\Rightarrow$ (4) Vamos provar a recíproca. Sejam $M, N \in \mathcal{M}$ tais que $M \not\doteq \varnothing$ e $N \not\doteq \varnothing$. Por (3), sabemos que

$$m\left(\bigcup_{n \in \mathbb{I}_+} f^{-n}(M)\right) = 1,$$

e disso segue que

$$0 < m(N) = m\left(\bigcup_{n \in \mathbb{I}_+} N \cap f^{-n}(M)\right) \leq \sum_{n \in \mathbb{I}_+} m(N \cap f^{-n}(M)).$$

Logo, existe $n \in \mathbb{I}_+$ tal que $m(N \cap f^{-n}(M)) > 0$.

(4) $\Rightarrow$ (1) Seja $M \in \mathcal{M}$ tal que $f^{-1}(M) = M$. Então segue que $f^{-n}(M) = M$ para todo $n \in \mathbb{I}_+$, logo

$$0 = m(M \cap M^{\complement}) = m(f^{-n}(M) \cap M^{\complement})$$

para todo $n \in \mathbb{I}_+$, o que implica por (4) que $M \doteq \varnothing$ ou $M^{\complement} \doteq \varnothing$, logo $M \doteq \varnothing$ ou $M \doteq X$, e a ergodicidade está provada. ■

Proposição 4.4. *1. A rotação R_α no círculo é ergódica se, e somente se, α é irracional.*

2. O shift de Bernoulli é ergódico.

3. A expansão no círculo é ergódica.

Demonstração. 1. Se $\alpha \in \mathbb{Q}$, escrevemos $\alpha = \frac{p}{q}$ fração reduzida. Então, tomando um conjunto mensurável $A \subseteq \mathbb{T}$ tal que $) < m(A) < \frac{1}{q}$, segue que o conjunto

$$M := A \cup R_\alpha(A) \cup R_\alpha^2(A) \cap \cdots \cup R_\alpha^{q-1}(A)$$

é invariante sob R_α, pois $R_\alpha^q = id_{\mathbb{T}}$. No entanto, $0 < m(M) < 1$, o que mostra que $\mathbb{R}_\alpha$ não é ergódica.

Se $\alpha \notin \mathbb{Q}$, então a rotação é densa em $\mathbb{T}$. Sendo assim, seja $M \subseteq \mathbb{T}$ invariante sob R_α. Podemos escolher, para todo ε uma função f contínua no toro tal que $\|f - \chi_M\|_1 < \varepsilon$ e da invariância de M segue que $\|f \circ R_\alpha^n - f\|_1 < 2\varepsilon$. Da continuidade de f, segue que, para todo $t \in \mathbb{R}$, $\|f \circ R_t - f\|_1 \leq 2\varepsilon$. Agora, como a medida é invariante por rotação, usando o teorema de Fubini e a desigualdade triangular para integrais temos que

$$\begin{aligned}
\left\| f - \int f(t)dt \right\|_1 &= \int \left| \int (f(x) - f(x+t))dt \right| dx \\
&\leq \int \int |f(x) - f(x+t)|dxdt \\
&\leq 2\varepsilon.
\end{aligned}$$

Mas então

$$\|\chi_M - m(M)\|_1 \leq \|\chi_m - f\|_1 + \left\| f - \int f(t)dt \right\|_1 + \left\| \int f(t)dt - m(M) \right\|_1 < 4\varepsilon.$$

Concluímos, portanto, que χ_M é constante, pois isso vale para qualquer ε. Mas então $m(M)$ é 0 ou 1, logo R_α é ergódica.

2. Seja $N \in \mathcal{M}$ invariante. Para qualquer $\varepsilon > 0$, existe uma união finita de cilindros M tais que $m(M \triangle N) < \varepsilon$, o que implica $|m(M) - m(N)| < \varepsilon$. O conjunto M pode ser escrito como

$$M = \{x \in X : x_{-m} \cdots x_m \in F\}$$

para algum m e $F \subseteq \{0, 1, \cdots, n\}^{\{-m, \cdots, m\}}$. Assim, para todo $k > 2m$, temos que

$$\sigma^{-k}(M) = \{x \in X : x_{k-m} \cdots x_{k+m} \in F\}.$$

Como esse dois conjuntos são independentes, segue que

$$m\left(\sigma^{-k}(M) \setminus M\right) = m\left(\sigma^{-k}(M) \cap M^{\complement}\right) = m\left(\sigma^{-k}(M)\right) m\left(M^{\complement}\right) = m(M)m\left(M^{\complement}\right).$$

Como N é invariante, temos que $m(N \triangle \sigma^{-1}(M))$, e então

$$m(\sigma^{-k}(M) \triangle N) = m(\sigma^{-k}(M) \triangle \sigma^{-k}(N)) = m(M \triangle N) < \varepsilon$$

logo $m(\sigma^{-k}(M) \triangle M) < 2\varepsilon$ e

$$m(\sigma^{-k}(M) \setminus M) + m(M \setminus \sigma^{-k}(M)) = m(\sigma^{-k}(M) \triangle M) < 2\varepsilon.$$

Finalmente, obtemos que

$$\begin{aligned} m(N)m\left(N^{\complement}\right) &< (m(M)+\varepsilon)(m(M^{\complement})+\varepsilon) \\ &= m(M)m(M^{\complement}) + \varepsilon\mu(M) + \varepsilon\mu(M^{\complement}) + \varepsilon^2 \\ &< 2\varepsilon + \varepsilon + \varepsilon^2. \end{aligned}$$

Como isso vale para qualquer ε, segue que $m(N)m\left(N^{\complement}\right) = 0$; ou seja, que $m(N)$ é 0 ou 1, logo o sistema é ergódico.

3. Como os sistemas de Bernoulli Σ e a expansão E_2 são isomorfos por medida, segue que a expansão é ergódica.

■

4.4 Sistemas Misturadores

Nesta seção consideraremos duas noções de um sistema misturador: a de *misturador* e de *fracamente misturador*. Um sistema misturador também é conhecido como fortemente misturador para distinguí-lo dos fracamente misturadores. Existem, de fato, outras noções de misturador, que generalizam a ideia de misturador, como *misturador de ordem k* (cf. [?]), mas aqui consideraremos apenas essas duas. Antes de introduzir a definição, vale ressaltar uma outra definição equivalente de sistema ergódico, enunciada abaixo, cuja demonstração pode ser achada em [?]. Essa classificação é uma consequência do teorema da média ergódica.

Proposição 4.5. *Um sistema $(X, \mathcal{M}, m, T)$ é ergódico se, e somente se, para todos $M, N \in \mathcal{M}$,*

$$\frac{1}{n}\sum_{i=0}^{n-1} m(M \cap T^{-i}(N)) \to m(M)m(N)$$

quando $n \to \infty$.

A definição a seguir, de um sistema misturador, destaca a seguinte ideia. Em alguns sistema, se tomamos dois subconjuntos mensuráveis quaisquer dos espaços, notamos que sempre ocorre que um deles é independente da n-ésima imagem inversa do outro quando n é grande. O sentido de independência empregado é que, em espaços de probabilidade, eventos são pensados como conjuntos mensuráveis e dois eventos E, F sã independentes quando $m(E \cap F) = m(E)m(F)$; ou seja, a probabilidade dos dois ocorrerem é o produto das probabilidades de cada um deles ocorrer separadamente. Pensando em conjuntos, e agora considerando uma ideia geométrica que motiva o nome *misturador*, essa ideia é equivalente a pensar que os conjuntos, quando iterados, estão muito bem separados no sistema, misturados, de modo que se fixamos um deles, o primeiro, a medida das imagens inversas do segundo que estão nesses primeiro conjunto, dividiva pela medida desse primeiro, é igual à medida de do segundo; ou seja, a quantidade do segundo que está no pimeiro é proporcionalmente igual à quantidade desse segundo no espaço todo. A definição a seguir formaliza o que foi aqui explicado.

Definição 4.4. Um sistema *misturador* é um sistema $(X, \mathcal{M}, m, T)$ que preserva medida tal que, para todos $M, N \in \mathcal{M}$,

$$m(M \cap T^{-n}N) \to m(M)m(N)$$

quando $n \to \infty$

Proposição 4.6. *A rotação R_α no círculo não é misturadora.*

Demonstração. Tomando uma sequência $n_i \longrightarrow \infty$ tal que $n_i\alpha \mod 1 \to 0$. No caso de α irracional isso é possível pois a rotação é densa, e no caso de racional, basta tomar uma sequência constante igual a zero. Tomando os conjuntos $M = N = [0, \frac{1}{2}]$, segue que, quando $i \to \infty$,

$$m(M \cap R_\alpha^{-n_i}(N)) = m(M \cap R_{-n_i\alpha}(M)) \to m(M) = \frac{1}{2} \neq \left(\frac{1}{2}\right)^2 = m(M)m(N).$$

■

Embora o conceito de sistema misturador como definido acima é mais natural, a noção levemente modificada abaixo é, muitas vezes, mais útil e simples de se aplicar. Esse é o conceito de fracamente misturador.

Definição 4.5. Um sistema *fracamante misturador* é um sistema $(X, \mathcal{M}, m, T)$ tal que, para todos $M, N \in \mathcal{M}$,

$$\frac{1}{n} \sum_{i=0}^{n-1} \left| m(M \cap T^{-i}(N)) - m(M)m(N) \right| \to 0$$

quando $n \to \infty$

É interessante comentar algo sobre os diferentes tipos de convergência aqui considerados. Além da convegência tradicional, $a_n \to a$, existem outras noções de convergência de uma sequência. Uma delas é convergência de Cesàro, que considera a convergência da sequência $\frac{1}{n} \sum_{i=0}^{n-1} a_n$ das n-ésimas médias da sequência original. Pelas definições e proposições anteriores, considrando a sequência $a_i := \mu(M \cap T^{-i}(N)) - m(M)m(N)$, temos que um sistema é *ergódico* se, e somente se, a_n converge a 0 pela convergência de Cesàro, é *misturador* se, e somente se, a_n converge a 0 no sentido tradicional, e é *fracamente misturador* se $|a_n|$ converge a 0 pela convergência de Cesàro. Assim, como, para qualquer sequência a_n que converge a 0, vale que $|a_n|$ converge a 0 por Cesàro e, então, a_n converge a 0 por Cesàro, segue a proposição abaixo.

Proposição 4.7. *Seja $(X, \mathcal{M}, m, T)$ um sistema que preserva medida. Então*

1. *Se o sistema é misturador, é fracamente misturador.*
2. *Se o sistema é fracamente misturador, é ergódico.*

Demonstração. Basta considerar a sequência $a_i := \mu(M \cap T^{-i}(N)) - m(M)m(N)$ e os comentários do parágrafo anterior. ■

Definição 4.6. A *densidade* de um conjunto $C \subseteq \mathbb{N}$ é o número

$$d(C) \coloneqq \lim_{n \longrightarrow \infty} \frac{1}{n} \left| \{k \in C : 0 \leq k < n\} \right| = \frac{1}{n} \left| (C \cap [0, n) \right|$$

quando o limite existir.

Antes de demonstrar um teorema importante a seguir com várias equivalências envolvendo a propriedade de ergodicidade e mistura fraca, demonstraremos um lema sobre convergência de sequências.

Lema 4.8. *Seja $(a_n)_{n \in \mathbb{N}}$ uma sequência limitada de números reais positivos. São equivalentes as seguintes noções de convergência.*

1. $\lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} a_i = 0.$

2. *Existe um conjunto $C \subseteq \mathbb{N}$ com densidade zero tal que $a_n \to 0$ para $n \to \infty$, $n \notin C$.*

3. $\lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} a_i^2 = 0.$

Demonstração. Vamos mostrar que (1) implica (2). Definamos $C_k \coloneqq \{j \in \mathbb{N} : a_j > \frac{1}{k}\}$. Então temos que

$$C_1 \subseteq C_2 \subseteq C_3 \subseteq \cdots, \tag{2}$$

pois, se $x \in C_k$, então $x > \frac{1}{k} > \frac{1}{k+1}$, logo $x \in C_{k+1}$. Então temos que, para todo $k \geq 1$,

$$\frac{1}{k} |C_k \cap [0, n)| = \sum_{i \in C_k \cap [0,n)} \frac{1}{k} < \sum_{i \in C_k \cap [0,n)} a_i \leq \sum_{i=0}^{n} a_i$$

e então segue que

$$\frac{1}{n} |C_k \cap [0, n)| \leq k \frac{1}{n} \sum_{i=0}^{n} a_i \to 0$$

quando $n \to \infty$. Logo, para todo $k \geq 1$, C_k tem densidade 0. Como cada C_k tem densidade 0, podemos escolher inteiros $0 < l_1 < l_2 < \cdots$ tais que, para todo $k \geq 1$ e para todo $n \geq l_k$,

$$\frac{1}{n} |C_k \cap [0, n)| \leq \frac{1}{k}. \tag{3}$$

Basta escolher l_1 e tomar um l_2 suficientemente grande e assim por diante. Definimos então o conjunto

$$C \coloneqq \bigcup_{k \in \mathbb{N}} \left(C_k \cap [l_k, l_{k+1}) \right). \tag{4}$$

Vamos mostrar que a_i converge a 0 para $n \notin C$ e que C tem densidade 0. Primeiro, notamos que $C_k \cap [l_k, \infty) \subseteq C$, pois, se $x \in C_k$, então está em todo C_p para $p \geq k$, e, então, se estiver em $[l_p, l_{p+1})$, estará em $C_p \cap [l_p, l_{p+1})$, logo em C. Assim, para todo $n \geq l_k$, então se $n \notin C$, temos $n \notin C_k$ e, portanto, $a_n \leq \frac{1}{k}$, logo $a_n \to 0$. Agora, para mostrar que C tem densidade zero, notemos que, para todo $n \in [l + k, l_{k+1})$ temos que

$C \cap [0,n) \subseteq C_k \cap [0,n)$, pois se $x \in C \cap [0,n)$, está em algum $[l_p, l_{p+1})$ e, por consequência, em $C_k \cap [0,n)$, já que todo C_p está contido no C_k para $p \leq k$. Isso implica que

$$\frac{1}{n}|C \cap [0,n)| \leq \frac{1}{n}|C_k \cap [0,n)| \leq \frac{1}{k} \tag{5}$$

e, portanto, C tem densidade zero.

Agora, vamos mostrar que (2) implica (1). Como a sequêcia é limitada, seja $L > 0$ tal que $a_n \leq L$ para todo $n \geq 1$. Como a_n converge para zero fora de C e o conjunto C tem densidade zero, podemos escolher, para cada $k \geq 1$, $n_k \in \mathbb{N}$ tal que, para todo $n \notin C$, $n \geq n_k$, temos que $a_n < \frac{1}{k}$ e, para todo $n \geq n_k$, $\frac{1}{n}|C \cap [0,n)| \leq \frac{1}{k}$. Sendo assim, para todo $n \geq kn_k$, temos que

$$\begin{aligned}
\frac{1}{n}\sum_{i=0}^{n-1} a_i &= \frac{1}{n}\left(\sum_{i=0}^{n_k-1} a_i + \sum_{i \in C \cap [n_k,n)} a_i + \sum_{i \in C^{\complement} \cap [n_k,n)} a_i\right) \\
&< \frac{1}{n}\left(Ln_k + L|C \cap [0,n)| + n\frac{1}{k}\right) \\
&\leq \frac{2L+1}{k},
\end{aligned}$$

o que mostra que a média parcial converge para 0.

Para mostrar a equivalência de (3), basta notar que pela equivalência de (1) e (2), e como $a_n \to 0$ se, e somente se, $a_n^2 \to 0$, pois a sequência é positiva, temos a equivalência de (1) e (3). ■

Teorema 4.9. *Seja $(X, \mathcal{M}, m, T)$ um sistema que preserva medida. São equivalentes*

1. *T é fracamente misturadora.*
2. *$T \times T$ é ergódica com respeito a $m \times m$.*
3. *$T \times T$ é fracamente misturadora com respeito a $m \times m$.*
4. *Para todo sistema ergódico $(Y, \mathcal{Y}, \nu, S)$ que preserva medida, o sistema $(X \times Y, \mathcal{M} \otimes \mathcal{Y}, m \times \nu, T \times S)$ é ergódico.*
5. *Para todos $M, N \in \mathcal{M}$, existe conjunto $C \subseteq \mathbb{N}$ com densidade zero tal que $m(M \cap T^{-n}(N)) \to m(M)m(N)$ quando $n \to \infty$ e $n \notin C$.*
6. *Para todos $M, N \in \mathcal{M}$,*

   $$\frac{1}{n}\sum_{i=0}^{n-1}\left|m(M \cap T^{-i}(N)) - m(M)m(N)\right|^2 \to 0$$

   quando $n \to \infty$.

Demonstração. A equivalência de (1),(5) e (6) é consequência do lema anterior tomando como a_i a sequência $|m(M \cap T^{-i}(N)) - m(M)m(N)|$. A demonstração das demais equivalências seguirá o seguinte esquema: (5) $\Rightarrow$ (3) $\Rightarrow$ (1) $\Rightarrow$ (4) $\Rightarrow$ (2) $\Rightarrow$ (6) $\Rightarrow$ (5).

(5) $\Rightarrow$ (3). Sejam $M_1, M_2, N_1, N_2 \in \mathcal{M}$, existem conjuntos $C_1, C_2 \subseteq \mathbb{N}$ com densidade 0 tais que

$$\begin{aligned} m(M_1 \cap T^{-n}(N_1)) &\to m(M_1)m(N_1) \quad \text{para } n \notin C_1 \\ m(M_2 \cap T^{-n}(N_2)) &\to m(M_2)m(N_2) \quad \text{para } n \notin C_2. \end{aligned}$$

Seja $C := C_1 \cup C_2$. Então C tem densidade zero, pois $C \cap [0,n) = (C_2 \cap [0,n)) \cap (C_2 \cap [0,n))$, o que implica

$$\frac{1}{n}|C \cap [0,n)| \leq \frac{1}{n}(|C_1 \cap [0,n)| + |C_2 \cap [0,n)|) \to 0.$$

Assim, temos que, para todo $n \notin C$,

$$\begin{aligned} &\Big((m \times \mu)(M_1 \times M_2) \cap (T \times T)^{-n}(N_1 \times N_2) - (m \times \mu)(M_1 \times M_2)(m \times \mu)(N_1 \times N_2)\Big) \\ &= \Big(m(M_1 \cap T^{-n}(N_1))m(M_2 \cap T^{-n}(N_2)) - m(M_1)m(M_2)m(N_1)m(N_2)\Big) \to 0. \end{aligned}$$

Isso mostra que $T \times T$ é fracamente misturadora, pois os retângulos geram $\mathcal{M} \otimes \mathcal{M}$.

(3) $\Rightarrow$ (1). Sejam $M, N \in \mathcal{M}$. Então, para $M \times X, N \times X \in \mathcal{M} \otimes \mathcal{M}$ e um conjunto com densidade zero $C \subseteq \mathbb{N}$, e para todo $n \notin C$,

$$\begin{aligned} 0 &= \lim_{n \notin C} \Big((m \times \mu)(M \times X) \cap (T \times T)^{-n}(N \times X) - (m \times \mu)(M \times X)(m \times \mu)(N \times X)\Big) \\ &= \lim_{n \notin C} \Big(m(M \cap T^{-n}(N))m(X \cap T^{-n}(X)) - m(M)m(X)m(N)m(X)\Big) \\ &= \lim_{n \notin C} \Big(m(M \cap T^{-n}(N)) - m(M)m(N)\Big), \end{aligned}$$

pois $m(X) = 1$ e T é invariante, o que implica que T é fracamente misturadora.

(1) $\Rightarrow$ (4). Consideremos um sistema ergódico $(Y, \mathcal{Y}, \nu, S)$ que preserva medida. Sejam $M_1, N_1 \in \mathcal{M}$ e $M_2, N_2 \in \mathcal{Y}$. Então

$$\begin{aligned} &\lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \Big((m \times \nu)(M_1 \times M_2 \cap (T \times S)^{-i}(N_1 \times N_2)\Big) \\ &= \lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} m(M_1 \cap T^{-i}(N_1))\nu(M_2 \cap S^{-i}(N_2)) \\ &= \lim_{n \longrightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} m(M_1)m(M_2)\nu(M_2 \cap S^{-i}(N_2)) \\ &\qquad + \frac{1}{n} \sum_{i=0}^{n-1} \Big(m(M_1 \cap T^{-i}(N_1)) - m(M_1)m(M_2)\Big)\nu(M_2 \cap S^{-i}(N_2)). \end{aligned}$$

Mas a primeira parcela dessa soma final é $m(M_1)m(N_1)\nu(M_1)\nu(N_2)$, pois, pela ergodicidade de ν, temos que $\sum_{i=0}^{n-1} \nu(M_2 \cap S^{-i}(N_2)) \to \nu(M_2)\nu(N_2)$, e a segunda parcela da soma é zero, pois $\nu(M_2 \cap S^{-i})(N_2) \leq 1$ é limitado e T é fracamente misturadora. Como $m(M_1)m(N_1)\nu(M_1)\nu(N_2) = (m \times \nu)(M_1 \times M_2)(m \times \nu)(N_1 \times N_2)$, temos o resultado procurado.

(4) $\Rightarrow$ (2). Consideremos o sistema $(Y, \mathcal{Y}, \nu, S)$, com $Y = \{y\}$ e $S = id_Y$, que é ergódico e preserva medida. Nesse caso, $T \times S$ é isomórfico a T e, então, por (4), T é ergódico. Usando (4) novamente, e tomando $Y = X$, temos que $T \times T$ é ergódico.

(2) $\Rightarrow$ (6). Sejam $M, N \in \mathcal{M}$. Pela ergodicidade de $T \times T$, temos que

$$\begin{aligned}
&\frac{1}{n}\sum_{i=0}^{n-1} m(M \cap T^{-i}(N)) \\
&= \frac{1}{n}\sum_{i=0}^{n-1} \Big((m \times \mu)(M \times X \cap (T \times T)^{-i}(N \times X))\Big) \\
&\to (m \times \mu)(M \times X)(m \times \mu)(N \times X) = m(M)m(N)
\end{aligned}$$

e, analogamente,

$$\frac{1}{n}\sum_{i=0}^{n-1} m(M \times T^{-i}(N))^2 \to m(M)^2\mu(N)^2.$$

Disso, segue que

$$\begin{aligned}
&\frac{1}{n}\sum_{i=0}^{n-1} \Big(m(M \cap T^{-i}(N)) - m(M)m(N)\Big)^2 \\
&= \frac{1}{n}\sum_{i=0}^{n-1} m(M \cap T^{-i}(N))^2 - 2m(M)m(N)\frac{1}{n}\sum_{i=0}^{n-1} m(M \cap T^{-i}(N)) + m(M)^m(N)^2\frac{1}{n}\sum_{i=0}^{n-1} 1 \\
&\to m(M)^2\mu(N)^2 - 2\mu(M)m(N)m(M)m(N) + m(M)^2\mu(N)^2 = 0.
\end{aligned}$$

Isso completa a demostração do teorema, pois já sabemos que (6) $\Rightarrow$ (5). ■

4.5 Entropia

Definição 4.7. Seja $n \in \mathbb{N}$. O simplexo n-dimensional é o conjunto

$$\Delta^n := \left\{ p \in (\mathbb{R}_+)^{n+1} \;\middle|\; \sum_{i=0}^{n} p_i = 1 \right\}.$$

Proposição 4.10. *Seja $E : \bigcup_{n \in \mathbb{N}} \Delta^n \longrightarrow \mathbb{R}$ uma função que satisfaz*

1. *$E(p_0, \ldots, p_n) \geq 0$ para todo $(p_0, \ldots, p_n) \in \Delta^n$;*
2. *$E(p_0, \ldots, p_n) = 0$ se, e somente se, existe i tal que $p_i = 1$;*
3. *Para todo $n \in \mathbb{N}^*$,*
$$\max\Big(E|_{\Delta^{n-1}}(\Delta^{n-1})\Big) = E\left(\frac{1}{n}, \ldots, \frac{1}{n}\right);$$
4. *Para todo $(p_0, \ldots, p_n) \in \Delta^n$,*
$$E(p_0, \ldots, p_n) = E(p_0, \ldots, p_n, 0);$$
5. *Para todo $(q_{0,0}, \ldots, q_{0,k-1}, \ldots, q_{n-1,0}, \ldots, q_{n-1,k-1}) \in \Delta^{nk-1}$, seja $p_i := \sum_{j=0}^{k-1} q_{i,j}$; então*
$$\begin{aligned}
&E(q_{0,0}, \ldots, q_{0,k-1}, \ldots, q_{n-1,0}, \ldots, q_{n-1,k-1}) \\
&\qquad = E(p_0, \ldots, p_{n-1}) + \sum_{i=0}^{n-1} p_i E\left(\frac{q_{i,0}}{p_i}, \ldots, \frac{q_{i,k-1}}{p_i}\right).
\end{aligned}$$

6. Para todo $n \in \mathbb{N}$, $E|_{\Delta^n}$ é simétrica e contínua.

Então existe $\lambda \geq 0$ tal que (assumindo $0 \log 0 = 0$)

$$E(p_0, \ldots, p_n) = -\lambda \sum_{i=0}^{n} p_i \log p_i.$$

Demonstração. Definamos, para cada $n \in \mathbb{N}^*$ e $(\frac{1}{n}, \ldots, \frac{1}{n}) \in \Delta^{n-1}$,

$$L(n) := E\left(\frac{1}{n}, \ldots, \frac{1}{n}\right).$$

Mostraremos que $L(n) = \lambda \log n$ para algum real $\lambda \geq 0$. Pelas propriedades 4 e 3,

$$L(n) = E\left(\frac{1}{n}, \ldots, \frac{1}{n}, 0\right) \leq E\left(\frac{1}{n+1}, \ldots, \frac{1}{n+1}\right) = L(n+1),$$

o que mostra que L é uma função crescente. Sejam $n \in \mathbb{N}$ e $k \in \mathbb{N}^*$. Vamos mostrar que $L(n^k) = kL(n)$ por indução em k. Tomando $\left(\frac{1}{n^k}, \ldots, \frac{1}{n^k}\right) \in \Delta^{n^k - 1}$ e notando que $\frac{1}{n^{k-1}} = \sum_{j=0}^{n-1} \frac{1}{n^k}$, segue da propriedade 5 que

$$\begin{aligned}
E\left(\frac{1}{n^k}, \ldots, \frac{1}{n^k}\right) &= E\left(\frac{1}{n^{k-1}}, \ldots, \frac{1}{n^{k-1}}\right) + \sum_{i=0}^{n^{k-1}-1} \frac{1}{n^{k-1}} E\left(\frac{\frac{1}{n^k}}{\frac{1}{n^{k-1}}}, \ldots, \frac{\frac{1}{n^k}}{\frac{1}{n^{k-1}}}\right) \\
&= E\left(\frac{1}{n^{k-1}}, \ldots, \frac{1}{n^{k-1}}\right) + \sum_{i=0}^{n^{k-1}-1} \frac{1}{n^{k-1}} E\left(\frac{1}{n}, \ldots, \frac{1}{n}\right) \\
&= E\left(\frac{1}{n^{k-1}}, \ldots, \frac{1}{n^{k-1}}\right) + E\left(\frac{1}{n}, \ldots, \frac{1}{n}\right).
\end{aligned}$$

Isso mostra que $L(n^k) = L(n^{k-1}) + L(n)$ (Em particular, segue que $E(1) = 0$ tomando $k = 1$). Assim, por indução em k, a relação vale claramente vale para $k = 1$ e, supondo por indução que $L(n^{k-1}) = (k-1)L(n)$, segue que

$$L(n^k) = L(n^{k-1}) + L(n) = (k-1)L(n) + L(n) = kL(n).$$

Agora, sejam $n_0, n_1, k_0 \in \mathbb{N}^*$ e tomemos $k_1 \in \mathbb{N}^*$ tal que

$${n_1}^{k_1} \leq {n_0}^{k_0} < {n_1}^{k_1+1}.$$

Então segue que $k_1 \log n_1 \leq k_0 \log n_0 < (k_1 + 1) \log n_1$, portanto

$$\frac{k_1}{k_0} \leq \frac{\log n_0}{\log n_1} < \frac{k_1}{k_0} + \frac{1}{k_0}.$$

Segue também, pela monotonicidade de L, que

$$L({n_1}^{k_1}) \leq L({n_0}^{k_0}) \leq L({n_1}^{k_1+1}),$$

logo $k_1 L(n_1) \leq k_0 L(n_0) \leq (k_1 + 1)L(n_1)$; portanto

$$\frac{k_1}{k_0} \leq \frac{L(n_0)}{L(n_1)} < \frac{k_1}{k_0} + \frac{1}{k_0}.$$

Das desigualdades das frações, obtemos

$$\left|\frac{L(n_0)}{L(n_1)} - \frac{\log n_0}{\log n_1}\right| \leq \frac{1}{k_0}.$$

Como k_0 pode ser tomado arbitrariamente grande, segue que

$$\frac{L(n_0)}{L(n_1)} = \frac{\log n_0}{\log n_1},$$

e da arbitrariedade de n_0 e n_1, segue que $L(n) = \lambda \log n$ para uma constante $\lambda \in \mathbb{R}$. Da monotonicidade de L, temos $\lambda \geq 0$.

Consideremos agora um vetor de probabilidade $(p_0, \dots, p_{n-1}) \in \Delta^{n-1}$ cujas entradas são números racionais. Seja $p_i = \frac{g_i}{k}$, em que $k = \mathop{+}_{i=0}^{n-1} g_i$. Escolhemos $(q_{i,j}) \in \Delta^{nk-1}$ tal que

$$q_{i,j} := \begin{cases} \frac{1}{k}, & (\mathop{+}_{l=0}^{i-1} g_l) \leq j < (\mathop{+}_{l=0}^{i} g_l) \\ 0, & \text{caso contrário.} \end{cases}$$

ou seja,

$$\begin{aligned}
(q_{0,0}, \dots, q_{0,k-1}) &= \left(\frac{1}{k}, \dots, \frac{1}{k}, 0, \dots, 0\right) \\
&\vdots \\
(q_{i,0}, \dots, q_{i,k-1}) &= \left(\underbrace{0, \dots, 0}_{g_0 + \dots + g_{i-1}}, \underbrace{\frac{1}{k}, \dots, \frac{1}{k}}_{g_i}, \underbrace{0, \dots, 0}_{g_{i+1} + \dots + g_{n-1}}\right) \\
&\vdots \\
(q_{n-1,0}, \dots, q_{n-1,k-1}) &= \left(0, \dots, 0, \frac{1}{k}, \dots, \frac{1}{k}\right).
\end{aligned}$$

Observe que k das entradas de $(q_{i,j})$ são $\frac{1}{k}$ e as outras $(n-1)k$ são 0. Da simetria de E e da propriedade de que uma entrada 0 não muda o valor de E, temos portanto que

$$E(q_{0,0}, \dots, q_{0,k-1}, \dots, q_{n-1,0}, \dots, q_{n-1,k-1}) = E\left(\frac{1}{k}, \dots, \frac{1}{k}\right).$$

Como $\mathop{+}_{j=0}^{k-1} q_{i,j} = g_i \frac{1}{k} = p_i$, das mesmas propriedades segue que

$$\begin{aligned}
E\left(\frac{q_{i,0}}{p_i}, \dots, \frac{q_{i,k-1}}{p_i}\right) &= E\left(\underbrace{0, \dots, 0}_{g_0 + \dots + g_{i-1}}, \underbrace{\frac{\frac{1}{k}}{p_i}, \dots, \frac{\frac{1}{k}}{p_i}}_{g_i}, \underbrace{0, \dots, 0}_{g_{i+1} + \dots + g_{n-1}}\right) \\
&= E\left(\underbrace{0, \dots, 0}_{g_0 + \dots + g_{i-1}}, \underbrace{\frac{1}{g_i}, \dots, \frac{1}{g_i}}_{g_i}, \underbrace{0, \dots, 0}_{g_{i+1} + \dots + g_{n-1}}\right) \\
&= E\left(\frac{1}{g_i}, \dots, \frac{1}{g_i}\right)
\end{aligned}$$

Segue então, da propriedade 5 e da fórmula para $L(n)$, que

$$\begin{aligned} E\left(\frac{1}{k},\ldots,\frac{1}{k}\right) &= E\left(p_0,\ldots,p_{n-1}\right) + \sum_{i=0}^{n-1} p_i E\left(\frac{1}{g_i},\ldots,\frac{1}{g_i}\right) \\ &= E\left(p_0,\ldots,p_{n-1}\right) + \sum_{i=0}^{n-1} p_i \lambda \log g_i \\ &= E\left(p_0,\ldots,p_{n-1}\right) + \lambda \sum_{i=0}^{n-1} p_i \log p_i + \lambda \log k. \end{aligned}$$

Como $E\left(\frac{1}{k},\ldots,\frac{1}{k}\right) = \lambda \log k$, segue que

$$E\left(p_0,\ldots,p_{n-1}\right) = -\lambda \sum_{i=0}^{n-1} p_i \log p_i.$$

Pela continuidade da função, segue finalmente que essa fórmula vale para qualquer vetor de probabilidade real. ■

Essa é basicamente a demonstração apresentada por Khinchin, com algumas modificações para que a propriedade 5 não dependa da definição de um esquema ou partição, e com adaptações de enunciado e notação adotadas por Peter Walters.

Definição 4.8. Seja $\boldsymbol{X}$ um sistema dinâmico de medida e $\mathcal{P} = (P_i)_{i=0}^{n}$ uma partição finita de X. A *entropia* de $\mathcal{P}$ é o número

$$H(\mathcal{P}) := H(m(P_0),\ldots,m(P_n)) = -\sum_{i=0}^{n} m(P_i)\log m(P_i).$$

5 Exemplos

5.1 Deslocamento Simbólico (de Bernoulli)

Vou adotar a convenção de que o natural $N \in \mathbb{N}$ é o conjunto dos naturais menores que ele, seguindo a construção de ordinais do Von Neumann. Isso simplificará a notação. Assim temos que $N = \{0,\cdots,N-1\}$. Consideremos o conjunto $N^{\mathbb{I}}$ das funções

$$\begin{aligned} x\colon \mathbb{I} &\longrightarrow N \\ i &\longmapsto x_i. \end{aligned}$$

Denotaremos essas funções por $(x_i)_{i\in\mathbb{I}}$ e as chamaremos de *sequências bilaterais* e o elemento $x_i \in n$ será chamado de i-ésima entrada da sequência. Desse modo, o conjunto $N^{\mathbb{I}}$ é o conjunto de todas sequências bilaterias cujas entradas são números menores que N. Definimos no conjunto N a σ-álgebra discreta em que todo conjunto $\{i\}$ com $i \in N$ é mensurável. Ainda, a partir de uma N-upla $p = (p_0,\cdots,p_{N-1}) \in \mathbb{R}_+^N$ tal que $\sum_{i\in N} p_i = 1$, definimos a medida m na σ-álgebra de N como $m(\{i\}) = p_i$. Essa medida é uma probabilidade em N. A partir dessa probabilidade, definimos uma probabilidade em $N^{\mathbb{I}}$ da seguinte forma. Consiraremos conjuntos chamados de *cilindros* em $N^{\mathbb{I}}$. Esses conjuntos são formados a partir de uma função $r : F \subseteq \mathbb{I} \longrightarrow N$, em que F é um subconjunto finito de $\mathbb{I}$. Então definimos o cilindro

$$C(r) := \left\{ (x_i)_{i\in\mathbb{I}} \in N^{\mathbb{I}} \,\middle|\, \forall i \in F, \quad x_i = r(i) \right\}.$$

Vale notar que esses conjuntos também podem ser escritos como a interseção finita

$$C(r) = \bigcap_{i\in F} \pi_i^{-1}(\{r(i)\}),$$

em que π_i é a projeção canônica

$$\begin{aligned} \pi_i \colon N^{\mathbb{I}} &\longrightarrow N \\ (x_i)_{i\in\mathbb{I}} &\longmapsto x_i. \end{aligned}$$

5.2 Rotação no Círculo

Proposição 5.1. *1. A rotação R_α em $\mathbb{T}^1$ preserva a medida de Lesbegue.*

2. A expansão E_m em $\mathbb{T}^1$ preserva a medida de Lesbegue.

Demonstração. 1. É suficiente mostrar que a medida é preservada para intervalos. Mas isso é simples, pois dado $[a, b] \in \mathbb{T}$, como $R_\alpha^{-1}([a, b]) = [a - \alpha, b - \alpha]$, segue que

$$m(R_\alpha^{-1}([a, b])) = (b - \alpha) - (a - \alpha) = b - a = m([a, b]).$$

2. Como no caso anterior, é suficiente mostrar para intervalos. Dado $[a, b] \in \mathbb{T}$, como $E_m^{-1}([a, b]) = \bigsqcup_{i=0}^{m-1} [\frac{a+i}{m}, \frac{b+i}{m}]$, segue que

$$m(E_m^{-1}([a, b])) = \sum_{i=0}^{m-1} \frac{b+i}{m} - \frac{a+i}{m} = m\frac{b-a}{m} = m([a, b]).$$

■

Proposição 5.2. *Seja o sistema de Bernoulli $(\Sigma, \mathcal{B}, m, \sigma)$, em que $\Sigma := \{0, 1\}^{\mathbb{N}}$, $\mathcal{B}$ é a σ-álgebra de Borel de Σ, σ é o shift de Bernoullie m é a medida produto $m := \prod_{\mathbb{N}} \nu$, em que ν é a medida em $\{0, 1\}$ definida por $\nu(\{0\}) = \nu(\{1\}) = \frac{1}{2}$. Então o mapa*

$$\begin{aligned} \phi : \Sigma &\longrightarrow \mathbb{T} \\ (x_n)_{n\in\mathbb{N}} &\longmapsto \sum_{n\in\mathbb{N}} \frac{x_n}{2^{n+1}} \end{aligned}$$

é um isomorfismo entre o sistema de Bernoulli e a expansão E_2 no círculo.

Demonstração. Primeiro notemos que o sistema de Bernoulli deve preservar medida pelo comentário anterior. Assim, notemos que, para $(x_n)_{n\in N}$,

$$\phi \circ \sigma((x_n)) = \phi((x_{n+1})) = \sum_{n\in\mathbb{N}} \frac{x_{n+1}}{2^{n+1}} = 2\sum_{n\in\mathbb{N}} \frac{x_{n+1}}{2^{n+2}} = E_2 \circ \phi(x_n)).$$

A inversa de ϕ é definida a menos de o conjunto de racionais diádicos, que tem medida ula, e então segue o resultado. ■

6 Folheações

Definição 6.1. Seja $\boldsymbol{V}$ uma $\mathscr{C}^r$-variedade d-dimensional. Uma $\mathscr{C}^s$-*folheação* k-dimensional de $\boldsymbol{V}$, com $s \leq r$ e $k \leq d$, é um $\mathscr{C}^s$-atlas $\mathcal{F}$ de $\boldsymbol{V}$ que é maximal segundo as propriedades:

1. Para toda carta $(A, \varphi) \in \mathcal{F}$, existem discos abertos $D_1 \subseteq \mathbb{R}^k$ e $D_2 \subseteq \mathbb{R}^{d-k}$ tais que
$$\varphi(A) = D_1 \times D_2;$$

2. Para todas cartas (A_1, φ_1) e $(A_2, \varphi_2) \in \mathcal{F}$, a transição de coordenadas
$$\varphi_2 \circ {\varphi_1}^{-1} : \varphi_1(A_1 \cap A_2) \longrightarrow \varphi_2(A_1 \cap A_2)$$
é da forma
$$\varphi_2 \circ {\varphi_1}^{-1}(x, y) = (h_1(x, y), h_2(y)) \in \mathbb{R}^k \times \mathbb{R}^{d-k}$$
para todo $(x, y) \in \varphi_1(A_1 \cap A_2)$, com $x \in \mathbb{R}^k$ e $y \in \mathbb{R}^{d-k}$.

Nesse caso, a variedade $\boldsymbol{V}$ é *folheada* por $\mathcal{F}$ e as cartas de $\mathcal{F}$ são *cartas de folheação* (ou *cartas trivializadoras*).

Definição 6.2. Sejam $\boldsymbol{V}$ uma $\mathscr{C}^r$-variedade d-dimensional, $\mathcal{F}$ uma $\mathscr{C}^s$-folheação k-dimensional de $\boldsymbol{V}$ e $(A, \varphi) \in \mathcal{F}$ uma carta de folheação. Um *placa* de $\mathcal{F}$ baseada em (A, φ) é um conjunto da forma
$$\varphi^{-1}(D_1 \times \{c\}),$$
em que $\varphi(A) = D_1 \times D_2$ e $c \in D_2$.

Proposição 6.1. *1. A função*
$$\varphi^{-1}|_{D_1 \times \{c\}} : D_1 \times \{c\} \longrightarrow A$$
é um mergulho $\mathscr{C}^r$. Portanto as placas são $\mathscr{C}^r$-subvariedades k-dimensionais.

2. As placas baseadas em (A, φ) são disjuntas duas a duas.

Demonstração. 1.

2. Sejam α e β placas de $\mathcal{F}$.

■

A Sistemas Dinâmicos

A.1 Notação de Conjuntos Numéricos

O conjunto dos *números inteiros* será denotado por '$\mathbb{I}$' e o conjunto dos *números reais* será denotado por '$\mathbb{R}$'. O superíndice '$*$' indica que o conjunto não inclui o número 0; o subíndice '$+$' indica que os somente os números positivos estão excluídos do conjunto, inclusive o 0; o subíndice '$-$' indica que somente os números negativos são mantidos no conjunto, inclusive o 0.

A.2 Partição de Conjuntos

Definição A.1. Seja X um conjunto. Uma *partição* de X é um conjunto $\mathcal{P} \subseteq \wp(X)$ de subconjuntos de X que satisfaz

1. $\varnothing \notin \mathcal{P}$;
2. $\bigcup \mathcal{P} = X$;
3. Para todos conjuntos distintos $C_0, C_1 \in \mathcal{P}$, $C_0 \cap C_1 = \varnothing$.

Os conjuntos $C \in \mathcal{P}$ são as *células* de P.

Uma partição, se identidicamos um subconjunto de $\wp(X)$ com uma família de suncon-juntos de X, é uma cobertura de X por conjuntos disjuntos dois a dois (logo distintos) que não contém o conjunto vazio.

Definição A.2. Sejam X um conjunto e $\mathcal{P}$ uma partição de X. Um *refinamento* (*superpartição*) de $\mathcal{P}$ é uma partição $\mathcal{R}$ de X que satisfaz: para toda célula $D \in \mathcal{R}$, existe uma célula $C \in \mathcal{P}$ tal que $D \subseteq C$. Denota-se $\mathcal{P} \leq \mathcal{R}$. Diz que $\mathcal{P}$ é um *engrossamento* (*subpartição*) de $\mathcal{R}$.

Definição A.3. Sejam X um conjunto e $(\mathcal{P}_i)_{i \in I}$ uma família de partições de X. O *refinamento comum* a $(\mathcal{P}_i)_{i \in I}$ é o conjunto

$$\bigvee_{i \in I} \mathcal{P}_i := \left\{ \bigcap_{i \in I} C_i \;\middle|\; i \in I,\ C_i \in \mathcal{P}_i \ \text{ e } \ \bigcap_{i \in I} C_i \neq \varnothing \right\}.$$

Proposição A.1. *Sejam X um conjunto e $(\mathcal{P}_i)_{i \in I}$ uma família de partições de X. O refinamento comum $\bigvee_{i \in I} P_i$ a $(\mathcal{P}_i)_{i \in I}$ é a menor partição de que X que refina $\mathcal{P}_i$ para todo $i \in I$.*

A.3 Ações de Monoides

Definição A.4. Sejam $\boldsymbol{M}$ um monoide e X um conjunto. Uma *ação* de $\boldsymbol{M}$ em X é um homomorfismo de monoides

$$\begin{aligned} A\colon \boldsymbol{M} &\longrightarrow \boldsymbol{X}^{\boldsymbol{X}} \\ m &\longmapsto A^m. \end{aligned}$$

Denota-se $A : \boldsymbol{M} \curvearrowright X$. Diz-se que o monoide $\boldsymbol{M}$ *age* no conjunto X, e denota-se $\boldsymbol{M} \curvearrowright X$, se, e somente se, existe ação de $\boldsymbol{M}$ em X.

A definição acima é uma definição muito simplificada pois ela depende do conceito um pouco mais complexo de homomorfismo de monoides. Essa definição é equivalente a dizer que existe uma função

$$\begin{aligned} A\colon M \times X &\longrightarrow X \\ (m, x) &\longmapsto A^m(x) \end{aligned}$$

que satisfaz

1. (Identidade) Para todo $x \in X$,
$$A^e(x) = x;$$
2. (Compatibilidade) Para todos $m_0, m_1 \in M$ e $x \in X$,
$$A^{m_1 * m_0}(x) = A^{m_1}(A^{m_0}(x)).$$

Essa é geralmente considerada a definição de uma ação *à esquerda* de $\boldsymbol{M}$ em X. Analogamente, uma ação à direita pode ser definida. Quando o monoide é um grupo, é comum estudar ações em conjuntos X com alguma estrutura adicional, geralmente uma topologia. Nesse caso, exige-se que a ação seja uma função contínua, mas também é necessário que o grupo tenha uma estrutura topológica, portanto isso não será definido com cuidado agora.

B Sistemas Dinâmicos de Medida

B.1 Espaços Mensuráveis e de Medida

B.1.1 Morfismos Mensuráveis e de Medida

Definição B.1. Sejam $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ espaços mensuráveis. Uma *função mensurável* (*morfismo mensurável*) de $\boldsymbol{X}_1$ para $\boldsymbol{X}_2$ é uma função $f : X_1 \longrightarrow X_2$ que satisfaz, para todo $M \in \mathcal{M}_2$,

$$f^{-1}(M) \in \mathcal{M}_1.$$

Denota-se $f : \boldsymbol{X}_1 \longrightarrow \boldsymbol{X}_2$.

Definição B.2. Sejam $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ espaços de medida. Um *morfismo de medida* (*função que preserva medida*) é uma função mensurável $f : X_1 \longrightarrow X_2$ que satisfaz, para todo $M \in \mathcal{M}_2$,

$$m_1(f^{-1}(M)) = m_2(M).$$

Denota-se $f : \boldsymbol{X}_1 \longrightarrow \boldsymbol{X}_2$.

B.1.2 Medidas Empurradas e Puxadas

Definição B.3. Sejam $(X, \mathcal{M}_X, m)$ um espaço de medida, $(Y, \mathcal{M}_Y)$ um espaço mensurável e $f : (X, \mathcal{M}_X) \longrightarrow (Y, \mathcal{M}_Y)$ uma função mensurável. A *medida empurrada* de m por f é a função

$$\begin{aligned} m_\star\colon \mathcal{M}_Y &\longrightarrow \mathbb{R}_\infty \\ M &\longmapsto m(f^{-1}(M)), \end{aligned}$$

Se é necessário ressaltar a função f, denota-se $m_\star = f_\star(m)$.

Essa definição significa que uma função preservar medida é equivalente a dizer que ela empurra a medida de um sistema na do outro, que a medida do segundo é a medida empurrada do primeiro:

$$m_{1\star} = f_\star(m_1) = m_2.$$

Definição B.4. Sejam $(X, \mathcal{M}_X)$ um espaço mensurável, $(Y, \mathcal{M}_Y, m)$ um espaço de medida e $f : (X, \mathcal{M}_X) \longrightarrow (Y, \mathcal{M}_Y)$ uma função mensurável. A *medida puxada* de m por f é a função

$$\begin{aligned} m^\star \colon \mathcal{M}_X &\longrightarrow \mathbb{R}_\infty \\ M &\longmapsto m(f(M)), \end{aligned}$$

Se é necessário ressaltar a função f, denota-se $m^\star = f^\star(m)$.

B.2 Medidas em Espaços/Grupos Topológicos

Definição B.5. Seja $\boldsymbol{X}$ um espaço topológico. A *σ-álgebra topológica* de $\boldsymbol{X}$ é a σ-álgebra gerada por $\mathcal{T}$

$$\mathcal{M}_{\boldsymbol{X}} := \langle \mathcal{T} \rangle .$$

Definição B.6. Um *espaço topológico de medida* é um par $(\boldsymbol{X}, m)$ em que $\boldsymbol{X}$ é um espaço topológico e $(X, \mathcal{M}_{\boldsymbol{X}}, m)$ é um espaço de medida.

Definição B.7. Uma ação de grupo *que preserva medida* em $\boldsymbol{X}$ é uma ação $f : G \times X \longrightarrow X$ que é mensurável com respeito ao espaço de medida produto $\boldsymbol{G} \times \boldsymbol{X}$. Denota-se $f : \boldsymbol{G} \curvearrowright \boldsymbol{X}$.

B.3 Quase

B.3.1 Quase Todo

Definição B.8. Seja $\boldsymbol{X}$ um espaço de medida. Uma propriedade P de elementos de X vale para *quase todo* ponto se existe um conjunto $M \in \mathcal{M}$ com $m(M) = 0$ tal que, para todo $x \in X \setminus M$, vale a propriedade P:

$$\mathring{\forall} x \mathrm{P} x \equiv \exists M \forall x (M \in \mathcal{M} \wedge m(M) = 0 \wedge (x \in X \setminus M \to \mathrm{P} x)).$$

A partir desse quantificador, podemos definir o quantificar *quase existe* da seguinte forma

$$\mathring{\exists} x \mathrm{P} x \equiv \neg \mathring{\forall} x \neg \mathrm{P} x$$

$$\begin{aligned} \neg \mathring{\forall} x \neg \mathrm{P} x &\equiv \neg \exists M \forall x (M \in \mathcal{M} \wedge m(M) = 0 \wedge (x \in X \setminus M \to \neg \mathrm{P} x)) \\ &\equiv \forall M \exists x (M \notin \mathcal{M} \vee m(M) > 0 \vee (x \in X \setminus M \wedge \mathrm{P} x)) \\ &\equiv \forall M \exists x (\neg (M \notin \mathcal{M} \vee m(M) > 0) \to (x \in X \setminus M \wedge \mathrm{P} x)) \\ &\equiv \forall M \exists x (M \in \mathcal{M} \wedge m(M) = 0 \to (x \in X \setminus M \wedge \mathrm{P} x)) \end{aligned}$$

pois

$$\neg (A \to B) \equiv A \wedge \neg B$$

$$A \vee B \equiv \neg A \to B$$

Proposição B.1. *Sejam $\boldsymbol{X}$ um espaço de medida e* P *e* Q *propriedades de elementos de* X*. Então*

1. $\overset{\circ}{\forall} x \mathrm{P}x \wedge \overset{\circ}{\forall} x \mathrm{Q}x \to \overset{\circ}{\forall} x(\mathrm{P}x \wedge \mathrm{Q}x)$*;*

2. $\forall x \mathrm{P}x \to \overset{\circ}{\forall} x \mathrm{P}x$*;*

3. $\overset{\circ}{\exists} x \mathrm{P}x \to \exists x \mathrm{P}x$*;*

4. $\neg \overset{\circ}{\exists} x \neg \mathrm{P}x \equiv \overset{\circ}{\forall} x \mathrm{P}x$*;*

Demonstração. 1. Sejam $M, N \in \mathcal{M}$ os conjuntos de medida 0 em cujos complementares P e Q falham, respectivamente. Então $M \cup N \in \mathcal{M}$ tem medida 0, portanto P e Q falham no complementar de $M \cup N$ e segue o teorema.

2. Tomando $M = \varnothing$, temos $X \setminus M = X$ e concluímo que para todo $x \in X \setminus M$ vale $\mathrm{P}x$.

3. Tomando $M = \varnothing$, temos $X \setminus M = X$ e concluímos que existe $x \in X$ tal que $\mathrm{P}x$.

4.

■

B.3.2 Quase Igualdade

Definição B.9. Sejam $\boldsymbol{X} = (X, \mathcal{M}, m)$ um espaço de medida e $A \in \mathcal{M}$. Um conjunto *quase contido* em A em $\boldsymbol{X}$ é um conjunto $B \in \mathcal{M}$ tal que $m(B \setminus A) = 0$. Denota-se $B \overset{\circ}{\subseteq} A$.

Um conjunto *quase igual* a A em $\boldsymbol{X}$ é um conjunto $B \in \mathcal{M}$ tal que $B \overset{\circ}{\subseteq} A$ e $A \overset{\circ}{\subseteq} B$. Denota-se $B \overset{\circ}{=} A$. Conjuntos *quase vazios* em $\boldsymbol{X}$ são conjuntos quase iguais ao conjunto vazio ($A \overset{\circ}{=} \varnothing$) e conjuntos *quase totais* em $\boldsymbol{X}$ são conjuntos quase iguais a X ($A \overset{\circ}{=} X$).

Proposição B.2. *Seja $\boldsymbol{X}$ um espaço de medida.*

1. *A relação $\overset{\circ}{\subseteq}$ em $\mathcal{M}$ é uma relação de ordem parcial (com respeito a $\overset{\circ}{=}$);*

2. *A relação $\overset{\circ}{=}$ em $\mathcal{M}$ é uma relação de equivalência.*

Demonstração. 1. (Reflexividade) Seja $A \in \mathcal{M}$. Então $A \overset{\circ}{\subseteq} A$, pois $m(A \setminus A) = m(\varnothing) = 0$. (Antissimetria) Sejam $A, B \in \mathcal{M}$ tais que $A \overset{\circ}{\subseteq} B$ e $B \overset{\circ}{\subseteq} A$. Então $A \overset{\circ}{=} B$ por definição. (Transitividade) Sejam $A, B, C \in \mathcal{M}$ tais que $A \overset{\circ}{\subseteq} B$ e $B \overset{\circ}{\subseteq} C$. Então $m(A \setminus B) = 0$ e $m(B \setminus C) = 0$. Mas $A \setminus C = (A \cap B \setminus C) \cup (A \setminus (B \cup C))$; como $(A \cap B \setminus C) \subseteq B \setminus C$ e $(A \setminus (B \cup C)) \subseteq A \setminus B$, segue que

$$m(A \setminus C) = m(A \cap B \setminus C) + m(A \setminus (B \cup C)) \leq 0,$$

logo $A \overset{\circ}{\subseteq} C$;

2. (Reflexividade) Seja $A \in \mathcal{M}$. Então $A \doteq A$, pois $A \overset{\circ}{\subseteq} A$. (Simetria) Sejam $A, B \in \mathcal{M}$ tais que $A \doteq B$. Então $A \overset{\circ}{\subseteq} B$ e $B \overset{\circ}{\subseteq} A$, portanto $B \doteq A$. (Transitividade) Sejam $A, B, C \in \mathcal{M}$ tais que $A \doteq B$ e $B \doteq C$. Então $A \overset{\circ}{\subseteq} B$ e $B \overset{\circ}{\subseteq} C$, portanto $A \overset{\circ}{\subseteq} C$. Analogamente, $C \overset{\circ}{\subseteq} A$, portanto $A \doteq C$.

■

Proposição B.3. *Sejam $\boldsymbol{X}$ um espaço de medida e $A, B \in \mathcal{M}$.*

1. *$A \doteq B$ se, e somente se, $m(A \triangle B) = 0$;*

2. *Se $A \doteq B$, então $m(A) = m(B)$;*

3. *Se $A \subseteq B$, então $A \doteq B$ implica $m(A) = m(B)$; se $A \subseteq B$ e m é finita, então $m(A) = m(B)$ implica $A \doteq B$;*

4. *$A \doteq \varnothing$ se, e somente se, $m(A) = 0$;*

5. *$A \doteq X$ se, e somente se, $m(A^{\complement}) = 0$. Se m é finita, $A \doteq X$ se, e somente se, $m(A) = m(X)$;*

6. *$A \doteq \varnothing$ se, e somente se, $A^{\complement} \doteq X$;*

7. *Se $A \doteq X$ e $B \doteq X$, então $A \cap B \doteq X$.*

8. *Se $A \doteq X$ e m é finita, então, para todo $B \subseteq A$, vale $m(B) = m(A \cap B)$;*

Demonstração. 1. Como $A \setminus B$ e $B \setminus A$ são disjuntos, temos $m(A \triangle B) = m(A \setminus B) + m(B \setminus A)$, portanto $m(A \setminus B) = m(B \setminus A) = 0$ se, e somente se, $m(A \triangle B) = 0$;

2. Como $A \setminus B$ e $B \setminus A$ são disjuntos, temos $m(A \triangle B) = m(A \setminus B) + m(B \setminus A)$, portanto $m(A \setminus B) = m(B \setminus A) = 0$. Como $A = (A \cap B) \cup (A \setminus B)$ e $A \cap B$ e $A \setminus B$ são disjuntos, $m(A) = m(A \cap B) + m(A \setminus B) = m(A \cap B)$. Analogamente $m(B) = m(B \cap A)$, portanto $m(A) = m(B)$.

3. Como $A \subseteq B$, então $A \cup B = A$ e $A \cap B = B$, logo de $m(A \triangle B) = 0$ segue

$$m(A) = m(A \cup B) = m(A \cap B) + m(A \triangle B) = m(B).$$

Se m é finita, da mesma igualdade segue que

$$m(A \triangle B) = m(A \cup B) - m(A \cap B) = m(A) - m(B) = 0.$$

4. A ida segue do item acima. Como $A \cup \varnothing = A$ e $A \cap \varnothing = \varnothing$, segue que $A \triangle \varnothing = A$, logo $m(A \triangle \varnothing) = m(A)$. Assim segue de $m(A) = 0$ que $A \doteq \varnothing$.

5. A ida segue do item acima e de $m(X) = m(A) + m(A^{\complement})$. Como $A \cup X = X$ e $A \cap X = A$, segue que $A \triangle X = A^{\complement}$, logo $m(A \triangle X) = m(A^{\complement})$. Assim segue de $m(A^{\complement}) = 0$ que $A \doteq X$. Se m é finita, $m(A^{\complement}) = m(X) - m(A)$. Assim segue de $m(A) = m(X)$ se, e somente se, $m(A^{\complement}) = 0$.

6. Consequência dos itens anteriores.

7. Como $A \doteq X$ e $B \doteq X$, então $m(A^{\complement}) = m(B^{\complement}) = 0$, portanto

$$m\left((A \cap B)^{\complement}\right) = m\left(A^{\complement} \cup B^{\complement}\right) \leq m\left(A^{\complement}\right) + m\left(B^{\complement}\right) = 0,$$

   o que mostra que $A \cap B \doteq X$.

8. Como $A \subseteq A \cup B$, então $m(A \cup B) = m(A \cap B) + m(A \triangle B) = 1$ e $m(A) = m(A \cap B) + m(A \setminus B) = 1$, e, por m ser finita, igualando as expressões segue que

$$m(A \triangle B) = m(A \setminus B).$$

   Como

$$m(A \triangle B) = m(A \setminus B) = m(B \setminus A),$$

   segue que $m(B \setminus A) = 0$. Portanto, como

$$m(B) = m(B \cap A) + m(B \setminus A)$$

   concluímos que $m(B) = m(B \cap A)$.

■

Proposição B.4. *Seja $\boldsymbol{X}$ um espaço de medida.*

1. *Os conjuntos $\{M \in \mathcal{M} \mid M \doteq \varnothing\}$ e $\{M \in \mathcal{M} \mid M \doteq X\}$ são fechados sob união e interseção enumeráveis;*

2. *O conjunto $\{M \in \mathcal{M} \mid M \doteq \varnothing \;\text{ ou }\; M \doteq X\}$ é uma sigma-álgebra.*

Demonstração. 1. Seja $(M_i)_{i \in \mathbb{N}}$ uma sequência de conjuntos quase vazios. Então

$$m\left(\bigcup_{i \in \mathbb{N}} M_i\right) \leq \sum_{i \in \mathbb{N}} m(M_i) = 0$$

   e, para algum $i \in \mathbb{N}$

$$m\left(\bigcap_{i \in \mathbb{N}} M_i\right) \leq m(M_i) = 0.$$

   Seja $(M_i)_{i \in \mathbb{N}}$ uma sequência de conjuntos quase totais. O mesmo vale já que $M \doteq X$ se, e somente, $M^{\complement} \doteq \varnothing$.

2. Segue do item anterior, do fato que $M \doteq X$ se, e somente, $M^{\complement} \doteq \varnothing$, e de $\varnothing$ ser quase vazio.

■

Definição B.10. Sejam $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ espaços de medida. Um *quase função* de $\boldsymbol{X}_1$ para $\boldsymbol{X}_2$ é uma função $f : C \longrightarrow X_2$ tal que $C \doteq X_1$. Denota-se $f : \boldsymbol{X}_1 \overset{\circ}{\longrightarrow} \boldsymbol{X}_2$.

Definição B.11. Sejam X_1, C_1, X_2, C_2 e X_3 conjuntos tais que $C_1 \subseteq X_1$ e $C_2 \subseteq X_2$, e $f_1 : C_1 \longrightarrow X_2$ e $f_2 : C_2 \longrightarrow X_3$ funções. A *composição* (generalizada) de f_1 com f_2 é a função

$$\begin{aligned} f_2 \circ f_1 \colon C_1 \cap {f_1}^{-1}(C_2) &\longrightarrow X_3 \\ x &\longmapsto f_2(f_1(x)). \end{aligned}$$

Proposição B.5 (Composição de quase função)**.** *Sejam $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ e $\boldsymbol{X}_3$ espaços de medida e $f_1 : \boldsymbol{X}_1 \overset{\circ}{\longrightarrow} \boldsymbol{X}_2$ e $f_2 : \boldsymbol{X}_2 \overset{\circ}{\longrightarrow} \boldsymbol{X}_3$ quase funções, com respectivos domínios C_1 e C_2, que preservam medida. Então $f_2 \circ f_1 : \boldsymbol{X}_1 \overset{\circ}{\longrightarrow} \boldsymbol{X}_3$ é uma quase função, com domínio $C_1 \cap {f_1}^{-1}(C_2)$, que preserva medida.*

Demonstração. Como f_1 e f_2 preservam medida, $f_2 \circ f_1$ preserva medida. Basta mostrar que seu domínio é quase total. Como f_1 é quase função, $m_1({C_1}^{\complement}) = 0$, e como f_2 é quase função, $m_2({C_2}^{\complement}) = 0$. Assim, como f_1 preserva medida, $m_1({f_1}^{-1}((C_2)^{\complement})) = m_2((C_2)^{\complement}) = 0$. Por fim segue que

$$\begin{aligned} m_1\left(\left(C_1 \cap {f_1}^{-1}(C_2)\right)^{\complement}\right) &= m_1\left({C_1}^{\complement} \cup {f_1}^{-1}({C_2}^{\complement})\right) \\ &\leq m_1\left({C_1}^{\complement}\right) + m_1\left({f_1}^{-1}({C_2}^{\complement})\right) \\ &= 0 + 0 = 0, \end{aligned}$$

portanto $C_1 \cap {f_1}^{-1}(C_2) \overset{\circ}{=} X_1$, e concluímos que $f_2 \circ f_1$ é uma quase função. ■

Proposição B.6. *Sejam $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, $\boldsymbol{X}_3$ e $\boldsymbol{X}_4$ espaços de medida e $f_1 : \boldsymbol{X}_1 \overset{\circ}{\longrightarrow} \boldsymbol{X}_2$, $f_2 : \boldsymbol{X}_2 \overset{\circ}{\longrightarrow} \boldsymbol{X}_3$ e $f_3 : \boldsymbol{X}_3 \overset{\circ}{\longrightarrow} \boldsymbol{X}_4$ quase funções, com respectivos domínios C_1, C_2 e C_3, que preservam medida. Então*

$$f_3 \circ (f_2 \circ f_1) = (f_3 \circ f_2) \circ f_1.$$

Demonstração. O domínio de $f_3 \circ (f_2 \circ f_1)$ é o conjunto

$$(C_1 \cap {f_1}^{-1}(C_2)) \cap (f_2 \circ f_1)^{-1}(C_3)$$

e o domínio de $(f_3 \circ f_2) \circ f_1$ é o conjunto

$$C_1 \cap {f_1}^{-1}(C_2 \cap {f_2}^{-1}(C_3)).$$

Vamos mostrar que esses conjuntos são iguais.

$$\begin{aligned} (C_1 \cap {f_1}^{-1}(C_2)) \cap (f_2 \circ f_1)^{-1}(C_3) &= C_1 \cap {f_1}^{-1}(C_2) \cap (f_2 \circ f_1)^{-1}(C_3) \\ &= C_1 \cap {f_1}^{-1}(C_2) \cap {f_1}^{-1}({f_2}^{-1}(C_3)) \\ &= C_1 \cap {f_1}^{-1}(C_2 \cap {f_2}^{-1}(C_3)). \end{aligned}$$

■

Definição B.12. Sejam $\boldsymbol{X}_1$ e $\boldsymbol{X}_2$ espaços de medida. Quase funções de $\boldsymbol{X}_1$ para $\boldsymbol{X}_2$ *quase iguais* são quase funções $f_1 : \boldsymbol{X}_1 \overset{\circ}{\longrightarrow} \boldsymbol{X}_2$ e $f_2 : \boldsymbol{X}_1 \overset{\circ}{\longrightarrow} \boldsymbol{X}_2$ satisfazendo $f_1|_C = f_2|_C$ para algum conjunto $C \overset{\circ}{=} X_1$. Denota-se $f_1 \overset{\circ}{=} f_2$.

Proposição B.7. *A relação* $\doteq$ *de quase igualdade de quase funções é uma relação de equivalência.*

Demonstração. (Reflexividade) Claramente, $f \doteq f$. (Simetria) Claramente $f \doteq g$ se, e somente se $g \doteq f$. (Transitividade) Se $f \doteq g$ e $g \doteq h$, então existem conjuntos C e D quase totais tais que $f|_C = g|_C$ e $g|_D = h|_D$. Então, como $C \cap D$ é quase total, segue que $f|_{C\cap D} = g|_{C\cap D} = h|_{C\cap D}$, portanto $f \doteq h$. ∎

B.3.3 Quase Invariância

Definição B.13. Seja $\mathcal{S} = (\boldsymbol{X}, f)$ um sistema dinâmico de medida. Um conjunto *quase invariante* de $\mathcal{S}$ é um conjunto $M \in \mathcal{M}$ tal que $f^{-1}(M) \doteq M$.

Proposição B.8. *Sejam* $\mathcal{S} = (\boldsymbol{X}, f)$ *um sistema dinâmico de medida e* $M \in \mathcal{M}$ *um conjunto quase invariante de* $\mathcal{S}$*. Então existe um conjunto invariante* $N \in \mathcal{M}$ *tal que* $M \doteq N$*.*

Demonstração. Consideremos o limite superior

$$N := \limsup_{n\in\mathbb{N}} f^{-n}(M) = \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} f^{-n}(M),$$

o conjunto dos pontos que pertencem a infinitos dos conjuntos $f^{-n}(M)$. Notemos que, para todo $m \in \mathbb{N}$,

$$M \triangle \bigcup_{n=m}^{\infty} f^{-n}(M) \subseteq \bigcup_{n=m}^{\infty} M \triangle f^{-n}(M),$$

pois, se x é elemento do primeiro, então (1) é elemento de M mas não da união dos $f^{-n}(M)$ ou (2) é elemento da união mas não de M. No primeiro caso, ele não é elemento de nenhum dos $f^{-n}(M)$, logo é elemento de todos $M \triangle f^{-n}(M)$ e, portanto, da união; no segundo caso, é elemento de algum $f^{-n}(M)$ e, portanto, de algum $M \triangle f^{-n}(M)$, logo da união. Ainda, para todo $n \in \mathbb{N}$ vale

$$M \triangle f^{-n}(M) \subseteq \bigcup_{i=0}^{n-1} f^{-i}(M) \triangle f^{-(i+1)}(M).$$

Para ver isso, seja $x \in M \triangle f^{-n}(M)$. Suponhamos que $x \in M$. Então não pertence a $f^{-n}(M)$. Então, se $x \notin f^{-1}(M)$, temos que $x \in M \triangle f^{-1}(M)$ e, se $x \in f^{-1}(M)$, repetimos o processo com $f^{-2}(M)$ chegando, no máximo, até a $x \in f^{-(n-1)}(M)$, o que implicará $x \in f^{-(n-1)}(M) \triangle f^{-n}(M)$; em todos os casos, x pertence à união acima. Se $x \in f^{-n}(M)$, o mesmo processo se aplica, mas na ordem inversa. Isso implica que $m(M \triangle f^{-n}(M)) = 0$, pois o termo à direita tem medida zero.

Definamos $N_m := \bigcup_{n=m}^{\infty} f^{-n}(M)$, e temos que $N = \bigcap_{m=0}^{\infty} N_m$ e, claramente,

$$N_0 \supseteq N_1 \supseteq N_2 \supseteq \cdots .$$

Mas de $m(M \cap N_m) = 0$, então, segue que $m(N \triangle M) = 0$, logo que $m(N) = m(M)$. Por fim, notando que

$$f^{-1}(N) = \bigcap_{m=0}^{\infty} \bigcup_{n=m}^{\infty} f^{-(n+1)}(M) = \bigcap_{m=0}^{\infty} \bigcup_{n=m+1}^{\infty} f^{-n}(M) = N,$$

segue que N é invariante. ∎

B.4 Partição de Espaços de Medida

Quando temos uma medida, podemos enfraquecer a restrição de que as células de uma partição são iguais ou disjuntas para a de que são *quase-iguais* – a interseção tem medida zero – e a restrição de que cobrem para a restrição de que a união da partição tem medida total, e por fim, de que nenhum elemento da partição tem medida nula, e definir que isso é uma *partição por medida com respeito a m* ou uma *m-quase-partição*.

Definição B.14. Seja $\boldsymbol{X} = (X, \mathcal{M}, m)$ um espaço de medida. Uma *partição por medida* (ou *quase partição*) de $\boldsymbol{X}$ é um conjunto $\mathcal{P} \subseteq \mathcal{M}$ de subconjuntos mensuráveis de X que satisfaz

1. Para todo $C \in \mathcal{P}$, $C \not\mathrel{\mathring{=}} \varnothing$ (nenhuma célula é quase vazia);
2. $\bigcup \mathcal{P} \mathrel{\mathring{=}} X$ (a união das células é quase total);
3. Para todas células distintas $C_0, C_1 \in \mathcal{P}$, $C_0 \cap C_1 \mathrel{\mathring{=}} \varnothing$ (todas duas células distintas são quase disjuntas).

Os conjuntos $C \in P$ são as *células* de $\mathcal{P}$.

Definição B.15. Sejam $(\boldsymbol{X}, f)$ um sistema dinâmico de medida e $\mathcal{P}$ uma partição por medida de $\boldsymbol{X}$. A *partição por medida puxada* por f é o conjunto

$$f^{-1}(\mathcal{P}) := \left\{ f^{-1}(C) \,\middle|\, C \in \mathcal{P} \right\}.$$

Como f é mensurável, temos que $f^{-1}(C) \in \mathcal{M}$ para todo $C \in \mathcal{P}$. Como f preserva medida, temos que (1) $m(f^{-1}(C)) = m(C)$, portanto nenhuma célula da partição puxada é quase vazia; (2) sua união é

$$\bigcup_{C \in \mathcal{P}} f^{-1}(C) = f^{-1}\left(\bigcup_{C \in \mathcal{P}} C \right) \mathrel{\mathring{=}} X;$$

e (3) se $C_0, C_1 \in \mathcal{P}$ são distintas, $f^{-1}(C_0)$ e $f^{-1}(C_1)$ são distintas e

$$m(f^{-1}(C_0) \cap f^{-1}(C_1)) = m(f^{-1}(C_0 \cap C_1)) = m(C_0 \cap C_1) \mathrel{\mathring{=}} \varnothing.$$

$$A \mathrel{\mathring{=}} X \Rightarrow f^{-1}(A) \mathrel{\mathring{=}} f^{-1}(X) = X$$

$m(A \triangle B)$

Definição B.16. Sejam $\boldsymbol{X} = (X, \mathcal{M}, m)$ um espaço de medida e $(\mathcal{P}_n)_{n \in \mathbb{N}}$ uma sequência de partições por medida de $\boldsymbol{X}$. O *refinamento comum* a $(\mathcal{P}_n)_{n \in \mathbb{N}}$ é o conjunto

$$\bigvee_{n \in \mathbb{N}} \mathcal{P}_n := \left\{ \bigcap_{n \in \mathbb{N}} C_n \,\middle|\, n \in I,\ C_n \in \mathcal{P}_n \ \text{ e } \ \bigcap_{n \in \mathbb{N}} C_n \not\mathrel{\mathring{=}} \varnothing \right\}.$$