Análise espectral no estudo de sistemas dinâmicos

Raquel Brás Sá Couto
Mestrado em Matemática
Departamento de Matemática
2019

Orientador
Jorge Miguel Milhazes de Freitas, Professor Associado
Faculdade de Ciências da Universidade do Porto

U. PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ______/______/________

Agradecimentos

Ao meu orientador, Professor Jorge Freitas, todas as palavras são poucas. Agradeço pela dedicação, disponibilidade e paciência, mas, mais do que tudo, pelo incentivo e confiança transmitidos sempre nos momentos mais importantes.

À Lininha e à Mafalda, principais responsáveis pela minha decisão de arriscar na Licenciatura em Matemática.

À Inês, uma amiga para a vida, agradeço por muitas ideias, matemáticas e não matemáticas.

Aos meus amigos que foram chegando e ficando desde a Carlinha até à Joana, os visados identificar-se-ão.

À Joana por ter feito o meu 6.º ano na FCUP.

A todos os professores que me marcaram academicamente e, acima de tudo, pessoalmente.

Por fim, aos meus pais, que com mais ou menos crença me deram a liberdade de seguir este percurso, à minha irmã, que me permitiu testar os limites da minha paciência, e aos meus avós, que estiveram e estarão sempre presentes.

Resumo

A presente dissertação é o reflexo de um primeiro contacto com algumas ferramentas fundamentais para a investigação nas áreas de Sistemas Dinâmicos e Teoria Ergódica. Num primeiro momento, a passagem por uma classe particular de sistemas dinâmicos simbólicos, que estão relacionados com as cadeias de Markov, serviu como motivação. Seguidamente, o foco no operador de Perron-Frobenius e sua caracterização espectral permitiu um entendimento aprofundado das propriedades estatísticas de aplicações expansoras por pedaços definidas em intervalos fechados. Entre as mesmas propriedades, destacamos a existência de uma medida invariante absolutamente contínua com respeito à medida de *Lebesgue*, decaimento exponencial de correlações e validade do Teorema do Limite Central.

Palavras-chave: *shift*, cadeia de Markov, Perron-Frobenius, expansora por pedaços, variação limitada, Lasota-Yorke, medida invariante, decaimento de correlações, Teorema do Limite Central

Abstract

This dissertation consists on a study of spectral analytical tools, which reveal to be extremely useful for a better understanding of Dynamical Systems and Ergodic Theory. Motivated by symbolic dynamics and Markov chains we looked at the Perron Theorem and its connection with the Limit Theorem for Markov chains. Stimulated by these results we studied the Perron-Frobenius operator, its spectral characterisation and the impact on the understanding of the statistical behaviour of piecewise-expanding maps of a closed interval. Among the statistical properties of such systems, we highlight the existence of an absolutely continuous invariant measure, exponential decay of correlations and a Central Limit Theorem.

Keywords: shift, Markov chain, Perron-Frobenius, piecewise-expanding, bounded variation, Lasota-Yorke, invariant measure, decay of correlations, Central Limit Theorem

Conteúdo

Introdução

A Teoria Ergódica é a área que se dedica ao estudo dos Sistemas Dinâmicos usando uma panóplia de ferramentas probabilísticas para melhor compreender o comportamento estatístico dos sistemas.

A ferramenta principal nesta dissertação é o operador de Perron-Frobenius, que será introduzido no Capítulo 2 e que tem como pontos fixos as densidades invariantes de um determinado sistema. As suas propriedades espectrais revelam-se muito importantes para melhor compreender o comportamento estatístico dos sistemas e, em particular, a rapidez com que estes perdem memória. A motivação para esta abordagem será feita no Capítulo 1, onde, para dinâmica simbólica, começaremos por estudar as potencialidades da análise espectral em dimensão finita.

A introdução do operador de Perron-Frobenius, bem como de algumas das suas propriedades, constitui o foco do Capítulo 2, que é inspirado na exposição do Capítulo 4 de [BG97].

Com a dualidade medida invariante−densidade fixa em mente, no Capítulo 3, por sua vez baseado no Capítulo 3 de [Via97], é apresentada a prova da existência de tais medidas no caso particular em que a dinâmica é determinada por uma aplicação expansora por pedaços num intervalo fechado; para a unicidade, basta que a aplicação seja, adicionalmente, topologicamente misturadora, caso em que se deduz um decaimento exponencial de correlações e um Teorema do Limite Central para densidades com variação limitada. A conclusão do capítulo com a análise do espectro do operador de Perron-Frobenius remete-nos para o Capítulo 1 e serve de ponte para a generalização feita no Capítulo 4.

No Capítulo 4, o pressuposto é a generalização da própria definição de aplicação expansora por pedaços (permitindo-se uma quantidade de numerável de descontinuidades ao invés de uma quantidade finita) e com base nele vem uma descrição mais geral do espectro do operador de Perron-Frobenius, segundo [Ryc83].

Mas o ponto de partida, no Capítulo 1, tinha sido o estudo de uma classe revelante de sistemas dinâmicos simbólicos: os *shift*'s. Depois de uma passagem por algumas noções fundamentais e gerais na temática dos sistemas dinâmicos, os *shift*'s conduziram-nos às Cadeias de Markov e respetivo Teorema Limite que, como corolário do Teorema de Perron, acaba por ser a formulação do resultado sobre o espectro do operador de Perron-Frobenius num contexto bastante restrito ($\mathbb{R}^n$). Para o desenvolvimento do Capítulo 1, foi seguido o Capítulo 3 de [BS02].

Em cada capítulo as principais referências são as que aqui foram mencionadas. Na maior parte das demonstrações foram incluídos passos intermédios que não estão presentes nas mesmas referências.

Capítulo 1

Dinâmica Simbólica e Cadeias de Markov

Iniciamos este capítulo com o estudo da dinâmica das aplicações $g_a = ax \mod 1$ ($[0,1] \to [0,1]$) às quais, como veremos, o *shift* num alfabeto de a símbolos é semiconjugado.

As Cadeias de Markov (em tempo discreto) são um tipo particular de tais sistemas dinâmicos simbólicos. Ora, uma vez assim interpretadas, faz sentido que uma abordagem clássica de Sistemas Dinâmicos possa ser usada para o seu estudo.

No final do capítulo são preenchidas algumas entrelinhas na prova do Teorema de Perron apresentada em [BS02], da qual resulta o Teorema Limite para Cadeias de Markov.

1.1 Sistemas Dinâmicos: algum vocabulário

Começamos pela noção de sistema dinâmico e por algumas definições fundamentais.

Definição 1.1. Um *sistema dinâmico* (discreto) é descrito por um conjunto não vazio X e uma aplicação $f : X \to X$; estabelecendo que $f^0 = id$ e que $f^n = f^{n-1} \circ f$ definimos f^n para todo o $n \in \mathbb{N}$.

Definição 1.2. A *órbita* de $x \in X$, que denotamos por $\mathcal{O}_x(f)$, é o conjunto dos sucessivos iterados de x pela aplicação f i.e. $\mathcal{O}_x(f) = \{x, f(x), f^2(x), ..., f^n(x), ...\}$.

Definição 1.3. Um elemento $x \in X$ diz-se um *ponto periódico de período* n se n é o menor número natural tal que $f^n(x) = x$; se $x \in X$ é periódico de período 1, designa-se *ponto fixo*.

Definição 1.4. Um elemento $x \in X$ diz-se um *ponto pré periódico (de período* n*)* se existe $m \in \mathbb{N}$ tal que $f^m(x)$ é um ponto periódico (de período n).

Observação 1.5.

- pontos periódicos distintos que não têm a mesma órbita têm órbitas disjuntas: supomos que $z \in \mathcal{O}_x(f) \cap \mathcal{O}_y(f)$, pelo que $z = f^i(x)$ e $z = f^j(y)$ para alguns i, j; mas então, por um lado, $\mathcal{O}_z(f) = \{f^i(x), f^{i+1}(x), ..., x, f(x), ...\} = \mathcal{O}_x(f)$ e, por outro lado, $\mathcal{O}_z(f) = \{f^j(y), f^{j+1}(y), ..., y, f(y), ...\} = \mathcal{O}_y(f)$ – contradição com o pressuposto de que x e y têm órbitas distintas.

- é claro que para $x \in X$ periódico de período n se verifica $f^m(x) = x \;\; \forall m \in \mathbb{N} : n|m$.

Vejamos um exemplo, o qual será relevante mais à frente (secção 1.2).

Exemplo 1.6. Para $a \geq 2$ (e $a \in \mathbb{N}$), consideremos

$$\begin{aligned} g_a \colon \; [0,1] &\to [0,1] \\ x &\mapsto \begin{cases} ax, & 0 \leq x \leq \dfrac{1}{a} \\ ax - j, & \dfrac{j}{a} < x \leq \dfrac{j+1}{a} \quad j \in \{1, ..., a-1\} \end{cases} \end{aligned}$$

Daqui em diante referir-nos-emos a este sistema dinâmico simplesmente como g_a ou $ax \mod 1$.

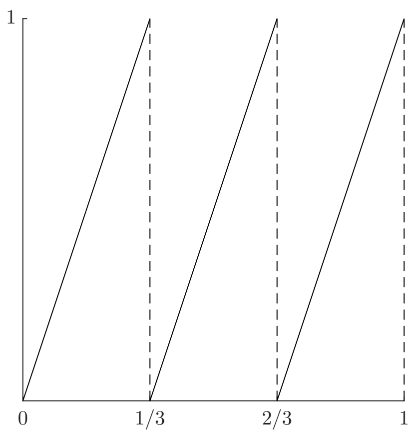

Figura 1.1: Gráfico de g_3 ($3x \mod 1$)

Propriedades:

1. g_a tem a^n pontos cujo período divide n.

2. $a^n = \sum_{m|n} \mathcal{N}_m(g_a)$, onde $\mathcal{N}_m(g_a)$ denota o número de pontos periódicos de período m de g_a.

3. os pontos em 1. são precisamente os $k/(a^n - 1)$ para $k \in \{0, ..., a^n - 1\}$.

4. Se $gcd(k, a^n - 1) = 1$, então $k/(a^n - 1)$ é um ponto periódico de período n de g_a.

5. Se $k_1/(a^n - 1)$ e $k_2/(a^n - 1)$ pertencem à mesma órbita, então $gcd(k_1, a^n - 1) = gcd(k_2, a^n - 1)$.

Para justificar 1. consideramos a aplicação g_a^n: pontos periódicos de g_a cujo período divide n são os pontos fixos de g_a^n; observando que o gráfico de g_a^n consiste em a^n retas de declive a^n (> 1), então há a^n interseções com a diagonal, ou seja, a^n pontos fixos.

Os pontos cujo período divide n são precisamente os pontos de período m tal que $m|n$, logo 2. segue.

3. é consequência de os a^n pontos de 1. estarem igualmente espaçados (o que pode ser verificado muito facilmente com argumentos de semelhanças de triângulos sobre o gráfico de g_a^n) e 0 e 1 pertencerem a tal conjunto de pontos.

Para 4., suponhamos que $gcd(k, a^n - 1) = 1$ e $k/(a^n - 1)$ é um ponto periódico de período m de g_a, com $m \neq n$ tal que $m|n$. Então, tal ponto é da forma $j/(a^m - 1)$ para $j \in \{0, ..., a^m - 1\}$ e, como tal, $j/(a^m - 1) = k/(a^n - 1)$ $-$ contradição, pois $gcd(k, a^n - 1) = 1$ implica que $k/(a^n - 1)$ não pode ser representado por uma fração com um denominador menor.

Temos 5. se notarmos que dado $x = \dfrac{k}{a^n - 1}$ então

$g_a(x) = a\dfrac{k}{a^n - 1} - l\dfrac{a^n - 1}{a^n - 1} = \dfrac{ak - l(a^n - 1)}{a^n - 1}$, $l \in \{0, ..., a - 1\}$ e

$gcd(ak - l(a^n - 1), a^n - 1) = gcd(ak, a^n - 1) = gcd(k, a^n - 1)$.

Definição 1.7. Uma *semiconjugação* de (Y, g) a (X, f) é uma aplicação sobrejetiva $h : Y \to X$ tal que $h \circ g = f \circ h$; no caso de h ser bijetiva então tem-se uma *conjugação*, e (Y, g) e (X, f) dizem-se conjugados.

Notemos que se $x = h(y)$ $(x \in X, y \in Y)$ então $f^n(x) = f^n(h(y)) = f^{n-1}(f(h(y))) = f^{n-1}(h(g(y))) = ... = h(g^n(y))$. Assim, se h é sobrejetiva então $h \circ g^n = f^n \circ h$, o que significa que a semiconjugação preserva as órbitas de (Y, g). Ora, tratando-se de uma conjugação então há uma identificação bijetiva entre as órbitas de (Y, g) e as órbitas de (X, f).

É imediato verificar que a relação $\sim$ dada por $(Y, g) \sim (X, f) \iff (Y, g)$ e (X, f) são conjugados, é uma relação de equivalência. Juntando este facto com o que acaba de ser observado acima, então vem que o estudo de um determinado sistema dinâmico pode ser reduzido ao estudo de qualquer outro na sua classe de equivalência; mais ainda, se apenas pretendermos deduzir propriedades do sistema, não temos sequer necessidade de conhecer explicitamente h.

1.2 Dinâmica Simbólica

Dado um sistema dinâmico (X, f) podemos tomar uma partição $X_0, ..., X_{n-1}$ de X e codificar as trajetórias dos seus pontos de acordo com a seguinte regra: $a_i = k \iff f^i(x) \in X_k$, de tal modo que x fica identificado com a sequência $a_1a_2a_3...$. Numa perspetiva de análise da evolução do sistema a longo prazo, então o que gostaríamos de observar seriam as caudas de todas as sequências acima mencionadas. Fica pois motivado o estudo de um caso particular de sistema dinâmico simbólico.

Definição 1.8. O *full one-sided shift* é o sistema dinâmico (Σ_n^+, σ):

- Σ_n^+ é o conjunto das sequências, indexadas por $\mathbb{N}$, formadas a partir do alfabeto de n símbolos $\{0, ..., n-1\}$
- $\begin{array}{rccc} \sigma\colon & \Sigma_n^+ & \to & \Sigma_n^+ \\ & (a_i)_i & \mapsto & (a_{i+1})_i \end{array}$

De facto, podemos introduzir em Σ_n^+ uma estrutura de espaço métrico.

Usamos em $\Sigma_n^+ = \{0, ..., n-1\}^{\mathbb{N}}$ a topologia produto, para a qual o conjunto de todos os cilindros

$$C_{j_1,...,j_k}^{m_1,...,m_k} = \{(a_i)_i : \forall i = 1, ..., k \;\; a_{m_i} = j_i, \; m_i < m_{i+1} \in \mathbb{N}, \; j_i \in \{0, ..., n-1\}\}$$

constitui uma base; tem-se $\sigma^{-1}(C^{n_1,\ldots,n_k}_{j_1,\ldots,j_k}) = C^{n_1+1,\ldots,n_k+1}_{j_1,\ldots,j_k}$ portanto, com a topologia produto, σ é contínua.

Verifica-se que a métrica $d((a_i)_i,(b_i)_i) = 2^{-l}$ para $l = \min\{i : a_i \neq b_i\}$ gera a topologia produto.

Nestas condições, (Σ_n^+, σ) adquire as seguintes propriedades:

1. Há pontos periódicos de todos os períodos.

2. O conjunto dos pontos periódicos é denso.

3. O conjunto dos pontos pré-periódicos é denso.

4. O conjunto dos pontos que são nem periódicos nem pré-periódicos é denso.

5. $\exists \varepsilon > 0 : \forall a, b \in \Sigma_n^+, a \neq b, \exists n \in \mathbb{N} : d(\sigma^n(a), \sigma^n(b)) = \varepsilon$

6. Há pelo menos um ponto com órbita densa.

1. é imediata e para 2. a 4. basta observar que qualquer sequência em Σ_n^+ pode ser arbitrariamente aproximada por uma sequência periódica, por uma sequência pré-periódica, ou por uma sequência que seja nem periódica nem pré-periódica, respetivamente. Em 5. $\epsilon = 1/2$. $a = 01...(n-1)0001...0(n-1)1011...1(n-1)...(n-1)0(n-1)1...(n-1)(n-1)...$ tem órbita densa, o que valida 6..

De facto, (Σ_n^+, σ) é semiconjugado a $([0,1], g_n)$ do **Exemplo 1.5.**:

Dado $x \in [0,1]$, temos uma sequência $(x_i)_i$ naturalmente associada à sua representação na base n:

$$x = \sum_{i=1}^{+\infty} \frac{x_i}{n^i}$$

e como procuramos $h : \Sigma_n^+ \to [0,1]$ sobrejetiva então tomamos

$$\begin{array}{rccl} h\colon & \Sigma_n^+ & \to & [0,1] \\ & (x_i)_i & \mapsto & \displaystyle\sum_{i=1}^{+\infty} \frac{x_i}{m^i} \end{array}$$

(sabemos que a representação na base n é única excetuando nos casos em que a expansão de x tem uma cauda da forma $0(0)...$ ou $n-1(n-1)...$, portanto, a menos desses mesmos casos, h é até bijetiva).

Além disso, para ser uma semiconjugação, h deve satisfazer $h \circ \sigma = g_n \circ h$; tal decorre imediatamente tendo em conta que se $x \in [0,1]$ então

$$g_n(x) = nx \bmod 1 = \sum_{i=1}^{+\infty} \frac{x_{i+1}}{n^i}$$

(com 1 admitindo apenas a representação 0.9(9)...).

Consequentemente, a propriedade 1. acima apontada para (Σ_n^+, σ) é também uma propriedade de $([0,1], g_n)$. Porque h é contínua, o mesmo se verifica para as propriedades 2. a 6..

Focamo-nos agora num certo tipo de subespaços dos espaços Σ_n^+, que estão relacionados com as Cadeias de Markov.

Definição 1.9. Um *subshift* é qualquer subespaço de um Σ_n^+ que seja fechado e $\sigma-$invariante.

A condição de $\sigma-$invariância legitima que um *subshift*, com a aplicação σ, seja ainda um sistema dinâmico.

Por ser fechado, tal espaço fica completamente determinado por um conjunto de sequências proibidas, no sentido de que nenhuma sequência que lhe pertença pode conter uma sequência proibida.

Exemplo 1.10. Em Σ_2^+, o conjunto de sequências que entre dois 1's contêm um número par de 0's é um *subshift*. As sequências proibidas são as do conjunto $\mathcal{F} = \{1\underbrace{0...0}_{2n+1}1,\ n \in \mathbb{N}_0\}$.

Definição 1.11. Um *subshift de tipo finito* é o caso particular de *subshift* onde as sequências proibidas têm comprimento finito e constituem um conjunto finito.

Exemplo 1.12. O **Exemplo 1.9.** não é de um *subshift* de tipo finito. No entanto, em Σ_2^+, o conjunto de sequências que não contêm 1's consecutivos é um *subshift* de tipo finito, com o conjunto de sequências proibidas $\mathcal{F} = \{11\}$.

Sempre que, tal como no **Exemplo 1.11.**, as sequências proibidas têm comprimento 2, podemos representar o *subshift* de tipo finito através de um grafo dirigido (finito) cujos vértices são os símbolos do respetivo alfabeto e as arestas correspondem a transições permitidas entre eles.

Definição 1.13. Um *subshift* de tipo finito em que as sequências proibidas têm comprimento 2 designa-se uma *Cadeia de Markov topológica.*

1.3 Cadeias de Markov

Notação: $\mathbb{P}(A|B)$ representa a probabilidade condicional de A dado B, i.e.

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \text{ sempre que } P(B) \neq 0.$$

A definição que apresentamos é a definição clássica de Cadeia de Markov enquanto processo estocástico:

Definição 1.14. Num espaço de probabilidade $(\Omega, \mathcal{B}, \mathbb{P})$, consideramos uma sequência de variáveis aleatórias (v.a.) $X_0, X_1, ...$ para as quais assumimos:

(i) o conjunto de valores admissíveis é finito, digamos $\{1, ..., m\}$

(ii) X_n representa o estado do sistema no instante n

(iii) $\mathbb{P}(X_n = i_n | X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}) = \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1})$

$X_0, X_1, ...$ diz-se uma *Cadeia de Markov* (discreta).

Restringir-nos-emos a processos estacionários, i.e. processos tais que as probabilidades de transição entre estados são independentes do valor de n.

Definição 1.15. $p_{ij} = \mathbb{P}(X_n = j | X_{n-1} = i)$ designa-se *probabilidade de transição do estado i para o estado j.*

Definição 1.16. A matriz $P = (p_{ij})_{ij}$, $i, j \in \{1, ..., m\}$, é chamada *matriz de transição.*

Definição 1.17. Uma matriz A $(k \times k)$ diz-se:

- *positiva (não negativa)* se todas as suas entradas são positivas (não negativas).
- *irredutível* se for não negativa, quadrada, e tal que $\forall i, j \; \exists n \in \mathbb{N} : (A^n)_{ij} > 0$.
- *primitiva* se for não negativa, quadrada, e tal que $\exists n \in \mathbb{N} \; \forall i, j : (A^n)_{ij} > 0$.

- *estocástica* se for real não negativa, quadrada, e tal que $\sum_{j=1}^{k} p_{ij} = 1 \ \forall i \in \{1, ..., k\}$ (equivalente a afirmar que o vetor coluna com todas as entradas iguais a 1 é um vetor próprio associado ao valor próprio 1).

Observação 1.18.

- P é não negativa e estocástica (decorre imediatamente da axiomática de probabilidade)

- $\mathbb{P}(X_0 = i_0, X_1 = i_1, ..., X_n = i_n) =$
  $\mathbb{P}(X_n = i_n | X_0 = i_0, ..., X_{n-1} = i_{n-1})\mathbb{P}(X_0 = i_0, ..., X_{n-1} = i_{n-1}) =$
  $p_{i_{n-1}i_n}\mathbb{P}(X_{n-1} = i_{n-1} | X_0 = i_0, ..., X_{n-2} = i_{n-2})\mathbb{P}(X_0 = i_0, ..., X_{n-2} = i_{n-2}) = ... =$
  $p_{i_{n-1}i_n} p_{i_{n-2}i_{n-1}} ... p_{i_0 i_1}\mathbb{P}(X_0 = i_0)$

Definição 1.19. $p_{ij}^{(n)} = \mathbb{P}(X_n = j | X_0 = i)$ designa-se *probabilidade de transição a n passos do estado i para o estado j.*

É consequência do segundo item da *Observação* 1.17. que a entrada correspondente à linha i e coluna j da matriz P^n é precisamente $p_{ij}^{(n)}$.

<u>Notação:</u> $x^{(n)} = \begin{pmatrix} \mathbb{P}(X_n = 1) & \mathbb{P}(X_n = 2) & \dots & \mathbb{P}(X_n = m) \end{pmatrix}$ (matriz linha)

Decorre que $x^{(n)} = x^{(0)}P^n$.

Teorema 1.20 (Teorema Limite para Cadeias de Markov)**.** *Se P é primitiva então*

$$\lim_{n \to +\infty} P^n = \begin{pmatrix} q_1 & q_2 & \cdots & q_m \\ q_1 & q_2 & \cdots & q_m \\ \vdots & \vdots & \ddots & \vdots \\ q_1 & q_2 & \cdots & q_m \end{pmatrix}$$

onde para $i = 1, ..., m$ $q_i > 0$ e $q_1 + q_2 + \cdots + q_m = 1$.

Uma demonstração é apresentada na secção seguinte.

Informalmente, o que o Teorema afirma é que as probabilidades de transição a n passos, à medida que n aumenta, convergem para valores que não são condicionados pelo ponto de partida.

Adicionalmente, também o vetor estocástico $x^{(n)}$ tem limite:

Corolário 1.21. $\lim_{n\to+\infty} x^{(n)} = q = \begin{pmatrix} q_1 & q_2 & \dots & q_m \end{pmatrix}$.

Demonstração. $\lim_{n\to+\infty} x^{(n)} = \lim_{n\to+\infty} x^{(0)} P^n = x^{(0)} \begin{pmatrix} q_1 & q_2 & \dots & q_m \\ q_1 & q_2 & \dots & q_m \\ \vdots & \vdots & \ddots & \vdots \\ q_1 & q_2 & \dots & q_m \end{pmatrix} =$

$$(\mathbb{P}(X_0 = 1)q_1 + \cdots + \mathbb{P}(X_0 = m)q_1 \quad \mathbb{P}(X_0 = 1)q_2 + \cdots + \mathbb{P}(X_0 = m)q_2$$
$$\dots \quad \mathbb{P}(X_0 = 1)q_m + \cdots + \mathbb{P}(X_0 = m)q_m) =$$

$$(\mathbb{P}(X_0 = 1) + \mathbb{P}(X_0 = 2) + ... + \mathbb{P}(X_0 = m)) \begin{pmatrix} q_1 & q_2 & \dots & q_m \end{pmatrix} = \begin{pmatrix} q_1 & q_2 & \dots & q_m \end{pmatrix} \quad \square$$

E tal limite é o único ponto fixo da aplicação que a um vetor estocástico x associa xP:

Corolário 1.22. *q é o único vetor estocástico tal que $qP = q$.*

Demonstração. Suponhamos que existe $q' \neq q$ tal que $q'P = q'$. Pelo **Corolário 1.20.**, sabemos que $\lim_{n\to+\infty} q'^{(n)} = q$. Mas $q'^{(n)} = q'$ $\forall n \in \mathbb{N}$, portanto $\lim_{n\to+\infty} q'^{(n)} = q'$ – contradição. $\square$

Relembrando agora a **Definição 1.10.** fica claro porquê que as Cadeias de Markov (tal como na **Definição 1.13.**) podem ser vistas como *subshifts* de tipo finito associados a palavras proibidas de comprimento 2: (iii) da respetiva definição justifica a representação de uma Cadeia de Markov por um grafo cujo conjunto de vértices é o conjunto dos valores admissíveis e onde as arestas têm pesos determinados pela matriz de transição.

Notemos que começamos por definir Cadeia de Markov como processo estocástico e acabámos de ver que se trata também de um caso particular de sistema dinâmico simbólico, onde, para além da possibilidade de um estudo *standard* no contexto da teoria de sistemas dinâmicos, se acrescenta a perspetiva de uma abordagem combinatória, à luz da teoria de grafos.

1.4 O Teorema de Perron e o Teorema Limite para Cadeias de Markov

Seja $L : \mathbb{R}^k \to \mathbb{R}^k$ uma aplicação linear. Por vezes indistintamente consideramos o sistema dinâmico linear associado a L.

Lema 1.23. *Se λ é um valor próprio de L ao qual está associado o vetor próprio v, então λ^i é um valor próprio de L^i ao qual está associado o vetor próprio v.*

Demonstração. Basta notar que $L^i v = L^{i-1} L v = L^{i-1} \lambda v = ... = \lambda^i v$. □

Definição 1.24. v^T diz-se um *vetor próprio esquerdo* de L associado ao valor próprio λ se $v^T L = v^T \lambda$.

Observação 1.25. Quando nada é dito em contrário, assumimos a representação de v através de uma matriz coluna, compatível com a multiplicação à direita por L, e apenas por esse facto surge na **Definição 1.24** v^T (ocorrendo a multiplicação à esquerda por L, o vetor tem que entrar sob a forma de matriz linha); como sugere a mesma definição, a designação *vetor próprio* corresponde, em rigor, a *vetor próprio direito.*

Lema 1.26. *Se v é um vetor próprio de L associado ao valor próprio λ, então v^T é um vetor próprio esquerdo de L^T associado ao mesmo valor próprio λ.*

Demonstração. Tem-se $v^T L^T = (Lv)^T = (\lambda v)^T = v^T \lambda$. □

Notação: $E_\lambda = \{v \in \mathbb{R}^k : Lv = \lambda v\}$ (subespaço próprio associado a λ)

Apenas no Lema que se segue vamos considerar $L : \mathbb{R}^2 \to \mathbb{R}^2$.

Lema 1.27. *Seja $L : \mathbb{R}^2 \to \mathbb{R}^2$. Suponhamos que existe λ valor próprio de L tal que $|\lambda| = 1$.*

(i) Se λ é uma raiz da unidade, então todos os pontos de E_λ têm órbita periódica.

(ii) Se λ não é uma raiz da unidade, então todos os pontos de E_λ têm órbita densa em $S^1 \equiv \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$.

Demonstração. O caso *(i)* é claro, pois, sendo $\lambda^i = 1$ para algum i, então dado $v \in E_\lambda$ tem-se, pelo **Lema 1.23**, $L^i v = \lambda^i v = v$.

Em *(ii)* trata-se de provar que dado $v \in E_\lambda$, o conjunto $\mathcal{O}_v(L) = \{v, Lv, L^2v, ...\}$ é denso em S^1. Ora, $\mathcal{O}_v(L)$ é uma sucessão contida num compacto (S^1), portanto admite uma subsucessão convergente que é, em particular, uma sucessão de *Cauchy*. Assim, fixando $\varepsilon > 0$, $\exists m, n \in \mathbb{N}$ $(m \neq n) : d(L^m v, L^n v) < \varepsilon$; definimos f como a rotação de arco $L^m v L^n v$, pelo que $\exists k \in \mathbb{N}$ tal que $\{v, f(v), ..., f^k(v)\}$ é $\varepsilon-$denso em S^1. Consequentemente, $\mathcal{O}_v(L)$ é $\varepsilon-$denso em S^1. Como ε é arbitrário, então $\mathcal{O}_v(L)$ é denso em S^1. $\square$

Proposição 1.28. *Se existe um conjunto compacto $C \neq \emptyset$ e tal que $0 \in int(C)$ e $L^i(C) \subset int(C)$ para algum $i > 0$, então todos os valores próprios de L têm módulo estritamente menor que 1.*

Demonstração. Vejamos que podemos, sem perda de generalidade, assumir que $L(C) \subset int(C)$. De facto, tomemos $K : \mathbb{R}^k \to \mathbb{R}^k$ tal que $K = L^i$ para i tal que $L^i(C) \subset int(C)$; ora, se o resultado é válido, então a conclusão é que todos os valores próprios de K têm módulo estritamente menor que 1, pelo que, como consequência do **Lema 1.23**, todos os valores próprios de L têm módulo estritamente menor que 1.

Notemos que o facto de se ter $0 \in int(C)$ permite-nos concluir que $E_\lambda \cap C \setminus \{0\} \neq \emptyset$, qualquer que seja λ valor próprio de L, pelo que podemos tomar $v \in E_\lambda \cap C$, ou mais particularmente, $v \in E_\lambda \cap int(C)$ ou $v \in E_\lambda \cap \partial(C)$.

Suponhamos que existe λ valor próprio de L tal que $|\lambda| > 1$. Então, $v \in E_\lambda \cap int(C)$ é tal que $\lim_{n\to+\infty} \|L^n v\| = \lim_{n\to+\infty} |\lambda^n| \|v\| = +\infty$; mas, por hipótese, $L^n v \in int(C)$ $\forall n \geq 1$ e $int(C)$ é limitado $-$ contradição.

Suponhamos agora que existe λ valor próprio de L tal que $|\lambda| = 1$. Tomemos $v \in E_\lambda \cap \partial(C)$. Se λ é uma raiz da unidade, então, pelo **Lema 1.27***(i)*, $L^j v = v$ para algum $j \in \mathbb{N}$ $-$ contradição, pois $L(C) \subset int(C)$. Se λ não é uma raiz da unidade, então, pelo **Lema 1.27***(ii)*, v tem órbita densa em alguma circunferência, D, de dimensão 1 contida em $\mathbb{R}^k$ (centrada na origem e de raio $\|v\|$); mas, assim, existe uma subsucessão $(L^{n_i} v)_i$ de pontos da órbita de v que converge para v e está contida em D $-$ contradição, pois sendo D fechado então $v \in D$, ou seja $v \in D \cap \partial(C)$, e $L^{n_i} v \in D \cap int(C)$ $\forall n_i$. $\square$

Teorema 1.29 (Perron)**.** *Seja A uma matriz $m \times m$ primitiva. Então, A tem um valor próprio $\lambda > 0$ com as seguintes propriedades:*

1. *λ é uma raiz simples do polinómio característico de A*
2. *λ tem associado um vetor próprio v positivo*
3. *qualquer outro valor próprio de A tem módulo estritamente menor que λ*
4. *qualquer vetor próprio de A não negativo é um múltiplo (positivo) de v*

Demonstração. O primeiro passo é justificar a existência de λ e *2.*

Seja $S = \{(x_1, ..., x_m) \in \mathbb{R}^m : \forall j = 1, ..., m \ x_j \geq 0, \sum_j x_j = 1\}$ (o simplexo unitário $(m-1)-$dimensional).
Consideramos $f : S \to S$ tal que $f(x) =$ projeção radial em S de Ax, que é contínua. Observemos que, porque A é não negativa, então Ax é não negativo.
Ora, f contínua e S compacto implicam, pelo Teorema do Ponto Fixo de Brouwer, que existe $v \in S$ tal que $f(v) = v$. Mas então, v é tal que Av (não negativo, pela observação anterior) está na semi-reta que parte da origem e passa por v, ou seja, Av é um múltiplo não negativo de v. De facto, porque A é primitiva, decorre que Av é um múltiplo positivo de v, e, porque adicionalmente se tem o **Lema 1.23**, então v tem todas as componentes positivas.

Prosseguimos no sentido de mostrar *3.* e *1.*

Vamos exibir uma matriz estocástica, M, cujo espectro nada mais é que o espectro de A dividido por λ; ora, o nosso problema reduzir-se-á, pois, a provar que 1 é uma raiz simples do polinómio característico de M e que os restantes valores próprios de M têm módulo estritamente menor que 1.

Seja $V = \begin{pmatrix} v_1 & & & \\ & v_2 & & \\ & & \ddots & \\ & & & v_m \end{pmatrix}$ a matriz diagonal onde $v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{pmatrix}$ é o vetor acima mencionado.

Definimos $M = \dfrac{1}{\lambda} V^{-1} A V$ (a verificação de que se trata de uma matriz estocástica é imediata pois $V(1\,1\,...\,1)^T = v$ e $Av = \lambda v$.). Notemos que M é primitiva, já que é claramente não negativa e A é primitiva.

Consideremos $g : S \to S$ tal que $g(x) = xM$. É fácil ver que g está bem definida, e g é contínua. Logo, pelo Teorema do Ponto Fixo de Brouwer, existe $w \in S$ tal que $g(w^T) = w^T$ e w tem todas as componentes positivas ($w^T = w^T M = w^T M^2 = ... = w^T M^j = ...$ e M é primitiva, logo, se j é tal que todas as entradas de M^j são positivas, então $w^T M^j$ tem todas as componentes positivas).

A conclusão sobre os valores próprios de M é agora consequência da **Proposição 1.28** e do **Lema 1.26**: $C = S - w \neq \emptyset$ é compacto (de um subespaço de $\mathbb{R}^m$ isomorfo a $\mathbb{R}^{m-1}$), contém 0 no seu interior e, se j é tal que todas as entradas de $(M^T)^j$ são positivas, então $(M^T)^j(C) = (M^T)^j(S - w) = (M^T)^j(S) - (M^T)^j(w) = (M^T)^j(S) - (w^T M^j)^T \subset int(S) - w = int(C)$, portanto, pela **Proposição 1.28**, vem que, na restrição ao subespaço $(m-1)-$dimensional gerado por C, M^T só admite valores próprios de módulo estritamente menor do que 1, de modo que M só admite valores próprios de módulo estritamente menor do que 1, pelo **Lema 1.26**; por fim, w^T ponto fixo de g implica que 1 é valor próprio de M, com vetor próprio esquerdo w^T, logo, devido à conclusão imediatamente anterior, $E_1 =< w >$.

Resta *4*.

Mas, na sequência do raciocínio acima, basta mostrar que qualquer vetor próprio esquerdo de M não negativo é um múltiplo (positivo) de w^T. Ora, dado que o subespaço gerado por C é $(M^T)^j-$invariante e que o seu único vetor não negativo é $0 \in \mathbb{R}^m$, então qualquer vetor próprio de M^T que seja linearmente independente de w tem alguma componente negativa, o que é equivalente a afirmar que qualquer vetor próprio esquerdo de M que seja linearmente independente de w^T tem alguma componente negativa. □

Lema 1.30. *Suponhamos que $\lambda_1, ..., \lambda_k$ (valores próprios de L, eventualmente repetidos) têm módulos estritamente menores que* 1. *Então existem $\mu < 1$ e uma norma, $\|.\|$, tais que $\forall v \;\; \|Lv\| \leq \mu\|v\|$.*

Demonstração. Consideramos as duas possibilidades:

(i) $\forall \lambda_1, ..., \lambda_k$ as multiplicidades algébrica e geométrica coincidem

(ii) $\exists \lambda_i$ cuja multiplicidade algébrica é superior à multiplicidade geométrica

Para (i) basta considerar $\mu = \max\{|\lambda_1|, ..., |\lambda_k|\}$: como existe uma base de $\mathbb{R}^k$ formada por vetores próprios de L então, nessa mesma base, a matriz de L é diagonal e tem-se $\|Lv\| = \sqrt{\lambda_1^2 v_1^2 + \lambda_2^2 v_2^2 + ... + \lambda_k^2 v_k^2} \leq \mu\|v\|$.

Para (ii) podemos supor, sem perda de generalidade, que λ_1 tem multiplicidade algébrica 2 e multiplicidade geométrica 1 e que $\forall \lambda_2, ..., \lambda_k$ as multiplicidades algébrica e geométrica coincidem. A forma canónica de Jordan conduz-nos à matriz

$$\begin{pmatrix} \lambda_1 & 1 & & & \\ 0 & \lambda_1 & & & \\ & & \lambda_2 & & \\ & & & \ddots & \\ & & & & \lambda_k \end{pmatrix}.$$

Assim, basta-nos obter $\rho < 1$ tal que $\|Kv\| \leq \rho\|v\|$ para $K = \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$ $(\mu = \max\{\rho, |\lambda_2|, ..., |\lambda_k|\})$.

Fixado $\varepsilon > 0$, seja $S_\varepsilon = \begin{pmatrix} \varepsilon & 0 \\ 0 & 1 \end{pmatrix}$. Ora, $S_\varepsilon K S_\varepsilon^{-1} = \begin{pmatrix} \lambda_1 & \varepsilon \\ 0 & \lambda_1 \end{pmatrix}$ e queremos provar que

$\|Kv\| = \sqrt{(\lambda_1 v_1 + \varepsilon v_2)^2 + (\lambda_1 v_2)^2} \leq \rho\sqrt{v_1^2 + v_2^2} = \rho\|v\|$.

Tem-se $\lambda_1^2 v_1^2 + 2\varepsilon\lambda_1 v_1 v_2 + \varepsilon^2 v_2^2 + \lambda_1^2 v_2^2 \leq (\lambda_1^2 + \varepsilon^2)(v_1^2 + v_2^2) + 2\varepsilon\lambda_1 v_1 v_2$

$$\leq (\lambda_1^2 + \varepsilon^2)(v_1^2 + v_2^2) + \varepsilon\lambda_1(v_1^2 + v_2^2) \leq (\lambda_1 + \varepsilon)^2(v_1^2 + v_2^2).$$

Logo, $\rho = |\lambda_1 + \varepsilon|$. □

Podemos, por fim, deduzir o Teorema Limite para Cadeias de Markov - **Teorema 1.20** - como corolário do Teorema de Perron - **Teorema 1.29**.

Observação 1.31. Todos os vetores na demonstração que se segue devem ser interpretados como matriz linha.

Demonstração. Notemos que no contexto das Cadeias de Markov nos interessa apenas considerar vetores do Simplexo Unitário ($(m-1)-$dimensional), S, espaço de todos os vetores estocásticos de $\mathbb{R}^m$.

Designemos por $h : S \to S$ a aplicação tal que $h(x) = xP$. Queremos provar que se u é qualquer um dos vetores da base canónica de $\mathbb{R}^m$, então $\lim_{n\to+\infty} h^n(u) = q$.

P é primitiva, o que nos deixa nas condições do Teorema de Perron, e é também estocástica. Assim, tal como foi visto para M na demonstração do mesmo Teorema, h tem um ponto fixo, seja q, com todas as componentes positivas; e como tal, q é vetor próprio esquerdo de P com valor próprio 1.

Ainda como na demonstração do Teorema de Perron, a restrição de P a $< q >^\perp$ só admite valores próprios de módulo estritamente menor do que 1, pelo que, pelo **Lema 1.30**, se $x \in < q >^\perp$ então $\lim_{n\to+\infty} \|xP^n\| = 0$, portanto, $\lim_{n\to+\infty} xP^n = 0$.

Ora, dado que qualquer u se escreve como $u = q + x$ onde $x \in < q >^\perp$, tem-se $\lim_{n\to+\infty} h^n(u) = \lim_{n\to+\infty} h^n(q + x) = \lim_{n\to+\infty}((q + x)P^n) = \lim_{n\to+\infty}(qP^n + xP^n) = \lim_{n\to+\infty} qP^n + \lim_{n\to+\infty} xP^n = q + 0 = q$. $\square$

Capítulo 2

O operador de Perron-Frobenius

O operador de Perron-Frobenius é uma ferramenta para o estudo dos Sistemas Dinâmicos de um ponto de vista probabilístico. Seja $(\Omega, \mathcal{B}, \lambda)$ um espaço de probabilidade e considere-se $X : \Omega \to \mathbb{R}$ uma variável aleatória (v.a.). O operador de Perron-Frobenius possibilita, em particular, o transporte (pela dinâmica) da densidade associada a X. Daqui em diante, escreveremos $\tau : \Omega \to \Omega$ para denotar a dinâmica.

2.1 Definição do Operador

Na terminologia acima, consideramos $\Omega = [a, b] \subset \mathbb{R}$, $\mathcal{B}$ a $\sigma-$álgebra de *Borel*, λ a medida de *Lebesgue* normalizada definida em $[a, b]$, e X a v.a. absolutamente contínua com respeito a λ à qual está associada a função densidade de probabilidade f, de modo que, se $A \in \mathcal{B}$,

$$\mathbb{P}(X \in A) = \int_A f d\lambda.$$

De facto, $\mu(A) \equiv \displaystyle\int_A f d\lambda$ é uma medida de probabilidade.

Seja $\tau : \Omega \to \Omega$ uma aplicação mensurável.

Definimos $\tau_* \mu(A) = \mu(\tau^{-1}(A))$ (o *pushforward* de μ por τ).

Se $\tau_* \mu \ll \mu$, portanto, por transitividade, $\tau_* \mu \ll \lambda$, então, pelo Teorema de Radon-Nikodym,

$$\exists \phi \in \mathcal{L}^1 : \tau_* \mu(A) = \int_A \phi d\lambda$$

e ϕ é única (*qtp*).

Ora, $X \circ \tau$ é uma v.a. e tem-se

$$\mathbb{P}(X \circ \tau \in A) = \mathbb{P}(X \in \tau^{-1}(A)) = \int_{\tau^{-1}(A)} f d\lambda = \mu(\tau^{-1}(A)) = \tau_* \mu(A) = \int_A \phi d\lambda.$$

Assim, nos termos da discussão que acaba de ser feita,

Definição 2.1. O *operador de Perron-Frobenius* é $P_\tau : \begin{array}{ccc} \mathcal{L}^1 & \to & \mathcal{L}^1 \\ f & \mapsto & \phi \end{array}$

Observação 2.2. $(\Omega, \mathcal{B}, \lambda)$ não tem que ser um espaço de probabilidade, apenas um espaço mensurável, e $X : \Omega \to \mathbb{R}$ uma aplicação mensurável (μ passa a ser uma medida com sinal e f a densidade correspondente).

2.2 Propriedades

P_τ goza das seguintes propriedades:

Proposição 2.3. *(Linearidade) $P_\tau : \mathcal{L}^1 \to \mathcal{L}^1$ é um operador linear.*

Demonstração. Dadas $f, g \in \mathcal{L}^1, \alpha, \beta \in \mathbb{R}$, queremos provar que

$$P_\tau(\alpha f + \beta g) = \alpha P_\tau(f) + \beta P_\tau(g).$$

Ora, $\displaystyle\int_A P_\tau(\alpha f + \beta g) d\lambda = \int_{\tau^{-1}(A)} (\alpha f + \beta g) d\lambda = \alpha \int_{\tau^{-1}(A)} f d\lambda + \beta \int_{\tau^{-1}(A)} g d\lambda$

$$= \alpha \int_A P_\tau(f) d\lambda + \beta \int_A P_\tau(g) d\lambda = \int_A (\alpha P_\tau(f) + \beta P_\tau(g)) d\lambda$$

pelo que, sendo A qualquer conjunto mensurável, então

$$P_\tau(\alpha f + \beta g) = \alpha P_\tau(f) + \beta P_\tau(g) \;\; qtp$$

(o que corresponde ao pretendido já que P_τ é definido a menos de equivalência *qtp*). □

Proposição 2.4. *(Positividade) Seja $f \in \mathcal{L}^1$ tal que $f \geq 0$. Então, $P_\tau(f) \geq 0$.*

Demonstração. $\displaystyle\int_A P_\tau(f) d\lambda = \int_{\tau^{-1}(A)} f d\lambda \geq 0$

pelo que, sendo A qualquer conjunto mensurável, então $P_\tau(f) \geq 0$. □

Proposição 2.5. *(Preservação de integrais)*

$$\int_\Omega P_\tau(f)d\lambda = \int_\Omega f d\lambda$$

Demonstração. $\displaystyle\int_\Omega P_\tau(f)d\lambda = \int_{\tau^{-1}(\Omega)} f d\lambda = \int_\Omega f d\lambda.$ □

Proposição 2.6. *(Contração) $P_\tau : \mathcal{L}^1 \to \mathcal{L}^1$ é uma contração i.e.*

$$\forall f \in \mathcal{L}^1 \ \|P_\tau(f)\|_1 \leq \|f\|_1.$$

Demonstração. Sejam $f^+ = \max\{f, 0\}$ e $f^- = -\min\{0, f\}$, de modo que $f \in \mathcal{L}^1 \implies f^+ \in \mathcal{L}^1, f^- \in \mathcal{L}^1$, $f = f^+ - f^-$ e $|f| = f^+ + f^-$.

Por linearidade, $P_\tau(f) = P_\tau(f^+ - f^-) = P_\tau(f^+) - P_\tau(f^-)$

logo, $|P_\tau(f)| \leq |P_\tau(f^+)| + |P_\tau(f^-)| = P_\tau(f^+) + P_\tau(f^-) = P_\tau(f^+ + f^-) = P_\tau(|f|)$.

Assim, $\displaystyle\|P_\tau(f)\|_1 = \int_\Omega |P_\tau(f)|d\lambda \leq \int_\Omega P_\tau(|f|)d\lambda = \int_\Omega |f|d\lambda = \|f\|_1$

(onde a penúltima igualdade vem da **Proposição 2.5**). □

Observação 2.7. Para a topologia induzida pela norma, $P_\tau : \mathcal{L}^1 \to \mathcal{L}^1$ é uma aplicação contínua.

Proposição 2.8. *(Composição) Seja $\sigma : \Omega \to \Omega$ nas mesmas condições de τ (i.e. σ mensurável e tal que $\sigma_*\mu \ll \mu$). Então, $P_{\sigma\circ\tau} = P_\sigma \circ P_\tau$.*

Demonstração. Comecemos por notar que $(\sigma \circ \tau)_*\mu \ll \mu$: se $\mu(A) = 0$ então, porque $\sigma_*\mu \ll \mu$, $\mu(\sigma^{-1}(A)) = 0$ pelo que, porque $\tau_*\mu \ll \mu$, $\mu(\tau^{-1}(\sigma^{-1}(A))) = 0$, ou seja, $\mu((\sigma \circ \tau)^{-1}(A)) = 0$.

Assim, como anteriormente, pelo Teorema de Radon-Nikodym vem $\displaystyle(\sigma\circ\tau)_*\mu(A) = \int_A \phi d\lambda$ para $\phi \in \mathcal{L}^1$ única *qtp*. Ora, $P_{\sigma\circ\tau}$ está, portanto, bem definido.

Agora, $\displaystyle\int_A P_{\sigma\circ\tau}(f)d\lambda = (\sigma\circ\tau)_*\mu(A) = \int_{(\sigma\circ\tau)^{-1}(A)} f d\lambda = \int_{\tau^{-1}(\sigma^{-1}(A))} f d\lambda = \int_{\sigma^{-1}(A)} P_\tau(f)d\lambda$

$$= \int_A P_\sigma(P_\tau(f))d\lambda = \int_A P_\sigma \circ P_\tau(f)d\lambda$$

de onde resulta $P_{\sigma\circ\tau}(f) = P_\sigma \circ P_\tau(f)$ *qtp*. □

Observação 2.9. Por indução tem-se que $P_{\tau^n} = P_\tau^n$.

Definição 2.10. Seja $\tau : \Omega \to \Omega$ uma aplicação mensurável. $\begin{array}{rccc} U_\tau : & \mathcal{L}^\infty & \to & \mathcal{L}^\infty \\ & f & \mapsto & f \circ \tau \end{array}$ é o *operador de Koopman.*

Proposição 2.11. *Sejam $f \in \mathcal{L}^1$, $g \in \mathcal{L}^\infty$. Então,*

$$\int_\Omega P_\tau(f) g d\lambda = \int_\Omega f U_\tau(g) d\lambda.$$

Demonstração. Seja f qualquer, $A \in \mathcal{B}$ e $g = \mathbb{1}_A$. Então,

$$\int_\Omega P_\tau(f) g d\lambda = \int_\Omega P_\tau(f) \mathbb{1}_A d\lambda = \int_A P_\tau(f) d\lambda = \int_{\tau^{-1}(A)} f d\lambda =$$

$$\int_\Omega f \mathbb{1}_{\tau^{-1}(A)} d\lambda = \int_\Omega f \mathbb{1}_A \circ \tau d\lambda = \int_\Omega f U_\tau(\mathbb{1}_A) d\lambda = \int_\Omega f U_\tau(g) d\lambda.$$

Uma vez que qualquer $g \in \mathcal{L}^\infty$ é limite de uma sucessão de funções simples (i.e. funções que são combinação linear de indicatrizes), o Teorema da Convergência Dominada de *Lebesgue* implica o resultado. □

Observação 2.12. A equação na **Proposição 2.11** é frequentemente usada como definição do operador P_τ.

Concluímos esta secção com uma condição necessária e suficiente para que uma densidade seja um ponto fixo de P_τ.

Proposição 2.13. *Seja $f \in \mathcal{L}^1$ tal que $f \geq 0$ e $\|f\|_1 = 1$. Então, $P_\tau(f) = f$ (qtp) se e só se μ tal que $\mu(A) = \int_A f d\lambda$ é $\tau-$invariante (i.e. $\mu(\tau^{-1}(A)) = \mu(A)$).*

Demonstração. Seja A um qualquer conjunto mensurável.

Suponhamos que $\mu(\tau^{-1}(A)) = \mu(A)$.

Ora, $\int_A P_\tau(f) d\lambda = \int_{\tau^{-1}(A)} f d\lambda = \mu(\tau^{-1}(A)) = \mu(A) = \int_A f d\lambda.$

Logo, $P_\tau(f) = f$ *qtp*.

Por outro lado, se $P_\tau(f) = f$ (*qtp*) então $\int_A P_\tau(f) d\lambda = \int_A f d\lambda$, e resulta

$$\mu(\tau^{-1}(A)) = \int_{\tau^{-1}(A)} f d\lambda = \int_A P_\tau(f) d\lambda = \int_A f d\lambda = \mu(A).$$

□

2.3 Aplicação

Depois de justificarmos uma fórmula para o cálculo de P_τ, válida no caso particular em que τ é uma aplicação monótona por pedaços, faremos uso da mesma através de um exemplo.

Notação: Se $[a,b] \subset \mathbb{R}$ e se tem $a = x_0 < x_1 < ... < x_{n-1} < x_n = b$, então $\mathcal{P}(x_0, ..., x_n)$ denota a partição $\{I_i = [x_{i-1}, x_i) : \ i = 1, ..., n-1, I_n = [x_{n-1}, x_n]\}$ de $[a,b]$.

De agora em diante, sempre que nos referirmos a $\tau : \Omega \to \Omega$ vamos assumir tratar-se de uma aplicação contínua à direita (o que, em particular, implica que existe pelo menos uma escolha de $\mathcal{P}(x_0, ..., x_n)$ tal que todas as restrições $\tau|_{I_i}$ são contínuas).

Observação 2.14. Não há qualquer perda de generalidade pelo que acabámos de assumir uma vez que caso existisse um conjunto finito de pontos nos quais τ não fosse contínua à direita então bastaria modificar as respetivas imagens de modo a obter τ contínua à direita; tal não acarreta qualquer prejuízo para as propriedades de integração de τ, das quais depende P_τ e, por conseguinte, toda a teoria em torno de si desenvolvida.

Definição 2.15. $\tau : \Omega \to \Omega$ diz-se *monótona por pedaços* se existe $\mathcal{P}(x_0, ..., x_n)$ tal que, para $i = 1, ..., n$, $\tau_i = \tau|I_i$ é monótona.

Vejamos então que se τ está nas condições da **Definição 2.15**, $P_\tau(f)$ pode ser descrito por:

$$P_\tau(f)(x) = \sum_{z \in \{\tau^{-1}(x)\}} \frac{f(z)}{|\tau'(z)|}.$$

De facto, por definição de P_τ, dado $A \in \mathcal{B}$

$$\int_A P_\tau(f) d\lambda = \int_{\tau^{-1}(A)} f d\lambda$$

e, como cada τ_i é invertível, então $\tau^{-1}(A) = \bigcup_{i=1}^{n} \tau_i^{-1}(A)$ (união disjunta) e resulta

$$\int_A P_\tau(f) d\lambda = \sum_{i=1}^{n} \int_{\tau_i^{-1}(A)} f d\lambda = \sum_{i=1}^{n} \int_{\tau(I_i) \cap A} f(\tau_i^{-1}(x)) |(\tau_i^{-1})'(x)| d\lambda =$$

$$\sum_{i=1}^{n} \int_A f(\tau_i^{-1}(x)) |(\tau_i^{-1})'(x)| \mathbb{1}_{\tau(I_i)}(x) d\lambda = \int_A \sum_{i=1}^{n} \frac{f(\tau_i^{-1}(x))}{|\tau_i'(\tau_i^{-1}(x))|} \mathbb{1}_{\tau(I_i)}(x) d\lambda.$$

Logo, $P_\tau(f)(x) = \sum_{i=1}^{n} \frac{f(\tau_i^{-1}(x))}{|\tau_i'(\tau_i^{-1}(x))|} \mathbb{1}_{\tau(I_i)}(x)$, o que pode ser reescrito como

$$P_\tau(f)(x) = \sum_{z \in \{\tau^{-1}(x)\}} \frac{f(z)}{|\tau'(z)|}.$$

Exemplo 2.16. Seja $\tau = g_a$, i.e. $\tau(x) = ax \mod 1$ (**Exemplo 1.6**).

Tem-se

$I_1 = [0, 1/a)$, $I_2 = [1/a, 2/a)$, ..., $I_a = [(a-1)/a, 1]$

$\tau_1(x) = ax$, $\tau_2(x) = ax - 1$, ..., $\tau_a(x) = ax - (a-1)$

de modo que

$$\begin{aligned} P_\tau(f)(x) &= \frac{f\left(\frac{x}{a}\right)}{\left|a\left(\frac{x}{a}\right)\right|} + \frac{f\left(\frac{x+1}{a}\right)}{\left|a\left(\frac{x+1}{a}\right)\right|} + \cdots + \frac{f\left(\frac{x+(a-1)}{a}\right)}{\left|a\left(\frac{x+(a-1)}{a}\right)\right|} \\ &= \frac{1}{a}\left(f\left(\frac{x}{a}\right) + f\left(\frac{x+1}{a}\right) + \cdots + f\left(\frac{x+(a-1)}{a}\right)\right). \end{aligned}$$

Capítulo 3

Dinâmica unidimensional expansora por pedaços

Neste capítulo dedicar-nos-emos em exclusivo a aplicações expansoras por pedaços definidas em intervalos fechados de números reais. Veremos que a tais aplicações estão sempre associadas medidas invariantes, absolutamente contínuas com respeito a *Lebesgue*. Além disso, mediante certas condições adicionais, a medida invariante é única e conduz a uma decomposição de um certo espaço vetorial de funções, o que constitui uma generalização do que foi visto no Capítulo 1 com o Teorema de Perron.

Em todo o capítulo continuaremos com $(\Omega, \mathcal{B}, \lambda)$ onde $\Omega = [a, b] \subset \mathbb{R}$, $\mathcal{B}$ a respetiva $\sigma-$álgebra de *Borel* e λ a medida de *Lebesgue* normalizada definida em $[a, b]$.

Na Secção 3.2, τ é sempre expansora por pedaços (mesmo que tal facto possa não ser mencionado nos enunciados de todos os resultados apresentados). Na Secção 3.3, τ será sempre expansora por pedaços e topologicamente misturadora.

3.1 Variação limitada

Antes de definirmos aplicação expansora por pedaços de domínio $[a, b] \subset \mathbb{R}$, vamos primeiro introduzir o conceito de função com variação limitada.

Definição 3.1. Dada uma função $h : [a,b] \to \mathbb{R}$, a *variação* de h é

$$var h = \sup \sum_{i=1}^{n} |h(x_{i-1}) - h(x_i)|$$

onde o supremo recai sobre todas as possíveis escolhas de $\mathcal{P}(x_0, ..., x_n)$ (conforme a notação fixada na secção 2.3).

Observação 3.2. Analogamente se define a variação de h em $\eta \subset [a,b]$, denotada por $var_\eta h$, considerando $\inf \eta = x_0 < x_1 < ... < x_{n-1} < x_n = \sup \eta$ e $\mathcal{P}(x_0, ..., x_n)$ a correspondente partição de $\overline{\eta}$.

Proposição 3.3. *Propriedades da variação:*

1. $var_\eta(h_1 + h_2) \leq var_\eta h_1 + var_\eta h_2$
2. $var_\eta(h_1 h_2) \leq var_\eta h_1 \sup_\eta |h_2| + \sup_\eta |h_1| var_\eta h_2$
3. $var_\eta(\varphi\psi) \leq var_\eta \varphi \sup_\eta |\psi| + \sup_\eta |\psi'| \int_\eta |\varphi| d\lambda$, *se* $\psi \in C^1(\overline{\eta})$
4. $var_\eta |\varphi| \leq var_\eta \varphi$
5. $var_\eta(\varphi \circ h) = var_{h(\eta)} \varphi$, *se* $h : \eta \to h(\eta)$ *é um homeomorfismo*
6. $var_\eta \int \varphi(t,.) d\theta(t) \leq \int var_\eta \varphi(t,.) d\theta(t)$ *para* θ *uma qualquer medida de probabilidade definida em* T *(espaço mensurável), e* $\varphi : T \times [a,b] \to \mathbb{R}$ *com* $var \varphi(t,.) < \infty$ *para todo o* $t \in T$.

Demonstração.

1.
$$\begin{aligned} var_\eta(h_1 + h_2) &= \sup \sum_{i=1}^{n} |(h_1 + h_2)(x_{i-1}) - (h_1 + h_2)(x_i)| \\ &= \sup \sum_{i=1}^{n} |h_1(x_{i-1}) + h_2(x_{i-1}) - h_1(x_i) - h_2(x_i)| \\ &\leq \sup \sum_{i=1}^{n} |h_1(x_{i-1}) - h_1(x_i)| + |h_2(x_{i-1}) - h_2(x_i)| \\ &= \sup \sum_{i=1}^{n} |h_1(x_{i-1}) - h_1(x_i)| + \sup \sum_{i=1}^{n} |h_2(x_{i-1}) - h_2(x_i)| = var_\eta h_1 + var_\eta h_2 \end{aligned}$$

2. $var_\eta(h_1h_2) = \sup\sum_{i=1}^{n}|(h_1h_2)(x_{i-1}) - (h_1h_2)(x_i)| = \sup\sum_{i=1}^{n}|h_1(x_{i-1})h_2(x_{i-1}) - h_1(x_i)h_2(x_i)|$

$$= \sup\sum_{i=1}^{n}|h_1(x_{i-1})h_2(x_{i-1}) - h_1(x_i)h_2(x_{i-1}) + h_1(x_i)h_2(x_{i-1}) - h_1(x_i)h_2(x_i)|$$

$$= \sup\sum_{i=1}^{n}|(h_1(x_{i-1}) - h_1(x_i))h_2(x_{i-1}) + h_1(x_i)(h_2(x_{i-1}) - h_2(x_i))|$$

$$\leq \sup\sum_{i=1}^{n}|(h_1(x_{i-1}) - h_1(x_i))h_2(x_{i-1})| + \sup\sum_{i=1}^{n}|h_1(x_i)(h_2(x_{i-1}) - h_2(x_i))|$$

$$\leq \sup_\eta|h_2|\sup\sum_{i=1}^{n}|h_1(x_{i-1}) - h_1(x_i)| + \sup_\eta|h_1|\sup\sum_{i=1}^{n}|h_2(x_{i-1}) - h_2(x_i)|$$

$$= \sup_\eta|h_2|var_\eta h_1 + \sup_\eta|h_1|var_\eta h_2$$

3. Retomamos a prova de *2.* na primeira desigualdade com $h_1 = \varphi$ e $h_2 = \psi$, ou seja,

$$var_\eta(\varphi\psi) \leq \sup\sum_{i=1}^{n}|(\varphi(x_{i-1}) - \varphi(x_i))\psi(x_{i-1})| + \sup\sum_{i=1}^{n}|\varphi(x_i)(\psi(x_{i-1}) - \psi(x_i))|$$

$$\leq \sup_\eta|\psi|var_\eta\varphi + \sup\sum_{i=1}^{n}\frac{|\varphi(x_i)(\psi(x_{i-1}) - \psi(x_i))|}{|x_{i-1} - x_i|}|x_{i-1} - x_i|$$

$$\leq \sup_\eta|\psi|var_\eta\varphi + \sup_\eta|\psi'|\sup\sum_{i=1}^{n}|\varphi(x_i)||x_{i-1} - x_i| \ (*).$$

Ora, como $\varphi \in \mathcal{L}^1$, da definição de integral de *Riemann* vem que para qualquer $\varepsilon > 0$ existe $\mathcal{P}(z_0, ..., z_t)$ refinamento de $\mathcal{P}(x_0, ..., x_n)$ (i.e. tal que $\{x_0, ..., x_n\} \subseteq \{z_0, ..., z_t\}$) tal que

$$\sum_{i=1}^{t}|\varphi(z_i)||z_{i-1} - z_i| \leq \int_\eta|\varphi|d\lambda + \varepsilon \ (**).$$

Vamos supor que $\mathcal{P}(z_0, ..., z_t)$ difere de $\mathcal{P}(x_0, ..., x_n)$ em apenas um ponto, digamos z_k tal que $x_j < z_k < x_l$, caso em que se observa

$$\sum_{i=1}^{t}|\varphi(z_{i-1}) - \varphi(z_i)| =$$

$$\sum_{i=1}^{j}|\varphi(z_{i-1}) - \varphi(z_i)| + |\varphi(z_j) - \varphi(z_k)| + |\varphi(z_k) - \varphi(z_l)| + \sum_{i=l+1}^{t}|\varphi(z_{i-1}) - \varphi(z_i)|$$

$$\geq \sum_{i=1}^{j}|\varphi(z_{i-1}) - \varphi(z_i)| + |\varphi(z_j) - \varphi(z_k) + \varphi(z_k) - \varphi(z_l)| + \sum_{i=l+1}^{t}|\varphi(z_{i-1}) - \varphi(z_i)|$$

$$= \sum_{i=1}^{j}|\varphi(x_{i-1}) - \varphi(x_i)| + |\varphi(x_j) - \varphi(x_l)| + \sum_{i=l+1}^{n}|\varphi(x_{i-1}) - \varphi(x_i)| = \sum_{i=1}^{n}|\varphi(x_{i-1}) - \varphi(x_i)|.$$

Significa, portanto, que a substituição de uma dada partição por um seu refinamento conduz sempre a um aumento do somatório cujo supremo define a variação.

Assim, é sempre possível encontrar $\mathcal{P}(z_0, ..., z_t)$ refinamento de $\mathcal{P}(x_0, ..., x_n)$, que satisfaz $(**)$, mas que ainda não possui tantos pontos quanto a partição sobre a qual se deve tomar $\sup\sum_{i=1}^{n}|\varphi(x_i)||x_{i-1} - x_i|$ de $(*)$. Logo, como ε é arbitrário resulta o pretendido.

4. resulta de se ter $|\varphi(x_{k-1}) - \varphi(x_k)| > ||\varphi(x_{k-1})| - \varphi(x_k)|$ sempre que $\varphi(x_{k-1}) < 0$.

5. $$\begin{aligned} var_\eta(\varphi \circ h) &= \sup\sum_{i=1}^{n}|(\varphi \circ h)(x_{i-1}) - (\varphi \circ h)(x_i)| = \sup\sum_{i=1}^{n}|\varphi(h(x_{i-1})) - \varphi(h(x_i))| \\ &= \sup\sum_{i=1}^{n}|\varphi(y_{i-1}) - \varphi(y_i)| = var_{h(\eta)}\varphi \end{aligned}$$

onde $y_k = h(x_k)$ para todo o $k \in \{0, ..., n\}$, pois, se h é um homeomorfismo, então $h(\mathcal{P}(x_0, ..., x_n)) = \mathcal{P}(h(x_0), ..., h(x_n)) = \mathcal{P}(y_0, ..., y_n)$.

6. $$\begin{aligned} var_\eta \int \varphi(t, .)d\theta(t) &= \sup\sum_{i=1}^{n}\left|\int \varphi(t, x_{i-1})d\theta(t) - \int \varphi(t, x_i)d\theta(t)\right| \\ &= \sup\sum_{i=1}^{n}\left|\int (\varphi(t, x_{i-1}) - \varphi(t, x_i))d\theta(t)\right| \\ &\leq \sup\sum_{i=1}^{n}\int |\varphi(t, x_{i-1}) - \varphi(t, x_i)|d\theta(t) \\ &= \sup\int \sum_{i=1}^{n}|\varphi(t, x_{i-1}) - \varphi(t, x_i)|d\theta(t) \\ &\leq \int \sup\sum_{i=1}^{n}|\varphi(t, x_{i-1}) - \varphi(t, x_i)|d\theta(t) = \int var_\eta\varphi(t, .)d\theta(t) \end{aligned}$$

□

Definição 3.4. $h : [a, b] \to \mathbb{R}$ tem *variação limitada* se $var h < +\infty$.

Observação 3.5. Resulta que h com variação limitada tem, no máximo, uma quantidade numerável de pontos de descontinuidade: $h = f_1 - f_2$ onde $f_1(x) = var h|_{[a,x]}$ e $f_2(x) = f_1(x) - h(x)$; ora, f_1 e f_2 são ambas não decrescentes pelo que a afirmação segue do **Lema 3.6**.

Lema 3.6. *Se $f : [a, b] \to \mathbb{R}$ é monótona então tem, no máximo, uma quantidade numerável de pontos de descontinuidade.*

Demonstração. Consideremos apenas o caso de f não decrescente. Suponhamos que f tem um conjunto não numerável de pontos de descontinuidade, e seja $x \in [a, b]$ um tal ponto. Então, $f(x^-) < f(x^+)$ e existe um racional $q(x)$ tal que $f(x^-) < q(x) < f(x^+)$.

Mas, como f é não decrescente, se $y \in [a, b]$ é outro ponto de descontinuidade de f tal que $y > x$, então $q(x) < q(y)$ – obtém-se uma função injetiva de um conjunto não numerável em $\mathbb{Q}$, o que é absurdo. $\square$

3.2 $\tau : \Omega \to \Omega$ expansora por pedaços

Definição 3.7. Dado $m \geq 1$, $\mathcal{P}^{(m)}(x_0, ..., x_n)$, na forma abreviada $\mathcal{P}^{(m)}$, é a partição obtida de $\mathcal{P}(x_0, ..., x_n)$ quando se impõe $\mathcal{P}^{(m)}(x) = \mathcal{P}^{(m)}(y) \iff \mathcal{P}(\tau^j(x)) = \mathcal{P}(\tau^j(y))$ $\forall j \in \{0, ..., m-1\}$.[1]

Notação: $g_\eta^{(m)} = \dfrac{1}{|(\tau^m|\eta)'|}$, para $\eta \in \mathcal{P}^{(m)}$

Definição 3.8. $\tau : \Omega \to \Omega$ diz-se *expansora por pedaços* se existe $\mathcal{P}(x_0, ..., x_n)$ tal que

(i) para $i = 1, ..., n$, $\tau_i = \tau|int(I_i)$ é de classe C^1, $|\tau_i'(x)| > 0$ e $g_i(x) \equiv \dfrac{1}{|\tau_i'(x)|}$ tem variação limitada;

(ii) existem $C_1 > 0$ e $\lambda_1 < 1$ tais que $\sup g_\eta^{(m)} \leq C_1 \lambda_1^m$ para todo o $m \geq 1$.

Observação 3.9. Segundo (i) da **Definição 3.8**, porque todas as τ_i são de classe C^1 e as respetivas derivadas têm módulo positivo, então uma aplicação expansora por pedaços é, em particular, monótona por pedaços; a variação limitada das g_i é uma condição de regularidade. Por (ii) da mesma definição, vem que as derivadas laterais de τ são estritamente maiores do que 1 em todos os pontos do intervalo Ω.

Observação 3.10. Se τ é expansora por pedaços em $\mathcal{P}(x_0, ..., x_n)$, então τ^m é expansora por pedaços em $\mathcal{P}^{(m)}$.

[1] $\mathcal{P}^{(1)} = \mathcal{P}(x_0, ..., x_n)$, como seria desejável

Exemplo 3.11. Exemplos de aplicações expansoras por pedaços:

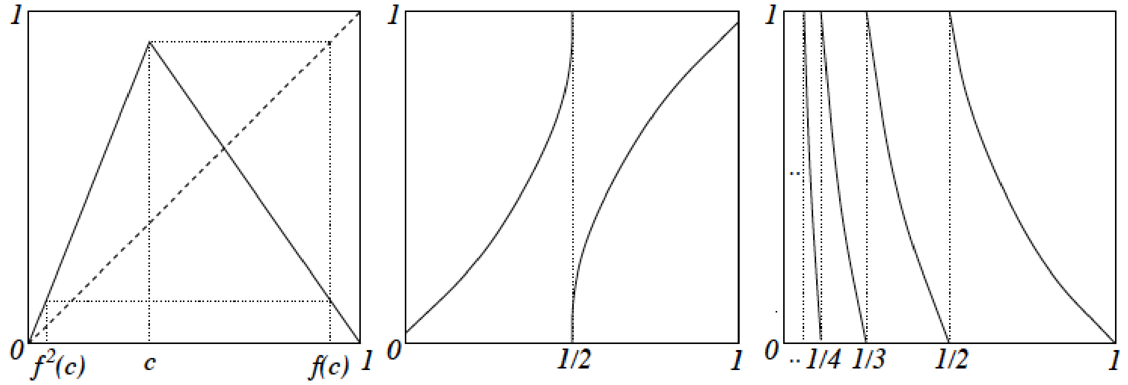

Figura 3.1: Gráficos das aplicações: tenda, tipo *Lorenz*, *Gauss* (da esquerda para a direita) [Via97]

Proposição 3.12. *Existem* $C_2 > 0$ *e* $\lambda_2 < 1$ *tais que* $var g_\eta^{(m)} \leq C_2 \lambda_2^m$ *para todo o* $m \geq 1$*; mais precisamente,* $C_2 = \sup\{(C_1^3/\lambda_1) m (\lambda_1/\lambda_2)^m : m \geq 1\}$ *para* $\lambda_2 \in (\lambda_1, 1)$ *previamente fixado.*

Demonstração. Comecemos por tomar $C_1 > 0$ tão grande quanto necessário de modo que $var g_i \leq C_1$ para todo o $i \in \{1, ..., n\}$.

Dados $\eta \in \mathcal{P}^{(m)}$ e $0 \leq j < m$, consideremos

- $\rho_j \in \mathcal{P}^{(j)}$ tal que $\eta \subset \rho_j$
- $\widetilde{\eta}_j \in \mathcal{P}(x_0, ..., x_n)$ tal que $\tau^j(\eta) \subset \widetilde{\eta}_j$
- $\zeta_j \in \mathcal{P}^{(m-(j+1))}$ tal que $\tau^{j+1}(\eta) \subset \zeta_j$

de modo que $g_\eta^{(m)} = g_{\zeta_j}^{(m-(j+1))}|_{\tau^{j+1}(\eta)} g_{\widetilde{\eta}_j}|_{\tau^j(\eta)} g_{\rho_j}^{(j)}|_\eta$.

Ora, por *2.* da **Proposição 3.3**, então

$$
\begin{aligned}
var g_\eta^{(m)} &\leq var g_{\zeta_j}^{(m-(j+1))}|_{\tau^{j+1}(\eta)} \sup |g_{\widetilde{\eta}_j}|_{\tau^j(\eta)} g_{\rho_j}^{(j)}|_\eta| + \sup |g_{\zeta_j}^{(m-(j+1))}|_{\tau^{j+1}(\eta)}| var g_{\widetilde{\eta}_j}|_{\tau^j(\eta)} g_{\rho_j}^{(j)}|_\eta \\
&\leq var g_{\zeta_j}^{(m-(j+1))}|_{\tau^{j+1}(\eta)} \sup |g_{\widetilde{\eta}_j}|_{\tau^j(\eta)}| \sup |g_{\rho_j}^{(j)}|_\eta| + \\
&\qquad \sup |g_{\zeta_j}^{(m-(j+1))}|_{\tau^{j+1}(\eta)}| \left(var g_{\widetilde{\eta}_j}|_{\tau^j(\eta)} \sup |g_{\rho_j}^{(j)}|_\eta| + \sup |g_{\widetilde{\eta}_j}|_{\tau^j(\eta)}| var g_{\rho_j}^{(j)}|_\eta \right) \\
&\leq \sum_{j=0}^{m-1} \sup |g_{\zeta_j}^{(m-(j+1))}|_{\tau^{j+1}(\eta)}| var g_{\widetilde{\eta}_j}|_{\tau^j(\eta)} \sup |g_{\rho_j}^{(j)}|_\eta| \leq \sum_{j=0}^{m-1} C_1 \lambda_1^{m-(j+1)} C_1 C_1 \lambda_1^j \\
&= m(C_1^3/\lambda_1)\lambda_1^m
\end{aligned}
$$

pelo que, fixando $\lambda_2 \in (\lambda_1, 1)$, se $C_2 = \sup\{m(C_1^3/\lambda_1)(\lambda_1/\lambda_2)^m : m \geq 1\}$ vem $varg_\eta^{(m)} \leq C_2\lambda_2^m$. $\square$

Proposição 3.13. *Seja $f : \Omega \to \mathbb{R}$ com variação limitada. Então, existem $C_0 > 0$ e $\lambda_0 < 1$ tais que*

$$varP_\tau^m(f) \leq C_0\lambda_0^m varf + C_0 \int |f| d\lambda$$

para todo o $m \geq 1$.

Observação 3.14. À desigualdade da **Proposição 3.13** chamamos desigualdade de Lasota-Yorke.

Demonstração. Sabemos que $P_\tau^m(f) = P_{\tau^m}(f)$ (**Proposição 2.8** e **Observação 2.9**). Assim,

$$P_\tau^m(f) = \sum_{\eta \in \mathcal{P}^{(m)}} g_\eta^{(m)} f \circ (\tau^m|\eta)^{-1} \mathbb{1}_{\tau^m(\eta)}$$

e, usando propriedades da variação (**Proposição 3.3**), vem

$$varP_\tau^m(f) \leq \sum_{\eta \in \mathcal{P}^{(m)}} varg_\eta^{(m)} f \circ (\tau^m|\eta)^{-1} \mathbb{1}_{\tau^m(\eta)} = \sum_{\eta \in \mathcal{P}^{(m)}} varg_\eta^{(m)} f|\eta \mathbb{1}_{\tau^m(\eta)}$$

$$\leq \sum_{\eta \in \mathcal{P}^{(m)}} varg_\eta^{(m)} f|\eta + 2 \sup g_\eta^{(m)} f|\eta \leq \sum_{\eta \in \mathcal{P}^{(m)}} varg_\eta^{(m)} \sup f|\eta + \sup g_\eta^{(m)} varf|\eta + 2 \sup g_\eta^{(m)} f|\eta$$

$$\leq \sum_{\eta \in \mathcal{P}^{(m)}} (varg_\eta^{(m)} + 2 \sup g_\eta^{(m)}) \sup f|\eta + \sup g_\eta^{(m)} varf|\eta.$$

Pelo Teorema do Valor Médio para integrais,

$$\exists z \in \eta : |f(z)| = \frac{1}{\lambda(\eta)} \int_\eta |f| d\lambda$$

portanto, porque $\sup|f||\eta \leq |f(z)| + var|f||\eta$,

$$\sup f|\eta \leq \sup|f||\eta \leq var|f||\eta + \frac{1}{\lambda(\eta)} \int_\eta |f| d\lambda \leq varf|\eta + \frac{1}{\lambda(\eta)} \int |f| d\lambda \quad (*)$$

logo, substituindo na expressão imediatamente anterior,

$$varP_\tau^m(f) \leq \sum_{\eta \in \mathcal{P}^{(m)}} (varg_\eta^{(m)} + 2 \sup g_\eta^{(m)}) \left(varf|\eta + \frac{1}{\lambda(\eta)} \int |f| d\lambda \right) + \sup g_\eta^{(m)} varf|\eta.$$

Temos ainda que $varg_\eta^{(m)} \leq C_2\lambda_2^m$ e $\sup g_\eta^{(m)} \leq C_1\lambda_1^m$, portanto,

$$varP_\tau^m(f) \leq \sum_{\eta \in \mathcal{P}^{(m)}} (C_2\lambda_2^m + 2C_1\lambda_1^m) \left(varf|\eta + \frac{1}{\lambda(\eta)} \int |f| d\lambda \right) + C_1\lambda_1^m varf|\eta$$

pelo que, como $C_2\lambda_2^m = \sup\{(C_1^3/\lambda_1)m(\lambda_1/\lambda_2)^m : m \geq 1\}\lambda_2^m \geq (C_1^3/\lambda_1)m(\lambda_1/\lambda_2)^m\lambda_2^m = mC_1^3\lambda_1^{m-1} \geq C_1\lambda_1^m$ [2], fica

$$\leq 4C_2\lambda_2^m \sum_{\eta\in\mathcal{P}^{(m)}} var f|\eta + 3C_2\lambda_2^m \sum_{\eta\in\mathcal{P}^{(m)}} \frac{1}{\lambda(\eta)} \int |f|d\lambda \leq C_3\lambda_3^m var(f) + K_3(m)\int |f|d\lambda$$

onde $C_3 = 4C_2$, $\lambda_3 = \lambda_2$, e $K_3(m) = 3C_2\lambda_2^m(\#\mathcal{P}^{(m)})\sup\left\{\frac{1}{\lambda(\eta)} : \eta \in \mathcal{P}^{(m)}\right\}$.

Ora, para chegarmos à fórmula pretendida, resta-nos remover a dependência de m em K_3.

Tomemos N tal que $C_3\lambda_3^N \leq 1/2$ e $\widehat{K} = \max\{K_3(m) : 1 \leq m \leq N\}$.

Como para qualquer $m \geq 1$ temos que $m = qN + r$, com $q \geq 0$ e $1 \leq r \leq N$, então

$$var P_\tau^m(f) \leq C_3\lambda_3^N var P_\tau^{m-N}(f) + K_3(N)\int |P_\tau^{m-N}(f)|d\lambda.$$

Relembrando que $|f| = f^+ + f^-$ [3], então

$$\begin{aligned} P_\tau^{m-N}(|f|) =& P_\tau^{m-N}(f^+ + f^-) \\ &= P_\tau^{m-N}(f^+) + P_\tau^{m-N}(f^-) \ (\textbf{Proposição 2.3}) \\ &= |P_\tau^{m-N}(f^+)| + |P_\tau^{m-N}(f^-)| \ (\textbf{Proposição 2.4}) \\ &\geq |P_\tau^{m-N}(f^+) - P_\tau^{m-N}(f^-)| = |P_\tau^{m-N}(f^+ - f^-)| = |P_\tau^{m-N}(f)| \end{aligned}$$

pelo que

$$\begin{aligned} var P_\tau^m(f) \leq& \frac{1}{2} var P_\tau^{m-N}(f) + \widehat{K}\int P_\tau^{m-N}(|f|)d\lambda \\ &= \frac{1}{2} var P_\tau^{m-N}(f) + \widehat{K}\int |f|d\lambda \ (\textbf{Proposição 2.5}). \end{aligned}$$

Recursivamente,

$$\begin{aligned} var P_\tau^m(f) \leq& \frac{1}{2^{q-1}}\left(\frac{1}{2} var P_\tau^{m-qN}(f) + \widehat{K}\int |f|d\lambda\right) + \left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{q-2}}\right)\widehat{K}\int |f|d\lambda \\ &= \frac{1}{2^q} var P_\tau^r(f) + \left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{q-1}}\right)\widehat{K}\int |f|d\lambda \\ &\leq \frac{1}{2^q}\left(C_3\lambda_3^r var(f) + \widehat{K}\int |f|d\lambda\right) + \left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{q-1}}\right)\widehat{K}\int |f|d\lambda \\ &= \frac{1}{2^q} C_3\lambda_3^r var(f) + \left(1 + \frac{1}{2} + \cdots + \frac{1}{2^q}\right)\widehat{K}\int |f|d\lambda \\ &\leq \frac{1}{2^q} C_3\lambda_3^r var(f) + 2\widehat{K}\int |f|d\lambda. \end{aligned}$$

[2] o que é válido já que podemos sempre tomar $C_1 \geq 1$

[3] f^+ e f^- conforme definidas na demonstração da **Proposição 2.6**

Ora, $C_0 > 0$ e $\lambda_0 < 1$ tais que

$$var P_\tau^m(f) \leq C_0 \lambda_0^m var f + C_0 \int |f| d\lambda$$

e acrescentando a informação da desigualdade que acabámos de deduzir, podem então ser tais que $C_0 \geq \max\{C_3, 2\widehat{K}\}$ e $\max\{\lambda_3, \frac{1}{\sqrt[N]{2}}\} \leq \lambda_0 < 1$. □

Corolário 3.15. *Seja $C(a)$ o cone de funções $f : \Omega \to \mathbb{R}$, não negativas, e tais que $var f \leq a \int f d\lambda$. Então, para a suficientemente grande, existe $N \geq 1$ tal que*

$$P_\tau^N(C(a)) \subset C(a/2).$$

Demonstração. Tomemos $N \geq 1$ tal que $C_0 \lambda_0^N \leq 1/4$. Assim, dada $f \in C(a)$,

$$var P_\tau^N(f) \leq \frac{1}{4} var f + C_0 \int f d\lambda \leq \left(\frac{a}{4} + C_0\right) \int f d\lambda.$$

Logo, basta que $C_0 \leq \frac{a}{4}$ e temos o pretendido, ou seja, $a \geq 4C_0$. □

Portanto, se a dinâmica é determinada por uma aplicação expansora por pedaços, então a propriedade de variação limitada é preservada para densidades e há subespaços do espaço vetorial das funções com variação limitada que são positivamente invariantes.

3.2.1 Medidas invariantes

Teorema 3.16 (Helly)**.** *Seja $\psi_n : \Omega \to \mathbb{R}$, $n \geq 1$, uma sucessão de funções. Suponhamos que existem $K_1, K_2 > 0$ tais que, para todo o $n \geq 1$, $\sup \psi_n \leq K_1$ e $var \psi_n \leq K_2$. Então, existe uma subsucessão $(\psi_{n_k})_k$ e uma função $\psi_0 : \Omega \to \mathbb{R}$, satisfazendo também $\sup \psi_0 \leq K_1$ e $var \psi_0 \leq K_2$, tais que $\lim_{k \to +\infty} \psi_{n_k} = \psi_0$ $\lambda-$qtp e em $\mathcal{L}^1$.*

Demonstração. Comecemos por listar os principais passos da demonstração:

(i) para cada $n \in \mathbb{N}$, escrevemos ψ_n como diferença entre ψ_n^+ e ψ_n^-, ambas funções não decrescentes e uniformemente limitadas

(ii) fixado $q \in [0,1] \cap \mathbb{Q}$, encontramos um subconjunto de índices $(n_k)_k$ tais que ambas as sucessões $\psi_{n_k}^+(q)$ e $\psi_{n_k}^-(q)$ são convergentes, para, respetivamente, $\psi_0^+(q)$ e $\psi_0^-(q)$

(iii) estendemos $\psi_0^{\pm}$ ao intervalo $[0,1]$, e observamos que a densidade dos racionais em $\mathbb{R}$ assegura a convergência de tal extensão em todos os pontos de continuidade de $\psi_0^{\pm}$, que são, precisamente, os pontos de $[0,1] \setminus \mathbb{Q}$

(iv) definimos ψ_0 como diferença de funções contínuas à direita obtidas das $\psi_0^{\pm}$

Assumimos, sem perda de generalidade, que $\Omega = [0,1]$.

Seja $\psi_n^+(x) = var\psi_n|_{[0,x]}$; $\psi_n^- = \psi_n^+ - \psi_n$.

Como $q_1 \leq q_2 \implies \psi_0^{\pm}(q_1) \leq \psi_0^{\pm}(q_2)$, então podemos estender $\psi_0^{\pm}$ a funções não decrescentes em $[0,1]$, tomando

$$\psi_0^{\pm}(x) = \inf\{\psi_0^{\pm}(q) : q \in [x,1] \cap \mathbb{Q}\}.$$

Seja $x \in [0,1]$ um ponto de continuidade de $\psi_0^{\pm}$. Para qualquer $\delta > 0$, existem $q_1 \leq x \leq q_2$ tais que

$$\psi_0^{\pm}(x) - \delta \leq \psi_0^{\pm}(q_1) \leq \psi_0^{\pm}(x) \leq \psi_0^{\pm}(q_2) \leq \psi_0^{\pm}(x) + \delta$$

de modo que, para k suficientemente grande,

$$\psi_0^{\pm}(x) - 2\delta \leq \psi_0^{\pm}(q_1) - \delta \leq \psi_{n_k}^{\pm}(q_1) \leq \psi_{n_k}^{\pm}(x) \leq \psi_{n_k}^{\pm}(q_2) \leq \psi_0^{\pm}(q_2) + \delta \leq \psi_0^{\pm}(x) + 2\delta$$

i.e. para todo o ponto de continuidade de $\psi_0^{\pm}$ tem-se que $\lim_{k\to+\infty} \psi_{n_k}^{\pm}(x) = \psi_0^{\pm}(x)$.

Designemos por $\widetilde{\psi}_0^{\pm}$ as funções contínuas à direita que em cada x ponto de continuidade de $\psi_0^{\pm}$ tomam o valor $\psi_0^{\pm}(x)$; definimos $\psi_0 = \widetilde{\psi}_0^{+} - \widetilde{\psi}_0^{-}$. Assim, ψ_{n_k} converge para ψ_0 em todo o ponto de continuidade de $\psi_0^{\pm}$; mas, sendo os pontos de descontinuidade de $\psi_0^{\pm}$ necessariamente números racionais (consequência da densidade dos racionais em $\mathbb{R}$), então $\lim_{k\to+\infty} \psi_{n_k} = \psi_0$ $\lambda-qtp$. Como todas as $\psi_{n_k}^{\pm}$ são uniformemente limitadas, também $\lim_{k\to+\infty} \psi_{n_k}^{\pm} = \psi_0^{\pm}$ em $\mathcal{L}^1$ (pelo Teorema da Convergência Dominada de *Lebesgue*), logo $\lim_{k\to+\infty} \psi_{n_k} = \psi_0$ em $\mathcal{L}^1$.

Por fim, uma vez que as condições $\sup \psi_0 \leq K_1$ e $var\psi_0 \leq K_2$ só poderão falhar em pontos de descontinuidade de ψ_0, sejam a e $a_1 \leq b_1 \leq \cdots \leq a_s \leq b_s$ tais pontos. Tem-se que

$$|\psi_0(a)| = \lim_k |\psi_{n_k}(a)| \leq K_1$$

$$\sum_{j=1}^{s} |\psi_0(a_j) - \psi_0(b_j)| = \lim_k \sum_{j=1}^{s} |\psi_{n_k}(a_j) - \psi_{n_k}(b_j)| \leq \sup_k var\psi_{n_k} \leq K_2.$$

Assim, como ψ_0 é contínua à direita, vem que $\sup \psi_0 \leq K_1$ e $var\psi_0 \leq K_2$. $\square$

Notação: Escrevemos $\mu = fd\lambda$ quando μ é dada por $\mu(A) = \displaystyle\int_A fd\lambda$, para A qualquer conjunto mensurável.

Corolário 3.17. *τ admite pelo menos uma medida de probabilidade, μ_0, absolutamente contínua e invariante. Além disso, a qualquer medida μ com tais propriedades está associada uma densidade f com variação limitada.*

Demonstração. Demonstremos primeiro a existência de μ_0, medida de probabilidade absolutamente contínua e $\tau-$invariante.

Para $n \geq 2$, consideremos $\varphi_n = \dfrac{1}{n}\displaystyle\sum_{j=0}^{n-1} \mathcal{P}_\tau^j(1)$. Notemos que se obtém imediatamente das propriedades do operador de Perron-Frobenius que $\varphi_n > 0$ e $\|\varphi_n\|_1 = 1$.

$var\varphi_n \leq \dfrac{1}{n}\displaystyle\sum_{j=0}^{n-1} var\mathcal{P}_\tau^j(1)$ e, pela **Proposição 3.13**, então $var\varphi_n \leq \dfrac{1}{n}\displaystyle\sum_{j=0}^{n-1} C_0 \int 1d\lambda = C_0$

$\sup \varphi_n \leq$ [4]$var\varphi_n + \dfrac{1}{\lambda(\Omega)}\displaystyle\int_\Omega \varphi_n d\lambda = C_0 + 1$

Estamos nas condições do **Teorema 3.16**, portanto existe uma subsucessão $(\varphi_{n_k})_k$ convergente (*qtp*) para φ_0, que, por sua vez, também satisfaz $var\varphi_0 \leq C_0$ e $\sup \varphi_0 \leq C_0 + 1$.

Claro que $\varphi_0 \geq 0$ e $\|\varphi_0\|_1 = \|\lim_{k\to+\infty} \varphi_{n_k}\|_1 = \lim_{k\to+\infty}\|\varphi_{n_k}\|_1 = 1$.

Escrevemos $\mathcal{P}_\tau(\varphi_{n_k}) = \varphi_{n_k} + \dfrac{1}{n_k}(\mathcal{P}_\tau^{n_k}(1) - 1)$. Como $\mathcal{P}_\tau$ é contínuo (**Observação 2.7**), $\mathcal{P}_\tau(\varphi_0) = \mathcal{P}_\tau(\lim_{k\to+\infty} \varphi_{n_k}) = \lim_{k\to+\infty} \mathcal{P}_\tau(\varphi_{n_k}) = \varphi_0$.

Ou seja, φ_0 é um ponto fixo de $\mathcal{P}_\tau$, logo $\mu_0 = \varphi_0 d\lambda$, medida de probabilidade, é $\tau-$invariante pela **Proposição 2.13**.

Para a segunda parte, seja então μ uma medida de probabilidade absolutamente contínua e $\tau-$invariante, portanto, à qual está associada uma densidade de probabilidade ψ que, pela **Proposição 2.13**, é um ponto fixo de $\mathcal{P}_\tau$.

O objetivo é provar que ψ coincide (*qtp*) com uma função com variação limitada.

Dado que qualquer função não negativa e mensurável pode ser aproximada por uma sucessão de funções simples que é, em particular, uma sucessão de funções com variação limitada, então podemos afirmar o seguinte: existe $(\psi_l)_l$ sucessão de funções com variação

[4]consequência do Teorema do Valor Médio para integrais, como na demonstração da **Proposição 3.13**

limitada que converge para ψ (qtp); podemos assumir que, para todo o l, $\|\psi_l\|_1 \leq 2$.

Para cada l, consideremos a sucessão $(\psi_n^{(l)})_n$ tal que $\psi_n^{(l)} = \frac{1}{n}\sum_{j=0}^{n-1}\mathcal{P}_\tau^j(\psi_l)$.

$var\psi_n^{(l)} \leq \frac{1}{n}\sum_{j=0}^{n-1} var\mathcal{P}_\tau^j(\psi_l) \leq \frac{1}{n}\sum_{j=0}^{n-1} C_0\left(\lambda_0^j var\psi_l + \int \psi_l d\lambda\right)$ (pela **Proposição 3.13**),

e então $\lim_{n\to+\infty} var\psi_n^{(l)} \leq C_0 \int \psi_l d\lambda \leq 2C_0$ o que garante que, quando muito excetuando um número finito de termos, os elementos de $(\psi_n^{(l)})_n$ têm variação menor ou igual que $2C_0$.

Como $\mathcal{P}_\tau$ é uma contração (**Proposição 2.6**), temos ainda que

$$\left\|\frac{1}{n}\sum_{j=0}^{n-1}\mathcal{P}_\tau^j(\psi_l)\right\|_1 \leq \|\psi_l\|_1 \leq 2$$

logo, conjuntamente com o facto que $var\psi_n^{(l)} \leq 2C_0$, resulta que $\sup\psi_n^{(l)} \leq 2(C_0+1)$.

Assim, pelo **Teorema 3.16**, existe uma subsucessão $(\psi_{n_k}^{(l)})_k$ convergente para $\widetilde{\psi}^{(l)}$, e $var\widetilde{\psi}^{(l)} \leq 2C_0$ e $\sup\widetilde{\psi}^{(l)} \leq 2(C_0+1)$. E agora, porque o **Teorema 3.16** pode ser aplicado a $(\widetilde{\psi}^{(l)})_l$, obtemos ρ, com $var\rho \leq 2C_0$ e $\sup\rho \leq 2(C_0+1)$, como limite de uma subsucessão de $(\widetilde{\psi}^{(l)})_l$.

Mas $\|\widetilde{\psi}^{(l)} - \psi\|_1 = \lim \left\|\frac{1}{n_k}\sum_{j=0}^{n_k-1}\mathcal{P}_\tau^j(\psi_l-\psi)\right\|_1 \leq \|\psi_l - \psi\|_1$ implica que $\widetilde{\psi}^{(l)}$ converge para ψ (qtp) e, portanto, $\psi = \rho$ (qtp), ou seja, $\mu = \psi d\lambda = \rho d\lambda$. □

Corolário 3.18. *Há apenas um número finito de medidas de probabilidade que são ergódicas, absolutamente contínuas e invariantes.*

Demonstração. Suponhamos que μ é uma medida que satisfaz as propriedades acima.

Como se trata de uma medida de probabilidade absolutamente contínua e invariante, então, pelo **Corolário 3.17**, existe f com variação limitada tal que $\mu = fd\lambda$. Sabemos que:

- f admite, no máximo, uma quantidade numerável de pontos de descontinuidade (consequência de f ter variação limitada), pelo que, designando por A o conjunto dos pontos de descontinuidade de f, então $\lambda(A) = 0$

- $\int_{\Omega} f d\lambda = 1$ e $f \geq 0$ (consequência de f ser densidade de probabilidade), pelo que

  se $B = \{x : f(x) > 0\}$ então $\lambda(B) > 0$

logo, $\lambda(B \cap A^c) > 0$ o que implica que $B \cap A^c \neq \emptyset$ i.e. existe y tal que $f(y) > 0$ e f é contínua em y; assim, $f > 0$ em algum intervalo $J \subset \Omega$.

Vamos ver que $P_{\tau^m}(f) > 0$ em $\tau^m(J)$ para todo o $m \in \mathbb{N}$.

Como μ é $\tau-$invariante, temos que f é um ponto fixo de P_τ, e portanto, $f = P_\tau^m(f) = P_{\tau^m}(f)$. Relembrando que

$$P_\tau^m(f) = \sum_{\eta \in \mathcal{P}^{(m)}} g_\eta^{(m)} f \circ (\tau^m|\eta)^{-1} \mathbb{1}_{\tau^m(\eta)}$$

de onde deduzimos $P_\tau^m(f)(x) = \displaystyle\sum_{z \in \{\tau^{-m}(x)\}} \frac{f(z)}{|(\tau^m)'(z)|}$, então, se $x \in J$,

$P_\tau^m(f)(\tau^m(x)) = \displaystyle\sum_{z \in \{\tau^{-m}(\tau^m(x))\}} \frac{f(z)}{|(\tau^m)'(z)|}$ pelo que, em particular, $\dfrac{f(x)}{|(\tau^m)'(x)|}$ constitui uma parcela da soma. Mas então, como $\dfrac{f(x)}{|(\tau^m)'(x)|} > 0$ e as restantes parcelas são garantidamente não negativas, resulta que $P_\tau^m(f)(\tau^m(x)) > 0$, como pretendíamos.

Em particular, acabámos de verificar que qualquer intervalo J de continuidade para f é tal que $\mu(\tau^m(J)) > 0$, i.e. $\lambda \ll \mu$ em cada $\tau^m(J)$. Assim, μ e λ são medidas equivalentes em cada $\tau^m(J)$.

Ora, como μ é ergódica, tem-se que $\displaystyle\frac{1}{n}\sum_{j=0}^{n-1} f(\tau^j(x))$ converge para $\displaystyle\int f d\mu$ para $\mu - qtp\ x$, logo $\displaystyle\frac{1}{n}\sum_{j=0}^{n-1} f(\tau^j(x))$ converge para $\displaystyle\int f d\mu$ para $x \in \tau^m(J)$ *Lebesgue-qtp.*

Por um lado, algum dos $x_1, ..., x_{n-1}$ (de $\mathcal{P}^{(m)}(x_0, ..., x_n)$) tem que pertencer a algum dos iterados de J pois, se assim não fosse, então a medida de tais iterados acabaria por ultrapassar a medida de Ω. Por outro lado, como medidas ergódicas distintas têm que ser suportadas em conjuntos disjuntos, então os iterados de J a que nos referimos estão contidos em conjuntos disjuntos, que, consequentemente, são, no máximo, $n - 1$. $\square$

3.3 $\tau : \Omega \to \Omega$ expansora por pedaços e topologicamente misturadora

Definição 3.19. $\tau : \Omega \to \Omega$ diz-se *topologicamente misturadora* se existe $I \subset \Omega$ tal que

(i) $\tau(I) = I$;

(ii) para todo o $x \in int(\Omega)$ existe $n \in \mathbb{N}$ tal que $\tau^n(x) \in I$;

(iii) para todo o $J \subset I$ existe $n \in \mathbb{N}$ tal que $\tau^n(J) = I$.

Corolário 3.20. *Existe uma única medida de probabilidade, μ_0, que é absolutamente contínua e invariante. Além disso, tal medida é ergódica e tem I como suporte.*

Demonstração. Suponhamos que existem μ_1, μ_2 medidas de probabilidade absolutamente contínuas e invariantes, portanto, tais que $\mu_1 = f_1 d\lambda$ e $\mu_2 = f_2 d\lambda$ com $\displaystyle\int f_1 d\lambda = 1 = \int f_2 d\lambda$ e f_1, f_2 pontos fixos de P_τ. Pelo **Corolário 3.17**, sabemos ainda que f_1 e f_2 têm variação limitada.

Consideremos os conjuntos $X_1 = \{x : f_1(x) \geq f_2(x)\}$ e $X_2 = \{x : f_1(x) < f_2(x)\}$.

Definimos medidas $\nu_1 = (f_1 - f_2)\mathbb{1}_{X_1} d\lambda$ e $\nu_2 = (f_2 - f_1)\mathbb{1}_{X_2} d\lambda$, que são absolutamente contínuas com respeito a *Lebesgue*.

Como $f_1 - f_2$ é um ponto fixo de P_τ, então

$$\begin{aligned}\int_{X_1} (f_1 - f_2) d\lambda &= \int_{X_1} P_\tau(f_1 - f_2) d\lambda = \int_{\tau^{-1}(X_1)} (f_1 - f_2) d\lambda \\ &\leq \int_{\tau^{-1}(X_1) \cap X_1} (f_1 - f_2) d\lambda \leq \int_{X_1} (f_1 - f_2) d\lambda\end{aligned}$$

(a primeira desigualdade deve-se ao facto de $f_1 - f_2$ tomar apenas valores negativos no complementar de X_1).

Assim, tem-se que $\lambda(\tau^{-1}(X_1)) = \lambda(\tau^{-1}(X_1) \cap X_1) = \lambda(X_1)$, o que implica que ν_1 é $\tau-$invariante.

De forma inteiramente análoga se chega às mesmas conclusões para X_2 e ν_2.

Ora, mais uma vez pelo **Corolário 3.17**, vem que tanto $(f_1 - f_2)\mathbb{1}_{X_1}$ como $(f_2 - f_1)\mathbb{1}_{X_2}$ têm variação limitada.

Suponhamos que $(f_1 - f_2)\mathbb{1}_{X_1}(x) \neq 0$ para todo o $x \in \Omega$ a menos de um conjunto de medida nula; então, existe um intervalo $J \subset I$ tal que $(f_1 - f_2)\mathbb{1}_{X_1}(x) > 0 \, \forall x \in J$[5] $-$ de facto, porque τ é topologicamente misturadora, se existe $J \subset I$ nas referidas condições, então $J = I$.

Portanto, se $(f_1 - f_2)\mathbb{1}_{X_1}(x) \neq 0 \ \lambda - qtp$, então I está contido no suporte de ν_1. Mas então, $(f_2 - f_1)\mathbb{1}_{X_2}(x) = 0 \ \lambda - qtp$ pois, caso contrário, também I estaria contido no suporte de ν_2 $-$ contradição com a hipótese de que ν_1 e ν_2 são distintas.

Logo, $f_1 \geq f_2$ *qtp* pelo que, como $\displaystyle\int f_1 d\lambda = \int f_2 d\lambda$, $f_1 = f_2$ *qtp* $-$ contradição com a hipótese de que μ_1 e μ_2 são distintas.

Supusemos a existência de medidas μ_1 e μ_2 distintas e chegamos a um absurdo. Portanto, só pode existir uma única medida de probabilidade absolutamente contínua e invariante.

Resta apenas ver que μ_0 é ergódica.

Consideremos $A \subset I$ tal que $\mu_0(A) > 0$ e $\tau(A) = A$ (o que implica $\tau^{-1}(A) = A$). Seja μ a medida de probabilidade tal que, para $B \in \mathcal{B}$,

$$\mu(B) = \frac{\mu_0(B \cap A)}{\mu_0(A)}.$$

É imediato que μ é absolutamente contínua com respeito a *Lebesgue*.

Vejamos que μ é $\tau-$invariante:

$$\begin{aligned}\mu(\tau^{-1}(B)) = & \frac{\mu_0(\tau^{-1}(B) \cap A)}{\mu_0(A)} = \frac{\mu_0(\tau^{-1}(B) \cap \tau^{-1}(A))}{\mu_0(A)} = \frac{\mu_0(\tau^{-1}(B \cap A))}{\mu_0(A)} \\ = & \frac{\mu_0(B \cap A)}{\mu_0(A)} = \mu(B)\end{aligned}$$

Logo, necessariamente $\mu_0 = \mu$. Em particular, $\mu_0(A) = 1$, e resulta que μ_0 é ergódica. $\square$

3.3.1 Decaimento de correlações e o Teorema do Limite Central

Vamos agora mostrar que no caso de a aplicação τ ser monótona por pedaços e topologicamente misturadora existe um decaimento exponencial de correlações entre funções com variação limitada e que as mesmas satisfazem o Teorema do Limite Central.

[5] a justificação deste facto vista aquando da demonstração do **Corolário 3.18**

Recordemos que, pelo **Corolário 3.20**, se τ satisfizer tais requisitos então há uma única medida absolutamente contínua e invariante, μ_0, que, além do mais, é ergódica. Também sabemos que lhe está associada uma densidade com variação limitada, isto pelo **Corolário 3.17**.

- **caso em que τ tem derivada limitada**

Consideremos a restrição adicional $\exists \delta > 0 : \forall i\ g_i(x) \geq \delta$ (nos termos da **Definição 3.8**), que nos permitirá chegar mais diretamente à conclusão pretendida. Posteriormente, será abordado o problema em toda a sua generalidade.

Vamos adotar a notação que se segue, onde I tal como na **Definição 3.19**.

Notação: $C(a)$ passa a designar o cone de funções $f : I \to \mathbb{R}$, não negativas, e tais que $var_I f \leq a \int_I f d\lambda$, e $\mathcal{P}^{(m)}$ a partição de I que se obtém pela restrição a I dos elementos de $\mathcal{P}^{(m)}$.

Observação 3.21. O **Corolário 3.15** permanece válido após a redefinição de $C(a)$ que acabámos de fazer.

Definição 3.22. $\theta_a : C(a) \times C(a) \to \mathbb{R}_0^+$ é a métrica em $C(a)$ dada por

$$\theta_a(\varphi_1, \varphi_2) = \log \frac{\beta(\varphi_1, \varphi_2)}{\alpha(\varphi_1, \varphi_2)}$$

com $\beta(\varphi_1, \varphi_2) = \inf\{s > 0 : s\varphi_1 - \varphi_2 \in C(a)\}$ e $\alpha(\varphi_1, \varphi_2) = \sup\{t > 0 : \varphi_2 - t\varphi_1 \in C(a)\}$.

Proposição 3.23. *Existe $k \geq 1$ tal que $D = \sup\{\theta_a(\psi_1, \psi_2) : \psi_1, \psi_2 \in P_\tau^{kN}(C(a))\}$ é finito.*[6]

Demonstração. Olhemos primeiro para $\sup\{\theta_a(\psi_1, \psi_2) : \psi_1, \psi_2 \in C(a/2)\}$.

Ora, sem perda de generalidade, podemos supor que $\int \psi_1 d\lambda = 1 = \int \psi_2 d\lambda$.

Como $var(\psi_2 - \frac{1}{3}\psi_1) \leq var\psi_2 + \frac{1}{3}var\psi_1 \leq \frac{a}{2}\int \psi_2 d\lambda + \frac{a}{6}\int \psi_1 d\lambda = \frac{2a}{3} = a\int(\psi_2 - \frac{1}{3}\psi_1)d\lambda$

então $\psi_2 - \frac{1}{3}\psi_1 \in C(a)$. Por outro lado, $(\psi_2 - t\psi_1)(x) \geq 0 \iff t \leq \frac{\psi_2}{\psi_1}(x)$.

Assim, $\alpha(\psi_1, \psi_2) \geq \inf\{\frac{1}{3}, \frac{\psi_2}{\psi_1}(x) : x \in I\}$.

[6] N tal como no **Corolário 3.15**

Concluímos também que $\beta(\psi_1, \psi_2) \leq \sup\{3, \frac{\psi_1}{\psi_2}(x) : x \in I\}$, por um raciocínio análogo.

Afirmamos agora que existem $k \geq 1$ e $\gamma > 0$ tais que, se $\psi \in P_\tau^{kN}(C(a))$,

$$\gamma \int \psi d\lambda \leq \inf \psi \leq \sup \psi \leq \frac{1}{\gamma} \int \psi d\lambda.$$

Portanto, já que $P_\tau^{kN}(C(a)) \subset C(a/2)$ (pelo **Corolário 3.15**), resulta que para quaisquer $\psi_1, \psi_2 \in P_\tau^{kN}(C(a))$ com $\int \psi_1 d\lambda = 1 = \int \psi_2 d\lambda$ se tem

$$\alpha(\psi_1, \psi_2) \geq \min\{\frac{1}{3}, \gamma^2\} \text{ e } \beta(\psi_1, \psi_2) \leq \max\{3, \frac{1}{\gamma^2}\}.$$

Ou seja, se a afirmação for válida, tomando $\gamma \leq 1/\sqrt{3}$ obtemos $D \leq 4\log(1/\gamma) = 4\log(\sqrt{3})$.

Resta, portanto, justificar a afirmação.

Seja $k = q + j$, onde:

- $q \geq 1$ é tal que $\lambda(\eta) \leq 1/2a$ para todo o $\eta \in \mathcal{P}^{(qN)}$ (q existe por (ii) da **Definição 3.8**);
- $j \geq 0$ é tal que $\tau^{(q+j)N}(\eta) = I$ (j existe porque τ é topologicamente misturadora)

Seja $\psi = P_\tau^{kN}(\varphi)$, para $\varphi \in C(a)$.

Como consequência do Teorema do Valor Médio, temos que, para qualquer $\eta \in \mathcal{P}^{(qN)}$,

$$var\varphi|\eta \geq \frac{1}{\lambda(\eta)} \int_\eta \varphi d\lambda - \inf \varphi|\eta \geq 2a \left(\int_\eta \varphi d\lambda - \lambda(\eta) \inf \varphi|\eta \right)$$

pelo que

$$var\varphi \geq 2a \int \varphi d\lambda - 2a \sum_{\eta \in \mathcal{P}^{(qN)}} \lambda(\eta) \inf \varphi|\eta \geq 2a \int \varphi d\lambda - 2a \max_{\eta \in \mathcal{P}^{(qN)}} \inf \varphi|\eta$$

logo, se $\varphi \in C(a)$,

$$\max_{\eta \in \mathcal{P}^{(qN)}} \inf \varphi|\eta \geq \frac{1}{2} \int \varphi d\lambda$$

i.e. existe $\eta_0 \in \mathcal{P}^{(qN)}$ tal que $\inf \varphi|\eta_0 \geq \frac{1}{2} \int \varphi d\lambda$.

Assim, se $y \in I$,

$$\psi(y) = P_\tau^{kN}(\varphi)(y) = \sum_{\zeta \in \mathcal{P}^{(kN)}} g_\zeta^{(kN)} \varphi \circ (\tau^{kN}|\zeta)^{-1}(y) \geq \sum_{\zeta \subset \eta_0} g_\zeta^{(kN)} \varphi \circ (\tau^{kN}|\zeta)^{-1}(y).$$

Mas $\sum_{\zeta\subset\eta_0} g_\zeta^{(kN)}\varphi\circ(\tau^{kN}|\zeta)^{-1}(y)\geq\delta^{kN}\frac{1}{2}\int\varphi d\lambda$ (porque $y\in\eta_0$), e concluímos que

$$\inf\psi\geq\delta^{kN}\frac{1}{2}\int\varphi d\lambda=\delta^{kN}\frac{1}{2}\int\psi d\lambda,$$

$$\sup\psi\leq\int\psi d\lambda+var\psi\leq(a+1)\int\psi d\lambda$$

o que prova a afirmação. □

Corolário 3.24. *$P_\tau^{kN}:C(a)\to P_\tau^{kN}(C(a))$ é uma $\Lambda-$contração, com $\Lambda=1-e^{-D}$.*

Demonstração. O objetivo é provar que $\theta_a(P_\tau^{kN}(\varphi_1),P_\tau^{kN}(\varphi_2))\leq\Lambda\theta_a(\varphi_1,\varphi_2)$, quaisquer que sejam $\varphi_1,\varphi_2\in C(a)$.

Ora, $\beta(\varphi_1,\varphi_2)=\inf\{s>0:s\varphi_1-\varphi_2\in C(a)\}$ e $\alpha(\varphi_1,\varphi_2)=\sup\{t>0:\varphi_2-t\varphi_1\in C(a)\}$, pelo que existem sucessões $(s_n)_n$ e $(t_n)_n$, respetivamente decrescente para $\beta(\varphi_1,\varphi_2)$ e crescente para $\alpha(\varphi_1,\varphi_2)$, tais que, para todo o n, $s_n\varphi_1-\varphi_2\in C(a)$ e $\varphi_2-t_n\varphi_1\in C(a)$.

Assim, para todo o n, tem-se que $\theta_a(P_\tau^{kN}(s_n\varphi_1-\varphi_2),P_\tau^{kN}(\varphi_2-t_n\varphi_1)\leq D$, pela **Proposição 3.23**.

E agora, existem sucessões $(S_n)_n$ e $(T_n)_n$, respetivamente decrescente para $\beta(P_\tau^{kN}(s_n\varphi_1-\varphi_2),P_\tau^{kN}(\varphi_2-t_n\varphi_1))$ e crescente para $\alpha(P_\tau^{kN}(s_n\varphi_1-\varphi_2),P_\tau^{kN}(\varphi_2-t_n\varphi_1))$, tais que $\lim_{n\to+\infty}\log\frac{S_n}{T_n}\leq D$ e, para todo o n,

$S_nP_\tau^{kN}(s_n\varphi_1-\varphi_2)-P_\tau^{kN}(\varphi_2-t_n\varphi_1)\in C(a)$ e $P_\tau^{kN}(\varphi_2-t_n\varphi_1)-T_nP_\tau^{kN}(s_n\varphi_1-\varphi_2)\in C(a)$.

Mas então, resulta que

$$(S_ns_n+t_n)P_\tau^{kN}(\varphi_1)-(S_n+1)P_\tau^{kN}(\varphi_2)\in C(a)\implies\beta(P_\tau^{kN}(\varphi_1),P_\tau^{kN}(\varphi_2))\leq\frac{S_ns_n+t_n}{S_n+1}$$

$$(1+T_n)P_\tau^{kN}(\varphi_2)-(t_n+T_ns_n)P_\tau^{kN}(\varphi_1)\in C(a)\implies\alpha(P_\tau^{kN}(\varphi_1),P_\tau^{kN}(\varphi_2))\geq\frac{t_n+T_ns_n}{1+T_n}$$

e, portanto,

$\theta_a(P_\tau^{kN}(\varphi_1),P_\tau^{kN}(\varphi_2))\leq\log\left(\frac{S_ns_n+t_n}{S_n+1}\frac{1+T_n}{t_n+T_ns_n}\right)\leq\log(S_ns_n+t_n)-\log(S_n+1)+\log(1+T_n)-\log(t_n+T_ns_n)=\log(S_ns_n+t_n)+\log\left(\frac{1}{s_n}\right)-\log(S_n+1)+\log(1+T_n)-\log(t_n+T_ns_n)-\log\left(\frac{1}{s_n}\right)=\log\left(S_n+\frac{t_n}{s_n}\right)-\log(S_n+1)+\log(1+T_n)-\log\left(\frac{t_n}{s_n}+T_n\right)=\int_1^{\frac{t_n}{s_n}}\frac{1}{S_n+x}dx-\int_1^{\frac{t_n}{s_n}}\frac{1}{x+T_n}dx=\int_0^{\log(t_n/s_n)}\frac{e^x}{S_n+e^x}dx-\int_0^{\log(t_n/s_n)}\frac{e^x}{e^x+T_n}dx=$

$$\int_0^{\log(t_n/s_n)} \frac{e^x}{S_n+e^x} - \frac{e^x}{e^x+T_n} dx = \int_0^{\log(t_n/s_n)} \frac{e^x(T_n-S_n)}{(S_n+e^x)(e^x+T_n)} dx \leq$$

$$\log\left(\frac{t_n}{s_n}\right) \sup_{x\in[0,\log(t_n/s_n)]} \frac{e^x(T_n-S_n)}{(S_n+e^x)(e^x+T_n)} =$$

$$\log\left(\frac{s_n}{t_n}\right) \sup_{x\in[0,\log(t_n/s_n)]} \frac{e^x(S_n-T_n)}{(S_n+e^x)(e^x+T_n)} \leq \log\left(\frac{s_n}{t_n}\right)\left(1-\frac{T_n}{S_n}\right).$$

Logo, se $n \to +\infty$, $\theta_a(P_\tau^{kN}(\varphi_1), P_\tau^{kN}(\varphi_2)) \leq (1-e^{-D})\theta_a(\varphi_1,\varphi_2)$. □

Notação: C_+ designa o cone de funções $f : \Omega \to \mathbb{R}$ não negativas.

Definição 3.25. $\theta_+ : C_+ \times C_+ \to \mathbb{R}_0^+$ é a métrica em C_+ dada por

$$\theta_+(\varphi_1,\varphi_2) = \log \frac{\sup(\varphi_2/\varphi_1)}{\inf(\varphi_2/\varphi_1)}.$$

Observação 3.26. $\theta_+|C(a) \times C(a)$ é mais fina do que θ_a.

Lema 3.27. *Seja φ uma função com variação limitada. Então, $\varphi = \varphi^+ - \varphi^-$ com $\varphi^+, \varphi^- \in C(a)$.*

Demonstração. Seja M tal que $var\varphi = M$. Ora, existem φ^+, φ^- com variação limitada tais que $\varphi = \varphi^+ - \varphi^-$ que, como $var\varphi = var(\varphi^+ - \varphi^-) \leq var\varphi^+ + var\varphi^-$, escolhemos satisfazendo $var\varphi^+ = \dfrac{M}{2}$ e $var\varphi^- = \dfrac{M}{2}$.

Resta encontrar a tal que $\dfrac{M}{2} \leq a \displaystyle\int \varphi^+ d\lambda$ e $\dfrac{M}{2} \leq a \displaystyle\int \varphi^- d\lambda$. Mas então, basta que $a \geq \dfrac{M}{2\max\left\{\displaystyle\int \varphi^+ d\lambda, \int \varphi^- d\lambda\right\}}$. □

Proposição 3.28. *Sejam φ uma função com variação limitada e ψ uma função não negativa e integrável. Então, existe $C(\varphi,\psi) > 0$ tal que*

$$\left|\int (\psi\circ\tau^n)\varphi d\lambda - \int \psi d\mu_0 \int \varphi d\lambda\right| \leq C(\varphi,\psi)\Lambda^n.$$

Demonstração. Depois do **Lema 3.27**, basta agora mostrar o resultado para $\varphi \in C(a)$.

Supomos que $\displaystyle\int \varphi d\lambda = 1$. Tem-se que

$$\left|\int (\psi\circ\tau^n)\varphi d\lambda - \int \psi d\mu_0\right| = \left|\int \psi(P_\tau^n(\varphi) - \varphi_0)d\lambda\right| \leq \sup|P_\tau^n(\varphi) - \varphi_0| \|\psi\|_1$$

e $\sup|P_\tau^n(\varphi) - \varphi_0| \leq \sup|\varphi_0| \sup\left|\dfrac{P_\tau^n(\varphi)}{\varphi_0} - 1\right|$.

Mas, como $\varphi_0 \in C(a)$, podemos afirmar que $\sup|\varphi_0| \leq \frac{1}{\gamma}$[7], e então

$\sup|P_\tau^n(\varphi) - \varphi_0| \leq \frac{1}{\gamma}(e^{\theta_+(\varphi_0, P_\tau^n(\varphi))} - 1)$.

Ora, $\theta_+(\varphi_0, P_\tau^n(\varphi)) \leq \theta_a(\varphi_0, P_\tau^n(\varphi)) = \theta_a(P_\tau^{n-kN}(P_\tau^{kN}(\varphi_0)), P_\tau^{n-kN}(P_\tau^{kN}(\varphi))) \leq$

$\Lambda^{n-kN}\theta_a(P_\tau^{kN}(\varphi_0), P_\tau^{kN}(\varphi)) \leq \Lambda^{n-kN}D \leq C'\Lambda^n$ $(C' = \Lambda^{-kN}D)$.

Assim, $\sup|P_\tau^n(\varphi) - \varphi_0| \leq \frac{1}{\gamma}(e^{C'\Lambda^n} - 1) = \frac{1}{\gamma}\left(\sum_{i=0}^{+\infty}\frac{(C'\Lambda^n)^i}{i!} - 1\right) = \frac{1}{\gamma}\sum_{i=1}^{+\infty}\frac{(C'\Lambda^n)^i}{i!} =$
$\frac{1}{\gamma}\Lambda^n\sum_{i=1}^{+\infty}\frac{(C')^i\Lambda^{n(i-1)}}{i!} = C''\Lambda^n$ $(C'' = \frac{1}{\gamma}\sum_{i=1}^{+\infty}\frac{(C')^i\Lambda^{n(i-1)}}{i!})$. □

O decaimento exponencial de correlações entre funções com variação limitada é agora uma consequência imediata.

Corolário 3.29. *Sejam φ_1, φ_2 funções com variação limitada.*
Então, existe $C(\varphi_1, \varphi_2) > 0$ tal que

$$\left|\int(\varphi_1 \circ \tau^n)\varphi_2 d\mu_0 - \int\varphi_1 d\mu_0 \int\varphi_2 d\mu_0\right| \leq C(\varphi_1, \varphi_2)\Lambda^n.$$

Demonstração. Basta aplicar a **Proposição 3.28** a $\psi = \varphi_1$ e $\varphi = \varphi_2\varphi_0$. □

Proposição 3.30 (Teorema do Limite Central)**.** *Seja φ uma função com variação limitada. Se $\phi = \varphi - \int\varphi d\mu_0$ e $\sigma^2 = \int\phi^2 d\mu_0 + 2\sum_{j=1}^{\infty}\int\phi(\phi\circ\tau^j)d\mu_0 \in (0, +\infty)$, então*

$$\lim_{n\to+\infty}\mu_0\{x : \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1}\varphi\circ\tau^j(x) - \int\varphi d\mu_0 \in A\} = \int_A \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{t^2}{2\sigma^2}}dt.$$

- **caso "geral"**

Notação: $BV \subset \mathcal{L}^1$ designa o espaço de funções com variação limitada; por abuso de linguagem, se $f \in BV$ dizemos que 'f é BV'.

Notação: $spec(P_\tau)$ designa o espectro de $P_\tau : BV \to BV$.

Se nenhuma restrição adicional for imposta a τ (expansora por pedaços e topologicamente misturadora), o decaimento de correlações e o Teorema do Limite Central para funções BV podem ser deduzidos após a análise de $spec(P_\tau)$ no mesmo espaço de funções.

[7] Demonstração da **Proposição** 3.23.

Mais precisamente, se $P_\tau : BV \to BV$ é tal que $spec(P_\tau) = \{1\} \cup \Sigma_0$ onde Σ_0 está contido num disco de raio $\Lambda < 1$, veremos que estamos perante uma generalização do que já foi visto para as cadeias de Markov, onde há um único ponto fixo para o qual convergem todos os restantes pontos do espaço.

Definição 3.31. $\|.\|_{BV} : BV \to \mathbb{R}_0^+$ é a norma em BV dada por

$$\|\varphi\|_{BV} = var\varphi + \int |\varphi| d\lambda.$$

Ora, já vimos que $P_\tau : BV \to BV$ está bem definido (pela **Proposição 3.13**) portanto, se φ é BV,

$$\|P_\tau^m(\varphi)\|_{BV} = \|P_{\tau^m}(\varphi)\|_{BV} = varP_{\tau^m}(\varphi) + \int |P_{\tau^m}(\varphi)| d\lambda.$$

Mas, como $|P_{\tau^m}(\varphi)| \leq P_{\tau^m}(|\varphi|)$[8] e $P_{\tau^m}(|\varphi|)$ é uma função positiva, então

$$\int |P_{\tau^m}(\varphi)| d\lambda \leq \int P_{\tau^m}(|\varphi|) d\lambda = \int |\varphi| d\lambda \text{ (pela } \textbf{Proposição 2.5}\text{)}$$

e $varP_{\tau^m}(\varphi) \leq C_0 \lambda_0^m var\varphi + C_0 \int |\varphi| d\lambda$ para alguns $C_0 > 0$ e $\lambda < 1$ (pela **Proposição 3.13**), pelo que

$$\|P_\tau^m(\varphi)\|_{BV} \leq C_0 \lambda_0^m var\varphi + (C_0 + 1) \int |\varphi| d\lambda \leq \max\{C_0 \lambda_0^m, C_0 + 1\} \|\varphi\|_{BV}.$$

Em particular, para qualquer m, $\|P_\tau^m\|_{BV} \leq \max\{C_0 \lambda_0, C_0 + 1\}$, portanto, dada φ, todas as normas $\|P_\tau^m(\varphi)\|_{BV}$ são uniformemente limitadas.

Mas então, nenhum dos elementos de $spec(P_\tau)$ pode ser, em módulo, superior a 1.

Relembrando que existe φ_0 com variação limitada tal que $\mu_0 = \varphi_0 d\lambda$, para μ_0 (a única) medida de probabilidade absolutamente contínua e invariante, que é necessariamente um ponto fixo de P_τ (**Proposição 2.13**), resulta que 1 é, de facto, valor próprio de P_τ.

Agora, como consequência de P_τ poder ser bem aproximado por um conjunto de operadores de dimensão finita[9], vem que $spec(P_\tau) = \{1, \sigma_1, ..., \sigma_k\} \cup \Sigma_0$ onde $\sigma_1, ..., \sigma_k$ são valores próprios de módulo 1 e multiplicidade finita e Σ_0 está contido num disco de raio $\Lambda < 1$.

O próximo objetivo é mostrar que não existem $\sigma_1, ..., \sigma_k$, o que resulta de μ_0 ser exata, como veremos.

[8] Demonstração da **Proposição** 3.13.

[9] ver [Via97]

Definição 3.32. Seja μ uma medida definida no espaço mensurável $(\Omega, \mathcal{B})$. Diz-se que μ é:

- *mixing* se $\lim_{n\to+\infty} \mu(\tau^{-n}(A) \cap B) = \mu(A)\mu(B)$ para quaisquer $A, B \in \mathcal{B}$.
- *exata* se $\bigcap_{n=0}^{+\infty} \tau^{-n}(\mathcal{B})$ ou tem medida nula ou tem medida total.

Lema 3.33. *Se μ é uma medida exata, então μ é mixing.*

Lema 3.34. *Dado $\beta > 0$ existe $N \geq 1$ tal que para todo o $n \geq 0$ existe um subconjunto $\mathcal{Q}_{n+N} \subset \mathcal{P}^{(n+N)}$ de intervalos de monotonia de τ^{n+N} com $\mu_0\left(\bigcup_{\eta\in\mathcal{Q}_{n+N}} \eta\right) \geq 1-\beta$ e*

1. *$\tau^n(\eta) \in \mathcal{P}^{(N)}$ para todo o $\eta \in \mathcal{Q}_{n+N}$*
2. *$e^{-\beta}\dfrac{\mu_0(\xi)}{\mu_0(\eta)} \leq \dfrac{\mu_0(\tau^n(\xi))}{\mu_0(\tau^n(\eta))} \leq e^{\beta}\dfrac{\mu_0(\xi)}{\mu_0(\eta)}$ para todo o $\xi \subset \eta \in \mathcal{Q}_{n+N}$*

Demonstração. Começaremos por construir um subconjunto de $\mathcal{P}^{(n+N)}$, $\mathcal{Q}^1_{n+N}$, tal que *1.* é válida para todos os seus elementos e a medida total dos restantes intervalos de $\mathcal{P}^{(n+N)}$ é pequena, para N suficientemente grande.

Notemos que da definição de $\mathcal{P}^{(m)}$ (**Definição 3.7**) resulta que se $\eta \in \mathcal{P}^{(N)}$ então

- $\forall x \in int(\eta), \tau^j(x) \notin \{x_0, ..., x_n\}$ qualquer que seja $j = 0, ..., N-1$
- se $x \in \partial(\eta)$ então $\tau^j(x) \in \{x_0, ..., x_n\}$ para algum $j = 0, ..., N-1$

pelo que, em particular, se $\eta \in \mathcal{P}^{(n+N)}$ é tal que $\tau^j(\partial(\eta)) \notin \{x_0, ..., x_n\}$ $\forall j \in \{0, ..., n-1\}$, então $\tau^j(\partial(\eta)) \in \{x_0, ..., x_n\}$ para algum $j \in \{n, ..., n+N-1\}$, i.e. $\tau^n(\eta) \in \mathcal{P}^{(N)}$; assim, tomamos $\mathcal{Q}^1_{n+N}$ o conjunto de todos os $\eta \in \mathcal{P}^{(n+N)}$ com a propriedade

$$\tau^j(\partial(\eta)) \notin \{x_0, ..., x_n\} \ \forall j \in \{0, ..., n-1\} \ (*).$$

Fixemos $j \in \{0, ..., n-1\}$, e seja $\eta \in \mathcal{P}^{(n+N)}$ tal que $\tau^j(\partial(\eta)) \in \{x_0, ..., x_n\}$. Então, $\tau^j(\eta) \subset \zeta \in \mathcal{P}^{(n+N-j)}$ com ζ tal que $\partial(\zeta) \cap \{x_0, ..., x_n\} \neq \emptyset$. Como, no máximo, existem $2n$ ζ's em tais condições, e cada um deles satisfaz $\lambda(\zeta) \leq C_1\lambda_1^{n+N-j}$, então

$$\mu_0(\cup\zeta) \leq 2n\,\text{sup}|\varphi_0|C_1\lambda_1^{n+N-j}$$

e, porque μ_0 é invariante, então também qualquer $\eta \in \mathcal{P}^{(n+N)} \setminus \mathcal{Q}^1_{n+N}$ é tal que $\mu_0(\eta) \leq \text{sup}|\varphi_0|C_1\lambda_1^{n+N-j}$, de onde

$$\mu_0(\cup\eta) \leq \sum_{j=0}^{n-1} 2n\,\text{sup}|\varphi_0|C_1\lambda_1^{n+N-j} \leq 2n\,\text{sup}|\varphi_0|C_1 \sum_{k=N+1}^{\infty} \lambda_1^k \leq \frac{\beta}{4}$$

para N suficientemente grande.

Prosseguimos no sentido de provar *2.* para todo o intervalo η de um subconjunto grande de $\mathcal{Q}^1_{n+N}$.

Justificaremos apenas a primeira desigualdade, já que a segunda se obtém de forma inteiramente análoga.

Tem-se

$$\frac{\mu_0(\tau^n(\xi))}{\mu_0(\tau^n(\eta))} = \frac{\int_{\tau^n(\xi)} \varphi_0 d\lambda}{\int_{\tau^n(\eta)} \varphi_0 d\lambda} = \frac{\int_\xi (\varphi_0 \circ \tau^n)|(\tau^n)'|d\lambda}{\int_\eta (\varphi_0 \circ \tau^n)|(\tau^n)'|d\lambda} \geq \frac{\inf_\xi((\varphi_0 \circ \tau^n)|(\tau^n)'|/\varphi_0)}{\sup_\eta((\varphi_0 \circ \tau^n)|(\tau^n)'|/\varphi_0)} \frac{\mu_0(\xi)}{\mu_0(\eta)}.$$

Denotemos por $g(x) = \dfrac{1}{|\tau'(x)|}$, desde que $x \notin \{x_0, ..., x_n\}$.

Ora, $varg = \sum_{i=1}^n varg_i$, nos termos da **Definição 3.8**. Portanto, em particular, como $varg_i \leq C_2\lambda_2$ (para todo o i), vem que $varg \leq nC_2\lambda_2 < \infty$.

Por outro lado, $\displaystyle\int \frac{1}{g} d\mu_0 \leq \sup|\varphi_0| \sum_{I_i \in \mathcal{P}^{(1)}} \int_{I_i} |\tau_i'| d\lambda \leq \sup|\varphi_0| \sum_{I_i \in \mathcal{P}^{(1)}} \lambda(\tau(I_i)) \leq n \sup|\varphi_0|$.

Seja $\mathcal{Q}^2_{n+N}$ o subconjunto dos intervalos de $\mathcal{P}^{(n+N)}$ tais que, para cada $j = 0, ..., n-1$, $\tau^j(\eta) \subset \zeta_j \in \mathcal{P}^{(n+N-j)}$ com ζ_j satisfazendo

$$\frac{\inf(g|\zeta_j)}{\sup(g|\zeta_j)} \geq 1 - \lambda_1^{(n+N-j)/4} \; (**).$$

Assim, se $\eta \in \mathcal{Q}^2_{n+N}$, tem-se

$$\frac{\inf_\eta |(\tau^n)'|}{\sup_\eta |(\tau^n)'|} \geq \prod_{j=0}^{n-1} \frac{\inf(g|\zeta_j)}{\sup(g|\zeta_j)} \geq \prod_{j=0}^{n-1} (1 - \lambda_1^{(n+N-j)/4}) \geq \prod_{k=N+1}^{\infty} (1 - \lambda_1^{k/4}) \geq e^{-\beta/2}$$

para N suficientemente grande.

Vejamos qual a medida total dos ζ's pertencentes a $\mathcal{P}^{(n+N-j)}$ que não verificam a propriedade $(**)$.

Tomemos, primeiro, os ζ's para os quais $\sup(g|\zeta) \leq \lambda_1^{(n+N-j)/4}$.

Ora, $\displaystyle\int \frac{1}{g} d\mu_0 \geq \lambda_1^{-(n+N-j)/4} \sum_\zeta \mu_0(\zeta)$.

Se, pelo contrário, $\sup(g|\zeta) \geq \lambda_1^{(n+N-j)/4}$, então

$$varg|\zeta \geq \sup(g|\zeta) \left(1 - \frac{\inf(g|\zeta)}{\sup(g|\zeta)}\right) \geq \lambda_1^{(n+N-j)/2}$$

de onde se conclui que não há mais do que $\lambda_1^{-(n+N-j)/2} varg$ ζ's em tais condições e, portanto,

$$\mu_0(\cup\zeta) \leq \lambda_1^{-(n+N-j)/2} varg \sup|\varphi_0| C_1 \lambda_1^{n+N-j} = C_1 \lambda_1^{(n+N-j)/2} varg \sup|\varphi_0|.$$

Logo, a medida total dos ζ's em $\mathcal{P}^{(n+N-j)}$ que não cumprem $(**)$ é majorada por

$$\lambda_1^{(n+N-j)/4} \int \frac{1}{g} d\mu_0 + C_1 \lambda_1^{(n+N-j)/2} varg \sup|\varphi_0| \leq C_1' \lambda_1^{(n+N-j)/2}$$

onde $C_1' = \int \frac{1}{g} d\mu_0 + C_1 varg \sup|\varphi_0|$.

Novamente, como μ_0 é invariante, então a medida total dos η's pertencentes a $\mathcal{P}^{(n+N)} \setminus \mathcal{Q}_{n+N}^2$ é majorada por

$$\sum_{j=0}^{n-1} C_1' \lambda_1^{(n+N-j)/2} \leq \sum_{k=N+1}^{\infty} C_1' \lambda_1^{k/2} \leq \frac{\beta}{4}$$

para N suficientemente grande.

Agora, notando que $var\varphi_0 < \infty$ e que $\int \frac{1}{\varphi_0} d\mu_0 = \int 1 d\lambda = 1$, todo o raciocínio que acabámos de fazer é válido para $g = \varphi_0$.

Logo, existe $\mathcal{Q}_{n+N}^3$ subconjunto de $\mathcal{P}^{(n+N)}$ tal que, para N suficientemente grande,

$$\frac{\inf(\varphi_0|\tau^j(\eta))}{\sup(\varphi_0|\tau^j(\eta))} \geq 1 - \lambda_1^{(n+N-j)/4} \geq e^{-\beta/4}$$

para cada $j = 0, ..., n-1$ e $\eta \in \mathcal{Q}_{n+N}^3$, e ainda os conjuntos em $\mathcal{P}^{(n+N)} \setminus \mathcal{Q}_{n+N}^3$ têm medida μ_0 total inferior a $\frac{\beta}{4}$.

Finalmente, $\mathcal{Q}_{n+N} = \mathcal{Q}_{n+N}^1 \cap \mathcal{Q}_{n+N}^2 \cap \mathcal{Q}_{n+N}^3$; claro que os elementos de $\mathcal{P}^{(n+N)} \setminus \mathcal{Q}_{n+N}$ têm medida μ_0 total inferior a β (por construção), e, de facto,

$$\frac{\mu_0(\tau^n(\xi))}{\mu_0(\tau^n(\eta))} \geq e^{-\beta} \frac{\mu_0(\xi)}{\mu_0(\eta)}$$

para todo o $\xi \subset \eta \in \mathcal{Q}_{n+N}$. □

Proposição 3.35. *Seja τ expansora por pedaços e topologicamente misturadora. Então, a única medida de probabilidade absolutamente contínua e invariante, μ_0, é exata.*

Demonstração. Seja I conforme a **Definição 3.19**; tomemos $Z \subset I$ tal que, para cada $j \geq 1$, existe $Z_j \in \mathcal{B}$ com $Z = \tau^{-j}(Z_j)$ (existe sempre Z em tais condições uma vez que, porque μ_0 está suportada em I, Z é mensurável, e τ – e todos os seus iterados – é expansora por pedaços, logo mensurável).

Vejamos, em primeiro lugar, que $\mu_0(Z) > 0 \implies \mu_0(Z) = 1$ é suficiente para μ_0 exata i.e. $\mu_0\left(\bigcap_{n=0}^{+\infty} \tau^{-n}(\mathcal{B})\right) > 0 \implies \mu_0\left(\bigcap_{n=0}^{+\infty} \tau^{-n}(\mathcal{B})\right) = 1$.

Ora, tal decorre diretamente das definições de $\bigcap_{n=0}^{+\infty} \tau^{-n}(\mathcal{B})$ e Z, pois de

$$A \in \bigcap_{n=0}^{+\infty} \tau^{-n}(\mathcal{B}) \iff A \in \tau^{-n}(\mathcal{B})\ \forall n \in \mathbb{N}$$

concluímos que $Z \in \bigcap_{n=0}^{+\infty} \tau^{-n}(\mathcal{B})$.

Assim, recordando que μ_0 tem I como suporte, basta provar que se $\mu_0(Z) > 0$ então $\mu_0(I \setminus Z) = 0$.

Notemos que, por ser topologicamente misturadora, então τ é sobrejetiva: de facto, supondo que $\tau : X \to Y$ é tal que $\tau(X) \subset Y$ (inclusão estrita) i.e. $\exists y \in Y : \tau(x) \neq y\ \forall x \in X$, então, porque existe $n \in \mathbb{N} : \tau^n(y) \in I$ e $\tau^n(y) = \tau(\tau^{n-1}(y))$ com $\tau^{n-1}(y) \in I$ (porque $\tau(I) = I$), obtém-se, por indução, que $y \in I$, logo $y = \tau(x)$ para algum $x \in I$ – contradição.

Assim, $Z_j = \tau^j(\tau^{-j}(Z_j)) = \tau^j(Z)$.

Pretendemos colocar-nos em condições para aplicar o **Lema 3.34**.

Fixemos $\beta \leq \mu_0(Z)/3$. Ora, pelo **Lema 3.34**, temos $N(= N(\beta))$.

Assumimos agora o seguinte:[10]

- $\exists Z_0 \subset Z$ com $\mu_0(Z_0) = \mu_0(Z) \geq 3\beta$
- dado $a \in Z_0$ qualquer, então J denota um intervalo fechado contendo a
- $\forall a \in Z_0$, $\lim_{r\to 0} \inf_{|J|<r} \dfrac{\mu_0(Z \cap J)}{\mu_0(J)} = 1$

i.e. é impossível encontrar um intervalo fechado (que pode ser tão pequeno quanto se queira) contendo qualquer ponto de Z_0 que não esteja "quase totalmente" contido em Z.

Assim, em particular, para qualquer $\varepsilon > 0$ existem $r > 0$ e $Z_\varepsilon \subset Z_0$ tais que $\mu_0(Z_\varepsilon) \geq 2\beta$

[10] Teorema da Densidade de *Lebesgue*

e

$$\inf_{|J|<r} \frac{\mu_0(Z \cap J)}{\mu_0(J)} \geq 1 - \varepsilon \ \ (*)$$

para todo o $a \in Z_\varepsilon$.

Tomemos $n \geq 0$ tão grande quanto necessário para que $|\eta| < r$ para todo o $\eta \in \mathcal{P}^{(n+N)}$, e seja $\mathcal{Q}_{n+N}$ conforme o **Lema 3.34**. Ora, como por hipótese $\mu_0(Z_\varepsilon) \geq 2\beta$, então necessariamente Z_ε interseta algum elemento de $\mathcal{Q}_{n+N}$, designemo-lo η_ε. Assim, por $(*)$,

$\dfrac{\mu_0(\eta_\epsilon \setminus Z)}{\mu_0(\eta_\epsilon)} \leq \varepsilon$, e, pelo **Lema 3.34** *2.*, vem $\dfrac{\mu_0(\tau^n(\eta_\epsilon) \setminus Z_n)}{\mu_0(\tau^n(\eta_\epsilon))} < e^\beta \varepsilon$.

Por outro lado, pelo **Lema 3.34** *1.*, sabemos que $\zeta_\varepsilon = \tau^n(\eta_\varepsilon) \in \mathcal{P}^{(N)}$.

Mas, porque τ é topologicamente misturadora, então existe $q \geq 1$ (que depende apenas de N) tal que $\tau^q(\zeta_\varepsilon) = I$.

Como há apenas um número finito de interseções entre ζ_ε e os elementos de $\mathcal{P}^{(q)}$ (porque, desde logo, $\mathcal{P}^{(q)}$ é composta por um número finito de elementos), sejam ζ_i, $i = 1, ..., k$ tais interseções. Ora, τ^q é monótona em cada um dos ζ_i's e de $\tau^q(\zeta_\varepsilon) = I$ resulta que $\bigcup_{i=1}^k \zeta_i \supseteq I$.

Afirmamos:

$$\forall \delta > 0 \ \exists \varepsilon_1 : \mu_0(A) \leq \varepsilon_1 \implies \mu_0(\tau^q(A)) \leq \delta$$

qualquer que seja $A \in \mathcal{B}$ tal que $A \subset \zeta_i$ para algum $i \in \{1, ..., k\}$ $-$ de facto, a afirmação é válida para a medida de *Lebesgue*, porque $\tau^q|_{\zeta_i}$ é $\mathcal{C}^\infty$, e tem-se que μ_0 e λ são equivalentes em I.

Escolhemos ε acima tal que $e^\beta \varepsilon \leq \varepsilon_1$.

Ora,

$$\mu_0(\zeta_i \setminus Z_n) \leq \mu_0(\zeta_\varepsilon \setminus Z_n) \leq e^\beta \varepsilon \mu_0(\zeta_\varepsilon) \leq \varepsilon_1$$

pelo que

$$\mu_0(\tau^q(\zeta_i) \setminus Z_{n+q}) = \mu_0(\tau^q(\zeta_i \setminus Z_n)) \leq \delta$$

de onde, uma vez que μ_0 é invariante, resulta

$$\mu_0(I \setminus Z) = \mu_0(I \setminus Z_{n+q}) \leq \sum_{i=1}^k \mu_0(\tau^q(\zeta_i) \setminus Z_{n+q}) \leq k\delta.$$

Mas δ é arbitrário e $k \leq \#\mathcal{P}^{(q)}$ (que não depende de δ), logo $\mu_0(I \setminus Z) = 0$. $\square$

Corolário 3.36. *$spec(P_\tau) = \{1\} \cup \Sigma_0$, com 1 valor próprio simples.*

Demonstração. Suponhamos que existe $\varphi_1 \neq 0$ tal que $P_\tau(\varphi_1) = \sigma_1\varphi_1$ com $|\sigma_1| = 1$.

Como φ_0 tem I como suporte (**Corolário 3.20**), então φ_1/φ_0 está definida *qtp* em I e

$$\int_I |\varphi_1/\varphi_0| d\mu_0 = \int_I |\varphi_1| d\lambda \leq \sup|\varphi_1| < +\infty$$

ou seja, $\varphi_1/\varphi_0 \in \mathcal{L}^1$. Dada $\varphi \in \mathcal{L}^1$, escrevemos

$$\int \varphi P_\tau^n(\varphi_1) d\lambda = \int (\varphi \circ \tau^n)\varphi_1 d\lambda = \int (\varphi \circ \tau^n)(\varphi_1/\varphi_0) d\mu_0.$$

A conclusão resulta agora de se ter que μ_0 é *mixing*, pela **Proposição 3.35** e pelo **Lema 3.33**, pois, sendo assim,

$$\int (\varphi \circ \tau^n)(\varphi_1/\varphi_0) d\mu_0 \to \int \varphi d\mu_0 \int (\varphi_1/\varphi_0) d\mu_0$$

e, como $\int \varphi d\mu_0 \int (\varphi_1/\varphi_0) d\mu_0 = \int \varphi(\varphi_0 \int \varphi_1 d\lambda) d\lambda$, então $\sigma_1^n \varphi_1 = P_\tau^n(\varphi_1)$ converge (*qtp*) para $\varphi_0 \int \varphi_1 d\lambda$. Mas tal só é possível caso $\sigma_1 = 1$ e, consequentemente, $\varphi_1 = \varphi_0 \int \varphi_1 d\lambda$.

Ou seja, $P_\tau : BV \to BV$ não admite valores próprios de módulo 1 para além do próprio 1, e qualquer vetor próprio associado a 1 é um múltiplo de φ_0. Além disso, 1 tem multiplicidade algébrica 1, já que, se assim não fosse, existiria $\psi_0 \in BV$ tal que $P_\tau^n(\psi_0) = \varphi_0 + n\psi_0$ e, então, $\lim_{n\to+\infty}\|P_\tau^n(\psi_0)\| = +\infty$ – contradição com o facto de que a mesma norma é uniformemente limitada. □

Corolário 3.37. *BV admite a decomposição espectral $BV = \mathbb{R}\varphi_0 \oplus X_0$, onde $X_0 = \{\varphi : \int \varphi d\lambda = 0\}$.*

Demonstração. Basta notar que ambos os subespaços $\mathbb{R}\varphi_0$ e X_0 são invariantes para a ação de P_τ e, como tal, os vetores próprios associados aos valores próprios de Σ_0, que são todos os valores próprios de $P_\tau|X_0$, têm que pertencer a X_0. Mas então, se $\psi \in X_0$, $\lim_{n\to+\infty} \int P_\tau^n(\psi) d\lambda = \lim_{n\to+\infty} \int \sigma^n \psi d\lambda = 0$. Como P_τ preserva integrais (**Proposição 2.5**), então $\int \psi d\lambda = 0$. □

Voltamos agora a deduzir o decaimento exponencial de correlações (**Corolário 3.29**).

Queremos provar que existe $C(\varphi_1, \varphi_2) > 0$ tal que

$$\left| \int (\varphi_1 \circ \tau^n)\varphi_2 d\mu_0 - \int \varphi_1 d\mu_0 \int \varphi_2 d\mu_0 \right| \leq C(\varphi_1, \varphi_2)\Lambda^n$$

(recordemos que $\Lambda < 1$ diz respeito ao raio de um disco onde está esteja contido Σ_0, conforme especificado no início da discussão deste caso).

Ora, $\int (\varphi_1 \circ \tau^n)\varphi_2 d\mu_0 - \int \varphi_1 d\mu_0 \int \varphi_2 d\mu_0 = \int (\varphi_1 \circ \tau^n)\varphi_2\varphi_0 d\lambda - \int \varphi_1\varphi_0 d\lambda \int \varphi_2\varphi_0 d\lambda$

$$= \int \varphi_1 \left(P_\tau^n(\varphi_2\varphi_0) - \varphi_0 \int \varphi_2\varphi_0 d\lambda \right) d\lambda = \int \varphi_1 P_\tau^n(\varphi_2\varphi_0 - \varphi_0 \int \varphi_2\varphi_0 d\lambda) d\lambda$$

pois, pelas Proposições **2.3** e **2.13**, $\varphi_0 \int \varphi_2\varphi_0 d\lambda = P_\tau^n(\varphi_0 \int \varphi_2\varphi_0 d\lambda)$.

Mas, denotando por π_0 a projeção em X_0, i.e. $\pi_0(\psi) = \psi - \varphi_0 \int \psi d\lambda$, então vem

$$\int (\varphi_1 \circ \tau^n)\varphi_2 d\mu_0 - \int \varphi_1 d\mu_0 \int \varphi_2 d\mu_0 = \int \varphi_1 P_\tau^n(\pi_0(\varphi_2\varphi_0)) d\lambda.$$

Logo, $\left| \int (\varphi_1 \circ \tau^n)\varphi_2 d\mu_0 - \int \varphi_1 d\mu_0 \int \varphi_2 d\mu_0 \right| = \left| \int \varphi_1 P_\tau^n(\pi_0(\varphi_2\varphi_0)) d\lambda \right|$

$$\leq \int |\varphi_1 P_\tau^n(\pi_0(\varphi_2\varphi_0))| d\lambda \leq \sup |P_\tau^n(\pi_0(\varphi_2\varphi_0))| \int |\varphi_1| d\lambda \leq \|P_\tau^n(\pi_0(\varphi_2\varphi_0))\|_{BV} \int |\varphi_1| d\lambda$$

$\leq C'\Lambda^n \|\pi_0(\varphi_2\varphi_0)\|_{BV} \int |\varphi_1| d\lambda \leq C''\Lambda^n \|\varphi_2\varphi_0\|_{BV} \int |\varphi_1| d\lambda$ para alguns $C', C'' > 0$.

Portanto, basta que $C(\varphi_1, \varphi_2) = C''\|\varphi_2\varphi_0\|_{BV} \int |\varphi_1| d\lambda$.

Para o Teorema do Limite Central, se φ tem variação limitada, $\phi = \varphi - \int \varphi d\mu_0$, e $\sigma^2 = \int \phi^2 d\mu_0 + 2 \sum_{j=1}^{\infty} \int \phi(\phi \circ \tau^j) d\mu_0 \in (0, +\infty)$, então

$$\lim_{n \to +\infty} \mu_0\{x : \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} \phi(\tau^j(x)) \in A\} = \int_A \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{t^2}{2\sigma^2}} dt.$$

Capítulo 4

Generalização: um acréscimo numerável à quantidade de descontinuidades

O objetivo deste capítulo é chegar a uma decomposição espectral de P_τ nos termos do que foi apresentado na parte final da secção anterior (**Corolário 3.36** e **Corolário 3.37**), para τ expansora por pedaços mas admitindo agora uma quantidade numerável de descontinuidades. Assim, o resultado central é o **Teorema 4.10**, de [Ryc83], cuja demonstração completa não será exibida ainda que sejam discutidas as principais ideias por detrás da mesma.

Numa fase inicial, iremos desviar-nos um pouco da abordagem de [Ryc83], por forma a estabelecermos o paralelo com os conceitos introduzidos no Capítulo 3 e, consequentemente, podermos aproveitar resultados já obtidos.

Comecemos, portanto, pelas noções de partição e aplicação expansora por pedaços a adotar neste contexto.

Relembremos que temos como referência o espaço de probabilidade $(\Omega, \mathcal{B}, \lambda)$ onde $\Omega = [a, b] \subset \mathbb{R}$, $\mathcal{B}$ a $\sigma-$álgebra de *Borel* e λ a medida de *Lebesgue* normalizada definida em $[a, b]$.

Definição 4.1. Seja $\mathcal{S} = \{I_i\}_{i\in\mathbb{N}}$ tal que, para todo o i, I_i é um subintervalo de Ω fechado à esquerda, e $\overline{\bigcup_{i\in\mathbb{N}} I_i} = \Omega$. Se $I_i \cap I_j = \emptyset \ \forall i \neq j$, então $\mathcal{S}$ diz-se uma *partição* de Ω.

Novamente, não há perda de generalidade em assumir que $\tau : \Omega \to \Omega$ é contínua à direita (o que implica que existe alguma partição $\mathcal{S}$ tal que todas as restrições $\tau|_{I_i}$ são contínuas).

Obtemos $\mathcal{S}^{(m)}$ a partir de $\mathcal{S}$ exatamente da mesma forma que obtivemos $\mathcal{P}^{(m)}(x_0, ..., x_n)$ a partir de $\mathcal{P}(x_0, ..., x_n)$, de acordo com a **Definição 3.7**.

E também, $g_\eta^{(m)} \equiv \dfrac{1}{|(\tau^m|\eta)'|}$, para $\eta \in \mathcal{S}^{(m)}$.

Definição 4.2. $\tau : \Omega \to \Omega$ diz-se *expansora por pedaços* se existe $\mathcal{S} = \{I_i\}_{i\in\mathbb{N}}$ tal que

(i) para todo o $i \in \mathbb{N}$, $\tau_i = \tau|int(I_i)$ é de classe C^1, $|\tau_i'(x)| > 0$ e $g_i(x) \equiv \dfrac{1}{|\tau_i'(x)|}$ tem variação limitada;

(ii) existem $C_1 > 0$ e $\lambda_1 < 1$ tais que $\sup g_\eta^{(m)} \leq C_1\lambda_1^m$ para todo o $m \geq 1$.

Observação 4.3. Se $x \notin \bigcup_{i\in\mathbb{N}} int(I_i)$, então x pode ser um dos pontos de descontinuidade de τ, onde τ' não está definida. Assim,

$$g(x) \equiv \begin{cases} 0 & \text{se } x \text{ é ponto de descontinuidade de } \tau \\ \dfrac{1}{|\tau'(x)|} & \text{caso contrário} \end{cases}.$$

Da mesma forma,

$$g^{(m)}(x) \equiv \begin{cases} 0 & \text{se } x \text{ é ponto de descontinuidade de } \tau^m \\ \dfrac{1}{|(\tau^m)'(x)|} & \text{caso contrário} \end{cases}.$$

Observação 4.4. Tal como no Capítulo 3, são agora consequência da **Definição 4.2**:

- τ^m é expansora por pedaços em $\mathcal{S}^{(m)}$
- existem $C_2 > 0$ e $\lambda_2 < 1$ tais que $var g_\eta^{(m)} \leq C_2\lambda_2^m$ para todo o $m \geq 1$

Grande parte do trabalho apresentado na primeira secção de [Ryc83] tem como propósito a dedução da desigualdade de Lasota-Yorke, como enunciada na **Proposição 3.13**.

Também de acordo com a nossa abordagem, apesar da garantia de que todas as $g_\eta^{(m)}$ têm variação limitada, que é facto essencial para a obtenção do majorante para a variação de

P_τ fornecido pela mesma proposição, se a partição onde as $g_\eta^{(m)}$ estão definidas não for finita então a aplicação da propriedade *1.* da variação (**Proposição 3.3**) não é válida e, consequentemente, a desigualdade de Lasota-Yorke não resulta de forma imediata.

A chave estará em considerarmos, em vez de $\mathcal{S}$, uma partição finita. Ora, apesar de existir uma infinidade de descontinuidades de τ em algum dos elementos de tal partição, pelo lema que se segue não há prejuízo para a propriedade de variação limitada das funções $g_\eta^{(m)}$. Assim, a desigualdade de Lasota-Yorke para $\mathcal{S}$ obtém-se à custa da sua 'aproximação' por essa mesma partição finita.

Lema 4.5. *Para todo o* $\varepsilon > 0$ *existe uma partição finita,* $\mathcal{A}$*, tal que*

$$\max_{\alpha \in \mathcal{A}} var_\alpha g \leq \sup g + \varepsilon.$$

Demonstração. Notando que g é não negativa e limitada, então os seus saltos não podem exceder $\sup g$. Assim, para cada $x \in \Omega$, existe um intervalo aberto U_x onde $var_{U_x} g \leq \sup g + \varepsilon$. Ora, $\{U_x\}_x$ é uma cobertura (aberta) de Ω, que é compacto, logo admite uma subcobertura finita. $\square$

Observação 4.6. Resulta do **Lema 4.5** que existe uma escolha de $\mathcal{S}$ tal que $\max_{\alpha \in \mathcal{A}} var_\alpha g \leq C_1\lambda_1 + \varepsilon$: se $\mathcal{S}$ da **Definição 4.2** for tal que as descontinuidades de τ são exatamente os extremos dos intervalos I_i, então, de acordo com a mesma definição, $\sup g_i \leq C_1\lambda_1$ e ficam apenas a faltar os pontos de descontinuidade de τ onde, por imposição, $g = 0$.

Observação 4.7. O **Lema 4.5** é, de facto, importante pois com algum dos elementos de $\mathcal{A}$ contendo, necessariamente, uma infinidade de descontinuidades de τ, não seria de todo garantido que nesse conjunto em particular g tivesse variação limitada.

Notando que $\max_{I \in \mathcal{S}} var_I g \leq \max_{\alpha \in \mathcal{A}} var_\alpha g$, então temos a desigualdade de Lasota-Yorke para $\mathcal{S}$.

Observação 4.8. Tal como no Capítulo 3, são agora consequência da **Proposição 3.13**:

- τ admite pelo menos uma medida de probabilidade, μ_0, absolutamente contínua e invariante, à qual está associada uma densidade, f, com variação limitada (**Corolário 3.17**)

- para qualquer m, $\|P_\tau^m\|_{BV} \leq \max\{C_0\lambda_0, C_0+1\}$, o que implica, em particular, que para qualquer m, $\|P_\tau^m\|_{BV} \leq 2C_0+1$

Para alicerçar o **Teorema 4.10** fica apenas a faltar a proposição que se segue.

Proposição 4.9. *Se τ, admitindo uma quantidade numerável de descontinuidades, é expansora por pedaços, então para algum $m \geq 1$ existe um operador de dimensão finita em BV, K, tal que $\|P_\tau^m - K\|_{BV} < 1$.*

Demonstração. Tomamos $\mathcal{A}$ nas condições do **Lema 4.5** e m tal que a desigualdade de Lasota-Yorke é válida com $C_0\lambda_0^m < 1/4$ e também $C_1\lambda_1^m < 1/4$ (nos termos da **Definição 4.2**).

Seja $E : BV \to BV$ tal que $E(f) = \mathbb{E}_\lambda(f|\mathcal{A}) = \sum_{\alpha\in\mathcal{A}} \int_\alpha f d\lambda$.

Vamos ver que $K = P_\tau^m E$.

Ora, porque há apenas um número finito de α's em $\mathcal{A}$, então E tem dimensão finita e, consequentemente, K tem dimensão finita; logo, resta verificar que K satisfaz $\|P_\tau^m - K\|_{BV} < 1$.

Vamos provar que, para qualquer $f \in BV$, $\|(P_\tau^m - K)(f)\|_{BV} < \|f\|_{BV}$.

Se $f \in BV$, seja $h = f - E(f) = (I-E)(f)$.

Como $\|(P_\tau^m - K)(f)\|_{BV} = \|P_\tau^m(I-E)(f)\|_{BV} = \|P_\tau^m(h)\|_{BV}$, então trata-se de provar que $\|P_\tau^m(h)\|_{BV} < \|f\|_{BV}$.

Ora, para qualquer $\alpha \in \mathcal{A}$,

$$\int_\alpha h d\lambda = \int_\alpha (f - E(f))d\lambda = \int_\alpha f d\lambda - \int_\alpha E(f)d\lambda = 0.$$

Logo, pela **Proposição 3.13**, $varP_\tau^m(h) \leq C_0\lambda_0^m varh < \frac{1}{4}varh$ e, como $h = f - E(f)$,

$varh \leq varf + varE(f) \leq 2varf$, portanto $varP_\tau^m(h) < \frac{1}{4}2varf = \frac{1}{2}varf$.

Por outro lado, $\|P_\tau^m(h)\|_1 \leq \|h\|_1$ (**Proposição 2.6**); mas $\|h\|_1 = \sum_{\beta\in\mathcal{A}^m} \int_\beta |h| d\lambda$.

Fixemos $\beta \in \mathcal{A}^m$. De facto,

$$\int_\beta |h| d\lambda \leq \sup_\beta |h| \lambda(\beta) = \sup_\beta |f - E(f)| \lambda(\beta) \leq \sup_\beta |f| \lambda(\beta) + \sup_\beta |E(f)| \lambda(\beta)$$
$$\leq var_\beta f \lambda(\beta) + var_\beta E(f) \lambda(\beta) \leq 2 var_\beta f \lambda(\beta).$$

Assim, $\displaystyle\sum_{\beta \in \mathcal{A}^m} \int_\beta |h| d\lambda \leq 2 \max_{\beta \in \mathcal{A}^m} \lambda(\beta) \sum_{\beta \in \mathcal{A}^m} var_\beta f \leq 2 \max_{\beta \in \mathcal{A}^m} \lambda(\beta) var f$; mas $\max_{\beta \in \mathcal{A}^m} \lambda(\beta) \leq \sup_\beta g^{(m)} \leq C_1 \lambda_1^m < 1/4$, portanto $\|h\|_1 < 2\frac{1}{4} var f = \frac{1}{2} var f$.

Tem-se $\|P_\tau^m(h)\|_{BV} < var f$ e, consequentemente, $\|P_\tau^m(h)\|_{BV} < \|f\|_{BV}$. □

Teorema 4.10. *Suponhamos que existem $m \geq 1$ e K operador de dimensão finita em BV tais que $\|P_\tau^m - K\|_{BV} < 1$. Então, $spec(P_\tau) = \{\sigma_1, ..., \sigma_k\} \cup \Sigma_0$ onde*

1. *$\sigma_1, ..., \sigma_k$ têm módulo 1, são pólos simples, e, para $i = 1, ..., k$, E_{σ_i} tem dimensão finita*
2. *Σ_0 está contido num disco de raio $r \in (0, 1)$*
3. *P_τ admite a decomposição*
$$P_\tau = \bigoplus_{i=1}^k \sigma_i Q_i \oplus R$$
com Q_i o projetor correspondente a σ_i (i.e. a restrição de P_τ ao subespaço E_{σ_i}) e $R : BV \to BV$ tal que $\inf_m \|R^m\|^{1/m} < r$; todos os Q_i's comutam entre si e com R e a decomposição é ortogonal ($\forall i \neq j\ Q_i Q_j = 0$ e $\forall i\ Q_i R = 0$)

Observação 4.11. Resulta que por iteração do sistema a parte correspondente aos valores próprios de módulo inferior a 1 deixa de ter relevância, em rigor se $f \in BV$ então $\lim_{m \to +\infty} R^m(f) = 0$ pelo que, em particular, a menos de iteração, P_τ pode ser arbitrariamente aproximado por um número finito de operadores de dimensão finita (os Q_i's). Como a decomposição é ortogonal, vale ainda que se $m = mmc\{dim(Q_1), ..., dim(Q_k)\}$ então $spec(P_\tau^m) = \{1\} \cup \Sigma_0$, portanto, para τ nas condições do **Teorema 4.10** há um número finito de medidas invariantes.

O facto de se ter $(\|P_\tau^m\|_{BV})_m$ uniformemente limitada (**Observação 4.8**) implica, em particular, que $P_\tau : BV \to BV$ é um operador (linear) limitado que, portanto, não pode ter valores próprios de módulo superior a 1.

Notação: Z denota o disco de raio 1 no plano complexo i.e. $\{z \in \mathbb{C} : |z| \leq 1\}$.

Assim, o **Teorema 4.10** é uma versão do **Teorema 4.12**, válido para qualquer operador linear positivo definido num espaço de medida $\sigma-$finito, cujo espectro esteja contido no disco unitário e que possa ser aproximado por um operador compacto.

Teorema 4.12. *Seja (S, Σ, μ) um espaço de medida, com μ uma medida $\sigma-$finita. Suponhamos que T é um operador linear positivo em $\mathcal{L}^1(S, \Sigma, \mu)$ tal que T^n/n converge para 0 na topologia fraca dos operadores e seja K um operador compacto tal que $|T^N - K| < 1$ para algum $N \in \mathbb{N}$. Então,*

1. *$spec(T)$ pode ser decomposto na união de um conjunto fechado, Σ_0, contido num disco de raio r estritamente menor que 1, com um número finito de pólos simples $e^{2\pi i\theta_k}$, onde para $k = 1, ..., q$, θ_k é racional*

2. *se $E_k = E(e^{2\pi i\theta_k})$ e $E_D = E(\Sigma_0)$, com $D = T(E_D)$, então cada E_k tem dimensão finita e, para $m \geq 1$,*
$$T^m = \bigoplus_{k=1}^{q} e^{2\pi i\theta_k m} E_k \oplus D^m$$

3. *existe $M > 0$ tal que, para todo o $m \geq 1$, $|D^m| \leq Mr^m$*

Observação 4.13. Podemos substituir a hipótese de T^n/n convergir para 0 na topologia fraca dos operadores pela hipótese de $spec(T) \subseteq Z$, que é mais forte: se todos os valores próprios de T têm módulo menor ou igual que 1 então, dada qualquer função $f \in \mathcal{L}^1(S, \Sigma, \mu)$, a sucessão $\left\| \frac{T^n(f)}{n} \right\|$ é convergente para 0.

Por fim, vejamos que é conhecida uma descrição completa do caso em que 1 é o único valor próprio de P_τ com módulo 1.

Relembremos que este caso já surgiu no Capítulo 3: se τ, com um número finito de descontinuidades, for expansora por pedaços e topologicamente misturadora, então $spec(P_\tau) = \{1\} \cup \Sigma_0$ onde $\Sigma_0 \subset int(Z)$.

Portanto, em particular, as conclusões que iremos retirar são válidas para aplicações expansoras por pedaços e topologicamente misturadoras com um número finito de descontinuidades.

De acordo com o **Teorema 4.10**,

$$P_\tau = Q \oplus R$$

(com Q o projetor correspondente ao valor próprio 1).

Teorema 4.14. *Existem funções não negativas* $\varphi_1, ..., \varphi_s \in BV$ *e* $\psi_1, ..., \psi_s \in \mathcal{L}^\infty$ *tais que:*

1. *dada* $f \in \mathcal{L}^1$, $Q(f) = \sum_{i=1}^{s} \varphi_i \int f \psi_i d\lambda$
2. *para* $i = 1, ..., s$, φ_i *é ponto fixo de* P_τ *e* ψ_i *é ponto fixo de* U_τ *i.e.* $P_\tau(\varphi_i) = \varphi_i$ *e* $U_\tau(\psi_i) = \psi_i$
3. $\int \varphi_i \psi_j d\lambda = \delta_{ij}$, *se* $i \neq j$ $\min(\varphi_i, \varphi_j) = 0 = \min(\psi_i, \psi_j)$ *e,*
   para $i = 1, ..., s$, $\int \varphi_i d\lambda = 1$
4. *existem conjuntos mensuráveis* $C_1, ..., C_s \subset \Omega$ *tais que, para* $i = 1, ..., s$, $\psi_i = 1\!\!1_{C_i}$ *(qtp) e* $\Omega = \bigcup_{i=1}^{s} C_i$ *(qtp)*
5. $\bigcap_{n=0}^{\infty} U_\tau^n(\mathcal{L}^1) = \bigcap_{n=0}^{\infty} U_\tau^n(\mathcal{L}^\infty) = < \psi_1, ..., \psi_s >$ *(espaço gerado por* $\psi_1, ..., \psi_s$*)*
6. *dada* $f \in \mathcal{L}^1$, $\lim_{n\to+\infty} U_\tau^n(f) = Q^*(f)$ *na topologia* $\sigma(\mathcal{L}^1, BV)$*; dada* $f \in \mathcal{L}^\infty$, $\lim_{n\to+\infty} U_\tau^n(f) = Q^*(f)$ *na topologia* $\sigma(\mathcal{L}^\infty, \mathcal{L}^1)$ *e*

$$Q^*(f) = \sum_{i=1}^{s} \psi_i \int f \varphi_i d\lambda$$

Observação 4.15. Algumas considerações a respeito do **Teorema 4.14**:

(i) revisitando o **Teorema 4.10**, sabemos que 1 é um pólo simples e que Q é a restrição de P_τ a E_1 cuja dimensão é finita; ora, assim sendo, claro que E_1 é gerado por um conjunto finito de funções BV, $\varphi_1, ..., \varphi_s$, que são pontos fixos de P_τ, ou seja, densidades de probabilidade com respetivas medidas de probabilidade invariantes

(ii) o facto de $\varphi_1, ..., \varphi_s$ gerarem E_1 justifica a representação de Q descrita em *1.*: a ação de Q em qualquer função de $\mathcal{L}^1$ fica completamente descrita através das projeções nos φ_i's

(iii) pela informação acrescentada por *3.* e *4.*, deduzimos que os ψ_i's correspondem a indicatrizes dos conjuntos que definem uma partição de Ω e que cada φ_i está suportada no conjunto C_i; logo, $\varphi_1, ..., \varphi_s$ constituem uma base ortogonal de E_1 e, para cada i, $\int f\psi_i d\lambda$ é a projeção ortogonal em φ_i

(iv) porque os ψ_i's são pontos fixos de U_τ, então os conjuntos C_i são invariantes, e, assim, $\varphi_1 + ... + \varphi_s$ é uma densidade invariante

A demonstração do ponto *1.* requer conhecimentos sobre reticulados de *Banach.* Vamos assumir o resultado de *1.* e três factos que advêm da sua demonstração, e fazer a prova dos pontos *2.* a *6.* do **Teorema 4.14**.

Demonstração.

2. O facto de que, para $i = 1, ..., s$, $P_\tau(\varphi_i) = \varphi_i$ obtém-se da demonstração de *1.*; portanto, fica apenas a faltar a verificação de que, para $i = 1, ..., s$, $U_\tau(\psi_i) = \psi_i$.

Comecemos por ver que $P_\tau(Q(f)) = Q(f) = Q(P_\tau(f))$.

Ora, se $P_\tau = Q \oplus R$, então

$$\begin{aligned} P_\tau(Q(f)) =&(Q \oplus R)(Q(f)) = Q(Q(f)) \oplus R(Q(f)) = Q(Q(f)) \oplus 0 = Q^2(f) = Q(f) \\ &= Q^2(f) = Q(Q(f)) \oplus 0 = Q(Q(f)) \oplus Q(R(f)) = Q((Q \oplus R)(f)) = Q(P_\tau(f)). \end{aligned}$$

Assim,

$$\begin{aligned} P_\tau(Q(f)) =&P_\tau\left(\sum_{i=1}^{s} \varphi_i \int f\psi_i d\lambda\right) = \sum_{i=1}^{s} P_\tau\left(\varphi_i \int f\psi_i d\lambda\right) = \sum_{i=1}^{s} P_\tau(\varphi_i) \int f\psi_i d\lambda \\ &= \sum_{i=1}^{s} \varphi_i \int f\psi_i d\lambda = Q(f) = Q(P_\tau(f)) = \sum_{i=1}^{s} \varphi_i \int P_\tau(f)\psi_i d\lambda \\ &= \sum_{i=1}^{s} \varphi_i \int fU_\tau(\psi_i) d\lambda \quad (\textbf{Proposição 2.11}) \end{aligned}$$

pelo que, em particular, $\int f\psi_i d\lambda = \int fU_\tau(\psi_i) d\lambda \implies U_\tau(\psi_i) = \psi_i$.

3. Também da demonstração de *1.* resulta que se $i \neq j$ $\min(\varphi_i, \varphi_j) = 0$ e, para $i = 1, ..., s$, $\int \varphi_i d\lambda = 1$. Limitar-nos-emos a provar que $\int \varphi_i\psi_j d\lambda = \delta_{ij}$ uma vez que depois de provarmos *6.* obteremos que se $i \neq j$ $\min(\psi_i, \psi_j) = 0$.

Mas basta notar que

$$\begin{aligned}\sum_{i=1}^{s}\varphi_i\int f\psi_i d\lambda =&Q(f)=Q^2(f)=Q\left(\sum_{i=1}^{s}\varphi_i\int f\psi_i d\lambda\right)\\ &=\sum_{j=1}^{s}\varphi_j\int\left(\sum_{i=1}^{s}\varphi_i\int f\psi_i d\lambda\right)\psi_j d\lambda=\sum_{i,j=1}^{s}\varphi_j\int\varphi_i\psi_j d\lambda\int f\psi_i d\lambda\end{aligned}$$

de onde, em particular, resulta que $\int \varphi_i\psi_j d\lambda=\delta_{ij}$.

6. Vejamos, em primeiro lugar, que $Q^*(f)=\sum_{i=1}^{s}\psi_i\int f\varphi_i d\lambda$.

Ora, admitindo a fórmula apresentada, tal corresponde a verificar que se tem a igualdade $\int Q(f)hd\lambda=\int fQ^*(h)d\lambda$ (sempre que h é tal que os integrais são convergentes).

De facto,

$$\begin{aligned}\int Q(f)hd\lambda &=\int\left(\sum_{i=1}^{s}\varphi_i\int f\psi_i d\lambda\right)hd\lambda=\sum_{i=1}^{s}\int h\varphi_i d\lambda\int f\psi_i d\lambda\\ &=\int f\left(\sum_{i=1}^{s}\psi_i\int h\varphi_i d\lambda\right)d\lambda=\int fQ^*(h)d\lambda.\end{aligned}$$

Agora, dada $f\in\mathcal{L}^1$, $\lim_{n\to+\infty}\int hU_\tau^n(f)d\lambda=\lim\limits_{n\to+\infty}\int P_\tau^n(h)fd\lambda=\int Q(h)fd\lambda=\int hQ^*(f)d\lambda$ para qualquer $h\in BV$; portanto, $\lim_{n\to+\infty}U_\tau^n(f)=Q^*(f)$ na topologia $\sigma(\mathcal{L}^1,BV)$. Exatamente a mesma justificação é valida para $f\in\mathcal{L}^\infty$ e, portanto, $\lim_{n\to+\infty}U_\tau^n(f)=Q^*(f)$ na topologia $\sigma(\mathcal{L}^\infty,\mathcal{L}^1)$.

Pelo que acabámos de mostrar, para $i=1,...,s$,

$$\lim_{n\to+\infty}U_\tau^n(\varphi_i)=Q^*(\varphi_i)=\sum_{j=1}^{s}\psi_j\int\varphi_i\varphi_j d\lambda=\psi_i\int\varphi_i^2 d\lambda\ \text{ (na topologia } \sigma(\mathcal{L}^1,BV));$$

por outro lado, para $i\neq j$, $\min(\psi_i,U_\tau^n(\varphi_j))=\min(U_\tau^n(\psi_i),U_\tau^n(\varphi_j))=\min(\psi_i,\varphi_j)\circ\tau^n=0$ para todo o n.

Logo,

$$0=\lim_{n\to+\infty}\int\psi_i U_\tau^n(\varphi_j)d\lambda=\int\psi_i\left(\psi_j\int\varphi_j^2 d\lambda\right)d\lambda=\int\psi_i\psi_j d\lambda\int\varphi_j^2 d\lambda$$

e, como $\int\varphi_j^2 d\lambda\neq 0$, então $\int\psi_i\psi_j d\lambda=0\implies\min(\psi_i,\psi_j)=0$, o que completa a prova de *3.*.

4. Basta notar que $U_\tau(1) = 1 \implies Q^*(1) = 1 \iff \sum_{i=1}^{s} \psi_i = 1$; o pretendido resulta acrescentando as hipóteses de $\psi_1, ..., \psi_s \in \mathcal{L}^\infty$ não negativas e tais que $\min(\psi_i, \psi_j) = 0$ se $i \neq j$.

5. Vamos provar que $\bigcap_{n=0}^{\infty} U_\tau^n(\mathcal{L}^\infty) =< \psi_1, ..., \psi_s >$; de facto, por um argumento de aproximação, $\bigcap_{n=0}^{\infty} U_\tau^n(\mathcal{L}^1) = \bigcap_{n=0}^{\infty} U_\tau^n(\mathcal{L}^\infty)$.

Veremos que se $h \in \bigcap_{n=0}^{\infty} U_\tau^n(\mathcal{L}^\infty)$ então é igual (*qtp*) a $h_0 \in< \psi_1, ..., \psi_s >$.

Tomemos uma sucessão, $(h_n)_n$, de funções de $\mathcal{L}^\infty$ tal que, para cada n, $h = U_\tau^n(h_n)$ e $\|h_n\|_\infty \leq \|h\|_\infty$.

Ora, $(Q^*(h_n))_n$ é uma sucessão limitada em $< \psi_1, ..., \psi_s >$, portanto existe $(Q^*(h_{n_k}))_{n_k}$, subsucessão de $(Q^*(h_n))_n$, convergente para h_0.

Agora, dada $f \in \mathcal{L}^1$,

$$\int fh d\lambda = \int f U_\tau^{n_k}(h_{n_k}) d\lambda = \int P_\tau^{n_k}(f) h_{n_k} d\lambda = \int (P_\tau^{n_k} - Q)(f) h_{n_k} d\lambda + \int Q(f) h_{n_k} d\lambda$$

que, como $\int (P_\tau^{n_k} - Q)(f) h_{n_k} d\lambda \to 0$, então converge para $\int Q(f) h_{n_k} d\lambda = \int f Q^*(h_{n_k}) d\lambda = \int f h_0 d\lambda$ i.e. $h = h_0$ (*qtp*). □

Corolário 4.16. *Seja $\mathcal{B}_\infty = \bigcap_{n=0}^{\infty} \tau^{-n}(\mathcal{B})$. Então, $\mathcal{B}_\infty$ é gerado pela decomposição $(C_1, ..., C_s)$ e o subespaço próprio associado ao valor próprio 1, E_1, é precisamente $< \psi_1, ..., \psi_s >$.*

Observação 4.17. Os sistemas dinâmicos (τ_i, ν_i), onde $\tau_i = \tau|_{C_i}$ e $\nu_i = \varphi_i d\lambda$, são exatos e a única medida invariante (para τ_i) e absolutamente contínua com respeito a $\lambda|_{C_i}$ é ν_i.

Assim, vem que $\varphi_1, ..., \varphi_s$ são as únicas densidades invariantes em $C_1, ..., C_s$, respetivamente, e, como tal, $\varphi = \varphi_1 + ... + \varphi_s$ é a única densidade invariante em Ω.

Bibliografia

[BG97] Abraham Boyarsky and Pawel Gora. *Laws of chaos: invariant measures and dynamical systems in one dimension.* Springer Science & Business Media, 1997.

[BS02] Michael Brin and Garrett Stuck. *Introduction to dynamical systems.* Cambridge university press, 2002.

[Dev89] Robert Devaney. *An introduction to chaotic dynamical systems.* Addison-Wesley, 1989.

[DS57] Nelson Dunford and Jacob T Schwartz. *Linear Operators. Part 1: General Theory.* New York Interscience, 1957.

[Ryc83] Marek Rychlik. Bounded variation and invariant measures. *Studia mathematica*, 76(1):69–80, 1983.

[Via97] Marcelo Viana. *Stochastic dynamics of deterministic systems*, volume 21. IMPA Rio de Janeiro, 1997.