



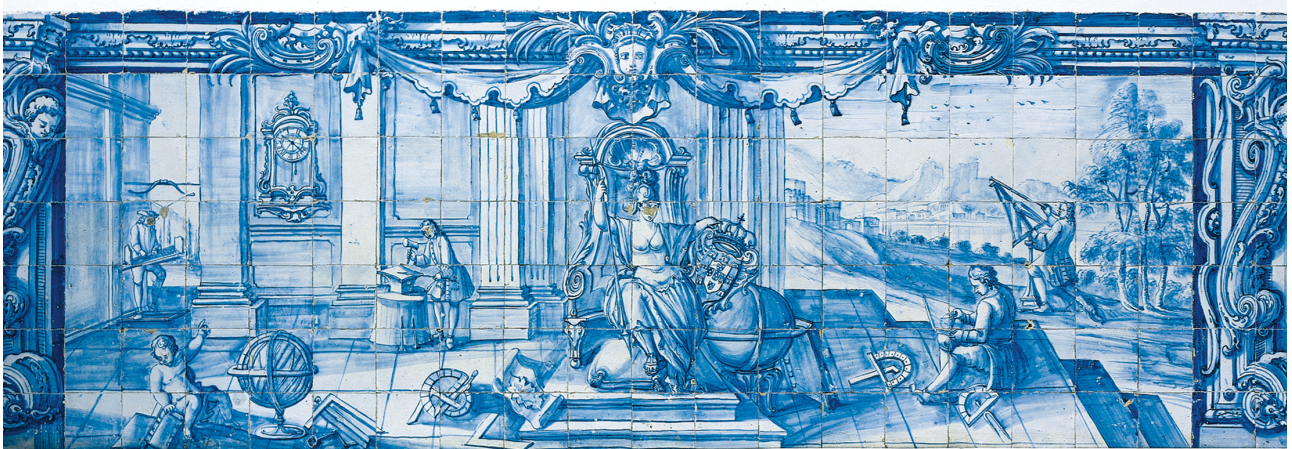
EQUILÍBRIO E TAXAS DE CONVERGÊNCIA EM SISTEMAS DINÂMICOS DISCRETOS NÃO AUTÓNOMOS

Maria Teresa Morais de Paiva Martins e Silva

Tese apresentada à Universidade de Évora
para obtenção do Grau de Doutor em Matemática
Especialidade: Análise Matemática

ORIENTADOR: *Luís Manuel Ferreira da Silva*
CO-ORIENTADOR: *Sara Luísa Dimas Fernandes*

ÉVORA, JULHO DE 2015



INSTITUTO DE INVESTIGAÇÃO E FORMAÇÃO AVANÇADA



EQUILÍBRIO E TAXAS DE CONVERGÊNCIA EM SISTEMAS DINÂMICOS DISCRETOS NÃO AUTÓNOMOS

Maria Teresa Morais de Paiva Martins e Silva

Tese apresentada à Universidade de Évora
para obtenção do Grau de Doutor em Matemática
Especialidade: Análise Matemática

ORIENTADOR: *Luís Manuel Ferreira da Silva*
CO-ORIENTADOR: *Sara Luísa Dimas Fernandes*

ÉVORA, JULHO DE 2015

INSTITUTO DE INVESTIGAÇÃO E FORMAÇÃO AVANÇADA

Resumo

Na primeira parte deste trabalho estudamos a convergência para a distribuição de equilíbrio em grafos não autónomos periódicos. Introduzimos a noção de equilíbrio em grafos não autónomos e apresentamos uma estimativa superior para a distância ao equilíbrio, à custa do segundo valor próprio em módulo das matrizes produto, supondo todas as matrizes ergódicas e pelo menos uma delas reversível. A estimativa obtida depende explicitamente da dimensão das matrizes consideradas. Estabelecemos a relação entre grafos autónomos e cadeias de Markov não homogêneas. Ilustramos, com um exemplo, a aplicação a sistemas dinâmicos não autónomos. Estendemos o estudo a matrizes não reversíveis demonstrando um resultado análogo ao caso reversível, no contexto autónomo e não autónomo, que utiliza a factorização da matriz através da forma normal de Jordan. Finalmente discutimos a pertinência dos resultados originais obtidos comparando-os com resultados conhecidos. Esta parte corresponde aos capítulos 2 e 3.

A segunda parte, capítulo 4, é dedicada ao estudo detalhado de uma família de sistemas não autónomos de período 2, gerados pela iteração sequencial de duas funções do tipo *tenda cortada*. Apresentamos os conceitos de dinâmica simbólica, renormalização e produto estrela no contexto não autónomo, com o objectivo de calcular a taxa de convergência de sucessões de pontos no espaço de parâmetros, construídas através de produtos estrela/renormalizações consecutivas, generalizando assim as sequências de Feigenbaum. Concluimos que as taxas de convergência são independentes do ponto inicial, mostrando assim que o contexto não autónomo exhibe propriedades universais do tipo encontrado por Feigenbaum em famílias de sistemas autónomos.

Palavras-Chave

Sistemas dinâmicos não autónomos, grafos não autónomos, cadeias de Markov não homogêneas, distribuição de equilíbrio, dinâmica simbólica, renormalização, produto estrela.

Abstract

Equilibrium and Convergence Rates in Nonautonomous Discrete Dynamical Systems

In the first part of this thesis we study the convergence for the equilibrium distribution in periodic non autonomous graphs. We introduce the notion of equilibrium in non autonomous graphs and give an upper bound for the distance to the equilibrium, using the second largest eigenvalue in modulus of the product matrices, assuming all of them ergodic and, at least, one reversible. The estimate obtained depends explicitly on the dimension of the considered matrices. We set the relation between non autonomous graphs and non homogeneous Markov chains. We illustrate the applications to the non autonomous dynamical systems with an example. We extend the study to non reversible matrices proving an analogous result, in both autonomous and non autonomous settings, using the matrix factorization with the Jordan normal form. Finally we discuss the relevance of the results obtained comparing them with the previously known results. This is the subject of chapters 2 and 3.

The second part, chapter 4, is dedicated to studying a family of non autonomous systems with period 2, generated by the sequential iteration of two *stunted sawtooth* maps. We present the concepts of symbolic dynamics, renormalization and star product in the non autonomous setting, in order to compute the convergence rates of sequences of points in the parameter space. These sequences are obtained through consecutive star products/renormalizations, generalizing in this way the Feigenbaum sequences. We show that the convergence rates are independent of the initial point, thus, concluding that the non autonomous setting has universal properties of the type founded by Feigenbaum in families of autonomous systems.

Keywords

Nonautonomous dynamical systems, nonautonomous graphs, non homogeneous Markov chains, equilibrium distribution, symbolic dynamics, renormalization, star product.

Agradecimentos

Agradeço ao Luís Silva e à Sara Fernandes pela orientação científica, mas também pela disponibilidade, encorajamento e motivação que foram fundamentais na elaboração deste trabalho.

Agradeço aos colegas da ADM pelo esforço na gestão da actividade docente que possibilitou uma maior dedicação a este trabalho e aos colegas da ADEETC pelo apoio logístico.

À Celina Belém agradeço a generosidade e a hospitalidade em Évora.

Agradeço aos meus pais e ao Pedro, a sua confiança é, e será sempre, fundamental.

Agradeço aos amigos que me têm acompanhado neste percurso.

Ao Luís agradeço tudo. Esta tese nunca teria existido sem o seu apoio incondicional.

À Catarina.

Índice

1	Introdução	1
1.1	Sistemas dinâmicos discretos não autónomos no intervalo	1
1.1.1	Cadeias de Markov não homogéneas enquanto sistemas dinâmicos não autónomos	2
1.1.2	Cadeias de Markov	3
1.1.3	Dinâmica simbólica	5
1.1.4	Sobre o operador de renormalização	6
1.2	Descrição do trabalho efectuado por capítulo	7
1.3	Trabalho futuro	9
2	Grafos nao autónomos	11
2.1	Resultados preliminares sobre grafos autónomos	11
2.2	Equilíbrio em grafos não autónomos periódicos	15
2.3	Convergência em cadeias de Markov não homogéneas	22
2.3.1	Tempo de <i>mixing</i>	24
2.3.2	Interpretação da convergência em cadeias de Markov não homogéneas	26
2.3.3	Uma estimativa sem valores próprios	27
2.4	Aplicação aos sistemas dinâmicos	28
2.4.1	O Operador de Perron-Frobenius	28
2.4.2	Exemplo	29
3	Matrizes de adjacência não reversíveis	33
3.1	Motivação	33
3.2	Estimativas conhecidas	34
3.2.1	Caso autónomo	34
3.2.2	Caso não autónomo	36
3.3	Estimativas na aproximação ao equilíbrio em grafos autónomos	41
3.3.1	Forma normal de Jordan de uma matriz	41
3.3.2	Estimativas obtidas	42
3.4	Estimativas na aproximação ao equilíbrio em grafos não autónomos periódicos	48

3.5	Exemplos	50
4	Taxas de convergência universais numa família de sistemas não autónomos	53
4.1	Família de funções <i>tenda cortada</i>	53
4.2	Produto cruzado	54
4.3	Dinâmica simbólica	55
4.3.1	Sistema não autónomo de período 2	60
4.4	Esqueleto no espaço de parâmetros	64
4.5	Renormalização e produto-*	70
4.6	Taxas de convergência	74
	Índice Remissivo	87
	Bibliografia	89

Capítulo 1

Introdução

1.1 Sistemas dinâmicos discretos não autônomos no intervalo

O presente trabalho insere-se no âmbito dos sistemas dinâmicos discretos, autônomos e não autônomos, de uma dimensão, originados pela composição sucessiva de sequências de funções num intervalo.

Consideremos a equação às diferenças

$$x_{n+1} = F(x_n),$$

onde $F : I \rightarrow \mathbb{R}$. Definimos um sistema dinâmico discreto autônomo através da equação anterior, obtendo uma sucessão de valores por iteração de F . Do mesmo modo, definimos um sistema dinâmico discreto não autônomo através de uma equação às diferenças em que a função F depende explicitamente de n , ou seja,

$$x_{n+1} = F(n, x_n). \tag{1.1}$$

Na palavras de Elaydi e Sacker, em [ES05b], existem duas escolas de equações às diferenças, uma delas vê as equações às diferenças como a versão discreta das equações diferenciais; a outra escola vê as equações às diferenças como sistemas dinâmicos discretos e os seus autores estudam predominantemente questões de estabilidade, bifurcação e caos. Ainda que a classificação sugerida seja artificial e surjam cada vez mais pontes entre as duas abordagens, neste trabalho sentimo-nos muito mais próximos da segunda escola e as nossas referências estão a ela associadas.

O sistema não autônomo originado pela equação às diferenças não autónoma (1.1) deve ser formulado de modo a generalizar as propriedades de semi-grupo satisfeitas no caso autônomo. Existem para isso duas opções, (ver [KR11]). A generalização mais óbvia

considera um semi-grupo a dois parâmetros, a outra hipótese é uma formulação usando o produto cruzado, com a vantagem de se interpretar o sistema como autónomo, num outro domínio. No Capítulo 4 usaremos este formalismo.

As equações às diferenças aparecem em muitos contextos, nomeadamente nas ciências naturais. Na Biologia, há muito tempo que a modelação do comportamento dinâmico de populações, usando dados experimentais para estimar os parâmetros da equação, é usada. R. B. May, em 1976, apresenta os sistemas dinâmicos discretos, do ponto de vista das aplicações à biologia e epidemiologia, (dando também exemplos na economia e nas ciências sociais), usando uma linguagem simples mas rigorosa, num artigo, [May76], que veio a ser amplamente divulgado, tornando-se um estímulo à investigação nesta área.

Em particular os sistemas dinâmicos não autónomos têm sido alvo de grande interesse nos últimos anos, e aqui incluímos sistemas discretos, provenientes de equações às diferenças não autónomas; sistemas contínuos, provenientes de equações diferenciais não autónomas; e também sistemas que não são gerados por equações como os provenientes de cadeias de Markov não homogéneas.

Alguns conceitos bem definidos no caso autónomo precisam de ser modificados adequadamente para que tenham interesse no caso não autónomo. É o caso da noção de equilíbrio que abordamos nos Capítulos 2 e 3. Do ponto de vista da convergência, a definição apresentada na Secção 2.1, do equilíbrio num sistema autónomo como vector próprio esquerdo associado ao valor próprio 1 da matriz de adjacência estocástica, deixa de fazer sentido no caso não autónomo genérico. Num sistema p -periódico, (gerado por uma sequência p -periódica de funções), é expectável que, sob certas condições, se obtenha convergência para um p -uplo ordenado de vectores. O caso particular, analisado em [SCZ07], em que todas as matrizes de adjacência partilham a mesma distribuição de equilíbrio, apesar de ter aplicações, parece-nos restritivo tendo em mente as aplicações aos sistemas dinâmicos gerados por sequências de funções.

1.1.1 Cadeias de Markov não homogéneas enquanto sistemas dinâmicos não autónomos

O enquadramento deste trabalho foi desde o início o dos sistemas dinâmicos discretos não autónomos. Durante o percurso que levou à escrita desta tese, e numa fase já avançada, com dois artigos escritos ([SSF14a] e [SSF14b]), demo-nos conta de que os objectos que estávamos a estudar tinham lugar no contexto das cadeias de Markov não homogéneas. Tanto quanto é do nosso conhecimento, as literaturas sobre sistemas dinâmicos não autónomos e cadeias não homogéneas pouco se cruzam. As diferentes motivações e aplicações parecem justificar o distanciamento. Enquanto a teoria das cadeias de Markov não homogéneas está mais focada na aplicação a métodos estatísticos, por exemplo para a obtenção de aproximações numéricas de funções, como o método de Monte Carlo, os

sistemas dinâmicos não autónomos, discretos ou contínuos, estão historicamente ligados às equações às diferenças ou às equações diferenciais que evidentemente inspiram muitas das aplicações dos sistemas. Este trabalho, em particular nos Capítulos 2 e 3, pretende dar um contributo na aproximação dos dois assuntos. Num artigo recente, [BM14], os autores constroem um sistema dinâmico discreto cuja evolução é dada por uma cadeia de Markov não homogénea, no que será a ligação mais imediata entre os dois conceitos.

1.1.2 Cadeias de Markov

Uma cadeia de Markov é essencialmente definida através das suas probabilidades de transição $P_{(t_0,t)}(x, A)$, onde $x \in \mathbb{S}$, ($\mathbb{S} = \mathbb{R}$ ou \mathbb{N} , ou \mathbb{Z}), $A \subset \mathbb{S}$ boreliano e $t \geq t_0 \geq 0$ (ver [Bré99] e [Eva]). $P_{(t_0,t)}(x, A)$ é interpretada como a probabilidade de transitar de x para um elemento do conjunto A , no instante t_0 , em tempo $t - t_0$, e satisfaz as seguintes propriedades:

1. $x \rightarrow P(x, A)$ é mensurável para cada A ;
2. $P(x, \cdot)$ é uma medida de Borel para cada x ;
3. (Igualdades de Chapman – Kolmogorov)
 - $P_{(t_0,t_0)}(x, \cdot) = \delta_x(\cdot)$, $\forall t_0$, onde $\delta_x(A) = 1$ se $x \in A$ e $\delta_x(A) = 0$ se $x \notin A$,
 - $P_{(t_0,t)}(x, A) = \int_{\mathbb{R}} P_{(t_0,s)}(y, A)P_{(s,t)}(x, dy)$, $t \geq s \geq t_0 \geq 0$.

Doravante consideramos cadeias de Markov em tempo discreto, $t \in \mathbb{N}$, com espaço de estados finito, de dimensão N , e fixamos $t_0 = 0$. Assim toda a informação pode ser encapsulada numa sucessão de matrizes de transição, $P_n(i, \{j\}) = P_n(i, j)$, com $(i, j) \in N \times N$, onde $P_n(i, j)$ é a probabilidade de transitar do estado i para o estado j no passo n , donde $P_n(i, A) = \sum_{j \in A} P_n(i, j)$. Se $P_n = P$ para todo o $n \in \mathbb{N}$ a cadeia diz-se homogénea, caso contrário diz-se não homogénea. Serão as cadeias deste último tipo o nosso principal objecto de interesse. Neste contexto a propriedade (3) toma a forma matricial $P_0 = Id$, (o que corresponde à ideia intuitiva da probabilidade de ir de um estado para outro em zero passos) e

$$P_{n+m}(i, j) = \sum_{k=1}^N P_n(i, k)P_m(k, j), \quad n, m \in \mathbb{N}$$

ou, matricialmente

$$P_{n+m} = P_n P_m, \quad n, m \in \mathbb{N}.$$

Das propriedades acima resultam as conhecidas propriedades das matrizes de transição das cadeias de Markov:

1. $\sum_{j=1}^N P_n(i, j) = 1$, $i = 1, \dots, N$, $\forall n \in \mathbb{N}$

$$2. P_n(i, j) \geq 0, \quad i, j = 1, \dots, N, \quad \forall n \in \mathbb{N}.$$

Voltando ao contexto das cadeias de Markov genéricas, com transições $P_t(x, A)$, (em que fixamos $t_0 = 0$ por facilidade de notação), define-se para cada $t \in \mathbb{R}$ o operador P_t no espaço das funções mensuráveis limitadas, do seguinte modo

$$P_t(f(x)) = \int_{\mathbb{R}} f(y)P_t(x, y)dy$$

Resulta das equações de Chapman – Kolmogorov que os operadores $(P_t)_{t \in \mathbb{R}}$ formam um grupo:

$$P_0 = Id, \quad P_{t+s} = P_t P_s.$$

Por outro lado, os operadores adjuntos P_t^* definidos por

$$P_t^*(\mu)(A) = \int_{\mathbb{R}} P_t(1_A(y))\mu(dy)$$

formam um grupo no espaço das medidas de probabilidade em \mathbb{R} , uma vez que

$$P_0^* = Id \quad P_{t+s}^* = P_t^* P_s^*.$$

Na linguagem das cadeias de Markov discretas num espaço de estados finito (CMDF), o operador adjunto P^* é representado pela matriz transposta P^T e uma medida μ corresponde a um vector coluna π^T com todas as componentes positivas e soma 1. Uma medida μ diz-se invariante para a cadeia de Markov $P_t(x, A)$ se $P_t^*(\mu) = \mu$ para todo o t , que no caso particular de CMDF se traduz na existência de um vector de distribuição π tal que $P_t^T \pi^T = \pi^T$ para todo o t , ou de outra forma, $\pi P_t = \pi$ para todo o t . No caso das CMDF homogéneas a existência de uma medida invariante é equivalente à existência de um vector próprio esquerdo normalizado de P , (em ℓ^1), associado ao valor próprio. Veremos que a irreduzibilidade da cadeia homogénea implica unicidade da medida invariante e a ergodicidade implica convergência por iteração de P , para a medida invariante.

Definidas as matrizes de transição como *estocásticas por linhas*, a iteração de P_t dá-se através de produtos matriciais à direita. A soma das entradas de cada linha vale 1 o que torna os produtos à esquerda ou à direita significativamente diferentes. Esta versão é a mais natural, tendo em conta a construção efectuada, mas na vasta literatura sobre produtos de matrizes encontramos resultados de produtos nos dois sentidos. Naturalmente, tomando matrizes estocásticas por colunas, os produtos correspondentes seriam à esquerda, mais coerentes com a ideia de iteração de funções.

Do conjunto das CMDF irreduzíveis começamos por estudar no Capítulo 2 as que satisfazem a propriedade de reversibilidade (ver definição 2.1.7). Estas matrizes representam

operadores autoadjuntos em relação ao produto escalar

$$(u, v)_\pi = \sum_{i=1}^N u_i v_i \pi_i$$

o que as torna especialmente *tratáveis*, pelo que, na literatura, muitos resultados se restringem a este caso. No Capítulo 3 apresentamos estimativas para a distância ao equilíbrio, usando o segundo maior valor próprio em módulo, sem a hipótese da reversibilidade, em ambos os casos, autónomo e não autónomo.

Encontrámos na literatura poucos resultados *quantitativos* sobre a convergência para o equilíbrio em cadeias de Markov discretas não homogéneas. Mossel, Peres e Sinclair em [MPS08] fornecem uma estimativa para o tempo de convergência de certas cadeias não homogéneas, a qual não utiliza valores próprios. Em [SCZ11], Saloff-Coste e Zúñiga discutem a ergodicidade fraca (*merging*) e estabilidade (generalização da existência de uma distribuição de equilíbrio) em cadeias não homogéneas em que a sucessão de matrizes é construída a partir de um conjunto finito de matrizes satisfazendo algumas propriedades. Os mesmo autores apresentam em [SCZ11] uma estimativa usando valores próprios, que é uma generalização da estimativa apresentada em [Bré99] para cadeias homogéneas. Em [SCZ07], num contexto mais genérico mas sob a forte hipótese de existência de uma distribuição de equilíbrio π comum a todas as matrizes $(P_n)_{n \in \mathbb{N}}$, (onde P_n corresponde ao passo n da cadeia), é fornecida uma estimativa para a distância entre a cadeia e o equilíbrio. Num artigo seguinte, [SCZ09], os autores relaxam a condição de existência de uma distribuição de equilíbrio comum para todas as matrizes e apresentam resultados para a velocidade de convergência em CMDP não homogéneas *c-estáveis*. Neste trabalho não supomos nenhuma hipótese desse tipo. Consideramos uma cadeia onde a sucessão das matrizes de transição (ergódicas no sentido homogéneo) é periódica e no Capítulo 2 apresentamos, sob certas condições, estimativas para a aproximação ao equilíbrio em função do tempo em grafos não autónomos periódicos, (definição mais à frente), que corresponde precisamente a cadeias de Markov não homogéneas cuja sucessão de matrizes de transição é periódica.

De seguida fazemos uma breve revisão da literatura no que diz respeito ao Capítulo 4, que envolve ferramentas distintas das apresentadas até agora.

1.1.3 Dinâmica simbólica

Quando em 1988 surgiu o célebre artigo de Milnor e Thurston [MT88b], (na verdade, divulgado desde 1977 através de *pre-prints*), a história da dinâmica simbólica já contava largas décadas. O primeiro registo ligado à ideia de dinâmica simbólica atribui-se ao trabalho de Hadamard ([Had98]) sobre o fluxo geodésico em superfícies de curvatura negativa, no fim do século XIX. Morse e Hedlund ([MH21]) foram responsáveis pelos

primeiros resultados *sistemáticos* em dinâmica simbólica com o objectivo de os aplicar ao estudo de sistemas contínuos. Ligações a outras áreas, como a teoria da informação de Shannon, espoletaram consideráveis avanços no desenvolvimento da dinâmica simbólica. A ideia base é dividir o espaço de estados num número finito de regiões e atribuir a letra de um alfabeto a cada uma delas. Assim, em vez de se seguir a trajectória de um ponto através da iteração de uma função, segue-se a ordem determinada pelas letras criando uma sequência simbólica. Uma iteração de f corresponde a um *shift* na sequência. Sendo evidente a perda de informação nesta abordagem, é surpreendente que uma parte substancial dessa informação resista, como a periodicidade das órbitas ou o comportamento caótico. No contexto das aplicações do intervalo, a dinâmica simbólica apareceu inicialmente associada a partições de Markov, no entanto, a partir dos anos 70, um artigo de Metropolis, Stein e Stein ([MSS73]) e sobretudo a *teoria do amassamento* de Milnor e Thurston ([MT88b]), trouxeram-lhe novo fôlego, aplicando-a de um modo mais genérico. Considerando partições definidas pelos pontos críticos, cria-se a possibilidade de estudar famílias de funções com o mesmo ponto crítico de modo unificado (ver [CH03]), o que fornece a descrição de toda uma família de funções à custa do mesmo alfabeto permitindo a comparação entre elas. Em [MT88b], os autores introduziram o conceito de invariante de amassamento de funções contínuas do intervalo, mais tarde o invariante foi generalizado para funções descontínuas, nomeadamente funções de Lorenz (ver [GH96] e [SR02]). A *teoria do amassamento*, longe de se ter esgotado, continua a permitir avanços no estudo dos sistemas dinâmicos. Este trabalho, em particular o Capítulo 4, pretende contribuir na sua aplicação aos sistemas discretos não autónomos, no seguimento do recente artigo [FSS13] que, tanto quanto é do nosso conhecimento, introduziu a dinâmica simbólica no caso não autónomo. Um dos conceitos centrais no último capítulo é a renormalização aplicada ao caso não autónomo e a sua ligação à dinâmica simbólica através do produto-*. De seguida apresentamos um pequeno apanhado sobre a sua recente história.

1.1.4 Sobre o operador de renormalização

Em 1978, Feigenbaum [Fei78] e, independentemente, Couillet e Tresser [CT78], observaram um fenómeno comum na transição da dinâmica *ordenada* para a dinâmica caótica através da duplicação de período, em famílias de funções unimodais a um parâmetro: para valores do parâmetro λ_n correspondentes a sucessivas bifurcações de duplicação do período tem-se

$$\lim_n \frac{\lambda_n - \lambda_{n-1}}{\lambda_{n+1} - \lambda_n} = 4.6692\dots$$

A geometria do diagrama de bifurcação teria assim propriedades universais nas famílias de funções a um parâmetro, que seriam amplamente exploradas. Para explicar a universalidade do fenómeno introduziu-se um operador no espaço das funções unimodais,

o operador da duplicação do período, e conjecturou-se que este operador teria um único ponto fixo hiperbólico e que as constantes encontradas seriam valores próprios da derivada do operador no ponto fixo. Esta conjectura foi provada por Lanford, em [Lan82], usando métodos numéricos e um computador mas cedo se identificaram as mesmas propriedades universais, encontradas na estrutura de bifurcação da duplicação do período noutros contextos, o que levou a uma reformulação da conjectura para um operador mais geral — o operador de renormalização. Em 1999, Lyubich ([Lyu99]), na sequência de determinantes contribuições de vários autores, nomeadamente Sullivan e McMullen, provou a conjectura para uma vasta classe de funções reais, com tipo combinatório limitado, sem recurso a um computador.

Dada uma função f definida num intervalo I , não é expectável que a função f^p restrin- gida a um certo intervalo $J \subset I$ seja do mesmo tipo que a função inicial, no entanto este fenómeno verifica-se para certas classes de funções. Nesse caso o processo pode ser iterado obtendo-se uma sucessão de funções do mesmo tipo definidas em regiões cada vez meno- res. O operador de renormalização é definido após um reescalonamento que transforma cada termo da sucessão numa função de I em I , de modo a que a dinâmica do operador ocorra num mesmo intervalo, o que permite estudar propriedades interessantes como a existência de pontos fixos [Avi10]. A unidade de tempo após renormalização corresponde a um número finito de unidades de tempo na dinâmica original. A renormalização para funções do intervalo pode ser formalizada do seguinte modo: Uma função $f : I \rightarrow I$ é re- normalizável no intervalo $[a, b] \subset I$ se existe n tal que $f^n : [a, b] \rightarrow [a, b]$ é topologicamente conjugada de outra função g . A generalização do conceito de renormalização num sistema não autónomo periódico foi introduzida em [FSS13].

1.2 Descrição do trabalho efectuado por capítulo

A sequência dos capítulos corresponde, essencialmente, à ordem cronológica do estudo realizado.

Quando demos início ao trabalho que veio a ser publicado no artigo [SSF14a], que é a base do Capítulo 2, tentámos perceber, através de exemplos e usando ferramentas compu- tacionais, o que se passava com os produtos cíclicos de matrizes de transição estocásticas de grafos não autónomos introduzidos em [AS12], do ponto de vista da convergência. Proposemo-nos estabelecer critérios para a existência de *um* estado de equilíbrio e para a convergência dos produtos de matrizes, com a motivação original de aplicar os resultados a sistemas discretos não autónomos periódicos. A questão seguinte surgiu imediatamente: nas hipóteses de convergência para o equilíbrio, que estimativas podemos obter para a taxa de convergência, e de que parâmetros dependem. Do caso autónomo, admitindo que a matriz é ergódica e reversível (as definições serão apresentadas oportunamente), sabemos que a convergência depende do espectro da matriz de adjacência, em particular do segundo

maior valor próprio em módulo. No caso não autónomo periódico mostrámos (Teorema 2.2.10) que depende igualmente do segundo maior valor próprio em módulo de matrizes que são produtos cíclicos B_i das matrizes individuais associadas a cada *etiqueta* do grafo não autónomo, desde que pelo menos uma das matrizes seja reversível. Recordados os conceitos de cadeias de Markov homogéneas e não homogéneas, na Secção 1.1.2, estabelecemos na Secção 2.3.2 a relação entre os resultados obtidos e os resultados de convergência em cadeias de Markov não homogéneas conhecidos. Na Secção 2.4 reconstituímos o caminho até aos sistemas dinâmicos, através do operador de Perron-Frobenius.

A condição mais forte com que nos confrontámos e da qual não conseguimos prescindir em [SSF14a] é a reversibilidade de pelo menos uma das matrizes B_i . Uma cuidadosa revisão na literatura não revelou existência de resultados sobre taxas de convergência para matrizes não reversíveis no contexto autónomo à custa dos seus valores próprios, tendo sido esse o ponto de partida para o artigo [SSF14b], cujos resultados são apresentados no Capítulo 3. Os exemplos com que havíamos trabalhado revelaram que, em certos casos, apesar de as matrizes não serem reversíveis, as estimativas para a convergência, usando o segundo valor próprio em módulo, eram válidas. Noutros casos isto não acontecia e claramente a velocidade de convergência estava relacionada com a presença de Blocos de Jordan de dimensão maior do que 1 na factorização que utiliza a forma normal de Jordan da matriz. Esta ideia levou-nos à demonstração do Teorema 3.3.3 e à sua generalização ao caso não autónomo, Teorema 3.4.1. Dos resultados originais obtidos destacamos o facto de não dependerem de nenhuma estrutura, tipo espaço de Hilbert. Na Secção 3.2 demonstramos, com base em [SCZ09], uma estimativa para a aproximação ao equilíbrio para cadeias não autónomas, sem considerar hipóteses de reversibilidade, que utiliza valores singulares de cada uma das matrizes individuais da cadeia. Note-se que a definição de valor singular pressupõe a existência de um operador adjunto, pelo que assenta numa estrutura de espaço de Hilbert. Terminamos o capítulo com a ilustração das estimativas obtidas através de alguns exemplos.

No Capítulo 4 estudamos um sistema dinâmico gerado pela iteração de duas funções do tipo *tenda cortada*. Tanto quanto é do nosso conhecimento, a descrição deste sistema, apresentada nos artigos [SSF15] e [SSR15], é original. Nestes dois artigos complementares, é fornecida uma descrição detalhada sobre um sistema dinâmico particular usando ferramentas desenvolvidas em [Rad04] e [FSS13], na linha do que foi feito em [Sil07] para funções de Lorenz cortadas. O objectivo é obter resultados que mostram a existência de estruturas de semelhança e universalidade no diagrama de bifurcação do sistema não autónomo. Para isso calculamos a taxa de convergência de uma sucessão de pontos no diagrama de bifurcação correspondentes a sequências simbólicas construídas com o produto-* e portanto relacionadas com a renormalização, introduzida em [FSS13] para sistemas não autónomos periódicos. Provamos que a taxa de convergência, à semelhança do que se passa no caso autónomo, não depende do ponto inicial. Nas secções iniciais do capítulo

munimo-nos de toda a artilharia para obter as taxas de convergência referidas, desde a dinâmica simbólica, ao conceito de esqueleto no contexto não autónomo, produto-* e renormalização.

1.3 Trabalho futuro

Um dos aspectos que pretendemos desenvolver no futuro tem a ver com as aplicações dos Capítulos 2 e 3 a sistemas dinâmicos não autónomos gerados pela iteração de funções, generalizando o que foi apresentado no exemplo da Secção 2.4.2, em que a partição de Markov é comum às três funções. Interessa-nos modelar o comportamento de um sistema em que as partições de Markov são distintas, digamos, satisfazendo apenas uma propriedade de Markov do tipo da definida em [KMS99], onde, dado um sistema não autónomo $(f_i, I_i)_{i \in \{1, \dots\}}$, a propriedade de Markov é satisfeita se existe uma sucessão de conjuntos finitos $C_i \subset I_i$, $i = 1, \dots$ nas seguintes condições:

- Os extremos de I_i pertencem a C_i , $i = 1, \dots$
- Para qualquer componente J de $I_i \setminus C_i$ a função f_i é monótona em J ;
- $f_i(C_i) \subset C_{i+1}$, para $i \geq 1$.

Em [AS15] é proposta uma ligação entre sistemas não autónomos periódicos gerados por funções de Lorenz com itinerários críticos finitos e grafos não autónomos, em que as arestas são definidas através da admissibilidade das sequências simbólicas. Assim podemos construir partições de Markov, baseadas nas sequências simbólicas, que não são necessariamente comuns às funções do sistema periódico. Parece-nos que uma generalização do nosso trabalho a este contexto seria pertinente. Nesse caso podemos construir matrizes $(B_i)_{i \in \{0, \dots, p-1\}}$ ergódicas com diferentes dimensões e a teoria das cadeias de Markov não homogéneas não fornece respostas tão claras. Ainda assim é expectável que no caso periódico se consiga manter a definição de vector equilíbrio Π com vectores de diferentes dimensões e obter resultados análogos quanto à convergência.

Como trabalho de fundo, gostaríamos de explorar a teoria das cadeias de Markov não homogéneas e aplicá-la aos sistemas não autónomos.

Em relação ao Capítulo 4 parece-nos expectável a obtenção de resultados análogos noutros sistemas não autónomos, em famílias de tendas cortadas com dois declives diferentes. Esta alteração levanta desde logo questões na definição de Φ , por isso não é uma generalização trivial. Sistemas definidos com famílias de funções quadráticas, seguindo mais uma vez o trabalho de [Rad04], é também uma direcção possível.

Capítulo 2

Grafos não autônomos

2.1 Resultados preliminares sobre grafos autônomos

A teoria de grafos surgiu inicialmente com o objectivo de modelar relações binárias entre objectos de um sistema. Mais tarde, de modo a incluir diferentes relações entre os objectos, muniu-se cada aresta com um valor, *peso*, que quantifica a relação entre os dois vértices e que pode ser interpretado como probabilidade, custo, comprimento, etc, construindo assim um grafo *pesado*. Para dar resposta a problemas de diferentes áreas, das ciências sociais à biologia, cedo se reconheceu a necessidade de introduzir modelos onde as relações entre os objectos poderiam sofrer alterações ao longo do tempo (ver [Sil08] e as referências aí apresentadas). Segundo [NL08], um grafo é dinâmico quando algumas entidades (vértices, arestas, pesos) variam no tempo. Neste capítulo focamo-nos em grafos que apresentam evolução dos pesos ao longo do tempo, definindo uma versão *pesada* de grafo não autónomo introduzido por [AS12]. A um grafo pesado está associada uma matriz de adjacência a partir da qual construímos uma matriz estocástica que define uma cadeia de Markov discreta. Uma cadeia de Markov representada por uma matriz ergódica (a definir mais à frente) converge para uma distribuição de equilíbrio e em certas condições é possível obter estimativas para a distância ao equilíbrio em função do número de iteradas, ou *passos da cadeia*, à custa de valores próprios ([Beh00]).

Nas próximas secções estabelecemos um enquadramento análogo para grafos não autónomos periódicos representados por p matrizes, que nos remete imediatamente para algumas questões:

1. Como definir uma versão da distribuição de equilíbrio num sistema não autónomo?
2. Definido o vector de equilíbrio Π , em que condições:
 - Existe convergência para o equilíbrio?
 - É possível obter resultados análogos para a taxa de convergência usando valores espectrais de cada uma das matrizes?

O nosso objectivo é discutir e responder a estas questões, definindo equilíbrio em grafos não autônomos periódicos como um vector de equilíbrios e obter uma estimativa para o erro sob certas condições sobre as matrizes de adjacência. Enquanto no caso autónomo obtemos estimativas usando valores próprios da matriz estocástica construída a partir da matriz de adjacência A , no caso não autónomo periódico esse papel é desempenhado pelo produto das p matrizes estocásticas $B_i = P_i \cdots P_{(i+p-1) \bmod p}$, $i = 1, \dots, p$.

Definição 2.1.1. *Um grafo dirigido pesado D é um par ordenado $D = (V, E)$ onde V é um conjunto não vazio de N vértices e $E \subset V \times V$ é o conjunto das arestas orientadas. Cada dois vértices v_i e v_j podem estar ligados por arestas orientadas $(v_i, v_j) \in E$, às quais associamos pesos $e_{ij} \in \mathbb{R}^+$. Se $(v_i, v_j) \notin E$ tomamos $e_{ij} = 0$.*

O peso e_{ij} pode ser interpretado como a probabilidade ou o custo de passar do estado i para o estado j . A um grafo D associamos uma matriz A , dita matriz de adjacência ou de transição, com entradas não negativas, tais que $A(i, j) = e_{ij}$ se $(v_i, v_j) \in E$ e $A(i, j) = 0$ se $(v_i, v_j) \notin E$ (ver [Sil08]).

Definimos a matriz P tal que

$$P(i, j) = \frac{A(i, j)}{\sum_{k=1}^N A(i, k)}.$$

P é uma matriz estocástica uma vez que satisfaz as propriedades:

1. $P(i, j) \geq 0$, para $i, j = 1, \dots, N$
2. $\sum_{j=1}^N P(i, j) = 1$ para cada $i = 1, \dots, N$.

Os seguintes resultados sobre matrizes estocásticas podem ser encontrados por exemplo em [Beh00].

Dada uma matriz estocástica P , 1 é um valor próprio, ao qual está associado o vector próprio direito $(1, \dots, 1) \in \mathbb{R}^N$ e um vector próprio esquerdo v , i.e.,

$$\exists v \in \mathbb{R}^N : vP = v.$$

Qualquer valor próprio λ_i de uma matriz estocástica satisfaz $|\lambda_i| \leq 1$.

Definição 2.1.2. *Uma matriz P não negativa diz-se irredutível se existe $k \in \mathbb{N}$ tal que as entradas da matriz $P + P^2 + \dots + P^k$ são estritamente positivas.*

Definição 2.1.3. *Uma matriz P diz-se aperiódica se, para cada (i, j) , ou $P^k(i, j) = 0, \forall k \in \mathbb{N}$, ou existe N tal que para $k > N$ se tem $P^k(i, j) > 0$.*

Definição 2.1.4. *Se existir $k \in \mathbb{N}$ tal que todas as entradas da matriz P^k são estritamente positivas então a matriz P diz-se ergódica.*

Das definições anteriores resulta que uma matriz ergódica é irredutível e aperiódica. Dizemos que uma cadeia de Markov é irredutível, aperiódica ou ergódica se a matriz que a representa é irredutível, aperiódica ou ergódica, respectivamente.

Uma matriz de adjacência irredutível corresponde a um grafo conexo, se além disso a matriz é aperiódica (logo ergódica) então o grafo não pode ser bipartido [AF14]. O resultado seguinte é central na teoria das cadeias de Markov. Apresentamos uma versão para matrizes estocásticas cuja prova pode ser encontrada por exemplo em [Sen06]:

Teorema 2.1.5 (Perron-Frobenius). *Uma matriz estocástica ergódica P satisfaz as seguintes condições:*

1. $\lambda = 1$ é o único valor próprio de módulo 1, ou seja $1 = \lambda_1 > |\lambda_2| \geq \dots \geq |\lambda_N|$;
2. A multiplicidade algébrica de $\lambda = 1$ é 1.

Nota. Se a multiplicidade algébrica de um valor próprio λ é 1, — ou seja, λ é uma raiz simples do polinómio característico $\det(A - \lambda Id)$, — então a multiplicidade geométrica é também 1, — ou seja, o subespaço próprio gerado pelos vectores próprios associados a $\lambda = 1$ tem dimensão 1.

Assim, o subespaço próprio esquerdo associado a $\lambda = 1$ tem um único vector π normalizado em ℓ^1 , com todas as componentes positivas, isto é:

1. $\pi P = \pi$;
2. $\sum_{i=1}^N \pi_i = 1$.

Veremos que este vector desempenha um papel fundamental no comportamento da cadeia.

Nota. Recordemos que a norma de um vector $v = (v_1, \dots, v_N)$ em ℓ^1 é dada por $\|v\| = \sum_{i=1}^N |v_i|$.

Definição 2.1.6. *Dada uma matriz ergódica P , chamamos distribuição de equilíbrio ou distribuição estacionária ao único vector próprio esquerdo de P normalizado em ℓ^1 e representamo-lo por π .*

Nota. Por conveniência de notação representamos os vectores linha por v e os vectores coluna por v^T .

Definição 2.1.7. *Dizemos que uma matriz P e um vector π satisfazem as equações de detailed balance¹ se*

$$\pi_i P(i, j) = \pi_j P(j, i), \quad \forall i, j. \quad (2.1)$$

Se P e π satisfazem (2.1) a matriz P diz-se reversível.

¹Não encontramos nenhuma tradução que nos pareça captar a ideia do termo inglês.

É fácil verificar que, se um vector v satisfaz esta igualdade então v é o vector de equilíbrio, ou seja, $v = \pi$, daí a escolha de notação para o vector na definição anterior.

A partir de uma matriz estocástica reversível P podemos construir um grafo pesado não dirigido com pesos $w_j = \pi_j p(i, j)$, (ver [AF14]).

Teorema 2.1.8. *Dada uma matriz estocástica ergódica P , as potências P^k convergem componente a componente para uma matriz estocástica W cujas linhas são iguais à distribuição de equilíbrio π . Isto é,*

$$\lim_{k \rightarrow \infty} P^k = \lim_{k \rightarrow \infty} \begin{bmatrix} P(1,1) & \dots & P(1,N) \\ \vdots & \ddots & \vdots \\ P(N,1) & \dots & P(N,N) \end{bmatrix}^k = \begin{bmatrix} \pi_1 & \dots & \pi_N \\ \dots & \dots & \dots \\ \pi_1 & \dots & \pi_N \end{bmatrix} = W.$$

A prova pode ser encontrada por exemplo em [Beh00]. Num espaço de estados finito a convergência componente a componente é equivalente à convergência em norma.

Garantida a convergência das potências $P^n(i, j)$, é natural, e relevante do ponto de vista das aplicações (algoritmos computacionais, etc), procurar estimativas para a velocidade da convergência, nomeadamente obter majorações para a *distância* entre P^n e W em função de n . A distância entre P^n e W é equivalente à distância entre o estado da cadeia vP^n ao fim de n passos com determinada distribuição de probabilidade inicial v , a π , pretendendo-se majorar nesse caso o erro na aproximação ao estado de equilíbrio a partir de uma distribuição inicial. A rapidez com que a cadeia se aproxima do estado de equilíbrio pode ser quantificada através do número de iteradas necessário para atingir determinada distância ao equilíbrio, conhecido na literatura como *tempo de mixing*². O nosso objectivo neste capítulo é a obtenção de uma estimativa para o erro maximal, no caso não autónomo, à custa dos valores próprios de P . Em particular apresentaremos um resultado que generaliza o Teorema 2.1.11, para o qual precisamos das seguintes definições:

Definição 2.1.9. *Dada uma matriz estocástica e ergódica P representamos por λ^* o maior dos módulos dos valores próprios de P diferentes de 1, ou seja*

$$\lambda^* = \max\{|\lambda| : \lambda \text{ é um valor próprio de } P \wedge \lambda \neq 1\}.$$

Sabemos que $\lambda^* < 1$ pelo Teorema 2.1.5.

Definição 2.1.10. *Definimos o erro maximal relativo das componentes de P^k em relação às componentes de π , por*

$$\delta(k) = \max_{i,j} \left(\frac{|P^k(i, j) - \pi_j|}{\pi_j} \right),$$

²ou tempo de decaimento das correlações.

onde $P^k(i, j)$ é a entrada (i, j) de P^k .

O teorema seguinte fornece uma estimativa para $\delta(k)$ usando λ^* :

Teorema 2.1.11. *Se P é uma matriz estocástica ergódica e reversível então*

1. $\delta(k) \leq \frac{(\lambda^*)^k}{\min_i \pi_i}$, para qualquer k ;

2. $\delta(k) \geq (\lambda^*)^k$, para k par.

Se todos os valores próprios são maiores ou iguais a zero então a desigualdade é válida para todos os k .

Uma matriz reversível é semelhante a uma matriz simétrica, logo todos os valores próprios são reais. A demonstração deste teorema pode ser encontrada por exemplo em [Beh00].

2.2 Equilíbrio em grafos não autônomos periódicos

De seguida introduzimos uma generalização dos resultados anteriores para o caso não autônomo periódico, provados em [SSF14a]. Começamos por apresentar algumas definições:

Definição 2.2.1. *Um grafo dirigido não autônomo é um par $\mathcal{G} = (V, (E_i)_{i=0}^\infty)$ onde $V = \{1, \dots, N\}$ e $E_i \subset V \times V$, para todo o i inteiro. Os elementos de V dizem-se vértices e os elementos (v_j, v_l) de E_i chamam-se i -arestas. A cada par $(v_j, v_l) \in E_i$ associamos um peso $e_i(j, l) > 0$ se $(v_j, v_l) \in E_i$ e definimos $e_i(j, l) = 0$ se $(v_j, v_l) \notin E_i$.*

Os grafos não autônomos são interpretados de forma simples como grafos cujos pesos (probabilidades, ou outros) mudam ao longo do tempo.

Definição 2.2.2. *Dizemos que um grafo \mathcal{G} é p -periódico, com p natural, se $E_{i+p} = E_i$, para qualquer $i \in \mathbb{N}$.*

Para cada $i \in \mathbb{N}_0$, seja A_i a matriz de adjacência associada a $D_i = (V, E_i)$, ou seja, a matriz com entradas $A_i(j, l) \in \mathbb{R}_0^+$ tais que $A_i(j, l) = e_i(j, l)$.

A sucessão de matrizes $(A_i)_{i=0}^\infty$ é periódica de período p se \mathcal{G} é p -periódico. Fazemos corresponder a cada matriz A_i uma matriz estocástica P_i do seguinte modo,

$$P_i(j, l) = \frac{A_i(j, l)}{\sum_{k=1}^N A_i(j, k)}.$$

Nota 2.2.2.1. *A sucessão de matrizes $(P_i)_{i=0}^\infty$ associada ao grafo \mathcal{G} pode ser interpretada como uma cadeia de Markov não homogénea.*

Um caminho num grafo não autônomo é uma sucessão infinita de vértices $(v_i)_{i \in \mathbb{N}}$ tais que $A_i(v_i, v_{i+1}) > 0$ para todo $i \in \mathbb{N}_0$.

Dado um grafo não autônomo \mathcal{G} , p -periódico, com $p \in \mathbb{N}$, matrizes de adjacência (A_0, \dots, A_{p-1}) e correspondentes matrizes estocásticas (P_0, \dots, P_{p-1}) , construímos o conjunto de p -matrizes, (B_0, \dots, B_{p-1}) do seguinte modo

$$B_i = P_i \cdots P_{p-1} P_0 \cdots P_{i-1}. \quad (2.2)$$

Lema 2.2.3. B_i é uma matriz estocástica para qualquer $i \in \{0, \dots, p-1\}$.

Demonstração. O produto de quaisquer duas matrizes estocásticas A, B é uma matriz estocástica, pois

$$\sum_k (AB)_{ik} = \sum_k \sum_j A_{ij} B_{jk} = \sum_j \left(A_{ij} \sum_k B_{jk} \right) = \sum_j A_{ij} = 1$$

□

Teorema 2.2.4. Dado um grafo não autônomo p -periódico \mathcal{G} com matrizes $(B_j)_{j=0, \dots, p-1}$ definidas em (2.2), se existe $i \in \{0, \dots, p-1\}$ tal que B_i é ergódica, então para cada $j \in \{0, \dots, p-1\}$ as potências B_j^k , $k \in \mathbb{N}$, convergem componente a componente para uma matriz estocástica W_j onde todas as linhas são iguais.

Se representarmos cada linha de W_j por $\pi(j) = [\pi_1(j) \dots \pi_N(j)]$ então $\pi(j)$ é o único vector tal que $\pi(j)B_j = \pi(j)$ e $\sum_k \pi_k(j) = 1$.

Demonstração. Se B_i é ergódica, as potências B_i^k convergem componente a componente para W_i como vimos no Teorema 2.1.8. Consideremos agora $j \neq i$. Por simplicidade e sem perda de generalidade tomemos $j = i+1$, para $j \neq i+1$ podemos usar os mesmos argumentos. Temos

$$\begin{aligned} \lim_{k \rightarrow \infty} B_{i+1}^k &= \lim_k (P_{i+1} \cdots P_{i-1} P_i)^k = \lim_k \left(P_{i+1} \cdots P_{i-1} B_i^{k-1} P_i \right) = P_{i+1} \cdots P_{i-1} \lim_k B_i^{k-1} P_i = \\ &= P_{i+1} \cdots P_{i-1} W_i P_i. \end{aligned}$$

Uma vez que $P_{i+1} \cdots P_{i-1}$ é estocástica, a soma das entradas de cada linha é 1, por outro lado todas as linhas de W são iguais, então

$$P_{i+1} \cdots P_{i-1} W_i = W_i$$

donde

$$P_{i+1} \cdots P_{i-1} W_i P_i = W_i P_i.$$

Fazendo $W_{i+1} = W_i P_i$, como nenhuma entrada da matriz depende de j a matriz tem todas

as linhas iguais:

$$W_{i+1}(j, k) = \sum_n W_i(j, n)P_i(n, k) = \sum_n \pi_n(i)P_i(n, k)$$

representando cada linha por $\pi(i+1) = [\pi_1(i+1) \dots \pi_N(i+1)]$ temos

$$\pi_k(i+1) = \sum_n \pi_n(i)P_i(n, k), \quad k = 0, \dots, N \quad (2.3)$$

e este vector é normalizado, no sentido de ℓ^1 :

$$\sum_k \pi_k(i+1) = \sum_k \left(\sum_n \pi_n(i)P_i(n, k) \right) = \sum_n \left(\pi_n(i) \sum_k P_i(n, k) \right) = \sum_n \pi_n(i) = 1$$

O facto de que $\pi(i+1)B_{i+1} = \pi(i+1)$ e a sua unicidade são consequência imediata do caso autónomo. \square

Corolário 2.2.5. *Nas condições do Teorema 2.2.4, qualquer que seja o vector v tal que $v_i \geq 0$ e $\sum_{i=1}^N v_i = 1$, tem-se para todo o $i \in \{0, \dots, p\}$, $l, k \in \{1, \dots, N\}$,*

$$\lim_{k \rightarrow \infty} \max_l \left| (vB_i^k)_l - \pi_l(i) \right| = 0.$$

Demonstração. O corolário anterior é consequência directa do caso autónomo, basta ver que, para qualquer i e j :

$$\begin{aligned} |(vB_i^k)_j - \pi_j(i)| &= \left| \sum_{l=1}^N v_l B_i^k(l, j) - \pi_j(i) \right| = \left| \sum_{l=1}^N v_l (B_i^k(l, j) - \pi_j(i)) \right| \\ &\leq \sum_{l=1}^N v_l \left| (B_i^k(l, j) - \pi_j(i)) \right| \leq \sum_{l=1}^N v_l \max_l \left| (B_i^k(l, j) - \pi_j(i)) \right| \\ &= \max_l \left| (B_i^k(l, j) - \pi_j(i)) \right| \end{aligned}$$

\square

Terminamos esta série de resultados apresentando o seguinte teorema sobre o espectro de produtos cíclicos de matrizes o qual será usado na obtenção das majorações:

Proposição 2.2.6. *Se existe $i \in \{0, \dots, p-1\}$ tal que λ é um valor próprio de B_i então λ é valor próprio de B_j para qualquer $j \in \{0, \dots, p-1\}$. Em particular, se existe i tal que B_i é ergódica então $\lambda = 1$ é um valor próprio com multiplicidade geométrica 1 para todas as matrizes B_i .*

Demonstração. Se λ é um valor próprio de B_i então existe v tal que $vB_i = \lambda v$, multipli-

cando ambos os membros por P_i temos

$$\begin{aligned} vB_iP_i &= \lambda vP_i \\ \Leftrightarrow vP_iP_{i+1} \cdots P_{i-1}P_i &= \lambda vP_i \\ \Leftrightarrow vP_iB_{i+1} &= \lambda vP_i \\ \Leftrightarrow uB_{i+1} &= \lambda u \end{aligned}$$

Então $u = vP_i$ é um vector próprio esquerdo de B_{i+1} correspondente ao valor próprio λ . Pelo Teorema 2.2.4 sabemos que o espaço próprio esquerdo associado a $\lambda = 1$ é gerado por $\pi(i)$, então λ tem multiplicidade geométrica 1. Repetindo o argumento p vezes temos o resultado para todos os $j \in \{0, \dots, n-1\}$. \square

Os resultados anteriores permitem estabelecer a seguinte definição:

Definição 2.2.7. *Dado um grafo \mathcal{G} não autónomo p -periódico, a distribuição de equilíbrio Π é dada pelo vector*

$$\Pi = (\pi(0), \dots, \pi(p-1)),$$

onde, para cada i , $\pi(i)$ é a distribuição de equilíbrio associada a B_i .

Proposição 2.2.8. *A distribuição de equilíbrio Π pode escrever-se na forma*

$$\Pi = (\pi(0), \pi(0)P_0, \dots, \pi(0)P_0 \cdots P_{p-1})$$

Demonstração. A igualdade $\pi(i+1) = \pi(i)P_i$ resulta directamente da igualdade (2.3) na prova do Teorema 2.2.4. \square

A partir desta definição, considerando os pesos como probabilidades, cada componente $\pi_i(j)$ pode ser interpretada como a probabilidade de o sistema estar na posição i depois de um grande número n de iteradas tal que $n \bmod p = j$.

Como o segundo valor próprio em módulo, λ^* , é o mesmo para todas as matrizes B_i , é natural esperar que os resultados na convergência das potências $(B_i)^k$ sejam uniformes.

Definição 2.2.9. *O erro maximal na aproximação das componentes da distribuição de equilíbrio Π por $(B_1^k, \dots, B_{p-1}^k)$ é dado por*

$$\delta(k) = \max_{l,i,j} \frac{|B_l^k(i,j) - \pi_j(l)|}{\pi_j(l)}.$$

O teorema seguinte é a versão do Teorema 2.1.11 para sistemas não autónomos, fornecendo uma estimativa superior baseada em λ^* para o tempo de convergência quando as matrizes B_i , com $i \in \{0, \dots, p-1\}$ são ergódicas e pelo menos uma delas é reversível (ver [SSF14a]).

Teorema 2.2.10. *Consideremos um grafo \mathcal{G} , não autónomo p -periódico com matrizes estocásticas, de dimensão $N \times N$, P_0, \dots, P_{p-1} e respectivas matrizes associadas B_0, \dots, B_{p-1} tais que B_l é ergódica para todos os $l \in \{0, \dots, p-1\}$. Se, para certo $j \in \{0, \dots, p-1\}$, uma das matrizes B_j é reversível então temos*

$$\delta(k) \leq \frac{(\lambda^*)^{k-1}(N-1)}{\min_i \pi_i(j)} \quad (2.4)$$

Demonstração. Consideremos as p matrizes B_i ergódicas para $i \in \{0, \dots, p-1\}$, construídas a partir do grafo \mathcal{G} e supomos que existe $j \in \{0, \dots, p-1\}$ tal que B_j é reversível. Tomamos $j = 0$, por exemplo. Se B_0 é reversível satisfaz as equações de *detailed balance* com $\pi(0)$, e existe uma matriz simétrica A tal que

$$A = DB_0D^{-1} \quad (2.5)$$

onde D é a matriz diagonal:

$$D = \begin{bmatrix} \sqrt{\pi_1(0)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{\pi_N(0)} \end{bmatrix}.$$

Podemos escolher uma matriz ortogonal S tal que

$$C = SAS^{-1} \quad (2.6)$$

é diagonal. As linhas $e_1 \dots e_N$ de S , assim como as colunas, são ortogonais (em relação ao produto escalar usual).

Uma vez que A é simétrica todos os valores próprios são reais. Assumimos, sem perda de generalidade,

$$C = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{bmatrix}$$

onde $1 > |\lambda_2| \geq \dots \geq |\lambda_N|$. Cada e_i , $i \in \{1, \dots, N\}$, é um vector próprio esquerdo de A associado a λ_i , em particular

$$e_1A = e_1,$$

logo

$$e_1DB_0 = e_1D,$$

o que significa que e_1D é um vector próprio esquerdo de B_0 associado a 1. Assim existe c

tal que $\pi(0) = ce_1D$, donde $\pi(0)D^{-1} = ce_1$.

Como ambos os vectores $\pi(0)D^{-1}$ e e_1 são normalizados (em relação ao produto escalar usual) temos $\pi(0) = \pm e_1D$. Assumindo $\pi(0) = e_1D$, vem

$$e_1 = \left(\sqrt{\pi_1(0)}, \dots, \sqrt{\pi_N(0)} \right)$$

Mais, da igualdade (2.6) podemos obter

$$A^k = \sum_{i=1}^N \lambda_i^k e_i^T e_i$$

que assegura uma útil representação de B_0^k :

$$B_0^k = (D^{-1}AD)^k = D^{-1}A^kD = \sum_{i=1}^N \lambda_i^k D^{-1}e_i^T e_iD$$

Por outro lado

$$B_1^{k+1} = P_1 \dots P_{p-1} (P_0 \dots P_{p-1})^k P_0 = P_1 \dots P_{p-1} B_0^k P_0$$

então

$$B_1^{k+1} = \sum_{i=1}^N \lambda_i^k P_1^* D^{-1} e_i^T e_i D P_0$$

com $P_1^* = P_1 \dots P_{p-1}$.

Uma vez que conhecemos e_1 e D podemos calcular a primeira parcela da soma anterior:

$$\lambda_1^k P_1^* D^{-1} e_1^T e_1 D P_0 = P_1^* D^{-1} e_1^T e_1 D P_0 = P_1^* W_0 P_0 = P_1^* W_1$$

onde cada linha de W_i coincide com $\pi(i)$, $i = 0, 1, \dots, p-1$, logo $W_0 P_0 = W_1$.

Estamos finalmente em condições de escrever explicitamente a entrada (i, j) da matriz B_1^{k+1} :

$$\begin{aligned} B_1^{k+1}(i, j) &= \\ &= \sum_{m=1}^N P_1^*(i, m) W_1(m, j) + \sum_{l=2}^N \lambda_l^k \sum_{m=1}^N \sum_{n=1}^N \frac{\sqrt{\pi_m(0)}}{\sqrt{\pi_n(0)}} P_1^*(i, n) S(l, m) S(l, n) P_0(m, j) \\ &= \pi_j(1) + \sum_{l=2}^N \lambda_l^k \sum_{m=1}^N S(l, m) \sqrt{\pi_m(0)} P_0(m, j) \sum_{n=1}^N \frac{1}{\sqrt{\pi_n(0)}} P_1^*(i, n) S(l, n). \end{aligned}$$

Usando o facto de P_1^* ser estocástica, $|S(l, n)| \leq 1$ and $\pi(1) = \pi(0)P_0$, temos

$$\begin{aligned}
& \left| \frac{B_1^{k+1}(i, j) - \pi_j(1)}{\pi_j(1)} \right| = \\
&= \frac{1}{\pi_j(1)} \left| \sum_{l=2}^N \left(\lambda_i^k \sum_{m=1}^N S(l, m) \sqrt{\pi_m(0)} P_0(m, j) \cdot \sum_{n=1}^N \frac{1}{\sqrt{\pi_n(0)}} P_1^*(i, n) S(l, n) \right) \right| \\
&\leq \frac{(\lambda^*)^k}{\pi_j(1)} \left| \sum_{l=2}^N \left(\max_m S(l, m) \sum_{m=1}^N \frac{\pi_m(0)}{\sqrt{\pi_m(0)}} P_0(m, j) \cdot \sum_{n=1}^N \frac{1}{\sqrt{\pi_n(0)}} P_1^*(i, n) S(l, n) \right) \right| \\
&\leq \frac{(\lambda^*)^k}{\pi_j(1) \min_i \pi_i(0)} \sum_{l=2}^N \left(\left| \max_m S(l, m) \right| \cdot \sum_{m=1}^N \pi_m(0) P_0(m, j) \left| \max_n S(l, n) \right| \sum_{n=1}^N P_1^*(i, n) \right) \\
&\leq \frac{(\lambda^*)^k}{\min_i \pi_i(0)} \sum_{l=2}^N \left| \max_m S(l, m) \right| \left| \max_m S(l, m) \right| = \frac{(\lambda^*)^k}{\min_i \pi_i(0)} \sum_{l=2}^N \left| \max_m S(l, m) \right|^2 \\
&\leq \frac{(\lambda^*)^k (N-1)}{\min_i \pi_i(0)}
\end{aligned}$$

Usando os mesmos argumentos provamos a desigualdade para qualquer B_i , com $i \in \{0, \dots, p-1\}$. \square

A estimativa obtida no teorema anterior depende explicitamente da dimensão N da matriz. No caso autónomo a dependência não é explícita mas naturalmente existe. Por exemplo, se π é a distribuição de probabilidade uniforme temos $\pi = (\frac{1}{N}, \dots, \frac{1}{N})$.

Nota. A ergodicidade de cada uma das matrizes B_i é essencial na obtenção da estimativa para o erro relativo, garantindo que todas as componentes da distribuição de equilíbrio Π são não nulas, esta não poderia ser substituída pela ergodicidade das matrizes P_i , como se pode concluir do exemplo de sistema de período 2 com matrizes de transição:

$$P_0 = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad P_1 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix},$$

P_0 e P_1 são ambas ergódicas sem que os produtos B_0 e B_1 o sejam. Este exemplo é paradigmático dos problemas em definir condições suficientes de convergência em grafos não autónomos.

Corolário 2.2.11. *Consideremos um vector $v = (v_1, \dots, v_N)$, com todas as componentes positivas, normalizado em ℓ^1 , e o conjunto de p matrizes B_l , com $l \in \{0, \dots, p-1\}$, nas condições do Teorema 2.2.10. Se existe $n \in \{0, \dots, p-1\}$ tal que B_n é reversível, tem-se*

$$\max_{j,l} \left| \frac{(vB_l^k)_j - \pi_j(l)}{\pi_l(l)} \right| \leq \frac{3(\lambda^*)^{k-1}(N-1)}{\min_j \pi_j(n)} \quad (2.7)$$

Demonstração. Para qualquer l, j temos

$$\begin{aligned}
\frac{|(vB_l^k)_j - \pi_j(l)|}{\pi_j(l)} &= \frac{|v_1 B_l^k(1, j) + v_2 B_l^k(2, j) + \dots + v_N B_l^k(N, j) - \pi_j(l)|}{\pi_j(l)} \\
&= \frac{1}{\pi_j(l)} \left| (v_1 + \dots + v_n) B_l^k(1, j) + v_2 (B_l^k(2, j) - B_l^k(1, j)) + \right. \\
&\quad \left. + \dots + v_N (B_l^k(N, j) - B_l^k(1, j)) - \pi_j(l) \right| \\
&= \frac{|B_l^k(1, j) - \pi_j(l)|}{\pi_j(l)} + \\
&\quad + \frac{|v_2 (B_l^k(2, j) - B_l^k(1, j)) + \dots + v_N (B_l^k(N, j) - B_l^k(1, j))|}{\pi_j(l)}.
\end{aligned}$$

Pelo teorema anterior, para l, j fixos, temos

$$\begin{aligned}
\max_{i, m} \frac{|B_l^k(i, j) - B_l^k(m, j)|}{\pi_j(l)} &\leq \max_{i, j, l} \left(\frac{|B_l^k(i, j) - \pi_j(l)|}{\pi_j(l)} + \frac{|B_l^k(m, j) - \pi_j(l)|}{\pi_j(l)} \right) \\
&\leq 2 \frac{(\lambda^*)^{k-1} (N-1)}{\min_j (\pi_j(n))}
\end{aligned}$$

e uma vez que $0 \leq v_2 + \dots + v_N \leq 1$ concluímos

$$\max_{i, j, l} \frac{|(vB_l^k)(i, j) - \pi_j(l)|}{\pi_j(l)} \leq \frac{3(\lambda^*)^{k-1} (N-1)}{\min_j \pi_j(n)}.$$

□

2.3 Convergência em cadeias de Markov não homogêneas

Em grande parte das referências bibliográficas na área das Cadeias de Markov, apesar de referências à existência de cadeias não homogêneas, rapidamente o estudo é restringido a cadeias homogêneas. Muitas das questões respondidas para Cadeias de Markov homogêneas permanecem em aberto no caso não homogêneo, mesmo em exemplos simples.

Cadeias de Markov não homogêneas serão bons candidatos por exemplo para modelos da emigração entre países ao longo do tempo. A *heterogeneidade* temporal pode resultar também de factores sazonais que darão origem a cadeias periódicas. Estamos interessados na questão da convergência para o equilíbrio, sobretudo em quantificar a velocidade dessa convergência. Uma revisão da literatura diz-nos que nos anos 50, o demógrafo J. Hajnal em dois artigos (1956 e 1958), deu uma forte contribuição no estudo de propriedades ergódicas em cadeias de Markov não homogêneas, em especial no desenvolvimento de ferramentas para a análise da ergodicidade fraca ([Sen14]). Para além disso, uma das primeiras referências sobre o comportamento assintótico de cadeias de Markov não homogêneas

deve-se a E. Borel em [Bor72]. No artigo [SCZ11] encontramos uma colecção dos resultados conhecidos, e respectivas referências, para cadeias não homogêneas, do ponto de vista da ergodicidade e do comportamento assintótico, — propriedades essenciais para os objectivos deste trabalho.

Para cadeias de Markov homogêneas ($P_t = P, \forall t$) sabemos que P é ergódica se e só se a cadeia converge componente a componente para uma matriz W com linhas idênticas, *perdendo a memória*, isto é, ao fim de um número grande de iteradas não conseguimos restituir o percurso da cadeia e conhecer o estado inicial. Assim, a perda de memória é uma propriedade inerente à ergodicidade. Porém, no caso não homogêneo podemos estabelecer duas propriedades distintas.

Nota. No contexto das CMDF homogêneas a convergência em norma (em ℓ^1 , ℓ^2 , do máximo, ...) ou a convergência componente a componente são equivalentes. No caso das CMDF não homogêneas surgem algumas diferenças (a norma em ℓ^1 e a norma do sup relativo não são equivalentes, sendo a segunda mais forte, ver [SCZ09] para discussão e exemplos sobre esse assunto),

Dado um vector de distribuição de probabilidade u , a sua norma em ℓ^1 é dada por

$$\|u\| = \sum_{i=1}^N |u_i|.$$

Temos as seguintes definições (ver [IM76] ou [Bré99]):

Definição 2.3.1. *Uma cadeia de Markov não homogênea $(P_n)_{n \in \mathbb{N}}$ diz-se fracamente ergódica se, para qualquer m :*

$$\lim_{k \rightarrow \infty} \sup_{v, w} \|vP_m \cdots P_k - wP_m \cdots P_k\| = 0 \quad (2.8)$$

onde v, w são quaisquer vectores que representam densidades de probabilidade, ou seja, todas as componentes são não negativas com soma 1.

Intuitivamente, uma cadeia fracamente ergódica apresenta perda de memória mas não se tem necessariamente $vP_m \cdots P_k$ próximo de $vP_m \cdots P_{k+1}$, e portanto não há convergência para um estado de equilíbrio, esse será o caso de uma cadeia fortemente ergódica:

Definição 2.3.2. *Uma cadeia de Markov não homogênea $(P_n)_{n \in \mathbb{N}}$ diz-se fortemente ergódica se existe um vector v com $\sum_i v_i = 1$ e $v_i \geq 0, \forall i \in \{1 \dots N\}$, tal que, para todo o m*

$$\lim_{k \rightarrow \infty} \sup_w \|wP_m \cdots P_k - v\| = 0 \quad (2.9)$$

onde w representa uma densidade de probabilidade, ou seja, tem componentes não negativas com soma 1.

Numa cadeia homogénea estes dois tipos de ergodicidade coincidem ([IM76]), isto é, não existem cadeias de Markov homogéneas que exibam perda de memória sem convergência para um estado de equilíbrio. Numa cadeia homogénea em dimensão finita os dois tipos de ergodicidade, fraca e forte, coincidem com a ergodicidade definida em 2.1.4.

As definições apresentadas acima são, em geral, difíceis de testar directamente. Existem então vários critérios para provar a ergodicidade fraca e também alguns para a ergodicidade forte (ver por exemplo [IM76], [Bré99], [NS98]). Alguns desses critérios envolvem o chamado *coeficiente ergódico*, introduzido por Dobrushin [Dob56], que fornece uma medida da *distância entre uma dada matriz e uma matriz constante*. No nosso caso, conseguimos aplicar as definições directamente para provar a ergodicidade fraca, portanto usaremos o coeficiente ergódico enquanto estimador da taxa de convergência.

Definição 2.3.3. *Dada uma matriz estocástica P de dimensão $N \times N$, o coeficiente ergódico $\delta(P)$ é definido por*

$$\delta(P) = \frac{1}{2} \sup_{i,j} \sum_{k=1}^N |P(i,k) - P(j,k)|$$

Na Secção 2.3.3 voltaremos ao coeficiente ergódico. Apresentamos agora a noção de tempo de *mixing*, uma medida do tempo que a cadeia demora a atingir uma certa distância ao equilíbrio.

2.3.1 Tempo de *mixing*

A distância de variação total entre dois vectores de distribuição de probabilidade μ e ν num espaço de probabilidade Ω é dada por

$$\|\mu - \nu\|_{TV} = \max_{A \subset \Omega} |\mu(A) - \nu(A)|.$$

Num espaço de dimensão finita N , dados dois vectores de distribuição de probabilidade u e v , a distância de variação total traduz-se por (ver por exemplo [LPW09]):

$$\|u - v\|_{TV} = \frac{1}{2} \sum_{i=1}^N |v_i - u_i|.$$

Dada uma cadeia de Markov P , de dimensão finita N , ergódica, uma distribuição de probabilidade inicial v e a distribuição de equilíbrio π , definimos a distância, $d(t)$, entre o estado da cadeia ao fim de t passos e o equilíbrio π por

$$d(t) = \|vP^t - \pi\|_{TV} = \frac{1}{2} \sum_{i=1}^N |(vP^t)_i - \pi_i|$$

Dada uma CMDF homogênea, com matriz de transição P , o tempo de convergência ou tempo de *mixing*, $t_{\text{mix}}(\varepsilon)$, é um parâmetro definido por:

$$t_{\text{mix}}(\varepsilon) = \max_v \min_t \{t : d(t) < \varepsilon\},$$

onde v percorre todos os vectores de distribuição de probabilidade inicial (ver [LPW09]).

O tempo de *mixing* aparece como medida do número de iteradas necessário para atingir uma certa distância ao equilíbrio π . Definimos também

$$t_{\text{mix}} = t_{\text{mix}}\left(\frac{1}{4}\right).$$

A segunda definição é pouco natural mas a escolha de $\varepsilon = \frac{1}{4}$, sendo aleatória, (por exemplo em [FR07b] a escolha é $\frac{1}{e}$), tem uma justificação parcial: de modo que a desigualdade (provada em [LPW09])

$$d(lt_{\text{mix}}(\varepsilon)) \leq (2\varepsilon)^l$$

não seja trivial, a escolha de ε terá de ser $\varepsilon < \frac{1}{2}$. Em particular para $\varepsilon = \frac{1}{4}$ obtemos

$$d(lt_{\text{mix}}) \leq 2^{-l}$$

e resulta a relação entre t_{mix} e $t_{\text{mix}}(\varepsilon)$

$$t_{\text{mix}}(\varepsilon) \leq \log_2(\varepsilon^{-1}) t_{\text{mix}}.$$

Assim, o conhecimento de t_{mix} fornece imediatamente uma medida da velocidade de convergência para o equilíbrio. Naturalmente o tempo de *mixing* e a distância ao equilíbrio em função do número de iteradas estão estreitamente relacionados, no caso dos grafos autónomos/CMDF homogêneas. Sendo nosso objectivo definir uma versão de tempo de *mixing* para grafos não autónomos iremos primeiro definir a distância a usar. Na literatura encontrámos ferramentas que desempenham, essencialmente, o papel que queremos dar ao tempo de *mixing*, nomeadamente em [SCZ07] define-se tempo de *merging*, no qual nos vamos inspirar.

Definição 2.3.4. *Dada uma sucessão de matrizes de transição estocásticas de dimensão $N \times N$ chamamos tempo de merging para $\varepsilon > 0$*

$$t_{\text{mer}}(\varepsilon) = \inf_n \{ \max_{u,v} \|uP_1 \cdots P_n - vP_1 \cdots P_n\|_{TV} < \varepsilon \} \quad (2.10)$$

onde a norma, da variação total, é dada por $\|u - v\|_{TV} = \frac{1}{2} \sum_{i=1}^N |v_i - u_i|$.

O tempo de *merging* para ε fornece o menor número de passos da cadeia que garante uma distância inferior a ε entre os estados alcançados pela cadeia com vectores iniciais u

e v , quaisquer que sejam u e v .

2.3.2 Interpretação da convergência em cadeias de Markov não homogêneas

Dado um grafo não autónomo $\mathcal{G} = (V, (E_i)_{i=0}^\infty)$, p -periódico, consideremos a sucessão de matrizes estocásticas associadas $(P_i)_{i=0}^\infty$. Esta sucessão de matrizes define uma cadeia de Markov discreta, em espaço finito, não homogênea. Os resultados obtidos podem assim ser aplicados à teoria das cadeias de Markov não homogêneas.

De seguida justificamos que o resultado obtido no Teorema 2.2.4 garante que a cadeia não homogênea, definida através de uma sucessão de matrizes de adjacência ergódicas associadas a um grafo não autónomo p -periódico, é fracamente ergódica mas não é fortemente ergódica, segundo as definições apresentadas na Secção 2.3.

Teorema 2.3.5. *Consideremos um grafo não autónomo $\mathcal{G} = (V, (E_i)_{i=0}^\infty)$, p -periódico tal que cada uma das matrizes P_i , da sucessão de matrizes estocásticas $(P_i)_{i=0}^\infty$, é ergódica. A cadeia de Markov não homogênea definida por $(P_i)_{i=0}^\infty$ é fracamente ergódica.*

Demonstração. Consideremos dois vectores v, w , ambos com componentes não negativas, de norma 1. Fixado m , temos

$$P_m \cdots P_k = P_m \cdots P_{m+(k-m) \bmod p} P_{m+(k-m) \bmod p+1} \cdots P_k = P_{m \bmod p} \cdots P_{k \bmod p} B_{(k+1) \bmod p}^{\lfloor \frac{k-m}{p} \rfloor}$$

onde $\lfloor \frac{k-m}{p} \rfloor$ é o quociente da divisão inteira de $k-m$ por p . Ou seja, temos o membro esquerdo escrito à custa de produtos das matrizes B_i . Donde, fazendo $v_1 = v P_{m \bmod p} \cdots P_{k \bmod p}$ e $w_1 = w P_{m \bmod p} \cdots P_{k \bmod p}$, temos

$$\begin{aligned} & \limsup_{k \rightarrow \infty} \sup_{v, w} \|v P_m \cdots P_k - w P_m \cdots P_k\| \\ &= \limsup_{k \rightarrow \infty} \sup_{v, w} \left\| v_1 B_{(k+1) \bmod p}^{\lfloor \frac{k-m}{p} \rfloor} - w_1 B_{(k+1) \bmod p}^{\lfloor \frac{k-m}{p} \rfloor} \right\| \end{aligned}$$

Mas, qualquer que seja v_1, w_1 ,

$$\begin{aligned} & \left\| v_1 B_{(k+1) \bmod p}^{\lfloor \frac{k-m}{p} \rfloor} - w_1 B_{(k+1) \bmod p}^{\lfloor \frac{k-m}{p} \rfloor} \right\| \\ & \leq \left\| v_1 B_{(k+1) \bmod p}^{\lfloor \frac{k-m}{p} \rfloor} - \pi((k+1) \bmod p) \right\| + \left\| w_1 B_{(k+1) \bmod p}^{\lfloor \frac{k-m}{p} \rfloor} - \pi((k+1) \bmod p) \right\| \end{aligned}$$

então, tomando a norma do máximo e usando o Corolário 2.2.5, o limite anterior existe e é 0. \square

A condição de ergodicidade forte é muito restritiva e em geral não é verificada para cadeias de Markov não homogêneas periódicas ainda que cada matriz de adjacência seja

ergódica, uma vez que não se tem necessariamente $\pi(j) = \pi(i)$ para $i \neq j$.

2.3.3 Uma estimativa sem valores próprios

No início desta secção introduzimos (segundo [IM76]) o coeficiente ergódico de uma matriz, o qual será usado na obtenção de uma estimativa superior para a distância ao equilíbrio, em função do número de iteradas. Vamos usar a distância entre dois vectores de distribuição de probabilidade dada pela distância de variação total:

$$\|u - v\|_{TV} = \frac{1}{2} \sum_{i=1}^N |u_i - v_i|$$

Temos o seguinte resultado para cadeias de Markov homogêneas:

Teorema 2.3.6. *Dada uma matriz P estocástica e dois vectores de distribuição de probabilidade u e v , então, para qualquer $k \in \mathbb{N}$*

$$\|uP^k - vP^k\|_{TV} \leq \|u - v\|_{TV} \delta(P)^k. \quad (2.11)$$

A prova utiliza a propriedade $\delta(PQ) \leq \delta(P)\delta(Q)$ e pode ser encontrada em [Bré99]. A propriedade que torna o resultado relevante é que o coeficiente ergódico para qualquer matriz estocástica satisfaz $0 \leq \delta(P) \leq 1$.

Pretendemos encontrar uma versão deste resultado no contexto das cadeias de Markov não homogêneas correspondentes a grafos não autónomos periódicos, no espírito do trabalho apresentado na Secção 2.2. Nesse sentido apresentamos o seguinte teorema:

Teorema 2.3.7. *Dado um grafo não autónomo \mathcal{G} , p -periódico, com matrizes estocásticas P_0, \dots, P_{p-1} ergódicas e correspondentes matrizes produto B_0, \dots, B_{p-1} definidas pela igualdade (2.2), com vector equilíbrio $\Pi = (\pi(0), \dots, \pi(p-1))$, temos, para qualquer vector de distribuição inicial u*

$$\|uB_i^n - \pi(i)\| \leq \|u - \pi(i)\| \delta(B_i)^n. \quad (2.12)$$

O teorema anterior oferece uma estimativa da distância ao equilíbrio para qualquer matriz B_i . Tomando o máximo em i ficamos a conhecer uma majoração da distância a que a cadeia se encontra do seu vector equilíbrio Π , em função dos coeficientes ergódicos de cada uma das matrizes B_i .

Demonstração. Na desigualdade (2.11) tomamos $v = \pi(i)$. Uma vez que $\pi(i)$ é vector próprio esquerdo de B_i temos $\pi(i) = \pi(i)B_i$ e logo $\pi(i) = \pi(i)B_i^n$. Donde

$$\|uB_i^n - \pi(i)\| \leq \|uB_i^n - \pi(i)\| \delta(B_i)^n.$$

□

Corolário 2.3.8. *Dado um grafo não autônomo \mathcal{G} , p -periódico, e um vector u ambos nas condições do Teorema 2.3.7, temos, para qualquer $j, n \in \mathbb{N}$ e $i = n \bmod p$*

$$\|uP_j \cdots P_{j+n-1} - \pi(i)\| \leq \|v - \pi(i)\| \delta(B_i)^{\lfloor \frac{n-j}{p} \rfloor},$$

onde $v = uP_{j \bmod p} \cdots P_{i-1}$.

Demonstração. Basta escrever $P_j \cdots P_{j+n-1}$ à custa do produto B_i adequado, ou seja

$$uP_j \cdots P_{j+n-1} = uP_{j \bmod p} \cdots P_{i-1} P_i \cdots P_i \cdots P_i \cdots P_{i-1} \cdots = uP_{j \bmod p} \cdots P_{i-1} B_i^{\lfloor \frac{n-j}{p} \rfloor}$$

□

Para finalizar esta secção propomos uma definição de tempo de *mixing* para um grafo não autônomo periódico e discutimos a sua aplicação:

Definição 2.3.9. *Dado um grafo \mathcal{G} não autônomo, p -periódico, com p matrizes ergódicas estocásticas $(P_i)_{i \in \{0, \dots, p-1\}}$ e matrizes $(B_i)_{i \in \{0, \dots, p-1\}}$, o tempo de *mixing* para ε é dado por*

$$t_{mix}(\varepsilon) = \max_l \inf_n \{n : \max_u \|uB_l^n - \pi(l)\| < \varepsilon\},$$

onde u percorre todos os vectores de distribuição de probabilidade e a norma é a da variação total.

Independentemente da distribuição inicial, a partir de $t_{mix}(\varepsilon)$ iteradas/passos sabemos que a distribuição no grafo se encontra a uma distância de todos os $\pi(i)$, inferior a ε . Claramente o tempo de *mixing* é finito quando a CMDF não homogénea formada pelas matrizes de transição é fracamente ergódica.

2.4 Aplicação aos sistemas dinâmicos

2.4.1 O Operador de Perron-Frobenius

Consideremos uma função $\tau : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$. No estudo dos sistemas dinâmicos através da teoria da medida, em vez de seguirmos o comportamento da órbita de um ponto $x \in [a, b]$ pela função τ ao longo do tempo,

$$x, \tau(x), \tau^2(x), \dots$$

consideramos uma certa densidade de probabilidade inicial f e estudamos a evolução das densidades de probabilidade pela iteração de τ . Assim, a cada função τ associamos um

operador, P_τ , — operador de Perron-Frobenius, — definido no espaço das densidades de probabilidade e estudamos as suas iteradas,

$$f, P_\tau f, P_\tau^2 f \dots$$

(ver [BG97] e [LM94]). Uma medida, μ , invariante para τ , (isto é, tal que $\mu(\tau^{-1}(A)) = \mu(A), \forall A \subset [a, b]$), corresponde assim a um ponto fixo do operador P_τ . Estas relações permitem usar as ferramentas da análise funcional e teoria da medida para estudar o comportamento do sistema ao fim de um número grande de iteradas. De particular interesse é o caso em que a função τ satisfaz a propriedade de Markov e é linear por troços numa partição \mathcal{P} do intervalo $[a, b]$. Nesse caso o operador de Perron-Frobenius P_τ tem dimensão finita e pode ser representado por uma matriz quadrada P de dimensão N igual ao número de intervalos da partição, que actua no espaço S das densidades de probabilidade constantes por troços na partição \mathcal{P} . Obviamente as funções em S podem ser representadas por vectores de dimensão N e a aplicação sucessiva do operador P corresponde à multiplicação sucessiva da matriz P por si própria pelo que recuperamos as ferramentas desenvolvidas neste capítulo. A restrição P do operador de Perron-Frobenius ao espaço das densidades de probabilidade constantes por troços, permite explorar propriedades do operador: todos os valores próprios da matriz pertencem ao espectro do operador e o maior deles em módulo é também o maior valor próprio em módulo do operador ([DFS00]), o que justifica a sua relação com a taxa de convergência para o equilíbrio. Em [FR07a] o espectro do operador de Perron-Frobenius é estudado em famílias de funções unimodais e bimodais obtendo-se, para essas famílias, uma relação com a taxa de *mixing*.

2.4.2 Exemplo

Um caso particular que motivou este trabalho é o dos *subshifts* de tipo finito obtidos a partir de sistemas dinâmicos gerados pela iteração de uma função $x_{n+1} = F(x_n)$ onde $F : I \subset \mathbb{R} \rightarrow I$ satisfaz a propriedade de Markov (ver [LM95]), ou seja, se existe uma partição finita I_1, \dots, I_N de um intervalo I tal que a imagem $F(I_i)$ de cada intervalo I_i é uma união de intervalos $\cup_j I_j$. Consideremos um grafo tal que cada vértice v_i corresponde a um intervalo I_i e o peso e_{ij} de cada aresta (i, j) é definido por $e_{ij} = \mu(F^{-1}(I_j) \cap I_i)$, onde μ é a medida de Lebesgue em \mathbb{R} . O peso e_{ij} é a medida do conjunto de pontos que estão em I_i e cujas imagens por F estão em I_j .

De acordo com [KS96], dado um espaço métrico compacto X e uma sucessão de funções $(F_i)_{i=1}^\infty$ com $F_i : X \rightarrow X, i \in \mathbb{N}$, temos um sistema dinâmico não autónomo $(X, (F_i)_{i=1}^\infty)$ definido por $F_i^n = F_{i+n-1} \circ \dots \circ F_{i+1} \circ F_i$. O sistema não autónomo é periódico com período p se $F_i = F_{i+p}, \forall i$.

Consideremos o sistema não autónomo de período 3 dado por $(I, (F_i)_{i=1}^\infty)$, onde $F_i = F_{i+3}, \forall i$ e F_0, F_1, F_2 são as funções apresentadas na Figura 2.1. Estas funções, do in-

tervalo $I = [0, 10]$ satisfazem a propriedade de Markov e partilham a mesma partição $I_1 = [0, 3], I_2 = [3, 5], I_3 = [5, 10]$.

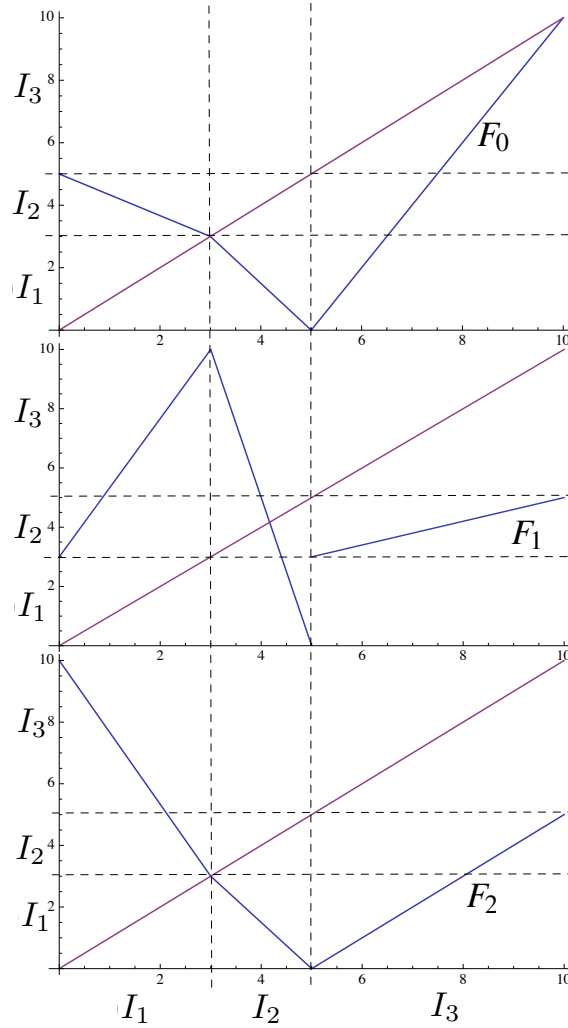


Figura 2.1: Gráficos de F_0, F_1, F_2 a azul e os intervalos da partição de Markov I_1, I_2, I_3 .

O grafo não autônomo de período 3 está representado na Figura 2.2.

Os pesos, $e_{jl}(i) = \mu(F_i^{-1}(I_l) \cap I_j)$, $i = 0, 1, 2, j, l = 1, 2, 3$, são dados pelas seguintes matrizes

$$A_0 = [e_{ij}(1)] = \begin{bmatrix} 0 & 3 & 0 \\ 2 & 0 & 0 \\ \frac{3}{2} & 1 & \frac{5}{2} \end{bmatrix} \quad A_1 = [e_{ij}(2)] = \begin{bmatrix} 0 & \frac{15}{7} & \frac{6}{7} \\ \frac{3}{5} & \frac{2}{5} & 1 \\ 0 & 5 & 0 \end{bmatrix} \quad A_2 = [e_{ij}(3)] = \begin{bmatrix} 0 & \frac{15}{7} & \frac{6}{7} \\ 2 & 0 & 0 \\ 3 & 2 & 0 \end{bmatrix}$$

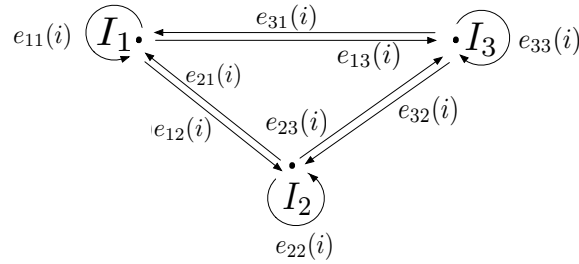


Figura 2.2: Grafo não autônomo de período 3 associado ao sistema.

e as matrizes estocásticas correspondentes:

$$P_0 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ \frac{3}{10} & \frac{1}{5} & \frac{1}{2} \end{bmatrix} \quad P_1 = \begin{bmatrix} 0 & \frac{5}{7} & \frac{2}{7} \\ \frac{3}{10} & \frac{1}{5} & \frac{1}{2} \\ 0 & 1 & 0 \end{bmatrix} \quad P_2 = \begin{bmatrix} 0 & \frac{5}{7} & \frac{2}{7} \\ 1 & 0 & 0 \\ \frac{3}{5} & \frac{2}{5} & 0 \end{bmatrix}$$

Na notação das secções anteriores temos

$$B_0 = \begin{bmatrix} \frac{1}{2} & \frac{29}{70} & \frac{3}{35} \\ \frac{31}{35} & \frac{4}{35} & 0 \\ \frac{303}{350} & \frac{41}{350} & \frac{3}{175} \end{bmatrix}, \quad B_1 = \begin{bmatrix} \frac{4}{35} & \frac{31}{35} & 0 \\ \frac{11}{25} & \frac{181}{350} & \frac{3}{70} \\ 0 & 1 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} \frac{3}{175} & \frac{127}{175} & \frac{9}{35} \\ \frac{3}{10} & \frac{1}{5} & \frac{1}{2} \\ \frac{9}{50} & \frac{71}{175} & \frac{29}{70} \end{bmatrix}.$$

Da definição apresentada para a reversibilidade de uma matriz, concluímos facilmente que B_i é reversível se DB_iD^{-1} é uma matriz simétrica, onde D representa a matriz diagonal com entradas $D(j, j) = \sqrt{\pi_j(i)}$. De facto temos

$$DB_1D^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{3} & 0 \\ \frac{\sqrt{3}}{3} & \frac{1}{6} & \frac{\sqrt{6}}{6} \\ 0 & \frac{\sqrt{6}}{6} & 0 \end{bmatrix}$$

então B_1 é reversível. $\pi(1) = [\frac{1078}{3341}, \frac{2170}{3341}, \frac{93}{3341}]$ e $\lambda^* = \frac{129 + \sqrt{19041}}{700} \approx 0.3814$, pelo que $\varepsilon(k) = \frac{(\lambda^*)^{k-1}(N-1)}{\min_i \pi_i(1)} \approx 188.377(0.3814)^k$. A Figura 2.3 mostra os gráficos de $\delta(k)$ e $\varepsilon(k)$. Podemos confirmar que $\varepsilon(k)$ fornece uma estimativa superior de $\delta(k)$.

A distribuição de equilíbrio é

$$\Pi \approx ((0.6384, 0.3060, 0.0557), (0.3227, 0.6495, 0.0278), (0.194852, 0.388207, 0.416941)).$$

Tomemos por exemplo um sistema dado pela iteração sucessiva das três funções começando com f_0 , isto é, $f_0, f_1 \circ f_0, f_2 \circ f_1 \circ f_0, \dots$, representado por $\mathbf{P}_0^k = P_0 P_1 P_2 \dots P_{k \bmod 3}^k$, e que se encontra inicialmente com probabilidade 1 no intervalo I_1 . Ao fim de 41 iteradas,

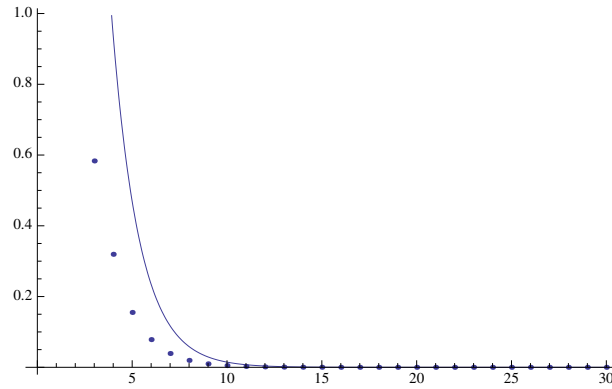


Figura 2.3: A linha a cheio representa $\varepsilon(k)$ e o gráfico ponteadado representa $\delta(k)$.

esperamos que se distribua nos intervalos da partição com densidade de aproximadamente $\left[\begin{array}{ccc} 0.194852 & 0.388278 & 0.416941 \end{array} \right]$ e com erro maximal conhecido. Com efeito,

$$\left[\begin{array}{ccc} 1 & 0 & 0 \end{array} \right] P_0 P_1 B_2^{13} = \left[\begin{array}{ccc} \frac{1}{10} & \frac{1}{5} & \frac{1}{2} \end{array} \right] \left[\begin{array}{ccc} \frac{3}{175} & \frac{127}{175} & \frac{9}{35} \\ \frac{3}{10} & \frac{1}{5} & \frac{1}{2} \\ \frac{9}{50} & \frac{71}{175} & \frac{29}{70} \end{array} \right]^{13} = \left[\begin{array}{ccc} 0.194851 & 0.388208 & 0.416941 \end{array} \right]$$

Mais, pelo Corolário 2.2.11, temos

$$\max_j \frac{|v_0 \mathbf{P}_0^{41} - \pi_j(2)|}{\pi_j(2)} \leq \frac{6(\lambda^*)^{12}}{\min_j \pi_j(1)} \approx \frac{6(0.3814)^{12}}{0.0278} \approx 0.002045.$$

que constitui uma majoração do erro relativo da densidade de probabilidade ao fim de 41 iterações em relação ao equilíbrio.

Nota. Os resultados desta secção foram obtidos usando o programa Mathematica 8.

Capítulo 3

Matrizes de adjacência não reversíveis

3.1 Motivação

Neste capítulo vamos estender os resultados apresentados no Capítulo 2 para grafos autônomos e não autônomos periódicos retirando a hipótese de reversibilidade. Se no caso autônomo a reversibilidade é uma condição restritiva, no caso não autônomo torna-se ainda mais pertinente encontrar estimativas para a velocidade de convergência para o equilíbrio que não exijam reversibilidade das matrizes envolvidas. É nesse campo que este capítulo pretende dar um contributo.

Existem na literatura estimativas para a distância ao equilíbrio em função do número de iterações de cadeias de Markov não reversíveis, (ver [Fil91] e [SCZ09]) que no entanto usam o segundo valor próprio de uma matriz reversível relacionada com a matriz P , que pode não garantir boas estimativas, como veremos mais à frente.

Por outro lado, as estimativas para a taxa de aproximação ao equilíbrio em grafos autônomos e não autônomos que usam o segundo maior valor próprio em módulo da matriz de adjacência estocástica, limitam-se a matrizes ergódicas reversíveis, isto é, matrizes P , com distribuição de equilíbrio π , que satisfazem

$$\pi_i p(i, j) = \pi_j p(j, i).$$

Neste capítulo apresentamos estimativas para o erro na aproximação ao equilíbrio, no espírito dos resultados apresentados no Capítulo 2, para grafos autônomos em que a matriz de adjacência é ergódica. Além disso, estendemos os resultados de [SSF14a], onde introduzimos a noção de distribuição de equilíbrio em grafos não autônomos periódicos e demos uma estimativa para a velocidade de aproximação ao equilíbrio, retirando agora a hipótese de reversibilidade. Concretamente, dado um grafo não autônomo p -periódico

$\mathcal{G} = (V, (E_i)_{i=0}^\infty)$ com matrizes de adjacência $(P_i)_{i=0}^\infty$ consideramos o produto cíclico $B_i = P_i \cdots P_{(i+p-1) \bmod p}$, com $i = 1, \dots, p$ e $p \in \mathbb{N}$. Vimos que, se uma das matrizes B_i é reversível, então existe uma estimativa para o erro maximal na aproximação ao equilíbrio usando o segundo maior valor próprio em módulo, λ^* . Apresentaremos uma versão desse resultado exigindo apenas a ergodicidade de cada uma das matrizes B_i , com $i = 0, \dots, p-1$.

Em ambos os contextos, autónomo e não autónomo, a restrição da reversibilidade levanta alguns problemas: nas aplicações em que estamos interessados aparecem matrizes de adjacência não necessariamente reversíveis, por exemplo as matrizes que resultam da restrição do operador de Perron-Frobenius a um espaço de dimensão finita (ver Secção 2.4 e [BG97]); Por outro lado, no caso não autónomo, como escolher matrizes estocásticas $P_i, i = 0, \dots, p-1$ tais que um dos produtos cíclicos $B_i, i = 0, \dots, p-1$ seja reversível?

Os resultados deste capítulo aplicam-se tanto a grafos autónomos como não autónomos e configuram estimativas superiores para a velocidade de convergência para o equilíbrio, *grosso modo*, dependendo do segundo maior valor próprio em módulo e da dimensão dos blocos de Jordan a ele associados.

Na segunda secção apresentamos algumas estimativas conhecidas, na terceira secção demonstramos os resultados de [SSF14a] para o caso autónomo, distinguindo matrizes diagonalizáveis e não diagonalizáveis. Na secção seguinte discutimos o caso não autónomo e na última secção apresentamos alguns exemplos.

Nota. Existem diferentes hipóteses de associar uma cadeia de Markov a um grafo. Por exemplo, na teoria ergódica, se a matriz de adjacência A é irredutível existe uma forma canónica de associar probabilidades às arestas, construindo a chamada matriz de Parry associada à medida de máxima entropia (ver [Wal82] e [LM95]). À semelhança do capítulo anterior, definimos a matriz P atribuindo a cada entrada (i, j) o valor $p(i, j) = \frac{A(i, j)}{\sum_{k=1}^N A(i, k)}$, mas sublinhamos que os resultados apresentados são válidos para quaisquer matrizes estocásticas, logo, para a todas as cadeias de Markov associadas a um grafo.

3.2 Estimativas conhecidas

3.2.1 Caso autónomo

No caso não reversível, é possível obter estimativas para a distância ao equilíbrio usando os valores próprios de uma matriz reversível construída a partir de P ([Fil91], [Bré99]). Com efeito, consideremos uma matriz P ergódica, não reversível. Definimos a matriz \tilde{P} por

$$\tilde{P}(i, j) = \frac{\pi_j P(j, i)}{\pi_i}$$

onde π é o vector equilíbrio. Matricialmente temos

$$\tilde{P} = D^{-1} P^T D$$

em que D é a matriz diagonal

$$D = \begin{bmatrix} \sqrt{\pi_1(0)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{\pi_N(0)} \end{bmatrix}.$$

A matriz $M = P\tilde{P}$ diz-se matriz de *reversibilização* multiplicativa de P .

Lema 3.2.1. *A matriz M tem vector de equilíbrio π e é reversível.*

Demonstração. π é distribuição de equilíbrio de $P\tilde{P}$ uma vez que

$$\pi P\tilde{P} = \pi D^{-1}P^T D = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} P^T D = \begin{bmatrix} 1 & \dots & 1 \end{bmatrix} D = \pi$$

dado que π é a distribuição de equilíbrio de P e P^T é estocástica por colunas.

Quanto à reversibilidade, vejamos que

$$M = P\tilde{P} = PD^{-1}P^T D,$$

logo

$$M(i, j) = \sum_k \frac{P(i, k)}{\pi_k} P(j, k) \pi_j$$

donde

$$\pi_i M(i, j) = \pi_i \sum_k \frac{P(i, k)}{\pi_k} \pi_j P(j, k) = \pi_j \sum_k P(j, k) \frac{P(i, k)}{\pi_k} \pi_i = \pi_j M(j, i)$$

□

Definimos ainda a seguinte distância à distribuição de equilíbrio π :

$$d_2(v, \pi) = \|v - \pi\|_{2, \pi} = \sum_{i=1}^N \frac{(v_i - \pi_i)^2}{\pi_i}.$$

Temos a seguinte estimativa, devida a Fill [Fil91]:

Teorema 3.2.2. *Dada uma matriz P ergódica, com distribuição de equilíbrio π , seja $\lambda(M)$ o segundo maior valor próprio em módulo da matriz M . Então, dado um vector de distribuição de probabilidade v ,*

$$4\|vP^k - \pi\|_{TV}^2 \leq \lambda(M)^k \|v - \pi\|_{2, \pi} \quad (3.1)$$

A demonstração baseia-se num resultado obtido por M. Mihail em 1989 que usa argumentos da linguagem combinatória e probabilística que não nos parece pertinente incluir

neste trabalho. Em [SCZ11], uma desigualdade equivalente a 3.1 também é referida como ponto de partida para uma generalização ao caso não homogéneo.

Nota. O Teorema 3.2.2 apresenta uma estimativa à custa dos valores singulares de P , o que, por um lado constitui uma estimativa alternativa à proposta em [SSF14b], mas que pode tornar-se desinteressante uma vez que, ao contrário de P , o segundo maior valor próprio de M pode ser 1, (essencialmente porque M pode não ser irredutível), o que confere maior relevância aos resultados obtidos na Secção 3.3. Outra diferença importante é que as nossas estimativas não supõem a existência da estrutura de espaço de Hilbert.

Apresentamos em seguida alguns comentários pertinentes:

- O segundo maior valor próprio em módulo da matriz M é precisamente o segundo maior valor próprio, por se tratar de uma matriz simétrica semidefinida positiva.
- Pode considerar-se a chamada matriz de *reversibilização* aditiva, $A = \frac{P+\tilde{P}}{2}$, e obter-se a partir do teorema anterior, sob hipóteses mais fortes, uma estimativa em função do segundo maior valor próprio de A . Em [Fil91] é discutida a utilização do segundo valor próprio da matriz A .

De uma outra perspectiva, interpretemos a matriz P como um operador no espaço de Hilbert \mathbb{R}^N , com o produto interno definido por $(u, v)_\pi = \sum_{i=1}^N u_i v_i \pi_i$. Assim, M é uma matriz autoadjunta e as raízes quadradas dos seus valores próprios são os valores singulares de P . (Se P é autoadjunta, $M = P^2$, e os valores singulares de M coincidem com os valores próprios de P .)

Repare-se que os valores singulares de uma matriz P , ao contrário dos valores próprios, dependem do espaço de Hilbert considerado, que neste caso está associado ao vector de equilíbrio π . Considerando operadores em espaços de Hilbert associados aos equilíbrios das diferentes matrizes da cadeia não homogénea, em [SCZ09] obtém-se uma estimativa do tipo apresentado no Teorema 3.2.2, usando os *valores singulares* das matrizes P_i da cadeia de Markov. Com esta interpretação podemos usar as ferramentas robustas dos espaços de Hilbert, como veremos de seguida.

3.2.2 Caso não autónomo

Nesta secção fazemos a demonstração do resultado de [SCZ09] no caso das CMDF não homogéneas, associadas a grafos não autónomos. Começamos por enunciar um resultado da Teoria dos operadores que, *grosso modo*, generaliza a decomposição de uma matriz no espaço dos seus vectores próprios quando estes formam uma base do espaço e pode ser encontrado em [Rin71].

Teorema 3.2.3 (Decomposição em valores singulares). *Consideremos dois espaços de Hilbert reais, H_1 e H_2 de dimensão N . Seja $P : H_1 \rightarrow H_2$ um operador compacto.*

Então existem bases ortonormadas, ϕ_i de H_1 e ψ_i de H_2 , e valores reais não negativos σ_i tais que $P\phi_i = \sigma_i\psi_i$ e $P^*\psi_i = \sigma_i\phi_i$.

Os valores σ_i são os valores singulares de P e coincidem com as raízes quadradas dos valores próprios de PP^* e P^*P ¹. No nosso contexto o operador P é representado por uma matriz $N \times N$ e a matriz P^* é a matriz adjunta de P . Consideramos a seguinte ordenação dos valores singulares:

$$1 = \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_N.$$

O próximo lema é a base da demonstração do teorema seguinte.

Lema 3.2.4. *Dada uma matriz P , de dimensão $N \times N$, estocástica, consideremos o operador $P : \ell^2(\pi(0)) \rightarrow \ell^2(\pi(1))$ onde $\ell^2(\pi(k))$ é o espaço de Hilbert \mathbb{R}^N , munido do produto interno definido entre dois vectores $v, w \in \mathbb{R}^N$ por $(v, w)_\pi = \sum_{i=1}^N v_i w_i \pi_i(k)$, sendo $\pi(k)$, $k = 0, 1$ vectores de distribuição de probabilidade. Temos*

$$d_2(P(l, \cdot), \pi(1))^2 = \sum_{i=2}^N \sigma_i^2 (\psi_i)_l^2 \quad (3.2)$$

onde $(\psi_i)_{i \in \{1, \dots, N\}}$ é a base ortonormada correspondente aos valores singulares $(\sigma_i)_{i \in \{1, \dots, N\}}$ de P .

Demonstração. Consideremos uma matriz P estocástica. Interessa-nos escrever a distância d_2 à custa do produto interno $(\cdot, \cdot)_\pi$ de modo a usar as propriedades do espaço de Hilbert, assim

$$d_2(P(l, \cdot), \pi(1))^2 = \sum_{j=1}^N \frac{(P(l, j) - \pi_j(1))^2}{\pi_j(1)} = \sum_{j=1}^N \left(\frac{P(l, j)}{\pi_j(1)} - 1 \right)^2 \pi_j(1) = \quad (3.3)$$

$$= \left(\frac{P(l, \cdot)}{\pi(1)} - \mathbf{1}, \frac{P(l, \cdot)}{\pi(1)} - \mathbf{1} \right)_{\pi(1)} \quad (3.4)$$

onde adoptámos a notação $\frac{P(l, \cdot)}{\pi(1)} = \left(\frac{P(l, 1)}{\pi_1(1)}, \dots, \frac{P(l, N)}{\pi_n(1)} \right)$ e $(1, \dots, 1) = \mathbf{1}$. Usando o facto de $\pi(1)$ ser uma distribuição de probabilidade e P ser estocástica (soma das entradas de cada linha igual a 1) obtemos

$$d_2(P(l, j), \pi(1))^2 = \left(\frac{P(l, \cdot)}{\pi(1)}, \frac{P(l, \cdot)}{\pi(1)} \right)_{\pi(1)} - 1 \quad (3.5)$$

¹Seguindo a terminologia introduzida por [Fil91], a matriz PP^* é a reversibilização multiplicativa de P , (ver secção anterior).

Veamos que, pelo Teorema 3.2.3, para qualquer vector v , temos

$$Pv = \sum_{i=1}^N (Pv, \psi_i)_{\pi(1)} \psi_i = \sum_{i=1}^N (v, P^* \psi_i)_{\pi(1)} \psi_i = \sum_{i=1}^N (v, \sigma_i \varphi_i)_{\pi(1)} \psi_i \quad (3.6)$$

$$= \sum_{i=1}^N \sigma_i (v, \varphi_i)_{\pi(1)} \psi_i \quad (3.7)$$

definindo o vector $\delta(j)$ tal que todas as suas coordenadas valem zero excepto a coordenada j , que vale 1, $\delta(j) = (0, \dots, 0, \underbrace{1}_j, 0, \dots, 0)$, temos $\frac{P(l, \cdot)}{\pi(1)} = P \frac{\delta(l)}{\pi(1)}$. Então, fazendo $v = \frac{\delta(l)}{\pi(1)}$

na igualdade (3.7) obtemos

$$\frac{P(l, \cdot)}{\pi(1)} = \sum_{i=1}^N \sigma_i \left(\frac{\delta(l)}{\pi(1)}, \varphi_i \right)_{\pi(1)} \psi_i \quad (3.8)$$

logo

$$\frac{P(l, j)}{\pi_j(1)} = \sum_{i=1}^N \sigma_i \left(\frac{\delta(l)}{\pi(1)}, \varphi_i \right)_{\pi(1)} (\psi_i)_j = \sum_{i=1}^N \sigma_i \frac{\delta_i(l)}{\pi_l(1)} (\varphi_i)_l \pi_l(1) (\psi_i)_j = \sum_{i=1}^N \sigma_i (\varphi_i)_l (\psi_i)_j. \quad (3.9)$$

Aplicando (3.9) na igualdade (3.5), temos

$$d_2(P(l, \cdot), \pi(1))^2 = \left(\frac{P(l, \cdot)}{\pi(1)}, \frac{P(l, \cdot)}{\pi(1)} \right)_{\pi(1)} - 1 = \sum_{j=1}^N \left(\sum_{i=1}^N \sigma_i (\varphi_i)_l (\psi_i)_j \sum_{i=1}^N \sigma_i (\varphi_i)_l (\psi_i)_j \pi_j(1) \right) - 1$$

que se pode decompor numa soma de parcelas do tipo:

$$\sum_{j=1}^N \sigma_i (\psi_i)_l (\varphi_i)_l \sigma_m (\psi_m)_l (\varphi_m)_j \pi_j(1).$$

Ora, se $(\varphi_i)_{i=1, \dots, N}$ é uma base ortonormada de $\ell^2(\pi)$, as únicas parcelas não nulas são tais que $i = m$, obtendo-se nesse caso

$$\sum_{j=1}^N \sigma_i (\psi_i)_l (\varphi_i)_l \sigma_m (\psi_m)_l (\varphi_m)_j \pi_j(1) = \sigma_i^2 (\psi_i)_l^2 \sum_{j=1}^N (\varphi_i)_j (\varphi_i)_j \pi_j(1) = \sigma_i^2 (\psi_i)_l^2.$$

Por outro lado, $\sigma_1 = 1$, e $\psi_1 = (1, \dots, 1)$, logo $\sigma_1 (\psi_1)_l^2 = 1$, donde se conclui,

$$d_2(P(l, \cdot), \pi(1))^2 = \sum_{i=2}^N \sigma_i^2 (\psi_i)_l^2 \quad (3.10)$$

Onde $(\psi_i)_{i \in \{1, \dots, N\}}$ é uma base ortonormada de $\ell^2(\pi(0))$. \square

Nota. Uma vez que os valores singulares e as bases fornecidas pelo Teorema 3.2.3 dependem fortemente da escolha do produto interno, B_i e B_j (definição no Capítulo 2) terão, em geral, valores singulares diferentes se $i \neq j$ e portanto não é expectável que se obtenha uma estimativa uniforme para todos os i , na linha do que obtivemos usando valores próprios, que coincidem para todas as matrizes $(B_i)_{i \in \{0, \dots, p-1\}}$ (Teorema 3.4.1).

Temos então o seguinte teorema ([SCZ09]):

Teorema 3.2.5. *Consideremos um grafo não autónomo \mathcal{G} p -periódico, com matrizes de transição estocásticas $(P_i)_{i \in \mathbb{N}}$. Dado um vector de distribuição de probabilidade u , definimos $u(i) = uP_1 \cdots P_i$, ($u(0) = u$). Então*

$$\sum_{l=1}^N d_2((P_1 \cdots P_n)(l, \cdot), u(n))^2 u_l(0) \leq \sum_{i=2}^N \prod_{j=1}^n \sigma_i(P_j, u(j-1))^2$$

Além disso, para cada linha l da matriz produto $P_1 \cdots P_n$,

$$d_2((P_1 \cdots P_n)(l, \cdot), u(n))^2 \leq \left(\frac{1}{u_l(0)} - 1 \right) \prod_{i=1}^n \sigma_2(P_i, u(i-1))^2 \quad (3.11)$$

Demonstração. Neste caso, cada matriz P_i está definida num espaço de Hilbert $\ell^2(u(i-1))$, e portanto, para uma identificação clara dos valores singulares teremos de incluir mais informação na notação, escrevemos $\sigma_i(P_n, u(n-1))$ para representar um valor singular da matriz $P_n : \ell^2(u(n-1)) \rightarrow \ell^2(u(n))$ e $\sigma_i(P_1 \cdots P_n, u(0))$ para representar um valor singular de $P_1 \cdots P_n : \ell^2(u(0)) \rightarrow \ell^2(u(n))$. Apliquemos agora a igualdade (3.10), fazendo $P = P_1 \cdots P_n$, $\pi(1) = u(n)$ e $\pi(0) = u(0)$:

$$d_2((P_1 \cdots P_n)(l, \cdot), u(n))^2 = \sum_{i=2}^N \sigma_i(P_1 \cdots P_n, u(0))^2 (\psi_i)_l^2 \quad (3.12)$$

Consideremos $(\psi_i)_{i=1, \dots, N}$ a base dada pelo Teorema 3.2.3 para $\ell^2(u(0))$ e $P_1 \cdots P_n$.

A desigualdade seguinte relaciona os valores singulares de uma matriz que resulta do produto de matrizes, com o produto dos valores singulares dos factores, e obtém-se a partir de [HJ91]:

$$\forall k = 2, \dots, N, \quad \sum_{i=2}^k \sigma_i(P_1 \cdots P_n, u(0))^2 \leq \sum_{i=2}^k \prod_{j=1}^n \sigma_i(P_j, u(j-1))^2 \quad (3.13)$$

Usando a igualdade (3.12) e seguidamente a desigualdade (3.13) com $k = N$ temos

$$\begin{aligned}
\sum_{l=1}^N d_2((P_1 \cdots P_n)(l, \cdot), u(n))^2 u_l(0) &= \sum_{l=1}^N \sum_{i=2}^N \sigma_i(P_1 \cdots P_n, u(0))^2 (\psi_i)_l^2 u_l(0) \\
&= \sum_{i=2}^N \sigma_i(P_1 \cdots P_n, u(0))^2 \sum_{l=1}^N (\psi_i)_l^2 u_l(0) \\
&= \sum_{i=2}^N \sigma_i(P_1 \cdots P_n, u(0))^2 \\
&\leq \sum_{i=2}^N \prod_{j=1}^n \sigma_i(P_j, u(j-1))^2
\end{aligned}$$

Para a segunda desigualdade do teorema, escrevemos, para cada $l \in \{1, \dots, N\}$, o vector $\frac{\delta(l)}{u_l(0)}$ como uma combinação linear dos vectores da base:

$$\frac{\delta(l)}{u_l(0)} = \sum_{i=1}^N \left(\frac{\delta(l)}{u_l(0)}, \psi_i \right)_{\pi(0)} \psi_i = \sum_{i=1}^N (\psi_i)_l \psi_i$$

então, uma vez que a base (ψ_i) é ortonormada, temos

$$\left(\frac{\delta(l)}{u_l(0)}, \frac{\delta(l)}{u_l(0)} \right)_{\pi(0)} = \left(\sum_{i=1}^N (\psi_i)_l \psi_i, \sum_{i=1}^N (\psi_i)_l \psi_i \right)_{\pi(0)} = \sum_{i=1}^N (\psi_i)_l^2$$

e por outro lado

$$\left(\frac{\delta(l)}{u_l(0)}, \frac{\delta(l)}{u_l(0)} \right)_{\pi(0)} = \left\| \frac{\delta(l)}{u_l(0)} \right\|_{\ell^2(u(0))}^2 = \frac{1}{u_l(0)}. \quad (3.14)$$

Usamos a relação $\sigma_2(P_1 \cdots P_n, u(0)) \geq \sigma_i(P_1 \cdots P_n, u(0))$ para qualquer $1 \leq i \leq N$, a desigualdade (3.13) com $k = 2$ e a igualdade (3.14) temos

$$\begin{aligned}
d_2((P_1 \cdots P_n)(l, \cdot), u(n))^2 &= \sum_{i=2}^N \sigma_i(P_1 \cdots P_n, u(0))^2 (\psi_i)_l^2 \leq \sigma_2(P_1 \cdots P_n, u(0))^2 \sum_{i=2}^N (\psi_i)_l^2 \\
&= \sigma_2(P_1 \cdots P_n, u(0))^2 \left(\sum_{i=1}^N (\psi_i)_l^2 - (\psi_1)_l \right) \\
&\leq \prod_{j=1}^n \sigma_2(P_j, u(j-1))^2 \left(\frac{1}{u_l(0)} - 1 \right)
\end{aligned}$$

Obtendo-se a segunda desigualdade, (3.11), do teorema. \square

Nota. A utilização da distância d_2 permite tirar conclusões também para a distância de

variação total, uma vez que

$$\|(P_1 \cdots P_n)(l, \cdot) - (P_1 \cdots P_n)(m, \cdot)\|_{TV} \leq \max_l \{d_2((P_1 \cdots P_n)(l, \cdot, u(n)))\}$$

O teorema anterior é o resultado mais geral que conhecemos no contexto não autónomo e é essa a razão fundamental para o incluirmos, aliado ao facto de o lema que serve a demonstração fornecer o esqueleto da prova de um resultado análogo para o caso autónomo. No entanto, o teorema revela-se de difícil aplicação, uma vez que se torna necessário calcular uma sucessão de valores singulares associados a uma sucessão de vectores $u(n)$, difícil de controlar. Além disso, não podemos garantir $\sigma_2 < 1$ sem juntar outras hipóteses. Para uma discussão neste sentido remetemos para o artigo [SCZ09].

3.3 Estimativas na aproximação ao equilíbrio em grafos autónomos

3.3.1 Forma normal de Jordan de uma matriz

Alguns dos teoremas obtidos nesta secção usam a decomposição de Jordan de uma matriz, o que justifica uma breve revisão desse resultado e das ferramentas da álgebra linear envolvidas, (para detalhes e prova ver por exemplo [Str88]). Uma matriz quadrada A é diagonalizável se é semelhante a uma matriz diagonal D , ou seja, se A admite uma factorização do tipo $A = M^{-1}DM$. Uma matriz A genérica admite uma factorização do mesmo tipo, em que a matriz D é substituída por uma matriz J diagonal por blocos, chamada forma canónica de Jordan de A . Cada bloco é uma matriz B triangular superior, cuja diagonal acima da principal está preenchida com o valor 1. A matriz J apresenta uma estrutura simples e *próxima* de uma matriz diagonal e nesse sentido generaliza o papel desempenhado pela matriz D no caso diagonalizável.

Definição 3.3.1. *Dada uma matriz quadrada A , dizemos que v é um vector próprio generalizado associado ao valor próprio λ se existe $n \in \mathbb{N}$ tal que*

$$(A - \lambda Id)^n v = 0.$$

Se $n = 1$ recuperamos a definição usual de vector próprio. O teorema seguinte formaliza os comentários anteriores.

Teorema 3.3.2. *Dada uma matriz A com R vectores próprios, existem matrizes J e M tais que as colunas de M são vectores próprios generalizados de A e $A = MJM^{-1}$ onde*

$$J = \begin{bmatrix} J_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_R \end{bmatrix}$$

é uma matriz diagonal por blocos, cada bloco J_i de J tem um valor próprio λ_i , um vector próprio correspondente, e as entradas acima da diagonal principal valem 1, ou seja

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda_i \end{bmatrix}$$

3.3.2 Estimativas obtidas

Apresentamos de seguida os resultados originais, obtidos em [SSF14b], para grafos autónomos a cujas matrizes de adjacência ergódicas não se exige nenhuma hipótese de reversibilidade. Recordamos algumas definições apresentadas no Capítulo 2, $\lambda^* = \max\{|\lambda| : \lambda \text{ é um valor próprio de } P \wedge \lambda \neq 1\}$, π é o vector próprio esquerdo, normalizado em ℓ^1 , associado ao valor próprio 1 e $\delta(k) = \max_{i,j} \left(\frac{|P^k(i,j) - \pi_j|}{\pi_j} \right)$.

Teorema 3.3.3. *Dada uma matriz P estocástica e ergódica, existe uma constante M tal que, o erro maximal no passo k , $\delta(k)$, verifica a desigualdade*

$$\delta(k) \leq \frac{M \binom{k}{m^* - 1} (\lambda^*)^{k - m^* + 1}}{\min_l \pi_l}, \quad (3.15)$$

para k a partir de certa ordem, onde m^* é a maior das dimensões dos blocos de Jordan associados aos valores próprios λ_j tais que $|\lambda_j| = \lambda^*$.

Corolário 3.3.4. *Em particular, se P é ergódica e diagonalizável então existe M tal que*

$$\delta(k) \leq \frac{M(\lambda^*)^k}{\min_l \pi_l}, \quad (3.16)$$

para todos os $k > 0$.

Demonstração. A desigualdade (3.16) obtém-se tomando $m^* = 1$ e observando que a desigualdade (3.23) é verificada para $k > k_0 = 2(m^* - 1)$. Na primeira parte da demonstração do teorema fornecemos uma prova independente para a desigualdade (3.16). \square

Nota 3.3.4.1. *Se P é uma matriz reversível então obtemos*

$$\delta(k) \leq \frac{(\lambda^*)^k}{\min_l \pi_l}.$$

As matrizes reversíveis são semelhantes a matrizes simétricas, que são um caso particular de matrizes diagonalizáveis. Assim, se P é reversível, então $P = SAS^{-1}$, com A simétrica, logo $P = SQDQ^{-1}S^{-1}$, onde D é diagonal e Q pode ser escolhida de modo a ser ortogonal. Usando a desigualdade de Cauchy-Schwarz obtemos $M \leq 1$, (ver [Beh00]).

Demonstração. (Teorema 3.3.3) Seja P uma matriz de dimensão $N \times N$, estocástica e ergódica. Ordenamos os valores próprios de P de modo que $1 > |\lambda_2| \geq \dots \geq |\lambda_N|$. A cada λ_L corresponde um vector próprio, então existem tantas repetições do valor próprio λ_L quanto o valor da sua multiplicidade geométrica. R é o número de diferentes blocos de Jordan ou vectores próprios independentes. Para uma leitura mais clara distinguimos três casos, começando pelas matrizes diagonalizáveis.

Suponhamos de momento que P é diagonalizável, então existem N valores próprios, ($R = N$), $1 > |\lambda_2| \geq \dots \geq |\lambda_N|$, e N vectores próprios correspondentes, v_1, \dots, v_N . Repare-se que a diagonalização tem lugar em \mathbb{C} , logo os valores e vectores próprios podem ser complexos. Temos

$$P = SDS^{-1}$$

onde D é uma matriz diagonal cujas entradas são os valores próprios de P , as colunas u_i de S são os vectores próprios direitos de P e as linhas v_i de S^{-1} são os vectores próprios esquerdos de P associados a λ_i .

Uma vez que π é um vector próprio esquerdo associado a 1 podemos escolher a primeira linha de S^{-1} tal que $v_1 = \pi$. Resulta que a primeira coluna de S é $u_1 = (1, \dots, 1)$, i. e., o vector próprio u_1 associado a 1 satisfazendo $u_1 \cdot v_1 = 1$, de modo a garantir $S^{-1}S = Id$. Para k arbitrário temos,

$$\begin{aligned} P^k &= (SDS^{-1})^k = SD^kS^{-1} = \\ &= \begin{bmatrix} 1 & u_2(1) & \dots & u_N(1) \\ 1 & u_2(2) & \dots & u_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & u_2(N) & \dots & u_N(N) \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & (\lambda_2)^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & (\lambda_N)^k \end{bmatrix} \begin{bmatrix} \pi_1 & \pi_2 & \dots & \pi_N \\ v_2(1) & v_2(2) & \dots & v_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ v_N(1) & v_N(2) & \dots & v_N(N) \end{bmatrix} \end{aligned}$$

o que nos permite escrever as entradas de P^k da seguinte forma:

$$P^k(i, j) = \pi_j + \sum_{L=2}^N \lambda_L^k u_L(i) v_L(j), \quad \text{for } 1 \leq i, j \leq N.$$

Donde

$$\begin{aligned} \left| P^k(i, j) - \pi_j \right| &= \left| \sum_{L=2}^N \lambda_L^k u_L(i) v_L(j) \right| \\ &\leq (\lambda^*)^k \sum_{L=2}^N |u_L(i)| |v_L(j)|. \end{aligned}$$

Uma vez que $M = \max_{i,j} \sum_{L=2}^N |u_L(i)| |v_L(j)|$ não dependem de k , obtemos:

$$\frac{|P^k(i, j) - \pi_j|}{\pi_j} \leq \frac{(\lambda^*)^k M}{\min_l \pi_l}, \quad 1 \leq i, j \leq N.$$

Finalmente,

$$\max_{i,j} \left(\frac{|P^k(i, j) - \pi_j|}{\pi_j} \right) \leq \frac{(\lambda^*)^k M}{\min_l \pi_l}.$$

Para matrizes diagonalizáveis provamos o resultado, suponhamos agora que P não é diagonalizável e representemos por m^* a dimensão do maior bloco de Jordan associado a $\lambda_j : |\lambda_j| = \lambda^*$.

Repare-se que, é possível que existam mais do que um valor próprio λ_c com $|\lambda_c| = \lambda^*$, de momento consideramos um caso particular: A matriz P tal que λ_2 é o único valor próprio com módulo λ^* e para todos os valores próprios excepto λ_2 as multiplicidades geométrica e algébrica coincidem. Isto significa que os blocos de Jordan associados a todos os valores próprios excepto λ_2 têm dimensão um, logo $R = N - m^* + 1$. Além disso, é possível que existam blocos de Jordan associados a λ_2 com dimensão $n \leq m^*$, assumimos por agora que a dimensão de cada um desses eventuais outros blocos é um. A ideia de considerar este caso particular intermédio é evidenciar a questão principal da existência de um bloco de Jordan que controla a convergência dos outros blocos, simplificando toda a restante estrutura da matriz.

Nas condições definidas acima, a decomposição de Jordan de P é dada por

$$P = SDS^{-1}$$

onde as colunas de S são vectores próprios, ou vectores próprios generalizados, direitos de P , e as linhas de S^{-1} são vectores próprios, ou vectores próprios generalizados, e a matriz

$$D = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & J_2 & \ddots & & \vdots \\ \vdots & \ddots & \lambda_3 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_{N-m^*+1} \end{bmatrix}$$

é a forma normal de Jordan de P contendo um bloco de Jordan J_2 , associado a λ_2 com

dimensão m^* , tal que

$$J_2 = \begin{bmatrix} \lambda_2 & 1 & \dots & 0 \\ 0 & \lambda_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda_2 \end{bmatrix}.$$

Usando a decomposição de Jordan de P temos para P^k :

$$P^k = (SDS^{-1})^k = \begin{bmatrix} 1 & u_2(1) & \dots & u_N(1) \\ 1 & u_2(2) & \dots & u_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & u_2(N) & \dots & u_N(N) \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & (J_2)^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & (\lambda_{N-m^*+1})^k \end{bmatrix} \begin{bmatrix} \pi_1 & \pi_2 & \dots & \pi_N \\ v_2(1) & v_2(2) & \dots & v_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ v_N(1) & v_N(2) & \dots & v_N(N) \end{bmatrix},$$

onde

$$(J_2)^k = \begin{bmatrix} (\lambda_2)^k & \binom{k}{1}(\lambda_2)^{k-1} & \dots & \dots & \binom{k}{m^*-1}(\lambda_2)^{k-m^*+1} \\ 0 & (\lambda_2)^k & \binom{k}{1}(\lambda_2)^{k-1} & \dots & \binom{k}{m^*-2}(\lambda_2)^{k-m^*+2} \\ \vdots & 0 & (\lambda_2)^k & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \binom{k}{1}(\lambda_2)^{k-1} \\ 0 & 0 & \dots & 0 & (\lambda_2)^k \end{bmatrix}.$$

Então podemos escrever a entrada (i, j) de P^k através da seguinte expressão:

$$P^k(i, j) = \pi_j + \sum_{r=0}^{m^*-1} \binom{k}{r} (\lambda_2)^{k-r} \sum_{l=1}^{m^*-r} u_{l+1}(i) v_{l+r+1}(j) + \sum_{l=3}^{N-m^*+1} (\lambda_l)^k v_{l+m^*-1}(i) u_{l+m^*-1}(j) \quad (3.17)$$

que nos permite obter para $k > 2(m^* - 1)$:

$$\begin{aligned} |P^k(i, j) - \pi_j| \leq & \binom{k}{m^*-1} (\lambda^*)^{k-m^*+1} \sum_{r=0}^{m^*-1} \sum_{l=1}^{m^*-r} |u_{l+1}(i) v_{l+r+1}(j)| + \\ & + (\lambda^*)^k \sum_{l=m^*+2}^N |v_l(i) u_l(j)|. \end{aligned} \quad (3.18)$$

Escrevendo

$$M = \max_{i,j} \sum_{r=0}^{m^*-1} \sum_{l=1}^{m^*-r} |u_{l+1}(i)v_{l+r+1}(j)|$$

e

$$Q = \max_{i,j} \sum_{l=m^*+2}^N |u_l(i)v_l(j)|$$

temos

$$\max_{i,j} \left(\frac{|P^k(i,j) - \pi_j|}{\pi_j} \right) \leq \frac{\binom{k}{m^*-1} (\lambda^*)^{k-m^*+1} M + (\lambda^*)^k Q}{\min_l \pi_l}. \quad (3.19)$$

Vejam agora o caso geral, ou seja, uma matriz tal que todos os valores próprios podem ter blocos de Jordan de dimensão $m_L \geq 1$. Recordemos que os valores próprios estão numerados de modo que $1 > |\lambda_2| \geq \dots \geq |\lambda_R|$. A forma normal de Jordan para P é dada por

$$D = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & J_2 & \ddots & & \vdots \\ \vdots & \ddots & J_3 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & J_R \end{bmatrix}$$

com

$$J_L = \begin{bmatrix} \lambda_L & 1 & \dots & 0 \\ 0 & \lambda_L & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda_L \end{bmatrix}$$

definindo $S_L = \sum_{l=1}^{L-1} m_l$ e tendo em conta a construção da igualdade (3.17), obtemos facilmente

$$P^k(i,j) = \pi_j + \sum_{L=2}^R \sum_{r=0}^{m_L-1} \binom{k}{r} (\lambda_L)^{k-r} \sum_{l=1}^{m_L-r} u_{l+S_L}(i)v_{l+r+S_L}(j). \quad (3.20)$$

Tomando $m = \max_{L \in \{0, \dots, P\}} m_L$, temos para $k > 2(m-1)$

$$|P^k(i,j) - \pi_j| \leq \sum_{L=2}^R \binom{k}{m_L-1} |\lambda_L|^{k-m_L+1} \sum_{r=0}^{m_L-1} \sum_{l=1}^{m_L-r} |u_{l+S_L}(i)v_{l+r+S_L}(j)|, \quad (3.21)$$

Então

$$\left| P^k(i, j) - \pi_j \right| \leq \sum_{L=2}^R \binom{k}{m_L - 1} (\lambda^*)^{k-m_L+1} \sum_{r=0}^{m_L-1} \sum_{l=1}^{m_L-r} |u_{l+S_L}(i) v_{l+r+S_L}(j)| \quad (3.22)$$

De modo a obter uma estimativa superior de $\delta(k)$ que envolva apenas λ^* , observamos o seguinte:

Se $m^* \geq m_L$, para qualquer $L \in \{0, \dots, R\}$, isto é, $m^* = m$, então temos, para $k > 2(m-1)$,

$$\binom{k}{m_L - 1} |\lambda_L|^{k-m_L+1} \leq \binom{k}{m^* - 1} (\lambda^*)^{k-m^*+1}, \quad (3.23)$$

Se $m^* < m$, existe k_0 tal que, para todo o $L \in \{0, \dots, R\}$, para $k > k_0$

$$\binom{k}{m_L - 1} |\lambda_L|^{k-m_L+1} \leq \binom{k}{m^* - 1} (\lambda^*)^{k-m^*+1}. \quad (3.24)$$

Assim, obtemos

$$\left| P^k(i, j) - \pi_j \right| \leq \binom{k}{m^* - 1} (\lambda^*)^{k-m^*+1} \sum_{L=2}^R \sum_{r=0}^{m_L-1} \sum_{l=1}^{m_L-r} |u_{l+S_L}(i) v_{l+r+S_L}(j)|$$

para todo o $k > k_0$, onde k_0 satisfaz $k_0 = 2(m^* - 1)$ se $m^* = m$, ou, se $m^* < m$, k_0 é tomado de modo a satisfazer (3.24).

Finalmente, escrevendo

$$M = \max_{i,j} \sum_{L=2}^R \sum_{r=0}^{m_L-1} \sum_{l=1}^{m_L-r} |u_{l+S_L}(i) v_{l+r+S_L}(j)|$$

obtemos

$$\max_{i,j} \frac{|P^k(i, j) - \pi_j|}{\pi_j} \leq \frac{M \binom{k}{m^* - 1} (\lambda^*)^{k-m^*+1}}{\min_l \pi_l},$$

para k a partir de certa ordem. □

O teorema anterior mostra essencialmente que, no caso geral, o erro relativo na aproximação à distribuição de equilíbrio em função do número de iteradas é limitado por uma potência de λ^* mas não $(\lambda^*)^k$ como no caso diagonalizável. A convergência é *atrasada* pela presença de blocos de Jordan de dimensão $m^* > 1$ associados a λ_j tais que $|\lambda_j| = \lambda^*$. Assim, a dimensão máxima dos subespaços próprios associados aos valores próprios com módulo λ^* assume um papel fundamental na determinação da velocidade de convergência

para o equilíbrio. Este teorema fornece uma estimativa útil e de fácil aplicação quando se trata de matrizes com blocos de Jordan de dimensão razoável, no entanto, se a desigualdade (3.24) da demonstração é necessária, então o resultado torna-se mais difícil de aplicar.

Nota 3.3.4.2. *Da demonstração do Teorema 3.3.3, nomeadamente da condição (3.22), tiramos facilmente a seguinte conclusão: Nas condições do Teorema 3.3.3, existe uma constante M tal que*

$$\delta(k) \leq \frac{M \binom{k}{m-1} (\lambda^*)^{k-m+1}}{\min_l \pi_l} \quad (3.25)$$

para $k > 2(m-1)$, onde m é a maior das dimensões dos blocos de Jordan presentes na forma normal de P . A desigualdade (3.25) é mais fraca do que a desigualdade (3.15) mas é verificada para $k > k_0 = 2(m-1)$ enquanto no Teorema 3.3.3 a ordem k_0 não está definida.

3.4 Estimativas na aproximação ao equilíbrio em grafos não autónomos periódicos

Nesta secção regressamos ao contexto dos grafos não autónomos definidos na Secção 2.2 do Capítulo 2 e introduzidos em [AS12]. No seguimento do que apresentámos para o caso autónomo, fornecemos agora uma estimativa para a aproximação ao equilíbrio em função do número de iteradas para grafos não autónomos periódicos, sem nenhuma hipótese de reversibilidade sobre as matrizes de adjacência, generalizando assim as estimativas obtidas no capítulo anterior. Mantemos a hipótese de ergodicidade para todas as matrizes $(B_i)_{i=0, \dots, p-1}$. Os teoremas seguintes, demonstrados em [SSF14b], utilizam definições e resultados apresentados no capítulo anterior, a partir de [AS12] e [SSF14a], nomeadamente a definição de grafo não autónomo $\mathcal{G} = (V, (E_i)_{i=0}^\infty)$ e distribuição de equilíbrio Π .

Recordamos que, pela Proposição 2.2.6, λ^* é comum a todas as matrizes produto B_j , $j \in \{0, \dots, p-1\}$ e o erro relativo maximal na aproximação das componentes de Π por $(B_1^k, \dots, B_{p-1}^k)$ é dado por (Definição 2.2.9):

$$\delta(k) = \max_{l, i, j} \frac{|B_l^k(i, j) - \pi_j(l)|}{\pi_j(l)}.$$

Estamos então em condições de apresentar uma versão do Teorema 2.2.4 para grafos não autónomos à custa do segundo maior valor próprio em módulo λ^* , supondo cada uma das matrizes B_i ergódica.

Teorema 3.4.1. *Consideremos um grafo não autônomo G , p -periódico, com matrizes estocásticas P_0, \dots, P_{p-1} , e matrizes associadas B_0, \dots, B_{p-1} tais que cada B_l é ergódica. Então existe M_0 tal que*

$$\delta(k) \leq \frac{(\lambda^*)^{k-m^*} \binom{k-1}{m^*-1} M_0}{\min_t \pi_t(0)}$$

para k a partir de certa ordem, onde m^* é a maior das dimensões dos blocos de Jordan associados aos λ_j tais que $|\lambda_j| = \lambda^*$.

Nota 3.4.1.1. 1. *Uma vez que a multiplicidade geométrica dos valores próprios coincide em todas as matrizes produto B_i , então também coincide a dimensão dos blocos de Jordan correspondentes.*

2. *O Teorema 3.4.1 generaliza o Teorema 2.2.10 onde pelo menos uma das matrizes B_i é reversível.*

3. *Tal como no caso autônomo, obtemos uma estimativa mais simples se alguma das matrizes B_i é diagonalizável, (e consequentemente todas), obtendo nesse caso a desigualdade:*

$$\delta(k) \leq \frac{(\lambda^*)^{k-1} M_0}{\min_t \pi_t(0)}, \quad \text{para qualquer } k.$$

Demonstração. (Teorema 3.4.1) Sabemos da secção anterior que para cada B_l existe uma estimativa superior de $\delta(k)$ que depende de λ^* . Vamos provar que essa mesma estimativa funciona para todos os B_l . Aplicando a igualdade (3.20) a $B_0(i, j)$ temos

$$B_0^k(i, j) = \pi_j(0) + \sum_{L=2}^P \sum_{r=0}^{m_L-1} \binom{k}{r} (\lambda_L)^{k-r} \sum_{l=1}^{m_L-r} v_{l+S_L}(i) u_{l+r+S_L}(j).$$

Por outro lado

$$B_1^{k+1} = P_1 \dots P_{p-1} (P_0 \dots P_{p-1})^k P_0 = P_1 \dots P_{p-1} B_0^k P_0.$$

Definindo $P_1^* = P_1 \dots P_{p-1}$ e usando $\pi(0)P_0 = \pi(1)$, estamos em condições de escrever a entrada (i, j) da matriz B_1^{k+1} :

$$\begin{aligned} B_1^{k+1}(i, j) &= \sum_{t=1}^N \sum_{s=1}^N P_1^*(i, s) B_0^k(s, t) P_0(t, j) \\ &= \sum_{t=1}^N \sum_{s=1}^N P_1^*(i, s) \left(\pi_t(0) + \sum_{L=2}^R \sum_{r=0}^{m_L-1} \binom{k}{r} (\lambda_L)^{k-r} \sum_{l=1}^{m_L-r} v_{l+S_L}(s) u_{l+r+S_L}(t) \right) P_0(t, j) \\ &= \pi_j(1) + \sum_{t=1}^N \sum_{s=1}^N P_1^*(i, s) \left(\sum_{L=2}^R \sum_{r=0}^{m_L-1} \binom{k}{r} (\lambda_L)^{k-r} \sum_{l=1}^{m_L-r} v_{l+S_L}(s) u_{l+r+S_L}(t) \right) P_0(t, j) \end{aligned}$$

E finalmente, usando a mesma notação da prova do Teorema 2.2.4, obtemos

$$\begin{aligned}
\frac{|B_1^{k+1}(i, j) - \pi_j(1)|}{\pi_j(1)} &\leq \frac{(\lambda^*)^{k-m^*+1} \binom{k}{m^*-1}}{\pi_j(1)} \sum_{t=1}^N P_0(t, j) \sum_{s=1}^N P_1^*(i, s) \times \\
&\quad \times \max_{s,t} \sum_{L=2}^P \sum_{r=0}^{m_L-1} \sum_{l=1}^{m_L-r} |v_{l+S_L}(s) u_{l+r+S_L}(t)| \\
&\leq \frac{(\lambda^*)^{k-m^*+1} \binom{k}{m^*-1} M_0}{\pi_j(1) \min_t \pi_0(t)} \sum_{t=1}^N \pi_t(0) P_0(t, j) \\
&= \frac{(\lambda^*)^{k-m^*+1} \binom{k}{m^*-1} M_0}{\min_t \pi_t(0)}
\end{aligned}$$

para k a partir de certa ordem. Usando os mesmos argumentos provamos a desigualdade para todos os B_l , com $l \in \{0, \dots, p-1\}$. \square

3.5 Exemplos

Nesta secção apresentamos exemplos do comportamento de matrizes ergódicas, ou seja cadeias de Markov ergódicas (homogéneas e não homogéneas), quanto à convergência para o equilíbrio, ilustrando as estimativas obtidas anteriormente.

Exemplo 3.5.1. *Seja*

$$P = \begin{bmatrix} 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{6} & 0 & \frac{1}{6} & \frac{2}{3} \end{bmatrix}$$

a matriz de adjacência estocástica associada a certo grafo \mathcal{G} . P é diagonalizável, com $\lambda^* = \frac{1+\sqrt{5}}{12}$ e $M = \frac{745+333\sqrt{5}}{145}$. Temos

$$\delta(k) \leq \frac{29}{18} \frac{745 + 333\sqrt{5}}{145} \left(\frac{1 + \sqrt{5}}{12} \right)^k \approx 1.61(0.2697)^k,$$

para qualquer $k > 0$.

Exemplo 3.5.2. Consideremos agora uma matriz não diagonalizável

$$P = \begin{bmatrix} \frac{775}{2004} & \frac{941}{8016} & \frac{3739}{8016} & \frac{59}{2004} \\ \frac{21}{668} & \frac{681}{2672} & \frac{3397}{8016} & \frac{581}{2004} \\ \frac{141}{668} & \frac{6367}{40080} & \frac{3341}{8016} & \frac{2137}{10020} \\ \frac{201}{668} & \frac{1757}{8016} & \frac{307}{8016} & \frac{295}{668} \end{bmatrix}.$$

Então $\lambda^* = \frac{1}{6}$ tem multiplicidade algébrica 3 e geométrica 1, pelo que o bloco de Jordan associado a $\frac{1}{6}$ tem dimensão $m = 3$. Temos também $M = 675/334$, então, de acordo com o Teorema 3.3.3, existe uma estimativa para o erro relativo maximal das entradas da matriz P^k :

$$\delta(k) \leq \frac{675 \binom{k}{2}}{81} \left(\frac{1}{6}\right)^{k-2} \approx \frac{4.167 \cdot k(k-1)}{6^{k-2}}, \text{ para todo o } k > 4.$$

Neste caso a convergência para a distribuição de equilíbrio é mais lenta quando comparada com a convergência de uma matriz diagonalizável. Com efeito, para $k > 4$, temos $\delta(k) > \frac{M(\lambda^*)^k}{\min_l \pi_l}$, logo, a estimativa esperada para matrizes diagonalizáveis não seria verificada neste exemplo.

Exemplo 3.5.3. Consideremos agora um grafo não autónomo de período 2, cujas matrizes de adjacência estocásticas, de dimensão 3×3 , são as matrizes P_0 e P_1 definidas por:

$$P_0 = \begin{bmatrix} \frac{14}{45} & \frac{11}{360} & \frac{79}{120} \\ \frac{23}{180} & \frac{113}{360} & \frac{67}{120} \\ \frac{1}{12} & \frac{1}{24} & \frac{7}{8} \end{bmatrix}, \quad P_1 = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{6} & \frac{1}{12} & \frac{3}{4} \end{bmatrix}$$

Então

$$B_0 = P_0 P_1 = \begin{bmatrix} \frac{191}{720} & \frac{101}{1440} & \frac{319}{480} \\ \frac{113}{720} & \frac{293}{1440} & \frac{307}{480} \\ \frac{3}{16} & \frac{3}{32} & \frac{23}{32} \end{bmatrix}, \quad B_1 = P_1 P_0 = \begin{bmatrix} \frac{71}{360} & \frac{13}{360} & \frac{23}{30} \\ \frac{19}{180} & \frac{8}{45} & \frac{43}{60} \\ \frac{1}{8} & \frac{1}{16} & \frac{13}{16} \end{bmatrix}$$

são ambas diagonalizáveis com $\lambda^* = \frac{1}{8}$, $M_0 = \frac{5}{3}$, pelo que $\delta(k) \leq \frac{50}{3} \left(\frac{1}{8}\right)^{k-1}$ para qualquer k .

Capítulo 4

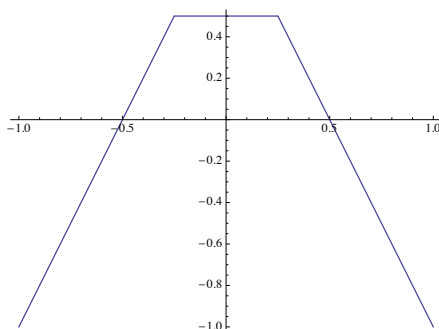
Taxas de convergência universais numa família de sistemas não autónomos

4.1 Família de funções *tenda cortada*

Seja $I = [-1, 1]$ e $f : I \rightarrow I$ uma função contínua com m pontos críticos, $-1 < c_1 < \dots < c_m < 1$, monótona nos $m + 1$ intervalos I_j entre os pontos de viragem. Nestas condições f diz-se uma função m -modal. Os exemplos mais simples de funções m -modais são as chamadas funções *dentes de serra*. As funções da família *dentes de serra cortada* constituem bons modelos na teoria de amassamento uma vez que são suficientemente ricos para conter todas as sequências de amassamento, ou seja, todas as sequências simbólicas admissíveis são realizadas como itinerário crítico para alguma função desta família. Neste capítulo consideramos uma família de funções *dentes de serra cortada*, unimodal — a família das *tendas cortadas* — e estudaremos o sistema dinâmico discreto não autónomo de período 2 que se obtém pela composição sucessiva alternada de dois elementos, f_0, f_1 , dessa família. Trata-se da família de funções a um parâmetro $v \in [-1, 1]$, $f^v : [-1, 1] \rightarrow [-1, 1]$, definida por

$$f^v(x) = \begin{cases} 1 - 2x, & x > \frac{1-v}{2} \\ v, & \frac{v-1}{2} \leq x \leq \frac{1-v}{2} \\ 1 + 2x, & x < \frac{v-1}{2} \end{cases} \quad (4.1)$$

com $v \in [-1, 1]$. O gráfico de $f^{\frac{1}{2}}$ está representado na Figura 4.1. O parâmetro v desta

Figura 4.1: Gráfico de $f^{\frac{1}{2}}$.

família de funções corresponde à altura onde a função tenda, definida de seguida, é *cortada*:

$$T(x) = \begin{cases} T^-(x) = 1 - 2x, & x \geq 0 \\ T^+(x) = 1 + 2x, & x < 0 \end{cases} \quad (4.2)$$

A função f_v obtém-se a partir de T tomando esta constante com valor v no intervalo $[-\frac{v-1}{2}, \frac{v-1}{2}]$. Ao intervalo $[-\frac{v-1}{2}, \frac{v-1}{2}]$ chamamos patamar, assim v coincide com a altura do patamar.

4.2 Produto cruzado

Neste capítulo optamos por usar a definição de sistema dinâmico não autónomo usando a ideia de *produto cruzado*, introduzida por Elaydi e Sacker em [ES05a] e [ES05b] para lidar com sistemas não autónomos periódicos. Apesar de a estrutura do sistema dinâmico ser bastante intuitiva — a iteração sequencial de uma sucessão f_n de funções, — esta escolha surge naturalmente para superar dificuldades de notação e definir com rigor os conceitos envolvidos, nomeadamente a renormalização. Começamos por fazer a analogia com os sistemas dinâmicos autónomos.

Dado um intervalo $I \in \mathbb{R}$, chamamos sistema dinâmico discreto autónomo a uma função $F : I \times \mathbb{Z}_0^+ \rightarrow I$, contínua, que satisfaz

1. $F(x, 0) = x, \forall x \in I$;
2. $F(F(x_0, n), m) = F(x_0, n + m), \forall x_0 \in I, n, m \in \mathbb{Z}$.

ou seja, F satisfaz as propriedades de semigrupo, (ou grupo, se f é invertível).

Claramente a igualdade $F(x, n) = f^n(x)$, com $x \in I$, onde f é contínua, satisfaz as condições acima e define um sistema dinâmico discreto autónomo. Consideremos uma sucessão de funções $f_n : I \rightarrow I, n \in \mathbb{Z}_0^+$ e o sistema não autónomo gerado por $F :$

$I \times \mathbb{Z}_0^+ \times \mathbb{Z}_0^+ \rightarrow I \times \mathbb{Z}_0^+$ tal que $F((x, i), 0) = (x, i)$ e para $n \geq 0$,

$$F((x, i), n) = (f_{i+n-1} \circ \dots \circ f_i(x), n + i). \quad (4.3)$$

O sistema assim definido satisfaz

$$\begin{aligned} F(F((x, i), n), m) &= F((f_{i+n-1} \circ \dots \circ f_i(x), n + i), m) \\ &= (f_{m+n+i-1} \circ \dots \circ f_{n+i}(f_{i+n-1} \circ \dots \circ f_i(x)), m + n + i) \\ &= (f_{m+n+i-1} \circ \dots \circ f_{n+i} \circ f_{i+n-1} \circ \dots \circ f_i(x), m + n + i) \\ &= F((x, i), m + n) \end{aligned}$$

para $n, m \in \mathbb{Z}_0^+$. O formalismo do *produto cruzado* surge assim para assegurar a generalização das propriedades dos sistemas autónomos. Por outro lado, essas propriedades permitem interpretar o sistema não autónomo como um sistema autónomo num espaço de dimensão 2. *Grosso modo* a primeira coordenada de F indica o resultado da iteração e a segunda coordenada de F guarda a posição onde a iteração começa.

Usaremos a notação simplificada $F((x, i), n) = F^n(x, i)$. Se a sucessão de funções f_n é periódica de período k , isto é, $f_{n+k}(x) = f_n(x), \forall x \in I, k \in \mathbb{N}$, então basta tomar $i \in \{0, \dots, k-1\}$ e a igualdade (4.3) escreve-se

$$F^n(x, i) = (f_{(i+n-1) \bmod k} \circ \dots \circ f_i(x), (n + i) \bmod k). \quad (4.4)$$

Nas próximas secções estudaremos o sistema não autónomo F , definido através do produto cruzado, gerado pela iteração alternada de duas funções f^v e f^w , ambas da família das *tendas cortadas*, onde $(v, w) \in [-1, 1]^2$, fazendo $f_0 = f^v$ e $f_1 = f^w$.

No que se segue os valores $i + n$ deverão ser entendidos $\bmod k$. Chamamos órbita de (x, i) através de F à sucessão dos pares

$$(x, i), (f_i(x), i + 1), (f_{i+1}(f_i(x)), i + 2), \dots$$

Chamamos órbita de x através de $F(\cdot, i)$ à sucessão das iteradas

$$x, f_i(x), f_{i+1}(f_i(x)), f_{i+2}(f_{i+1}(f_i(x))), \dots$$

4.3 Dinâmica simbólica

Definição 4.3.1. *O itinerário $i_T(x)$ de um ponto x sob iteração da função tenda T é uma sucessão de símbolos em $\{L, 0, R\}$,*

$$i_T(x) = X_1 X_2 \dots X_n$$

com $n = \min\{n \geq 1 : T^n(x) = 0\}$, tal que $X_i = ad(T^i(x))$, onde

$$ad(x) = \begin{cases} L & \text{se } x < 0 \\ 0 & \text{se } x = 0 \\ R & \text{se } x > 0 \end{cases} . \quad (4.5)$$

O itinerário será uma sequência simbólica finita terminada em 0 se x for uma pré-imagem de 0, caso contrário será uma sequência infinita nos símbolos L e R .

Nota 4.3.1.1. *As sequências finitas — correspondentes a pré-imagens de 0 — podem ser codificadas como sequências infinitas apenas com os símbolos L e R , mais, cada sequência finita $X = X_1 \dots X_{n-1}0$ tem dois representantes no espaço simbólico das sequências infinitas em L e R , a saber, $X = X_1 \dots X_{n-1}LRL^\infty$ e $X = X_1 \dots X_{n-1}RRL^\infty$. Não se trata de sequências distintas mas apenas de três possíveis representações da mesma sequência que identificaremos sempre por $X = X_1 \dots X_{n-1}0$. É esta identificação, (do mesmo tipo que se faz com 0.(9) e 1), que permite definir uma conjugação entre a função tenda em $[-1, 1]$ e a função deslocamento em Σ (ver [Dev89]).*

Consideremos o conjunto Σ contendo todas as sequências simbólicas nos símbolos $\{L, 0, R\}$ definidas anteriormente, isto é, contendo as sequências $X = X_1 \dots X_n$ tais que $X_i \neq 0$ para $i < n$ e $n = +\infty$ ou $X_n = 0$.

O número de símbolos de uma sequência X diz-se comprimento da sequência e é representado por $|X|$. Assim $|X_1 \dots X_n| = n$, com $n \in \mathbb{N}$ ou $n = \infty$.

Definimos em Σ a aplicação deslocamento:

$$\begin{aligned} \sigma : \Sigma \setminus \{0\} &\longrightarrow \Sigma \\ \sigma(X_1 X_2 X_3 \dots) &= X_2 X_3 \dots \end{aligned}$$

Consideremos o sinal associado a cada símbolo X_i , para $i \in \{1, \dots, |X| - 1\}$, da sequência X :

$$\varepsilon(X_i) = \begin{cases} 1 & \text{se } X_i = L \\ 0 & \text{se } X_i = 0 \\ -1 & \text{se } X_i = R \end{cases}$$

O sinal de cada símbolo X_i está relacionado com a monotonia do troço da função no intervalo correspondente a X_i e, assim, o produto dos sinais $\varepsilon_i(X) = \prod_{j=1}^i \varepsilon(X_j)$ indica a monotonia da função T^i numa vizinhança de x .

Dado $X \in \Sigma$, definimos a função

$$\Phi(X) = - \sum_{i=1}^{|X|} \frac{\varepsilon_i(X)}{2^i},$$

onde $\varepsilon_i(X) = \prod_{j=1}^i \varepsilon(X_j)$.

Definimos $L < 0 < R$ e $-R < 0 < -L$ seguindo a ordem induzida naturalmente pelo intervalo I , de modo a munir o conjunto das seqüências Σ com uma relação de ordem:

Dadas duas seqüências X e Y em Σ , dizemos que $X < Y$ se existe $0 \leq r \leq \min\{|X|, |Y|\}$ tal que $X_i = Y_i$, se $i < r \wedge \varepsilon_{r-1}(X)X_r < \varepsilon_{r-1}(Y)Y_r$.

Dotamos Σ da topologia induzida pela métrica:

$$d(X, Y) = \sum_{i=1}^{\infty} \frac{|\varepsilon(X_i) - \varepsilon(Y_i)|}{2^i}.$$

Esta métrica vai de encontro à ideia de que duas seqüências com um *grande* bloco inicial de símbolos coincidentes estão *próximas*.

Temos os seguintes resultados:

Proposição 4.3.2. *Dadas duas seqüências X e Y em Σ :*

1. Se $X < Y$ então $\Phi(X) < \Phi(Y)$;
2. $\Phi(\Sigma) \subset [-1, 1]$.

Demonstração. 1. Se $X < Y$ existe $n \leq \max\{|X|, |Y|\}$ tal que $X_1 \dots X_{n-1} = Y_1 \dots Y_{n-1}$ e $X_n \neq Y_n$ então, se $i \leq n-1$, $\varepsilon_i(X) = \varepsilon_i(Y)$. Suponhamos que $\varepsilon_n(Y) - \varepsilon_n(X) = -2$. Temos $\varepsilon_{k+n}(Y) - \varepsilon_{k+n}(X) \leq 2$ qualquer que seja $k \geq 0$, então

$$\begin{aligned} \Phi(X) - \Phi(Y) &= \left(\sum_{i=n}^{|Y|} \frac{\varepsilon_i(Y)}{2^i} - \sum_{i=n}^{|X|} \frac{\varepsilon_i(X)}{2^i} \right) \\ &= \left(\frac{\varepsilon_n(Y) - \varepsilon_n(X)}{2^n} + \sum_{i=n+1}^{|Y|} \frac{\varepsilon_i(Y)}{2^i} - \sum_{i=n+1}^{|X|} \frac{\varepsilon_i(X)}{2^i} \right) \\ &< \left(\frac{\varepsilon_n(Y) - \varepsilon_n(X)}{2^n} + \frac{1}{2^n} \sum_{i=1}^{\infty} \frac{2}{2^i} \right) \\ &= \left(\frac{\varepsilon_n(Y) - \varepsilon_n(X)}{2^n} + \frac{1}{2^{n-1}} \right) = \left(\frac{-2}{2^n} + \frac{1}{2^{n-1}} \right) = 0 \end{aligned}$$

Justificação da desigualdade estrita:

Verifica-se

$$\sum_{i=n+1}^{|Y|} \frac{\varepsilon_i(Y)}{2^i} - \sum_{i=n+1}^{|X|} \frac{\varepsilon_i(X)}{2^i} \leq \frac{1}{2^n} \sum_{i=1}^{\infty} \frac{2}{2^i},$$

e a igualdade só é realizada se $\sigma^{n+1}(X) = \sigma^{n+1}(Y) = RL^\infty$. Juntando as hipóteses $\varepsilon_n(Y) - \varepsilon_n(X) = -2$, $X_1 \dots X_{n-1} = Y_1 \dots Y_{n-1}$, estaríamos na presença de duas

seqüências que são representantes da mesma seqüência finita, mais precisamente $X_1 \dots X_{n-1}0$, o que contraria a hipótese inicial.

Suponhamos agora que $\varepsilon_n(Y) - \varepsilon_n(X) = -1$, uma das seqüências é necessariamente finita e assim temos

$$\begin{aligned} \Phi(X) - \Phi(Y) &< \left(\frac{\varepsilon_n(Y) - \varepsilon_n(X)}{2^n} + \frac{1}{2^n} \sum_{i=1}^{\infty} \frac{1}{2^i} \right) \\ &= \left(\frac{\varepsilon_n(Y) - \varepsilon_n(X)}{2^n} + \frac{1}{2^n} \right) \\ &= \left(\frac{-1}{2^n} + \frac{1}{2^n} \right) = 0 \end{aligned}$$

Também neste caso a igualdade só seria atingida se $\sigma^{n+1}(X) = RL^\infty$ (supondo X infinita e Y finita) o que faria de X um representante da seqüência finita Y .

2. Tomemos X a menor e Y a maior seqüência em Σ , respectivamente $X = L^\infty$ e $Y = RL^\infty$, temos $\Phi(X) = -\sum_{i=1}^{\infty} \frac{1}{2^i} = -1$ e $\Phi(Y) = -\left(-\frac{1}{2} + \sum_{i=2}^{\infty} \frac{-1}{2^i}\right) = 1$. Uma vez que no ponto anterior se provou a monotonia de Φ , o resultado segue. \square

Proposição 4.3.3. *A função Φ é uma conjugação entre a aplicação deslocamento σ em $\Sigma \setminus \{0\}$ e a função tenda T em $[-1, 1]$, assim, $\Phi(\sigma(X)) = T(\Phi(X))$ e o seguinte diagrama é comutativo:*

$$\begin{array}{ccc} \Sigma \setminus \{0\} & \xrightarrow{\sigma} & \Sigma \\ \Phi \downarrow & & \Phi \downarrow \\ I & \xrightarrow{T} & I \end{array}$$

Demonstração. Continuidade de Φ em Σ : Seja X uma seqüência em Σ , dado $\delta > 0$ tomemos n tal que $2^{-n} < \delta$ e consideremos uma seqüência Y tal que $d(X, Y) = \sum_{i=1}^{\infty} \frac{|\varepsilon(X_i) - \varepsilon(Y_i)|}{2^i} < 2^{-n}$. Assim

$$\begin{aligned} |\Phi(X) - \Phi(Y)| &= \left| \frac{\sum_{i=1}^{\infty} \varepsilon(X_i)}{2^i} - \frac{\sum_{i=1}^{\infty} \varepsilon(Y_i)}{2^i} \right| \leq \sum_{i=1}^{\infty} \frac{|\varepsilon(X_i) - \varepsilon(Y_i)|}{2^i} \\ &< 2^{-n} < \delta \end{aligned}$$

Seja $X = X_1 X_2 \dots$, suponhamos que $X_1 = L$, então $\Phi(X) \leq 0$.

$$\begin{aligned} \Phi(\sigma(X)) &= \Phi(X_2 \dots X_n \dots) = - \sum_{i=1}^{|\sigma(X)|} \frac{\varepsilon_i(X_2 \dots)}{2^i} = - \sum_{i=1}^{|X|-1} \frac{\varepsilon_{i+1}(X)}{2^i} \\ &= -2 \sum_{i=1}^{|X|-1} \frac{\varepsilon_{i+1}(X)}{2^{i+1}} = -2 \sum_{i=2}^{|X|} \frac{\varepsilon_i(X)}{2^i}. \end{aligned}$$

Por outro lado,

$$\begin{aligned} T(\Phi(X)) &= T\left(-\sum_{i=1}^{|X|} \frac{\varepsilon_i(X)}{2^i}\right) = 2\left(-\sum_{i=1}^{|X|} \frac{\varepsilon_i(X)}{2^i}\right) + 1 = -2\left(\sum_{i=1}^{|X|} \frac{\varepsilon_i(X)}{2^i} - \frac{1}{2}\right) \\ &= -2 \sum_{i=2}^{|X|} \frac{\varepsilon_i(X)}{2^i}. \end{aligned}$$

Se $X_1 = R$ a prova é análoga. \square

Notação. Dada a sequência $X \in \Sigma$, identificamos $X_{[1,p]}$ com o bloco simbólico $X_1 \dots X_p$.

De forma análoga ao que foi feito para a função tenda, vamos definir o itinerário $i_F(x, i)$. Neste caso usaremos o mesmo conjunto de símbolos $\{L, R, 0\}$ para a primeira componente do itinerário, enquanto a segunda componente guarda a informação sobre a primeira função a aplicar:

$$i_F(x, i) = (ad(f_i(x)), ad(f_{i+1} \circ f_i(x)), \dots, ad(f_{i+n} \circ \dots \circ f_i(x)), i+1)$$

usando ad definido em (4.5) e $n = \min\{n \geq 1 : f_{n+i} \circ \dots \circ f_i(x) = 0\}$.

O itinerário de (x, i) por F é um par cuja primeira componente é uma sequência finita de símbolos L e R e último símbolo 0 se e só se (x, i) é uma pré-imagem de $(0, k)$, para algum k , caso contrário é uma sequência infinita de símbolos L e R .

Nota. O itinerário assim definido traduz simbolicamente a órbita de $(f_i(x), i+1)$ e não de (x, i) .

Definição 4.3.4. Chamamos projecção à função π que faz corresponder a um par (y, z) a sua primeira componente y . Aplicaremos π em ambos os casos (x, i) ou (X, i) , ou seja, $\pi(x, i) = x$ e $\pi(X, i) = X$.

No nosso contexto as sequências finitas são identificadas com sequências infinitas de forma distinta consoante o comprimento da sequência seja par ou ímpar, como veremos mais à frente.

Definição 4.3.5. Dadas duas sequências X, Y em Σ e $i \in \mathbb{N}$ dizemos que $(X, i) < (Y, i)$ se $X < Y$.

Ou seja, a relação de ordem na fibra é induzida pela projecção.

As definições anteriores permitem estabelecer as seguintes relações:

Proposição 4.3.6. *Consideremos o sistema dinâmico não autônomo F e os pontos $x, y \in I$,*

1. *se $f_i(x) < f_i(y)$ então $i_F(x, i) \leq i_F(y, i)$.*
2. *se $i_F(x, i) < i_F(y, i)$ então $f_i(x) < f_i(y)$.*

Demonstração. 1. Consideremos x, y tais que $f_i(x) < f_i(y)$ e $i_F(x, i) \neq i_F(y, i)$, existe $n \geq 1$ tal que $X_i = Y_i$, para $i < n$ e $X_n \neq Y_n$. Assim temos $\pi(F^{n-1}(x, i))$ e $\pi(F^{n-1}(y, i))$ num mesmo intervalo da partição de I , pelo que $\pi(F^n(\cdot, i))$ é monótona no intervalo $[f_i(x), f_i(y)]$, e $\pi(F^n(x, i))$ e $\pi(F^n(y, i))$ pertencem a intervalos distintos da partição. Se $\varepsilon_{n-1}(X) = 1$ então $\pi(F^n(\cdot, i))$ é crescente, donde $\pi(F^n(x, i)) < 0 < \pi(F^n(y, i))$ logo $i_F(x, i) < i_F(y, i)$. Se $\varepsilon_{n-1}(X) = -1$ então $\pi(F^n(\cdot, i))$ é decrescente, donde $\pi(F^n(x, i)) > 0 > \pi(F^n(y, i))$, logo $i_F(x, i) < i_F(y, i)$. Invertendo o argumento prova-se a segunda desigualdade. \square

4.3.1 Sistema não autônomo de período 2

Doravante consideramos um sistema F , definido como anteriormente, de período 2, tal que $(f_0, f_1) = (f^v, f^w)$.

Evidentemente uma sequência arbitrária não é necessariamente admissível como (primeira componente do) itinerário de um ponto (x, i) . Pretendendo saber quais as sequências simbólicas realizadas por F , começamos por apresentar algumas definições:

Definição 4.3.7. *Dada uma sequência $X \in \Sigma$, dizemos que X é maximal se $\sigma^{2k}(X) \leq \sigma(X)$, $\forall k \in \mathbb{N}$.*

Analogamente dizemos que X é minimal se $\sigma^{2k}(X) \geq \sigma(X)$, $\forall k \in \mathbb{N}$.

Se $|X| < \infty$ então as desigualdades anteriores são estritas.

Dizemos informalmente que (x, i) é uma pré-imagem (de ordem) par de 0 se (x, i) é tal que $\pi(F^{2k}(x, i)) = 0$, ou seja se $|\pi(i_F(x, i))|$ é par, e (x, i) é uma pré-imagem ímpar de 0 se $\pi(F^{2k-1}(x, i)) = 0$, ou seja se $|\pi(i_F(x, i))|$ é ímpar, para $k \geq 1$ inteiro .

Definição 4.3.8. *Designamos $K_p = (K^0, K^1) = (\pi(i_F(0, 0)), \pi(i_F(0, 1)))$ como par de amassamento associado ao sistema F , sendo F o sistema gerado pelo par de aplicações $(f_0, f_1) = (f^v, f^w)$ com $p = (v, w)$. (K^0, K^1) diz-se invariante de amassamento.*

Definição 4.3.9. *Dizemos que $(X^0, X^1) \in \Sigma \times \Sigma$ é admissível se existe um par $p = (v, w)$ tal que $K_p = (X^0, X^1)$.*

Representamos por Σ_2^+ o conjunto dos pares de amassamento admissíveis.

Teorema 4.3.10. *Um par $(X^0, X^1) \in \Sigma \times \Sigma$ é admissível se e só se*

$$\sigma^n(X^i) \leq X^{(i+n) \bmod 2}, \quad \text{para } i = 0, 1, \quad n \in \mathbb{N}. \quad (4.6)$$

Nota. As condições de admissibilidade implicam em particular que cada uma das sequências X^0 e X^1 seja maximal.

Demonstração. Suponhamos que (X^0, X^1) verifica a condição (4.6), isto é, para $k \geq 1$ inteiro,

$$\begin{aligned} \sigma^{2k}(X^0) &\leq X^0, \\ \sigma^{2k-1}(X^0) &\leq X^1 \\ \sigma^{2k}(X^1) &\leq X^1, \\ \sigma^{2k-1}(X^1) &\leq X^0, \end{aligned}$$

consideremos $v = \Phi(X^0)$, $w = \Phi(X^1)$ e o sistema $F(\cdot, i)$ definido usualmente, com $(f_0, f_1) = (f^v, f^w)$. Usando as proposições 4.3.2 e 4.3.3 temos para qualquer k

$$\begin{aligned} \sigma^{2k}(X^0) \leq X^0 &\Rightarrow \Phi(\sigma^{2k}(X^0)) \leq \Phi(X^0) \Rightarrow T^{2k}(\Phi(X^0)) \leq \Phi(X^0) \Leftrightarrow \\ &\Leftrightarrow T^{2k}(v) \leq v. \end{aligned}$$

Do mesmo modo temos $T^{2k-1}(v) \leq w$ e assim a órbita de $x = 0$ através de $F(\cdot, 0)$ não entra em nenhum patamar a não ser eventualmente num dos extremos. Usando os mesmos argumentos provamos que a órbita de $x = 0$ através de $F(\cdot, 1)$ não entra em nenhum patamar a não ser eventualmente nos extremos, pelo que o itinerário de $(0, i)$, com $i = 0, 1$, através de $F(\cdot, i)$ é realizado por X^i . A recíproca é imediata. \square

Nota. A prova do teorema anterior fornece um método explícito para obter parâmetros que realizam cada par de sequências admissíveis.

O Teorema 4.3.10 permite identificar as sequências K^0 que tornam o par (K^0, K^1) admissível, fixada uma sequência K^1 e vice-versa, como estabelece o seguinte corolário:

Corolário 4.3.11. *Dada uma sequência X em Σ , maximal, o par (X, Y) é um par de amassamento admissível se e só se Y é maximal, $\sigma^{2k-1}(Y) \leq X$ e $\sigma^{2k-1}(X) \leq Y$, para $k \in \mathbb{N}$. (Analogamente o par (Y, X) é admissível se e só se Y é maximal, $\sigma^{2k-1}(Y) \leq X$ e $\sigma^{2k-1}(X) \leq Y$, para $k \in \mathbb{N}$.)*

Exemplo 4.3.1. *Apresentamos as sequências maximais de comprimento par menor ou igual a 6:*

- $L0, R0$;

- $LRL0, RRL0, RLR0, RLL0$;
- $LRLLL0, RRLLL0, RRLRL0, RRRRL0, RLRRR0, RLRRLO, RLLRLO, RLLRR0, RLLLR0, RLLLL0$.

Os pares $(R0, RLR0)$ e $(RLRRL0, RLL0)$ são admissíveis enquanto o par $(R0, LRL0)$ não é admissível.

Repare-se que no contexto autônomo da dinâmica simbólica gerada pela iteração de uma tenda cortada, a sequência $X = RLRRLO$ não seria admissível como invariante de amassamento uma vez que $\sigma^3(X) = RL0 > X$, no nosso caso esta desigualdade é irrelevante para as condições de admissibilidade.

Definição 4.3.12. Dado $p = (v, w)$ consideremos o sistema dinâmico não autônomo F , de período 2, com $(f_0, f_1) = (f^v, f^w)$, com invariante de amassamento $K_p = (K^0, K^1)$. Dizemos que o par (X, i) é p -admissível se existe $x \in [-1, 1]$ tal que $X = \pi(i_F(x, i))$. O conjunto das sequências p -admissíveis é representado por Σ_p^+ .

O resultado seguinte mostra que, tal como no caso autônomo, o invariante de amassamento $K_p = (K^0, K^1) = (\pi(i_F(0, 0)), \pi(i_F(0, 1)))$ determina o conjunto de todas as sequências p -admissíveis:

Teorema 4.3.13. Um par $(X, i) \in \Sigma \times \{0, 1\}$ é p -admissível, com $K_p = (K^0, K^1)$, se e só se verifica a seguinte condição

$$\sigma^n(\pi(X, i)) \leq K^{(n+i) \bmod 2}, \quad n \in \mathbb{N} \quad (4.7)$$

Demonstração. Suponhamos que a condição (4.7) é verificada para $p = (v, w) \in [-1, 1]^2$ e consideremos o par de amassamento $K_p = (K^0, K^1)$. O par $p' = (v', w') = (\Phi(K^0), \Phi(K^1))$ realiza o par de sequências de amassamento correspondente a p mas a órbita de 0 através de $F(\cdot, i)$, com $i = 0, 1$, não entra nos patamares a não ser eventualmente nos extremos. Assim podemos supor $(v, w) = (v', w')$.

Tomemos uma sequência X que verifica a desigualdade (4.7), vamos mostrar que existe x , tal que $f_0(x) = v$ e $v = \Phi(X)$, que realiza $\pi(X, 0)$ através de F com $(f_0, f_1) = (f^v, f^w)$, (para $\pi(X, 1)$ é análogo). Pela Proposição 4.3.3, $X = i_T(x)$. Podemos escolher x num dos extremos do patamar, para fixar ideias escolhemos o extremo esquerdo. Se a órbita de x não entra em nenhum patamar temos imediatamente $X = i_T(x) = \pi(i_F(x, 0))$. Suponhamos agora que existe k par, tal que, $\pi(F^k(x, 0))$ entra num patamar, e suponhamos que não é no extremo, então $\sigma^k(X) = it_T(T^k(x)) > \pi(it_F(F^k(0), 0)) = K^0$, o que contraria a hipótese, pelo que $\pi(F^k(x, 0))$ só poderá igualar um extremo do patamar. Se k é ímpar teremos a mesma conclusão por $\sigma^k(X) = it_T(T^k(x)) > \pi(it_F(F^k(0), 0)) = K^1$. Em ambos os casos temos $X = i_T(x) = \pi(i_F(x, 0))$. A recíproca é imediata. \square

Exemplo 4.3.2. Consideremos o par de amassamento $(K^0, K^1) = (RLL0, R0)$. Temos $\Phi(K^0) = 0.875$ e $\Phi(K^1) = 0.5$. No gráfico seguinte (Figura 4.2) está representado o par de funções (f^v, f^w) com $(v, w) = (0.875, 0.5)$. Podemos verificar que as projecções dos itinerários de 0 por $F(\cdot, 0)$ e por $F(\cdot, 1)$ coincidem com as seqüências $(RLL0, R0)$.

Se mantivermos $v = 0.875$ fixo e subirmos o patamar w , o itinerário de 0 por $F(\cdot, 0)$ mantém-se. Então, no espaço de parâmetros $[-1, 1]^2$ vamos encontrar um segmento de recta $(v = 0.875, 0.5 \leq w \leq 1)$ correspondente à seqüência $RLL0$, veremos na próxima secção que o segmento referido é efectivamente parte do osso correspondente à seqüência $RLL0$.

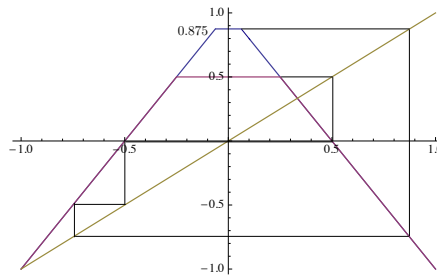


Figura 4.2: Gráficos de $f_0 = f^v$ e $f_1 = f^w$, para $(v, w) = (0.875, 0.5)$.

É usual em dinâmica simbólica identificar uma seqüência finita $X_{[1,n]}0$ com as seqüências infinitas periódicas $X_{[1,n]}(LK)^\infty$ ou $X_{[1,n]}(RK)^\infty$, onde K é o invariante de amassamento. No entanto, no contexto não autónomo, dado o par (X, i) temos de observar qual a paridade de $|X|$, de $|K^0|$ e de $|K^1|$. Com efeito, a interpretação do par $(X, i) = (X_{[1,n]}0, i)$, num sistema com invariante de amassamento (K^0, K^1) , depende da paridade da posição do símbolo 0, ou seja, se $|X|$ é par, trata-se do itinerário de uma pré-imagem de $(0, 0)$, se $|X|$ é ímpar, é o itinerário de uma pré-imagem de $(0, 1)$. Para $(X_{[1,n]}0, 0)$ temos os seguintes cenários:

- Se $|X|$ é par e $|K^0|$ é par, a seqüência $X_{[1,n]}0$ pode ser identificada com uma das seqüências $X_{[1,n]}(LK^0)^\infty$ ou $X_{[1,n]}(RK^0)^\infty$
- Se $|X|$ é par, $|K^0|$ é ímpar e $|K^1|$ é par, a seqüência $X_{[1,n]}0$ pode ser identificada com uma das seqüências $X_{[1,n]}K^0(LK^1)^\infty$ ou $X_{[1,n]}K^0(RK^1)^\infty$
- Se $|X|$ é par, $|K^0|$ é ímpar e $|K^1|$ é ímpar, a seqüência $X_{[1,n]}0$ pode ser identificada com uma das seqüências $X_{[1,n]}(LK^0LK^1)^\infty$ ou $X_{[1,n]}(RK^0RK^1)^\infty$
- Se $|X|$ é ímpar e $|K^1|$ é par, a seqüência $X_{[1,n]}0$ pode ser identificada com uma das seqüências $X_{[1,n]}(LK^1)^\infty$ ou $X_{[1,n]}(RK^1)^\infty$

- Se $|X|$ é ímpar, $|K^0|$ é ímpar e $|K^1|$ é ímpar, a sequência $X_{[1,n]}0$ pode ser identificada com uma das sequências $X_{[1,n]}(LK^1LK^0)^\infty$ ou $X_{[1,n]}(RK^1RK^0)^\infty$
- Se $|X|$ é ímpar, $|K^0|$ é par e $|K^1|$ é ímpar, a sequência $X_{[1,n]}0$ pode ser identificada com uma das sequências $X_{[1,n]}LK^1(LK^0)^\infty$ ou $X_{[1,n]}RK^1(RK^0)^\infty$

De forma análoga podemos estabelecer cenários correspondentes para o par $(X_{[1,n]}0, 1)$.

4.4 Esqueleto no espaço de parâmetros

A ideia de representar curvas no espaço de parâmetros relacionadas com invariantes de amassamento e estudar a sua estrutura foi introduzida por MacKay e Tresser em [MT88a] para funções bimodais. Vários autores como A. Radulescu, ([Rad04]), e L. Silva e Sousa Ramos, ([SR01]), descreveram estruturas deste tipo para outras famílias de funções a dois parâmetros. Nesta secção vamos descrever brevemente o *esqueleto* associado à família de sistemas não autónomos F de período 2, que resultam da iteração alternada das tendas cortadas (f^v, f^w) , família a dois parâmetros $p = (v, w) \in [-1, 1]^2$. Dada uma sequência maximal X com $|X| < \infty$, quais os pares (v, w) no espaço de parâmetros $[-1, 1]^2$ para os quais o invariante de amassamento K_p tem uma das coordenadas X ? A descrição de *osso* em [Rad04] é dada apenas para ossos de comprimento par e baseia-se em permutações dos termos das órbitas críticas. Havendo diferenças em relação aos ossos definidos pelos itinerários críticos, — por exemplo um osso correspondente a determinada permutação corresponde a dois itinerários críticos, — a estrutura é essencialmente a mesma.

Definição 4.4.1. *Dada uma sequência X maximal de comprimento finito, chamamos osso esquerdo associado a X ao conjunto de pontos $p = (v, w)$ em $[-1, 1]^2$ tais que X coincide com a primeira coordenada do invariante de amassamento K_p e chamamos osso direito associado a X ao conjunto de pontos (v, w) em $[-1, 1]^2$ tais que K_p tem como segunda coordenada X . Escrevemos $B^0(X)$ e $B^1(X)$ para osso esquerdo e osso direito respectivamente, isto é, $B^0(X) = \{p = (v, w) : K_p^0 = X\}$ e $B^1(X) = \{p = (v, w) : K_p^1 = X\}$. Ao conjunto de todos os ossos chamamos *esqueleto*.*

Da definição de osso esquerdo e da descrição apresentada no fim da secção anterior resulta que um osso de comprimento par corresponde a um conjunto de pontos onde o itinerário simbólico de $(0, 0)$ é constante, pelo que podemos tirar conclusões sobre a existência de bifurcações nas vizinhanças dos ossos, embora esse não seja o objectivo deste trabalho. Se o comprimento do osso é ímpar, o itinerário de $(0, 0)$ sofre alterações ao longo do osso, designadamente quando é intersectado por ossos direitos. Assim, do ponto de vista da dinâmica, os ossos de comprimento ímpar fornecem informação que tem de ser complementada com informação sobre os ossos direitos que os cruzam. Na Figura 4.4 estão representados os ossos de comprimento par menor ou igual a 6.

Definição 4.4.2. Chamamos *órbita crítica* a uma órbita que contenha o ponto $(0, 0)$ ou o ponto $(0, 1)$. Uma órbita diz-se *bicrítica* se contém $(0, 0)$ e $(0, 1)$.

No resultado seguinte propomos uma descrição construtiva do esqueleto simbólico, da qual resultam as Figuras 4.3 e 4.4:

Teorema 4.4.3. Dada uma sequência X maximal, com $|X| = 2n$, existem $w_1, v_1 \in [-1, 1]$ e um intervalo I_{v_1} tais que $B^0(X) = \{\Phi(X)\} \times [w_1, 1] \cup I_{v_1} \times \{w_1\}$.

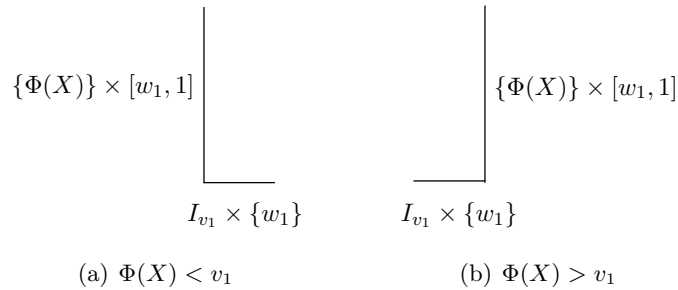


Figura 4.3: Representação dos dois formatos de ossos esquerdos.

Demonstração. Fazendo $w_1 = \Phi(\max_k(\sigma^{2k-1}(X)))$ para $k \in \mathbb{N}$ e $2k - 1 < |X|$, resulta que o sistema F correspondente ao par $(\Phi(X), w_1)$ tem itinerário crítico de $(0, 0)$ dado por X , evidentemente *deslizando* o patamar de f_1 no intervalo $[w_1, 1]$ o itinerário de $(0, 0)$ não é alterado, o que justifica o segmento vertical do osso. Por outro lado, tomemos v_1 tal que $v_1 = \Phi(X_1 \dots X_{n-1}0)$ onde n é tal que $\sigma^n(X) = \max_k(\sigma^{2k-1}X)$ e $k < \frac{|X|+1}{2}$. A sequência $X_1 \dots X_{n-1}0$ atinge o ponto 0 numa iteração ímpar pelo que, sob iteração do par (f^{v_1}, f^{w_1}) , 0 realiza uma órbita bicrítica, claramente $\Phi(X_1 \dots X_{n-1}0)$ não pertence ao osso $B^0(X)$. Se $X < X_1 \dots X_{n-1}0$ então

$$\Phi(X) < \Phi(X_1 \dots X_{n-1}0) = v_1.$$

Nesse caso, ao longo do segmento $\{w_1\} \times I_{v_1} =]\Phi(X), v_1[$ a órbita de $(0, 0)$ percorre, na n -ésima iteração ímpar, o patamar de f_1 entre um dos seus extremos e o ponto $x = 0$, prosseguindo em seguida com o itinerário de $f_1(0)$, $X_{n+1} \dots X_{|X|-1}0$, o que justifica a parte horizontal do osso. Se $X > X_1 \dots X_{n-1}0$ então $I_{v_1} =]v_1, \Phi(X)[$ e o resultado segue. \square

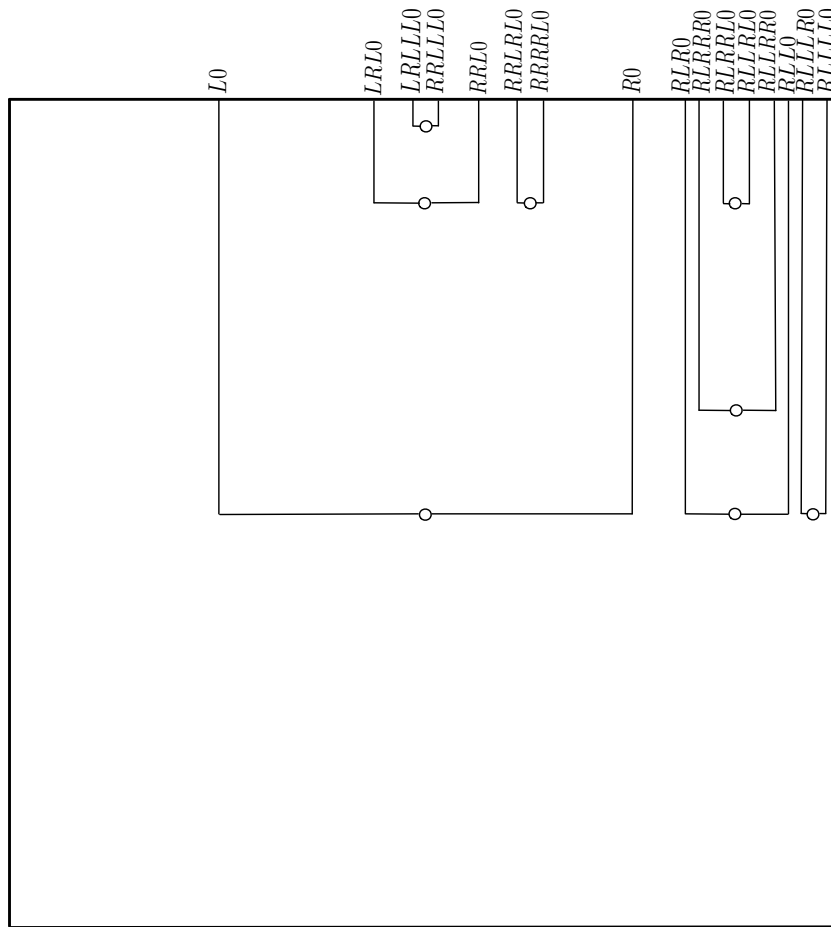


Figura 4.4: Esqueleto esquerdo de ordem par menor ou igual a 6.

Assim, cada osso esquerdo de comprimento par é a união de dois segmentos, um vertical com um extremo na recta $w = 1$ e outro horizontal. Ao longo do segmento vertical a órbita de $(0,0)$ não atinge patamares a não ser eventualmente nos extremos. Ao longo do segmento horizontal a órbita de $(0,0)$ percorre o patamar de f_1 entre um dos seus extremos e o ponto $x = 0$; nesse ponto o itinerário deixa de realizar a mesma sequência pelo que cada osso tem um extremo aberto que é fronteira com outro osso, Figura 4.4. Cada um destes pontos corresponde a uma órbita bicrítica e portanto a um osso ímpar.

recta $w = \frac{1}{3}$ para $F(\cdot, 0)$, representado na Figura 4.6(a). Para valores do parâmetro v tais que $-1 < v < -\frac{1}{3}$, existe um ciclo atractivo de período 2, $\{v, f^{\frac{1}{3}}(v)\}$, que corresponde a um ponto fixo de $f^{\frac{1}{3}} \circ f^v$. Para $-\frac{1}{3} \leq v \leq \frac{1}{3}$ a órbita de 0 entra no patamar de $f^{\frac{1}{3}}$ e o ciclo atractivo de período 2 é $\{v, \frac{1}{3}\}$. Naturalmente, para $v = \frac{1}{3}$ o ciclo de período 2 colapsa num único ponto fixo. Para $\frac{1}{3} < v < 0.6$ mantém-se o ciclo atractivo de período 2, $\{v, f^{\frac{1}{3}}(v)\}$. Em $v = 0.6$, o ciclo de período 2 torna-se repulsivo, surgindo um ciclo atractivo de período 4 para $v > 0.6$, $\{v, f^{\frac{1}{3}}(v), f^v(f^{\frac{1}{3}}(v)), f^{\frac{1}{3}}(f^v(f^{\frac{1}{3}}(v)))\}$. A criação do período 4 pode ser interpretada a partir dos gráficos de $F^4(\cdot, 0)$ para valores de v próximos de 0.6, representados na Figura 4.7.

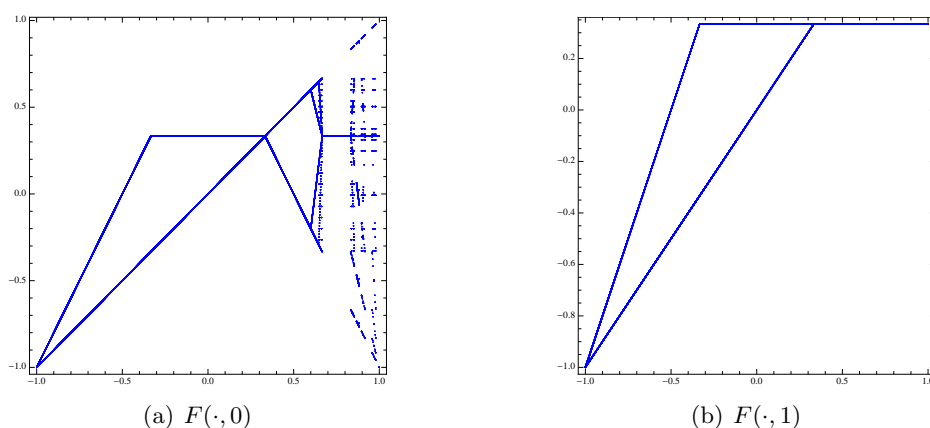


Figura 4.6: Diagramas de bifurcação ao longo da recta $w = \frac{1}{3}$.

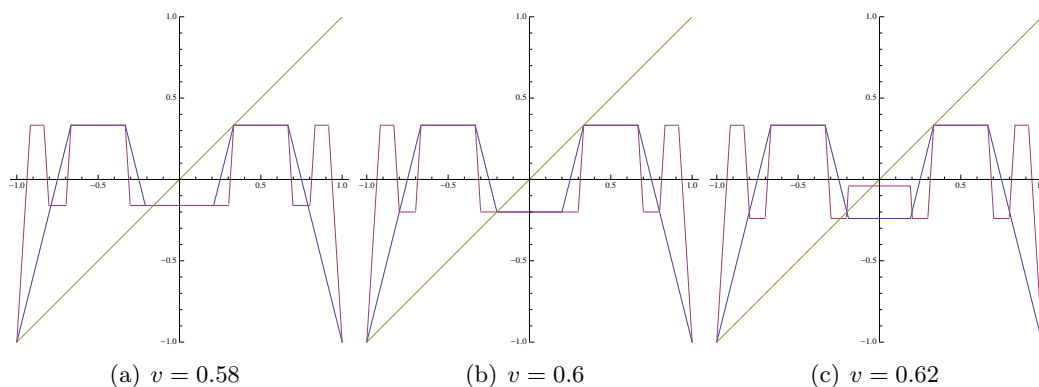


Figura 4.7: Comportamento de $F^2(\cdot, 0)$, a azul, e $F^4(\cdot, 0)$, a roxo, para diferentes valores de v próximos de $v = 0.6$.

$F(\cdot, 1)$ tem uma dinâmica mais simples (Figura 4.6(b)). Para valores do parâmetro v próximos de -1 existe igualmente um ciclo de período 2 atractivo, $\{v, f^{\frac{1}{3}}(v)\}$. Este ciclo mantém-se até $v = \frac{1}{3}$, a partir desse valor o sistema tem um único ponto fixo atractivo

uma vez que $f^v(\frac{1}{3}) = f^{\frac{1}{3}}(\frac{1}{3}) = \frac{1}{3}$.

Podemos concluir que o comportamento dinâmico do sistema depende em grande medida do comportamento dinâmico dos patamares. A descrição apresentada neste exemplo baseia-se no trabalho desenvolvido em [SSR15].

Na Figura 4.8 apresentamos diagramas de bifurcação unidimensionais segundo algumas rectas horizontais do espaço de parâmetros.

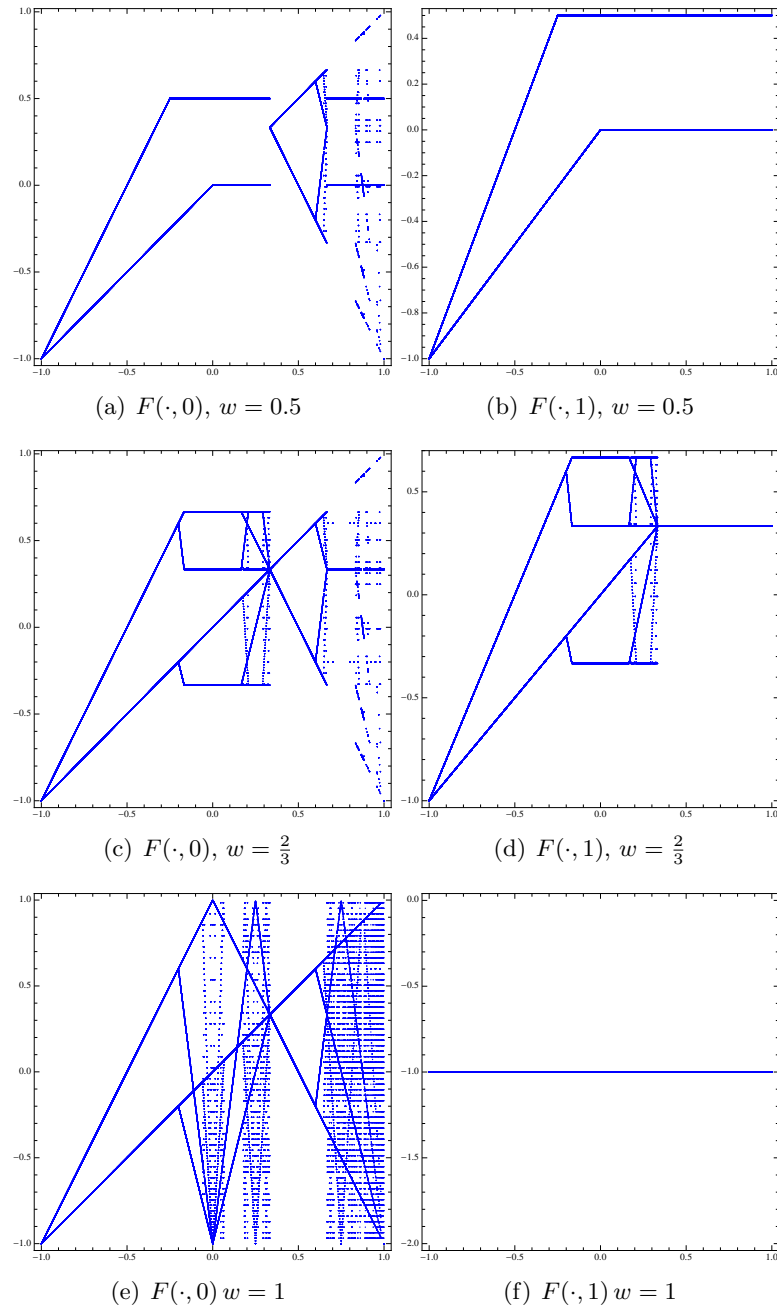


Figura 4.8: Diagramas de bifurcação segundo algumas rectas do tipo $w = a$.

4.5 Renormalização e produto-*

Em 1978, Derrida, Gervois e Pomeau [DGP78] introduziram o produto-* aplicado aos invariantes de amassamento de famílias unimodais. Esta ferramenta algébrica permite gerar sequências admissíveis novas a partir de sequências conhecidas, está directamente relacionada com a renormalização das aplicações e, conseqüentemente, com a existência de auto-semelhanças no diagrama de bifurcação. Em [FSS13] introduziu-se o conceito de renormalização e definiu-se produto-* no contexto dos sistemas não autónomos periódicos, em particular no caso das funções monótonas por troços. Provou-se que, à semelhança do que é conhecido para sistemas autónomos, (ver [SR02]), há uma correspondência entre a criação de zonas de renormalização no espaço de parâmetros e a decomposição das sequências simbólicas em factores-*. Seguindo de perto [FSS13] estabelecemos as seguintes definições:

Definição 4.5.1. *Consideremos o sistema dinâmico não autónomo de período 2, F , gerado pela iteração de $(f_0, f_1) = (f^v, f^w)$. Dizemos que F é i -renormalizável, $i \in \{0, 1\}$, se existe n par e $a > 0$, tais que, a função $R_i^n(F) : [-a, a] \rightarrow [-a, a]$ definida por*

$$R_i^n(F)(x) = \pi(F^n(x, i))$$

é conjugada de uma função da família das funções tenda cortada. A função $R_i^n(F)$ diz-se uma i -renormalização de F . Dizemos que F é renormalizável se é i -renormalizável para ambos $i = 0, 1$.

Exemplo 4.5.1. *Consideremos o par $(v, w) = (\frac{5}{8}, \frac{3}{4})$ e o sistema dinâmico não autónomo F , de período 2, gerado pela iteração de $(f_0, f_1) = (f^v, f^w)$. Neste caso só há i -renormalização para $i = 0$ como mostra a Figura 4.9 onde estão representados os gráficos de $\pi(F^2(\cdot, 0)) = f^w \circ f^v$ e $\pi(F^2(\cdot, 1)) = f^v \circ f^w$ para o par $(v, w) = (\frac{5}{8}, \frac{3}{4})$.*

Exemplo 4.5.2. *Na Figura 4.10 estão representados os gráficos de $\pi(F^4(\cdot, 0)) = f^w \circ f^v \circ f^w \circ f^v$ e $\pi(F^4(\cdot, 1)) = f^v \circ f^w \circ f^v \circ f^w$ para o par $(v, w) = (\frac{23}{128}, \frac{113}{128})$. Observa-se que o intervalo de renormalização $I_0 = [-a, a]$ é invariante por $F^4(\cdot, 0)$ e $I_1 = [-b, b]$ é invariante por $F^4(\cdot, 1)$, isto é, $\pi(F^4(I_0, 1)) \subset I_0$ e $\pi(F^4(I_1, 0)) \subset I_1$. Uma vez que F é i -renormalizável para $i = 0$ e $i = 1$, e ambas as renormalizações são de ordem $n = 4$, existem outros intervalos onde $F^4(\cdot, 0)$ e $F^4(\cdot, 1)$ são invariantes, neste caso existem quatro intervalos, com interiores separados, em cada uma das fibras. Note-se que no exemplo anterior isto não acontecia.*

Para a construção do produto-* e resultados seguintes precisamos de definir a noção de admissibilidade no contexto autónomo. Em [FSS13] a noção de admissibilidade é desde logo definida num contexto mais abrangente, que aqui seria desadequado. Assim optámos

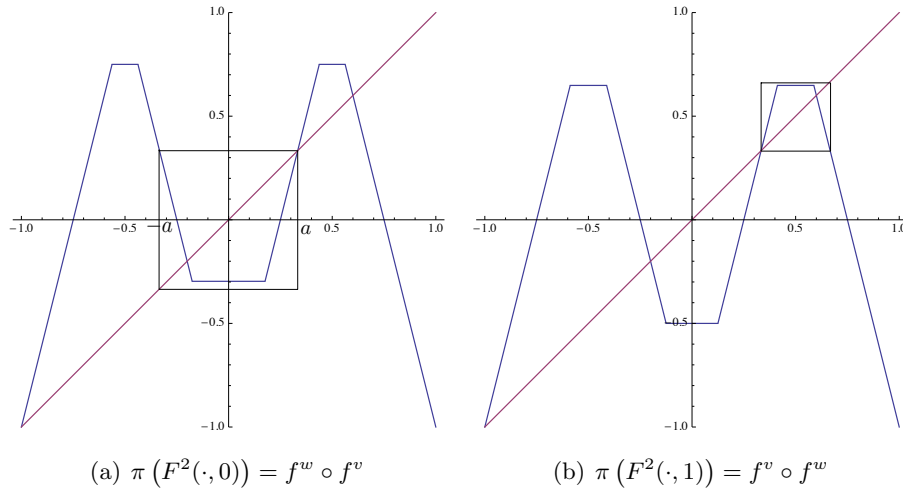


Figura 4.9: Intervalo de renormalização para $(v, w) = (\frac{5}{8}, \frac{3}{4})$.

por apresentar inicialmente as definições no contexto não autónomo e nesta secção estabelecemos as definições no contexto autónomo, para a iteração de uma função *tenda cortada* e de uma função *tenda cortada invertida* definida por

$$f_v(x) = \begin{cases} -1 + 2x, & x > \frac{1-v}{2} \\ -v, & \frac{v-1}{2} \leq x \leq \frac{1-v}{2} \\ -1 - 2x, & x < \frac{v-1}{2} \end{cases} \quad (4.8)$$

com $v \in [-1, 1]$.

Definição 4.5.2. Dizemos que uma sequência $X \in \Sigma$ é *admissível-(+, -)* se é o itinerário crítico, $i_f(0)$, pela iteração de uma função *tenda cortada* definida por 4.1. Dizemos que X é *admissível-(-, +)* se é o itinerário crítico, $i_f(0)$, pela iteração de uma *tenda cortada invertida*, ou seja, uma função do tipo 4.8. Representamos por $\Sigma_{(\alpha, -\alpha)}$ o conjunto das sequências *admissíveis*-($\alpha, -\alpha$), com $\alpha \in \{-, +\}$.

As funções tipo *tenda cortada* e *tenda cortada invertida* também são chamadas funções de tipo (+, -) e (-, +) respectivamente.

Definição 4.5.3. Dada uma sequência Y em $\Sigma_{(\alpha, -\alpha)}$, com $\alpha \in \{-, +\}$, dizemos que X é *Y-admissível* se para uma função f de tipo $(\alpha, -\alpha)$ com invariante de amassamento Y , existe algum ponto $x \in [-1, 1]$ tal que $it_f(x) = X$ e representamos o conjunto das sequências *Y-admissíveis* por $\Sigma_{(\alpha, -\alpha)}(Y)$.

Nota. Identificamos naturalmente os sinais $\{-, +\}$ com os valores $\{-1, 1\}$ sempre que seja conveniente, nomeadamente quando α resulta de produtos de sinais associados a símbolos, $\varepsilon(X_i)$, nas sequências simbólicas.

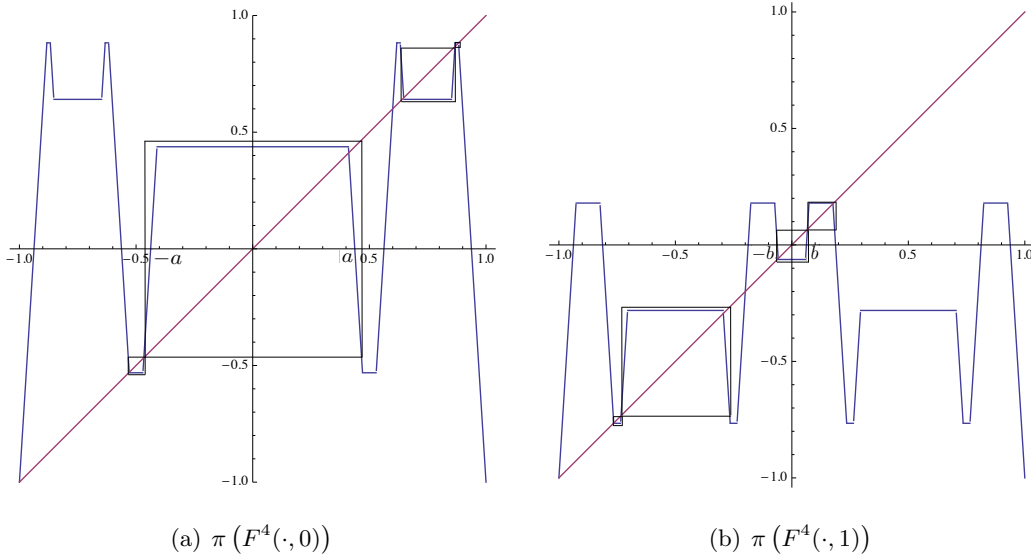


Figura 4.10: Intervalos de renormalização para $(v, w) = \left(\frac{23}{128}, \frac{113}{128}\right)$

Recordamos a definição de produto-*, no contexto da iteração de funções do tipo tenda cortada, entre uma seqüência finita $X = X_1 \dots X_{n-1} 0$ e uma seqüência $Y = Y_1 \dots Y_k \dots$ em Σ :

$$X * Y = X_{[1, n-1]} Y_1 X_{[1, n-1]} Y_2 \dots X_{[1, n-1]} Y_k X_{[1, n-1]} \dots$$

Seguindo [FSS13] temos a seguinte generalização do produto-* no contexto não autônomo:

Definição 4.5.4. *O produto- $*$ _j entre um par de seqüências finitas (X^0, X^1) , com $|X^j| \bmod 2 = 0$, e uma seqüência $Y = Y_1 \dots Y_k \dots$ é o par*

$$(X^0, X^1) *_{j} Y = (X^0 *_{j} Y, X^1 *_{j} Y)$$

onde $X^i *_{j} Y = X^i$ se $i \neq j$, e $X^i *_{j} Y = X^i * Y$ se $i = j$.

Naturalmente coloca-se a questão de saber quais as seqüências admissíveis resultantes do produto-*. O resultado seguinte responde cabalmente a esta questão:

Teorema 4.5.5. *Dados um par $(X^0, X^1) \in \Sigma_2^+$ e uma seqüência $Y \in \Sigma$, temos $(X^0, X^1) *_{j} Y \in \Sigma_2^+$, com j tal que $|X^j| \bmod 2 = 0$, se e só se $Y = Y_1 \dots Y_n 0 \in \Sigma_{(\alpha, -\alpha)}^+$ onde $\alpha = \varepsilon_{|X^j|-1}(X^j)$.*

Os argumentos da prova são totalmente análogos aos apresentados em [FSS13].

Definição 4.5.6. *O produto-* entre um par de seqüências finitas (X^0, X^1) , com $|X^j| \bmod 2 = 0$, para $j \in \{0, 1\}$, e um par de seqüências $(Y^0, Y^1) \in \Sigma$ é o par*

$$(X^0, X^1) * (Y^0, Y^1) = (X^0 * Y^0, X^1 * Y^1).$$

Corolário 4.5.7 (Teorema 4.5.5). *Dados os pares de sequências $(X^0, X^1) \in \Sigma$, com $|X^j| \bmod 2 = 0$ e $(Y^0, Y^1) \in \Sigma$, o produto $(X^0, X^1) * (Y^0, Y^1) = (X^0 * Y^0, X^1 * Y^1)$ é admissível se e só se $(X^0, X^1) \in \Sigma_2^+$ e $(Y^0, Y^1) \in \Sigma_{(\alpha(0), -\alpha(0))} \times \Sigma_{(\alpha(1), -\alpha(1))}$ onde $\alpha(i) = \varepsilon_{|X^i|-1}(X^i)$.*

Demonstração. Resulta imediatamente de $(X^0, X^1) * (Y^0, Y^1) = ((X^0, X^1) *_0 Y^0) *_1 Y^1$. \square

O resultado seguinte relaciona a existência de intervalos de renormalização num sistema dinâmico não autónomo F , de período 2, com invariante de amassamento (X^0, X^1) , com a redutibilidade de cada uma das sequências em factores-*. Os argumentos da demonstração sobre o resultado análogo no contexto não autónomo, para funções estritamente monótonas por troços, (ver [FSS13]), podem ser reproduzidos no nosso contexto.

Teorema 4.5.8. *Consideremos o sistema dinâmico não autónomo F , de período 2, gerado pela iteração de $(f_0, f_1) = (f^v, f^w)$, com invariante de amassamento (K^0, K^1) . F é i -renormalizável no intervalo $[-a, a]$ com renormalização $R_i^n(F)$ se e só se existem $(X^0, X^1) \in \Sigma_2^+$ com $|X^i| = n$, par, e $Y \in \Sigma_{(\alpha, -\alpha)}$ onde $\alpha = \varepsilon_{|X^j|-1}(X^j)$ tal que $(K^0, K^1) = (X^0, X^1) *_i Y$. Mais, Y é o invariante de amassamento de $R_i^n(F)$.*

O resultado anterior permite obter uma decomposição do espaço de fases em relação à combinatória da dinâmica.

Corolário 4.5.9. *Consideremos um sistema dinâmico não autónomo F , de período 2, renormalizável em $[-a, a]$ com invariante de amassamento $K[F] = (K^0, K^1) = (X^0, X^1) *_j Y$. Então $\pi(F^{|X^j|}([-a, a], j)) \subset [-a, a]$ pelo que*

$$\mathcal{R} = \bigcup_{i=0}^{|X^j|-1} F^i([-a, a], (i+j) \bmod 2)$$

*é uma união de intervalos nas fibras, invariantes por F . Mais, se $x \in [-a, a]$ então existe $Z \in \Sigma_{(\alpha, -\alpha)}^+(Y)$, com $\alpha = \varepsilon(Y)$ tal que $i_F(x, j) = (X^0, X^1) *_j Z$. Se $(x, i) \in [-1, 1] \times \{0, 1\} \setminus \mathcal{R}$ e $F^k(x, i) \notin \mathcal{R}$ para nenhum k então $i_F(x, i) \in \Sigma^+(X^0, X^1)$.*

Exemplo 4.5.3. *Consideremos o sistema não autónomo F com invariante de amassamento $K = (RLR0, R0)$, a que correspondem o par de funções $(f_0, f_1) = (f^v, f^w)$ com $(v, w) = (\frac{5}{8}, \frac{1}{2})$. Temos $RLR0 = R0 * L0$, logo $(RLR0, R0) = (R0, R0) *_0 L0$. Pelo Teorema 4.5.8, F é 0-renormalizável com renormalização $R_0^2[F]$ cujo invariante de amassamento é $L0$. Na Figura 4.11(a) temos o gráfico de $\pi(F^2(\cdot, 0))$ onde se identifica o intervalo de renormalização $[-\frac{1}{3}, \frac{1}{3}]$. Na Figura 4.11(b) temos também o gráfico de $R_0^2[F]$ no intervalo $[-\frac{1}{3}, \frac{1}{3}]$. Da análise gráfica verifica-se que o itinerário crítico por $R_0^2[F]$ é $L0$.*

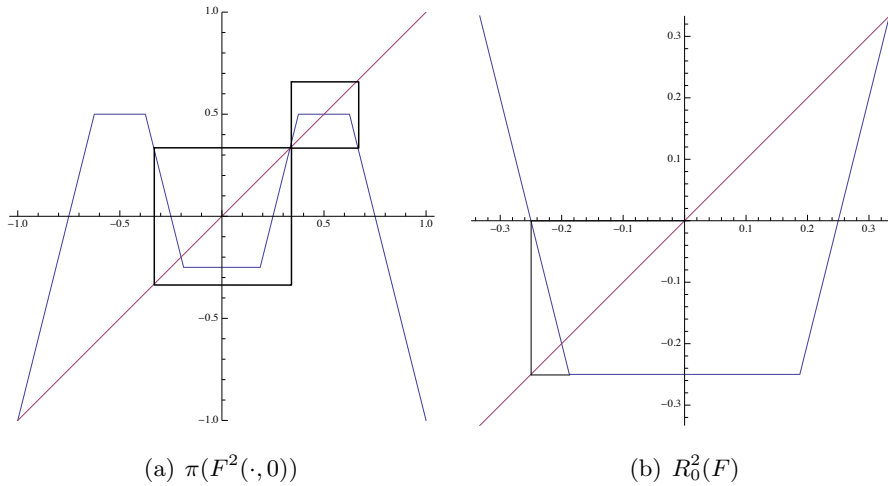


Figura 4.11: Renormalização

Exemplo 4.5.4. Consideremos novamente o sistema do exemplo 4.5.2. O invariante de amassamento correspondente é $(X^0, X^1) = (RRLRRRL0, RLLLRL0)$. De acordo com o Teorema 4.5.8 ambas as seqüências são decomponíveis em factores-*, com efeito $(X^0, X^1) = (RRL0, RLLLRL0) *_0 L0$ e $(X^0, X^1) = (RRLRRRL0, RRL0) *_1 R0$. Pela leitura do gráfico a renormalização na fibra I_0 , $R_0^4(F)$, admite itinerário crítico $L0$ e a renormalização na fibra 1, $R_1^4(F)$, admite itinerário crítico $R0$, além disso, em cada uma das fibras, os intervalos invariantes representados têm uma relação compatível com a seqüência simbólica do factor esquerdo do produto-*.

Exemplo 4.5.5. De regresso ao exemplo 4.5.1, à luz do Teorema 4.5.8, observamos que o par de seqüências de amassamento correspondentes a $(\frac{5}{8}, \frac{3}{4})$, $(X^0, X^1) = (RLR0, RL0)$, se factoriza no produto-*: $(RLR0, RL0) = (R0, RL0) *_0 L0$. O teorema garante a existência do intervalo de renormalização $[-a, a]$ na fibra $i = 0$, mas não de um segundo intervalo invariante como mostra este exemplo. (A existência de um intervalo de renormalização numa fibra implica que existam intervalos invariantes na outra fibra.)

4.6 Taxas de convergência

Nesta secção apresentamos o resultado central do capítulo, sobre a convergência de sucessões construídas a partir do produto-* de seqüências simbólicas, que estão directamente relacionadas com a estrutura de bifurcação do sistema não autónomo estudado nas secções anteriores. Provamos a convergência e calculamos explicitamente o limite de sucessões de parâmetros correspondentes a pontos de bifurcação, mostrando que não depende do ponto inicial onde se inicia a iteração. Em [Sil07], L. Silva calculou o limite para sucessões do mesmo tipo num sistema autónomo de famílias de funções de Lorenz a dois parâmetros,

nesta secção seguimos de perto a abordagem desse artigo. Os resultados apresentados são originais e podem ser encontrados em [SSF14a].

Como referimos no Capítulo 1, Feigenbaum conjecturou e deu fortes evidências experimentais de que, para uma vasta classe de famílias unimodais, a sucessão λ_n de parâmetros correspondentes a bifurcações de duplicação de período tem uma taxa de convergência universal. Para a função logística, $f(x) = \lambda x(1 - x)$, o limite

$$\lim \frac{\lambda_n - \lambda_{n-1}}{\lambda_{n+1} - \lambda_n} = 4.669 \dots$$

corresponde a uma taxa de convergência de tipo linear, dita geométrica, uma vez que satisfaz

$$\lim \frac{\lambda_n - \lambda_{n-1}}{(\lambda_{n+1} - \lambda_n)^p} = \begin{cases} 0 & \text{se } p > l \\ +\infty & \text{se } p < l \end{cases} \quad (4.9)$$

para $l = 1$.

Em [BS82a], Beyer e Stein estudaram a convergência numa família de tendas cortadas a um parâmetro, λ , definidas no intervalo $[0, 2]$ por

$$\lambda g_e(x) = \begin{cases} \frac{x}{e} & \text{se } x < e \\ 1 & \text{se } e \leq x \leq 2 - e \\ 2 - \frac{x}{e} & \text{se } x > 2 - e \end{cases}$$

com $e \in [0, 0.99]$. Usando estimativas obtidas por computador, provaram que, se $\lambda_n(e)$ é o menor valor para o qual o itinerário crítico, $it(1)$, é um ciclo de período 2^n então

$$\lim \frac{\log \left(\frac{\lambda_{n+2}(e) - \lambda_{n+1}(e)}{\lambda_{n+1}(e) - \lambda_n(e)} \right)}{\log \left(\frac{\lambda_{n+1}(e) - \lambda_n(e)}{\lambda_n(e) - \lambda_{n-1}(e)} \right)} = 2.$$

Em [BS82b], os mesmos autores provaram que a sucessão $\lambda_n(e)$ satisfaz a condição (4.9) para $l = 2$, ou seja, apresenta taxa de convergência de tipo quadrático.

Por outro lado, sabemos que uma bifurcação de duplicação de período corresponde à criação de um intervalo de renormalização, ou seja, à existência de um intervalo $J \subset I$, contendo o ponto crítico, onde F^2 é unimodal. Nesse sentido, e estabelecida a relação entre renormalização e produto-* no contexto não autónomo, vamos construir uma sucessão de pontos $(u_n, v_n)_{n \in \mathbb{N}}$ no espaço de parâmetros correspondente a uma sucessão de pares de potências (de produto-*) de sequências simbólicas, dando origem a uma sucessão de pares de funções tendas cortadas, renormalizáveis em cada uma das fibras.

Podemos considerar valores de x cujas órbitas não intersectam os patamares, a não ser eventualmente nos extremos, uma vez que os pares de parâmetros a considerar cor-

respondem a intersecções de ossos pares, que acontecem nos segmentos verticais dos ossos esquerdos e segmentos horizontais dos ossos direitos.

Consideremos $u \in [-1, 1]$, $j \in \{0, 1\}$, e o itinerário (X, j) de (x, j) , com $T(x) = u$, onde $X = X_1 \dots X_{n-1}0$ de comprimento $|X| = n$. Temos

$$T^X(u) = T^{X|X|-1} \circ \dots \circ T^{X_1}(u)$$

onde $T^R = T^+$ e $T^L = T^-$, definidas em 4.2.

Recordemos que o sinal $\varepsilon_i(X)$ é dado por $\varepsilon_i(X) = \prod_{j=1}^i \varepsilon(X_j)$, onde $\varepsilon(X_j)$ é o sinal associado ao símbolo X_j , tendo em conta a monotonia das funções da família das *tendas cortadas* nos intervalos da partição, ou seja

$$\varepsilon(X_j) = \begin{cases} 1, & X_j = L \\ -1, & X_j = R \end{cases}$$

Escrevendo $\bar{\varepsilon}_i(X) = \prod_{j=i}^{|X|-1} \varepsilon(X_j)$, temos

$$T^X(u) = 1 + \sum_{i=2}^{|X|-1} 2^{|X|-i} \bar{\varepsilon}_i(X) + 2^{|X|-1} \varepsilon_{|X|-1}(X)u, \quad (4.10)$$

Se $T^X(u) = 0$, temos

$$u = -\frac{1 + \sum_{i=2}^{|X|-1} 2^{|X|-i} \bar{\varepsilon}_i(X)}{2^{|X|-1} \varepsilon_{|X|-1}(X)}. \quad (4.11)$$

Definição 4.6.1. Dada uma sequência $X = X_1 X_2 \dots X_n$, com $n \in \mathbb{N}$ ou $n = +\infty$, a sequência $-X$ é definida por $-X = (-X)_1 (-X)_2 \dots (-X)_n$, onde

$$\begin{cases} -L = R \\ -R = L \\ -0 = 0 \end{cases}$$

Dado um par de sequências finitas $(X^0, X^1) \in \Sigma_2^+$, uma sequência $Y \in \Sigma_{(\alpha, -\alpha)}$, onde $\alpha = \varepsilon_{|X^j|-1}(X^j)$ e $j \in \{0, 1\}$ construímos uma sucessão de pares $(X^0(n), X^1(n))_j$ do seguinte modo:

$$(X^0(0), X^1(0))_j = (X^0, X^1), \quad (4.12)$$

$$(X^0(n+1), X^1(n+1))_j = \begin{cases} (X^0(n), X^1(n))_j * Y & \text{se } Y \in \Sigma_{(\alpha_n, -\alpha_n)} \\ (X^0(n), X^1(n))_j * (-Y) & \text{se } Y \in \Sigma_{(-\alpha_n, \alpha_n)} \end{cases} \quad (4.13)$$

onde $(\alpha_n, -\alpha_n)$ é tal que $\alpha_n = \varepsilon_{|X^j(n)|-1}(X^j(n))$.

Da definição de produto- \ast_j resulta que para $i \neq j$ temos $X^i(n) = X^i$, para todos os $n \in \mathbb{N}$.

Lema 4.6.2. *A sucessão $X^j(n)$ definida acima é estritamente crescente.*

Demonstração. Dada uma sucessão $X^j(n)$, suponhamos que $\varepsilon_{|X^j(n)|-1}(X^j(n)) = +$, (ou seja, o número de símbolos R em $X^j(n)$ é par), então, se $Y_1 = R$ temos $X^j(n+1) = X^j(n) \ast Y$, caso contrário temos $X^j(n+1) = X^j(n) \ast (-Y)$, em qualquer dos casos $X^j(n+1) = X^j(n)_1 \dots X^j(n)_{|X^j(n)|-1} R \dots > X^j(n)_1 \dots X^j(n)_{|X^j(n)|-1} 0 = X(n)$. Se o número de símbolos R é ímpar o argumento é análogo. \square

Lema 4.6.3. *Sejam $(X^0, X^1) \in \Sigma_2^+$, $Y \in \Sigma_{(\alpha, -\alpha)}$, onde $\alpha = \varepsilon_{|X^j|-1}(X^j)$ e $j \in \{0, 1\}$, então a sucessão $X^j(n)$ definida acima satisfaz*

$$\frac{|X^j(n+1)|}{|X^j(n)|} = |Y|.$$

Demonstração. Basta ver que

$$|X^j(n+1)| = |X^j(n)_{[1, |X^j(n)|-1]} Y_1 X^j(n)_{[1, |X^j(n)|-1]} Y_2 \dots Y_{|Y|-1} X^j(n)| = |Y| |X^j(n)|.$$

\square

Exemplo 4.6.1. *Consideremos o par de amassamento $(R0, X^1)$ e a sequência $Y = L0$. Definindo a sucessão de sequências simbólicas com sucessivos produtos- \ast através da igualdade (4.12), consideremos as primeiras coordenadas dos primeiros quatro termos da sucessão $(X^0(n), X^1(n))_0$ e os respectivos parâmetros apresentados na Tabela 4.6.1:*

$X^0(n)$	$x_n = \Phi(X^0(n))$
$R0$	0.5
$R0 \ast L0 = RLR0$	0.625
$RLR0 \ast R0 = RLRRRLR0$	0.648438
$RLRRRLR0 \ast L0 = RLRRRLRLRLRRRLR0$	0.649811

Tabela 4.1:

A sucessão de sequências simbólicas $(X^0(n), X^1(n))_0$ corresponde a uma sucessão de pontos (x_n, w) localizados na recta horizontal do espaço de parâmetros, $w = \Phi(X^1)$.

Consideremos $w = 1$. Na Figura 4.8(e) está representado o diagrama de bifurcação de $F(\cdot, 0)$ ao longo da recta $w = 1$. No entanto, para melhor identificarmos os pontos de duplicação de período correspondentes às sequências simbólicas da Tabela 4.2 apresentamos algumas ampliações do diagrama na Figura 4.12.

$Z(n)$	$z_n = \Phi(Z(n))$	Período
$(RL)^\infty$	0.6	4
$(RLRR)^\infty$	0.6470588	8
$(RLRRRLRL)^\infty$	0.64980544708	16

Tabela 4.2: Alguns pontos de duplicação de período e respectivas sequências simbólicas.

Em $v = 0.6$ é criado um ciclo atractivo de período 4, em $v = 0.6470588$ é criado um ciclo atractivo de período 8, em 0.64980544708 é criado um ciclo atractivo de período 16, como se pode ver ampliando a respectiva faixa do espaço de parâmetros, Figura 4.12.

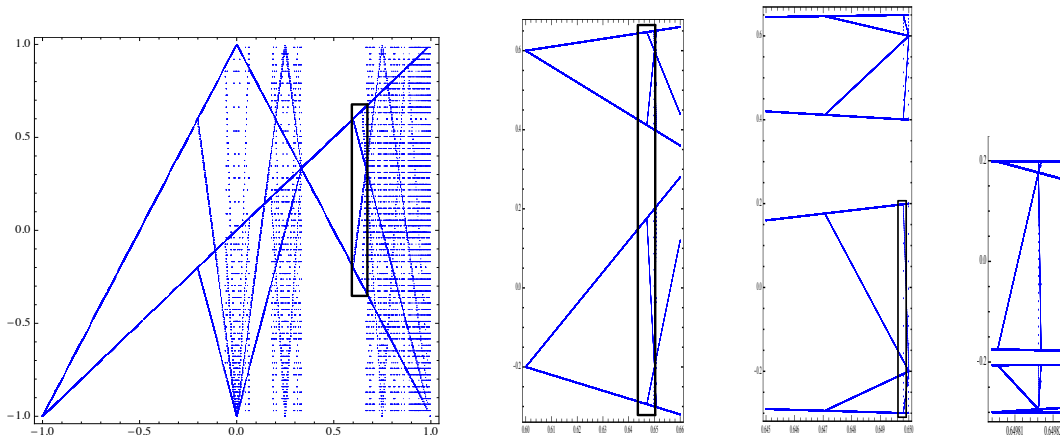


Figura 4.12: Detalhe do diagrama de bifurcação de $F(\cdot, 0)$, $w = 1$.

Na primeira imagem vemos uma zona com um ciclo de período 2 atractivo entre 0.5 e 0.6. Na segunda representação, intervalo $[0.6, 0.66]$, vemos claramente a formação de um ciclo de período 4. Na terceira representação, intervalo $[0.645, 0.65]$, vemos a criação de um ciclo de período 8 e na última imagem, intervalo $[0.649804, 0.64982]$, vemos a formação de um ciclo de período 16, embora só esteja representada a parte inferior do diagrama. Os pontos x_n e os pontos v_n estão relacionados mas não coincidem. Entre cada dois pontos consecutivos x_n e x_{n+1} da sucessão de parâmetros correspondentes a dois valores da sucessão de simbólica, $X^0(n)$ e $X^0(n+1)$, existe um valor, v_n , de bifurcação. No teorema seguinte veremos que a sucessão x_n converge com uma determinada taxa de convergência, poderemos assim concluir que os pontos de bifurcação convergem com a mesma taxa de convergência. Na Figura 4.13 conseguimos detectar esta relação para os primeiros valores.

A criação da duplicação do período pode ser entendida através dos gráficos das composições de $F(\cdot, 0)$. Uma vez que $w = 1$ está fixo, escrevemos F_v para designar $F(\cdot, 0)$

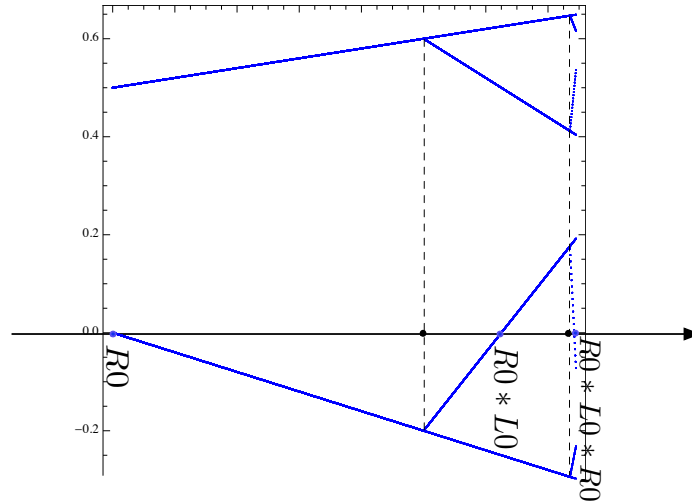


Figura 4.13: Representação, no intervalo de parâmetros $[0.5, 0.649]$, dos primeiros termos da sequência $X^0(n)$.

correspondente aos parâmetros $(v, w) = (v, 1)$. Quando a intersecção da função F_v^2 com a diagonal acontece no patamar existe um ciclo de período 2 atractivo. Em particular, se a intersecção acontece no centro do patamar central, ou seja, em $x = 0$, o parâmetro v corresponde à sequência simbólica de comprimento 2, $R0$. Fazendo variar a altura do patamar, existe um valor a partir do qual a diagonal deixa de intersectar a função F_v^2 no patamar, e a intersecção passa a dar-se num ponto onde $|(F^2)'(x)| > 1$ tornando o ciclo de período 2 repulsivo. Isso acontece no valor v em que a altura do patamar de F_v^4 deixa de coincidir com a altura do patamar de F_v^2 e a diagonal mantém a intersecção com F_v^4 no seu patamar, criando assim um ciclo atractivo de período 4, ver Figura 4.15.

O mesmo fenómeno pode ser interpretado nos gráficos de F_v^4 e F_v^8 , desta vez para a criação de um ciclo atractivo de período 8.

Para finalizar, apresentamos o resultado obtido em [SSF15] sobre as taxas de convergência:

Teorema 4.6.4. *Dados $(X^0, X^1) \in \Sigma_2^+$, $Y \in \Sigma_{(\alpha, -\alpha)}$ e $j \in \{0, 1\}$ definimos como anteriormente a sucessão $X^j(n)$. A sucessão u_n^j de pontos no espaço de parâmetros tais que $T^{X^j(n)}(u_n^j) = 0$ é convergente e tem-se*

$$\lim_{n \rightarrow \infty} \frac{\log \left(\frac{u_{n+1}^j - u_n^j}{u_{n+2}^j - u_{n+1}^j} \right)}{\log \left(\frac{u_n^j - u_{n-1}^j}{u_{n+1}^j - u_n^j} \right)} = |Y|.$$

Nota. 1. Segundo [BEQ90], a sucessão u_n satisfaz a condição (4.9) para $l = |Y|$, donde

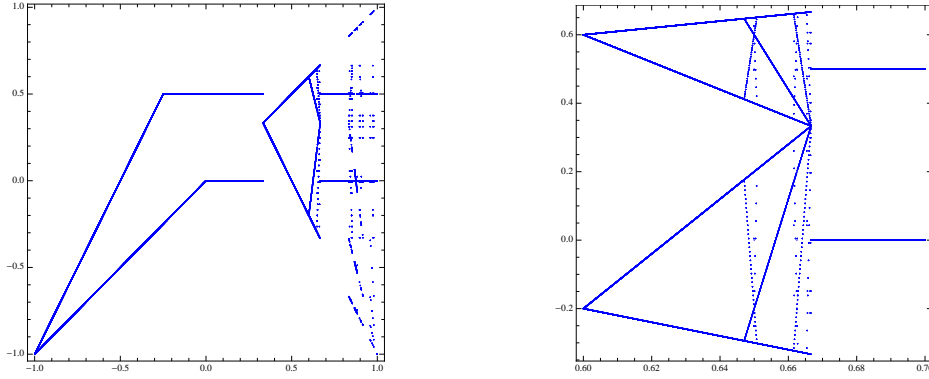


Figura 4.14: Diagramas de bifurcação de $F(\cdot, 0)$, com $w = 0.5$, para v no intervalo $[-1, 1]$, à esquerda e no intervalo $[0.6, 0.7]$ à direita.

a taxa de convergência é de ordem $|Y|$.

2. O sistema não autónomo $F(n)$, correspondente a um par de parâmetros (u_n, v) , é 0-renormalizável, com renormalização $R_j^{|Y|}(F)$. De facto, $F(n)$ é n vezes 0-renormalizável, com renormalizações $R_0^{i|Y|}(F)$, $i = 1 \dots n$.

Demonstração. Pelos lemas anteriores podemos concluir que a sucessão $(u_n^j)_{n \in \mathbb{N}}$ é monótona e limitada logo convergente, calculemos o seu limite.

A sucessão $(u_n^j)_{n \in \mathbb{N}}$ de pontos no espaço de parâmetros $[-1, 1]$ tais que $T^{X^j(n)}(u_n^j) = 0$, é dada por

$$u_n^j = - \frac{\sum_{i=2}^{|X^j(n)|-1} 2^{|X^j(n)|-i} \bar{\varepsilon}_i(X^j(n)) + 1}{2^{|X^j(n)|-1} \varepsilon_{|X^j(n)|-1}(X^j(n))}$$

logo

$$u_{n+1}^j = - \frac{\sum_{i=2}^{|X^j(n+1)|-1} 2^{|X^j(n+1)|-i} \bar{\varepsilon}_i(X^j(n+1)) + 1}{2^{|X^j(n+1)|-1} \varepsilon_{|X^j(n+1)|-1}(X^j(n+1))}.$$

Temos

$$X^j(n+1) = X^j(n) * Y = X^j(n)_{[1, |X^j(n)|-1]} Y_1 X^j(n)_{[1, |X^j(n)|-1]} Y_2 \dots X^j(n)$$

ou

$$X^j(n+1) = X^j(n) * (-Y) = X^j(n)_{[1, |X^j(n)|-1]} (-Y_1) X^j(n)_{[1, |X^j(n)|-1]} (-Y_2) \dots X^j(n),$$

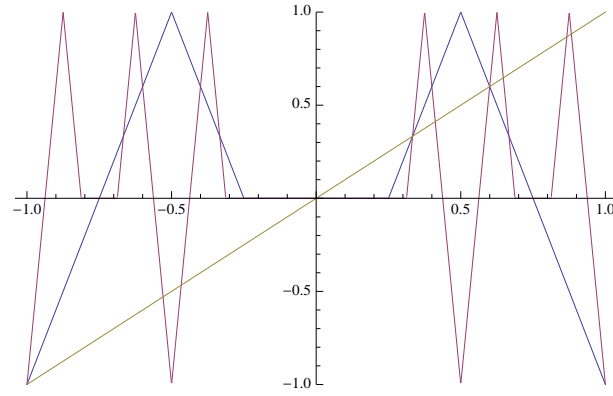
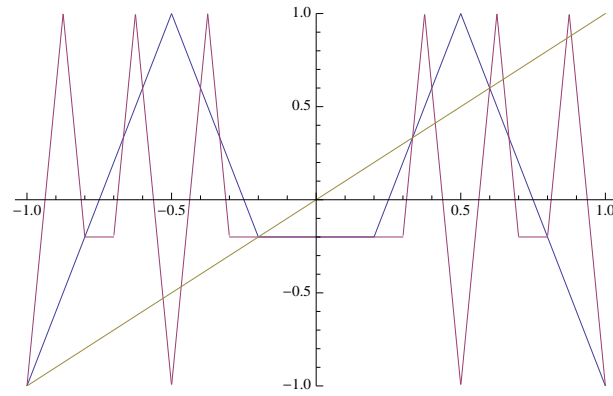
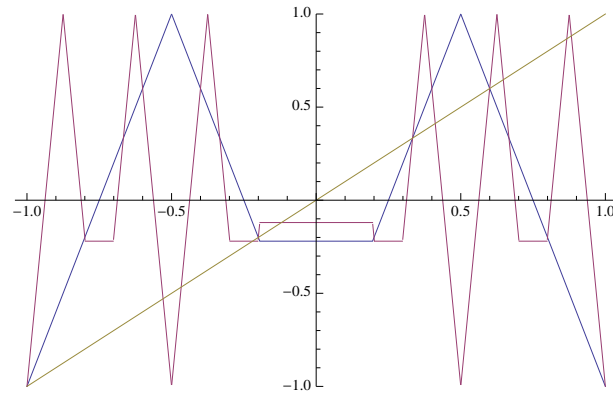
(a) $v = 0.5$ (b) $v = 0.6$ (c) $v = 0.61$

Figura 4.15: Duplicação de período vista através dos gráficos F_v^2 , representado a azul, e F_v^4 , representado a roxo.

em qualquer dos casos podemos escrever

$$\begin{aligned}
 u_{n+1}^j = & -\frac{1}{2^{|X^j(n+1)|-1} \varepsilon_{|X^j(n+1)|-1}(X^j(n+1))} \left(\sum_{i=2}^{|X^j(n)|-1} 2^{|X^j(n+1)|-i} \bar{\varepsilon}_i(X^j(n+1)) + \right. \\
 & + 2^{|X^j(n+1)|-|X^j(n)|} \bar{\varepsilon}_{|X^j(n)|}(X^j(n+1)) + \\
 & \left. + \sum_{i=2}^{|X^j(n+1)|-|X^j(n)|} 2^{|X^j(n+1)|-|X^j(n)|+1-i} \bar{\varepsilon}_{|X^j(n)|+i-1}(X^j(n+1)) + 1 \right)
 \end{aligned}$$

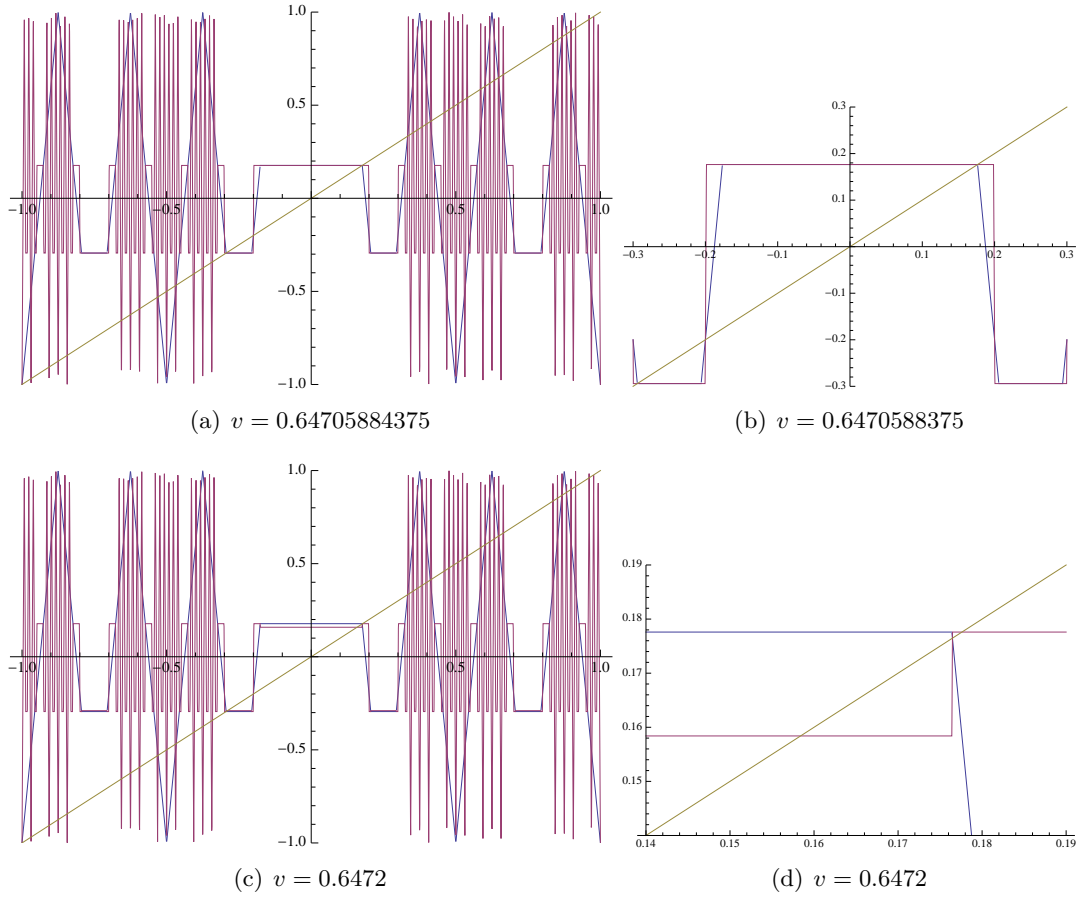


Figura 4.16: Duplicação de período vista através dos gráficos F_v^4 , representado a azul, e F_v^8 , representado a roxo, no intervalo $[-1, 1]$ à esquerda e num intervalo em torno do ponto fixo de F_v^8 à direita.

$$\begin{aligned}
&= - \frac{\sum_{i=2}^{|X^j(n)|-1} 2^{|X^j(n)|-i} \bar{\varepsilon}_i(X^j(n)) \bar{\varepsilon}_{|X^j(n)|}(X^j(n+1)) + \bar{\varepsilon}_{|X^j(n)|}(X^j(n+1))}{2^{|X^j(n)|-1} \varepsilon_{|X^j(n)|-1}(X^j(n+1)) \bar{\varepsilon}_{|X^j(n)|}(X^j(n+1))} - \\
&\quad \frac{\sum_{i=2}^{|X^j(n+1)|-|X^j(n)|} 2^{|X^j(n+1)|-|X^j(n)|+1-i} \bar{\varepsilon}_{|X^j(n)|-1+i}(|X^j(n+1)|) + 1}{2^{|X^j(n+1)|} \varepsilon_{|X^j(n+1)|-1}(X^j(n+1))} \\
&= - \frac{\sum_{i=2}^{|X^j(n)|-1} 2^{|X^j(n)|-i} \bar{\varepsilon}_i(X^j(n)) + 1}{2^{|X^j(n)|-1} \varepsilon_{|X^j(n)|-1}(X^j(n))} - \\
&\quad \frac{\sum_{i=2}^{|X^j(n+1)|-|X^j(n)|} 2^{|X^j(n+1)|-|X^j(n)|+1-i} \bar{\varepsilon}_{|X^j(n)|-1+i}(X^j(n+1)) + 1}{2^{|X^j(n+1)|-1} \varepsilon_{|X^j(n+1)|-1}(X^j(n+1))},
\end{aligned}$$

temos então

$$u_{n+1}^j = u_n^j - \frac{\sum_{i=2}^{|X^j(n+1)|-|X^j(n)|} 2^{|X^j(n+1)|-|X^j(n)|+1-i} \bar{\varepsilon}_{|X^j(n)|-1+i}(X^j(n+1)) + 1}{2^{|X^j(n+1)|-1} \varepsilon_{|X^j(n+1)|-1}(X^j(n+1))}. \quad (4.14)$$

Consideremos

$$X^j(n+1) = X^j(n)_{[1,|X^j(n)|-1]} Y_1 X^j(n)_{[1,|X^j(n)|-1]} Y_2 \dots,$$

(o caso $X^j(n+1) = X^j(n) * (-Y) = X^j(n)_{[1,|X^j(n)|-1]} (-Y_1) X^j(n)_{[1,|X^j(n)|-1]} (-Y_2) \dots$ é totalmente análogo).

Definimos a sequência $Z^j(n) = (X^j(n) * \sigma(Y) = X^j(n) Y_2 X^j(n) \dots$, e a sequência $W^j(n) = Y_1 Z^j(n)$. Temos $|W^j(n)| = |X^j(n+1)| - |X^j(n)| + 1$. Seja z_n tal que $T^{W^j(n)}(z_n) = 0$, segundo a igualdade (4.11), z_n é dado por

$$z_n = - \frac{\sum_{i=2}^{|X^j(n+1)|-|X^j(n)|} 2^{|X^j(n+1)|-|X^j(n)|+1-i} \bar{\varepsilon}_i(W^j(n)) + 1}{2^{|X^j(n+1)|-|X^j(n)|} \varepsilon_{|X^j(n+1)|-|X^j(n)|}(W^j(n))} \quad (4.15)$$

$$= - \frac{\sum_{i=2}^{|X^j(n+1)|-|X^j(n)|} 2^{|X^j(n+1)|-|X^j(n)|+1-i} \bar{\varepsilon}_{|X^j(n)|-1+i}(X^j(n+1)) + 1}{2^{|X^j(n+1)|-|X^j(n)|} \varepsilon_{|X^j(n)|-1}(X^j(n)) \varepsilon_{|X^j(n+1)|-1}(X^j(n+1))} \quad (4.16)$$

onde usamos a relação entre os sinais de $W^j(n)$ e $X^j(n+1)$, isto é, tendo $W^j(n)_i = X^j(n+1)_{|X(n)|-1+i}$, para todo $1 \leq i \leq |X^j(n+1)| - |X^j(n)| + 1$, vem

$$\varepsilon_i(W^j(n)) = \varepsilon_{|X^j(n)|-1}(X^j(n)) \varepsilon_{|X^j(n)|-1+i}(X^j(n+1))$$

para $1 \leq i \leq |X^j(n+1)| - |X^j(n)| + 1$.

Finalmente,

$$u_{n+1}^j = u_n^j + \frac{z_n^j}{2^{|X^j(n)|} \varepsilon_{|X^j(n)|-1}(X^j(n))}. \quad (4.17)$$

Pelo Lema 4.6.2 e pela Proposição 4.3.6, $(u_n^j)_{n \in \mathbb{N}}$ é crescente, donde $\frac{z_n}{\varepsilon_{|X^j(n)|-1}(X^j(n))} \geq 0, \forall n$. Temos

$$\lim_{n \rightarrow \infty} \frac{\log \left(\frac{u_{n+1}^j - u_n^j}{u_{n+2}^j - u_{n+1}^j} \right)}{\log \left(\frac{u_n^j - u_{n-1}^j}{u_{n+1}^j - u_n^j} \right)} = \lim_{n \rightarrow \infty} \frac{\log \frac{z_n^j 2^{|X^j(n+1)|} \varepsilon_{|X^j(n+1)|-1}(X^j(n+1))}{z_{n+1}^j 2^{|X^j(n)|} \varepsilon_{|X^j(n)|-1}(X^j(n))}}{\log \frac{z_{n-1}^j 2^{|X^j(n)|} \varepsilon_{|X^j(n)|-1}(X^j(n))}{z_n^j 2^{|X^j(n-1)|-1} \varepsilon_{|X^j(n-1)|-1}(X^j(n-1))}}$$

$$\begin{aligned}
&= \lim_{n \rightarrow \infty} \frac{(|X^j(n+1)| - |X^j(n)|) \log 2 + \log \left(\frac{z_n^j}{\varepsilon_{|X^j(n)|-1}(X^j(n))} \right) - \log \left(\frac{z_{n+1}^j}{\varepsilon_{|X^j(n+1)|-1}(X^j(n+1))} \right)}{(|X^j(n)| - |X^j(n-1)|) \log 2 + \log \left(\frac{z_{n-1}^j}{\varepsilon_{|X^j(n-1)|-1}(X^j(n-1))} \right) - \log \left(\frac{z_n^j}{\varepsilon_{|X^j(n)|-1}(X^j(n))} \right)} \\
&= \lim_{n \rightarrow \infty} \frac{(|X^j(n+1)| - |X^j(n)|) \log 2 + \log |z_n^j| - \log |z_{n+1}^j|}{(|X^j(n)| - |X^j(n-1)|) \log 2 + \log |z_{n-1}^j| - \log |z_n^j|}
\end{aligned}$$

Temos $T(z_n^j) = \Phi(\sigma(Y_1 Z^j(n))) = \Phi(Z^j(n))$. Assim, uma vez que $X^j(n)$ é finito, existe k tal que $X^j(n) < RL^k$, logo $X^j(n) * K < RL^k$. Donde,

$$T(z_n^j) < \Phi(RL^k) < \Phi(Rl^\infty) = 1, \forall n, \quad (4.18)$$

e assim $|z_n^j| > (T^+)^{-1}(\Phi(RL^k)) > 0$, logo $|z_n^j| \rightarrow 0$. Então,

$$\lim_{n \rightarrow \infty} \frac{\log \left(\frac{u_{n+1}^j - u_n^j}{u_{n+2}^j - u_{n+1}^j} \right)}{\log \left(\frac{u_n^j - u_{n-1}^j}{u_{n+1}^j - u_n^j} \right)} = \lim_{n \rightarrow \infty} \frac{|X^j(n+1)| - |X^j(n)|}{|X^j(n)| - |X^j(n-1)|} = \lim_{n \rightarrow \infty} \frac{\frac{|X^j(n+1)|}{|X^j(n)|} - 1}{1 - \frac{|X^j(n-1)|}{|X^j(n)|}} = |Y|.$$

□

Consideremos agora a sucessão de pares de seqüências definida por

$$\begin{aligned}
(X^0(0), X^1(0)) &= (X^0, X^1), \\
X^j(n+1) &= \begin{cases} X^j(n) * Y^j & \text{se } Y^j \in \Sigma_{(\alpha(j)_n, -\alpha(j)_n)} \\ X^j(n) * (-Y^j) & \text{se } Y^j \in \Sigma_{(-\alpha(j)_n, \alpha(j)_n)} \end{cases}
\end{aligned}$$

onde $\alpha(j)_n = \varepsilon_{|X^j(n)|-1}(X^j(n))$. Pelo Lema 4.6.2, a sucessão $X^j(n)$ é crescente para $i \in \{0, 1\}$. Consideramos a sucessão dos pontos do plano $[-1, 1]^2$, $(u_n, v_n)_{n \in \mathbb{N}}$, tais que $T^{X^0(n)}(u_n) = 0$ e $T^{X^1(n)}(v_n) = 0$. Usando a distância usual em \mathbb{R}^2 ,

$$d((u, v), (w, z)) = \sqrt{(u - w)^2 + (v - z)^2},$$

definimos a sucessão

$$d_n = d((u_{n+1}, v_{n+1}), (u_n, v_n)) = \sqrt{(u_{n+1} - u_n)^2 + (v_{n+1} - v_n)^2},$$

com (u_n, v_n) , onde $T^{X^0(n)}(u_n) = 0$ e $T^{X^1(n)}(v_n) = 0$.

Temos uma generalização do teorema anterior:

Teorema 4.6.5. *Dados $(X^0, X^1) \in \Sigma_2^+$, $(Y^0, Y^1) \in \Sigma_{(\alpha(0), -\alpha(0))} \times \Sigma_{(\alpha(1), -\alpha(1))}$ e $j \in \{0, 1\}$, consideremos a sucessão de pares $(X^0(n), X^1(n))$. A sucessão d_n definida anteri-*

ormente é convergente e tem-se

$$\lim \frac{\log \left(\frac{d_n}{d_{n+1}} \right)}{\log \left(\frac{d_{n-1}}{d_n} \right)} = \min_{i \in \{0,1\}} |Y^i|. \quad (4.19)$$

Demonstração. Usando a igualdade (4.17) temos

$$\begin{aligned} \lim \frac{\log \left(\frac{d_n}{d_{n+1}} \right)}{\log \left(\frac{d_{n-1}}{d_n} \right)} &= \lim \frac{\log \left(\frac{\left(\frac{z_n^0}{2^{|X^0(n)|}} \right)^2 + \left(\frac{z_n^1}{2^{|X^1(n)|}} \right)^2}{\left(\frac{z_{n+1}^0}{2^{|X^0(n+1)|}} \right)^2 + \left(\frac{z_{n+1}^1}{2^{|X^1(n+1)|}} \right)^2} \right)}{\log \left(\frac{\left(\frac{z_{n-1}^0}{2^{|X^0(n-1)|}} \right)^2 + \left(\frac{z_{n-1}^1}{2^{|X^1(n-1)|}} \right)^2}{\left(\frac{z_n^0}{2^{|X^0(n)|}} \right)^2 + \left(\frac{z_n^1}{2^{|X^1(n)|}} \right)^2} \right)} \\ &= \lim \frac{\log \left(\left(\frac{z_n^0}{2^{|X^0(n)|}} \right)^2 + \left(\frac{z_n^1}{2^{|X^1(n)|}} \right)^2 \right) - \log \left(\left(\frac{z_{n+1}^0}{2^{|X^0(n+1)|}} \right)^2 + \left(\frac{z_{n+1}^1}{2^{|X^1(n+1)|}} \right)^2 \right)}{\log \left(\left(\frac{z_{n-1}^0}{2^{|X^0(n-1)|}} \right)^2 + \left(\frac{z_{n-1}^1}{2^{|X^1(n-1)|}} \right)^2 \right) - \log \left(\left(\frac{z_n^0}{2^{|X^0(n)|}} \right)^2 + \left(\frac{z_n^1}{2^{|X^1(n)|}} \right)^2 \right)} = \\ \lim \frac{(|X^0(n+1)| - |X^0(n)|) \log 4 + \log \left((z_n^0)^2 + \frac{(z_n^1)^2}{2^{2(|X^1(n)| - |X^0(n)|)}} \right) - \log \left((z_{n+1}^0)^2 + \frac{(z_{n+1}^1)^2}{2^{2(|X^1(n+1)| - |X^0(n+1)|)}} \right)}{(|X^0(n)| - |X^0(n-1)|) \log 4 + \log \left((z_{n-1}^0)^2 + \frac{(z_{n-1}^1)^2}{2^{2(|X^1(n-1)| - |X^0(n-1)|)}} \right) - \log \left((z_n^0)^2 + \frac{(z_n^1)^2}{2^{2(|X^1(n)| - |X^0(n)|)}} \right)}. \end{aligned}$$

Suponhamos que $|Y^0| < |Y^1|$, então, a partir de certa ordem, temos $|X^0(n)| < |X^1(n)|$, uma vez que

$$|X^i(n)| = |X^i(n)| |Y^i|^n, i \in \{0, 1\},$$

logo, $\lim 2^{|X^1(n)| - |X^0(n)|} = +\infty$. Além disso, com os cálculos da demonstração do resultado anterior, temos

$$\lim \frac{\log \left(\frac{d_n}{d_{n+1}} \right)}{\log \left(\frac{d_{n-1}}{d_n} \right)} = \lim \frac{(|X^0(n+1)| - |X^0(n)|)}{(|X^0(n)| - |X^0(n-1)|)} = |Y^0|$$

□

O resultado anterior permite concluir que a taxa de convergência é independente do ponto inicial, depende apenas do menor comprimento entre $|Y^0|$ e $|Y^1|$. Assim, a ordem de convergência é $l = \min\{|Y^0|, |Y^1|\}$.

Índice Remissivo

- admissibilidade, 60, 71
- cadeia de Markov, 3
 - não homogénea, 3
 - fortemente ergódica, 23
 - fracamente ergódica, 23
 - homogénea, 3
- coeficiente ergódico, 24
- conjugação, 58
- convergência
 - de ordem l , 75
 - geométrica, 75
- diagrama de bifurcação, 67
- duplicação de período, 75
- equilíbrio, 13, 18
- esqueleto, 64
- função
 - deslocamento, 56
 - logística, 75
 - tenda, 53
 - tenda cortada, 53
 - unimodal, 53
- grafo, 11
 - dinâmico, 11
 - não autónomo, 11
- invariante de amassamento, 6, 60, 63
- itinerário, 55, 59
- Jordan
 - bloco de, 41
 - forma normal de, 41
- matriz
 - de adjacência, 15
 - aperiódica, 12
 - ergódica, 12
 - irreduzível, 12
 - reversível, 13
- medida
 - de máxima entropia, 34
 - invariante, 4
- órbita, 55
 - bicrítica, 65
 - crítica, 65
- osso, 64
- Perron-Frobenius
 - operador de, 29
 - teorema de, 13
- produto cruzado, 54
- produto-*, 72
- projecção, 59
- renormalização, 70
- sequência
 - comprimento da, 56
 - maximal, 61
- sinal associado a um símbolo, 56
- tempo de mixing, 14, 28

valor

 próprio generalizado, 41

 singular, 37

Bibliografia

- [AF14] D. Aldous and J. Fill. *Reversible Markov chains and random walks on graphs*. unfinished monograph, 2014.
- [AS12] J. Alves and L. Silva. Periodic paths on nonautonomous graphs. *Linear Algebra and Its Applications*, 437(3):1003–1015, 2012.
- [AS15] J. Alves and L. Silva. Nonautonomous graphs and topological entropy of nonautonomous Lorenz systems. *a aparecer em International Journal of Bifurcation and Chaos*, 2015.
- [Avi10] A. Avila. Dynamics of renormalization operator. In *Proceedings of the International Congress of Mathematicians*, 2010.
- [Beh00] E. Behrends. *Introduction to Markov chains with special emphasis on rapid mixing*. Vieweg, 2000.
- [BEQ90] W. A. Beyer, B. R. Ebanks, and C. R. Qualls. Convergence rates and convergence-order profiles for sequences. *Acta Applicae Mathematica*, 30(3):267–284, 1990.
- [BG97] A. Boyarsky and P. Góra. *Laws of chaos*. Birkhauser, 1997.
- [BM14] M. Bujoriano and R. S. MacKay. Perturbation and sensitivity of inhomogeneous Markov chains in dynamic environments. In *21st International Symposium on Mathematical Theory of Networks and Systems*, pages 184–191, July 7–11, 2014.
- [Bor72] E. Borel. Sur le battage des cartes. *Oeuvres de Émile Borel, Tome II, Centre National de la Recherche Scientifique*, 1972.
- [Bré99] P. Brémaud. *Markov Chains, Gibbs fields, Monte Carlo simulation and queues*. Springer-Verlag, 1999.
- [BS82a] W. A. Beyer and P. R. Stein. Further results on periods and period doubling for iterates of the trapezoid function. *Advances in Applied Mathematics*, 3(3):265–287, 1982.
- [BS82b] W. A. Beyer and P. R. Stein. Period doubling for trapezoid function iteration: metric theory. *Advances in Applied Mathematics*, 3(1):1–17, 1982.
- [CH03] A. Carvalho and T. Hall. Symbolic dynamics and topological models in dimensions 1 and 2. *Topics in Dynamics and Ergodic Theory*, 2003.

- [CT78] P. Couillet and C. Tresser. Itérations d'endomorphismes et groupe de renormalization. *Journal de Physique Colloques*, 39(5):C5–25 – C5–28, 1978.
- [Dev89] R. Devaney. *An introduction to Chaotic dynamical systems*. Addison–Wesley Publishing Company, 1989.
- [DFS00] M. Dellnitz, G. Froyland, and S. Sertl. On the isolated spectrum of the perron-frobenius operator. *Nonlinearity*, 13(4):1171–1188, 2000.
- [DGP78] B. Derrida, A. Gervois, and Y. Pomeau. Iterations of endomorphisms on the real axis and representation of numbers. *Annales de l'I.H.P., section A*, 29(3):305–356, 1978.
- [Dob56] R. Dobrushin. Central limit theorems for non-stationary Markov chains ii. *Theory of Probability and its Applications (english translation)*, 1(4):329–383, 1956.
- [ES05a] S. Elaydi and R. Sacker. Global stability of periodic orbits of non autonomous difference equations and population biology. *Journal of Differential Equations*, 208:258–273, 2005.
- [ES05b] S. Elaydi and R. Sacker. Skew-product dynamical systems: Applications to difference equations. In *Proceedings of the Second Annual Celebration of Mathematics, United emirates*, 2005.
- [Eva] L. C. Evans. *Invariant measures for finite-state, discrete-time Markov chains*. University of Missouri.
- [Fei78] M. J. Feigenbaum. Quantitative universality for a class of nonlinear transformations. *Journal of Statistical Physics*, 19(1), 1978.
- [Fil91] J. Fill. Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains with an application to the exclusion process. *The Annals of Applied Probability*, 1(1):62–87, 1991.
- [FR07a] S. Fernandes and J. Sousa Ramos. Second eigenvalue of transition matrix associated to iterated maps. *Chaos, Solitons and Fractals*, 31:316–326, 2007.
- [FR07b] N. Fountoulakis and B.A. Reed. Faster mixing and small bottlenecks. *Probability Theory and Related Fields*, 137:475–486, 2007.
- [FSS13] N. Franco, L. Silva, and P. Simões. Symbolic dynamics and renormalization of non autonomous k periodic dynamical systems. *Journal of difference equations and Applications*, 19(1):27–38, 2013.
- [GH96] P. Glendinning and T. Hall. Zeros of the kneading invariant and topological entropy for Lorenz maps. *Nonlinearity* 9, 999–1014, 1996.
- [Had98] M. Hadamard. Les surfaces à courbures opposées et leurs lignes géodésiques. *Journal des mathématiques pures e appliquées*, 5serie(tome 4):27–74, 1898.
- [HJ91] R. Horn and C. Jonhson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

- [IM76] D. Isaacson and R. Madsen. *Markov Chains Theory and Applications*. John Wiley and Sons, 1976.
- [KMS99] S. Kolyada, M. Misiurewicz, and L. Snoha. Topological entropy of nonautonomous piecewise monotone dynamical systems of the interval. *Fundamenta mathematicae*, 160, 1999.
- [KR11] P. Kloeden and M. Rasmussen. *Nonautonomous dynamical systems*. American Mathematical Society, 2011.
- [KS96] S. Kolyada and L. Snoha. Topological entropy of non autonomous dynamical systems. *Random & Computational Dynamics*, 4:205–233, 1996.
- [Lan82] O. Lanford. A computer-assisted proof of the feigenbaum conjectures. *Bulletin (New Series) of the American mathematical society*, 6(3):427–434, 1982.
- [LM94] A. Lasota and M. Mackey. *Chaos, Fractals, and Noise*. Springer-Verlag, 1994.
- [LM95] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, 1995.
- [LPW09] D. Levin, Y. Peres, and E. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2009.
- [Lyu99] M. Lyubich. Feigenbaum-Couillet-Tresser universality and Milnor’s hairiness conjecture. *Annals of Mathematics*, 149:319–420, 1999.
- [May76] R. B. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459–467, 1976.
- [MH21] M. Morse and G. Hedlund. Symbolic dynamics. *American Journal of Mathematics*, 60(4):815–866, 1921.
- [MPS08] E. Mossel, Y. Peres, and A. Sinclair. Shuffling by semi-random transpositions. [arXiv.org/pdf/math/0404438v1](https://arxiv.org/pdf/math/0404438v1), 2008.
- [MSS73] N. Metropolis, M. Stein, and P. Stein. On finite limit sets for transformations on the unit interval. *Journal of combinatorial Theory (A)*, 15(25–44), 1973.
- [MT88a] R. S. MacKay and C. Tresser. Some flesh on the skeleton: the bifurcation structure of bimodal maps. *Physica D*, 27(3):412–422, 1988.
- [MT88b] J. Milnor and W. Thurston. On iterated maps of the interval. *Lecture notes in mathematics 1342*, 1988.
- [NL08] G. Nannicini and L. Liberi. Shortest paths on dynamic graphs. *Transactions in Operational research*, 15:1–13, 2008.
- [NS98] M. Neumann and H. Schneider. The convergence of general products of matrices and the weak ergodicity of Markov chains. *Linear Algebra and Its Applications*, 1998.

- [Rad04] A. Radulescu. *The Connected Isentropes Conjecture in a Space of Quartic*. PhD thesis, Stony Brook, 2004.
- [Rin71] J. Ringrose. *Compact non-self-adjoint operators*. Van Nostrand Reinhold Company, 1971.
- [SCZ07] L. Saloff-Coste and J. Zúñiga. Convergence of some time inhomogeneous Markov chains via spectral techniques. *Stochastic Processes and their Applications*, 117(961–979), 2007.
- [SCZ09] L. Saloff-Coste and J. Zúñiga. Merging for time inhomogeneous finite Markov chains, part i: singular values and stability. *Electronic Journal of Probabilities*, 14(1456–1494), 2009.
- [SCZ11] L. Saloff-Coste and J. Zúñiga. Merging and stability for time inhomogeneous finite Markov chains. In Zurich EMS Series of Congress Reports, editor, *Surveys in stochastic processes*, pages 127–151, 2011.
- [Sen06] E. Seneta. *Non negative matrices and Markov chains*. Springer–Verlag, 2006.
- [Sen14] E. Seneta. Inhomogeneous Markov chains and ergodicity coefficients: John Hajnal (1924–2008). In Taylor and Francis, editors, *Communications in Statistics - Theory and Methods 43:7*, pages 1296–1308, 2014.
- [Sil07] L. Silva. Universal convergence rates for stunted sawtooth Lorenz maps. In *Communications of the Laufen Colloquium on Science*. Shaker Publishing, 2007.
- [Sil08] D. D. Siljak. Dynamic graphs. *Nonlinear Analysis: Hybrid Systems*, 2(2):544–567, 2008.
- [SR01] L. Silva and J. Sousa Ramos. A unified renormalization scheme for two piecewise monotonous maps of the interval. *International Journal of Bifurcation and Chaos*, 13(7):1711–1719, 2001.
- [SR02] L. Silva and J. Sousa Ramos. Topological invariants and renormalization of Lorenz maps. *Physica D*, 162(3–4):233–243, 2002.
- [SSF14a] T. Silva, L. Silva, and S. Fernandes. Equilibrium distributions of discrete nonautonomous graphs. *Journal of difference equations and applications*, 20(8):1190–1200, 2014.
- [SSF14b] T. Silva, L. Silva, and S. Fernandes. Convergence time to equilibrium distributions of autonomous and periodic non-autonomous graphs. *Submetido*, Oct 2014.
- [SSF15] L. Silva, T. Silva, and S. Fernandes. Universal convergence rates for a non autonomous family of dynamical systems generated by stunted sawtooth maps. *em preparação*, 2015.
- [SSR15] L. Silva, T. Silva, and J. L. Rocha. Bifurcation skeleton of a non autonomous family of dynamical systems generated by stunted sawtooth maps. *em preparação*, 2015.

- [Str88] G. Strang. *Linear Algebra and its Applications*. Brooks/Cole, 1988.
- [Wal82] Peter Walters. *Introduction to ergodic theory*. Springer-Verlag, 1982.



UNIVERSIDADE DE ÉVORA
INSTITUTO DE INVESTIGAÇÃO
E FORMAÇÃO AVANÇADA

Contactos:

Universidade de Évora
Instituto de Investigação e Formação Avançada - IIFA
Palácio do Vimioso — Largo Marquês de Marialva, Apart. 94
7002-554 Évora — Portugal
Tel: (+351) 266 706 581
Fax: (+351) 266 744 677
email: iifa@uevora.pt