

Universidade Federal do Rio Grande do Sul
Instituto de Matemática
Programa de Pós Graduação em Matemática

**Operador de Ruelle para Cadeias de Markov a Tempo
Contínuo**

Dissertação de mestrado

Luísa Bürgel Borsato

Porto Alegre, 5 de março de 2018

Dissertação submetida por Luísa Bürgel Borsato¹, como requisito parcial para a obtenção do grau de Mestre em Matemática pelo Programa de Pós-Graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professor Orientador:

Dr. Rafael Rigão Souza

Banca examinadora:

Dr. Artur Oscar Lopes

Dr. Elismar da Rosa Oliveira

Dra. Joana Mohr

¹ Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

Agradecimentos

Agradeço aos meus pais, Maristela e Vanderlei, e aos meus irmãos, Gabriel e Júlia, por todo o suporte dado ao longo desta jornada. Sem vocês, ela não teria sido possível, tão menos prazerosa. Agradeço também pela constante torcida e pelos momentos de descontração.

Agradeço ao Thiago, meu parceiro de vida pelos maravilhosos momentos que dividimos e pelo amor que multiplicamos. Agradeço também aos seus pais, que já fazem parte da minha família.

Agradeço ao meu orientador, Rafael Rigão Souza, por toda a disponibilidade, pelo comprometimento com a minha formação e por todos os conselhos e ensinamentos.

Agradeço aos meus professores, que foram essenciais para uma formação completa dentro e fora de sala de aula. Gostaria de agradecer, especialmente, aos professores Adriana Neumann de Oliveira, Artur Oscar Lopes, Eduardo Henrique de Mattos Brietzke e Virgínia Maria Rodrigues.

Agradeço aos meus amigos, que estiveram ao meu lado e proporcionaram bons momentos e muitas risadas. Agradeço também aos meus colegas, e, acima disto, amigos, Eduardo Horta, Eduardo Longa, Jader Brasil, Josué Knosrt, Marcus Vinícius da Silva e Matheus Bohrer.

Resumo

Este trabalho divide-se em três partes. Na primeira parte fazemos uma breve descrição de cadeias de Markov a tempo discreto e tempo contínuo. Na segunda parte, seguindo o artigo [5], introduzimos o formalismo termodinâmico no espaço de Bernoulli com símbolos dados em um espaço métrico compacto, generalizando a teoria usual onde o espaço de estados é finito. Após, seguindo o artigo [1], introduziremos uma versão do *Operador de Ruelle* para cadeias de Markov a tempo contínuo. Ainda, a partir de uma função V que funcionará como uma perturbação, definiremos um operador de Ruelle modificado e, para este operador, mostraremos a existência de uma auto-função e uma auto-medida.

Palavras-chave: Cadeias de Markov, Formalismo Termodinâmico, Operador de Ruelle.

Abstract

This work is divided in three parts. In the first one, we give a brief description of Markov chains in both discrete time and continuous time. In the second one, following the article [5], we introduce the thermodynamic formalism in the Bernoulli space with symbols in a compact metric space, generalizing the usual theory, where the space of states is finite. Then, following the article [1], we will introduce a version of *Ruelle Operator* for Markov chains in continuous time. Also, using a V function, which will be seen as a perturbation, we will define a modified Ruelle operator and, for this operator, we will show the existence of a eigenfunction and a eigenmeasure.

Key-words: Markov Chains, Thermodynamic Formalism, Ruelle Operator.

Introdução

Desde a década de 60, com trabalhos pioneiros de Bowen, Ruelle e Sinai, o formalismo termodinâmico vem desempenhando um papel importante na teoria de sistemas dinâmicos e na mecânica estatística. O *Operador de Ruelle* é uma ferramenta de extrema utilidade em formalismo termodinâmico. Um dos principais objetivos em formalismo termodinâmico é garantir a existência de uma autofunção para este operador e também a existência de uma auto-medida para seu operador dual, facilitando assim o estudo dos estados de equilíbrio do sistema, e também, via um processo limite adequado, o estudo de medidas maximizadoras (dentro da teoria usualmente chamada de otimização ergódica). Atualmente, tenta-se usar este operador em contextos mais gerais, e dessa forma, estudar estados de equilíbrio para outras classes. Neste presente trabalho, exporemos uma versão do operador de Ruelle para cadeias de Markov a tempo contínuo e, a partir disso, demonstraremos diversas propriedades importantes.

Esta dissertação está estruturada em três capítulos. Os dois primeiros visam apresentar ao leitor os principais conceitos de *Cadeias de Markov* e de *Formalismo Termodinâmico*. Mais especificamente, no capítulo 2 trataremos sobre *Cadeias de Markov* a tempo discreto e também a tempo contínuo, trazendo conceitos básicos desta teoria, juntamente com propriedades relevantes e exemplos. No capítulo 3, desenvolveremos a Teoria de Formalismo Termodinâmico para espaços métricos compactos com medida a priori qualquer, definindo o *Operador de Ruelle* neste caso. Além disso demonstraremos uma versão do *Teorema de Ruelle-Perron-Fröbenius* para funções α -Hölder, que é primordial para que encontrarmos uma autofunção para o operador de

Ruelle e uma automedida para o seu dual.

Finalmente, no capítulo 4, seguindo o artigo *A Ruelle Operator for continuous time Markov Chains*, dos autores *A. Baraviera, A. O. Lopes e R. Exel*, aliaremos os conceitos introduzidos nos dois capítulos anteriores, definindo uma versão do Operador de Ruelle para Cadeias de Markov a tempo contínuo. Para tal, consideraremos um *espaço de estados* S finito, uma cadeia de Markov tomando valores em S e uma medida *a priori*. Mostraremos que esta medida é um ponto fixo para o dual do operador de Ruelle e diversas propriedades relacionadas a este operador. Depois, a partir de uma perturbação, definiremos um Operador de Ruelle modificado e, com o auxílio de uma versão do *Teorema de Ruelle-Perron-Fröbenius*, construiremos uma autofunção para este novo operador e uma automedida para o seu dual.

Cadeias de Markov

Nosso objetivo, neste capítulo, é introduzir o conceito de Cadeias de Markov a tempo contínuo, que serão essenciais para o bom entendimento do Capítulo 4, onde apresentaremos uma versão do Operador de Ruelle para Cadeias de Markov a tempo contínuo. Para tal, abordaremos inicialmente Cadeias de Markov a tempo discreto e depois estenderemos este conceito para tempo contínuo. Apresentaremos importantes propriedades das Cadeias de Markov, juntamente a exemplos que ajudarão a ilustrar a teoria.

2.1 Preliminares

Como preparação para este capítulo, trataremos brevemente nesta seção de conceitos básicos da teoria de probabilidade, que embasa o estudo de Cadeias de Markov.

Definição 2.1. *Seja Ω um conjunto e \mathcal{A} uma família de subconjuntos de Ω . Dizemos que \mathcal{A} é uma σ -álgebra em Ω quando:*

- (i) $\emptyset \in \mathcal{A}$;
- (ii) se $A \in \mathcal{A}$, então $A^C \in \mathcal{A}$;
- (iii) se $A_1, A_2, \dots \in \mathcal{A}$, então $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Um par (Ω, \mathcal{A}) , constituído de um conjunto e uma σ -álgebra sobre este conjunto, é chamado de *espaço mensurável*. Ainda, sobre uma σ -álgebra, podemos definir uma probabilidade:

Definição 2.2. Uma probabilidade \mathbb{P} sobre uma σ -álgebra \mathcal{A} em Ω é uma função σ -aditiva tomando valores em $[0, 1]$, ou seja, $\mathbb{P}: \mathcal{A} \rightarrow [0, 1]$ é tal que:

- (i) $\mathbb{P}(\emptyset) = 0$;
- (ii) $\mathbb{P}(\Omega) = 1$;
- (iii) Se $\{A_n\}_{n \in \mathbb{N}}$ é uma sequência de elementos de \mathcal{A} , com $A_i \cap A_j = \emptyset$ sempre que $i \neq j$, então

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{+\infty} \mathbb{P}(A_n).$$

Uma tripla $(\Omega, \mathcal{A}, \mathbb{P})$ é chamada de *espaço de probabilidade*. A próxima proposição traz algumas propriedades básicas sobre espaços de probabilidade e não será demonstrada:

Proposição 2.3. Seja $(\Omega, \mathcal{A}, \mathbb{P})$ um espaço de probabilidade. São válidas as seguintes propriedades:

- (i) $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, para todo $A \in \mathcal{A}$;
- (ii) Se $A, B \in \mathcal{A}$ e $A \subset B$, então $\mathbb{P}(A) \leq \mathbb{P}(B)$;
- (iii) Dados $A_1, A_2, \dots \in \mathcal{A}$, tem-se $\mathbb{P}(\cup_{i=1}^{+\infty} A_i) \leq \sum_{i=1}^{+\infty} \mathbb{P}(A_i)$;
- (iv) Se $A_1, A_2, \dots \in \mathcal{A}$ é tal que $A_i \subset A_{i+1}$ e $A = \cup_{i=1}^{+\infty} A_i$, então $\mathbb{P}(A) = \lim_{i \rightarrow +\infty} \mathbb{P}(A_i)$;
- (v) Se $A_1, A_2, \dots \in \mathcal{A}$ é tal que $A_{i+1} \subset A_i$ e $A = \cap_{i=1}^{+\infty} A_i$, então $\mathbb{P}(A) = \lim_{i \rightarrow +\infty} \mathbb{P}(A_i)$

Definição 2.4. Sejam (Ω, \mathcal{A}) , (S, \mathcal{B}) dois espaços mensuráveis. Uma função $\varphi: \Omega \rightarrow S$ é dita função mensurável se para todo $B \in \mathcal{B}$ vale $\varphi^{-1}(B) \in \mathcal{A}$.

S é chamado de *espaço de estados*, ou seja, é o conjunto de possíveis valores que podem ser assumidos pela função mensurável. Quando $S \subset \mathbb{R}$, comumente as funções mensuráveis são chamadas de *variáveis aleatórias*.

Para as próximas seções, um conceito muito importante em teoria da probabilidade é o conceito de *probabilidade condicional*, que mede a probabilidade de que um evento A ocorra dada a ocorrência de um evento B , desde que $\mathbb{P}(B) > 0$.

Definição 2.5. *Sejam A, B eventos, com $\mathbb{P}(B) > 0$. A probabilidade condicional de A dado B é o número $\mathbb{P}(A|B)$ definido por*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Geometricamente, temos a seguinte interpretação para probabilidade condicional: considere a figura a seguir, onde A e B estão desenhados de modo que suas áreas sejam proporcionais às suas probabilidades.

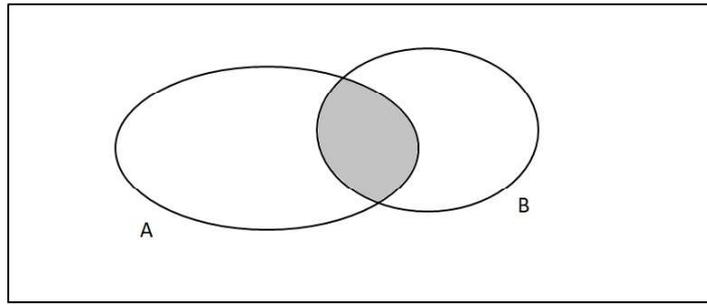


Fig. 2.1

Então, $\mathbb{P}(A|B)$ é a proporção de B ocupada por A .

Proposição 2.6. *Sejam $A_1, A_2, \dots \in \mathcal{A}$ eventos que formam uma partição de Ω , isto é, $A_i \cap A_j = \emptyset$, se $i \neq j$, e $\bigcup_{n=1}^{\infty} A_n = \Omega$. Então, dado um evento $B \in \mathcal{A}$, vale que*

$$\mathbb{P}(B) = \sum_{n=1}^{+\infty} \mathbb{P}(A_n) \mathbb{P}(B|A_n).$$

Definição 2.7. *Dois eventos $A, B \in \mathcal{A}$ são independentes se*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B).$$

Observação 2.8. Se A e B são independentes, então A e B^c também são independentes, assim como A^c e B e também A^c e B^c .

Definição 2.9. *Os eventos $A_1, A_2, \dots, A_n \in \mathcal{A}$ são independentes dois a dois se*

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \mathbb{P}(A_j),$$

para todo $i \neq j$.

Definição 2.10. Os eventos $A_1, A_2, \dots, A_n \in \mathcal{A}$ são independentes se

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_i\right) = \prod_{i=1}^n \mathbb{P}(A_i).$$

Observação 2.11. Independência dois a dois não implica independência conjunta, mas independência conjunta implica independência dois a dois.

Definição 2.12. Uma coleção qualquer de eventos $\{A_i\}_{i \in I}$ é dita independente se A_{i_1}, \dots, A_{i_n} são independentes para qualquer $n \in \mathbb{N}$ e $i_1, \dots, i_n \in I$.

2.2 Cadeias de Markov a Tempo Discreto

As cadeias de Markov recebem esse nome em homenagem ao matemático russo Andrei Markov, que formalizou esta teoria no início do século XX.

Informalmente, dizemos que um processo estocástico tem a propriedade Markoviana se, dado o estado presente, podemos fazer previsões, em termos de distribuição dos estados futuros, de maneira independente do passado. Formalizaremos estes conceitos ao longo desta seção.

A partir desta seção, trataremos apenas o caso em que S é um espaço de estados finito. Sem perda de generalidade, escreveremos $S = \{1, 2, \dots, d\}$.

2.2.1 Definições e Propriedades básicas

Definição 2.13. Dado um espaço de probabilidade $(\Omega, \mathcal{A}, \mathbb{P})$ e (S, \mathcal{G}) um espaço mensurável, um processo estocástico a tempo discreto é uma família de variáveis aleatórias $\{X_m: (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (S, \mathcal{G})\}_{m \in \mathbb{N}}$, onde S é o espaço de estados.

Definição 2.14. Seja $\mathcal{P} = (P_{ij})$ uma matriz $d \times d$. \mathcal{P} é dita **coluna estocástica** se, $P_{ij} \geq 0$ e, fixado $j \in \{1, \dots, d\}$, temos que $\sum_{i=1}^d P_{ij} = 1$.

Podemos definir, analogamente, uma matriz *linha-estocástica*.

Exemplo 2.15. Tomando $S = \{1, 2, 3, 4\}$, considere

$$\mathcal{P} = \begin{pmatrix} 1/4 & 0 & 1 & 1/2 \\ 1/4 & 1/3 & 0 & 0 \\ 1/4 & 2/3 & 0 & 1/2 \\ 1/4 & 0 & 0 & 0 \end{pmatrix}.$$

Então \mathcal{P} é coluna estocástica, mas não linha-estocástica.

Exemplo 2.16. Se $S = \{1, 2\}$, uma matriz coluna estocástica \mathcal{P} é sempre da forma $\mathcal{P} = \begin{pmatrix} \alpha & \beta \\ 1 - \alpha & 1 - \beta \end{pmatrix}$, com $0 \leq \alpha, \beta \leq 1$.

Definição 2.17. Um vetor $\pi = (\pi_i)$, $i \in \{1, \dots, d\}$, é um vetor de probabilidade se $\pi_i \geq 0$ e $\sum_{i=1}^d \pi_i = 1$.

Exemplo 2.18. Se $S = \{1, 2\}$, um vetor de probabilidade é sempre da forma $\pi = (\gamma, 1 - \gamma)$, com $0 \leq \gamma \leq 1$.

Há duas possíveis formas de definirmos cadeias de Markov; uma usando matrizes linha estocástica e outra usando matrizes coluna estocásticas. Apesar da maioria dos textos sobre cadeias de Markov ou processos estocásticos fazer a construção com matrizes linha estocásticas, aqui escolheremos a menos usual destas formas: matrizes coluna estocásticas, visando preparar o leitor para o capítulo 4.

Definição 2.19. Seja $\{X_t : t \in \mathbb{N}\}$ um processo estocástico com espaço de estados $S = \{1, \dots, d\}$. Dizemos que X_n é uma cadeia de Markov a tempo discreto se, para qualquer $n \geq 0$ e qualquer conjunto de estados $i_0, i_1, \dots, i_n, i_{n+1}$ tal que $\mathbb{P}(X_0 = i_0, \dots, X_n = i_n) > 0$, temos

$$\mathbb{P}(X_{n+1} = i_{n+1} | X_0 = i_0, \dots, X_n = i_n) = \mathbb{P}(X_{n+1} = i_{n+1} | X_n = i_n).$$

Definição 2.20. $\mathbb{P}(X_{n+1} = j | X_n = i)$ é dita probabilidade de transição do estado i para o estado j em uma etapa, isto é, é a probabilidade do processo estar no estado j no tempo $n + 1$ dado que estava no estado i no tempo n .

Nesta dissertação, trataremos apenas do caso de Cadeias de Markov com probabilidades de transição estacionárias, ou homogêneas no tempo, isto é:

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \mathbb{P}(X_1 = j | X_0 = i).$$

Como consequência,

$$\mathbb{P}(X_{n+m} = j | X_n = i) = \mathbb{P}(X_m = j | X_0 = i).$$

(homogeneidade no tempo nas probabilidades de transição em m passos).

Em outras palavras, isso significa que as probabilidades de transição não dependem de n . Com isto, construímos a *matriz de transição* \mathcal{P} para *Cadeias de Markov Estacionárias*, que a partir de agora chamaremos sem distinção de *Cadeias de Markov*, da seguinte maneira:

$$\mathcal{P}_{ij} = \mathbb{P}(X_1 = i | X_0 = j).$$

Essa matriz assim construída descreve as probabilidades de transição entre os estados em uma etapa da Cadeia de Markov associada e cumpre a propriedade de ser coluna estocástica.

De fato, fixado $j \in \{1, \dots, d\}$,

$$\begin{aligned}
 \sum_{i=1}^d \mathcal{P}_{ij} &= \sum_{i=1}^d \mathbb{P}(X_1 = i | X_0 = j) \\
 &= \sum_{i=1}^d \frac{\mathbb{P}(X_1 = i, X_0 = j)}{\mathbb{P}(X_0 = j)} \\
 &= \frac{\sum_{i=1}^d \mathbb{P}(X_1 = i, X_0 = j)}{\mathbb{P}(X_0 = j)} \\
 &= \frac{\mathbb{P}(\cup_{i=1}^d \{X_1 = i, X_0 = j\})}{\mathbb{P}(X_0 = j)} \\
 &= \frac{\mathbb{P}(X_0 = j)}{\mathbb{P}(X_0 = j)} \\
 &= 1.
 \end{aligned}$$

Teorema 2.21. *Se $\pi^0 = (\pi_1^0, \dots, \pi_s^0)^T$ é a distribuição inicial, então a distribuição no tempo 1 é dada por $\pi^1 = \mathcal{P}\pi^0$.*

Demonstração. Fixado $k \in S = \{1, \dots, d\}$, temos que

$$\begin{aligned}
 \mathbb{P}[X_1 = k] &= \sum_{j=1}^d \mathbb{P}[X_1 = k, X_0 = j] \\
 &= \sum_{j=1}^d \mathbb{P}[X_1 = k | X_0 = j] \mathbb{P}[X_0 = j] \\
 &= \sum_{j=1}^d \mathcal{P}_{kj} \pi_j^0 \\
 &= (\mathcal{P}\pi^0)_k.
 \end{aligned}$$

□

Como consequência imediata, note que, se $\pi^0 = (\pi_1^0, \dots, \pi_s^0)^T$ é a distribuição inicial, para descobrirmos a distribuição após n passos basta olharmos para $\pi^n = \mathcal{P}^n \pi^0$.

Definição 2.22. *Um vetor de probabilidade $\pi = (\pi_1, \dots, \pi_d)$ é dito estacionário, ou invariante, para \mathcal{P} se $\mathcal{P}\pi = \pi$.*

Neste caso, dizemos que π é estacionário para a cadeia de Markov associada a matriz coluna estocástica \mathcal{P} . Vejamos agora que sempre podemos encontrar um vetor de probabilidade π tal que π é estacionário para \mathcal{P} .

Teorema 2.23. *Para toda matriz estocástica \mathcal{P} , existe π vetor de probabilidade tal que π é invariante para \mathcal{P} .*

Demonstração. Considere $\Sigma = \{p = (p_1, \dots, p_d) : p_i \geq 0, \sum_{i=1}^d p_i = 1\} \subset \mathbb{R}^d$ o conjunto dos vetores de probabilidade. Σ é compacto e convexo. Ainda, considere a aplicação $T: \Sigma \rightarrow \Sigma$, com $T(p) = \mathcal{P}p$. Esta aplicação é contínua, pois está definida a partir de produtos e somas finitas.

Pelo *Teorema do Ponto Fixo de Brouwer*, existe $\pi \in \Sigma$ tal que π é ponto fixo para T , isto é, $\mathcal{P}\pi = \pi$. \square

Uma observação interessante: se todas as entradas da matriz \mathcal{P} forem estritamente positivas, pode-se provar que o vetor estacionário é único e, ainda, é um atrator para a dinâmica definida por \mathcal{P} , ou seja, $\mathcal{P}^n q \rightarrow \pi$ quando $n \rightarrow \infty$, para qualquer estado inicial q . (De fato basta que uma potência de \mathcal{P} tenha entradas todas estritamente positivas.) Para tratarmos das probabilidades de transição em m etapas, devemos estudar o comportamento de \mathcal{P}^m . Primeiramente, vejamos que \mathcal{P}^m também é coluna estocástica.

Proposição 2.24. *Se \mathcal{P} é matriz $d \times d$ coluna estocástica, então \mathcal{P}^m é matriz $d \times d$ coluna estocástica, para todo $m \in \mathbb{N}$.*

Demonstração. Demonstraremos de duas maneiras:

A primeira maneira é baseada no fato de que a k -ésima coluna de \mathcal{P}^m coincide com $\mathcal{P}^m e_k$, onde e_k é o k -ésimo elemento da base canônica do \mathbb{R}^n . Pelo teorema 2.21, sabemos que $\mathcal{P}^m e_k$ é a distribuição da cadeia de Markov no tempo m , e portanto é um vetor de probabilidade.

A segunda maneira é por indução. Para $m = 1$ o resultado segue trivialmente. Suponha que o resultado é válido para m , isto é, \mathcal{P}^m é tal que, fixado $j \in \{1, \dots, d\}$, $\sum_{i=1}^d (\mathcal{P}^m)_{ij} = 1$ e mostremos que o mesmo vale para $m + 1$.

Fixe $j \in \{1, \dots, d\}$. Então

$$\begin{aligned} \sum_{i=1}^d (\mathcal{P}^{m+1})_{ij} &= \sum_{i=1}^d \sum_{k=1}^d \mathcal{P}_{ik}^m \mathcal{P}_{kj} \\ &= \sum_{k=1}^d \sum_{i=1}^d \mathcal{P}_{ik}^m \mathcal{P}_{kj} \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^d \mathcal{P}_{kj} \sum_{i=1}^d \mathcal{P}_{ik}^m \\
&= \sum_{k=1}^d \mathcal{P}_{kj} \\
&= 1,
\end{aligned}$$

onde na quarta igualdade usamos o fato de que \mathcal{P}^m é coluna estocástica e na quinta igualdade usamos o fato de que \mathcal{P} é coluna estocástica. \square

A próxima proposição é conhecida como *Equação de Chapman-Kolmogorov*.

Proposição 2.25. *Fixados $i, j \in S$, tem-se*

$$\mathcal{P}_{ij}^{l+m} = \sum_{k=1}^d \mathcal{P}_{ik}^l \mathcal{P}_{kj}^m$$

para quaisquer $l, m \in \mathbb{N}$.

Demonstração. Temos

$$\begin{aligned}
\mathcal{P}_{ij}^{l+m} &= \mathbb{P}(X_{l+m} = i | X_0 = j) \\
&= \sum_{k=1}^d \mathbb{P}(X_{l+m} = i, X_m = k | X_0 = j) \\
&= \sum_{k=1}^d \frac{\mathbb{P}(X_{l+m} = i, X_m = k, X_0 = j)}{\mathbb{P}(X_0 = j)} \\
&= \sum_{k=1}^d \frac{\mathbb{P}(X_{l+m} = i, X_m = k, X_0 = j)}{\mathbb{P}(X_m = k, X_0 = j)} \frac{\mathbb{P}(X_m = k, X_0 = j)}{\mathbb{P}(X_0 = j)} \\
&= \sum_{k=1}^d \mathbb{P}(X_{l+m} = i | X_m = k, X_0 = j) \mathbb{P}(X_m = k | X_0 = j) \\
&= \sum_{k=1}^d \mathbb{P}(X_{l+m} = i | X_m = k) \mathbb{P}(X_m = k | X_0 = j) \\
&= \sum_{k=1}^d \mathbb{P}(X_l = i | X_0 = k) \mathbb{P}(X_m = k | X_0 = j) \\
&= \sum_{k=1}^d \mathcal{P}_{ik}^l \mathcal{P}_{kj}^m.
\end{aligned}$$

onde na sexta igualdade usamos a propriedade Markoviana e na sétima igualdade usamos a homogeneidade no tempo. \square

Com isto, fica natural a interpretação de \mathcal{P}^m : é a matriz cujas entradas são as probabilidades de transição em m passos, ou seja,

$$\mathcal{P}_{ij}^m = \mathbb{P}(X_m = i | X_0 = j).$$

2.3 Cadeias de Markov a Tempo Contínuo

Agora vamos estender o conceito de Cadeias de Markov para tempo contínuo. A primeira dificuldade em que esbarramos é como definimos probabilidade de transição a tempo 1, já que, agora, as potências da matriz de transição não definem completamente as probabilidades de transição para qualquer tempo $t \in \mathbb{R}$. Por exemplo, se quisermos saber a probabilidade de transição de j para i no tempo 3.45, isto não está definido. Devemos, então, encontrar outra maneira de definirmos probabilidades de transição, mas de modo que quando particularizarmos para tempo discreto, a definição coincida com a definição para tempo discreto, anteriormente tratada na seção 2.2. Para tal, as probabilidades de transição serão definidas a partir de um *semigrupo* \mathcal{P}^t que é construído da seguinte forma:

Definição 2.26. *Uma matriz Q de ordem $d \times d$ é dita coluna soma zero se*

- (i) $Q_{ii} < 0$, para todo $i \in S$;
- (ii) $Q_{ij} \geq 0$, para todo $i \neq j \in S$;
- (iii) $\sum_{i=1}^d Q_{ij} = 0$, para todo $j \in S$ fixado.

Q_{ij} é interpretado como a taxa de migração do estado j para o estado i e Q_{ii} é interpretado como a taxa de saída do estado i .

Exemplo 2.27. $Q = \begin{pmatrix} -3 & \frac{1}{2} & 1 \\ 2 & -1 & 0 \\ 1 & \frac{1}{2} & -1 \end{pmatrix}$ é uma matriz 3×3 coluna soma zero, pois sua diagonal é composta por números negativos, os números fora da diagonal são todos não-negativos e a soma de cada coluna é 0.

Dado $p \in [0, +\infty)$, uma maneira de interpolarmos a sequência discreta $(p^n : n = 0, 1, 2, \dots)$ é definirmos a sequência $(e^{tq} : t \geq 0)$, com $q = \log p$. Faremos uma construção análoga para matrizes.

Para tal, precisamos definir exponencial de matrizes.

Definição 2.28. Dada uma matriz $A = (q_{ij})_{i,j \in I}$ qualquer, com I conjunto finito, definimos a exponencial de A , e^A , como a série $\sum_{k=0}^{+\infty} A^k/k!$.

Observação 2.29. Como $\sum_{k=0}^{+\infty} \frac{A^k}{k!}$ converge componente a componente, a exponencial de matrizes está bem definida.

Uma propriedade importante sobre exponencial de matrizes é que se A e B são duas matrizes que comutam, ou seja, se $AB = BA$, então $e^{A+B} = e^A e^B$.

Para mais detalhes sobre exponencial de matrizes, o leitor pode consultar [14], seção 2.10.

Dito isso, dada uma matriz $\mathcal{P}_{d \times d}$, gostaríamos que existisse uma matriz $Q_{d \times d}$ de modo que $\mathcal{P} = e^Q$ e, portanto, teríamos que $\mathcal{P}^m = (e^Q)^m = e^{mQ}$, pela comutatividade. Dessa forma, podemos repetir o raciocínio feito para números e, ao considerarmos $(\mathcal{P}(t) := e^{tQ} : t \geq 0)$, conseguiremos interpolar a sequência discreta $(\mathcal{P}^m : m = 0, 1, 2, \dots)$. Veremos, mais adiante, que $\mathcal{P}(t) = e^{tQ}$ é coluna estocástica se, e somente se, Q é matriz coluna soma zero. O próximo teorema dará base para provarmos este resultado.

Teorema 2.30. Sejam $Q_{d \times d}$ uma matriz e $\mathcal{P}(t) = e^{tQ}$. Então

(i) $\mathcal{P}(t+s) = \mathcal{P}(t)\mathcal{P}(s)$;

(ii) $(\mathcal{P}(t) : t \geq 0)$ é a única solução de $\begin{cases} \frac{d}{dt}\mathcal{P}(t) = \mathcal{P}(t)Q \\ \mathcal{P}(0) = I \end{cases}$ (Forward Equation).

(iii) $(\mathcal{P}(t) : t \geq 0)$ é a única solução de $\begin{cases} \frac{d}{dt}\mathcal{P}(t) = Q\mathcal{P}(t) \\ \mathcal{P}(0) = I \end{cases}$ (Backward Equation).

(iv) Para $k = 0, 1, 2, \dots$, $\left(\frac{d}{dt}\right)^k \Big|_{t=0} \mathcal{P}(t) = Q^k$.

Demonstração.

(i) Dados quaisquer $s, t \in \mathbb{R}$, temos que $(sQ)(tQ) = (tQ)(sQ)$, de onde segue que

$$\mathcal{P}(t+s) = e^{(t+s)Q} = e^{tQ} e^{sQ} = \mathcal{P}(t)\mathcal{P}(s)$$

(ii) Já que raio de convergência da série $\mathcal{P}(t) = e^{tQ} = \sum_{k=0}^{+\infty} \frac{(tQ)^k}{k!}$ é infinito, podemos derivar cada componente e a derivação é feita termo a termo, de onde obtemos

$$\mathcal{P}'(t) = \sum_{k=1}^{+\infty} \frac{k t^{k-1} Q^k}{k!} = \sum_{k=1}^{+\infty} \frac{t^{k-1} Q^k}{(k-1)!} = \left(\sum_{k=1}^{+\infty} \frac{t^{k-1} Q^{k-1}}{(k-1)!} \right) Q = \mathcal{P}(t)Q.$$

Além disso, $\mathcal{P}(0) = I$, trivialmente.

Com isto, provamos que $\mathcal{P}(t)$ é solução de $\begin{cases} \frac{d}{dt}\mathcal{P}(t) = \mathcal{P}(t)Q \\ \mathcal{P}(0) = I \end{cases}$. Falta provarmos que esta é a única solução.

Seja $M(t)$ tal que $M(t)$ é solução de $\begin{cases} \frac{d}{dt}\mathcal{P}(t) = \mathcal{P}(t)Q \\ \mathcal{P}(0) = I \end{cases}$. Então

$$\begin{aligned} \frac{d}{dt}(M(t)e^{-tQ}) &= \left(\frac{d}{dt}M(t)\right)e^{-tQ} + M(t)\left(\frac{d}{dt}e^{-tQ}\right) \\ &= M(t)Qe^{-tQ} - M(t)Qe^{-tQ} \\ &= 0 \end{aligned}$$

onde na segunda igualdade utilizamos o fato de que $M(t)$ satisfaz a *forward equation*.

Logo, $M(t)e^{-tQ}$ é constante e, como $M(0) = I$, segue que $M(t) = e^{tQ} = \mathcal{P}(t)$, como desejávamos.

(iii) Análogo ao item anterior.

(iv) Basta derivarmos termo a termo k vezes e tomarmos $t = 0$.

□

Corolário 2.31. Para todo $n \in \mathbb{N}$, $\mathcal{P}(t) = \mathcal{P}\left(\frac{t}{n}\right)^n$.

Demonstração. Consequência imediata do item (i). □

Teorema 2.32. Seja $Q_{d \times d}$ uma matriz. Então Q é uma matriz coluna soma zero se, e somente se, $\mathcal{P}(t) = e^{tQ}$ é uma matriz coluna estocástica, para todo $t \geq 0$.

Observação 2.33. O mesmo resultado vale trocando *coluna* por *linha* no teorema acima.

Demonstração. (\implies) Primeiramente, vejamos que se Q é uma matriz cujas colunas somam zero, então Q^k também terá esta propriedade, para todo $k \geq 1$. Para $k = 0$, convencionamos que $Q^0 = I$. Mostraremos este resultado por indução.

Para $k = 1$ o resultado é válido trivialmente. Suponha agora válido para k e mostraremos que vale para $k + 1$.

De fato, fixado $j \in \{1, \dots, d\}$, temos

$$\sum_{i=1}^d (Q^{k+1})_{ij} = \sum_{i=1}^d \sum_{m=1}^d Q_{im} (Q^k)_{mj} = \sum_{m=1}^d (Q^k)_{mj} \sum_{i=1}^d Q_{im} = 0,$$

já que $\sum_{i=1}^d Q_{im} = 0$, pois estamos assumindo que as colunas de Q somam zero.

Agora vejamos que $\sum_{i=1}^d (e^{tQ})_{ij} = 1$, concluindo assim que $\mathcal{P}(t) = e^{tQ}$ é coluna estocástica.

$$\begin{aligned} \sum_{i=1}^d (e^{tQ})_{ij} &= \sum_{i=1}^d \left(\sum_{k=0}^{+\infty} \frac{(tQ)^k}{k!} \right)_{ij} \\ &= \sum_{i=1}^d \sum_{k=0}^{+\infty} \frac{t^k}{k!} (Q^k)_{ij} \\ &= 1 + \sum_{i=1}^d \sum_{k=1}^{+\infty} \frac{t^k}{k!} (Q^k)_{ij} \\ &= 1 + \sum_{k=1}^{+\infty} \frac{t^k}{k!} \sum_{i=1}^d (Q^k)_{ij} \\ &= 1 \end{aligned}$$

pois as colunas de Q^k somam zero, como foi mostrado anteriormente.

(\Leftarrow) Para tratarmos sobre a recíproca, precisamos antes estabelecermos certa notação: diremos que $f(t) = \mathcal{O}(t)$ quando $t \rightarrow 0$ se, e somente se, $\left| \frac{f(t)}{t} \right| \leq C$, para algum $C < +\infty$ e para todo t suficientemente pequeno. Assim, se t é suficientemente pequeno, podemos escrever

$$\mathcal{P}(t) = I + tQ + \mathcal{O}(t^2).$$

Dessa forma, para $i \neq j$ e t suficientemente pequeno, $Q_{ij} \geq 0$ se, e somente se, $(\mathcal{P}(t))_{ij} \geq 0$. Gostaríamos que esta propriedade valesse para todo $t \geq 0$ e não apenas para t suficientemente pequeno. Para estendermos, basta aplicarmos o corolário 2.31. De fato, dado $t \geq 0$, escolhamos $m \in \mathbb{N}$ suficientemente grande de modo que $\frac{t}{m}$ cumpra a propriedade referida e aplicamos o corolário. Como $\mathcal{P}(t) = e^{tQ}$ é coluna estocástica e, pelo teorema 2.30,

$$\sum_{i=1}^d Q_{ij} = \left. \frac{d}{dt} \right|_{t=0} \sum_{i=1}^d (\mathcal{P}(t))_{ij} = 0.$$

□

Juntando os últimos resultados, temos que $\mathcal{P}(t)$ é coluna estocástica para todo $t \geq 0$ e cumpre as propriedades de semigrupo, ou seja, $\mathcal{P}(t+s) = \mathcal{P}(t)\mathcal{P}(s)$ e $\mathcal{P}(0) = I$. Diremos que Q é seu gerador infinitesimal.

Exemplo 2.34. No caso geral, em $S = \{1, 2\}$, uma matriz Q coluna soma zero é dada por $Q = \begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix}$.

Neste caso,

$$\begin{aligned} Q^2 &= \begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} \begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} \\ &= \begin{pmatrix} \mu^2 + \lambda\mu & -\lambda\mu - \lambda^2 \\ -\mu^2 - \lambda\mu & \lambda\mu + \lambda^2 \end{pmatrix} \\ &= -(\mu + \lambda) \begin{pmatrix} -\mu & \lambda \\ \mu & -\lambda \end{pmatrix} \\ &= -(\mu + \lambda)Q \end{aligned}$$

Seguindo por indução, temos que

$$Q^n = (-1)^{n-1}(\mu + \lambda)^{n-1}Q.$$

Assim, para $t \in \mathbb{R}^+$ fixo,

$$\mathcal{P}(t) = \frac{1}{\mu + \lambda} \begin{pmatrix} \lambda + \mu e^{-t(\mu+\lambda)} & \lambda - \lambda e^{-t(\mu+\lambda)} \\ \mu - \mu e^{-t(\mu+\lambda)} & \mu + \lambda e^{-t(\mu+\lambda)} \end{pmatrix}.$$

Exemplo 2.35. Seja $S = \{1, 2\}$ e $Q = \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix}$. Então, pelo exemplo anterior, segue que

$$e^{tQ} = \begin{pmatrix} \frac{2}{3} + \frac{e^{-3t}}{3} & \frac{2}{3} - \frac{2e^{-3t}}{3} \\ \frac{1}{3} - \frac{e^{-3t}}{3} & \frac{1}{3} + \frac{2e^{-3t}}{3} \end{pmatrix}.$$

Exemplo 2.36. Seja $S = \{1, 2, 3\}$ e $Q = \begin{pmatrix} -2 & 0 & 2 \\ 2 & -1 & 1 \\ 0 & 1 & -3 \end{pmatrix}$. Então pode-se

mostrar que

$$e^{tQ} = \begin{pmatrix} \frac{1}{5} + e^{-3t} \left(\frac{4}{5} \cos(t) + \frac{2}{5} \operatorname{sen}(t) \right) & \frac{1}{5} + e^{-3t} \left(-\frac{1}{5} \cos(t) - \frac{3}{5} \operatorname{sen}(t) \right) & \frac{1}{5} + e^{-3t} \left(-\frac{1}{5} \cos(t) + \frac{7}{5} \operatorname{sen}(t) \right) \\ \frac{3}{5} + e^{-3t} \left(-\frac{3}{5} \cos(t) + \frac{1}{5} \operatorname{sen}(t) \right) & \frac{3}{5} + e^{-3t} \left(\frac{2}{5} \cos(t) + \frac{1}{5} \operatorname{sen}(t) \right) & \frac{3}{5} + e^{-3t} \left(-\frac{3}{5} \cos(t) - \frac{4}{5} \operatorname{sen}(t) \right) \\ \frac{1}{5} + e^{-3t} \left(-\frac{1}{5} \cos(t) - \frac{3}{5} \operatorname{sen}(t) \right) & \frac{1}{5} + e^{-3t} \left(-\frac{1}{5} \cos(t) + \frac{2}{5} \operatorname{sen}(t) \right) & \frac{1}{5} + e^{-3t} \left(\frac{4}{5} \cos(t) - \frac{3}{5} \operatorname{sen}(t) \right) \end{pmatrix}.$$

Definição 2.37. Um processo estocástico a tempo contínuo é uma família de variáveis aleatórias $\{X_t : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (S, \mathcal{G})\}_{t \in \mathbb{R}}$, onde $(\Omega, \mathcal{A}, \mathbb{P})$ é um espaço de probabilidade e (S, \mathcal{G}) é um espaço mensurável.

Trabalharemos nesta dissertação apenas com processos estocásticos que são contínuos à direita. No caso em que o espaço de estados é discreto, isto significa que, para todo $\omega \in \Omega$ e $t \geq 0$, existe $\delta > 0$ tal que $X_s(\omega) = X_t(\omega)$, para todo $s \in [t, t + \delta]$. Ainda, deve existir o limite à esquerda. Tais processos são usualmente chamados de processos CADLAG (do francês “continue à droite avec limite à gauche”).

Exemplo 2.38. Para $S = \{1, 2, 3\}$ um possível caminho está retratado abaixo

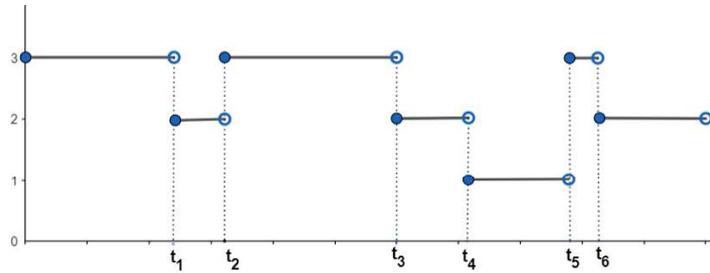


Fig. 2.2

Ao considerarmos apenas caminhos desta forma, a probabilidade de qualquer evento em Ω pode ser determinada pelas suas distribuições finito-dimensionais.

Definição 2.39. Seja $\{X_t : t \in \mathbb{R}\}$ um processo estocástico com espaço de estados $S = \{1, \dots, n\}$. Dizemos que X_t é um processo estocástico de Markov a tempo contínuo se para qualquer conjunto finito de tempos $0 \leq t_1 < t_2 < \dots < t_n < t_{n+1}$ e qualquer conjunto de estados $i_1, \dots, i_n, i_{n+1} \in S$ tais que $\mathbb{P}(X_{t_n} = i_n, \dots, X_{t_1} = i_1) > 0$, temos

$$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n, \dots, X_{t_1} = i_1) = \mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n).$$

Nesta dissertação, trataremos apenas do caso em que o processo é homogêneo no tempo, isto é, para quaisquer $0 < s, t < +\infty$, vale que $\mathbb{P}(X_{t+s} = i | X_s = j) = \mathbb{P}(X_t = i | X_0 = j)$. Com isto, definimos as probabilidades de transição de j para i da cadeia de Markov no tempo t , para qualquer $t \in \mathbb{R}$ da seguinte maneira:

$$\mathbb{P}(X_t = i | X_0 = j) = (\mathcal{P}(t))_{ij}.$$

Note que a equação acima realmente define uma probabilidade de transição e, além disso, se $t \in \mathbb{N}$, esta definição coincide com a definição para tempo discreto, como gostaríamos.

Definição 2.40. Fixada uma matriz coluna soma zero Q , um vetor de probabilidade π é dito estacionário, ou invariante, para o semigrupo $\mathcal{P}(t)$ se, para todo $t \in \mathbb{R}^+$, vale que

$$\mathcal{P}(t)\pi = \pi$$

Teorema 2.41. Seja Q o gerador infinitesimal de $\mathcal{P}(t)$. Então π é vetor estacionário para $\mathcal{P}(t)$ se, e somente se, $Q\pi = 0$.

Demonstração. (\implies) Considere a equação diferencial linear

$$\begin{cases} x'(t) = Qx(t) \\ x(0) = \pi \end{cases}$$

Como, por hipótese, $\pi = e^{tQ}\pi$, temos que a solução do sistema é constante e, portanto, $0 = x'(t) = Qx(t)$.

(\impliedby) Supondo que $Q\pi = 0$, temos que a solução de

$$\begin{cases} x'(t) = Qx(t) \\ x(0) = \pi \end{cases}$$

é dada por $x(t) = \pi$, para todo $t \in \mathbb{R}$. Por outro lado, $x(t)$ satisfaz $e^{tQ}\pi$. Portanto, $e^{tQ}\pi = \pi$

□

Teorema 2.42. Se Q é uma matriz coluna soma zero, então existe π vetor de probabilidade tal que π é estacionário para $\mathcal{P}(t) = e^{tQ}$.

Demonstração. Tal teorema pode ser demonstrado usando o teorema de Perron-Fröbenius para matrizes positivas. Aqui, contudo, optaremos por uma demonstração direta.

Para todo $n \in \mathbb{N}$, temos que $\mathcal{P}(\frac{1}{n})$ é estocástica. Logo, pelo teorema 2.23 existe π_n vetor de probabilidade tal que $\mathcal{P}(\frac{1}{n})\pi_n = \pi_n$.

Logo, para todo $n \in \mathbb{N}$, temos que

$$\frac{(P(\frac{1}{n}) - I)\pi_n}{\frac{1}{n}} = \frac{\mathcal{P}(\frac{1}{n})\pi_n - \pi_n}{\frac{1}{n}} = \frac{\pi_n - \pi_n}{\frac{1}{n}} = 0.$$

Como $\Sigma = \{p = (p_1, \dots, p_d) : p_i \geq 0, \sum_{i=1}^d p_i = 1\}$ é compacto, existe algum ponto de acumulação para a sequência π_n , ou seja, passando para uma subsequência se necessário, temos a existência de $\pi \in \Sigma$ tal que $\pi = \lim_{n \rightarrow +\infty} \pi_n$.

Além disso, temos que $\lim_{t \rightarrow 0} \frac{\mathcal{P}(t) - I}{t} = Q$ uniformemente. Portanto,

$$0 = \lim_{n \rightarrow +\infty} \frac{(P(\frac{1}{n}) - I) \pi_n}{\frac{1}{n}} = \lim_{n \rightarrow +\infty} \frac{(P(\frac{1}{n}) - I) \pi}{\frac{1}{n}} = Q\pi.$$

□

Temos, também, uma versão da *Equação de Chapman-Kolmogorov* para tempo contínuo:

Teorema 2.43. *Para quaisquer t e s positivos e $i, j \in S = \{1, \dots, d\}$ fixados, vale que*

$$(\mathcal{P}(t+s))_{ij} = \sum_{k=1}^d (\mathcal{P}(t))_{ik} (\mathcal{P}(s))_{kj}.$$

Podemos reescrever este teorema como

$$\mathcal{P}(t+s) = \mathcal{P}(t)\mathcal{P}(s).$$

Demonstração.

$$\begin{aligned} (\mathcal{P}(t+s))_{ij} &= \mathbb{P}(X_{t+s} = i | X_0 = j) \\ &= \sum_{k=1}^d \mathbb{P}(X_{t+s} = i, X_s = k | X_0 = j) \\ &= \sum_{k=1}^d \frac{\mathbb{P}(X_{t+s} = i, X_s = k, X_0 = j)}{\mathbb{P}(X_0 = j)} \\ &= \sum_{k=1}^d \frac{\mathbb{P}(X_{t+s} = i, X_s = k, X_0 = j)}{\mathbb{P}(X_s = k, X_0 = j)} \frac{\mathbb{P}(X_s = k, X_0 = j)}{\mathbb{P}(X_0 = j)} \\ &= \sum_{k=1}^d \mathbb{P}(X_{t+s} = i | X_s = k, X_0 = j) \mathbb{P}(X_s = k | X_0 = j) \\ &= \sum_{k=1}^d \mathbb{P}(X_{t+s} = i | X_s = k) \mathbb{P}(X_s = k | X_0 = j) \\ &= \sum_{k=1}^d \mathbb{P}(X_t = i | X_0 = k) \mathbb{P}(X_s = k | X_0 = j) \\ &= \sum_{k=1}^d \mathcal{P}(t)_{ik} \mathcal{P}(s)_{kj}, \end{aligned}$$

onde na sexta igualdade foi usada a propriedade de Markov e na sétima a homogeneidade no tempo. □

Por indução, obtemos que dados $t_1, t_2, \dots, t_n \geq 0$,

$$\mathcal{P}(t_1 + t_2 + \dots + t_n) = \mathcal{P}(t_1)\mathcal{P}(t_2) \dots \mathcal{P}(t_n).$$

Formalismo Termodinâmico

Neste capítulo, definiremos o *Operador de Ruelle*, que foi introduzido por David Ruelle na década de 60. Em *Mecânica Estatística*, este operador também é conhecido por operador de transferência.

Em Formalismo Termodinâmico, tem-se como um dos principais objetivos encontrar uma autofunção para o *Operador de Ruelle* e uma automedida para seu dual. Esta automedida está associada à medida que satisfaz o *princípio variacional*, um princípio da física que diz que a medida que maximiza a soma da entropia com o valor médio de um observável (que codifica informações físicas de relevância) está associada a *estados de equilíbrio* na natureza. Para isso, utiliza-se fortemente o conhecido *Teorema de Ruelle-Perron-Fröbenius*, o qual enunciaremos e demonstraremos aqui.

3.1 Operador de Ruelle

Sejam (M, d) espaço métrico compacto e \mathcal{A} a σ -álgebra de Borel sobre M . Como M é compacto, podemos considerar sem perda de generalidade que $d(x, y) \leq 1$ para todo par de pontos x e y em M . Consideremos o espaço das sequências $\Omega = M^{\mathbb{N}} = \{x = (x_0, x_1, x_2, \dots) \mid x_i \in M, i \in \mathbb{N}\}$ com \mathcal{B} σ -álgebra de Borel sobre Ω . Ao equiparmos Ω com distância $\tilde{d}(x, y) = \sum_{n=0}^{+\infty} \frac{d(x_n, y_n)}{2^{n+1}}$, temos, pelo *Teorema de Tychonoff* (ver [13]), que Ω é compacto. Considere o conjunto $C(\Omega) = \{g : \Omega \rightarrow \mathbb{R} \mid g \text{ é contínua}\}$. Neste conjunto, trabalharemos com a norma $\|\cdot\|_0$ dada por

$$\|g\|_0 = \max_{x \in \Omega} |g(x)|.$$

Esta norma também é conhecida como norma do máximo, ou da convergência uniforme, e faz de $(C(\Omega), \|\cdot\|_0)$ um espaço de Banach (ver [13]).

Definição 3.1. Fixadas uma probabilidade $\nu: \mathcal{A} \rightarrow [0, 1]$ a priori e uma aplicação $A: \Omega \rightarrow \mathbb{R}$ contínua, o operador de Ruelle $\mathcal{L}_A: C(\Omega) \rightarrow C(\Omega)$ é dado por

$$\mathcal{L}_A(\varphi)(x) = \int_M e^{A(ax)} \varphi(ax) d\nu(a),$$

para todo $x = (x_0, x_1, \dots)$, onde $ax = (a, x_0, x_1, \dots)$.

Para maiores detalhes sobre formalismo termodinâmico com medidas a priori quaisquer, o leitor pode consultar [5].

Observação 3.2. No caso em que $M = \{1, \dots, d\}$ é espaço finito, e a medida a priori é a medida da contagem, a definição acima se reduz a

$$\mathcal{L}_A(\varphi)(x) = \frac{1}{d} \sum_{a=1}^d e^{A(ax)} \varphi(ax).$$

Usualmente, a aplicação A é chamada de *potencial*. A próxima proposição garante que o operador \mathcal{L}_A está bem definido.

Proposição 3.3. $\mathcal{L}_A(\varphi) \in C(\Omega)$, para toda $\varphi \in C(\Omega)$.

Demonstração. Dados $x, y \in \Omega$, vale a seguinte desigualdade

$$\begin{aligned} |\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| &= \left| \int_M e^{A(ax)} \varphi(ax) d\nu(a) - \int_M e^{A(ay)} \varphi(ay) d\nu(a) \right| \\ &= \left| \int_M (e^{A(ax)} \varphi(ax) - e^{A(ay)} \varphi(ay)) d\nu(a) \right| \\ &\leq \int_M |e^{A(ax)} \varphi(ax) - e^{A(ay)} \varphi(ay)| d\nu(a) \\ &= \int_M |e^{A(ax)} \varphi(ax) - e^{A(ax)} \varphi(ay) + e^{A(ax)} \varphi(ay) - e^{A(ay)} \varphi(ay)| d\nu(a) \\ &\leq \int_M |e^{A(ax)} \varphi(ax) - e^{A(ax)} \varphi(ay)| + |e^{A(ax)} \varphi(ay) - e^{A(ay)} \varphi(ay)| d\nu(a) \\ &= \int_M |e^{A(ax)} (\varphi(ax) - \varphi(ay))| + |\varphi(ay) (e^{A(ax)} - e^{A(ay)})| d\nu(a) \end{aligned}$$

Seja $w = \max_{z \in \Omega} |A(z)|$. Então, como consequência do *Teorema do Valor Médio*, temos que

$$|e^{A(ax)} - e^{A(ay)}| \leq e^w |A(ax) - A(ay)|,$$

e, portanto,

$$\begin{aligned} |\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| &\leq \int_M |e^{A(ax)}(\varphi(ax) - \varphi(ay))| + |\varphi(ay)e^w(A(ax) - A(ay))| d\nu(a) \\ &\leq \int_M |e^w(\varphi(ax) - \varphi(ay))| + |e^w\varphi(ay)(A(ax) - A(ay))| d\nu(a) \\ &\leq \int_M e^w \left(|\varphi(ax) - \varphi(ay)| + \left| \max_{z \in \Omega} \varphi(z)(A(ax) - A(ay)) \right| \right) d\nu(a). \end{aligned}$$

Como Ω é compacto e φ e A são contínuas, φ e A são uniformemente contínuas. Logo, dado $\epsilon > 0$, existe $\delta > 0$ tal que

$$|\varphi(ax) - \varphi(ay)| < \epsilon$$

e

$$|A(ax) - A(ay)| < \epsilon$$

sempre que $\tilde{d}(ax, ay) = \tilde{d}(x, y)/2 < \delta$.

Sendo assim, se $\tilde{d}(x, y) < 2\delta$, temos

$$\begin{aligned} |\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| &< \int_M e^w \left(\epsilon + \left| \max_{z \in \Omega} \varphi(z) \right| \epsilon \right) d\nu \\ &= K\epsilon, \end{aligned}$$

onde $K = \nu(M)(e^w + e^w |\max_{z \in \Omega} \varphi(z)|) = e^w + e^w |\max_{z \in \Omega} \varphi(z)|$. Logo, $\mathcal{L}_A(\varphi)$ é uniformemente contínua e o resultado segue. \square

Proposição 3.4. *Seja $\mathcal{L}_A: C(\Omega) \rightarrow C(\Omega)$ o operador de Ruelle associado ao potencial $A: \Omega \rightarrow \mathbb{R}$. São válidas as seguintes propriedades:*

- (i) *O operador \mathcal{L}_A é linear, isto é, $\mathcal{L}_A(\varphi + \psi) = \mathcal{L}_A(\varphi) + \mathcal{L}_A(\psi)$ e $\mathcal{L}_A(\beta\varphi) = \beta\mathcal{L}_A(\varphi)$, para quaisquer $\varphi, \psi \in C(\Omega)$ e $\beta \in \mathbb{R}$.*
- (ii) *\mathcal{L}_A é um operador positivo, isto é, se $\varphi \in C(\Omega)$ é tal que $\varphi(x) \geq 0$, para todo $x \in \Omega$, então $\mathcal{L}_A(\varphi)(x) \geq 0$, para todo $x \in \Omega$.*

Demonstração.

- (i) Esta propriedade segue diretamente da linearidade da integral.
- (ii) Se $\varphi(x) \geq 0$, para todo $x \in \Omega$, temos que $e^{A(ax)}\varphi(ax) \geq 0$, para todo $x \in \Omega$ e o resultado segue.

\square

Definição 3.5. A aplicação *shift*, ou *deslocamento de Bernoulli*, é a dinâmica $\sigma: \Omega \rightarrow \Omega$ definida por

$$\sigma((x_0, x_1, x_2, \dots)) = (x_1, x_2, \dots)$$

Denotaremos por σ^i o i -ésimo iterado de σ , ou seja, $\sigma^i((x_0, x_1, x_2, \dots)) = (x_i, x_{i+1}, \dots)$. Com isso, uma pergunta natural que surge é como se comportam os n -ésimos iterados do Operador de Ruelle, que denotaremos por \mathcal{L}_A^n . Considere a aplicação $S_n(A)(x) = \sum_{i=0}^{n-1} A \circ \sigma^i(x)$. Então é de fácil verificação que

$$\mathcal{L}_A^n(\varphi)(x) = \int_{M^n} \varphi(a_1 a_2 \dots a_n x) e^{S_n(A)(a_1 a_2 \dots a_n x)} d\nu(a_1) \dots d\nu(a_n).$$

Observação 3.6. No caso em que $M = \{1, \dots, d\}$, temos que

$$\mathcal{L}_A^n(\varphi)(x) = \frac{1}{d^n} \sum_{a_n=1}^d \dots \sum_{a_1=1}^d e^{A(a_1 \dots a_n x) + \dots + A(a_n x)} \varphi(a_1 \dots a_n x).$$

Proposição 3.7. $\mathcal{L}_A(\psi(\varphi \circ \sigma))(x) = \varphi(x) \mathcal{L}_A(\psi)(x)$, para quaisquer $\varphi, \psi \in C(\Omega)$.

Demonstração.

$$\begin{aligned} \mathcal{L}_A(\psi(\varphi \circ \sigma))(x) &= \int_M e^{A(ax)} \psi(ax) \varphi \circ \sigma(ax) d\nu(a) \\ &= \int_M e^{A(ax)} \psi(ax) \varphi(x) d\nu(a) \\ &= \varphi(x) \int_M e^{A(ax)} \psi(ax) d\nu(a) \\ &= \varphi(x) \mathcal{L}_A(\psi)(x). \end{aligned}$$

□

Um caso importante é quando $\psi = 1$ e está enunciado abaixo:

Corolário 3.8. $\mathcal{L}_A(\varphi \circ \sigma)(x) = \varphi(x) \mathcal{L}_A(1)(x)$, para toda $\varphi \in C(\Omega)$.

A partir de agora, assumiremos que A é uma função α -Hölder, ou seja, existem constantes $0 < \alpha \leq 1$ e $K > 0$ tais que

$$|A(x) - A(y)| \leq K \tilde{d}^\alpha(x, y),$$

para todo $x, y \in \Omega$. A menor das constantes K satisfazendo a equação acima é chamada constante de Hölder e denotada por $Hol(A)$. Denotaremos $C^\alpha(\Omega) = \{g: \Omega \rightarrow \mathbb{R} : g \text{ é } \alpha\text{-Hölder}\}$.

Observação 3.9. Toda função α -Hölder é claramente contínua.

Exemplo 3.10. Se $\Omega = M^{\mathbb{N}}$, com $M = \{1, \dots, d\}$, e $A: \Omega \rightarrow \mathbb{R}$ depende de apenas de finitas coordenadas, isto é, existe $k \in \mathbb{N}$ tal que $A(x_0, x_1, x_2, \dots) = A(x_0, x_1, \dots, x_k)$, então A é α -Hölder.

Como Ω é compacto, pelo Teorema de Representação de Riesz-Markov temos que o espaço dual de $C(\Omega)$, denotado por $(C(\Omega))^*$, é isomorfo ao conjunto das medidas $\mathcal{M}(\Omega)$. Sendo assim, podemos definir o dual do operador de Ruelle, \mathcal{L}_A^* , como sendo o único operador linear agindo nas medidas a satisfazer a seguinte identidade

$$\int_{\Omega} \mathcal{L}_A(\varphi) d\mu = \int_{\Omega} \varphi d(\mathcal{L}_A^*(\mu))$$

para toda $\varphi \in C(\Omega)$.

Definição 3.11. Se $\mathcal{L}_A(1) = 1$, dizemos que \mathcal{L}_A está normalizado.

Proposição 3.12. Se \mathcal{L}_A está normalizado e μ é uma probabilidade, então $\mathcal{L}_A^*(\mu)$ é probabilidade.

Demonstração. Dada $\varphi \geq 0$, temos que

$$\int \varphi d\mathcal{L}_A^*(\mu) = \int \mathcal{L}_A(\varphi) d\mu \geq 0,$$

dado que \mathcal{L}_A é operador positivo. Também temos que

$$\int 1 d\mathcal{L}_A^*(\mu) = \int \mathcal{L}_A(1) d\mu = \int 1 d\mu = 1,$$

pois o operador está normalizado.

Logo, $\mathcal{L}_A^*(\mu)$ é medida de probabilidade. \square

Com o intuito de preparar o leitor para o *Teorema de Ruelle-Perron-Fröbenius*, enunciaremos e demonstraremos um resultado técnico necessário à prova deste importante teorema.

Lema 3.13. Seja $A: \Omega \rightarrow \mathbb{R}$ potencial α -Hölder e considere o operador de Ruelle associado $\mathcal{L}_A: C^\alpha(\Omega) \rightarrow C^\alpha(\Omega)$. Se \mathcal{L}_A está normalizado, então para toda $\varphi \in C^\alpha(\Omega)$, temos que $\mathcal{L}_A^n(\varphi) \in C^\alpha(\Omega)$ e vale

$$Hol(\mathcal{L}_A^n(\varphi)) \leq C \|\varphi\|_0 + \frac{1}{2^{n\alpha}} Hol(\varphi),$$

onde a constante $C > 0$ não depende de n .

Demonstração. Demonstraremos este resultado por indução.

Para $n = 1$, temos

$$\begin{aligned}
|\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| &= \left| \int_M \varphi(ax)e^{A(ax)} d\nu(a) - \int_M \varphi(ay)e^{A(ay)} d\nu(a) \right| \\
&= \left| \int_M \varphi(ax)e^{A(ax)} - \varphi(ax)e^{A(ay)} + \varphi(ax)e^{A(ay)} - \varphi(ay)e^{A(ay)} d\nu(a) \right| \\
&\leq \int_M |\varphi(ax)e^{A(ax)} - \varphi(ax)e^{A(ay)} + \varphi(ax)e^{A(ay)} - \varphi(ay)e^{A(ay)}| d\nu(a) \\
&\leq \int_M |\varphi(ax)e^{A(ax)} - \varphi(ax)e^{A(ay)}| + |\varphi(ax)e^{A(ay)} - \varphi(ay)e^{A(ay)}| d\nu(a) \\
&\leq \int_M \|\varphi\|_0 |e^{A(ax)} - e^{A(ay)}| d\nu(a) + \int_M |e^{A(ay)}| |\varphi(ax) - \varphi(ay)| d\nu(a)
\end{aligned}$$

Como φ é α -Hölder, temos que $|\varphi(ax) - \varphi(ay)| \leq Hol(\varphi)d^\alpha(ax, ay) = Hol(\varphi)\frac{1}{2^\alpha}d^\alpha(x, y)$. Substituindo este fato na desigualdade acima, obtemos:

$$\begin{aligned}
|\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| &\leq \int_M \|\varphi\|_0 |e^{A(ax)} - e^{A(ay)}| d\nu(a) + Hol(\varphi)\frac{1}{2^\alpha}d^\alpha(x, y) \int_M |e^{A(ay)}| d\nu(a) \\
&= \int_M \|\varphi\|_0 |e^{A(ax)} - e^{A(ay)}| d\nu(a) + Hol(\varphi)\frac{1}{2^\alpha}d^\alpha(x, y) \int_M e^{A(ay)} d\nu(a) \\
&= \|\varphi\|_0 \int_M |e^{A(ax)} - e^{A(ay)}| d\nu(a) + Hol(\varphi)\frac{1}{2^\alpha}d^\alpha(x, y)\mathcal{L}_A(1)(y) \\
&= \|\varphi\|_0 \int_M |e^{A(ax)} - e^{A(ay)}| d\nu(a) + Hol(\varphi)\frac{1}{2^\alpha}d^\alpha(x, y),
\end{aligned}$$

onde na última igualdade utilizamos que \mathcal{L}_A está normalizado.

Ainda, como consequência do *Teorema do Valor Médio*, temos que existe C_1 tal que

$$\int_M |e^{A(ax)} - e^{A(ay)}| d\nu(a) \leq C_1 \int_M |A(ax) - A(ay)| d\nu(a).$$

Além disso, como $A \in C^\alpha(\Omega)$,

$$\begin{aligned}
\int_M |e^{A(ax)} - e^{A(ay)}| d\nu(a) &\leq C_1 \int_M |A(ax) - A(ay)| d\nu(a) \\
&\leq C_1 Hol(A)d^\alpha(ax, ay) \int_M d\nu(a) \\
&= C_1 Hol(A)d^\alpha(ax, ay) \\
&= C_1 Hol(A)\frac{1}{2^\alpha}d^\alpha(x, y)
\end{aligned}$$

Portanto,

$$|\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| \leq \|\varphi\|_0 C_1 Hol(A) \frac{1}{2^\alpha} d^\alpha(x, y) + Hol(\varphi) \frac{1}{2^\alpha} d^\alpha(x, y)$$

Pondo $C_0 = C_1 Hol(A) \frac{1}{2^\alpha}$, chegamos a

$$|\mathcal{L}_A(\varphi)(x) - \mathcal{L}_A(\varphi)(y)| \leq \left(C_0 \|\varphi\|_0 + Hol(\varphi) \frac{1}{2^\alpha} \right) d^\alpha(x, y).$$

Assim, está provado que $\mathcal{L}_A(\varphi) \in C^\alpha(\Omega)$ e $Hol(\mathcal{L}_A(\varphi)) \leq C_0 \|\varphi\|_0 + \frac{1}{2^\alpha} Hol(\varphi)$.

Suponha agora o resultado válido para n , isto é, suponha que $\mathcal{L}_A^n(\varphi) \in C^\alpha(\Omega)$ e que $Hol(\mathcal{L}_A^n(\varphi)) \leq C_n \|\varphi\|_0 + \frac{1}{2^{n\alpha}} Hol(\varphi)$. Provemos que o mesmo vale para $n + 1$.

$$\begin{aligned} Hol(\mathcal{L}_A^{n+1}(\varphi)) &= Hol(\mathcal{L}_A^n(\mathcal{L}_A(\varphi))) \\ &\leq C_n \|\mathcal{L}_A(\varphi)\|_0 + \frac{1}{2^{n\alpha}} Hol(\mathcal{L}_A(\varphi)) \\ &\leq C_n \|\varphi\|_0 + \frac{1}{2^{n\alpha}} \left(C_0 \|\varphi\|_0 + \frac{1}{2^\alpha} Hol(\varphi) \right) \\ &= \|\varphi\|_0 \left(C_n + C_0 \frac{1}{2^{n\alpha}} \right) + \frac{1}{2^{(n+1)\alpha}} Hol(\varphi). \end{aligned}$$

Tomando $C_{n+1} = C_n + C_0 \frac{1}{2^{n\alpha}}$, obtemos o resultado.

Resta verificarmos que a sequência $\{C_n\}_{n \in \mathbb{N}}$ é limitada.

De fato,

$$C_n = C_0 \left(\sum_{k=0}^{n-1} \left(\frac{1}{2^\alpha} \right)^k \right) \leq C_0 \frac{1}{1 - \frac{1}{2^\alpha}}.$$

Sendo assim, tomando $C = C_0 \frac{1}{1 - \frac{1}{2^\alpha}}$, obtemos

$$Hol(\mathcal{L}_A^n(\varphi)) \leq C \|\varphi\|_0 + \frac{1}{2^{n\alpha}} Hol(\varphi),$$

como queríamos. □

O próximo teorema, Teorema de Ruelle-Perron-Fröbenius, é de fundamental importância ao nos fornecer a existência de uma automedida para o dual do operador de Ruelle, e também uma forma computacionalmente eficiente de calcular esta automedida. Como dissemos no início deste capítulo, tal automedida está associada a medidas de equilíbrio: o produto da automedida pela autofunção do operador satisfaz o princípio variacional (que não será detalhado nesta dissertação).

Teorema 3.14. (*Ruelle-Perron-Fröbenius*) *Seja A potencial α -Hölder para o operador de Ruelle \mathcal{L}_A . Então*

- (i) \mathcal{L}_A possui um autovalor positivo, λ_A , com h_A autofunção associada, isto é, $\mathcal{L}_A(h_A) = \lambda_A h_A$. Além disso, h_A é α -Hölder e positiva.
- (ii) Existe uma única probabilidade μ_A em \mathcal{A} tal que $\mathcal{L}_A^* \mu_A = \lambda_A \mu_A$;
- (iii) $\lim_{n \rightarrow +\infty} \left\| \frac{1}{\lambda_A^n} \mathcal{L}_A^n \varphi - h_A \int_{\Omega} \varphi d\mu_A \right\|_0 = 0$, para toda $\varphi \in C(\Omega)$.

Antes de demonstrar o teorema de Ruelle-Perron-Fröbenius, observamos que ele permite "normalizar" um potencial:

Observação 3.15. No caso em que o operador \mathcal{L}_A não está normalizado, isto é, $\mathcal{L}_A(1) \neq 1$, podemos considerar um novo potencial $\bar{A} = A + \log h_A - \log h_A \circ \sigma - \log \lambda_A$ e teremos que $\mathcal{L}_{\bar{A}}$ está normalizado.

De fato,

$$\begin{aligned}
 \mathcal{L}_{\bar{A}}(1)(x) &= \int_M e^{\bar{A}(ax)} 1(ax) d\nu \\
 &= \int_M e^{A(ax) + \log h_A(ax) - \log h_A \circ \sigma(ax) - \log \lambda_A} d\nu \\
 &= \int_M e^{A(ax)} e^{\log h_A(ax)} e^{-\log h_A(x)} e^{-\log \lambda_A} d\nu \\
 &= \int_M e^{A(ax)} \frac{h_A(ax)}{\lambda_A h_A(x)} d\nu \\
 &= \frac{1}{\lambda_A h_A(x)} \int_M e^{A(ax)} h_A(ax) d\nu \\
 &= \frac{1}{\lambda_A h_A(x)} \mathcal{L}_A(h_A)(x) \\
 &= \frac{1}{\lambda_A h_A(x)} \lambda_A h_A(x) \\
 &= 1.
 \end{aligned}$$

Agora, demonstraremos o teorema de Ruelle-Perron-Fröbenius, adaptando uma demonstração contida em [15] (a demonstração de [15] foi feita originalmente para o operador de Ruelle em finitos símbolos, veja a observação 3.2).

Demonstração.

(i) Fixado $c > 0$, considere o conjunto

$$\Lambda_c = \{g: \Omega \rightarrow \mathbb{R} \mid 0 \leq g(x) \leq 1, \forall x \in \Omega \text{ e } g(x) \leq g(y)e^{c\tilde{d}^\alpha(x,y)}, \forall x, y \in \Omega\}.$$

Note que se $b < c$, então temos que $\Lambda_b \subset \Lambda_c$. É fácil ver que Λ_c é fechado e convexo. Provemos agora que $\Lambda_c \subset C^\alpha(\Omega)$ e Λ_c é uma família uniformemente contínua: dados $x, y \in \Omega$, temos

$$\begin{aligned} |g(x) - g(y)| &\leq \left| g(y)e^{c\tilde{d}^\alpha(x,y)} - g(y) \right| \\ &= \left| g(y) \left(e^{c\tilde{d}^\alpha(x,y)} - 1 \right) \right| \\ &\leq \|g\|_0 \left| e^{c\tilde{d}^\alpha(x,y)} - 1 \right| \\ &< \|g\|_0 c e^c \tilde{d}^\alpha(x,y), \end{aligned}$$

onde na última desigualdade utilizamos a expansão em série de Taylor da função exponencial e o fato de que $\tilde{d}(x,y) \leq 1$:

$$\begin{aligned} e^{c\tilde{d}^\alpha(x,y)} - 1 &= c\tilde{d}^\alpha(x,y) + \frac{c^2}{2}\tilde{d}^{2\alpha}(x,y) + \frac{c^3}{6}\tilde{d}^{3\alpha}(x,y) + \dots \\ &\leq \left(c + \frac{c^2}{2} + \frac{c^3}{6} + \dots \right) \tilde{d}^\alpha(x,y) < c e^c \tilde{d}^\alpha(x,y) \end{aligned}$$

Logo, tomando $K = \|g\|_0 c e^c$ temos

$$|g(x) - g(y)| \leq K \tilde{d}^\alpha(x,y).$$

Com isso, podemos aplicar o *Teorema de Arzelá-Áscoli*, de modo a concluir que Λ_c é compacto na norma $\|\cdot\|_0$.

Além disso, vejamos que se c é suficientemente grande, $\frac{\mathcal{L}_A(g)}{\|\mathcal{L}_A(g)\|_0} \in \Lambda_c$ se $g \in \Lambda_c$. De fato, seja $g \in \Lambda_c$. Então

$$\begin{aligned} \mathcal{L}_A(g)(x) &= \int_M e^{A(ax)} g(ax) d\nu(a) \\ &\leq \int_M e^{A(ax)} g(ay) e^{c\tilde{d}^\alpha(ax,ay)} d\nu(a) \\ &= \int_M e^{A(ay)} e^{A(ax)-A(ay)} g(ay) e^{c\tilde{d}^\alpha(ax,ay)} d\nu(a) \end{aligned}$$

Como o potencial A é Hölder, temos que $|A(ax) - A(ay)| \leq \text{Hol}(A)\tilde{d}(ax, ay)$. Além disso, $\tilde{d}^\alpha(ax, ay) = \left(\frac{\tilde{d}(x,y)}{2}\right)^\alpha$. Substituindo estas informações na inequação acima, obtemos

$$\begin{aligned}
\mathcal{L}_A(g)(x) &\leq \int_M e^{A(ay)} e^{A(ax)-A(ay)} g(ay) e^{c\tilde{d}^\alpha(ax,ay)} d\nu(a) \\
&\leq \int_M e^{A(ay)} g(ay) e^{Hol(A)\tilde{d}(ax,ay)} e^{c\tilde{d}^\alpha(ax,ay)} d\nu(a) \\
&\leq \int_M e^{A(ay)} g(ay) e^{Hol(A)\left(\frac{\tilde{d}(x,y)}{2}\right)^\alpha} e^{c\left(\frac{\tilde{d}(x,y)}{2}\right)^\alpha} d\nu(a) \\
&= e^{Hol(A)\left(\frac{\tilde{d}(x,y)}{2}\right)^\alpha} e^{c\left(\frac{\tilde{d}(x,y)}{2}\right)^\alpha} \int_M e^{A(ay)} g(ay) d\nu(a) \\
&= e^{\left(\frac{\tilde{d}^\alpha(x,y)}{2^\alpha}(Hol(A)+c)\right)} \mathcal{L}_A(g)(y)
\end{aligned}$$

Concluimos assim que $\frac{\mathcal{L}_A(g)}{\|\mathcal{L}_A(g)\|_0} \in \Lambda_{c^*}$, onde $c^* = \frac{Hol(A)+c}{2^\alpha}$. Assim, se c é suficientemente grande para que $c^* = \frac{Hol(A)+c}{2^\alpha} < c$, teremos que $\frac{\mathcal{L}_A(g)}{\|\mathcal{L}_A(g)\|_0} \in \Lambda_c$. Usando o fato de que, para qualquer $n \in \mathbb{N}$, a função constante $f_n(x) = \frac{1}{n} \in \Lambda_c$, para todo $c \in \mathbb{R}$, podemos aplicar o que foi feito acima para $g + f_n = g + 1/n$, e obtemos

$$\mathcal{L}_A(g + 1/n)(x) \leq e^{c\tilde{d}(x,y)} \mathcal{L}_A(g + 1/n)(y), \quad (3.1)$$

para c suficientemente grande. Agora, para cada $n \geq 1$, defina o operador contínuo L_n da seguinte maneira

$$L_n(g) = \frac{\mathcal{L}_A(g + 1/n)}{\|\mathcal{L}_A(g + 1/n)\|_0}.$$

É claro que $\|L_n\|_0 = 1$. Dividindo a equação 3.1 por $\|\mathcal{L}_A(g + 1/n)\|_0$, obtemos

$$\frac{\mathcal{L}_A(g + 1/n)(x)}{\|\mathcal{L}_A(g + 1/n)\|_0} \leq e^{c\tilde{d}(x,y)} \frac{\mathcal{L}_A(g + 1/n)(y)}{\|\mathcal{L}_A(g + 1/n)\|_0},$$

ou seja,

$$L_n(g)(x) \leq e^{c\tilde{d}(x,y)} L_n(g)(y)$$

Isso posto, temos que $L_n(\Lambda_c) \subset \Lambda_c$, para c suficientemente grande. Agora estamos aptos a aplicar o *Teorema de Tychonov-Schauder* e, assim, para cada $n \in \mathbb{N}$, existe um ponto fixo h_n para o operador L_n , ou seja, $L_n(h_n) = h_n$. Em outras palavras, vale que

$$\mathcal{L}_A(h_n + 1/n) = \|\mathcal{L}_A(h_n + 1/n)\|_0 h_n.$$

Como Λ_c é compacto, existe $h_A \in \Lambda_c$ ponto de acumulação da sequência $\{h_n\}_{n \in \mathbb{N}}$. Pela continuidade de \mathcal{L}_A , obtemos $\mathcal{L}_A(h_A) = \lambda_A h_A$, com $\lambda_A =$

$\|\mathcal{L}_A(h_A)\|_0$. Falta verificarmos ainda que $\lambda_A > 0$ e que h_A é estritamente positiva (já sabemos que h_A é positiva pois é limite de funções positivas). Para verificarmos que $\lambda_A > 0$, note que

$$\begin{aligned} \|\mathcal{L}_A(h_n + 1/n)\|_0 h_n &= \mathcal{L}_A(h_n + 1/n) \\ &= \int_M e^{A(ax)} (h_n(ax) + 1/n) d\nu(a) \\ &\geq e^{-\|A\|_0} \int_M (h_n(ax) + 1/n) d\nu(a) \\ &\geq e^{-\|A\|_0} \left(\inf_{z \in \Omega} h_n(z) + 1/n \right) \end{aligned}$$

Portanto,

$$\begin{aligned} \|\mathcal{L}_A(h_n + 1/n)\|_0 \inf_{z \in \Omega} h_n &\geq e^{-\|A\|_0} \left(\inf_{z \in \Omega} h_n(z) + 1/n \right) \\ &\geq e^{-\|A\|_0} \inf_{z \in \Omega} h_n(z). \end{aligned}$$

Dividindo a inequação por $\inf_{z \in \Omega} h_n(z)$, chegamos a

$$\|\mathcal{L}_A(h_n + 1/n)\|_0 \geq e^{-\|A\|_0}.$$

Logo, $\lambda_A = \|\mathcal{L}_A(h)\|_0 = \lim_{n \rightarrow +\infty} \|\mathcal{L}_A(h_n + 1/n)\|_0 > e^{-\|A\|_0} > 0$.

Mostremos agora que h_A é estritamente positiva. Suponha, por absurdo, que exista $x \in \Omega$ tal que $h_A(x) = 0$. Então

$$\begin{aligned} 0 &= \lambda_A h_A \\ &= \mathcal{L}_A(h_A)(x) \\ &= \int_M e^{A(ax)} h_A(ax) d\nu(a), \end{aligned}$$

visto que h_A é autofunção de \mathcal{L}_A associada a λ_A . Seguindo indutivamente, temos que, para todo $n \geq 1$, vale

$$0 = \int_{M^n} e^{S_n(A)(a_1 a_2 \dots a_n x)} h_A(a_1 a_2 \dots a_n x) d\nu(a_1) \dots d\nu(a_n).$$

Como o conjunto $\{a_1 \dots a_n x \mid a_i \in M, n = 1, 2, \dots, n\}$ é denso em Ω e h_A é uma função contínua, teríamos que $h_A \equiv 0$. Mas isto é absurdo, pois $0 < \lambda_A = \|\mathcal{L}_A(h_A)\|_0$.

(ii) Suponhamos, sem perda de generalidade (pela observação 3.15), que \mathcal{L}_A está normalizado. Seja $\mathcal{M}_\sigma(\Omega) = \{\mu: \mathcal{B} \rightarrow [0, 1] \mid \mu \text{ é invariante por } \sigma\}$. É um fato conhecido que $\mathcal{M}_\sigma(\Omega)$ é compacto e convexo. Vejamos que o operador dual de \mathcal{L}_A , $\mathcal{L}_A^*: C(\Omega)^* \rightarrow C(\Omega)^*$, preserva $\mathcal{M}_\sigma(\Omega)$, isto é, dada $m \in \mathcal{M}_\sigma(\Omega)$, temos que $\mathcal{L}_A^*(m) \in \mathcal{M}_\sigma(\Omega)$. Para tal, basta mostrarmos que $\mathcal{L}_A^*(m)$ é invariante, ou seja, dada $\varphi \in C(\Omega)$, vale que $\int \varphi \circ \sigma d[\mathcal{L}_A^*(m)] = \int \varphi dm$.

De fato,

$$\begin{aligned} \int \varphi \circ \sigma d[\mathcal{L}_A^*(m)] &= \int \mathcal{L}_A(\varphi \circ \sigma) dm \\ &= \int \varphi \mathcal{L}_A(1) dm \\ &= \int \varphi dm, \end{aligned}$$

onde na segunda igualdade utilizamos o corolário 3.8 e na última igualdade utilizamos que \mathcal{L}_A está normalizado.

Ainda, da proposição 3.12, temos que $\mathcal{L}_A^*(m)$ é probabilidade.

Consequentemente, pelo *Teorema de Tychonov-Shauder*, garantimos a existência de $\mu_A \in \mathcal{M}_\sigma(\Omega)$ tal que $\mathcal{L}_A^*(\mu_A) = \mu_A$.

(iii) Suponhamos, novamente, que \mathcal{L}_A está normalizado. Assim, para provarmos este item, basta provarmos que

$$\lim_{n \rightarrow +\infty} \mathcal{L}_A^n(\varphi) = \int \varphi d\mu_A,$$

para toda $\varphi \in C(\Omega)$.

Pelo lema 3.13, temos que $\{\mathcal{L}_A^n(\varphi)\}_{n \in \mathbb{N}}$ é uma sequência equicontínua e, como \mathcal{L}_A é normalizado, temos que $\{\mathcal{L}_A^n\}_{n \in \mathbb{N}}$ é uniformemente limitada. Portanto, pelo *Teorema de Arzelà-Áscoli*, existe uma subsequência $\{\mathcal{L}_A^{n_k}\}_{k \in \mathbb{N}}$ convergente na norma da convergência uniforme. Seja $\psi = \lim_{k \rightarrow +\infty} \mathcal{L}_A^{n_k}$. Mostremos que ψ é uma função constante. Como \mathcal{L}_A é normalizado, para qualquer $\varphi \in C^\alpha(\Omega)$, vale a seguinte cadeia de desigualdades

$$\min \varphi \leq \min \mathcal{L}_A(\varphi) \leq \dots \leq \min \mathcal{L}_A^{n_k}(\varphi) \leq \dots \leq \min \psi.$$

Logo, para todo $m \in \mathbb{N}$, vale

$$\min \mathcal{L}_A^m(\psi) = \min \mathcal{L}_A^m\left(\lim_{k \rightarrow +\infty} \mathcal{L}_A^{n_k}(\varphi)\right) = \min \left(\lim_{k \rightarrow +\infty} \mathcal{L}_A^{m+n_k}(\varphi)\right) = \min \psi,$$

onde a última igualdade segue da monotonicidade da sequência $\mathcal{L}_A^m(\varphi)$ e de $\lim_{k \rightarrow +\infty} \mathcal{L}_A^{n_k}(\varphi) = \psi$. Dado $\epsilon > 0$, seja $N \in \mathbb{N}$ e $x \in \Omega$ tais que $\min \mathcal{L}_A^N(\psi) =$

$\mathcal{L}_A(\psi)(x)$ (a existência deste x é garantida pela continuidade de $\mathcal{L}_A(\psi)$ e pela compacidade de Ω) e $\{a_1 \dots a_N x \mid a_i \in M, i \in \{1, \dots, n\}\}$ é ϵ -denso em Ω . Como $\min \mathcal{L}_A^m(\psi) = \min \psi$, para todo $m \in \mathbb{N}$, temos que $\min \psi = \min \mathcal{L}_A^m(\psi) = \min \mathcal{L}_A(\psi) = \mathcal{L}_A(\psi)(x)$, para todo $m \in \mathbb{N}$. Seja $z \in \Omega$ tal que $\psi(z) = \min \psi$. Então

$$\psi(z) = \min \psi = \mathcal{L}_A(\psi)(x) = \mathcal{L}_A^N(\psi)(x) = \int_{M^N} e^{S_N(A)(a_1 \dots a_N x)} \psi(a_1 \dots a_N x) d\nu(a_1) \dots d\nu(a_N).$$

Como \mathcal{L}_A está normalizado, temos que $\mathcal{L}_A^N(1) = 1$ e, dessa forma, podemos reescrever

$$\begin{aligned} 0 &= \int_{M^N} e^{S_N(A)(a_1 \dots a_N x)} \psi(a_1 \dots a_N x) d\nu(a_1) \dots d\nu(a_N) - \psi(z) \\ &= \int_{M^N} e^{S_N(A)(a_1 \dots a_N x)} \psi(a_1 \dots a_N x) d\nu(a_1) \dots d\nu(a_N) - \psi(z) \mathcal{L}_A^N(1)(x) \\ &= \int_{M^N} e^{S_N(A)(a_1 \dots a_N x)} \psi(a_1 \dots a_N x) d\nu(a_1) \dots d\nu(a_N) - \psi(z) \int_{M^N} e^{S_N(A)(a_1 \dots a_N x)} d\nu(a_1) \dots d\nu(a_N) \\ &= \int_{M^N} e^{S_N(A)(a_1 \dots a_N x)} (\psi(a_1 \dots a_N x) - \psi(z)) d\nu(a_1) \dots d\nu(a_N) \end{aligned}$$

Pela continuidade de ψ , temos que $\psi(a_1 \dots a_N x) = \psi(z)$, para qualquer escolha de $a_i \in M$. Como $\{a_1 \dots a_N x \mid a_i \in M, i \in \{1, \dots, n\}\}$ é ϵ -denso em Ω , segue que ψ é constante. Seja μ_A tal que $\mathcal{L}_A^*(\mu_A) = \mu_A$ (cuja existência foi garantida no item (ii)). Então, para todo $n_k \in \mathbb{N}$, vale que $(\mathcal{L}_A^*)^{n_k}(\mu_A) = \mu_A$. Com isso, temos que

$$\lim_{k \rightarrow +\infty} \int \mathcal{L}_A^{n_k}(\varphi) d\mu_A = \int \varphi d\mu_A,$$

para toda $\varphi \in C^\alpha(\Omega)$. Por outro lado, como ψ é contante, vale que

$$\lim_{k \rightarrow +\infty} \int \mathcal{L}_A^{n_k}(\varphi) d\mu_A = \int \lim_{k \rightarrow +\infty} \mathcal{L}_A^{n_k}(\varphi) d\mu_A = \int \psi d\mu_A = \psi = \lim_{k \rightarrow +\infty} \mathcal{L}_A^{n_k}(\varphi)$$

Portanto, $\lim_{k \rightarrow +\infty} \mathcal{L}_A^{n_k}(\varphi) = \int \varphi d\mu_A$, para toda $\varphi \in C^\alpha(\Omega)$. Como $C^\alpha(\Omega)$ é denso em $C(\Omega)$, podemos assumir que esta igualdade é válida para toda $\varphi \in C(\Omega)$. Mostremos agora que $\{\mathcal{L}_A^{n_k}(\varphi)\}$ é convergente. Suponha que outra subsequência $\{\mathcal{L}_A^{l_k}(\varphi)\}$ convirja uniformemente para ψ^* . Pelo mesmo raciocínio feito anteriormente, temos que ψ^* é constante. Logo, temos que

$$\lim_{k \rightarrow +\infty} \int \mathcal{L}_A^{l_k}(\varphi) d\mu_A = \int \varphi d\mu_A$$

e, por outro lado,

$$\lim_{k \rightarrow +\infty} \int \mathcal{L}_A^{l_k}(\varphi) d\mu_A = \int \lim_{k \rightarrow +\infty} \mathcal{L}_A^{l_k}(\varphi) d\mu_A = \int \psi^* d\mu_A = \psi^* = \lim_{k \rightarrow +\infty} \mathcal{L}_A^{l_k}(\varphi),$$

já que ψ^* é constante. Sendo assim, $\lim_{k \rightarrow +\infty} \mathcal{L}_A^{l_k}(\varphi) = \int \varphi d\mu_A$. Dessa forma, concluímos que toda subsequência de $\{\mathcal{L}_A^n\}$ tem uma subsequência convergente e estas convergem todas para o mesmo valor, $\int \varphi d\mu_A$, ou seja,

$$\lim_{n \rightarrow +\infty} \mathcal{L}_A^n(\varphi) = \int \varphi d\mu_A,$$

para toda $\varphi \in C(\Omega)$.

□

Observação 3.16. Se não tivéssemos assumido, no item (ii) do teorema anterior, que o operador \mathcal{L}_A estava normalizado, então teríamos garantido a existência de uma medida μ_A tal que $\mathcal{L}_A^*(\mu_A) = \lambda_A \mu_A$. Este fato está formalizado na próxima proposição.

Proposição 3.17. *Seja A potencial α -Hölder não necessariamente normalizado e sejam h_A e λ_A a autofunção e o autovalor associados, cuja existência é garantida pelo teorema 3.14. Seja \bar{A} o potencial normalizado associado e seja m a única probabilidade satisfazendo $\mathcal{L}_{\bar{A}}^*(m) = m$. Então a medida $\mu_A = m/h_A$ satisfaz $\mathcal{L}_A^*(\mu_A) = \lambda_A \mu_A$.*

Demonstração. Como $\mathcal{L}_{\bar{A}}^*(m) = m$, temos que para qualquer $\varphi \in C(\Omega)$ vale

$$\begin{aligned} \int_{\Omega} \varphi dm &= \int_{\Omega} \varphi d(\mathcal{L}_{\bar{A}}^*(m)) \\ &= \int_{\Omega} \mathcal{L}_{\bar{A}}(\varphi) dm \\ &= \int_{\Omega} \left(\int_M e^{\bar{A}(ax)} \varphi(ax) d\nu(a) \right) dm \\ &= \int_{\Omega} \left(\int_M e^{A(ax) + \log h_A(ax) - \log h_A \circ \sigma(ax) - \log \lambda_A} \varphi(ax) d\nu(a) \right) dm \\ &= \int_{\Omega} \left(\int_M e^{A(ax)} h_A(ax) \frac{1}{h_A(x)} \frac{1}{\lambda_A} \varphi(ax) d\nu(a) \right) dm \end{aligned}$$

Como a igualdade acima vale para qualquer $\varphi \in C(\Omega)$, substitua φ por

$\frac{\varphi}{h_A}$, de modo a obter

$$\begin{aligned}
 \int_{\Omega} \frac{\varphi}{h_A} dm &= \int_{\Omega} \left(\int_M e^{A(ax)} h_A(ax) \frac{1}{h_A(x)} \frac{1}{\lambda_A} \frac{\varphi(ax)}{h_A(ax)} d\nu(a) \right) dm \\
 &= \int_{\Omega} \left(\int_M e^{A(ax)} \frac{1}{h_A(x)} \frac{1}{\lambda_A} \varphi(ax) d\nu(a) \right) dm \\
 &= \int_{\Omega} \left(\int_M e^{A(ax)} \varphi(ax) d\nu(a) \right) \frac{1}{\lambda_A h_A(x)} dm \\
 &= \frac{1}{\lambda_A} \int_{\Omega} \mathcal{L}_A(\varphi) \frac{1}{h_A} dm \\
 &= \frac{1}{\lambda_A} \int_{\Omega} \mathcal{L}_A(\varphi) d\mu_A.
 \end{aligned}$$

Portanto, $\lambda_A \int_{\Omega} \varphi d\mu_A = \int_{\Omega} \frac{\varphi}{h_A} dm = \int_{\Omega} \mathcal{L}_A(\varphi) d\mu_A$, concluindo o resultado. \square

Operador de Ruelle para Cadeias de Markov a Tempo Contínuo

Neste capítulo, seguindo [1], nós estenderemos a noção de *Operador de Ruelle*, agora para *Cadeias de Markov a Tempo Contínuo*. E, dada uma perturbação, definiremos um *Operador de Ruelle modificado* e mostraremos a existência de uma auto-função e uma auto-medida associadas a este operador.

4.1 Definições e resultados básicos

Considere um conjunto de espaços de estados finito $S = \{1, \dots, n\}$ e seja Ω o conjunto de caminhos ω de \mathbb{R}^+ tomando valores em S , localmente constantes, contínuos à direita e com limite à esquerda. Em Ω , consideraremos a σ -álgebra \mathcal{B} gerada pelos cilindros da forma $\{w_0 = a_0, w_{t_1} = a_1, \dots, w_{t_r} = a_r\}$, onde $t_i \in \mathbb{R}$, com $0 < t_1 < \dots < t_r$, $r \in \mathbb{N}$ e $a_i \in S$.

Exemplo 4.1. Exemplos de caminhos, com $S = \{1, 2, 3, 4, 5\}$, estão ilustrados a seguir:

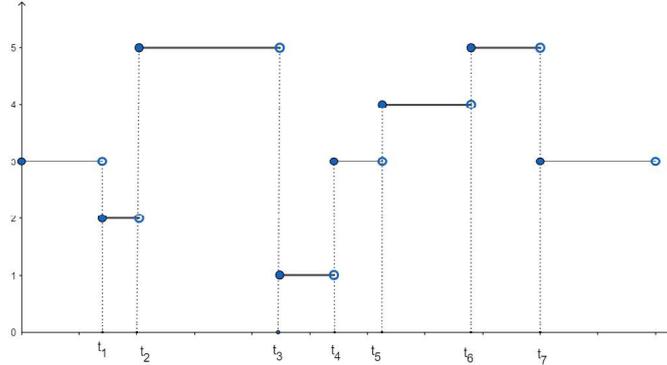


Fig. 4.1: Elemento de Ω com $S = \{1, 2, 3, 4, 5\}$.

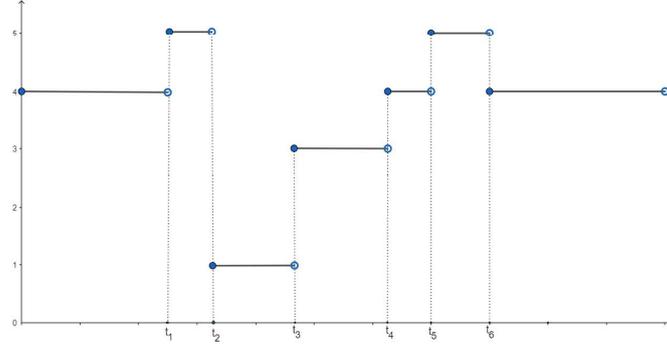


Fig. 4.2: Elemento de Ω com $S = \{1, 2, 3, 4, 5\}$.

A fim de definirmos uma nova versão do Operador de Ruelle, precisamos de uma medida a priori, que será construída da seguinte maneira: seja L uma matriz coluna soma zero, com elementos negativos em sua diagonal e não-negativos fora dela. Vimos que $P^t = e^{tL}$ é um semigrupo com gerador infinitesimal L e que e^{tL} é uma matriz coluna estocástica para todo $t \in \mathbb{R}^+$. Ainda, note que se $L(v) = 0$, então $e^{tL}(v) = v$, para todo $t \in \mathbb{R}^+$, ou seja, v é um autovetor de e^{tL} com autovalor associado 1. A partir deste semigrupo, construímos, em (Ω, \mathcal{B}) , a probabilidade \mathbb{P} definida nos cilindros da seguinte forma:

$$\mathbb{P}(\{w_0 = a_0, w_{t_1} = a_1, \dots, w_{t_r} = a_r\}) = P_{a_r a_{r-1}}^{t_r - t_{r-1}} \cdots P_{a_2 a_1}^{t_2 - t_1} P_{a_1 a_0}^{t_1} p_0^{a_0},$$

onde p_0 é o vetor de probabilidade tal que $L(p_0) = 0$.

Observação 4.2. Lembre que se Q é matriz coluna soma zero, sempre p vetor de probabilidade tal que $Q(p) = 0$. Além disso, nestas condições, $e^{tQ}(p) = p$, para todo $t \in \mathbb{R}$. (vide teorema 2.41 e teorema 2.42).

Também é necessário introduzirmos uma versão a tempo contínuo do *shift*, que é uma transformação contínua com respeito a métrica de Skorohod-Stone (para mais detalhes sobre esta métrica, o leitor pode consultar [2]). Para cada $s \in \mathbb{R}^+$ fixado, definimos a transformação \mathcal{B} -mensurável da seguinte maneira:

$$\begin{aligned}\Theta_s: \Omega &\rightarrow \Omega \\ \omega_t &\mapsto \omega_{t+s}.\end{aligned}$$

Exemplo 4.3. Ação do *shift* em um caminho de Ω :

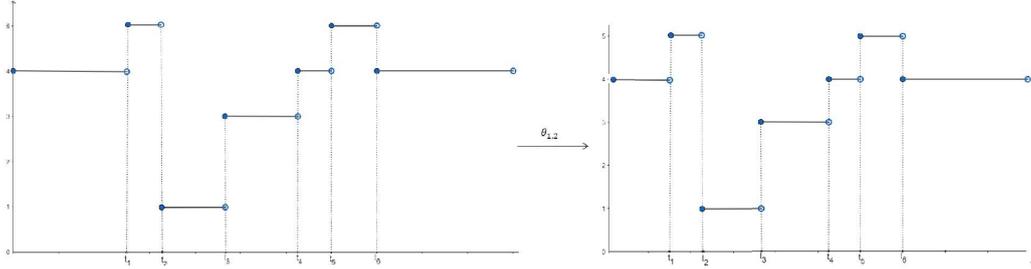


Fig. 4.3: Exemplo da ação do *shift* para $s = 1, 2$

Observação 4.4. A probabilidade \mathbb{P} definida anteriormente é estacionária com respeito ao *shift*, isto é, para qualquer função integrável $f: \Omega \rightarrow \mathbb{R}^+ \in L^\infty(\mathbb{P})$ e qualquer $s > 0$ vale que

$$\int f(w) d\mathbb{P}(w) = \int (f \circ \Theta_s)(w) d\mathbb{P}(w).$$

4.2 Operador de Ruelle a Tempo Contínuo e Propriedades

Consideraremos, a partir de agora, uma família de medidas μ_t^z , indexada pelos elementos $z \in \Omega$ e tempos $t > 0$, que funcionará como uma desintegração da probabilidade \mathbb{P} anteriormente definida. Seja $t > 0$ e considere uma seqüência $0 = t_0 < t_1 < \dots < t_{j-1} < t \leq t_j < \dots < t_r$. Então, para $z \in \Omega$, definimos a medida μ_t^z nos cilindros por:

$$\mu_t^z([X_0 = a_0, \dots, X_{t_r} = a_r]) = \begin{cases} \frac{1}{p_0(z(t))} P_{z(t)a_{j-1}}^{t-t_{j-1}} \dots P_{a_2 a_1}^{t_2-t_1} P_{a_1 a_0}^{t_1} p_0(a_0), & \text{se } a_j = z(t_j), \dots, a_r = z(t_r) \\ 0, & \text{caso contrário.} \end{cases}$$

Com estas ferramentas em mãos, podemos definir o Operador de Ruelle a tempo contínuo da seguinte maneira: fixado $t > 0$, pomos

$$\mathcal{L}^t(f)(z) = \int_{y \in D[0,t]} f(yz) d\mu_t^z(yz),$$

onde $D[0, t) = \{y : [0, t) \rightarrow S : y \text{ é contínua à direita e com limite à esquerda}\}$ e yz representa a concatenação dos caminhos y e z .

Observação 4.5. $\lim_{s \rightarrow t} y(s)$ não precisa ser igual a $z(0)$.

Exemplo 4.6. Para ilustrar a ação do operador de Ruelle acima definido, considere uma sequência $0 = t_0 < t_1 < \dots < t_{j-1} < t \leq t_j < \dots < t_r$ e seja $f = \mathbb{1}_{[X_0=a_0, X_{t_1}=a_1, \dots, X_{t_r}=a_r]}$. Dividiremos em dois casos:

- (1) $z \in \Omega$ tal que $z_{t_j-t} = a_j, \dots, z_{t_r-t} = a_r$, ou seja, z é tal que se $y \in D[0, t)$, então $(yz)_{t_i} = z_{t_i-t} = a_i$, para $i \in \{j, \dots, r\}$. Então

$$\begin{aligned} \mathcal{L}^t(f)(z) &= \int_{y \in D[0, t)} f(yz) d\mu_t^z(yz) \\ &= \int_{y \in D[0, t)} \mathbb{1}_{[X_0=a_0, X_{t_1}=a_1, \dots, X_{t_r}=a_r]}(yz) d\mu_t^z(yz) \\ &= \frac{1}{p_0^{z_0}} P_{z(0)a_{j-1}}^{t-t_{j-1}} \cdots P_{a_2 a_1}^{t_2-t_1} P_{a_1 a_0}^{t_1} p_0^{a_0} \end{aligned}$$

- (2) $z \in \Omega$ não é tal que $z_{t_j-t} = a_j, \dots, z_{t_r-t} = a_r$.

Neste caso, $\mathcal{L}^t(f)(z) = 0$

Portanto, $\mathcal{L}^t(f)(z) = \frac{1}{p_0^{z_0}} P_{z(0)a_{j-1}}^{t-t_{j-1}} \cdots P_{a_2 a_1}^{t_2-t_1} P_{a_1 a_0}^{t_1} p_0^{a_0} \mathbb{1}_{[z_{t_j-t}=a_j, \dots, z_{t_r-t}=a_r]}$.

A próxima proposição afirma que o Operador de Ruelle a tempo contínuo definido está normalizado.

Proposição 4.7. $\mathcal{L}^t(1) = 1$.

Demonstração. Temos

$$\begin{aligned} \mathcal{L}^t(1)(z) &= \int_{y \in D[0, t)} 1(yz) d\mu_t^z(yz) \\ &= \int_{y \in D[0, t)} d\mu_t^z(yz) \end{aligned}$$

$$\begin{aligned}
&= \mu_t^z(\{X_t = z(0)\}) \\
&= \sum_{j=1}^n \mu_t^z(\{X_0 = j, X_t = z(0)\}) \\
&= \frac{1}{p_0^{z(0)}} \sum_{j=0}^n P_{z(0)j}^t p_0^j \\
&= 1
\end{aligned}$$

□

Podemos definir o operador dual de \mathcal{L}^t , que denotaremos por $(\mathcal{L}^t)^*$, agindo nas medidas. Este operador é tal que $\int f d(\mathcal{L}^t)^*(\mu) = \int \mathcal{L}^t(f) d\mu$. A principal consequência da proposição acima é que a imagem de uma probabilidade pelo operador dual é também uma probabilidade. Este resultado está formalizado abaixo.

Proposição 4.8. *Se μ é probabilidade, $(\mathcal{L}^t)^*(\mu)$ é probabilidade.*

Demonstração. Devemos mostrar que $(\mathcal{L}^t)^*(\mu)(\Omega) = 1$.

$$\begin{aligned}
(\mathcal{L}^t)^*(\mu)(\Omega) &= \int 1 d(\mathcal{L}^t)^*(\mu) \\
&= \int \mathcal{L}^t(1) d\mu \\
&= \int 1 d\mu \\
&= \mu(\Omega) \\
&= 1.
\end{aligned}$$

□

A próxima proposição traz um resultado importante: de que \mathbb{P} é um ponto fixo para este operador.

Proposição 4.9. $(\mathcal{L}^t)^*(\mathbb{P}) = \mathbb{P}$

Demonstração. Devemos mostrar que, dado E conjunto \mathcal{B} -mensurável, vale a igualdade $(\mathcal{L}^t)^*(\mathbb{P})(E) = \mathbb{P}(E)$, ou seja, $\int_E d(\mathcal{L}^t)^*\mathbb{P}(z) = \int_E d\mathbb{P}(z)$. Como os cilindros geram a σ -álgebra \mathcal{B} , basta mostrarmos a igualdade para conjuntos

da forma $E = [X_0 = a_0, \dots, X_{t_r} = a_r]$, com $0 = t_0 < t_1 < \dots < t_{j-1} < t \leq t_j < \dots < t_r$. Então

$$\begin{aligned}
\int_E d(\mathcal{L}^t)^* \mathbb{P}(z) &= \int \mathbf{1}_E d(\mathcal{L}^t)^* \mathbb{P}(z) \\
&= \int \mathcal{L}^t(\mathbf{1}_E)(z) d\mathbb{P}(z) \\
&= \sum_{b=1}^n \int_{[X_0=b]} \mathcal{L}^t(\mathbf{1}_E)(z) d\mathbb{P}(z) \\
&= \sum_{b=1}^n \int \mathbf{1}_{[X_0=b]} \mathbf{1}_{[X_{t_j-t}=a_j, \dots, X_{t_r-t}=a_r]} \frac{1}{p_0^{yz(t)}} P_{z(0)a_{j-1}}^{t-t_{j-1}} \dots P_{a_{2a_1}}^{t_2-t_1} P_{a_1 a_0}^{t_1} p_0^{a_0} d\mathbb{P}(z) \\
&= \sum_{b=1}^n \int \mathbf{1}_{[X_0=b, X_{t_j-t}=a_j, \dots, X_{t_r-t}=a_r]} \frac{1}{p_0^b} P_{z(0)a_{j-1}}^{t-t_{j-1}} \dots P_{a_{2a_1}}^{t_2-t_1} P_{a_1 a_0}^{t_1} p_0^{a_0} d\mathbb{P}(z)
\end{aligned}$$

onde a penúltima igualdade segue do exemplo 4.6. Aplicando a definição de \mathbb{P} e Chapman-Kolmogorov, obtemos que

$$\begin{aligned}
\int_E d(\mathcal{L}^t)^* \mathbb{P}(z) &= \sum_{b=1}^n P_{a_r a_{r-1}}^{t_r-t-(t_{r-1}-t)} \dots P_{a_j b}^{t_j-t} p_0^b \frac{1}{p_0^b} P_{z(0)a_{j-1}}^{t-t_{j-1}} \dots P_{a_{2a_1}}^{t_2-t_1} P_{a_1 a_0}^{t_1} p_0^{a_0} \\
&= P_{a_r a_{r-1}}^{t_r-t_{r-1}} \dots P_{a_j a_{j-1}}^{t_j-t_{j-1}} \dots P_{a_1 a_0}^{t_1} p_0^{a_0} \\
&= \mathbb{P}[X_0 = a_0, \dots, X_{t_r} = a_r] \\
&= \int \mathbf{1}_{[X_0=a_0, \dots, X_{t_r}=a_r]} d\mathbb{P} \\
&= \int_E d\mathbb{P},
\end{aligned}$$

concluindo assim o resultado. \square

A próxima proposição é especialmente útil no caso em que $f = 1$. Ela também será utilizada na proposição 4.11, que trata do dual do operador de Ruelle.

Proposição 4.10. *Dados $t \in \mathbb{R}^+$ e funções $f, g \in L^\infty(\mathbb{P})$, temos*

$$\mathcal{L}^t(f(g \circ \Theta_t))(z) = g(z) \mathcal{L}^t(f)(z)$$

Demonstração.

$$\begin{aligned}
\mathcal{L}^t(f(g \circ \Theta_t))(z) &= \int_{y \in D[0,t]} f(yz)(g \circ \Theta_t)(yz) d\mu_t^z(yz) \\
&= \int_{y \in D[0,t]} f(yz)g(z) d\mu_t^z(yz) \\
&= g(z) \int_{y \in D[0,t]} f(yz) d\mu_t^z(yz) \\
&= g(z) \mathcal{L}^t(f)(z).
\end{aligned}$$

□

Considere o endomorfismo $\alpha_t : L^\infty(\mathbb{P}) \rightarrow L^\infty(\mathbb{P})$, $\alpha_t(f) = f \circ \Theta_t$. Então podemos reescrever a proposição acima como $\mathcal{L}^t(f\alpha_t(g)) = g\mathcal{L}^t(f)$.

Proposição 4.11. α_t é o dual de \mathcal{L}^t em $\mathcal{L}^2(\mathbb{P})$, ou seja, $\langle \mathcal{L}^t(f), g \rangle = \langle f, \alpha_t(g) \rangle$.

Demonstração.

$$\begin{aligned}
\langle \mathcal{L}^t(f), g \rangle &= \int \mathcal{L}^t(f)g d\mathbb{P} \\
&= \int \mathcal{L}^t(f\alpha_t(g)) d\mathbb{P} \\
&= \int f\alpha_t(g) d\mathbb{P} \\
&= \langle f, \alpha_t(g) \rangle
\end{aligned}$$

onde na terceira igualdade utilizamos o fato de que \mathbb{P} é ponto fixo para o operador $(\mathcal{L}^t)^*$, ou seja, a proposição 4.9. □

4.3 Operador de Ruelle Modificado

Consideraremos, a partir desta seção, funções $V : \Omega \rightarrow \mathbb{R}$ constantes nos cilindros da forma $\{X_0 = c\}$. Analisaremos a perturbação do Operador de Ruelle \mathcal{L}^t por V . Esta perturbação se dá da seguinte maneira: consideremos a aplicação $G_t : \Omega \rightarrow \mathbb{R}$ dada por $G_t(x) = e^{\int_0^t V(x(s)) ds}$. Então, a partir desta aplicação, definimos o Operador de Ruelle modificado $\mathcal{L}_V^t : L^\infty(\mathbb{P}) \rightarrow L^\infty(\mathbb{P})$

por:

$$\begin{aligned}\mathcal{L}_V^t(f)(z) &= \mathcal{L}^t(G_t f)(z) \\ &= \mathcal{L}^t\left(e^{\int_0^t (V \circ \Theta_s)(\cdot) ds} f\right)(z) \\ &= \sum_{b=1}^n \mathcal{L}^t\left(e^{\int_0^t (V \circ \Theta_s)(\cdot) ds} \mathbf{1}_{\{X_t=b\}} f\right)(z).\end{aligned}$$

O próximo Teorema é uma versão do Teorema de Perron-Fröbenius para tempo contínuo. Devido às suas consequências, este resultado é de suma importância para que possamos encontrar uma auto-função para o Operador de Ruelle modificado e uma auto-medida para seu dual.

Teorema 4.12. *Dados L matriz $n \times n$ coluna soma zero, p_0 vetor de probabilidade tal que $L(p_0) = p_0$ e $V: \Omega \rightarrow \mathbb{R}$ tal que V é constante em $\{X_0 = c\}$, então existem:*

a) *uma única função positiva $u_V: \Omega \rightarrow \mathbb{R}$ constante e igual aos valores u_V^i em cada cilindro $\{X_0 = i\}$, $i \in \{1, \dots, n\}$;*

b) *um único vetor de probabilidade μ_V em \mathbb{R}^n tal que*

b1) $\mu_V(i) > 0$;

b2) $\sum_{i=1}^n (u_V)_i (\mu_V)_i = 1$;

c) *um número real λ_V tal que:*

c1) *para todo $s > 0$*

$$e^{-s\lambda_V} u_V e^{s(L+V)} = u_V;$$

c2) *dado $s > 0$,*

$$e^{-s\lambda_V} e^{s(L+V)} \mu_V = \mu_V.$$

A demonstração deste teorema baseia-se no teorema de Perron-Fröbenius para matrizes positivas (ver [11]) e no seguinte fato elementar:

$$Av = \lambda v \iff e^{tA} v = e^{t\lambda} v \quad \forall t \in \mathbb{R}.$$

O Teorema de Perron-Fröbenius para matrizes positivas pode ser aplicado neste caso, pois dado $L + V$ existe $k > 0$ suficientemente grande de modo que $L + V + kI$ terá todas as entradas positivas. A soma desta constante não prejudica o argumento.

Algumas consequências deste teorema são:

(i) De c1) obtemos que $u_V e^{s(L+V)} = e^{s\lambda_V} u_V$ e, portanto, $u_V(L+V) = \lambda_V u_V$. Ou seja,

$$u_V(L+V - \lambda_V I) = 0.$$

(ii) De c2), obtemos

$$(L+V - \lambda_V I)\mu_V = 0.$$

Agora, seja a σ -álgebra $\mathcal{F}_t^- = \sigma(\{X_s | s \leq t\})$. Em \mathcal{F}_t^- , para cada $i_0 \in S$ e $t \in \mathbb{R}^+$ fixados, consideremos a probabilidade $\mu_{i_0}^t$ com suporte em $\{X_0 = i_0\}$ definida nos cilindros da seguinte forma:

$$\mu_{i_0}^t(\{X_0 = i_0, X_{t_1} = a_1, \dots, X_{t_r} = a_r, X_t = j_0\}) = P_{j_0 a_r}^{t-t_r} \dots P_{a_2 a_1}^{t_2-t_1} P_{a_1 i_0}^{t_1}.$$

Observação 4.13. Esta probabilidade não é estacionária.

Seja $Q(j, i)_t$ a entrada ij da matriz $e^{t(L+V)}$, isto é, $Q(j, i)_t = (e^{t(L+V)})_{ij}$. Então, por Feynman-Kac, obtemos que

$$\begin{aligned} Q(j_0, i_0)_t &= \mathbb{E}_{\{X_0=i_0\}} \left(e^{\int_0^t (V \circ \Theta_s)(w) ds}, X_t = j_0 \right) \\ &= \int \mathbb{1}_{\{X_t=j_0\}} e^{\int_0^t (V \circ \Theta_s)(w) ds} d\mu_{i_0}^t(w) \end{aligned}$$

Para maiores detalhes, o leitor pode consultar [9], Lemma 5.5

Proposição 4.14. A função $f_V(w) = \frac{\mu_V(w)}{p^0(w)} = \frac{(\mu_V)_{w(0)}}{(p^0)_{w(0)}}$ é uma auto-função para \mathcal{L}_V^t com autovalor $e^{t\lambda_V}$.

Demonstração. Primeiramente, note que podemos escrever

$$\mathcal{L}_V^t \left(\frac{\mu_V}{p^0} \right) (w) = \sum_{i=1}^n \sum_{j=1}^n \mathcal{L}_V^t \left(\frac{\mu_V(i)}{p_0(i)} \mathbb{1}_{\{X_0=i\}} \mathbb{1}_{\{X_t=j\}} \right).$$

Seja $w(0) = j_0$. Fixado $i \in S$, temos dois casos:

(i) Caso $j = j_0$.

$$\begin{aligned} \mathcal{L}_V^t (\mathbb{1}_{\{X_0=i\}} \mathbb{1}_{\{X_t=j_0\}}) (w) &= \mathcal{L}^t \left(e^{\int_0^t (V \circ \Theta_s)(\cdot) ds} \mathbb{1}_{\{X_0=i\}} \mathbb{1}_{\{X_t=j_0\}} \right) (w) \\ &= \int_{y \in D[0,t]} e^{\int_0^t (V \circ \Theta_s)(y) ds} \mathbb{1}_{\{X_0=i\}} \mathbb{1}_{\{X_t=j_0\}}(yw) d\mu_t^w(yw) \\ &= \frac{Q(j_0, i)_t p_i^0}{p_{j_0}^0} \end{aligned}$$

(ii) Caso $j \neq j_0$.

$$\mathcal{L}_V^t \left(\mathbb{1}_{\{X_0=i\}} \mathbb{1}_{\{X_t=j\}} \right) (w) = 0$$

Dessa forma,

$$\begin{aligned} \mathcal{L}_V^t \left(\frac{\mu_V}{p^0} \right) (w) &= \sum_{i=1}^n \sum_{j=1}^n \mathcal{L}_V^t \left(\frac{\mu_V(i)}{p_0(i)} \mathbb{1}_{\{X_0=i\}} \mathbb{1}_{\{X_t=j\}} \right) \\ &= \sum_{i=1}^n \frac{\mu_V(i)}{p_0(i)} Q(j_0, i)_t \frac{p_0(c)}{p_0(j_0)} \\ &= \sum_{i=1}^n \frac{\mu_V(i)}{p_0(i)} \left(e^{t(L+V)} \right)_{j_0 i} \frac{p_0(c)}{p_0(j_0)} \\ &= e^{t\lambda_V} \frac{(\mu_V)_{j_0}}{p_0(j_0)} \\ &= e^{t\lambda_V} \left(\frac{\mu_V}{p_0} \right) (w). \end{aligned}$$

□

A partir de agora, consideraremos um novo operador, $\hat{\mathcal{L}}_V^t$, definido da seguinte forma: para cada $t \in \mathbb{R}^+$, $\hat{\mathcal{L}}_V^t: L^\infty(\mathbb{P}) \rightarrow L^\infty(\mathbb{P})$ é o operador dado por

$$\hat{\mathcal{L}}_V^t(g)(w) = \frac{1}{f_V} \mathcal{L}^t \left(e^{\int_0^t (V - \lambda_V) \circ \Theta_s(\cdot) ds} g f_V \right) (w)$$

Observação 4.15. Podemos reescrever este operador como

$$\hat{\mathcal{L}}_V^t(g)(w) = \frac{1}{f_V} \frac{1}{e^{\lambda_V t}} \mathcal{L}_V^t(g f_V)(w).$$

De fato,

$$\begin{aligned} \hat{\mathcal{L}}_V^t(g)(w) &= \frac{1}{f_V} \mathcal{L}^t \left(e^{\int_0^t (V - \lambda_V) \circ \Theta_s(\cdot) ds} g f_V \right) (w) \\ &= \frac{1}{f_V} \mathcal{L}^t \left(e^{\int_0^t V \circ \Theta_s(\cdot) ds - \int_0^t \lambda_V ds} g f_V \right) (w) \\ &= \frac{1}{f_V} \mathcal{L}^t \left(\frac{e^{\int_0^t V \circ \Theta_s(\cdot) ds}}{e^{t\lambda_V}} g f_V \right) (w) \\ &= \frac{1}{f_V} \frac{1}{e^{\lambda_V t}} \mathcal{L}^t \left(e^{\int_0^t V \circ \Theta_s(\cdot) ds} g f_V \right) (w) \\ &= \frac{1}{f_V} \frac{1}{e^{\lambda_V t}} \mathcal{L}_V^t(g f_V)(w). \end{aligned}$$

A próxima proposição mostra que $\hat{\mathcal{L}}_V^t$ está normalizado.

Proposição 4.16. $\hat{\mathcal{L}}_V^t(1) = 1$, para todo $t \in \mathbb{R}^+$.

Demonstração. De fato,

$$\begin{aligned}
\hat{\mathcal{L}}_V^t(1)(z) &= \left[\frac{p_o}{\mu_V} \mathcal{L}^t \left(\frac{e^{\int_0^t (V - \lambda_V) \circ \Theta_s(\cdot) ds} \mu_V}{p_0} \right) \right] (z) \\
&= \left[\frac{p_o}{\mu_V} \mathcal{L}^t \left(\frac{e^{\int_0^t V \circ \Theta_s(\cdot) ds} \mu_V}{e^{\int_0^t \lambda_V ds} p_0} \right) \right] (z) \\
&= \left[\frac{p_o}{\mu_V} \mathcal{L}^t \left(\frac{e^{\int_0^t V \circ \Theta_s(\cdot) ds} \mu_V}{e^{\lambda_V t} p_0} \right) \right] (z) \\
&= \left[\frac{p_o}{\mu_V} \frac{1}{e^{\lambda_V t}} \mathcal{L}^t \left(\frac{e^{\int_0^t V \circ \Theta_s(\cdot) ds} \mu_V}{p_0} \right) \right] (z) \\
&= \left[\frac{p_o}{\mu_V} \frac{1}{e^{\lambda_V t}} \mathcal{L}_V^t \left(\frac{\mu_V}{p_0} \right) \right] (z) \\
&= \left[\frac{p_o}{\mu_V} \frac{1}{e^{\lambda_V t}} \mathcal{L}_V^t (f_V) \right] (z) \\
&= \left[\frac{p_o}{\mu_V} \frac{1}{e^{\lambda_V t}} e^{\lambda_V t} \frac{\mu_V}{p_0} \right] (z) \\
&= 1,
\end{aligned}$$

para todo $z \in \Omega$. □

Além disso, definimos, para cada $t \in \mathbb{R}^+$ fixado, o operador dual de $(\hat{\mathcal{L}}_V^t)^*$, que age no conjunto das probabilidades μ em Ω . Analogamente a proposição 4.8, temos que se μ é probabilidade, então $\nu = (\hat{\mathcal{L}}_V^t)^*(\mu)$ também é probabilidade, já que $\hat{\mathcal{L}}_V^t$ está normalizado.

Exemplo 4.17. Sejam $0 < t_1 < t_2 < t$ e $g = \mathbb{1}_{[X_0=a_0, X_{t_1}=a_1, X_{t_2}=a_2, X_t=a_3]}$. Então

$$\begin{aligned}
\hat{\mathcal{L}}_V^t(g)(w) &= \frac{1}{f_V} \frac{1}{e^{\lambda_V t}} \mathcal{L}_V^t(g f_V)(w) \\
&= \frac{1}{f_V} \frac{1}{e^{\lambda_V t}} \int_{y \in D[0, t]} e^{\int_0^t (V \circ \Theta_s)(yx)} \mathbb{1}_{[X_0=a_0, X_{t_1}=a_1, X_{t_2}=a_2, X_t=a_3]} \frac{\mu_V(y\omega(0))}{p_0(y\omega(0))} d\mu_x^t
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{f_V} \frac{1}{e^{\lambda_V t}} \int_{y_3 \in D[t_2, t]} \int_{y_2 \in D[t_1, t_2]} \int_{y_1 \in D[0, t_1]} e^{\int_0^{t_1} (V \circ \Theta_s)(\cdot)} e^{\int_{t_1}^{t_2} (V \circ \Theta_s)(\cdot)} e^{\int_{t_2}^t (V \circ \Theta_s)(\cdot)} \\
&\quad \mathbb{1}_{[X_0=a_0, X_{t_1}=a_1, X_{t_2}=a_2, X_t=a_3]} \frac{\mu_V(a_0) p_0(a_0)}{p_0(a_0) p_0(a_1)} d\mu_{a_0}^{t_1}(y_1 x) \frac{p_0(a_1)}{p_0(a_2)} d\mu_{a_1}^{t_2-t_1}(y_2 y_1 x) \\
&\quad \frac{p_0(a_2)}{p_0(a_3)} d\mu_{a_3}^{t-t_2}(y_3 y_2 y_1 x) \\
&= \frac{1}{f_V} \frac{1}{e^{\lambda_V t}} \int_{y_3 \in D[t_2, t]} \int_{y_2 \in D[t_1, t_2]} \int_{y_1 \in D[0, t_1]} e^{\int_0^{t_1} (V \circ \Theta_s)(\cdot)} e^{\int_{t_1}^{t_2} (V \circ \Theta_s)(\cdot)} e^{\int_{t_2}^t (V \circ \Theta_s)(\cdot)} \\
&\quad \mathbb{1}_{[X_0=a_0, X_{t_1}=a_1]} \mathbb{1}_{[X_{t_1}=a_1, X_{t_2}=a_2]} \mathbb{1}_{[X_{t_2}=a_2, X_t=a_3]} \frac{\mu_V(a_0) p_0(a_0)}{p_0(a_0) p_0(a_1)} d\mu_{a_0}^{t_1}(y_1 x) \\
&\quad \frac{p_0(a_1)}{p_0(a_2)} d\mu_{a_1}^{t_2-t_1}(y_2 y_1 x) \frac{p_0(a_2)}{p_0(a_3)} d\mu_{a_3}^{t-t_2}(y_3 y_2 y_1 x) \\
&= \frac{1}{f_V} \frac{1}{e^{\lambda_V t}} Q(a_1, a_0)_{t_1} Q(a_2, a_1)_{t_2-t_1} Q(a_3, a_2)_{t-t_2},
\end{aligned}$$

se w é tal que $w_0 = a_3$ e $\hat{\mathcal{L}}_V^t(g)(w) = 0$, caso contrário.

O próximo teorema garante a existência de uma probabilidade ν_V tal que $\int \mathcal{L}_V^t d\nu_V = \int g d\nu_V$, para toda $g \in L^\infty(\mathbb{P})$, ou seja, ν_V é ponto fixo para o operador $(\mathcal{L}_V^t)^*$.

Teorema 4.18. *Para cada $t \in \mathbb{R}^+$ existe uma probabilidade ν_V em (Ω, \mathcal{B}) que é ponto fixo para a transformação $(\hat{\mathcal{L}}_V^t)^*$. Ainda, esta probabilidade não depende de t .*

Consideremos ν_V a probabilidade obtida da seguinte maneira: dados $0 < t_1 < t_2 < \dots < t_s < \dots < t_r$ e $g = \mathbb{1}_{\{X_0=a_0, X_{t_1}=a_1, \dots, X_{t_r}=a_r\}}$, definimos a integral de g com respeito a probabilidade ν_V como

$$\begin{aligned}
\int g(w) d\nu(w) &= e^{(t_r-t_{r-1})(L+V+\lambda_V I)_{a_r a_{r-1}}} \dots e^{(t_1)(L+V+\lambda_V I)_{a_1 a_0}} \mu_V(a_0) \\
&= Q(a_{r-1}, a_r)_{t_r-t_{r-1}} \dots Q(a_0, a_1)_{t_1} \mu_V(a_0).
\end{aligned}$$

Vejamos que ν é, de fato, uma probabilidade.

$$\begin{aligned}
\nu(\Omega) &= \int 1 d\nu \\
&= \sum_{i=1}^n \sum_{j=1}^n \int \mathbb{1}_{\{X_0=j, X_t=i\}} d\nu
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n \sum_{j=1}^n e^{t(L+V-\lambda_V I)} \mu_V(j) \\
&= \sum_{i=1}^n \mu_V(i) \\
&= 1
\end{aligned}$$

Com isto, construímos a probabilidade ν_V nos cilindros. Pelo *Teorema da Extensão de Kolmogorov*, podemos estender a probabilidade ν_V a σ -álgebra \mathcal{B} .

Agora, vejamos a ação do operador $\hat{\mathcal{L}}_V^t$ na função g supracitada, com t é tal que $0 < t_1 < \dots < t_{s-1} < t \leq t_s < \dots < t_r$. Temos

$$\hat{\mathcal{L}}_V^t(g)(w) = \frac{1}{\mu_V(i)} e^{(t-t_{s-1})(L+V-\lambda_V I)} \dots e^{t_1(L+V-\lambda_V I)} \mu_V(a_0)$$

se $w(0) = i, w(t_s - t) = a_s, \dots, w(t_r - t) = a_r$ e $\hat{\mathcal{L}}_V^t(g)(w) = 0$, caso contrário. Para concluirmos o resultado, devemos mostrar que

$$\int \hat{\mathcal{L}}_V^t(g)(w) d\nu_V = \int g d\nu_V.$$

De fato,

$$\begin{aligned}
\int \hat{\mathcal{L}}_V^t(g)(w) d\nu_V &= \int \mathbb{1}_{\{X_{t_s-t}=a_s, \dots, X_{t_r-t}=a_r\}} \hat{\mathcal{L}}_V^t(g)(w) d\nu_V \\
&= \int \sum_{i=1}^n \mathbb{1}_{\{X_0=i, X_{t_s-t}=a_s, \dots, X_{t_r-t}=a_r\}} \hat{\mathcal{L}}_V^t(g)(w) d\nu_V,
\end{aligned}$$

onde na primeira igualdade usamos o fato que $\hat{\mathcal{L}}_V^t(g)(w) \neq 0$ apenas quando $w(0) = i, w(t_s - t) = a_s, \dots, w(t_r - t) = a_r$.

Pelos cálculos anteriores, obtemos

$$\begin{aligned}
\int \hat{\mathcal{L}}_V^t(g)(w) d\nu_V &= \int \sum_{i=1}^n \mathbb{1}_{\{X_{t_s-t}=a_s, \dots, X_{t_r-t}=a_r\}} \frac{1}{\mu_V(i)} e^{(t-t_{s-1})(L+V-\lambda_V I)} \dots \\
&\quad \dots e^{t_1(L+V-\lambda_V I)} \mu_V(a_0) d\nu_V
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n \nu(\{X_0 = i, X_{t_s-t} = a_s, \dots, X_{t_r-t} = a_r\}) \frac{1}{\mu_V(i)} e^{(t-t_{s-1})(L+V-\lambda_V I)} \dots \\
&\quad \dots e^{t_1(L+V-\lambda_V I)} \mu_V(a_0) \\
&= \sum_{i=1}^n e^{(t_r-t-(t_{r-1}-t))(L+V-\lambda_V I)} \dots e^{((t_{s+1}-t)-(t_s-t))(L+V-\lambda_V I)} e^{(t_s-t)(L+V-\lambda_V I)} \mu_V(i) \\
&\quad \frac{1}{\mu_V(i)} e^{(t-t_{s-1})(L+V-\lambda_V I)} \dots e^{t_1(L+V-\lambda_V I)} \mu_V(a_0) \\
&= \sum_{i=1}^n e^{(t_r-t_{r-1})(L+V-\lambda_V I)} \dots e^{(t_s-t)(L+V-\lambda_V I)} e^{(t-t_{s-1})(L+V-\lambda_V I)} \dots \\
&\quad \dots e^{t_1(L+V-\lambda_V I)} \mu_V(a_0)
\end{aligned}$$

Utilizando o fato que $\sum_{i=1}^n e^{(t_s-t)(L+V-\lambda_V I)} e^{(t-t_{s-1})(L+V-\lambda_V I)} = e^{(t-t_{s-1})(L+V-\lambda_V I)}$, chegamos a

$$\begin{aligned}
\int \hat{\mathcal{L}}_V^t(g)(w) d\nu_V &= e^{(t_r-t_{r-1})(L+V-\lambda_V I)} \dots e^{(t-t_{s-1})(L+V-\lambda_V I)} \dots e^{t_1(L+V-\lambda_V I)} \mu_V(a_0) \\
&= \int \mathbf{1}_{\{X_0=a_0, \dots, X_{t_r}=a_r\}} d\nu_V \\
&= \int g d\nu_V,
\end{aligned}$$

concluindo o resultado.

Referências Bibliográficas

- [1] A. BARAVIERA, A. O. LOPES, R. EXEL, *A Ruelle Operator for continuous time Markov Chains*, The São Paulo J. of Math. Sci. 4(1), 1-16(2010).
- [2] A. NEUMANN, *A métrica de Skorohod*, Dissertação de Mestrado, Universidade Federal do Rio Grande do Sul, 2007.
- [3] A. O. LOPES, A. NEUMANN, *Large Deviations for stationary probabilities of a family of continuous time Markov chains via Aubry-Mather theory*, Journ. of Statistical Physics. Vol. 159 - Issue 4 pp 797-822 (2015).
- [4] A. O. LOPES, A. NEUMANN, PH. THIEULLEN, *A Thermodynamic Formalism for continuous time Markov chains with values on the Bernoulli Space: entropy, pressure and large deviations*, Journ. of Statist. Phys. Volume 152, Issue 5, Page 894-933, (2013).
- [5] A. O. LOPES, J. K. MENGUE, J. MOHR, R. R. SOUZA, *Entropy and Variational Principle for one-dimensional lattice systems with a general a priori probability: positive and zero temperature*, Ergodic Theory and Dynamical Systems, 35, pp 1925-1961(2015).
- [6] A. O. LOPES, S. R. C. LOPES, *Introdução aos Processos Estocásticos para estudantes de Matemática*, Notas de Aula, 2016.
- [7] A. O. LOPES, *Thermodynamic Formalism, Maximizing Probabilities and Large Deviations*, Notas de Aula, 2017.
- [8] B. R. JAMES, *Probabilidade: um curso em nível intermediário*, Sociedade Brasileira de Matemática - IMPA, 2015.

-
- [9] D. W. STROOCK, *An Introduction to the Theory of Large Deviations*, Springer, 1984.
- [10] E. A. SILVA, *O Operador de Ruelle em Espaços de Estados Compactos*, Tese de Doutorado, Universidade de Brasília, 2016.
- [11] E. SENETA, *Non-negative matrices and Markov chains*, Springer, 1981.
- [12] K. OLIVEIRA E M. VIANA, *Fundamentos da Teoria Ergódica*, Sociedade Brasileira de Matemática - IMPA, Rio de Janeiro, 2014.
- [13] J. R. MUNKRES, *Topology*, Prentice-Hall, 2000.
- [14] J. R. NORRIS, *Markov Chains*, University of Cambridge, 1997.
- [15] M. POLLICOTT E W. PARRY, *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics*, Astérisque, 1990.
- [16] R. DURRET, *Probability: Theory and Examples*, Cambridge University Press, 2010.
- [17] R. G. BARTLE, *The Elements of Integration and Lebesgue Measure*, John Wiley & Sons, 1995.
- [18] S. KARLIN, H. TAYLOR, *A First Course in Stochastic Processes*, Academic Press, 1975.
- [19] S. KARLIN, H. TAYLOR, *A Second Course in Stochastic Processes*, Academic Press, 1975.