



Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Departamento de Estatística
Programa de Pós-graduação em Estatística

LEON TARQUINO DA COSTA

DINÂMICA DE EVOLUÇÃO EM PROCESSOS COM OPERADORES DE SUBSTITUIÇÃO

Recife
2020

LEON TARQUINO DA COSTA

DINÂMICA DE EVOLUÇÃO EM PROCESSOS COM OPERADORES DE SUBSTITUIÇÃO

Tese apresentada ao programa de Pós-graduação em Estatística da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de Doutor em Estatística.

Área de Concentração: Probabilidade.

Orientador: **Prof. Dr. Alex Dias Ramos**

Recife
2020

Catálogo na fonte
Bibliotecária Arabelly Ascoli CRB4-2068

C837d Costa, Leon Tarquino da
Dinâmica de evolução em processos com operadores de substituição / Leon Tarquino da Costa . – 2020.
108 f.: il., fig., tab.

Orientador: Alex Dias Ramos
Tese (Doutorado) – Universidade Federal de Pernambuco.
CCEN. Estatística. Recife, 2020.
Inclui referências e apêndices.

1. Probabilidade. 2. Sistemas de partículas interagentes com comprimento variável. 3. Interação não local. 4. Tempo médio de convergência. I. Ramos, Alex Dias (orientador). II. Título.

519.2 CDD (22. ed.) UFPE-CCEN 2020-38

Dinâmica de evolução em processos com operadores de substituição

por

Leon Tarquino da Costa

Tese apresentada ao programa de Pós-graduação em Estatística da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Doutor em Estatística.

Área de Concentração: Probabilidade

Aprovada em 28 de janeiro de 2020.

Banca Examinadora:

Prof. Dr. Alex Dias Ramos – DE - UFPE
(Orientador)

Prof. Dr. Abraão David Costa do Nascimento – DE - UFPE
(Examinador interno)

Prof^a. Dr^a. Maria do Carmo Soares de Lima – DE - UFPE
(Examinador Interno)

Prof. Dr César Augusto Rodrigues Castilho – DMAT - UFPE
(Examinador externo)

Prof. Dr. Sérgio Galvão Coutinho – DF - UFPE
(Examinador externo)

Dedico esta tese

Aos meus pais, Maria do Socorro Padilha da Costa e José Tarquino da Costa Filho e a minha esposa Ivanilde Carlos Tarquino Moureira Neta.

AGRADECIMENTOS

Quero agradecer primeiramente a Deus, pois sem ele nada sou.

A minha família, em especial a minha mãe Maria do Socorro Padilha da Costa, que é sem dúvida a pessoa que mais quis que eu fizesse um doutorado, ao meu pai José Tarquino da Costa Filho, que apesar de seu pouco estudo é a pessoa mais inteligente que conheço e é também a pessoa em quem me inspiro todos os dias. À minha esposa Ivanilde Carlos Tarquino Moureira Neta, que é meu ponto de equilíbrio, meu ombro amigo, minha companheira. A estas pessoas mais que especiais para mim, agradeço pela compreensão, força, confiança e carinho que tiveram por mim nestes últimos anos. Sem estas pessoas certamente não estaria onde estou.

Ao prof. Dr. Alex Dias Ramos por ter orientado este trabalho, pelo conhecimento transmitido, pelo tempo que dedicou a mim, por me motivar, transmitindo sua experiência de forma sábia e precisa, e principalmente, pela confiança depositada. Sou um homem abençoado, mas além disso, tive muita sorte de tê-lo escolhido para me orientar. Uma pessoa humilde e sincera, que me cobrou muito, mas sempre forneceu os meios para que eu pudesse realizar meu trabalho.

Aos professores do Programa de Pós Graduação em Estatística da Universidade Federal de Pernambuco (UFPE), em especial àqueles que tive o prazer de cursar disciplinas.

Aos colegas que conheci na minha caminhada acadêmica, em especial àqueles que transpassaram a barreira do coleguismo e se tornaram amigos como Suelena, Mariana, Thiago e não poderia deixar de citar um que considero como irmão Victor José Araujo de Carvalho.

A banca examinadora: Prof. Dr. Abraão David Costa do Nascimento, Prof^a. Dr^a Maria do Carmo Soares de Lima, Prof. Dr César Augusto Rodrigues Castilho, Prof. Dr. Sérgio Galvão Coutinho, Prof. Dr. Francisco José de Azevêdo Cysneiros e Prof^a. Dr^a. Calitéia Santana de Sousa por aceitarem participar da avaliação deste trabalho e pelas contribuições que enriquecem o nosso trabalho. Agradeço também ao Prof. Dr. Getúlio José Amorim do Amaral por ter participado do meu exame de qualificação.

Ao CODAI-UFRPE, instituição a qual pertencço, que me concedeu afastamento remunerado, me liberando das atividades internas no ano de 2019, para que eu pudesse me dedicar exclusivamente ao doutorado nesta reta final.

À todos aqueles que contribuíram direta ou indireta para meu crescimento e para que eu pudesse concluir este curso.

“Aqueles que abrem mão da liberdade essencial por um pouco de segurança temporária não merecem nem liberdade nem segurança.”

(FRANKLIN, 1818, p. 142)

RESUMO

Em nosso trabalho, estudamos um *sistema de partículas interagentes com comprimento variável* (SPICV) unidimensional com tempo discreto, em que cada partícula pode ter dois possíveis estados, chamados de *mais* e *menos*. Para cada passo de tempo duas transformações ocorrem: a primeira chamada de *flip* transforma qualquer menos em mais com probabilidade β independentemente do que ocorre nas outras posições; a segunda é chamada *murder*, sob sua ação sempre que um mais é vizinho esquerdo de um menos, este mais desaparece com probabilidade α , independentemente do que ocorre nas outras posições. Mostramos que para esse processo há um tipo de transição de fase de primeira ordem. E também, que o espaço de parâmetros (β, α) exibe uma região para a qual o processo é não ergódico e outra para a qual o processo é ergódico. Apresentamos condições para que nosso processo Flip-Murder, começando em uma medida de probabilidade uniforme, sempre convirja para a configuração todos mais. Com o objetivo de encontrar uma curva ótima que separa as regiões de ergodicidade e não ergodicidade, fizemos um estudo por meio de uma Aproximação de Campo Médio e também um estudo computacional por meio de simulação de Monte Carlo. Estudamos também um outro SPICV unidimensional com tempo discreto, em que cada partícula pode estar nos estados *mais* e *menos* e para cada passo de tempo duas transformações ocorrem: a primeira chamada de *flip* transforma qualquer menos em mais com probabilidade β ; a segunda, chamada *aniquilação* é imparcial, sob sua ação sempre que um mais é vizinho esquerdo de um menos, ambos desaparecem com probabilidade α , independentemente do que ocorre nas outras posições. Este processo foi proposto por Toom (2004). Em um dos resultados trazidos para esse processo, ele mostra no espaço de parâmetros, (α, β) , a existência de duas regiões, uma de não ergodicidade e outra de ergodicidade do processo. Em nosso trabalho, ampliamos a região de não ergodicidade. Provamos para uma dada classe de distribuições iniciais, chamadas de arquipélagos de menos, que o processo Flip-Aniquilação sempre convergirá em distribuição para a medida concentrada na configuração em que todos os estados são mais. Além disso, obtivemos um limite superior para o tempo médio dessa convergência. Por fim, sempre que tomamos uma distribuição inicial dessa classe, descrevemos o limite superior do número médio de menos. Mostramos que esses limites superiores são funções da distribuição inicial.

Palavras-chaves: Sistemas de partículas interagentes com comprimento variável. Interação não local. Tempo médio de convergência.

ABSTRACT

In our work we study a *interacting particle systems with variable length* (IPSVL) one-dimensional with discrete time, where each particle can have two possible states, called *plus* and *minus*. For each time step two transformations occur: the first called *flip* which transforms any minus into plus with probability β independently of what happens in the other positions; the second is called *murder*, under its action whenever a plus is a left neighbor of a minus, this plus disappears with probability α , independently of what happens in the other positions. We show for this process that there is a first order phase transition type. We show in the parameter space (β, α) that there is a region for which the process is non ergodic and a for which the process is ergodic. We present conditions for our Flip-Murder process, starting at a uniform probability measure, always converge to the all plus configuration. In order to find an optimal curve that separates the ergodicity and non ergodicity regions, we made a study using a Midfield Approximation and also a computational study by means of Monte Carlo simulation. We also study another one-dimensional with discrete time IPSVL, where each particle can be in the *plus* and *minus* states and for each time step two transformations occur: the first called *flip* transforms any minus into plus with probability β ; the second, called *annihilation*, is impartial, under its action whenever a plus is a left neighbor of a minus, both disappear with probability α , independently of what happens in other positions. This process was proposed by A. Toom. In one of the results brought to this process, he shows in the parameter space, (α, β) , the existence of two regions, one of non-ergodicity and another of ergodicity of the process. In our work, we expanded the non-ergodicity region. We proved for a given initial distributions class, which we call archipelagos of minus, that the Flip-Annihilation process will always converge on distribution to the configuration concentrated measure, where all states are plus. In addition, we obtained an upper bound for the mean time of this convergence. Finally, whenever we take an initial distribution of this class, we describe the upper bound of the mean number of minus. We show that these upper bonds are functions of the initial distribution.

Key-words: Interagent particle systems with variable length. Non-local interaction. Convergence mean time.

LISTA DE FIGURAS

Figura 1	Ilustração de parte de um grafo G e seu grafo dual \overline{G}	35
Figura 2	Ilustração do contorno $\text{tour}(\omega)$	40
Figura 3	Ilustração do contorno $\text{tour}(\omega)$ com o tipo 3 tendo deslocamento $(0, 1)$ e $(-1, 1)$	48
Figura 4	Uma circular C com $ C = n$	56
Figura 5	Este gráfico mostra ambas as curvas rigorosas (em verde) e duas aproximações para a curva verdadeira que separa as regiões de ergodicidade e não ergodicidade: a aproximação de Campo Médio (campo médio em azul) e a obtida por simulação de Monte Carlo (M.C.).	58
Figura 6	As cores representam os valores de $\overline{s(\beta, \alpha)}$, onde o processo sugere não ergodicidade. A paleta de cores no lado direito, correspondem ao valor de $\overline{s(\beta, \alpha)}$. Quando $\overline{s(\beta, \alpha)} = 1$ a cor associada é branca.	59
Figura 7	Em ambas as figuras descrevemos as curvas em que $\beta = \frac{\alpha^2}{250}$ e $\beta = \frac{9\alpha^2}{1000}$. Em 7(b), plotamos os dados obtidos por simulação computacional do processo. Esses dados estimam a curva crítica. Se (α, β) estão abaixo da curva crítica estimada, o processo μ_t é não ergódico, caso contrário, o processo μ_t é ergódico.	69
Figura 8	Ilustração da ação dos operadores N_α e A_α numa ilha cujo $\text{Pop}(x) = 3$	72
Figura 9	Diagrama das probabilidades de transição do processo (4.18).	76
Figura 10	Diagrama das probabilidades de transição do processo (4.19).	76
Figura 11	Gráficos das estimativas e dos limites superiores de $\mathbb{E}(\tau_\mu)$ para ilhas de menos em que $\text{Pop}(x) = 10$	88
Figura 12	Gráficos da estimativa e dos limites superiores de $\mathbb{E}(\tau_\mu)$ para ilhas de menos em que $\text{Pop}(x) = 50$	89
Figura 13	Gráficos da estimativa e dos limites superiores de $\mathbb{E}(\tau_\mu)$ para ilhas de menos em que $\text{Pop}(x) = 100$	90
Figura 14	Gráficos das estimativas de $\mathbb{E}(\tau_\mu)$ para ilhas de menos em que $\text{Pop}(x) \in \{10, 50, 100\}$	91

LISTA DE TABELAS

Tabela 1	Definição de tipos, eventos associados e variáveis associadas a passos em G começando em um \oplus -vértice.	38
Tabela 2	Definição de tipos, chances e deslocamentos de passos em \bar{G} tendo uma \oplus -face em seu lado esquerdo.	39

LISTA DE SÍMBOLOS

\mathcal{A}	alfabeto
Λ	palavra vazia
$dic(\mathcal{A})$	dicionário
\mathbb{Z}	conjunto dos números inteiros
$\mathcal{A}^{\mathbb{Z}}$	conjunto das sequências bi-infinitas cujos termos são elementos de \mathcal{A}
\mathbb{A}	topologia discreta sob \mathcal{A}
$\mathbb{A}^{\mathbb{Z}}$	σ -álgebra do espaço produto $\mathcal{A}^{\mathbb{Z}}$
\mathcal{M}	conjunto das medidas de probabilidade uniformes em $\mathcal{A}^{\mathbb{Z}}$
\mathcal{M}_u	conjunto das medidas normalizadas e uniformes em $\mathcal{A}^{\mathbb{Z}}$
Ω ou $\Omega(\mathcal{A})$	conjunto das palavras aleatórias em \mathcal{A}
$\mathcal{M}_{\{\ominus, \oplus\}}$	conjunto das medidas normalizadas, uniformes em $\{\ominus, \oplus\}^{\mathbb{Z}}$
Δ	conjunto das ilhas de menos
\mathcal{A}_{\ominus}	conjunto das medidas normalizadas na σ -álgebra gerada pelos cilindros em Δ
\mathcal{F}	σ -álgebra gerada pelos cilindros em Δ
$\overline{\Omega}$	é o espaço $(\{\text{mover, ficar}\} \times \{\text{atacar, parar}\})^{\mathbb{Z} \cdot \mathbb{Z}_+}$
Ω_0	conjunto dos $\omega \in \overline{\Omega}$ tal que $\text{estado}(x, 2T) = \ominus$
Ω_k para $k \in \{1, 2, 3, \dots\}$	conjunto dos $\omega \in \Omega_0$ para os quais o número de flores é $\geq k$
Ω_{\odot}	conjunto das configurações bi-infinitas, formadas por \oplus, \ominus e \odot
$\text{real}(C)$	conjunto dos $\omega \in \Omega_0$ que realizam C
CL_k	conjunto dos códigos k -legais
CPL_k	conjunto dos códigos principais k -legais
Δ_{\odot}	extensão do conjunto das ilhas de menos
\mathcal{A}_{\odot}	conjunto das medidas normalizadas na σ -álgebra gerada pelos cilindros em Δ_{\odot}
\mathbb{N}	conjunto dos números naturais
\mathbb{Z}_+	conjunto dos números inteiros não negativos
\mathbb{Z}^2	conjunto dos pares ordenados em que cada coordenada é um número inteiro
\mathbb{Z}_n	conjunto dos inteiros módulo n
Ω_n	espaço dos periódicos
F_{β}	operador flip
M_{α}	operador murder
A_{α}	operador aniquilação

Trap_α	operador armadilha
Hide	operador ocultador
longo	aplicação longo
curto	aplicação curto
G	grafo
\overline{G}	grafo dual
HD	deslocamento horizontal
VD	deslocamento vertical
$\text{chance}(C)$	chance de um código C
OS	operador de substituição
$\mathbb{E}(X)$	valor esperado da variável aleatória X

SUMÁRIO

1	INTRODUÇÃO	15
2	DEFINIÇÕES GERAIS	18
2.1	Operadores de Substituição	19
3	O OPERADOR FLIP-MURDER	22
3.1	Introdução ao operador Flip-Murder	22
3.2	Prova dos Teoremas 2, 3 e 4	26
3.3	O processo ν e sua representação gráfica	30
3.4	Uma cadeia de igualdades e desigualdades	34
3.5	Aproximação de Campo Médio do Processo Flip-Murder	53
3.6	Simulação de Monte Carlo	56
3.6.1	Estimação de $s(\beta, \alpha)$	58
4	MODELO FLIP-ANIQUILAÇÃO: NOVA ESTIMATIVA E CARACTERIZAÇÕES	60
4.1	Apresentação dos Teorema 5 e 6	60
4.2	Prova dos Teoremas 5 e 6	62
4.3	Ilhas e Arquipélagos	68
4.4	Distância de Variação Total	71
4.5	Os Operadores Neutralização N_α , e Eliminação E_α	72
4.6	Algumas Cadeias de Markov em \mathbb{Z}_+	75
4.7	Uma composição de duas Cadeias de Markov em \mathbb{Z}_+	80
4.8	Operador $F_\beta E_\alpha$ e o processo Z	82
4.9	Prova dos Teoremas 7 e 8	83
4.10	Estudo computacional para $\mathbb{E}(\tau_\mu)$	85
5	CONCLUSÃO E TRABALHOS FUTUROS	92
	REFERÊNCIAS	94
	APÊNDICE A - PROVA DA SOBREJETIVIDADE DA APLI- CAÇÃO longo.	97

APÊNDICE B - LEMAS UTILIZADOS NAS DEMONSTRAÇÕES DOS TEOREMAS 1 E 5.

Capítulo 1

INTRODUÇÃO

A partir da metade do século XX, teve início o desenvolvimento de uma nova parte da teoria de processos estocásticos, chamada teoria de processos estocásticos com interação local (MAES *et al.*, 2005; DE SANTANA *et al.*, 2015). Esta teoria é atualmente mais conhecida por *sistemas com partículas interagentes*, SPI por simplicidade. Novos tipos e modelos de SPI têm surgido. Estes novos modelos não são explicados completamente na estrutura matemática da teoria padrão. Os SPI que são tradicionalmente estudados, geralmente não podem ser propostos para modelar fenômenos investigados em teorias formuladas recentemente (DE SANTANA *et al.*, 2015; MAES *et al.*, 2005; RAMOS *et al.*, 2017). Em nosso trabalho, estudamos um SPI pertencente a uma nova classe de processos. Diferente dos modelos tradicionais normalmente estudados, o qual considera que o conjunto de partículas, também chamados espaço, não muda no processo de interação, em nosso processo o espaço muda no processo de interação. Estes processos são conhecidos por *sistemas de partículas interagentes com comprimento variável*, SPICV por simplicidade, (RAMOS *et al.*, 2017; ROCHA *et al.*, 2011; TOOM, 2004; TOOM *et al.*, 2011).

Inicialmente, nosso estudo visa fornecer entendimento teórico através do desenvolvimento de técnicas analíticas e modelagens computacionais, as quais nos forneçam melhor compreensão desta nova classe. Esta fundamentação é crucial para fornecer uma melhor perspectiva de possíveis aplicações.

Um modelo de sistemas de partículas interagentes com comprimento variável considerando-se o tempo contínuo, foi proposto para estudar aspectos da gravitação quântica (MALYSHEV, 2000; MALYSHEV, 2002).

O modelo de Ising (MCCOY, 2014) é uma ferramenta importante para o estudo de materiais com propriedades magnéticas. Seu objetivo principal foi estudar o ferromagnetismo. Ele apresenta um tipo de transição de fase em temperatura finita, quando a dimensão do sistema, d , é maior ou igual a 2; e não apresenta transição quando $d = 1$.

O estudo de percolação começou em 1957, motivado por algumas considerações físicas e muito progresso ocorreu ao longo dos anos, principalmente no caso bidimensional (GRIMMETT, 1999). Um dos modelos mais simples na teoria de probabilidade que exhibe o *parâmetro crítico* α^* , é o modelo de percolação. Isso significa que há um parâmetro de

ordem no modelo que apresenta um tipo de comportamento para valores menores que α^* e outro para valores maiores que α^* . Nos modelos de percolação os resultados são fáceis de serem enunciados, mas suas soluções, quando conhecidas, normalmente requerem métodos sofisticados (STEIF, 2011; STAUFFER e AHARONY, 2018).

Uma classe relacionada de processos estocásticos, chamada de sistema de expansão-modificação, foi estudada em (SALGADO-GARCÍA e UGALDE, 2013). O sistema de expansão-modificação pertence à classe de sistemas dinâmicos de substituição aleatória, que tem atraído atenção nas últimas décadas pelas possíveis aplicações nos estudos de evolução dos genomas. Isso, devido ao fato de que vários modelos foram introduzidos (dentre eles, (LI, 1997; MA, *et al*, 2008; SAAKIAN, 2008; SOBOTTKA e HART, 2011)) para descrever a evolução das sequências de nucleotídeos, bem como os padrões e correlações que ocorrem no genoma. Este modelo pertence ao SPICV.

Nosso trabalho é direcionado ao estudo de SPICV unidimensional com tempo discreto, onde cada partícula pode ter dois possíveis estados, chamados de *mais* e *menos*. Temos interesse em transição de fase, e no entendimento teórico e computacional destes processos.

No Capítulo 2, traremos algumas definições e resultados gerais que serão utilizados no desenvolvimento desse trabalho.

No Capítulo 3, estudamos um SPICV onde para cada passo de tempo duas transformações ocorrem: a primeira chamada de *flip* transforma qualquer menos em mais com probabilidade β , independentemente do que ocorre nas outras posições; a segunda é chamada *murder*, sob sua ação sempre que um mais é vizinho esquerdo de um menos, este mais desaparece com probabilidade α , independentemente do que ocorre nas outras posições. Mostramos para esse processo que há um tipo de transição de fase de primeira ordem. Mostramos no espaço de parâmetros (β, α) que existe uma região para a qual o processo é não ergódico e outra para a qual o processo é ergódico. Para mostrar o resultado de não ergodicidade, utilizamos duas ideias conhecidas na literatura: o método do contorno de Peierls e a dualidade dos grafos planares. Apresentamos condições para que nosso processo Flip-Murder, começando em uma medida de probabilidade uniforme, sempre convirja para a configuração todos mais, na qual todas as partículas estão no estado mais. Mostramos também a existência de uma outra medida invariante para o operador Flip-Murder, cuja frequência de mais é estritamente menor que 1. Com o objetivo de encontrar uma curva ótima que separa as regiões de ergodicidade e não ergodicidade desse processo, fizemos um estudo por meio de uma Aproximação de Campo Médio e também um estudo computacional por meio de simulação de Monte Carlo e obtivemos assim duas curvas que nos permitem conjecturar sobre o comportamento da curva verdadeira (que separa as regiões de ergodicidade e não ergodicidade), caso ela exista.

No Capítulo 4, revisitamos um outro SPICV unidimensional com tempo discreto, em que cada partícula pode estar nos estados *mais* e *menos* e para cada passo de tempo duas transformações ocorrem: a primeira chamada de *flip* transforma qualquer menos em mais com probabilidade β , independentemente do que ocorre nas outras posições; a segunda, chamada *aniquilação* é imparcial, sob sua ação sempre que um mais é vizinho esquerdo de

um menos, ambos desaparecem com probabilidade α , independentemente do que ocorre nas outras posições. Esse processo foi proposto por Toom (2004), para o qual ele mostrou que há um tipo de transição de fase de primeira ordem. Num dos resultados trazidos para esse processo, ele mostra no espaço de parâmetros, (α, β) , a existência de uma região para a qual o processo é não ergódico e outra para o qual o processo é ergódico. Em nosso trabalho, ampliamos a região de não ergodicidade. Provamos que para uma dada classe de distribuições iniciais, que chamamos de arquipélagos de menos, o processo Flip-Aniquilação sempre convergirá fracamente para a medida concentrada na configuração, em que todos os estados são mais. Não havendo sob essas condições uma transição de fase. Além disso, obtivemos um limite superior para o tempo médio dessa convergência, esse limite superior é função da distribuição inicial e apresentamos a forma explícita dessa função. Por fim, sempre que tomamos uma distribuição inicial dessa classe, descrevemos o limite superior do número médio de menos, o qual é função da distribuição inicial.

Capítulo 2

DEFINIÇÕES GERAIS

Neste capítulo, relembremos algumas definições e resultados gerais, que serão utilizados no decorrer deste trabalho, com o objetivo de fomentar uma melhor compreensão dos nossos resultados.

Chamaremos de *alfabeto* qualquer conjunto finito e não-vazio e denotaremos por \mathcal{A} . Os elementos desse conjunto, chamaremos de *letras*, e qualquer sequência finita de letras chamaremos de *palavra*. O número de letras em uma palavra W é chamado *comprimento* e denotado por $|W|$. Qualquer letra pode ser considerada como uma palavra de comprimento um. Existe a palavra vazia, denotada por Λ , cujo comprimento é zero. O conjunto das palavras no alfabeto \mathcal{A} é chamado de *dicionário* e denotado por $\text{dic}(\mathcal{A})$. Denotemos por \mathbb{Z} o conjunto dos números inteiros e por $\mathcal{A}^{\mathbb{Z}}$ o conjunto das sequências bi-infinitas, cujos termos são elementos de \mathcal{A} . Essas sequências bi-infinitas podem ser encaradas como funções $x : \mathbb{Z} \rightarrow \mathcal{A}$, sendo $x(i)$ denotado por x_i .

Lembremos o enunciado do Teorema de Tychonoff, cujo estudo mais detalhado, pode ser encontrado em (MUNKRES, 2000) e (LIMA, 1970).

Teorema de Tychonoff. *Um produto cartesiano arbitrário de espaços topológicos compactos, é compacto na topologia produto.*

Denotemos por \mathbb{A} a topologia discreta sob \mathcal{A} . Consideremos medidas de probabilidades na σ -álgebra $\mathbb{A}^{\mathbb{Z}}$ do espaço produto $\mathcal{A}^{\mathbb{Z}}$ dotado com a topologia, que é o produto das topologias discretas sob todas as cópias de \mathcal{A} . Uma vez que \mathcal{A} é finito, ele é compacto na topologia discreta, e pelo Teorema de Tychonoff, $\mathcal{A}^{\mathbb{Z}}$ também é compacto.

Como sempre, translações em \mathbb{Z} geram translações em $\mathcal{A}^{\mathbb{Z}}$ e conseqüentemente geram translações em $\mathbb{A}^{\mathbb{Z}}$. Denotemos por \mathcal{M} o conjunto das medidas de probabilidade uniformes em $\mathcal{A}^{\mathbb{Z}}$. Chamamos uma medida $\mu \in \mathcal{M}$ uniforme, se ela é invariante sob todas as translações. Isto é, para toda palavra $W = (a_1, \dots, a_n)$,

$$\mu(W) = \mu(a_1, \dots, a_n) = \mu(s_{i+1} = a_1, \dots, s_{i+n} = a_n),$$

para todo $i \in \mathbb{Z}$. Para formar uma medida uniforme, o número $\mu(W)$ deve ser não negativo,

e para qualquer palavra W (incluindo a palavra vazia)

$$\mu(W) = \sum_{a \in \mathcal{A}} \mu(W, a) = \sum_{a \in \mathcal{A}} \mu(a, W)$$

onde (W, a) e (a, W) são concatenações da palavra W e da letra a nas duas possíveis ordens.

2.1 Operadores de Substituição

Neste ponto trataremos algumas definições e resultados descritos em (TOOM *et al.*, 2011; ROCHA *et al.*, 2011).

Dadas duas palavras $W = (a_1, \dots, a_m)$ e $V = (b_1, \dots, b_n)$, onde $|W| \leq |V|$, chamamos os números inteiros pertencentes ao intervalo $[0, n - m]$, de *posições* de W em V . Dizemos que W *entra* em V na posição k se

$$\forall i \in \mathbb{Z} : 1 \leq i \leq m \Rightarrow a_i = b_{i+k}.$$

Chamamos uma palavra W de *auto-sobreposição* se existe uma palavra V tal que $|V| < 2|W|$ e W entra em V em duas diferentes posições. Uma palavra é chamada *auto-evitando* se ela não é auto-sobreposição. Em particular, a palavra vazia, toda palavra consistindo de uma letra, e toda palavra consistindo de duas letras diferentes são auto-evitando.

Um *Operador de Substituição* genérico, ou OS por simplicidade, atua de \mathcal{M} em \mathcal{M} da seguinte maneira: dadas duas palavras G e H , onde G é auto-evitando, e um número real $\rho \in [0, 1]$, um operador de substituição, falando informalmente, substitui toda ocorrência da palavra G em qualquer configuração, pela palavra H com probabilidade ρ (ou a deixa inalterada com probabilidade $1 - \rho$), independentemente do que acontece nas outras posições. Denotamos este tipo de operador por $(G \xrightarrow{\rho} H)$.

Definimos uma *palavra aleatória* X no alfabeto \mathcal{A} , como uma variável aleatória sobre $\text{dic}(\mathcal{A})$ que é concentrada em um subconjunto finito de $\text{dic}(\mathcal{A})$. Uma palavra aleatória é determinada por suas componentes $P(X = V)$, isto é, a probabilidade de $X = V$, cuja soma é 1, o conjunto $\{V : P(X = V) > 0\}$ é finito.

Seja Ω o conjunto das palavras aleatórias em \mathcal{A} . Vamos definir algumas classes de operadores atuando sob palavras aleatórias, que chamamos *operadores básicos*.

Operador Básico 1: Conversão ($g \xrightarrow{\rho} h$). Para qualquer duas letras diferentes $g, h \in \mathcal{A}$, definimos a *conversão* de g para h como uma aplicação de Ω em Ω . O operador conversão transforma cada ocorrência da letra g na letra h com probabilidade $\rho \in [0, 1]$, ou a deixa inalterada com probabilidade $1 - \rho$, independentemente das outras ocorrências.

Observação 1. *O operador de conversão é o único operador linear, dentre os 5 que vamos apresentar neste momento.*

Operador Básico 2: Inserção ($\Lambda \xrightarrow{\rho} h$). O operador de inserção introduz uma letra $h \notin \mathcal{A}$ entre todas duas letras vizinhas em uma palavra aleatória no alfabeto \mathcal{A} com probabilidade $\rho \in [0, 1]$, independentemente das posições das letras.

Operador Básico 3: Eliminação ($g \xrightarrow{\rho} \Lambda$). A eliminação de uma letra $g \in \mathcal{A}$ com probabilidade $\rho \in [0, 1]$, em uma palavra aleatória, significa que cada ocorrência de g desaparece com probabilidade ρ ou permanece inalterada com probabilidade $1 - \rho$, independentemente das outras ocorrências.

Operador Básico 4: Compressão ($G \xrightarrow{1} h$). Dada uma palavra não-vazia e auto-evitando G no alfabeto \mathcal{A} e uma letra $h \notin \mathcal{A}$, a *compressão* de G em h é a seguinte aplicação de $\Omega(\mathcal{A})$ em $\Omega(\mathcal{A}')$, onde $\mathcal{A}' = \mathcal{A} \cup \{h\}$ e $\Omega(\mathcal{A})$ é o conjunto das palavras aleatórias no alfabeto \mathcal{A} tal que cada ocorrência da palavra G é substituída pela palavra h com probabilidade 1.

Operador Básico 5: Descompressão ($g \xrightarrow{1} H$). Dada uma palavra não-vazia e auto-evitando H no alfabeto \mathcal{A} e uma letra $g \notin \mathcal{A}$, a *descompressão* de g em H é uma aplicação de $\Omega(\mathcal{A}')$ em $\Omega(\mathcal{A})$, onde $\mathcal{A}' = \mathcal{A} \cup \{g\}$ dada por: toda ocorrência da letra g é substituída pela palavra H com probabilidade 1.

Note que os operadores básicos de 1 à 5 são OS.

Definição 1. *Seja $\nu \in \mathcal{M}$ e $P = P_1 \circ \dots \circ P_j$, onde P_1, \dots, P_j é uma sequência finita de operadores de substituição. Então, definimos o processo de substituição discreto generalizado (ν_t) , onde $\nu_0 = \nu$, como segue,*

$$\nu_t(W) = \nu P^t(W) \quad \text{para toda palavra } W,$$

e P^t denota a t -ésima composição do operador P .

Definição 2. *Uma medida μ é chamada invariante para um operador P , se $\mu P = \mu$.*

Os seguintes resultados, provados em (ROCHA *et al.*, 2011), serão utilizados por nós;

Resultado 1 (Teorema (ROCHA *et al.*, 2011)). *Consideremos um processo de substituição discreto generalizado $\nu_n = \nu P^n$, onde $P = P_1 P_2 \dots P_j$ como na definição 1. Seja $R \subset \mathbb{A}^{\mathbb{Z}}$ algum subconjunto da σ -álgebra $\mathbb{A}^{\mathbb{Z}}$. Então, se $\nu_n(c) \leq \delta$ (respectivamente $\nu_n(c) \geq \epsilon$) para todo $c \in R$, então P tem uma medida invariante μ tal que $\mu(c) \leq \delta$ (respectivamente $\mu(c) \geq \epsilon$) para todo $c \in R$, onde $\delta, \epsilon > 0$ são constantes positivas.*

Resultado 2 (Teorema (ROCHA *et al.*, 2011)). *Dadas duas palavras G e H , considere o operador $(G \xrightarrow{\rho} H)$ atuando sobre medidas, onde G é auto-evitando e $\rho \in [0, 1]$ ($\rho < 1$, se $H = \Lambda$). Então as seguintes identidades valem para qualquer $s, g, h \notin \mathcal{A}$:*

$$\begin{aligned} \mu(G \xrightarrow{\rho} H) &= \mu(G \xrightarrow{1} h)(h \xrightarrow{\rho} s)(s \xrightarrow{1} H)(h \xrightarrow{1} G) \\ \mu(G \xrightarrow{\rho} \Lambda) &= \mu(G \xrightarrow{1} g)(g \xrightarrow{\rho} \Lambda)(g \xrightarrow{1} G) \\ \mu(\Lambda \xrightarrow{\rho} H) &= \mu(\Lambda \xrightarrow{\rho} h)(h \xrightarrow{1} H) \end{aligned}$$

Definição 3. Dizemos que um processo $\mu\mathbf{P}^t$, onde μ é alguma medida inicial e \mathbf{P} é um operador, é *érgódico*, se $\lim_{t \rightarrow \infty} \mu\mathbf{P}^t$ existe e é o mesmo para qualquer medida inicial μ .

Capítulo 3

O OPERADOR FLIP-MURDER

3.1 Introdução ao operador Flip-Murder

Em (TOOM, 2004) foi proposto um processo de comprimento variável composto pelos operadores F_β e A_α , onde F_β é o operador que transforma cada componente em uma configuração do estado menos para o estado mais com probabilidade β , independentemente uma das outras, e A_α é o operador que elimina toda palavra $\oplus\ominus$ com probabilidade α , ou seja, sempre que um mais é um vizinho esquerdo de um menos, essa palavra é eliminada com probabilidade α , independentemente de outras ocorrências dessa palavra. Dentre os resultados, foi exibida a existência de uma transição de fase de primeira ordem. Contudo, para melhorar a tratabilidade do operador A_α , algumas considerações foram necessárias; em uma, o conjunto de medidas teve que ser uniforme, e em outra, que ele denotou por caso problemático, ocorre quando $\alpha = 1$, o que levou a necessidade de se estudar esse operador nesse caso particular no trabalho (RAMOS e TOOM, 2008).

Neste trabalho, propomos um outro processo de comprimento variável unidimensional substituindo o operador A_α . Diferente de A_α , nosso operador fará nosso processo mais equilibrado pois, ao atuar em uma palavra do tipo $\oplus\ominus$, ele irá eliminar somente a letra \oplus com probabilidade α . Eliminando a necessidade de se estudar este modelo com distintos cenários, $\alpha \in (0, 1)$ e $\alpha = 1$. Em nosso modelo, conseguimos reproduzir de forma qualitativa, todos os resultados obtidos no modelo Flip-Aniquilação, com a vantagem de trabalhar com um modelo menos complexo.

Vamos mostrar alguns resultados considerando que nosso alfabeto \mathcal{A} tem apenas dois elementos, que denotamos por \ominus e \oplus e chamamos de *menos* e *mais*. Neste caso, nossos operadores atuam em $\mathcal{M}_{\{\ominus, \oplus\}}$, o conjunto das medidas normalizadas, uniformes no espaço de configuração $\{\ominus, \oplus\}^{\mathbb{Z}}$. A seguir, vamos definir dois operadores atuando em $\mathcal{M}_{\{\ominus, \oplus\}}$ e que dependem dos parâmetros α e β , onde $\alpha, \beta \in [0, 1]$.

O operador que chamamos de *flip*, e denotamos por F_β , é bastante conhecido. Este operador é de comprimento constante e linear. Sob sua ação, qualquer menos se transforma em mais com probabilidade β , independentemente uns dos outros. Precisamos representar o operador flip usando variáveis auxiliares independentes. Denotemos por $x_i \in \{\ominus, \oplus\}$,

para todo $i \in \mathbb{Z}$, as coordenadas do espaço $\{\ominus, \oplus\}^{\mathbb{Z}}$, onde a medida inicial μ é dada. Usaremos também, variáveis mutuamente independentes F_i , para todo $i \in \mathbb{Z}$, cada uma delas tomando dois valores chamados *mover* e *ficar*, distribuídas de acordo com a medida produto π , definidas da seguinte maneira:

$$F_i = \begin{cases} \text{mover,} & \text{com probabilidade } \beta, \\ \text{ficar,} & \text{com probabilidade } 1 - \beta. \end{cases}$$

Para finalizar a definição desse operador, tomamos um terceiro conjunto de variáveis $y_i \in \{\ominus, \oplus\}$, para todo $i \in \mathbb{Z}$, em que a medida μF_β é induzida pelo produto das medidas μ e π com a aplicação

$$y_i = \begin{cases} \ominus, & \text{se } x_i = \ominus \text{ e } F_i = \text{ficar,} \\ \oplus, & \text{em todos os outros casos.} \end{cases}$$

O operador F_β pode ser aplicado a uma medida $\mu \in \mathcal{M}_{\{\ominus, \oplus\}}$ e produz uma medida em $\mathcal{M}_{\{\ominus, \oplus\}}$, preservando o conjunto das medidas uniformes e normadas.

O segundo operador que chamaremos de *murder* é denotado por $M_\alpha : \mathcal{M}_{\{\ominus, \oplus\}} \rightarrow \mathcal{M}_{\{\ominus, \oplus\}}$. Este operador é de comprimento variável. Sob a ação deste operador, sempre que um mais é vizinho esquerdo de um menos, o mais desaparece com probabilidade α , independentemente de outras ocorrências deste tipo. Falando informalmente, sempre que uma palavra $\oplus\ominus$ ocorre na configuração infinita, o mais desaparece com probabilidade α independentemente de todas as outras ocorrências. Note que sob a ação de M_α os sítios¹ eliminados desaparecem completamente ao invés de ir para outro estado. Se considerássemos um processo similar em uma configuração finita, toda ação do operador murder em uma palavra $\oplus\ominus$ decresceria o tamanho da configuração em uma unidade. Devemos definir M_α como uma superposição de dois operadores: $M_\alpha = \text{Trap}_\alpha \text{Hide}$; primeiro a ação de Trap_α e depois a ação de Hide . Podemos imaginar que quando Trap_α é aplicado, uma armadilha ocorre com todo par de \oplus e \ominus ocupando a i -ésima e a $(i + 1)$ -ésima posição, respectivamente. Se um comando “*atacar!*” é dado, o que ocorre para todo tal par independentemente uns dos outros, com probabilidade α , a componente no estado menos “assassina” a componente no estado mais. Caso contrário, um comando “*parar!*” é dado, e nada acontece. Quando Hide é aplicado, os “cadáveres” são ocultados e os sítios vivos fecham a configuração, tornando-se vizinhos.

Agora, vamos definir o operador Trap_α (que chamaremos de operador *armadilha*), um operador linear e de comprimento constante que transforma qualquer medida em $\{\ominus, \oplus\}^{\mathbb{Z}}$ em uma medida em $\{\ominus, \oplus, \odot\}^{\mathbb{Z}}$, onde \odot é um terceiro estado, introduzido especialmente para esta ocasião e chamado de *morto*. Estados diferentes de morto, isto é, os estados mais e menos, são chamados *vivos*. Mais uma vez, vamos chamar $x_i \in \{\ominus, \oplus\}$, $i \in \mathbb{Z}$, as coordenadas do espaço $\{\ominus, \oplus\}^{\mathbb{Z}}$, onde a medida original μ está definida. Usaremos também variáveis mutuamente independentes A_i para todo $i \in \mathbb{Z}$, cada uma tomando dois valores chamados *atacar* e *parar* distribuídas de acordo com a medida produto π , definidas como

¹também chamados de componentes ou posições

segue

$$A_i = \begin{cases} \text{atacar,} & \text{com probabilidade } \alpha, \\ \text{parar,} & \text{com probabilidade } 1 - \alpha, \end{cases}$$

para qualquer $i \in \mathbb{Z}$ independentemente de todas as outras componentes e da medida μ . Denotemos por $y_i \in \{\ominus, \oplus, \odot\}$ as coordenadas do espaço, onde a medida μTrap_α é induzida pelo produto de μ e π , com a seguinte aplicação:

$$y_i = \begin{cases} \odot, & \text{se } x_i = \oplus, x_{i+1} = \ominus \text{ e } A_{i+1} = \text{atacar,} \\ x_i, & \text{em todos os outros casos.} \end{cases}$$

Note que, por um lado, $\mu(\oplus) = \mu(\oplus, \oplus) + \mu(\oplus, \ominus)$ e, por outro lado, $\mu(\oplus) = \mu(\oplus, \oplus) + \mu(\ominus, \oplus)$. Daí, $\mu(\oplus, \ominus) = \mu(\ominus, \oplus)$. Além disso, $\mu(\oplus, \ominus) + \mu(\ominus, \oplus) \leq 1$. Logo, $\mu(\oplus, \ominus) \leq \frac{1}{2}$. Portanto, sendo $\alpha < 1$, e da definição de Trap_α , temos

$$\mu\text{Trap}_\alpha(\odot) = \alpha\mu(\oplus, \ominus) < \frac{1}{2}. \quad (3.1)$$

Agora vamos definir o operador de comprimento variável $\text{Hide} : \mathcal{M}_{\{\ominus, \oplus, \odot\}} \rightarrow \mathcal{M}_{\{\ominus, \oplus\}}$. Para qualquer $\mu \in \mathcal{M}_{\{\ominus, \oplus, \odot\}}$, expressamos os valores de μHide sob todas as palavras no alfabeto $\{\ominus, \oplus\}$ em termos dos valores de μ sob todas as palavras no alfabeto $\{\ominus, \oplus, \odot\}$. Definiremos $\mu\text{Hide}(\Lambda) = 1$. Para qualquer palavra não vazia $W = (a_0, \dots, a_k) \in \text{dic}(\ominus, \oplus)$, definimos

$$\begin{aligned} \mu\text{Hide}(W) &= \mu\text{Hide}(a_0, \dots, a_k) \\ &= \frac{1}{1 - \mu(\odot)} \sum_{n_1, \dots, n_k=0}^{\infty} \mu(a_0 \odot^{n_1} a_1 \odot^{n_2} a_2 \cdots \odot^{n_{k-1}} a_{k-1} \odot^{n_k} a_k), \end{aligned} \quad (3.2)$$

onde \odot^n significa uma palavra consistindo de n letras, todas as quais são iguais a \odot , em particular $\odot^0 = \Lambda$. Assim,

$$a_0 \odot^{n_1} a_1 \odot^{n_2} a_2 \cdots \odot^{n_{k-1}} a_{k-1} \odot^{n_k} a_k$$

significa a palavra que começa com a letra a_0 , em seguida tem n_1 letras \odot , depois a letra a_1 , em seguida tem n_2 letras \odot , depois a letra a_2 , e assim sucessivamente até n_k letras \odot , e finalmente a letra a_k . Note que a fórmula (3.2) é não linear, daí a teoria bem desenvolvida de operadores lineares não pode ser aplicada aqui, o que adiciona uma dificuldade para lidar com processos de comprimento variável. No caso particular em que $k = 0$, da equação (3.2) temos

$$\mu\text{Hide}(\ominus) = \frac{\mu(\ominus)}{1 - \mu(\odot)}, \quad \mu\text{Hide}(\oplus) = \frac{\mu(\oplus)}{1 - \mu(\odot)}. \quad (3.3)$$

Note que, para qualquer $\alpha < 1$, o operador $M_\alpha = \text{Trap}_\alpha\text{Hide}$ pode ser aplicado a qualquer medida em $\mathcal{M}_{\{\ominus, \oplus\}}$ e retorna uma medida em $\mathcal{M}_{\{\ominus, \oplus\}}$.

Usando a notação do Capítulo 2, o operador murder é denotado por $(\oplus \ominus \xrightarrow{\alpha} \ominus)$.

Denotemos a medida

$$\mu_t = \delta_{\ominus}(\mathbf{F}_{\beta}\mathbf{M}_{\alpha})^t, \quad (3.4)$$

onde δ_{\ominus} é a medida concentrada na configuração todos menos.

Vejamos a seguir alguns casos triviais para nosso processo (3.4), os quais não deveremos nos preocupar no desenvolvimento das provas de nossos resultados.

- 1) Se $\alpha = 0$ e $\beta \in (0, 1)$, então $\delta_{\ominus}(\mathbf{F}_{\beta}\mathbf{M}_{\alpha})^t \rightarrow \delta_{\oplus}$ com $t \rightarrow \infty$.
- 2) Se $\beta = 1$ e $\alpha \in [0, 1]$, então $\delta_{\ominus}\mathbf{F}_{\beta}\mathbf{M}_{\alpha} = \delta_{\oplus}$.
- 3) Se $\beta = 0$ e $\alpha \in [0, 1]$, então $\delta_{\ominus}\mathbf{F}_{\beta}\mathbf{M}_{\alpha} = \delta_{\ominus}$.

Teorema 1. *Seja $\alpha \in (0, 1)$. Se $\beta < \frac{\alpha^2}{55}$,*

(A. 1) *então, para todo valor natural t , $\mu_t(\oplus) < 1$.*

(B. 1) *então, existe uma medida, ν , tal que $\nu(\mathbf{F}_{\beta}\mathbf{M}_{\alpha}) = \nu$, onde $\nu(\oplus) < 1$.*

Note que, se considerarmos inicialmente a configuração todos mais, então nem o operador flip nem o operador murder mudam essa configuração, logo δ_{\oplus} é invariante para $\mathbf{F}_{\beta}\mathbf{M}_{\alpha}$ para quaisquer α e β . Assim, o Teorema 1 implica que o operador $\mathbf{F}_{\beta}\mathbf{M}_{\alpha}$ não pode ser ergódico quando $\beta < \alpha^2/55$, porque neste caso, μ_t não pode convergir para δ_{\oplus} .

Teorema 2. *Se $\alpha < h(\beta)$, então a medida μ_t tende para δ_{\oplus} , quando $t \rightarrow \infty$, onde*

$$h(\beta) = \begin{cases} 2\beta, & \text{se } \beta \leq \beta^*, \\ \frac{\beta}{(1-\beta)^2 + (1-\beta)\beta^2}, & \text{se } \beta > \beta^*, \end{cases}$$

$$\text{em que } \beta^* = -\frac{1}{6} \left(26 + 6\sqrt{33} \right)^{1/3} + \frac{4}{3 \left(26 + 6\sqrt{33} \right)^{1/3}} + \frac{2}{3}.$$

Os Teoremas 1 e 2 mostram que a sequência de medidas μ_t tem dois diferentes modos de comportamento. No primeiro modo ($\beta < \alpha^2/55$) estas medidas não tendem para δ_{\oplus} quando $t \rightarrow \infty$ e no segundo modo ($h(\beta) > \alpha$) elas tendem para δ_{\oplus} . Na Figura 5 resumimos nossos resultados dos Teoremas 1 e 2, juntamente com o Lema 6 e uma simulação de Monte Carlo que fizemos para o processo Flip-Murder.

Teorema 3. *Sejam $\alpha, \beta \in (0, 1)$ e $\mu \in \mathcal{M}_{\{\ominus, \oplus\}}$. Se $\mu(\ominus) \in \left(0, \frac{\beta}{\alpha(1-\beta)} \right)$, então $\mu(\mathbf{F}_{\beta}\mathbf{M}_{\alpha})^t$ tende para δ_{\oplus} , quando $t \rightarrow \infty$.*

Denotemos por $s(\beta, \alpha)$ o supremo das densidades de \oplus na medida μ_t para todo natural t .

Teorema 4. *Para todo $\alpha \in \left(\frac{\beta}{(1-\beta)}, 1 \right)$, $s(\beta, \alpha)$ não é contínua como função de β .*

Nossos teoremas mostram semelhanças e diferenças entre nosso processo e o bem conhecido processo de contato (ver (LIGGETT, 2012)). Uma vez que nosso processo é discreto, é melhor compará-lo com o bem conhecido processo de Stavskaya, uma versão com tempo discreto do processo de contato (ver (STAVSKAYA e PIATETSKI-SHAPIRO, 1971)). Usando nossa notação, o processo de Stavskaya é uma sequência de medidas $\delta_{\ominus}(\mathbf{F}_{\beta}\mathbf{Stav})^t$, onde o operador determinístico, de comprimento constante, $\mathbf{Stav} : \{\ominus, \oplus\}^{\mathbb{Z}} \rightarrow \{\ominus, \oplus\}^{\mathbb{Z}}$ é definido pela regra

$$\forall s \in \{\ominus, \oplus\}^{\mathbb{Z}}, \quad k \in \mathbb{Z} : (s\mathbf{Stav})_k = \begin{cases} \oplus, & \text{se } s_k = s_{k+1} = \oplus, \\ \ominus, & \text{nos outros casos.} \end{cases}$$

O operador \mathbf{Stav} favorece os menos em comparação aos mais, porque ele torna qualquer mais em um menos sempre que seu vizinho direito é um menos, mas nunca torna menos em mais. O operador \mathbf{F}_{β} , pelo contrário, torna menos em mais com uma taxa β . É natural que a composição destes operadores se comporte de maneira diferente para valores grandes de β e para valores pequenos de β . A saber, quando β é grande, os menos desaparecem e quando β é pequeno eles não desaparecem. Em nosso caso, o comportamento é mais inesperado: \mathbf{F}_{β} favorece os mais para qualquer $\beta > 0$, de modo que os menos ainda sobrevivem para β/α^2 suficientemente pequeno.

Ao contrário do nosso processo, como mostrado pelo Teorema 3, o processo de Stavskaya não tende para δ_{\oplus} de qualquer medida inicial em que os menos e os mais estão misturados de forma aleatória em qualquer proporção, se considerarmos que a densidade inicial de menos é positiva e β é suficientemente pequeno.

3.2 Prova dos Teoremas 2, 3 e 4

Nesta seção iremos provar os Teoremas 2, 3 e 4, pois a prova do Teorema 1 é bastante construtiva e depende de algumas definições e resultados que apresentaremos nas seções 3.3 e 3.4. Iniciaremos provando o Teorema 3, pois, sua prova é independente, e o utilizaremos para provar o Teorema 2.

Notemos que os Teoremas 2, 3 e 4 apenas fazem sentido quando $\frac{\beta}{\alpha(1-\beta)} \leq 1$.

Prova do Teorema 3. Vamos mostrar inicialmente que $\mu\mathbf{F}_{\beta}\mathbf{M}_{\alpha}(\ominus) < \mu(\ominus)$. Note que para toda $\mu \in \mathcal{M}_{\{\ominus, \oplus\}}$, temos que $\mu\mathbf{Trap}_{\alpha}(\ominus) = \mu(\ominus)$. Este fato, juntamente com (3.1) e (3.3) implicam que

$$\mu\mathbf{F}_{\beta}\mathbf{M}_{\alpha}(\ominus) = \frac{\mu\mathbf{F}_{\beta}(\ominus)}{1 - \alpha\mu\mathbf{F}_{\beta}(\oplus, \ominus)}. \quad (3.5)$$

Da consistência da medida temos, $\mu(\ominus) = \mu(\ominus, \ominus) + \mu(\ominus, \oplus)$. Uma vez que $\mu(\ominus, \ominus) \geq 0$, segue que $\mu(\ominus) \geq \mu(\oplus, \ominus)$. Consequentemente,

$$\frac{\mu\mathbf{F}_{\beta}(\ominus)}{1 - \alpha\mu\mathbf{F}_{\beta}(\oplus, \ominus)} \leq \frac{\mu\mathbf{F}_{\beta}(\ominus)}{1 - \alpha\mu\mathbf{F}_{\beta}(\ominus)}.$$

Da definição de F_β temos que $\mu F_\beta(\Theta) = (1 - \beta)\mu(\Theta)$. Logo,

$$\frac{\mu F_\beta(\Theta)}{1 - \alpha \mu F_\beta(\Theta, \Theta)} \leq \frac{(1 - \beta)\mu(\Theta)}{1 - \alpha(1 - \beta)\mu(\Theta)}. \quad (3.6)$$

Deste modo, é suficiente mostrarmos que

$$\frac{(1 - \beta)\mu(\Theta)}{1 - \alpha(1 - \beta)\mu(\Theta)} < \mu(\Theta). \quad (3.7)$$

Com efeito,

$$\begin{aligned} & \frac{(1 - \beta)\mu(\Theta)}{1 - \alpha(1 - \beta)\mu(\Theta)} < \mu(\Theta) \\ \iff & (1 - \beta)\mu(\Theta) < \mu(\Theta) - \alpha(1 - \beta)[\mu(\Theta)]^2 \\ \iff & [-\beta + \alpha(1 - \beta)\mu(\Theta)]\mu(\Theta) < 0 \\ \iff & -\beta + \alpha(1 - \beta)\mu(\Theta) < 0 \\ \iff & \mu(\Theta) < \frac{\beta}{\alpha(1 - \beta)}. \end{aligned}$$

Portanto, se $\mu(\Theta) \in \left(0, \frac{\beta}{\alpha(1 - \beta)}\right)$, juntamente com (3.5), (3.6) e (3.7), segue que

$$\mu F_\beta M_\alpha(\Theta) < \mu(\Theta).$$

Note que $\mu(F_\beta M_\alpha)^t(\Theta) > 0$, para todo $t \in \mathbb{N}$. Se $\mu(\Theta) \in \left(0, \frac{\beta}{\alpha(1 - \beta)}\right)$, então $\mu(F_\beta M_\alpha)^t(\Theta) \in \left(0, \frac{\beta}{\alpha(1 - \beta)}\right)$. Vamos mostrar este resultado usando indução em t .

- Para $t = 1$ temos da parte inicial que: $\mu F_\beta M_\alpha(\Theta) < \mu(\Theta)$. Portanto, $\mu F_\beta M_\alpha(\Theta) \in \left(0, \frac{\beta}{\alpha(1 - \beta)}\right)$.
- Suponha que o resultado é válido para t e vamos mostrar para $t + 1$. Note que,

$$\mu(F_\beta M_\alpha)^{t+1}(\Theta) = \mu(F_\beta M_\alpha)^t(F_\beta M_\alpha)(\Theta).$$

Como por hipótese de indução $\mu(F_\beta M_\alpha)^t(\Theta) \in \left(0, \frac{\beta}{\alpha(1 - \beta)}\right)$, segue como consequência da parte inicial que

$$\mu(F_\beta M_\alpha)^t(F_\beta M_\alpha)(\Theta) < \mu(F_\beta M_\alpha)^t(\Theta).$$

Portanto, $\mu(F_\beta M_\alpha)^{t+1}(\Theta) \in \left(0, \frac{\beta}{\alpha(1 - \beta)}\right)$.

Se $\mu(\Theta) \in \left(0, \frac{\beta}{\alpha(1-\beta)}\right)$, então da parte inicial e do resultado acima temos que a sequência $\mu(\mathbf{F}_\beta \mathbf{M}_\alpha)^t(\Theta)$ é monótona decrescente. Pois,

$$\mu(\mathbf{F}_\beta \mathbf{M}_\alpha)^{t+1}(\Theta) = \mu(\mathbf{F}_\beta \mathbf{M}_\alpha)^t(\mathbf{F}_\beta \mathbf{M}_\alpha)(\Theta) < \mu(\mathbf{F}_\beta \mathbf{M}_\alpha)^t(\Theta).$$

Como a sequência $\mu(\mathbf{F}_\beta \mathbf{M}_\alpha)^t(\Theta)$ é monótona decrescente e limitada inferiormente por zero, segue que $\mu(\mathbf{F}_\beta \mathbf{M}_\alpha)^t(\Theta) \rightarrow 0$, quando $t \rightarrow \infty$. Ou equivalentemente, $\mu(\mathbf{F}_\beta \mathbf{M}_\alpha)^t \rightarrow \delta_\Theta$, quando $t \rightarrow \infty$. \square

Agora vamos provar o Teorema 2.

Prova do Teorema 2. Note que, o caso $\beta = 0$ é impossível, e $h(\beta)$ não está definida para $\beta = 1$, deste modo vamos considerar $\beta \in (0, 1)$. Se existe $t \in \mathbb{N}$ tal que $\mu_t(\Theta) \in \left(0, \frac{\beta}{\alpha(1-\beta)}\right)$, o Teorema 2, segue como consequência imediata do Teorema 3. Se $\beta > \beta^*$, então estamos no caso em que

$$\bullet \alpha < \frac{\beta}{(1-\beta)^2 + (1-\beta)\beta^2}$$

e neste caso, para $t = 1$ temos que $\mu_1(\Theta) \in \left(0, \frac{\beta}{\alpha(1-\beta)}\right)$, pois das definições de \mathbf{F}_β e \mathbf{M}_α ,

$$\begin{aligned} \mu_1(\Theta) &= \delta_\Theta \mathbf{F}_\beta \mathbf{M}_\alpha(\Theta) = \frac{\delta_\Theta \mathbf{F}_\beta(\Theta)}{1 - \alpha \delta_\Theta \mathbf{F}_\beta(\Theta, \Theta)} \\ &= \frac{1 - \beta}{1 - \alpha \delta_\Theta \mathbf{F}_\beta(\Theta) \delta_\Theta \mathbf{F}_\beta(\Theta)} = \frac{1 - \beta}{1 - \alpha \beta (1 - \beta)}. \end{aligned}$$

Além disso,

$$\frac{1 - \beta}{1 - \alpha \beta (1 - \beta)} < \frac{\beta}{\alpha(1 - \beta)} \iff \alpha < \frac{\beta}{(1 - \beta)^2 + (1 - \beta)\beta^2}.$$

Com isso, o teorema está provado.

Por outro lado, suponha que $\beta \leq \beta^*$, então estamos no caso em que

$$\bullet \alpha < 2\beta.$$

Neste caso, suponhamos que $\mu_t(\Theta) \geq \frac{\beta}{\alpha(1-\beta)}$ para todo $t \in \mathbb{N}$. Vamos mostrar que isso é impossível. Note que,

$$\mu_{t+1}(\Theta) = \delta_\Theta(\mathbf{F}_\beta \mathbf{M}_\alpha)^{t+1}(\Theta) = \delta_\Theta(\mathbf{F}_\beta \mathbf{M}_\alpha)^t(\mathbf{F}_\beta \mathbf{M}_\alpha)(\Theta) = \mu_t(\mathbf{F}_\beta \mathbf{M}_\alpha)(\Theta).$$

Da definição de \mathbf{F}_β e \mathbf{M}_α temos que,

$$\mu_{t+1}(\Theta) = \frac{(1 - \beta)\mu_t(\Theta)}{1 - \alpha \mu_t \mathbf{F}_\beta(\Theta, \Theta)}. \quad (3.8)$$

Lembremos que para toda $\mu \in \mathcal{M}_{\{\ominus, \oplus\}}$, $\mu(\oplus, \ominus) \leq 1/2$, logo $\alpha\mu(\oplus, \ominus) \leq \alpha/2$, usando este fato em (3.8) obtemos,

$$\mu_{t+1}(\ominus) \leq \frac{(1-\beta)\mu_t(\ominus)}{1-\alpha/2}.$$

Portanto,

$$\begin{aligned} \mu_{t+1}(\ominus) - \mu_t(\ominus) &\leq \frac{(1-\beta)\mu_t(\ominus)}{1-\alpha/2} - \mu_t(\ominus) \\ &= \frac{(1-\beta)\mu_t(\ominus) - \mu_t(\ominus) + (\alpha/2)\mu_t(\ominus)}{1-\alpha/2} \\ &= \frac{(\alpha/2 - \beta)\mu_t(\ominus)}{1-\alpha/2}. \end{aligned}$$

Esta última igualdade é uma função linear com relação a $\mu_t(\ominus)$. Note que:

- Se $\mu_t(\ominus) = \frac{\beta}{\alpha(1-\beta)}$, então

$$\frac{(\alpha/2 - \beta)\mu_t(\ominus)}{1-\alpha/2} = \frac{(\alpha/2 - \beta)(\beta/\alpha(1-\beta))}{1-\alpha/2} \leq 0, \text{ pois estamos no caso em que } \alpha < 2\beta.$$

- Se $\mu_t(\ominus) = 1$, então

$$\frac{(\alpha/2 - \beta)\mu_t(\ominus)}{1-\alpha/2} = \frac{(\alpha/2 - \beta)}{1-\alpha/2} \leq 0, \text{ pois estamos no caso em que } \alpha < 2\beta.$$

Portanto, se $\alpha < 2\beta$, ambos os valores dos extremos do intervalo onde a função linear $\frac{(\alpha/2 - \beta)\mu_t(\ominus)}{1-\alpha/2}$ está definida são negativos, deste modo,

$$\mu_{t+1}(\ominus) - \mu_t(\ominus) \leq m,$$

onde m é uma constante negativa. Sendo assim, podemos observar que

$$\mu_t(\ominus) - \mu_0(\ominus) = \mu_t(\ominus) - \mu_{t-1}(\ominus) + \mu_{t-1}(\ominus) - \mu_{t-2}(\ominus) + \cdots + \mu_1(\ominus) - \mu_0(\ominus) \leq tm,$$

logo, $\mu_t(\ominus) \leq tm + 1$, onde m é uma constante negativa, portanto, $\mu_t(\ominus) \rightarrow -\infty$, quando $t \rightarrow \infty$, o que é impossível porque a probabilidade não pode ser negativa. Portanto, o Teorema 2 está provado. \square

Suponhamos provado o Teorema 1 e vamos provar o Teorema 4.

Prova do Teorema 4. Note que $s(\beta, \alpha)$ não pode tomar valores no intervalo $(1 - (\beta/\alpha(1-\beta)), 1)$, porque se $s(\beta, \alpha)$ tomasse valor nesse intervalo, então existiria um número natural t tal que $\mu_t(\oplus) > 1 - (\beta/\alpha(1-\beta))$, o que é equivalente a $\mu_t(\ominus) < \beta/\alpha(1-\beta)$. Então, devido ao Teorema 3, $\mu_t(\oplus)$ tende a 1, quando $t \rightarrow \infty$, daí, $s(\beta, \alpha) = 1$.

Note que, inicialmente, estamos considerando $\alpha, \beta \in (0, 1)$, todavia o Teorema 2

apenas faz sentido quando $h(\beta) < 1$. Mas $h(\beta) = 1$ e $\beta \in (0, 1)$ se, e somente se, $\beta = 1/3 \left(2 - 5\sqrt[3]{2/(3\sqrt{69} - 11)} + \sqrt[3]{(3\sqrt{69} - 11)/2} \right) \cong 0.4301$. Portanto, uma vez que $h(\beta)$ é crescente, podemos considerar para a validade deste teorema que $\beta < 1/3 \left(2 - 5\sqrt[3]{2/(3\sqrt{69} - 11)} + \sqrt[3]{(3\sqrt{69} - 11)/2} \right)$. Assim, para qualquer $\alpha \in (\beta/(1 - \beta), 1)$, $s(\beta, \alpha)$

- é igual a 1 se $\alpha < h(\beta)$, devido ao Teorema 2;
- tende a 0, quando $\beta \rightarrow 0$, devido ao Teorema 1;
- não pode assumir valores no intervalo $(1 - (\beta/\alpha(1 - \beta)), 1)$, devido ao Teorema 3.

Portanto, $s(\beta, \alpha)$ não pode ser contínua como função de β . \square

3.3 O processo ν e sua representação gráfica

Nossa prova do Teorema 1 é baseada em duas ideias bem conhecidas na literatura: o método do contorno de Peierls e a dualidade dos grafos planares. Para mais detalhes sobre esses métodos, ver (GRIMMETT, 1999). Vamos introduzir um processo ν , que difere do nosso processo original no seguinte sentido: não necessitamos ocultar as partículas mortas em cada passo de tempo. Podemos deixá-las onde elas estão, mas, neste caso, temos que sacrificar a localidade, a saber, temos que organizar a interação das partículas vivas como se as partículas assassinadas estivessem escondidas. Denotemos por $x \in \mathbb{Z}$ as coordenadas do espaço. Usaremos também um parâmetro natural y , que é igual a zero no começo e cresce de um em um após toda aplicação de F_β ou M_α . Assim, y cresce para dois quando t na equação (3.4) cresce para um. Denotemos por $F(x, t)$ e $A(x, t)$ e chamamos de *variáveis básicas* aquelas variáveis F_i e A_i , que participaram na t -ésima aplicação de $F_\beta M_\alpha$. Portanto, nosso espaço básico é

$$\bar{\Omega} = (\{\text{mover, ficar}\} \times \{\text{atacar, parar}\})^{\mathbb{Z} \cdot \mathbb{Z}_+}$$

com coordenadas

$$(F(x, t), A(x, t)), \quad \text{onde } x \in \mathbb{Z}, t \in \mathbb{Z}_+$$

e com uma medida produto π , de acordo com o que para todos x, t

$$F(x, t) = \begin{cases} \text{mover,} & \text{com probabilidade } \beta, \\ \text{ficar,} & \text{com probabilidade } 1 - \beta, \end{cases} \tag{3.9}$$

$$A(x, t) = \begin{cases} \text{atacar,} & \text{com probabilidade } \alpha, \\ \text{parar,} & \text{com probabilidade } 1 - \alpha. \end{cases}$$

Denotemos por

$$V = \{(x, y), x \in \mathbb{Z}, y \in \mathbb{Z}_+\}.$$

O conjunto dos pares $(x, y) \in V$ com um dado y , são chamados de y -níveis ou apenas níveis. Todo par $(x, y) \in V$ tem um estado denotado por $\text{estado}(x, y)$, que é igual a \ominus , \oplus ou \odot e todos esses estados são funções dos $\omega \in \bar{\Omega}$ definidos da seguinte maneira indutiva.

Base de Indução. $\text{estado}(x, 0) = \ominus$, para todo $x \in \mathbb{Z}$.

Passo de indução quando y é par, isto é, $y = 2t$, onde $t \in \mathbb{Z}_+$ (imitando a ação de F_β). Para todo $x \in \mathbb{Z}$:

$$\text{estado}(x, 2t + 1) = \begin{cases} \oplus, & \text{se } \text{estado}(x, 2t) = \ominus \text{ e } F(x, t) = \text{mover}, \\ \text{estado}(x, 2t), & \text{em todos os outros casos.} \end{cases}$$

Passo de indução quando y é ímpar, isto é, $y = 2t + 1$ onde $t \in \mathbb{Z}_+$ (imitando a ação de M_α , mas sem a localidade). Para todo $x \in \mathbb{Z}$:

$$\text{estado}(x, 2t+2) = \begin{cases} \odot, & \text{se } \text{estado}(x, 2t + 1) = \oplus \text{ e existe } x' > x \text{ tal que} \\ & \text{estado}(x', 2t + 1) = \ominus \text{ e } A(x', t) = \text{atacar e para todo} \\ & x'' \in \mathbb{Z} : x < x'' < x' \Rightarrow \text{estado}(x'', 2t + 1) = \odot; \\ \text{estado}(x, 2t + 1), & \text{em todos os outros casos.} \end{cases}$$

Falando informalmente, neste processo as partículas nunca desaparecem e permanecem com os mesmos índices inteiros que elas tinham no início. Se uma partícula é assassinada, ela vai para o estado morto, \odot , e permanece neste estado para sempre. Partículas vivas interagem como se as partículas mortas não existissem. Deste modo, temos uma aplicação definida indutivamente de $\bar{\Omega}$ em $\{\ominus, \oplus, \odot\}^V$. Denotemos por ν a medida em $\{\ominus, \oplus, \odot\}^V$ induzida pela distribuição π das variáveis básicas (3.9) com esta aplicação e ν_y a distribuição dos estados no y -ésimo nível. O processo ν é útil para nós porque

$$\nu_{2t}\text{Hide} = \mu_t \tag{3.10}$$

para todo t .

De fato, vamos mostrar por indução.

- Para $t = 0$, temos que, por definição do processo ν , $\text{estado}(x, 0) = \ominus$, para todo $x \in \mathbb{Z}$, logo ν_0 é a configuração todos menos, portanto imune a ação do operador **Hide**, além disso, $\mu_0 = \delta_\ominus$, portanto,

$$\nu_0\text{Hide} = \mu_0.$$

- Suponha que o resultado é válido para t , isto é, $\nu_{2t}\text{Hide} = \mu_t$ e vamos mostrar para $t + 1$.

$$\mu_{t+1} = \delta_\ominus(\mathbf{F}_\beta \mathbf{M}_\alpha)^{t+1} = \delta_\ominus(\mathbf{F}_\beta \mathbf{M}_\alpha)^t(\mathbf{F}_\beta \mathbf{M}_\alpha) = \mu_t(\mathbf{F}_\beta \mathbf{M}_\alpha).$$

Como por hipótese de indução $\mu_t = \nu_{2t}\text{Hide}$, então

$$\begin{aligned}\mu_t(\mathbf{F}_\beta \mathbf{M}_\alpha) &= \nu_{2t}\text{Hide}(\mathbf{F}_\beta \mathbf{M}_\alpha) = \nu_{2t}\text{Hide}(\mathbf{F}_\beta \text{Trap}_\alpha \text{Hide}) \\ &= \nu_{2t}\text{Hide}(\nu_2\text{Hide}) = \nu_{2t}\nu_2\text{Hide} \\ &= \nu_{2t+2}\text{Hide} = \nu_{2(t+1)}\text{Hide}.\end{aligned}$$

Pode-se verificar também que

$$\left\{ \begin{array}{l} \text{(a)} \quad \nu(\text{estado}(x, y) = \ominus) > 0 \text{ para todo } (x, y) \in V. \\ \text{(b)} \quad \text{Para qualquer inteiro } x_0 \text{ e qualquer natural } y \\ \quad \nu(\forall x \geq x_0 : \text{estado}(x, y) \neq \ominus) = \nu(\forall x \leq x_0 : \text{estado}(x, y) \neq \ominus) = 0. \\ \text{(c)} \quad \mu_t(\ominus) > 0 \text{ para todo natural } t. \\ \text{(d)} \quad \text{Para qualquer inteiro } x_0 \text{ e qualquer natural } y \\ \quad \mu_t(\forall x \geq x_0 : s_x \neq \ominus) = \mu_t(\forall x \leq x_0 : s_x \neq \ominus) = 0. \end{array} \right. \quad (3.11)$$

Agora vamos partir para a representação gráfica do processo ν . No desenvolvimento do texto devemos ignorar alguns eventos, cuja probabilidade é zero. Assim, lendo eles, devemos mentalmente inserir “quase”, “quase todos”, “quase certamente”, sempre que necessário. Para qualquer $\omega \in \bar{\Omega}$ definimos um grafo G . Além de descrever o grafo G , devemos descrever como desenhá-lo no plano, representando vértices por pontos e arestas por curvas (na verdade, seguimentos de reta). O conjunto dos vértices de G é

$$V_G = \{(x, y) \in V, \text{estado}(x, y) \neq \odot\}$$

e todo vértice (x, y) é colocado em um ponto (x, y) do plano, onde x e y são as coordenadas ortogonais usuais, o eixo x é horizontal e o eixo y é vertical. O grafo G tem dois tipos de arestas, que chamamos *arestas verticais* e *arestas horizontais*. Vamos descrevê-las.

Arestas verticais. Quaisquer dois vértices (x, y_1) e (x, y_2) de G , onde $y_2 - y_1 = 1$, são conectados com uma aresta vertical. A direção desta aresta de (x, y_1) para (x, y_2) é chamada *norte*, a outra direção é chamada *sul*. Chamamos (x, y_1) de vizinho sul de (x, y_2) e (x, y_2) de vizinho norte de (x, y_1) .

Arestas horizontais. Quaisquer dois vértices (x_1, y) e (x_2, y) de G , onde $x_1 < x_2$, estão conectados por uma aresta horizontal se

$$\forall x \in \mathbb{Z} : x_1 < x < x_2 \implies \text{estado}(x, y) = \odot.$$

A direção desta aresta de (x_1, y) para (x_2, y) é chamada *leste*, a outra direção é chamada *oeste*. Chamamos (x_1, y) de vizinho oeste de (x_2, y) e (x_2, y) de vizinho leste de (x_1, y) . Assim, G , que tem apenas estas arestas, que estão especificadas acima, está definido. Estas arestas são representadas por segmentos de reta conectando os pontos representando o fim das arestas.

Um vértice de G , cujo nível y é par, sempre tem exatamente um vizinho oeste, exatamente um vizinho leste e exatamente um vizinho norte. Também tem exatamente um vizinho sul, exceto o caso $y = 0$, quando ele não tem vizinho sul. Um vértice de G , cujo nível é ímpar, sempre tem exatamente um vizinho oeste, exatamente um vizinho leste e exatamente um vizinho sul. Além disso, ele pode ter um vizinho norte ou não. Devido a definição de G , todo vértice dele deve estar no estado \oplus ou \ominus ; no primeiro caso, o chamamos de um \oplus -vértice e no segundo caso de um \ominus -vértice.

É evidente que diferentes arestas de G não se interceptam, exceto nas extremidades em comum. Devemos chamar a *figura* de G sua representação no plano. Esta figura corta o plano em partes, que chamamos de *faces*. Assumimos que todas as faces são fechadas. Dizemos que duas faces são vizinhas se elas tem uma aresta em comum. Nossa figura de G tem exatamente uma face ilimitada, a saber, a metade inferior do plano. Todas as outras faces de G são limitadas e a chamamos de *caixas*. Toda caixa tem a forma de um retângulo, intercalada entre duas linhas paralelas nos níveis y_1 e $y_1 + 1$, onde y_1 é natural; assim, podemos denotá-las por

$$\{(x, y) \in \mathbb{R}^2 : x_1 \leq x \leq x_2, y_1 \leq y \leq y_1 + 1\}. \quad (3.12)$$

Para todo natural y_1 , as caixas intercaladas entre as linhas paralelas nos níveis y_1 e $y_1 + 1$ formam uma sequência bi-infinita, em que todos dois termos próximos tem um lado em comum, e que chamamos de *corredor horizontal* no sub-nível $y_1 + 1$. Qualquer caixa tem pelo menos quatro vértices posicionados em suas quinas e não tem mais vértices em suas paredes oeste, leste e norte, assim elas tem exatamente um vizinho oeste, um vizinho leste e um vizinho norte. Se y_1 é par, a caixa (3.12) não tem mais vértices em sua parede sul. Se y_1 é ímpar, temos dois casos a considerar: se houve um assassinato entre os sítios x_1 e x_2 do nível y_1 para o nível $y_1 + 1$, então a caixa (3.12) tem dois vizinhos sul. Por outro lado, se não houve aplicação do operador M_α entre os sítios x_1 e x_2 , então a caixa (3.12) tem exatamente um vizinho sul.

Como em (DE LIMA *et al.*, 2008; GRIMMETT, 1999), usaremos a bem conhecida dualidade das figuras dos grafos. Vamos descrever um grafo, que denotaremos por \overline{G} , e sua figura, que será dual da figura de G . Colocaremos um vértice de \overline{G} , que é dual da caixa (3.12), no ponto

$$\left(x_2 - \frac{1}{2}, y_1 + 1 - \varepsilon\right), \quad (3.13)$$

onde $\varepsilon > 0$ é escolhido de forma diferente para diferentes caixas, mas deve ser suficientemente pequeno em cada caso. Dizemos que o vértice (3.13) está no subnível $y_1 + 1$. Dizemos que ele está em um subnível par, se $y_1 + 1$ é par, e está em um subnível ímpar se $y_1 + 1$ é ímpar. Existe apenas uma sutileza: que o vértice de \overline{G} , que é dual da única face ilimitada da figura de G , é colocada “infinitamente longe” na direção negativa do eixo y e as arestas que vão até ele são raios na mesma direção. Todas as outras arestas de \overline{G} são segmentos de reta conectando os pontos representando suas extremidades. Assim, o grafo

\overline{G} e sua figura estão definidos. É fácil ver que, para qualquer caixa, o ε correspondente pode ser escolhido tão pequeno que a condição usual da figura dual é realizada.

Devemos chamar de *horizontais* aquelas arestas de \overline{G} , que são dual das arestas verticais de G e *verticais* aquelas arestas de \overline{G} , que são dual das arestas horizontais de G . Note que arestas horizontais de \overline{G} são aproximadamente horizontais porque os valores de ε para todos os vértices de \overline{G} são aproximadamente iguais a zero. Para qualquer natural y , os vértices de \overline{G} que estão no subnível $y + 1$ e as arestas horizontais conectando eles formam um caminho bi-infinito, que chamamos *caminho horizontal* no subnível $y + 1$ e que é dual do subcorredor $y + 1$. Qualquer face limitada de \overline{G} está intercalada entre caminhos horizontais nos subníveis y e $y + 1$. As faces ilimitadas de \overline{G} são duais dos vértices de G no nível zero. Elas são meias-tiras ilimitadas, que preenchem toda a metade do plano abaixo do caminho horizontal no subnível 1. Note que, na Figura 1, todas as faces de \overline{G} , do nível 1 em diante, tem uma parede sul, porém as faces de \overline{G} no nível 0, têm apenas as paredes oeste, norte e leste. Por isso, as chamamos de meias-tiras. Além disso, as paredes nortes dessas faces no nível 0, são as que formam o subcorredor horizontal no subnível 1. Uma face de \overline{G} é chamada de vizinho oeste (respectivamente, leste, norte e sul) de outra face de \overline{G} se seus vértices correspondentes de G têm a mesma relação.

De acordo com o que dissemos sobre os vértices de G em um nível par, qualquer face de \overline{G} em um nível par tem exatamente um vizinho oeste, exatamente um vizinho leste e exatamente um vizinho norte. Também tem exatamente um vizinho sul, exceto o caso $y = 0$, quando ela não tem vizinho sul. Sempre que $y > 0$, chamamos estas faces de \overline{G} de *retângulos*. De fato, todas elas são aproximadamente retângulos. De acordo com o que dissemos sobre os vértices de G em um nível ímpar, qualquer face de \overline{G} em um nível ímpar tem no máximo um vizinho norte. Se ela tem um vizinho norte, chamamos esta face de *trapézio*, caso contrário, se a face não tem nenhum vizinho norte, a chamamos de *triângulo*. De fato, estas faces são aproximadamente trapézios e triângulos.

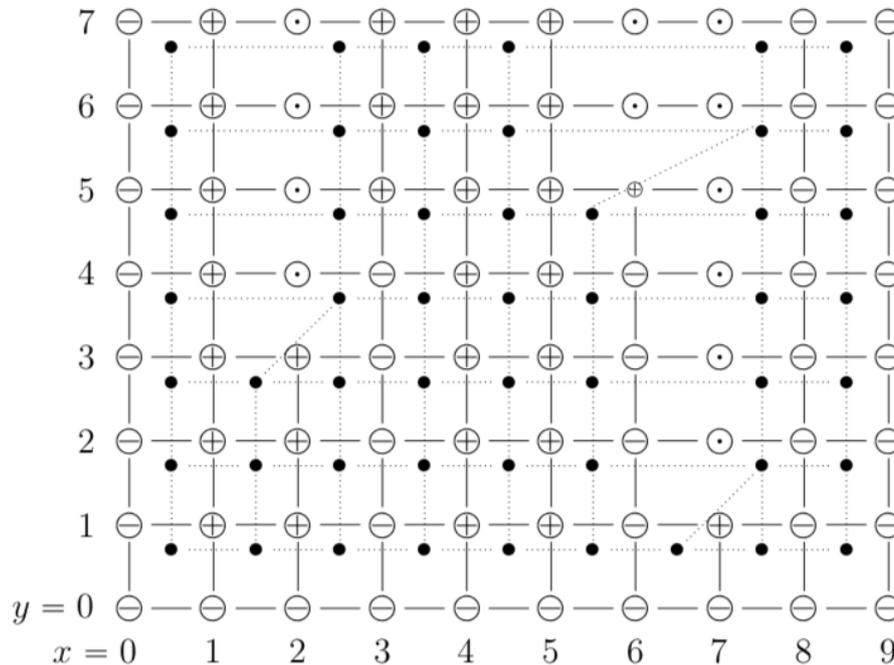
Na Figura 1, apresentamos parte de um grafo G juntamente com seu grafo dual \overline{G} .

3.4 Uma cadeia de igualdades e desigualdades

Inicialmente, fixemos um número natural, porém arbitrário, T . Nosso objetivo geral é estimar $\mu_T(\oplus)$ uniformemente em T . De (3.11), $\mu_T(\ominus)$ é positivo, deste modo, a fração $\frac{\mu_T(\oplus)}{\mu_T(\ominus)}$ faz sentido e é suficiente estimar esta fração. Para reduzir nossa tarefa um pouco mais, vamos provar que

$$\mu_T(\oplus) = \sum_{k=1}^{\infty} \mu_T(\ominus, \oplus^k). \quad (3.14)$$

De fato, vamos considerar o evento da presença de um mais em um certo site e cortá-lo em pedaços de acordo com o número de mais no lado esquerdo deste site. Pelo item (d) de (3.11), temos que este número é finito quase certamente, daí obtemos (3.14).

Figura 1 – Ilustração de parte de um grafo G e seu grafo dual \bar{G} .

Fonte: O autor (2020).

Então, de (3.14),

$$\mu_T(\oplus) \leq \frac{\mu_T(\oplus)}{\mu_T(\ominus)} = \sum_{k=1}^{\infty} \frac{\mu_T(\ominus, \oplus^k)}{\mu_T(\ominus)}. \quad (3.15)$$

Para reduzir nossa tarefa ainda mais, vamos concentrar nossa atenção em Ω_0 , o conjunto dos $\omega \in \bar{\Omega}$ tal que $\text{estado}(0, 2T) = \ominus$. Para qualquer $\omega \in \Omega_0$ denotamos por $x_{\max}(\omega)$ o menor x positivo tal que $\text{estado}(x, 2T) = \ominus$. Devido ao item (b) de (3.11), $x_{\max}(\omega)$ existe quase certamente pois, fixado um nível y , a probabilidade de que todas as componentes estejam no estado diferente de menos é zero. Vamos chamar de *flores* todos aqueles pares $(x, 2T)$, onde $0 < x < x_{\max}(\omega)$, para os quais $\text{estado}(x, 2T) = \oplus$ (Na Figura 2, as flores, no nível $2T = 6$, são indicadas com a letra “F”). Denotemos por $\phi(\omega)$ o número de flores. Uma vez que $x_{\max}(\omega)$ existe quase certamente, $\phi(\omega)$ é finito quase certamente. Para qualquer $k = 1, 2, 3, \dots$, denotemos por Ω_k o conjunto dos $\omega \in \Omega_0$ para os quais $\phi(\omega) \geq k$. Note que $\Omega_0 \supseteq \Omega_1 \supseteq \Omega_2 \supseteq \dots$. Vamos provar que, para todo k ,

$$\frac{\pi(\Omega_k)}{\pi(\Omega_0)} = \frac{\mu_T(\ominus, \oplus^k)}{\mu_T(\ominus)}. \quad (3.16)$$

Fixado um número natural arbitrário T , calcular $\pi(\Omega_0)$ significa calcular a probabilidade de encontrarmos uma partícula no estado menos na posição $(0, 2T)$. Como a medida ν é induzida pela medida π , então, calcular $\pi(\Omega_0)$ é o mesmo que calcular $\nu_{2T}(x_0 = \ominus)$, ou seja, é o mesmo que calcular a probabilidade de encontrarmos uma partícula no estado

menos na posição x_0 no nível $2T$ pela medida ν , e, como esta é uniforme, segue que $\nu_{2T}(x_0 = \ominus) = \nu_{2T}(\ominus)$. Portanto, $\pi(\Omega_0) = \nu_{2T}(\ominus)$. De (3.10) e (3.3)

$$\mu_T(\ominus) = \nu_{2T} \text{Hide}(\ominus) = \frac{\nu_{2T}(\ominus)}{1 - \nu_{2T}(\odot)},$$

daí,

$$\pi(\Omega_0) = \nu_{2T}(\ominus) = \mu_T(\ominus)(1 - \nu_{2T}(\odot)). \quad (3.17)$$

Por outro lado, Ω_k é o conjunto daqueles $\omega \in \Omega_0$, para os quais a configuração no nível $2T$ contém uma das palavras

$$\ominus \odot^{n_1} \oplus \odot^{n_2} \dots \oplus \odot^{n_{k-1}} \oplus \odot^{n_k} \oplus$$

começando na componente em que $x = 0$, isto é, $\text{estado}(0, 2T) = \ominus$. Deste modo, calcular $\pi(\Omega_k)$ significa calcular a probabilidade de se encontrar uma das palavras acima no nível $2T$ pela medida ν , pois esta é uniforme. Logo,

$$\pi(\Omega_k) = \sum_{n_1, \dots, n_k=0}^{\infty} \nu_{2T}(\ominus \odot^{n_1} \oplus \dots \oplus \odot^{n_k} \oplus).$$

Mas, de (3.10) e (3.2) resulta,

$$\begin{aligned} \mu_T(\ominus, \oplus^k) &= \nu_{2T} \text{Hide}(\ominus, \oplus^k) \\ &= \frac{1}{1 - \nu_{2T}(\odot)} \sum_{n_1, \dots, n_k=0}^{\infty} \nu_{2T}(\ominus \odot^{n_1} \oplus \dots \oplus \odot^{n_k} \oplus). \end{aligned}$$

Logo,

$$\mu_T(\ominus, \oplus^k) = \frac{\pi(\Omega_k)}{1 - \nu_{2T}(\odot)}.$$

Assim,

$$\pi(\Omega_k) = \mu_T(\ominus, \oplus^k)(1 - \nu_{2T}(\odot)).$$

Dividindo este resultado por (3.17), obtemos

$$\frac{\pi(\Omega_k)}{\pi(\Omega_0)} = \frac{\mu_T(\ominus, \oplus^k)}{\mu_T(\ominus)}.$$

Somando essa equação para todo $k \geq 1$,

$$\sum_{k=1}^{\infty} \frac{\pi(\Omega_k)}{\pi(\Omega_0)} = \sum_{k=1}^{\infty} \frac{\mu_T(\ominus, \oplus^k)}{\mu_T(\ominus)},$$

e usando (3.15) obtemos

$$\frac{\mu_T(\oplus)}{\mu_T(\ominus)} = \sum_{k=1}^{\infty} \frac{\pi(\Omega_k)}{\pi(\Omega_0)}. \quad (3.18)$$

Lembrando da sabedoria de um velho jardineiro: não existem flores sem raízes. Tomemos qualquer $\omega \in \Omega_1$ e vamos chamar um caminho em G de *noroeste* se todos os seus passos vão para norte ou oeste. Chamaremos um vértice de G de *raíz* se existe um caminho noroeste deste vértice para alguma flor, e todos os vértices deste caminho devem estar no estado \oplus . Em particular, todas as flores são raízes. Os vértices de G que não são raízes, são chamados de *não raízes*. Como T é fixo, $x_{\max}(\omega)$ existe quase certamente. Logo, o conjunto das flores é finito quase certamente e, portanto, o conjunto das raízes é finito quase certamente. Nossa estimação é baseada na construção de um “contorno” ao redor de todas as raízes. Dizemos que um conjunto S de vértices de um grafo é *conectado* neste grafo se para quaisquer dois elementos deste conjunto existe um caminho neste grafo conectando eles, em que todos os vértices pertencem a S .

Lema 1. *Para qualquer $\omega \in \Omega_1$*

- (a) *O conjunto das raízes é não vazio, finito e conectado em G ;*
- (b) *O conjunto das não raízes é infinito e conectado em G .*

Prova de (a). Como $\omega \in \Omega_1$ temos que o número de flores é maior ou igual a 1, e como as flores são raízes, concluímos que o conjunto das raízes é não vazio. Além disso, como $\Omega_1 \subset \Omega_0$, então para todo $\omega \in \Omega_1$, temos que o número de flores é finito, digamos $\phi(\omega)$. Então para um determinado T fixado, existem no máximo $2T(x_{\max}(\omega) - 1)$ raízes. Portanto, o conjunto das raízes é finito. Finalmente, vejamos que o conjunto das raízes é conectado em G . Com efeito, sejam (x_1, y_1) e (x_2, y_2) duas raízes, então existe um caminho noroeste ligando (x_1, y_1) para alguma flor de modo que todos os vértices deste caminho estão no estado mais, portanto todos são raízes. Analogamente, existe um caminho noroeste ligando (x_2, y_2) para alguma flor (não necessariamente a mesma flor do caso anterior), em que todos os vértices deste caminho estão no estado mais e, portanto, são raízes. Como as duas flores estão no mesmo nível $2T$, existe um caminho oeste, conseqüentemente noroeste, ligando essas flores em que todos os vértices deste caminho são também flores, portanto raízes. Logo, fazendo a justaposição destes três caminhos, obtemos um caminho que liga (x_1, y_1) a (x_2, y_2) em que todos os vértices são raízes. Portanto, o conjunto das raízes é conectado em G .

Prova de (b). Como provado no item (a), o conjunto das raízes é finito, portanto o conjunto das não raízes é infinito. Resta-nos mostrar que o conjunto das não raízes é conectado em G . Para isto, denotemos por S_{in} o conjunto dos (x, y) para os quais $0 < x < x_{\max}(\omega)$ e $0 < y < 2T$. Denotemos por S_{out} o conjunto dos vértices de G que não pertencem a S_{in} . Pelas próprias definições vemos que todas as raízes pertencem a S_{in} e que S_{out} é conectado em G . Deste modo, precisamos mostrar que qualquer não raíz em S_{in} está conectada com S_{out} por um caminho sem raízes. Seja $(x_0, y_0) \in S_{\text{in}}$ uma

não raíz qualquer e defina o conjunto S_0 da seguinte maneira: um vértice (x, y) pertence a S_0 se existe um caminho em G conectando (x, y) com (x_0, y_0) em que todos os vértices são não raízes. Denotemos por t_0 o nível minimal dos elementos de S_0 . Isto é possível pois todo vértice $(x, y) \in G$ é tal que $x \in \mathbb{Z}$ e $y \in \mathbb{Z}_+$. Agora, vamos olhar para o elemento mais ocidental de S_0 que está no nível t_0 . Se este elemento não existe, então (x_0, y_0) está conectado com S_{out} por um caminho sem raízes. Vamos assumir que este elemento existe. Se seu estado é \oplus , então ele é uma raíz, porque seu vizinho oeste é uma raíz. Se seu estado é \ominus , então o estado de seu vizinho sul também é menos, logo este é uma não raíz, o que contradiz nossa escolha de t_0 . Em ambos os casos, temos uma contradição. \square

Vamos chamar de *raíz dual* aquelas faces de \overline{G} , que são dual das raízes, e denotar por U a união das raízes duais. Uma vez que toda raíz dual é limitada, U também é limitado e fechado uma vez que assumimos que todas as faces são fechadas. Daí, como no Lema 1 provamos que o conjunto das raízes é finito, segue que U é homeomorfo a um disco fechado. Então a fronteira de U é uma curva fechada que inclui o lado leste do retângulo dual do vértice $(0, 2T)$. Assim, esta curva fechada inclui V_0 , a extremidade norte deste lado, e podemos assumir que ela começa e termina em V_0 e contorna U na direção anti-horária. Esta curva pode ser representada como um caminho em \overline{G} , que denotaremos por $\text{tour}(\omega)$, pois ele é determinado por ω . Na Figura 2 ilustramos um tal caminho. Agora precisamos classificar todas as possíveis formas de $\text{tour}(\omega)$. Para isto, precisamos classificar todos os passos que $\text{tour}(\omega)$ pode incluir, isto é, alguns passos em \overline{G} . Devemos começar classificando alguns passos no grafo original G . Vamos chamar de *tipos* os elementos do conjunto

$$\{1, 1', 2, 2', 2'', 3, 4, 4', 5\}.$$

Tabela 1 – Definição de tipos, eventos associados e variáveis associadas a passos em G começando em um \oplus -vértice.

Passo em G começando em um \oplus -vértice	Tipo	Evento associado	Variável associada
Passo oeste em um nível par	1	trivial	nenhuma
Passo oeste em um nível ímpar	1'	trivial	nenhuma
Passo de $(x, 2t + 1)$ para $(x, 2t)$ se $F(x, t) = \text{mover}$	2	$F(x, t) = \text{mover}$	$F(x, t)$
Passo de $(x, 2t + 1)$ para $(x, 2t)$ se $F(x, t) = \text{ficar}$	2'	$F(x, t) = \text{ficar}$	$F(x, t)$
Passo sul de um nível par para um nível ímpar	2''	trivial	nenhuma
Passo de $(x, 2t + 1)$ para seu vizinho leste se $A(x, t) = \text{atacar}$	3	$A(x, t) = \text{atacar}$	$A(x, t)$
Passo de $(x, 2t + 1)$ para seu vizinho leste se $A(x, t) = \text{parar}$	4	$A(x, t) = \text{parar}$	$A(x, t)$
Passo leste em nível par	4'	trivial	nenhuma
Passo norte	5	trivial	nenhuma

Fonte: O autor (2020).

Devemos atribuir tipos a aqueles, e apenas aqueles, passos em G que começam em um \oplus -vértice. Os casos que podem ocorrer estão listados na Tabela 1.

Os passos que tem a palavra “trivial” na terceira coluna são chamados *trivial*, os outros passos são chamados de *não trivial*. Para todo passo em G , que tem um tipo, atribuímos a eles um *evento associado*. Para todo passo trivial o evento associado é $\bar{\Omega}$ e é também chamado trivial. Os eventos não trivial são representados na Tabela 1 por suas condições. Para todo passo não trivial também definimos uma *variável básica associada*, que é mostrada na última coluna. Além disso, todo passo em G , que tem um tipo, tem uma *chance*. Por razões tipográficas as chances serão mostradas na próxima tabela, mas podemos facilmente inserí-las à direita agora, porque as chances sempre são iguais a probabilidade do evento associado. Devemos usar a mesma correspondência biunívoca entre os passos em G e passos em \bar{G} como foi definido em (TOOM, 1968) e com mais detalhes em (DE LIMA *et al.*, 2008). Aqui é:

$$\left. \begin{array}{l} \text{Se uma aresta } \bar{e} \text{ de } \bar{G} \text{ é dual de uma aresta } e \text{ de } G, \text{ então} \\ \text{para cada direção de } e \text{ a direção dual de } \bar{e} \text{ é a direção da} \\ \textbf{direita} \text{ para } \textbf{esquerda} \text{ quando vamos ao longo de } e \text{ na} \\ \text{direção dada.} \end{array} \right\} \quad (3.19)$$

Tabela 2 – Definição de tipos, chances e deslocamentos de passos em \bar{G} tendo uma \oplus -face em seu lado esquerdo.

Passo em \bar{G} tendo uma \oplus -face em seu lado esquerdo	tipo	chance	deslocamento
Passo sul através de um nível par	1	1	(0, -1)
Passo sul através de um nível ímpar	1'	1	(0, -1)
“mover” passo leste em um subnível ímpar	2	β	(1, 0)
“ficar” passo leste em um subnível ímpar	2'	$1 - \beta$	(1, 0)
Passo leste em um subnível par	2''	1	(1, 0)
“atacar” passo norte passando por um nível ímpar	3	α	(0, 1)
“parar” passo norte passando por um nível ímpar	4	$1 - \alpha$	(0, 1)
Passo norte através de um nível par	4'	1	(0, 1)
Passo oeste	5	1	(-1, 0)

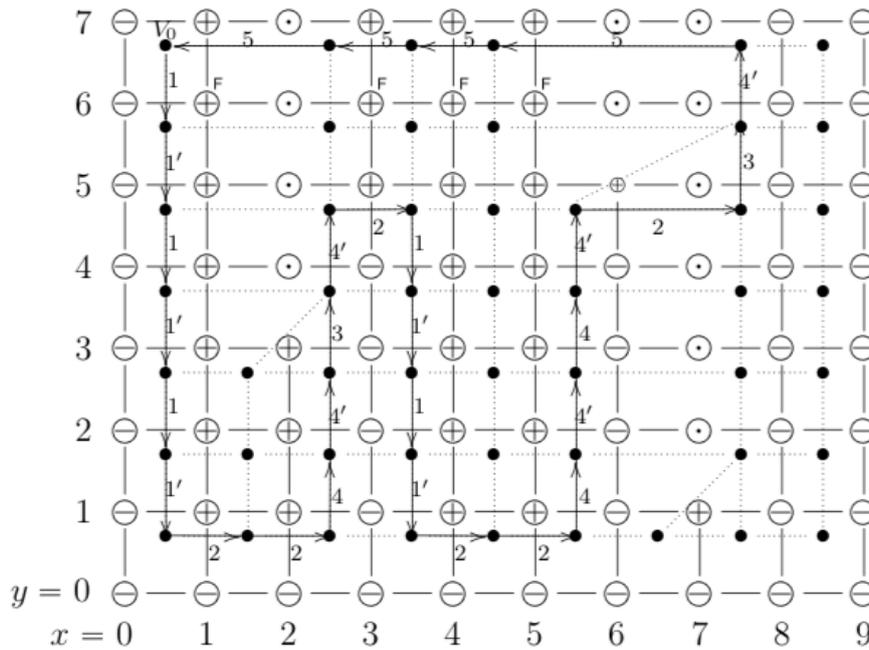
Fonte: O autor (2020).

Tipos, eventos e chances, atribuídos a um passo em G , são atribuídos a seus passos dual em \bar{G} também. Uma vez que um passo em G tem um tipo se, e somente se, ele começa em um \oplus -vértice, um passo em \bar{G} tem um tipo se, e somente se, ele tem uma \oplus -face em seu lado esquerdo.

Podemos imaginar as Tabelas 1 e 2 como uma única tabela, que foi cortada em duas partes por razões tipográficas. A última coluna da Tabela 2 mostra os *deslocamentos* definidos para todos os tipos. O deslocamento é um vetor bidimensional, cujas componentes

são chamadas de HD e VD (abreviações para deslocamento horizontal e deslocamento vertical). A primeira coluna da Tabela 2 é formalmente redundante porque decorre do que foi dito na primeira coluna da Tabela 1; no entanto, ela ajuda a entender porque os deslocamentos são definidos desta maneira. As chances mostradas na terceira coluna são iguais as probabilidades dos eventos na tabela anterior.

Figura 2 – Ilustração do contorno $\text{tour}(\omega)$.



Fonte: O autor (2020).

Lema 2. Para qualquer $\omega \in \Omega_1$ temos:

- (a) Todos os passos do caminho $\text{tour}(\omega)$ tem tipos;
- (b) O caminho $\text{tour}(\omega)$ é uma concatenação de dois caminhos, que denotamos por $\text{bag}(\omega)$ e $\text{lid}(\omega)$, com as seguintes propriedades: todos os passos de $\text{bag}(\omega)$ tem tipos diferentes de 5; $\text{lid}(\omega)$ tem $\phi(\omega)$ passos, todos os quais tem tipo 5.

Prova de (a). Uma vez que todo passo em $\text{tour}(\omega)$ tem uma \oplus -face em seu lado esquerdo, ele tem um tipo.

Prova de (b). Vamos chamar de *flor dual* aquelas faces de \overline{G} que são dual das flores. Note que todas elas são retângulos. Vamos fazer o contorno de U na direção horária começando em V_0 . Então, nossos primeiros $\phi(\omega)$ passos passam no lado norte das flores dual na direção leste. Portanto, os últimos $\phi(\omega)$ passos de $\text{tour}(\omega)$ estão do lado norte das flores dual passando na direção oeste. Vamos chamar este caminho com $\phi(\omega)$ passos de $\text{lid}(\omega)$, e vamos verificar que todos os seus tipos são 5. Note que se um passo em $\text{tour}(\omega)$ tem tipo 5, ele tem uma flor dual em seu lado esquerdo. De fato, se um passo em $\text{tour}(\omega)$

tem tipo 5, ele tem uma raiz em seu lado esquerdo, que é sul, e uma não raiz em seu lado direito, que é norte. Uma vez que F_β não transforma mais em menos, isso só é possível se estes passos tem uma flor em seu lado esquerdo. Logo, os passos de $\text{tour}(\omega)$ que tem tipo 5 tem uma flor dual em seu lado esquerdo. Note também que se um passo de $\text{lid}(\omega)$ tem tipo 5, então todos os passos seguintes a ele, caso existam, também tem tipo 5. De fato, se um passo de $\text{lid}(\omega)$ tem tipo 5, então ele tem uma flor dual em seu lado esquerdo. Mas ele está em um subnível $(2t + 1)$, daí o próximo passo, caso exista, tem que estar na mesma direção, que é oeste, e no mesmo nível, isto é, ele também está no lado norte de uma flor dual, assim seu tipo é 5. Portanto $\text{lid}(\omega)$ tem $\phi(\omega)$ passos, todos os quais tem tipo 5 e o restante do caminho para completar o contorno $\text{tour}(\omega)$, vamos chamar de $\text{bag}(\omega)$, este caminho tem todos os passos diferentes de 5. \square

Vamos examinar o caminho $\text{bag}(\omega)$. Começemos fazendo duas observações:

Observação 2. *Se $\text{bag}(\omega)$ tem um passo tipo 2, então este passo tem uma \ominus -face em seu lado direito (que é sul).*

Observação 3. *Se $\text{bag}(\omega)$ tem um passo tipo 3, 4 ou 4', então este passo tem uma \ominus -face em seu lado direito (que é leste).*

De fato, em ambos os casos, se existesse uma \oplus -face do lado direito dos passos tipos 2, 3, 4 ou 4' ela deveria ser uma raiz dual, pois o vértice correspondente seria uma raiz, uma vez que, dando um passo oeste no segundo caso ou um passo norte no primeiro, chegaria em uma raiz. Mas, o contorno ao redor de U não pode separar as raízes duais umas das outras.

Chamaremos de *código* qualquer sequência de tipos. Definiremos o deslocamento de um código como o somatório dos deslocamentos de seus termos e definiremos a chance de um código como o produto das chances de seus termos. Se todos os passos de um caminho p tem tipos, chamaremos de *código de p* , e denotaremos por $\text{código}(p)$, a sequência dos tipos dos passos de p . Desse modo, quando falarmos em deslocamento e chance de um caminho p , estaremos nos referindo ao deslocamento e a chance de seu código. Do Lema 2, $\text{bag}(\omega)$ tem um código e precisamos estudá-lo. Chamaremos um caminho em \bar{G} de *bem localizado* se ele começa em V_0 , todos os seus passos tem tipos e todas as variáveis básicas associadas com seus passos são independentes umas das outras e de Ω_0 . Dado algum $\omega \in \Omega_0$ e um código C , dizemos que ω realiza C se o grafo \bar{G} contém um caminho bem localizado p tal que o código de p é igual a C .

Lema 3. *Todo $\omega \in \Omega_1$ realiza o código de $\text{bag}(\omega)$.*

Prova. Vamos classificar todas as variáveis básicas $F(x, t)$ e $A(x, t)$ cujo parâmetro t está entre zero e T , em duas classes: *variáveis básicas esquerdas*, aquelas com $x \leq 0$ e *variáveis básicas direitas*, aquelas com $x > 0$. Denotemos por $\mathcal{A}_{\text{left}}$ a σ -álgebra gerada pelas variáveis básicas esquerdas e $\mathcal{A}_{\text{right}}$ a σ -álgebra gerada pelas variáveis básicas direitas. Note que, qualquer evento em $\mathcal{A}_{\text{right}}$ é independente de qualquer evento em $\mathcal{A}_{\text{left}}$ porque os eventos em $\mathcal{A}_{\text{right}}$ dependem das variáveis básicas $F(x, t)$ e $A(x, t)$ com $x > 0$, enquanto que

os eventos em $\mathcal{A}_{\text{left}}$ dependem apenas das variáveis básicas $F(x, t)$ e $A(x, t)$ com $x \leq 0$. Vamos provar que $\Omega_0 \in \mathcal{A}_{\text{left}}$. Isto é consequência de uma afirmação mais geral: para qualquer $y \geq 0$, qualquer $n \geq 0$ e quaisquer $a_{-n}, a_{-n+1}, \dots, a_{-1} \in \{\ominus, \oplus, \odot\}$, o evento

$$\text{estado}(-n, y) = a_{-n}, \text{estado}(-n+1, y) = a_{-n+1}, \dots, \text{estado}(-1, y) = a_{-1}, \text{estado}(0, y) = \ominus \quad (3.20)$$

pertence a $\mathcal{A}_{\text{left}}$. Vamos provar esta afirmação por indução em y .

Base de indução. Para $y = 1$, temos que o evento

$$\text{estado}(-n, 1) = a_{-n}, \text{estado}(-n+1, 1) = a_{-n+1}, \dots, \text{estado}(-1, 1) = a_{-1}, \text{estado}(0, 1) = \ominus$$

pertence a $\mathcal{A}_{\text{left}}$, pois para este evento ocorrer precisamos apenas das variáveis básicas esquerdas $F(x, 1) \in \{\text{mover}, \text{ficar}\}$ para $-n \leq x \leq -1$ e $F(0, 1) = \text{ficar}$.

Hipótese de indução. Suponha que o resultado é válido para y e vamos mostrar para $y+1$.

Passo de indução. Suponha, sem perda de generalidade, que $y \leq 2T$, pois fixado T , Ω_0 é o conjunto dos $\omega \in \bar{\Omega}$ tal que $\text{estado}(0, 2T) = \ominus$. Logo, apenas faz sentido para nós estudarmos $\text{estado}(0, y)$ para $y \leq 2T$. Desta maneira, se y é par, o evento

$$\begin{aligned} \text{estado}(-n, y+1) &= a_{-n}, \text{estado}(-n+1, y+1) \\ &= a_{-n+1}, \dots, \text{estado}(-1, y+1) = a_{-1}, \text{estado}(0, y+1) = \ominus \end{aligned}$$

pertence a $\mathcal{A}_{\text{left}}$, pois para este evento ocorrer precisamos apenas das variáveis básicas esquerdas $F(x, y/2) \in \{\text{mover}, \text{ficar}\}$ com $-n \leq x \leq -1$ e $F(0, y/2) = \text{ficar}$. Por outro lado, se y é ímpar o evento

$$\begin{aligned} \text{estado}(-n, y+1) &= a_{-n}, \text{estado}(-n+1, y+1) \\ &= a_{-n+1}, \dots, \text{estado}(-1, y+1) = a_{-1}, \text{estado}(0, y+1) = \ominus \end{aligned}$$

pertence a $\mathcal{A}_{\text{left}}$, pois para este evento ocorrer precisamos apenas das variáveis básicas esquerdas $A(x, (y-1)/2) \in \{\text{atacar}, \text{parar}\}$ com $-n \leq x \leq -1$ e $A(0, (y-1)/2) = \text{parar}$.

Assim, provamos (3.20).

Agora, note que para qualquer $\omega \in \Omega_1$ todas as variáveis básicas associadas com os passos de $\text{tour}(\omega)$ pertencem a $\mathcal{A}_{\text{right}}$ e são diferentes umas das outras, pois devido ao Lema 2, o contorno de U não pode passar em uma mesma aresta duas vezes. Daí, $\text{tour}(\omega)$ é bem localizado. Portanto, todo $\omega \in \Omega_1$ realiza o código de $\text{bag}(\omega)$. \square

Para qualquer código C , denotemos por $\text{real}(C)$ o conjunto dos $\omega \in \Omega_0$, que realiza C .

Lema 4. Para qualquer código C temos que,

$$\frac{\pi(\text{real}(C))}{\pi(\Omega_0)} \leq \text{chance}(C). \quad (3.21)$$

Prova. Seja C um código de comprimento n , isto é, $C = (c_1, \dots, c_n)$, então $\text{real}(C)$ é o conjunto dos $\omega \in \Omega_0$ que realiza C . Se não existe $\omega \in \Omega_0$ que realiza C , então $\text{real}(C) = \emptyset$ e portanto (3.21) é satisfeito.

Agora, suponhamos $\text{real}(C) \neq \emptyset$. Estamos usando o fato de que: se C_1 e C_2 são dois códigos tais que $C_1 \neq C_2$ e, além disso, p_1 e p_2 são dois caminhos bem localizados tais que $\text{código}(p_1) = C_1$ e $\text{código}(p_2) = C_2$, então $p_1 \neq p_2$. Mesmo que esses caminhos coincidam em seus primeiros passos. Desta maneira, podemos observar que para todo $\omega \in \text{real}(C)$, existe um único caminho p tal que $\text{código}(p) = C$. Logo, calcular $\pi(\text{real}(C))$ é igual a calcular a probabilidade de ocorrência do caminho p tal que $\text{código}(p) = C$. Ou seja, $\pi(\text{real}(C)) = \mathbb{P}(p)$. Onde, $\mathbb{P}(p)$ é a probabilidade de ocorrência do caminho p .

Seja $\omega \in \text{real}(C)$, então o grafo dual \overline{G} contém um caminho bem localizado p tal que $\text{código}(p) = C$. Como p é um caminho bem localizado, ele começa em V_0 e todos os seus passos tem tipos. Sejam p_1, \dots, p_n os passos de p ; ao passo p_i associamos o tipo c_i , para todo $i = 1, \dots, n$, de modo que a probabilidade de ocorrência do passo p_i , dada pelas variáveis básicas associadas aos passos do caminho p , é menor do que ou igual à $\text{chance}(c_i) \cdot \pi(\Omega_0)$. Como p é bem localizado, as variáveis básicas associadas aos seus passos são independentes umas das outras e de Ω_0 . Logo,

$$\begin{aligned} \mathbb{P}(p) &= \mathbb{P}(p_1) \cdot \dots \cdot \mathbb{P}(p_n) \\ &\leq \text{chance}(c_1) \cdot \dots \cdot \text{chance}(c_n) \cdot (\pi(\Omega_0))^n \\ &= \text{chance}(C) \cdot (\pi(\Omega_0))^n. \end{aligned} \tag{3.22}$$

Portanto, desde que $0 < \pi(\Omega_0) < 1$ e, usando (3.22), temos que

$$\frac{\pi(\text{real}(C))}{\pi(\Omega_0)} \leq \frac{\mathbb{P}(p)}{(\pi(\Omega_0))^n} \leq \text{chance}(C).$$

□

Devido aos Lemas 3 e 4, para qualquer k ,

$$\frac{\pi(\Omega_k)}{\pi(\Omega_0)} \leq \frac{\sum \pi(\text{real}(\text{código}(\text{bag}(\omega))))}{\pi(\Omega_0)} \leq \sum \text{chance}(\text{código}(\text{bag}(\omega))), \tag{3.23}$$

onde ambos os somatórios são tomados sob todos os diferentes $\text{código}(\text{bag}(\omega))$, para $\omega \in \Omega_k$. Com o intuito de estimar essa última soma, para todo $k \in \mathbb{N}$, definimos um conjunto de códigos, que denotamos por CL_k e cujos elementos chamamos de *códigos k -legais*. Um

código $C = (c_1, \dots, c_n)$ pertence a CL_k se satisfaz as seguintes condições:

$$\left\{ \begin{array}{l} (CL - a) \quad c_1 = 1 \text{ e } c_n = 4'. \\ (CL - b) \quad \text{Todos os elementos de } C \text{ pertencem a lista } 1, 1', 2, 3, 4, 4'. \\ (CL - c) \quad \text{Todos os pares } (c_i, c_{i+1}) \text{ pertencem a lista} \\ \quad \quad \quad 11', 1'1, 1'2, 21, 22, 23, 24, 34', 44', 4'2, 4'3, 4'4 \\ (CL - d) \quad \text{HD}(C) \geq k \text{ e } \text{VD}(C) = 0. \end{array} \right. \quad (3.24)$$

Note que $CL_k \supseteq CL_{k+1}$, para todo $k \geq 1$. Denotemos CL_1 por CL e chamemos os elementos de CL apenas de *códigos legais*. Naturalmente, a definição de códigos legais é escolhida para ajustar o código de $\text{bag}(\omega)$, como mostrado no Lema 5.

Lema 5. *Para todo $\omega \in \Omega_k$, o código de $\text{bag}(\omega)$ pertence a CL_k .*

Prova. Vamos provar que o código de $\text{bag}(\omega)$ satisfaz todas as condições de (3.24).

Prova da condição (CL - a): O código de $\text{bag}(\omega)$ começa com tipo 1, pois $\text{bag}(\omega)$ começa do lado leste da face dual ao vértice $(0, 2T)$, indo de norte para sul, ou seja, é um passo sul passando por um nível par, assim, de acordo com a Tabela 2, seu tipo é 1. O código de $\text{bag}(\omega)$ termina com tipo $4'$, porque como as flores estão em um nível par, o último passo de $\text{bag}(\omega)$ é um passo norte através de um nível par, assim, seu tipo é $4'$.

Prova da condição (CL - b): Do Lema 2, o código de $\text{bag}(\omega)$ não pode conter tipo 5. Resta-nos mostrar que $\text{código}(\text{bag}(\omega))$ não pode conter tipos $2'$ ou $2''$. Suponha que algum passo de $\text{bag}(\omega)$ tem tipo $2'$ ou $2''$, então este passo tem uma raiz em seu lado esquerdo, que é norte. Mas se os passos são tipo $2'$ ou $2''$, então em seu lado sul também tem uma raiz, pois existe um caminho noroeste deste vértice (que está no estado \oplus) para alguma flor, o que é impossível uma vez que os passos em $\text{tour}(\omega)$ não podem separar as raízes umas das outras.

Prova da condição (CL - c): Vamos apresentar alguns argumentos, devido aos quais todas as combinações de tipos (c_i, c_{i+1}) , não inclusas em nossa lista, são impossíveis no código de $\text{bag}(\omega)$.

- Pares, em que o primeiro termo está no conjunto $\{1, 3, 4\}$ e o segundo termo está no conjunto $\{1, 2, 3, 4\}$, são impossíveis, porque o primeiro termo termina em um subnível par, mas o segundo termo começa em um subnível ímpar.
- Pares, em que o primeiro termo está no conjunto $\{1', 2, 4'\}$ e o segundo termo está no conjunto $\{1', 4'\}$, são impossíveis, porque o primeiro termo termina em um subnível ímpar, mas o segundo termo começa em um subnível par.
- Pares, em que o primeiro termo está no conjunto $\{1, 1'\}$ e o segundo termo está no conjunto $\{3, 4, 4'\}$, são impossíveis. Se pudesse ocorrer os pares decorrentes desses conjuntos, então como o primeiro tipo é um passo sul e o segundo tipo é um passo norte, então eles estariam na mesma aresta vetical de \overline{G} . Mas os termos do primeiro conjunto precisam de uma \oplus -face em seu lado leste, enquanto que os termos do

segundo conjunto precisam de uma \ominus -face em seu lado leste devido à **Observação 3**. O que é um absurdo.

- Pares em que o primeiro termo está no conjunto $\{3, 4\}$ e o segundo termo é $1'$ são impossíveis, porque eles tem que estar na mesma aresta, mas a face no lado leste de um passo tipo 3 ou 4 deve estar no estado \ominus , enquanto a face no lado leste de passo tipo $1'$ deve estar no estado \oplus .
- Pares, em que o primeiro termo é $4'$ e o segundo termo é 1, são impossíveis, porque eles tem que estar na mesma aresta, mas a face no lado leste de um passo tipo $4'$ deve estar no estado \ominus , enquanto a face no lado leste de um passo tipo 1 deve estar no estado \oplus .

Prova da condição (CL-d): Do Lema 2, temos que o caminho $\text{tour}(\omega)$ é uma concatenação dos caminhos $\text{bag}(\omega)$ e $\text{lid}(\omega)$, onde $\text{lid}(\omega)$ tem $\phi(\omega)$ passos tipo 5, todos os quais tem deslocamento $(-1, 0)$, logo o deslocamento de $\text{lid}(\omega)$ é $(-\phi(\omega), 0)$. Desta maneira, como o contorno $\text{tour}(\omega)$ começa e termina em V_0 , segue que $\text{bag}(\omega)$ tem x passos leste, onde $x \geq \phi(\omega)$ (se não houver assassinatos $x = \phi(\omega)$). Da condição (CL-b) o único passo leste que $\text{bag}(\omega)$ pode ter é o tipo 2, cujo deslocamento é $(1, 0)$, logo $\text{HD}(\text{bag}(\omega)) = x \geq \phi(\omega) \geq k$.

Como o contorno $\text{tour}(\omega)$ começa e termina em V_0 e o caminho $\text{lid}(\omega)$ só tem passos oeste (tipo 5), então o número de passos sul (tipos 1 e $1'$) é igual ao número de passos norte (tipos 3, 4 e $4'$) no caminho $\text{bag}(\omega)$. Vamos chamar esse número de passos norte, consequentemente o número de passos sul, de y . Uma vez que todos os tipos sul tem deslocamento $(0, -1)$ e todos os tipos norte tem deslocamento $(0, 1)$ segue que $\text{VD}(\text{bag}(\omega)) = -y + y = 0$. □

Segue, do Lema 5 e de (3.23), que para todo natural k ,

$$\frac{\pi(\Omega_k)}{\pi(\Omega_0)} \leq \sum_{C \in \text{CL}_k} \text{chance}(C). \quad (3.25)$$

Para finalizar nossos argumentos, precisamos fazer uma estimativa numérica, mas será trabalhoso fazer isto com tantos tipos. Para reduzir este número para quatro, vamos chamar de *tipos principais* os elementos do conjunto $\{1, 2, 3, 4\}$. Todas as quantidades definidas para tipos são válidas para tipos principais. Em particular, todo tipo principal tem um deslocamento e uma chance (ver na Tabela 2).

Um *código principal* é uma sequência finita, em que todos os termos são tipos principais. Para qualquer código C , denotamos por $\text{curto}(C)$ o código principal obtido de C deletando todos os tipos não principais. Devemos simplificar nossa tarefa lidando com $\text{curto}(\text{código}(\text{bag}(\omega)))$ ao invés de $\text{código}(\text{bag}(\omega))$. Para todo natural k definimos o conjunto CPL_k , cujos elementos chamamos de *códigos principais k -legais*. Por definição, um código principal k -legal é um código principal $C = (c_1, \dots, c_n)$, que satisfaz as seguintes

condições:

$$\left\{ \begin{array}{l} (CPL - a) \quad c_1 = 1. \\ (CPL - b) \quad \text{Para todo } i = 1, \dots, n - 1 \text{ é impossível que} \\ \quad (c_i = 1, c_{i+1} = 3) \text{ ou } (c_i = 3, c_{i+1} = 1) \text{ ou} \\ \quad (c_i = 1, c_{i+1} = 4) \text{ ou } (c_i = 4, c_{i+1} = 1). \\ (CPL - c) \quad c_n \text{ é igual a 3 ou 4.} \\ (CPL - d) \quad \text{HD}(C) \geq k \\ (CPL - e) \quad \text{VD}(C) = 0. \end{array} \right. \quad (3.26)$$

Note que $CPL_k \supseteq CPL_{k+1}$, para todo $k \geq 1$. Denotemos CPL_1 por CPL e chamaremos os elementos de CPL apenas de *códigos principais legais*. Note que de $(CPL - a)$, $(CPL - b)$ e $(CPL - c)$ qualquer código principal legal tem comprimento de pelo menos 3.

Denotemos nesta parte do texto, C para códigos principais legais e C' para códigos legais. Sendo assim, para qualquer código principal legal C denotaremos por $\text{longo}(C)$ o *código legal* C' tal que $C = \text{curto}(C')$.

Vamos estabelecer uma bijeção entre os códigos principais legais e os códigos legais. Denotemos esta bijeção por

$$\text{longo} : CPL \rightarrow CL.$$

Dado $C \in CPL$, $\text{longo}(C) = C'$ é obtido por meio do seguinte procedimento:

$$\left\{ \begin{array}{l} (a) \quad \text{Começamos com } C. \\ (b) \quad \text{Após todo 1 inserimos } 1'. \\ (c) \quad \text{Após todo 3 inserimos } 4'. \\ (d) \quad \text{Após todo 4 inserimos } 4'. \end{array} \right. \quad (3.27)$$

A aplicação longo é injetiva, pois dados dois códigos principais legais C_1 e C_2 , com $C_1 \neq C_2$, então pelo procedimento (3.27) temos que $\text{longo}(C_1) \neq \text{longo}(C_2)$. A sobrejetividade segue do Lema 14 (ver Apêndice 5).

Note que a inversa da aplicação longo é a aplicação curto que pega um código legal C' e elimina os tipos 1' e 4'.

Agora podemos estimar o somatório no lado direito de (3.25). Lembremos que dado $C' \in CL_k$ existe único $C \in CPL_k$ tal que $C' = \text{longo}(C)$. Vimos pela condição (3.27) que C e C' diferem pelos tipos 1' e 4', os quais tem chance 1. Daí, se $C' = \text{longo}(C)$ temos que $\text{chance}(C') = \text{chance}(C)$. Além disso, vemos que a aplicação longo é bijetora. Logo,

$$\begin{aligned} \sum_{C' \in CL_k} \text{chance}(C') &= \sum_{C \in CPL_k} \text{chance}(\text{longo}(C)) \\ &= \sum_{C \in CPL_k} \text{chance}(C). \end{aligned}$$

Resta-nos mostrar que

$$\sum_{k=1}^{\infty} \sum_{C \in CPL_k} \text{chance}(C) \leq \frac{\alpha^2}{55}.$$

Note que

$$\sum_{k=1}^{\infty} \sum_{C \in CPL_k} \text{chance}(C) = \sum_{C \in CPL} \text{HD}(C) \cdot \text{chance}(C),$$

portanto, devemos provar que

$$\sum_{C \in CPL} \text{HD}(C) \cdot \text{chance}(C) \leq \frac{\alpha^2}{55}.$$

Para quaisquer inteiros x e y , qualquer natural z e $k \in \{1, 2, 3, 4\}$ denotemos por $S_k(x, y, z)$ o somatório das chances dos códigos principais que satisfazem as condições $(CPL - a)$ e $(CPL - b)$ da definição de códigos principais legais, cujo HD é igual a x , cujo VD é igual a y e que tem z termos, o último dos quais é k . Os códigos principais que satisfazem $(CPL - e)$ tem VD igual a zero. Segue da definição de $S_k(x, y, z)$ e das condições $(CPL - c)$, $(CPL - d)$ e $(CPL - e)$ de (3.26) que

$$\sum_{C \in CPL} \text{HD}(C) \cdot \text{chance}(C) \leq \sum_{x=1}^{\infty} \sum_{y=0}^{\infty} \sum_{z=1}^{\infty} x(S_3(x, y, z) + S_4(x, y, z)). \quad (3.28)$$

Para auxiliar nossas contas, vamos redefinir o tipo 3 da Tabela 2 de modo que seu deslocamento seja da forma $(-1, 1)$. Note que este procedimento não altera a quantidade de contornos $\text{tour}(\omega)$, tampouco a “quantidade de mais” dentro destes contornos. Na Figura 3 a seguir traremos uma ilustração do contorno da Figura 2 e também o novo contorno após a alteração do tipo 3.

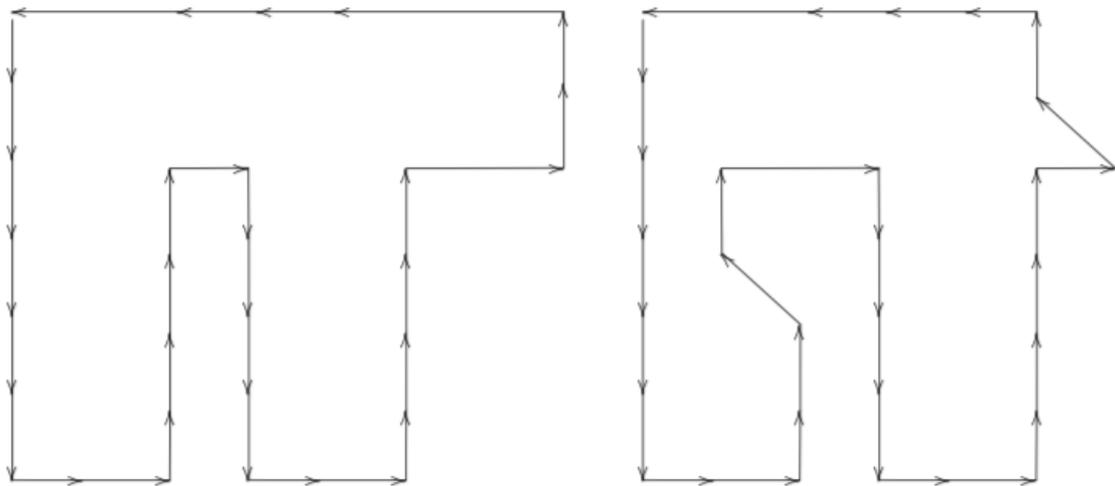
Devido a condição $(CPL - a)$ de (3.26), os números $S_k(x, y, z)$ satisfazem a condição inicial

$$S_k(x, y, 1) = \begin{cases} 1, & \text{se } x = 0, y = -1 \text{ e } k = 1, \\ 0, & \text{em todos os outros casos,} \end{cases}$$

pois se um código principal legal só tem um termo, ele deve ser tipo 1, logo seu deslocamento horizontal é zero e seu deslocamento vertical é -1 e sua chance é 1. E, devido à condição $(CPL - b)$, os números $S_k(x, y, z)$ satisfazem às equações de transição

$$\begin{cases} S_1(x, y, z + 1) = S_1(x, y + 1, z) + S_2(x, y + 1, z) \\ S_2(x, y, z + 1) = \beta[S_1(x - 1, y, z) + S_2(x - 1, y, z) + S_3(x - 1, y, z) + S_4(x - 1, y, z)] \\ S_3(x, y, z + 1) = \alpha[S_2(x + 1, y - 1, z) + S_3(x + 1, y - 1, z) + S_4(x + 1, y - 1, z)] \\ S_4(x, y, z + 1) = (1 - \alpha)[S_2(x, y - 1, z) + S_3(x, y - 1, z) + S_4(x, y - 1, z)]. \end{cases} \quad (3.29)$$

Figura 3 – Ilustração do contorno $\text{tour}(\omega)$ com o tipo 3 tendo deslocamento $(0, 1)$ e $(-1, 1)$.



Fonte: O autor (2020).

Para estimar (3.28), vamos usar as somas

$$S_i(z) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} p^{-x} q^{-y} S_i(x, y, z), \quad \text{para todo } i = 1, 2, 3, 4, \quad (3.30)$$

onde p e q são parâmetros positivos tais que $x < p^{-x}$ e $1 < q^{-y}$. Desta maneira, o lado direito de (3.28) é limitado superiormente por

$$\sum_{z=1}^{\infty} [S_3(z) + S_4(z)]. \quad (3.31)$$

As somas em (3.30) satisfazem à condição inicial

$$S_1(1) = q \quad \text{e} \quad S_2(1) = S_3(1) = S_4(1) = 0$$

e às condições de recorrência

$$\begin{cases} S_1(z+1) = q[S_1(z) + S_2(z)], \\ S_2(z+1) = \beta/p [S_1(z) + S_2(z) + S_3(z) + S_4(z)], \\ S_3(z+1) = p\alpha/q [S_2(z) + S_3(z) + S_4(z)], \\ S_4(z+1) = (1-\alpha)/q [S_2(z) + S_3(z) + S_4(z)]. \end{cases}$$

Note que há uma relação entre $S_3(z)$ e $S_4(z)$. Dessa maneira vamos definir as seguintes

quantidades

$$S_1^*(z) = S_1(z), \quad S_2^*(z) = S_2(z) \quad \text{e} \quad S_3^*(z) = S_3(z) + S_4(z)$$

que tem a seguinte condição inicial

$$S_1^*(1) = q \quad \text{e} \quad S_2^*(1) = S_3^*(1) = 0, \quad (3.32)$$

e as seguintes condições de recorrência

$$\begin{cases} S_1^*(z+1) = q[S_1^*(z) + S_2^*(z)] \\ S_2^*(z+1) = \beta/p[S_1^*(z) + S_2^*(z) + S_3^*(z)] \\ S_3^*(z+1) = r[S_2^*(z) + S_3^*(z)] \end{cases}$$

onde

$$r = \frac{p\alpha + (1 - \alpha)}{q}. \quad (3.33)$$

Introduzindo o vetor $S^*(z) = (S_1^*(z), S_2^*(z), S_3^*(z))$ podemos reescrever as equações de recorrência como $S^*(z+1) = S^*(z) \cdot B$. Portanto, $S^*(z) = S^*(1) \cdot B^{z-1}$, onde B é a matriz

$$B = \begin{pmatrix} q & \beta/p & 0 \\ q & \beta/p & r \\ 0 & \beta/p & r \end{pmatrix}.$$

Considere a matriz

$$C = \begin{pmatrix} q & \beta/p & 0 \\ q & \beta/p & r \\ q & \beta/p & r \end{pmatrix}.$$

Note que $B \leq C$, então, pelo Lema 16 (ver Apêndice),

$$S^*(z) \leq S^*(1) \cdot C^{z-1}.$$

A matriz C possui três autovalores distintos, a saber,

$$\lambda_{PF} = \frac{r + q + \beta/p + \sqrt{\Delta}}{2}, \quad \lambda_2 = \frac{r + q + \beta/p - \sqrt{\Delta}}{2}, \quad \lambda_3 = 0,$$

onde $\Delta = (r + q + \beta/p)^2 - 4qr$ e λ_{PF} é o autovalor de Perron-Frobenius (ver (SENETA, 2006)). Para que $\lambda_{PF} < 1$, devemos ter (ver Lema 19 no Apêndice)

$$\beta < p(1 + qr - q - r). \quad (3.34)$$

Os valores de p e q que maximizam o lado direito de (3.34) são

$$p = \frac{4\alpha - 3 + \sqrt{9 - 8\alpha}}{8\alpha} \quad \text{e} \quad q = \sqrt{p\alpha + 1 - \alpha}.$$

Note que p e q são decrescentes como função de $\alpha \in (0, 1)$. Além disso, para cada α , $p < q < 1$.

Note, ainda, que fixados p , q e α , λ_{PF} é uma função monótona crescente com relação a β . Desse modo, se considerarmos

$$\beta < \frac{\alpha^2}{55} \tag{3.35}$$

ainda teremos que $\lambda_{PF} < 1$. E esse, é o limite superior que usaremos para β .

Os autovetores associados a esses autovalores são da forma

$$\begin{aligned} v_{\lambda_{PF}} &= \left(v_{11}, \frac{\beta}{pq} v_{11}, \frac{r\beta}{(\lambda_{PF} - r)pq} v_{11} \right), \quad \text{com } v_{11} \in \mathbb{R} \setminus \{0\}, \\ v_{\lambda_2} &= \left(v_{21}, \frac{\beta}{pq} v_{21}, \frac{r\beta}{(\lambda_2 - r)pq} v_{21} \right), \quad \text{com } v_{21} \in \mathbb{R} \setminus \{0\}, \\ v_{\lambda_3} &= (0, v_{31}, -v_{31}), \quad \text{com } v_{31} \in \mathbb{R} \setminus \{0\}. \end{aligned}$$

Como a matriz C possui três autovalores distintos, segue que ela é diagonalizável, logo, pelo Lema 15 (ver Apêndice), podemos escrever

$$C = V^{-1}DV$$

onde D é a matriz diagonal formada pelos autovalores e V é a matriz dos autovetores, ou seja,

$$D = \begin{pmatrix} \lambda_{PF} & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad V = \begin{pmatrix} 1 & a & b \\ 1 & a & c \\ 0 & 1 & -1 \end{pmatrix}$$

onde $a = \beta/pq$, $b = r\beta/((\lambda_{PF} - r)pq)$ e $c = r\beta/((\lambda_2 - r)pq)$. Logo,

$$V^{-1} = \begin{pmatrix} -\frac{a+c}{b-c} & \frac{a+b}{b-c} & -a \\ \frac{1}{b-c} & -\frac{1}{b-c} & 1 \\ \frac{1}{b-c} & -\frac{1}{b-c} & 0 \end{pmatrix}.$$

Queremos calcular $S^*(1)C^{z-1}$. Para simplificar a notação, vamos chamar $k = z - 1$, sendo assim, queremos calcular $S^*(1)C^k$, onde $S^*(1) = \begin{pmatrix} q & 0 & 0 \end{pmatrix}$. Note que,

$$S^*(1)C^k = S^*(1)V^{-1}D^kV$$

$$\begin{aligned}
&= \begin{pmatrix} q & 0 & 0 \end{pmatrix} \begin{pmatrix} -\frac{a+c}{b-c} & \frac{a+b}{b-c} & -a \\ \frac{1}{b-c} & -\frac{1}{b-c} & 1 \\ \frac{1}{b-c} & -\frac{1}{b-c} & 0 \end{pmatrix} D^k V \\
&= q \begin{pmatrix} -\frac{a+c}{b-c} & \frac{a+b}{b-c} & -a \end{pmatrix} \begin{pmatrix} \lambda_{PF}^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & a & b \\ 1 & a & c \\ 0 & 1 & -1 \end{pmatrix} \\
&= q \begin{pmatrix} -\frac{a+c}{b-c} & \frac{a+b}{b-c} & -a \end{pmatrix} \begin{pmatrix} \lambda_{PF}^k & a\lambda_{PF}^k & b\lambda_{PF}^k \\ \lambda_2^k & a\lambda_2^k & c\lambda_2^k \\ 0 & 0 & 0 \end{pmatrix} \\
&= q \begin{pmatrix} -\frac{a+c}{b-c}\lambda_{PF}^k + \frac{a+b}{b-c}\lambda_2^k & -\frac{a+c}{b-c}a\lambda_{PF}^k + \frac{a+b}{b-c}a\lambda_2^k & -\frac{a+c}{b-c}b\lambda_{PF}^k + \frac{a+b}{b-c}c\lambda_2^k \end{pmatrix}
\end{aligned}$$

Desta forma, (usando o fato que $k = z - 1$)

$$S_3^*(z) \leq q \left(-\frac{a+c}{b-c}b\lambda_{PF}^{z-1} + \frac{a+b}{b-c}c\lambda_2^{z-1} \right).$$

Logo, usando (3.35),

$$\begin{aligned}
\sum_{z=1}^{\infty} S_3^*(z) &\leq \sum_{z=1}^{\infty} q \left(-\frac{a+c}{b-c}b\lambda_{PF}^{z-1} + \frac{a+b}{b-c}c\lambda_2^{z-1} \right) \\
&= -\frac{a+c}{b-c}bq \sum_{z=0}^{\infty} \lambda_{PF}^z + \frac{a+b}{b-c}cq \sum_{z=0}^{\infty} \lambda_2^z \\
&= -\frac{a+c}{b-c}bq \frac{1}{1-\lambda_{PF}} + \frac{a+b}{b-c}cq \frac{1}{1-\lambda_2}.
\end{aligned}$$

Note que,

$$a+b = \frac{\beta\lambda_{PF}}{pq(\lambda_{PF}-r)}, \quad a+c = \frac{\beta\lambda_2}{pq(\lambda_2-r)}, \quad b-c = \frac{r\beta(\lambda_2-\lambda_{PF})}{pq(\lambda_{PF}-r)(\lambda_2-r)}.$$

Logo,

$$-\frac{a+c}{b-c} = -\frac{\lambda_2(\lambda_{PF}-r)}{r(\lambda_2-\lambda_{PF})} \quad \text{e} \quad \frac{a+b}{b-c} = \frac{\lambda_{PF}(\lambda_2-r)}{r(\lambda_2-\lambda_{PF})}$$

daí,

$$-\frac{a+c}{b-c}bq = -\frac{\beta\lambda_2}{p(\lambda_2-\lambda_{PF})} \quad \text{e} \quad \frac{a+b}{b-c}cq = \frac{\beta\lambda_{PF}}{p(\lambda_2-\lambda_{PF})}.$$

Então,

$$\sum_{z=1}^{\infty} S_3^*(z) \leq -\frac{\beta\lambda_2}{p(\lambda_2-\lambda_{PF})} \frac{1}{1-\lambda_{PF}} + \frac{\beta\lambda_{PF}}{p(\lambda_2-\lambda_{PF})} \frac{1}{1-\lambda_2}$$

$$\begin{aligned}
&= \frac{\beta}{p(\lambda_2 - \lambda_{PF})} \left[\frac{\lambda_{PF}}{1 - \lambda_2} - \frac{\lambda_2}{1 - \lambda_{PF}} \right] \\
&= \frac{\beta}{p(\lambda_{PF} - \lambda_2)} \left[\frac{\lambda_2}{1 - \lambda_{PF}} - \frac{\lambda_{PF}}{1 - \lambda_2} \right] \\
&\leq \frac{\beta}{p(\lambda_{PF} - \lambda_2)} \left[\frac{\lambda_{PF}}{1 - \lambda_{PF}} - \frac{\lambda_{PF}}{1 - \lambda_2} \right] \\
&= \frac{\beta}{p(\lambda_{PF} - \lambda_2)} \left[\frac{\lambda_{PF} - \lambda_{PF}\lambda_2 - \lambda_{PF} + \lambda_{PF}^2}{(1 - \lambda_{PF})(1 - \lambda_2)} \right] \\
&= \frac{\beta}{p(\lambda_{PF} - \lambda_2)} \left[\frac{\lambda_{PF}(\lambda_{PF} - \lambda_2)}{(1 - \lambda_{PF})(1 - \lambda_2)} \right] \\
&= \frac{\beta\lambda_{PF}}{p(1 - \lambda_{PF})(1 - \lambda_2)}. \tag{3.36}
\end{aligned}$$

Denotemos (3.36) por

$$LS_1(p, q, \alpha, \beta). \tag{3.37}$$

Prova do item (A. 1) do Teorema 1. Do Lema 20 (ver Apêndice), temos que $LS_1(p, q, \alpha, \beta)$ é crescente como função de β , para $\beta \in (0, \alpha^2/55)$. Assim, se $LS_1(p, q, \alpha, \alpha^2/55)$ for menor que 1, então $LS_1(p, q, \alpha, \beta)$ será menor que 1 para $\beta \in (0, \alpha^2/55)$. Uma vez que $(1 - \lambda_{PF})(1 - \lambda_2) = 1 + qr - q - r - \beta/p$, segue que:

$$LS_1(p, q, \alpha, \beta) = \frac{\beta\lambda_{PF}}{p(1 + qr - q - r) - \beta}.$$

Desta maneira,

$$\begin{aligned}
LS_1(p, q, \alpha, \beta) < 1 &\iff \frac{\beta\lambda_{PF}}{p(1 + qr - q - r) - \beta} < 1 \\
&\iff \beta\lambda_{PF} < p(1 + qr - q - r) - \beta \\
&\iff \beta(\lambda_{PF} + 1) < p(1 + qr - q - r).
\end{aligned}$$

Usando que $p = (4\alpha - 3 + \sqrt{9 - 8\alpha})/8\alpha$, $q = \sqrt{p\alpha + 1 - \alpha}$ e $r = (p\alpha + 1 - \alpha)/q$, então, do Lema 21 (ver Apêndice),

$$p(1 + qr - q - r) \geq \frac{\alpha^2}{27}.$$

Logo, quando $\beta = \alpha^2/55$, temos $2\beta < p(1 + qr - q - r)$. Ou seja, $\beta < p(1 + qr - q - r)/2$, daí $0 < \lambda_{PF} < 1$. Portanto,

$$\beta(\lambda_{PF} + 1) < p(1 + qr - q - r).$$

Assim,

$$LS_1\left(\frac{4\alpha - 3 + \sqrt{9 - 8\alpha}}{8\alpha}, \sqrt{p\alpha + 1 - \alpha}, \alpha, \frac{\alpha^2}{55}\right) < 1.$$

□

Prova do item (B. 1) do Teorema 1. Utilizaremos o Resultado 1, assumindo $R = \{\oplus\} \subset \mathbb{A}^{\mathbb{Z}}$, isto é, $R = \{x \in \mathcal{A}^{\mathbb{Z}} : x_{i_0} = \oplus \text{ para algum } i_0 \in \mathbb{Z}\}$. Seja $\delta = LS_1(p, q, \alpha, \beta)$, onde $p = (4\alpha - 3 + \sqrt{9 - 8\alpha})/8\alpha$, $q = \sqrt{p\alpha + 1 - \alpha}$ e $\beta \in (0, \alpha^2/55)$. Da demonstração do item (A. 1) do Teorema 1,

$$\mu_t(\oplus) = \delta_{\ominus}(\mathbf{F}_{\beta}\mathbf{M}_{\alpha})^t(\oplus) \leq LS_1(p, q, \alpha, \beta), \quad \text{para todo } t.$$

Logo, pelo Resultado 1, o operador $\mathbf{F}_{\beta}\mathbf{M}_{\alpha}$ tem uma medida invariante ν , isto é, $\nu(\mathbf{F}_{\beta}\mathbf{M}_{\alpha}) = \nu$ tal que

$$\nu(\oplus) \leq LS_1(p, q, \beta, \alpha) < 1.$$

□

3.5 Aproximação de Campo Médio do Processo Flip-Murder

Como em (RAMOS e TOOM, 2008), denotemos por $\mathcal{C} : \mathcal{M}_u \rightarrow \mathcal{M}_u$, o bem conhecido *operador de Campo Médio*, onde \mathcal{M}_u é o conjunto das medidas normalizadas e uniformes em $\mathcal{A}^{\mathbb{Z}}$. Sua ação equivale a misturar aleatoriamente todas as componentes. Em outras palavras, para cada $\mu \in \mathcal{M}_u$ a medida $\mu\mathcal{C}$ é uma medida produto com a mesma frequência de todas as letras que μ tinha. O operador de Campo Médio nos permite aproximar um dado processo $\mu\mathbf{P}^t$ no espaço de configuração $\mathcal{A}^{\mathbb{Z}}$ por outro processo $\mu(\mathcal{CP})^t$ no mesmo espaço. Assim, ao invés do processo original, cujo conjunto de incógnitas é infinito ou muito grande, lidamos com a evolução das densidades das letras, que é um conjunto finito e limitado de incógnitas. Uma vez que a densidade das letras somam um, o número de incógnitas independentes na aproximação de Campo Médio é igual ao número de letras no alfabeto menos um. Em nosso caso, com apenas duas letras, lidamos com apenas uma incógnita: para tal, vamos escolher a densidade de mais.

Da definição dos operadores flip e murder temos

$$\begin{aligned} \mu\mathbf{M}_{\alpha}(\oplus) &= \frac{\mu(\oplus) - \alpha\mu(\oplus, \ominus)}{1 - \alpha\mu(\oplus, \ominus)}, \\ \mu\mathbf{F}_{\beta}(\oplus) &= \mu(\oplus) + \beta\mu(\ominus) = \mu(\oplus) + \beta(1 - \mu(\oplus)) = (1 - \beta)\mu(\oplus) + \beta, \\ \mu\mathbf{F}_{\beta}(\ominus) &= 1 - \mu\mathbf{F}_{\beta}(\oplus). \end{aligned}$$

Logo, $\mu\mathbf{F}_{\beta}(\oplus, \ominus) = \mu\mathbf{F}_{\beta}(\oplus)\mu\mathbf{F}_{\beta}(\ominus) = \mu\mathbf{F}_{\beta}(\oplus)(1 - \mu\mathbf{F}_{\beta}(\oplus))$. Daí,

$$\begin{aligned} \mu\mathbf{F}_{\beta}\mathbf{M}_{\alpha}(\oplus) &= \frac{\mu\mathbf{F}_{\beta}(\oplus) - \alpha\mu\mathbf{F}_{\beta}(\oplus, \ominus)}{1 - \alpha\mu\mathbf{F}_{\beta}(\oplus, \ominus)} \\ &= \frac{\mu\mathbf{F}_{\beta}(\oplus) - \alpha\mu\mathbf{F}_{\beta}(\oplus)(1 - \mu\mathbf{F}_{\beta}(\oplus))}{1 - \alpha\mu\mathbf{F}_{\beta}(\oplus)(1 - \mu\mathbf{F}_{\beta}(\oplus))}. \end{aligned}$$

Devido às propriedades do operador de Campo Médio, a densidade de mais na medida $\mu\mathcal{CF}_\beta\mathbf{M}_\alpha$ depende apenas da densidade de mais na medida μ , e esta dependência pode ser expressada pela fórmula

$$f(x) = \frac{b - \alpha b(1 - b)}{1 - \alpha b(1 - b)}, \quad (3.38)$$

onde x denota a frequência de mais na medida μ , $f(x)$ denota a densidade de mais na medida $\mu\mathcal{CF}_\beta\mathbf{M}_\alpha$ e $b = \mu\mathbf{F}_\beta(\oplus) = (1 - \beta)\mu(\oplus) + \beta$. Assim, o estudo do operador $\mathbf{F}_\beta\mathbf{M}_\alpha$ é substituído pelo estudo do operador $\mathcal{CF}_\beta\mathbf{M}_\alpha$, que se resume ao estudo do sistema dinâmico $f : [0, 1] \rightarrow [0, 1]$ com parâmetros $\alpha, \beta \in [0, 1]$. Chamamos um *ponto fixo* deste sistema, um valor $x \in [0, 1]$ tal que $f(x) = x$. Chamamos nosso sistema dinâmico *ergódico* se ele tem um único ponto fixo x_{fixo} e

$$\forall x \in [0, 1] : \lim_{t \rightarrow \infty} f^t(x) = x_{\text{fixo}},$$

onde f^t significa a t -ésima iteração de f .

Lema 6. *A aproximação de Campo Médio $\mathcal{CF}_\beta\mathbf{M}_\alpha$ é ergódico se $\alpha < 4\beta$ e é não ergódico se $\alpha \geq 4\beta$.*

Prova. Inicialmente vamos encontrar os pontos fixos de (3.38). Note que,

$$\begin{aligned} f(x) = x &\Leftrightarrow \frac{b - \alpha b(1 - b)}{1 - \alpha b(1 - b)} = \mu(\oplus) \\ \Leftrightarrow b - \alpha b(1 - b) &= \mu(\oplus) - \alpha b(1 - b)\mu(\oplus) \\ \Leftrightarrow b - \mu(\oplus) + \alpha b(1 - b)(\mu(\oplus) - 1) &= 0 \\ \Leftrightarrow (1 - \beta)\mu(\oplus) + \beta - \mu(\oplus) + \alpha[(1 - \beta)\mu(\oplus) + \beta][1 - (1 - \beta)\mu(\oplus) - \beta](\mu(\oplus) - 1) &= 0 \\ \Leftrightarrow -\beta(\mu(\oplus) - 1) + \alpha\{\mu(\oplus)[(1 - \beta)(1 - 2\beta)] - [(1 - \beta)\mu(\oplus)]^2 + \beta - \beta^2\}(\mu(\oplus) - 1) &= 0 \\ \Leftrightarrow \{\alpha\{\mu(\oplus)[(1 - \beta)(1 - 2\beta)] - [(1 - \beta)\mu(\oplus)]^2 + \beta - \beta^2\} - \beta\}(\mu(\oplus) - 1) &= 0 \\ \Leftrightarrow \mu(\oplus) = 1 \\ \text{ou } -\alpha(1 - \beta)^2[\mu(\oplus)]^2 + \alpha(1 - \beta)(1 - 2\beta)\mu(\oplus) + \alpha\beta - \alpha\beta^2 - \beta &= 0 \end{aligned} \quad (3.39)$$

Vamos resolver esta equação do segundo grau em (3.39). Note que $\Delta = [\alpha(1 - \beta)(1 - 2\beta)]^2 + 4\alpha(1 - \beta)^2(\alpha\beta - \alpha\beta^2 - \beta)$. Deste modo,

$$\begin{aligned} \mu(\oplus) &= \frac{-\alpha(1 - \beta)(1 - 2\beta) \pm \sqrt{(1 - \beta)^2[(\alpha(1 - 2\beta))^2 + 4\alpha(\alpha\beta - \alpha\beta^2 - \beta)]}}{-2\alpha(1 - \beta)^2} \\ &= \frac{\alpha(2\beta - 1) \pm \sqrt{\alpha(\alpha - 4\beta)}}{2\alpha(\beta - 1)}. \end{aligned}$$

Portanto, os pontos fixos de (3.38) são $\mu(\oplus) \in \left\{ 1, \frac{2\alpha\beta - \alpha + \sqrt{\alpha(\alpha - 4\beta)}}{2\alpha(\beta - 1)}, \frac{2\alpha\beta - \alpha - \sqrt{\alpha(\alpha - 4\beta)}}{2\alpha(\beta - 1)} \right\}$.

Vamos examinar os pontos fixos de $\mathcal{CF}_\beta \mathbf{M}_\alpha$. Denotemos por:

$$\lambda_0 = 1, \lambda_1 = \frac{2\alpha\beta - \alpha - \sqrt{\alpha(\alpha - 4\beta)}}{2\alpha(\beta - 1)} \text{ e } \lambda_2 = \frac{2\alpha\beta - \alpha + \sqrt{\alpha(\alpha - 4\beta)}}{2\alpha(\beta - 1)}.$$

- Se $\alpha(\alpha - 4\beta) > 0$, então $2\alpha\beta - \alpha + \sqrt{\alpha(\alpha - 4\beta)} > 2\alpha\beta - \alpha - \sqrt{\alpha(\alpha - 4\beta)}$. Como $2\alpha(\beta - 1) < 0$, pois $\beta < 1$, segue que $\lambda_2 < \lambda_1$. Neste caso nosso operador é não ergódico.
- Se $\alpha(\alpha - 4\beta) < 0$, teremos λ_1 e λ_2 complexos. Logo, a única raiz real de (3.39) é $\lambda_0 = 1$. Neste caso nosso operador é ergódico. Para isto precisamos mostrar que para todo $x_0 \in [0, 1]$, $f^t(x_0)$ tende para 1 quando t tende para infinito.

Uma vez que a função f é contínua em $[0, 1]$, temos que a função $g(x) = f(x) - x$ é também contínua neste intervalo. Note que

$$g(0) = f(0) - 0 = \frac{\beta - \alpha\beta(1 - \beta)}{1 - \alpha\beta(1 - \beta)} > 0,$$

uma vez que $\alpha(1 - \beta) < 1$. Além disso,

$$g(1) = f(1) - 1 = \frac{1}{1} - 1 = 0.$$

Uma vez que g é contínua, $g(1) = 0$ e $g(0) > 0$, concluímos que $g(x) > 0$ para todo $x < 1$, daí, $f(x) > x$ para todo $x < 1$. Então para todo $x_0 < 1$ a sequência $f^t(x_0)$ é crescente e limitada por 1 e portanto tem um limite, o qual é um ponto fixo de f . Como 1 é o único ponto fixo de f , então $f^t(x_0) \rightarrow 1$ quanto $t \rightarrow \infty$. \square

Note que, $\alpha(\alpha - 4\beta) < 0 \Leftrightarrow \alpha - 4\beta < 0 \Leftrightarrow \alpha < 4\beta$. Logo, nosso processo é ergódico se $\alpha < 4\beta$ e é não ergódico se $\alpha \geq 4\beta$.

Na região de não ergodicidade temos que $\lambda_1 < 1$. De fato,

$$\begin{aligned} & \lambda_1 < 1 \\ \Leftrightarrow & \frac{2\alpha\beta - \alpha - \sqrt{\alpha(\alpha - 4\beta)}}{2\alpha(\beta - 1)} < 1 \\ \Leftrightarrow & 2\alpha\beta - \alpha - \sqrt{\alpha(\alpha - 4\beta)} > 2\alpha(\beta - 1) \\ \Leftrightarrow & \alpha - \sqrt{\alpha(\alpha - 4\beta)} > 0 \\ \Leftrightarrow & \alpha > \sqrt{\alpha(\alpha - 4\beta)} \\ \Leftrightarrow & \alpha^2 > \alpha^2 - 4\alpha\beta \\ \Leftrightarrow & 0 > -4\alpha\beta. \end{aligned}$$

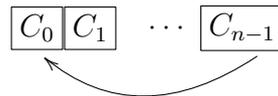
Portanto, na região de não ergodicidade temos que $\lambda_2 < \lambda_1 < \lambda_0$.

3.6 Simulação de Monte Carlo

Junto com o processo em $\mathcal{A}^{\mathbb{Z}}$, devemos considerar seus análogos em espaços finitos. Quando nossos operadores atuam em configurações finitas, toda ação do assassinato em uma palavra $\oplus\ominus$ decresce o comprimento da configuração por uma unidade, daí, em média, o número de componentes decresce e o processo se degenera em uma sequência finita de mais (desde que $\beta > 0$), a qual permanece a mesma para sempre. No entanto, o tempo necessário para isto pode depender drasticamente dos valores de nossos parâmetros α e β .

Vamos considerar a seguir o análogo do processo Flip-Murder, porém em configuração finita. Ele é uma Cadeia de Markov com um conjunto contável Ω de estados chamados *circulares*. As circulares são similares as palavras, pois também são sequências finitas de mais \oplus e menos \ominus , mas agora imaginemos essas sequências tendo a forma circular. Denotemos por $|C|$ o número de componentes em uma circular C . Os índices destas componentes são resto módulo $|C|$ (ver Figura 4 onde $|C| = n$).

Figura 4 – Uma circular C com $|C| = n$.



Fonte: O autor (2020).

(Poderíamos utilizar palavras em vez de circulares, mas isto exigiria definições especiais no fim quando as transformamos.) Na maioria das nossas simulações de Monte Carlo, a circular inicial C consiste de 1000 menos. Em todo experimento simples, o tempo inteiro t cresce de zero a, no máximo, 100000. A circular obtida no tempo t foi denotada por C^t e sua i -ésima componente foi denotada por C_i^t , onde $i = 0, \dots, |C^t| - 1$.

Dizemos que uma palavra $W = (a_1, a_2, \dots, a_n)$ aparece em um lugar i em uma circular $C = (c_1, c_2, \dots, c_m)$ se

$$c_{i+1} = a_1, c_{i+2} = a_2, \dots, c_{i+n} = a_n.$$

Denotemos por $\text{quant}(W|C)$ a quantidade de diferentes lugares onde a palavra W aparece na circular C . Após isto, definimos a *frequência* de W em C como segue:

$$\text{freq}(W|C) = \frac{\text{quant}(W|C)}{|C|}. \quad (3.40)$$

Vamos descrever um procedimento, que chamaremos *Imitação*, e que é uma imitação de Monte Carlo do nosso processo. Este procedimento gera uma sequência de circulares da seguinte maneira indutiva.

Base de indução. A circular inicial C^0 consiste de 1000 menos.

t -ésimo passo de indução. Dada uma circular C^t , onde $t = 0, 1, 2, \dots$ realizamos três procedimentos:

Primeiro procedimento imitando a ação de flip: Toda componente de C^t , que é um menos, torna-se um mais com probabilidade β independentemente de outras componentes.

(Em detalhes mais técnicos, para todo menos em C^t geramos uma nova variável aleatória distribuída uniformemente em $(0, 1)$ e tornamos este menos em um mais se esta variável é menor que β .) Denotamos a circular resultante por $(C')^t$.

Segundo procedimento imitando a ação de murder: Sempre que uma componente de $(C')^t$, que é um mais, é um vizinho esquerdo de uma componente, que é um menos, a componente que é um mais é eliminada da circular com probabilidade α independentemente de outras componentes. (Em detalhes técnicos, para todo tal par geramos uma nova variável aleatória distribuída uniformemente em $(0, 1)$ e realizamos essa eliminação se essa variável é menor que α .) Denotamos a circular resultante por $(C'')^t$.

Terceiro procedimento que ajuda imitar o processo infinito: Dado $(C'')^t$, geramos uma nova circular, a saber C^{t+1} , da seguinte maneira: se $|(C'')^t| < N_{\min}$, onde $N_{\min} = 500$, então C^{t+1} é obtido de $(C'')^t$ concatenando-a com sua cópia e duplicando, assim, seu comprimento; caso contrário $C^{t+1} = (C'')^t$.

Quando paramos: Dada uma constante $T = 100000$, paramos quando $t = T$ ou não existir nenhum menos na circular C^t .

Vamos explicar porque precisamos do terceiro procedimento. Lembremos que, sob a ação do nosso operador, componentes podem desaparecer, mas não podem aparecer; assim, para qualquer $\beta > 0$, o comprimento de qualquer circular finita decresce em média. Além disso, ela conterá todas as suas componentes no estado mais quando o tempo tende para infinito. O terceiro procedimento nos permite adiar isso, e assim, nos ajuda a tornar nossa simulação mais semelhante ao processo infinito.

Assim, o procedimento Imitação está descrito. Usamos isso para vários propósitos em nosso estudo, mas agora o usamos apenas para uma finalidade: atribuir o valor apropriado para a variável Booleana denotada por E (que significa ergodicidade). A saber, E assume o valor *sim* se a última circular C^t não contém menos; caso contrário, E assume o valor *não*. Se $E = \text{sim}$, interpretamos isto como uma sugestão que o processo, com os dados valores de α e β , é ergódico; o resultado $E = \text{não}$ é tomado como uma sugestão que nosso processo é não ergódico.

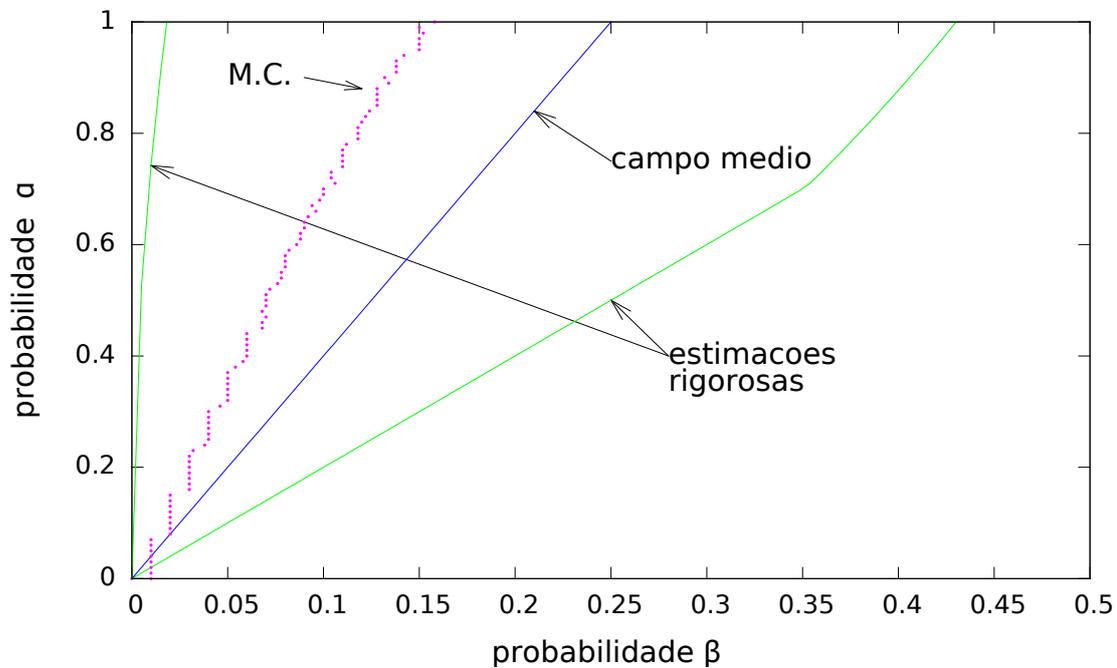
De fato, usamos a Imitação dentro de um ciclo com o crescimento de β : começamos com $\beta = 0$ e então realizamos iterativamente a Imitação e atualizamos β somando 0.001 e repetimos isto até β atingir o valor 1 ou E tenha o valor *sim*, isto é, ergodicidade foi sugerida. Assim, obtemos um certo valor de β . De fato, realizamos este ciclo 5 vezes e registramos a média aritmética dos 5 valores de β assim obtidos.

Lembremos que tudo isso foi feito com um certo valor de α . Na verdade, consideramos 1000 valores de α , a saber, os valores $\alpha_i = 0.001 \cdot i$ para $i = 1, \dots, 1000$. O valor registrado correspondente de β foi denotado por β_i . Assim, obtivemos 1000 pares (α_i, β_i) . O gráfico chamado M. C. na Figura 5 consiste desses pares plotados.

Podemos ver que a “curva” de Monte Carlo não é exatamente uma curva, ela é uma disposição de pontos. No entanto, dá uma ideia do comportamento do nosso processo.

Vemos que as curvas campo médio e M. C. na Figura 5, estão muito mais próximas uma da outra do que as curvas rigorosas. Conjecturamos que elas estão mais próximas da

Figura 5 – Este gráfico mostra ambas as curvas rigorosas (em verde) e duas aproximações para a curva verdadeira que separa as regiões de ergodicidade e não ergodicidade: a aproximação de Campo Médio (campo médio em azul) e a obtida por simulação de Monte Carlo (M.C.).



Fonte: O autor (2020).

curva verdadeira também.

Note que, no gráfico apresentado na Figura 5, abaixo da primeira curva verde (dada pelo Teorema 2) temos a região de ergodicidade do nosso processo, enquanto que acima da segunda curva verde (obtida do Teorema 1) temos a região de não ergodicidade do nosso processo μ_t .

3.6.1 Estimação de $s(\beta, \alpha)$

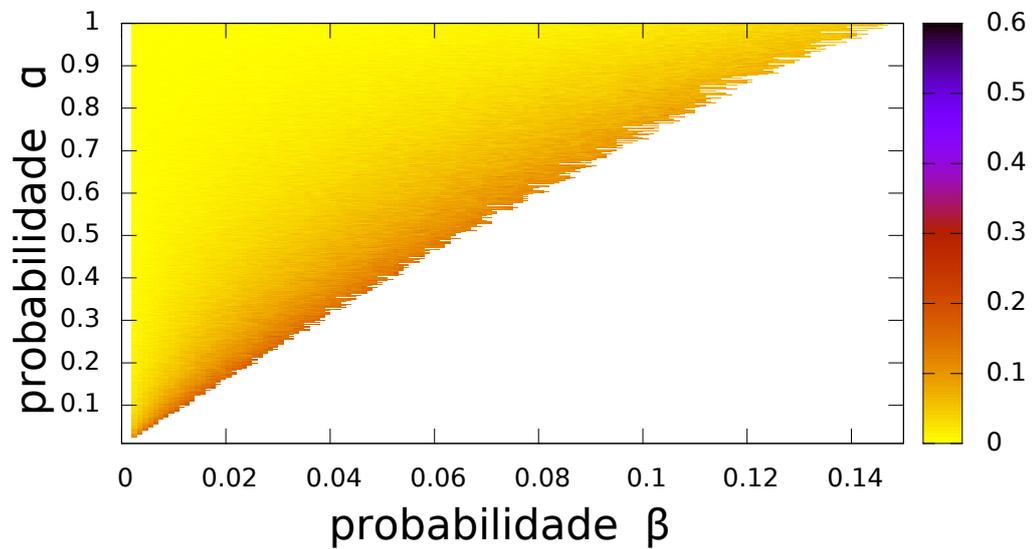
De forma análoga a (TOOM, 2004), definimos a função $s(\beta, \alpha)$ como o supremo da densidade de mais na medida μ_t para todo natural t . Para cada $\alpha \in (\beta/(1 - \beta), 1)$ provamos que a função $s(\beta, \alpha)$ não é contínua como função de β . Queremos estimar $s(\beta, \alpha)$ numericamente, mas estimá-la diretamente seria difícil, assim, em vez disso, estimamos

$$\overline{s(\beta, \alpha)} = \text{freq}(\oplus | C^t) \text{ dado que } t = 100000.$$

A Figura 6 ilustra os valores de $\overline{s(\beta, \alpha)}$. Na área de ergodicidade (área “branca”) $\overline{s(\beta, \alpha)} = 1$. Na outra área, onde nossa aproximação finita sugere não ergodicidade, os valores de $\overline{s(\beta, \alpha)}$, são representados por cores de acordo com a referência apresentada na paleta de cores localizada no lado direito da Figura 6. Todos os valores de $\overline{s(\beta, \alpha)}$, que obtivemos para todas as áreas não ergódicas, foram menor ou igual a 0.6, o que ilustra a

descontinuidade de $s(\beta, \alpha)$ como uma função de β .

Figura 6 – As cores representam os valores de $\overline{s(\beta, \alpha)}$, onde o processo \overline{s} sugere não ergodicidade. A paleta de cores no lado direito, correspondem ao valor de $\overline{s(\beta, \alpha)}$. Quando $\overline{s(\beta, \alpha)} = 1$ a cor associada é branca.



Fonte: O autor (2020).

Capítulo 4

MODELO FLIP-ANIQUILAÇÃO: NOVA ESTIMATIVA E CARACTERIZAÇÕES

4.1 Apresentação dos Teorema 5 e 6

Revisitamos o processo descrito em (TOOM, 2004) e estudado analítica e computacionalmente em (RAMOS e TOOM, 2008; RAMOS e TOOM, 2008; RAMOS *et al.*, 2017; TOOM *et al.*, 2011). Nesse processo, toda partícula tem dois possíveis estados, chamados de *mais* e *menos*. Para cada passo de tempo, duas transformações ocorrem:

- a primeira chamada de *flip* e denotada por F_β transforma qualquer menos em mais com probabilidade β independentemente uns dos outros.
- A segunda, chamada *aniquilação* e denotada por A_α , é imparcial, sob sua ação sempre que um mais é vizinho esquerdo de um menos, ambos desaparecem com probabilidade α , independentemente de outras ocorrências deste tipo.

O operador F_β é um caso especial do operador básico 1, apresentado no Capítulo 2, que chamamos de conversão. Pela notação presente no Capítulo 2, o operador aniquilação é denotado por $(\oplus\ominus \xrightarrow{\alpha} \Lambda)$. Além disso, usando o Resultado 2, apresentado no Capítulo 2, podemos escrever este operador da seguinte maneira:

$$\begin{aligned} \mu A_\alpha &= \mu(\oplus\ominus \xrightarrow{\alpha} \Lambda) \\ &= \mu(\oplus\ominus \xrightarrow{1} \odot)(\odot \xrightarrow{\alpha} \Lambda)(\odot \xrightarrow{1} \oplus\ominus), \end{aligned}$$

note que $\odot \notin \{\oplus, \ominus\}$. Ou seja, o operador aniquilação pode ser escrito como a composição de dois operadores básicos, compressão e descompressão, e um terceiro operador. Como em (TOOM, 2004), denotamos

$$\mu_t = \delta_\ominus (F_\beta A_\alpha)^t, \quad (4.1)$$

onde δ_{\ominus} é a medida concentrada na configuração todos menos, na qual todas as partículas assumem o estado menos. Dentre os resultados descritos em (TOOM, 2004), com um melhoramento em (RAMOS e TOOM, 2008), mostrou-se que existe algum tipo de transição de fase, a qual foi dada pelos resultados: (i) Se $2\beta > \alpha$, a medida μ_t tende para δ_{\oplus} quando $t \rightarrow \infty$; (ii) Para todo natural t a frequência de mais na medida μ_t não excede $250 \cdot \beta/\alpha^2$.

O Teorema 5, apresentado a seguir, nos dá uma região onde μ_t é não ergódico. Na verdade, contém a região de não ergodicidade, descrita em (ii) por (TOOM, 2004) e, posteriormente, melhorada por (RAMOS e TOOM, 2008). Mostraremos que:

Teorema 5. *Seja $\alpha \in (0, 1)$. Se $\beta < \frac{9\alpha^2}{1000}$,*

(A. 5) *então, para todo valor natural t , $\mu_t(\oplus) < 1$, e*

(B. 5) *existe uma medida, ν , tal que $\nu(F_{\beta}A_{\alpha}) = \nu$, onde $\nu(\oplus) < 1$.*

O Teorema 5 vem colaborar no sentido de apresentar uma melhor aproximação para a curva crítica (caso exista), que foi estimada por meio de uma simulação computacional (RAMOS e TOOM, 2008), e que separa as regiões de ergodicidade e não ergodicidade do processo μ_t . Esta curva é apresentada neste trabalho na Figura 4.8(b).

Teorema 6. *Seja $\mu \in \mathcal{M}_{\{\ominus, \oplus\}}$. Se $\mu(\ominus) \in \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$, então $\mu(F_{\beta}A_{\alpha})^t \rightarrow \delta_{\oplus}$, quando $t \rightarrow \infty$.*

Em (TOOM, 2004) temos o seguinte resultado:

Resultado 3. *Tome qualquer $\mu \in \mathcal{M}_{\{\ominus, \oplus\}}$ e suponha que $\beta > 0$ e $(1-\beta)\mu(\ominus) \leq 1/2$. Então as medidas $\mu(F_{\beta}A_{\alpha})^t$ tendem para δ_{\oplus} , quando $t \rightarrow \infty$.*

O Teorema 6 vem colaborar no sentido de abranger novas medidas de probabilidades que não eram contempladas pelo Resultado 3. No entanto, há duas situações em que esses resultados têm alguma relação, a saber:

- Se $\alpha \leq \beta$, então o Resultado 3 implica em nosso Teorema 6.
- Se $\beta < \alpha$, então nosso Teorema 6 implica no Resultado 3.

Vejamos a seguir dois exemplos que mostram que esses resultados não são equivalentes. Estamos considerando valores de α e β de modo que: se $\alpha \leq \beta$, então $\frac{1}{2(1-\beta)} < 1$; se $\beta < \alpha$, então $\frac{\beta}{2\alpha(1-\beta)} < 1$.

Exemplo 1. *No caso em que $\alpha \leq \beta$, consideremos o caso em que $\alpha = 0.3$ e $\beta = 0.33$, daí $\frac{1}{2(1-\beta)} \simeq 0.7462$ e $\frac{\beta}{2\alpha(1-\beta)} \simeq 0.8208$. Assim, se $\mu(\ominus) = 0.8$, então $\mu(\ominus) \in \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$, mas $\mu(\ominus) \notin \left(0, \frac{1}{2(1-\beta)}\right)$.*

Exemplo 2. No caso em que $\beta < \alpha$, consideremos o caso em que $\alpha = 0.4$ e $\beta = 0.34$, daí $\frac{\beta}{2\alpha(1-\beta)} \simeq 0.5357$ e $\frac{1}{2(1-\beta)} \simeq 0.7142$. Assim, se $\mu(\Theta) = 0.6$, então $\mu(\Theta) \in \left(0, \frac{1}{2(1-\beta)}\right)$, mas $\mu(\Theta) \notin \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$.

4.2 Prova dos Teoremas 5 e 6

Prova do Teorema 6. Inicialmente, vejamos que $\mu F_\beta A_\alpha(\Theta) < \mu(\Theta)$. Da definição de A_α ,

$$\mu F_\beta A_\alpha(\Theta) = \frac{\mu F_\beta(\Theta) - \alpha \mu F_\beta(\oplus, \Theta)}{1 - 2\alpha \mu F_\beta(\oplus, \Theta)},$$

e, da consistência da medida, segue que, $\mu(\Theta) = \mu(\Theta, \Theta) + \mu(\oplus, \Theta)$. Uma vez que $\mu(\Theta, \Theta) \geq 0$, concluímos que $\mu(\Theta) \geq \mu(\oplus, \Theta)$. Consequentemente, para todo $\beta \in (0, 1)$,

$$\mu F_\beta(\oplus, \Theta) \leq \mu F_\beta(\Theta)$$

e

$$1 - 2\alpha \mu F_\beta(\oplus, \Theta) \geq 1 - 2\alpha \mu F_\beta(\Theta).$$

Portanto,

$$\begin{aligned} \frac{\mu F_\beta(\Theta) - \alpha \mu F_\beta(\oplus, \Theta)}{1 - 2\alpha \mu F_\beta(\oplus, \Theta)} &\leq \frac{\mu F_\beta(\Theta) - \alpha \mu F_\beta(\oplus, \Theta)}{1 - 2\alpha \mu F_\beta(\Theta)} \\ &< \frac{\mu F_\beta(\Theta)}{1 - 2\alpha \mu F_\beta(\Theta)} \\ &= \frac{(1-\beta)\mu(\Theta)}{1 - 2\alpha(1-\beta)\mu(\Theta)}. \end{aligned}$$

Logo é suficiente mostrarmos que $\frac{(1-\beta)\mu(\Theta)}{1 - 2\alpha(1-\beta)\mu(\Theta)} < \mu(\Theta)$. Note que,

$$\begin{aligned} \frac{(1-\beta)\mu(\Theta)}{1 - 2\alpha(1-\beta)\mu(\Theta)} < \mu(\Theta) &\Leftrightarrow (1-\beta)\mu(\Theta) < \mu(\Theta) - 2\alpha(1-\beta)[\mu(\Theta)]^2 \\ &\Leftrightarrow \mu(\Theta)[2\alpha(1-\beta)\mu(\Theta) - \beta] < 0. \end{aligned} \quad (4.2)$$

Como $\mu(\Theta) > 0$, (4.2) só é possível quando $2\alpha(1-\beta)\mu(\Theta) - \beta < 0$. Portanto, se $\mu(\Theta) \in \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$, então $\mu F_\beta A_\alpha(\Theta) < \mu(\Theta)$.

Note que $\mu(F_\beta A_\alpha)^t(\Theta) \in \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$ para todo $t \in \mathbb{N}$.

- Primeiro vamos mostrar para $t = 1$. Como $\mu(\Theta) \in \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$, $\mu F_\beta A_\alpha(\Theta) < \mu(\Theta)$, logo $\mu F_\beta A_\alpha(\Theta) \in \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$.

- Agora, supondo que o resultado é válido para t , vamos mostrar para $t + 1$. Note que,

$$\mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^{t+1}(\Theta) = \mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\mathbf{F}_\beta \mathbf{A}_\alpha)(\Theta).$$

Como, por hipótese de indução $\mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\Theta) \in \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$, temos, da parte inicial, que $\mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\mathbf{F}_\beta \mathbf{A}_\alpha)(\Theta) < \mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\Theta)$, portanto $\mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^{t+1}(\Theta) \in \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$.

Deste fato, segue que a sequência $\mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\Theta)$ é monótona decrescente. Com efeito, como $\mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\Theta) \in \left(0, \frac{\beta}{2\alpha(1-\beta)}\right)$ para todo $t \in \mathbb{N}$, segue, da parte inicial, que

$$\mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^{t+1}(\Theta) = \mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\mathbf{F}_\beta \mathbf{A}_\alpha)(\Theta) < \mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\Theta).$$

Como a sequência $\mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\Theta)$ é monótona decrescente e limitada inferiormente por zero, então $\mu(\mathbf{F}_\beta \mathbf{A}_\alpha)^t(\Theta) \rightarrow 0$, quando $t \rightarrow \infty$. \square

Assim como no Capítulo 3, para provar o item (A. 5) do Teorema 5, utilizamos duas ideias bem conhecidas na literatura: o método de contorno de Peierls e a dualidade dos grafos planares, (GRIMMETT, 1999), as quais são utilizadas nos processos de contato (LIGGETT, 1997) e processos de Stavskaya (TOOM *et al.*, 2001). Após implementada esta metodologia, a seguinte desigualdade foi atingida, a qual será nosso ponto de partida.

$$\mu_t(\oplus) \leq \sum_{x=1}^{\infty} \sum_{y=0}^{\infty} \sum_{z=1}^{\infty} x \cdot [S_3(x, y, z) + S_4(x, y, z)]. \quad (4.3)$$

O número $S_k(x, y, z)$ satisfaz à condição inicial

$$S_k(x, y, 1) = \begin{cases} 1 & \text{se } x = 0, y = -1, \text{ e } k = 1, \\ 0 & \text{em todos os outros casos,} \end{cases}$$

e satisfaz também às equações de transição

$$\begin{cases} S_1(x, y, z+1) &= S_1(x, y+1, z) + S_2(x, y+1, z) + S_3(x, y+1, z), \\ S_2(x, y, z+1) &= 2\beta[S_1(x-1, y, z) + S_2(x-1, y, z) + S_3(x-1, y, z) + S_4(x-1, y, z)], \\ S_3(x, y, z+1) &= \alpha \cdot [S_2(x+1, y-1, z) + S_3(x+1, y-1, z) + S_4(x+1, y-1, z)], \\ S_4(x, y, z+1) &= (1-\alpha) \cdot [S_2(x, y-1, z) + S_3(x, y-1, z) + S_4(x, y-1, z)]. \end{cases} \quad (4.4)$$

Para obter um limite superior para a soma (4.3), utilizamos

$$\left\{ \begin{array}{l} S_1(z) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} p^{-x} q^{-y} S_1(x, y, z), \\ S_2(z) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} p^{-x} q^{-y} S_2(x, y, z), \\ S_3(z) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} p^{-x} q^{-y} S_3(x, y, z), \\ S_4(z) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} p^{-x} q^{-y} S_4(x, y, z), \end{array} \right. \quad (4.5)$$

onde p e q são parâmetros positivos e p e q são tais que $x < p^{-x}$ e $1 < q^{-y}$. Note que o lado direito de (4.3) é limitado superiormente por

$$\sum_{z=1}^{\infty} (S_3(z) + S_4(z)), \quad (4.6)$$

Assim, é suficiente estimar (4.6).

Como visto em (RAMOS e TOOM, 2008; TOOM *et al.*, 2011) as somas em (4.5) satisfazem a condição inicial

$$S_1(1) = q, \quad S_2(1) = S_3(1) = S_4(1) = 0,$$

o que é uma correção do que é apresentado em (TOOM, 2004).

Observe que de (4.4) e (4.5) as condições de recorrência são:

$$\left\{ \begin{array}{l} S_1(z+1) = q(S_1(z) + S_2(z) + S_3(z)), \\ S_2(z+1) = \frac{2\beta}{p}(S_1(z) + S_2(z) + S_3(z) + S_4(z)), \\ S_3(z+1) = \frac{p\alpha}{q}(S_2(z) + S_3(z) + S_4(z)), \\ S_4(z+1) = \frac{1-\alpha}{q}(S_2(z) + S_3(z) + S_4(z)). \end{array} \right. \quad (4.7)$$

Com o intuito de escrever (4.7) em forma matricial, introduzimos o vetor $S(z) = (S_1(z), S_2(z), S_3(z), S_4(z))$, o que nos permite colocar as condições de recorrência da seguinte maneira $S(z+1) = S(z) \cdot M$, logo, $S(z) = S(1) \cdot M^{z-1}$, onde M é a matriz

$$M = \begin{pmatrix} q & 2\beta/p & 0 & 0 \\ q & 2\beta/p & p\alpha/q & (1-\alpha)/q \\ q & 2\beta/p & p\alpha/q & (1-\alpha)/q \\ 0 & 2\beta/p & p\alpha/q & (1-\alpha)/q \end{pmatrix}.$$

Note que nossa matriz M não é a mesma descrita por (TOOM, 2004). Naquele trabalho

utilizou-se a relação entre os elementos da terceira e quarta colunas da matriz M que adotamos.

Pelo Lema 16 (ver Apêndice),

$$S(z) \leq S(1) \cdot N^{z-1}, \quad (4.8)$$

em que

$$N = \begin{pmatrix} q & 2\beta/p & 0 & 0 \\ q & 2\beta/p & p\alpha/q & (1-\alpha)/q \\ q & 2\beta/p & p\alpha/q & (1-\alpha)/q \\ q & 2\beta/p & p\alpha/q & (1-\alpha)/q \end{pmatrix}.$$

A matriz N possui 4 autovalores onde dois deles são iguais a zero, e possui autovetores gerados por $(0, 1, 0, -1)$ e $(0, 0, 1, -1)$. O autovalor de Perron-Frobenius é denotado por λ_{PF} , (para mais detalhes sobre o autovalor de Perron-Frobenius ver Teorema 1.1 em (SENETA, 2006)) e λ_2 é o último destes 4 autovalores.

$$\lambda_{PF} = \frac{q + 2\beta/p + (p\alpha + 1 - \alpha)/q + \sqrt{\Delta}}{2} \quad \text{e} \quad \lambda_2 = \frac{q + 2\beta/p + (p\alpha + 1 - \alpha)/q - \sqrt{\Delta}}{2},$$

$$\text{onde } \Delta = \left(q + \frac{2\beta}{p} + \frac{\alpha(p-1)+1}{q} \right)^2 - 4(\alpha(p-1)+1).$$

Se $\lambda_{PF} < 1$, então $\beta < 11\alpha^2/200(3-\alpha)$. Além disso, fixados p, q e α , λ_{PF} é uma função crescente em relação a β . Logo, se assumirmos

$$\beta < \frac{9\alpha^2}{1000}, \quad (4.9)$$

então $\lambda_{PF} < 1$. E esse é o limite superior de β que adotaremos de agora em diante.

A matriz N é diagonalizável, o que nos permite escrever $N^z = P^{-1}D^zP$ (ver Lema 15 no Apêndice). Escrevendo $N^z = (n_{ij}^z)_{i,j=1,2,3,4}$ obtemos que

$$n_{1j}^z = \lambda_{PF}^z \tilde{p}_{11} p_{1j} + \lambda_2^z \tilde{p}_{12} p_{2j}, \quad \forall j = 1, 2, 3, 4, \quad (4.10)$$

onde $P = (p_{ij})_{i,j=1,2,3,4}$ é a matriz dos autovetores e $P^{-1} = (\tilde{p}_{ij})_{i,j=1,2,3,4}$ é sua inversa.

Os autovetores associados a λ_{PF} e λ_2 são,

$$\left(\frac{\left(\lambda_i - \frac{\alpha(p-1)+1}{q} \right) q^2}{2\alpha\beta} v_3, \frac{\left(\lambda_i - \frac{\alpha(p-1)+1}{q} \right) q}{p\alpha} v_3, v_3, \frac{1-\alpha}{p\alpha} v_3 \right) \quad \text{para } v_3 \in \mathbb{R} \setminus \{0\} \text{ e } i \in \{PF, 2\}.$$

Tomaremos representantes unitários² positivos dos autovetores, o que certamente é possível para λ_{PF} e também possível para λ_2 , tendo-se em vista sua expressão algébrica. Devido a

²Um vetor $p = (p_1, \dots, p_n)$ é unitário, se $\sum_{i=1}^n p_i = 1$.

nossa escolha dos autovetores, sabemos que $\sum_{j=1}^4 p_{1j} = \sum_{j=1}^4 p_{2j} = 1$. Podemos verificar que

$$\tilde{p}_{11} = \frac{p_{22} + p_{23} + p_{24}}{\xi} \quad \text{e} \quad \tilde{p}_{12} = -\frac{p_{12} + p_{13} + p_{14}}{\xi},$$

em que $\xi = p_{11}(p_{22} + p_{23} + p_{24}) - p_{21}(p_{12} + p_{13} + p_{14}) = p_{11} - p_{21}$.

Substituindo o valor de ξ em \tilde{p}_{11} e \tilde{p}_{12} temos

$$\tilde{p}_{11} = \frac{1 - p_{21}}{p_{11} - p_{21}} \quad \text{e} \quad \tilde{p}_{12} = \frac{p_{11} - 1}{p_{11} - p_{21}}. \quad (4.11)$$

Assim, $\tilde{p}_{11} + \tilde{p}_{12} = 1$.

Para que os autovetores associados a λ_{PF} e λ_2 sejam unitários é suficiente usarmos

$$v_3 = \frac{2\alpha\beta/q}{(\lambda_i - (\alpha(p-1) + 1)/q)(q + 2\beta/p) + 2\beta/p(\alpha(p-1) + 1)/q} \quad \text{para } i \in \{PF, 2\}.$$

Assumindo (4.9) e usando (4.8) e (4.10),

$$\begin{aligned} \sum_{z=1}^{\infty} (S_3(z) + S_4(z)) &\leq q \left(\sum_{z=0}^{\infty} (n_{13}^z + n_{14}^z) \right) \\ &= q \left(\sum_{z=0}^{\infty} (\lambda_{PF}^z \tilde{p}_{11} p_{13} + \lambda_2^z \tilde{p}_{12} p_{23}) + \sum_{z=0}^{\infty} (\lambda_{PF}^z \tilde{p}_{11} p_{14} + \lambda_2^z \tilde{p}_{12} p_{24}) \right) \\ &= q \left(\frac{\tilde{p}_{11}(p_{13} + p_{14})}{1 - \lambda_{PF}} + \frac{\tilde{p}_{12}(p_{23} + p_{24})}{1 - \lambda_2} \right) \\ &= q \left(\frac{\tilde{p}_{11}(p_{13} + p_{14})}{1 - \lambda_{PF}} + \frac{(1 - \tilde{p}_{11})(p_{23} + p_{24})}{1 - \lambda_2} \right), \end{aligned} \quad (4.12)$$

em que,

$$\begin{aligned} p_{11} &= \frac{q\lambda_{PF} - (\alpha(p-1) + 1)}{\lambda_{PF}(q + 2\beta/p) - (\alpha(p-1) + 1)}, & p_{21} &= \frac{q\lambda_2 - (\alpha(p-1) + 1)}{\lambda_2(q + 2\beta/p) - (\alpha(p-1) + 1)}, \\ p_{13} &= \frac{2\alpha\beta/q}{\lambda_{PF}(q + 2\beta/p) - (\alpha(p-1) + 1)}, & p_{23} &= \frac{2\alpha\beta/q}{\lambda_2(q + 2\beta/p) - (\alpha(p-1) + 1)}, \\ p_{14} &= \frac{2\beta(1 - \alpha)/pq}{\lambda_{PF}(q + 2\beta/p) - (\alpha(p-1) + 1)}, & p_{24} &= \frac{2\beta(1 - \alpha)/pq}{\lambda_2(q + 2\beta/p) - (\alpha(p-1) + 1)}. \end{aligned}$$

Denotamos (4.12) por

$$LS(p, q, \beta, \alpha). \quad (4.13)$$

Prova do item (A. 5) do Teorema 5. Considerando $p = \frac{3}{10}$, $q = 1 - \frac{\alpha}{3}$ com $\alpha \in (0, 1)$, temos pelo Lema 18 (ver Apêndice) que (4.13) é uma função monótona crescente em relação a β , para $\beta \in \left[0, \frac{9\alpha^2}{1000}\right]$. Assim, se $LS\left(\frac{3}{10}, 1 - \frac{\alpha}{3}, \frac{9\alpha^2}{1000}, \alpha\right)$ for menor que 1, temos que

$LS(\frac{3}{10}, 1 - \frac{\alpha}{3}, \beta, \alpha)$ é menor que 1 para $\beta \in (0, \frac{9\alpha^2}{1000})$.

Fazendo $\beta = \frac{9\alpha^2}{1000}$ temos que (4.13) é função somente de α . Denotamos esta função, para a qual foi preciso utilizar uma certa computação algébrica, por $g(\alpha)$, e a descreveremos por:

$$g(\alpha) = \frac{-(36(-3 + \alpha)\alpha^2(9\alpha^3 - 77\alpha^2 + 465\alpha - 450))}{(-9\alpha^3 + f(\alpha)\alpha + 77\alpha^2 - 3f(\alpha) - 315\alpha)(9\alpha^3 + f(\alpha)\alpha - 77\alpha^2 - 3f(\alpha) + 315\alpha)}.$$

Então, $g(\alpha)$ é uma função contínua, monótona decrescente para $\alpha \in [0, 1]$ e assume valor menor que 1 para $\alpha = 0$ e maior que zero para $\alpha = 1$ (ver Lema 17 no Apêndice). \square

Prova do item (B. 5) do Teorema 5. Utilizaremos o Resultado 1, do Capítulo 2, assumindo $R = \{\oplus\} \subset \mathbb{A}^{\mathbb{Z}}$, isto é, $R = \{x \in \mathcal{A}^{\mathbb{Z}} : x_{i_0} = \oplus \text{ para algum } i_0 \in \mathbb{Z}\}$. Seja $\delta = LS(p, q, \beta, \alpha)$, onde $p = \frac{3}{10}$, $q = 1 - \frac{\alpha}{3}$ e $\beta \in (0, \frac{9\alpha^2}{1000})$. Da demonstração de (A. 5) temos que

$$\mu_t(\oplus) = \delta_{\ominus}(\mathbf{F}_{\beta}\mathbf{A}_{\alpha})^t(\oplus) < LS(p, q, \beta, \alpha), \quad \text{para todo } t.$$

Logo, pelo Resultado 1, o operador $\mathbf{F}_{\beta}\mathbf{A}_{\alpha}$ tem uma medida invariante ν , isto é, $\nu(\mathbf{F}_{\beta}\mathbf{A}_{\alpha}) = \nu$ tal que

$$\nu(\oplus) < LS(p, q, \beta, \alpha) < 1.$$

\square

Poderíamos utilizar a técnica apresentada no Teorema 3 em (TOOM, 2007) para provar (B. 5) de outra forma. Contudo, isto traz a necessidade de uma gama de definições, o que não torna viável, pois foge do objetivo deste trabalho.

Note que, para $\beta < \frac{9\alpha^2}{1000}$, o operador $\mathbf{F}_{\beta}\mathbf{A}_{\alpha}$ tem pelo menos duas medidas invariantes distintas, a saber, a medida δ_{\oplus} e a medida ν de (B. 5).

Como apresentado em nosso texto, o processo $\mu_t = \delta_{\ominus}(\mathbf{F}_{\beta}\mathbf{A}_{\alpha})^t$, é ergódico para $\beta > \alpha/2$, e provamos que ele é não ergódico para $\beta < \frac{9\alpha^2}{1000}$. Supor que exista alguma curva que delimite a transição entre estes comportamentos é uma hipótese bem natural, chamamos de *curva verdadeira* a estimação computacional que obtemos para essa curva de transição de fase entre as regiões de ergodicidade e não ergodicidade desse processo. Os resultados, citados anteriormente, embora rigorosos, deixam uma lacuna entre tais regiões. Sendo assim, (RAMOS e TOOM, 2008) apresentaram estimacões computacionais da curva verdadeira. Fizeram isso de duas maneiras: em uma utilizaram a aproximação de campo médio, e na outra, simulação de Monte Carlo. Assim, obtiveram duas curvas, que são muito mais próximas uma da outra.

Com o objetivo de ilustrar nossos resultados, $\beta < \frac{9\alpha^2}{1000}$, que melhora a região de não ergodicidade do processo μ_t , em relação ao resultado encontrado em (TOOM, 2004), o qual posteriormente foi melhorado em (RAMOS e TOOM, 2008), a saber, $\beta < \frac{\alpha^2}{250}$, exibiremos as Figuras 7(a) e 7(b) a seguir. Também mostramos a curva verdadeira obtida

por meio de simulações computacionais, a qual apresenta uma melhor estimativa sobre as regiões de ergodicidade vs. não ergodicidade do processo μ_t . Na Figura 7(a) apresentada a seguir, observamos a região de não ergodicidade do processo μ_t . A curva tracejada, a saber $\beta = \frac{\alpha^2}{250}$, foi obtida em (TOOM, 2004; RAMOS e TOOM, 2008; TOOM *et al.*, 2011) e a curva contínua, a saber $\beta = \frac{9\alpha^2}{1000}$, foi obtida por nós neste trabalho. Obtivemos uma curva teórica que melhor amplia a região de não ergodicidade de μ_t . Na Figura 7(b), além das curvas apresentadas na Figura 7(a), colocamos a curva verdadeira obtida por meio de simulação computacional, as explicações detalhadas da simulação podem ser obtidas em (RAMOS e TOOM, 2008).

4.3 Ilhas e Arquipélagos

Sejam $x, y \in \{\ominus, \oplus\}^{\mathbb{Z}}$. Dizemos que estas duas configurações, x e y , são *próximas* uma da outra se o conjunto $\{i \in \mathbb{Z} : x_i \neq y_i\}$ é finito. Uma configuração é chamada uma *ilha de menos* se ela está próxima da configuração “todos mais”. Denotamos o conjunto das ilhas de menos, Δ . Se $x \in \Delta$, definimos a *população* de x , e denotamos por $\text{Pop}(x)$, como a quantidade de menos na ilha x , isto é,

$$\text{Pop}(x) = \#\{i \in \mathbb{Z} : x_i = \ominus\},$$

onde $\#(\cdot)$ denota a cardinalidade do conjunto.

Denotemos por \mathcal{A}_{\ominus} o conjunto das medidas normalizadas em Δ . Uma medida pertencente a \mathcal{A}_{\ominus} é chamada um *arquipélago de menos*. Deste ponto em diante, se não for feita menção contrária, μ denota um arquipélago de menos.

O operador $F_{\beta} : \mathcal{A}_{\ominus} \rightarrow \mathcal{A}_{\ominus}$ está bem definido. O operador A_{α} é definido para atuar em medidas uniformes, isto é, invariantes por translação. Mas, sua ação em ilhas de menos é bastante natural. Logo, sua ação em arquipélagos de menos também. Assim, temos $A_{\alpha} : \mathcal{A}_{\ominus} \rightarrow \mathcal{A}_{\ominus}$.

Dado um arquipélago de menos μ , definiremos a variável aleatória

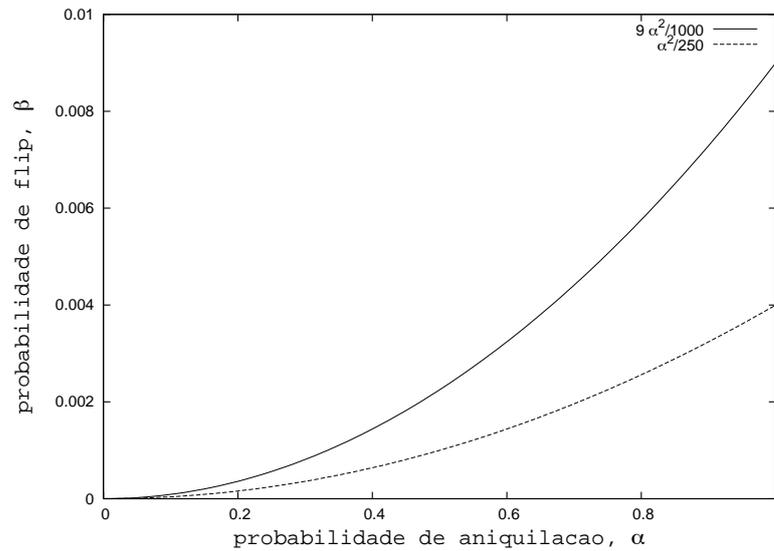
$$\tau_{\mu} = \inf \{t \geq 0 : \mu(F_{\beta}A_{\alpha})^t = \delta_{\oplus}\}.$$

Sendo o ínfimo do conjunto vazio, infinito. A variável aleatória τ_{μ} denota o tempo para atingir a configuração “todos mais”, dado que começamos nosso processo, $F_{\beta}A_{\alpha}$, com a medida μ .

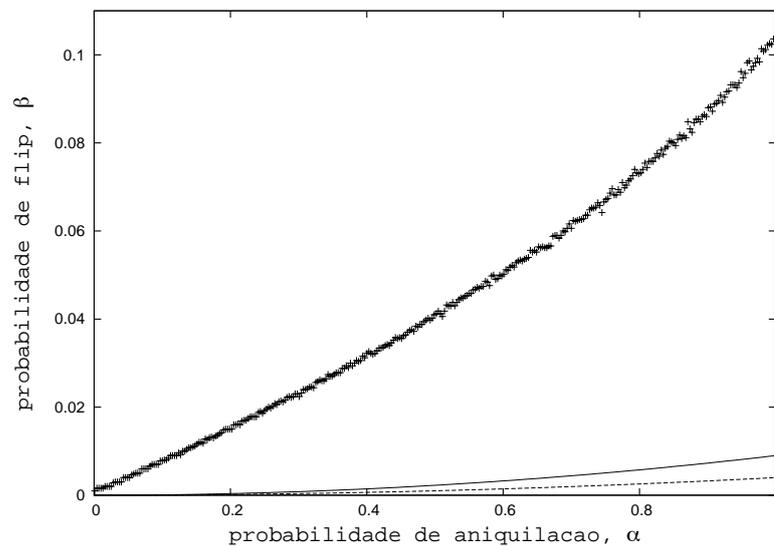
Note que $\mu = \sum_{i=1}^{\infty} k_i \delta_{x^i}$, onde $k_1 > 0, k_2 > 0, \dots, k_1 + k_2 + \dots = 1$ e $x^i \in \Delta$ para $i \in \mathbb{N}$. Definimos a *máxima população* de μ por

$$\text{Max}(\mu) = \max \left\{ \text{Pop}(x^i) : \mu = \sum_{i=1}^{\infty} k_i \delta_{x^i} \right\}.$$

Figura 7 – Em ambas as figuras descrevemos as curvas em que $\beta = \frac{\alpha^2}{250}$ e $\beta = \frac{9\alpha^2}{1000}$. Em 7(b), plotamos os dados obtidos por simulação computacional do processo. Esses dados estimam a curva crítica. Se (α, β) estão abaixo da curva crítica estimada, o processo μ_t é não ergódico, caso contrário, o processo μ_t é ergódico.



(a)



(b)

Fonte: O autor (2020).

Se não existe tal máximo, dizemos que $\text{Max}(\mu) = \infty$.

Seja δ_x a medida concentrada em $x \in \Delta$. Neste caso, é claro que se $\alpha > 0$ ou $\beta > 0$ teremos que $\delta_x(\mathbf{F}_\beta \mathbf{A}_\alpha)^t$ tende para δ_\oplus quando t tende para infinito. Contudo, não fica evidente sobre o quão veloz ocorre esta convergência. Também não sabemos como esta quantidade de menos se comporta em cada passo de tempo. Os Teoremas 7 e 8 nos fornecerão informações nessa direção.

Geralmente, quando falamos em autômatos celulares probabilísticos, \mathbf{P} , chamamos de processo a sequência de medidas

$$\mu, \mu\mathbf{P}, \mu\mathbf{P}^2, \dots, \mu\mathbf{P}^t, \dots$$

para alguma medida inicial μ . Mas podemos definir processos considerando uma sequência

$$x^0, x^1, x^2, \dots, x^t, \dots \quad (4.14)$$

onde x^0 tem distribuição μ e x^t tem distribuição $\mu\mathbf{P}^t$. A forma (4.14) de definir processo é natural e pode ser encontrada em (MAIRESSE e MARCOVICI, 2014).

Para $x \in \Delta$ considere (4.14) para descrever o processo. Seja δ_x a distribuição concentrada em x e $\delta_x(\mathbf{F}_\beta \mathbf{A}_\alpha)^t$ a distribuição de x^t , definimos a população de x no tempo t por $\text{Pop}(x^t)$, esta variável aleatória representa a quantidade de menos na ilha x , após a aplicação do operador $\mathbf{F}_\beta \mathbf{A}_\alpha$, t vezes. Dada $\mu \in \mathcal{A}_\ominus$, definimos a máxima população de μ no tempo t por

$$\text{Max}(\mu, t) = \max \left\{ \text{Pop}((x^i)^t) : \mu\mathbf{P}^t = \sum_{i=1}^{\infty} k_i \delta_{(x^i)^t} \right\},$$

onde $(x^i)^t$ denota a i -ésima ilha no tempo t e, se não existe tal máximo, dizemos que $\text{Max}(\mu, t) = \infty$.

Teorema 7. *Seja $\mu \in \mathcal{A}_\ominus$ tal que $\text{Max}(\mu) < \infty$. Então,*

$$\mathbb{E}(\tau_\mu) \leq \begin{cases} f_\mu(\alpha), & \text{se } \alpha \geq \frac{\text{Max}(\mu)}{g_\mu(\beta)}, \\ g_\mu(\beta), & \text{se } \alpha < \frac{\text{Max}(\mu)}{g_\mu(\beta)} \end{cases}$$

$$\text{onde } f_\mu(\alpha) = \frac{\text{Max}(\mu)}{\alpha} \text{ e } g_\mu(\beta) = \sum_{i=1}^{\text{Max}(\mu)} \frac{1}{1 - (1 - \beta)^i}.$$

Teorema 8. *Dado um valor natural t e $\mu \in \mathcal{A}_\ominus$. Se $\text{Max}(\mu, t)$ é finito, então*

$$\mathbb{E}(\text{Max}(\mu, t)) \leq (1 - \beta)^{t-1} \left(\text{Max}(\mu) + \alpha\beta^{\text{Max}(\mu)} - \alpha \right).$$

A prova dos Teoremas 7 e 8 serão apresentadas na seção 4.9.

4.4 Distância de Variação Total

Lembremos que Δ é o conjunto das ilhas de menos, isto é, se $y \in \Delta$, então $\#\{i \in \mathbb{Z} : y_i \neq \oplus\}$ é finito. Lembremos também que \mathcal{A}_\ominus é o conjunto das medidas normalizadas na σ -álgebra gerada pelos cilindros em Δ , de modo que, se $\mu \in \mathcal{A}_\ominus$, então

$$\mu = \sum_{y \in \Delta} k_y \delta_y.$$

Desta maneira, mostraremos a distância de variação total entre δ_\oplus e $\mu \in \mathcal{A}_\ominus$. A qual é expressa, como em (LEVIN *et al.*, 2008), por

$$\|\delta_\oplus - \mu\|_{VT} = \sup_{A \in \mathcal{F}} |\delta_\oplus(A) - \mu(A)|,$$

onde \mathcal{F} é a σ -álgebra gerada pelos cilindros em Δ . Esta definição da distância de variação total é um supremo sobre todos os subconjuntos de \mathcal{F} , portanto usá-la nem sempre é a maneira mais conveniente de estimar a distância.

Como em (LEVIN *et al.*, 2008), vamos dar uma caracterização alternativa bastante útil. Uma vez que Δ é enumerável,

$$\|\delta_\oplus - \mu\|_{VT} = \frac{1}{2} \sum_{x \in \Delta} |\delta_\oplus(x) - \mu(x)|. \quad (4.15)$$

Por meio desta caracterização podemos mostrar o seguinte Lema.

Lema 7. *Seja $\mu \in \mathcal{A}_\ominus$, então*

$$\|\delta_\oplus - \mu\|_{VT} = 1 - \mu(\text{“todos mais”})$$

Prova. Usando (4.15),

$$\begin{aligned} \|\delta_\oplus - \mu\|_{VT} &= \frac{1}{2} \sum_{x \in \Delta} \left| (1 - k_\oplus) \delta_\oplus(\text{todos mais}) - \sum_{y \in \Delta \setminus \{\text{todos mais}\}} k_y \delta_y(x) \right| \\ &= \frac{1}{2} \left((1 - k_\oplus) + \sum_{x \in \Delta \setminus \{\text{todos mais}\}} k_x \right) \\ &= \frac{1}{2} \left(1 - \mu(\text{todos mais}) + \sum_{x \in \Delta \setminus \{\text{todos mais}\}} \mu(x) \right) \\ &= \frac{1}{2} (1 - \mu(\text{todos mais}) + 1 - \mu(\text{todos mais})) \\ &= 1 - \mu(\text{todos mais}). \end{aligned}$$

□

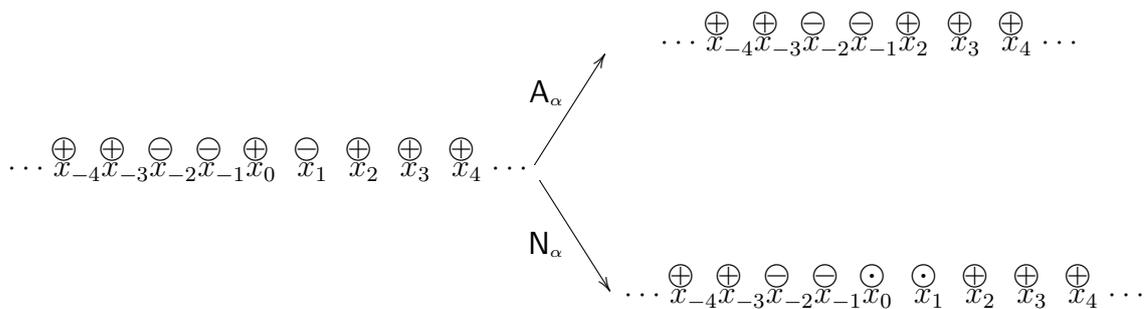
Este resultado nos fornece uma caracterização das nossas distribuições iniciais, $\mu \in \mathcal{A}_\ominus$, em relação a nossa distribuição invariante δ_\oplus , por meio da sua distância de variação total.

4.5 Os Operadores Neutralização N_α , e Eliminação E_α

Seja $\Omega_\ominus = \{\oplus, \ominus, \odot\}^{\mathbb{Z}}$, onde \odot é chamado de *ponto*. Denotamos por Δ_\ominus o conjunto que contém todo $x \in \Omega_\ominus$ tal que $\text{Pop}(x)$ é finita. Logo, há uma cópia de Δ em Δ_\ominus , esse fato será utilizado nesta seção. Vale notar que, se $x \in \Delta_\ominus$, então $\text{Pop}(x)$ é finito, no entanto podem haver infinitas posições que assumem os estados mais e ponto. Ou seja, se $x \in \Delta_\ominus$, então existem índices i_0, j_0 , com $i_0 < j_0$ tal que $x_{i_0} = x_{j_0} = \ominus$ e $x_k \neq \ominus$ para todos $k < i_0$ e $k > j_0$. Agora, definimos \mathcal{A}_\ominus como sendo o conjunto das medidas normalizadas na σ -álgebra gerada pelos cilindros em Δ_\ominus . Vamos considerar um operador, que chamaremos *neutralização*, $N_\alpha : \mathcal{A}_\ominus \rightarrow \mathcal{A}_\ominus$, que age da seguinte maneira: toda palavra $\oplus \odot^n \ominus$ é transformada na palavra \odot^{n+2} com probabilidade α , ou permanece inalterada com probabilidade $1 - \alpha$, para todo $n \in \mathbb{N}$. Onde \odot^n é a palavra consistindo de n letras iguais a \odot . Em particular, denotamos $\oplus \odot^0 \ominus$ por $\oplus \ominus$.

Note que N_α não é de comprimento variável, pois, neste processo nossas partículas nunca desaparecem. Se uma partícula vai para o estado \odot , ela permanece neste estado para sempre. Além disso, sob a ação de N_α as partículas nos estados mais e menos interagem como se as partículas no estado ponto não existissem.

Figura 8 – Ilustração da ação dos operadores N_α e A_α numa ilha cujo $\text{Pop}(x) = 3$.



Fonte: O autor (2020).

Na Figura 8, ilustramos uma possível ação, isto é, tem probabilidade positiva de ocorrer, dos operadores aniquilação e neutralização num fragmento de uma configuração $x \in \Delta$. O operador aniquilação elimina as posições. Por outro lado, o operador neutralização não elimina as posições, ele atua mudando o estado das componentes.

Seja $x \in \Delta$ e δ_x a medida degenerada concentrada em x . Pelas definições de N_α e A_α , vemos que há uma relação, a qual é expressa a seguir,

$$\delta_x A_\alpha(\ominus) = \frac{\delta_x N_\alpha(\ominus)}{1 - \delta_x N_\alpha(\odot)}.$$

Daí, para cada $t \in \mathbb{N}$,

$$\delta_x \mathbf{A}_\alpha^t(\ominus) = \frac{\delta_x \mathbf{N}_\alpha^t(\ominus)}{1 - \delta_x \mathbf{N}_\alpha^t(\odot)}. \quad (4.16)$$

Definimos agora o operador *eliminação*, $\mathbf{E}_\alpha : \mathcal{A}_\odot \rightarrow \mathcal{A}_\odot$, este é um caso particular do operador neutralização. A saber, dada $x \in \Delta$, (ou seja, sua cópia em Δ_\odot) o operador eliminação procura o menos mais à esquerda e transforma a palavra $\oplus \odot^n \ominus$ em \odot^{n+2} com probabilidade α , ou nada acontece com probabilidade $1 - \alpha$. Vale ressaltar que diferentemente do operador neutralização, o operador eliminação age em apenas uma palavra $\oplus \odot^n \ominus$, transformando-a em \odot^{n+2} .

Lema 8. *O operador \mathbf{E}_α é linear.*

Prova. Seja $x \in \Delta_\odot$, tal que, existem inteiros i_0, j_0 com $i_0 < j_0$ tal que $x_{i_0} = x_{j_0} = \ominus$ e $x_k \neq \ominus$ para todo $k < i_0$ e $k > j_0$.

Se $x \in \Delta_\odot$ for tal que $x_k = \odot$ para todo $k < i_0$ teremos

$$\delta_x \mathbf{E}_\alpha = \delta_x.$$

Neste caso é evidente que \mathbf{E}_α é linear. Por isso, consideraremos o caso que há $l_0 < i_0$ tal que $x_{l_0} = \oplus$, $x_{l_0+1} = \dots = x_{i_0-1} = \odot$ e $x_{i_0} = \ominus$. Agora, seja $y \in \Delta_\odot$, definida da seguinte forma

$$y_i = \begin{cases} \odot, & \text{se } i \in \{l_0, i_0\}, \\ x_i, & \text{caso contrário.} \end{cases} \quad (4.17)$$

Por isso, podemos escrever

$$\delta_x \mathbf{E}_\alpha = \alpha \delta_y + (1 - \alpha) \delta_x.$$

Assim, para cada $x^i \in \Delta_\odot$, denotemos

$$\delta_{x^i} \mathbf{E}_\alpha = \alpha \delta_{y^i} + (1 - \alpha) \delta_{x^i},$$

onde y^i segue a relação para x^i dada em (4.17). Logo,

$$\begin{aligned} \left(\sum_{i=1}^{\infty} k_i \delta_{x^i} \right) \mathbf{E}_\alpha &= \alpha \left(\sum_{i=1}^{\infty} k_i \delta_{y^i} \right) + (1 - \alpha) \left(\sum_{i=1}^{\infty} k_i \delta_{x^i} \right) \\ &= \sum_{i=1}^{\infty} (\alpha k_i \delta_{y^i} + (1 - \alpha) k_i \delta_{x^i}) \\ &= \sum_{i=1}^{\infty} k_i (\alpha \delta_{y^i} + (1 - \alpha) \delta_{x^i}) \\ &= \sum_{i=1}^{\infty} k_i (\delta_{x^i} \mathbf{E}_\alpha). \end{aligned}$$

□

Provamos o Lema 8 de forma construtiva. Para provar o Lema 9 utilizamos conceitos

padrões de Cadeias de Markov com um estado contável de estados. Esses conceitos podem ser estudados em (FERRARI e GALVES, 1997; KARLIN, 2014; DURRETT e DURRETT, 1999; SCHINAZI, 1999), assim como em livros clássicos tais como (SENETA, 2006; NORRIS, 1998). Também utilizamos a ideia de enumerabilidade, que pode ser encontrada em (BARTLE e SHERBERT, 2011), (LIMA, 2004) e (LIMA, 2006). Podemos utilizar a mesma ideia para provar o Lema 8.

Lema 9. *O operador N_α é linear.*

Prova. É evidente que Δ_\odot é contável. Logo, dados $x, y \in \Delta_\odot$, se denotarmos p_{xy} a probabilidade de transição da ilha x para a ilha y , teremos que há $K_x \subset \Delta_\odot$ finito tal que

$$p_{xy} = \begin{cases} \text{positivo,} & \text{se } y \in K_x, \\ \text{nulo,} & \text{se } y \notin K_x. \end{cases}$$

Além disso, podemos escrever

$$\Delta_\odot = \{x^1, x^2, x^3, \dots\}.$$

Portanto, podemos definir uma matriz de transição

$$P = (p_{x^i x^j})_{i, j \in \mathbb{N}}.$$

Assim, definimos nossa Cadeia de Markov, a qual certamente é linear. Uma vez que o operador neutralização está associado com esta Cadeia de Markov, segue o resultado. \square

Lema 10. *Para toda $\mu \in \mathcal{A}_\ominus$ temos*

$$\mu A_\alpha(\ominus) \leq \frac{\mu E_\alpha(\ominus)}{1 - \mu E_\alpha(\odot)}.$$

Prova. Note que há uma cópia de \mathcal{A}_\ominus em \mathcal{A}_\odot , logo podemos falar em μE_α , para toda $\mu \in \mathcal{A}_\ominus$. Além disso, dada $\mu \in \mathcal{A}_\ominus$, podemos escrever $\mu = \sum_{i=1}^{\infty} k_i \delta_{x^i}$, onde $k_i \geq 0$ para todo $i \in \mathbb{N}$, $\sum_{i=1}^{\infty} k_i = 1$ e $x^i \in \Delta$ para todo $i \in \mathbb{N}$. De modo que μ será a δ_{x^i} -medida com probabilidade k_i , para algum i . Consequentemente, de (4.16) segue que

$$\mu A_\alpha(\ominus) = \frac{\mu N_\alpha(\ominus)}{1 - \mu N_\alpha(\odot)}.$$

Como E_α é um caso particular de N_α , em que apenas uma ocorrência da palavra $\oplus \odot^n \ominus$ é transformada em \odot^{n+2} , para algum $n \in \mathbb{N}$, com probabilidade α , segue que

$$\mu N_\alpha(\ominus) \leq \mu E_\alpha(\ominus) \quad \text{e} \quad \mu N_\alpha(\odot) \geq \mu E_\alpha(\odot).$$

Daí,

$$\frac{\mu N_\alpha(\ominus)}{1 - \mu N_\alpha(\odot)} \leq \frac{\mu E_\alpha(\ominus)}{1 - \mu E_\alpha(\odot)}.$$

De onde segue o resultado. \square

4.6 Algumas Cadeias de Markov em \mathbb{Z}_+

Neste seção, definiremos e daremos algumas características de duas Cadeias de Markov, cujo conjunto de estado é \mathbb{Z}_+ . Denotaremos a sua probabilidade de transição de um estado i para um estado j , por p_{ij} . Por simplicidade, algumas vezes utilizaremos a notação C.M., no lugar de: Cadeia de Markov.

Nossa primeira C. M. será obtida do seguinte cenário: suponha que há uma urna contendo exatamente n bolas no momento inicial. Em cada momento, cada bola pode ser retirada da urna com probabilidade β ou permanecer na urna com probabilidade $1 - \beta$, de forma independente uma das outras.

Esta C. M. tem probabilidade de transição

$$p_{mn} = \begin{cases} 0, & \text{se } m < n; \\ \binom{m}{n} (1 - \beta)^n \beta^{m-n}, & \text{se } m \geq n, \end{cases} \quad (4.18)$$

onde $\beta \in [0, 1]$.

Nossa segunda C. M. tem probabilidades de transição

$$\begin{aligned} p_{00} &= 1, \\ p_{nn-1} &= q_n, \quad p_{nn} = 1 - q_n, \quad p_{nj} = 0 \text{ se } j \notin \{n, n-1\}, \end{aligned} \quad (4.19)$$

com $q_n \in [0, 1]$.

Nas Figuras 9 e 10 exibimos os diagramas correspondentes à probabilidade de transição dos processos (4.18) e (4.19), respectivamente.

Note que em ambas as Cadeias de Markov apresentadas, se $q_n \in (0, 1)$ e $\beta \in (0, 1)$, então a probabilidade de absorção é 1, isto é, dada uma C.M. $(X_t)_{t \in \mathbb{N}}$

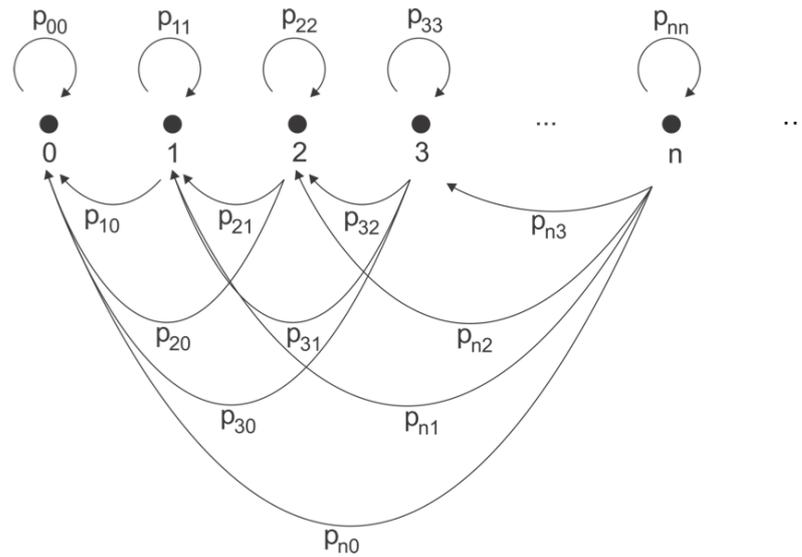
$$\mathbb{P}(H_n < \infty) = 1,$$

onde

$$H_n = \inf\{t \in \mathbb{N} : X_t = 0 \text{ e } X_0 = n\}$$

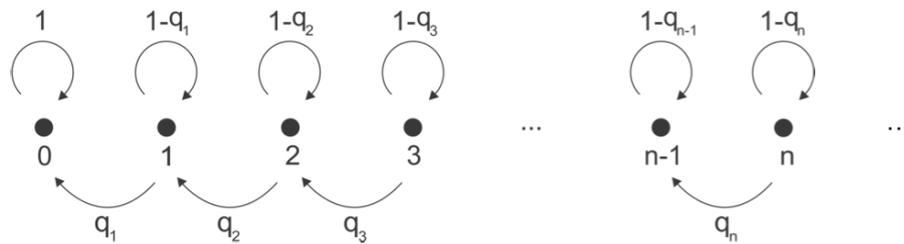
é a variável aleatória, a qual descreve o tempo para alcançar o estado zero dado que a C.M. iniciou no estado n . Estamos interessados na $\mathbb{E}(H_n)$ para ambas as Cadeias de Markov, isto é, o tempo médio para alcançar o estado zero dado que este iniciou no estado n . Por questões de simplicidade, denotaremos $k_n = \mathbb{E}(H_n)$. Deste ponto em diante, denotaremos

Figura 9 – Diagrama das probabilidades de transição do processo (4.18).



Fonte: O autor (2020).

Figura 10 – Diagrama das probabilidades de transição do processo (4.19).



Fonte: O autor (2020).

por $X = (X_t)_{t \in \mathbb{Z}_+}$ e $Y = (Y_t)_{t \in \mathbb{Z}_+}$ as Cadeias de Markov com probabilidades de transição dadas por (4.18) e (4.19) respectivamente.

Agora, vamos enunciar um resultado, que pode ser encontrado em (NORRIS, 1998), pois o utilizaremos para provar o Lema 11 a seguir.

Teorema 1.3.5 em (NORRIS, 1998). *O Vetor do tempo médio de alcance $k^A = (k_i^A : i \in I)$ é uma solução não negativa minimal para o sistema de equações lineares*

$$\begin{cases} k_i^A = 0 & \text{para, } i \in A, \\ k_i^A = 1 + \sum_{j \notin A} p_{ij} k_j^A & \text{para, } i \notin A. \end{cases}$$

Lema 11. Para a Cadeia de Markov Y , temos

$$k_n = \sum_{i=1}^n \frac{1}{q_i} \quad \text{para } n \geq 1. \quad (4.20)$$

Prova. Para obter k_n , utilizaremos o Teorema 1.3.5 dado em (NORRIS, 1998). Necessitando obter a solução não negativa minimal para o sistema de equações

$$\begin{cases} k_0 = 0, \\ k_n = 1 + q_n k_{n-1} + (1 - q_n) k_n. \end{cases}$$

É fácil verificar que para todo $n \geq 1$,

$$k_n - k_{n-1} = \frac{1}{q_n}.$$

Vamos mostrar o lema usando indução em n .

- Para $n = 1$, $k_1 - k_0 = \frac{1}{q_1}$, e como $k_0 = 0$, segue que $k_1 = \frac{1}{q_1}$;
- Suponha que o resultado é válido para n , isto é, $k_n = \sum_{i=1}^n \frac{1}{q_i}$;
- Vamos mostrar o resultado para $n + 1$.

Note que, $k_{n+1} - k_n = \frac{1}{q_{n+1}}$, e como por hipótese de indução $k_n = \sum_{i=1}^n \frac{1}{q_i}$, segue que

$$k_{n+1} = \left(\sum_{i=1}^n \frac{1}{q_i} \right) + \frac{1}{q_{n+1}}, \text{ ou seja, } k_{n+1} = \sum_{i=1}^{n+1} \frac{1}{q_i}.$$

□

Utilizando o Teorema 1.3.5 dado em (NORRIS, 1998), com o intuito de obter k_n , na Cadeia de Markov X , teremos que resolver a seguinte equação de diferença

$$k_n = \frac{1 + \sum_{j=1}^{n-1} p_{nj} k_j}{1 - (1 - \beta)^n}, \quad \text{para todo } n > 1 \quad \text{e} \quad k_1 = \frac{1}{\beta}. \quad (4.21)$$

Onde $p_{nj} = \binom{n}{j} (1 - \beta)^j \beta^{n-j}$.

Obter a solução exata de (4.21) é uma tarefa difícil. Por isso, faremos outras considerações para essa Cadeia de Markov, as quais nos darão a solução exata de (4.21). Vamos nomear as n bolas da urna como $b_1, b_2, b_3, \dots, b_n$. Para cada bola b_i , associamos a variável aleatória T^i , a qual indica o momento em que a bola b_i é retirada da urna.

Lema 12. A variável aleatória T^i tem distribuição geométrica com parâmetro β , isto é,

$$\mathbb{P}(T^i = k) = (1 - \beta)^{k-1} \beta,$$

onde $\beta \in (0, 1)$ e $k \in \mathbb{N}$.

Prova. De fato, dada uma bola b_i pertencente a urna, calcular a probabilidade dessa bola ser retirada da urna no momento $k \in \mathbb{N}$, significa que ela permaneceu na urna nos $k - 1$ primeiros momentos e só foi retirada no momento k , isto é,

$$\mathbb{P}(T^i = k) = (1 - \beta)^{k-1}\beta, \text{ para } k \in \mathbb{N}.$$

Portanto, T^i tem distribuição geométrica com parâmetro β . \square

Pela construção da nossa C. M., as variáveis aleatórias T^1, T^2, \dots, T^n são independentes e identicamente distribuídas.

Dado $x > 0$, calcular $\mathbb{P}(H_n \leq x)$, significa calcular a probabilidade de que todas as bolas b_i , para $i \in \{1, 2, \dots, n\}$, sejam retiradas da urna até o momento x , isto é,

$$\mathbb{P}(H_n \leq x) = \mathbb{P}(T^1 \leq x, T^2 \leq x, \dots, T^n \leq x),$$

uma vez que as variáveis aleatórias T^i são independentes temos que

$$\mathbb{P}(H_n \leq x) = \mathbb{P}(T^1 \leq x) \cdot \mathbb{P}(T^2 \leq x) \cdot \dots \cdot \mathbb{P}(T^n \leq x),$$

desde que cada variável aleatória T^i tem mesma distribuição geométrica com parâmetro β , segue que

$$\mathbb{P}(H_n \leq x) = (\mathbb{P}(T \leq x))^n, \quad (4.22)$$

onde T é uma variável aleatória que tem distribuição geométrica com parâmetro β . Assim,

$$\mathbb{P}(T \leq x) = \sum_{j=1}^{\lfloor x \rfloor} \beta(1 - \beta)^{j-1},$$

onde $\lfloor x \rfloor$ é o maior inteiro menor ou igual a x .

Deste modo, para todo $i \in \mathbb{N}$, se $x \in [i, i + 1)$, então

$$\mathbb{P}(T \leq x) = 1 - (1 - \beta)^i.$$

Portanto usando (4.22),

$$\mathbb{P}(H_n \leq x) = \begin{cases} 0, & \text{se } x < 1; \\ (1 - (1 - \beta)^i)^n, & \text{se } x \in [i, i + 1), \text{ com } i \in \mathbb{N}. \end{cases} \quad (4.23)$$

Proposição 1. A solução da equação (4.21) é

$$k_n = \sum_{i=0}^{\infty} [1 - (1 - (1 - \beta)^i)^n],$$

Prova. Como H_n assume valores no conjunto dos números naturais então,

$$k_n = \mathbb{E}(H_n) = \sum_{i=1}^{\infty} i\mathbb{P}(H_n = i) = \sum_{i=1}^{\infty} \mathbb{P}(H_n \geq i) = \sum_{i=0}^{\infty} \mathbb{P}(H_n > i)$$

e como

$$\sum_{i=0}^{\infty} \mathbb{P}(H_n > i) = \sum_{i=0}^{\infty} (1 - \mathbb{P}(H_n \leq i)),$$

usando (4.23),

$$k_n = \sum_{i=0}^{\infty} [1 - (1 - (1 - \beta)^i)^n].$$

□

Obtivemos na Proposição 1, uma outra representação analítica para (4.21). Infelizmente, ela ainda é apresentada por uma série, a qual não foi possível obter de uma forma elementar simples.

Na teoria de probabilidades, *acoplamento* é uma técnica de prova que permite comparar duas variáveis aleatórias não relacionadas X e Y , criando um vetor aleatório W , cujas distribuições marginais correspondem a X e a Y , respectivamente.

Proposição 2. *Sejam $X = (X_t)_{t \in \mathbb{Z}_+}$ e $Y = (Y_t)_{t \in \mathbb{Z}_+}$ as Cadeias de Markov, com $q_n = 1 - (1 - \beta)^n$. Se $X_0 = Y_0 = n$, então $\mathbb{P}(X_t \leq Y_t) = 1$ para todo $t \in \mathbb{Z}_+$.*

Prova. Sejam $X_0 = Y_0 = n$, e $(U_t)_{t \in \mathbb{N}}$ uma sequência de variáveis aleatórias independentes e identicamente distribuídas, onde U_t tem distribuição uniforme no intervalo $[0, 1]$ para todo $t \in \mathbb{N}$. Denotemos por p_{ij}^X e p_{ij}^Y as probabilidades de transição de X e Y , respectivamente.

Note que

$$p_{nn}^X = p_{nn}^Y = 1 - q_n \quad \text{e} \quad \sum_{j=0}^{n-1} p_{nj}^X = p_{nn-1}^Y = q_n.$$

Usaremos agora o conceito de acoplamento, o qual pode ser encontrado com mais detalhes em (FERRARI e GALVES, 1997). Considere o seguinte acoplamento de (X_t, Y_t) :

- Se $U_t \in [0, 1 - q_n]$, então $(X_t, Y_t) = (X_{t-1}, Y_{t-1})$, neste caso nada muda;
- Se $U_t \in (1 - q_n, 1]$ então $Y_t = Y_{t-1} - 1$ se $Y_{t-1} > 0$, e X_t assume um valor no conjunto $\{0, 1, \dots, X_{t-1} - 1\}$ se $X_{t-1} > 0$. Mais especificamente, dado $k \in \{0, 1, \dots, n - 1\}$. Se $U_t \in \left(\sum_{j=0}^k p_{nn-j}^X, \sum_{j=0}^{k+1} p_{nn-j}^X \right)$, então $X_t = n - k - 1$ para $X_{t-1} > 0$.

Logo,

$$\mathbb{P}(X_t \leq Y_t) = 1 \text{ para todo } t \in \mathbb{N}.$$

□

Pelo Lema 11, o tempo médio de absorção da C. M. Y , k_n^Y , com $q_i = 1 - (1 - \beta)^i$ é dado por

$$k_n^Y = \sum_{i=1}^n \frac{1}{1 - (1 - \beta)^i}. \quad (4.24)$$

Como consequência da Proposição 2, temos que o tempo médio para absorção de Y , com $q_i = 1 - (1 - \beta)^i$, é maior ou igual ao tempo médio de absorção de X , k_n^X , pois nesta, estando em um estado inicial n , podemos ir para qualquer outro estado j , com $j \leq n$, em cada passo de tempo. Já, no caso de Y , estando no mesmo estado inicial n , apenas podemos ir para o estado $n - 1$ ou permanecer no estado n , em cada passo de tempo. Portanto,

$$k_n^X = \sum_{i=0}^{\infty} [1 - (1 - (1 - \beta)^i)^n] \leq \sum_{i=1}^n \frac{1}{1 - (1 - \beta)^i} = k_n^Y. \quad (4.25)$$

4.7 Uma composição de duas Cadeias de Markov em \mathbb{Z}_+

Comumente, quando estudamos Cadeias de Markov descrevemos algumas de suas características, a saber: tempo de retorno, classificação dos estados entre transiente ou recorrente, dentre outras. Contudo, se temos duas Cadeias de Markov e conhecemos bem suas características, pouco podemos informar sobre uma C. M. dada pela composição destas outras duas. Nesta seção definiremos uma C. M. como um tipo de composição entre X e Y . O seguinte cenário descreve o que acontece em um passo de tempo: no primeiro momento nossa urna tem n bolas, cada bola será retirada com probabilidade β ou permanecerá na urna com probabilidade $1 - \beta$, no segundo momento, tomemos uma e apenas uma das bolas que permaneceram na urna, ela será retirada com probabilidade α ou permanecerá com probabilidade $1 - \alpha$. Após concluído o segundo momento, teremos obtido um passo de tempo nesse novo processo, o qual denotaremos por $Z = (Z_t)_{t \in \mathbb{Z}_+}$.

Assim como ocorre no caminhante aleatório, vamos descrever o processo Z por meio da soma de variáveis aleatórias.

Vamos considerar duas sequências de variáveis aleatórias independentes $(a_i^t)_{i, t \in \mathbb{N}}$ e $(c^t)_{t \in \mathbb{N}}$, onde

$$a_i^t = \begin{cases} 0, & \text{com probabilidade } 1 - \beta; \\ -1, & \text{com probabilidade } \beta. \end{cases} \quad \text{e} \quad c^t = \begin{cases} 0, & \text{com probabilidade } 1 - \alpha; \\ -1, & \text{com probabilidade } \alpha. \end{cases}$$

Pela definição de Z , temos que Z_t determina a quantidade de bolas na urna no tempo t dado que $Z_0 = n$. Definimos Z_t indutivamente por

$$Z_{t+1} = Z_t + \sum_{i=1}^{Z_t} a_i^{t+1} + c^{t+1} \mathbb{1}_{\{\Psi_t > 0\}}(\Psi_t), \quad (4.26)$$

onde $\mathbb{1}_{\{\cdot\}}(\cdot)$ é a função indicadora, e $\Psi_t = Z_t + \sum_{i=1}^{Z_t} a_i^{t+1}$. Falando informalmente, esta equação nos diz que a quantidade de bolas na urna no tempo $t + 1$ será igual a quantidade de bolas na urna no tempo t menos as bolas retiradas no primeiro e no segundo momento pelos processos (4.18) e (4.19) respectivamente. Nosso objetivo é calcular o número médio

de bolas no tempo $t + 1$ dado que iniciamos com n bolas.

Proposição 3. *Considere o processo Z e a forma indutiva de Z_t em (4.26), temos*

$$\mathbb{E}(Z_{t+1}) = (1 - \beta)\mathbb{E}(Z_t) + \alpha[\mathbb{E}(\beta^{Z_t}) - 1]. \quad (4.27)$$

Prova. Note que,

$$\begin{aligned} \mathbb{E}(Z_{t+1}) &= \sum_{k=0}^{\infty} \mathbb{E}(Z_{t+1}|Z_t = k)\mathbb{P}(Z_t = k) \\ &= \sum_{k=0}^{\infty} \left[\mathbb{E} \left(k + \sum_{i=1}^k a_i^{t+1} \right) \mathbb{P}(Z_t = k) + \mathbb{E} \left(c^{t+1} \mathbb{1}_{\{\Psi_t > 0\}}(\Psi_t) \right) \mathbb{P}(Z_t = k) \right], \end{aligned}$$

daí, como $\mathbb{E}(a_i^{t+1}) = -\beta$ e $\mathbb{E}(c^{t+1}) = -\alpha$ temos

$$\begin{aligned} \mathbb{E}(Z_{t+1}) &= \sum_{k=0}^{\infty} \left[(k - k\beta)\mathbb{P}(Z_t = k) - \alpha \mathbb{P} \left(k + \sum_{i=1}^k a_i^{t+1} > 0 \right) \mathbb{P}(Z_t = k) \right] \\ &= (1 - \beta) \sum_{k=0}^{\infty} k\mathbb{P}(Z_t = k) - \alpha \sum_{k=0}^{\infty} \mathbb{P} \left(-\sum_{i=1}^k a_i^{t+1} < k \right) \mathbb{P}(Z_t = k). \end{aligned}$$

Mas, $-\sum_{i=1}^k a_i^{t+1}$ tem distribuição binomial com parâmetros k e β , portanto,

$$\mathbb{P} \left(-\sum_{i=1}^k a_i^{t+1} < k \right) = 1 - \beta^k.$$

Logo,

$$\begin{aligned} \mathbb{E}(Z_{t+1}) &= (1 - \beta)\mathbb{E}(Z_t) - \alpha \sum_{k=0}^{\infty} (1 - \beta^k)\mathbb{P}(Z_t = k) \\ &= (1 - \beta)\mathbb{E}(Z_t) + \alpha \left[\sum_{k=0}^{\infty} \beta^k \mathbb{P}(Z_t = k) - 1 \right]. \end{aligned}$$

Uma vez que $\mathbb{E}(\beta^{Z_t}) = \sum_{k=0}^{\infty} \beta^k \mathbb{P}(Z_t = k)$, conclui-se a prova desta proposição. □

Lema 13. *A expressão (4.27) pode ser reescrita por:*

$$\mathbb{E}(Z_{t+1}) = (1 - \beta)^{t+1} Z_0 + \alpha \sum_{i=0}^t (1 - \beta)^i (\mathbb{E}(\beta^{Z_{t-i}}) - 1). \quad (4.28)$$

Prova. Vamos mostrar usando indução em t .

- Para $t = 0$ é trivial.
- Suponha que o resultado é válido para t , ou seja,

$$\mathbb{E}(Z_t) = (1 - \beta)^t Z_0 + \alpha \sum_{i=0}^{t-1} (1 - \beta)^i (\mathbb{E}(\beta^{Z_{t-i-1}}) - 1)$$

e vamos mostrar que é válido para $t + 1$. Note que,

$$\begin{aligned} \mathbb{E}(Z_{t+1}) &= (1 - \beta)\mathbb{E}(Z_t) + \alpha[\mathbb{E}(\beta^{Z_t}) - 1] \\ &= (1 - \beta) \left[(1 - \beta)^t Z_0 + \alpha \sum_{i=0}^{t-1} (1 - \beta)^i (\mathbb{E}(\beta^{Z_{t-i-1}}) - 1) \right] + \alpha[\mathbb{E}(\beta^{Z_t}) - 1] \\ &= (1 - \beta)^{t+1} Z_0 + \alpha \sum_{i=0}^{t-1} (1 - \beta)^{i+1} (\mathbb{E}(\beta^{Z_{t-i-1}}) - 1) + \alpha[\mathbb{E}(\beta^{Z_t}) - 1] \\ &= (1 - \beta)^{t+1} Z_0 + \alpha \sum_{j=1}^t (1 - \beta)^j (\mathbb{E}(\beta^{Z_{t-j}}) - 1) + \alpha[\mathbb{E}(\beta^{Z_t}) - 1] \\ &= (1 - \beta)^{t+1} Z_0 + \alpha \sum_{j=0}^t (1 - \beta)^j (\mathbb{E}(\beta^{Z_{t-j}}) - 1). \end{aligned}$$

□

Como $\beta \in [0, 1]$, segue que $\beta^k \leq 1$ para todo $k \in \mathbb{N}$, daí $\beta^k \mathbb{P}(Z_t = k) \leq \mathbb{P}(Z_t = k)$, de onde segue que $\sum_{k=0}^{\infty} \beta^k \mathbb{P}(Z_t = k) \leq \sum_{k=0}^{\infty} \mathbb{P}(Z_t = k) = 1$, ou seja, $\mathbb{E}(\beta^{Z_t}) - 1 \leq 0$, para todo $t \in \mathbb{N}$. Sendo assim, considerando apenas a parcela em que $i = t$ no lado direito de (4.28), obtemos

$$\mathbb{E}(Z_{t+1}) \leq (1 - \beta)^{t+1} Z_0 + \alpha(1 - \beta)^t (\mathbb{E}(\beta^{Z_0}) - 1) \leq (1 - \beta)^t (Z_0 + \alpha\beta^{Z_0} - \alpha). \quad (4.29)$$

4.8 Operador $F_\beta E_\alpha$ e o processo Z

Inicialmente, apresentaremos um acoplamento entre o operador $F_\beta E_\alpha$ agindo em uma δ -medida, δ_x , onde x pertence à cópia de Δ em Δ_\odot e o processo Z definido na seção 4.7.

- No primeiro momento, temos a ação do operador F_β , o qual transforma cada componente no estado menos para o estado mais com probabilidade β independentemente do que ocorre nas outras posições. Isto é o mesmo que ocorre no primeiro momento do processo Z .
- No segundo momento, temos a ação do operador E_α , o qual transforma uma única componente no estado menos para o estado ponto com probabilidade α . Isto é o mesmo que ocorre no segundo momento do processo Z .

Assim, associamos o operador $F_\beta E_\alpha$ (atuando primeiro F_β e depois E_α , nesta ordem) à ação do primeiro momento e segundo momento do processo Z .

Os operadores F_β e E_α são lineares (ver Lema 8). Daí, $F_\beta E_\alpha$ é linear. A medida μ pertence a uma cópia de \mathcal{A}_\ominus em \mathcal{A}_\odot e é uma combinação convexa de δ -medidas. Uma vez que definimos $F_\beta E_\alpha$ agindo em δ -medidas, temos então definido $\mu F_\beta E_\alpha$.

Para $x \in \Delta_\odot$ temos $Z_0 = \text{Pop}(x)$ e Z_t determina a quantidade de menos da ilha x no tempo t , x^t , a qual tem distribuição $\delta_x(F_\beta E_\alpha)^t$. Já para um arquipélago de menos μ , onde $\mu = \sum_{i=1}^{\infty} k_i \delta_{x^i}$ temos

$$\mathbb{P}(Z_0 = \text{Pop}(x^i)) = k_i.$$

4.9 Prova dos Teoremas 7 e 8

Prova do Teorema 7. Como $\mu \in \mathcal{A}_\ominus$, podemos escrever $\mu = \sum_{i=1}^{\infty} k_i \delta_{x^i}$, onde $k_1 > 0, k_2 > 0, \dots, k_1 + k_2 + \dots = 1$ e x^i é uma ilha de menos para todo $i \in \mathbb{N}$. Assim,

$$\begin{aligned} \tau_\mu &= \inf\{t \geq 0 : \mu(F_\beta A_\alpha)^t = \delta_\oplus\} \\ &= \inf\{t \geq 0 : \mu(F_\beta A_\alpha)^t(\ominus) = 0\} \end{aligned} \quad (4.30)$$

Sabemos que há uma cópia de \mathcal{A}_\ominus em \mathcal{A}_\odot . Logo podemos escrever a seguinte relação decorrente do Lema 10

$$\mu(F_\beta A_\alpha)^t(\ominus) \leq \frac{\mu(F_\beta E_\alpha)^t(\ominus)}{1 - \mu(F_\beta E_\alpha)^t(\odot)}, \quad \text{para cada } t \in \mathbb{N}.$$

Daí, retornando a (4.30) e usando o Lema 8 temos,

$$\begin{aligned} \tau_\mu &\leq \inf\{t \geq 0 : \mu(F_\beta E_\alpha)^t(\ominus) = 0\} \\ &= \inf\{t \geq 0 : \delta_{x^1}(F_\beta E_\alpha)^t(\ominus) = \delta_{x^2}(F_\beta E_\alpha)^t(\ominus) = \dots = 0\} \\ &= \inf\{t \geq 0 : \delta_\xi(F_\beta E_\alpha)^t(\ominus) = 0\} \end{aligned}$$

onde $\xi = x^{\text{Max}(\mu)}$. Denotemos esta quantidade na última igualdade por τ_ξ .

Vimos que

$$\delta_\xi(F_\beta E_\alpha)^t(\ominus), \quad (4.31)$$

tem uma relação com a Cadeia de Markov $Z = (Z_t)_{t \in \mathbb{Z}_+}$ da seção 4.7.

Z é a composição das Cadeias de Markov X e Y descritas na seção 4.6, considerando que a urna começa com $\text{Max}(\mu)$ bolas. Definimos o tempo de absorção de X e Y por

$$\tau_{\text{Max}(\mu)}^X = \inf\{t \geq 0 : X_t = 0 \text{ e } X_0 = \text{Max}(\mu)\},$$

e

$$\tau_{\text{Max}(\mu)}^Y = \inf\{t \geq 0 : Y_t = 0 \text{ e } Y_0 = \text{Max}(\mu)\}.$$

Como Z é a composição de X e Y , agindo nessa ordem, primeiro X e depois Y , em cada passo de tempo, segue que $\mathbb{E}\left(\tau_{\text{Max}(\mu)}^Z\right) \leq \min\left\{\mathbb{E}\left(\tau_{\text{Max}(\mu)}^X\right), \mathbb{E}\left(\tau_{\text{Max}(\mu)}^Y\right)\right\}$. Assim, de (4.25),

$$\mathbb{E}(\inf\{t \geq 0 : \delta_\xi F_\beta^t = \delta_\oplus\}) = \mathbb{E}\left(\tau_{\text{Max}(\mu)}^X\right) \leq \sum_{i=1}^{\text{Max}(\mu)} \frac{1}{1 - (1 - \beta)^i} = g_\mu(\beta),$$

e

$$\mathbb{E}(\inf\{t \geq 0 : \delta_\xi E_\alpha^t = \delta_\oplus\}) = \mathbb{E}\left(\tau_{\text{Max}(\mu)}^Y\right) = \frac{\text{Max}(\mu)}{\alpha} = f_\mu(\alpha).$$

Note que, se $\alpha \geq \frac{\text{Max}(\mu)}{g_\mu(\beta)}$, então $\min\{g_\mu(\beta), f_\mu(\alpha)\} = f_\mu(\alpha)$. Por outro lado, se $\alpha < \frac{\text{Max}(\mu)}{g_\mu(\beta)}$, então $\min\{g_\mu(\beta), f_\mu(\alpha)\} = g_\mu(\beta)$.

Assim, temos

$$\mathbb{E}(\tau_\mu) \leq \mathbb{E}(\tau_\xi) = \mathbb{E}\left(\tau_{\text{Max}(\mu)}^Z\right) \leq \begin{cases} f_\mu(\alpha), & \text{se } \alpha \geq \frac{\text{Max}(\mu)}{g_\mu(\beta)}, \\ g_\mu(\beta), & \text{se } \alpha < \frac{\text{Max}(\mu)}{g_\mu(\beta)}. \end{cases}$$

□

Prova do Teorema 8. Vamos considerar a distribuição inicial, δ_x , com $x \in \Delta$. Neste caso,

$$\text{Max}(\mu, t) = \text{Pop}(x^t).$$

Da associação entre a Cadeia de Markov $Z = (Z_t)_{t \in \mathbb{Z}_+}$ da seção 4.7, o processo $F_\beta E_\alpha$ e o processo $F_\beta A_\alpha$ segue que

$$\mathbb{P}(\text{Pop}(x^t) \leq Z_t) = 1.$$

Daí,

$$\mathbb{E}(\text{Max}(\mu, t)) = \mathbb{E}(\text{Pop}(x^t)) \leq \mathbb{E}(Z_t).$$

Usando que $Z_0 = \text{Max}(\mu) = \text{Pop}(x)$ e (4.29),

$$\mathbb{E}(\text{Max}(\mu, t)) \leq (1 - \beta)^{t-1} \left(\text{Max}(\mu) + \alpha \beta^{\text{Max}(\mu)} - \alpha \right).$$

Provar no caso em que μ não é uma δ -medida é análogo. Basta usar o fato que μ é uma combinação convexa de δ -medidas. □

4.10 Estudo computacional para $\mathbb{E}(\tau_\mu)$

Simulações computacionais nos proporcionam teorizar o que ocorre em determinadas situações, as quais são difíceis de serem descritas formalmente. Como, por exemplo, o valor exato da percolação não orientada em \mathbb{Z}^2 , a qual foi descrita computacionalmente bem antes de sua prova formal (STEIF, 2011).

Estudos computacionais são amplamente utilizados na busca para obter um melhor entendimento em alguns sistemas de partículas interagentes, (RAMOS e TOOM, 2008; BUŠIĆ *et al.*, 2013; FERRARI *et al.*, 2002; TOOM, 1995; SŁOWIŃSKI e MACKAY, 2015; RAMOS *et al.*, 2017).

Ao simularmos o processo, assumimos que sua configuração terá um número finito de componentes. Seja \mathbb{Z}_n o conjunto dos inteiros módulo n . Chamamos *periódico* o elemento $x \in \{\oplus, \ominus\}^{\mathbb{Z}_n}$ e $\Omega_n = \{\oplus, \ominus\}^{\mathbb{Z}_n}$ é o *espaço dos periódicos*. Note que, se $x \in \Omega_n$, então $x = (x_i)_{i \in \mathbb{Z}_n}$.

Nosso intuito aqui é apresentar o comportamento do limite superior descrito no Teorema 7 e sua correspondente estimativa computacional.

Com o objetivo de estimar o tempo médio para que nosso processo $\mu F_\beta A_\alpha$, com $\mu \in \mathcal{A}_\ominus$, atinja a configuração todos mais, faremos um estudo computacional, por meio de simulação. Consideraremos o caso particular em que $\mu = \delta_x$, onde $x \in \Delta$. Fizemos o estudo para $\text{Max}(\mu) \in \{10, 20, \dots, 100\}$.

Vejam a seguir um pseudo-código que escrevemos para estimar o tempo médio de convergência do nosso processo.

- Passo 1.** Consideramos inicialmente um periódico com 1000 componentes em que n estão no estado menos e as outras $1000 - n$ componentes estão no estado mais.
- Passo 2.** Consideramos um gerador de números aleatórios uniforme pertencente ao intervalo $(0, 1)$.
- Passo 3.** Fizemos um laço para α começando em 0.00 e atualizando em 0.02 até $\alpha < 1$. Dentro deste laço, fizemos outro laço para β começando em 0.01 e atualizando em 0.02 até $\beta < 1$.
- Passo 4.** Dentro dos laços acima, chamamos nosso gerador e comparamos o número gerado com o α e com o β de modo que:
 - Se o número gerado for menor do que β , atualizamos cada componente do estado menos para o estado mais, independentemente umas das outras;
 - Em seguida, se o número gerado for menor que α , cada par $\oplus\ominus$ existente em nosso periódico será eliminado, de forma independente uns dos outros.
- Passo 5.** Paramos o processo quando todas as componentes são mais, ou o processo tiver sido executado 100000 passos de tempo.

Passo 6. Para um mesmo (α, β) repetimos este procedimento 100 vezes, gerando um conjunto $\{t^1, t^2, \dots, t^{100}\}$, onde t^i indica o tempo de parada do programa na i -ésima repetição.

Passo 7. Daí, calculamos

$$\widehat{\mathbb{E}}(\tau_\mu) = \frac{\sum_{i=1}^{100} t^i}{100},$$

que representa a estimativa do tempo médio para atingir o periódico no qual todos os componentes são mais.

Nas Figuras 11(a), 12(a) e 13(a) vemos os limites superiores para $\mathbb{E}(\tau_\mu)$, os quais são descritos no Teorema 7. Nestas figuras, consideramos o caso particular em que $\mu = \delta_x$, onde $x \in \Delta$ e é tal que $\text{Max}(\mu) = \text{Pop}(x) \in \{10, 50, 100\}$, respectivamente. Percebemos nestes casos uma diferença de tonalidade que vai do vermelho, quando (α, β) está próximo de $(0, 0)$, até o laranja, quando (α, β) se afasta da origem. As funções $f_\mu(\alpha)$ e $g_\mu(\beta)$ que são os limites superiores da $\mathbb{E}(\tau_\mu)$ assumem valores grandes na região vermelha, a saber, $\max\{f_\mu(\alpha), g_\mu(\beta)\} = 296.4759$ na Figura 11(a), $\max\{f_\mu(\alpha), g_\mu(\beta)\} = 473.7327$ na Figura 12(a) e $\max\{f_\mu(\alpha), g_\mu(\beta)\} = 570.3338$ na Figura 13(a) e esses valores diminuem à medida que as cores vão ficando mais fracas, sendo a região laranja, onde a $f_\mu(\alpha)$ e a $g_\mu(\beta)$ assumem os menores valores, a saber, $\min\{f_\mu(\alpha), g_\mu(\beta)\} = 10.0000$ na Figura 11(a), $\min\{f_\mu(\alpha), g_\mu(\beta)\} = 50.0000$ na Figura 12(a) e $\min\{f_\mu(\alpha), g_\mu(\beta)\} = 100.0000$ na Figura 13(a). Notamos ainda que a medida que a população da ilha de menos cresce, a região na cor vermelha cresce horizontalmente, isto quer dizer que, quanto maior for a quantidade de menos em uma ilha, para valores pequenos de β , o tempo de alcance da configuração todos mais tende a aumentar, mesmo para valores grandes de α . Por outro lado, temos as Figuras 11(b), 12(b) e 13(b), que representam a simulação computacional para o tempo médio de alcance da configuração todos mais. Note que, para as “ilhas de menos” cuja população pertence ao conjunto $\{10, 50, 100\}$, quando estamos próximos da origem, percebemos uma pequena região na cor vermelha, onde o tempo que o processo leva para atingir a configuração todos mais é muito longo, a saber, o tempo médio máximo é 285.6263, 478.6979 e 519.0707 passos de tempo, respectivamente. Notamos ainda que para valores de β muito pequenos, independente dos valores de α , a estimativa de $\mathbb{E}(\tau_\mu)$ ainda é muito alta. É o que está acontecendo nas regiões azul, verde e amarelo. Mas, quando nos afastamos da origem percebemos uma região consideravelmente grande na cor laranja, onde o tempo que leva para o processo atingir a configuração todos mais é pequeno, a saber, o tempo médio mínimo é 1.0101, 1.0202 e 1.0101 passos de tempo, respectivamente. Mesmo considerando ilhas de menos, com outras populações, os gráficos praticamente não se alteram. Todavia, o tempo médio máximo aumenta ou diminui de acordo com o tamanho da ilha, porém o tempo médio mínimo se mantém quase estável.

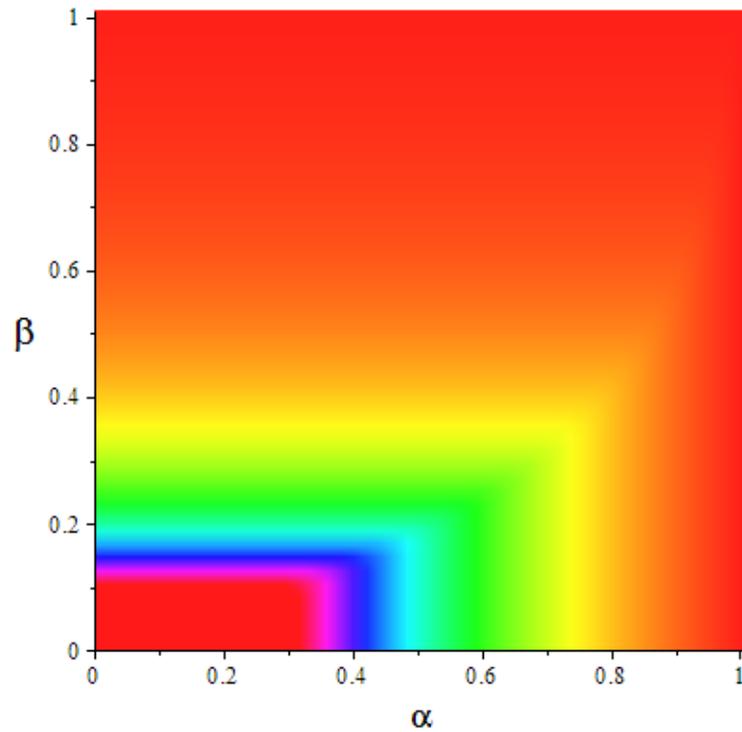
Note que se $\beta \in (0, 1)$ então $\frac{1}{1 - (1 - \beta)^i} > 1$ para todo $i \in \mathbb{N}$, logo

$$\sum_{i=1}^{\text{Max}(\mu)} \frac{1}{1 - (1 - \beta)^i} > \sum_{i=1}^{\text{Max}(\mu)} 1,$$

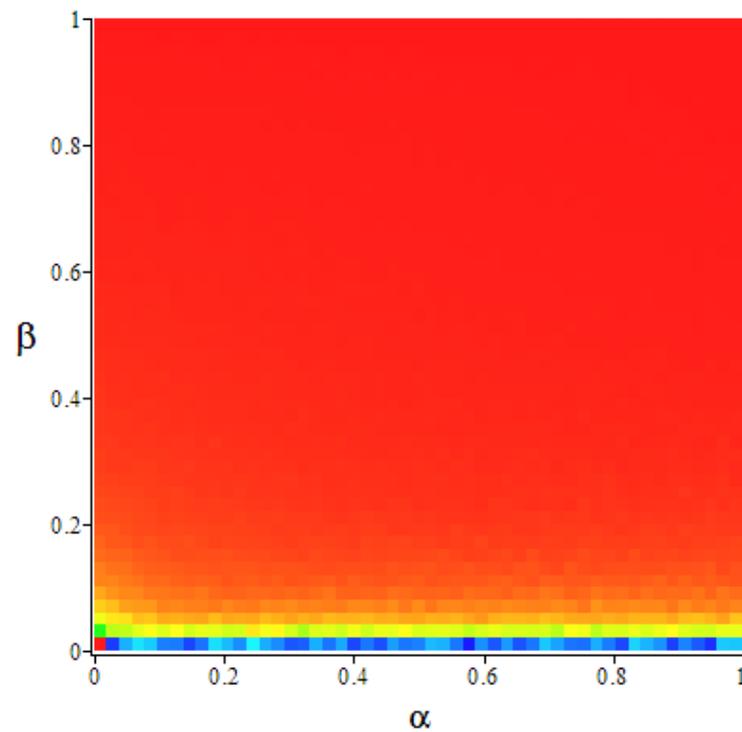
ou seja, $g_\mu(\beta) > \text{Max}(\mu)$, portanto $\frac{\text{Max}(\mu)}{g_\mu(\beta)} < 1$, para todo $\beta \in (0, 1)$. Deste modo, se $\text{Max}(\mu) \rightarrow \infty$, então $g_\mu(\beta) \rightarrow \infty$, portanto $\frac{\text{Max}(\mu)}{g_\mu(\beta)} \rightarrow 1$. Logo, para valores grandes de $\text{Max}(\mu)$, temos pouca influência de $f_\mu(\alpha)$ como limite superior de $\mathbb{E}(\tau_\mu)$. Consequentemente, temos uma grande região em que olharemos apenas para o caso em que $\mathbb{E}(\tau_\mu) \leq g_\mu(\beta)$. Deixando assim, os gráficos dos limites superiores mais uniforme e tornando-os cada vez mais semelhante aos gráficos 11(b), 12(b) e 13(b) obtidos por meio de simulação computacional.

Na Figura 14, exibiremos os gráficos, para $\beta \in (0, 0.2)$, semelhantes aos gráficos de simulação apresentados nas Figuras 11(b), 12(b) e 13(b). Com isso podemos enxergar melhor a variação de valores expressos por tonalidades.

Figura 11 – Gráficos das estimativas e dos limites superiores de $\mathbb{E}(\tau_\mu)$ para ilhas de menos em que $\text{Pop}(x) = 10$.



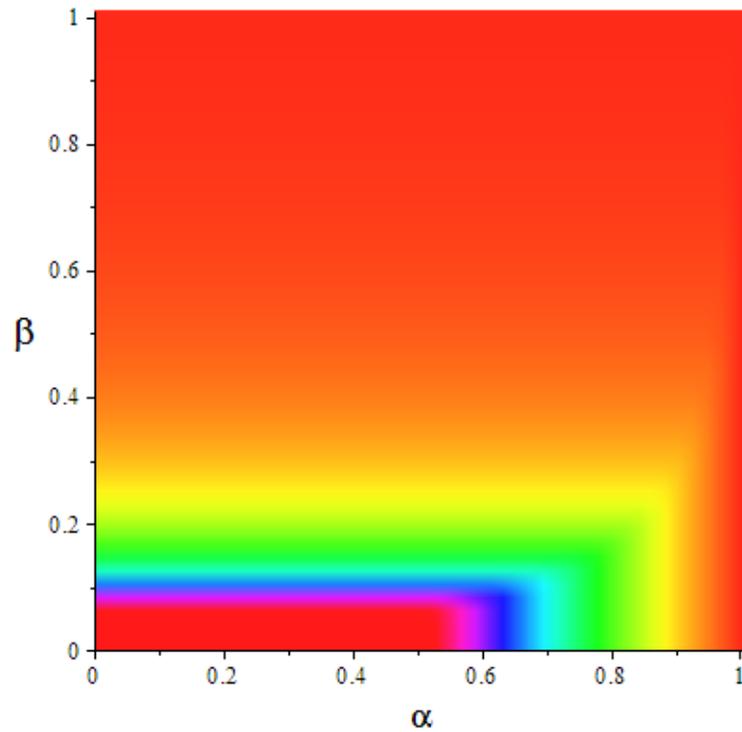
(a) Limites superiores de $\mathbb{E}(\tau_\mu)$ com $\text{Pop}(x) = 10$. Onde o tempo médio mínimo = 10.0000 e o tempo médio máximo = 296.4759.



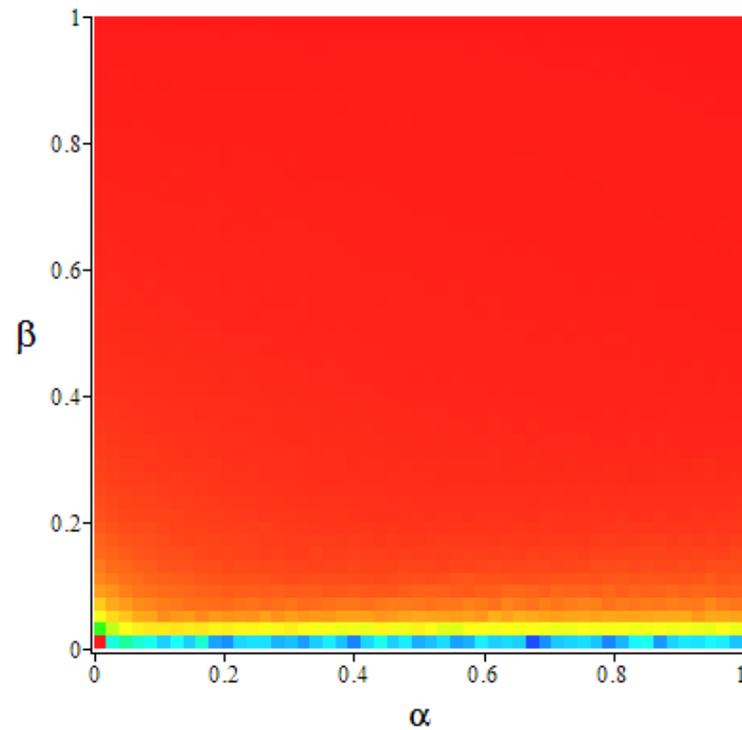
(b) Estimativa da $\mathbb{E}(\tau_\mu)$ com $\text{Pop}(x) = 10$. Onde o tempo médio mínimo = 1.0101 e o tempo médio máximo = 285.6263

Fonte: O autor (2020).

Figura 12 – Gráficos da estimativa e dos limites superiores de $\mathbb{E}(\tau_\mu)$ para ilhas de menos em que $\text{Pop}(x) = 50$.



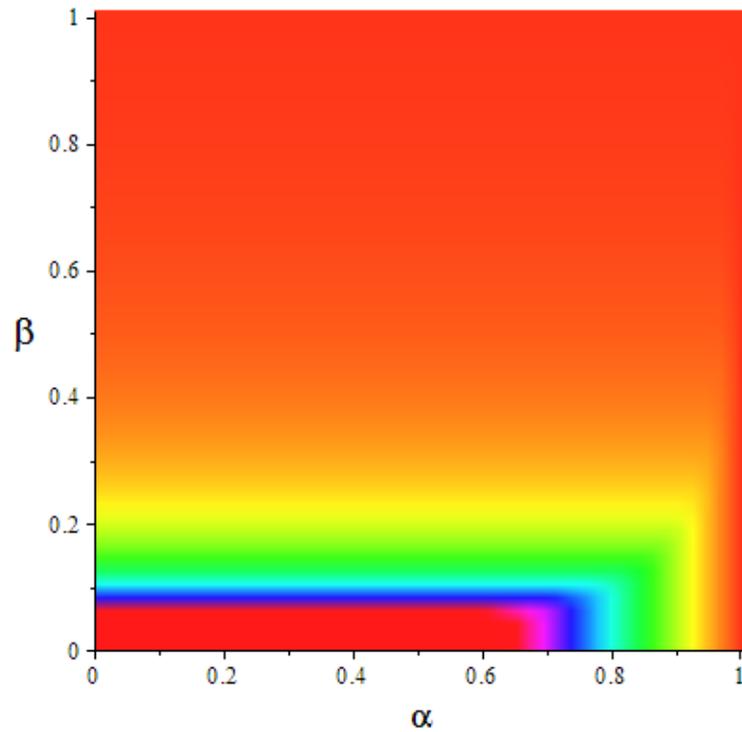
(a) Limites superiores de $\mathbb{E}(\tau_\mu)$ com $\text{Pop}(x) = 50$. Onde o tempo médio mínimo = 50.0000 e o tempo médio máximo = 473.7327.



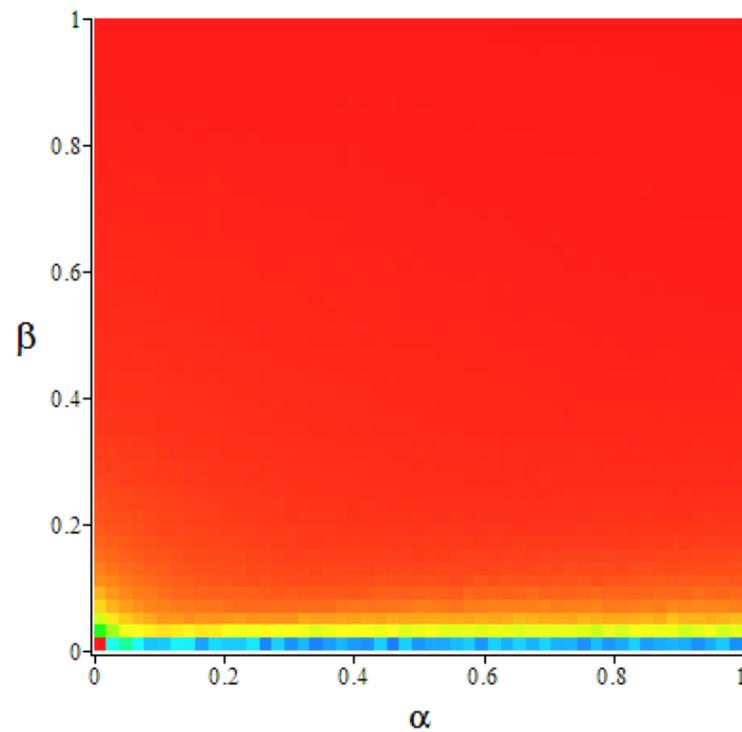
(b) Estimativa da $\mathbb{E}(\tau_\mu)$ com $\text{Pop}(x) = 50$. Onde o tempo médio mínimo = 1.0202 e o tempo médio máximo = 478.6970

Fonte: O autor (2020).

Figura 13 – Gráficos da estimativa e dos limites superiores de $\mathbb{E}(\tau_\mu)$ para ilhas de menos em que $\text{Pop}(x) = 100$.



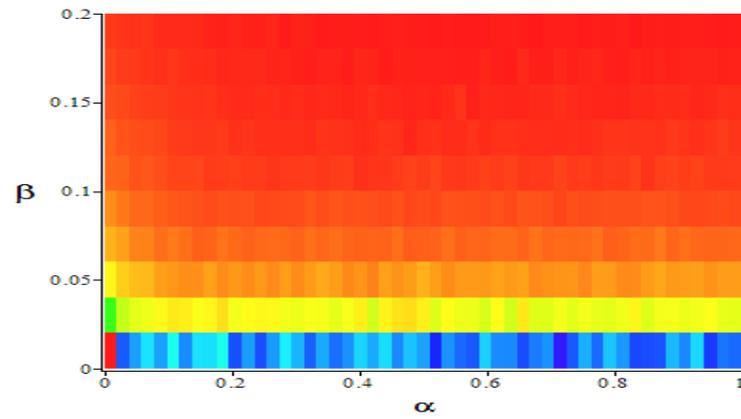
(a) Limites superiores de $\mathbb{E}(\tau_\mu)$ com $\text{Pop}(x) = 100$. Onde o tempo médio mínimo = 100.0000 e o tempo médio máximo = 570.3338.



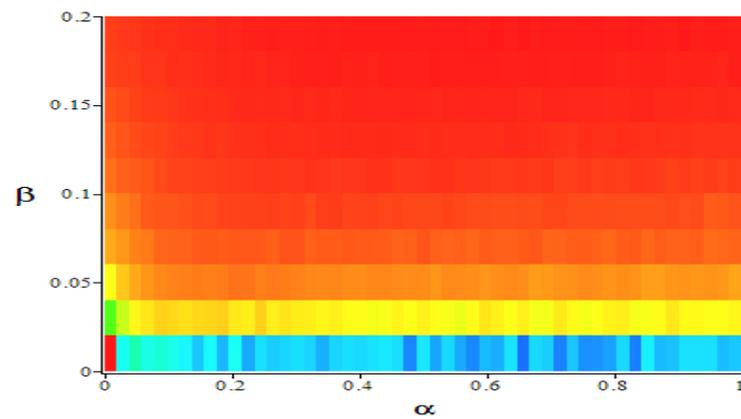
(b) Estimativa da $\mathbb{E}(\tau_\mu)$ com $\text{Pop}(x) = 100$. Onde o tempo médio mínimo = 1.0101 e o tempo médio máximo = 519.0707

Fonte: O autor (2020).

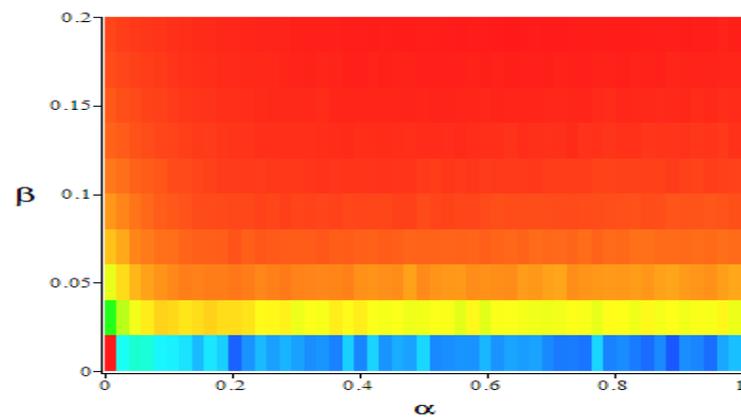
Figura 14 – Gráficos das estimativas de $\mathbb{E}(\tau_\mu)$ para ilhas de menos em que $\text{Pop}(x) \in \{10, 50, 100\}$.



(a) Estimativa da $\mathbb{E}(\tau_\mu)$ com $\text{Pop}(x) = 10$



(b) Estimativa da $\mathbb{E}(\tau_\mu)$ com $\text{Pop}(x) = 50$



(c) Estimativa da $\mathbb{E}(\tau_\mu)$ com $\text{Pop}(x) = 100$

Fonte: O autor (2020).

Capítulo 5

CONCLUSÃO E TRABALHOS FUTUROS

Em nosso trabalho, estudamos dois processos de uma nova classe de processos de partículas unidimensionais com tempo discreto, os quais chamamos de Flip-Murder e Flip-Aniquilação. Para o primeiro, mostramos que há uma transição de fase de primeira ordem entre as regiões de ergodicidade vs. não ergodicidade. Apresentamos duas curvas que delimitam estas regiões. Mostramos também que existe uma medida invariante para este processo, cuja frequência de mais é estritamente maior que zero e menor que um. Fizemos um estudo por meio de aproximação de Campo Médio e também um estudo computacional por meio de simulação de Monte Carlo.

Para o processo Flip-Aniquilação, apresentamos uma curva que melhora, no sentido de ampliar, a região de não ergodicidade. O que é uma contribuição bastante significativa, uma vez que ainda temos uma grande região no espaço de parâmetros, em que não podemos afirmar se este processo é ou não ergódico.

Quando pensamos nesse processo iniciando numa certa classe de medidas, chamadas arquipélagos de menos, μ , apresentamos funções que dependem dos parâmetros α , β e da máxima população de μ . Essas funções atuam como limites superiores para o tempo médio de alcance da configuração todos mais e para a quantidade média de menos no processo em cada passo de tempo.

Pretendemos definir, se possível, e estudar a metaestabilidade dos processos Flip-Murder e Flip-Aniquilação. Queremos também estudar as características da medida invariante, para esses processos, cuja frequência de mais está estritamente entre zero e um. Como por exemplo, como se comporta o decaimento da correlação nessa medida, o que traz uma relação com as medidas de Gibbs.

Provamos que, sob certas condições, o processo Flip-Murder converge fracamente para a medida concentrada na configuração todos mais, δ_{\oplus} . E, sob outras condições, esse processo não converge para δ_{\oplus} . Uma pergunta natural que pretendemos responder é: Quando esse processo é ergódico, ele apresenta um comportamento metaestável antes de alcançar o equilíbrio δ_{\oplus} ? Ou ainda, mesmo na região de não ergodicidade nosso processo apresenta

um comportamento metaestável?

Outros modelos que gostaríamos de estudar, são variantes do modelo Flip-Murder com uma certa simetria no operador. No primeiro caso, se modificarmos o operador Murder. Suponha que ele é de alguma forma simétrico, isto é, sempre que uma componente no estado mais for vizinho esquerdo de uma componente no estado menos, este operador irá eliminar a letra mais com probabilidade α ou a letra menos com probabilidade $1 - \alpha$. No segundo caso, pretendemos tornar o operador flip simétrico, isto é, toda componente no estado menos será transformada em mais com probabilidade β e toda componente no estado mais será transformada em menos com probabilidade β , independentemente umas das outras.

Utilizamos a distância de variação total entre nossas distribuições iniciais, a saber, os arquipélagos de menos, e nossa distribuição invariante δ_{\oplus} . Dessa maneira, apresentamos uma caracterização das nossas distribuições iniciais em relação a distribuição δ_{\oplus} . Nessa parte do texto, fizemos uma análise tímida, todavia traz a possibilidade para estudarmos melhor esse tema em um trabalho futuro. Poderemos ir além, estudando por exemplo o tempo de mistura, ou seja, um parâmetro que mede o tempo necessário por uma Cadeia de Markov para que a distância de estacionariedade seja pequena.

REFERÊNCIAS

- BARTLE, Robert G.; SHERBERT, Donald R. *Introduction to real analysis*, Hoboken, NJ: Wiley, 2011.
- BUŠIĆ, Ana; MAIRESSE, Jean; MARCOVICCI, Irene. Probabilistic cellular automata, invariant measures, and perfect sampling. *Advances in Applied Probability*, v. 41, n. 4, p. 960-980, 2013.
- DE LIMA, Pedro Ferreira et al. Dualities useful in bond percolation. *CUBO, A Mathematical Journal*, v. 10, n. 3, p. 93–102, 2008.
- DE SANTANA, L. H.; RAMOS, A. D.; TOOM, A. Eroders on a plane with three states at a point. Part I: Deterministic. *Journal of Statistical Physics*, v. 159, n. 5, p. 1175-1195, 2015.
- DURRETT, R.; DURRETT, R. *Essenciais de processos estocásticos. Vol. 1* Nova Iorque: Springer, 1999.
- FERRARI, P. A.; GALVES, A. *Acoplamento e processos estocásticos*. Rio de Janeiro: IMPA, 1997.
- FERRARI, Pablo A.; FERNÁNDEZ, Roberto; GARCIA, Nancy L. Perfect simulation for interacting point processes, loss networks and Ising models. *Stochastic Processes and their Applications* v. 102, n. 1, p. 63-88, 2002.
- FRANKLIN, Benjamin; COLBURN, Henry. *Memoirs of the Life and Writings of Benjamin Franklin*, LL. DFRS &c. Minister Plenipotentiary from the United States of America at the Court of France, and for the Treaty of Peace of Peace and Independence with Great Britain, &c. &c. 1818.
- GRIMMETT, Geoffrey. *What is percolation?*. In: *Percolation*. Springer, Berlin, Heidelberg, p. 1-31, 1999.
- KARLIN, S. *A first course in stochastic processes*. Academic press, 2014.
- LEVIN, David A.; PERES, Yuval; WILMER, Elizabeth L. *Markov Chains and Mixing Times*, American Mathematical Society, 2008.
- LI, Wentian. The study of correlation structures of DNA sequences: a critical review. *Computers & chemistry*, v. 21, n. 4, p. 257-271, 1997.
- LIGGETT, Thomas M. Stochastic models of interacting systems. *The Annals of Probability*, v. 25, n. 1, p. 1-29, 1997.

- LIGGETT, Thomas Milton. *Interacting particle systems*. Springer Science & Business Media, v. 276, 2012.
- LIMA, Elon L. *Elementos de topologia geral*. Ao Livro Técnico, Editora da Universidade de São Paulo, 1970.
- LIMA, Elon L. *Curso de análise, Volume 1*. Projeto Euclides, IMPA, décima primeira edição, 2004.
- LIMA, Elon L. *Análise Real: Funções de uma variável vol. 1*. Coleção Matemática Universitária, IMPA, 2006.
- MA, Jian, et al. The infinite sites model of genome evolution. *Proceedings of the National Academy of Sciences*, v. 105, n. 38, p.14254-14261, 2008.
- MAES, Christian et. al. New trends in interacting particle systems. *In: Markov Proc. Rel. Fields*, p. 283-288, 2005.
- MAIRESSE, J.; MARCOVICI, I. Around probabilistic cellular automata. *Theoretical Computer Science*, v. 559, p.42-72, 2014.
- MALYSHEV, Vadim A. Quantum grammars. *Journal of Mathematical Physics* v. 41, n. 7, p. 4508-4520, 2000.
- MALYSHEV, Vadim A. Quantum evolution of words. *Theoretical computer science* v. 273, n. 1-2, p. 263-269 2002.
- MCCOY, Barry M.; WU, Tai Tsun. *The two-dimensional Ising model*. Courier Corporation, 2014.
- MUNKRES, James R. *Topology*. Prentice Hall, 2000.
- NORRIS, J. R. *Markov chains*, n. 2, Cambridge University Press, 1998.
- RAMOS, Alex D.; TOOM, André. An error correction. Letter to the editor. *Journal of Statistical Physics*, v. 131, n. 1, p. 167-168, 2008.
- RAMOS, Alex D.; TOOM, André. Chaos and Monte Carlo Approximations of the Flip-Annihilation process. *Journal of Statistical Physics*, v. 133, n. 4, p. 761-771, 2008.
- RAMOS, A. D.; SILVA F. S. G.; SOUSA, C. S.; TOOM, A. Variable-length analog of Stavskaya process: A new example of misleading simulation. *Journal of Mathematical Physics* v. 58, n. 5, p. 053304, 2017.
- ROCHA, Andréa V.; SIMAS, Alexandre B.; TOOM, André. Substitution operators. *Journal of Statistical Physics*, v. 148, n. 3, p. 585-618, 2011.
- SAAKIAN, D. B. Evolution models with base substitutions, insertions, deletions, and selection. *Physical Review E*, v. 78, n. 6, 061920, 2008.
- SALGADO-GARCÍA, Raúl; UGALDE, Edgardo. Exact scaling in the expansion-modification system. *Journal of Statistical Physics*, v. 153, n. 5, p. 842-863, 2013.

- SCHINAZI, R. B. *Classical and spatial stochastic processes*. Boston: Birkhauser, 1999.
- SENETA, Eugene. *Non-negative matrices and Markov chains*. Springer Science & Business Media, 2006.
- SŁOWIŃSKI, Piotr; MACKAY, Robert S. Phase diagrams of majority voter probabilistic cellular automata. *Journal of Statistical Physics* v. 159, n. 1, p. 43-61, 2015.
- SOBOTTKA, Marcelo; HART, Andrew G. A model capturing novel strand symmetries in bacterial DNA. *Biochemical and biophysical research communications*, v. 410, n. 4, p. 823-828, 2011.
- STAUFFER, Dietrich; AHARONY, Ammon. *Introduction to percolation theory*. Taylor & Francis, 2018.
- STAVSKAYA, Olga N.; PIATETSKI-SHAPIRO, Ilya I. On homogeneous nets of spontaneously active elements. *Systems Theory Res*, v. 20, p. 75-88, 1971.
- STEIF, Jeffrey E. A mini course on percolation theory. *Jyväskylä lectures in mathematics*, n. 3, 2011.
- TOOM, A. A family of uniform nets of formal neurons. *DOKLADY AKADEMII NAUK SSSR*, v. 183, n. 1, p. 49, 1968.
- TOOM, Andrei. Cellular automata with errors: problems for students of probability *Topics in contemporary probability and its applications*. p. 117-157, 1995.
- TOOM, André; ROCHA, Andréa Vanessa; DE LIMA, Pedro Ferreira. *Contornos, conjuntos convexos e autômatos celulares*. IMPA, 2001.
- TOOM, André. Non-ergodicity in a 1-D particle process with variable length. *Journal of Statistical Physics*, v. 115, n. 3-4, p. 895-924, 2004.
- TOOM, A. Every continuous operator has an invariant measure. *Journal of Statistical Physics*, v. 129, n. 3, p. 555-566, 2007.
- TOOM, André.; RAMOS, Alex D.; ROCHA, Andréa V.; SIMAS, Alexandre B. *Random Processes with Variable Length 28º Colóquio Brasileiro de Matemática*, IMPA, 2011.

APÊNDICE A – PROVA DA SOBREJETIVIDADE DA APLICAÇÃO longo.

Usando a notação do Capítulo 3, vamos provar que para qualquer código principal legal C e $C' = \text{longo}(C)$

$$\text{HD}(C') = \text{HD}(C), \tag{A.1}$$

$$\text{VD}(C') = 2 \cdot \text{VD}(C). \tag{A.2}$$

A equação (A.1) é verdadeira, pois C' é obtido de C inserindo apenas os tipos 1' e 4', ambos têm $\text{HD} = 0$. Para provar (A.2), vamos classificar os tipos principais em, horizontal, a saber, o tipo 2, cujo $\text{VD} = 0$, e vertical, a saber os tipos 1, 3 e 4. Devido a regra (3.27), podemos estabelecer uma correspondência 1-para-1 entre os termos verticais de C e os termos inseridos após eles no decorrer do processo. Vamos chamar de A' os termos inseridos após os termos verticais de C . Então $\text{VD}(A') = \text{VD}(C)$, pois só foram inseridos os tipos 1' e 4'. Além disso, $\text{VD}(C') = \text{VD}(C) + \text{VD}(A')$, de onde segue (A.2).

Lema 14. *Para qualquer k , se $C' \in CL_k$, então $\text{curto}(C') \in CPL_k$.*

Prova. Vamos mostrar que se $C' \in CL_k$, então $\text{curto}(C')$ satisfaz todas as condições de (3.26).

Prova de (CPL – a). Como $C' \in CL_k$, então de $(CL - a)$, $c_1 = 1$, e como 1 é um tipo principal, então o primeiro termo de $\text{curto}(C')$ também é $c_1 = 1$.

Prova de (CPL – b). Impossibilidade de (1, 3). Se existisse uma tal combinação em $\text{curto}(C')$, então do item (b) de (3.27) existiria uma combinação (1, 1', 3) em C' , mas a combinação (1', 3) é impossível de acordo com o item $(CL - c)$ de (3.24). Impossibilidade de (3, 1). Se existisse uma tal combinação em $\text{curto}(C')$, então do item (c) de (3.27)

existiria uma combinação $(3, 4', 1)$ em C' , mas a combinação $(4', 1)$ é impossível de acordo com o item $(CL - c)$ de (3.24). Impossibilidade de $(1, 4)$. Se existisse uma tal combinação em $\text{curto}(C')$, então do item (b) de (3.27) existiria uma combinação $(1, 1', 4)$ em C' , mas a combinação $(1', 4)$ é impossível de acordo com o item $(CL - c)$ de (3.24). Impossibilidade de $(4, 1)$. Se existisse uma tal combinação em $\text{curto}(C')$, então do item (d) de (3.27) existiria uma combinação $(4, 4', 1)$ em C' , mas a combinação $(4', 1)$ é impossível de acordo com o item $(CL - c)$ de (3.24).

Prova de $(CPL - c)$. Devido ao item $(CL - a)$ de (3.24), o último termo de C' é $4'$. Devido ao item $(CL - c)$ de (3.24) o termo que precede $4'$ em C' é 3 ou 4, portanto este é o último termo de $\text{curto}(C')$ quando $4'$ é eliminado.

Prova de $(CPL - d)$. De (A.1), $\text{HD}(C') = \text{HD}(\text{curto}(C'))$. Como $C' \in CL_k$, então do item $(CL - d)$ de (3.24) $\text{HD}(C') \geq k$. Portanto, $\text{HD}(\text{curto}(C')) \geq k$.

Prova de $(CPL - e)$. Do item $(CL - d)$ de (3.24), $\text{VD}(C') = 0$. De (A.2) temos $\text{VD}(C') = 2 \cdot \text{VD}(\text{curto}(C'))$. Portanto $\text{VD}(\text{curto}(C')) = 0$.

O Lema 14 está provado. □

APÊNDICE B – LEMAS UTILIZADOS NAS DEMONSTRAÇÕES DOS TEOREMAS 1 E 5

Lema 15. *Seja A uma matriz $n \times n$ diagonalizável, e suponha que queremos calcular autovetores “pela esquerda”, isto é, queremos resolver a equação $vA = \lambda v$. Então existe uma matriz inversível P tal que*

$$A = P^{-1}DP$$

onde D é uma matriz diagonal formada pelos autovalores de A .

Prova.

Suponhamos que A é uma matriz $n \times n$ diagonalizável, mas queremos calcular os autovetores “pela esquerda”. Isto é, para cada autovalor $\lambda_1, \lambda_2, \dots, \lambda_n$ de A (não necessariamente distintos), queremos encontrar os autovetores V_1, \dots, V_n correspondentes tal que $V_i A = \lambda_i V_i$. Neste caso, temos que os autovetores são vetores linha, isto é, $V_i = \begin{bmatrix} v_{i1} & v_{i2} & \cdots & v_{in} \end{bmatrix}$. Note que, neste contexto

$$\begin{aligned} V_1 A &= \lambda_1 V_1 &= \begin{pmatrix} \lambda_1 v_{11} & \lambda_1 v_{12} & \cdots & \lambda_1 v_{1n} \end{pmatrix} \\ V_2 A &= \lambda_2 V_2 &= \begin{pmatrix} \lambda_2 v_{21} & \lambda_2 v_{22} & \cdots & \lambda_2 v_{2n} \end{pmatrix} \\ & & \vdots \\ V_n A &= \lambda_n V_n &= \begin{pmatrix} \lambda_n v_{n1} & \lambda_n v_{n2} & \cdots & \lambda_n v_{nn} \end{pmatrix} \end{aligned}$$

ou seja,

$$\begin{bmatrix} V_1 A \\ V_2 A \\ \vdots \\ V_n A \end{bmatrix} = \begin{bmatrix} \lambda_1 V_1 \\ \lambda_2 V_2 \\ \vdots \\ \lambda_n V_n \end{bmatrix}$$

Sendo assim, denotemos por $P = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix}$ a matriz $n \times n$ onde cada linha é composta por um autovetor de A . Daí temos que,

$$\begin{aligned} PA &= \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} A = \begin{bmatrix} V_1 A \\ V_2 A \\ \vdots \\ V_n A \end{bmatrix} = \begin{bmatrix} \lambda_1 V_1 \\ \lambda_2 V_2 \\ \vdots \\ \lambda_n V_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_{11} & \lambda_1 v_{12} & \cdots & \lambda_1 v_{1n} \\ \lambda_2 v_{21} & \lambda_2 v_{22} & \cdots & \lambda_2 v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_n v_{n1} & \lambda_n v_{n2} & \cdots & \lambda_n v_{nn} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \cdot \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{bmatrix} = D \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} \\ &= DP \end{aligned}$$

Como as linhas da matriz P são formadas por vetores linearmente independentes, segue que $\det(P) \neq 0$, logo P é invertível. Sendo assim, multiplicando a equação acima por P^{-1} pela esquerda obtemos

$$A = P^{-1}DP, \quad (\text{B.3})$$

onde P é a matriz cujas linhas são formadas por autovetores de A e D é a matriz diagonal, onde na diagonal principal estão os autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ de A . \square

Observe que sob mesmas hipóteses do Lema 15, quando queremos calcular o autovetor pela direita, temos $A = PDP^{-1}$.

Lema 16. *Sejam A e B matrizes não negativas. Se $A \leq B$, então para todo natural m , $A^m \leq B^m$.*

Prova. Vamos mostrar este resultado usando indução sobre m .

•**Base de indução.** O resultado é válido para $m = 1$ por hipótese.

•**Hipótese de indução.** Suponha que o resultado é válido para $m - 1$ e vamos mostrar para m .

Passo de indução. Note que $A^m = A^{m-1} \cdot A$, mas por hipótese de indução, $A^{m-1} \leq B^{m-1}$, logo, uma vez que A e B são não-negativas, $A^{m-1} \cdot A \leq B^{m-1} \cdot A$ e como por hipótese $A \leq B$, segue que $B^{m-1} \cdot A \leq B^{m-1} \cdot B$. Portanto,

$$A^m \leq B^m.$$

□

Lema 17. Consideremos a função $g : [0, 1] \rightarrow [0, 1]$ dada por

$$g(\alpha) = \frac{-(36(-3 + \alpha)\alpha^2(9\alpha^3 - 77\alpha^2 + 465\alpha - 450))}{(-9\alpha^3 + f(\alpha)\alpha + 77\alpha^2 - 3f(\alpha) - 315\alpha)(9\alpha^3 + f(\alpha)\alpha - 77\alpha^2 - 3f(\alpha) + 315\alpha)},$$

onde

$$f(\alpha) = \sqrt{\frac{\alpha^2(81\alpha^4 - 1386\alpha^3 + 16999\alpha^2 - 47910\alpha + 48825)}{(-3 + \alpha)^2}}.$$

Então, $g(\alpha)$ é estritamente decrescente e além disso $g(0) < 1$ e $g(1) > 0$.

Prova. Inicialmente, vamos analisar o denominador de $g(\alpha)$.

$$\begin{aligned} & (-9\alpha^3 + f(\alpha)\alpha + 77\alpha^2 - 3f(\alpha) - 315\alpha)(9\alpha^3 + f(\alpha)\alpha - 77\alpha^2 - 3f(\alpha) + 315\alpha) \\ &= -81\alpha^6 + 1386\alpha^5 - 11599\alpha^4 + 48510\alpha^3 - 99225\alpha^2 + (f(\alpha))^2(\alpha^2 - 6\alpha + 9) \\ &= -81\alpha^6 + 1386\alpha^5 - 11599\alpha^4 + 48510\alpha^3 - 99225\alpha^2 \\ &\quad + 81\alpha^6 - 1386\alpha^5 + 16999\alpha^4 - 47910\alpha^3 + 48825\alpha^2 \\ &= \alpha^2(5400\alpha^2 + 600\alpha - 50400). \end{aligned}$$

Logo,

$$g(\alpha) = \frac{-(36(-3 + \alpha)\alpha^2(9\alpha^3 - 77\alpha^2 + 465\alpha - 450))}{\alpha^2(5400\alpha^2 + 600\alpha - 50400)},$$

ou seja,

$$g(\alpha) = \frac{3(3 - \alpha)(9\alpha^3 - 77\alpha^2 + 465\alpha - 450)}{50(9\alpha^2 + \alpha - 84)}.$$

Agora vamos derivar a função g . Para isto, denotemos por

$$\varphi(\alpha) = 3(3 - \alpha)(9\alpha^3 - 77\alpha^2 + 465\alpha - 450) \quad \text{e} \quad \psi(\alpha) = 50(9\alpha^2 + \alpha - 84).$$

Sendo assim, como $g(\alpha) = \frac{\varphi(\alpha)}{\psi(\alpha)}$, então $g'(\alpha) = \frac{\varphi'(\alpha)\psi(\alpha) - \varphi(\alpha)\psi'(\alpha)}{(\psi(\alpha))^2}$. E consequentemente,

$$g'(\alpha) = \frac{150[-165\alpha^5 + 909\alpha^4 + 3232\alpha^3 - 47559\alpha^2 + 140778\alpha - 115830]}{[50(9\alpha^2 + \alpha - 84)]^2}.$$

Queremos mostrar que $g'(\alpha) < 0$. Vamos considerar o numerador de g' dividido por 150, denotemos esta quantidade por $h(\alpha)$.

$$h(\alpha) = -165\alpha^5 + 909\alpha^4 + 3232\alpha^3 - 47559\alpha^2 + 140778\alpha - 115830,$$

então,

$$h'(\alpha) = -825\alpha^4 + 3636\alpha^3 + 9696\alpha^2 - 95118\alpha + 140778.$$

Como, $\forall \alpha \in [0, 1]$, $3636\alpha^3 - 825\alpha^5 > 0$, $140778 - 95118\alpha > 0$ e $9696\alpha^2 > 0$, segue que, $h'(\alpha) > 0$. Logo, a função h é estritamente crescente $\forall \alpha \in [0, 1]$. Além disso $h(1) = -18635 < 0$, logo, $h(\alpha) < 0 \forall \alpha \in [0, 1]$. Consequentemente, o numerador de $g'(\alpha)$ é menor que zero e como o denominador é maior que zero, segue que,

$$g'(\alpha) < 0 \quad \forall \alpha \in [0, 1].$$

Portanto, a função $g(\alpha)$ é estritamente decrescente $\forall \alpha \in [0, 1]$. Além disso, $g(0) \cong 0,964$ e $g(1) \cong 0,086$. □

Lema 18. *Se considerarmos $p = \frac{3}{10}$, $q = 1 - \frac{\alpha}{3}$ e $\alpha \in (0, 1)$, então (4.13) é crescente como função de β , onde $\beta \in \left(0, \frac{9\alpha^2}{1000}\right]$.*

Prova. A função $LS(3/10, 1 - \alpha/3, \beta, \alpha)$ é dada por:

$$f_1(\beta) = \frac{800(-3+\alpha)\beta(10\alpha^2-200\alpha\beta-93\alpha+600\beta+90)}{[g_1(\beta)\alpha+10\alpha^2-200\alpha\beta-3g_1(\beta)-63\alpha+600\beta][g_1(\beta)\alpha-10\alpha^2+200\alpha\beta-3g_1(\beta)+63\alpha-600\beta]},$$

onde,

$$g_1(\beta) = \left(\frac{100\alpha^4 - 4000\alpha^3\beta + 40000\alpha^2\beta^2 + 60\alpha^3 + 61200\alpha^2\beta - 240000\alpha\beta^2 + 9\alpha^2}{(-3 + \alpha)^2} - \frac{219600\alpha\beta + 360000\beta^2 + 216000\beta}{(-3 + \alpha)^2} \right)^{1/2}.$$

Note que, podemos reescrever o denominador de $f_1(\beta)$ da seguinte forma

$$\begin{aligned} & [g_1(\beta)\alpha+10\alpha^2-200\alpha\beta-3g_1(\beta)-63\alpha+600\beta][g_1(\beta)\alpha-10\alpha^2+200\alpha\beta-3g_1(\beta)+63\alpha-600\beta] \\ &= \alpha^2(g_1(\beta))^2 - 3\alpha(g_1(\beta))^2 - 3\alpha(g_1(\beta))^2 + 9(g_1(\beta))^2 - 100\alpha^4 + 4000\beta\alpha^3 - 40000\beta^2\alpha^2 \\ & \quad + 1260\alpha^3 - 37200\beta\alpha^2 + 240000\beta^2\alpha + 75600\beta\alpha - 3969\alpha^2 - 360000\beta^2 \\ &= (\alpha^2 - 6\alpha + 9)(g_1(\beta))^2 - 100\alpha^4 + 4000\beta\alpha^3 - 40000\beta^2\alpha^2 + 1260\alpha^3 - 37200\beta\alpha^2 \\ & \quad + 240000\beta^2\alpha - 3969\alpha^2 + 75600\beta\alpha - 360000\beta^2 \\ &= 100\alpha^4 + 4000\alpha^3\beta + 40000\alpha^2\beta^2 + 60\alpha^3 + 61200\alpha^2\beta - 240000\alpha\beta^2 + 9\alpha^2 - 219600\alpha\beta \\ & \quad + 360000\beta^2 + 216000\beta - 100\alpha^4 + 4000\beta\alpha^3 - 40000\beta^2\alpha^2 + 1260\alpha^3 - 37200\beta\alpha^2 \\ & \quad + 240000\beta^2\alpha - 3969\alpha^2 + 75600\beta\alpha - 360000\beta^2 \\ &= 1320\alpha^3 + 24000\alpha^2\beta - 3960\alpha^2 - 144000\alpha\beta + 216000\beta. \end{aligned}$$

Logo,

$$f_1(\beta) = \frac{20\beta(-3 + \alpha)(10\alpha^2 - 200\alpha\beta - 93\alpha + 600\beta + 90)}{5400\beta + 600\alpha^2\beta - 3600\alpha\beta + 33\alpha^3 - 99\alpha^2}.$$

Queremos mostrar que para todo $\beta \in \left(0, \frac{9\alpha^2}{1000}\right]$, $f_1(\beta)$ é uma função crescente com relação a β . Para isto, vamos derivar esta função com relação a β . Assim, obtemos

$$f_1'(\beta) = \frac{20}{3} \frac{110\alpha^4 - 4400\alpha^3\beta - 40000\alpha^2\beta^2 - 1023\alpha^3 + 13200\alpha^2\beta + 240000\alpha\beta^2 + 990\alpha^2 - 360000\beta^2}{(11\alpha^2 + 200\alpha\beta - 600\beta)^2}$$

Note que,

$$0 < \beta < \frac{9\alpha^2}{1000} \Rightarrow 0 < \beta^2 < \frac{36\alpha^4}{1000000},$$

e conseqüentemente,

$$-4400\alpha^3\beta > -\frac{396\alpha^5}{10}, \quad -40000\alpha^2\beta^2 > -\frac{144\alpha^6}{100} \quad \text{e} \quad -360000\beta^2 > -\frac{1296\alpha^4}{100}.$$

Agora vamos analisar algumas parcelas do numerador de $f_1'(\beta)$. Note que,

$$\begin{aligned} & 990\alpha^2 + 110\alpha^4 - 4400\alpha^3\beta - 40000\alpha^2\beta^2 - 1023\alpha^3 - 360000\beta^2 \\ & > 990\alpha^2 + 110\alpha^4 - \frac{396}{10}\alpha^5 - \frac{144}{100}\alpha^6 - 1023\alpha^3 - \frac{1296}{100}\alpha^4 \\ & > 990\alpha^2 + 110\alpha^4 - 39,6\alpha^4 - 1,44\alpha^4 - 1023\alpha^3 - 12,96\alpha^4 \\ & = 990\alpha^2 + 56\alpha^4 - 1023\alpha^3 \\ & = \alpha^2(56\alpha^2 - 1023\alpha + 990). \end{aligned}$$

Claro que esta última igualdade é positiva para todo $\alpha \in (0, 1)$. Agora, analisaremos as parcelas restantes do numerador de $f_1'(\beta)$. Como para todo $\alpha \in (0, 1)$ e para todo $\beta \in \left(0, \frac{9\alpha^2}{1000}\right]$, $13200\alpha^2\beta > 0$ e $240000\alpha\beta^2 > 0$, segue que $f_1'(\beta) > 0$. Portanto, $f_1(\beta)$ é uma função crescente de β para todo $\beta \in \left(0, \frac{9\alpha^2}{1000}\right]$. \square

Considere

$$\lambda_{PF}(p, q, \alpha, \beta) = \frac{r + q + \beta/p + \sqrt{\Delta}}{2}, \quad (\text{B.4})$$

onde $r = (p\alpha + (1 - \alpha))/q$, $\Delta = (r + q + \beta/p)^2 - 4qr$, $p, q, \alpha, \beta \in (0, 1)$.

Lema 19. *Seja $\lambda_{PF}(p, q, \alpha, \beta)$ como definido em (B.4).*

(i) *Se $\beta < p(1 + qr - q - r)$, então $\lambda_{PF}(p, q, \alpha, \beta) < 1$;*

(ii) *Os valores de p, q que maximizam β sob a condição (i) são:*

$$p = \frac{4\alpha - 3 + \sqrt{9 - 8\alpha}}{8\alpha} \quad \text{e} \quad q = \frac{\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}}{4}.$$

Prova de (i).

$$\begin{aligned}
\lambda_{PF}(p, q, \alpha, \beta) < 1 &\iff r + q + \beta/p - 2 < -\sqrt{\Delta} \\
&\iff (r + q + \beta/p - 2)^2 > \Delta \\
&\iff (r + q + \beta/p)^2 - 4(r + q + \beta/p) + 4 > (r + q + \beta/p)^2 - 4qr \\
&\iff r + q + \beta/p - 1 < qr \\
&\iff \beta < p(1 + qr - q - r).
\end{aligned}$$

Prova de (ii). Seja

$$f(p, q) = p(1 + qr - q - r) \quad (\text{B.5})$$

então, $f(p, q) = p + p^2\alpha + p(1 - \alpha) - pq - (p^2\alpha + p(1 - \alpha))/q$. Nosso objetivo é maximizar $f(p, q)$. Note que

$$\frac{\partial f}{\partial p}(p, q) = 1 + 2p\alpha + 1 - \alpha - q - \frac{2p\alpha + 1 - \alpha}{q} \quad \text{e} \quad \frac{\partial f}{\partial q}(p, q) = -p + \frac{p^2\alpha + p(1 - \alpha)}{q^2}.$$

Daí, $\nabla f(p, q) = 0 \iff \left(\frac{\partial f}{\partial p}(p, q), \frac{\partial f}{\partial q}(p, q) \right) = (0, 0)$. Ou seja,

$$\begin{cases} 1 + 2p\alpha + 1 - \alpha - q - \frac{2p\alpha + 1 - \alpha}{q} = 0, \\ -p + \frac{p^2\alpha + p(1 - \alpha)}{q^2} = 0. \end{cases} \quad (\text{B.6})$$

Resolvendo o sistema (B.6). Da segunda equação,

$$\begin{aligned}
\frac{p^2\alpha + p(1 - \alpha)}{q^2} = p &\iff q^2 = \frac{p^2\alpha + p(1 - \alpha)}{p} \\
&\iff q^2 = p\alpha + 1 - \alpha \\
&\iff q = \pm\sqrt{p\alpha + 1 - \alpha}.
\end{aligned}$$

Uma vez que q é um número positivo, devemos considerar $q = \sqrt{p\alpha + 1 - \alpha}$. Substituindo esse valor de q na primeira equação de (B.6),

$$2p\alpha + 2 - \alpha = \sqrt{p\alpha + 1 - \alpha} + \frac{2p\alpha + 1 - \alpha}{\sqrt{p\alpha + 1 - \alpha}}.$$

Após algumas manipulações algébricas, obtemos que esta igualdade é equivalente a:

$$4p^3\alpha^3 - 8p^2\alpha^3 + 3p^2\alpha^2 + 5p\alpha^3 - 4p\alpha^2 - \alpha^3 + \alpha^2 = 0. \quad (\text{B.7})$$

As raízes deste polinômio do terceiro grau em relação a p são:

$$p_1 = 1, \quad p_2 = \frac{4\alpha - 3 + \sqrt{9 - 8\alpha}}{8\alpha} \quad \text{e} \quad p_3 = \frac{4\alpha - 3 - \sqrt{9 - 8\alpha}}{8\alpha}.$$

Para $\alpha \in (0, 1)$ temos que $p_3 < 0$, logo desconsideramos esta solução. Note que a relação $x < p^{-x}$ não é satisfeita para $p_1 = 1$. Como $p_2 \in (0, 1)$, vamos considerar esta solução.

Logo, a solução da equação (B.7) que vamos considerar é:

$$p = \frac{4\alpha - 3 + \sqrt{9 - 8\alpha}}{8\alpha}.$$

Dai,

$$q = \frac{\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}}{4}$$

Portanto, o ponto crítico para $f(p, q)$ é

$$(p_0, q_0) = \left(\frac{4\alpha - 3 + \sqrt{9 - 8\alpha}}{8\alpha}, \frac{\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}}{4} \right).$$

Resta-nos verificar se este é um ponto de máximo. Note que

$$\frac{\partial^2 f}{\partial p^2}(p, q) = 2\alpha - \frac{2\alpha}{q}, \quad \frac{\partial^2 f}{\partial q^2}(p, q) = -\frac{2(p^2\alpha + p(1 - \alpha))}{q^3}$$

e

$$\frac{\partial^2 f}{\partial p \partial q}(p, q) = \frac{\partial^2 f}{\partial q \partial p}(p, q) = -1 + \frac{2p\alpha + 1 - \alpha}{q^2}.$$

Seja

$$H(p, q) = \det \begin{pmatrix} \frac{\partial^2 f}{\partial p^2}(p, q) & \frac{\partial^2 f}{\partial q \partial p}(p, q) \\ \frac{\partial^2 f}{\partial p \partial q}(p, q) & \frac{\partial^2 f}{\partial q^2}(p, q) \end{pmatrix},$$

daí,

$$\begin{aligned} H(p_0, q_0) &= \\ \det &\begin{pmatrix} \frac{2\alpha\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}} - 8\alpha}{\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}} & \frac{-8\alpha - 6 + 2\sqrt{9 - 8\alpha}}{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}} \\ \frac{-8\alpha - 6 + 2\sqrt{9 - 8\alpha}}{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}} & -\frac{2(4\alpha - 3 + \sqrt{9 - 8\alpha})^2 + 16(4\alpha - 3 + \sqrt{9 - 8\alpha})(1 - \alpha)}{\alpha(-8\alpha + 10 + 2\sqrt{9 - 8\alpha})^{3/2}} \end{pmatrix} \\ &= \left(\frac{2\alpha\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}} - 8\alpha}{\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}} \right) \times \\ &\quad \left(-\frac{2(4\alpha - 3 + \sqrt{9 - 8\alpha})^2 + 16(4\alpha - 3 + \sqrt{9 - 8\alpha})(1 - \alpha)}{\alpha(-8\alpha + 10 + 2\sqrt{9 - 8\alpha})^{3/2}} \right) \end{aligned}$$

$$- \left(\frac{-8\alpha - 6 + 2\sqrt{9 - 8\alpha}}{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}} \right)^2 > 0, \text{ para todo } \alpha \in (0, 1).$$

Além disso,

$$\frac{\partial^2 f}{\partial p^2}(p_0, q_0) = \frac{2\alpha\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}} - 8\alpha}{\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}} < 0, \text{ para todo } \alpha \in (0, 1).$$

Assim, concluímos que (p_0, q_0) é um ponto de máximo para $f(p, q)$. \square

Note que, para p_0 e q_0 definidos anteriormente,

$$f(p_0, q_0) = \frac{1}{8\alpha} \left((4\alpha - 3 + \sqrt{9 - 8\alpha}) \left(-\frac{\alpha}{2} + \frac{13 + \sqrt{9 - 8\alpha}}{8} - \frac{\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}}{4} - \frac{4(-\alpha/2 + (5 + \sqrt{9 - 8\alpha})/8)}{\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}} \right) \right)$$

Para provar o Lema 20, utilizaremos o seguinte resultado. Sejam $a, b \in \mathbb{R}$, então

$$\begin{aligned} \left(1 - \frac{a+b}{2}\right) \left(1 - \frac{a-b}{2}\right) &= 1 - \frac{a-b}{2} - \frac{a+b}{2} + \frac{a+b}{2} \frac{a-b}{2} \\ &= 1 - \frac{a-b+a+b}{2} + \frac{a^2-ab+ab-b^2}{4} \\ &= 1 - a + \frac{a^2-b^2}{4} \end{aligned} \tag{B.8}$$

Lema 20. *Dados $p, q, r, \alpha \in (0, 1)$. A função $LS_1(p, q, \alpha, \beta)$ em (3.37) é crescente como função de β , para $\beta \in (0, \alpha^2/55)$.*

Prova. Usando (B.8) temos que

$$\begin{aligned} (1 - \lambda_{PF})(1 - \lambda_2) &= 1 - (r + q + \beta/p) + \frac{(r + q + \beta/p)^2 - [(r + q + \beta/p)^2 - 4qr]}{4} \\ &= 1 - r - q - \beta/p + qr, \end{aligned}$$

daí, $p(1 - \lambda_{PF})(1 - \lambda_2) = p(1 + qr - q - r - \beta/p)$. Logo,

$$LS_1(p, q, \alpha, \beta) = \frac{\beta\lambda_{PF}}{p(1 + qr - q - r - \beta/p)}.$$

Usaremos λ'_{PF} para denotar a derivada de λ_{PF} em relação a β . Sabemos que: $\lambda_{PF} > 0$, λ_{PF} é crescente como função de β . Logo, $\lambda'_{PF} > 0$. Além disso

$$p(1 + qr - q - r - \beta/p) = p(1 - \lambda_{PF})(1 - \lambda_2) > 0.$$

Derivando LS_1 com relação a β obtemos

$$LS'_1(p, q, \alpha, \beta) = \frac{[\lambda_{PF} + \beta\lambda'_{PF}][p(1 + qr - q - r - \beta/p)] + \beta\lambda_{PF}}{[p(1 + qr - q - r - \beta/p)]^2},$$

o qual é positivo pelas observações anteriores.

Assim, $LS_1(p, q, \alpha, \beta)$ é crescente com relação a β . \square

Considerando $p = (4\alpha - 3 + \sqrt{9 - 8\alpha})/8\alpha$, $q = \sqrt{p\alpha + 1 - \alpha}$ e $r = (p\alpha + 1 - \alpha)/q$, então

$$p(1 + qr - q - r) = \frac{1}{64\alpha\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}} \left[(4\alpha - 3 + \sqrt{9 - 8\alpha}) \times \right. \\ \left. \times \left(\sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}}(-4\alpha + 13 + \sqrt{9 - 8\alpha}) + 32\alpha - 40 - 8\sqrt{9 - 8\alpha} \right) \right].$$

Denotemos esta função por $f_2(\alpha)$, isto é,

$$f_2(\alpha) = p(1 + qr - q - r)$$

Vamos provar o seguinte lema,

Lema 21. Para todo $\alpha \in (0, 1)$,

$$f_2(\alpha) \geq \frac{\alpha^2}{27}.$$

Mas, antes faremos algumas observações.

Seja $V_\alpha = \sqrt{-8\alpha + 10 + 2\sqrt{9 - 8\alpha}} = \sqrt{9 - 8\alpha + 2\sqrt{9 - 8\alpha} + 1} = \sqrt{9 - 8\alpha} + 1$, então

$$\begin{aligned} f_2(\alpha) &= \frac{1}{64\alpha V_\alpha} \left[(4\alpha - 3 + \sqrt{9 - 8\alpha}) \left((13 - 4\alpha + \sqrt{9 - 8\alpha})V_\alpha - 4V_\alpha^2 \right) \right] \\ &= \frac{1}{64\alpha} \left[(4\alpha - 3 + \sqrt{9 - 8\alpha}) \left(13 - 4\alpha + \sqrt{9 - 8\alpha} - 4(\sqrt{9 - 8\alpha} + 1) \right) \right] \\ &= \frac{1}{64\alpha} \left[(4\alpha - 3 + \sqrt{9 - 8\alpha}) (9 - 4\alpha - 3\sqrt{9 - 8\alpha}) \right] \\ &= \frac{1}{64\alpha} \left[36\alpha - 16\alpha^2 - 12\alpha\sqrt{9 - 8\alpha} - 27 + 12\alpha + 9\sqrt{9 - 8\alpha} + 9\sqrt{9 - 8\alpha} \right. \\ &\quad \left. - 4\alpha\sqrt{9 - 8\alpha} - 3(9 - 8\alpha) \right] \\ &= \frac{1}{64\alpha} \left[-16\alpha^2 + 72\alpha - 16\alpha\sqrt{9 - 8\alpha} + 18\sqrt{9 - 8\alpha} - 54 \right] \\ &= \frac{-2\alpha + 9}{8} + \frac{1}{32\alpha} \left[(9 - 8\alpha)\sqrt{9 - 8\alpha} - 27 \right]. \end{aligned}$$

Prova do Lema 21.

$$\begin{aligned} f_2(\alpha) &\geq \frac{\alpha^2}{27} \\ \iff \frac{1}{32\alpha} \left[(9 - 8\alpha)\sqrt{9 - 8\alpha} - 27 \right] &\geq \frac{\alpha^2}{27} - \frac{-2\alpha + 9}{8} \end{aligned}$$

$$\begin{aligned}
&\iff \frac{1}{32\alpha} [(9 - 8\alpha)\sqrt{9 - 8\alpha} - 27] \geq \frac{8\alpha^2 + 54\alpha - 243}{216} \\
&\iff \frac{216}{32\alpha} [(9 - 8\alpha)\sqrt{9 - 8\alpha} - 27] \geq 8\alpha^2 + 54\alpha - 243 \\
&\iff \frac{27}{4\alpha} [(9 - 8\alpha)\sqrt{9 - 8\alpha} - 27] \geq 8\alpha^2 + 54\alpha - 243 \\
&\iff 27 [(9 - 8\alpha)\sqrt{9 - 8\alpha} - 27] \geq 4\alpha (8\alpha^2 + 54\alpha - 243) \\
&\iff 27(9 - 8\alpha)\sqrt{9 - 8\alpha} \geq 4\alpha (8\alpha^2 + 54\alpha - 243) + 729. \tag{B.9}
\end{aligned}$$

Uma vez que $4\alpha (8\alpha^2 + 54\alpha - 243) + 729 > 0$ para todo $\alpha \in (0, 1)$, colocando ambos os lados da desigualdade (B.9) ao quadrado, obtemos:

$$\begin{aligned}
&27^2(9 - 8\alpha)^2(9 - 8\alpha) \geq (4\alpha (8\alpha^2 + 54\alpha - 243) + 729)^2 \\
&\iff -373248\alpha^3 + 1259712\alpha^2 - 1417176\alpha + 531441 \geq 1024\alpha^6 + 13824\alpha^5 - 15552\alpha^4 \\
&\quad - 373248\alpha^3 + 1259712\alpha^2 - 1417176\alpha + 531441 \\
&\iff 1024\alpha^6 + 13824\alpha^5 - 15552\alpha^4 \leq 0 \\
&\iff 64\alpha^4(16\alpha^2 + 216\alpha - 243) \leq 0 \\
&\iff 16\alpha^2 + 216\alpha - 243 \leq 0.
\end{aligned}$$

Considere a função $g_2 : [0, 1] \rightarrow \mathbb{R}$ dada por $g_2(\alpha) = 16\alpha^2 + 216\alpha - 243$. Então $g_2'(\alpha) = 32\alpha + 216$. Logo, $g_2'(\alpha) > 0$ para todo $\alpha \in [0, 1]$. Assim concluímos que $g_2(\alpha)$ é uma função estritamente crescente. Além disso, $g_2(1) = 16 + 216 - 243 = -11 < 0$. Portanto, $g_2(\alpha) < 0$, para todo $\alpha \in [0, 1]$. Logo, $16\alpha^2 + 216\alpha - 243 \leq 0$ para todo $\alpha \in (0, 1)$, ou equivalentemente, $f_2(\alpha) \geq \alpha^2/27$. \square