Capítulo 1

Teorema Ergódico de Birkhoff

Seja $x \in M$ e um conjunto mensurável $E \subset M$, vamos tomar os n primeiros iterados da orbita de x, e vamos considerar a fração desses iterados que estão em E:

$$\begin{aligned} \tau_n(E,x) &= \frac{1}{n}\#\{f^j(x) \in E : 0 \leq j \leq n-1\} \\ &= \frac{1}{n}\sum_{j=0}^{n-1} \mathcal{X}_E(f^j(x)) \end{aligned} \tag{1.1}$$

Onde $\mathcal{X}_E$ é a função característica 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 o *tempo médio de permanência* da orbita de x em E, sendo o limite dessas frações fazendo n tender ao infinito:

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

Em geral, esse limite pode não existir.

Lema 1.1. *Se o tempo médio de permanência existe para um ponto $x \in M$, então*

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

Demonstração. De fato, por definição temos:

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

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

Teorema 1.2. *Seja $f : M \to M$ uma transformação mensurável e μ uma probabilidade invariante por f. Dado qualquer conjunto mensurável $E \subset M$, o tempo médio de permanência $\tau(E,x)$ existe em μ-quase todo ponto $x \in M$. Além disso,*

$$\int \tau(E,x)d\mu(x) = \mu(E).$$

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

$$\begin{aligned}
\overline{\tau}(E,x) &= \limsup_{n\to+\infty}\frac{1}{n}\#\big\{f^j(x)\in E : 0\le j\le n-1\big\} \\
\underline{\tau}(E,x) &= \liminf_{n\to+\infty}\frac{1}{n}\#\big\{f^j(x)\in E : 0\le j\le n-1\big\}
\end{aligned}$$

Para todo $x \in M$ temos que

$$\overline{\tau}(E,f(x)) = \overline{\tau}(E,x) \quad \text{e} \quad \underline{\tau}(E,f(x)) = \underline{\tau}(E,x) \tag{1.2}$$

De fato, a demonstração é análoga a do lema 1.1.

Para demonstrar a existência do tempo médio $\tau(E,x)$, basta mostrar que

$$\overline{\tau}(E,x) = \underline{\tau}(E,x) \tag{1.3}$$

para μ-quase todo ponto $x \in M$.

Como $\overline{\tau}(E,x) \geq \underline{\tau}(E,x)$ para todo $x \in M$, então pra mostrar a igualdade, basta mostrar a seguinte desigualdade

$$\int \overline{\tau}(E,x)d\mu(x) \leq \mu(E) \leq \int \underline{\tau}(E,x)d\mu(x) \tag{1.4}$$

Vamos provar a primeira desigualdade em (1.4), e a segunda segue de argumento inteiramente análogo.

Dado $\varepsilon > 0$ arbitrário. Por definição de $\limsup$, para cada $x \in M$ existem inteiros $t \geq 1$, tais que

$$\tau_t(E,x) \geq \overline{\tau}(E,x) - \varepsilon \tag{1.5}$$

Definamos $t : M \to \mathbb{N}$ uma função que leva o ponto $x \in M$ ao primeiro t que satisfaça (1.5). 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 $x \in M$. Então fixemos um certo $n \in \mathbb{N}$.

Dado $x \in M$, definamos uma sequencia $x_0, x_1, \cdots, x_s$ de pontos de M e uma sequencia $t_0, t_1, \cdots, t_s$ 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 $t_0 + t_1 + \cdots + t_s \geq n$.

Note que aplicando $x_i = f^{t_0+t_1+\cdots+t_{i-1}}(x)$ em (1.2), temos $\overline{\tau}(E,x_i) = \overline{\tau}(E,x)$ para todo i. Pela definição de $\tau_t(E,x)$ em (1.1), temos

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

Como essa equação vale para todo $x \in M$ em particular vale para todo x_i, e aplicando em (1.5), temos

$$\begin{aligned}\sum_{j=0}^{t_i-1} \mathcal{X}_E(f^j(x_i)) &= t_i \tau_{t_i}(E, x_i) \\ &\geq t_i(\overline{\tau}(E, x) - \varepsilon) \qquad (1.6)\end{aligned}$$

Pela definição da sequencia x_i temos que

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

onde

$$t_0 + t_1 + \cdots + t_{s-1} < n - 1$$

E como pelo item 3 da definição da sequencia temos $t_0 + t_1 + \cdots + t_{s-1} \geq n - t_s \geq n - K$, podemos reescrever (1.6), colocando x_i em função de x para todo i, e somar todos os eles

$$\begin{aligned}\sum_{j=0}^{n-1} \mathcal{X}_E(f^j(x)) &\geq (t_0 + t_1 + \cdots + t_{s-1})\,(\overline{\tau}(E, x) - \varepsilon) \\ &\geq (n - t_s)(\overline{\tau}(E, x) - \varepsilon) \\ &\geq (n - K)(\overline{\tau}(E, x) - \varepsilon) \qquad (1.7)\end{aligned}$$

Essa desigualdade vale para todo $x \in M$, e como as funções características são integráveis, então

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

Esse resultado vale para todo $n \in \mathbb{N}$, então passamos ao limite quando $n \to +\infty$

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

Como $\varepsilon > 0$ é arbitrário, podemos tendê-lo a 0, e temos

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

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

Caso Geral: Dado $\varepsilon > 0$ fixemos $K >> 1$ suficientemente grande, de modo que

$$\mu\left(\left\{y \in M; t(y) > K\right\}\right) < \varepsilon$$

Vamos mostrar que esse K de fato existe. Definamos $A_n \subset M$ da seguinte maneira

$$A_j = \left\{y \in M; t(y) \leq K\right\}$$

É claro que $A_j \subset A_{j+1}$ e $\cup_{j\in\mathbb{N}} A_j = M$, então $\lim \mu(A_j) = \mu\left(\cup_{j\in\mathbb{N}} A_j\right) = \mu(M) = 1$. Ou seja, para todo $\varepsilon > 0$ dado, existe $j_0 \in \mathbb{N}$, tal que $j > j_0$ implica $\mu(A_j) > 1 - \varepsilon$. Tomemos $K > j_0$, então $\mu(A_k) > 1 - \varepsilon$, isso implica $\mu(M - A_k) < \varepsilon$. Agora observe que $(M - A_k) = \big\{y \in M; t(y) \leq K\big\}$. Vamos chamar esse conjunto de B.

De maneira similar ao cado particular, fixemos um certo $n \in \mathbb{N}$. Dado $x \in M$, definamos uma sequencia $x_0, x_1, \cdots, x_s$ de pontos de M e uma sequencia $t_0, t_1, \cdots, t_s$ 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, tal que $t(x_i) \leq K$, a desigualdade (1.6) continua valendo

$$\sum_{j=0}^{t_i-1} \mathcal{X}_E(f^j(x_i)) \geq t_i(\overline{\tau}(E,x) - \varepsilon) \tag{1.8}$$

A desigualdade acima, implica na seguinte

$$\sum_{j=0}^{t_i-1} \mathcal{X}_E(f^j(x_i)) \geq t_i(\overline{\tau}(E,x) - \varepsilon) - \sum_{j=0}^{t_i-1} \mathcal{X}_B(f^j(x_i)) \tag{1.9}$$

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 (1.8), e quando $t(x_i) > K$, temos que $t_i = 1$ e esses somatórios terão apenas um elemento, ficando

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

Como $0 \leq \mathcal{X}_E(f(x_i)) \leq 1$ por definição de função característica, $(\overline{\tau}(E,x) - \varepsilon) < 1$ pois $\overline{\tau}(E,x) \leq 1$ e $\varepsilon > 0$, e por fim $\mathcal{X}_B(f(x_i)) = 1$ 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, e o da direita é menor que zero.

Agora usando o mesmo método que usamos pra concluir (1.7), fazendo novamente $x_i = f^{t_0+t_1+\cdots+t_{i-1}}(x)$ para todo i, $x_s = f^{t_0+t_1+\cdots+t_{s-1}}(x)$ onde $t_0+t_1+\cdots+t_{s-1} < n-1$, e tendo em mente que pelo item 4 da definição da sequencia temos $t_0+t_1+\cdots+t_{s-1} \geq n-t_s \geq n-K$, e essa ultima desigualdade é verdade pois $t_i \leq K$ para todo i, então podemos generalizar

$$\begin{aligned}
\sum_{j=0}^{n-1} \mathcal{X}_E(f^j(x)) &\geq \sum_{j=0}^{t_0+\cdots+t_{s-1}-1} \mathcal{X}_E(f^j(x)) \\
&= \sum_{j=0}^{t_0-1} \mathcal{X}_E(f^j(x_i)) + \sum_{j=t_0}^{t_0+t_1-1} \mathcal{X}_E(f^j(x_i)) + \sum_{j=t_0+t_1}^{t_0+t_1+t_2-1} \mathcal{X}_E(f^j(x_i)) + \cdots \\
&\quad \cdots + \sum_{t_0+\cdots+t_{s-2}}^{t_0+\cdots+t_{s-1}-1} \mathcal{X}_E(f^j(x_i)) \\
&\geq \big[t_0(\overline{\tau}(E,x) - \varepsilon) - R(t_0)\big] + \big[t_1(\overline{\tau}(E,x) - \varepsilon) - R(t_2)\big] + \cdots \\
&\quad \cdots + \big[t_{s-1}(\overline{\tau}(E,x) - \varepsilon) - R(t_{s-1})\big] \qquad (1.10)
\end{aligned}$$

onde $R(t_0) = \sum_{j=0}^{t_1-1} \mathcal{X}_B(f^j(x_i))$ e $R(t_i) = \sum_{j=t_0+\cdots+t_{i-1}}^{t_0+\cdots+t_i-1} \mathcal{X}_B(f^j(x_i))$, logo somando todos os $R(t_i)$ temos

$$\begin{aligned}
\sum_{j=0}^{t_0+\cdots+t_{s-1}-1} R(t_j) &= \sum_{j=0}^{t_0+\cdots+t_i-1} \mathcal{X}_B(f^j(x_i)) \\
&\leq \sum_{j=0}^{n-1} \mathcal{X}_B(f^j(x))
\end{aligned}$$

Podemos voltar para (1.10), e concluir

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

Essa desigualdade vale para todo $x \in M$, e como as funções características são integráveis, então

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

Esse resultado vale para todo $n \in \mathbb{N}$, então passamos ao limite quando $n \to +\infty$, e como $\mu(B) < 3$, temos

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

Como $\varepsilon > 0$ é arbitrário, podemos tendê-lo a 0, e temos

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

Isso completa a demonstração do caso geral do teorema.

$\square$

Teorema 1.3. (Teorema Ergódico de Birkhoff) *Seja $f : M \to M$ uma transformação mensurável e μ uma probabilidade invariante por f. Dada qualquer função integrável $\varphi : M \to \mathbb{R}$ o limite*

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

existe em μ-quase todo ponto $x \in M$. Além disso,

$$\int \tilde{\varphi}(x) d\mu(x) = \int \varphi(x) d\mu(x).$$

Demonstração. Este enunciado mais geral pode ser provado usando uma versão um pouco mais elaborada do argumento usado pra provar 1.2, que não apresentaremos aqui. $\square$

Observação. *O Teorema 1.2 é o caso particular do Teorema Ergódico de Birkhoff* (1.3) *quando $\varphi = \mathcal{X}_E$, a função característica do conjunto E.*