Universidade Federal do Rio de Janeiro

Matrizes Aleatórias e Lei do Semicírculo de Wigner

Felipe Bottega Diniz

Rio de Janeiro

Julho de 2014

Universidade Federal do Rio de Janeiro

Matrizes Aleatórias e Lei do Semicírculo de Wigner

Felipe Bottega Diniz

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Matemática Aplicada, Instituto de Matemática da Universidade Federal do Rio de Janeiro (UFRJ), como parte dos requisitos necessários à obtenção do título de Mestre em Matemática Aplicada.

Orientador: Prof. Gregorio Malajovich.

Rio de Janeiro
Julho de 2014

Universidade Federal do Rio de Janeiro

Matrizes Aleatórias e Lei do Semicírculo de Wigner

Felipe Bottega Diniz

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Matemática, Instituto de Matemática da Universidade Federal do Rio de Janeiro (UFRJ), como parte dos requisitos necessários à obtenção do título de Mestre em Matemática.

Aprovada por:

_____

Presidente, Prof. Gregorio Malajovich Munoz

_____

Prof. Bernardo Freitas Paulo da Costa

_____

Prof. Glauco Valle da Silva Coelho

_____

Prof. Leandro Pinto Rodrigues Pimentel

Rio de Janeiro
Julho de 2014

CIP - Catalogação na Publicação

BD585m

Bottega Diniz, Felipe
Matrizes Aleat órias e Lei do Semic írculo de Wigner / Felipe Bottega Diniz. -- Rio de Janeiro, 2014.
55 f.

Orientador: Gregorio Malajovich.
Dissertação (mestrado) - Universidade Federal do Rio de Janeiro, Instituto de Matemática, Programa de Pós-Graduação em Matemática, 2014.

1. Matrizes Aleatórias. 2. Probabilidade. 3. Distribui ção de Autovalores. 4. Matrizes de Wigner. 5. Lei do Semicí rculo. I. Malajovich, Gregorio , orient. II. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a).

A Deus e ao universo.

Agradecimentos

A minha família, por acreditarem em mim e me apoiarem sempre.

Aos meus amigos, por terem sempre me incentivado e me ajudado nas dificuldades, que não foram poucas.

Ao professor Gregorio Malajovich, por toda a ajuda, pela paciência, pelos ótimos cursos e pelas conversas (sem estas conversas eu nunca iria conhecer o fascinante mundo das matrizes aleatórias).

Aos professores da banca: Bernardo Costa, Glauco Valle Coelho e Leandro Pimentel, pelos comentários, sugestões e animação que demonstraram em relação ao meu trabalho.

Aos meus colegas da matemática aplicada e da pura, que tiveram a boa vontade de conhecer e discutir um pouco sobre o meu trabalho.

A todos os professores que tive a sorte de conhecer, todos influenciaram o meu aprendizado de maneira significativa.

A todos os matemáticos e físicos que contribuíram para o crescimento desta magnífica área, que é a teoria de matrizes aleatórias.

Felipe Bottega Diniz
Julho de 2014

"Descobrimos o segredo que capacita os matemáticos a lidar com a matemática: um matemático é alguém para quem a matemática é uma série de televisão... Os "personagens" numa série de matemática não são gente, mas sim objetos matemáticos. Os fatos e as relações que são o foco de atenção não são nascimentos e mortes, casamentos, casos amorosos e relações de negócios, mas sim fatos e relações matemáticas sobre objetos matemáticos."

Keith Devlin - O Gene da Matemática
página 284

Resumo

Matrizes Aleatórias e Lei do Semicírculo de Wigner

Felipe Bottega Diniz

Resumo da dissertação de Mestrado submetida ao Programa de Pós-graduação em Matemática Aplicada, Instituto de Matemática da Universidade Federal do Rio de Janeiro (UFRJ), como parte dos requisitos necessários à obtenção do título de Mestre em Matemática Aplicada.

Resumo: A teoria de matrizes aleatórias é uma área nova da matemática, que apenas começou a ganhar destaque nos anos 1950, com os trabalhos de Wigner. Essa dissertação tem como objetivo apresentar os conceitos básicos da teoria de matrizes aleatórias, provar um dos teoremas clássicos da área - a Lei do Semicírculo de Wigner - e por fim, mostrar algumas aplicações que decorrem deste teorema.

Palavras–chave. Matrizes Aleatórias, Probabilidade, Distribuição de Autovalores, Matrizes de Wigner, Lei do Semicírculo.

Rio de Janeiro
Abril de 2014

Abstract

Random Matrices and the Wigner's Semicircle Law

Felipe Bottega Diniz

Abstract da dissertação de Mestrado submetida ao Programa de Pós-graduação em Matemática Aplicada, Instituto de Matemática da Universidade Federal do Rio de Janeiro (UFRJ), como parte dos requisitos necessários à obtenção do título de Mestre em Matemática Aplicada.

Abstract: The theory of of random matrices is a new area of mathematics, which only began to gain prominence in the 1950s with the work of Wigner. The present work aims to present the basics of random matrices theory, to prove one of the classic theorems of this area - the Wigner's Semicircle Law - and finally, show some applications stemming from this theorem.

Keywords. Random Matrices, Probability, Eigenvalues Distribution, Wigner Matrices, Semicircle Law.

Rio de Janeiro
Julho 2014

Sumário

1 Introdução **3**
1.1 História . . . 3
1.2 Matrizes aleatórias . . . 3
1.3 Ensembles . . . 7

2 Lei do Semicírculo de Wigner **11**
2.1 Matrizes de Wigner . . . 11
2.2 Método dos Momentos . . . 14
2.3 Palavras e Multigrafos . . . 18
2.4 Passeios de Bernoulli e Caminhos de Dyck . . . 25
2.5 Sentenças e Multigrafos . . . 29
2.6 Lei do Semicírculo de Wigner . . . 35
2.7 Resultados Extras . . . 39
2.8 Lei do Semicírculo de Wigner Hermitiano . . . 42

3 Aplicações **49**
3.1 Consequências Básicas . . . 49
3.2 Estimativas de Erro . . . 50
3.3 Determinante . . . 52

Capítulo 1

Introdução

1.1 História

Nos anos 1950, a teoria de matrizes aleatórias começou a ganhar destaque graças aos trabalhos do físico Eugene Wigner. O objetivo de Wigner era descrever as propriedades dos níveis de energia de núcleos pesados, quando estes eram excitados. Este sistema nuclear pode ser representado por H, um operador hermitiano num espaço de Hilbert, chamado de *Hamiltoniano*. De acordo com a mecânica quântica, os níveis de energia são representados pelos autovalores de H, de modo que os espaços entre os níveis de energia dos núcleos sejam dados pelos espaços entre autovalores consecutivos de H. O problema é que geralmente H não é conhecido. Mesmo que seja determinado, provavelmente será intratável, tanto do ponto de vista matemático quanto do computacional. Tentar computar seus autovalores diretamente não é a melhor opção. Wigner propôs que seria melhor considerar H como uma matriz aleatória, que estaria dentro de uma classe de matrizes aleatórias, chamada de *ensemble*. O ensemble seria tal que quase todos os seus membros teriam propriedades semelhantes às de H. Então podemos aproximar o sistema por um ensemble de matrizes aleatórias e analisar as estatísticas dos autovalores deste ensemble. Como H tem ordem infinita, podemos tentar determinar o comportamento assintótico dos autovalores de matrizes aleatórias, além de fazer aproximações finitas e analisar suas estatísticas. Mais detalhes sobre a história de matrizes aleatórias podem ser encontrados em [9], o principal artigo de Wigner está em [10].

No tempo de Wigner, a ideia de se trabalhar com matrizes aleatórias não era inteiramente nova, mas foi o seu trabalho que chamou atenção para a área. Desde então esta área cresceu ainda mais e se mostrou ser aplicável em uma variedade de situações distintas, como física nuclear, física de materiais, análise numérica e teoria dos números. A teoria de matrizes aleatórias nasceu a partir de aplicações, mas independente disto, é uma área matematicamente rica, com conexões em diversas outras áreas como combinatória enumerativa, análise clássica, álgebra linear, teoria dos números, probabilidade e teoria da probabilidade livre[1].

1.2 Matrizes aleatórias

Sempre que nos referirmos a alguma norma de matriz, estaremos considerando a norma de operador, definida por

$$\|\mathbf{X}\|_2 = \sup_{\substack{v\in\mathbb{C}^N \\ \|v\|_2=1}} \|\mathbf{X}v\|_2$$

onde $\|v\|_2$ é a norma usual em $\mathbb{C}^N$. Denotaremos por $N(m,\sigma^2)_{\mathbb{R}}$ a distribuição gaussiana real e por $N(m,\sigma^2)_{\mathbb{C}}$ a distribuição gaussiana complexa, onde $m \in \mathbb{C}$ é o valor esperado e $\sigma^2 > 0$ é a variância.

[1] Do inglês de "free probability", mas não encontrei tradução alguma para o português.

Definição 1.1. Para $1 \leq i, j \leq N$, sejam X_{ij} variáveis aleatórias, reais ou complexas, em um mesmo espaço de probabilidade $(\Omega, \mathcal{F}, P)$. A matriz $\mathbf{X}$ dada por $(\mathbf{X})_{ij} = X_{ij}$ é chamada de *matriz aleatória*. Então podemos escrever

$$\mathbf{X} = \begin{bmatrix} X_{11} & \dots & X_{1N} \\ \vdots & \ddots & \vdots \\ X_{N1} & \dots & X_{NN} \end{bmatrix}.$$

Fixadas as distribuições para os X_{ij}, estamos interessados no comportamento das matrizes geradas pela variável aleatória $\mathbf{X} : \Omega \to \mathcal{M}(\mathbb{R})$ no espaço das matrizes reais de ordem $N \times N$, dada por

$$\mathbf{X}(\omega) = \begin{bmatrix} X_{11}(\omega) & \dots & X_{1N}(\omega) \\ \vdots & \ddots & \vdots \\ X_{N1}(\omega) & \dots & X_{NN}(\omega) \end{bmatrix}.$$

Seja $\mathbf{A} = \mathbf{X}(\omega)$ uma dessas matrizes, suponha que seja simétrica, e sejam $\lambda_1(\mathbf{A}) \leq \ldots \leq \lambda_N(\mathbf{A})$ os seus autovalores.

Definição 1.2. Para cada autovalor $\lambda_i(\mathbf{A})$, temos que a função $\delta_{\lambda_i(\mathbf{A})} : \mathcal{L}(\mathbb{R}) \to [0, 1]$, dada por[2]

$$\delta_{\lambda_i(\mathbf{A})}\big(E\big) = \begin{cases} 1, \text{ se } \lambda_i(\mathbf{A}) \in E \\ 0, \text{ se } \lambda_i(\mathbf{A}) \notin E \end{cases}$$

é uma medida de probabilidade (chamada de *medida de Dirac*).

Definição 1.3. A partir da medida de Dirac, podemos definir a função de distribuição $\delta_{\lambda_i(\mathbf{A})} : \mathbb{R} :\to [0, 1]$, dada por

$$\delta_{\lambda_i(\mathbf{A})}(x) = \begin{cases} 1, \text{ se } x \leq \lambda_i(\mathbf{A}) \\ 0, \text{ se } x > \lambda_i(\mathbf{A}). \end{cases}$$

Definição 1.4. A partir das medidas $\delta_{\lambda_i(\mathbf{A})}$, podemos definir a medida $\mu_{\mathbf{A}} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(\mathbf{A})}$, que é uma medida de probabilidade aleatória (supondo que $\mathbf{A}$ foi obtida aleatoriamente).

Deste modo, temos que a matriz $\mathbf{A}$ induz o espaço de probabilidade $(\mathbb{R}, \mathcal{L}(\mathbb{R}), \mu_{\mathbf{A}})$.

Definição 1.5. No espaço de probabilidade dado acima, podemos definir a função de distribuição $\mu_{\mathbf{A}} : \mathbb{R} \to [0, 1]$, que também pode ser escrita como

$$\mu_{\mathbf{A}}(x) = \frac{\text{quantidade de autovalores } \leq x}{N}.$$

Esta distribuição é muito importante para a análise espectral de matrizes aleatórias, ela é chamada de *Distribuição Espectral Empírica*. Além do espaço de probabilidade citado acima, temos que cada matriz $\mathbf{A}$ induz um outro espaço de probabilidade $\big(U, \mathcal{P}(U), P\big)$, onde $U = \{\lambda_1(\mathbf{A}), \ldots, \lambda_N(\mathbf{A})\}$ é o espectro de $\mathbf{A}$ e $\mathcal{P}\big(U\big)$ é o conjunto das partes de U. Neste espaço, definimos a variável aleatória $\Lambda : U \to \mathbb{R}$, dada por $\Lambda(x) = x$, ela representa a escolha aleatória de um autovalor. Temos que $\mu_{\mathbf{A}}$ é a função de distribuição

[2] Estamos denotando por $\mathcal{L}(\mathbb{R})$ a σ-álgebra de Lebesgue na reta.

associada a este espaço, sendo que $\mu_{\mathbf{A}}(x) = P(\Lambda \leq x)$ é a probabilidade de que um autovalor de $\mathbf{A}$, escolhido aleatoriamente, seja menor ou igual a x. Então a probabilidade de que um autovalor escolhido aleatoriamente esteja no intervalo $(a, b]$ é dada por $\mu_{\mathbf{A}}(b) - \mu_{\mathbf{A}}(a) = \int_a^b d\mu_{\mathbf{A}}$, e é essa a informação que nos interessa, de modo que não precisamos realmente conhecer Λ para trabalhar com as probabilidades dos autovalores. Associado à uma matriz aleatória, temos também os seus autovalores aleatórios, que são variáveis aleatórias $\lambda_i : \Omega \to \mathbb{R}$ tal que $\lambda_i(\omega) = \lambda_i(\mathbf{X}(\omega))$, tal que $\lambda_i(\omega)$ é o i-ésimo autovalor de $\mathbf{X}(\omega)$.

Note que não é necessário tentar definir explicitamente o espaço amostral Ω, podemos trabalhar diretamente com os valores tomados pelos X_{ij} quando as distribuições são conhecidas. Além disso, são os X_{ij} que nos interessam mesmo. Por exemplo, se quisermos saber a probabilidade de X_{11} estar num certo intervalo, integramos a função de densidade de X_{11} neste intervalo. Podemos fazer isso sem haver necessidade alguma de mencionar o espaço amostral. Neste contexto, a distribuição empírica faz o papel de medida de probabilidade enquanto que algum conjunto especificado de matrizes faz o papel de espaço amostral. Classificamos as matrizes aleatórias pelo do tipo de distribuição dos X_{ij} e pelo conjunto de matrizes considerado. Fixado o par (Conjunto de Matrizes**,** Distribuição), temos o que chamamos de *ensemble de matrizes aleatórias*. Se o conjunto de matrizes faz o papel de espaço amostral e a distribuição faz o papel de medida de probabilidade, então o ensemble faz o papel de espaço de probabilidade. Ele descreve como um sistema funciona, probabilisticamente.[3]

Exemplo 1.6. Para mostrar como os conceitos definidos aqui podem ser usados, considere uma matriz aleatória $\mathbf{X} : \Omega \to S_2(\mathbb{R})$, onde $S_2(\mathbb{R})$ é o espaço das matrizes simétricas reais 2×2. Podemos escrever

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{bmatrix}.$$

Considere também que as entradas X_{11}, X_{12}, X_{22} são i.i.d. com distribuição $N(0,1)_{\mathbb{R}}$. Poderíamos usar esta matriz aleatória para modelar algum fenômeno de interesse, cujo comportamento seja mais acessível estatisticamente que deterministicamente.

Podemos calcular diretamente os autovalores $\lambda_1 \leq \lambda_2$ de $\mathbf{X}$, obtendo

$$\lambda_1 = \frac{1}{2}\Big(X_{11} + X_{22} - \sqrt{X_{11}^2 + 4X_{12}^2 - 2X_{11}X_{22} + X_{22}^2}\Big)$$

e

$$\lambda_2 = \frac{1}{2}\Big(X_{11} + X_{22} + \sqrt{X_{11}^2 + 4X_{12}^2 - 2X_{11}X_{22} + X_{22}^2}\Big).$$

Então os autovalores esperados são dados por $E(\lambda_1) = E\Big(\frac{1}{2}\big(X_{11}+X_{22}-\sqrt{X_{11}^2 + 4X_{12}^2 - 2X_{11}X_{22} + X_{22}^2}\big)\Big)$ e $E(\lambda_2) = E\Big(\frac{1}{2}\big(X_{11} + X_{22} + \sqrt{X_{11}^2 + 4X_{12}^2 - 2X_{11}X_{22} + X_{22}^2}\big)\Big)$. Obter uma expressão exata para estes valores pode ser complicado, mas podemos obter boas aproximações numéricas. Fazendo isso, obtemos $E(\lambda_1) \approx -1.07766$ e $E(\lambda_2) \approx 1.07766$. Note que também podemos determinar facilmente o valor esperado do determinante desta matriz.

$$\det \mathbf{X} = X_{11}X_{22} - X_{12}^2 \implies E(\det \mathbf{X}) = E(X_{11}X_{22} - X_{12}^2) = EX_{11}EX_{22} - E(X_{12}^2) = -1$$

Mostramos a seguir os resultados de experimentos computacionais, onde geramos diversas matrizes aleatoriamente, de acordo com a definição de $\mathbf{X}$. O parâmetro t é a quantidade de amostras tomadas, a linha vermelha é o valor teórico, e os círculos são os valores empíricos, obtidos pela média das t amostras.

[3] Ensemble é a formalização da ideia de repetirmos algum experimento diversas vezes para observar os resultados em larga escala. Este termo é mais usado na física no lugar de "espaço de probabilidade", podendo ser usado para outros sistemas, além do das matrizes.

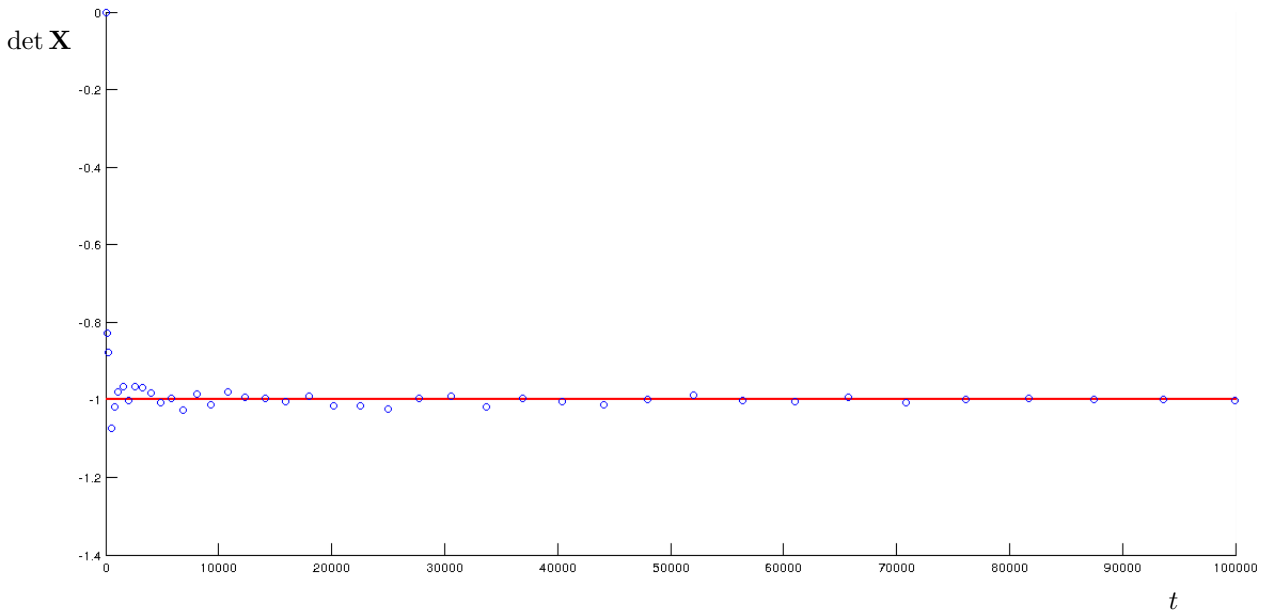

Figura 1.1: Cada círculo representa a média dos determinantes de t matrizes simétricas, 2×2, tomadas aleatoriamente, de acordo com a distribuição normal em cada entrada. Quanto mais amostras são tomadas, mais fica claro que o determinante esperado converge para -1.

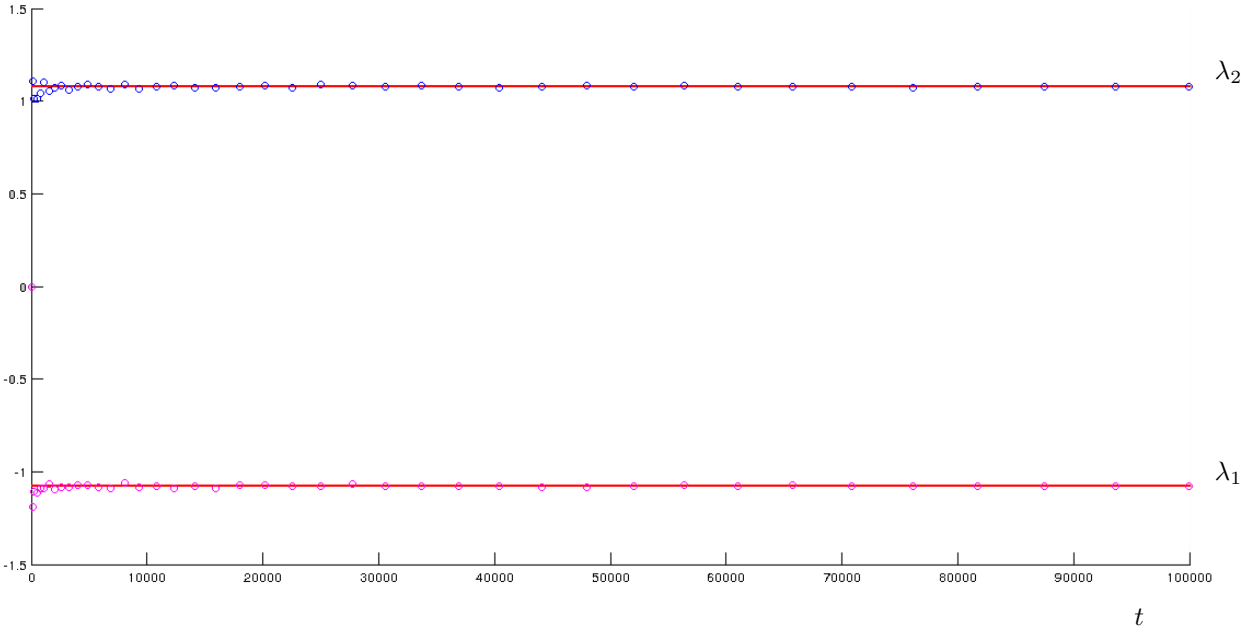

Figura 1.2: Cada círculo representa a média dos autovalores λ_1 (em magenta) e λ_2 (em azul) de t matrizes simétricas, 2×2, tomadas aleatoriamente, de acordo com a distribuição normal padrão em cada entrada.

Exemplo 1.7. Há um resultado bem interessante, que será apenas comentado aqui, a título de curiosidade. Considere que $\mathbf{X}$ é uma matriz aleatória $N \times N$, em que todas as suas entradas são i.i.d. com distribuição $N(0,1)_{\mathbb{R}}$. É esperado que $\mathbf{X}$ tenha mais autovalores estritamente complexos que reais, ainda assim, a proporção entre as quantidades não é evidente. Esse problema é abordado e resolvido em [12], com uma solução elegante. Outros resultados bem interessantes são mostrados neste artigo, mas mostraremos apenas um dos principais. Denote por E_N o número esperado de autovalores reais de $\mathbf{X}$, então temos que

$$E_N = \sqrt{2} \sum_{k=0}^{\frac{N}{2}-1} \frac{(4k-1)!!}{(4k)!!},$$

se k for par, e

$$E_N = 1 + \sqrt{2} \sum_{k=1}^{\frac{N-1}{2}} \frac{(4k-3)!!}{(4k-2)!!},$$

se k for ímpar.

A notação " !! " é para o duplo fatorial (ou semi fatorial), que é definido por

$$k!! = \begin{cases} 1 \cdot 3 \cdot 5 \cdot \ldots \cdot k, & \text{se } k \text{ é ímpar} \\ 2 \cdot 4 \cdot 6 \cdot \ldots \cdot k, & \text{se } k \text{ é par} \\ 1, & \text{se } k = 0 \text{ ou } k = -1. \end{cases}$$

A tabela abaixo mostra uma comparação entre os valores empíricos e teóricos, os resultados são bastante precisos.

N	Valor empírico	E_N	Erro relativo
1	1	1	0
2	1.33200	1.41421	0.00581
3	1.68400	1.70711	0.01353
4	1.94600	1.94454	0.00075
5	2.0940	2.14905	0.02561
10	2.96200	2.92799	0.01161
50	6.05800	6.09953	0.00680
255	13.17800	13.22245	0.00336
511	18.60600	18.52319	0.00447
1000	25.6000	25.72186	0.00473

Tabela 1.1: Cada valor empírico foi obtido através de 1000 matrizes $N \times N$ geradas aleatoriamente.

1.3 Ensembles

Na teoria de matrizes aleatórias, escolhemos algumas propriedades para as distribuições das entradas e para as matrizes em si, com isso se tem um ensemble. A partir daí, analisamos as estatísticas do ensemble, e se possível, tentamos generalizar os resultados obtidos para mais ensembles. Definimos a seguir, alguns dos principais ensembles de matrizes aleatórias.

Ensemble i.i.d. São matrizes aleatórias em que todas as entradas X_{ij} são i.i.d..

Um tipo de ensemble i.i.d. importante é o *ensemble gaussiano real*, em que cada X_{ij} é real e possui distribuição normal padrão $N(0,1)_{\mathbb{R}}$, de modo análogo temos o *ensemble gaussiano complexo*, onde cada entrada possui distribuição $N(0,1)_{\mathbb{C}}$.

Ensemble de Wigner. São matrizes aleatórias hermitianas em que as entradas da diagonal são i.i.d. com média 0, e as entradas acima da diagonal são i.i.d, com média 0 e variância 1. Além disso, todos os momentos de todas as entradas são finitos.

No ensemble de Wigner, ainda temos temos outros ensembles importantes a considerar.

Ensemble de Wigner Simétrico. Ensembles de Wigner onde os X_{ij} são reais.

Ensemble de Wigner Hermitiano. Ensembles de Wigner onde X_{ij} são complexos.

Ensemble Gaussiano de Wigner. Ensembles de Wigner onde os X_{ij} tem distribuição gaussiana.

Ensemble Gaussiano Ortogonal. Conhecido como GOE, de *gaussian orthogonal ensemble*. Ele faz parte do ensemble de Wigner simétrico e do gaussiano, além disso, a diagonal possui distribuição $N(0,2)_{\mathbb{R}}$ e o restante das entradas possui distribuição $N(0,1)_{\mathbb{R}}$.

Ensemble Gaussiano Unitário. Conhecido como GUE, de *gaussian unitary ensemble*. Ele faz parte do ensemble de Wigner hermitiano e do gaussiano, além disso, a diagonal possui distribuição $N(0,1)_{\mathbb{R}}$ e o restante possui distribuição $N(0,1)_{\mathbb{C}}$.

A figura 1.3 mostra um esquema com alguns dos principais ensembles de matrizes aleatórias.

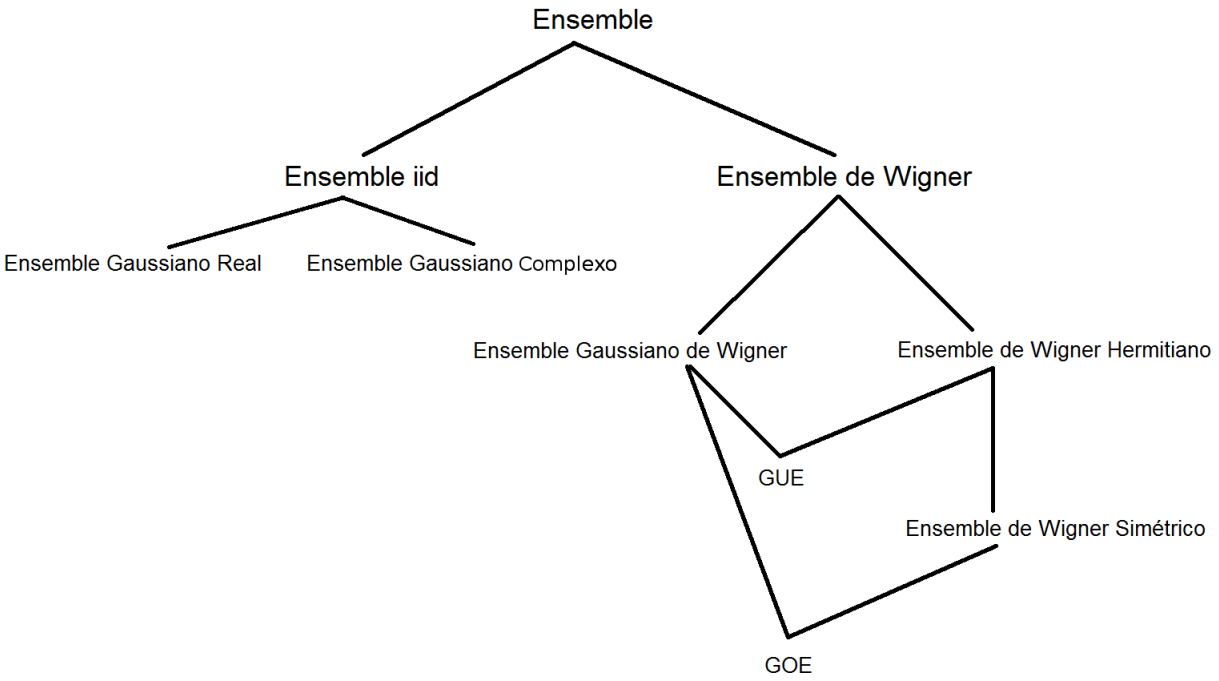

Figura 1.3: Alguns dos principais ensembles de matrizes aleatórias, iremos nos focar nos ensembles de Wigner.

Temos então o conceito de ensemble, que nos permite construir e estudar matrizes aleatórias. Será que podemos perceber algum padrão nestas matrizes aleatórias? Algumas perguntas relevantes podem ser feitas:

Como se distribuem os autovalores destas matrizes?
Há alguma diferença relevante na distribuição dos autovalores entre matrizes aleatórias reais e complexas?
Como é o histograma dos autovalores?
Quando aumentamos indefinidamente o tamanho das matrizes, qual a forma assintótica do histograma?
Qual a quantidade de autovalores reais esperados?

Também é de interesse considerar normas, determinantes, valores singulares, número de condicionamento, etc, de matrizes aleatórias. Nós iremos nos focar no ensemble de Wigner, afim de analisar a distribuição dos autovalores.

Capítulo 2

Lei do Semicírculo de Wigner

2.1 Matrizes de Wigner

Em todo este capítulo, iremos considerar apenas matrizes do ensemble de Wigner. Se $\mathbf{X}$ tem ordem $N \times N$, iremos dizer que $\frac{1}{\sqrt{N}}\mathbf{X}$ é a matriz $\mathbf{X}$ *normalizada*. Isto se deve ao fato de que as matrizes do ensemble de Wigner tipicamente tem norma de tamanho $O(\sqrt{N})$, de modo que seus autovalores também tem tamanho $O(\sqrt{N})$, então dividir as matrizes por $\sqrt{N}$ é uma escolha natural, no intuito de analisar as estatísticas dos autovalores em um espaço limitado. Não iremos provar este resultado, mas ele pode ser encontrado em [2].

Definição 2.1. Seja $\mathbf{X}$ uma matriz do ensemble de Wigner Simétrico, de ordem $N \times N$, então a matriz normalizada $\frac{1}{\sqrt{N}}\mathbf{X}$ é chamada de *matriz de Wigner*.

Sempre que nos referirmos a uma matriz aleatória $\mathbf{X}$, considere que ela é uma matriz do ensemble de Wigner Simétrico, mas não necessariamente normalizada, a normalização será tratada de modo explícito.

Como estamos interessados no comportamento assintótico das estatísticas dos autovalores, poderíamos simplesmente ir gerando matrizes de ordem $N \times N$ grande o suficiente e fazer o histograma dos seus autovalores. Teríamos então um gráfico que nos daria uma aproximação da função densidade dos autovalores desta matriz, cuja função distribuição empírica associada é a distribuição espectral empírica. Por esse motivo, iremos trabalhar constantemente com a distribuição espectral empírica afim de obter resultados a respeito das matrizes de Wigner.

Através de experimentos computacionais, geramos aleatoriamente várias matrizes de Wigner, com a mesma ordem, e então plotamos os histogramas normalizados dos seus autovalores. Estes histogramas são uma aproximação para a função densidade associada à distribuição espectral empírica. Abaixo, mostramos quatro histogramas que obtivemos pelos experimentos feitos.

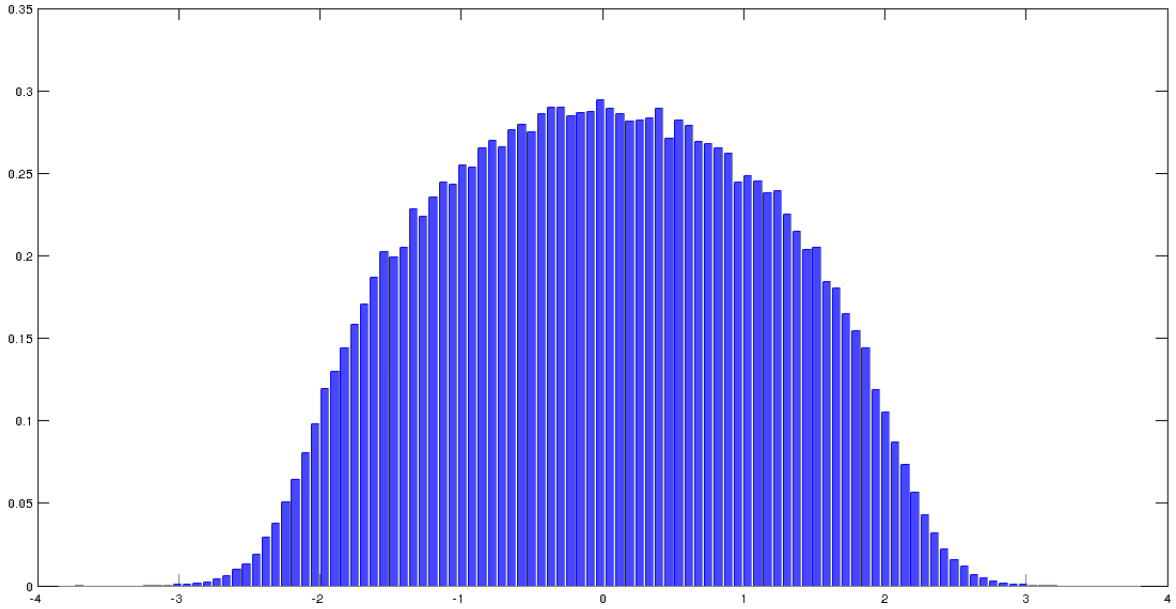

Figura 2.1: 20000 matrizes de ordem 10×10

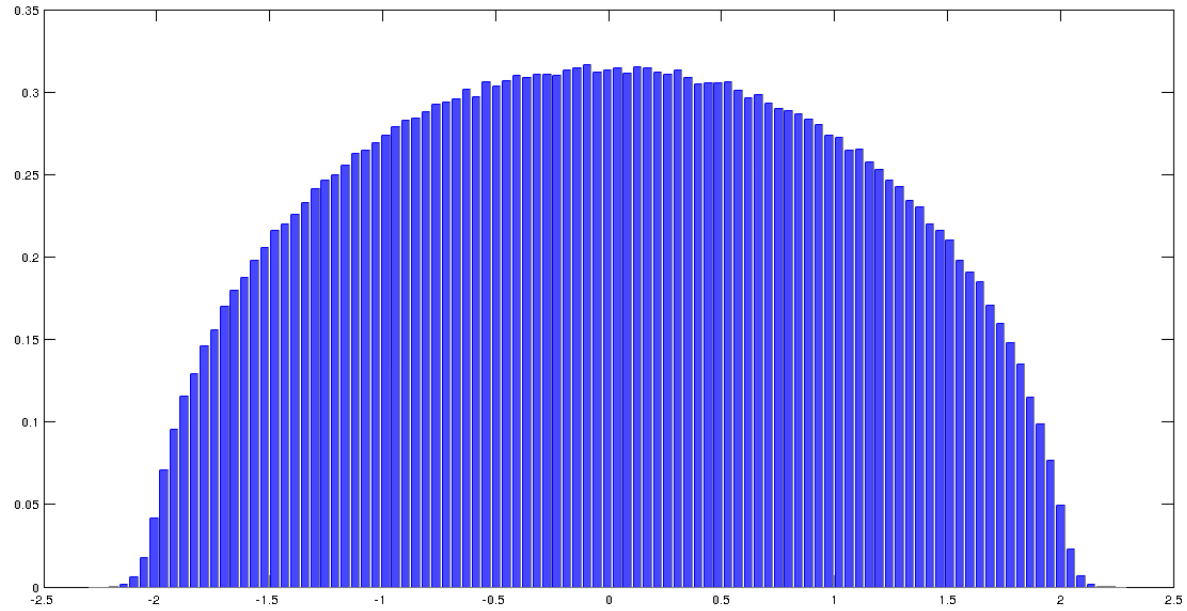

Figura 2.2: 10000 matrizes de ordem 100×100

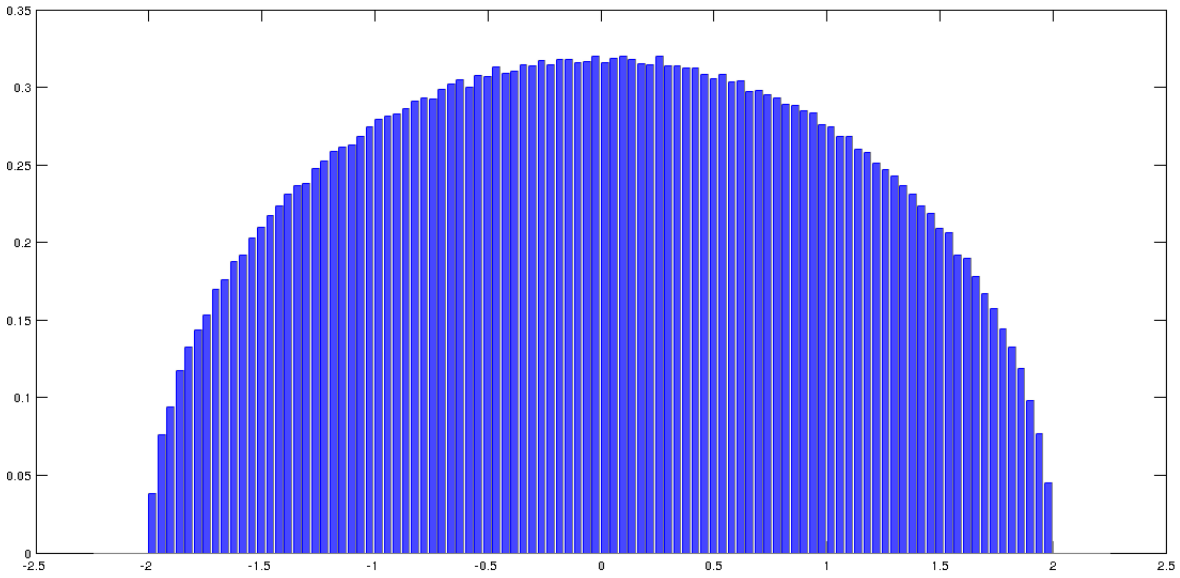

Figura 2.3: 10 matrizes de ordem 5000×5000

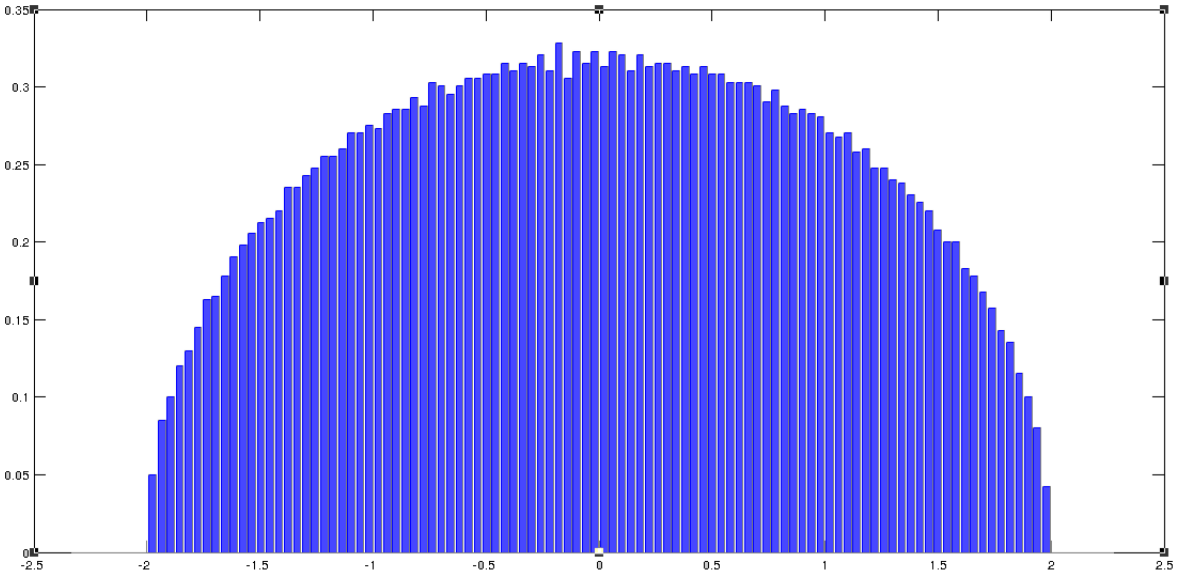

Figura 2.4: 1 matriz de ordem 10000×10000

Todos os histogramas foram gerados no MATLAB, através do código mostrado abaixo (variando os parâmetros de acordo). A variável t é a quantidade de vezes que o experimento foi repetido (de modo equivalente, é a quantidade de matrizes que foram geradas) e N é a ordem da matriz.

```
t=10000;
N=100;
E=zeros(0);
for i=1:t
    a= triu(random('Normal', 0, 1, N));
    a=(1/sqrt(N))*(a+a');
    E=[E,(eig(a))'];
end
[n,x]=hist(E,150);
delta = x(2)-x(1);
bar(x, n/(t*N*delta))
```

Pelo que podemos ver, o histograma tende para o formato de um semicírculo quando aumentamos a ordem das matrizes, ainda mais, se a ordem for alta o suficiente, o histograma dos autovalores de uma única matriz já terá o formato bem próximo ao de um semicírculo. Isso sugere que podemos prever como os autovalores se distribuem, ainda que as entradas sejam aleatórias. Veremos adiante que este fenômeno sempre ocorre quando usamos matrizes de Wigner, mais precisamente, temos o seguinte teorema.

Lei do Semicírculo de Wigner. Dada uma sequência aleatória de distribuições espectrais empíricas μ_N, associadas a matrizes de ordem $N \times N$ do ensemble de Wigner, temos que μ_N converge fracamente, em probabilidade, para μ, onde μ é uma distribuição de probabilidade cuja densidade é $\mu'(x) = \frac{1}{2\pi}\sqrt{4-x^2}\ \mathbf{I}_{|x|\leq 2}$.

2.2 Método dos Momentos

O método dos momentos permite obter informações da distribuição de uma variável aleatória através dos seus momentos. Para trabalhar com momentos, usaremos a notação $\langle \mu, f\rangle = \int_{-\infty}^{\infty} f(x)\, d\mu$, onde μ é uma medida em $\mathbb{R}$ e f é uma função real μ-integrável. Seja $S(\mathbb{R})$ o espaço das matrizes simétricas reais e seja $\Pr(\mathbb{R})$ o espaço das medidas de probabilidades em $\mathbb{R}$, agora considere $\Gamma : S(\mathbb{R}) \to \Pr(\mathbb{R})$ dada por $\Gamma(M_N) = \mu_{M_N}$, onde M_N é uma matriz ordem $N \times N$ escolhida aleatoriamente e μ_{M_N} é a distribuição espectral empírica associada à M_N. Temos que Γ é uma variável aleatória em $\Pr(\mathbb{R})$ e cada $\Gamma(M_N)$ é uma medida de probabilidade aleatória, ou seja, cada matriz aleatória está associada a uma probabilidade aleatória. A nossa ideia é tomar uma sequência aleatória de distribuições espectrais empíricas μ_{M_N} e analisar o que acontece quando $N \to \infty$.

Mais para frente, iremos ver que $\langle \mu_{M_N}, x\rangle$ é a média dos autovalores de M_N. De modo que podemos considerar a variável aleatória[1] $\omega \mapsto \langle \mu_{\mathbf{X}(\omega)}, x\rangle = \frac{1}{N}\mathrm{tr}\big(\mathbf{X}(\omega)\big)$, onde $\mathbf{X}(\omega)$ tem ordem $N \times N$, iremos denotá-la simplesmente por $\langle \mu_N, x\rangle$. De maneira mais geral, podemos considerar $\langle \mu_N, x^k\rangle$ para todo $k \in \mathbb{N}$. Temos que $\langle \mu_N, x^k\rangle$ nos fornece a média dos autovalores de $\mathbf{X}^k$. Os momentos $E\langle \mu_N, x^k\rangle$ nos dão o valor esperado da média dos autovalores de $\mathbf{X}^k$ sobre todo o espaço de matrizes considerado. Iremos usar estes momentos para obter informações sobre a distribuição do autovalores de $\mathbf{X}$.

[1] Podemos considerar também $M_N \mapsto \langle \mu_{M_N}, x\rangle = \frac{1}{N}\mathrm{tr}\big(M_N\big)$ como uma variável aleatória em $S(\mathbb{R})$. Em geral, considerar as coisas deste modo é mais conveniente.

Agora iremos dar algumas definições que nos serão úteis no decorrer do texto.

Definição 2.2. Seja X um conjunto, denotaremos por $C_b(X)$ o conjunto das funções $f : X \to \mathbb{R}$ que são contínuas e limitadas.

Definição 2.3. Seja X_N uma sequência de variáveis aleatórias no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ e X uma outra variável aleatória neste mesmo espaço. Seja $A = \{\omega \in \Omega : \lim_{N\to\infty} X_N(\omega) = X(\omega)\}$, se $P(A) = 1$, então dizemos que a sequência X_N converge *quase certamente* para X, e denotamos $X_N \xrightarrow{qc} X$.

Definição 2.4. Seja X_N uma sequência de variáveis aleatórias no espaço de probabilidade $(\Omega, \mathcal{F}, P)$ e X uma outra variável aleatória neste mesmo espaço. Se $\lim_{N\to\infty} P(|X_N - X| > \varepsilon) = 0$ para todo $\varepsilon > 0$, então dizemos que a sequência X_N converge *em probabilidade* para X, e denotamos $X_N \xrightarrow{P} X$.

Antes de começar a trabalhar com matrizes aleatórias, iremos introduzir a noção de convergência fraca. Comecemos considerando um espaço métrico (X, d) e a σ-álgebra de Borel $\mathcal{B}(X)$ em X induzida por d. Suponha também que (X, d) seja separável, ou seja, existe um subconjunto enumerável $A \subset X$ que é denso em X.

Definição 2.5. Seja $\mu : \mathcal{B}(X) \to [0, \infty)$ uma medida finita de Borel em X, tal que, para todo $\varepsilon > 0$ existe um compacto $K \subset X$ satisfazendo $\mu(X) - \mu(K) \leq \varepsilon$. Então μ é chamada de uma *medida de Radon* em X.

Definição 2.6. Seja μ_N uma sequência de medidas finitas de Borel em X e seja μ também uma medida finita de Borel em X, dizemos que μ_N converge fracamente para μ se $\lim_{N\to\infty} \int_X f \, d\mu_N = \int_X f \, d\mu$, para toda função $f \in C_b(X)$.

Dado $A \subset X$, usaremos as notações $\overline{A}$ para o fecho de A, $\overset{\circ}{A}$ para o interior de A e $\partial A = \overline{A} - \overset{\circ}{A}$ para o bordo de A. Então temos o seguinte teorema, que iremos apenas enunciar (sua prova pode ser encontrada em [13]).

Teorema 2.7. Seja μ_N uma sequência de medidas finitas de Borel em X e suponha que esta sequência converge fracamente para μ. Então $\lim_{N\to\infty} \mu_N(A) = \mu(A)$, para todo $A \in \mathcal{B}(X)$ com $\mu(\partial A) = 0$.

Este teorema nos mostra que a convergência fraca coincide com a convergência usual quando lidamos com conjuntos de bordo com medida nula. Apesar desta convergência ser um pouco restrita, na prática ela é o suficiente para os nossos propósitos.

Como estamos interessados no caso em que $X = \mathbb{R}$, temos que $(\mathbb{R}, |\ |)$ é um espaço métrico separável, além disso, toda medida finita de Borel em $\mathbb{R}$ é uma medida de Radon (em particular, toda distribuição espectral empírica é uma medida de Radon). Restringindo nossa atenção às medidas em $\Pr(\mathbb{R})$, temos que o teorema acima é um importante resultado a respeito da convergência de medidas de probabilidade em $\mathbb{R}$, pois ele nos diz em quais tipos de conjuntos podemos esperar uma convergência mais forte. No contexto de matrizes aleatórias, dizemos que uma sequência μ_N de distribuições espectrais empíricas converge fracamente para $\mu \in \Pr(\mathbb{R})$ se $\lim_{N\to\infty} \langle \mu_N, f \rangle = \langle \mu, f \rangle$, para toda função $f \in C_b(\mathbb{R})$. Neste texto, iremos dar mais foco para os dois modos de convergência definidos abaixo.

Definição 2.8. Seja μ_N uma sequência de distribuições espectrais empíricas, onde cada μ_N está associada a uma matriz de ordem $N \times N$. Se $\langle \mu_N, f \rangle \xrightarrow{P} \langle \mu, f \rangle$, para toda função $f \in C_b(\mathbb{R})$, dizemos que μ_N converge fracamente, em probabilidade, para $\mu \in \Pr(\mathbb{R})$.

Definição 2.9. Seja μ_N uma sequência de distribuições espectrais empíricas, onde cada μ_N está associada a uma matriz de ordem $N \times N$. Se $\langle \mu_N, f \rangle \xrightarrow{qc} \langle \mu, f \rangle$, para toda função $f \in C_b(\mathbb{R})$, dizemos que μ_N converge fracamente, quase certamente, para $\mu \in \Pr(\mathbb{R})$.

Nos resultados que se seguem, faremos uso de um número muito importante em combinatória.

Definição 2.10. Para $k \in \mathbb{N}$, o valor $\binom{2k}{k} \cdot \frac{1}{k+1}$ é chamado de *número de Catalan*, será denotado por C_k.

Antes de provar a Lei do Semicírculo de Wigner, precisamos provar dois lemas.

Lema 2.11. Seja μ_N uma sequência de distribuições espectrais empíricas de matrizes do ensemble de Wigner, de ordem $N \times N$, e seja $k \in \mathbb{N}$, então

$$\lim_{N\to\infty} E\langle \mu_N, x^k\rangle = \begin{cases} C_{\frac{k}{2}}, \text{ se } k \text{ é par} \\ 0, \text{ se } k \text{ é ímpar.} \end{cases}$$

Lema 2.12. Seja μ_N uma sequência de distribuições espectrais empíricas de matrizes do ensemble de Wigner, de ordem $N \times N$, então para todo $k \in \mathbb{N}$ e $\varepsilon > 0$, temos que

$$\lim_{N\to\infty} P\Big(\big|\langle \mu_N, x^k\rangle - E\langle \mu_N, x^k\rangle\big| > \varepsilon\Big) = 0.$$

Para provar estes lemas, precisamos de alguns resultados preliminares. Comecemos com um resultado importante na teoria de matrizes aleatórias, que enunciaremos sem prová-lo.

Proposição 2.13. Dada uma matriz $\mathbf{X}(\omega)$, de ordem $N \times N$, com autovalores $\lambda_1 \leq \ldots \leq \lambda_N$, seja μ_N a sua distribuição espectral empírica, então $\langle \mu_N, f\rangle = \frac{1}{N}\sum_{i=1}^N f(\lambda_i)$ para toda função $f \in C_b(\mathbb{R})$.

Corolário 2.14. Dada uma matriz $\mathbf{X}(\omega)$, de ordem $N \times N$, seja μ_N a sua distribuição espectral empírica. Então $\langle \mu_N, x^k\rangle = \frac{1}{N}\mathrm{tr}(\mathbf{X}^k(\omega))$.

Prova: Sejam $\lambda_1 \leq \ldots \leq \lambda_N$ os autovalores de $\mathbf{X}(\omega)$, então os autovalores de $\mathbf{X}^k(\omega)$ são $\lambda_1^k, \ldots, \lambda_N^k$. Como o traço de uma matriz é a soma dos seus autovalores, temos que $\mathrm{tr}(\mathbf{X}^k(\omega)) = \sum_{i=1}^N \lambda_i^k$. Por outro lado, a Proposição 2.13 nos garante que $\langle \mu_N, x^k\rangle = \frac{1}{N}\sum_{i=1}^N \lambda_i^k$, daí temos que $\langle \mu_N, x^k\rangle = \frac{1}{N}\mathrm{tr}(\mathbf{X}^k(\omega))$. □

Esta proposição pode ser naturalmente estendida para matrizes aleatórias, de modo que também podemos escrever $\langle \mu_N, x^k\rangle = \frac{1}{N}\mathrm{tr}(\mathbf{X}^k)$. No caso de termos uma matriz normalizada $\frac{1}{\sqrt{N}}\mathbf{X}$, temos que $\langle \mu_N, x^k\rangle = \frac{1}{N^{\frac{k}{2}+1}}\mathrm{tr}(\mathbf{X}^k)$. Então tudo que nos resta fazer é procurar um modo de reescrever $\mathrm{tr}(\mathbf{X}^k)$ e então calcular $E\Big(\frac{1}{N^{\frac{k}{2}+1}}\mathrm{tr}(\mathbf{X}^k)\Big)$. Começamos com a seguinte proposição.

Proposição 2.15. Seja $\mathbf{X}$ uma matriz $N \times N$, então para todo $k \in \mathbb{N}$, com $k \geq 2$,

$$\mathrm{tr}(\mathbf{X}^k) = \sum_{i_1,\ldots,i_k=1}^N X_{i_1,i_2}X_{i_2,i_3}\ldots X_{i_{k-1},i_k}X_{i_k,i_1}.$$

Prova: Primeiramente, vamos provar que $(\mathbf{X}^k)_{i,j} = \sum_{i_1,\ldots,i_{k-1}=1}^N X_{i,i_1}X_{i_1,i_2}\ldots X_{i_{k-2},i_{k-1}}X_{i_{k-1},j}$ para todo $k \geq 2$. Faremos a prova por indução em k. Para $k = 2$, temos que

$$(\mathbf{X}^2)_{i,j} = \sum_{l=1}^N X_{i,l}X_{l,j} = \sum_{i_1=1}^N X_{i,i_1}X_{i_1,j}.$$

Então vale o resultado para $k = 2$. Suponha que ele é válido um certo k, ou seja, $(\mathbf{X}^k)_{i,j} = \sum_{i_1,\ldots,i_{k-1}=1}^{N} X_{i,i_1}X_{i_1,i_2}\ldots X_{i_{k-2},i_{k-1}}X_{i_{k-1},j}$. Agora consideramos o caso $k+1$, então temos que

$$(\mathbf{X}^{k+1})_{i,j} = (\mathbf{X}^k \cdot \mathbf{X})_{i,j} = \sum_{l=1}^{N}(\mathbf{X}^k)_{i,l}X_{l,j} = \sum_{l=1}^{N}\left(\sum_{i_1,\ldots,i_{k-1}=1}^{N} X_{i,i_1}X_{i_1,i_2}\ldots X_{i_{k-2},i_{k-1}}X_{i_{k-1},l}\right)X_{l,j}.$$

Escrevendo $l = i_k$, obtemos

$$(\mathbf{X}^{k+1})_{i,j} = \sum_{i_1,\ldots,i_k=1}^{N} X_{i,i_1}X_{i_1,i_2}\ldots X_{i_{k-1},i_k}X_{i_k,j}.$$

Portanto, o resultado vale para todo $k \geq 2$, pelo princípio de indução.
Agora podemos calcular diretamente o traço de $\mathbf{X}^k$.

$$\mathrm{tr}(\mathbf{X}^k) = \sum_{i=1}^{N}(\mathbf{X}^k)_{i,i} = \sum_{i=1}^{N}\left(\sum_{i_1,\ldots,i_{k-1}=1}^{N} X_{i,i_1}X_{i_1,i_2}\ldots X_{i_{k-2},i_{k-1}}X_{i_{k-1},i}\right)$$

Podemos reescrever os índices deste somatório, de modo a obtermos

$$\mathrm{tr}(\mathbf{X}^k) = \sum_{i_1,\ldots,i_k=1}^{N} X_{i_1,i_2}X_{i_2,i_3}\ldots X_{i_{k-1},i_k}X_{i_k,i_1}$$

como queríamos demonstrar. □

Os resultados acima são bastante úteis no sentido de simplificar $E\langle\mu_N, x^k\rangle$, que é o que estamos interessados em calcular.

Exemplo 2.16. Suponha uma matriz $\mathbf{X}$ de ordem $N \times N$, tal que $N = 3$, e seja $k = 2$, daí temos que

$$\langle\mu_N, x^k\rangle = \frac{1}{3^{\frac{2}{2}+1}}\mathrm{tr}(\mathbf{X}^2) = \frac{1}{9}\sum_{i,j=1}^{3} X_{i,j}X_{j,i} =$$

$$= \frac{1}{9}\big(X_{1,1}^2 + X_{1,2}X_{2,1} + X_{1,3}X_{3,1} + X_{2,1}X_{1,2} + X_{2,2}^2 + X_{2,3}X_{3,2} + X_{3,1}X_{1,3} + X_{3,2}X_{2,3} + X_{3,3}^2\big).$$

Lembrando que estamos considerando apenas matrizes de Wigner, temos que

$$E\langle\mu_N, x^k\rangle = E\frac{1}{9}\big(X_{1,1}^2 + 2X_{1,2}^2 + 2X_{1,3}^2 + X_{2,2}^2 + 2X_{2,3}^2 + X_{3,3}^2\big) =$$

$$= \frac{1}{9}\big(EX_{1,1}^2 + 2EX_{1,2}^2 + 2EX_{1,3}^2 + EX_{2,2}^2 + 2EX_{2,3}^2 + EX_{3,3}^2\big).$$

Sabemos que a variância dos elementos fora da diagonal é 1, sendo σ^2 a variância dos elementos da diagonal, a expressão acima se reduz à

$$\frac{1}{9}\big(\sigma^2 + 2\cdot 1 + 2\cdot 1 + \sigma^2 + 2\cdot 1 + \sigma^2\big) = \frac{1}{9}\cdot(3\sigma^2+6) = \frac{\sigma^2+2}{3}.$$

Quando a matriz tem ordem maior, temos muitos termos do tipo $X_{i_1,i_2}X_{i_2,i_3}\dots X_{i_{k-1},i_k}X_{i_k,i_1}$ em que existe um $X_{r,s}$ que ocorre apenas uma vez (levando em conta o fato de $\mathbf{X}$ ser simétrica). Como todos os termos são variáveis aleatórias independentes, temos que[2]

$$E(X_{i_1,i_2}X_{i_2,i_3}\dots X_{i_{k-1},i_k}X_{i_k,i_1}) = EX_{r,s}\cdot E\Big(X_{i_1,i_2}X_{i_2,i_3}\dots \widehat{X_{r,s}}\dots X_{i_{k-1},i_k}X_{i_k,i_1}\Big) = 0,$$

pois o valor esperado de $X_{r,s}$ é 0.

É esperado que haja muitos anulamentos devido a termos deste tipo, então podemos tentar organizar melhor os índices de modo a estimar o valor de $E\langle \mu_N, x^k\rangle$.

2.3 Palavras e Multigrafos

Gostaríamos de organizar os termos em $\mathrm{tr}(\mathbf{X}^k)$ antes de calcular o valor esperado. Seria interessante poder separar os termos da diagonal do restante, já que possuem distribuições diferentes. Além disso, gostaríamos de juntar termos de índices iguais ou invertidos (lembrando que estamos lidando apenas com matrizes simétricas). Por fim, gostaríamos de destacar termos que aparecem apenas uma vez, pois neste caso eles vão zerar todo o produtório, como foi mencionado acima. Para alcançar este objetivo, devemos organizar os índices, e para isto, vamos usar a noção de *palavra*, que iremos definir agora.

Definição 2.17. Seja $N \in \mathbb{N}$ fixado. *N-palavra* é uma sequência finita e não-vazia $s_1\dots s_k$, tal que cada s_i pertence a $\{1,\dots,N\}$. Neste contexto, cada s_i é chamado de *letra*.

Por questão de simplicidade, chamaremos uma N-palavra apenas por *palavra*, e apenas iremos explicitar N quando houver necessidade. Seja $w = s_1\dots s_k$ uma palavra, então temos as seguintes definições.

Definição 2.18 Uma palavra é dita *fechada* se a sua primeira letra é igual à última letra, ou seja, w é fechada se $s_1 = s_k$.

Definição 2.19. O *comprimento* de w é a quantidade de letras de w, denota-se por $\ell(w)$. Temos que $\ell(w) = k$.

Definição 2.20. *Suporte* é o conjunto das letras de w, denota-se por $\mathrm{supp}(w)$. Temos que $\mathrm{supp}(w) = \{s_i :\ i = 1\dots k\}$.

Definição 2.21. *Peso* é a quantidade de letras distintas de w, denota-se por $\mathrm{wt}(w)$. Temos que $\mathrm{wt}(w) = \sharp\,\mathrm{supp}(w)$.

Podemos associar um multigrafo à palavra w, este multigrafo é dado por $G_w = (V_w, E_w)$, tal que $V_w = \mathrm{supp}(w)$ é o conjunto dos vértices e $E_w = \big\{\{s_i, s_{i+1}\} :\ i = 1\dots k-1\big\}$ é o conjunto[3] das arestas. Este multigrafo trata cada letra s_i como um vértice e cada par $\{s_i, s_{i+1}\}$ como uma aresta, ele também

[2] A notação de "chapéu" em $X_{i_1,i_2}X_{i_2,i_3}\dots\widehat{X_{r,s}}\dots X_{i_{k-1},i_k}X_{i_k,i_1}$, significa que $X_{r,s}$ é o único termo que não deve ser contabilizado nesta expressão.

[3] Não estamos considerando E_w como multiconjunto, pois as arestas repetidas serão determinadas pela palavra w.

pode ter arestas indo de um vértice para o mesmo vértice e pode ter mais de uma aresta entre dois vértices, daí concluímos que o multigrafo possui $\mathrm{wt}(w)$ vértices e $\ell(w)-1$ arestas. Note também que G_w não é um multigrafo orientado. Em relação à este multigrafo, temos as seguintes definições.

Definição 2.22. O conjunto $\big\{e \in E_w : \ e = \{u,u\},\ u \in V_w\big\}$ é chamado de conjunto das *auto arestas*, é denotado por E_w^s.

Definição 2.23. O conjunto $E_w \backslash E_w^s$ é chamado de conjunto das *arestas conectivas*, é denotado por E_w^c.

Definição 2.24. Para dar conta das arestas repetidas, denotamos por N_e^w a quantidade de vezes que ocorre a aresta $e \in E_w$ no multigrafo G_w.

Exemplo 2.25. Considere a palavra $w = 212443312$, temos que $\ell(w) = 9$, $\mathrm{supp}(w) = \{1,2,3,4\}$ e $\mathrm{wt}(w) = 4$. Além disso, se $G_w = (V_w, E_w)$ é o multigrafo associado à w, então temos que $V_w = \{1,2,3,4\}$, $E_w = \big\{\{1,2\},\{2,4\},\{4,4\},\{3,4\},\{3,3\},\{1,3\}\big\}$, $E_w^s = \big\{\{3,3\},\{4,4\}\big\}$, $E_w^c = \big\{\{1,2\},\{2,4\},\{3,4\},\{1,3\}\big\}$. Mostramos abaixo o multigrafo G_w.

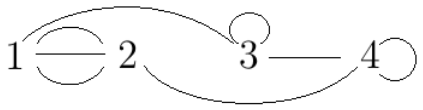

Figura 2.6: Multigrafo G_w gerado pela palavra $w = 212443312$.

Neste exemplo, temos que $N_{\{1,2\}}^w = 3$ e $N_e^w = 1$ quando $e \neq \{1,2\}$. Note ainda que $\sum_{e\in E_w} N_e^w$ é a quantidade total de arestas de G_w (não confundir com $\sharp E_w$), portanto $\sum_{e\in E_w} N_e^w = \ell(w) - 1$.

Usaremos estes conceitos para representar os índices de $X_{i_1,i_2}X_{i_2,i_3}\dots X_{i_{k-1},i_k}X_{i_k,i_1}$ através da palavra $w = i_1i_2i_3\dots i_{k-1}i_ki_1$. Fazendo isso, podemos escrever

$$X_{i_1,i_2}X_{i_2,i_3}\dots X_{i_{k-1},i_k}X_{i_k,i_1} = \prod_{e\in E_w} X_e^{N_e^w}.$$

Como $E_w = E_w^s \cup E_w^c$ e $E_w^s \cap E_w^c = \emptyset$, então temos que

$$\prod_{e\in E_w} X_e^{N_e^w} = \prod_{e\in E_w^s} X_e^{N_e^w} \cdot \prod_{e\in E_w^c} X_e^{N_e^w}.$$

Dado um índice $i = (i_1, i_2\dots, i_k) \in \{1,\dots,N\}^k$, podemos definir a palavra $w_i = i_1i_2\dots i_ki_1$, que associaremos a este índice. Deste modo, temos que

$$\langle \mu_N, x^k\rangle = \frac{1}{N^{\frac{k}{2}+1}}\mathrm{tr}(\mathbf{X}^k) = \frac{1}{N^{\frac{k}{2}+1}} \sum_{i\in\{1,\dots,N\}^k} \left(\prod_{e\in E_{w_i}^s} X_e^{N_e^{w_i}} \cdot \prod_{e\in E_{w_i}^c} X_e^{N_e^{w_i}} \right).$$

Esta maneira de escrever $\mathrm{tr}(\mathbf{X}^k)$ nos permite separar os elementos da diagonal dos outros, além disso nos permite saber quantas vezes aparece cada elemento. A partir disso fica mais fácil trabalhar com o valor esperado. Sabemos que $EX_e^{N_e^{w_i}} = 0$ quando $N_e^{w_i} = 1$, então é suficiente considerar apenas os índices $i \in \{1, \ldots, N\}^k$ tais que $N_e^{w_i} \geq 2$ para todo $e \in E_{w_i}$. Esta restrição nos conduz à seguinte proposição.

Proposição 2.26. Sejam $i \in \{1, \ldots, N\}^k$ e w_i a palavra associada a este índice, se $N_e^{w_i} \geq 2$ para todo $e \in E_{w_i}$, então $\mathrm{wt}(w_i) \leq \frac{k}{2} + 1$.

Prova: Temos que $w_i = i_1 i_2 i_3 \ldots i_{k-1} i_k i_1$ e $E_{w_i} = \big\{\{i_1, i_2\}, \{i_2, i_3\}, \ldots, \{i_{k-1}, i_k\}, \{i_k, i_1\}\big\}$. Vamos fazer a prova para dois casos: quando k é par e quando k é ímpar.

Quando k é par, temos que $\sum_{e \in E_{w_i}} N_e^{w_i}$ é par, pois $\sum_{e \in E_{w_i}} N_e^{w_i} = \ell(w_i) - 1 = k$. Note que, quanto mais arestas distintas, maior é $\mathrm{wt}(w_i)$, sendo que o máximo é atingido quando $N_e^{w_i} = 2$, para todo $e \in E_{w_i}$. Neste caso, temos que $\sum_{e \in E_{w_i}} N_e^{w_i} = 2 \cdot \sharp E_{w_i}$, e portanto, $\sharp E_{w_i} = \frac{k}{2}$. Sabemos que $\sharp V_{w_i} - 1 \leq \sharp E_{w_i}$ e $\sharp V_{w_i} = \mathrm{wt}(w_i)$, então $\mathrm{wt}(w_i) - 1 \leq \frac{k}{2}$, logo, $\mathrm{wt}(w_i) \leq \frac{k}{2} + 1$.

Quando k é ímpar, temos que $\sum_{e \in E_{w_i}} N_e^{w_i}$ é ímpar. O valor máximo para $\mathrm{wt}(w_i)$ é atingido quando temos exatamente uma aresta $e \in E_{w_i}$, com $N_e^{w_i} = 3$, e o restante é 2. Então temos que $\sum_{e \in E_{w_i}} N_e^{w_i} = 3 + 2 \cdot (\sharp E_{w_i} - 1) = k \implies \sharp E_{w_i} = \frac{k-1}{2}$, portanto, $\mathrm{wt}(w_i) - 1 \leq \frac{k-1}{2} \implies \mathrm{wt}(w_i) \leq \frac{k+1}{2} < \frac{k}{2} + 1$. □

Exemplo 2.27. Considere que $X_{1,1}X_{2,1}X_{1,1}X_{1,2}X_{2,3}X_{3,2}X_{2,1}$ e $X_{2,2}X_{6,2}X_{2,2}X_{2,6}X_{6,5}X_{5,6}X_{6,2}$ são termos de algum $\langle \mu_N, x^k \rangle$, quando $k = 7$. Para calcularmos $E\langle \mu_N, x^k \rangle$, primeiramente devemos calcular $E(X_{1,1}X_{2,1}X_{1,1}X_{1,2}X_{2,3}X_{3,2}X_{2,1})$ e $E(X_{2,2}X_{6,2}X_{2,2}X_{2,6}X_{6,5}X_{5,6}X_{6,2})$. Temos que

$$E(X_{1,1}X_{2,1}X_{1,1}X_{1,2}X_{2,3}X_{3,2}X_{2,1}) = E(X_{1,1}^2)E(X_{1,2}^3)E(X_{2,3}^2) = 1 \cdot E(X_{1,2}^3) \cdot 1 = E(X_{1,2}^3)$$

e

$$E(X_{2,2}X_{6,2}X_{2,2}X_{2,6}X_{6,5}X_{5,6}X_{6,2}) = E(X_{2,2}^2)E(X_{2,6}^3)E(X_{5,6}^2) = 1 \cdot E(X_{2,6}^3) \cdot 1 = E(X_{2,6}^3).$$

Como $X_{1,2}$ e $X_{2,6}$ tem a mesma distribuição, sabemos que $E(X_{1,2}^3) = E(X_{2,6}^3)$. De um modo geral, temos que $E(X_{i,j}^k) = E(X_{i',j'}^k)$, se os dois estão na diagonal ou os dois estão fora da diagonal, pois nestes casos, eles tem a mesma distribuição. Então o que nos importa é saber quantas vezes ocorre cada termo e se ele está na diagonal ou não, pois havendo igualdade destas duas propriedades, os valores esperados são iguais. Estas ideias nos levam à seguinte definição.

Definição 2.28. Duas N-palavras, w_1, w_2, são ditas *equivalentes* se existir uma bijeção em $\{1, \ldots, N\}$ que leva uma palavra na outra. Denota-se por $w_1 \sim w_2$.

Exemplo 2.29. Sejam $w_1 = 123431$ e $w_2 = 213732$, então $f : \{1, \ldots, 7\} \to \{1, \ldots, 7\}$ dada por $f(1) = 2$, $f(2) = 1$, $f(3) = 3$, $f(4) = 7$, $f(5) = 6$, $f(6) = 4$, $f(7) = 5$, é uma bijeção que leva w_1 em w_2, portanto, $w_1 \sim w_2$. Note que é necessário termos $\ell(w_1) = \ell(w_2)$ e $\mathrm{wt}(w_1) = \mathrm{wt}(w_2)$ para que as palavras sejam equivalentes.

Proposição 2.30. Sejam as palavras $w_1 = i_1 i_2 i_3 \ldots i_{k-1} i_k i_1$ e $w_2 = j_1 j_2 j_3 \ldots j_{k-1} j_k j_1$ tais que $w_1 \sim w_2$, então $E(X_{i_1,i_2}X_{i_2,i_3} \ldots X_{i_{k-1},i_k}X_{i_k,i_1}) = E(X_{j_1,j_2}X_{j_2,j_3} \ldots X_{j_{k-1},j_k}X_{j_k,j_1})$.

Prova: Temos que

$$E(X_{i_1,i_2}X_{i_2,i_3} \ldots X_{i_{k-1},i_k}X_{i_k,i_1}) = E \prod_{e \in E_{w_1}} X_e^{N_e^{w_1}} = E\left(\prod_{e \in E_{w_1}^s} X_e^{N_e^{w_1}} \cdot \prod_{e \in E_{w_1}^c} X_e^{N_e^{w_1}} \right).$$

De modo análogo, temos que

$$E(X_{j_1,j_2}X_{j_2,j_3}\dots X_{j_{k-1},j_k}X_{j_k,j_1}) = E\left(\prod_{e\in E^s_{w_2}} X_e^{N_e^{w_2}} \cdot \prod_{e\in E^c_{w_2}} X_e^{N_e^{w_2}}\right).$$

A bijeção entre w_1 e w_2 nos garante que para cada auto aresta $e \in E^s_{w_1}$ existe uma outra auto aresta $e' \in E^s_{w_2}$ tal que $N_e^{w_1} = N_{e'}^{w_2}$, de modo que $EX_e^{N_e^{w_1}} = EX_{e'}^{N_{e'}^{w_2}}$. O mesmo vale para cada aresta conectiva. Daí temos que

$$\prod_{e\in E^s_{w_1}} EX_e^{N_e^{w_1}} = \prod_{e'\in E^s_{w_2}} EX_{e'}^{N_{e'}^{w_2}} \implies E\prod_{e\in E^s_{w_1}} X_e^{N_e^{w_1}} = E\prod_{e'\in E^s_{w_2}} X_{e'}^{N_{e'}^{w_2}}$$

De modo análogo, temos que

$$E\prod_{e\in E^c_{w_1}} X_e^{N_e^{w_1}} = E\prod_{e'\in E^c_{w_2}} X_{e'}^{N_{e'}^{w_2}},$$

portanto,

$$E\prod_{e\in E^s_{w_1}} X_e^{N_e^{w_1}} \cdot E\prod_{e\in E^c_{w_1}} X_e^{N_e^{w_1}} = E\prod_{e'\in E^s_{w_2}} X_{e'}^{N_{e'}^{w_2}} \cdot E\prod_{e'\in E^c_{w_2}} X_{e'}^{N_{e'}^{w_2}} \implies$$

$$\implies E\left(\prod_{e\in E^s_{w_1}} X_e^{N_e^{w_1}} \cdot \prod_{e\in E^c_{w_1}} X_e^{N_e^{w_1}}\right) = E\left(\prod_{e'\in E^s_{w_2}} X_{e'}^{N_{e'}^{w_2}} \cdot \prod_{e'\in E^c_{w_2}} X_{e'}^{N_{e'}^{w_2}}\right) \qquad \square$$

Proposição 2.31. Seja $w = s_1 \dots s_k$ uma N-palavra tal que $\mathrm{wt}(w) = t$, com $t \le k \le N$, então existem $\frac{N!}{(N-t)!}$ palavras equivalentes a w, incluindo w.

Prova: Seja $\mathrm{supp}(w) = \{s_{i_1}, \dots, s_{i_t}\}$ o conjunto das letras distintas de w. Podemos trocar algumas letras s_{i_j} por outras letras $s'_{i_j} \in \{1, \dots, N\}$ e obter o conjunto $\{s'_{i_1}, \dots, s'_{i_t}\}$, de modo que este conjunto ainda tenha t letras distintas. Consideramos a bijeção $f : \{1, \dots, N\} \to \{1, \dots, N\}$ tal que $f(s_{i_j}) = s'_{i_j}$. Deste modo, temos que $w' = f(s_1) \dots f(s_k)$ é uma N-palavra equivalente à w.

Podemos obter todas as palavras equivalentes à w a partir desta construção, além disso, podemos contar todas as possíveis palavras que podemos obter deste modo. Queremos escolher $\{s'_{i_1}, \dots, s'_{i_t}\}$ de modo que todas as letras sejam distintas entre si, além disso, só há N possíveis escolhas para cada letra. Então temos N escolhas para s'_{i_1}, $N-1$ escolhas para s'_{i_2}, e assim por diante, até chegarmos à s'_{i_t}, onde temos $N - t + 1$ escolhas. Portanto, há $N \cdot (N-1) \dots (N-t+1) = \frac{N!}{(N-t)!}$ possíveis N-palavras equivalentes à w, incluindo a própria palavra w nesta contagem. $\square$

As Proposições 2.30 e 2.31 nos permitem uma grande simplificação na expressão para $E\langle \mu_N, x^k\rangle$, pois o somatório passa por todos os índices, então teremos uma grande quantidade de palavras equivalentes. Mas antes disso, note que podem existir duas (ou mais) N-palavras com peso t que não são equivalentes, por exemplo, $w_1 = 122151$ e $w_2 = 121131$. Apesar de não serem equivalentes, a Proposição 2.31 vale para cada uma delas, então existem $\frac{N!}{(N-3)!}$ palavras equivalentes à w_1 e mais outras $\frac{N!}{(N-3)!}$ palavras equivalentes à w_2, sendo que todas estas $2 \cdot \frac{N!}{(N-3)!}$ palavras são distintas entre si. Por fim, note que esta equivalência que definimos é de fato uma relação de equivalência, então podemos considerar todas as

palavras equivalentes à w_1 como sendo a mesma palavra, o mesmo vale para as palavras equivalentes à w_2. A ideia é que podemos pensar em termos de classes de equivalência. Levando em conta que estamos interessados apenas em palavras w tais que $N_e^w \geq 2$ para todo $e \in E_w$, somos levados a fazer a seguinte definição.

Definição 2.32. Denotamos por $W_{k,t}$ o conjunto dos representantes de cada classe de equivalência das N-palavras que são fechadas, com $\ell(w) = k+1$, $\mathrm{wt}(w) = t$, e $N_e^w \geq 2$ para todo $e \in E_w$.

Exemplo 2.33. Sejam $N = 3$, $k = 3$, e considere as N-palavras fechadas com $\ell(w) = 4$. Podemos agrupar as palavras equivalentes.

$$A = \{1111, 2222, 3333\}$$
$$B = \{1121, 1131, 2212, 2232, 3313, 3323\}$$
$$C = \{1211, 1311, 2122, 2322, 3133, 3233\}$$
$$D = \{1221, 1331, 2112, 2332, 3113, 3223\}$$
$$E = \{1231, 1321, 2132, 2312, 3123, 3213\}$$

Neste caso, apenas o conjunto A é tal que $N_e^w \geq 2$ para todo $e \in E_w$, então, escolhendo 1111 para ser o representante de A, temos que $W_{3,1} = \{1111\}$. Daí também podemos concluir que $W_{3,2}$ e $W_{3,3}$ são conjuntos vazios.

Voltando ao caso geral, note que, dada uma palavra w_1, com $\mathrm{wt}(w_1) = t$, as Proposições 2.30 e 2.31 nos dizem que

$$\sum_{w \sim w_1} \prod_{e \in E_w^s} EX_e^{N_e^w} \cdot \prod_{e \in E_w^c} EX_e^{N_e^w} = \frac{N!}{(N-t)!} \prod_{e \in E_{w_1}^s} EX_e^{N_e^{w_1}} \cdot \prod_{e \in E_{w_1}^c} EX_e^{N_e^{w_1}}.$$

Variando t de 1 a k, podemos simplificar o somatório de $E\langle \mu_N, x^k \rangle$ com a ideia de classes de equivalência, como mostrado acima. Daí temos que

$$E\langle \mu_N, x^k \rangle = \frac{1}{N^{\frac{k}{2}+1}} \sum_{t=1}^{k} \sum_{w \in W_{k,t}} \left(\frac{N!}{(N-t)!} \prod_{e \in E_w^s} EX_e^{N_e^w} \cdot \prod_{e \in E_w^c} EX_e^{N_e^w} \right).$$

Pela Proposição 2.26, temos que é suficiente considerar $t \leq \lfloor \frac{k}{2} \rfloor + 1$, então

$$E\langle \mu_N, x^k \rangle = \frac{1}{N^{\frac{k}{2}+1}} \sum_{t=1}^{\lfloor \frac{k}{2} \rfloor + 1} \sum_{w \in W_{k,t}} \left(\frac{N!}{(N-t)!} \prod_{e \in E_w^s} EX_e^{N_e^w} \cdot \prod_{e \in E_w^c} EX_e^{N_e^w} \right) =$$

$$= \frac{1}{N^{\frac{k}{2}+1}} \sum_{t=1}^{\lfloor \frac{k}{2} \rfloor + 1} \frac{N!}{(N-t)!} \sum_{w \in W_{k,t}} \left(\prod_{e \in E_w^s} EX_e^{N_e^w} \cdot \prod_{e \in E_w^c} EX_e^{N_e^w} \right).$$

Por fim, note que a quantidade de elementos de $W_{k,t}$ não depende de N, uma vez que os elementos de $W_{k,t}$ são todas as possíveis combinações que usam t letras distintas. Essas combinações não dependem

de quais letras usamos, elas dependem apenas do posicionamento de cada uma das letras[4]. Aumentando a quantidade de letras, teremos apenas novas palavras que são equivalentes a algumas das combinações já definidas. Para ilustrar isso, considere o exemplo 2.33, onde determinamos $W_{3,1}$. Se fizermos $N = 4$ e agruparmos as palavras equivalentes, teremos o seguinte.

$$A = \{1111, 2222, 3333, 4444\}$$
$$B = \{1121, 1131, 1141, 2212, 2242, 2232, 3313, 3323, 3343, 4414, 4424, 4434\}$$
$$C = \{1211, 1311, 1411, 2122, 2322, 2422, 3133, 3233, 3433, 4144, 4244, 4344\}$$
$$D = \{1221, 1331, 1441, 2112, 2332, 2442, 3113, 3223, 3443, 4114, 4224, 4334\}$$
$$E = \{1231, 1321, 1241, 1421, 1341, 1431, 2132, 2312, 2142, 2412, 2342, 2432,$$
$$3123, 3213, 3143, 3413, 3243, 3423, 4124, 4214, 4134, 4314, 4234, 4324\}$$

Podemos ver que $W_{3,1}$ possui apenas um elemento, assim como antes. Em resumo, aumentar N não aumenta a quantidade de classes de equivalência, uma vez que a quantidade de classes possíveis depende do tamanho da palavra e não da quantidade de letras disponíveis.

Como $W_{k,t}$ não depende de N, podemos considerar o limite

$$\lim_{N\to\infty} E\langle \mu_N, x^k\rangle = \lim_{N\to\infty} \frac{1}{N^{\frac{k}{2}+1}} \sum_{t=1}^{\lfloor \frac{k}{2}\rfloor+1} \frac{N!}{(N-t)!} \sum_{w\in W_{k,t}} \left(\prod_{e\in E_w^s} EX_e^{N_e^w} \cdot \prod_{e\in E_w^c} EX_e^{N_e^w} \right).$$

Proposição 2.34.

$$\lim_{N\to\infty} E\langle \mu_N, x^k\rangle = \begin{cases} 0, \text{ se } k \text{ é ímpar} \\ \sum_{w\in W_{k,\frac{k}{2}+1}} \left(\prod_{e\in E_w^s} EX_e^{N_e^w} \cdot \prod_{e\in E_w^c} EX_e^{N_e^w} \right), \text{ se } k \text{ é par} \end{cases}$$

Prova:

$$\lim_{N\to\infty} \frac{1}{N^{\frac{k}{2}+1}} \sum_{t=1}^{\lfloor \frac{k}{2}\rfloor+1} \frac{N!}{(N-t)!} \sum_{w\in W_{k,t}} \left(\prod_{e\in E_w^s} EX_e^{N_e^w} \cdot \prod_{e\in E_w^c} EX_e^{N_e^w} \right) =$$
$$= \sum_{t=1}^{\lfloor \frac{k}{2}\rfloor+1} \lim_{N\to\infty} \frac{N!}{N^{\frac{k}{2}+1}(N-t)!} \sum_{w\in W_{k,t}} \left(\prod_{e\in E_w^s} EX_e^{N_e^w} \cdot \prod_{e\in E_w^c} EX_e^{N_e^w} \right) =$$
$$= \sum_{t=1}^{\lfloor \frac{k}{2}\rfloor+1} \sum_{w\in W_{k,t}} \left(\prod_{e\in E_w^s} EX_e^{N_e^w} \cdot \prod_{e\in E_w^c} EX_e^{N_e^w} \right) \cdot \lim_{N\to\infty} \frac{N!}{N^{\frac{k}{2}+1}(N-t)!}$$

Vamos calcular $\lim_{N\to\infty} \frac{N!}{N^{\frac{k}{2}+1}(N-t)!}$, para $t \le \lfloor \frac{k}{2} \rfloor$. Temos que

$$0 \le \lim_{N\to\infty} \frac{N!}{N^{\frac{k}{2}+1}(N-t)!} \le \lim_{N\to\infty} \frac{N!}{N^{\frac{k}{2}+1}\left(N-\lfloor \frac{k}{2}\rfloor\right)!} =$$
$$= \lim_{N\to\infty} \frac{N(N-1)\dots(N-\lfloor \frac{k}{2}\rfloor+1)}{N^{\frac{k}{2}+1}} \le \lim_{N\to\infty} \frac{N^{\lfloor \frac{k}{2}\rfloor}}{N^{\frac{k}{2}+1}} \le \lim_{N\to\infty} \frac{N^{\frac{k}{2}}}{N^{\frac{k}{2}+1}} = \lim_{N\to\infty} \frac{1}{N} = 0.$$

[4]Poderia depender de N se tivéssemos $N < t$, mas não é este o caso.

Daí temos que

$$\lim_{N\to\infty} E\langle \mu_N, x^k\rangle = \sum_{w\in W_{k,\lfloor \frac{k}{2}\rfloor+1}} \left(\prod_{e\in E_w^s} EX_e^{N_e^w} \cdot \prod_{e\in E_w^c} EX_e^{N_e^w} \right) \cdot \lim_{N\to\infty} \frac{N!}{N^{\frac{k}{2}+1}\Big(N-(\lfloor \frac{k}{2}\rfloor+1)\Big)!}.$$

Caso k seja ímpar, podemos usar o mesmo procedimento que usamos acima para calcular o limite, obtendo

$$0 \leq \lim_{N\to\infty} \frac{N!}{N^{\frac{k}{2}+1}\Big(N-(\lfloor \frac{k}{2}\rfloor+1)\Big)!} \leq \lim_{N\to\infty} \frac{N^{\lfloor \frac{k}{2}\rfloor+1}}{N^{\frac{k}{2}+1}}.$$

Como $\lfloor \frac{k}{2}\rfloor + 1 < \frac{k}{2}+1$, temos que $\lim_{N\to\infty} \frac{N^{\lfloor \frac{k}{2}\rfloor+1}}{N^{\frac{k}{2}+1}} = 0$, portanto, $\lim_{N\to\infty} E\langle \mu_N, x^k\rangle = 0$.

Caso k seja par, então $\lfloor \frac{k}{2}\rfloor + 1 = \frac{k}{2}+1$, daí temos que

$$\lim_{N\to\infty} \frac{N!}{N^{\frac{k}{2}+1}\Big(N-(\frac{k}{2}+1)\Big)!} = \lim_{N\to\infty} \frac{1}{N^{\frac{k}{2}+1}} \prod_{i=0}^{\frac{k}{2}} (N-i) =$$

$$= \lim_{N\to\infty} \prod_{i=0}^{\frac{k}{2}} \frac{N-i}{N} = \prod_{i=0}^{\frac{k}{2}} \lim_{N\to\infty} \frac{N-i}{N} = \prod_{i=0}^{\frac{k}{2}} 1 = 1,$$

portanto

$$\lim_{N\to\infty} E\langle \mu_N, x^k\rangle = \sum_{w\in W_{k,\frac{k}{2}+1}} \left(\prod_{e\in E_w^s} EX_e^{N_e^w} \cdot \prod_{e\in E_w^c} EX_e^{N_e^w} \right). \qquad \square$$

Proposição 2.35. Sejam k par e $w \in W_{k,\frac{k}{2}+1}$, então $E_w^s = \emptyset$ e $N_e^w = 2$ para todo $e \in E_w$.

Prova: Como $w \in W_{k,\frac{k}{2}+1}$, sabemos que $\mathrm{wt}(w) = \sharp V_w = \frac{k}{2}+1$. Temos também que $\sharp V_w - 1 \leq \sharp E_w$, então $\sharp E_w \geq \frac{k}{2}$. Por outro lado, como $N_e^w \geq 2$ para todo $e \in E_w$, temos que $\sum_{e\in E_w} N_e^w \geq 2 \cdot \sharp E_w$, portanto $\sharp E_w \leq \frac{1}{2} \cdot \sum_{e\in E_w} N_e^w = \frac{1}{2} \cdot (\ell(w)-1) = \frac{k}{2}$. Concluímos então que $\sharp E_w = \frac{k}{2}$, portanto $\sharp E_w = \sharp V_w - 1$. Daí temos que G_w é uma árvore, então $\sharp E_w^c = \sharp V_w - 1$ e $\sharp E_w^s = 0$, portanto $E_w^s = \emptyset$.

Agora note que, se $N_e^w = 2$ para todo $e \in E_w$, então $\sum_{e\in E_w} N_e^w = 2 \cdot \sum_{e\in E_w} 1 = 2 \cdot \sharp E_w = 2 \cdot \frac{k}{2} = k = \ell(w) - 1$. Então não é possível termos nenhum $N_e^w > 2$, uma vez que não podemos diminuir outros N_e^w para compensar e manter a igualdade $\sum_{e\in E_w} N_e^w = k$. Então podemos concluir que $N_e^w = 2$ para todo $e \in E_w$. $\square$

2.4 Passeios de Bernoulli e Caminhos de Dyck

Antes de provar o Lema 2.11, precisamos de alguns resultados a respeito de *Passeios de Bernoulli*, mais especificamente, o conjunto dos *Caminhos de Dyck*. Nesta seção iremos fazer uso do número de Catalan, que é dado por $C_k = \binom{2k}{k} \cdot \frac{1}{k+1}$.

Definição 2.36. Um passeio de Bernoulli de comprimento $k \in \mathbb{N}$ é uma sequência de inteiros $s_0, \dots s_k$, tal que $s_0 = 0$ e $|s_{i+1} - s_i| = 1$ para todo $i = 0 \dots k-1$.

Podemos representar um passeio de Bernoulli no plano certesiano pelos pares ordenados (i, s_i). Mostramos abaixo o gráfico do passeio de Bernoulli $0, 1, 0, -1, -2, -3, -2, -3$.

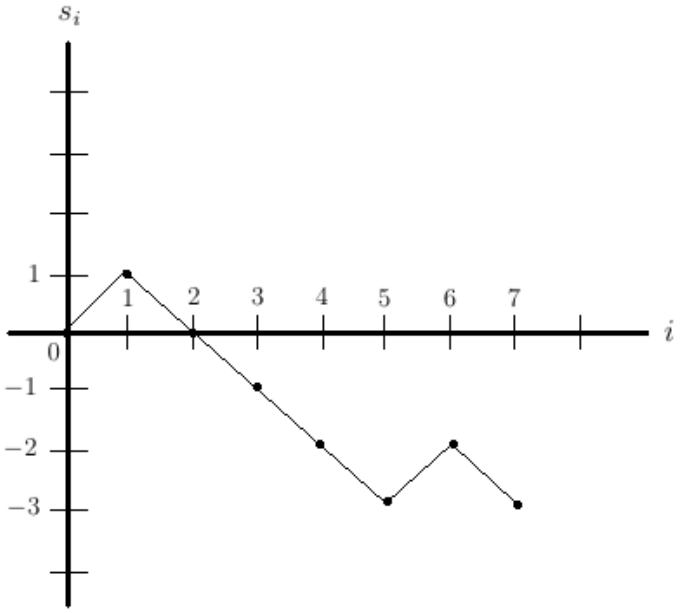

Figura 2.7: O passeio de Bernoulli $0, 1, 0, -1, -2, -3, -2, -3$ pode ser representado no plano cartesiano.

Definição 2.37. Um caminho de Dyck é um passeio de Bernoulli que não possui elementos negativos e que termina em 0.

Mostramos abaixo o caminho de Dyck $0, 1, 2, 1, 2, 3, 2, 1, 0$.

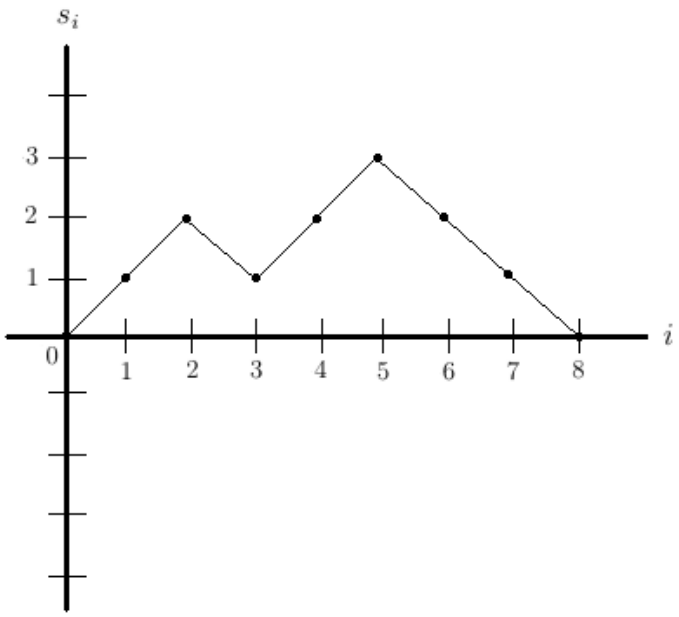

Figura 2.8: Caminho de Dyck $0, 1, 2, 1, 2, 3, 2, 1, 0$ representado no plano cartesiano.

Denotaremos por B_k o conjunto dos passeios de Bernoulli de comprimento k e por $\mathcal{D}_k$ o conjunto dos caminhos de Dyck de comprimento k.

Proposição 2.38. $\sharp B_k = 2^k$, para todo $k \in \mathbb{N}$.

Prova: Cada s_i, $i = 1, \ldots, k$, pode ser associado a 1 ou -1, dependendo se $s_i = s_{i-1} \pm 1$. Como temos 2 possibilidades para cada uma das k escolhas, concluímos que existem 2^k passeios de Bernoulli. □

Dado um passeio de Bernoulli $s_0, \ldots, s_k$, denotamos por D (direita) a quantidade de termos s_i tal que $s_i = s_{i-1} + 1$, e por E (esquerda) a quantidade de termos s_i tal que $s_i = s_{i-1} - 1$. A ideia de direita e esquerda vem do fato que o passeio é feito na reta, e cada novo passo é dado em uma dessas direções. Com esta notação, note que a sequência $s_0, \ldots, s_k$ deve terminar em $s_k = D - E$. A proposição a seguir diz quantos caminhos de Bernoulli terminando em $D - E$ existem, quando temos D e E fixados.

Proposição 2.39. Dados $D, E \in \mathbb{N}$, existem $\binom{k}{E}$ (note que $k = D + E$) caminhos de B_k terminando em $D - E$.

Prova: Denote por "-1" os elementos s_i que são um passo para a esquerda. Podemos obter todos os caminhos possíveis ao considerar a quantidade de maneiras de se colocar "-1" em E posições de $s_0, \ldots, s_k$ de todas as maneiras possíveis. Como isso é o arranjo de k elementos tomados E a E, então podemos concluir que existem $\binom{k}{E}$ caminhos de Bernoulli terminando em $D-E$. □

Poderíamos ter colocado $\binom{k}{D}$ no teorema, usando o mesmo argumento para a prova, isto vem do fato que $\binom{k}{D} = \binom{k}{E}$, já que $k = D+E$. Note também que todo caminho de Dyck tem que ter comprimento par.

A proposição a seguir diz quantos caminhos há em $\mathcal{D}_{2k}$.

Proposição 2.40. $\sharp\mathcal{D}_{2k} = C_k$, para todo $k \in \mathbb{N}$.

Prova: Seja $\overline{B_{2k}} \subset B_{2k}$ o conjunto dos passeios de Bernoulli de comprimento $2k$ que terminam em 0 e seja $\overline{b_{2k}} \subset \overline{B_{2k}}$ o conjunto dos passeios de $\overline{B_{2k}}$ que possuem pelo menos um elemento negativo. Note que $\overline{B_{2k}}$ é o conjunto em que $D = E = k$, então pela Proposição 2.39, temos que $\sharp\overline{B_{2k}} = \binom{2k}{k}$. Além disso, note que $\overline{B_{2k}} = \mathcal{D}_{2k} \cup \overline{b_{2k}}$ e $\mathcal{D}_{2k} \cap \overline{b_{2k}} = \emptyset$, então $\sharp\overline{B_{2k}} = \sharp\mathcal{D}_{2k} + \sharp\overline{b_{2k}}$, e portanto $\sharp\mathcal{D}_{2k} = \binom{2k}{k} - \sharp\overline{b_{2k}}$.

Para calcular $\sharp\overline{b_{2k}}$, vamos estabelecer uma bijeção entre $\overline{b_{2k}}$ e o conjunto dos passeios de Bernoulli que terminam em -2. Fazemos isso do seguinte modo: considere $s_0, \ldots, s_{2k} \in \overline{b_{2k}}$ e seja s_t o primeiro elemento negativo da sequência, ou seja, s_t é o primeiro elemento igual a -1. Agora refletimos todos os pontos que vem depois de s_t. Temos que essa reflexão é uma bijeção entre $\overline{b_{2k}}$ e um outro conjunto de passeios de Bernoulli. Podemos ver que a partir de s_{t+1}, o ponto refletido de 0 é sempre -2, e como a sequência original termina em 0, a sequência refletida termina em -2. Concluímos então que $\overline{b_{2k}}$ possui uma bijeção com o conjunto dos passeios de Bernoulli que terminam em -2. Neste conjunto, temos que $E = k+1$ e $D = k-1$, portanto $\sharp\overline{b_{2k}} = \binom{2k}{k-1}$. Então $\sharp\mathcal{D}_{2k} = \binom{2k}{k} - \binom{2k}{k-1} = \binom{2k}{k} \cdot \frac{1}{k+1} = C_k$. □

Agora podemos provar o Lema 2.11.

Prova do Lema 2.11: Se k é ímpar, a Proposição 2.34 garante que $\lim_{N\to\infty} E\langle \mu_N, x^k \rangle = 0$. Caso k seja par, as Proposições 2.34 e 2.35 garantem que

$$\lim_{N\to\infty} E\langle \mu_N, x^k \rangle = \sum_{w \in W_{k,\lfloor \frac{k}{2} \rfloor + 1}} \left(\prod_{e \in E_w^c} EX_e^{N_e^w} \right) = \sum_{w \in W_{k,\frac{k}{2}+1}} \left(\prod_{e \in E_w^c} EX_e^2 \right) =$$

$$= \sum_{w \in W_{k,\frac{k}{2}+1}} \prod_{e \in E_w^c} 1 = \sum_{w \in W_{k,\frac{k}{2}+1}} 1 = \sharp W_{k,\frac{k}{2}+1}.$$

Então tudo que precisamos fazer é mostrar que $W_{k,\frac{k}{2}+1}$ possui $C_{\frac{k}{2}}$ elementos. Faremos a escolha dos representantes em $W_{k,\frac{k}{2}+1}$ do seguinte modo: se $s_1 s_2 \ldots s_k s_{k+1} \in W_{k,\frac{k}{2}+1}$ é um representante, temos que $s_1 = 1$, e os termos novos vem em ordem crescente, usando os menores números possíveis (aumenta-se uma unidade caso a próxima letra seja nova, e repete-se uma letra quando tiver que repetir). Evidentemente, só há uma palavra em cada classe satisfazendo tais condições. Então dado $w = s_1 s_2 \ldots s_k s_{k+1} \in W_{k,\frac{k}{2}+1}$, pela Proposição 2.35 sabemos que G_w é uma árvore. Para cada $i = 1 \ldots k$, defina $x_i = d(s_{i+1}, s_1)$ como sendo o comprimento do menor caminho entre os vértices s_1 e s_{i+1} em G_w. Como G_w é uma árvore com arestas $\{s_i, s_{i+1}\}$, temos que os vértices s_i e s_{i+1} devem ser adjacentes, portanto $x_i = x_{i-1} \pm 1$, para cada $i = 1 \ldots k$. Deste modo, temos que a sequência $(0, x_1, x_2, \ldots, x_k)$ é um caminho de Dyck de comprimento k. Usando a notação $\mathcal{D}_k$ para os caminhos de Dyck de comprimento k, temos que o mapa $\phi : W_{k,\frac{k}{2}+1} \to \mathcal{D}_k$, tal que $\phi(s_1 s_2 \ldots s_k s_{k+1}) = (0, x_1, x_2, \ldots, x_k)$, do modo como descrevemos acima, é um mapa que está bem definido.

Agora vamos definir o mapa $\tau : \mathcal{D}_k \to W_{k,\frac{k}{2}+1}$, de modo que um caminho de Dyck $(0, x_1, x_2, \ldots, x_k)$ seja levado numa palavra $w = s_1 s_2 \ldots s_k s_{k+1}$, onde $x_i = d(s_{i+1}, s_1)$, para todo $i = 1 \ldots k$. A definição de τ se baseia em obter uma sequência crescente de multigrafos $G_{w_1}, G_{w_2}, \ldots, G_{w_k}$ a partir da sequência $x_1, x_2, \ldots, x_k$, a figura abaixo mostra um exemplo da ideia desta construção.

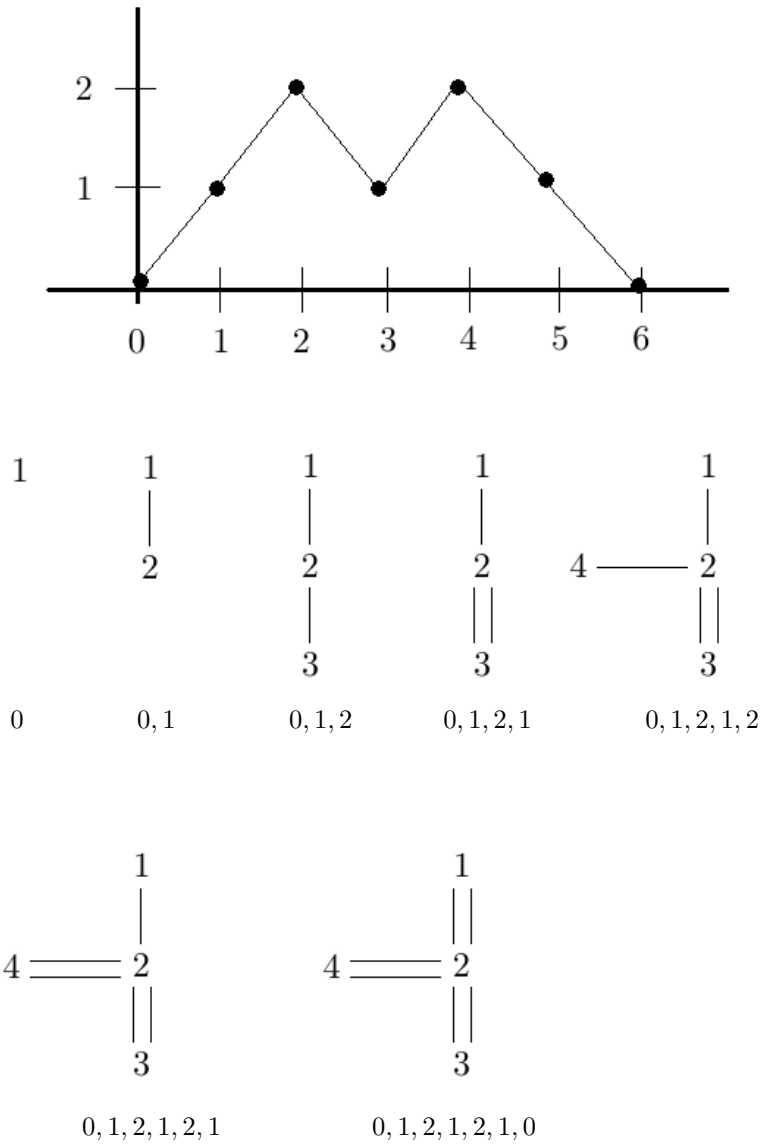

Figura 2.9: Sempre que $x_i = x_{i-1} + 1$, acrescentamos um novo ramo à árvore, indexado com o próximo número natural não indexado ainda. Caso tenhamos $x_i = x_{i-1} - 1$, devemos retroceder para o vértice pai. Neste exemplo, os números embaixo da sequência de multigrafos são os caminhos parciais de Dyck. A sequência de palavras é obtida pela sequência de vértices que vão aparecendo. Aqui temos que $w_1 = 1, 2$, $w_2 = 1, 2, 3$, $w_3 = 1, 2, 3, 2$, $w_4 = 1, 2, 3, 2, 4$, $w_5 = 1, 2, 3, 2, 4, 2$, $w_6 = 1, 2, 3, 2, 4, 2, 1$.

Seja $(0, x_1, x_2, \ldots, x_k) \in \mathcal{D}_k$ um caminho de Dyck de comprimento k, defina $w_1 = 1, 2$ e considere G_{w_1} o grafo com a única aresta $\{1, 2\}$, consideramos que o vértice indexado por "1" é o vértice pai. Cada vértice acrescentado ao multigrafo é também uma letra acrescentada à palavra (cada vértice irá ocorrer exatamente 2 vezes neste processo, o mesmo irá ocorrer com as letras). Para $i \geq 2$, obtemos G_{w_i} a partir de $G_{w_{i-1}}$ da seguinte maneira: seja $w_{i-1} = s_1 \ldots s_{i-1}$ a palavra associada a G_{w_i}. Se $x_i = x_{i-1} + 1$, acrescentamos um novo ramo à $G_{w_{i-1}}$, indexando-o com o menor número natural que ainda não foi indexado em $G_{w_{i-1}}$, neste caso definimos que $w_i = s_1 \ldots s_{i-1} s_i$, onde s_i é este novo índice. E caso tenhamos $x_i = x_{i-1} - 1$, retrocedemos para o vértice pai, neste caso definimos que $w_i = s_1 \ldots s_{i-1} s_i$,

onde s_i é o índice do vértice para onde retrocedemos. Com esta construção, obtemos uma sequência $G_{w_1}, G_{w_2}, \ldots, G_{w_k}$, associada às palavras $w_1, w_2, \ldots, w_k$, respectivamente. Escrevendo $w = w_k$, definimos que $\tau(0, x_1, x_2, \ldots, x_k) = w$.

Este processo apenas adiciona novos ramos e retrocede arestas, portanto todo G_{w_i} de fato será uma árvore. Além disso, para todo $i = 1 \ldots k$, temos que $d(s_{i+1}, s_1) = x_i$, pela própria construção que fizemos. Portanto, se $w = s_1 s_2 \ldots s_k s_1$ é a palavra obtida de $(0, x_1, x_2, \ldots, x_k)$ através desta construção, então o mapa $\tau : \mathcal{D}_k \to W_{k,\frac{k}{2}+1}$, dado por $\tau(0, x_1, x_2, \ldots, x_k) = w$ está bem definido, e é tal que $\phi(\tau(0, x_1, x_2, \ldots, x_k)) = \phi(w) = (0, x_1, x_2, \ldots, x_k)$. Note também que, dado $w \in W_{k,\frac{k}{2}+1}$, com $\phi(w) = (0, x_1, x_2, \ldots, x_k)$, temos que $\tau(\phi(w)) = \tau(0, x_1, x_2, \ldots, x_k) = w$. Daí podemos concluir que $\tau = \phi^{-1}$, portanto existe uma bijeção entre $\mathcal{D}_k$ e $W_{k,\frac{k}{2}+1}$, logo, $\sharp W_{k,\frac{k}{2}+1} = \sharp \mathcal{D}_k = C_{\frac{k}{2}}$. □

2.5 Sentenças e Multigrafos

Prosseguimos agora para o Lema 2.12. Começaremos com a seguinte proposição.

Proposição 2.41. Se

$$\lim_{N\to\infty} \mathrm{Var}\big(\langle \mu_N, x^k \rangle\big) = 0,$$

então

$$\lim_{N\to\infty} P\Big(\big|\langle \mu_N, x^k \rangle - E\langle \mu_N, x^k \rangle\big| > \varepsilon\Big) = 0.$$

Prova: A Desigualdade de Chebyshev nos diz que, dada uma variável aleatória X tal que $E(X) = m$, então para todo $\varepsilon > 0$,

$$P\Big(\big|X - m\big| > \varepsilon\Big) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2}.$$

Como $\langle \mu_N, x^k \rangle$ é uma variável aleatória, então para todo $\varepsilon > 0$, temos que

$$P\Big(\big|\langle \mu_N, x^k \rangle - E\langle \mu_N, x^k \rangle\big| > \varepsilon\Big) \leq \frac{\mathrm{Var}\big(\langle \mu_N, x^k \rangle\big)}{\varepsilon^2}.$$

Por outro lado, temos que $\mathrm{Var}\big(\langle \mu_N, x^k \rangle\big) = E\big(\langle \mu_N, x^k \rangle^2\big) - \big(E\langle \mu_N, x^k \rangle\big)^2$, então

$$P\Big(\big|\langle \mu_N, x^k \rangle - E\langle \mu_N, x^k \rangle\big| > \varepsilon\Big) \leq \frac{E\big(\langle \mu_N, x^k \rangle^2\big) - \big(E\langle \mu_N, x^k \rangle\big)^2}{\varepsilon^2}.$$

Por hipótese, temos que

$$\lim_{N\to\infty} \frac{E\big(\langle \mu_N, x^k \rangle^2\big) - \big(E\langle \mu_N, x^k \rangle\big)^2}{\varepsilon^2} = 0,$$

portanto

$$\lim_{N\to\infty} P\Big(\big|\langle \mu_N, x^k \rangle - E\langle \mu_N, x^k \rangle\big| > \varepsilon\Big) = 0. \qquad \square$$

Com esta proposição, sabemos que para provarmos o Lema 2.12 é suficiente provarmos que

$$\lim_{N\to\infty} E\big(\langle \mu_N, x^k\rangle^2\big) - \big(E\langle \mu_N, x^k\rangle\big)^2 = 0.$$

Iremos proceder de maneira análoga à prova do Lema 2.11. Começamos desenvolvendo a expressão $E\big(\langle \mu_N, x^k\rangle^2\big) - \big(E\langle \mu_N, x^k\rangle\big)^2$ em termos dos índices dos $X_{i,j}$, depois iremos usar técnicas de combinatória para simplificar a expressão e então calcularemos o limite quando $N \to \infty$.

Como já sabemos, $\langle \mu_N, x^k\rangle = \frac{1}{N}\mathrm{tr}\big(\frac{1}{\sqrt{N}}\mathbf{X}^k\big) = \frac{1}{N^{\frac{k}{2}+1}}\mathrm{tr}\big(\mathbf{X}^k\big)$, daí temos que

$$E\big(\langle \mu_N, x^k\rangle^2\big) - \big(E\langle \mu_N, x^k\rangle\big)^2 = E\left(\left(\frac{\mathrm{tr}(X^k)}{N^{\frac{k}{2}+1}}\right)^2\right) - \left(E\left(\frac{\mathrm{tr}(X^k)}{N^{\frac{k}{2}+1}}\right)\right)^2 =$$

$$= E\left(\left(\frac{\sum_{i_1,\ldots,i_k=1}^N X_{i_1,i_2}\ldots X_{i_k,i_1}}{N^{\frac{k}{2}+1}}\right)^2\right) - \left(E\left(\frac{\sum_{i_1,\ldots,i_k=1}^N X_{i_1,i_2}\ldots X_{i_k,i_1}}{N^{\frac{k}{2}+1}}\right)\right)^2 =$$

$$= \frac{1}{N^{k+2}}\left[\, E\left(\left(\sum_{i_1,\ldots,i_k=1}^N X_{i_1,i_2}\ldots X_{i_k,i_1}\right)^2\right) - \left(E\left(\sum_{i_1,\ldots,i_k=1}^N X_{i_1,i_2}\ldots X_{i_k,i_1}\right)\right)^2 \,\right] =$$

$$= \frac{1}{N^{k+2}}\left[\, E\left(\sum_{\substack{i_1,\ldots,i_k=1\\ j_1,\ldots,j_k=1}}^N X_{i_1,i_2}\ldots X_{i_k,i_1}\cdot X_{j_1,j_2}\ldots X_{j_k,j_1}\right) - \sum_{\substack{i_1,\ldots,i_k=1\\ j_1,\ldots,j_k=1}}^N E(X_{i_1,i_2}\ldots X_{i_k,i_1})\cdot E(X_{j_1,j_2}\ldots X_{j_k,j_1}) \,\right] =$$

$$= \frac{1}{N^{k+2}}\left[\, \sum_{\substack{i_1,\ldots,i_k=1\\ j_1,\ldots,j_k=1}}^N E(X_{i_1,i_2}\ldots X_{i_k,i_1}\cdot X_{j_1,j_2}\ldots X_{j_k,j_1}) - \sum_{\substack{i_1,\ldots,i_k=1\\ j_1,\ldots,j_k=1}}^N E(X_{i_1,i_2}\ldots X_{i_k,i_1})\cdot E(X_{j_1,j_2}\ldots X_{j_k,j_1}) \,\right] =$$

$$= \frac{1}{N^{k+2}} \sum_{\substack{i_1,\ldots,i_k=1\\ j_1,\ldots,j_k=1}}^N E(X_{i_1,i_2}\ldots X_{i_k,i_1}\cdot X_{j_1,j_2}\ldots X_{j_k,j_1}) - E(X_{i_1,i_2}\ldots X_{i_k,i_1})\cdot E(X_{j_1,j_2}\ldots X_{j_k,j_1}).$$

Para organizar os índices, iremos trabalhar com pares de N-palavras $i_1 i_2 \ldots i_k i_1$ e $j_1 j_2 \ldots j_k j_1$.

Definição 2.42. Seja $N \in \mathbb{N}$ fixado, uma *N-sentença* é uma sequência finita de N-palavras.

Considere as N-palavras $w_1, \ldots, w_k$, denotaremos por $a = (w_1, \ldots, w_k)$ uma N-sentença com k palavras. Por questão de simplicidade, chamaremos a apenas por *sentença*, e apenas iremos explicitar N quando houver alguma necessidade. As sentenças possuem definições semelhantes às definições de palavras.

Definição 2.43. O *suporte* de a é o conjunto das letras de todas as palavras w_i, é dado por $\mathrm{supp}(a) = \bigcup_{i=1}^k \mathrm{supp}(w_i)$.

Definição 2.44. O *peso* de a é a quantidade de letras distinas de todas as palavras w_i, é dado por $\mathrm{wt}(a) = \sharp\,\mathrm{supp}(a)$.

Podemos escrever cada palavra de a como $w_i = s_1^i s_2^i \dots s_{\ell(w_i)}^i$. Também podemos associar um multigrafo à sentença a, este multigrafo é dado por $G_a = (V_a, E_a)$, tal que $V_a = \mathrm{supp}(a)$ é o conjunto dos vértices e $E_a = \big\{\{s_j^i, s_{j+1}^i\} : \ i = 1 \dots k, \ j = 1 \dots \ell(w_i)\}\big\}$ é o conjunto das arestas.

Definição 2.45. O conjunto $\big\{e \in E_a : \ e = \{u, u\}, \ u \in V_a\big\}$ é chamado de conjunto das *auto arestas*, é denotado por E_a^s.

Definição 2.46. O conjunto $E_a \backslash E_a^s$ é chamado de conjunto das *arestas conectivas*, é denotado por E_a^c.

Definição 2.47. Para dar conta das arestas repetidas, denotamos por N_e^a a quantidade de vezes que ocorre a aresta $e \in E_a$ no multigrafo G_a.

Exemplo 2.48. Seja $N = 4$ e considere uma N-sentença $a = (w_1, w_2, w_3)$, com $w_1 = 1232141$, $w_2 = 13431$ e $w_3 = 31333123$. Temos que $\mathrm{supp}(a) = \{1, 2, 3, 4\}$ e $\mathrm{wt}(a) = 4$. Além disso, temos que o multigrafo associado é dado por $G_a = (V_a, E_a)$, tal que $V_a = \{1, 2, 3, 4\}$, $E_a = \big\{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{3, 3\}, \{3, 4\}\big\}$, $E_a^s = \big\{\{3, 3\}\big\}$ e $E_a^c = \big\{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{3, 4\}\big\}$. Por fim, note que $N_{\{1,3\}}^a = 5$, $N_{\{1,2\}}^a = N_{\{2,3\}}^a = 3$ e $N_{\{1,4\}}^a = N_{\{3,3\}}^a = N_{\{3,4\}}^a = 2$. Mostramos abaixo os multigrafos das palavras w_1, w_2, w_3 e o multigrafo G_a.

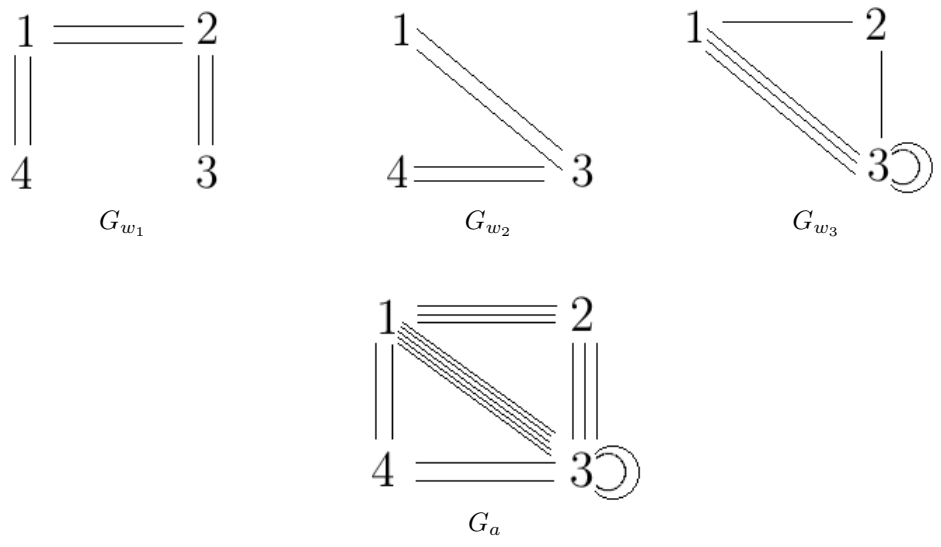

Figura 2.10: Multigrafo da N-sentença $a = (w_1, w_2, w_3)$, em geral, toda N-sentença $a = (w_1, \dots, w_N)$ dá origem à um multigrafo G_a que é obtido ao sobrepormos os multigrafos das N-palavras $G_{w_1}, \dots, G_{w_N}$.

Note que o multigrafo de uma sentença $a = (w_1, \dots, w_k)$ não é necessariamente igual ao multigrafo da concatenação das palavras $w_1, \dots, w_k$. Inclusive, é possível que o multigrafo de uma sentença seja um multigrafo desconexo, algo que não ocorreria se estivéssemos concatenando palavras. Além disso,

note que se $a = (w_1, \ldots, w_k)$ é uma N-sentença, então G_a possui exatamente $\sum_{i=1}^{k}(\ell(w_i) - 1)$ arestas, contando arestas repetidas.

Para cada par de índices $i, j \in \{1, \ldots, N\}^k$, com $i = (i_1, i_2, \ldots, i_k)$ e $j = (j_1, j_2, \ldots, j_k)$, definimos as palavras $w_i = i_1 i_2 \ldots i_k i_1$ e $w_j = j_1 j_2 \ldots j_k j_1$, a partir disso definimos $X_{w_i} = X_{i_1,i_2} X_{i_2,i_3} \ldots X_{i_{k-1},i_k} X_{i_k,i_1}$ e $X_{w_j} = X_{j_1,j_2} X_{j_2,j_3} \ldots X_{j_{k-1},j_k} X_{j_k,j_1}$. Seja a sentença $a_{i,j} = (w_i, w_j)$ e defina $X_{a_{i,j}} = X_{w_i} \cdot X_{w_j}$. Fazendo isso, temos que

$$E\big(\langle \mu_N, x^k \rangle^2\big) - \big(E\langle \mu_N, x^k \rangle\big)^2 = \frac{1}{N^{k+2}} \sum_{i,j \in \{1,\ldots,N\}^k} EX_{a_{i,j}} - EX_{w_i} EX_{w_j} =$$

$$= \frac{1}{N^{k+2}} \sum_{i,j \in \{1,\ldots,N\}^k} \left(E \prod_{e \in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e \in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e \in E_{w_j}} X_e^{N_e^{w_j}} \right).$$

Pelos mesmos motivos de antes, temos que é suficiente considerar apenas as palavras w_i e w_j tais que $N_e^{w_i} \geq 2$ para todo $e \in E_{w_i}$, e $N_e^{w_j} \geq 2$ para todo $e \in E_{w_j}$. Também é suficiente considerar apenas as sentenças $a_{i,j}$ tais que $N_e^{a_{i,j}} \geq 2$ para todo $e \in E_{a_{i,j}}$. Notando que, se $N_e^w = 1$ para alguma aresta e, tanto o produtório da sentença quanto o das palavras se anulam, portanto as duas condições mencionadas devem valer simultaneamente. Essas restrições simplificam a expressão acima. A seguinte proposição nos fornece mais uma simplificação.

Proposição 2.49. Se $E_{w_i} \cap E_{w_j} = \emptyset$, então $EX_{a_{i,j}} - EX_{w_i} EX_{w_j} = 0$.

Prova: Como G_{w_i} e G_{w_j} não possuem arestas em comum, temos que os índices de X_{w_i} são diferentes dos índices de X_{w_j}, levando em conta os índices invertidos (note que os índices invertidos devem ser levados em conta, pois $\mathbf{X}$ é simétrica). Portanto $EX_{a_{i,j}} = EX_{w_i} EX_{w_j}$, daí temos que $EX_{a_{i,j}} - EX_{w_i} EX_{w_j} = EX_{w_i} EX_{w_j} - EX_{w_i} EX_{w_j} = 0$. □

Também semelhante ao que fizemos com palavras, podemos considerar uma equivalência entre sentenças e pensar em termos de classes de equivalência.

Definição 2.50. Duas N-sentenças a_1, a_2 são ditas *equivalentes* se existir uma bijeção em $\{1, \ldots, N\}$ que leva uma sentença na outra. Denota-se por $a_1 \sim a_2$.

Note que se $a_1 = (w_1, \ldots, w_k)$ e $a_2 = (w'_1, \ldots, w'_{k'})$ são tais que $a_1 \sim a_2$, então[5] $w_i \sim w'_i$ para todo $i = 1 \ldots k$.

Exemplo 2.51. Seja $N = 7$ e considere as duas N-sentenças $a_1 = (12345671, 7167)$ e $a_2 = (76543217, 1721)$. Podemos definir $f : \{1, \ldots, 7\} \to \{1, \ldots, 7\}$, tal que $f(1) = 7$, $f(2) = 6$, $f(3) = 5$, $f(4) = 4, f(5) = 3$, $f(6) = 2$ e $f(7) = 1$. Temos que f é uma bijeção que leva a_1 em a_2 e vice-versa, portanto $a_1 \sim a_2$.

Vale também uma proposição análoga à Proposição 2.30.

Proposição 2.52. Sejam as sentenças $a_{i,j} = (w_i, w_j)$ e $a_{r,s} = (w_r, w_s)$ tais que $a_{i,j} \sim a_{r,s}$, então $EX_{a_{i,j}} - EX_{w_i} EX_{w_j} = EX_{a_{r,s}} - EX_{w_r} EX_{w_s}$.

Prova: Como $a_{i,j} \sim a_{r,s}$, então $w_i \sim w_r$ e $w_j \sim w_s$. Pela Proposição 2.30, temos que $EX_{w_i} = EX_{w_r}$ e $EX_{w_j} = EX_{w_s}$, portanto $EX_{w_i} EX_{w_j} = EX_{w_r} EX_{w_s}$. Nos resta apenas mostrar que $EX_{a_{i,j}} = EX_{a_{r,s}}$.

[5] Para esta conclusão ser válida, estamos considerando que a ordem das palavras é relevante. De modo que se $a_1 = (w_1, w_2)$ e $a_2 = (w_2, w_1)$, ainda podemos ter $a_1 \neq a_2$. Esta consideração faz sentido, uma vez que a matriz $\mathbf{X}$ pode nos dar sentenças invertidas e as nossas manipulações não vão levar essas inversões em conta.

Temos que

$$EX_{a_{i,j}} = \prod_{e\in E_{a_{i,j}}} EX_e^{N_e^{a_{i,j}}} = \prod_{e\in E^s_{a_{i,j}}} EX_e^{N_e^{a_{i,j}}} \cdot \prod_{e\in E^c_{a_{i,j}}} EX_e^{N_e^{a_{i,j}}}.$$

De modo análogo, temos que

$$EX_{a_{r,s}} = \prod_{e\in E_{a_{r,s}}} EX_e^{N_e^{a_{r,s}}} = \prod_{e\in E^s_{a_{r,s}}} EX_e^{N_e^{a_{r,s}}} \cdot \prod_{e\in E^c_{a_{r,s}}} EX_e^{N_e^{a_{r,s}}}.$$

A bijeção entre $a_{i,j}$ e $a_{r,s}$ nos garante que para cada auto aresta $e \in E^s_{a_{i,j}}$ existe uma outra auto aresta $e' \in E^s_{a_{r,s}}$ tal que $N_e^{a_{i,j}} = N_{e'}^{a_{r,s}}$, de modo que $EX_e^{N_e^{a_{i,j}}} = EX_{e'}^{N_{e'}^{a_{r,s}}}$. O mesmo vale para cada aresta conectiva. Daí temos que

$$\prod_{e\in E_{a^s_{i,j}}} EX_e^{N_e^{a_{i,j}}} = \prod_{e'\in E^s_{a_{r,s}}} EX_{e'}^{N_{e'}^{a_{r,s}}} \quad \text{e} \quad \prod_{e\in E_{a^c_{i,j}}} EX_e^{N_e^{a_{i,j}}} = \prod_{e'\in E^c_{a_{r,s}}} EX_{e'}^{N_{e'}^{a_{r,s}}},$$

portanto,

$$\prod_{e\in E_{a^s_{i,j}}} EX_e^{N_e^{a_{i,j}}} \cdot \prod_{e\in E_{a^c_{i,j}}} EX_e^{N_e^{a_{i,j}}} = \prod_{e'\in E^s_{a_{r,s}}} EX_{e'}^{N_{e'}^{a_{r,s}}} \cdot \prod_{e'\in E^c_{a_{r,s}}} EX_{e'}^{N_{e'}^{a_{r,s}}} \implies EX_{a_{i,j}} = EX_{a_{r,s}}.$$

Daí concluímos que $EX_{a_{i,j}} - EX_{w_i}EX_{w_j} = EX_{a_{r,s}} - EX_{w_r}EX_{w_s}$. □

Proposição 2.53. Seja $a = (w_1, w_2)$ uma N-sentença tal que $\mathrm{wt}(a) = t \leq N$, então existem $\frac{N!}{(N-t)!}$ sentenças equivalentes à a, incluindo a.

Prova: A prova é praticamente igual à prova da Proposição 2.31, pois a equivalência entre sentenças também depende unicamente das letras da sentença, não das palavras, pois estas também são determinadas unicamente pelas letras. Seja $\mathrm{supp}(a) = \{s_1, \ldots, s_t\}$ o conjunto das letras distintas de a. Podemos trocar algumas letras s_{i_j} por outras letras $s'_{i_j} \in \{1, \ldots, N\}$ e obter o conjunto $\{s'_{i_1}, \ldots, s'_{i_t}\}$, de modo que este conjunto ainda tenha t letras distintas. Consideramos a bijeção $f : \{1, \ldots, N\} \to \{1, \ldots, N\}$ tal que $f(s_{i_j}) = s'_{i_j}$. Deste modo, temos que $a' = f(s_1) \ldots f(s_k)$ é uma N-sentença equivalente à a.

Podemos obter todas as sentenças equivalentes à a a partir desta construção, além disso, podemos contar todas as possíveis sentenças que podemos obter deste modo. Queremos escolher $\{s'_{i_1}, \ldots, s'_{i_t}\}$ de modo que todas as letras sejam distintas entre si, além disso, só há N possíveis escolhas para cada letra. Então temos N escolhas para s'_{i_1}, $N-1$ escolhas para s'_{i_2}, e assim por diante, até chegarmos à s'_{i_t}, onde temos $N-t+1$ escolhas. Portanto, há $N \cdot (N-1) \ldots (N-t+1) = \frac{N!}{(n-t)!}$ possíveis N-palavras equivalentes à a, incluindo a própria palavra a nesta contagem. □

Para a definição a seguir, considere palavras w de comprimento $k+1$ tais que $w = s_1 s_2 s_3 \ldots s_{k-1} s_k s_1$, ou seja, palavras fechadas de comprimento $k+1$. Iremos trabalhar com palavras deste tipo durante todo o restante do capítulo. Também consideraremos sempre que $t \leq N$.

Definição 2.54. Denotamos por $W'_{k,t}$ o conjunto dos representantes de cada classe de equivalência das N-sentenças $a = (w_1, w_2)$ de peso t, constituída de duas palavras fechadas, de comprimento $k+1$, com $E_{w_1} \cap E_{w_2} \neq \emptyset$, e $N_e^a \geq 2$ para todo $e \in E_a$.

Com esta definição, podemos escrever

$$E\big(\langle\mu_N, x^k\rangle^2\big) - \big(E\langle\mu_N, x^k\rangle\big)^2 =$$
$$= \frac{1}{N^{k+2}} \sum_{t=1}^{N} \frac{N!}{(N-t)!} \sum_{a_{i,j}=(w_i,w_j)\in W'_{k,t}} \left(E \prod_{e\in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e\in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e\in E_{w_j}} X_e^{N_e^{w_j}} \right).$$

Como cada palavra w_1, w_2 possui no máximo k letras distintas, temos que t nao ultrapassa $2k$, então podemos reescrever a expressão acima como

$$\frac{1}{N^{k+2}} \sum_{t=1}^{2k} \frac{N!}{(N-t)!} \sum_{a_{i,j}=(w_i,w_j)\in W'_{k,t}} \left(E \prod_{e\in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e\in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e\in E_{w_j}} X_e^{N_e^{w_j}} \right).$$

Prova do Lema 2.12. Primeiramente, note que para toda sentença $a = (w_1, w_2) \in W'_{k,t}$, G_a é conexo pois $E_{w_1} \cap E_{w_2} \neq \emptyset$. Além disso, o total de arestas (contando arestas repetidas) é $\sum_{e\in E_a} N_e^a = \ell(w_1) - 1 + \ell(w_2) - 1 = 2k$.

No caso em que $t \geq k+2$, temos que G_a possui $t \geq k+2$ vértices, logo, possui no mínimo $k+1$ arestas distintas, e como $N_e^a \geq 2$ para todo $e \in E_a$, concluímos que G_a possui no mínimo $2(k+1)$ arestas (contando arestas repetidas). Isto não é possível, pois já vimos que o total de arestas é exatamente $2k$, logo, $W'_{k,t} = \emptyset$ quando $t \geq k+2$.

No caso em que $t \leq k+1$, note que

$$0 \leq \lim_{N\to\infty} \frac{N!}{(N-t)!N^{k+2}} \leq \lim_{N\to\infty} \frac{N!}{(N-k-1)!N^{k+2}} = 0.$$

Portanto,

$$\lim_{N\to\infty} E\big(\langle\mu_N, x^k\rangle^2\big) - \big(E\langle\mu_N, x^k\rangle\big)^2 =$$
$$= \lim_{N\to\infty} \frac{1}{N^{k+2}} \sum_{t=1}^{k+1} \frac{N!}{(N-t)!} \sum_{a_{i,j}=(w_i,w_j)\in W'_{k,t}} \left(E \prod_{e\in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e\in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e\in E_{w_j}} X_e^{N_e^{w_j}} \right) =$$
$$= \sum_{t=1}^{k+1} \left(\lim_{N\to\infty} \frac{N!}{(N-t)!N^{k+2}} \right) \sum_{a_{i,j}=(w_i,w_j)\in W'_{k,t}} \left(E \prod_{e\in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e\in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e\in E_{w_j}} X_e^{N_e^{w_j}} \right) = 0.$$

Temos que $\lim_{N\to\infty} \Big| E\big(\langle\mu_N, x^k\rangle^2\big) - \big(E\langle\mu_N, x^k\rangle\big)^2 \Big| = 0$, pois $\lim_{N\to\infty} \frac{N!}{(N-t)!N^{k+2}} = 0$, para todo $t = 1 \ldots k+1$, e o restante da expressão é constante. Então pela Proposição 2.41, concluímos que $\lim_{N\to\infty} P\Big(\big|\langle\mu_N, x^k\rangle - E\langle\mu_N, x^k\rangle\big| > \varepsilon\Big) = 0$. □

O Lema 2.12 nos garante que $\langle\mu_N, x^k\rangle - E\langle\mu_N, x^k\rangle \xrightarrow{P} 0$, isso nos mostra que as variáveis aleatórias λ_i^k satisfazem[6] a Lei Fraca dos Grandes Números, pois podemos escrever essa convergência como

$$\frac{\lambda_1^k + \ldots + \lambda_N^k}{N} - E\left(\frac{\lambda_1^k + \ldots + \lambda_N^k}{N}\right) \xrightarrow{P} 0.$$

[6]Estamos considerando $\lambda_i : \Omega \to \mathbb{R}$ tal que $\lambda_i(\omega) = \lambda_i(\mathbf{X}(\omega))$, ou seja, o i-ésimo autovalor de $\mathbf{X}(\omega)$.

O Lema 2.11 nos garante que

$$E\langle \mu_N, x^k \rangle \to \begin{cases} C_{\frac{k}{2}}, \text{ se } k \text{ é par} \\ 0, \text{ se } k \text{ é ímpar,} \end{cases}$$

com isto, podemos concluir que

$$\langle \mu_N, x^k \rangle \xrightarrow{P} \begin{cases} C_{\frac{k}{2}}, \text{ se } k \text{ é par} \\ 0, \text{ se } k \text{ é ímpar.} \end{cases}$$

Este resultado é relevante, pois nos permite estimar o autovalor médio esperado de $\mathbf{X}^k$. Quanto maior for a ordem da matriz, mais provável que o seu autovalor médio esteja próximo de $C_{\frac{k}{2}}$, se k for par, ou de zero, se k for ímpar. Abaixo, mostramos uma tabela com alguns experimentos computacionais. O parâmetro t representa a quantidade de amostras tomadas, o valor empírico é a média dos autovalores médios esperados[7] das amostras (o que irá nos fornecer uma boa aproximação para $\langle \mu_N, x^k \rangle$), e o valor teórico é $C_{\frac{k}{2}}$, se k for par, ou zero, se k for ímpar.

N	k	t	Valor empírico	Valor teórico
100	2	1000	1.0293	1
100	3	1000	0.00002	0
100	10	1000	50.9169	42
100	15	1000	-1.1185	0
100	20	1000	26699.9208	16796
4000	2	2	0.9997	1
4000	3	2	0.0007	0
4000	10	2	42.3169	42
4000	15	2	5.1147	0
4000	20	2	16829.2201	16796
8000	2	1	1.0005	1
8000	3	1	-0.0009	0
8000	10	1	42.0968	42
8000	15	1	3.3187	0
8000	20	1	16879.3136	16796

Tabela 2.1: Podemos estimar o autovalor médio esperado para potências k-ésimas de uma matriz de Wigner. O autovalor médio esperado de uma matriz é a média dos seus autovalores. É interessante que o número de Catalan seja este autovalor médio esperado quando a potência é um número par.

Os resultados são satisfatórios e concordam plenamente com a teoria desenvolvida até aqui. Agora podemos prosseguir para o principal teorema deste capítulo.

2.6 Lei do Semicírculo de Wigner

Definição 2.55. Seja μ uma distribuição de probabilidade cuja densidade é $\mu'(x) = \frac{1}{2\pi}\sqrt{4-x^2}\ \mathbf{I}_{|x|\leq 2}$, então μ é chamada de *distribuição do semicírculo.*

[7] Cada autovalor médio esperado é calculado pela média dos autovalores de cada matriz, o que nos dá uma aproximação para $\langle \mu_N, x^k \rangle$, então a média dos autovalores médios esperados de fato é uma boa aproximação para $\langle \mu_N, x^k \rangle$.

Junto desta definição, definimos $\langle \mu, x^k \rangle$ como sendo o *k-ésimo momento da distribuição do semicírculo.* Antes de provarmos a Lei do Semicírculo de Wigner, precisamos de um resultado a respeito desta distribuição.

Lema 2.56 (Fórmula da duplicação da função gamma). Seja $x \in \mathbb{R}$, $x > 0$, e considere $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\ dt$ a função gamma, então

$$\Gamma(2x) = \frac{2^{2x-1}\Gamma(x)\Gamma\big(x+\frac{1}{2}\big)}{\sqrt{\pi}}.$$

Prova: Inicialmente, considere a função beta, dada por $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = \int_0^1 (1-t)^{x-1}t^{y-1}\ dt$. Note que, fazendo a substituição $u^2 = t$ nesta integral, obtemos

$$B(x,y) = \int_0^1 (1-u^2)^{x-1}u^{2y-2}2u\ du = 2\int_0^1 (1-u^2)^{x-1}u^{2y-1}\ du. \qquad (*)$$

Temos que $B(x,x) = \int_0^1 (1-t)^{x-1}t^{x-1}\ dt$, agora fazemos a substituição $u = 2t-1$, obtendo

$$B(x,x) = \frac{1}{2}\int_{-1}^1 \Big(1-\frac{u+1}{2}\Big)^{x-1}\Big(\frac{u+1}{2}\Big)^{x-1}\ du = \frac{1}{2^{2x-1}}\int_{-1}^1 (1-u^2)^{x-1}\ du =$$
$$= \frac{1}{2^{2x-1}}\cdot 2\int_0^1 (1-u^2)^{x-1}\ du = \frac{1}{2^{2x-1}}B\Big(x,\frac{1}{2}\Big).$$

Note que esta última passagem é uma aplicação da equação $(*)$. A partir daí temos o seguinte.

$$B(x,x) = \frac{1}{2^{2x-1}}B\Big(x,\frac{1}{2}\Big) \implies \frac{\Gamma(x)\Gamma(x)}{\Gamma(2x)} = \frac{1}{2^{2x-1}}\frac{\Gamma(x)\Gamma(\frac{1}{2})}{\Gamma(x+\frac{1}{2})} \implies$$
$$\implies \frac{\Gamma(x)}{\Gamma(2x)} = \frac{1}{2^{2x-1}}\frac{\sqrt{\pi}}{\Gamma(x+\frac{1}{2})} \implies \Gamma(2x) = \frac{2^{2x-1}\Gamma(x)\Gamma\big(x+\frac{1}{2}\big)}{\sqrt{\pi}}. \qquad \square$$

Proposição 2.57.

$$\langle \mu, x^k \rangle = \begin{cases} C_{\frac{k}{2}}, \text{ se } k \text{ é par} \\ 0, \text{ se } k \text{ é ímpar.} \end{cases}$$

Prova: Faremos a prova em duas partes, primeiro para k ímpar e depois para k par. Primeiramente, note que $\langle \mu, x^k \rangle = \int_{-\infty}^\infty x^k\ d\mu = \int_{-\infty}^\infty x^k\mu'(x)\ dx = \int_{-2}^2 x^k\frac{1}{2\pi}\sqrt{4-x^2}\ dx$. Além disso, μ' é uma função par, de modo que, se k é ímpar, então x^k é ímpar e portanto $x^k\sqrt{4-x^2}$ também é ímpar, logo,

$$\int_{-2}^2 x^k\frac{1}{2\pi}\sqrt{4-x^2}\ dx = 0.$$

No caso em que k é par, temos que $x^k\sqrt{4-x^2}$ também é par, logo,

$$\int_{-2}^2 x^k\frac{1}{2\pi}\sqrt{4-x^2}\ dx = 2\int_0^2 x^k\frac{1}{2\pi}\sqrt{4-x^2}\ dx = \frac{1}{\pi}\int_0^2 x^k\sqrt{4-x^2}\ dx.$$

Fazendo a substituição $x = 2\sqrt{y}$, obtemos

$$\frac{1}{\pi}\int_0^1 (2\sqrt{y})^k\sqrt{4-4y}\frac{1}{\sqrt{y}}\,dy = \frac{2^{k+1}}{\pi}\int_0^1 y^{\frac{k-1}{2}}\sqrt{1-y}\,dy =$$

$$= \frac{2^{k+1}}{\pi}\int_0^1 y^{\frac{k+1}{2}-1}(1-y)^{\frac{3}{2}-1}\,dy = \frac{2^{k+1}\Gamma\left(\frac{k+1}{2}\right)\Gamma\left(\frac{3}{2}\right)}{\pi\cdot\Gamma\left(\frac{k+1}{2}+\frac{3}{2}\right)} = \frac{2^{k+1}\Gamma\left(\frac{k+1}{2}\right)\frac{\sqrt{\pi}}{2}}{\pi\cdot\Gamma\left(\frac{k+1}{2}+\frac{3}{2}\right)} = \frac{2^k\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi}\cdot\Gamma\left(\frac{k+4}{2}\right)}.$$

Podemos usar a fórmula da duplicação da função gamma, mostrada no Lema 2.56, daí temos que

$$\frac{2^{k-1}\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi}} = \Gamma(k) \implies \langle\mu, x^k\rangle = \frac{2^k\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi}\cdot\Gamma\left(\frac{k+4}{2}\right)} = \frac{2\Gamma(k)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{k+4}{2}\right)} =$$

$$= \frac{2(k-1)!}{\left(\frac{k}{2}-1\right)!\left(\frac{k+4}{2}-1\right)!} = \frac{k!}{\frac{k}{2}!\frac{k}{2}!}\cdot\frac{2^{\frac{k}{2}}}{k\left(\frac{k}{2}+1\right)} = \binom{k}{\frac{k}{2}}\cdot\frac{1}{\frac{k}{2}+1} = C_{\frac{k}{2}} \qquad \square$$

Para o teorema a seguir, considere que a sequência de distribuições μ_N é tal que cada μ_N está associada a uma matriz de ordem $N \times N$.

Teorema 2.58 (Lei do Semicírculo de Wigner). Dada a sequência aleatória de distribuições espectrais empíricas μ_N, de matrizes do ensemble de Wigner, temos que μ_N converge fracamente, em probabilidade, para μ, onde μ é a distribuição do semicírculo.

Prova: Sejam $B \geq 1$ e $\varepsilon > 0$, então pela Desigualdade de Chebyshev, temos que

$$P\Big(\langle\mu_N, |x|^k\mathbf{I}_{|x|>B}\rangle > \varepsilon\Big) \leq \frac{E\langle\mu_N, |x|^k\mathbf{I}_{|x|>B}\rangle}{\varepsilon}.$$

Note que $|x| > B \implies x^2 > B|x| \implies x^{2k} > B^k|x|^k \implies \frac{x^{2k}}{B^k} > |x|^k$. Portanto $|x|^k\mathbf{I}_{|x|>B} \leq \frac{x^{2k}}{B^k}$, e como estas duas funções são mensuráveis, podemos concluir que $\langle\mu_N, |x|^k\mathbf{I}_{|x|>B}\rangle \leq \langle\mu_N, \frac{x^{2k}}{B^k}\rangle$. Essa é uma desigualdade de variáveis aleatórias, portanto[8]

$$E\langle\mu_N, |x|^k\mathbf{I}_{|x|>B}\rangle \leq E\Big\langle\mu_N, \frac{x^{2k}}{B^k}\Big\rangle,$$

logo,

$$\frac{E\langle\mu_N, |x|^k\mathbf{I}_{|x|>B}\rangle}{\varepsilon} \leq \frac{E\langle\mu_N, x^{2k}\rangle}{B^k\varepsilon}.$$

Como $|x| > B \geq 1$, sabemos que $|x| < |x|^2 < \ldots < |x|^k < \ldots$, e pelo mesmo motivo de antes, temos que $\langle\mu_N, |x|^k\mathbf{I}_{|x|>B}\rangle$ é crescente em k. Então podemos concluir que

$$\limsup_{N\to\infty} P\Big(\langle\mu_N, |x|^k\mathbf{I}_{|x|>B}\rangle > \varepsilon\Big)$$

[8] Se X, Y são duas variáveis aleatórias não-negativas tais que $X \leq Y$, então $EX \leq EY$.

é não-decrescente em k. Fazendo $B = 5$ e notando que $C_k \leq \binom{2k}{k} \leq 2^{2k} = 4^k$, o Lema 2.11 nos garante que

$$\limsup_{N\to\infty} P\Big(\langle \mu_N, |x|^k \mathbf{I}_{|x|>B}\rangle > \varepsilon\Big) \leq \lim_{N\to\infty} \frac{E\langle \mu_N, x^{2k}\rangle}{5^k\varepsilon} = \frac{\langle \mu, x^{2k}\rangle}{5^k\varepsilon} = \frac{C_k}{5^k\varepsilon} \leq \frac{4^k}{5^k\varepsilon},$$

onde μ é a distribuição do semicírculo. Como o lado esquerdo desta desigualdade é não-decrescente em k, temos que

$$\limsup_{N\to\infty} P\Big(\langle \mu_N, |x|^k \mathbf{I}_{|x|>B}\rangle > \varepsilon\Big) \leq \lim_{k\to\infty} \frac{4^k}{5^k\varepsilon} = 0$$

portanto,

$$\limsup_{N\to\infty} P\Big(\langle \mu_N, |x|^k \mathbf{I}_{|x|>B}\rangle > \varepsilon\Big) = 0. \qquad (*)$$

Agora considere uma função $f \in C_b(\mathbb{R})$ que se anule fora do intervalo $[-5, 5]$. Seja $\delta > 0$, pelo Teorema da Aproximação de Weierstrass, existe um polinômio $Q(x)$ tal que

$$\sup_{x\in[-5,5]} |Q(x) - f(x)| \leq \frac{\delta}{8}.$$

Daí segue que

$$|\langle \mu_N, f\rangle - \langle \mu, f\rangle| = |\langle \mu_N, f\rangle - \langle \mu, f\rangle + \langle \mu_N, Q\rangle - \langle \mu_N, Q\rangle + E\langle \mu_N, Q\rangle - E\langle \mu_N, Q\rangle + \langle \mu, Q\rangle - \langle \mu, Q\rangle| \leq$$
$$|\langle \mu_N, f\rangle - \langle \mu_N, Q\rangle| + |\langle \mu, Q\rangle - \langle \mu, f\rangle| + |E\langle \mu_N, Q\rangle - \langle \mu, Q\rangle| + |\langle \mu_N, Q\rangle - E\langle \mu_N, Q\rangle|.$$

Note que

$$|\langle \mu_N, f\rangle - \langle \mu_N, Q\rangle| = |\langle \mu_N, (f-Q)\mathbf{I}_{x\leq|5|}\rangle + \langle \mu_N, (f-Q)\mathbf{I}_{x>|5|}\rangle| \leq$$
$$\langle \mu_N, |f-Q|\mathbf{I}_{x\leq|5|}\rangle + |\langle \mu_N, f\mathbf{I}_{x>|5|}\rangle - \langle \mu_N, Q\mathbf{I}_{x>|5|}\rangle| \leq \frac{\delta}{8} + |\langle \mu_N, Q\mathbf{I}_{x>|5|}\rangle|$$

e que

$$|\langle \mu, Q\rangle - \langle \mu, f\rangle| = |\langle \mu, (Q-f)\mathbf{I}_{x\leq|5|}\rangle + \langle \mu, (Q-f)\mathbf{I}_{|x|>5}\rangle| = |\langle \mu, (Q-f)\mathbf{I}_{x\leq|5|}\rangle| \leq \langle \mu, |Q-f|\mathbf{I}_{x\leq|5|}\rangle \leq \frac{\delta}{8},$$

portanto

$$|\langle \mu_N, f\rangle - \langle \mu, f\rangle| \leq \frac{\delta}{4} + |\langle \mu_N, Q\mathbf{I}_{x>|5|}\rangle| + |E\langle \mu_N, Q\rangle - \langle \mu, Q\rangle| + |\langle \mu_N, Q\rangle - E\langle \mu_N, Q\rangle|.$$

Desta desigualdade, segue que

$$P\Big(|\langle \mu_N, f\rangle - \langle \mu, f\rangle| > \delta\Big) \leq P\Big(|\langle \mu_N, Q\rangle - E\langle \mu_N, Q\rangle| > \frac{\delta}{4}\Big) + P\Big(|E\langle \mu_N, Q\rangle - \langle \mu, Q\rangle| > \frac{\delta}{4}\Big) + P\Big(|\langle \mu_N, Q\mathbf{I}_{|x|>5}\rangle| > \frac{\delta}{4}\Big).$$

Pelo Lema 2.12, temos que $\lim_{N\to\infty} P\Big(|\langle \mu_N, Q\rangle - E\langle \mu_N, Q\rangle| > \frac{\delta}{4}\Big) = 0$.

Pelo Lema 2.11 e Proposição 2.57, temos que $\lim_{N\to\infty} P\Big(|E\langle\mu_N, Q\rangle - \langle\mu, Q\rangle| > \frac{\delta}{4}\Big) = 0$.
Como $Q(x)$ é um polinômio, usamos $(*)$ em Q para concluir que

$$\lim_{N\to\infty} P\Big(|\langle\mu_N, Q\mathbf{I}_{|x|>5}\rangle| > \frac{\delta}{4}\Big) \leq \limsup_{N\to\infty} P\Big(\langle\mu_N, Q\mathbf{I}_{|x|>5}\rangle > \frac{\delta}{4}\Big) = 0.$$

A partir destas observações, segue que

$$\lim_{N\to\infty} P\Big(|\langle\mu_N, f\rangle - \langle\mu, f\rangle| > \delta\Big) = 0.$$

Portanto μ_N converge fracamente, em probabilidade, para μ. □

2.7 Resultados Extras

Em relação à Lei do Semicírculo de Wigner, há ainda um resultado mais forte: podemos trocar a convergência em probabilidade por convergência quase certa. Para provar isto, necessitamos de alguns resultados preliminares. Primeiramente, considere a expressão

$$\sum_{a_{i,j}=(w_i,w_j)\in W'_{k,t}} \left(E \prod_{e\in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e\in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e\in E_{w_j}} X_e^{N_e^{w_j}}\right)$$

tomada da prova do Lema 2.12. A partir dela, vamos obter alguns resultados mais específicos.

Proposição 2.59. Seguindo o mesmo contexto da prova do Lema 2.12, temos que se $t = k + 1$, então

$$\sum_{a_{i,j}=(w_i,w_j)\in W'_{k,t}} \left(E \prod_{e\in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e\in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e\in E_{w_j}} X_e^{N_e^{w_j}}\right) = 0.$$

Prova: Seja $a = (w_1, w_2)$, onde cada palavra da sentença tem comprimento $k + 1$, e o peso de a também é $k + 1$. Temos que G_a possui $k + 1$ vértices e $2k$ arestas. Por outro lado, sabemos que $2k = \ell(w_1) + \ell(w_2) - 2 = \sum_{e\in E_{w_1}} N_e^{w_1} + \sum_{e\in E_{w_2}} N_e^{w_2}$. O valor mínimo que esta expressão assume é quando $N_e^{w_i} = 2$ para todo $e \in E_{w_i}$, para $i = 1, 2$, neste caso temos que $\sum_{e\in E_{w_1}} N_e^{w_1} + \sum_{e\in E_{w_2}} N_e^{w_2} = 2k$, que é justamente o valor que devemos ter, portanto toda aresta deve ocorrer exatamente duas vezes. Como G_a é conexo por definição, temos que é uma árvore.

Sabemos que a primeira e última letra de w_1 são iguais, então o caminho que w_1 faz em G_a começa e termina no mesmo vértice, de modo que w_1 visita cada aresta exatamente duas vezes. Então o conjunto das arestas de w_1 é disjunto do conjunto das arestas de w_2, o que contradiz a definição de $W'_{k,t}$, logo, $W'_{k,t} = \emptyset$ quando $t = k + 1$. Daí temos que

$$\sum_{a_{i,j}=(w_i,w_j)\in W'_{k,k+1}} \left(E \prod_{e\in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e\in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e\in E_{w_j}} X_e^{N_e^{w_j}}\right) =$$

$$= \sum_{\emptyset} \left(E \prod_{e\in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e\in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e\in E_{w_j}} X_e^{N_e^{w_j}}\right) = 0.$$ □

Para o que segue, iremos definir a sequência de constantes $C(k)$, dadas por

$$C(k) = \sup_{N\in\mathbb{N}} N^2\big|E\big(\langle\mu_N, x^k\rangle^2\big) - \big(E\langle\mu_N, x^k\rangle\big)^2\big|.$$

Note que

$$\sup_{N\in\mathbb{N}} N^2\big|E\big(\langle\mu_N, x^k\rangle^2\big) - \big(E\langle\mu_N, x^k\rangle\big)^2\big| =$$

$$= \sup_{N\in\mathbb{N}}\left|\sum_{t=1}^{k+1}\frac{N^2N!}{(N-t)!N^{k+2}}\sum_{a_{i,j}=(w_i,w_j)\in W'_{k,t}}\left(E\prod_{e\in E_{a_{i,j}}}X_e^{N_e^{a_{i,j}}} - E\prod_{e\in E_{w_i}}X_e^{N_e^{w_i}}\cdot E\prod_{e\in E_{w_j}}X_e^{N_e^{w_j}}\right)\right| =$$

$$= \sup_{N\in\mathbb{N}}\left|\sum_{t=1}^{k}\frac{N^2N!}{(N-t)!N^{k+2}}\sum_{a_{i,j}=(w_i,w_j)\in W'_{k,t}}\left(E\prod_{e\in E_{a_{i,j}}}X_e^{N_e^{a_{i,j}}} - E\prod_{e\in E_{w_i}}X_e^{N_e^{w_i}}\cdot E\prod_{e\in E_{w_j}}X_e^{N_e^{w_j}}\right)\right|,$$

em que o cancelamento do termo em $t = k+1$ se deve à Proposição 2.59. Além disso, como

$$\begin{cases} \lim_{N\to\infty}\frac{N^2N!}{(N-t)!N^{k+2}} = 0, & \text{t=1...k-1} \\ \lim_{N\to\infty}\frac{N^2N!}{(N-t)!N^{k+2}} = 1, & \text{t=k,} \end{cases}$$

concluímos que $\{N^2\big|E\big(\langle\mu_N, x^k\rangle^2\big) - \big(E\langle\mu_N, x^k\rangle\big)^2\big|;\ N\in\mathbb{N}\}$ é um conjunto limitado, então sempre temos que $C(k) < \infty$. Por fim, note que $\big|E\big(\langle\mu_N, x^k\rangle^2\big) - \big(E\langle\mu_N, x^k\rangle\big)^2\big| \le \frac{C(k)}{N^2}$, para todo $k, N\in\mathbb{N}$, pela própria definição de $C(k)$.

Lema 2.60. Para todo $\varepsilon > 0$, temos que

$$P\Big(\bigcap_{j=1}^{\infty}\bigcup_{N=j}^{\infty}\{\omega\in\Omega:\ \big|\langle\mu_N, x^k\rangle - E\langle\mu_N, x^k\rangle\big| > \varepsilon\}\Big) = 0.$$

Prova: Dado $\varepsilon > 0$, a Desigualdade de Chebyshev nos garante que

$$P\Big(\big|\langle\mu_N, x^k\rangle - E\langle\mu_N, x^k\rangle\big| > \varepsilon\Big) \le \frac{\big|E\big(\langle\mu_N, x^k\rangle^2\big) - \big(E\langle\mu_N, x^k\rangle\big)^2\big|}{\varepsilon^2} \le \frac{C(k)}{\varepsilon^2N^2}.$$

Definindo $A_N = \{\omega\in\Omega:\ \big|\langle\mu_N, x^k\rangle - E\langle\mu_N, x^k\rangle\big| > \varepsilon\}$, temos que

$$\sum_{N=1}^{\infty}P(A_N) \le \sum_{N=1}^{\infty}\frac{C(k)}{\varepsilon^2N^2} = \frac{C(k)}{\varepsilon^2}\sum_{N=1}^{\infty}\frac{1}{N^2} = \frac{C(k)}{\varepsilon^2}\cdot\frac{\pi^2}{6} < \infty.$$

Pelo Lema Borel-Cantelli, podemos concluir que $P\Big(\limsup_{N\to\infty}A_N\Big) = 0$, portanto

$$P\Big(\bigcap_{j=1}^{\infty}\bigcup_{N=j}^{\infty}A_N\Big) = P\Big(\bigcap_{j=1}^{\infty}\bigcup_{N=j}^{\infty}\{\omega\in\Omega:\ \big|\langle\mu_N, x^k\rangle - E\langle\mu_N, x^k\rangle\big| > \varepsilon\}\Big) = 0. \qquad \square$$

Teorema 2.61. A Lei do Semicírculo de Wigner continua válida ao trocarmos a convergência fraca, em probabilidade, pela convergência fraca, quase certa.

Prova: A prova é essencialmente a mesma, vamos começar a partir da desigualdade mostrada abaixo, que obtivemos da prova anterior da Lei do Semicírculo de Wigner.

$$|\langle \mu_N, f\rangle - \langle \mu, f\rangle| \leq \frac{\delta}{4} + |\langle \mu_N, Q\mathbf{I}_{x>|5|}\rangle| + |E\langle \mu_N, Q\rangle - \langle \mu, Q\rangle| + |\langle \mu_N, Q\rangle - E\langle \mu_N, Q\rangle|$$

Sabemos que também que $\lim_{N\to\infty}\langle \mu_N, |x|^k\mathbf{I}_{|x|>B}\rangle = 0$ em um conjunto de medida total, além disso, pelo Lema 2.11, temos que $\lim_{N\to\infty}|E\langle \mu_N, Q\rangle - \langle \mu, Q\rangle| = 0$, por fim, o Lema 2.60 nos garante que, dado $\varepsilon > 0$ qualquer, existe um conjunto de medida total onde $\lim_{N\to\infty}|\langle \mu_N, Q\rangle - E\langle \mu_N, Q\rangle| < \varepsilon$. Com estes resultados, podemos concluir que

$$|\langle \mu_N, f\rangle - \langle \mu, f\rangle| \leq \frac{\delta}{4} + |\langle \mu_N, Q\mathbf{I}_{x>|5|}\rangle| + |E\langle \mu_N, Q\rangle - \langle \mu, Q\rangle| + |\langle \mu_N, Q\rangle - E\langle \mu_N, Q\rangle| \leq$$

$$\leq \frac{\delta}{4} + \frac{\delta}{4} + \frac{\delta}{4} + \frac{\delta}{4} = \delta$$

em um conjunto de medida total. Portanto μ_N tem convergência fraca, quase certa, para μ. □

Cabe fazer uma última observação, a Lei do Semicírculo de Wigner ainda pode ser generalizada para outras variâncias, estamos nos referindo à variância das entradas acima da diagonal. Nós a tomamos como sendo 1, mas se a tomarmos como sendo algum $\sigma^2 > 0$ qualquer, apenas o Lema 2.11 é afetado. Em particular, no início da prova do Lema 2.11 teremos

$$\lim_{N\to\infty} E\langle \mu_N, x^k\rangle = \sum_{w\in W_{k,\lfloor \frac{k}{2}\rfloor+1}}\left(\prod_{e\in E_w^c} EX_e^{N_e^w}\right) = \sum_{w\in W_{k,\frac{k}{2}+1}}\left(\prod_{e\in E_w^c} EX_e^2\right) =$$

$$= \sum_{w\in W_{k,\frac{k}{2}+1}}\prod_{e\in E_w^c}\sigma^2 = \sum_{w\in W_{k,\frac{k}{2}+1}}(\sigma^2)^{\frac{k}{2}} = \sigma^k\cdot \sharp W_{k,\frac{k}{2}+1} = \sigma^k\cdot C_{\frac{k}{2}}$$

quando k é par. Enquanto que o caso de k ímpar permanece inalterado.

Neste caso, a distribuição do semicírculo μ é tal que $\mu'(x) = \frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\,\mathbf{I}_{|x|\leq 2\sigma}$, com o seu k-ésimo momento dado por

$$\langle \mu, x^k\rangle = \begin{cases} \sigma^k\cdot C_{\frac{k}{2}}, & \text{se } k \text{ é par} \\ 0, & \text{se } k \text{ é ímpar.}\end{cases}$$

Deste modo, temos uma família de curvas descrevendo a distribuição dos autovalores para cada ensemble. Apesar do teorema original ser chamado de "semicírculo", note que, na verdade, estas curvas são semielipses, incluindo a curva de quando a variância é 1. Mostramos abaixo o gráfico de algumas destas curvas, tomando a variância como parâmetro.

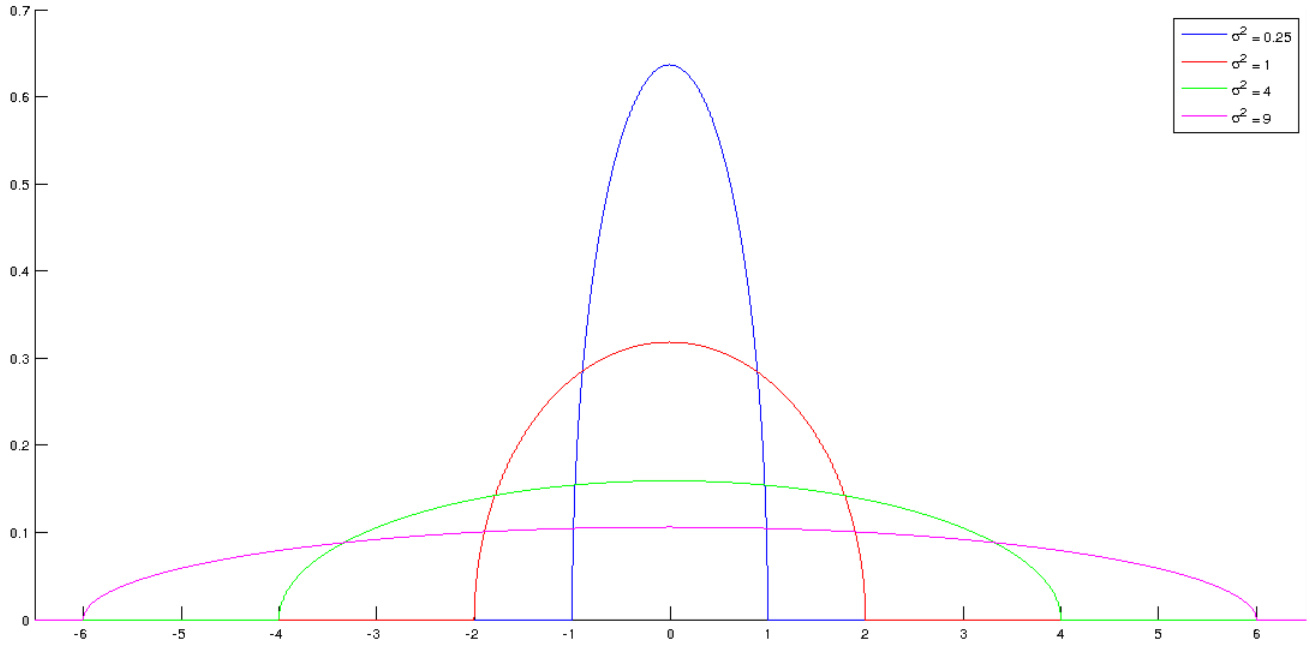

Figura 2.11: As semielipses obtidas pela Lei do Semicírculo de Wigner para outras vâriancias são dadas pela função $\frac{1}{2\pi\sigma^2}\sqrt{4\sigma^2 - x^2}\ \mathbf{I}_{|x|\leq 2\sigma}$. Nesta figura, temos as semielpises para $\sigma^2 = 0.25,\ 1,\ 4,\ 9$.

2.8 Lei do Semicírculo de Wigner Hermitiano

No início do capítulo 1, falamos um pouco da história por trás da Lei do Semicírculo de Wigner, que basicamente foi a tentativa de Wigner de modelar estatisticamente o comportamento de núcleos pesados. O modelo original tratava de matrizes hermitianas, mas em todo o texto tratamos apenas de matrizes simétricas. Esta limitação não foi sem motivo, pois a Lei do Semicírculo de Wigner ainda é válida para matrizes hermitianas, e a prova deste fato é essencialmente igual à prova do caso simétrico, com algumas pequenas alterações que comentaremos agora.

Sejam X, Y variáveis aleatórias reais, com valores esperados m_X, m_Y e variâncias σ_X^2, σ_Y^2, respectivamente. Então definimos a variável aleatória complexa Z dada por $Z = X + iY$, com valor esperado definido por $m_Z = m_X + i \cdot m_Y$ e variância $\sigma_Z^2 = E\big((Z - m_Z)\overline{(Z - m_Z)}\big) = E\big(|Z - m_Z|^2\big) = E\big(|(X - m_X) + i(Y - m_Y)|^2\big) = E\big((X - m_X)^2 + (Y - m_Y)^2\big) = E\big((X - m_X)^2\big) + E\big((Y - m_Y)^2\big) = \sigma_X^2 + \sigma_Y^2$.

Nesta seção, iremos tratar apenas do ensemble de Wigner Hermitiano, onde as entradas da diagonal são variáveis aleatórias reais, i.i.d., com valor esperado zero, e as entradas acima da diagonal são variáveis aleatórias complexas Z, i.i.d., com valor esperado $m_Z = EZ = 0$ e variância $\sigma_Z^2 = E|Z|^2 = 1$. Escrevendo $Z = X + iY$, temos que X, Y são independentes, mas não possuem necessariamente a mesma distribuição, além disso, $m_X = m_Y = 0$. Também iremos considerar que $\sigma_X^2 = \sigma_Y^2$, deste modo, se Z for uma entrada fora da diagonal, então $EZ^2 = E(X + iY)^2 = \sigma_X^2 - \sigma_Y^2 = 0$. Afim de manter a coerência com os resultados

anteriores, denotaremos por $\mathbf{X}$ as matrizes deste ensemble e por $X_{i,j}$ as suas entradas.

Definição 2.62. Seja $\mathbf{X}$ uma matriz do ensemble de Wigner Hermitiano, de ordem $N \times N$, então a matriz normalizada $\frac{1}{\sqrt{N}}\mathbf{X}$ é chamada de *matriz de Wigner Hermitiana*.

Sempre que nos referirmos a uma matriz aleatória $\mathbf{X}$, considere que ela é uma matriz do ensemble de Wigner Hermitiano, mas não necessariamente normalizada, a normalização será tratada de modo explícito.

Através de experimentos computacionais, geramos aleatoriamente várias matrizes de Wigner Hermitianas, com a mesma ordem, e então plotamos os histogramas normalizados dos seus autovalores. Abaixo, mostramos três histogramas que obtivemos pelos experimentos feitos.

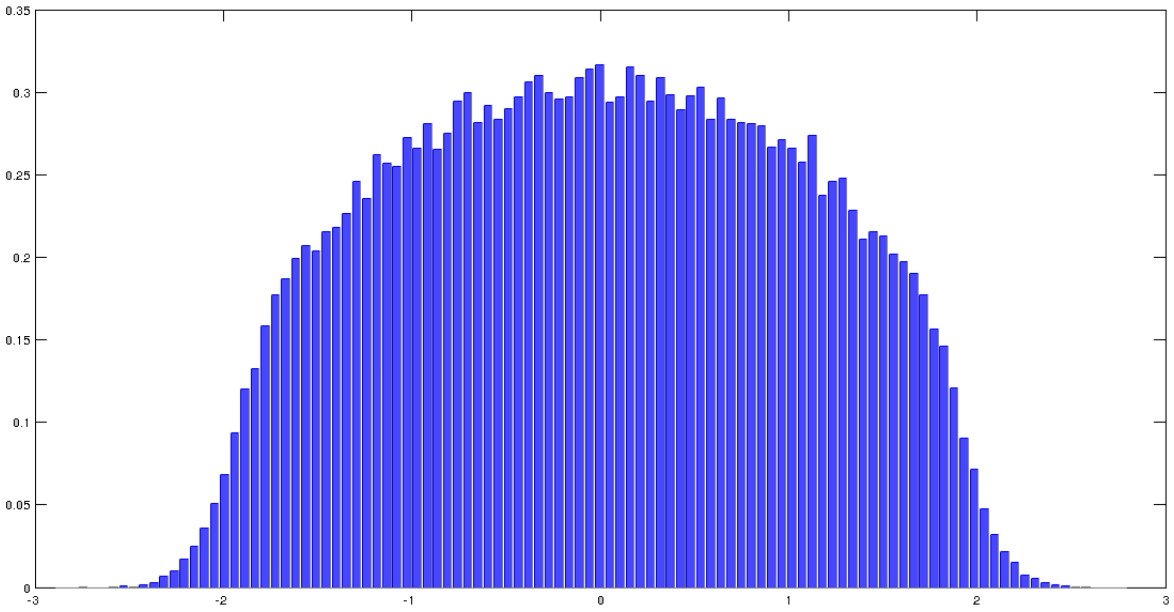

Figura 2.12: 10000 matrizes de ordem 10×10

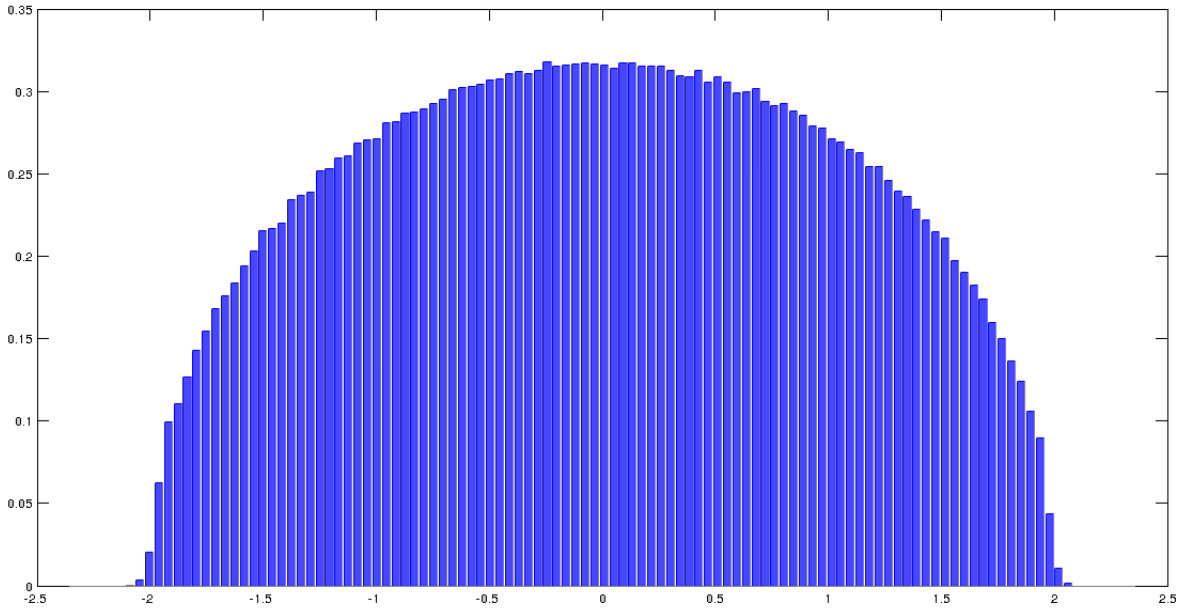

Figura 2.13: 5000 matrizes de ordem 100×100

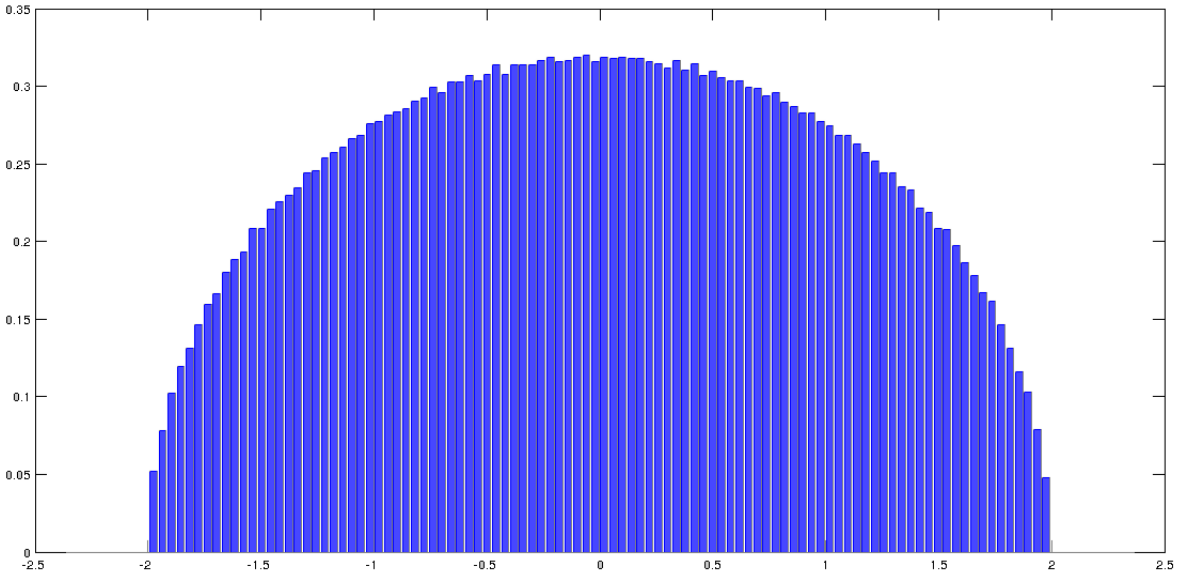

Figura 2.14: 5 matrizes de ordem 5000×5000

Todos os histogramas foram gerados no MATLAB, através do código mostrado abaixo. A variável t é a quantidade de vezes que o experimento foi repetido (de modo equivalente, é a quantidade de matrizes que foram geradas) e N é a ordem da matriz.

```
t=10000;
N=100;
E=zeros(0);
for i=1:t
    a= triu(random('Normal', 0, 1/sqrt(2), N));
    b= 1i*triu(random('Normal', 0, 1/sqrt(2), N));
    a=a+b;
    a=(1/sqrt(N))*(a+a');
    E=[E,(eig(a))'];
end
[n,x]=hist(E,100);
delta = x(2)-x(1);
bar(x, n/(t*N*delta))
```

Pelo que podemos ver, o histograma tende para o formato de um semicírculo, obedecendo à Lei do Semicírculo de Wigner mesmo quando as matrizes são hermitianas.

Prosseguindo do mesmo modo que no caso simétrico, sejam $\lambda_1 \leq \ldots \leq \lambda_N$ os autovalores (reais) de $\mathbf{X}$ e seja $\mu_{\mathbf{X}} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}$ a sua distribuição espectral empírica. Todas as considerações iniciais são análogas às feitas anteriormente, além disso, não é difícil ver que a maioria dos resultados preliminares continuam valendo no caso hermitiano. Para provar que a Lei do Semicírculo de Wigner ainda vale, usaremos a mesma ideia: tomar uma sequência crescente de matrizes aleatórias e analisar a convergência de suas distribuições espectrais empíricas, e para isso usaremos o método dos momentos. Devemos provar que os Lemas 2.11 e 2.12 continuam valendo para matrizes hermitianas, para depois provar a Lei do Semicírculo de Wigner. Nosso objetivo nesta seção não é obter provas rigorosas, mas sim comentar as principais alterações nas provas e como lidar com elas, pois o restante é essencialmente idêntico ao caso simétrico.

Lema 2.11'. Usando as mesmas ideias de combinatória, podemos ver que a manipulação dos índices através de palavras só não funciona exatamente igual por causa dos índices invertidos. No caso anterior usamos o fato de que $X_{i,j} = X_{j,i}$, desta vez apenas podemos garantir que $X_{i,j} = \overline{X_{j,i}}$.

Seja a palavra $w = i_1 i_2 i_3 \ldots i_{k-1} i_k i_1$, definimos o *caminho* induzido por w no multigrafo G_w como sendo a sequência de vértices $i_1, i_2, i_3, \ldots, i_{k-1}, i_k, i_1$. Então dada uma aresta $e = \{u, u'\} \in E_w$, com $u < u'$, definimos N_e^{w+} como sendo a quantidade de vezes que o par u, u' ocorre no caminho induzido por w, de modo análogo[9], definimos N_e^{w-} como sendo a quantidade de vezes que o par u', u ocorre no caminho induzido por w (de modo equivalente, podemos interpretar N_e^{w-} como sendo a quantidade de vezes que o par u, u' ocorre quando percorremos o caminho no sentido contrário). Note que não há necessidade de falar em caminhos quando $u = u'$, pois neste caso estamos nos referindo à um elemento da diagonal da matriz, que é real, então as mesmas conclusões do caso simétrico permanecem válidas. Por fim, note que $N_e^{w+} + N_e^{w-} = N_e^{w}$.

Exemplo 2.63. Considere novamente o exemplo 2.25, onde $w = 212443312$. Temos que o caminho induzido por w é $2, 1, 2, 4, 4, 3, 3, 1, 2$, note também que $N_{\{1,2\}}^{w+} = 2$, $N_{\{1,2\}}^{w-} = N_{\{1,3\}}^{w-} = N_{\{2,4\}}^{w+} = N_{\{3,4\}}^{w-} = 1$ e $N_{\{1,3\}}^{w+} = N_{\{1,4\}}^{w+} = N_{\{1,4\}}^{w-} = N_{\{2,3\}}^{w+} = N_{\{2,3\}}^{w-} = N_{\{2,4\}}^{w-} = N_{\{3,4\}}^{w+} = 0$, portanto

[9] No caso de matrizes simétricas, poderíamos ter definido N_e^w como sendo a soma da quantidade de vezes que u, u' e u', u ocorrem no caminho induzido por w, as definições são equivalentes. Note também que nós usamos intuitivamente a ideia de caminho na prova da Proposição 2.59. Naquele momento o uso era mais simples e não foi necessário dar uma definição formal, mas veremos que aqui essa definição mais formal será útil.

$$\prod_{e\in E_w} X_e = X_{2,1}X_{1,2}X_{2,4}X_{4,4}X_{4,3}X_{3,3}X_{3,1}X_{1,2} = X_{4,4}X_{3,3}\cdot\overline{X_{3,4}}X_{2,4}\overline{X_{1,3}}X_{1,2}^2\overline{X_{1,2}} =$$

$$= X_{4,4}^{N^w_{\{4,4\}}}X_{3,3}^{N^w_{\{3,3\}}}\cdot\overline{X_{3,4}}^{N^{w-}_{\{3,4\}}}X_{2,4}^{N^{w+}_{\{2,4\}}}\overline{X_{1,3}}^{N^{w-}_{\{1,3\}}}X_{1,2}^{N^{w+}_{\{1,2\}}}\overline{X_{1,2}}^{N^{w-}_{\{1,2\}}}.$$

Generalizando, podemos usar estas notações para escrever

$$\langle\mu_N, x^k\rangle = \frac{1}{N^{\frac{k}{2}+1}}\mathrm{tr}(\mathbf{X}^k) = \frac{1}{N^{\frac{k}{2}+1}}\sum_{i\in\{1,\dots,N\}^k}\prod_{e\in E_w} X_e^{N_e^w} =$$

$$= \frac{1}{N^{\frac{k}{2}+1}}\sum_{i\in\{1,\dots,N\}^k}\left(\prod_{e\in E^s_{w_i}} X_e^{N_e^{w_i}}\cdot\prod_{e\in E^c_{w_i}} X_e^{N_e^{w_i+}}\cdot\overline{X_e}^{N_e^{w_i-}}\right).$$

Pelo mesmo motivo do caso real, devemos ter $N_e^{w_i}\geq 2$, para todo $e\in E_{w_i}$. Além disso, dado X_e, com $e\in E^c_{w_i}$, temos que $X_e^{N_e^{w_i+}}$ e $X_e^{N_e^{w_i-}}$ não são independentes, então devemos considerar os termos do tipo $E(X_e^{N_e^{w_i+}}\cdot\overline{X_e}^{N_e^{w_i-}})$. Sabemos que $EX_e^2 = 0$ quando $e\in E^c_{w_i}$, então se $N_e^{w_i}=2$, não queremos considerar os casos $N_e^{w_i+}=2$ e $N_e^{w_i-}=2$. Segue que se $N_e^{w_i}=2$, então $N_e^{w_i+}=N_e^{w_i-}=1$.

Uma diferença relevante aqui é que a equivalência entre palavras não garante valores esperados iguais, ou seja, a Proposição 2.30 não é necessariamente válida. Mais especificamente, sejam $w_i\neq w_j$ duas palavras tais que $w_i\sim w_j$, não é necessário que $E(X_e^{N_e^{w_i+}}\cdot\overline{X_e}^{N_e^{w_i-}}) = E(X_e^{N_e^{w_j+}}\cdot\overline{X_e}^{N_e^{w_j-}})$. Apesar disto, temos que

$$|E(X_e^{N_e^{w_i+}}\cdot\overline{X_e}^{N_e^{w_i-}})| \leq E|X_e^{N_e^{w_i+}}\cdot\overline{X_e}^{N_e^{w_i-}}| = E\big(|X_e|^{N_e^{w_i}}\big),$$

então podemos nos valer desta desigualdade em cada classe de equivalência $W_{k,t}$. Daí temos que

$$|E\langle\mu_N, x^k\rangle| = \left|\frac{1}{N^{\frac{k}{2}+1}}\sum_{i\in\{1,\dots,N\}^k}\left(\prod_{e\in E^s_{w_i}} EX_e^{N_e^{w_i}}\cdot\prod_{e\in E^c_{w_i}} E\Big(X_e^{N_e^{w_i+}}\cdot\overline{X_e}^{N_e^{w_i-}}\Big)\right)\right| \leq$$

$$\leq \frac{1}{N^{\frac{k}{2}+1}}\sum_{i\in\{1,\dots,N\}^k}\left(\prod_{e\in E^s_{w_i}} E\big(|X_e^{N_e^{w_i}}|\big)\cdot\prod_{e\in E^c_{w_i}} E\big(|X_e|^{N_e^{w_i}}\big)\right)$$

A partir disto temos que

$$\lim_{N\to\infty}|E\langle\mu_N, x^k\rangle| \leq \lim_{N\to\infty}\frac{1}{N^{\frac{k}{2}+1}}\sum_{t=1}^{\lfloor\frac{k}{2}\rfloor+1}\frac{N!}{(N-t)!}\sum_{w\in W_{k,t}}\left(\prod_{e\in E^s_w} E|X_e|^{N_e^w}\cdot\prod_{e\in E^c_w} E|X_e|^{N_e^w}\right).$$

Portanto vale um resultado análogo ao da Proposição 2.34.

$$\lim_{N\to\infty} E\langle\mu_N, x^k\rangle = 0, \text{ se } k \text{ é ímpar}$$

$$\lim_{N\to\infty}|E\langle\mu_N, x^k\rangle| \leq \sum_{w\in W_{k,\frac{k}{2}+1}}\left(\prod_{e\in E^s_w} E\big(|X_e|^{N_e^w}\big)\cdot\prod_{e\in E^c_w} E\big(|X_e|^{N_e^w}\big)\right), \text{ se } k \text{ é par}$$

Considerando k par e a classe $W_{k,\frac{k}{2}+1}$, pela Proposição 2.35, sabemos que $E_w^s = \emptyset$. Além disso, G_w é uma árvore, com $N_e^w = 2$ para todo $e \in E_w$, portanto toda aresta ocorre duas vezes. Concluímos então que $N_e^{w+} = N_e^{w-} = 1$ para todo $e \in E_e^c$, mas neste caso, temos que $E(X_e \cdot \overline{X_e}) = E(|X_e|^2)$ para todo $e \in E_w^c$, então podemos escrever

$$\sum_{w \in W_{k,\frac{k}{2}+1}} \Big(\prod_{e \in E_w^c} E\big(|X_e|^2\big) \Big) = \sum_{w \in W_{k,\frac{k}{2}+1}} \Big(\prod_{e \in E_w^c} E\big(X_e \overline{X_e}\big) \Big) = \lim_{N \to \infty} E\langle \mu_N, x^k \rangle.$$

Por outro lado, temos que

$$\sum_{w \in W_{k,\frac{k}{2}+1}} \Big(\prod_{e \in E_w^c} E\big(|X_e|^2\big) \Big) = \sum_{w \in W_{k,\frac{k}{2}+1}} \prod_{e \in E_w^c} 1 = \sum_{w \in W_{k,\frac{k}{2}+1}} 1 = \sharp W_{k,\frac{k}{2}+1} = C_{\frac{k}{2}},$$

portanto $\lim_{N\to\infty} E\langle \mu_N, x^k \rangle = C_{\frac{k}{2}}$ se k for par, e zero se k for ímpar.

Lema 2.12'. Nós iremos usar as mesmas ideias do caso simétrico, mas dando atenção aos termos do tipo $E(X_e^{N_e^{w_i+}} \cdot \overline{X_e}^{N_e^{w_i-}})$ que irão aparecer quando tentarmos calcular $\lim_{N\to\infty} \big|E\big(\langle \mu_N, x^k \rangle^2\big) - \big(E\langle \mu_N, x^k \rangle\big)^2\big|$. A Proposição 2.49 ainda vale, mas a 2.52 não necessariamente, pelo mesmos motivos que comentamos no Lema 2.11. A equivalência entre sentenças não é suficiente para garantir a igualdade entre os valores esperados da proposição. Ainda assim, temos o seguinte.

$$\big|E\big(\langle \mu_N, x^k \rangle^2\big) - \big(E\langle \mu_N, x^k \rangle\big)^2\big| =$$

$$= \left| \frac{1}{N^{k+2}} \sum_{\substack{i_1,\dots,i_k=1 \\ j_1,\dots,j_k=1}}^{N} E(X_{i_1,i_2} \dots X_{i_k,i_1} \cdot X_{j_1,j_2} \dots X_{j_k,j_1}) - E(X_{i_1,i_2} \dots X_{i_k,i_1}) \cdot E(X_{j_1,j_2} \dots X_{j_k,j_1}) \right| =$$

$$= \left| \frac{1}{N^{k+2}} \sum_{i,j \in \{1,\dots,N\}^k} \left(E \prod_{e \in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e \in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e \in E_{w_j}} X_e^{N_e^{w_j}} \right) \right| \leq$$

$$\leq \frac{1}{N^{k+2}} \sum_{i,j \in \{1,\dots,N\}^k} \left| \left(E \prod_{e \in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e \in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e \in E_{w_j}} X_e^{N_e^{w_j}} \right) \right| =$$

$$= \frac{1}{N^{k+2}} \sum_{t=1}^{2k} \frac{N!}{(N-t)!} \sum_{a_{i,j}=(w_i,w_j) \in W'_{k,t}} \left| \left(E \prod_{e \in E_{a_{i,j}}} X_e^{N_e^{a_{i,j}}} - E \prod_{e \in E_{w_i}} X_e^{N_e^{w_i}} \cdot E \prod_{e \in E_{w_j}} X_e^{N_e^{w_j}} \right) \right|.$$

Fazendo $N \to \infty$, temos que a prova é a mesma, e o resultado é o mesmo.

Lei do Semicírculo de Wigner. A prova para o caso hermitiano é exatamente a mesma, tanto a da convergência em probabilidade quanto a convergência quase certa. Portanto a Lei do Semicírculo de Wigner também vale para matrizes de Wigner Hermitianas.

Capítulo 3

Aplicações

3.1 Consequências Básicas

Seja $\frac{1}{\sqrt{N}}\mathbf{X}$ uma matriz de Wigner, com autovalores $\lambda_1 \leq \ldots \leq \lambda_N$, e seja $E(I)$ a quantidade esperada de autovalores $\frac{1}{\sqrt{N}}\mathbf{X}$ em um intervalo I. O teorema 2.7 nos garante[1] que é possível usar a Lei do Semicírculo de Wigner para concluir que

$$\lim_{N\to\infty} \frac{E(I)}{N} = \int_I \frac{\sqrt{4-x^2}}{2\pi}\mathbf{I}_{|x|\leq 2}\, dx.$$

Agora seja $\mathbf{X}(\omega) = \mathbf{A}$ uma matriz hermitiana $N \times N$, com autovalores $\lambda_1 \leq \ldots \leq \lambda_N$. Se N for grande o suficiente, podemos usar a Lei do Semicírculo de Wigner para estimar a quantidade esperada de autovalores de $\frac{1}{\sqrt{N}}\mathbf{A}$ em um intervalo I dado. Estimando pela Lei do Semicírculo de Wigner, obtemos

$$\frac{E(I)}{N} \approx \int_I \frac{\sqrt{4-x^2}}{2\pi}\mathbf{I}_{|x|\leq 2}\, dx.$$

Outro modo de interpretar isto é notar que $\int_I \frac{\sqrt{4-x^2}}{2\pi}\mathbf{I}_{|x|\leq 2}\, dx$ também é a probabilidade de $\mathbf{A}$ possuir autovalores em I, ou seja,

$$P(\lambda_i \in I) \approx \int_I \frac{\sqrt{4-x^2}}{2\pi}\mathbf{I}_{|x|\leq 2}\, dx.$$

Usamos a notação $P(\lambda_i \in I)$ no mesmo sentido de $P(\Lambda \in I)$, onde $\Lambda : \{\lambda_1, \ldots \lambda_N\} \to \mathbb{R}$ é a variável aleatória que escolhe um autovalor (ela foi definida na seção 1.2).

Exemplo 3.1. Considere que $I = [-1, 1]$, neste caso temos que

$$\int_I \frac{\sqrt{4-x^2}}{2\pi}\mathbf{I}_{|x|\leq 2}\, dx \approx 0.6.$$

Isto nos indica que é esperado que 60% dos autovalores de $\mathbf{A}$ estejam em $[-1, 1]$, daí temos que $E([-1, 1]) \approx 0.6 \cdot N$. Iremos dar preferência à quantidade E em vez da probabilidade, por esta ter uma interpretação mais prática. Vale a pena notar também que as estimativas pela integral possuem erro

[1] Pois todo intervalo $I \subset \mathbb{R}$ possui bordo com medida nula.

bastante pequeno mesmo para matrizes de ordem baixa. A seguir, mostramos alguns experimentos computacionais comprovando este fato.

O parâmetro t representa a quantidade de amostras tomadas, N é a ordem da matriz, I é o intervalo tomado, o valor empírico é a média dos $E(I)$ de cada matriz, e o valor teórico é $\int_I \frac{\sqrt{4-x^2}}{2\pi}\mathbf{I}_{|x|\leq 2}\ dx$.

N	I	t	Valor empírico	Valor teórico
10	[-2, 2]	10000	0.9414	1
10	[-1.5, 1.5]	10000	0.7827	0.8557
10	[-1, 1]	10000	0.5552	0.6089
10	[-0.1, 0.1]	10000	0.0576	0.0636
10	[-0.01, 0.01]	10000	0.0057	0.0064
100	[-2, 2]	1000	0.9951	1
100	[-1.5, 1.5]	1000	0.8462	0.8557
100	[-1, 1]	1000	0.6024	0.6089
100	[-0.1, 0.1]	1000	0.0626	0.0636
100	[-0.01, 0.01]	1000	0.0061	0.0064
4000	[-2, 2]	10	0.99995	1
4000	[-1.5, 1.5]	10	0.8554	0.8557
4000	[-1, 1]	10	0.6089	0.6089
4000	[-0.1, 0.1]	10	0.0636	0.0636
4000	[-0.01, 0.01]	10	0.0062	0.0064

Tabela 3.1: Atráves de experimentos computacionais, obtemos a quantidade esperada de autovalores $E(I)$ no intervalo I, para matrizes de Wigner. Os resultados mostram que a teoria, além de consistente, já dá estimativas muito boas, mesmo para matrizes de ordem baixa.

Se quisermos considerar $\mathbf{A}$ sem estar normalizada, podemos notar que

$$\frac{E(\sqrt{N}\cdot I)}{N} \approx \int_I \frac{\sqrt{4-x^2}}{2\pi}\mathbf{I}_{|x|\leq 2}\ dx,$$

onde $E(\sqrt{N}\cdot I)$ é a quantidade esperada de autovalores de $\mathbf{A}$ em $[\sqrt{N}a, \sqrt{N}b]$, sendo $I=[a,b]$.

Estes resultados nos mostram, claramente, como a Lei do Semicírculo de Wigner é uma poderosa ferramenta para se estimar onde se encontram os autovalores de matrizes hermitianas.

3.2 Estimativas de Erro

Quando procuramos reproduzir certos fenômenos diversas vezes e tomar medidas deles, pequenos erros nas medidas são esperados. É comum termos erros que são gaussianos. Se tivermos um modelo que se baseie matrizes, isso pode conduzir a experimentos onde os resultados variam de acordo com certos erros que são gaussianos. Sendo estes erros refletidos nas entradas das matrizes, um modelo probabilístico pode ser de ajuda para se fazer estimativas. No caso de matrizes simétricas ou hermitianas, a Lei do Semicírculo de Wigner é uma ferramenta ideal.

Suponha uma matriz simétrica $\mathbf{A}$, que será a solução exata de algum problema ou fenômeno. Fazendo-se diversos experimentos, iremos obter matrizes $\mathbf{A}+\frac{1}{\sqrt{N}}\mathbf{E}$, onde as matrizes $\frac{1}{\sqrt{N}}\mathbf{E}$ são as perturbações. Podemos supor que estas perturbações são descritas por matrizes de Wigner simétrica e gaussianas. Além disso, vamos supor que as variâncias são σ^2, com $\sigma > 0$ bem pequeno. Podemos repetir o experimento anterior, mas desta vez o nosso valor teórico é $\int_I \frac{\sqrt{4\sigma^2-x^2}}{2\pi\sigma^2}\mathbf{I}_{|x|\leq 2\sigma}\ dx$. Por exemplo, para $\sigma^2 = 0.01$ temos

os seguintes resultados.

N	I	t	Valor empírico	Valor teórico
10	[-0.2, 0.2]	10000	0.9422	1
10	[-0.15, 0.15]	10000	0.7824	0.8557
10	[-0.1, 0.1]	10000	0.5529	0.6089
10	[-0.01, 0.01]	10000	0.0573	0.0636
10	[-0.001, 0.001]	10000	0.0057	0.0064
100	[-0.2, 0.2]	1000	0.9949	1
100	[-0.15, 0.15]	1000	0.8473	0.8557
100	[-0.1, 0.1]	1000	0.6015	0.6089
100	[-0.01, 0.01]	1000	0.0634	0.0636
100	[-0.001, 0.001]	1000	0.0057	0.0064
4000	[-0.2, 0.2]	10	0.99997	1
4000	[-0.15, 0.15]	10	0.8555	0.8557
4000	[-0.1, 0.1]	10	0.6087	0.6089
4000	[-0.01, 0.01]	10	0.0636	0.0636
4000	[-0.001, 0.001]	10	0.0063	0.0064

Tabela 3.2: Mudando a variância, podemos repetir o experimento anterior, obtendo a quantidade de autovalores esperados $E(I)$ no intervalo I. Podemos notar que os resultados são tão bons quanto os obtidos anteriormente. Com estas estimativas, esperamos poder controlar o erro em perturbações gaussianas para matrizes simétricas.

Iremos fazer uso do teorema enunciado abaixo. Não iremos prová-lo, mas a sua prova pode ser encontrada em [3].

Teorema 3.1 (Weyl). Sejam $\mathbf{A}$ e $\mathbf{E}$ matrizes simétricas de ordem $N \times N$. Sejam $\lambda_1 \leq \ldots \leq \lambda_N$ os autovalores de $\mathbf{A}$ e sejam $\tilde{\lambda}_1 \leq \ldots \leq \tilde{\lambda}_N$ os autovalores de $\mathbf{A}+\mathbf{E}$. Então $|\lambda_i - \tilde{\lambda}_i| \leq \|E\|_2$, para todo $i = 1 \ldots N$.

Note que, no caso de termos $\mathbf{A}+\frac{1}{\sqrt{N}}\mathbf{E}$, a estimativa dada pelo teorema 3.1 é $\left|\lambda_i - \frac{\tilde{\lambda}_i}{\sqrt{N}}\right| \leq \left\|\frac{1}{\sqrt{N}}\mathbf{E}\right\|_2$. Além disso, como $\mathbf{E}$ é simétrica, com autovalores $e_1 \leq \ldots \leq e_N$, então temos que

$$\left\|\frac{1}{\sqrt{N}}\mathbf{E}\right\|_2 = \frac{\max\{|e_i|\}}{\sqrt{N}}.$$

Aproximando a distribuição dos autovalores de $\frac{1}{\sqrt{N}}\mathbf{E}$ através da Lei do Semicírculo de Wigner, temos que seu autovalor máximo é aproximadamente 2σ e o seu mínimo é -2σ. Como já observamos, a tendência para os autovalores se distribuírem de acordo com a Lei do Semicírculo de Wigner é rápida, já valendo para matrizes de ordem baixa. Daí temos que é razoável supor que $\left\|\frac{1}{\sqrt{N}}\mathbf{E}\right\|_2 \approx 2\sigma$. A figura 3.1 mostra alguns experimentos computacionais, que reforçam a conclusão destas observações.

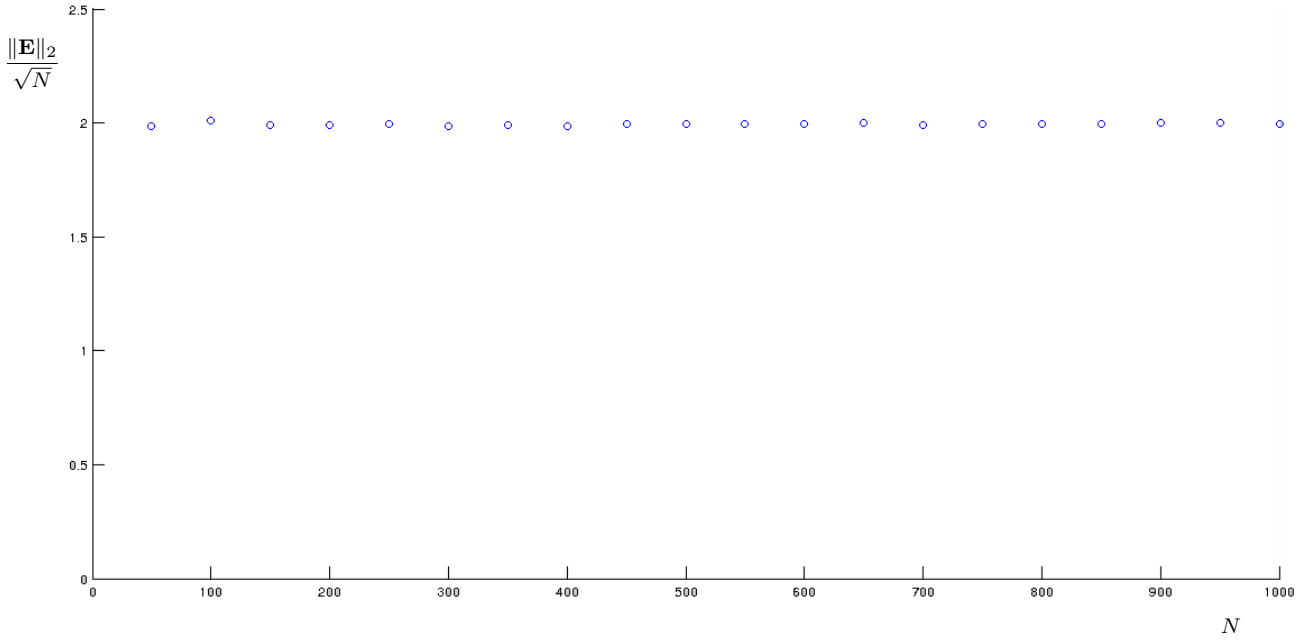

Figura 3.1: Cada círculo representa a média das normas de 20 matrizes de Wigner $N \times N$, tomadas aleatoriamente, com $N = 50, 100, 150, \ldots 1000$.

Do teorema de Weyl temos que $|\lambda_i - \tilde{\lambda}_i| \leq 2\sigma$, para todo $i = 1 \ldots N$. Portanto, temos um controle do erro para perturbações gaussianas em matrizes simétricas. Deve-se observar que esta cota aproximada é probabilística, o que podemos garantir é que as perturbações vão ter norma aproximadamente 2σ com probabilidade 1. Ainda assim, é um resultado forte, que pode ser aplicado em diversas situações. Note também que, a partir destas observações, temos ainda um outro resultado relevante: se tomarmos uma sequência $\frac{1}{\sqrt{N}}\mathbf{E}_N$ de matrizes aleatórias de Wigner, de ordem $N \times N$ e com variância σ^2, então temos que

$$\lim_{N\to\infty} \frac{\|E_N\|_2}{\sqrt{N}} = 2\sigma.$$

Convém observar que estamos cometendo um certo abuso de linguagem ao dizer que uma matriz de Wigner tem variância σ^2, uma vez que estas matrizes foram definidas como tendo variância 1. Já vimos que a mudança de variância nos leva a resultados análogos, e que o tratamento é basicamente o mesmo, por isso não há motivo para preocupação com esta alteração no uso do termo. Note também que não provamos este fato, ele é uma consequência natural de nossas observações e da Lei do Semicírculo de Wigner, a sua prova pode ser encontrada em [1] e [11].

3.3 Determinante

A Lei do Semicírculo de Wigner nos permite estimar o determinante de uma matriz de Wigner quando a ordem cresce indefinidamente. Pelo que vimos até aqui, não é surpresa que a estimativa assintótica seja também uma boa estimativa também para matrizes de ordem baixa. Seja $\frac{1}{\sqrt{N}}\mathbf{X}$ uma matriz de Wigner,

tal que $\lambda_1 \leq \ldots \leq \lambda_N$ são os autovalores de $\mathbf{X}$, então sabemos que

$$\left|\det\left(\frac{1}{\sqrt{N}}\mathbf{X}\right)\right| = \prod_{i=1}^{N}\left|\frac{\lambda_i}{\sqrt{N}}\right|.$$

Mas podemos escrever esta expressão de outra maneira, note que

$$\prod_{i=1}^{N}\left|\frac{\lambda_i}{\sqrt{N}}\right| = \exp\left(\ln\left(\prod_{i=1}^{N}\left|\frac{\lambda_i}{\sqrt{N}}\right|\right)\right) = \exp\left(\sum_{i=1}^{N}\ln\left|\frac{\lambda_i}{\sqrt{N}}\right|\right) = \exp\Big(\langle\mu_N, \ln|x|\rangle\Big).$$

Temos que $\langle\mu_N, \ln|x|\rangle = \int_0^\infty \ln x \; d\mu_N$, mas sabemos que se tomarmos N grande o suficiente, então os autovalores estarão concentrados em $[-2, 2]$. Fixemos N grande o suficiente, então podemos dizer que $\langle\mu_N, \ln x\rangle \approx \int_0^2 \ln x \; d\mu_N$. Temos que

$$\int_0^2 \ln x \; d\mu_N = \frac{1}{N}\sum_{i=1}^{r}\ln(\lambda_i),$$

onde λ_r é o maior autovalor em $[0, 2]$. Pelo que foi visto na seção 3.1, a proporção de autovalores esperados em $[0, 1]$ pode ser aproximada por $\int_0^1 \frac{\sqrt{4-x^2}}{2\pi}\; dx \approx 0.3$, enquanto que em $[1, 2]$ essa proporção é $\int_1^2 \frac{\sqrt{4-x^2}}{2\pi}\; dx \approx 0.2$. Portanto, esperamos ter mais autovalores entre $[0, 1]$ que em $[1, 2]$, isso significa que $\int_0^2 \ln x \; d\mu_N$ será negativo, mais ainda, a tendência é que $\int_0^2 \ln x \; d\mu_N$ apenas diminua com o aumento de N, pois a proporção de autovalores irá convergir, quase certamente, para $0,3$ em $[0, 1]$ e para $0,2$ em $[1, 2]$. Não apenas isso, mas ao olharmos o semicírculo, podemos perceber que a tendência é termos muitos autovalores bem próximos de zero, de modo que $\lim_{N\to\infty}\sum_{i=1}^{N}\ln\left|\frac{\lambda_i}{\sqrt{N}}\right| = -\infty$, portanto

$$\left|\det\left(\frac{1}{\sqrt{N}}\mathbf{X}\right)\right| = \exp\left(\sum_{i=1}^{N}\ln\left|\frac{\lambda_i}{\sqrt{N}}\right|\right) \approx 0.$$

Note que a estimativa apenas melhora com o aumento de N. As figuras 3.2 e 3.3 mostram alguns experimentos computacionais, que reforçam a conclusão destas observações.

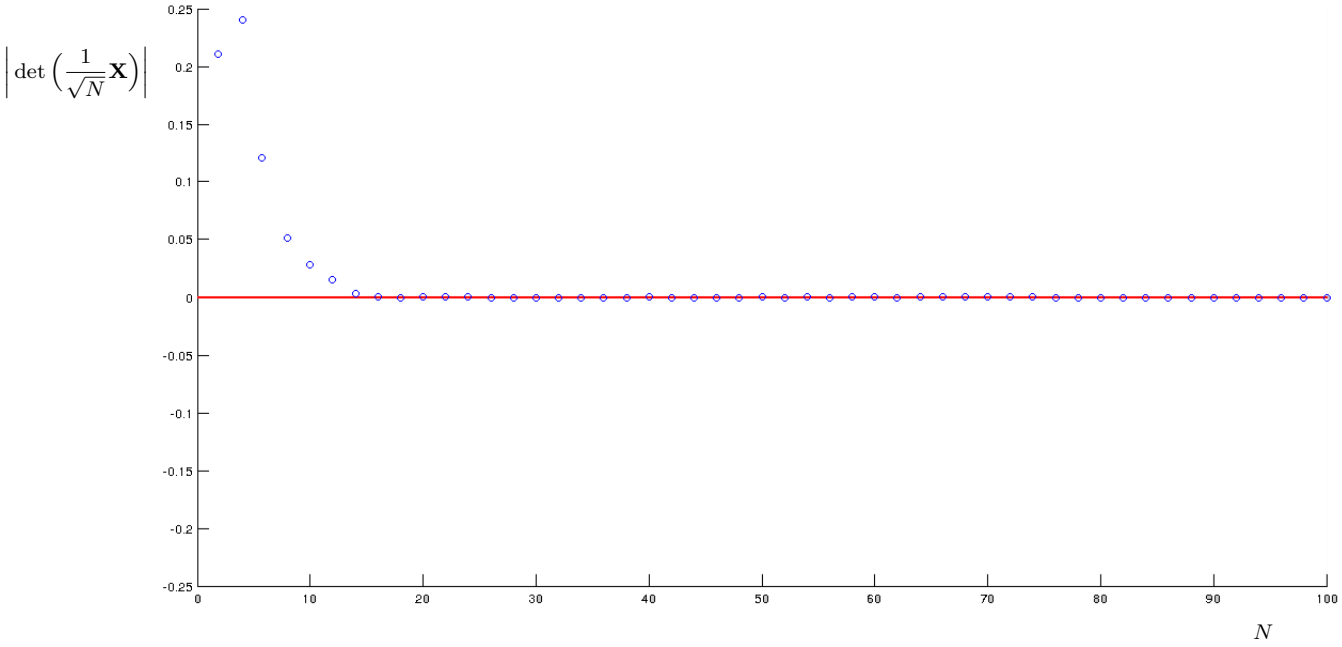

Figura 3.2: Cada círculo representa a média dos determinantes (em módulo) de 100 matrizes de Wigner $N \times N$, tomadas aleatoriamente, com $N = 2, 4, 6, \ldots 100$. Podemos ver que o determinante converge bastante rápido para zero.

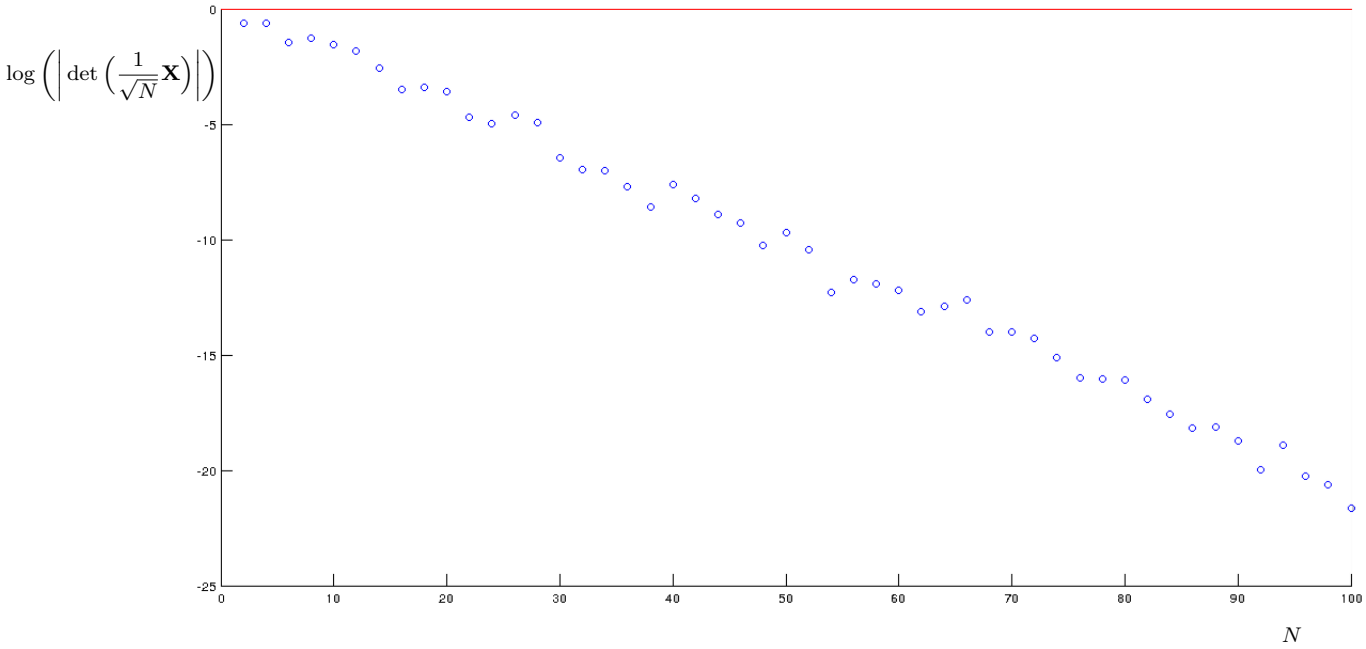

Figura 3.3: Cada círculo representa o logaritmo da média dos determinantes (em módulo) das mesmas matrizes mostradas na figura 3.2. Podemos ver que o decrescimento aparenta ser linear em N.

Referências Bibliográficas

[1] Greg W. Anderson, Alice Guinnet, Ofer Zeitouni. An Introduction to Random Matrices. Cambridge University Press, New York, 2010.

[2] Terence Tao. Topics in Random Matrix Theory. American Mathematical Society, Rhode Island, 2012.

[3] James W. Demmel. Applied Numerical Linear Algebra. Society for Industrial and Applied Mathematics, Philadelphia, 1997.

[4] Gerald B. Folland. Real Analysis: modern techniques and their applications . John Wiley & Sons, Canada, 1999.

[5] Lloyd N. Trefethen, Mark Embree. Spectra and Pseudospectra. Princeton University Press, New Jersey, 2005.

[6] Marcos Nascimento Magalhães. Probabilidade e Variáveis Aleatórias. Universidade de São Paulo, São Paulo, 2011.

[7] Slava Kargin, Elena Yudovina. Random Matrix Theory. University of Cambridge, Cambridge, 2011.

[8] Zhidong Bai, Jack W. Silverstein. Spcetral Analysis of Large Dimensional Random Matrices. Springer, New York, 2010.

[9] Madan Lal Metha. Random Matrices. Academic Press, San Diego, 1991.

[10] Eugene P. Wigner. Characteristic Vectors of Bordered Matrices With Infinite Dimensions. Annals of Mathematics, USA, 1955.

[11] Stuart Geman. A Limit Theorem for the Norm of Random Matrices. Annals of Probability, Brown University, 1980.

[12] Alan Edelman, Eric Kostlan, Michael Shub. How Many Eigenvalues of a Random Matrix are Real?. American Mathematical Society, Rhode Island, 1994.

[13] Patrick Billingsley. Convergence of Probability Measures. John Wiley & Sons, USA, 1999.