

UNIVERSIDADE PRESBITERIANA MACKENZIE
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E COMPUTAÇÃO

Brian Lee Mayer

EXPLORANDO SEQUÊNCIAS DE d -PRIMOS
E DE DIVISORES COM ALGORITMOS DE VISIBILIDADE

Orientador: Dr. Luiz Henrique Alves Monteiro

São Paulo
2023

UNIVERSIDADE PRESBITERIANA MACKENZIE
PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA ELÉTRICA E COMPUTAÇÃO

Brian Lee Mayer

EXPLORANDO SEQUÊNCIAS DE d -PRIMOS
E DE DIVISORES COM ALGORITMOS DE VISIBILIDADE

Tese de doutorado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Computação da Universidade Presbiteriana Mackenzie como requisito para a obtenção do grau de Doutor em Engenharia Elétrica e Computação

Orientador: Dr. Luiz Henrique Alves Monteiro

São Paulo
2023

Elaborado pelo Sistema de Geração Automática de Ficha Catalográfica da Mackenzie
com os dados fornecidos pelo(a) autor(a)

M468e Mayer, Brian Lee.
Explorando sequências de d-primos e de divisores com algoritmos
de visibilidade : [recurso eletrônico] / Brian lee Mayer.
20000 KB ; il.

Tese (Doutorado em Engenharia Elétrica e Computação) -
Universidade Presbiteriana Mackenzie, São Paulo, 2023.
Orientador(a): Prof(a). Dr(a). Luiz Henrique Alves Monteiro.
Referências Bibliográficas: f. 52-54.

1. Algoritmo De Visibilidade. 2. D-primo. 3. Entropia
Informacional. 4. Função Quantidade De Divisores. 5. Rede Complexa.
I. Monteiro, Luiz Henrique Alves, *orientador(a)*. II. Título.

Bibliotecário(a) Responsável: Maria Gabriela Brandi Teixeira - CRB 8/6339

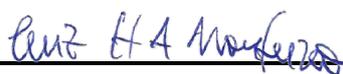
BRIAN LEE MAYER

INVESTIGANDO SEQUÊNCIAS DE D-PRIMOS E DE DIVISORES
COM ALGORITMOS DE VISIBILIDADE

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e Computação da Universidade Presbiteriana Mackenzie como requisito parcial à obtenção de título de Doutor em Engenharia Elétrica e Computação.

Aprovado em 03/08/2023

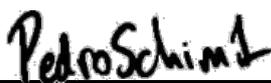
BANCA EXAMINADORA



Prof. Dr. Luiz Henrique Alves Monteiro
Universidade Presbiteriana Mackenzie



Prof. Dr. Marco Eisencraft
Universidade de São Paulo



Prof. Dr. Pedro Henrique Triguís Schimit
Universidade Nove de Julho



Prof. Dr. Pedro Paulo Balbi de Oliveira
Universidade Presbiteriana Mackenzie



Prof. Dr. Rogério de Oliveira
Universidade Presbiteriana Mackenzie

Folha de Identificação da Agência de Financiamento

Autor: Brian Lee Mayer

Programa de Pós-Graduação *Stricto Sensu* em Engenharia Elétrica e Computação

Título do Trabalho: Explorando sequências divisores e de d-primos com algoritmos de visibilidade

O presente trabalho foi realizado com o apoio de ¹:

- CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
- CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico
- FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo
- Instituto Presbiteriano Mackenzie/Isenção integral de Mensalidades e Taxas
- MACKPESQUISA - Fundo Mackenzie de Pesquisa
- Empresa/Indústria:
- Outro:

¹ **Observação:** caso tenha usufruído mais de um apoio ou benefício, selecione-os.

AGRADECIMENTOS

A Deus, por ter criado o universo através da Matemática.

À Simone, minha mãe e meu falecido pai, Robert, por terem me dado a vida e me criado com todo amor do mundo. Por terem me ensinado o valor do estudo, da ética e do amor.

Ao meu irmão Brenno, por ser meu parceiro desde sempre, que me ajuda mesmo com as ideias mais absurdas que eu tenho.

À minha família, por me amarem e me darem suporte para conseguir estudar e escrever essa tese. Por me incentivarem na vida e pelos sábios conselhos e ensinamentos.

À minha namorada Vanessa, mãe do meu filho, por me amar do jeito que eu sou. Por me aguentar no dia-a-dia, muitas vezes se sacrificando para que eu pudesse me dedicar aos estudos. E ainda por amar tudo isso.

Ao meu filho William, por já me trazer tanta alegria sem nem sequer falar uma palavra.

Aos meus queridos professores de matemática do EF e EM, que, devido ao enorme talento, despertaram o meu amor por essa ciência.

Ao meu orientador Luiz, por fazer parte dos estudos mais desafiadores da minha vida, por ter aceito, mais uma vez, me orientar. Por todo o conhecimento ensinado e pela paciência comigo.

Aos professores do PPGEEC, que lecionaram as melhores aulas que já assisti.

À Universidade Presbiteriana Mackenzie, pela bolsa de estudos e pela oportunidade de continuar meu estudos com pessoas muito bem qualificadas.

Aos meus amigos, por me motivarem e pela ajuda. Agradeço à Gina por todo o suporte com os estudos e pelas entusiasmadas trocas de conhecimento.

RESUMO

Seja d -primo um número natural com exatamente d divisores. De acordo com essa definição, os números primos usuais correspondem ao caso particular $d = 2$. Parte desta tese consiste em investigar computacionalmente as sequências numéricas que correspondem às lacunas entre d -primos consecutivos, com $d \in \{2, 3, \dots, 16\}$. A partir dessas sequências, são construídos grafos usando algoritmos de visibilidade natural e de visibilidade horizontal e, então, a topologia desses grafos é analisada. Nessas análises, são também calculadas a entropia informacional das sequências de lacunas e a densidade dos d -primos. As simulações computacionais mostram que os grafos gerados a partir das lacunas entre d -primos consecutivos têm, em geral, uma distribuição de graus que é livre-de-escala. Essas simulações mostram que a densidade de d -primos para d par é muito maior do que para d ímpar e que a entropia informacional é maior para d ímpar do que para d par, para $d \in \{2, 3, \dots, 16\}$. Nesta tese, também calcula-se a entropia informacional e constróem-se grafos de visibilidade natural e de visibilidade horizontal a partir das sequências formadas pela quantidade de divisores dos números naturais e pela aplicação recursiva da função quantidade de divisores sobre essa sequência. Para efeito de comparação, o mesmo trabalho é feito considerando os números felizes, que são números naturais que acabam convergindo para 1 ao se calcular recursivamente a soma dos quadrados de seus dígitos. A topologia desses grafos é caracterizada, sugerindo que as sequências que os geraram têm um caráter quasi-aleatório.

Palavras-chave: *algoritmo de visibilidade, d -primo, entropia informacional, função quantidade de divisores, rede complexa, sequência numérica.*

ABSTRACT

Let d -prime be a natural number with exactly d divisors. According to this definition, the usual prime numbers correspond to the particular case $d = 2$. Part of this thesis consists of computationally investigating the numerical sequences obtained from the gaps between consecutive d -primes, with $d \in \{2, 3, \dots, 16\}$. From these sequences, graphs are constructed by using natural visibility and horizontal visibility algorithms and, then, the topology of these graphs is analyzed. In these analyses, the informational entropy of the gap sequences and the density of the d -primes are also calculated. Computer simulations show that the graphs generated from the gaps between consecutive d -primes have, in general, a scale-free degree distribution. These simulations show that the density of d -primes for d even is much greater than for d odd and that the informational entropy is greater for d odd than for d even, for $d \in \{2, 3, \dots, 16\}$. In this thesis, informational entropy is calculated and graphs of natural visibility and horizontal visibility are also built from the sequences formed by the number of divisors of the natural numbers and by the recursive application of the divisor function on this sequence. For comparison purposes, the same work is done by considering the happy numbers, which are natural numbers that eventually converge to 1, by recursively computing the sum of the squares of their digits. The topology of these graphs is characterized, suggesting that the sequences that generated them have a quasi-random character.

Keywords: *complex network, d -prime, divisor function, informational entropy, numerical sequence, visibility algorithm.*

Sumário

| | | |
|----------|---|-----------|
| 1 | INTRODUÇÃO | 1 |
| 2 | DEFINIÇÕES E MÉTODOS | 8 |
| 2.1 | Definições | 8 |
| 2.1.1 | Sequência | 8 |
| 2.1.2 | Número primo | 8 |
| 2.1.3 | Teorema Fundamental da Aritmética | 9 |
| 2.1.4 | Número d -primo | 9 |
| 2.1.5 | Número feliz | 11 |
| 2.1.6 | Função quantidade de divisores | 11 |
| 2.1.7 | Entropia informacional | 12 |
| 2.1.8 | Lei de potência | 12 |
| 2.2 | Métodos | 13 |
| 2.2.1 | Visibilidade natural | 13 |
| 2.2.2 | Visibilidade horizontal | 15 |
| 2.3 | Algumas medidas em grafos | 16 |
| 2.4 | Relação entre d -primos e d primo | 17 |
| 3 | LACUNAS ENTRE d-PRIMOS SUCESSIVOS | 18 |
| 3.1 | Resultados | 20 |
| 3.2 | Discussão | 27 |
| 4 | DENSIDADE DE d-PRIMOS | 30 |
| 4.1 | Resultados | 30 |
| 4.2 | Discussão | 33 |
| 5 | QUANTIDADE DE DIVISORES DOS NÚMEROS NATURAIS E FELIZES | 36 |
| 5.1 | Resultados | 37 |
| 5.1.1 | Para os números naturais | 38 |

| | | |
|----------|--|-----------|
| 5.1.2 | Para os números felizes | 41 |
| 5.2 | Discussão | 45 |
| 6 | CONCLUSÃO | 48 |
| 6.1 | Trabalhos feitos não relatados | 49 |
| 6.2 | Trabalhos futuros | 50 |
| | REFERÊNCIAS BIBLIOGRÁFICAS | 54 |

Lista de Figuras

| | | |
|-----|--|----|
| 2.1 | A sequência $\{x(n)\}$ em função de n , com $x(1) = 5$, $x(2) = 19$, $x(3) = 24$ e $x(4) = 72$. As setas em azul mostram as conexões permitidas. A seta em laranja indica uma conexão proibida. | 13 |
| 2.2 | Grafo de VN gerado a partir da sequência mostrada na Figura 2.1. | 14 |
| 2.3 | A sequência $\{x(n)\}$ em função de n , com $x(1) = 5$, $x(2) = 19$, $x(3) = 24$ e $x(4) = 72$. As setas em azul mostram as conexões permitidas. A seta em laranja indica uma conexão proibida. | 15 |
| 2.4 | Grafo de VH gerado a partir da sequência mostrada na Figura 2.3. | 15 |
| 3.1 | As primeiras 100 arestas dos grafos para $d \in \{2, \dots, 6\}$. Esquerda: VN; direita: VH. | 22 |
| 3.2 | As primeiras 100 arestas dos grafos para $d \in \{7, \dots, 11\}$. Esquerda: VN; direita: VH. | 23 |
| 3.3 | As primeiras 100 arestas dos grafos para $d \in \{12, \dots, 16\}$. Esquerda: VN; direita: VH. | 24 |
| 3.4 | A distribuição de graus encontrada nos grafos de VN (pontos) e a função ajustada $Ak^{-\gamma}$ (linha sólida) para $d \in \{2, 3, \dots, 16\}$ | 25 |
| 3.5 | A distribuição de graus encontrada nos grafos de visibilidade horizontal (pontos) e a função ajustada $Ak^{-\gamma}$ (linha sólida). A linha pontilhada representando $P_{rand}(k) = (1/3)(2/3)^{k-2}$ foi incluída apenas por comparação. | 26 |
| 4.1 | Densidades de d -primos para $d = 2, 3, \dots, 16$. No eixo horizontal, estão os naturais e no vertical, o valor de $\Pi_d(n)/n$: a fração de d -primos em relação a n | 32 |
| 4.2 | Gráfico de $R(n) = \frac{\Pi_d(n)/n}{f_d(n)}$ em função de n para $d = 2, 3, \dots, 16$ | 34 |
| 5.1 | Grafos gerados usando o algoritmo de visibilidade natural. Esquerda: $\{D(n)\}$; direita: $\{T(n)\}$ | 39 |

| | | |
|------|---|----|
| 5.2 | A entropia informacional normalizada $\Delta = H/H_{max}$ para $\{D(n)\}$ (a) and $\{T(n)\}$ (b) em função do tamanho da sequência N | 39 |
| 5.3 | Distribuição de graus $P(k)$ para $\{n, D(n)\}$ (a) e $\{n, T(n)\}$ (b) para os grafos obtidos pela visibilidade natural. A linha sólida é a curva ajustada $P(k) = Ak^{-\gamma}e^{-\delta k}$ | 40 |
| 5.4 | Grafos gerados usando o algoritmo de visibilidade horizontal. Esquerda: $\{D(n)\}$; direita: $\{T(n)\}$ | 40 |
| 5.5 | Distribuição de graus $P(k)$ para $\{n, D(n)\}$ (a) e $\{n, T(n)\}$ (b) para os grafos obtidos pela visibilidade horizontal. A linha sólida é a curva ajustada $P(k) = Ak^{-\gamma}e^{-\delta k}$ | 41 |
| 5.6 | Grafos gerados usando o algoritmo de visibilidade natural. Esquerda: $\{h, D(h)\}$; direita: $\{h, T(h)\}$ | 42 |
| 5.7 | A entropia informacional normalizada $\Delta = H/H_{max}$ para $\{D(h)\}$ (a) e $\{T(h)\}$ (b) em função do tamanho da sequência N | 43 |
| 5.8 | Distribuição de graus $P(k)$ para $\{h, D(h)\}$ (a) e $\{h, T(h)\}$ (b) para os grafos obtidos pela visibilidade natural. A linha sólida é a curva ajustada $P(k) = Ak^{-\gamma}e^{-\delta k}$ | 43 |
| 5.9 | Grafos gerados usando o algoritmo de visibilidade horizontal. Esquerda: $\{h, D(h)\}$; direita: $\{h, T(h)\}$ | 44 |
| 5.10 | Distribuição de graus $P(k)$ para $\{h, D(h)\}$ (a) e $\{h, T(h)\}$ (b) para os grafos obtidos pela visibilidade horizontal. A linha sólida é a curva ajustada $P(k) = Ak^{-\gamma}e^{-\delta k}$ | 44 |

Lista de Tabelas

| | | |
|-----|--|----|
| 2.1 | Sequências de d -primos denotados por $p_d(n)$ | 10 |
| 3.1 | Sequências das diferenças de primeira ordem de d -primos consecutivos, dados por $x_d(n) = p_d(n + 1) - p_d(n)$ com $n \in \mathbb{N}^*$, para $d \in \{2, 3, \dots, 16\}$ | 19 |
| 3.2 | Entropia informacional normalizada Δ de $x_d(n)$ para $d \in \{2, 3, \dots, 16\}$. . . | 20 |
| 3.3 | Valores de A e γ correspondentes ao ajuste da distribuição de graus $P(k) = Ak^{-\gamma}$ e o grau médio $\langle k \rangle$ em função de d para os grafos de VN. | 27 |
| 3.4 | Valores de A e γ correspondentes ao ajuste da distribuição de graus $P(k) = Ak^{-\gamma}$ e o grau médio $\langle k \rangle$ em função de d para os grafos de VH. | 28 |
| 4.1 | Valores de a_d e b_d e os correspondentes desvios padrões σ da função ajustada. | 31 |
| 5.1 | Valores do grau médio $\langle k \rangle$, caminho mínimo médio $\langle l \rangle$, coeficiente de clusterização médio $\langle c \rangle$, e A , γ , e δ da distribuição de graus $P(k) = Ak^{-\gamma}e^{-\delta k}$ (e o erro quadrático médio (EQM) da função ajustada) dos grafos obtidos empregando os algoritmos de visibilidades natural e horizontal. | 38 |
| 5.2 | Porcentagens de nós ímpares (χ_1), pares (χ_2), ligações entre nós ímpares (m_{11}), ligações de um ímpar para um nó par (m_{12}), ligações de um nó par para um ímpar (m_{21}), e ligações entre nós pares (m_{22}) nos quatro grafos. O número total de ligações (M) é dado na última coluna. | 38 |
| 5.3 | Valores do grau médio $\langle k \rangle$, caminho mínimo médio $\langle l \rangle$, coeficiente de clusterização médio $\langle c \rangle$, e A , γ , e δ da distribuição de graus $P(k) = Ak^{-\gamma}e^{-\delta k}$ (e o erro quadrático médio (EQM) da função ajustada) dos grafos obtidos empregando os algoritmos de visibilidades natural e horizontal. | 41 |

5.4 Porcentagens de nós ímpares (χ_1), pares (χ_2), ligações entre nós ímpares (m_{11}), ligações de um ímpar para um nó par (m_{12}), ligações de um nó par para um ímpar (m_{21}), e ligações entre nós pares (m_{22}) nos quatro grafos. O número total de ligações (M) é dado na última coluna. 42

Capítulo 1

INTRODUÇÃO

Número é um conceito matemático usado para representar uma quantidade ou uma posição numa lista ordenada. Números permitem contar (as horas do dia), medir (a área de um terreno), ordenar (os atendimentos na fila do banco). O conjunto dos números naturais $\{0, 1, 2, 3, 4, 5, \dots\}$ foi o primeiro a surgir na História e é a base desta tese.

Um *número primo* é um número natural maior do que 1 que possui apenas dois divisores: o número 1 e ele próprio. Desde a Antiguidade, os números primos têm fascinado os matemáticos e os entusiastas da Matemática e atraído a atenção de grandes mentes, como Eratóstenes, Erdős, Euclides, Euler, Fermat, Gauss e Riemann (Diamond, 1982; Hardy; Wright, 2008). A primeira referência a um algoritmo para descobrir os números primos dentro de uma sequência numérica foi o crivo de Eratóstenes (Hoche, 1866), proposto por volta de 250 aC. Esse crivo funciona da seguinte maneira. Suponha que se deseja encontrar todos os primos menores ou iguais a n . Para isso:

1. crie uma lista de inteiros consecutivos de 2 até n ;
2. inicialmente, tome $p = 2$, o primeiro número primo;
3. elimine todos os múltiplos de p até o fim da lista;
4. considere o próximo número que não foi eliminado (no caso, $p = 3$) e repita o passo anterior;
5. repita o passo 3, considerando o próximo número da lista (no caso, $p = 5$);
6. repita o passo 3 até que não haja número eliminado. Quando isso ocorrer, pare. Ao final da execução desse algoritmo, os números que não foram eliminados são primos.

Os números primos são conhecidos por serem fundamentais em estudos em muitas áreas, como a Teoria dos Números (Diamond, 1982; Hardy; Wright, 2008). Na Matemática

Aplicada, os números primos são usados em criptografia (Maurer, 1995). Um exemplo clássico é o sistema de criptografia conhecido como RSA (em homenagem aos autores), que se baseia na dificuldade prática de se fatorarem números com muitos dígitos (da ordem de centenas ou milhares de bits) (Rivest; Shamir; Adleman, 1978). Na Biologia, algumas espécies de cigarras possuem ciclo de vida cuja duração é um número primo de anos, como 13 ou 17 anos, provavelmente como estratégia de sobrevivência (Yoshimura; Hayashi; Tanaka; Tainaka; Simon, 2009). Na Física, os primos caracterizam o espectro de energia de sistemas quânticos caóticos (Berry; Keating, 1999).

Se um número natural não é primo, então ele é composto. O Teorema Fundamental da Aritmética, que remete aos trabalhos de Euclides feitos por volta de 300 aC, estabelece que qualquer número natural possui uma única decomposição em fatores primos. Por exemplo, $4044 = 2^2 \times 3^1 \times 337^1$. Talvez, os números primos tenham ganhado esse nome porque são números primários, ou seja, os números a partir dos quais se escrevem os demais. Acredita-se que o conceito de número primo se deve a Euclides (Garbi, 2009); entretanto, esse conceito também aparece em escritos antigos chineses e indianos. Em meados do século XIX, Kulik determinou, a mão, os primos até 100.000.000 (Eves; Kulik, 1983). As tabelas anteriores listavam os primos até 100.000.

Por volta de 1800, Gauss, um dos maiores matemáticos de todos os tempos, paralelamente a Legendre (Dickson, 2005), conjecturou a respeito da função contagem de números primos, denotada por $\Pi(n)$, que expressa a quantidade de primos menores ou iguais a n . Para Gauss, $\Pi(n)$ se aproxima de:

$$\frac{n}{\ln(n)}$$

para $n \rightarrow \infty$. Ou seja:

$$\lim_{n \rightarrow \infty} \frac{\Pi(n)}{n/\ln(n)} = 1$$

Por exemplo, para $n = 10$, $\Pi(10) = 4$ (pois há quatro primos menores do que 10) e $10/\ln(10) \simeq 4.34$; logo, $\Pi(10)/[10/\ln(10)] \simeq 0.92$; para $n = 10^{10}$, $\Pi(10^{10}) = 455052511$ e $10^{10}/\ln(10^{10}) \simeq 4.34 \times 10^8$; logo $\Pi(10^{10})/[10^{10}/\ln(10^{10})] \simeq 1.05$. Para Legendre, no lugar de $\ln(n)$, escreve-se $\ln(n) - 1.08$. Desde esses trabalhos de Gauss e Legendre, a fórmula para $\Pi(n)$ continuou sendo investigada, por exemplo, por Chebyshev (Ford; Green; Konyagin; Tao, 2016). Hadamard provou, por volta de 1900, que a fórmula proposta por Gauss é a correta (Ireland; Rosen, 1998). Erdős, nos anos 50 do século passado, apresentou uma

prova mais simples desse resultado (Burton, 2012).

Em 1737, Euler encontrou uma identidade que relaciona os números primos com a função $\zeta(s)$, hoje conhecida como função zeta de Euler-Riemman, dada por:

$$\zeta(s) = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \dots$$

em que s é um número real. Euler mostrou que $\zeta(s)$ também se escreve como:

$$\zeta(s) = \frac{1}{(1 - (1/2^s))} \frac{1}{(1 - (1/3^s))} \frac{1}{(1 - (1/5^s))} \frac{1}{(1 - (1/7^s))} \dots$$

Essa identidade é chamada de fórmula do produto para a função zeta. Tomando $s = 1$, Euler provou que há infinitos números primos (Burton, 2012).

Em 1859, Riemman publicou um artigo que considera s em $\zeta(s)$ como sendo um número complexo (Burton, 2012). Esse artigo também contém a famosa Hipótese de Riemann, um dos problemas (ainda) abertos da Matemática Pura mais importantes de todos os tempos. Essa hipótese relaciona as raízes complexas da função $\zeta(s)$ com a possibilidade de se calcular $\Pi(n)$ a partir de $\zeta(s)$.

Um dos sonhos dos matemáticos é encontrar uma fórmula que preveja corretamente a sequência de todos os números primos. Inspirado por esse objetivo, Euler propôs uma função polinomial $P(n)$ capaz de gerar muitos números primos. Essa função é dada por:

$$P(n) = n^2 - n + 41$$

Essa fórmula gera primos para $n = 1, \dots, 40$, falha para $n = 41$ e 42 , mas continua gerando alguns primos para $n > 42$.

Existem muitos outros polinômios geradores de primos; no entanto, nenhum deles gera apenas números primos. Há, por exemplo, os primos de Mersenne que são obtidos de (Burton, 2012):

$$M(n) = 2^n - 1$$

Por exemplo, $M(2) = 3$, $M(3) = 7$ e $M(5) = 31$ são primos, mas $M(4) = 15$ não é. Os maiores números primos conhecidos são primos de Mersenne. Até meados de 2023, o maior primo foi encontrado em 2018 e é dado por $M(82589933)$ (Laroche, 2018).

Por volta de 1650, Fermat conjecturou numa carta para Mersenne que $2^{2^n} + 1$ daria

sempre números primos, mas essa fórmula falha para $n = 5$, como mostrado por Euler (Burton, 2012).

No início da década de 60 do século passado, Ulam representou os números naturais num plano sobre uma espiral. Ele percebeu que os primos tendiam a se agrupar ao longo de linhas retas. Esse curioso padrão espacial é chamado de espiral de Ulam (Stein; Ulam; Wells, 1964).

A lacuna entre números primos, ou seja, o intervalo entre primos consecutivos, é outro assunto que tem recebido atenção ao longo dos anos (Baker; Harman; Pintz, 2001; Erdős, 1935). Nesses estudos, busca-se caracterizar as propriedades estatísticas da distribuição das lacunas entre primos sucessivos, da distribuição das lacunas entre as lacunas, ou seja, das lacunas de segunda ordem, e das lacunas de ordens superiores (Szpiro, 2004; Szpiro, 2007). Em outro estudo (Ford; Green; Konyagin; Tao, 2016), deduziu-se uma fórmula para a lacuna máxima entre dois primos.

As distribuições dos números primos também têm sido estudadas usando grafos. Por exemplo, considere que os nós de um grafo são números primos. De acordo com a conjectura de Goldbach, qualquer número par pode ser escrito como a soma de dois primos. Assuma que dois nós (dois primos) estão ligados se a soma deles representa um número par (Chandra; Dasgupta, 2005). Nesse trabalho, analisou-se o grafo assim gerado, considerando que diversos números pares podem ser representados pela soma de diferentes primos (por exemplo $22=3+19=5+17$). Em outro estudo com grafos, os números naturais são os nós e uma ligação é adicionada entre dois nós se eles possuem um divisor comum que é primo (Corso, 2004).

Nesta tese, analisam-se sequências de d -primos por meio de grafos. Aqui, define-se um d -primo como um número natural que tem exatamente d -divisores naturais. Essa definição é do autor desta tese. Usando essa definição, é evidente que os números primos são o caso particular $d = 2$. A intenção deste estudo baseado em d -primos é explorar como as sequências de lacunas variam com d , para $d \in \{2, 3, \dots, 16\}$. Para isso, convertem-se essas sequências em grafos por meio de algoritmos de visibilidade (Lacasa; Luque; Ballesteros; Luque; Nuno, 2008; Luque; Lacasa; Ballesteros; Luque, 2009). Esses algoritmos transformam um número de uma sequência em um nó de um grafo, de modo que as conexões entre os nós refletem propriedades da sequência analisada. Para os grafos assim gerados, determinam-se a distribuição de grau e o grau médio. Ainda, a variabilidade dessas

sequências é avaliada calculando a entropia informacional (Shannon; Weaver, 1949). A densidade de d -primos, dada por $\Pi_d(n)/n$, também é computada e uma função modelo é ajustada. Nesta tese, outras sequências numéricas são também investigadas segundo esse mesmo método; ou seja, via cálculo de entropia informacional e análise de grafos obtidos por algoritmos de visibilidade.

Considere a *sequência de divisores* $\{x, D(x)\}$, em que $D(x) = d$ é o número de divisores de x , sendo $D(x)$ a *função quantidade de divisores* (Apostol, 1988; Broidó; Clauset, 2019; Burton, 2012; Dickson, 2005; Hardy; Wright, 2008; Ore, 1988). Para $\{x\} = \{n\}$, sendo n um número natural, a *sequência dos divisores dos números naturais* $\{n, D(n)\}_{n \in \mathbb{N}}$ começa assim:

$$\{(1, 1), (2, 2), (3, 2), (4, 3), (5, 2), (6, 4), (7, 2), \dots\} \quad (1.1)$$

Por exemplo, $D(6) = 4$, pois 6 possui quatro divisores: os números 6, 3, 2 e 1; portanto, 6 é um 4-primo. A lista $\{D(n)\}_{n \in \mathbb{N}}$ tem o código A000005 na “On-Line Encyclopedia of Integer Sequences” (OEIS) (OEIS, 2023).

Qualquer número inteiro positivo x pode ser escrito como $x = \sum_{p=0}^{\infty} \delta_p 10^p$, com $\delta_p \in \{0, 1, 2, \dots, 9\}$. Agora, troque esse número pela soma do quadrado de seus dígitos; isto é, por $\sum_{p=0}^{\infty} (\delta_p)^2$. Repita esse processo. Se o número 1 for atingido em alguma iteração, então x é chamado de *número feliz* na Teoria de Números (El-sedy; Siksek, 2000; Guy, 2004). Por exemplo, 23 é feliz, pois $2^2 + 3^2 = 13$, $1^2 + 3^2 = 10$ e $1^2 + 0^2 = 1$. Seja um número feliz denotado por h e o conjunto dos números felizes por \mathbb{H} . Na base de dados OEIS, o código da lista $\{h\}_{h \in \mathbb{H}}$ é A007770 (OEIS, 2023). Para $\{x\} = \{h\}$, o início da sequência de divisores dos números felizes $\{h, D(h)\}_{h \in \mathbb{H}}$ é:

$$\{(1, 1), (7, 2), (10, 4), (13, 2), (19, 2), (23, 2), (28, 6), \dots\} \quad (1.2)$$

Considere também a *sequência de trajetórias* $\{x, T(x)\}$. Para $x > 1$, cada $(x, T(x))$ é obtido de $D(d_i) = d_{i+1}$ com $i = 0, 1, 2, \dots, T$, em que $d_0 = x$ e $d_T = 2$. Ou seja, parte-se de x e termina-se em 2 após T aplicações sucessivas da função D . Perceba que $D(1) = 1$ e $D(2) = 2$, então 1 e 2 são pontos fixos do mapa $D(d_i) = d_{i+1}$ (pois o número 1 tem um divisor e o número 2 tem dois divisores); entretanto, o ponto fixo 2 é um atrator global para $x > 1$. O valor de T especifica o número de passos necessários para se atingir esse

ponto fixo a partir de qualquer $x > 1$. Por exemplo, tome $x = 60$, a fim de ilustrar como a sequência de trajetórias é formada. Observe que $D(60) = 12$, $D(12) = 6$, $D(6) = 4$, $D(4) = 3$, e $D(3) = 2$. Como o ponto fixo 2 foi atingido após 5 iterações a partir de $x = 60$, então $T(60) = 5$. Para os números naturais $\{n\}$, o código da lista $\{T(n)\}_{n \in \mathbb{N}}$ na base OEIS é A036459 (OEIS, 2023). A *sequência de trajetórias dos números naturais* $\{n, T(n)\}_{n \in \mathbb{N}}$ é dada por:

$$\{(1, 0), (2, 0), (3, 1), (4, 2), (5, 1), (6, 3), (7, 1), \dots\} \quad (1.3)$$

A *sequência de trajetórias dos números felizes* $\{h, T(h)\}_{h \in \mathbb{H}}$ é assim escrita (essa sequência ainda não está listada na base OEIS):

$$\{(1, 0), (7, 1), (10, 3), (13, 1), (19, 1), (23, 1), (28, 4), \dots\} \quad (1.4)$$

E, por fim, a *sequência dos números primos*, ou *2-primos*, $\{n, p_2(n)\}_{n \in \mathbb{N}}$ que possui o código A000040 na OEIS (OEIS, 2023):

$$\{(1, 2), (2, 3), (3, 5), (4, 7), (5, 11), (6, 13), (7, 17), \dots\} \quad (1.5)$$

Essa é a sequência que inspirou a proposta de d -primo, que é o primeiro tema de pesquisa desta tese. É difícil comparar o trabalho aqui feito sobre esse tema com outros trabalhos, por causa da novidade do conceito de d -primo (nós não encontramos essa ideia na literatura). Além disso, não foram encontrados trabalhos em que se aplicam algoritmos de visibilidade para analisar a sequência de primos.

A escolha dessas sequências deveu-se à oportunidade de contribuição científica, pois algumas sequências não foram estudadas em trabalhos anteriores, tampouco possuem cadastro na OEIS (OEIS, 2023). A sequência dos números primos, apesar de muito investigada, carecia de estudos que a relacionassem com sequências de números com mais de dois divisores. A proposta foi analisar essas sequências de natureza puramente matemática por meio de grafos.

Utilizou-se o *software Mathematica 12* para a execução de todos os cálculos ¹.

O restante deste texto está assim organizado. No capítulo 2, apresentam-se definições

¹Os arquivos-fonte e os resultados obtidos nesta tese podem ser encontrados na internet no endereço: <https://github.com/blmayer/thesis>.

básicas e os algoritmos utilizados nos experimentos numéricos com d -primos, divisores de números naturais e divisores de números felizes. No capítulo 3, descrevem-se os estudos numéricos realizados sobre as lacunas entre d -primos sucessivos. No capítulo 4, discorre-se sobre o estudo da densidade dos d -primos. No capítulo 5, trata-se dos estudos realizados sobre a quantidade de divisores dos números naturais e felizes. No capítulo 6, listam-se as conclusões e possíveis continuações deste trabalho.

A maioria dos resultados aqui apresentados foram publicados em (Mayer; Monteiro, 2020) e (Mayer; Monteiro, 2023). A primeira publicação relata os resultados obtidos a partir das lacunas dos d -primos para $d \in \{2, 3, \dots, 11\}$. Nesta tese, as análises foram feitas para $d \in \{2, 3, \dots, 16\}$. A segunda publicação relata os resultados obtidos considerando os divisores dos números naturais e dos números felizes.

As simulações foram feitas num computador pessoal com processador com 4 núcleos e *clock* de 3Gz e memória RAM de 8GB. Além disso, foi usado *swap* com adicionais 8GB. A principal limitação computacional foi a memória (alguns dos grafos analisados contêm mais de 1 milhão de arestas).

Capítulo 2

DEFINIÇÕES E MÉTODOS

Neste capítulo, são apresentados os conceitos, as fórmulas e os algoritmos necessários para os estudos numéricos realizados neste trabalho. Alguns desses conceitos já foram mencionados no capítulo anterior.

2.1 Definições

2.1.1 Sequência

Ao longo deste trabalho, o termo sequência aparece diversas vezes, sendo uma definição necessária para um melhor entendimento desse conceito.

Definição 2.1: Uma sequência numérica é uma função $\{x(n)\} : \mathbb{N} \rightarrow \mathbb{R}$ (Lima, 1976) que atribui um número real x a cada posição de uma lista, sendo essa posição identificada por um número natural n . Assim, $x(n)$ corresponde ao n -ésimo valor de x da sequência $\{x(1), x(2), x(3), \dots\}$. Note que esse conjunto possui ordem definida.

2.1.2 Número primo

A definição de número primo é apresentada em diversos livros de matemática básica.

Definição 2.2: Um número primo é um número natural maior do que 1 e que possui nenhum outro divisor além de 1 e de si mesmo (Knuth, 1997).

O número 1 não é considerado primo, a fim de não se violar o Teorema Fundamental da Aritmética. Se 1 fosse primo, os números compostos não teriam uma única representação.

Por exemplo, $12 = 2^2 \times 3 = 1 \times 2^2 \times 3 = 1^2 \times 2^2 \times 3$. Como multiplicar por 1 não altera o resultado do produto, então não se teria uma única representação de um número composto em termos de fatores primos. Para evitar esse problema, diz-se que 1 não é primo.

2.1.3 Teorema Fundamental da Aritmética

O Teorema Fundamental da Aritmética fornece o alicerce central para a construção dos números naturais e para a fatoração prima.

Teorema 2.1: Todo natural $n > 1$ ou é primo ou é um produto único de números primos. Assim, n possui uma representação única, desconsiderando a ordem dos fatores, e pode ser escrito como $n = p_1^a \times p_2^b \times p_3^c, \dots$ sendo p_1, p_2, p_3, \dots seus fatores primos e a, b, c, \dots as respectivas multiplicidades desses fatores (Burton, 2012).

A noção apresentada a seguir está intimamente relacionada com o teorema exposto acima. Apesar de simples e natural, a ideia de d -primo não foi encontrada na literatura.

2.1.4 Número d -primo

Assim como um número primo possui exatamente 2 divisores, pode-se estabelecer que um número chamado de 3-primo possui exatamente 3 divisores. Por exemplo, 4 é 3-primo, pois seus divisores são 1, 2 e 4. Generalizando, um d -primo possui exatamente d divisores naturais.

Definição 2.3: Um d -primo é um número natural que possui exatamente d divisores naturais.

Note que, pela definição acima, os 2-primos correspondem aos números primos usuais. A sequência de 2-primos começa com 2, 3, 5, 7, 11, 13, 17, 19, \dots , pois esses números possuem apenas dois divisores. Note 1 é 1-primo, segundo a definição de d -primo. Portanto, a definição aqui proposta de d -primo nada mais é do que uma generalização baseada na quantidade de divisores de um número natural.

Aqui, usa-se a notação $p_d(n)$ para designar o n -ésimo d -primo, com $n \in \mathbb{N}^*$. Dessa forma, $p_3(4)$ é o quarto 3-primo, que é o número 49. A Tabela 2.1 apresenta os primeiros d -primos para $d = \{2, 3, \dots, 16\}$.

Tabela 2.1: Sequências de d -primos denotados por $p_d(n)$.

| d | $p_d(n)$ |
|-----|---|
| 2 | 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, ... |
| 3 | 4, 9, 25, 49, 121, 169, 289, 361, 529, 841, ... |
| 4 | 6, 8, 10, 14, 15, 21, 22, 26, 27, 33, 34, 35, ... |
| 5 | 16, 81, 625, 2401, 14641, 28561, 83521, ... |
| 6 | 12, 18, 20, 28, 32, 44, 45, 50, 52, 63, 68, 75, ... |
| 7 | 64, 729, 15625, 117649, 1771561, 4826809, ... |
| 8 | 24, 30, 40, 42, 54, 56, 66, 70, 78, 88, 102, ... |
| 9 | 36, 100, 196, 225, 256, 441, 484, 676, 1089, ... |
| 10 | 48, 80, 112, 162, 176, 208, 272, 304, 368, ... |
| 11 | 1024, 59049, 9765625, 282475249, ... |
| 12 | 60, 72, 84, 90, 96, 108, 126, 132, 140, 150, ... |
| 13 | 4096, 531441, 244140625, 13841287201, ... |
| 14 | 192, 320, 448, 704, 832, 1088, 1216, 1458, ... |
| 15 | 16384, 4782969, 6103515625, 678223072849, ... |
| 16 | 120, 168, 210, 216, 264, 270, 280, 312, 330, ... |

As sequências acima estão indexadas na OEIS para valores de d até 101 (OEIS, 2023). A sequência com $d = 8$ possui uma propriedade única: os chamados números esfênicos (que são números dados pelo produto de três primos distintos) são uma subsequência dessa sequência, isto é, todo número esfênico é 8-primo.

O maior número utilizado nesta tese é o $p_{13}(10000)$ cujo valor é:

1741018879299188426700762715443778414040036179271771734964641

Esse número possui 61 dígitos e 61 é, curiosamente, um número 2-primo.

Contagem de d -primos

Assim como o conceito de d -primo é uma extensão do conceito de primo, a função contagem de d -primos $\Pi_d(n)$ é uma extensão da função contagem de primos $\Pi(n)$.

Definição 2.4: A função contagem de d -primos, representada por $\Pi_d(n)$, com $n \in \mathbb{N}$, é definida como o número de d -primos menores ou iguais a n .

Portanto, $\Pi_2(n)$ equivale à usual função contagem de primos. Por exemplo, $\Pi_2(10) = 4$, pois só há quatro 2-primos menores ou iguais a 10 (os números 2, 3, 5 e 7). Já

$\Pi_3(10) = 2$, pois apenas 4 e 9 são 3-primos menores ou iguais a 10.

2.1.5 Número feliz

Definição 2.5: Um inteiro positivo pode ser escrito como $x = \sum_{p=0}^{\infty} \delta_p 10^p$, com $\delta_p \in \{0, 1, 2, \dots, 9\}$, sendo δ_p os dígitos que compõem o inteiro x . Seja s a soma dos quadrados desses dígitos, isto é, $F(x) = \sum_{p=0}^{\infty} (\delta_p)^2 = s$, sendo F a função que calcula a soma dos quadrados de δ_p . Assuma que $y(i+1) = F(y(i))$, com $y(0) = s$ (ou seja, o primeiro termo dessa relação de recorrência vale s) e $F(y)$ é a função que representa a soma dos quadrados dos dígitos que compõem o número y . Um número é feliz se obtém-se $y = 1$ para algum valor de $i = 0, 1, 2, 3, \dots$ (El-sedy; Siksek, 2000; Guy, 2004).

Aqui, um número feliz é denotado por h e o conjunto dos números felizes por \mathbb{H} . Um número que não é feliz é um número triste. Os números tristes acabam convergindo para a sequência $\{4, 16, 37, 58, 89, 145, 42, 20\}$, que se repete indefinidamente. Por exemplo, 11 é infeliz pois $1^2 + 1^2 = 2$, $2^2 = 4$ e, partir daí, repete-se a sequência mencionada.

Permutar os dígitos de um número não altera sua felicidade. Portanto, como 203 é feliz, então 23, 32, 230, 302 e 320 também são. Incluir zeros também não altera a felicidade. Assim, 2003, 20003 e 200003 são felizes também.

2.1.6 Função quantidade de divisores

Definição 2.6: A função quantidade de divisores $D(n) : \mathbb{N} \rightarrow \mathbb{N}$ é tal que $D(n) = d$, sendo d o número de divisores do número natural n . Na literatura, essa função possui diferentes nomes e notações (Apostol, 1988; Broido; Clauset, 2019; Burton, 2012; Dickson, 2005; Hardy; Wright, 2008; Ore, 1988).

Por exemplo, Hardy e Wright (2008) definem a função quantidade de divisores $D(n)$ como:

$$D(n) = \sum_{x|n} 1 = d$$

sendo que $x|n$ significa que x divide n ; ou seja, a somatória é no conjunto dos divisores de n . Para cada divisor de n se adiciona 1 e o resultado dessa soma vale d . Alguns exemplos dos valores de $D(n)$ são $D(2) = 2$ e $D(4) = 3$.

2.1.7 Entropia informacional

A variabilidade de uma sequência $\{x(n)\}$ pode ser avaliada calculando a entropia H proposta por Shannon. Quanto maior a variabilidade dos valores de x , maior H .

Definição 2.7: A entropia informacional H é definida como $H = -\sum_{i=1}^q \rho_i \ln \rho_i$, sendo ρ_i a frequência relativa com que aparece o i -ésimo valor de x na sequência $\{x(n)\}$, sendo q o número de valores de x que são distintos. O máximo de H vale $H_{max} = \ln q$, que é obtido no caso em que $\rho_i = 1/q$ para $i = 1, \dots, q$; ou seja, a entropia máxima corresponde ao caso em que todos os valores distintos de x aparecem nessa sequência com mesma frequência relativa (Shannon; Weaver, 1949).

Por exemplo, a sequência $\{1, 1, 1, 1, 1\}$ possui $H = 0$, pois $\rho = 1$ para um único número; já a sequência $\{1, 5, 8, 17, 18\}$ possui entropia máxima, pois os valores que a compõem não se repetem. Nesse caso $\rho_i = 1/5$ para $i = 1, 2, 3, 4, 5$, de modo que $H = \ln(5) \simeq 1,61$. A sequência $\{1, 1, 1, 5, 18\}$ possui alguns valores repetidos, de modo que $\rho_1 = 3/5$, $\rho_2 = 1/5$ e $\rho_3 = 1/5$; portanto, $H \simeq 0,67$, um valor que está obviamente entre o mínimo e o máximo.

Essa entropia tem sido utilizada em diversos trabalhos envolvendo, por exemplo, a dinâmica de sistemas biológicos e sociais (Piqueira; Serboncini; Monteiro, 2006; Netto; Meirelles; Ribeiro, 2017). Ela ainda pode ser empregada como uma medida de complexidade. Por exemplo, a complexidade C_{SDL} (Shiner; Davison; Landsberg, 1999) de uma sequência é calculada por $C_{SDL} = \Delta(1 - \Delta)$, sendo $\Delta = H/H_{max}$. Portanto, C_{SDL} é máxima para $\Delta = 1/2$. Ou seja, a complexidade máxima ocorre entre a entropia mínima $H = 0$ ($\Delta = 0$) e a entropia máxima H_{max} ($\Delta = 1$).

2.1.8 Lei de potência

Definição 2.8: Uma lei de potência é uma função $f(x)$ que se relaciona com sua variável x por meio da fórmula $f(x) = Ax^{-\gamma}$, com A e γ constantes (Newman, 2005).

Uma lei de potência é livre-de-escala (*scale-free*), pois trocando x por Bx , sendo B uma constante, a relação entre $f(x)$ e x não muda. De fato, $f(Bx) = A(Bx)^{-\gamma} = A'x^{-\gamma}$, com $A' = AB^{-\gamma}$. Portanto, a mudança de escala em x altera apenas a constante multiplicativa (de A para A'), sem alterar a dependência com x (que, no caso, é do tipo $x^{-\gamma}$).

Num gráfico log-log, uma lei de potência corresponde a uma equação de reta, pois $\log f(x) = \log A - \gamma \log x$. Assim, $\log f(x)$ varia linearmente com $\log x$, de modo que o

coeficiente angular dessa reta vale $-\gamma$.

Nesta tese, usa-se uma lei de potência para ajustar a distribuição de graus dos grafos obtidos por meio dos algoritmos de visibilidade descritos a seguir.

2.2 Métodos

Aqui, são apresentados dois algoritmos que convertem séries temporais em grafos não direcionados (Lacasa; Luque; Ballesteros; Luque; Nuno, 2008; Luque; Lacasa; Ballesteros; Luque, 2009). Esses algoritmos têm sido usados, por exemplo, em análises de ações de bolsa de valores (Chen; Liu; Zhang; Zhang, 2019) e de eletroencefalogramas (Wang; Long; Arends; Aarts, 2017). Eles são chamados de algoritmo de visibilidade natural e algoritmo de visibilidade horizontal. Ambos geram grafos cujas propriedades topológicas refletem características da séries utilizadas. Por isso, é interessante analisar os grafos gerados por esses algoritmos.

2.2.1 Visibilidade natural

Seja a sequência (ou a série temporal) $\{x(1), x(2), x(3), \dots\}$. Assuma que $a < i < b$. No grafo de visibilidade natural (VN) (Lacasa; Luque; Ballesteros; Luque; Nuno, 2008), obtido a partir da sequência $\{x(n)\}$, com $n = 1, 2, 3, \dots$, os nós $x(a)$ e $x(b)$ estão conectados se todos os pontos intermediários $(i, x(i))$ estão abaixo da reta que liga os pontos $(a, x(a))$ e $(b, x(b))$.

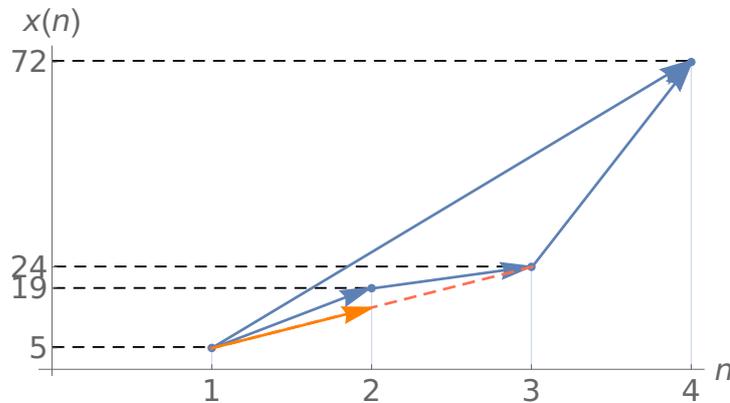


Figura 2.1: A sequência $\{x(n)\}$ em função de n , com $x(1) = 5$, $x(2) = 19$, $x(3) = 24$ e $x(4) = 72$. As setas em azul mostram as conexões permitidas. A seta em laranja indica uma conexão proibida.

Por exemplo, suponha que os quatro primeiros pontos de uma sequência $\{x(n)\}$ sejam $x(1) = 5$, $x(2) = 19$, $x(3) = 24$ e $x(4) = 72$, como ilustra a Figura 2.1. No grafo de VN correspondente a essa sequência, existe uma conexão entre os nós 19 e 72, pois o único ponto intermediário $x(3) = 24$ está abaixo da reta que liga $x(2) = 19$ e $x(4) = 72$. Já os pontos $x(1) = 5$ e $x(3) = 24$ não estão conectados, pois $x(2) = 19$ está acima da reta que passa por esses dois pontos.

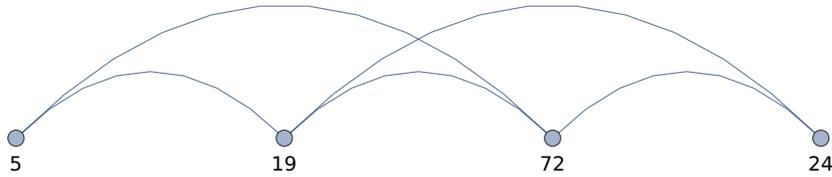


Figura 2.2: Grafo de VN gerado a partir da sequência mostrada na Figura 2.1.

Método 2.1: Sejam a , b e i índices de uma sequência $\{x(n)\}$, com $a < i < b$. No grafo de visibilidade natural (VN), $x(a)$ e $x(b)$ estão conectados se todos os pontos intermediários $(i, x(i))$ satisfazem a desigualdade (Lacasa; Luque; Ballesteros; Luque; Nuno, 2008):

$$x(i) < x(a) + (x(b) - x(a)) \left(\frac{i - a}{b - a} \right) \quad (2.1)$$

Em outras palavras, os nós $x(a)$ e $x(b)$ estão ligados se existe uma reta passando por (a, x_a) e (b, x_b) no gráfico $x(n) \times n$, tal que nenhum ponto intermediário está acima dessa reta ou coincide com ela. No exemplo dado da Figura 2.1, como $x(3) = 24 < x(2) + (x(4) - x(2))(3 - 2)/(4 - 2) = 19 + (72 - 19)/2 = 45,5$, então os nós 19 e 72 estão conectados. Por outro lado, como $x(2) = 19 > x(1) + (x(3) - x(1))(2 - 1)/(3 - 1) = 5 + (24 - 5)/2 = 14,5$, então os nós 5 e 24 não estão conectados. Nesse caso, o ponto $(2, 19)$ é alto o suficiente para impedir que os pontos $(1, 5)$ e $(3, 24)$ se “vejam”.

A Figura 2.2 mostra o grafo gerado usando a Equação (2.1) para a sequência correspondente à Figura 2.1.

2.2.2 Visibilidade horizontal

No grafo de visibilidade horizontal (VH) (Luque; Lacasa; Ballesteros; Luque, 2009), os nós $x(a)$ e $x(b)$ estão conectados se todos os pontos intermediários $(i, x(i))$ no gráfico $x(n) \times n$ estão abaixo da linha horizontal que une $(a, x(a))$ e $(b, x(b))$.

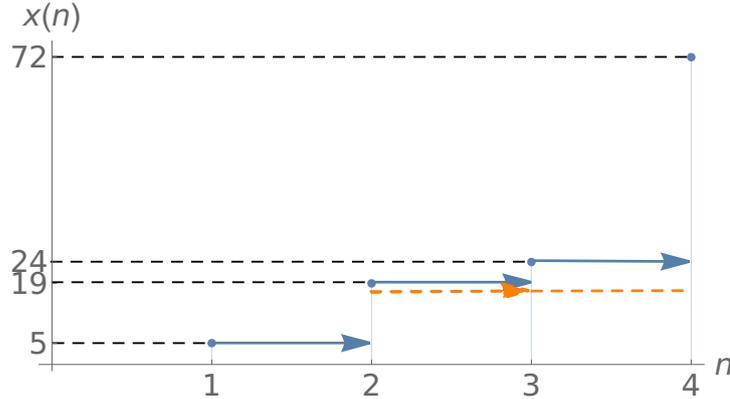


Figura 2.3: A sequência $\{x(n)\}$ em função de n , com $x(1) = 5$, $x(2) = 19$, $x(3) = 24$ e $x(4) = 72$. As setas em azul mostram as conexões permitidas. A seta em laranja indica uma conexão proibida.

Tomando a mesma sequência usada para exemplificar a VN, os nós $x = 19$ e $x = 72$ não estão conectados, pois $x(2) = 19 < x(3) = 24$; isto é, $(2, 19)$ e $(4, 72)$ não se veem se olharem apenas na direção horizontal, pois o ponto intermediário $(3, 24)$ é suficientemente alto para bloquear a visibilidade horizontal. Esse exemplo ilustra a diferença entre as duas visibilidades, pois pontos que estão conectados no grafo de VN podem não estar conectados no grafo de VH. Na verdade, o grafo de VH é um subgrafo do grafo de VN (Luque; Lacasa; Ballesteros; Luque, 2009).



Figura 2.4: Grafo de VH gerado a partir da sequência mostrada na Figura 2.3.

Método 2.2: Sejam a , b e i índices de uma sequência $\{x(n)\}$, com $a < i < b$. No grafo de visibilidade horizontal (VH), obtido a partir da sequência $\{x(n)\}$, com $n = 1, 2, 3, \dots$, os nós $x(a)$ e $x(b)$ estão conectados se todos os pontos intermediários $(i, x(i))$ satisfazem

a desigualdade (Luque; Lacasa; Ballesteros; Luque, 2009):

$$\{x_a, x_b\} > x_i \quad (2.2)$$

Assim, dois nós estão conectados se nenhum ponto intermediário está acima deles.

A Figura 2.4 mostra o grafo gerado usando a Equação (2.2) para a sequência correspondente à Figura 2.3.

2.3 Algumas medidas em grafos

A estrutura topológica dos grafos construídos por meio dos algoritmos de visibilidade pode ser caracterizada pelas medidas citadas a seguir (Monteiro, 2014).

A *distribuição de graus* $P(k)$ expressa a fração de nós com grau k . A forma com que $P(k)$ varia com k diferencia os tipos de grafos. Por exemplo, para um grafo aleatório, $P(k)$ obedece a uma distribuição de Poisson; para redes livre-de-escala, $P(k)$ segue uma lei de potência (Monteiro, 2014). O *grau médio* de um grafo, denotado por $\langle k \rangle$, é calculado a partir de:

$$\langle k \rangle = \frac{\sum_{i=1}^N k_i}{N} = \frac{\sum_{k=k_{min}}^{k_{max}} kP(k)}{N} = \frac{2M}{N} \quad (2.3)$$

sendo N o número de nós, M o número de arestas e k_{min} e k_{max} os graus mínimo e máximo encontrados no grafo.

O *menor caminho médio*, denotado por $\langle \ell \rangle$, é a média dos caminhos mais curtos entre todos os pares de nós do grafo. Para grafos não direcionados, seu valor é obtido de:

$$\langle \ell \rangle = \frac{2 \sum_{i=1}^{N-1} \sum_{j=i+1}^N \ell_{ij}}{N(N-1)} \quad (2.4)$$

sendo ℓ_{ij} o caminho mais curto (o menor número de arestas) entre os nós i e j .

Para o nó i , o *coeficiente de agregação* (*clustering coefficient*) c_i é definido como:

$$c_i = \frac{2b_i}{k_i(k_i - 1)} \quad (2.5)$$

sendo b_i o número de arestas que existem entre seus k_i vizinhos, para $k_i > 1$. Se $k_i = 1$,

então $c_i = 0$. O *coeficiente de agregação médio* $\langle c \rangle$ é obtido de:

$$\langle c \rangle = \frac{\sum_{i=1}^N c_i}{N} \quad (2.6)$$

2.4 Relação entre d -primos e d primo

Hardy e Wright (2008) dão um resultado que aumenta a eficiência dos geradores de números d -primos para o caso em que d é um 2-primo. Verifica-se que:

$$p_d(n) = p_2(n)^{d-1} \quad (2.7)$$

Esse resultado decorre diretamente da fatoração em produtos de primos; no entanto, Hardy e Wright (2008) e diversos outros autores não usam essa propriedade para classificar os números compostos.

Neste trabalho, a equação (2.7) foi utilizada para a geração de números com quantidade prima de divisores. Como é conhecida a sequência de 2-primos (2,3,5,7,11,...), esse resultado facilita gerar a sequência de d -primos. Por exemplo, a sequência de 3-primos pode ser calculada usando 2-primos assim:

$$2^2, 3^2, 5^2, 7^2, 11^2, 13^2, \dots$$

A sequência de 5-primos pode ser calculada de forma análoga:

$$2^4, 3^4, 5^4, 7^4, 11^4, 13^4, \dots$$

E assim por diante. Entretanto, nenhuma relação foi encontrada entre $p_2(n)$ e $p_j(n)$ com $j = \{4, 6, 8, 9, 10\}$; isto é, nenhuma relação foi encontrada entre números primos e d -primos quando d não é um número primo. Por isso, a geração dos outros d -primos foi feita por força bruta. Esse método não é eficiente, e juntamente com o poder computacional do nosso “hardware”, limita os estudos numéricos do presente trabalho.

No próximo capítulo, apresentam-se as análises numéricas realizadas com base nos métodos apresentados neste capítulo.

Capítulo 3

LACUNAS ENTRE d -PRIMOS SUCESSIVOS

Neste estudo, analisam-se as lacunas entre d -primos consecutivos. Para isso, inicialmente calcula-se a diferença de sucessivos d -primos para $d \in \{2, 3, \dots, 16\}$. Com isso, para cada valor de d , obtém-se uma nova sequência, que pode ser considerada como uma série temporal discreta. A ideia de analisar as lacunas, e não os número d -primos em si, vem do fato de que as lacunas não geram sequências necessariamente crescentes, enquanto que os d -primos obviamente apenas aumentam de valor, conforme se caminha pelas sequências correspondentes.

Para determinar a variabilidade de cada uma dessas sequências de lacunas, foi calculada a entropia informacional (Shannon; Weaver, 1949). Seu valor normalizado, denotado por Δ , é dado por:

$$\Delta = \frac{H}{H_{max}} \tag{3.1}$$

O valor de Δ indica o nível de repetição de valores das sequências consideradas.

Em seguida, essas sequências são transformadas em grafos não direcionados por meio da aplicação dos algoritmos de visibilidade mencionados (Lacasa; Luque; Ballesteros; Luque; Nuno, 2008; Luque; Lacasa; Ballesteros; Luque, 2009). Então, a distribuição de graus $P(k)$ e o grau médio $\langle k \rangle$ desses grafos são determinados. Lembre que o grau k de um nó é o número de arestas conectadas a esse nó e que a distribuição de graus expressa como a porcentagem $P(k)$ de nós com grau k varia com k (Monteiro, 2014). Usualmente, $P(k)$

é interpretado como a probabilidade de se encontrar um nó com grau k . Para os grafos obtidos, assume-se que $P(k)$ obedece a uma lei de potência, ou seja, $P(k) = Ak^{-\gamma}$, com A e γ constantes. Assim, assume-se que a distribuição de graus dos grafos é livre-de-escala.

Aqui, define-se:

$$x_d(n) = p_d(n+1) - p_d(n) \tag{3.2}$$

como a lacuna entre dois d -primos consecutivos, com $n = 1, 2, 3, \dots$. Por exemplo, $x_5(1) = 65$, pois $p_5(2) = 81$ e $p_5(1) = 16$.

A Tabela 3.1 mostra os primeiros valores de $x_d(n)$ para $d \in \{2, 3, \dots, 16\}$. Cada nó dos grafos de VN e VH representa um valor distinto de x_d .

Tabela 3.1: Sequências das diferenças de primeira ordem de d -primos consecutivos, dados por $x_d(n) = p_d(n+1) - p_d(n)$ com $n \in \mathbb{N}^*$, para $d \in \{2, 3, \dots, 16\}$.

| d | $x_d(n)$ |
|-----|--|
| 2 | 1, 2, 2, 4, 2, 4, 2, 4, 6, 2, 6, 4, 2, 4, 6, 6, 2, 6, 4, ... |
| 3 | 5, 16, 24, 72, 48, 120, 72, 168, 312, 120, 408, ... |
| 4 | 2, 2, 4, 1, 6, 1, 4, 1, 6, 1, 1, 3, 1, 7, 5, 4, 2, 1, 4, ... |
| 5 | 65, 544, 1776, 12240, 13920, 54960, 46800, ... |
| 6 | 6, 2, 8, 4, 12, 1, 5, 2, 11, 5, 7, 1, 16, 6, 1, 17, 1, ... |
| 7 | 665, 14896, 102024, 1653912, 3055248, ... |
| 8 | 6, 10, 2, 12, 2, 10, 4, 8, 10, 14, 2, 1, 5, 4, 14, 2, ... |
| 9 | 64, 96, 29, 31, 185, 43, 192, 413, 67, 69, 219, ... |
| 10 | 32, 32, 50, 14, 32, 64, 32, 64, 37, 59, 32, 16, ... |
| 11 | 58025, 9706576, 272709624, 25654949352, ... |
| 12 | 12, 12, 6, 6, 12, 18, 6, 8, 10, 6, 4, 38, 2, ... |
| 13 | 527345, 243609184, 13597146576, ... |
| 14 | 128, 128, 256, 128, 256, 128, 242, 14, 384, ... |
| 15 | 180, 76, 384, 1152, 89, 475, 204, 1265, 655, ... |
| 16 | 48, 42, 6, 48, 6, 10, 32, 18, 48, 6, 6, 18, 32, ... |

Os experimentos numéricos foram feitos tomando-se $n = 1, 2, \dots, 10000$, isto é, os

primeiros 10000 d -primos para cada valor de d . Com esses números, 9999 diferenças $x_d(n)$ foram calculadas. Os valores correspondentes de $n = 1, 2, 9998, 9999$ foram descartados, a fim de desprezar o transiente inicial ($n = 1, 2$) e efeitos do truncamento da sequência ($n = 9998, 9999$). O tamanho máximo das sequências aqui analisadas foi afetado pela capacidade de processamento do nosso computador. A geração dessas sequências e dos grafos correspondentes consomem cerca de um mês.

Das sequências acima, apenas as com $d = \{2, 3, 5\}$ estão registradas na OEIS. Elas são denotadas por A001223, A069482 e A276963, respectivamente (OEIS, 2023).

3.1 Resultados

A Tabela 3.2 mostra os valores da entropia normalizada Δ calculada a partir da Equação (3.1). Obtém-se $\Delta = 1$ para $d = 5, 7, 11$ e 13 , pois, para esses valores de d , não há valores repetidos de x_d nas sequências correspondentes. Para $d = 3, d = 9$, e $d = 15$, $\Delta \simeq 1$. Para d par, $\Delta \approx 0.7$. Então, para $d \in \{2, 3, \dots, 16\}$, d ímpar favorece mais a ocorrência de valores equiprováveis de x_d do que d par.

Tabela 3.2: Entropia informacional normalizada Δ de $x_d(n)$ para $d \in \{2, 3, \dots, 16\}$.

| d | Δ |
|-----|----------|
| 2 | 0.713 |
| 3 | 0.997 |
| 4 | 0.653 |
| 5 | 1.000 |
| 6 | 0.797 |
| 7 | 1.000 |
| 8 | 0.696 |
| 9 | 0.999 |
| 10 | 0.700 |
| 11 | 1.000 |
| 12 | 0.765 |
| 13 | 1.000 |
| 14 | 0.577 |
| 15 | 0.999 |
| 16 | 0.766 |

Nota-se valores muito próximos de 1 para $d = 9$ e $d = 15$. Isso se dá pelo fato de existirem 690 e 12 valores repetidos, respectivamente. Por exemplo, para $d = 9$ o primeiro valor repetido é 1085 e ocorre nas posições 33 e 162 da sequência. Esse fato diferencia as

sequências com d ímpar de d primo. Para $d = 15$ os valores repetidos são 12672, 52608, 56448, 557928, 573480 e 7776120 e se repetem apenas uma vez.

A partir das sequências de lacunas $x_d(n)$ para $d \in \{2, 3, \dots, 16\}$, grafos não direcionados de VN e VH foram construídos usando as equações (2.1) e (2.2). As Figuras 3.1, 3.2 e 3.3 exibem as primeiras 100 arestas dos grafos de VN e VH gerados.

As Figuras 3.1, 3.2 e 3.3 não são mostradas aqui apenas a título de curiosidade. Essas figuras lembram os grafos de redes do tipo mundo-pequeno (*small-worlds*) que aparecem no clássico artigo de Watts e Strogatz (1998).

As Figuras (3.4) e (3.5) exibem os gráficos log-log de $P(k)$ para os grafos de VN e VH, respectivamente, com $d \in \{2, 3, \dots, 16\}$. Observe que, nesses gráficos, a maioria das curvas $P(k)$ decai como uma lei de potências; ou seja, $P(k) \approx Ak^{-\gamma}$. Exceções são os gráficos de VN para $d = \{7, 11, 13$ e $15\}$.

A Tabela 3.3 apresenta os valores de A , γ e $\langle k \rangle$ para os 15 grafos construídos utilizando o algoritmo de visibilidade natural. A Tabela 3.4 apresenta os valores de A , γ e $\langle k \rangle$ para os 15 grafos de visibilidade horizontal. Para os ambos os casos, a função $Ak^{-\gamma}$ foi ajustada nos gráficos de $P(k)$ usando o método dos mínimos quadrados (Ljung, 1998).

Essas simulações numéricas mostram que os valores de $\langle k \rangle$ para os grafos de VH são menores do que os valores correspondentes nos grafos de VN. Esse resultado é esperado, pois grafos de VH são subgrafos dos respectivos grafos de VN. Como grafos de VH têm menos arestas, então o grau médio é menor, em comparação com os grafos de VN.

Nos grafos de VN, $\langle k \rangle \approx 6$, com a exceção de $d = \{5, 7, 11, 13, 15\}$, pois esses apresentam valores maiores.

Nos grafos de VH, $\langle k \rangle > 3.90$ para d ímpar e $\langle k \rangle < 3.90$ para d par. Note também que os valores de A e γ apresentam menor variabilidade nos grafos de VH do que nos de VN. Adicionalmente, nos grafos de VN, $\gamma \simeq 1.1$ e para os grafos de VH, $\gamma \simeq 1.7$.

Para nós, é um tanto surpreendente a regularidade dos resultados mostrados na Figura (3.5) e na Tabela 3.4, referente aos grafos de VH. Já os resultados exibidos na Figura (3.4) e na Tabela 3.3, relacionados aos grafos de VN, são menos uniformes.

Parte desses resultados culminaram na publicação (Mayer; Monteiro, 2020). Nessa publicação, porém, foram apresentados os resultados apenas para $d = \{2, 3, \dots, 11\}$.

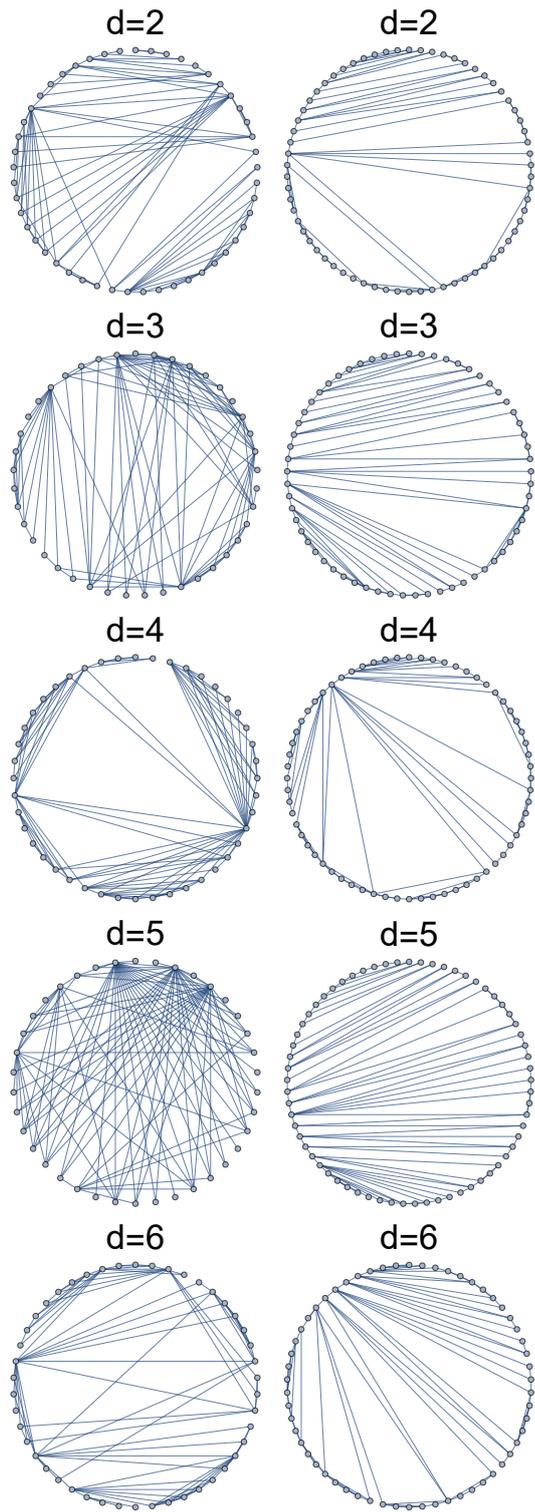


Figura 3.1: As primeiras 100 arestas dos grafos para $d \in \{2, \dots, 6\}$. Esquerda: VN; direita: VH.

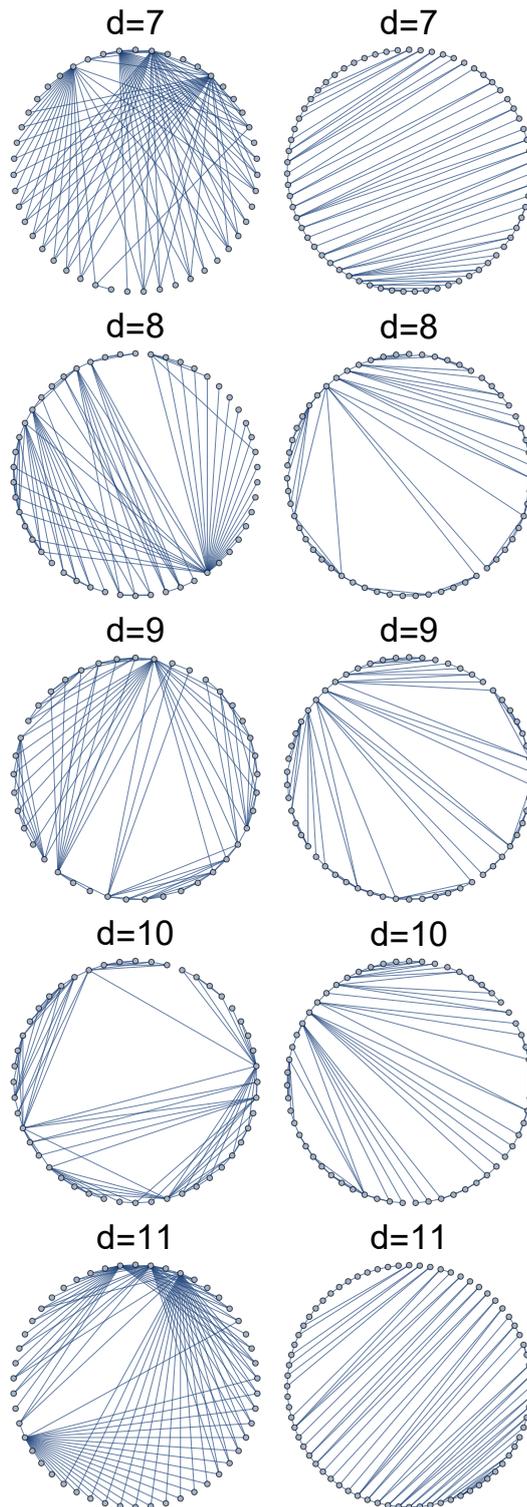


Figura 3.2: As primeiras 100 arestas dos grafos para $d \in \{7, \dots, 11\}$. Esquerda: VN; direita: VH.

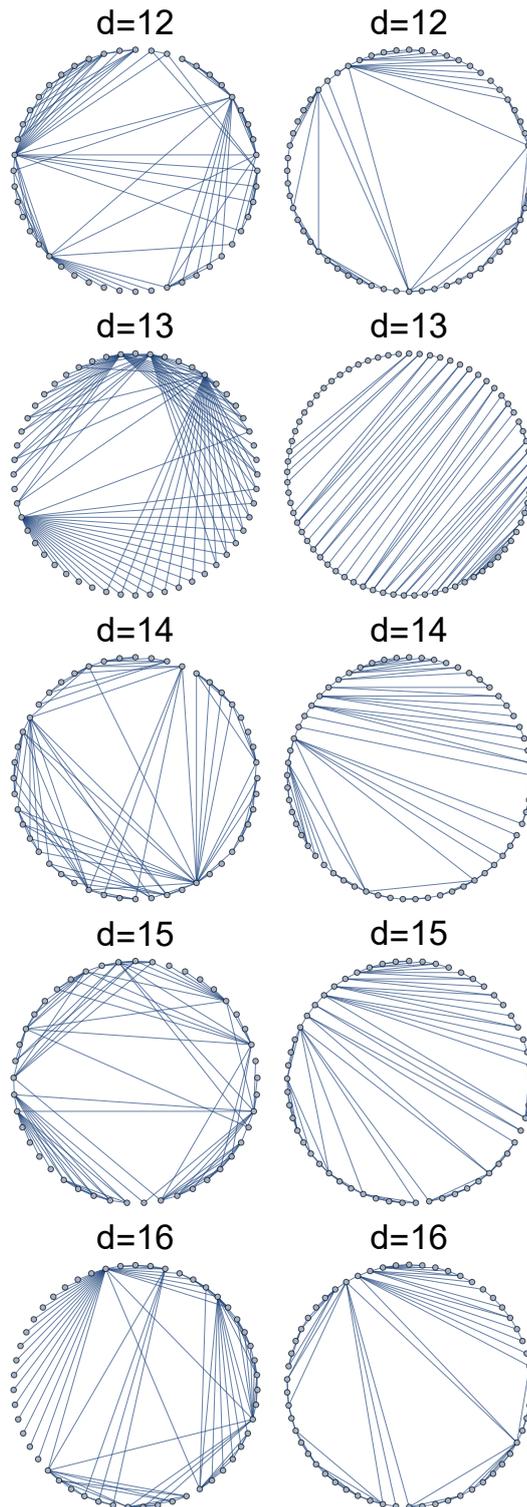


Figura 3.3: As primeiras 100 arestas dos grafos para $d \in \{12, \dots, 16\}$. Esquerda: VN; direita: VH.

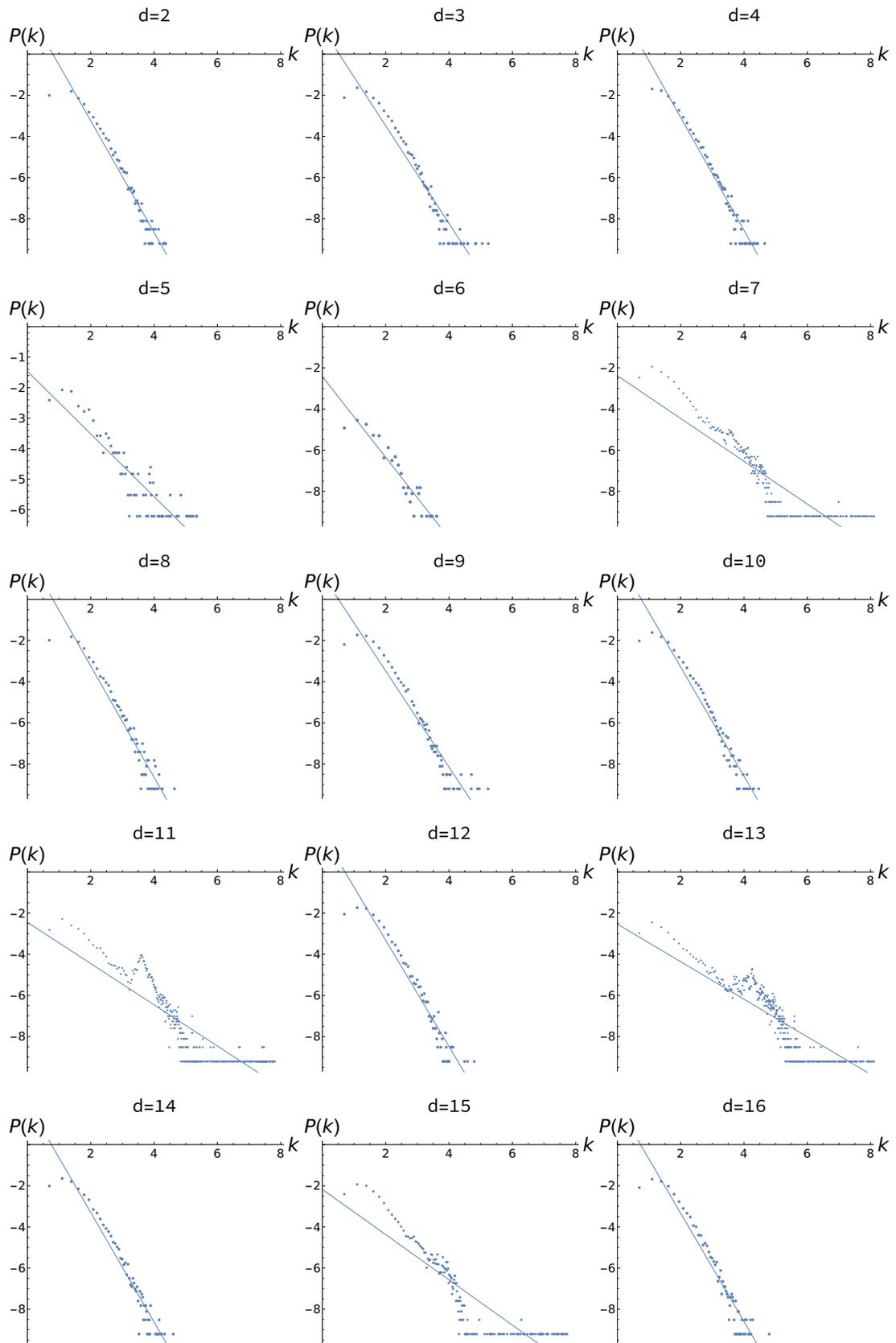


Figura 3.4: A distribuição de graus encontrada nos grafos de VN (pontos) e a função ajustada $Ak^{-\gamma}$ (linha sólida) para $d \in \{2, 3, \dots, 16\}$.

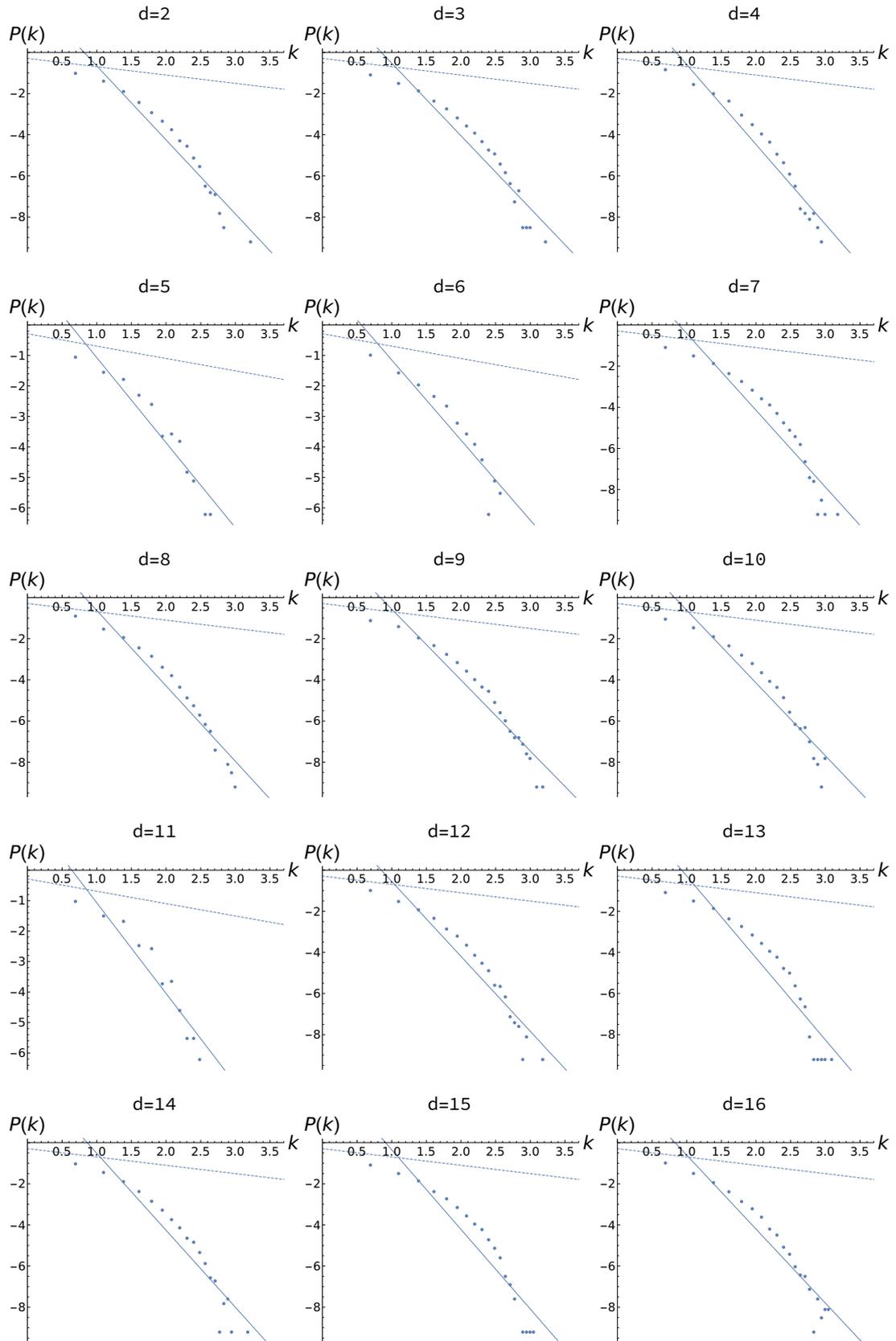


Figura 3.5: A distribuição de graus encontrada nos grafos de visibilidade horizontal (pontos) e a função ajustada $Ak^{-\gamma}$ (linha sólida). A linha pontilhada representando $P_{rand}(k) = (1/3)(2/3)^{k-2}$ foi incluída apenas por comparação.

Tabela 3.3: Valores de A e γ correspondentes ao ajuste da distribuição de graus $P(k) = Ak^{-\gamma}$ e o grau médio $\langle k \rangle$ em função de d para os grafos de VN.

| visibilidade | d | A | γ | $\langle k \rangle$ |
|--------------|-----|-----------------|-----------------|---------------------|
| natural | 2 | 0.4 ± 0.1 | 1.1 ± 0.2 | 6.17 |
| natural | 3 | 0.4 ± 0.1 | 1.1 ± 0.2 | 6.52 |
| natural | 4 | 1.2 ± 0.2 | 1.6 ± 0.1 | 6.40 |
| natural | 5 | 1.03 ± 0.07 | 1.53 ± 0.05 | 14.3 |
| natural | 6 | 0.5 ± 0.1 | 1.1 ± 0.2 | 6.33 |
| natural | 7 | 0.29 ± 0.03 | 1.02 ± 0.06 | 27.7 |
| natural | 8 | 0.4 ± 0.1 | 1.1 ± 0.2 | 6.28 |
| natural | 9 | 0.12 ± 0.05 | 0.6 ± 0.2 | 6.66 |
| natural | 10 | 0.4 ± 0.1 | 1.1 ± 0.2 | 6.23 |
| natural | 11 | 0.19 ± 0.02 | 0.91 ± 0.04 | 45.2 |
| natural | 12 | 0.42 ± 0.11 | 1.08 ± 0.18 | 6.40 |
| natural | 13 | 0.16 ± 0.01 | 0.90 ± 0.03 | 84.4 |
| natural | 14 | 0.44 ± 0.11 | 1.10 ± 0.18 | 6.25 |
| natural | 15 | 0.31 ± 0.05 | 1.01 ± 0.03 | 16.99 |
| natural | 16 | 0.42 ± 0.11 | 1.08 ± 0.19 | 6.31 |

Os resultados obtidos nesses experimentos numéricos são discutidos na próxima seção.

3.2 Discussão

Como mostra a Tabela 3.2, $\Delta \simeq 1$ para $d = \{3, 5, 7, 9, 11, 13, 15\}$, pois repetições nas sequências de x_d são raras, ou seja, essas sequências são aperiódicas. Adicionalmente $\Delta \approx 0.7$ para $d = \{2, 4, 6, 8, 10, 12, 14, 16\}$; portanto, a variabilidade do tamanho das lacunas para d par é menor do que para d ímpar. Então, uma possível conjectura é: a entropia normalizada Δ distingue d par ($\Delta \approx 0.7$) de d ímpar ($\Delta \simeq 1$) (Mayer; Monteiro, 2020).

As Figuras 3.4 e 3.5 e as Tabelas 3.3 e 3.4 mostram que a maior parte das distribuições de graus $P(k)$ dos grafos de VN e VH construídos a partir das lacunas entre d -primos sucessivos para $d \in \{2, 3, \dots, 16\}$ segue, aproximadamente, uma lei de potências dada por $P(k) = Ak^{-\gamma}$. As flutuações observadas próximas às linhas retas mostradas nessas figuras podem ser causadas pelo tamanho finito das sequências x_d usadas nos estudos numéricos (Broido; Clauset, 2019). Distribuições que seguem uma lei de potência na conectividade $P(k)$ estão associadas com redes complexas conhecidas como livres-de-escala (Broido;

Clauset, 2019; Keller, 2005; Newman, 2005). A invariância de escala na distribuição de graus $P(k)$ implica autossimilaridade, isto é, $P(k)$ de redes renormalizadas, obtidas por um processo de empacotamento (Song; Havlin; Makse, 2005), também seguem uma lei de potência.

Tabela 3.4: Valores de A e γ correspondentes ao ajuste da distribuição de graus $P(k) = Ak^{-\gamma}$ e o grau médio $\langle k \rangle$ em função de d para os grafos de VH.

| visibilidade | d | A | γ | $\langle k \rangle$ |
|--------------|-----|---------------|---------------|---------------------|
| horizontal | 2 | 1.2 ± 0.3 | 1.7 ± 0.3 | 3.67 |
| horizontal | 3 | 1.1 ± 0.2 | 1.6 ± 0.2 | 3.98 |
| horizontal | 4 | 1.7 ± 0.2 | 1.9 ± 0.2 | 3.47 |
| horizontal | 5 | 1.1 ± 0.2 | 1.6 ± 0.2 | 3.97 |
| horizontal | 6 | 1.1 ± 0.2 | 1.6 ± 0.2 | 3.89 |
| horizontal | 7 | 1.1 ± 0.2 | 1.6 ± 0.2 | 3.96 |
| horizontal | 8 | 1.5 ± 0.2 | 1.8 ± 0.2 | 3.59 |
| horizontal | 9 | 1.0 ± 0.2 | 1.6 ± 0.2 | 3.99 |
| horizontal | 10 | 1.1 ± 0.2 | 1.6 ± 0.2 | 3.84 |
| horizontal | 11 | 1.1 ± 0.2 | 1.6 ± 0.2 | 3.95 |
| horizontal | 12 | 1.3 ± 0.2 | 1.7 ± 0.2 | 3.78 |
| horizontal | 13 | 1.1 ± 0.2 | 1.6 ± 0.2 | 3.94 |
| horizontal | 14 | 1.2 ± 0.2 | 1.6 ± 0.2 | 3.77 |
| horizontal | 15 | 1.1 ± 0.2 | 1.6 ± 0.2 | 3.98 |
| horizontal | 16 | 1.3 ± 0.2 | 1.7 ± 0.2 | 3.77 |

Nota-se que o valor de γ nos gráficos de VN e VH não é um bom parâmetro para mostrar a influência do valor de d nas redes geradas pelas sequências de lacunas; entretanto $\langle k \rangle$ pode mostrar essa influência. De fato, nos gráficos de VN, para d primo, $\langle k \rangle$ aumenta com d ; para d não primo, $\langle k \rangle \approx 6$; nos gráficos de VH, $\langle k \rangle > 3.90$ para d ímpar e $\langle k \rangle < 3.90$ para d par. Portanto, outra conjectura possível é: $\langle k \rangle$ distingue d primos dos d não primos para a VN; e $\langle k \rangle$ distingue d ímpar de d par nos gráficos de VH (Mayer; Monteiro, 2020).

É sabido que para grafos de VH obtidos de sequências periódicas de período T (sem

números repetidos dentro de um período), o grau médio é dado por $\langle k \rangle = 4(1 - 1/2T)$ (Nunez; Lacasa; Valero; Gomez; Luque, 2012). Como consequência, para sequências aperiódicas, $\langle k \rangle = 4$ (pois $T \rightarrow \infty$). Na Tabela 3.4, esse é o valor aproximado de $\langle k \rangle$ encontrado para d -primos com d ímpar, o que está de acordo com o valor de Δ apresentado na Tabela 4.1. Também é conhecido que para uma sequência não correlacionada aleatória, o grafo de VH tem $P_{rand}(k) = (1/3)(2/3)^{k-2}$ (Nunez; Lacasa; Valero; Gomez; Luque, 2012). Portanto, desvios dessa distribuição de graus revelam que a sequência estudada não foi gerada por um processo aleatório não correlacionado. A curva correspondente a $P_{rand}(k)$ é mostrada como uma linha pontilhada na Figura 3.5. Note que $P(k)$ para d -primos tem menor inclinação do que $P_{rand}(k)$. Essa menor inclinação e $\langle k \rangle = 4$ podem ser indicações de uma sequência caótica (Nunez; Lacasa; Valero; Gomez; Luque, 2012).

No próximo capítulo, investiga-se a densidade de d -primos.

Capítulo 4

DENSIDADE DE d -PRIMOS

Munidos da função contagem de d -primos $\Pi_d(n)$, calculou-se a densidade de d -primos, isto é, $\Pi_d(n)/n$, para $d = 2, 3, \dots, 16$ e $n \in \mathbb{N}$. A intenção é determinar como a abundância de d -primos varia conforme se caminha no espaço dos números naturais. Além disso, essa métrica pode ser útil na estimativa das lacunas entre d -primos consecutivos.

A fim de obter uma expressão analítica aproximada para essa função, supôs-se que:

$$\frac{\Pi_d(n)}{n} \approx f_d(n) = \frac{a_d}{\ln(n) + b_d} \quad \text{para } n \rightarrow \infty \quad (4.1)$$

sendo a_d e b_d constantes que dependem do valor de d . Essas constantes são ajustadas para cada valor de d usando o método dos mínimos quadrados (Ljung, 1998). Obviamente, essa escolha de $f_d(n)$ foi inspirada no caso $d = 2$ para $n \rightarrow \infty$, para o qual $a_2 = 1$ e $b_2 = 0$, como mencionado na Introdução desta tese. O objetivo, então, é determinar os valores das constantes a_d e b_d que gerem um comportamento assintótico similar ao caso $d = 2$.

4.1 Resultados

Nesse estudo, a densidade de d -primos $\Pi_d(n)/n$ foi calculada para os diferentes valores de d . Para $d = 2$, esse assunto foi motivo de pesquisa de muitos matemáticos famosos, como comentado na Introdução.

A Figura 4.1 mostra como $\Pi_d(n)/n$ varia para $n \in \mathbb{N}$, sendo $\Pi_d(n)$ a quantidade de d -primos menores ou iguais a n . Essa figura revela que há dois grupos de resultados: o grupo com d par e grupo com d ímpar. Para os casos com d par, as curvas podem apresentar concavidade para cima ou para baixo. Para d ímpar, as curvas tendem rapidamente para

valores próximos de zero.

Uma possível explicação para a concavidade invertida para o caso $d = 8$, quando comparada com a concavidade dos primos usuais (2-primos), é que essa sequência contém os números esfênicos, que são números formados pelo produto de 3 primos distintos e, talvez por isso, sejam mais abundantes (se um número é formado pelo produto dos primos α , β e γ , então esse número possui oito divisores: α , β , γ , $\alpha\beta$, $\alpha\gamma$, $\beta\gamma$, $\alpha\beta\gamma$ e o número 1). Porém, para $d = \{12, 16\}$, o motivo da concavidade invertida não é sequer especulado. Para valores ímpares de d , a densidade é muito baixa, o que dificulta a visualização.

Combinando os resultados mostrados na Figura 4.1, percebe-se que, conforme n cresce, converge-se para uma distribuição em que cerca de 10% são 2-primos, 25% são 4-primos, 5% são 6-primos, 20% são 8-primos, 5% são 10-primos, 10% são 12-primos, 2% são 14-primos e 10% são 16-primos. A soma dessas porcentagens dá 92%. Os 8% restantes são d -primos com d ímpar ou com d par para $d > 16$. Essas porcentagens podem se alterar conforme n cresce; contudo, essas porcentagens podem convergir, como acontece para os 2-primos, para os quais $\Pi_d(n)/n \rightarrow 1/\ln(n)$.

É muito curioso perceber, por exemplo, que os 4-primos são os números (compostos) mais abundantes e que a abundância de 14-primos é muito menor do que a abundância de 12-primos ou 16-primos, no intervalo de números naturais analisado.

A Tabela 4.1 mostra os resultados obtidos pelo ajuste dos valores dos parâmetros que aparecem na função 4.1.

| d | a_d | σ | b_d | σ |
|-----|----------------------|----------------------|----------------------|----------------------|
| 2 | 1.017 | 0.001 | 1.229 | 0.008 |
| 3 | 0.0012 | 0.0002 | -0.6878 | 0.0009 |
| 4 | 5.83 | 0.02 | 14.55 | 0.06 |
| 5 | 0.00025 | 0.00002 | 0.0040 | 0.0003 |
| 6 | 0.620 | 0.002 | 2.02 | 0.02 |
| 7 | 0.000025 | 0.000002 | 0.0016 | 0.0001 |
| 8 | 8.2×10^{12} | 1.2×10^{22} | 3.9×10^{13} | 5.7×10^{22} |
| 9 | 0.00050 | 0.00003 | -1.585 | 0.002 |
| 10 | 0.0909 | 0.0002 | 1.82 | 0.02 |
| 11 | 3.0×10^{-7} | 3×10^{-8} | 0.00031 | 0.00003 |
| 12 | 75 | 15 | 734 | 151 |
| 13 | 3.3×10^{-8} | 3×10^{-9} | 0.00014 | 0.00001 |
| 14 | 0.01929 | 0.00003 | 1.78 | 0.02 |
| 15 | 0.00025 | 0.00001 | 0.036 | 0.002 |
| 16 | 7.8×10^{12} | 8.6×10^9 | 8.1×10^{13} | 8.3×10^8 |

Tabela 4.1: Valores de a_d e b_d e os correspondentes desvios padrões σ da função ajustada.

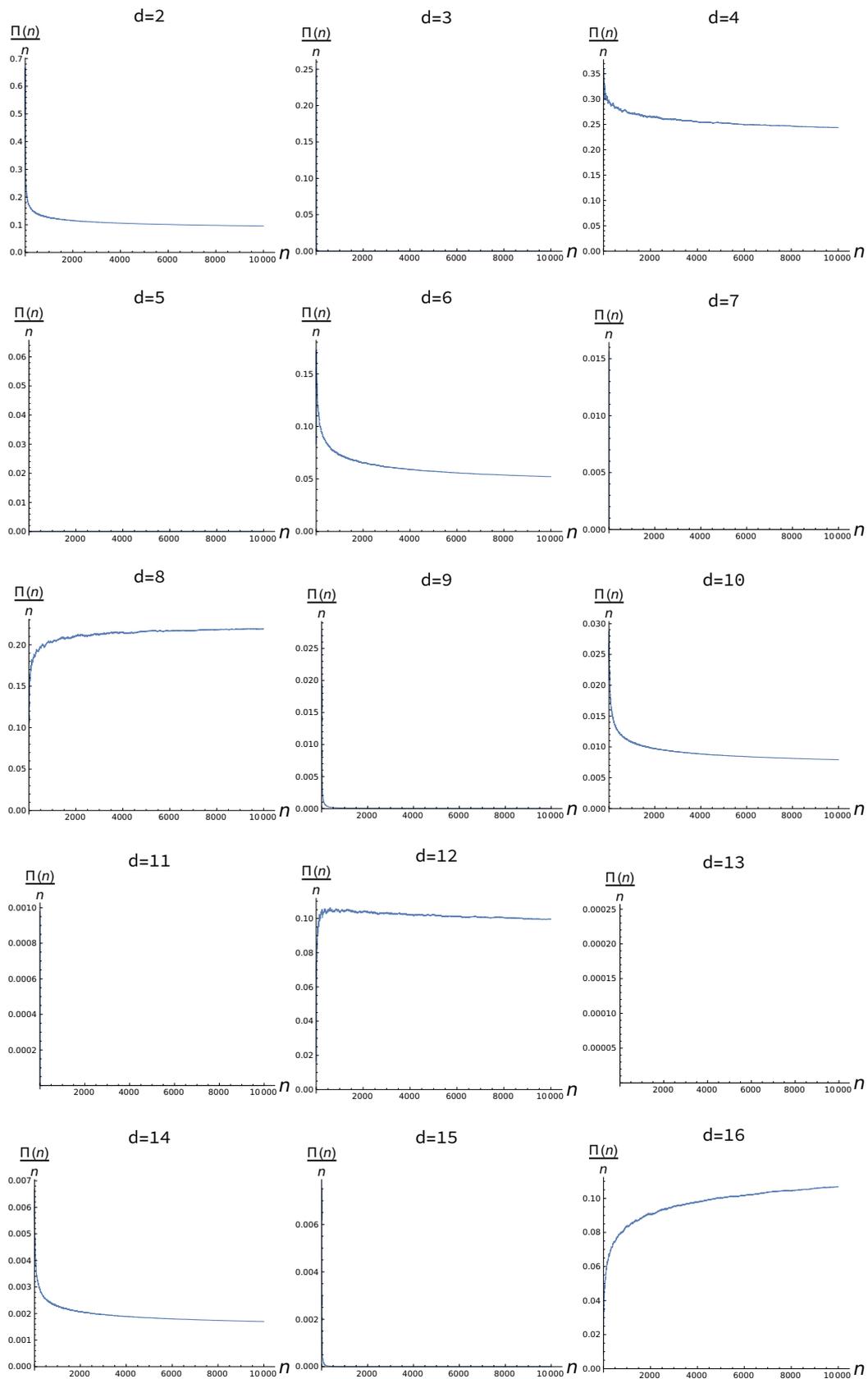


Figura 4.1: Densidades de d -primos para $d = 2, 3, \dots, 16$. No eixo horizontal, estão os naturais e no vertical, o valor de $\Pi_d(n)/n$: a fração de d -primos em relação a n .

A seguir, apresenta-se a figura 4.2. Essa figura mostra como o desvio relativo $R(n)$, definido como:

$$R(n) = \frac{\Pi_d(n)/n}{f_d(n)} \quad (4.2)$$

varia com n . Quanto mais próximo de 1 for o valor de R , melhor é o ajuste propiciado pela função $f_d(n)$ escolhida. Essa medida é importante para o trabalho, pois é com ela que a equivalência assintótica é determinada. Diz-se que as funções $F(n)$ e $G(n)$ são assintoticamente equivalentes se:

$$\lim_{n \rightarrow \infty} \frac{F(n)}{G(n)} = 1 \quad (4.3)$$

4.2 Discussão

Nota-se na Figura 4.1 que existe uma distinção de comportamento na densidade dos d -primos: para d par, a densidade parece tender a um número fixo entre 0.002 e 0.3; para d ímpar, esses valores caem rapidamente e se aproximam de zero.

Dentre os pares, para $d = \{8, 12, 16\}$ tem-se a densidade com concavidade voltada para baixo; as demais densidades para d par têm concavidade voltada para cima. Esse é o comportamento dos números primos usuais, ou seja, do caso em que $d = 2$. Esse resultado para $d = 2$ já era conhecido, pois é o mesmo estudo realizado por Gauss e Legendre, como apresentado na Introdução; ou seja, é sabido que $1/\log(n)$ é a solução assintoticamente equivalente para o caso $d = 2$.

De acordo com a Tabela 4.1, pode-se descartar a função f_d escolhida para $d = 8, 12, 16$, pois seus parâmetros possuem um desvio padrão muito grande. De fato, os valores de a e b são absurdos. Já a Figura 4.2 indica que os casos $d = 2, 4, 6, 10, 14$ possuem comportamento similar; isto é, a concavidade para cima, valores próximos e, visualmente, as curvaturas se assemelham. Para os casos com a concavidade voltada para baixo, a função f_d não é um bom ajuste.

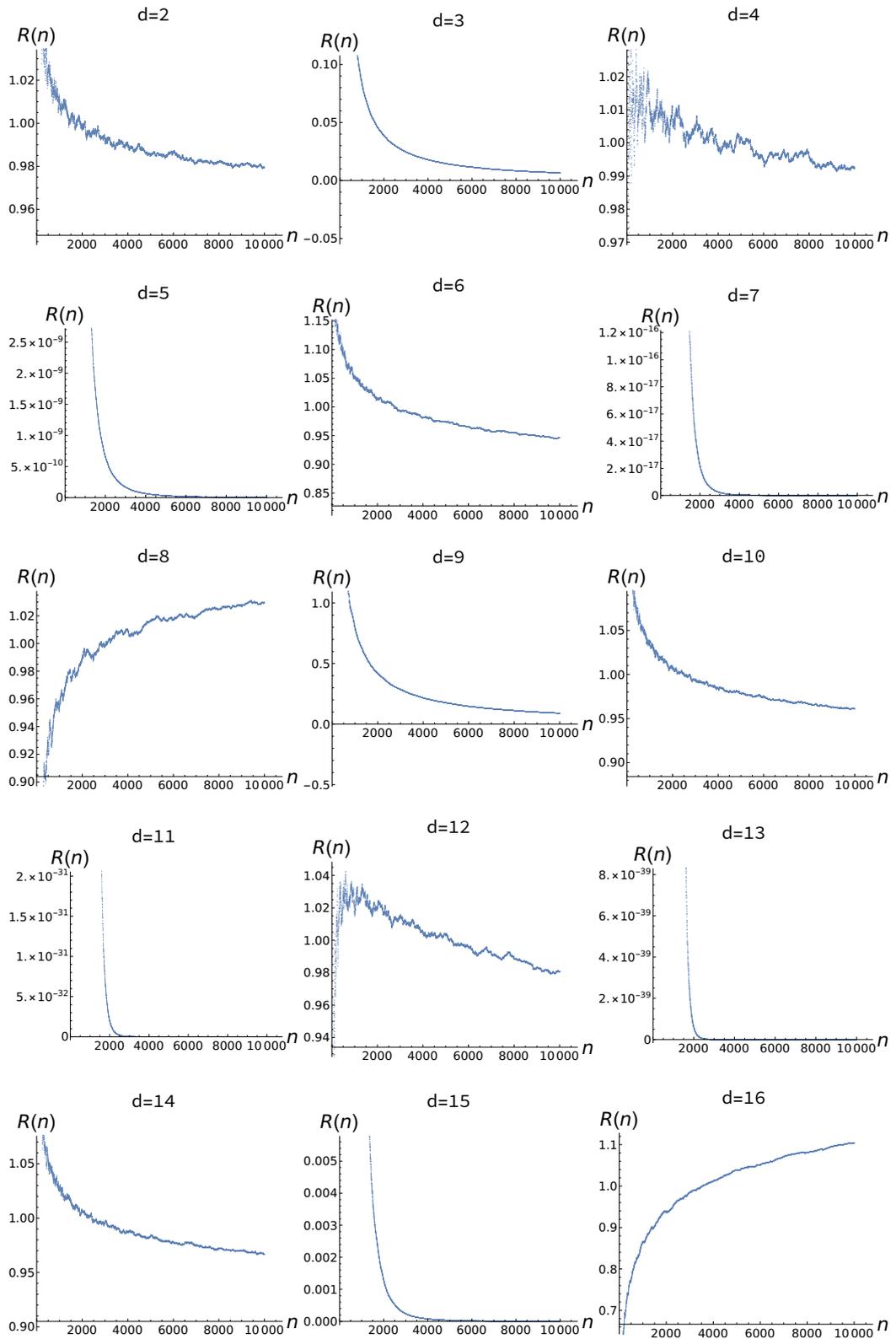


Figura 4.2: Gráfico de $R(n) = \frac{\Pi_d(n)/n}{f_d(n)}$ em função de n para $d = 2, 3, \dots, 16$.

No caso de d ímpar, a Tabela 4.1 e a Figura 4.2 se contradizem: de acordo com a tabela, os parâmetros têm erros muito baixos, o que pode indicar um ótimo ajuste; porém, a figura mostra os valores de R indo para zero rapidamente. Isso significa que, conforme n aumenta, os valores de f_d e os valores reais se tornam relativamente distantes, de modo que não são assintoticamente equivalentes.

Pode-se conjecturar que a dependência da densidade $\Pi_d(n)/n$ com n é similar ao caso $d = 2$ para d par e com concavidade voltada para baixo, o que corresponde a 62.5% dos pares ou 32.25% do total. Essa porcentagem pode se alterar com a análise de outros valores de d .

Para d 2-primo inesperadamente o modelo não é um bom ajuste, mesmo estando diretamente relacionados com os 2-primos, como foi mostrado pela relação (2.7). A densidade desses d -primos não é proporcional a dos 2-primos pois possui densidade muito baixa.

Esses resultados podem ser úteis na pesquisa de lacunas entre números primos, um campo de estudo famoso na teoria dos números. Aqui, a pesquisa abrangeu valores de $d > 2$. Isso permite classificar os d -primos, em termos de $\Pi_d(n)/n$, em 3 classes:

d ímpar: densidade muito baixa;

d par e concavidade para cima: densidade semelhante aos números primos;

d par e concavidade para baixo: diferente dos anteriores.

Para valores pares de d , mais pesquisas são necessárias, pois da mesma forma que $d = 8$ fornece um comportamento distinto, por conter os números esfênicos, outros valores porerão mostrar o mesmo comportamento anômalo. Isso levanta questões interessantes: para quais valores de d o gráfico possui a mesma concavidade de $d = \{8, 12, 16\}$? Seriam os múltiplos de 4 os candidatos? Ou outras potências de 2? Por exemplo, será que o caso $d = 20$ também possui essa característica? Ou apenas $d = 32$, que é a próxima potência de 2? Em ambos casos, por que $d = 4$, que é potência de 2 e múltiplo de 4, não possui o mesmo comportamento?

No próximo capítulo, volta-se a usar entropia informacional e algoritmos de visibilidade para analisar sequências matemáticas.

Capítulo 5

QUANTIDADE DE DIVISORES DOS NÚMEROS NATURAIS E FELIZES

Considere a *sequência dos divisores* $(n, D(n))$; isto é, a sequência da quantidade de divisores $D(n) = d$ do número natural n (OEIS A000005). Neste estudo, são levados em conta os 30000 primeiros termos dessa sequência:

$$\{(2, 2), (3, 2), (4, 3), (5, 2), (6, 4) \dots, (30000, 50)\}$$

Essa sequência recebe a notação $\{D(n)\}$.

Considere também outra sequência denominada de *sequência das trajetórias dos divisores*. Para ilustrar novamente a construção dessa sequência, tome, por exemplo, o número 10. Note que:

- 10 tem 4 divisores (os números 1, 2, 5 e 10); essa é a primeira iteração da aplicação da função quantidade de divisores D ;
- 4 tem 3 divisores; segunda iteração de D ;
- 3 tem 2 divisores; terceira e última iteração de D . Lembre que a iteração final é aquela em que se atinge o número 2.

Nesse caso, a função quantidade de divisores D foi aplicada três vezes. De fato, $D(10) = 4$, $D(4) = 3$ e $D(3) = 2$. Aqui, diz-se que a trajetória do número 10 tem tamanho

3. Isso corresponde ao par $(10, 3)$. Aplicando esse procedimento para os números naturais entre 1 e 30000, gera-se a sequência com código A036459 (OEIS, 2023) dada por:

$$\{(2, 1), (3, 1), (4, 2), (5, 1), (6, 3), (7, 1), \dots, (30000, 5)\}$$

Aqui essa sequência é denotada por $\{T(n)\}$.

Nesse estudo, inicialmente, foi calculado o valor da entropia informacional normalizada Δ . Em seguida, os grafos de VN e VH foram construídos, a partir dessas duas sequências relacionadas aos divisores dos números naturais. Para esses grafos, foram calculados a distribuição de graus $P(k)$ e o grau médio $\langle k \rangle$. Ainda, assumiu-se que $P(k) = Ak^{-\gamma}e^{-\delta k}$. Assim, seguindo a proposta de Newman (2001), introduziu-se δ , um parâmetro a mais em relação aos ajustes de $P(k)$ feitos para os d -primos, a fim de se verificar se esse parâmetro extra propicia ajustes melhores. Os coeficientes A , γ e δ dos modelos para $P(k)$ de cada grafo foram determinados pelo método dos mínimos quadrados (Ljung, 1998).

Para efeito de comparação, esse mesmo procedimento foi repetido para a sequência dos números felizes, gerando as sequências dos divisores e das trajetórias, denotadas por $\{D(h)\}$ e $\{T(h)\}$, respectivamente. Os termos de $\{D(h)\}$ considerados são:

$$\{(1, 1), (7, 2), (10, 4), (13, 2), (19, 2), (23, 2), (28, 6), \dots, (204480, 84)\}$$

e os termos de $\{T(h)\}$ são:

$$\{(1, 0), (7, 1), (10, 3), (13, 1), (19, 1), (23, 1), (28, 4), \dots, (204480, 6)\}$$

Os resultados desse estudo foram publicados em (Mayer; Monteiro, 2023).

5.1 Resultados

Nesta seção, apresentam-se os resultados referentes aos grafos gerados usando a sequência de divisores e a sequência das trajetórias dos divisores, para os número naturais e para os números felizes, como explicado acima.

5.1.1 Para os números naturais

A Tabela 5.1 reúne os dados da função $Ak^{-\gamma}e^{-\delta k}$ que ajusta a distribuição de graus $P(k)$ dos grafos de VH e VN, além do grau médio, do menor caminho médio e do coeficiente de agregação médio. A Tabela 5.2 apresenta outras medidas relacionadas a esses grafos. Essas medidas se referem à paridade dos nós que estão conectados. Nessa tabela, χ_1 é a porcentagem de nós pares, χ_2 , a porcentagem de nós ímpares, m_{11} a porcentagem de ligações entre nós ímpares, m_{22} a porcentagem de ligações entre nós pares, m_{12} a porcentagem de ligações de um nó ímpar para um nó par, m_{21} a porcentagem de ligações de um nó par para um nó ímpar e M o número total de arestas. Esses resultados são discutidos no capítulo final desta tese.

A Figura 5.1 mostra os grafos gerados utilizando os dois algoritmos de visibilidade para uma sequência com apenas as 100 primeiras arestas (números maiores dificultam a visualização).

Tabela 5.1: Valores do grau médio $\langle k \rangle$, caminho mínimo médio $\langle l \rangle$, coeficiente de clusterização médio $\langle c \rangle$, e A , γ , e δ da distribuição de graus $P(k) = Ak^{-\gamma}e^{-\delta k}$ (e o erro quadrático médio (EQM) da função ajustada) dos grafos obtidos empregando os algoritmos de visibilidades natural e horizontal.

| sequência | visibilidade | $\langle k \rangle$ | $\langle l \rangle$ | $\langle c \rangle$ | A | γ | δ | EQM |
|------------|--------------|---------------------|---------------------|---------------------|------|----------|----------|----------------------|
| $\{D(n)\}$ | natural | 5.78 | 7.92 | 0.33 | 0.17 | -2.27 | 0.78 | 1.6×10^{-5} |
| $\{T(n)\}$ | natural | 5.04 | 28.3 | 0.06 | 0.38 | -0.76 | 0.48 | 2.4×10^{-5} |
| $\{D(n)\}$ | horizontal | 3.58 | 16.3 | 0.27 | 1.68 | 1.53 | 0.14 | 3.3×10^{-5} |
| $\{T(n)\}$ | horizontal | 2.98 | 257 | 0.26 | 1.87 | 0.60 | 0.46 | 1.6×10^{-4} |

Tabela 5.2: Porcentagens de nós ímpares (χ_1), pares (χ_2), ligações entre nós ímpares (m_{11}), ligações de um ímpar para um nó par (m_{12}), ligações de um nó par para um ímpar (m_{21}), e ligações entre nós pares (m_{22}) nos quatro grafos. O número total de ligações (M) é dado na última coluna.

| sequência | visibilidade | χ_1 | χ_2 | m_{11} | m_{12} | m_{21} | m_{22} | M |
|------------|--------------|----------|----------|----------|----------|----------|----------|-------|
| $\{D(n)\}$ | natural | 50.0 | 50.0 | 3.4 | 26.7 | 26.7 | 43.2 | 86645 |
| $\{T(n)\}$ | natural | 50.0 | 50.0 | 3.7 | 28.7 | 28.4 | 39.2 | 75579 |
| $\{D(n)\}$ | horizontal | 50.0 | 50.0 | 1.0 | 31.8 | 31.8 | 35.4 | 53723 |
| $\{T(n)\}$ | horizontal | 50.0 | 50.0 | 1.3 | 37.2 | 37.2 | 24.3 | 44655 |

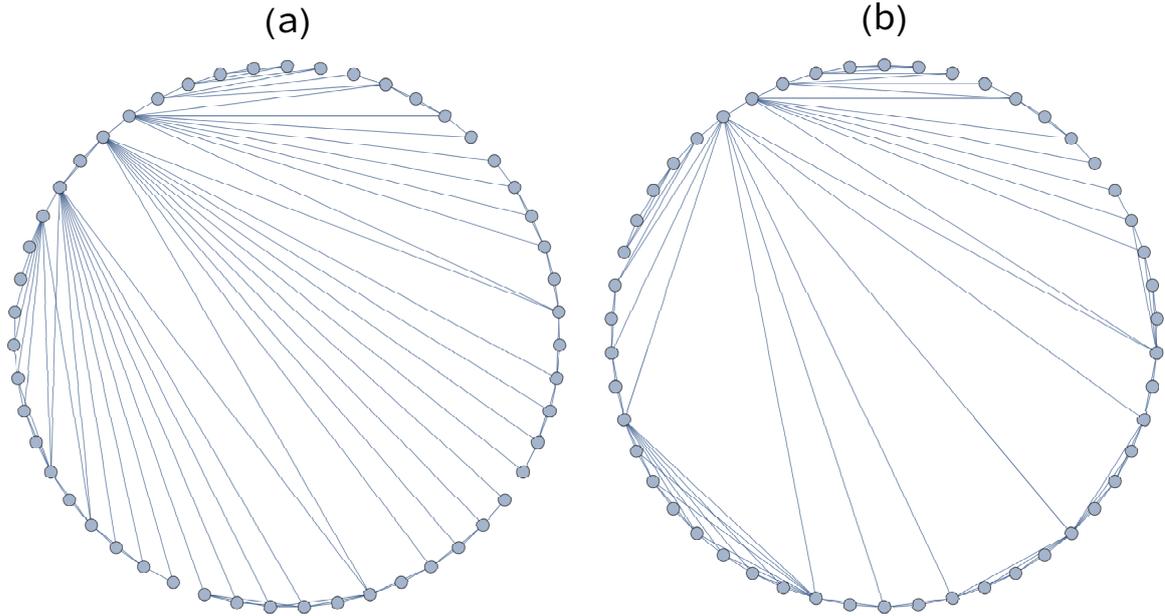


Figura 5.1: Grafos gerados usando o algoritmo de visibilidade natural. Esquerda: $\{D(n)\}$; direita: $\{T(n)\}$.

A Figura 5.2 mostra como entropia informacional normalizada Δ varia com o tamanho da sequência considerada.

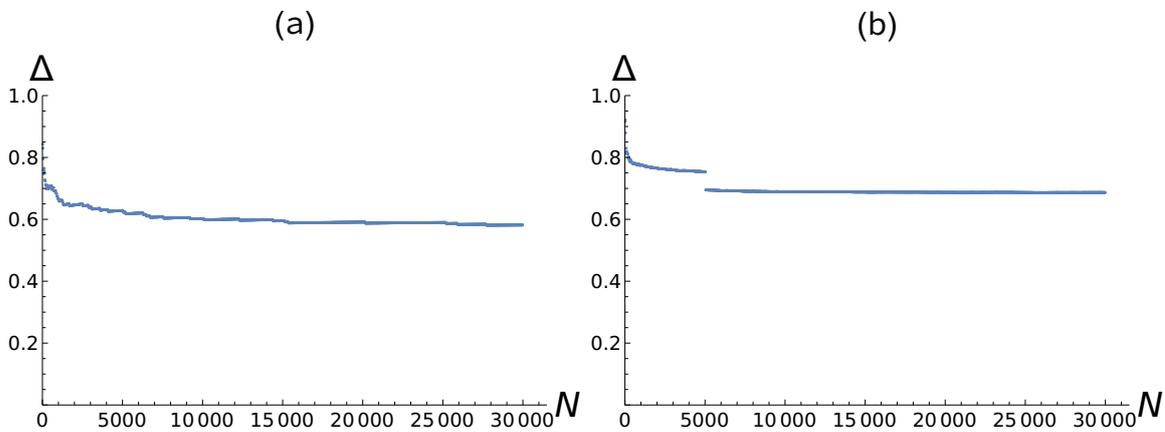


Figura 5.2: A entropia informacional normalizada $\Delta = H/H_{max}$ para $\{D(n)\}$ (a) and $\{T(n)\}$ (b) em função do tamanho da sequência N .

Os valores para $N = 30000$ são $\Delta = 0.582$ para $\{D(n)\}$ e $\Delta = 0.687$ para $\{T(n)\}$. Os saltos em Δ são mais evidentes em $\{T(n)\}$, pois correspondem à introdução de um novo valor na trajetória. O mais proeminente acontece com o número 5040, que possui trajetória de tamanho 6, um tamanho que não existia para valores menores, fazendo a entropia saltar de 0.753 para 0.694.

A Figura 5.3 exibe a distribuição de graus para os grafos de VN. A linha cheia corresponde à curva ajustada; os pontos são os dados obtidos nos experimentos computacionais.

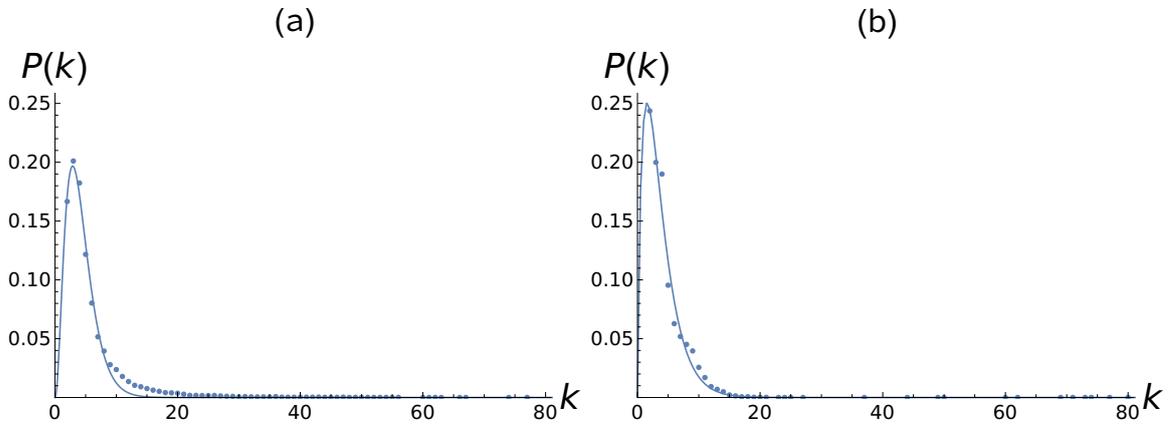


Figura 5.3: Distribuição de graus $P(k)$ para $\{n, D(n)\}$ (a) e $\{n, T(n)\}$ (b) para os grafos obtidos pela visibilidade natural. A linha sólida é a curva ajustada $P(k) = Ak^{-\gamma}e^{-\delta k}$.

Uma amostra dos grafos gerados com a VH é mostrada na Figura 5.4.

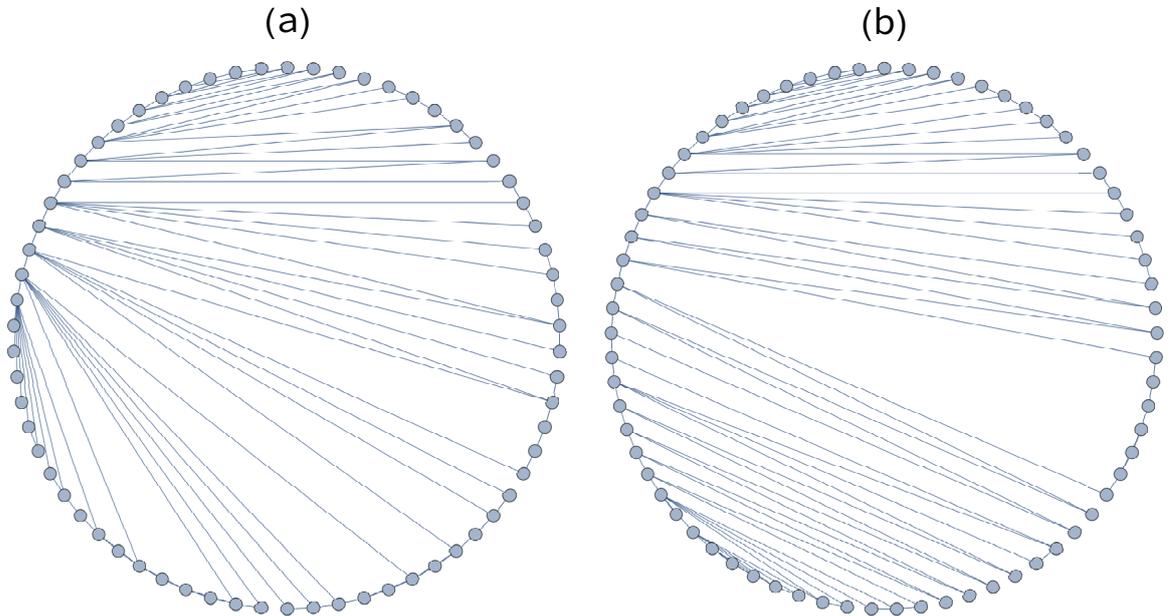


Figura 5.4: Grafos gerados usando o algoritmo de visibilidade horizontal. Esquerda: $\{D(n)\}$; direita: $\{T(n)\}$.

As Figuras 5.1 e 5.4 sugerem que os grafos gerados pelas sequências de divisores refletem são mais fortemente conectados do que os grafos gerados pelas sequências das trajetórias. Esse resultado é comprovado comparando-se os graus médios desses grafos, como mostrado na Tabela 5.1.

A Figura 5.5 exibe a distribuição de graus para os grafos de VH.

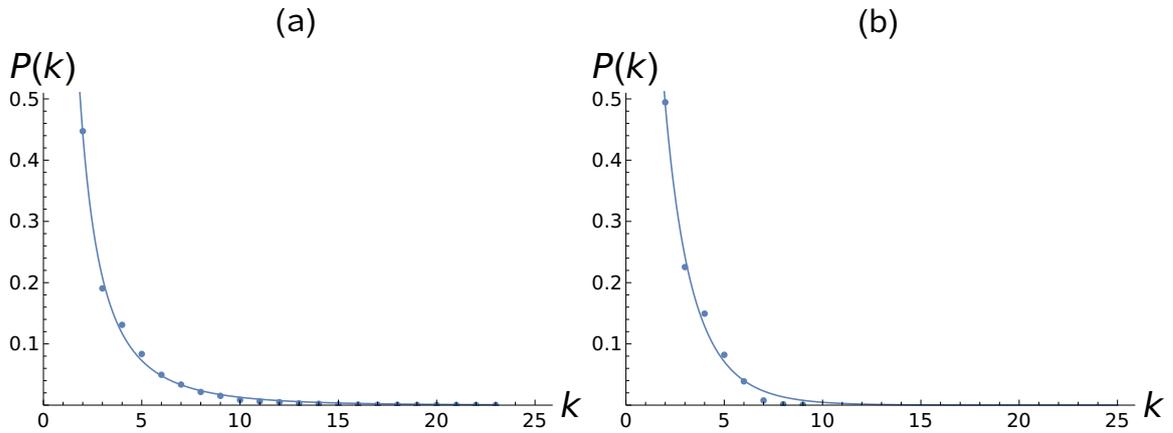


Figura 5.5: Distribuição de graus $P(k)$ para $\{n, D(n)\}$ (a) e $\{n, T(n)\}$ (b) para os grafos obtidos pela visibilidade horizontal. A linha sólida é a curva ajustada $P(k) = Ak^{-\gamma}e^{-\delta k}$.

5.1.2 Para os números felizes

Nesta seção, apresentam-se os resultados referentes às sequências dos números felizes.

A Tabela 5.3 reúne os dados da função $Ak^{-\gamma}e^{-\delta k}$ que ajusta a distribuição de graus $P(k)$ dos grafos de VH e VN.

Tabela 5.3: Valores do grau médio $\langle k \rangle$, caminho mínimo médio $\langle l \rangle$, coeficiente de clusterização médio $\langle c \rangle$, e A , γ , e δ da distribuição de graus $P(k) = Ak^{-\gamma}e^{-\delta k}$ (e o erro quadrático médio (EQM) da função ajustada) dos grafos obtidos empregando os algoritmos de visibilidades natural e horizontal.

| sequência | visibilidade | $\langle k \rangle$ | $\langle l \rangle$ | $\langle c \rangle$ | A | γ | δ | EQM |
|------------|--------------|---------------------|---------------------|---------------------|------|----------|----------|----------------------|
| $\{D(h)\}$ | natural | 6.61 | 7.39 | 0.35 | 0.10 | -2.25 | 0.66 | 9.6×10^{-6} |
| $\{T(h)\}$ | natural | 5.33 | 100 | 0.14 | 0.19 | -2.19 | 0.77 | 7.2×10^{-6} |
| $\{D(h)\}$ | horizontal | 3.51 | 15.4 | 0.27 | 1.34 | 0.71 | 0.33 | 9.8×10^{-6} |
| $\{T(h)\}$ | horizontal | 2.92 | 157 | 0.26 | 1.71 | -0.39 | 0.77 | 1.4×10^{-4} |

A Tabela 5.4 apresenta medidas relacionadas à paridade dos nós que estão conectados. Esses resultados são discutidos no próximo capítulo.

Os grafos para os números felizes apresentam estrutura similar aos gerados para os números naturais: há nós com muitas ligações e ligações que cruzam os grafos, numa estrutura típica de rede mundo-pequeno.

A Figura 5.6 mostra os grafos de VN com apenas 100 arestas.

Tabela 5.4: Porcentagens de nós ímpares (χ_1), pares (χ_2), ligações entre nós ímpares (m_{11}), ligações de um ímpar para um nó par (m_{12}), ligações de um nó par para um ímpar (m_{21}), e ligações entre nós pares (m_{22}) nos quatro grafos. O número total de ligações (M) é dado na última coluna.

| sequência | visibilidade | χ_1 | χ_2 | m_{11} | m_{12} | m_{21} | m_{22} | M |
|------------|--------------|----------|----------|----------|----------|----------|----------|-------|
| $\{D(h)\}$ | natural | 49.6 | 50.4 | 11.1 | 23.6 | 22.6 | 42.7 | 99143 |
| $\{T(h)\}$ | natural | 49.6 | 50.4 | 13.2 | 23.0 | 23.1 | 40.7 | 79940 |
| $\{D(h)\}$ | horizontal | 49.6 | 50.4 | 14.6 | 23.3 | 22.7 | 39.4 | 52626 |
| $\{T(h)\}$ | horizontal | 49.6 | 50.4 | 17.4 | 25.8 | 25.6 | 31.2 | 43755 |

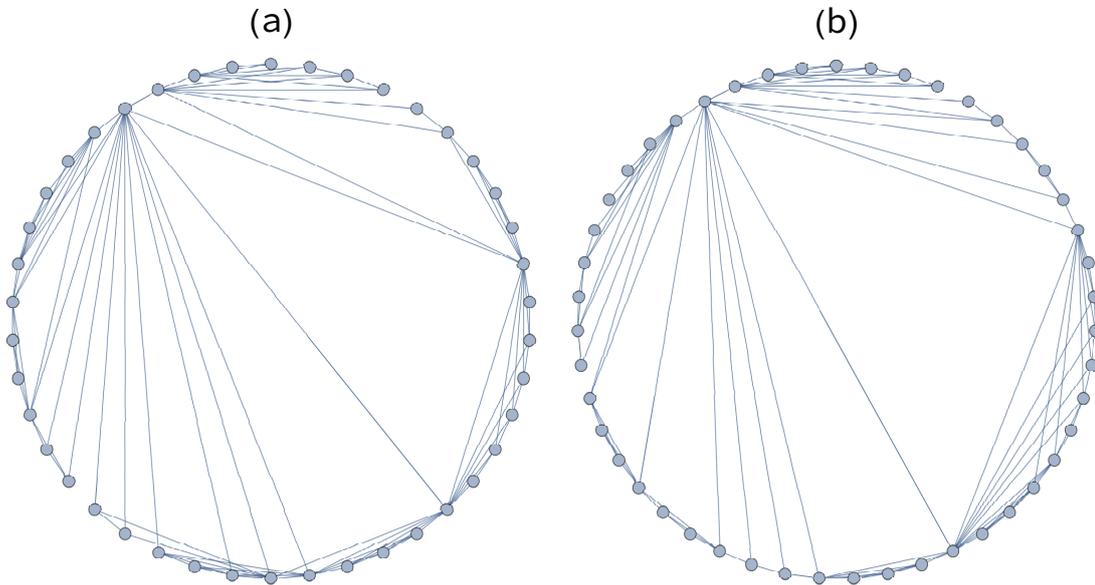


Figura 5.6: Grafos gerados usando o algoritmo de visibilidade natural. Esquerda: $\{h, D(h)\}$; direita: $\{h, T(h)\}$.

Esses grafos sugerem que há um maior número de conexões em menos nós, isto é, menor variabilidade.

O gráfico da entropia informacional, mostrado na Figura 5.7, é similar ao dos naturais. Entretanto, a descontinuidade mais evidente ocorre para um valor de N diferente. Nesse caso, esse salto corresponde ao número feliz 11088, pois é o primeiro com trajetória de tamanho 6. Como consequência, a entropia salta de 0.746 para 0.689.

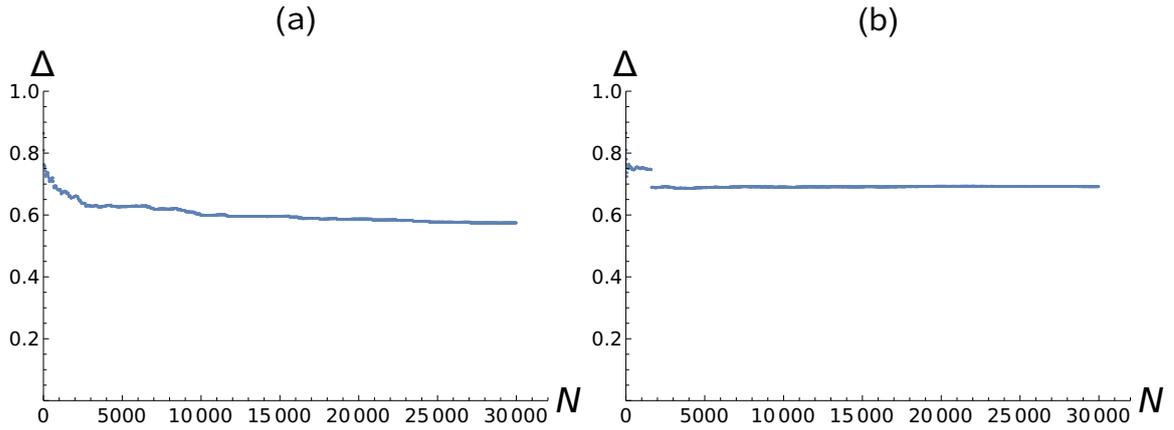


Figura 5.7: A entropia informacional normalizada $\Delta = H/H_{max}$ para $\{D(h)\}$ (a) e $\{T(h)\}$ (b) em função do tamanho da sequência N .

Os valores para $N = 30000$ são: $\Delta = 0.575$ para $\{D(n)\}$ e $\Delta = 0.692$ para $\{T(n)\}$.

A Figura 5.8 mostra a distribuição de graus para VN.

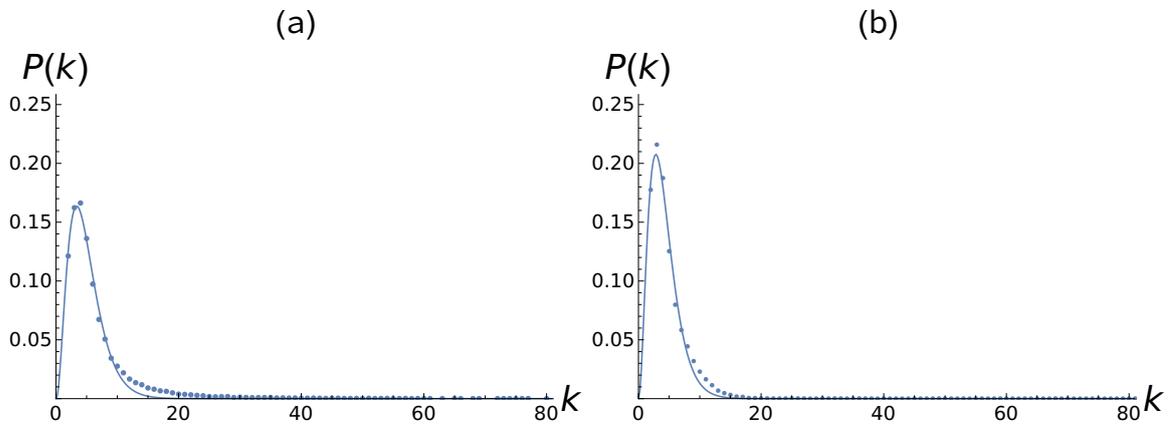


Figura 5.8: Distribuição de graus $P(k)$ para $\{h, D(h)\}$ (a) e $\{h, T(h)\}$ (b) para os grafos obtidos pela visibilidade natural. A linha sólida é a curva ajustada $P(k) = Ak^{-\gamma}e^{-\delta k}$.

Finalmente, a Figura 5.9 apresenta o grafo de VH com as primeiras 100 arestas.

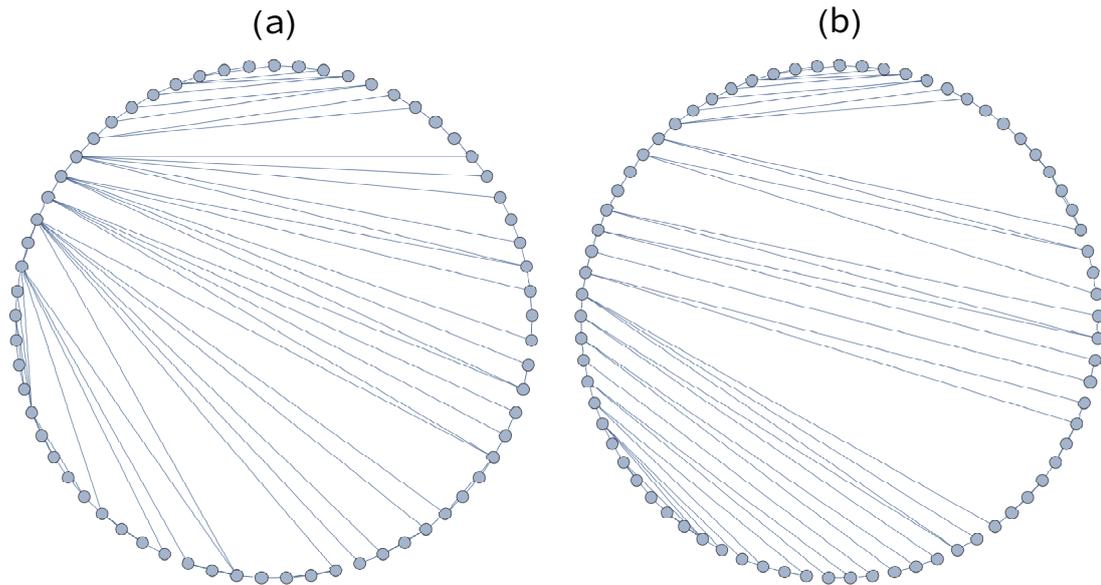


Figura 5.9: Grafos gerados usando o algoritmo de visibilidade horizontal. Esquerda: $\{h, D(h)\}$; direita: $\{h, T(h)\}$.

As distribuições de graus para os grafos de VH são mostradas na Figura 5.10.

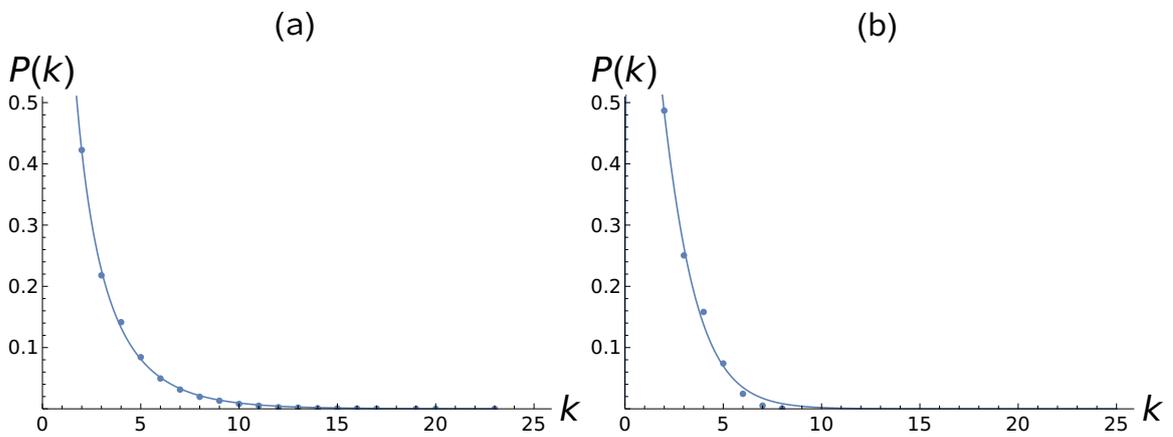


Figura 5.10: Distribuição de graus $P(k)$ para $\{h, D(h)\}$ (a) e $\{h, T(h)\}$ (b) para os grafos obtidos pela visibilidade horizontal. A linha sólida é a curva ajustada $P(k) = Ak^{-\gamma}e^{-\delta k}$

A seguir, discutem-se os resultados mostrados nas tabelas e figuras deste capítulo.

5.2 Discussão

Neste estudo, observou-se que, apesar da similaridade do valor de entropia normalizada, o que implica variabilidade comparável, e apesar da proporção similar de números ímpares e pares em $D(n)$ e $D(h)$, as sequências dos números naturais e felizes são mapeadas em grafos com medidas topológicas distintas. Essa afirmação também é válida para as sequências das trajetórias. Observa-se que $\Delta \simeq 0.7$ para $T(n)$ e $T(h)$, como pode ser visto nas Figuras 5.2 e 5.7; no entanto, essas sequências são transformadas em grafos com características diferentes.

Para $N = 30000$, $1 \leq D(n) \leq 96$, $1 \leq D(h) \leq 160$ e $0 \leq T(n), T(h) \leq 6$. Assim, o número de valores distintos nas sequências $\{D\}$ é muito maior do que nas sequências $\{T\}$. Surpreendentemente, a entropia normalizada Δ é maior para as sequências $\{T\}$ do que para as sequências $\{D\}$.

O grafo VH é um subgrafo do grafo de VN correspondente (Luque; Lacasa; Ballesteros; Luque, 2009; Lacasa; Luque; Ballesteros; Luque; Nuno, 2008; Zou; Donner; Marwan; Donges; Kurths, 2019); portanto, como esperado, as Tabelas 5.2 e 5.4 mostram que M é maior nos grafos de VN do que nos grafos de VH, e as Tabelas 5.1 e 5.3 mostram que $\langle k \rangle$ é maior nos grafos de NV do que nos grafos de VH, e $\langle l \rangle$ é menor nos grafos VN do que nos grafos VH. Por exemplo, para $\{D(n)\}$, $M_{VN} = 86645 > M_{VH} = 53723$, $\langle k \rangle_{VN} = 5.78 > \langle k \rangle_{VH} = 3.58$, $\langle l \rangle_{VN} = 7.92 < \langle l \rangle_{VH} = 16.3$. No entanto, a maior quantidade de arestas nos grafos VN não implica em um $\langle c \rangle$ maior. Por exemplo, $\langle c \rangle_{VN} = 0.33 > \langle c \rangle_{VH} = 0.27$ para $D(n)$, mas $\langle c \rangle_{VN} = 0.06 < \langle c \rangle_{VH} = 0.26$ para $\{T(n)\}$.

Ao empregar o algoritmo de VN, os grafos construídos a partir da sequência de divisores $\{D\}$ apresentaram $\langle k \rangle$ maior, $\langle l \rangle$ menor e $\langle c \rangle$ maior do que os respectivos grafos construídos a partir da sequência de trajetórias $\{T\}$, para os números naturais e os felizes. Por exemplo, para os grafos VN, $\langle k \rangle_{D(h)} = 6.61 > \langle k \rangle_{T(h)} = 5.33$, $\langle l \rangle_{D(h)} = 7.39 < \langle l \rangle_{T(h)} = 100$, e $\langle c \rangle_{D(h)} = 0.35 > \langle c \rangle_{T(h)} = 0.14$. Ao utilizar o algoritmo de VH, os grafos para $\{D\}$ apresentaram $\langle k \rangle$ maior e $\langle l \rangle$ menor do que os respectivos grafos para $\{T\}$, mas $\langle c \rangle$ é similar em ambos os grafos.

A fórmula $\langle l \rangle_{random} = \log N / \log \langle k \rangle$, geralmente utilizada para estimar o caminho mínimo médio em grafos puramente aleatórios, fornece valores menores do que os mostrados nas Tabelas 5.1 e 5.3. Por exemplo, para o grafo VN de $\{D(h)\}$, $\langle l \rangle = 7.39 > \langle l \rangle_{random} = 5.49$; para o grafo VH de $\{D(h)\}$, $\langle l \rangle = 15.4 > \langle l \rangle_{random} = 8.21$. De fato, os

grafos obtidos a partir de algoritmos de visibilidade não são puramente aleatórios; portanto, $\langle l \rangle$ é maior do que $\langle l \rangle_{random}$. No entanto, os gráficos de $P(k)$ sugerem que as quatro sequências podem ser consideradas como sequências quase aleatórias.

Além disso, a função ajustada $P(k)$ muda qualitativamente do grafo VN para o de VH, para os números naturais e para os números felizes. De maneira geral, o algoritmo VN cria redes aleatórias/mundo pequeno (Newman, 2003; Boccaletti; Latora; Moreno; Chavez; Hwanga, 2006) e o algoritmo VH cria redes livres de escala (Newman, 2003; Boccaletti; Latora; Moreno; Chavez; Hwanga, 2006); ou seja, $P(k)$ nos grafos VN é uma distribuição tipo Poisson (Figuras 5.3 e 5.8) e $P(k)$ nos grafos VH é uma distribuição livre de escala com uma cauda exponencial (Figuras 5.5 e 5.10). Uma vez que uma sequência aleatória é mapeada pelo algoritmo VN em um grafo com uma distribuição de graus de Poisson (Lacasa; Luque; Ballesteros; Luque; Nuno, 2008) e pelo algoritmo VH em um grafo com uma distribuição de graus exponencial (Luque; Lacasa; Ballesteros; Luque, 2009), os gráficos de $P(k)$ sugerem que as quatro sequências analisadas se comportam como sequências quase aleatórias. No entanto, essas sequências não são geradas por um processo puramente aleatório, pois tal processo apresentaria $\Delta = 1$ e essas sequências possuem $\Delta \simeq 0.6 - 0.7$.

As Tabelas 5.1 e 5.3 revelam que, em todos os grafos, o tipo de conexão menos frequente é aquele entre números ímpares (ou seja, $m_{11} < m_{12}, m_{21}, m_{22}$); no entanto, esse tipo de conexão é mais prevalente nos grafos derivados de números felizes do que nos derivados de números naturais. Por exemplo, considerando o algoritmo VN, $m_{11, \{D(h)\}} = 0.111 > m_{11, \{D(n)\}} = 0.034$. Essa tabela também revela que conexões de números ímpares para números pares e conexões de números pares para números ímpares aparecem em percentagens semelhantes (ou seja, $m_{12} \simeq m_{21}$) em todos os grafos. Observe que as ligações entre números ímpares e entre números pares conectam nós não consecutivos nos grafos para os naturais; porém isso não é verdade nos grafos para os números felizes.

Portanto as quatro sequências de natureza puramente matemática foram mapeadas em grafos com $\langle l \rangle \gtrsim \langle l \rangle_{random}$ e $0 \ll \langle c \rangle < 1$, que são características de conectividade geralmente encontradas em redes do mundo real (Watts; Strogatz, 1998; Newman, 2003; Boccaletti; Latora; Moreno; Chavez; Hwanga, 2006). Assim, do ponto de vista topológico, redes matemáticas podem ser semelhantes a redes que representam sistemas biológicos, sociais e tecnológicos. Além disso, as distribuições de graus monotonamente decrescentes

obtidas pelo algoritmo de VH e as distribuições de graus em formato de sino obtidas pelo algoritmo VN podem ser resultados do mapeamento de sequências quase aleatórias. Este estudo também mostrou que os grafos $\{D\}$ são mais conectados do que os respectivos grafos de $\{T\}$; e números pares são mais conectados do que números ímpares.

No próximo capítulo, conclui-se esta tese, listando os resultados mais relevantes, os estudos realizados mas não relatados e possíveis continuações para este trabalho.

Capítulo 6

CONCLUSÃO

Estudos sobre divisores de números naturais fazem parte de um campo abrangente da ciência e a pesquisa sobre quantidade de divisores diferente de dois carece de investigações. A seguir, apresenta-se uma lista dos resultados considerados os mais relevantes desta tese.

Sobre a lacuna entre d -primos:

- $\Delta \simeq 0.7$ para d par e $\Delta \simeq 1$ para d ímpar; assim, a entropia normalizada distingue d par de d ímpar;
- nos grafos de VH, $\langle k \rangle > 3.9$ para d par e $\langle k \rangle < 3.9$ para d ímpar; assim, o grau médio dos grafos de visibilidade horizontal diferencia d par de d ímpar;
- para d -ímpar, $\langle k \rangle \simeq 4$ nos grafos de VH, o que é característica de sequências aperiódicas.

Sobre a densidade dos d -primos:

- o comportamento da curva da densidade para d ímpar e para d par são distintos, e dentro dos pares há uma divisão levando em conta a concavidade dessa curva;
- d par com a concavidade para cima é similar ao caso de $d = 2$, pois esses casos seguem uma distribuição assintótica do tipo $a/(\ln(n) + b)$;
- para d par, concavidade voltada para baixo ocorre em 62.5% dos casos aqui analisados.

Sobre a quantidade de divisores dos números naturais e felizes:

- as quatro sequências foram mapeadas em grafos com $\langle l \rangle \gtrsim \langle l \rangle_{random}$ e $0 \ll \langle c \rangle < 1$, que são características de conectividade geralmente encontradas em redes biológicas, sociais e tecnológicas;
- analisando o formato de $P(k)$, os grafos de VN têm formato de sino e os de VH decrescem monotonamente, tanto para os números naturais quanto para os números felizes;
- apesar de a quantidade de pares e ímpares serem iguais nos naturais e quase iguais nos felizes, em todos os grafos, as ligações entre ímpares são as menos frequentes.

6.1 Trabalhos feitos não relatados

Ao longo destes quatro anos de doutorado, vários trabalhos foram feitos sem, entretanto, se atingir um resultado interessante. A seguir, citam-se três desses trabalhos

Não foi relatado um estudo feito com dígitos de números irracionais. Esse estudo é similar àquele dos divisores dos naturais, porém usando os dígitos decimais dos números π , e , $\sqrt{2}$ e $\phi = (1 + \sqrt{5})/2$. Os dígitos consecutivos foram tomados formando grupos de tamanhos 1, 2 e 3. Os resultados não se mostraram interessantes, pois os dígitos decimais parecem ocorrer na mesma frequência relativa (ou seja, as quantidades de 0,1,2,...,9 são praticamente idênticas, quando se consideram os números depois da vírgula dos irracionais analisados).

No estudo das densidades de d -primos, foram consideradas outras funções, como $f_d(n) = a/(\ln(n) + bn + cn^2 + d)$. Essas, porém, levaram a ajustes piores que o descrito neste trabalho.

Foi feita uma comparação das lacunas entre os primos usuais com o tamanho das trajetórias da sequência de Collatz, também conhecida como sequência granizo. Essa sequência é gerada partindo de um número natural e calculando-se o próximo aplicando a seguinte regra: se for par, divida por 2; se for ímpar, multiplique por 3 e some 1. Conjectura-se que, a partir de qualquer número natural, converge-se para a sequência $4 \rightarrow 2 \rightarrow 1$, embora ainda não haja prova rigorosa dessa convergência. A intenção era analisar a estrutura dos grafos gerados pelos algoritmos de VN e VH, buscando por

similaridades com os grafos obtidos com os números naturais. Esse estudo, porém, foi apenas parcialmente desenvolvido.

6.2 Trabalhos futuros

Este trabalho abre a possibilidade de estudos sobre d -primos mais aprofundados. Alguns exemplos são: determinar a convergência das densidades, utilizar sequências maiores e considerar valores de d maiores. Além disso, estudar a relação entre a densidade e as lacunas pode trazer resultados importantes no estudo dos números primos. Ainda, d -primos talvez possam ter aplicações em criptografia.

Outra possibilidade de trabalho futuro é empregar os algoritmos de VN e VH em outras sequências que também são resultados de processos determinísticos. Por exemplo: as sequências geradas quando se escreve um número irracional y como uma fração contínua, ou seja, quando se escreve tal número da seguinte forma:

$$y = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots}}}$$

A sequência $\{a_0, a_1, a_2, \dots\}$ caracteriza y . Por exemplo, para o número ϕ , tem-se $a_j = 1$ para todo j .

Um outro possível estudo usando os algoritmos de visibilidade aplicados às lacunas entre d -primos é procurar os valores de máximo e mínimo de intervalos utilizando as propriedades desses grafos como, por exemplo, o caminho mínimo médio e o coeficiente de agregação. Ainda, nesta tese, os valores de A e γ para a sequência de divisores de naturais e felizes e dos d -primos foram determinados, porém não foi feito um estudo buscando relacioná-los.

Uma análise simbólica dos problemas tratados, visando obter relações analíticas por meio de demonstrações, é uma possível continuação para este trabalho, pois aumentaria o rigor dos resultados obtidos. Por exemplo, pode-se perguntar se existe uma fórmula explícita para as sequências de d -primos para d não primo. No caso dos d -primos, com d primo, como essas sequências são da forma x^k , pode ser conveniente utilizar cálculo diferencial para estudar as características de tais sequências. Como o algoritmo de VN se utiliza da inclinação da reta entre dois pontos, a determinação da derivada pode ser útil

para sequências com uma função geratriz explícita.

Pode-se também investigar as similaridades entre as sequências dos divisores e as sequências das trajetórias dos naturais com a sequência de Collatz; isto é, a quantidade de arestas percorridas até se chegar ao vértice 1, aplicando repetidamente a regra de Collatz. Por exemplo, partindo do número 10, obtém-se a seguinte sequência:

$$10 \rightarrow 5 \rightarrow 16 \rightarrow 8 \rightarrow 4 \rightarrow 2 \rightarrow 1$$

No exemplo acima, a trajetória de Collatz tem tamanho 7 e o número 10 tem 4 divisores. Partindo do número 11, a trajetória de Collatz tem tamanho 15 e o número 11 possui 2 divisores. A intenção é determinar se há alguma dependência entre a quantidade de divisores de um número e a trajetória de Collatz.

Outro trabalho possível diz respeito à contagem da quantidade de fatores primos em vez da quantidade de divisores, e analisar se essa abordagem traz novas informações (por exemplo, $12 = 2^2 \times 3$ tem dois fatores primos e é um 6-primo). Se os resultados forem diferentes, então tem-se duas operações similares que criam topologias estruturalmente diferentes. Se os resultados forem os mesmos, então essas duas operações, apesar de distintas, geram grafos com a mesma topologia. Esse trabalho poderia levar a resultados diferentes daqueles feitos até então sobre fatores primos.

Referências Bibliográficas

- APOSTOL, T. M. *Introduction to Analytic Number Theory*. New York: Springer, 1988.
- BAKER, R. C.; HARMAN, G.; PINTZ, J. The difference between consecutive primes. II. *Proceedings of the London Mathematical Society*, v. 83, p. 532–562, 2001.
- BERRY, M. V.; KEATING, J. P. The Riemann zeros and eigenvalue asymptotics. *SIAM Review*, v. 41, p. 236–266, 1999.
- BOCCALETTI, S.; LATORA, V.; MORENO, Y.; CHAVEZ, M.; HWANGA, D. U. Complex networks: structure and dynamics. *Physics Reports*, v. 424, p. 175–308, 2006.
- BROIDO, A. D.; CLAUSET, A. Scale-free networks are rare. *Nature Communications*, v. 10, n. ID 1017, 2019.
- BURTON, D. M. *Elementary Number Theory*. New York: McGraw Hill, 2012.
- CHANDRA, A. K.; DASGUPTA, S. A small world network of prime numbers. *Physica A*, v. 357, p. 436–446, 2005.
- CHEN, D. R.; LIU, C.; ZHANG, Y. C.; ZHANG, Z. K. Predicting financial extremes based on weighted visual graph of major stock indices. *Complexity*, v. 2019, n. ID 5320686, 2019.
- CORSO, G. Families and clustering in a natural numbers network. *Physical Review E*, v. 69, n. ID 036106, 2004.
- DIAMOND, H. G. Elementary methods in the study of the distribution of prime numbers. *Bulletin of the American Mathematical Society*, v. 7, p. 553–590, 1982.
- DICKSON, L. E. *History of the Theory of Numbers, Vol. 1: Divisibility and Primality*. New York: Dover, 2005.
- EL-SEDY, E.; SIKSEK, S. On happy numbers. *Rocky Mountain Journal of Mathematics*, v. 30, p. 565–570, 2000.
- ERDÖS, P. On the difference of consecutive primes. *The Quarterly Journal of Mathematics*, v. 6, p. 124–128, 1935.
- EVES, H.; KULIK, J. *An Introduction to the History of Mathematics*. Philadelphia: Saunders College, 1983.
- FORD, K.; GREEN, B.; KONYAGIN, S.; TAO, T. Large gaps between consecutive prime numbers. *Annals of Mathematics*, v. 183, p. 935–974, 2016.

- GARBI, G. G. *A Rainha das Ciências*. São Paulo: Livraria da Física, 2009.
- GUY, R. *Unsolved Problems in Number Theory*. New York: Springer, 2004.
- HARDY, G. H.; WRIGHT, E. M. *An Introduction to the Theory of Numbers*. Oxford: Oxford University Press, 2008.
- HOCHE, R. *Nicomachi Geraseni Pythagorei Introductionis arithmeticae libri II*. Leipzig: B.G. Teubner, 1866.
- IRELAND, K.; ROSEN, M. *A Classical Introduction to Modern Number Theory*. 2. ed. New York: Springer, 1998.
- KELLER, E. F. Revisiting “scale-free” networks. *BioEssays*, v. 27, p. 1060–1068, 2005.
- KNUTH, D. E. *The Art of Computer Programming*. Boston: Addison-Wesley, 1997.
- LACASA, L.; LUQUE, B.; BALLESTEROS, F.; LUQUE, J.; NUNO, J. C. From time series to complex networks: the visibility graph. *Proceedings of the National Academy of Sciences of the USA*, v. 105, p. 4972–4975, 2008.
- LAROCHE, P. *Mersenne Prime Discovery - $2^{82589933} - 1$ is Prime!* 2018. Disponível em: <https://www.mersenne.org/primes/?press=M82589933>.
- LIMA, E. L. *Curso de Análise*. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 1976.
- LJUNG, L. *System Identification: Theory for the User*. Upper Saddle River: Prentice-Hall, 1998.
- LUQUE, B.; LACASA, L.; BALLESTEROS, F.; LUQUE, J. Horizontal visibility graphs: exact results for random time series. *Physical Review E*, v. 80, n. ID 046103, 2009.
- MAURER, U. M. Fast generation of prime numbers and secure public-key cryptographic parameters. *Journal of Cryptology*, v. 8, p. 123–155, 1995.
- MAYER, B. L.; MONTEIRO, L. H. A. A numerical study on the regularity of d-primes via informational entropy and visibility algorithms. *Complexity*, v. 2020, n. ID 1480890, 2020.
- MAYER, B. L.; MONTEIRO, L. H. A. On the divisors of natural and happy numbers: a study based on entropy and graphs. *AIMS Mathematics*, v. 8, p. 13411–13424, 2023.
- MONTEIRO, L. H. A. *Sistemas Dinâmicos Complexos*. São Paulo: Livraria da Física, 2014.
- NETTO, V. M.; MEIRELLES, J.; RIBEIRO, F. L. Social interaction and the city: the effect of space on the reduction of entropy. *Complexity*, v. 2017, n. ID 6182503, 2017.
- NEWMAN, M. E. J. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the USA*, v. 98, p. 404–409, 2001.
- NEWMAN, M. E. J. The structure and function of complex networks. *SIAM Review*, v. 45, p. 167–256, 2003.

- NEWMAN, M. E. J. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, v. 46, p. 323–351, 2005.
- NUNEZ, A.; LACASA, L.; VALERO, E.; GOMEZ, J. P.; LUQUE, B. Detecting series periodicity with horizontal visibility graphs. *International Journal of Bifurcation and Chaos*, v. 22, n. ID 1250160, 2012.
- OEIS. *The On-Line Encyclopedia of Integer Sequences*. 2023. Publicado eletronicamente em <http://oeis.org> e fundado em 1964 por N. J. A. SLOANE.
- ORE, O. *Number Theory and Its History*. New York: Dover, 1988.
- PIQUEIRA, J. R. C.; SERBONCINI, F. A.; MONTEIRO, L. H. A. Biological models: measuring variability with classical and quantum information. *Journal of Theoretical Biology*, v. 242, p. 309–313, 2006.
- RIVEST, R. L.; SHAMIR, A.; ADLEMAN, L. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, v. 21, p. 120–126, 1978.
- SHANNON, C.; WEAVER, W. *The Mathematical Theory of Communication*. Chicago: University of Illinois Press, 1949.
- SHINER, J. S.; DAVISON, M.; LANDSBERG, P. T. Simple measure for complexity. *Physical Review E*, v. 59, p. 1459–1464, 1999.
- SONG, C. M.; HAVLIN, S.; MAKSE, H. A. Self-similarity of complex networks. *Nature*, v. 433, p. 392–395, 2005.
- STEIN, M. L.; ULAM, S. M.; WELLS, M. B. A visual display of some properties of the distribution of primes. *American Mathematical Monthly*, v. 71, p. 516–520, 1964.
- SZPIRO, G. G. The gaps between the gaps: some patterns in the prime number sequence. *Physica A*, v. 341, p. 607–617, 2004.
- SZPIRO, G. G. Peaks and gaps: spectral analysis of the intervals between prime numbers. *Physica A*, v. 384, p. 291–296, 2007.
- WANG, L.; LONG, X.; ARENDS, J. B. A. M.; AARTS, R. M. EEG analysis of seizure patterns using visibility graphs for detection of generalized seizures. *Journal of Neuroscience Methods*, v. 290, p. 85–94, 2017.
- WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘small-world’ networks. *Nature*, v. 393, p. 440–442, 1998.
- YOSHIMURA, J.; HAYASHI, T.; TANAKA, Y.; TAINAKA, K. I.; SIMON, C. Selection for prime-number intervals in a numerical model of periodical cicada evolution. *Evolution*, v. 63, p. 288–294, 2009.
- ZOU, Y.; DONNER, R. V.; MARWAN, N.; DONGES, J. F.; KURTHS, J. Complex network approaches to nonlinear time series analysis. *Physics Reports*, v. 787, p. 1–97, 2019.