Capítulo 1

Conceitos Básicos da Teoria de Sistemas Dinâmicos

Neste capítulo apresentaremos as noções básicas de Sistemas Dinâmicos. Usaremos sempre o conjunto dos naturais incluindo o 0, então $\mathbb{N} = \{0, 1, 2, \cdots\}$ e nos casos em que precisarmos excluir o 0, usaremos $\mathbb{N}_* = \{1, 2, \cdots\}$. Um espaço métrico M é um conjunto com uma métrica d que possibilita calcular a distância entre dois pontos quaisquer. Definimos em M a topologia ζ gerada pela métrica d, isto é, ζ é a família de abertos de M, pela métrica d. Definimos também em M a $\sigma-$álgebra de Borel $\mathcal{B}(M)$, que é gerada pela topologia ζ, ou seja, a menor $\sigma-$álgebra que contém todos os subconjuntos abertos de M.

O conjunto M será sempre um espaço topológico mensurável compacto, e quando ele tiver uma outra estrutura ou propriedade, como por exemplo ser uma variedade diferenciável ou um espaço de probabilidade, será devidamente caracterizado. A função $f : M \to M$ será sempre uma função contínua, e portanto mensurável. Na sequência vamos definir uma série de termos básicos da teoria.

1.1 Dinâmica Topológica

Vamos definir e mostrar alguns resultados de uma dinâmica sob um olhar topológico. Dado uma aplicação $f : M \to M$ qualquer, dizemos que f é um **Sistema Dinâmico** que associa um ponto $x \in M$ a um outro ponto $f(x) \in M$ que é a dinâmica de x uma unidade de tempo depois, e portanto f é uma dinâmica com tempo discreto. Definimos $f^0(x) = x$, $f^1(x) = f(x)$ e $f^n(x) = f\big(f^{n-1}(x)\big)$, que são os iterados futuros de x, e então chamamos de **órbita futura** de x o conjunto $\mathcal{O}^+(x) = \{f^n(x); n \in \mathbb{N}\}$. Se f for bijetora, definimos também a **órbita**

passada de x, o conjunto $\mathcal{O}^-(x) = \{f^{-n}(x); n \in \mathbb{N}\}$, em que $f^{-1}(x)$ é a pré imagem de x e $f^{-n}(x) = f^{-1}\big(f^{-(n-1)}(x)\big)$, que são os iterados passados de x, e nesse caso podemos definir de maneira generalizada o conjunto $\mathcal{O}(x) = \mathcal{O}^+(x) \cup \mathcal{O}^-(x) = \{f^n(x); n \in \mathbb{Z}\}$, que chamamos de **órbita** de x. Caso precisemos deixar claro a qual função estamos nos referindo, denotaremos tais órbitas por $\mathcal{O}^+(x,f)$, $\mathcal{O}^-(x,f)$ e $\mathcal{O}(x,f)$, respectivamente.

Dizemos que um ponto $p \in M$ é um **ponto periódico** de f de período $\tau(p)$ se $f^{\tau(p)}(p) = p$ e chamamos o conjunto finito $\mathcal{O}(p) = \{p, f(p), \cdots, f^{\tau(p)-1}(p)\}$ de **órbita periódica** de p. Denotamos por $Per_n(f)$ o conjunto dos pontos periódicos de f de período n e $Per(f) = \bigcup_{n\in\mathbb{N}} Per_n(f)$ o conjunto de todos os pontos periódicos de f. Particularmente quando $\tau(p) = 1$, ou seja, $f(p) = p$, chamamos p de **ponto fixo** de f e definimos $Fix(f)$ o conjunto dos pontos fixos de f.

Um conjunto $A \subseteq M$ é chamado **positivamente** $f-$**invariante** se $f(A) \subseteq A$ e dizemos que o conjunto A é **negativamente** $f-$**invariante** se $f^{-1}(A) \subseteq A$. Se $f(A) = A$ dizemos que A é um conjunto $f-$**invariante**. Observemos que o conjunto $Per(f)$ é um conjunto f-invariante. De fato, se $p \in Per(f)$ então $f^{\tau(p)}(p) = p$, isso implica que $f^{\tau(p)}\big(f(p)\big) = f^{\tau(p)+1}(x) = f^{1+\tau(p)}(x) = f\big(f^{\tau(p)}(p)\big) = f(p)$, logo $f(p) \in Per(f)$, ou seja, $f\big(Per(f)\big) \subseteq Per(f)$. Reciprocamente, se $p \in f^{-1}\big(Per(f)\big)$ significa que $f(p) \in Per(f)$ e com raciocínio análogo ao caso anterior concluímos que $p \in Per(f)$, ou seja, $f^{-1}\big(Per(f)\big) \subseteq Per(f)$. Portanto $f\big(Per(f)\big) = Per(f)$.

Um ponto $x \in M$ é um **ponto recorrente no futuro** (respectivamente **ponto recorrente no passado**, caso f seja bijetora) se para cada vizinhança V_x de x dada, existe $n \in \mathbb{N}$ tal que $f^n(x) \in V_x$ (respectivamente existe $m \in \mathbb{N}$ tal que $f^{-m}(x) \in V_x$). Se x é recorrente no passado e no futuro, isto é, para cada vizinhança V_x de x dada, existem $m, n \in \mathbb{N}$ tal que $f^{-m}(x), f^n(x) \in V_x$, dizemos que x é **recorrente**. Denotamos por $\mathcal{R}(f)$ o conjunto dos pontos recorrentes de f. Dizemos que um ponto $x \in M$ é um **ponto não errante** de f se para cada vizinhança V_x de x dada, existe $n \in \mathbb{N}$ tal que $f^n\big(V_x\big) \cap V_x \neq \emptyset$, ou seja, para cada vizinhança V_x de x dada, existem $n \in \mathbb{N}$ e $y \in V_x$ tais que $f^n(y) \in V_x$. Denotamos por $\Omega(f)$ o conjunto dos pontos não errantes de f.

Seja $\{x_n\}_{n\in\mathbb{Z}} \subseteq M$ uma sequência de pontos de M. Dizemos que $\{x_n\}_{n\in\mathbb{Z}}$ é uma $\varepsilon-$**pseudo-órbita** para f se $d\big(f(x_n), x_{n+1}\big) \leq \varepsilon$. Se $x \in \{x_n\}_{n\in\mathbb{Z}}$ e existe um $n_0 \in \mathbb{N}$ tal que $d\big(f(x_{n_0-1}), x\big) \leq \varepsilon$ e $x_{n_0} = x_1$, dizemos que $\{x_n\}_{n\in\mathbb{Z}}$ é uma $\varepsilon-$**pseudo-órbita periódica contendo** x. Um ponto $y \in M$ $\delta-$**sombreia** a sequência $\{x_n\}_{n\in\mathbb{Z}}$ se $d\big(f^n(y), x_n\big) \leq \delta$ para todo $n \in \mathbb{Z}$. Dizemos que um ponto $x \in M$ é **recorrente por cadeia** se para cada $\varepsilon > 0$ dado, existe uma $\varepsilon-$pseudo-órbita periódica contendo x e denotamos por $\mathcal{RC}(f)$ o conjunto de todos os pontos recorrentes por cadeia.

Esses conjuntos definidos acima possuem uma relação entre si, que podemos ver na seguinte proposição.

Proposição 1.1. *Seja $f : M \to M$ uma aplicação qualquer, então temos a seguinte sequência de inclusões:*

$$Per(f) \subseteq \mathcal{R}\,(f) \subseteq \Omega(f) \subseteq \mathcal{RC}\,(f)\,.$$

Demonstração. Vamos provar separadamente cada inclusão.

$Per(f) \subseteq \mathcal{R}\,(f)$: Seja $p \in Per(f)$, então para toda vizinhança V_p de p, temos $f^{\tau(p)}(p) = p \in V_x$ e portanto $Per(f) \subseteq \mathcal{R}\,(f)$.

$\mathcal{R}\,(f) \subseteq \Omega(f)$: Seja $x \in \mathcal{R}\,(f)$, então para cada vizinhança V_x de x dada, existe $n \in \mathbb{N}$ tal que $f^n(x) \in V_x$, em particular $f^n\big(V_x\big) \cap V_x \neq \emptyset$, como $\varepsilon > 0$ é um qualquer, temos que $x \in \Omega(f)$ e portanto $\mathcal{R}\,(f) \subseteq \Omega(f)$.

$\Omega(f) \subseteq \mathcal{RC}\,(f)$: Seja $x \in \Omega(f)$, tomemos $\varepsilon > 0$ tal que $V_x \subseteq B(x, \varepsilon)$, em que V_x é uma vizinhança de x. Então, existem $n_0 \in \mathbb{N}$ e $y \in V_x$ tais que $f^{n_0}(y) \in V_x$. Definamos a sequência $\{y_n\}_{n\in\mathbb{Z}} \subseteq M$ tal que $y_n = f^n(y)$, logo $d\big(f(y_{n_0-1}), x\big) < \varepsilon$. Como $\varepsilon > 0$ é um qualquer, então $x \in \mathcal{RC}\,(f)$ e portanto $\Omega(f) \subseteq \mathcal{RC}\,(f)$.

□

Corolário 1.2. *Se $\overline{Per(f)} = M$, então $\Omega(f) = M$.*

Demonstração. Aplicando $\overline{Per(f)} = M$ na Proposição 1.1 temos $\overline{\Omega(f)} = M$, então basta provar que $\Omega(f)$ é um conjunto fechado e concluímos a demonstração. De fato, se $(x_k)_{k=1}^{+\infty} \subseteq \Omega(f)$ é uma sequência que converge para x, então dada uma vizinhança V_x de x, existe um $k_0 \in \mathbb{N}$ tal que para todo $k > k_0$ temos $x_k \in V_x$. Fixemos um $k > k_0$ e tomemos $\delta > 0$ tal que $B(x_k, \delta) \subseteq V_x$. Daí, como x_k é não errante, existe $n \in \mathbb{N}$ tal que $f^n\big(B(x_k, \delta)\big) \cap B(x_k, \delta) \neq \emptyset$, em particular $f^n(V_x) \cap V_x \neq \emptyset$. Como V_x é uma vizinhança qualquer de x, então $x \in \Omega(f)$ e portanto $\Omega(f)$ é um conjunto fechado, concluindo a demonstração. □

Existem dinâmicas cuja órbita de qualquer aberto percorre todo o espaço, e para algum iterado futuro intersecta qualquer outro aberto.

Definição 1.3. *Seja $f : M \to M$ uma aplicação qualquer. Dizemos que f é **topologicamente transitiva**, ou apenas **transitiva**, se para quaisquer abertos $U, V \subseteq M$ dados, existe $n \in \mathbb{N}$ tal que $f^n(U) \cap V \neq \emptyset$.*

A transitividade de uma função pode ser definida, de forma equivalente, através da órbita de um ponto.

Proposição 1.4. *Sejam $f : M \to M$ uma aplicação contínua qualquer e M um espaço métrico compacto, então f é transitiva se, e somente se, existe $x_0 \in M$ tal que $\overline{\mathcal{O}(x_0)} = M$.*

Demonstração. Se f é transitiva, então dados $U, V \subseteq M$ existe $n \in \mathbb{N}$ tal que $f^n(U) \cap V \neq \emptyset$ o que implica $f^{-n}(V) \cap U \neq \emptyset$, em que $f^{-n}(V)$ também é aberto, pois f é contínua. Como M é compacto, então admite uma base enumerável de abertos $B = \{B_1, B_2, \cdots\}$.

Construamos a seguinte sequência de compactos encaixados. Por hipótese, existe $n_1 \in \mathbb{N}$ tal que o aberto $U_1 = B_1 \cap f^{-n_1}(B_2) \neq \emptyset$, então tomemos um compacto $K_1 \subseteq U_1$. Por hipótese, existe $n_2 \in \mathbb{N}$ tal que o aberto $U_2 = U_1 \cap f^{-n_2}(B_3) \neq \emptyset$, então tomemos um compacto $K_2 \subseteq U_2$. Repetindo esse processo indefinidamente, obtemos uma sequência de compactos encaixados

$$K_1 \supseteq K_2 \supseteq K_3 \supseteq \cdots$$

em que a órbita futura dos pontos de K_n começam em B_1 e intercecta B_i após n_i iterações, para $i \in \{1, 2, \cdots, n\}$.

O conjunto $K = \bigcap_{i=1}^{+\infty} K_i$ é compacto não vazio, pois é uma intersecção enumerável de compactos não vazios encaixados. Portanto, dado $x_0 \in K$ temos que $f^{n_i}(x_0) \in B_1$ para todo $i \in \mathbb{N}$, e como B é uma base de M, ou seja, para todo aberto $A \subseteq M$ existe um $B_i \in B$ tal que $B_i \subseteq A$, então $\mathcal{O}^+(x_0) \cap A \neq \emptyset$, isto é, $\overline{\mathcal{O}(x_0)} = M$.

Reciprocamente, dado dois abertos $U, V \subseteq M$. Se $\overline{\mathcal{O}(x_0)} = M$, então existe $n_1, n_2 \in \mathbb{N}$ tais que $f^{n_1}(x_0) \in U$ e $f^{n_2}(x_0) \in V$; sem perda de generalidade podemos supor $n_1 < n_2$, pois a órbita de x_0 retorna a U e V infinitas vezes. Daí, $f^{n_1}(x_0) \in U$ e $f^{n_2-n_1}(f^{n_1}(x_0)) = f^{n_2}(x_0) \in V$. Portanto tomando $n = n_2 - n_1$ temos $f^n(U) \cap V \neq \emptyset$, ou seja, f é transitiva. $\square$

Se $\overline{\mathcal{O}(x)} = M$, para todo $x \in M$, isto é, a órbita de todo ponto é densa em M dizemos que f é **minimal**; essa definição equivale a dizer que M não tem conjuntos próprios $f-$invariantes e fechados. De fato, observemos que dado $x \in M$, o conjunto $\mathcal{O}(x)$ é o menor conjunto $f-$invariante contendo x, pois dado qualquer conjunto $A \subseteq M$, $f-$invariante e contendo x, então $f^n(x) \in A$ para todo $n \in \mathbb{Z}$, logo $\mathcal{O}(x) \subseteq A$. Daí, se A for $f-$invariante e fechado, então $M = \overline{\mathcal{O}(x)} \subseteq A$ e portanto A não é próprio.

Corolário 1.5. *Seja $f : M \to M$ uma aplicação qualquer. Se f for minimal, então f é transitiva.*

Demonstração. Como f é minimal, então $\overline{\mathcal{O}(x)} = M$ para todo $x \in M$. Em particular existe um $x_0 \in M$ tal que $\overline{\mathcal{O}(x_0)} = M$, e portanto pela Proposição 1.4, f é transitiva. $\square$

Algumas dinâmicas possuem uma propriedade mais forte ainda do que a transitividade, em que além da órbita de um aberto percorrer todo o espaço, ela expande esse aberto de tal forma que a partir de um certo iterado, essa órbita continua intersectando consecutivamente qualquer outro aberto.

Definição 1.6. *Seja $f : M \to M$ uma aplicação qualquer. Dizemos que f é **topologicamente mixing**, se para quaisquer abertos $U, V \subseteq M$ dados, existe $n_0 \in \mathbb{N}$ tal que $f^n(U) \cap V \neq \emptyset$, para todo $n > n_0$.*

Proposição 1.7. *Seja $f : M \to M$ uma aplicação qualquer, se f for topologicamente mixing, então f é transitiva.*

Demonstração. Sejam $U, V \subseteq M$ dois abertos quaisquer, como f é topologicamente *mixing* então existe um $n_0 \in \mathbb{N}$ tal que para todo $n > n_0$ temos que $f^n(U) \cap V \neq \emptyset$. Em particular existe um $n \in \mathbb{N}$ tal que $f^n(U) \cap V \neq \emptyset$, ou seja, f é transitiva. $\square$

Esses comportamentos podem ser verificados em várias dinâmicas sobre um mesmo espaço ou em dinâmicas semelhantes sobre espaços diferentes. Algumas, apesar de possuírem leis de formação diferentes, possuem as mesmas propriedades que podem ser relacionadas através de conjugações.

Definição 1.8. *Sejam $f : M_1 \to M_1$ e $g : M_2 \to M_2$, aplicações quaisquer, em que M_1 e M_2 são espaços métricos. Dizemos que $h : M_1 \to M_2$ é uma **conjugação** de f e g, se h for um homeomorfismo tal que o seguinte diagrama seja comutativo:*

$$\begin{array}{ccc} M_1 & \xrightarrow{f} & M_1 \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h} \\ M_2 & \xrightarrow[g]{} & M_2 \end{array}$$

*Ou seja, $h \circ f = g \circ h$. Dizemos que f e g são **topologicamente conjugadas** quando existe uma conjugação h entre elas. Lembrando que um homeomorfismo é uma função bijetora contínua com inversa contínua.*

Veremos agora um resultado que mostra como duas dinâmicas topologicamente conjugadas possuem o mesmo comportamento assintótico.

Teorema 1.9. *Sejam M_1 e M_2 espaços métricos, $f : M_1 \to M_1$ e $g : M_2 \to M_2$ aplicações, e $h : M_1 \to M_2$ uma conjugação de f e g, então as seguintes afirmações são verdadeiras:*

i) $h \circ f^n = g^n \circ h$;

ii) $h\big(Per(f)\big) = Per(g)$;

iii) $h\big(\mathcal{R}\,(f)\,\big) = \mathcal{R}\,(g)$;

iv) $h\big(\Omega(f)\big) = \Omega(g)$;

v) *f é transitiva se, e somente se, g é transitiva;*

vi) *f é topologicamente mixing se, e somente se, g é topologicamente mixing.*

Demonstração. *i)* Temos que $f = h^{-1} \circ g \circ h$, pois h^{-1} existe porque h é um homeomorfismo. Então

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

Portanto $h \circ f^n = g^n \circ h$.

ii) Se $h(p) \in h\big(Per(f)\big)$ então $f^{\tau(p)}(p) = p$ e pelo item anterior temos $g^{\tau(p)}\big(h(p)\big) = h\big(f^{\tau(p)}(p)\big)$ $= h(p)$. Logo $h(p) \in Per(g)$, o que significa que $h\big(Per(f)\big) \subseteq Per(g)$. Reciprocamente, se $q \in Per(g)$ então $g^{\tau(q)}(q) = q$ e como h é um homeomorfismo podemos escrever $p = h^{-1}(q) \in M_1$ e pelo item anterior temos $f^{\tau(q)}(p) = f^{\tau(q)}\big(h^{-1}(q)\big) = h^{-1}\big(g^{\tau(q)}(q)\big) = h^{-1}(q) = p$. Logo $p = h^{-1}(q) \in Per(f)$ e então $q = h(p) \in h\big(Per(f)\big)$, o que significa que $Per(g) \subseteq h\big(Per(f)\big)$. Portanto $h\big(Per(f)\big) = Per(g)$.

iii) Se $h(x) \in h\big(\mathcal{R}\,(f)\,\big)$ então dado uma vizinhança $U_{h(x)}$ de $h(x)$, temos que $V_x = h^{-1}\big(U_{h(x)}\big)$ é uma vizinhança de x e existe $n \in \mathbb{N}$ tal que $f^n(x) \in V_x$ e pelo primeiro item temos $g^n\big(h(x)\big) = h\big(f^n(x)\big) \in h(V_x) = U_{h(x)}$, logo $h(x) \in \Omega(g)$, o que significa que $h\big(\Omega(f)\big) \subseteq \Omega(g)$. Reciprocamente, se $y \in \Omega(g)$ então dada uma vizinhança V_x de $x = h^{-1}(y)$, temos que $h(V_x)$ é uma vizinhança de y e existe $n \in \mathbb{N}$ tal que $g^n(y) \in h(V_x)$ e pelo primeiro item $g^n(y) = g^n\big(h(x)\big) = h\big(f^n(x)\big) \in h(V_x)$ o que significa que $y = h(x) \in h\big(\Omega(f)\big)$. Portanto $h\big(\Omega(f)\big) = \Omega(g)$.

iv) Se $h(x) \in h\big(\Omega(f)\big)$, então dado uma vizinhança $U_{h(x)}$ de $h(x)$, temos que $V_x = h^{-1}\big(U_{h(x)}\big)$ é uma vizinhança de x e existe $n \in \mathbb{N}$ tal que $f^n(V_x) \cap V_x \neq \emptyset$, e pelo primeiro item temos $h\big(f^n(V_x)\big) = g^n\big(h(V_x)\big) = g^n\big(U_{h(x)}\big)$, logo $g^n\big(U_{h(x)}\big) \cap U_{h(x)} \neq \emptyset$ e $h(x) \in \Omega(g)$, o que significa que $h\big(\Omega(f)\big) \subseteq \Omega(g)$. Reciprocamente, se $y \in \Omega(g)$ então dada uma vizinhança V_x de $x = h^{-1}(y)$, temos que $h(V_x)$ é uma vizinhança de y e existe $n \in \mathbb{N}$ tal que $g^n\big(h(V_x)\big) \cap h(V_x) \neq \emptyset$ e pelo primeiro item $g^n\big(h(V_x)\big) = h\big(f^n(V_x)\bigcap\big)$, logo $h\big(f^n(V_x)\big) \cap h(V_x) \neq \emptyset$ e $y = h(x) \in h\big(\Omega(f)\big)$, o que significa que $\Omega(g) \subseteq h\big(\Omega(f)\big)$. Portanto $h\big(\Omega(f)\big) = \Omega(g)$.

vi) Se f é transitiva então dado dois abertos $U_2, V_2 \subseteq M_2$ quaisquer, temos $U_1 = h^{-1}(U_2)$ e $V_1 = h^{-1}(V_2)$ são abertos em M_1, logo existe $n \in \mathbb{N}$ tal que $f^n(U_1) \cap V_1 \neq \emptyset$ e pelo primeiro item $g^n(U_2) = g^n\big(h(U_1)\big) = h\big(f^n(U_1)\big)$, portanto $g^n(U_2) \cap V_2 \neq \emptyset$, o que significa que g é transitiva. Reciprocamente, se g é transitiva então dado dois abertos $U_1, V_1 \subseteq M_1$ quaisquer, temos $U_2 = h(U_1)$ e $V_2 = h(V_1)$ são abertos em M_2, pois a inversa de h é contínua, logo existe $n \in \mathbb{N}$ tal que $g^n(U_2) \cap V_2 \neq \emptyset$ e como h^{-1} também é uma conjugação, pelo primeiro item $f^n(U_1) = f^n\big(h^{-1}(U_2)\big) = h^{-1}\big(g^n(U_2)\big)$, portanto $f^n(U_1) \cap V_1 \neq \emptyset$, o que significa que f é transitiva.

vii) Se f topologicamente mixing então dado dois abertos $U_2, V_2 \subseteq M_2$ quaisquer, temos $U_1 = h^{-1}(U_2)$ e $V_1 = h^{-1}(V_1)$ são abertos em M_1, logo existe $n_0 \in \mathbb{N}$ tal que para todo $n > n_0$ temos que $f^n(U_1) \cap V_1 \neq \emptyset$ e pelo primeiro item $g^n(U_2) = g^n\big(h(U_1)\big) = h\big(f^n(U_1)\big)$, portanto $g^n(U_2) \cap V_2 \neq \emptyset$, o que significa que g é topologicamente mixing. Reciprocamente, se g é topologicamente mixing então dado dois abertos $U_1, V_1 \subseteq M_1$ quaisquer, temos $U_2 = h(U_1)$ e $V_2 = h(V_1)$ são abertos em M_2, pois a inversa de h é contínua, logo existe $n_0 \in \mathbb{N}$ tal que para todo $n > n_0$ temos que $g^n(U_2) \cap V_2 \neq \emptyset$ e como h^{-1} também é uma conjugação, pelo primeiro item $f^n(U_1) = f^n\big(h^{-1}(U_2)\big) = h^{-1}\big(g^n(U_2)\big)$, portanto $f^n(U_1) \cap V_1 \neq \emptyset$, o que significa que f é topologicamente mixing.

$\square$

Corolário 1.10. *Nas condições do Teorema 1.9, f é minimal se, e somente se, g é minimal.*

Demonstração. Dado $x \in M_1$, pelo primeiro item do Teorema 1.9, temos $h\big(\mathcal{O}(x, f)\big) = \mathcal{O}\big(h(x), g\big)$. Como f é minimal então $\overline{\mathcal{O}(x, f)} = M_1$ e portanto $\overline{\mathcal{O}\big(h(x), g\big)} = h\big(\overline{\mathcal{O}(x, f)}\big) = h(M_1) = M_2$, e como x é um ponto qualquer e h é bijetora, concluímos que g é minimal. Analogamente, aplicando o mesmo raciocínio para h^{-1}, concluímos que g ser minimal implica em f ser minimal. $\square$

Um exemplo interessante de dinâmica, é a rotação irracional na esfera unitária $S^1 = \big\{z \in \mathbb{C}; |z| = 1\big\} = \big\{e^{2\pi i x};\ x \in [0, 2\pi)\big\}$ em $\mathbb{C}$. Definamos o conjunto $\mathbb{R}/\mathbb{Z} = \big\{[x] \in [0, 1);\ x \sim x' \Leftrightarrow x - x' \in \mathbb{Z}\big\}$, em que $[x]$ é a classe de equivalência pela relação $\sim$. Para facilitar a notação, ao invés de escrevermos $[x] \in \mathbb{R}/\mathbb{Z}$, escreveremos apenas $x \in \mathbb{R}/\mathbb{Z}$ onde o x estará representando sua classe de equivalência $x \pmod 1$. Dado um $\alpha \in \mathbb{R}$, definamos também sobre esses conjuntos as seguintes dinâmicas

$$\begin{array}{rccc} R_\alpha : & S^1 & \to & S^1 \\ & e^{2\pi i x} & \mapsto & e^{2\pi i(x+\alpha)} \end{array} \quad \text{e} \quad \begin{array}{rccc} T_\alpha : & \mathbb{R}/\mathbb{Z} & \to & \mathbb{R}/\mathbb{Z} \\ & x & \mapsto & x+\alpha \end{array}.$$

Note que $R_\alpha(x)$ é uma rotação do ponto $e^{2\pi i x} \in S^1$ pelo angulo α e $T_\alpha(x)$ é a parte decimal do ponto $x \in \mathbb{R}/\mathbb{Z}$ transladado por α. Vamos mostrar que R_α é minimal quando $\alpha \in (\mathbb{R} - \mathbb{Q})$, e pra isso vejamos que R_α e T_α são topologicamente conjugadas. Seja $h : \mathbb{R}/\mathbb{Z} \to S^1$ tal que $h(x) = e^{2\pi i x}$, então h é uma conjugação de T_α e R_α.

i) h é injetora: De fato, seja $x, y \in \mathbb{R}/\mathbb{Z}$ tal que $h(x) = h(y)$, então $e^{2\pi i x} = e^{2\pi i y} \Rightarrow e^{2\pi i(x-y)} = 1 \Rightarrow 2\pi i(x-y) = \ln(1) \Rightarrow 2\pi i(x-y) = 0 \Rightarrow x - y = 0 \Rightarrow x = y$.

ii) h é sobrejetora: De fato, seja $z \in S^1$, então $z = e^{2\pi i x}$ para algum $x \in [0, 2\pi)$, logo existe $x \in \mathbb{R}/\mathbb{Z}$ tal que $h(x) = e^{2\pi i x} = z$.

iii) h é contínua com inversa contínua: De fato, a função exponencial é contínua, e sua inversa h^{-1} definida por $h^{-1}(z) = x$, em que $z = e^{2\pi i x}$, é a função logarítmica e também é contínua.

iv) $h \circ T_\alpha = R_\alpha \circ h$: De fato, $\big(h \circ T_\alpha\big)(x) = h\big(T_\alpha(x)\big) = h(x+\alpha) = e^{2\pi i(x+\alpha)} = R_\alpha(e^{2\pi i x}) = R_\alpha\big(h(x)\big) = \big(R_\alpha \circ h\big)(x)$.

Então podemos trabalhar apenas com T_α e as propriedades do Teorema 1.9 se aplicam a R_α. Vamos ver abaixo dois resultados que caracterizam a dinâmica R_α em função do α.

Proposição 1.11. *Se α for um número racional, então $Per(R_\alpha) = S^1$.*

Demonstração. Temos que $T_\alpha = x + \alpha$, $T_\alpha^2 = (x+\alpha) + \alpha = x + 2\alpha$ e $T_\alpha^n = \big(x + (n-1)\alpha\big) + \alpha = x + n\alpha$. Daí, dado $x \in \mathbb{R}/\mathbb{Z}$ e tomando $\alpha = \dfrac{p}{q} \in \mathbb{Q}$ temos

$$T_\alpha^q(x) = x + q\alpha = x + q\frac{p}{q} = x + q = x.$$

Essa última igualdade é verdade porque q é um número inteiro e então $x + q$ tem a mesma parte decimal que x. Portando x é periódico, isto é, $Per(T_\alpha) = \mathbb{R}/\mathbb{Z}$ e pelo Teorema 1.9 item *ii*), concluímos que $Per(R_\alpha) = h\big(Per(T_\alpha)\big) = h\big(\mathbb{R}/\mathbb{Z}\big) = S^1$. $\square$

Proposição 1.12. *Se α for um número irracional, então R_α é minimal.*

Demonstração. Dado $x \in \mathbb{R}/\mathbb{Z}$, suponhamos que $\overline{\mathcal{O}(x)} \neq \mathbb{R}/\mathbb{Z}$, então $B = \mathbb{R}/\mathbb{Z} - \overline{\mathcal{O}(x)} \neq \emptyset$ é um conjunto $f-$invariante aberto e por isso pode ser escrito como a união de intervalos abertos e disjuntos. Tomemos $I \subseteq B$ o maior desses intervalos, e caso tenha mais de um sendo maior que os outros, tomemos I um dos maiores. Seja $m(I) > 0$ a medida desse intervalo, isto é, se $I = (a, b)$ então $m(I) = b - a$.

Afirmação 1: $T_\alpha(I)$ continua sendo o maior desses intervalos. De fato, se $I = (a, b)$ então $m\big(T_\alpha(I)\big) = m\big((a + \alpha, b + \alpha)\big) = b + \alpha - (a + \alpha) = b - a = m(I)$.

Afirmação 2: $T_\alpha^n(I) \neq I$ para todo $n \in \mathbb{N}$. De fato, suponhamos que $T_\alpha^n(I) = I$ para algum $n \in \mathbb{N}$, então pelo Teorema do Valor Intermediário, existe um $c \in I$ tal que $T_\alpha^n(c) = c$, logo $c + n\alpha = c \pmod 1$ o que significa que $d = n\alpha \in \mathbb{Z}$ e daí $\alpha = \frac{d}{n}$, o que é absurdo.

Afirmação 3: $T_\alpha^n(I)$ são dois a dois disjuntos, para todo $n \in \mathbb{N}$. De fato, suponhamos que exista $m, n \in \mathbb{N}$ tal que $T_\alpha^m(I) \cap T_\alpha^n(I) \neq \emptyset$, como B é $f-$invariante e pela Afirmação 2 $T_\alpha^m(I) \neq T_\alpha^n(I)$, então $T_\alpha^m(I) \cup T_\alpha^n(I) \subseteq B$ seria o maior intervalo de B, o que é absurdo.

Portanto, podemos concluir que

$$m\left(\bigcup_{n=0}^{+\infty} T_\alpha(I)\right) = \sum_{n=0}^{+\infty} m\big(T_\alpha^n(I)\big) = \sum_{n=0}^{+\infty} m(I) = +\infty.$$

Absurdo, pois $m\left(\bigcup_{n=0}^{+\infty} T_\alpha(I)\right) \leq m(\mathbb{R}/\mathbb{Z}) = m\big([0,1)\big) = 1$. Logo $\overline{\mathcal{O}(x)} = \mathbb{R}/\mathbb{Z}$ e como x é um ponto qualquer, concluímos que T_α é minimal. Daí, pelo Corolário 1.10 R_α é minimal. $\square$

1.1.1 Hiperbolicidade

Quando o espaço M possui certas propriedades, podemos mapear melhor a sua dinâmica. Então, sejam M uma variedade diferenciável, $f : M \to M$ um difeomorfismo de classe C^r e $p \in M$ um ponto periódico de período $\tau(p)$. Podemos decompor T_pM através dos autoespaços do operador derivada $Df_p^{\tau(p)} : T_pM \to T_pM$. Chamamos de **subespaço estável** de T_pM o subespaço E_p^s gerado pelos autovalores $|\lambda| < 1$, **subespaço instável** E_p^u o gerado pelos autovalores $|\lambda| > 1$, e **subespaço central** E_p^c o gerado pelos autovalores $|\lambda| = 1$. Temos portanto uma decomposição do espaço T_pM em soma direta

$$T_pM = E_p^s \oplus E_p^c \oplus E_p^u.$$

Dizemos que p é um ponto **periódico hiperbólico** se $E^c = \{0\}$, ou seja, $Df_p^{\tau(p)}$ não possui autovalor λ com $|\lambda| = 1$ e portanto

$$T_pM = E_p^s \oplus E_p^u.$$

Essa definição pode ser estendida para os pontos que não são periódicos. Existem conjuntos $\Lambda \subseteq M$ que também admitem uma boa decomposição do fibrado tangente, que chamaremos de conjuntos hiperbólicos. Mais precisamente:

Definição 1.13. *Seja $f : M \to M$ um difeomorfismo de classe C^r definido em uma variedade diferenciável M. Dizemos que um conjunto $\Lambda \subseteq M$ é um* ***conjunto hiperbólico*** *se Λ é $f-$invariante e existem $0 < \lambda < 1$ e $C \in \mathbb{R}$, tais que para cada $x \in \Lambda$:*

i) Existe a decomposição $T_xM = E_x^s \oplus E_x^u$;

ii) Essa decomposição é $Df-$invariante, ou seja, $Df(E_x^s) = E_{f(x)}^s$ e $Df(E_x^u) = E_{f(x)}^u$;

iii) E_x^s e E_x^u variam continuamente com x;

iv) $\|Df^n(x)|_{E_x^s}(v)\| \leq C\lambda^n\|v\|$ para todo $v \in E_x^s$ e para todo $n \in \mathbb{N}$;

v) $\|Df^{-n}(x)|_{E_x^u}(v)\| \leq C\lambda^n\|v\|$ para todo $v \in E_x^u$ e para todo $n \in \mathbb{N}$.

Quando M é um conjunto hiperbólico, chamamos $f : M \to M$ de **difeomorfismo de Anosov**. Dizemos que um conjunto hiperbólico Λ é **isolado** ou **maximal** se existir uma vizinhança U de Λ tal que

$$\Lambda = \bigcap_{n\in\mathbb{Z}} f^n(U).$$

Definimos o **conjunto estável local** $W_\varepsilon^s(x) \subseteq M$ de x e o **conjunto instável local** $W_\varepsilon^u(x) \subseteq M$ de x, respectivamente, por

$$\begin{aligned} W_\varepsilon^s(x) &= \left\{y \in M;\ \lim_{n\to+\infty} d\big(f^n(x), f^n(y)\big) = 0 \text{ e } d\big(f^n(x), f^n(y)\big) < \varepsilon, \forall n \in \mathbb{N}\right\}, \\ W_\varepsilon^u(x) &= \left\{y \in M;\ \lim_{n\to+\infty} d\big(f^{-n}(x), f^{-n}(y)\big) = 0 \text{ e } d\big(f^{-n}(x), f^{-n}(y)\big) < \varepsilon, \forall n \in \mathbb{N}\right\}. \end{aligned}$$

De modo geral, definimos o **conjunto estável** $W^s(x) \subseteq M$ de x e o **conjunto instável** $W^u(x) \subseteq M$ de x, respectivamente, por

$$\begin{aligned} W^s(x) &= \left\{y \in M;\ \lim_{n\to+\infty} d\big(f^n(x), f^n(y)\big) = 0\right\}, \\ W^u(x) &= \left\{y \in M;\ \lim_{n\to+\infty} d\big(f^{-n}(x), f^{-n}(y)\big) = 0\right\}. \end{aligned}$$

É claro que $W_\varepsilon^s(x) \subseteq W^s(x)$ e $W_\varepsilon^u(x) \subseteq W^u(x)$. Caso precisemos deixar claro qual a função estamos nos referindo, denotaremos tais conjuntos por $W_\varepsilon^s(x,f)$, $W_\varepsilon^u(x,f)$, $W^s(x,f)$ e $W^u(x,f)$, respectivamente. O próximo resultado relaciona melhor esses conjuntos, e mais, mostra que localmente esse conjuntos tem um comportamento parecido com E_x^s e E_x^u.

Teorema 1.14. ***(Teorema da Variedade Estável)*** *Sejam $f : M \to M$ um difeomorfismo de classe C^r definido em uma variedade diferenciável M, $\Lambda \subseteq M$ um conjunto hiperbólico e $x \in \Lambda$ um ponto qualquer, então $W^s_\varepsilon(x)$ e $W^u_\varepsilon(x)$ são subvariedades de M tangentes a E^s_x e E^u_x, respectivamente. Além disso,*

$$W^s(x) = \bigcup_{n=0}^{+\infty} f^{-n}\Big(W^s_\varepsilon\big(f^n(x)\big)\Big) \quad e \quad W^u(x) = \bigcup_{n=0}^{+\infty} f^{n}\Big(W^u_\varepsilon\big(f^{-n}(x)\big)\Big)$$

também são subvariedades de M de mesma dimensão que E^s_x e E^u_x, respectivamente, e variam continuamente com o x.

Demonstração. Pode ser encontrada em [**?**], na Seção 8.1.3, Teorema 1.2. □

Um exemplo de dinâmica hiperbólica é o automorfismo hiperbólico no toro $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$, a fim de simplificar os cálculos usaremos o toro de dimensão 2, $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, mas o raciocínio para dimensão $d > 2$ é inteiramente análogo. Seja A uma matriz de ordem 2 cujos elementos são números inteiros, $|det(A)| = 1$ e se λ for um autovalor de A então $|\lambda| \neq 1$. Definamos um operador linear $F : \mathbb{R}^2 \to \mathbb{R}^2$ em que $F(z) = Az$; e seja $\pi : \mathbb{R}^2 \to \mathbb{T}^2$ a projeção natural de $\mathbb{R}^2$ sobre o toro $\mathbb{T}^2$, isto é, $\pi(x, y) = [x, y]$ em que $[x, y]$ é a classe de equivalência de (x, y) (mod 1). Então A induz uma uma função F_A em $\mathbb{T}^2$ definida pelo seguinte diagrama

$$\begin{tikzcd} \mathbb{R}^2 \arrow[r, "F"] \arrow[d, "\pi"'] & \mathbb{R}^2 \arrow[d, "\pi"] \\ \mathbb{T}^2 \arrow[r, "F_A"'] & \mathbb{T}^2 \end{tikzcd}$$

Temos que F_A é um difeomorfismo. De fato, a matriz Jacobiana de F_A em todo ponto $x \in \mathbb{T}^2$ é A, e como $|\det(A)| = 1$ então F_A admite inversa e a Jacobiana da inversa é A^{-1}. Chamamos F_A de **automorfismo hiperbólico no toro**. Apesar do método de construção da função F_A, a proposição seguinte mostra que ela possui um comportamento bem diferente de sua equivalente linear.

Proposição 1.15. *O conjunto $Per(F_A)$ é denso em $\mathbb{T}^2$.*

Demonstração. Seja $p \in \mathbb{T}^2$ um ponto qualquer com coordenadas racionais. Encontrando um denominador comum se necessário, podemos assumir que $p = \left(\dfrac{\alpha}{k}, \dfrac{\beta}{k}\right)$, com $\alpha, \beta, k \in \mathbb{Z}$. Sabemos que os pontos com coordenadas racionais são densos em $\mathbb{T}^2$, então basta mostrarmos que eles são periódicos.

Afirmação: p é periódico, com período menor ou igual a k^2. De fato, assumindo $0 \leq \alpha < k$ e $0 \leq \beta < k$, sabemos que existem k^2 pontos da forma $p = \left(\dfrac{\alpha}{k}, \dfrac{\beta}{k}\right)$. Como a matriz A é inteira, podemos supor $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ com $a, b, c, d \in \mathbb{Z}$, então

$$F_A\left(\frac{\alpha}{k},\frac{\beta}{k}\right)=\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \frac{\alpha}{k} \\ \frac{\beta}{k} \end{bmatrix}=\begin{bmatrix} \frac{a\alpha}{k}+\frac{b\beta}{k} \\ \frac{c\alpha}{k}+\frac{d\beta}{k} \end{bmatrix}=\begin{bmatrix} \frac{a\alpha+b\beta}{k} \\ \frac{c\alpha+d\beta}{k} \end{bmatrix}.$$

Ou seja, F_A permuta os pontos da forma $p=\left(\frac{\alpha}{k},\frac{\beta}{k}\right)$. Logo, existem $m,n\in\mathbb{N}$ tais que $F_A^m(p)=F_A^n(p)$ com $|m-n|\leq k^2$. Supondo, sem perda de generalidade, que $m<n$ e fazendo $\tau(p)=n-m$ temos $F_A^{\tau(p)}(p)=F_A^{n-m}(p)=p$. Assim, p é periódico de período $\tau(p)\leq k^2$. $\square$

Os autovalores λ_1 e λ_2 de A são números reais. De fato, o polinômio característico de A é $p(\lambda)=\lambda^2-\lambda(a+d)+\det(A)$, sendo as raízes $\lambda_j=\dfrac{(a+d)+(-1)^j\sqrt{\Delta}}{2}$, $j=1,2$, em que $\Delta=(a+d)^2-4\det(A)$, então temos dois casos:

Caso 1: $\det(A)=-1$, então $\Delta=(a+d)^2-4\det(A)=(a+d)^2+4>0$ logo $\lambda_1,\lambda_2\in\mathbb{R}$

Caso 2: $\det(A)=ad-bc=1$, suponhamos $\Delta=(a+d)^2-4<0$, então $\lambda_1\in\mathbb{C}$ e $\lambda_1=\dfrac{(a+d)}{2}-\dfrac{i\sqrt{|\Delta|}}{2}$, daí temos

$$|\lambda_1|^2=\frac{(a+d)^2}{4}+\frac{\Delta}{4}=\frac{(a+d)^2}{4}+\frac{-(a+d)^2+4}{4}=1.$$

Absurdo, pois $|\lambda_1|\neq 1$. Portanto $\Delta>0$ e $\lambda_1,\lambda_2\in\mathbb{R}$.

Como $|\det(A)|=|\lambda_1\lambda_2|=1$ e $|\lambda_j|\neq 1$ para $j=1,2$ então um dos autovalores deve satisfazer $|\lambda_i|<1$ e $|\lambda_j|>1$ para $i\neq j$. Vamos denotar por λ_s o primeiro caso, e λ_u o segundo. Desta forma podemos concluir que os subespaços estáveis e instáveis, E_z^s e E_z^u, para todo $z\in\mathbb{T}^2$, devem ser retas passando pela origem em $\mathbb{R}^2$ e com inclinação λ_s e λ_u, respectivamente, e então $T_z\mathbb{T}^2=E_z^s\oplus E_z^u$. É clado que essa decomposição é $DF_{Az}-$invariante, pois $DF_{Az}=A$ para todo $z\in\mathbb{T}^2$ e portando $DF_{Az}(E_z^s)=AE_z^s=E_{F_A(z)}^s$, o mesmo vale para E_z^u. Como F é uma transformação linear, então E_z^s e E_z^u variam continuamente com z. O que nos leva a concluir que T^2 é um conjunto hiperbólico para F_A.

Dado $z\in\mathbb{T}^2$, então pelo Teorema da Variedade Estável 1.14, temos que $W^s(z)$ e $W^u(z)$ são subvariedades de $\mathbb{T}^2$ de dimensão 1 tangentes a E_z^s e E_z^u, respectivamente, e variam continuamente com o z. O que nos leva ao seguinte resultado.

Proposição 1.16. *Para todo ponto $z\in\mathbb{T}^2$, temos*

i) $W^s(z)=\pi(E_z^s)$;

ii) $W^u(z)=\pi(E_z^u)$.

Demonstração. *i)* Se $z'\in W^s(z)$, tomemos l como sendo o segmento de reta que liga z e z'. Suponhamos que $l\nsubseteq E_z^s$, logo l possui uma componente na direção que expande E_z^u. Absurdo, pois $\lim_{n\to+\infty} d\big(F_A^n(z'), F_A^n(z)\big) = 0$. Portanto $l \subseteq E_z^s$, e então $z' \in E_z^s$, ou seja, $W^s(z) \subseteq \pi(E_z^s)$. Reciprocamente, se $z' \in E_z^s$ tal que $z' \neq z$, tomemos l como sendo o segmento de reta que liga z e z'. Pela linearidade de F, temos que $F_A^n(l)$ é um segmento de reta paralelo a W^s e como $|\lambda_s| < 1$ então $\lim_{n\to+\infty} F^n(l) = \lambda_s^n l = 0$, isto é, o tamanho desse segmento tende a zero. Portanto $l \subseteq W^s(z)$, e então $x' \in W^s(z)$, ou seja, $\pi(E_z^s) \subseteq W^s(z)$.

ii) Essa demonstração é inteiramente análoga ao item anterior, usando a inversa da função F_A. □

E o resultado mais interessante desse exemplo, é que para todo $z \in \mathbb{T}^2$, as suas variedades estáveis e instáveis são densas em $\mathbb{T}^2$, o que chamaremos no próximo capítulo de $s-$minimalidade e $u-$minimalidade, respectivamente.

Proposição 1.17. *Dado $z \in \mathbb{T}^2$ um ponto qualquer, então as variedades estável $W^s(z)$ e instável $W^u(z)$ de z são densas em $\mathbb{T}^2$.*

Demonstração. Comecemos por verificar que E_z^s é uma reta com inclinação irracional em $\mathbb{R}^2$, ou seja, se $z = (x_0, y_0)$ e $E_z^s = \big\{(x,y) \in \mathbb{R}^2;\ (y - y_0) = \alpha(x - x_0)\big\}$, então $\alpha \in (\mathbb{R} - \mathbb{Q})$. De fato, suponhamos que $\alpha \in \mathbb{Q}$, então existe um ponto $k = (k_1, k_2) \in E_z^s$ tal que $k_1, k_2 \in \mathbb{Z}$. E como A é uma matriz de elementos inteiros, então $F_A^n(k) \in \mathbb{Z}^2$ para todo $n \in \mathbb{N}$. Absurdo, pois $\lim_{n\to+\infty} d\big(F_A^n(k), F_A^n(z)\big) = 0$.

Agora, tomando os pontos da forma $(k_j, j) \in E_x^s$ que são a intersecção das retas da forma $y = n$, paralelas ao eixo x em $\mathbb{R}^2$ com E_z^s, como a inclinação de E_z^s é irracional, então $k_j \in (\mathbb{R}-\mathbb{Q})$ para todo $j \in \mathbb{Z}$, e portando $\pi(k_j, j) = (\alpha_j, 0)$ para algum irracional $0 < \alpha < 1$. Os pontos da forma $(x, 0) \in \mathbb{T}^2$ define um circulo no toro e suas imagens por F_A são uma rotação por um ângulo irracional, então pela Proposição 1.12 temos que a intersecção de $W^s(z) = \pi(E_z^s)$ com esse círculo é denso nele, ou seja, $W^s(z)$ é densa verticalmente no toro. Basta verificarmos então para a segunda coordenada.

Tomando os pontos de $W^s(z)$ que interceptam as retas da forma $y = n + \beta$, em que $\beta \in \mathbb{Q} \cap (0,1)$, temos que para cada β a variedade estável $W^s(z)$ é densa no circulo formados pelos pontos da forma (x, β), e como os racionais são densos em $[0,1)$, os círculos também são densos no toro. Portanto $W^s(z)$ é denso em $\mathbb{T}^2$.

Para $W^u(z)$ a demonstração é análoga. □

1.2 Dinâmica Ergódica

Nessa seção vamos estudar uma dinâmica sob o olhar da teoria ergódica, cujo foco são as dinâmicas que preservam uma medida. Começaremos vendo as condições necessárias pra garantir a existência de pontos recorrentes e de medidas invariantes, para podermos definir ergodicidade e provar seus resultados.

O resultado que garante que quase todo ponto é recorrente, relativamente a uma medida finita $f-$invariante, é o Teorema da Recorrência de Poincaré. Usaremos com frequência, para estudarmos medida de um conjunto, a **função característica** $\mathcal{X}_E : M \to \mathbb{R}$ do conjunto E, isto é, $\mathcal{X}_E(x) = 1$ se $x \in E$, e $\mathcal{X}_E(x) = 0$ se $x \notin E$. Definimos $\mu : \mathcal{B}(M) \to \overline{\mathbb{R}}$ uma **medida** em M, onde $\overline{\mathbb{R}} = \mathbb{R} \cup \big\{-\infty, +\infty\big\}$ é a reta estendida. Denotaremos essa medida apenas por μ; dizemos que μ é uma **medida finita** se $\mu(M) < +\infty$ e μ é uma **probabilidade** se $\mu(M) = 1$. Dizemos que μ é $f-$**invariante** se para todo conjunto $E \subseteq M$ mensurável vale $\mu\big(f^{-1}(E)\big) = \mu(E)$.

Dizemos que uma propriedade P vale para $\mu-$**quase todo ponto** $x \in E$, quando existe um conjunto $N \subseteq E$ com $\mu(N) = 0$, tal que P vale para todo ponto $x \in (E - N)$. Finalmente podemos enunciar o primeiro resultado.

Teorema 1.18. ***(Teorema de Recorrência de Poincaré - versão mensurável)*** *Sejam $f : M \to M$ uma aplicação mensurável e μ uma medida $f-$invariante finita. Se $E \subseteq M$ é um conjunto mensurável qualquer com $\mu(E) > 0$, então para $\mu-$quase todo ponto $x \in E$, existe $n \in \mathbb{N}$ tal que $f^n(x) \in E$.*

Demonstração. Chamemos $E^0 \subseteq E$ o conjunto dos pontos $x \in E$ que não retornam a E, ou seja, $x \in E^0$ implica que para todo $n \in \mathbb{N}_*$ temos que $f^n(x) \notin E$. Basta provarmos que E^0 tem medida nula, e concluímos a demonstração.

Afirmação: As suas pré-imagens $f^{-n}\big(E^0\big)$ são duas a duas disjuntas. De fato, suponhamos que existam $m, n \in \mathbb{N}_*$ com $m > n$ tais que $f^{-m}\big(E^0\big) \cap f^{-n}\big(E^0\big) \neq \emptyset$, tomemos x um ponto dessa intersecção e seja $y = f^n(x)$. Então $y \in E^0$ e $f^{m-n}(y) = f^{m-n} \circ f^n(x) = f^m(x) \in E^0 \subseteq E$, isso significa que y retorna a E, o que é absurdo pois $y \in E^0$. Logo as pré-imagens de E^0 por f são duas a duas disjuntas. Então, pela $\sigma-$aditividade de μ, temos

$$\mu\left(\bigcup_{n=0}^{+\infty} f^{-n}\big(E^0\big)\right) = \sum_{n=0}^{+\infty} \mu\Big(f^{-n}\big(E^0\big)\Big) = \sum_{n=0}^{+\infty} \mu\big(E^0\big).$$

Na última igualdade usamos a hipótese de que μ é f-invariante, ou seja, $\mu\Big(f^{-n}\big(E^0\big)\Big) = \mu\big(E^0\big)$ para todo $n \in \mathbb{N}$. Como μ é finita, temos que $\mu\Big(\bigcup_{n=0}^{+\infty} f^{-n}\big(E^0\big)\Big) < +\infty$, por outro lado, à direita temos que $\sum_{n=0}^{+\infty}\mu\big(E^0\big)$ é uma soma infinita de termos constantes. A única forma dessa soma ser finita, é se todas as suas parcelas forem nulas. Portanto concluímos que $\mu\big(E^0\big)=0$, como queríamos provar. $\square$

Esse resultado têm uma consequência direta mais forte, em que além de $\mu-$ quase todo ponto voltar a E, eles continuam voltando uma infinidade de vezes.

Corolário 1.19. *Nas condições do Teorema 1.18, para $\mu-$quase todo ponto $x\in E$, existem infinitos valores $n\in\mathbb{N}$, tais que $f^n(x)\in E$.*

Demonstração. Chamemos $E_k\subseteq E$ o conjunto dos pontos $x\in E$ que retornam a E exatamente k vezes, ou seja, $x\in E_k$ implica que existem exatamente $n_1,n_2,\cdots,n_k\in\mathbb{N}$, tais que $f^{n_i}(x)\in E$, onde $i\in\big\{1,2,\cdots,k\big\}$. Logo o conjunto dos pontos que retornam a E um número finito de vezes é $\bigcup_{k=1}^{+\infty}E_k$ e então

$$\mu\left(\bigcup_{k=1}^{+\infty}E_k\right)\leq\sum_{k=1}^{+\infty}\mu\big(E_k\big).$$

Basta provarmos que E_k tem medida nula, para todo $k\in\mathbb{N}$, e concluímos a demonstração do corolário. Suponhamos que $\mu\big(E_k\big)>0$, então pelo Teorema 1.18 temos que para $\mu-$quase todo ponto $x\in E_k$ existe $n\in\mathbb{N}$ tal que $f^n(x)\in E_k$. Fixemos um desses x e denotemos $y=f^n(x)\in E_k$. Pela definição de E_k, temos que y tem exatamente k iterados em E_k, mas como y é um iterado de x, então x tem pelo menos $k+1$ iterados em E_k o que é absurdo. Logo $\mu\big(E_k\big)=0$, como queríamos provar. $\square$

Na sequência vamos mostrar uma versão topológica desse resultado, que é útil para relacionarmos com os resultados da seção de dinâmica topológica.

Teorema 1.20. ***(Teorema de Recorrência de Poincaré - versão topológica)*** *Sejam $f:M\to M$ uma aplicação mensurável e μ uma medida $f-$invariante finita. Se M admite uma base enumerável de abertos, então $\mu-$quase todo ponto $x\in M$ é recorrente.*

Demonstração. Seja $B=\{U_k;\ k\in\mathbb{N}\}$ uma base enumerável de abertos de M. Para cada $k\in\mathbb{N}$, representaremos por U_k^0 o conjunto dos pontos $x\in U_k$ que nunca regressam a U_k, ou seja, $x\in U_k^0$ implica que para todo $n\in\mathbb{N}$ temos que $f^n(x)\notin U_k$. Então, pelo Teorema 1.18, U_k^0 tem medida nula pra todo k, e portanto se $\tilde{U}=\bigcup_{k=1}^{+\infty}U_k^0$, temos que

$$0\leq\mu\big(\tilde{U}\big)=\mu\left(\bigcup_{k=1}^{+\infty}U_k^0\right)\leq\sum_{k=1}^{+\infty}\mu\big(U_k^0\big)=0.$$

Logo $\mu(\tilde{U}) = 0$. Então basta provarmos que para todo $x \in M$ tal que $x \notin \tilde{U}$, temos que x é recorrente, e assim concluímos a demonstração. Sejam $x \in (M - \tilde{U})$ e V_x uma vizinhança aberta qualquer de x, como B é uma base de M então existe $U_k \in B$ tal que $x \in U_k \subseteq V_x$. Como $x \notin \tilde{U}$, então $x \notin U_k^0$, ou seja, existe um $n \in \mathbb{N}$, tal que $f^n(x) \in U_k \subseteq V_x$. Como V_x é uma vizinhança arbitrária, então x é um ponto recorrente, como queríamos mostrar. □

Esses resultados acima dependem fortemente da medida ser $f-$invariantes, o que será garantido pelo próximo teorema.

Teorema 1.21. ***(Teorema da Existência de Medidas $f-$invariantes)*** *Seja M um espaço métrico compacto. Se $f : M \to M$ é uma aplicação contínua, então existe pelo menos uma probabilidade $f-$invariante.*

Demonstração. Pode ser encontrada em [?], no Capítulo 2, Teorema 2.1. □

Podemos então concluir que com hipóteses relativamente fracas, conseguimos sempre encontrar pontos recorrentes em uma dinâmica.

Corolário 1.22. ***(Teorema da Recorrência de Birkhoff)*** *Seja M um espaço métrico compacto. Se $f : M \to M$ é uma aplicação contínua, então f tem algum ponto recorrente.*

Demonstração. Pelo Teorema 1.21, existe uma probabilidade $f-$invariante μ. Como M é compacto, então admite uma base enumerável de abertos, logo pelo Teorema 1.20 $\mu-$quase todo ponto $x \in M$ é recorrente. Em particular o conjunto dos pontos recorrentes é não vazio, concluindo a demonstração. □

O próximo resultado é o Teorema Ergódico de Birkhoff que nos permitirá um bom entendimento das dinâmicas em termos da densidade de suas órbitas em relação a uma medida. Sejam $f : M \to M$ uma aplicação mensurável, $x \in M$ um ponto qualquer e $E \subseteq M$ um conjunto mensurável, vamos fixar $n \in \mathbb{N}$ e definir $I_n = \{0, 1, \cdots, n-1\} \subseteq \mathbb{N}$, o conjunto dos n primeiros números naturais, e então vamos considerar $\tau_n(E, x)$ como sendo a fração dos $j \in I_n$ tal que $f^j(x) \in E$, ou seja, $\tau_n(E, x) = \frac{1}{n}\#\{f^j(x) \in E;\ j \in I_n\}$, onde $\#\{f^j(x) \in E;\ j \in I_n\}$ é a cardinalidade do conjunto. Observe que podemos rescrever $\tau_n(E, x)$ da seguinte forma

$$\tau_n(E, x) = \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{X}_E(f^j(x)). \tag{1.1}$$

Definição 1.23. *Sejam $f : M \to M$ uma plicação mensurável e $x \in M$ um ponto qualquer. Definimos $\tau(E, x)$, o* ***tempo médio de permanência*** *da órbita de x em E, como sendo o limite de $\tau_n(E, x)$ quando n tende ao infinito, ou seja,*

$$\tau(E, x) = \lim_{n \to +\infty} \tau_n(E, x).$$

Em geral, esse limite pode não existir. Mas quando existe, podemos garantir que ele não varia na órbita do ponto.

Lema 1.24. *Sejam $f : M \to M$ uma plicação mensurável e $x \in M$ um ponto qualquer. Se o tempo médio de permanência $\tau(E,x)$ existe, então*

$$\tau\big(E, f(x)\big) = \tau(E,x).$$

Demonstração. Por definição temos

$$\begin{aligned}
\tau\big(E, f(x)\big) &= \lim_{n\to+\infty} \tau_n\big(E, f(x)\big) \\
&= \lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{X}_E\Big(f^j\big(f(x)\big)\Big) \\
&= \lim_{n\to+\infty} \frac{1}{n}\sum_{j=1}^{n} \mathcal{X}_E\big(f^j(x)\big) \\
&= \lim_{n\to+\infty} \frac{1}{n}\left(\sum_{j=0}^{n-1} \mathcal{X}_E\big(f^j(x)\big) - \frac{1}{n}\Big[\mathcal{X}_E(x) - \mathcal{X}_E\big(f^n(x)\big)\Big]\right) \\
&= \lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{X}_E\big(f^j(x)\big) - \lim_{n\to\infty} \frac{1}{n}\Big[\mathcal{X}_E(x) - \mathcal{X}_E\big(f^n(x)\big)\Big] \\
&= \tau(E,x) - \lim_{n\to+\infty} \frac{1}{n}\Big[\mathcal{X}_E(x) - \mathcal{X}_E\big(f^n(x)\big)\Big].
\end{aligned}$$

Como a função característica é limitada, esse ultimo limite é igual a zero, e o lema está demonstrado. □

Com esses resultados, podemos provar o próximo teorema, que garante a existência desse limite e dá um método para calcular sua integral.

Teorema 1.25. *Sejam $f : M \to M$ uma aplicação mensurável e μ uma probabilidade $f-$invariante. Dado qualquer conjunto mensurável $E \subseteq M$, o tempo médio de permanência $\tau(E,x)$ existe para $\mu-$quase todo ponto $x \in M$. Além disso,*

$$\int_M \tau(E,x)d\mu(x) = \mu(E). \tag{1.2}$$

Demonstração. Seja $E \subseteq M$ um conjunto mensurável qualquer. Para cada $x \in M$, definamos

$$\begin{aligned}
\overline{\tau}(E,x) &= \limsup_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{X}_E\big(f^j(x)\big) \text{ e} \\
\underline{\tau}(E,x) &= \liminf_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{X}_E\big(f^j(x)\big).
\end{aligned}$$

Para todo $x \in M$ temos que

$$\overline{\tau}\big(E, f(x)\big) = \overline{\tau}(E,x) \quad \text{e} \quad \underline{\tau}\big(E, f(x)\big) = \underline{\tau}(E,x). \tag{1.3}$$

A demonstração das equações em (1.3) é análoga a do Lema 1.24.

Para demonstrar a existência do tempo médio $\tau(E,x)$, basta mostrar que para $\mu-$quase todo ponto $x \in M$, temos

$$\overline{\tau}(E,x) = \underline{\tau}(E,x). \tag{1.4}$$

Como $0 \leq \underline{\tau}(E,x) \leq \overline{\tau}(E,x)$ para todo $x \in M$, então $\int_M \underline{\tau}(E,x)d\mu(x) \leq \int_M \overline{\tau}(E,x)d\mu(x)$. E caso $\tau(E,x)$ exista, pela definição de $\liminf$ e $\limsup$ teremos $0 \leq \underline{\tau}(E,x) \leq \tau(E,x) \leq \overline{\tau}(E,x)$ para todo $x \in M$, e então $\int_M \underline{\tau}(E,x)d\mu(x) \leq \int_M \tau(E,x)d\mu(x) \leq \int_M \overline{\tau}(E,x)d\mu(x)$. Logo, pra demonstrarmos a igualdade dada no teorema, basta provarmos a seguinte desigualdade (1.5), e concluímos a demonstração.

$$\int_M \underline{\tau}(E,x)d\mu(x) \geq \mu(E) \geq \int_M \overline{\tau}(E,x)d\mu(x). \tag{1.5}$$

Vamos provar a segunda desigualdade em (1.5). Seja $\varepsilon > 0$ dado, por definição de $\limsup$ existem $t \in \mathbb{N}$, tais que para todo $x \in M$, temos

$$\tau_t(E,x) \geq \overline{\tau}(E,x) - \varepsilon. \tag{1.6}$$

Definamos $t : M \to \mathbb{N}$ uma função que leva o ponto $x \in M$ ao primeiro t que satisfaça (1.6). Agora dividiremos a demonstração em dois casos.

Caso Particular: Suponhamos que a função t seja limitada, ou seja, existe um $K \in \mathbb{N}$ tal que $t(x) \leq K$ para todo $x \in M$. Fixando um $n \in \mathbb{N}$ e dado $x \in M$, definamos uma sequência $x_0, x_1, \cdots, x_s$ de pontos de M e uma sequência $t_0, t_1, \cdots, t_s$ de número naturais, do seguinte modo:

1. Tomemos $x_0 = x$.

2. Depois fazemos $t_i = t(x_i)$ e $x_{i+1} = f^{t_i}(x_i)$.

3. Terminamos quando encontrarmos x_s tal que $t_0 + t_1 + \cdots + t_s \geq n$.

Pela definição de $\tau_t(E,x)$ em (1.1), temos

$$\tau_{t_i}(E,x) = \frac{1}{t_i}\sum_{j=0}^{t_i-1} \mathcal{X}_E\big(f^j(x)\big) \quad \Longrightarrow \quad \sum_{j=0}^{t_i-1} \mathcal{X}_E\big(f^j(x)\big) = t_i\tau_{t_i}(E,x).$$

Como essa equação vale para todo $x \in M$ em particular vale para todo x_i, e aplicando em (1.6), temos

$$\begin{aligned}\sum_{j=0}^{t_i-1} \mathcal{X}_E\big(f^j(x_i)\big) &= t_i \tau_{t_i}(E, x_i) \\ &\geq t_i\big(\overline{\tau}(E, x_i) - \varepsilon\big). \end{aligned} \tag{1.7}$$

Pela definição da sequência x_i temos que

$$\begin{aligned} x_0 &= x \\ x_1 &= f^{t_0}(x_0) = f^{t_0}(x) \\ x_2 &= f^{t_1}(x_1) = f^{t_1}\big(f^{t_0}(x)\big) = f^{t_0+t_1}(x) \\ x_3 &= f^{t_2}(x_2) = f^{t_2}\big(f^{t_0+t_1}(x)\big) = f^{t_0+t_1+t_2}(x) \\ &\vdots \\ x_s &= f^{t_0+t_1+\cdots+t_{s-1}}(x). \end{aligned}$$

E pelo item 3 da definição das sequências, temos que $t_0 + t_1 + \cdots + t_s \geq n$, então $t_0 + t_1 + \cdots + t_{s-1} \geq n - t_s$, e como todo $t_i = t(x_i) \leq K$ temos que $t_0 + t_1 + \cdots + t_{s-1} \geq n - K$. Note que aplicando $x_i = f^{t_0+t_1+\cdots+t_i}(x)$ em (1.3), temos $\overline{\tau}(E, x_i) = \overline{\tau}(E, x)$ para todo x_i da sequência. Como t_s é o menor número tal que $t_0 + t_1 + \cdots + t_s \geq n$, se tirarmos ele dessa soma teremos, $t_0 + t_1 + \cdots + t_{s-1} < n$ o que implica na seguinte desigualdade $t_0 + t_1 + \cdots + t_{s-1} - 1 < n - 1$. Então podemos reescrever (1.7), colocando x_i em função de x para todo x_i da sequência, e somar todos eles, em que

$$\begin{aligned}\sum_{j=0}^{n-1} \mathcal{X}_E\big(f^j(x)\big) &\geq \sum_{j=0}^{t_0+\cdots+t_{s-1}-1} \mathcal{X}_E\big(f^j(x)\big) \\ &\geq (t_0 + t_1 + \cdots + t_{s-1})\big(\overline{\tau}(E, x) - \varepsilon\big) \\ &\geq (n - K)\big(\overline{\tau}(E, x) - \varepsilon\big). \end{aligned} \tag{1.8}$$

Como o $x \in M$ é um qualquer, então essa desigualdade vale para todo $x \in M$, e como as funções características são integráveis e não negativas, a integral preserva a desigualdade e temos

$$
\begin{aligned}
\int_M \sum_{j=0}^{n-1} \mathcal{X}_E\big(f^j(x)\big)d\mu(x) &\geq \int_M (n-K)\big(\overline{\tau}(E,x)-\varepsilon\big)d\mu(x) \\
\sum_{j=0}^{n-1} \int_M \mathcal{X}_E\big(f^j(x)\big)d\mu(x) &\geq (n-K)\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\int_M d\mu(x)\right) \\
\sum_{j=0}^{n-1} \mu(E) &\geq (n-K)\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\mu(M)\right) \\
n\mu(E) &\geq (n-K)\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\right) \\
\mu(E) &\geq \frac{(n-K)}{n}\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\right).
\end{aligned}
\tag{1.9}
$$

No membro da esquerda em (1.9), todas as parcelas da soma é igual a $\mu(E)$ pois μ é $f-$invariante. Esse resultado vale para todo $n \in \mathbb{N}$, então passando ao limite quando $n \to +\infty$, temos

$$
\begin{aligned}
\mu(E) &\geq \lim_{n\to+\infty}\left[\frac{(n-K)}{n}\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\right)\right] \\
&= \lim_{n\to+\infty}\left[\frac{(n-K)}{n}\right]\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\right) \\
&= \int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon.
\end{aligned}
$$

Esse resultado vale para todo $\varepsilon > 0$, então podemos passar ao limite quando $\varepsilon \to 0$, e temos

$$\mu(E) \geq \int_M \overline{\tau}(E,x)d\mu(x).$$

Terminando assim a demonstração para esse caso, onde a função t é limitada.

Caso Geral: Quando t for ilimitada, partindo do $\varepsilon > 0$ dado em (1.6), fixemos $K \in \mathbb{N}$, suficientemente grande, de modo que o conjunto $B = \big\{y \in M;\ t(y) > K\big\}$ seja tal que $\mu(B) < \varepsilon$.

Vamos mostrar que esse K de fato existe. Definamos, para todo $m \in \mathbb{N}$, $A_m = \big\{y \in M;\ t(y) \leq m\big\}$. Temos que $A_m \subseteq A_{m+1}$ e $\bigcup_{m=0}^{+\infty} A_m = M$, então $\lim_{m\to+\infty}\mu(A_m) = \mu\left(\bigcup_{m=0}^{+\infty} A_m\right)$ $= \mu(M) = 1$, ou seja, para todo $\delta > 0$ dado, existe $m_0 \in \mathbb{N}$, tal que $m > m_0$ implica $\mu(A_m) > 1-\delta$. Tomemos então $K > m_0$ e $\delta = \varepsilon$, então $\mu(A_K) > 1-\varepsilon$, isso implica $\mu(M - A_K) < \varepsilon$. Agora observe que $(M - A_K) = \big\{y \in M;\ t(y) \leq K\big\} = B$.

De maneira similar ao caso particular, fixando um $n \in \mathbb{N}$ e dado $x \in M$, definamos uma sequência $x_0, x_1, \cdots, x_s$ de pontos de M e uma sequência $t_0, t_1, \cdots, t_s$ de número naturais, do seguinte modo:

1. Tomemos $x_0 = x$.

2. Se $t(x_i) \leq K$, fazemos $t_i = t(x_i)$ e $x_{i+1} = f^{t_i}(x_i)$.

3. Se $t(x_i) > K$, fazemos $t_i = 1$ e $x_{i+1} = f(x_i)$.

4. Terminamos quando encontrarmos x_s tal que $t_0 + t_1 + \cdots + t_s \geq n$.

Do caso particular, temos que para todo $i \in \mathbb{N}$, tal que $t(x_i) \leq K$, a desigualdade (1.7) continua valendo

$$\sum_{j=0}^{t_i-1} \mathcal{X}_E\big(f^j(x_i)\big) \geq t_i\big(\overline{\tau}(E,x) - \varepsilon\big). \tag{1.10}$$

A partir da desigualdade acima podemos escrever a seguinte

$$\sum_{j=0}^{t_i-1} \mathcal{X}_E\big(f^j(x_i)\big) \geq t_i\big(\overline{\tau}(E,x) - \varepsilon\big) - \sum_{j=0}^{t_i-1} \mathcal{X}_B\big(f^j(x_i)\big). \tag{1.11}$$

Essa desigualdade tem a vantagem de valer para todos os x_i. De fato, basta vermos que quando x_i for tal que $t(x_i) \leq K$, o ultimo somatório fica igual a zero, e decorre diretamente de (1.10), e quando $t(x_i) > K$, temos que $t_i = 1$ e esses somatórios terão apenas um elemento, ficando

$$\begin{aligned} \sum_{j=0}^{t_i-1} \mathcal{X}_E\big(f^j(x_i)\big) &\geq t_i\big(\overline{\tau}(E,x_i) - \varepsilon\big) - \sum_{j=0}^{t_i-1} \mathcal{X}_B\big(f^j(x_i)\big) \\ \mathcal{X}_E\big(f(x_i)\big) &\geq \big(\overline{\tau}(E,x_i) - \varepsilon\big) - \mathcal{X}_B\big(f(x_i)\big). \end{aligned}$$

E temos que $\big(\overline{\tau}(E,x_i) - \varepsilon\big) < 1$ pois $\overline{\tau}(E,x_i) \leq 1$ e $\varepsilon > 0$, e como $\mathcal{X}_B\big(f(x_i)\big) = 1$ pela definição de B e pela escolha do x_i, então podemos concluir que a desigualdade é verdadeira, pois o membro da esquerda é maior do que ou igual a zero, pela definição de função característica, e o da direita é menor que zero.

Agora usando o mesmo método que usamos pra concluir (1.8), fazendo novamente $x_i = f^{t_0+t_1+\cdots+t_{i-1}}(x)$ para todo x_i da sequência. Temos que $t_0 + t_1 + \cdots + t_{s-1} \geq n - t_s \geq n - K$, pois pelo o item 4 da definição da sequência, $t_i \leq K$ para todo t_i. Aplicando novamente $x_i = f^{t_0+t_1+\cdots+t_i}(x)$ em (1.3), temos $\overline{\tau}(E,x_i) = \overline{\tau}(E,x)$ para todo x_i da sequência. E temos também que vale a desigualdade $t_0 + t_1 + \cdots + t_{s-1} - 1 < n - 1$, então podemos generalizar,

$$\begin{aligned} \sum_{j=0}^{n-1} \mathcal{X}_E(f^j(x)) &\geq \sum_{j=0}^{t_0+\cdots+t_{s-1}-1} \mathcal{X}_E(f^j(x)) \\ &\geq (t_0 + t_1 + \cdots + t_{s-1})\big(\overline{\tau}(E,x) - \varepsilon\big) - \sum_{j=0}^{t_0+\cdots+t_{s-1}-1} \mathcal{X}_B\big(f^j(x)\big) \\ &\geq (n-K)\big(\overline{\tau}(E,x) - \varepsilon\big) - \sum_{j=0}^{n-1} \mathcal{X}_B\big(f^j(x)\big). \end{aligned}$$

Como o $x \in M$ é um qualquer, então essa desigualdade vale para todo $x \in M$, e como as funções características são integráveis e não negativas, a integral preserva a desigualdade e temos

$$\begin{aligned}
\int_M \sum_{j=0}^{n-1} \mathcal{X}_E\big(f^j(x)\big)d\mu(x) &\geq \int_M \left((n-K)\big(\overline{\tau}(E,x)-\varepsilon\big) - \sum_{j=0}^{n-1} \mathcal{X}_B\big(f^j(x)\big)\right) d\mu(x) \\
\sum_{j=0}^{n-1} \int_M \mathcal{X}_E\big(f^j(x)\big)d\mu(x) &\geq (n-K)\int_M \big(\overline{\tau}(E,x)-\varepsilon\big)d\mu(x) - \sum_{j=0}^{n-1}\int_M \mathcal{X}_B\big(f^j(x)\big)d\mu(x) \\
\sum_{j=0}^{n-1} \mu(E) &\geq (n-K)\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\right) - \sum_{j=0}^{n-1}\mu(B) \\
n\mu(E) &\geq (n-K)\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\right) - n\mu(B) \\
\mu(E) &\geq \frac{(n-K)}{n}\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\right) - \mu(B).
\end{aligned}$$

Como esse resultado vale para todo $n \in \mathbb{N}$, então passando ao limite quando $n \to +\infty$, e lembrando que $\mu(B) < \varepsilon$, temos

$$\begin{aligned}
\mu(E) &\geq \lim_{n\to+\infty}\left[\frac{(n-K)}{n}\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\right)-\varepsilon\right] \\
&= \lim_{n\to+\infty}\left[\frac{(n-K)}{n}\right]\left(\int_M \overline{\tau}(E,x)d\mu(x)-\varepsilon\right)-\varepsilon \\
&= \int_M \overline{\tau}(E,x)d\mu(x) - 2\varepsilon.
\end{aligned}$$

Esse resultado vale para todo $\varepsilon > 0$, então podemos passar ao limite quando $\varepsilon \to 0$, e temos

$$\mu(E) \geq \int_M \overline{\tau}(E,x)d\mu(x). \tag{1.12}$$

Isso completa a demonstração do caso geral, para a segunda desigualdade em (1.5). E para a primeira, basta notarmos que $\underline{\tau}(E,x) = 1 - \overline{\tau}(M-E,x)$. De fato, pois $\mathcal{X}_{M-E} = 1 - \mathcal{X}_E$ o que implica $\mathcal{X}_E = 1 - \mathcal{X}_{M-E}$, logo

$$\begin{aligned}
\underline{\tau}(E,x) &= \liminf_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n-1}\mathcal{X}_E\big(f^j(x)\big) \\
\underline{\tau}(E,x) &= \liminf_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n-1}\Big(1-\mathcal{X}_{M-E}\big(f^j(x)\big)\Big) \\
\underline{\tau}(E,x) &= 1-\limsup_{n\to+\infty}\frac{1}{n}\sum_{j=0}^{n-1}\mathcal{X}_{M-E}\big(f^j(x)\big) \\
\underline{\tau}(E,x) &= 1-\overline{\tau}(M-E,x).
\end{aligned}$$

Então $\overline{\tau}(M-E,x)=1-\underline{\tau}(E,x)$, e aplicando o conjunto mensurável $M-E$ na desigualdade (1.12) temos

$$\begin{aligned}
\int_M \overline{\tau}(M-E,x)d\mu(x) &\leq \mu(M-E)\\
\int_M 1-\underline{\tau}(E,x)d\mu(x) &\leq 1-\mu(E)\\
1-\int_M \underline{\tau}(E,x)d\mu(x) &\leq 1-\mu(E)\\
\int_M \underline{\tau}(E,x)d\mu(x) &\geq \mu(E).
\end{aligned}$$

Portanto mostramos que $\int_M \tau(E,x)d\mu(x) = \mu(E)$, o que garante a existência do tempo médio para $\tau(E,x)$ para $\mu-$quase todo ponto $x \in M$, provando assim o teorema. $\square$

Para uma aplicação mensurável $f : M \to M$, uma função integrável $\varphi : M \to \mathbb{R}$, e um ponto qualquer $x \in M$. Definimos $\tilde{\varphi}(x)$, a **média temporal** da órbita de x pelo potencial φ, como sendo o seguinte limite

$$\tilde{\varphi}(x) = \lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi\big(f^j(x)\big).$$

Em geral esse limite pode não existir. Definimos também $\overline{\varphi}$, a **média espacial** da função φ em M, como sendo

$$\overline{\varphi} = \frac{1}{\mu(M)}\int_M \varphi \; d\mu.$$

Um caso mais geral do Teorema 1.25, conhecido como Teorema Ergódico de Birkhoff, é um dos resultados principais dessa seção, e será enunciado a seguir.

Teorema 1.26. (Teorema Ergódico de Birkhoff) *Sejam $f : M \to M$ uma aplicação mensurável e μ uma probabilidade $f-$invariante. Dada qualquer função integrável $\varphi : M \to \mathbb{R}$, a média temporal $\tilde{\varphi}(x)$ existe em $\mu-$quase todo ponto $x \in M$. Além disso,*

$$\int_M \tilde{\varphi}\; d\mu = \int_M \varphi \; d\mu.$$

Demonstração. Este enunciado mais geral pode ser provado usando uma versão um pouco mais elaborada do argumento usado pra provar o Teorema 1.25, que pode ser encontrada em [?], no Capítulo 3, Teorema 3.2.3. $\square$

O Teorema 1.25 é o caso particular do Teorema Ergódico de Birkhoff 1.26 quando $\varphi = \mathcal{X}_E$, a função característica do conjunto E.

1.2.1 Ergodicidade

Dizemos que uma aplicação $f : M \to M$ é **ergódica** para uma probabilidade $f-$invariante μ (também dizemos que a probabilidade μ é ergódica pra f, ou que o sistema (f, μ) é ergódico) se as médias temporais coincidirem $\mu-$quase todo ponto $x \in M$ com as respectivas médias espaciais, ou seja, $\tilde{\varphi}(x) = \overline{\varphi}$ para $\mu-$quase todo ponto $x \in M$ e para toda função integrável $\varphi : M \to \mathbb{R}$. Uma função $\psi : M \to \mathbb{R}$ é dita $f-$**invariante** se $(\psi \circ f)(x) = \psi(x)$ para $\mu-$quase todo ponto $x \in M$.

Proposição 1.27. *Sejam $f : M \to M$ uma aplicação mensurável e $\varphi : M \to \mathbb{R}$ uma função integrável. Então a média temporal $\tilde{\varphi}$ é $f-$invariante.*

Demonstração. Para demonstrarmos essa Proposição, precisaremos enunciar o seguinte Lema, cuja demonstração é encontrada em [?], no Capítulo 3, Lema 3.2.5.

Lema 1.28. *Sejam $f : M \to M$ uma aplicação mensurável e $\phi : M \to \mathbb{R}$ uma função integrável, então $\lim_{n\to+\infty} \dfrac{1}{n}\phi\big(f^n(x)\big) = 0$ para $\mu-$quase todo ponto $x \in M$.*

Então, sejam $\varphi : M \to \mathbb{R}$ um função integrável qualquer e a sua média temporal

$$\tilde{\varphi}(x) = \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\big(f^j(x)\big).$$

Fixemos um $x \in M$, e temos

$$\begin{aligned}
\big(\tilde{\varphi} \circ f\big)(x) &= \tilde{\varphi}\big(f(x)\big) \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\Big(f^j\big(f(x)\big)\Big) \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\big(f^{j+1}(x)\big) \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{j=1}^{n} \varphi\big(f^j(x)\big) \\
&= \lim_{n\to+\infty} \frac{1}{n} \left[\sum_{j=0}^{n-1} \varphi\big(f^j(x)\big) + \varphi\big(f^n(x)\big) - \varphi\big(f^0(x)\big)\right] \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi\big(f^j(x)\big) + \lim_{n\to+\infty} \frac{1}{n}\varphi\big(f^n(x)\big) - \lim_{n\to+\infty} \frac{1}{n}\varphi(x) \\
&= \tilde{\varphi}(x) + \lim_{n\to+\infty} \frac{1}{n}\varphi\big(f^n(x)\big).
\end{aligned}$$

E pelo Lema 1.28 temos $\lim_{n\to+\infty} \dfrac{1}{n}\varphi\big(f^n(x)\big) = 0$ para $\mu-$quase todo ponto $x \in M$. Portanto $\tilde{\varphi}$ é $f-$invariante. $\square$

Existem várias maneiras equivalentes de definir a ergodicidade de uma aplicação, o próximo teorema relaciona alguma delas.

Teorema 1.29. *Sejam $f : M \to M$ uma aplicação mensurável e μ uma probabilidade $f-$invariante. São equivalentes:*

i) O sistema (f, μ) é ergódico.

ii) Se $E \subseteq M$ é um conjunto mensurável $f-$invariante, então $\mu(E) = 0$ ou $\mu(E) = 1$.

iii) Se $\psi : M \to \mathbb{R}$ é uma função mensurável $f-$invariante, então ψ é constante para $\mu-$quase todo ponto $x \in M$.

Demonstração. $i) \Rightarrow ii)$ Seja $E \subseteq M$ um conjunto mensurável, como f é ergódica, então $\tilde{\varphi}(x) = \int_M \varphi \, d\mu$ para toda função φ integrável e para $\mu-$quase todo ponto $x \in M$. Tomemos $\varphi = \mathcal{X}_E$, a função característica do conjunto E, então

$$\tilde{\mathcal{X}}_E(x) = \int_M \mathcal{X}_E d\mu = \mu(E)$$

Temos também que $f(E) = E$, pois E é $f-$invariante, então

$$\begin{aligned}
\mu(E) &= \tilde{\mathcal{X}}_E(x) \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_E\big(f^j(x)\big) \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{X}_E(x) \\
&= \lim_{n\to+\infty} \frac{1}{n} \big(n\mathcal{X}_E(x)\big) \\
&= \mathcal{X}_E(x)
\end{aligned}$$

E como $\mathcal{X}_E(x)$ só assume valores em $\{0, 1\}$, então $\mu(E) = 0$ ou $\mu(E) = 1$.

$ii) \Rightarrow iii)$ Sejam $\psi : M \to \mathbb{R}$ uma função mensurável $f-$invariante e $c \in \mathbb{R}$ uma constante qualquer, então os conjuntos $\psi^{-1}(c) \subseteq M$ são $f-$invariantes. De fato, se $x \in \psi^{-1}(c)$ então $\psi(x) = c$, e como ψ é $f-$invariante temos que $\psi\big(f(x)\big) = \psi(x) = c$, logo $f(x) \in \psi^{-1}(c)$. O que implica que $f\big(\psi^{-1}(c)\big) \subseteq \psi^{-1}(c)$. Reciprocamente, se $f(x) \in f\big(\psi^{-1}(c)\big)$ então $\psi\big(f(x)\big) = c$, e como ψ é $f-$invariante temos que $\psi(x) = \psi\big(f(x)\big) = c$, logo $x \in \psi^{-1}(c)$. O que implica que $\psi^{-1}(c) \subseteq f\big(\psi^{-1}(c)\big)$. Portando $\psi^{-1}(c) = f\big(\psi^{-1}(c)\big)$.

Logo, por hipótese temos que $\mu\big(\psi^{-1}(c)\big) = 0$ ou $\mu\big(\psi^{-1}(c)\big) = 1$, e como c é uma constante qualquer, então existe $c \in \mathbb{R}$ tal que $\mu\big(\psi^{-1}(c)\big) = 1$. De fato, suponhamos que não exista $c \in \mathbb{R}$

tal que $\mu\big(\psi^{-1}(c)\big) = 1$. Então $\mu\big(\psi^{-1}(c)\big) = 0$ para todo $c \in \mathbb{R}$, logo

$$0 \leq \mu(M) = \mu\big(\psi^{-1}(\mathbb{R})\big) = \mu\left(\psi^{-1}\left(\bigcup_{c\in\mathbb{R}}\{c\}\right)\right) \leq \sum_{c\in\mathbb{R}} \mu\big(\psi^{-1}(c)\big) = 0,$$

o que é um absurdo pois $\mu(M) = 1$. Portanto existe uma constante $c \in \mathbb{R}$ tal que $\psi(x) = c$ para $\mu-$quase todo ponto $x \in M$.

iii) $\Rightarrow$ *i)* Seja $\varphi : M \to \mathbb{R}$ uma função integrável. Pela Proposição 1.27, $\tilde{\varphi}$ é uma função $f-$invariante. Então, por hipótese, $\tilde{\varphi}$ é constante para $\mu-$quase todo ponto $x \in M$ e, pelo Teorema Ergódico de Birkhoff 1.26, temos

$$\overline{\varphi} = \int_M \varphi \, d\mu = \int_M \tilde{\varphi} \, d\mu = \tilde{\varphi} \int_M 1 \, d\mu = \tilde{\varphi}\mu(M) = \tilde{\varphi}.$$

Portanto (f, μ) é ergódico. □

Dizemos que uma medida ν é **absolutamente contínua** em relação a outra medida μ, e denotamos por $\nu \ll \mu$, se para todo conjunto mensurável $E \subseteq M$ tal que $\mu(E) = 0$, então $\nu(E) = 0$. O próximo resultado mostra que duas medidas absolutamente contínua e $f-$invariante, se forem ergódicas então são iguais.

Proposição 1.30. *Sejam $f : M \to M$ uma aplicação mensurável e, μ e ν probabilidades invariantes. Se μ é ergódica pra f e ν é absolutamente contínua em relação a μ, então $\mu = \nu$.*

Demonstração. Sejam $E \subseteq M$ um conjunto mensurável qualquer e $\mathcal{X}_E : M \to \mathbb{R}$ a sua função característica. Como μ é ergódica pra f, então $\tilde{\mathcal{X}_E}(x) = \overline{\mathcal{X}_E} = \int_M \mathcal{X}_E d\mu = \mu(E)$ para $\mu-$quase todo ponto $x \in M$, e como $\nu \ll \mu$, então $\tilde{\mathcal{X}_E}(x) = \mu(E)$ para $\nu-$quase todo ponto $x \in M$. Logo,

$$\int_M \tilde{\mathcal{X}_E}(x) \, d\nu = \int_M \mu(E) \, d\nu = \mu(E) \int_M 1 \, d\nu = \mu(E).$$

E pelo Teorema Ergódico de Birkhoff 1.26, temos

$$\mu(E) = \int_M \tilde{\mathcal{X}_E}(x) \, d\nu = \int_M \mathcal{X}_E \, d\nu = \nu(E).$$

Como $E \subseteq M$ é um conjunto mensurável qualquer, então $\mu = \nu$. □