Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics

A Combinatorial Proof of Wigner's Semicircle Law

A thesis submitted to the
Delft Institute of Applied Mathematics
in partial fulfillment of the requirements

for the degree

BACHELOR OF SCIENCE
in
APPLIED MATHEMATICS

by

LISANNE VAN WIERINGEN

Delft, The Netherlands
May 2023

BSc thesis APPLIED MATHEMATICS

"A Combinatorial Proof of Wigner's Semicircle Law"

LISANNE VAN WIERINGEN

Delft University of Technology

Supervisor

Dr.ir. W. Groenevelt

Thesis Committee

Dr. C. Kraaikamp

May, 2023 Delft

Abstract

A combinatorial proof of Wigner's Semicircle Law for the Gaussian Unitary Ensemble (GUE) is presented in this report. The distribution of eigenvalues of different samples of general Wigner matrices is shown to converge to the semicircle distribution, with the aid of histograms created in Python. The type of convergence that is shown is that of the averaged moments of the eigenvalue distribution of sample GUE matrices to the moments of the semicircle distribution, as the size of the matrices grow large. This is done by using a method known as the 'method of moments'. The concepts of random matrices, Catalan numbers, mixed moments of standard Gaussian random variables, (non-crossing) pairings, Wick's formula and permutation cycles are introduced in this method. The aim of this report is to provide a detailed proof of Wigner's Semicircle Law in expectation, understandable for bachelor level mathematics students.

Acknowledgement

I would like to thank my supervisor Wolter Groenevelt, for his endless patience with me. Also, I would like to thank Cor Kraaikamp for agreeing to be on my committee multiple times over the past years. Lastly, i would like to express my gratitude for everyone that supported me in finishing my bachelor thesis.

Contents

1 Introduction

1.1 Random matrix theory in general

First, what are random matrices? Random matrices, basically, have random variables corresponding to some probability distribution as elements. A random matrix can be thought of as an object that can be used to model uncertain, hard-to quantify information. Random matrices have applications in a big variety of fields within mathematics. The first prominent instance in which random matrices were used explicitly was in the work by John Wishart in 1928. He used random matrices to estimate sample covariance matrices in order to statistically analyze large samples. Then, in 1950, Eugene Wigner and other physicists looked at random matrices in the context of particle interactions. He used random matrices in order to model energy levels of the nuclei of heavy atoms. His work is considered to be the starting point of random matrix theory. Now, among many other applications, random matrices are of great significance in for example signal processing and information theory by quantifying noise and serve as testers in numerical linear algebra by validating the performance of algorithms.

1.2 Aim

In his work on random matrices, Wigner introduced a theorem which is referred to as Wigner's Semicircle Law (WSL). Roughly, the theorem states that the distribution of eigenvalues of a certain type of $N \times N$ random matrix converges weakly in probability to the semicircle distribution when N grows. There exist many variations on this theorem, stating different classes of random matrices and different types of convergence. Therefore many different types of proofs of Wigner's Semicircle Law have been constructed. Some involve analytic tools, others use combinatorial tools. This report will focus on a combinatorial proof of the theorem, using the 'method of moments'. Wigner's Semicircle Law will be proven for the matrices in the Gaussian Unitary Ensemble, with all its elements distributed according to the Gaussian distribution, showing convergence in average of the probability measures.
The existing proof of Wigner's Semicircle Law of this kind is not suitable to read for the average student equipped with bachelor level knowledge in mathematics, because deductions are made that are not trivial to them. The aim of this report is to provide a detailed combinatorics-based proof of Wigner's Semicircle Law using deductions understandable at bachelor level. Existing proofs were studied in order to do so. [Anderson et al., 2009][Speicher, 2020][Wolf, 2021]

1.3 Outline

First, a foundation for Wigner's Semicircle Law will be set. In this part, Wigner matrices and a distribution for their eigenvalues are introduced. The way the eigenvalues are distributed is then analyzed by plotting them in histograms for a variety of sizes of the Wigner matrix and a variety of different conditions for the elements. It is then shown using the histograms that, as the size of such matrices grow, the distribution of the eigenvalues converges to the semicircle distribution. Then, the type of Wigner matrix for which the convergence will be proven is introduced, followed by how to correctly describe such convergence mathematically. Finally the type of convergence that will be used to prove Wigner's Semicircle Law is given and the theorem itself is stated. After the foundation is set and the theorem is introduced, a brief overview of the course of the proof is given. Then the basis of the proof, which includes various propositions, lemmas and theorems, are presented in a structured way provided with explanatory examples. Finally, with the aid of the basis that was laid, Wigner's Semicircle Law is proven.

2 Wigner's Semicircle Law

2.1 Random matrices, ESD and the semicircle distribution

As mentioned, a random matrix is a matrix whose elements are random variables corresponding to some probability distribution. A random matrix may also be constructed using different probability distributions for different elements of the matrix. There are types of random matrices with certain specifications which are so commonly used, they get their own name. One of those random matrices is called the Wigner random matrix. A Wigner random matrix is a square (complex) matrix which is equal to its conjugate transpose (a Hermitian matrix), whose elements on the diagonal are identically independently distributed (i.i.d) and whose elements above the diagonal are i.i.d as well. See the definition below. Here, the expectation is denoted by $\mathbb{E}$ and the variance is denoted by Var.

Definition 2.1 (Wigner random matrix)**.** Let $W_N = (w_{ij})_{1\leq i,j\leq N}$ be a $N \times N$ matrix where the w_{ij}'s are (complex) random variables corresponding to a certain distribution such that:

[1] For all i, j: $\mathbb{E}(w_{ij}) = 0$

[2] For all $i = j$: w_{ij}'s are independently identically distributed

[3] For all $i < j$: w_{ij}'s are independently identically distributed and $\text{Var}(w_{ij}) = 1$

[4] For all i, j: $w_{ij} = \overline{w_{ji}}$ (W_N is Hermitian)

Then W_N is a Wigner random matrix.

Remark 2.1.1. In Wigner's Semicircle Law, the Wigner matrix is scaled by $\frac{1}{\sqrt{N}}$ and redefined as $\bar{W}_N = \frac{1}{\sqrt{N}} W_N$. To explain why particularly this precise scaling is used, more information is needed. Therefore, the reason will be given later on in the report in Remark 3.5.1.

Remark 2.1.2. The distribution from which the diagonal elements are drawn may be different from the distribution from which the off-diagonal elements are drawn.

Remark 2.1.3. For $i \neq j$, the w_{ij}'s are complex random variables of the form $Z = X + Yi$, where X and Y are distributed according to some, but the same, distribution. By property [4], the random variables w_{ii} must be real, since $w_{ii} = \overline{w_{ii}}$ must hold. From now on, whenever w_{ij} for $i \neq j$ is said to correspond to a distribution, it is meant that its real and imaginary part correspond that distribution.

Remark 2.1.4. Note that in this report, there is interchangeably talked about elements being random variables corresponding to a certain distribution, and elements drawn from a distribution. When spoken about the first, the random variable has not taken on a value yet. Whereas spoken about the latter, some value has been given to it by drawing according to the mentioned distribution.

As eigenvalues of a matrix give valuable information about a system that is being observed, it can be interesting to look at the behaviour of the eigenvalues of the matrix $\bar{W}_N$. A great tool for doing so is the Empirical Spectral Distribution (ESD). The ESD is a statistical measure used in the field of random matrix theory to describe the distribution of eigenvalues of a matrix. Below follows its definition.

Definition 2.2 (Empirical Spectral Distribution (ESD))**.** Let A_N be an $N \times N$ Hermitian matrix and let $\lambda_1(A_N) \leq \lambda_2(A_N) \leq \cdots \leq \lambda_N(A_N)$ be the ordered real eigenvalues of A_N, counting multiplicity. For $i = 1, 2, \ldots, N$ let δ be the Dirac delta function given by:

$$\delta_{\lambda_i(A_N)}(x) = \begin{cases} 1 & ;\ x = \lambda_i(A_N) \\ \\ 0 & ;\ x \neq \lambda_i(A_N). \end{cases}$$

Then the Empiral Spectral Distribution of the eigenvalues of A_N is the measure μ_{A_N} with density:

$$\sigma_{A_N}(x) := \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i(A_N)}(x).$$

.

The ESD of $\bar{W}_N$ can be computed in a graphical form like a histogram as follows. Divide a domain $I \subset \mathbb{R}$ into any desired amount of intervals $I_1, I_2, \ldots, I_j \in I$. The ESD counts the amount of eigenvalues of $\bar{W}_N$ that have a value within a certain interval I_j, using the indicator function $\delta_{\lambda_i(\bar{W}_N)}$. The ESD then normalizes this count by dividing it by the total number of eigenvalues N of the matrix. The result can be portrayed in the form of a histogram. The I_j's are the columns of the histogram and the normalized count, or in other words the probability an eigenvalue lies within the bounds of I_j, as the column height.

Let the scaled Wigner matrix $\bar{W}_N$ be as in Definition 2.1. Then $\bar{W}_N$ is an $N \times N$ Hermitian matrix, so it has N real eigenvalues. Since $\bar{W}_N$ has random variables as its elements, each eigenvalue of W_N is random too. In spite of this randomness, some sort of structure seems to occur when plotting the eigenvalues of a random matrix of the form of $\bar{W}_N$. Take for example $\bar{W}_{50}$, where all elements of the Wigner random matrix are drawn from the standard (complex) Gaussian distribution $\mathcal{N}(0,1)$. Below, a representation of the distribution of eigenvalues of three samples of such random matrices is shown. All figures in this report are made in Python. The code for creating these figures can be found in Appendix A.

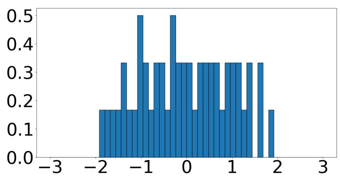


Figure 1: ESD histogram of a sample of Wigner matrix $\bar{W}_{50}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$.

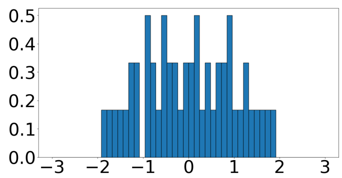


Figure 2: ESD histogram of a sample of Wigner matrix $\bar{W}_{50}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$.

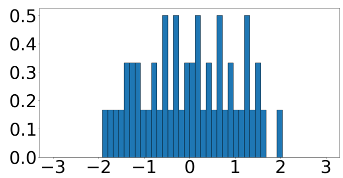


Figure 3: ESD histogram of a sample of Wigner matrix $\bar{W}_{50}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$.

When comparing figures 1, 2 and 3 already some sort of a pattern can be detected. Every time a Wigner matrix $\bar{W}_{50}$ is generated and the eigenvalues are computed, the probability that the eigenvalues lie around -2 and 2 seems to be systematically low. The highest probabilities of an eigenvalue existing within an interval shift between -1.5 and 1.5. An even more consistent pattern becomes visible when the size N of the scaled Wigner matrix is gradually increased. Now, the ESD of three samples of Wigner matrices of size $N = 150$ is compared.

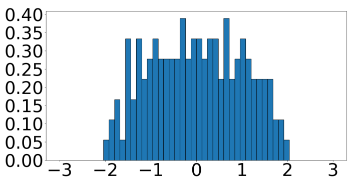


Figure 4: ESD histogram of a sample of Wigner matrix $\bar{W}_{150}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$.

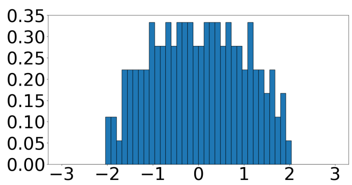


Figure 5: ESD histogram of a sample of Wigner matrix $\bar{W}_{150}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$.

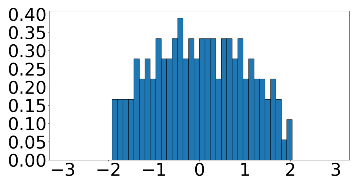


Figure 6: ESD histogram of a sample of Wigner matrix $\bar{W}_{150}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$.

Each time a $\bar{W}_{150}$ is generated, the corresponding eigenvalue distributions look even more similar to each other than the $\bar{W}_{50}$ case. With some outliers, an increase in probability of the eigenvalues existing within a certain interval can be seen from both -2 to 0 and 2 to 0. What happens when the size of the sample matrices are significantly increased to, for example, $N = 1500$?

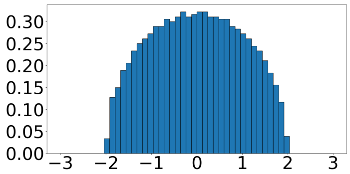


Figure 7: ESD histogram of a sample of Wigner matrix $\bar{W}_{1500}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$.

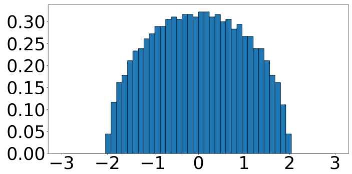


Figure 8: ESD histogram of a sample of Wigner matrix $\bar{W}_{1500}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$.

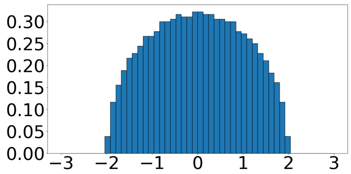


Figure 9: ESD histogram of a sample of Wigner matrix $\bar{W}_{1500}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$.

In figures 7, 8 and 9 barely any difference between the distribution of the eigenvalues is noticeable. A distinct shape occured in all them. Probabilities of eigenvalues appearing around -2 and 2 being the lowest, and going up with a descending increase towards 0. The shape to which the histograms of the ESD seem to converge towards to as N increases, is actually a deterministic distribution called Wigner's semicircle distribution.

Definition 2.3 (The Wigner semicircle distribution)**.** The Wigner semicircle distribution in Wigner's Semicircle Law is the measure μ_{sc}, with density:

$$\sigma_{sc}(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2} & ; \quad -2 \leq x \leq 2 \\ \\ 0 & ; \qquad |x| > 2. \end{cases}$$

Remark 2.3.1. The general Wigner semicircle distribution is actually a measure with density:

$$\sigma_R(x) := \frac{2}{\pi R^2}\sqrt{R^2 - x^2},$$

for $-R \leq x \leq R$ and $\sigma_R(x) = 0$ when $|x| > R$. In Wigner's Semicircle Law, the distribution with $R = 2$ is used.

σ_{sc} is better described as a semi-ellipse, with values between -2 and 2 and a maximum at $\frac{1}{\pi}$.

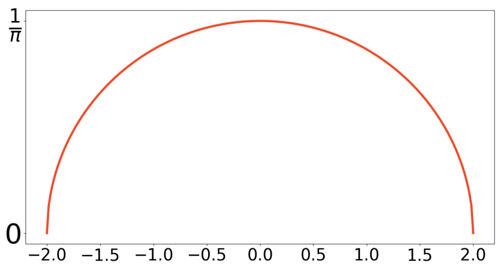


Figure 10: Wigner's semicircle distribution density function σ_{SC}.

As N increases, the eigenvalue distribution seems to get closer to Wigner's semicircle distribution. If again, a scaled Wigner random matrix where all elements of the Wigner random matrix are drawn from the standard (complex) Gaussian distribution $\mathcal{N}(0,1)$ is generated, and a dramatically bigger value for N is chosen, this presumption becomes almost undeniable (it still remains to be proven). See figure 11 below.

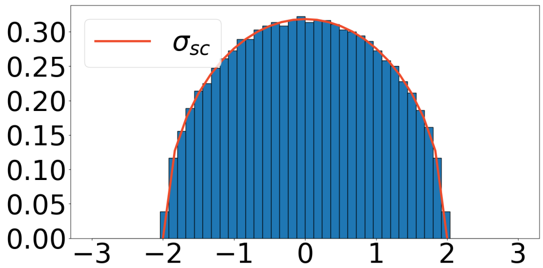


Figure 11: ESD histogram of a sample of Wigner matrix $\bar{W}_{3000}$, where the w_{ii}'s and w_{ij}'s are drawn from $\mathcal{N}(0,1)$, plotted against Wigner's semicircle distribution.

Up till now, only eigenvalue distributions of Wigner matrices have been shown where all the elements have been drawn from the standard (complex) Gaussian distribution. The way the Wigner matrix is defined in Definition 2.1, choosing other distributions for the elements of $\bar{W}_N$ should work too. As long as the mean of all the random variables is equal to 0 and the variance of the off-diagonal random variables is 1. As mentioned before, the Wigner matrix may be constructed using different distributions for the diagonal and off-diagonal elements as well. In the figures below, some results of the eigenvalue distribution where different distributions for the elements of $\bar{W}_{3000}$ are used, are shown.

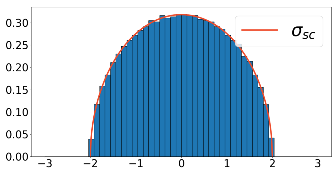


Figure 12: ESD histogram of a sample of Wigner matrix $\bar{W}_{1500}$, where the w_{ii}'s and w_{ij}'s are drawn from the Laplace distribution.

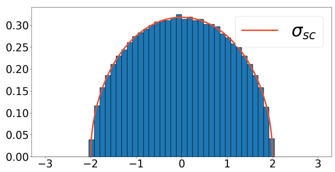


Figure 13: ESD histogram of a sample of Wigner matrix $\bar{W}_{1500}$, where the w_{ii}'s and w_{ij}'s are drawn from the logistic distribution.

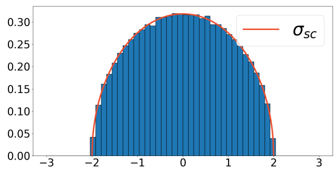


Figure 14: ESD histogram of a sample of Wigner matrix $\bar{W}_{1500}$, where the w_{ii}'s are drawn from the Laplace distribution and w_{ij}'s are drawn from the logistic distribution.

This is a remarkable result, since it turns out this convergence takes place regardless of which distribution is used on the entries of the Wigner matrix. That is very powerful, because even though the matrix is random, with random entries and therefore random eigenvalues, there is immense structure behind them. The eigenvalues may be random, but the group of them arrange themselves in a predictable pattern. The fundamental result that when the size of an $N \times N$ Wigner matrix $\bar{W}_N$ increases, the measure $\mu_{\bar{W}_N}$ of that matrix will approach the semicircle measure μ_{sc} is known as Wigner's Semicircle Law. The law yields an universal structure; it does not depend on the distribution of the entries of the matrix.

In this report, this convergence will be proven for a specific type of random matrix. This type of matrix belongs to the Gaussian Unitary Ensemble (GUE). The matrix is a Wigner random matrix like in Definition 2.1, where the elements on the diagonal are distributed following the standard Gaussian distribution $\mathcal{N}(0,1)$. The non-diagonal elements are distributed according to what will be referred to as the standard complex Gaussian distribution $\mathcal{CN}(0,1)$. Below its definition, followed by the definition of a GUE matrix.

Definition 2.4. Let $X, Y \in \mathcal{N}(0,1)$ be real standard Gaussian random variables. A standard complex Gaussian random variable $Z \in \mathcal{CN}(0,1)$ is of the form:

$$Z = \frac{X + iY}{\sqrt{2}}.$$

Remark 2.4.1. The division of $X + iY$ by $\sqrt{2}$ ensures the variance of Z being one, and therefore ensures that Z is a complex *standard* Gaussian random variable.

Definition 2.5 (Gaussian Unitary Ensemble matrix)**.** Let $W_N = (w_{ij})_{1 \le i,j \le N}$ be a $N \times N$ matrix where the w_{ij}'s are random variables such that:

[1] $w_{ij} = \overline{w_{ji}}$ (W_N is Hermitian)

[2] For all $i \le j$: w_{ij}'s are independently identically distributed

[3] For all $i < j$: $w_{ij} \in \mathcal{CN}(0, 1)$

[4] For all $i = j$: $w_{ij} \in \mathcal{N}(0, 1)$

Then W_N is a random matrix in the GUE.

With the scaled matrix $\bar{W}_N = \frac{1}{\sqrt{N}} W_N$, where W_N as in Definition 2.5, will be worked with from now on.

2.2 Types of convergence

How can the convergence observed in the previous section be described mathematically? As N goes to infinity, it is seen that the normalized number of eigenvalues within some interval $[s, t] \in \mathbb{R}$ approaches the semicircle distribution.

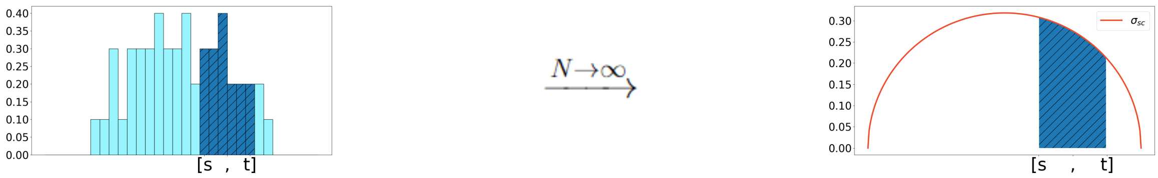


Figure 15: Convergence of the normalized amount of eigenvalues within an interval $[s, t]$ to the semicircle distribution at interval $[s, t]$ as N grows.

The convergence in figure 15 can be described mathematically as follows. For all intervals $[s, t] \in \mathbb{R}$ and for all characteristic functions $f = \mathbb{1}_{[s,t]}$:

$$\frac{1}{N} \sum_{i=1}^{N} f(\lambda_i(\bar{W}_N)) \xrightarrow{N \to \infty} \int_{-\infty}^{\infty} f(x)\, d\mu_{sc}(x). \tag{1}$$

The above convergence is difficult to prove since it is required to know the eigenvalues, which are random. So how can (1) be proven? Instead of looking at characteristic functions $f(x) = \mathbb{1}_{[s,t]}(x)$, the convergence can first be proven for continuous functions $f(x) = x^k$ for all $k = 1, 2, \ldots$:

$$\frac{1}{N} \sum_{i=1}^{N} \lambda_i(\bar{W}_N)^k \xrightarrow{N \to \infty} \int_{-\infty}^{\infty} x^k\, d\mu_{sc}(x). \tag{2}$$

It turns out by more theoretical results that the convergence in (2) is the same as the convergence in (1). In this report, those theoretical results will not be addressed. The limit in (2) is actually a convergence of moments of the Empirical Spectral Distribution to the moments of the semicircle distribution. To understand this, the definition of the k-th order moments of a measure is given below.

Definition 2.6 (k-th order moment)**.** Let X be a random variable with density $\sigma(x)$ corresponding to measure μ. The k-th order central moment of X about its mean $\mathbb{E}(X) = \alpha_X$ with respect to μ is defined as

$$m_{\mu,k}(x) := \mathbb{E}[(X - \alpha_X)^k] = \int_{-\infty}^{\infty} (x - \alpha_X)^k\, d\mu = \int_{-\infty}^{\infty} (x - \alpha_X)^k \sigma(x)\, dx.$$

Since the semicircle distribution has its mean at 0, it is clear that the expression for the k-th order moment of the semicircle distribution is exactly the expression where the limit converges to in (2):

$$m_{\mu_{sc},k}(x) := \int_{-\infty}^{\infty} (x)^k \, d\mu_{sc}.$$

Now, to see that the expression $\frac{1}{N}\sum_{i=1}^{N} \lambda_i(\bar{W}_N)^k$ in (2) is actually the k-th order moment of the Empirical Spectral Distribution, a little more explanation is needed. The mean of the ESD for the matrix described in Definition 2.5 is set at 0 as well, since the results indicated so as N grows. Using Definition 2.6, this is the k-th order moment of the Empirical Spectral measure:

$$m_{\mu_{\bar{W}_N},k}(x) := \int_{-\infty}^{\infty} x^k \, d\mu_{\bar{W}_N} = \int_{-\infty}^{\infty} x^k \sigma_{\bar{W}_N}(x)\, dx.$$

Integrating over all $x \in \mathbb{R}$, the integral only has non-zero values when x coincides with one of the N eigenvalues of $\bar{W}_N$, because of the indicator function $\delta_{\lambda_i(\bar{W}_N)}$ in $\sigma_{\bar{W}_N}(x)$. So the integral becomes the following finite sum, taking into account that the probability of $\lambda_i(\bar{W}_N)$ is equal to an x is $\frac{1}{N}$ by the Empirical Spectral measure:

$$\begin{aligned} m_{\mu_{\bar{W}_N},k}(x) &:= \int_{-\infty}^{\infty} x^k \sigma_{\bar{W}_N}(x)\, dx \\ &= \lambda^k_{1(\bar{W}_N)} \underbrace{\sigma_{sc}(\lambda_{1(\bar{W}_N)})}_{=\frac{1}{N}} + \lambda^k_{2(\bar{W}_N)} \underbrace{\sigma_{sc}(\lambda_{2(\bar{W}_N)})}_{=\frac{1}{N}} + \cdots + \lambda^k_{N(\bar{W}_N)} \underbrace{\sigma_{sc}(\lambda_{N(\bar{W}_N)})}_{=\frac{1}{N}}. \end{aligned} \tag{3}$$

So:

$$m_{\mu_{\bar{W}_N},k}(x) = \frac{1}{N}\sum_{i=1}^{N} \lambda_i(\bar{W}_N)^k. \tag{4}$$

So the convergence in (2) indeed describes the convergence of the moments of the Empirical Spectral Distribution to the moments of the semicircle distribution when N goes to infinity:

$$m_{\mu_{\bar{W}_N},k} \xrightarrow{N\to\infty} m_{\mu_{sc},k}. \tag{5}$$

The proof of this convergence can be split into two problems.

(1) Proof convergence in average: $\mathbb{E}(m_{\mu_{\bar{W}_N},k}) \to m_{\mu_{sc},k}$.

(2) Proof that deviation from the average will become small as $N \to \infty$.

First a weaker form of convergence is shown; convergence in averaged sense. There are infinite possible outcomes when generating a matrix of the form $\bar{W}_N$ like in Definition 2.5. In theory, in step (1) there is averaged over all the possibilities. This means summing over all possible matrices of the form $\bar{W}_N$, taking the sum of the k-th order moments of the ESD's of these matrices and dividing by the amount of possible outcomes, yielding averaged moments of the ESD. Then, in the second step, it will have to be shown that as N increases, the gap between the averaged moments and the 'typical' moments will become small, and the desired convergence in (5) is proven. With 'typical', it is meant that the outlyers, which will always be far from the average, will not be accounted for.

This report will focus on the first problem, giving only a brief overview of how to do the latter at the end. Below follows the theorem, a form of Wigner's Semicircle Law, which will be proven.

Theorem 2.1 (Wigner's Semicircle Law)**.** *Let $\bar{W}_N = \frac{1}{\sqrt{N}} W_N$ be a scaled GUE matrix. Let $m_{\mu_{\bar{W}_N},k}$ be the k-th order moment of the Empirical Spectral Distribution of $\bar{W}_N$ and let $m_{\mu_{sc},k}$ be the k-th order moment of the semicircle distribution. Then, for all $k = 1, 2, \ldots$, the moments of the Empirical Spectral Distribution converge in average to the moments of the semicircle distribution, when N grows:*

$$\lim_{N\to\infty} \mathbb{E}(m_{\mu_{\bar{W}_N},k}) = m_{\mu_{sc},k}.$$

3 Proof of Wigner's Semicircle Law

3.1 Overview

A global overview of the main aspects of the proof will be given before beginning the mathematical proof. First, the moments of the semicircle distribution are calculated. For this, the Catalan numbers are introduced. The odd moments turn out to be 0 and the even moments are shown to coincide with the Catalan numbers. Then, the moments of the Empirical Spectral Distribution of $\bar{W}_N$ are shown to represent the normalized trace of $\bar{W}_N^k$. Therefore, the initial convergence that needs to be proven is now split into two convergences. The limit of the expectation of the normalized trace of $\bar{W}_N^k$ needs to go to 0 for odd k and needs to go to the Catalan numbers for even k, when N is send to infinity. Next, the averaged moments of the ESD are rewritten several times. Starting with the expression that is the expectation of the normalized trace of $\bar{W}_N^k$, which is shown to represent the mixed moments of the elements of $\bar{W}_N$. To be able to work with mixed moments of complex standard Gaussian random variables, moments of real standard Gaussians, the concept of pairings and mixed moments of real standard Gaussians are linked to each other. It will turn out that mixed moments of complex Gaussians can be expressed by summing over pairings and taking the product over the second moments, which are known. This expression for the mixed moments is introduced as Wick's formula. Pairings are then interpreted as permutations, and the final expression for the averaged moments of the ESD of $\bar{W}_N$ is obtained. This expression requires the count of cycles of some combined permutation. When the limit of this expression is taken, the k-th order moments of the ESD turn out to coincide with the number of non-crossing pairings of a set of k elements. The odd moments are therefore concluded to be 0. In the last step of the proof, it is shown that the amount of non-crossing pairings that can be made from a set of an even number of elements, match the Catalan numbers. Therefore, the even moments converge to the even moments of the semicircle distribution and the proof is concluded.

3.2 The moments of the semicircle distribution

The first step of proving Wigner's theorem is determining the moments of the semicircle distribution. In order to do so, the concept of Catalan numbers is needed. The Catalan numbers are a sequence of numbers in $\mathbb{N}$.

Definition 3.1 (Catalan numbers)**.** For $k = 0, 1, 2, \dots$ the k-th Catalan number is defined by the following formula containing a binomial coefficient:

$$C_k := \frac{\binom{2k}{k}}{k+1} = \frac{(2k)!}{(k+1)!k!},$$

where $0! = 1$ is set and $C_0 = 1$. The first Catalan numbers are $1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, \dots$.

Lemma 3.1. *The Catalan numbers satisfy the following recurrence relations:*

[1] $C_k = \frac{4(2k-1)}{2k+2} C_{k-1}$

[2] $C_k = \sum_{l=1}^{k} C_{l-1} C_{k-l}$.

Proof. The first recurrence relation [1] of the Catalan number sequence can be found as follows:

$$\begin{aligned} C_k &= \frac{(2k)!}{(k+1)!k!} = \frac{(2k)(2k-1)(2k-2)!}{(k+1)(k)(k-1)!k!} = \frac{(2k)(2k-1)}{(k+1)(k)} \frac{(2(k-1))!}{k!(k-1)!} \\ &= \frac{(2k)(2k-1)}{(k+1)(k)} C_{k-1} = \frac{2(2k-1)}{k+1} C_{k-1} = \frac{4(2k-1)}{2k+2} C_{k-1}. \end{aligned}$$

The proof of the second recurrence relation [2] can be derived from the proof of the first recurrence relation [Z.,]. □

The Catalan numbers occur in a fair amount of problems in combinatorial mathematics. For example, they count polygon triangulations T_n (the amount of ways to cut an $(n+2)$-polygon into n triangles using $n-1$ non-crossing lines between vertices of a polygon), such that $T_n = C_n$. Catalan numbers are also of great importance in proving Wigner's Law, as they turn out to coincide with the even moments of the semicircle distribution. The remaining uneven moments of the semicircle distribution turn out to be zero.

Proposition 3.2. *For $k = 1, 2, \ldots$, let $m_{\mu_{sc},2k}$ be the even moments of the semicircle distribution and let $m_{\mu_{sc},2k+1}$ be the uneven moments. Then:*

i) $m_{\mu_{sc},2k} = C_k$,

ii) $m_{\mu_{sc},2k+1} = 0$.

Proof. To proof Proposition 3.2 *i)* it suffices to show that the even moments of the semicircle distribution and the Catalan numbers share the same starting value for $k = 1$ and the same recurrence relation. That is $m_{\mu_{sc},2} = C_1$ and by Lemma 3.1[1] the even moments of the semicircle distribution should satisfy $m_{\mu_{sc},2k} = \frac{4(2k-1)}{2k+2} m_{\mu_{sc},2k-2}$. The first condition is easily checked:

$$m_{\mu_{sc},2} = \int_{-\infty}^{\infty} x^2 \sigma_{sc}(x)\, dx = \int_{-\infty}^{\infty} x^2 \frac{1}{2\pi}\sqrt{4-x^2}\,\mathbb{1}_{|x|\leq 2}\, dx = \frac{1}{2\pi}\int_{-2}^{2} x^2\sqrt{4-x^2}\, dx = 1.$$

Now it will be checked whether the even moments of the semicircle distribution share the same recurrence relation with the Catalan numbers. Use Definition 2.6 for the even moments of the semicircle distribution:

$$m_{\mu_{sc},2k} := \int_{-\infty}^{\infty} x^{2k}\sigma_{sc}(x)\, dx = \frac{1}{2\pi}\int_{-2}^{2} x^{2k}\sqrt{4-x^2}\, dx.$$

Substituting x by $x = 2sin(\theta)$ and $dx = 2cos(\theta)d\theta$ gives:

$$\begin{aligned}
m_{\mu_{sc},2k} &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2\pi} 2^{2k}\sin^{2k}(\theta)\sqrt{4-4\sin^2(\theta)}\,2\cos(\theta)\, d\theta \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2\pi} 2^{2k}\sin^{2k}(\theta)\sqrt{4\cos^2(\theta)}\,2\cos(\theta)\, d\theta \\
&= \frac{2\cdot 2^{2k}}{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta) 2\cos(\theta)\cos(\theta)\, d\theta \overset{**}{=} \frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\cos^2(\theta)\, d\theta \\
&= \frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)(1-\sin^2(\theta))\, d\theta \\
&= \frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\, d\theta - \frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\sin^2(\theta)\, d\theta \\
&= \frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\, d\theta - \frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k+1}(\theta)\sin(\theta)\, d\theta \\
&\overset{I.P.}{=} \frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\, d\theta - \frac{2\cdot 2^{2k}}{\pi}\Big[\big[-\sin^{2k+1}(\theta)\cos(\theta)\big]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\
&\qquad\qquad + \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (2k+1)\sin^{2k}(\theta)\cos(\theta)\cos(\theta)\, d\theta\Big] \\
&= \frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\, d\theta - (2k+1)\frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\cos^2(\theta)\dot{d}\theta \\
&= \frac{2\cdot 2^{2k}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\, d\theta - (2k+1)m_{\mu_{sc},2k},
\end{aligned}$$

where integration by parts is used at $(^{I.P.})$ and where the last equality was deduced by $(^{**})$. So the following equality was found:

$$m_{\mu_{sc},2k} = \frac{2 \cdot 2^{2k}}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\, d\theta - (2k+1) m_{\mu_{sc},2k},$$

and rewriting gives:

$$m_{\mu_{sc},2k} \overset{*}{=} \frac{2 \cdot 2^{2k}}{\pi(2k+2)} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\, d\theta.$$

For deducing a recurrence relation of the even moments two different expressions of m_{2k-2} will be used:

(1) $m_{\mu_{sc},2k-2} = \frac{2 \cdot 2^{2k-2}}{\pi(2k-2+2)} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta)\, d\theta = \frac{2^{2k}}{\pi 4k} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta)\, d\theta,$

(2) $m_{\mu_{sc},2k-2} = \frac{2 \cdot 2^{2k-2}}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta) \cos^2(\theta)\, d\theta = \frac{2^{2k}}{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta) \cos^2(\theta)\, d\theta.$

Expression (1) was deduced by (*) and expression (2) was deduced by (**), which can both be found in the previous calculations. Now, a recurrence relation for the even moments of the semicircle distribution can be determined as follows:

$$\begin{aligned}
m_{\mu_{sc},2k} &= \frac{2 \cdot 2^{2k}}{\pi(2k+2)} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k}(\theta)\, d\theta = \frac{2 \cdot 2^{2k}}{\pi(2k+2)} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta) \sin^2(\theta)\, d\theta \\
&= \frac{2 \cdot 2^{2k}}{\pi(2k+2)} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta)(1 - \cos^2(\theta))\, d\theta \\
&= \frac{2 \cdot 2^{2k}}{\pi(2k+2)} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta)\, d\theta - \frac{2 \cdot 2^{2k}}{\pi(2k+2)} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta) \cos^2(\theta)\, d\theta \\
&= \frac{2}{\frac{1}{2} + \frac{1}{2k}} \frac{2^{2k}}{4k\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta)\, d\theta - \frac{2}{k+1} \frac{2^{2k}}{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^{2k-2}(\theta) \cos^2(\theta)\, d\theta \\
&\overset{(1),(2)}{=} \frac{2}{\frac{1}{2} + \frac{1}{2k}} m_{\mu_{sc},2k-2} - \frac{2}{k+1} m_{\mu_{sc},2k-2} = \frac{4(2k-1)}{2k+2} m_{\mu_{sc},2k-2}.
\end{aligned}$$

So $m_{\mu_{sc},2k} = \frac{4(2k-1)}{2k+2} m_{\mu_{sc},2k-2}$ and by Lemma 3.11 the recurrence relation of the even moments of the semicircle distribution is the same as the recurrence relation of the Catalan numbers. Now there may be concluded that $m_{\mu_{sc},2k} = C_k$; the even moments of the semicircle distribution coincide with the Catalan numbers. To proof 3.2 *ii*), use Definition 2.6 for the uneven moments of the semicircle distribution:

$$m_{\mu_{sc},2k+1} := \int_{-\infty}^{\infty} x^{2k+1} \sigma_{sc}(x)\, dx = \frac{1}{2\pi} \int_{-2}^{2} x^{2k+1} \sqrt{4 - x^2}\, dx.$$

One can easily check that $h(x) = x^{2k+1}\sqrt{4 - x^2}$ is an odd function. So by asymmetry, $m_{\mu_{sc},2k+1} = 0$. This concludes the proof. □

3.3 The moments of the Empirical Spectral Distribution

Now, the moments of the Empirical Spectral measure $\mu_{\bar{W}_N}$ will be rewritten. The moments $m_{\mu_{\bar{W}_N},k}$ were shown in Equation 3 to be expressed in terms of the eigenvalues $\lambda_i(\bar{W}_N)$ of the GUE Wigner matrix. However, since these are random, they are not useful to work with further. It is therefore more practical to express the k-th moment of the empirical density measure in terms of $\bar{W}_N$, of which more properties are known. To do this, some linear algebra is needed.

Definition 3.2 (Hermitian (self-adjoint) matrix)**.** A matrix A is Hermitian (or self-adjoint) if:

[1] A is a complex square matrix,

[2] $A = A^*$, where A^* is the conjugate transpose of A.

The Wigner random matrix $\bar{W}_N$ is a Hermitian (or self-adjoint) matrix, since it is a $N \times N$ complex square matrix with the property $w_{ij} = \overline{w_{ji}}$ ([4] in Definition 2.5). Hermitian matrices are known to be diagonalizable. So $\bar{W}_N$ can be unitarily diagonalized such that $\bar{W}_N = UDU^{-1}$, where

(1) U is a unitary matrix; a complex square matrix such that its conjugate transpose U^* is also its inverse; $U^* = U^{-1}$,

(2) D is a diagonal matrix with the eigenvalues of $\bar{W}_N$ in the diagonal;

$$D = \begin{pmatrix} \lambda_{1(\bar{W}_N)} & & & & \\ & \lambda_{2(\bar{W}_N)} & & \text{\Large 0} & \\ & & \ddots & & \\ & \text{\Large 0} & & \ddots & \\ & & & & \lambda_{N(\bar{W}_N)} \end{pmatrix}.$$

Doing so for $\bar{W}_N^k$ yields the following expression:

$$\bar{W_N^k} = (UDU^*)^k = \underbrace{UD\overbrace{U^*U}^{=I}DU^* \cdots UDU^*}_{k \text{ times}} = UD^kU^*$$

and therefore the trace of $\bar{W}_N^k$ becomes the expression that is needed:

$$Tr(\bar{W_N^k}) = Tr(UD^kU^*) = Tr(U^*UD^k) = Tr(D^k) = \sum_{i=1}^{N} \lambda_i(\bar{W}_N)^k,$$

where $Tr(ABC) = Tr(CAB)$ is used for square matrices A, B, C of the same size.[Fraleigh et al., 1990] Consequently, for the k-th order moment of the Empirical Spectral Distribution becomes:

$$m_{\mu_{\bar{W}_N},k} = \frac{1}{N}\sum_{i=1}^{N} \lambda_i(\bar{W}_N)^k = \frac{1}{N}Tr(\bar{W}_N^k). \tag{6}$$

Taking into account what was found for the moments of the semicircle distribution and what was found after redefining the moments of the ESD of $\bar{W}_N$, the limit in Theorem 2.1 that needs to be proven becomes:

$$\lim_{N\to\infty} \mathbb{E}\big(\frac{1}{N}Tr(\bar{W}_N^{2k})\big) = C_k, \tag{7}$$

$$\lim_{N\to\infty} \mathbb{E}\big(\frac{1}{N}Tr(\bar{W}_N^{2k+1})\big) = 0. \tag{8}$$

3.4 Wick's Formula

Up until now, it was found that in order to proof Wigner's Semicircle Law in Theorem 2.1, the limits in (7) and (8) need to be proven. To do so, the term $\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k))$ for all k needs to be examined further first. Since $\bar{W}_N^k = \frac{1}{\sqrt{N}^k}W_N^k$ we can simplify the term in the following way:

$$\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k)) = \frac{1}{N}\frac{1}{\sqrt{N}^k}\mathbb{E}(Tr(W_N^k)).$$

The trace of a $N \times N$ matrix A_N to some power has the following identity.

Lemma 3.3. *Let $A_N = (a_{ij})_{i,j=1}^N$ be a square matrix. Then,*

$$Tr(A_N^k) = \sum_{i_1,i_2,\ldots,i_k=1}^{N} a_{i_1i_2} \cdot a_{i_2i_3} \cdot \ldots \cdot a_{i_ki_1}.$$

Lemma 3.3 says that the trace of a matrix to the power k is the sum of a particular product over all possible N^k combinations of i_j's with $1 \leq j \leq k$ and $1 \leq i_1, i_2, \cdots, i_k \leq N$. This can be made more clear in an example.

Example 3.1 (The trace of a matrix A_3^2)**.** Let A_3 be the following 3×3 matrix:

$$A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

There are $3^2 = 9$ possible combinations of $1 \leq i_1, i_2 \leq 3$, namely

$$(i_1, i_2) = (1,1), (1,2), (2,1), (1,3), (3,1), (2,2)(2,3), (3,2), (3,3),$$

so the trace of A_3^2 becomes:

$$\begin{aligned} Tr(A_3^2) = \sum_{i_1,i_2=1}^{3} a_{i_1i_2}a_{i_2i_1} =& (a_{i_1i_2}a_{i_2i_1})_{i_1=1,i_2=1} + (a_{i_1i_2}a_{i_2i_1})_{i_1=1,i_2=2} + (a_{i_1i_2}a_{i_2i_1})_{i_1=2,i_2=1} + \\ & (a_{i_1i_2}a_{i_2i_1})_{i_1=1,i_2=3} + (a_{i_1i_2}a_{i_2i_1})_{i_1=3,i_2=1} + (a_{i_1i_2}a_{i_2i_1})_{i_1=2,i_2=2} + \\ & (a_{i_1i_2}a_{i_2i_1})_{i_1=2,i_2=3} + (a_{i_1i_2}a_{i_2i_1})_{i_1=3,i_2=2} + (a_{i_1i_2}a_{i_2i_1})_{i_1=3,i_2=3}. \end{aligned}$$

That is:

$$\begin{aligned} Tr(A_3^2) = \sum_{i_1,i_2=1}^{3} a_{i_1i_2}a_{i_2i_1} =& a_{11}a_{11} + a_{12}a_{21} + a_{21}a_{12} + \\ & a_{13}a_{31} + a_{31}a_{13} + a_{22}a_{22} + \\ & a_{23}a_{32} + a_{32}a_{23} + a_{33}a_{33}. \end{aligned}$$

Note that this is exactly the expression one would obtain for the trace of A^2, when calculating A^2 by hand and summing the diagonal elements.

So the expectation of the trace of the $N \times N$ Wigner matrix $W_N = (w_{ij})_{i,j=1}^N$ can be rewritten by using the identity in Lemma 3.3 and taking the expectation within the sum:

$$\mathbb{E}(\frac{1}{N} Tr(\bar{W}_N^k)) = \frac{1}{N}\frac{1}{\sqrt{N}^k} \sum_{i_1,i_2,\ldots,i_k=1}^{N} \mathbb{E}(w_{i_1i_2} \cdot w_{i_2i_3} \cdot \ldots w_{i_ki_1}). \tag{9}$$

Now, what does the expression $\mathbb{E}(w_{i_1i_2} \cdot w_{i_2i_3} \cdot \ldots w_{i_ki_1})$ within the sum mean? For the sake of understanding this term, an investigation of what that expression might look like when taking the trace of $\bar{W}_N$ to the power of $k = 8$ will be done in the example below.

Example 3.2 (A term of $\mathbb{E}(\frac{1}{N} Tr(\bar{W}_N^8))$)**.** Let the 8-th moment of the ESD of $\bar{W}_N$ with some $N \geq 9$ be:

$$\mathbb{E}(\frac{1}{N} Tr(\bar{W}_N^8)) = \frac{1}{N}\frac{1}{\sqrt{N}^8} \sum_{i_1,i_2,\ldots,i_8=1}^{N} \mathbb{E}(w_{i_1i_2} \cdot w_{i_2i_3} \cdot w_{i_3i_4} \cdot w_{i_4i_5} \cdot w_{i_5i_6} \cdot w_{i_6i_7} \cdot w_{i_7i_8} \cdot w_{i_8i_1}).$$

Take for example the following combination of i_j's:

$$(i_1, i_2, i_3, i_4, i_5, i_6, i_7, i_8) = (1, 1, 1, 7, 9, 7, 9, 1).$$

Then the term corresponding to that combination of i_j's in the above sum is:

$$\mathbb{E}(w_{11} \cdot w_{11} \cdot w_{17} \cdot w_{79} \cdot w_{97} \cdot w_{79} \cdot w_{91} \cdot w_{11}) = \mathbb{E}(w_{11}^3 \cdot w_{17}^1 \cdot w_{79}^2 \cdot w_{97}^1 \cdot w_{91}^1).$$

Remembering that each w_{ij} is a (complex) standard Gaussian random variable, it is seen by example 3.2 that the sum in (9) becomes a sum of mixed moments. In the example, the term shows the third moment of the random variable w_{11}, the first moment of the random variables w_{17}, w_{97} and w_{91} and the second moment of the random variable w_{79}. Therefore, the moments of (complex) standard Gaussian random variables are needed. First, the moments of only one real Gaussian random variable will be looked at. After that, how to calculate mixed moments of multiple real Gaussian random variables will be explained and finally mixed moments of multiple complex Gaussian random variables will become clear.

Definition 3.3. A real standard Gaussian random variable X, $X \sim \mathcal{N}(0,1)$ has the following probability density function:

$$\sigma_G(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}}.$$

It turns out that the k-th order moments of a real standard Gaussian random variable are positive integers that count all possible pairings of a set of k natural numbers. This will be shown by first calculating the moments of the real standard Gaussian random variables, then counting all pairings of a set of k elements and comparing the results.

The k-th order moment of a real standard Gaussian random variable, $m_{G,k}$, is:

$$m_{G,k} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^k e^{\frac{-t^2}{2}}\, dt.$$

Proposition 3.4. *The k-th order moments of a real standard Gaussian random variable $X \sim \mathcal{N}(0,1)$ are:*

$$m_{G,k} = \begin{cases} (k-1)!! & ; \ k \ \text{even} \\ \\ 0 & ; \ k \ \text{odd}. \end{cases}$$

Here, the double factorial for even k is the product $(k-1)!! = (k-1)(k-3)\cdots(1)$.

Proof. First, the odd moments will be checked. The odd moments of a real standard Gaussian random variable are:

$$m_{G,2k+1} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^{2k+1} e^{\frac{-t^2}{2}}\, dt.$$

One can easily check that $h(x) = t^{2k+1} e^{\frac{-t^2}{2}}$ is an odd function. In the same way that was shown in the proof of the odd semicircle distribution moments, it can be concluded that all odd moments of real standard Gaussian random variables are 0 due to symmetry.
Now, for the even moments. The even moments of a real standard Gaussian random variable are:

$$m_{G,2k} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^{2k} e^{\frac{-t^2}{2}}\, dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^{2k-1} t e^{\frac{-t^2}{2}}\, dt.$$

Using integration by parts:

$$\int u\, dv = \big[uv\big] - \int v\, du,$$

with $u = t^{2k-1}$, $du = (2k-1)t^{2k-2}dt$, $dv = te^{\frac{-t^2}{2}}dt$ and $v = -e^{\frac{-t^2}{2}}$ the desired result is acquired:

$$\begin{aligned} m_{G,2k} &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^{2k-1} t e^{\frac{-t^2}{2}}\, dt \overset{I.P.}{=} \frac{1}{\sqrt{2\pi}} \Big(\underbrace{\Big[-t^{2k-1} e^{\frac{-t^2}{2}} \Big]_{-\infty}^{\infty}}_{=0 \text{ by symmetry}} - \int_{\infty}^{\infty} (2k-1)t^{2k-2} \cdot -e^{\frac{-t^2}{2}}\, dt \Big) \\ &= (2k-1) \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} t^{2k-2} e^{\frac{-t^2}{2}}\, dt \\ &= (2k-1) m_{G,2k-2} = (2k-1)(2k-3) m_{G,2k-4} \\ &= (2k-1)(2k-3)(2k-5)\cdots(2k-(2k-3)) m_{G,2} \\ &= (2k-1)(2k-3)(2k-5)\cdots(3)(1) \\ &= (2k-1)!!. \end{aligned}$$

Where the second moment of the standard Gaussian distribution is equal to its variance, and therefore $m_{G,2} = 1$. This concludes the proof. □

From here, the combinatorial aspects of the proof of Wigner's Semicircle Law in Theorem 2.1 will be introduced. As mentioned, the moments of the Gaussian distribution turn out to count the number of ways to create a pairing of the elements of a set containing k elements. The definition of a pairing and a small example can be found below.

Definition 3.4 (Pairing)**.** Denote for $n \in \mathbb{N}$ the set $[n]{:=}\,\{1, 2, \ldots, n\}$. A pairing Π of $[n]$ is a collection of disjoint subsets (or pairs) of $[n]$, $\Pi = V_1, V_2, \ldots, V_k$, such that for all $i = 1, 2, \ldots, k$:

[1] $V_i \subseteq [n]$

[2] $\#V_i = 2$

[3] For $i \neq j$: $V_i \bigcap V_j = \emptyset$

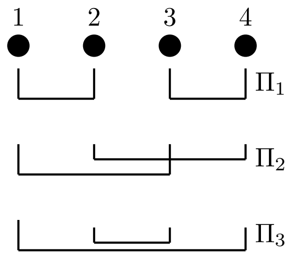
[4] $\bigcup_{i=1}^{k} V_i = [n]$

The set of all pairings of $[n]$ is denoted by:

$$\mathcal{P}_2(n) := \{\Pi : \Pi \text{ is a pairing of } [n]\}.$$

Remark 3.4.1. The pairs V_i of a pairing are unordered.

Example 3.3. Let $n = 4$. Then $[4] = \{1, 2, 3, 4\}$ and all possible pairings of $[4]$ are: $\Pi_1 = \{\{1, 2\}, \{3, 4\}\}$, $\Pi_2 = \{\{1, 3\}, \{2, 4\}\}$ and $\Pi_3 = \{\{1, 4\}, \{2, 3\}\}$. See below for a schematic representation of this.



The set of all pairings of $[4]$ is therefore:

$$\mathcal{P}_2(4) = \{\Pi_1, \Pi_2, \Pi_3\} = \{\{\{1, 2\}, \{3, 4\}\}, \{\{1, 3\}, \{2, 4\}\}, \{\{1, 4\}, \{2, 3\}\}\}.$$

Here, $\#\mathcal{P}_2(4) = 3$.

In the following proposition is stated what was said about the moments of a Gaussian random variable before.

Proposition 3.5. *For $k = 1, 2, \ldots$ let $\mathcal{P}_2(k)$ be the set of all pairings of $[k]$ and let $m_{G,k}$ be the k-th order moment of the real standard Gaussian distribution. Then:*

i) $m_{G,2k} = \#\mathcal{P}_2(2k)$,

ii) $m_{G,2k+1} = \#\mathcal{P}_2(2k + 1) = 0$.

Proof. To proof Proposition 3.5, by Proposition 3.4 it suffices to show that $\#\mathcal{P}_2(2k) = (2k - 1)!!$ and $\#\mathcal{P}_2(2k + 1) = 0$. For all k, count the elements of $\#\mathcal{P}_2(k)$ recursively in the following way. There are $k - 1$ possible pairs for the pair with 1 as its first element. After choosing the second element of the pair $\{1, \cdot\}$, there are $(k - 2)$ elements left to make create the rest of the pairing. The number of different pairings of $(k - 2)$ elements is $\#\mathcal{P}_2(k - 2)$. This generates the following equality:

$$\#\mathcal{P}_2(k) = (k - 1)\#\mathcal{P}_2(k - 2).$$

Now for the remaining $(k - 2)$ elements, again pick a different element that wasn't picked for the first or second element of the first pairing, and make it the first element of the second pairing. There can be $(k - 3)$ different pairs created after the first element of that pair is picked. Again, from the remaining

$(k-4)$ elements should be counted how many pairings can be made, that is $\#\mathcal{P}_2(k-4)$, Iterating this gives:

$$\begin{aligned}
\#\mathcal{P}_2(k) &= (k-1)\#\mathcal{P}_2(k-2) = (k-1)(k-3)\#\mathcal{P}_2(k-4) = (k-1)(k-3)(k-5)\#\mathcal{P}_2(k-6) \\
&= \begin{cases} (k-1)(k-3)\dots\#\mathcal{P}_2(2) & ;\ k \text{ even} \\ \\ (k-1)(k-3)\dots\#\mathcal{P}_2(1) & ;\ k \text{ odd} \end{cases} \\
&= \begin{cases} (k-1)(k-3)\dots(1) & ;\ k \text{ even} \\ \\ (k-1)(k-3)\dots(0) & ;\ k \text{ odd} \end{cases} \\
&= \begin{cases} (k-1)!! & ;\ k \text{ even} \\ \\ 0 & ;\ k \text{ odd.} \end{cases}
\end{aligned}$$

So indeed $m_{G,2k} = \#\mathcal{P}_2(2k)$ and $m_{G,2k+1} = \#\mathcal{P}_2(2k+1)$ by Proposition 3.4. □

Remember the objective is to prove the two limits in (7) and (8), which together represent the convergence of the averaged k-th order moment of the ESD of $\bar{W}_N$ to the k-th order moment of the semicircle distribution:

$$\lim_{N\to\infty} m_{\mu_{\bar{W}_N},2k} = \lim_{N\to\infty} \mathbb{E}\big(\frac{1}{N}Tr(\bar{W}_N^{2k})\big) = C_k = m_{\mu_{sc},2k},$$

$$\lim_{N\to\infty} m_{\mu_{\bar{W}_N},2k+1} = \lim_{N\to\infty} \mathbb{E}\big(\frac{1}{N}Tr(\bar{W}_N^{2k+1})\big) = 0 = m_{\mu_{sc},2k+1}.$$

Recall equation (9):

$$\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k)) = \frac{1}{N}\frac{1}{\sqrt{N}^k}\sum_{i_1,i_2,\dots,i_k=1}^{N} \mathbb{E}(w_{i_1i_2}\cdot w_{i_2i_3}\cdot \dots \cdot w_{i_ki_1}),$$

where the w_{ij}'s are the elements of the unscaled GUE matrix W_N. Example 3.2 gave an insight that the term $\mathbb{E}(w_{i_1i_2}\cdot w_{i_2i_3}\cdot \dots \cdot w_{i_ki_1})$ within the sum is actually a representation of mixed moments of independent standard complex Gaussian random variables. For the averaged k-th moment of the ESD of $\bar{W}_N$ any amount up to k of independent complex Gaussian random variables with orders up to k might appear in that term of the sum. In order to be able to work with this, first mixed moments of real standard Gaussian random variables will be analyzed, using Proposition 3.5.

3.4.1 Mixed moments of real Gaussians

Let $X_1, X_2, \dots, X_p$ be independent real standard Gaussian random variables with some finite order moment, respectively $k_1, k_2, \dots, k_p$. Then, by Proposition 3.4:

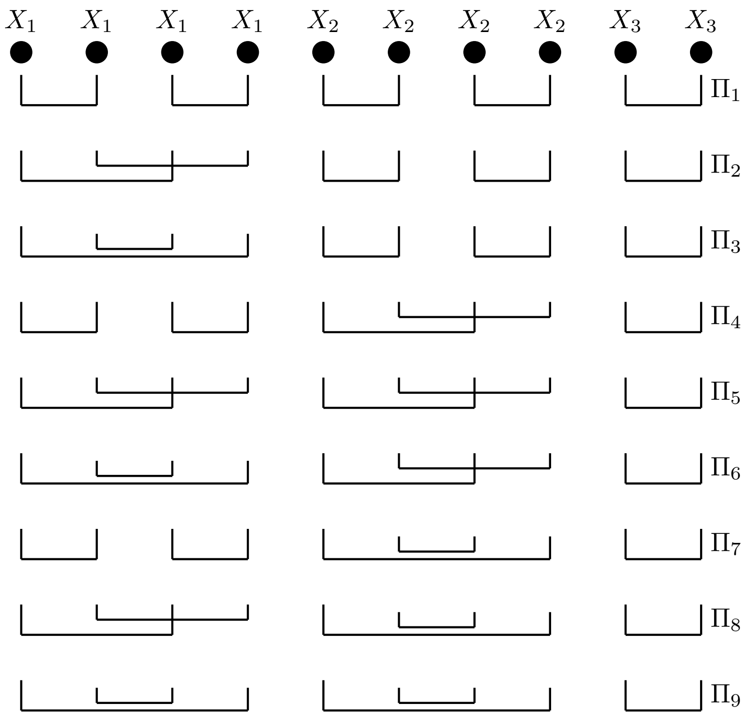
$$\mathbb{E}(X_1^{k_1}X_2^{k_2}\dots X_p^{k_p}) = \mathbb{E}(X_1^{k_1})\mathbb{E}(X_2^{k_2})\dots\mathbb{E}(X_p^{k_p}) = \#\mathcal{P}_2(k_1)\#\mathcal{P}_2(k_2)\dots\#\mathcal{P}_2(k_p).$$

This indicates that the mixed moments can be rewritten as counting the pairings that only contain pairs obtained from X_1's in $\#\mathcal{P}_2(k_1)$ different ways, obtained from X_2's in $\#\mathcal{P}_2(k_2)$ different ways, and so on. Pairings with pairs that contain for example one X_3 element and one X_7 element are not counted. This concept is explained in the example below.

Example 3.4. Let X_1, X_2 and X_3 be real independent standard Gaussian random variables. Let $x_1, x_2, \dots, x_{10}$ represent elements that are distributed in the following way: $x_1, x_2, x_3, x_4 = X_1$, $x_5, x_6, x_7, x_8 = X_2$ and $x_9, x_{10} = X_3$. Then:

$$\mathbb{E}(x_1\cdot x_2\cdot x_3\cdot x_4\cdot x_5\cdot x_6\cdot x_7\cdot x_8\cdot x_9\cdot x_{10}) = \mathbb{E}(X_1X_1X_1X_1X_2X_2X_2X_2X_3X_3),$$

and the pairings of $\{X_1, X_1, X_1, X_1, X_2, X_2, X_2, X_2, X_3, X_3\}$ that only pair the X_1's, the X_2's and the X_3's among themselves should be counted in order to calculate $\mathbb{E}(x_1 \cdot x_2 \cdot x_3 \cdot x_4 \cdot x_5 \cdot x_6 \cdot x_7 \cdot x_8 \cdot x_9 \cdot x_{10})$. See below for a schematic representation of those pairings.



So:

$$\mathbb{E}(x_1 \cdot x_2 \cdot x_3 \cdot x_4 \cdot x_5 \cdot x_6 \cdot x_7 \cdot x_8 \cdot x_9 \cdot x_{10}) = \mathbb{E}(X_1^4 X_2^4 X_3^2) = \#\mathcal{P}_2(4)\#\mathcal{P}_2(4)\#\mathcal{P}_2(2) = 3 \cdot 3 \cdot 1 = 9.$$

The next problem is how, for each term of the sum in equation (9), can be made sure pairings are counted that only pair X_j's among themselves? This is solved in the following way. First, all possible pairings of the elements within the expected value are made. Then, for each pairing is checked whether they contain pairs that are allowed. To verify whether a pair is allowed, the following analytical tool is used. For two independent real standard Gaussian random variables x_i and x_j the first and second moments are known. This gives the following information of a pair $\{x_i, x_j\}$:

$$\mathbb{E}(x_i x_j) = \begin{cases} \mathbb{E}(x_i^2) = 1 & ; \ x_i = x_j \\ \\ \mathbb{E}(x_i x_j) = \mathbb{E}(x_i)\mathbb{E}(x_j) = 0 & ; \ x_i \neq x_j. \end{cases} \tag{10}$$

So by summing over all possible pairings, and for every pairing multiplying the results of the expected value of the two elements of all pairs within that pairing, one can count the pairings that only connect the X_j's among themselves. This way, if there is one pair containing elements from two different X_j's, the product for that specific pairing will yield 0 and the pairing will not be included in the count. This formula for counting mixed moments for real independent standard Gaussian random variables is called Wick's formula.

Theorem 3.6 (Wick's formula for real Gaussians)**.** *Let $X_1, X_2, \ldots, X_p$ be real independent standard Gaussian random variables and let $x_1, x_2, \ldots, x_n \in \{X_1, X_2, \ldots, X_p\}$. Then Wick's formula states that mixed moments of n elements can be calculated using second moments:*

$$\mathbb{E}(x_1 \cdot x_2 \cdot \ldots \cdot x_n) = \sum_{\Pi \in \mathcal{P}_2(n)} \prod_{(i,j) \in \Pi} \mathbb{E}(x_i x_j),$$

where for $1 \leq s, t \leq p$

$$\mathbb{E}(x_i x_j) = \begin{cases} 1 & ;\ x_i, x_j \in X_s \\ \\ 0 & ;\ x_i \in X_s,\ x_j \in X_t \text{ for } s \neq t. \end{cases}$$

3.4.2 Mixed moments of complex Gaussians

Up until now, a way to calculate mixed moments of real Gaussian random variables has been obtained. The goal was to do this for the elements of the GUE matrix W_N in equation (9), which are complex Gaussian random variables. Note that the x_j's in Wick's formula can be replaced by complex z_j's because of its multi-linear structure. In Wick's formula in Theorem 3.6, when a pair $\{z_i, z_j\}$ gives a contribution needs to be determined. To do so, let Z be a standard complex Gaussian random variables as in Definition 2.4:

$$Z = \frac{X + iY}{\sqrt{2}},$$

and

$$\bar{Z} = \frac{X - iY}{\sqrt{2}}.$$

Then the first moments of Z and $\bar{Z}$ are:

$$\mathbb{E}(Z) = \frac{\mathbb{E}(X + iY)}{\sqrt{2}} = \frac{\mathbb{E}(X) + i\mathbb{E}(Y)}{\sqrt{2}} = \frac{0 + i0}{\sqrt{2}} = 0,$$
$$\mathbb{E}(\bar{Z}) = \frac{\mathbb{E}(X - iY)}{\sqrt{2}} = \frac{\mathbb{E}(X) - i\mathbb{E}(Y)}{\sqrt{2}} = \frac{0 - i0}{\sqrt{2}} = 0.$$

The second moments of Z and $\bar{Z}$ are:

$$\mathbb{E}(Z^2) = \mathbb{E}(\frac{X^2 + iXY + iYX + i^2Y^2}{2}) = \frac{1}{2}\big(\underbrace{\mathbb{E}(X^2)}_{=1} - \underbrace{\mathbb{E}(Y^2)}_{=1} + i(\underbrace{\mathbb{E}(XY)}_{=0} + \underbrace{\mathbb{E}(YX)}_{=0})\big) = 0,$$
$$\mathbb{E}(\bar{Z}^2) = \mathbb{E}(\frac{X^2 - iXY - iYX + i^2Y^2}{2}) = \frac{1}{2}\big(\underbrace{\mathbb{E}(X^2)}_{=1} - \underbrace{\mathbb{E}(Y^2)}_{=1} - i(\underbrace{\mathbb{E}(XY)}_{=0} + \underbrace{\mathbb{E}(YX)}_{=0})\big) = 0,$$

$$\mathbb{E}(Z\bar{Z}) = \mathbb{E}(\frac{X^2 - iXY + iYX - i^2Y^2}{2}) = \frac{1}{2}\big(\underbrace{\mathbb{E}(X^2)}_{=1} + \underbrace{\mathbb{E}(Y^2)}_{=1} - i(\underbrace{\mathbb{E}(XY)}_{=0} - \underbrace{\mathbb{E}(YX)}_{=0})\big) = 1.$$

Hence, for $z_i, z_j \in \{Z, \bar{Z}\}$ the pair $\{z_i, z_j\}$ will only give contribution 1 when $z_i \in Z$ and $z_j \in \bar{Z}$, or when $z_i \in \bar{Z}$ and $z_j \in Z$. Wick's formula for real Gaussians can now be rewritten for the complex case.

Theorem 3.7 (Wick's formula for complex Gaussians)**.** *Let $Z_1, Z_2, \dots, Z_p$ be complex independent standard Gaussian random variables and let $z_1, z_2, \dots, z_n \in \{Z_1, \bar{Z}_1, Z_2, \bar{Z}_2, \dots, Z_p, \bar{Z}_p\}$. Then Wick's formula states that mixed moments of n elements can be calculated using second moments:*

$$\mathbb{E}(z_1 \cdot z_2 \cdot \ldots \cdot z_n) = \sum_{\Pi \in \mathcal{P}_2(n)} \prod_{(i,j) \in \Pi} \mathbb{E}(z_i z_j),$$

where for $1 \leq s \leq p$

$$\mathbb{E}(z_i z_j) = \begin{cases} 1 & ;\ z_i \in Z_s,\ z_j \in \bar{Z}_s \text{ or } z_i \in \bar{Z}_s,\ z_j \in Z_s \\ \\ 0 & ;\ \text{otherwise.} \end{cases}$$

A combinatorical expression has been established for mixed moments of complex independent standard Gaussian random variables. This can now be applied to the mixed moments of elements of the matrix W_N, which is needed to be able to work with equation (9).

3.5 Counting cycles

In this section, equation (9) is further analyzed. According to Wick's formula for complex Gaussians, the equation can be rewritten in the following way:

$$\begin{aligned}\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k)) &= \frac{1}{N}\frac{1}{\sqrt{N}^k}\sum_{i_1,i_2,\ldots,i_k=1}^{N}\mathbb{E}(w_{i_1i_2}\cdot w_{i_2i_3}\cdot\ldots\cdot w_{i_ki_1})\\ &= \frac{1}{N}\frac{1}{\sqrt{N}^k}\sum_{i_1,i_2,\ldots,i_k=1}^{N}\sum_{\Pi\in\mathcal{P}_2(k)}\prod_{(a,b)\in\Pi}\mathbb{E}(w_{i_ai_{a+1}}w_{i_bi_{b+1}}).\end{aligned}$$

In Wick's formula for complex Gaussians is established that the second moment for standard complex Gaussian random variables only gives a contribution 1, when the two elements of a pair belong to each others conjugate distributions. Since the matrix W_N is Hermitian and therefore $w_{ij} = \bar{w}_{ji}$, the elements that belong to each others conjugate distribution are located in the mirrored positions from the diagonal. Also for elements on the diagonal $w_{ii} = \bar{w}_{ii}$ is true. So when does the term $\mathbb{E}(w_{i_ai_{a+1}}w_{i_bi_{b+1}})$ give a contribution unequal to 0 in both cases for elements in mirrored positions from the diagonal and for elements on the diagonal? When $i_a = i_{b+1}$ and $i_{a+1} = i_b$. Introduce the following function in order to summarize this in a clear way:

$$\delta_{xy} = \begin{cases} 1 & ;\ x = y \\ \\ 0 & ;\ x \neq y. \end{cases}$$

So $\mathbb{E}(w_{i_ai_{a+1}}w_{i_bi_{b+1}})$ only gives contribution 1 when $\mathbb{E}(w_{i_ai_{a+1}}w_{i_bi_{b+1}}) = \delta_{i_ai_{b+1}}\delta_{i_{a+1}i_b}$. Applying the above to equation (9) gives:

$$\begin{aligned}\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k)) &= \frac{1}{N}\frac{1}{\sqrt{N}^k}\sum_{i_1,i_2,\ldots,i_k=1}^{N}\sum_{\Pi\in\mathcal{P}_2(k)}\prod_{(a,b)\in\Pi}\mathbb{E}(w_{i_ai_{a+1}}w_{i_bi_{b+1}})\\ &= \frac{1}{N}\frac{1}{\sqrt{N}^k}\sum_{i_1,i_2,\ldots,i_k=1}^{N}\sum_{\Pi\in\mathcal{P}_2(k)}\prod_{(a,b)\in\Pi}\delta_{i_ai_{b+1}}\delta_{i_{a+1}i_b}.\end{aligned}$$

Here, $(a, b) \in \Pi$ is a more perspicuous expression for iterating over all pairs $\{w_{i_ai_{a+1}}, w_{i_bi_{b+1}}\}$ in a pairing Π. The two finite sums can be switched, yielding:

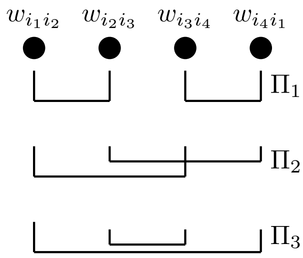
$$\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k)) = \frac{1}{N}\frac{1}{\sqrt{N}^k}\sum_{\Pi\in\mathcal{P}_2(k)}\sum_{i_1,i_2,\ldots,i_k=1}^{N}\prod_{(a,b)\in\Pi}\delta_{i_ai_{b+1}}\delta_{i_{a+1}i_b}. \tag{11}$$

So instead of first summing over all possible ways to assign the i_j's in $\{i_1, i_2, \ldots i_k\}$ to a number from 1 to N and then make pairings of the resulting $w_{i_1i_2}w_{i_2i_3}\cdots w_{i_ki_1}$, the finite sums are switched. Now k positions are set, all possible pairings are made for those positions. Then, the positions are filled with all the possible elements $w_{i_1i_2}w_{i_2i_3}\cdots w_{i_ki_1}$, where the $\{i_1, i_2, \ldots, i_k\}$ are assigned a number from 1 to N. An example on how in practise can be worked with the new obtained equation is given below.

Example 3.5. Examine $\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k))$ for $N = 3$ and $k = 4$:

$$\begin{aligned}\mathbb{E}(\frac{1}{3}Tr(\bar{W}_3^4)) &= \frac{1}{3}\frac{1}{\sqrt{3}^4}\sum_{i_1,i_2,i_3,i_4=1}^{3}\mathbb{E}(w_{i_1i_2}\cdot w_{i_2i_3}\cdot w_{i_3i_4}\cdot w_{i_4i_1})\\ &= \frac{1}{3}\frac{1}{\sqrt{3}^4}\sum_{\Pi\in\mathcal{P}_2(4)}\sum_{i_1,i_2,i_3,i_4=1}^{3}\prod_{(a,b)\in\Pi}\delta_{i_ai_{b+1}}\delta_{i_{a+1}i_b}.\end{aligned}$$

First, 4 positions are set and all possible pairings for those positions are made.



There are $3^4 = 81$ possible ways to assign i_1, i_2, i_3, i_4 to a number from 1 to 3. Below a short representation is given:

$$\begin{aligned}\{i_1, i_2, i_3, i_4\} =&\{1, 1, 1, 1\}, \{1, 1, 1, 2\}, \{1, 1, 2, 1\}, \{1, 2, 1, 1\},\\ &\{2, 1, 1, 1\}, \{1, 1, 2, 2\}, \dots, \{1, 2, 3, 2\}, \dots, \{2, 2, 2, 3\},\\ &\{2, 2, 3, 2\}, \{2, 3, 2, 2\}, \{3, 2, 2, 2\}, \dots, \{3, 3, 3, 3\}.\end{aligned}$$

Each of these assigned sets are applied to $w_{i_1i_2}w_{i_2i_3}w_{i_3i_4}w_{i_4i_1}$, and the resulting w_{ij}'s are placed on the 4 positions in the schematic above. Then pairings of the elements are made accordingly. Now, let the focus be on one of the assigned sets $\{i_1, i_2, i_3, i_4\} = \{1, 2, 3, 2\}$. Applying Wick's formula to $\mathbb{E}(w_{i_1i_2}w_{i_2i_3}w_{i_3i_4}w_{i_4i_1}) = \mathbb{E}(w_{12}w_{23}w_{32}w_{21})$ gives:

$$\mathbb{E}(w_{12}w_{23}w_{32}w_{21}) = \sum_{\Pi\in\mathcal{P}_2(4)} \prod_{(a,b)\in\Pi} \delta_{i_ai_{b+1}}\delta_{i_{a+1}i_b}.$$

So for each Π_1, Π_2 and Π_3, the pairs within are checked whether they are an allowed pairing.

$$\begin{aligned}\Pi_1 &= \{\{w_{i_1i_2}, w_{i_2i_3}\}, \{w_{i_3i_4}, w_{i_4i_1}\}\} = \{\{w_{12}, w_{23}\}, \{w_{32}, w_{21}\}\}\\ \Pi_2 &= \{\{w_{i_1i_2}, w_{i_3i_4}\}, \{w_{i_2i_3}, w_{i_4i_1}\}\} = \{\{w_{12}, w_{32}\}, \{w_{23}, w_{21}\}\}\\ \Pi_3 &= \{\{w_{i_1i_2}, w_{i_4i_1}\}, \{w_{i_2i_3}, w_{i_3i_4}\}\} = \{\{w_{12}, w_{21}\}, \{w_{23}, w_{32}\}\}\end{aligned}$$

It can already be seen that pairing Π_3 is the only pairing with pairs that contain elements of the matrix W_N on the mirrored positions. So this is the only pairing that will give a contribution. Wick's formula should come to the same conclusion:

$$\begin{aligned}\mathbb{E}(w_{12}w_{23}w_{32}w_{21}) &= \prod_{(a,b)\in\Pi_1} \delta_{i_ai_{b+1}}\delta_{i_{a+1}i_b} + \prod_{(a,b)\in\Pi_2} \delta_{i_ai_{b+1}}\delta_{i_{a+1}i_b} + \prod_{(a,b)\in\Pi_3} \delta_{i_ai_{b+1}}\delta_{i_{a+1}i_b}\\ &= \delta_{i_1i_3}\delta_{i_2i_2} \cdot \delta_{i_3i_1}\delta_{i_4i_4} + \delta_{i_1i_4}\delta_{i_2i_3} \cdot \delta_{i_2i_1}\delta_{i_3i_4} + \delta_{i_1i_1}\delta_{i_2i_4} \cdot \delta_{i_2i_4}\delta_{i_3i_3}\\ &= \delta_{13}\delta_{22} \cdot \delta_{31}\delta_{22} + \delta_{12}\delta_{23} \cdot \delta_{21}\delta_{32} + \delta_{11}\delta_{22} \cdot \delta_{22}\delta_{33}\\ &= 0 \cdot 1 \cdot 0 \cdot 1 + 0 \cdot 0 \cdot 0 \cdot 0 + 1 \cdot 1 \cdot 1 \cdot 1 = 1.\end{aligned}$$

Indeed, also by Wick's formula the number of allowed pairings for the particular case $\{i_1, i_2, i_3, i_4\} = \{1, 2, 3, 2\}$ is 1.

From here, some algebraic tools are introduced to rewrite the equation in (11) further. Let the pairing Π be looked at as a permutation in the symmetric group S_k of k elements, in which the pairs $(a, b) \in \Pi$ become cycles (ab) of the permutation that send a to b and b to a. Furthermore, let $\gamma = (12 \dots k) \in S_k$ be the permutation which represents a shift by 1 modulo k. Then, since $\Pi(a) = b$ and $\gamma(a) = a + 1$ in the condition $\delta_{i_ai_{b+1}}\delta_{i_{a+1}i_b}$, $i_{b+1} = i_{\Pi(a)+1} = i_{\gamma(\Pi(a))}$. Similarly, $i_{a+1} = i_{\gamma(\Pi(b))}$. So the product in (11) can be changed to a product over one variable:

$$\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k)) = \frac{1}{N}\frac{1}{\sqrt{N}^k} \sum_{\Pi\in\mathcal{P}_2(k)} \sum_{i_1,i_2,\dots,i_k=1}^{N} \prod_{a=1}^{k} \delta_{i_ai_{\gamma(\Pi(a))}}. \tag{12}$$

An example on how in practise can be worked with the new change in the equation is given below.

Example 3.6. Examine $\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k))$ for $N=3$ and $k=4$:

$$\begin{aligned}\mathbb{E}(\frac{1}{3}Tr(\bar{W}_3^4)) &= \frac{1}{3}\frac{1}{\sqrt{3}^4}\sum_{i_1,i_2,i_3,i_4=1}^{3}\mathbb{E}(w_{i_1i_2}\cdot w_{i_2i_3}\cdot w_{i_3i_4}\cdot w_{i_4i_1})\\ &= \frac{1}{3}\frac{1}{\sqrt{3}^4}\sum_{\Pi\in\mathcal{P}_2(4)}\sum_{i_1,i_2,i_3,i_4=1}^{3}\prod_{a=1}^{4}\delta_{i_a i_{\gamma(\Pi(a))}}.\end{aligned}$$

Recall example 3.5 and again, focus on the assigned set $\{i_1,i_2,i_3,i_4\}=\{1,2,3,2\}$. First the $\gamma\Pi$'s for the three different pairings $\Pi_1,\Pi_2,\Pi_3\in\mathcal{P}_2(4)$ need to be generated. Remember the pairings Π_1,Π_2,Π_3:

$$\begin{aligned}\Pi_1 &= \{\{w_{i_1i_2},w_{i_2i_3}\},\{w_{i_3i_4},w_{i_4i_1}\}\},\\ \Pi_2 &= \{\{w_{i_1i_2},w_{i_3i_4}\},\{w_{i_2i_3},w_{i_4i_1}\}\},\\ \Pi_3 &= \{\{w_{i_1i_2},w_{i_4i_1}\},\{w_{i_2i_3},w_{i_3i_4}\}\}.\end{aligned}$$

For $\gamma=(1234)$ and for pairs $(a,b)=\{w_{i_ai_{a+1}},w_{i_bi_{b+1}}\}$, the following permutations arise:

$$\begin{aligned}\Pi_1 = (12)(34) &\implies \gamma\Pi_1=(13)(2)(4),\\ \Pi_2=(13)(24) &\implies \gamma\Pi_2=(1432),\\ \Pi_3=(14)(23) &\implies \gamma\Pi_3=(1)(24)(3).\end{aligned}$$

Applying the new obtained equation to the specific case where $\{i_1,i_2,i_3,i_4\}=\{1,2,3,2\}$ like in example 3.5 gives:

$$\begin{aligned}\mathbb{E}(w_{12}w_{23}w_{32}w_{21}) &= \sum_{\Pi\in\mathcal{P}_2(4)}\prod_{a=1}^{4}\delta_{i_ai_{\gamma(\Pi(a))}}\\ &= \prod_{a=1}^{4}\delta_{i_ai_{\gamma(\Pi_1(a))}}+\prod_{a=1}^{4}\delta_{i_ai_{\gamma(\Pi_2(a))}}\prod_{a=1}^{4}\delta_{i_ai_{\gamma(\Pi_3(a))}}\\ &= \delta_{i_1i_3}\delta_{i_2i_2}\delta_{i_3i_1}\delta_{i_4i_4}+\delta_{i_1i_4}\delta_{i_2i_1}\delta_{i_3i_2}\delta_{i_4i_3}+\delta_{i_1i_1}\delta_{i_2i_4}\delta_{i_3i_3}\delta_{i_4i_2}\\ &= \delta_{13}\delta_{22}\delta_{31}\delta_{22}+\delta_{12}\delta_{21}\delta_{32}\delta_{23}+\delta_{11}\delta_{22}\delta_{33}\delta_{22}\\ &= 0\cdot1\cdot0\cdot1+0\cdot0\cdot0\cdot0+1\cdot1\cdot1\cdot1\\ &= 1.\end{aligned}$$

Analyzing equation (12) more closely, it can be observed that the product $\prod_{a=1}^{k}\delta_{i_ai_{\gamma(\Pi(a))}}\neq 0$ when the function $i:[k]\to[N]$ is constant under the cycles of $\gamma\Pi$. Read i_a for $a\in[k]$ as a function i of a. When generating $i_{\gamma\Pi(a)}\in[N]$, i_a and $i_{\gamma\Pi(a)}$ are only the same in $[N]=1,2,\ldots,N$ when i_a is constant under the cycles of $\gamma\Pi$. This yields a constraint on the sum $\sum_{i_1,i_2,\ldots,i_k=1}^{N}$, because the i_a's must be constant on each cycle. So there remain $\#\gamma\Pi$ free variables for which each can obtain N different values. So equation (12) becomes:

$$\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k))=\frac{1}{N}\frac{1}{\sqrt{N}^k}\sum_{\Pi\in\mathcal{P}_2(k)}N^{\#\gamma\Pi}. \tag{13}$$

Example 3.7. For $N=3$ and $k=4$, recall from example 3.6 the permutations and their amount of cycles:

$$\begin{aligned}\gamma\Pi_1=(13)(2)(4) &\implies \#\gamma\Pi_1=3,\\ \gamma\Pi_2=(1432) &\implies \#\gamma\Pi_2=1,\\ \gamma\Pi_3=(1)(24)(3) &\implies \#\gamma\Pi_3=3.\end{aligned}$$

Then $\sum_{i_1,i_2,i_3,i_4=1}^{3}\prod_{a=1}^{4}\delta_{i_ai_{\gamma(\Pi(a))}}$ only gives a contribution for those i_1,i_2,i_3,i_4 that are constant under the cycles of $\gamma\Pi$. So for all $a=1,2,3,4$:

$$\Pi_1 \ : \ \delta_{i_a i_{\gamma(\Pi_1(a))}} = 1 \iff \underbrace{\{i_1,i_2,i_3,i_4\} = \{1,1,1,1\},\{1,2,1,2\},\ldots,\{3,2,3,2\},\{3,3,3,3\}}_{N^{\#\gamma\Pi_1}=3^2=9}$$

$$\Pi_2 \ : \ \delta_{i_a i_{\gamma(\Pi_2(a))}} = 1 \iff \underbrace{\{i_1,i_2,i_3,i_4\} = \{1,1,1,1\},\{2,2,2,2\},\{3,3,3,3\}}_{N^{\#\gamma\Pi_2}=3^1=3}$$

$$\Pi_3 \ : \ \delta_{i_a i_{\gamma(\Pi_3(a))}} = 1 \iff \underbrace{\{i_1,i_2,i_3,i_4\} = \{1,1,1,1\},\{1,2,1,2\},\ldots,\{1,2,3,2\},\ldots,\{3,3,3,3\}.}_{N^{\#\gamma\Pi_3}=3^3=27}$$

Therefore:

$$\sum_{\Pi\in\mathcal{P}_2(4)} \sum_{i_1,i_2,i_3,i_4=1}^{3} \prod_{a=1}^{4} \delta_{i_a i_{\gamma(\Pi(a))}} = \sum_{\Pi\in\mathcal{P}_2(4)} 3^{\#\gamma\Pi} = 3^{\#\gamma\Pi_1} + 3^{\#\gamma\Pi_2} + 3^{\#\gamma\Pi_3} = 27+3+27 = 57.$$

So, the averaged k-th order moments of the ESD of $\bar{W}_N$ can be rewritten by identifying pairings as permutations and counting cycles like in equation (13).

Theorem 3.8. *Let $\bar{W}_N$ be a GUE matrix. Let the permutation $\gamma \in S_k$ be $\gamma = (12\ldots k)$ and let the permutation $\Pi \in S_k$ consist of 2-cycles containing the elements of pairs of pairings $\Pi \in \mathcal{P}_2(k)$. Then the averaged k-th order moment of the ESD of $\bar{W}_N$ is:*

$$\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^k)) = \sum_{\Pi\in\mathcal{P}_2(k)} N^{\#(\gamma\pi)-\frac{k}{2}-1}.$$

Example 3.8. Investigate the averaged 4-th order moment of the ESD of $\bar{W}_N$. According to Theorem 3.8:

$$\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^4)) = \sum_{\Pi\in\mathcal{P}_2(4)} N^{\#\gamma\Pi-3}.$$

By example 3.3 $\#\mathcal{P}_2(4) = 3$ and for all 3 pairings by example 3.6 $\#\gamma\Pi_1 = 2$, $\#\gamma\Pi_2 = 1$ and $\#\gamma\Pi_3 = 3$ So:

$$\begin{aligned}\mathbb{E}(\frac{1}{N}Tr(\bar{W}_N^4)) &= \sum_{\Pi\in\mathcal{P}_2(4)} N^{\#\gamma\Pi-3} = N^{\#\gamma\Pi_1-3} + N^{\#\gamma\Pi_2-3} + N^{\#\gamma\Pi_3-3} \\ &= N^{3-3} + N^{3-3} + N^{1-3} = 2 + \frac{1}{N^2} \xrightarrow{N\to\infty} 2 = C_2.\end{aligned}$$

So indeed:

$$\mathbb{E}\big(\frac{1}{N}Tr(\bar{W}_N^{2k})\big) \xrightarrow{N\to\infty} C_k, \text{ for } k=2.$$

The expression for the averaged k-th order moments of the ESD of $\bar{W}_N$ has now been reduced to an identity in which only the amount of cycles of the combined permutation $\gamma\Pi$, where Π is a pairing of k elements interpreted as a permuation, is needed. Now the question remains; what pairings yield a non N-term in the sum in Theorem 3.8? The answer is the non-crossing pairings. This will be proven in the next section.

3.6 Non-crossing pairings

Intuitively, the definition of a non-crossing pairing is of course a pairing in which the pair lines do not cross each other. In example 3.4 it is clear that the pairings Π_1, Π_3, Π_7 and Π_9 are non-crossing pairings. Below follows a more exact definition of non-crossing pairings.

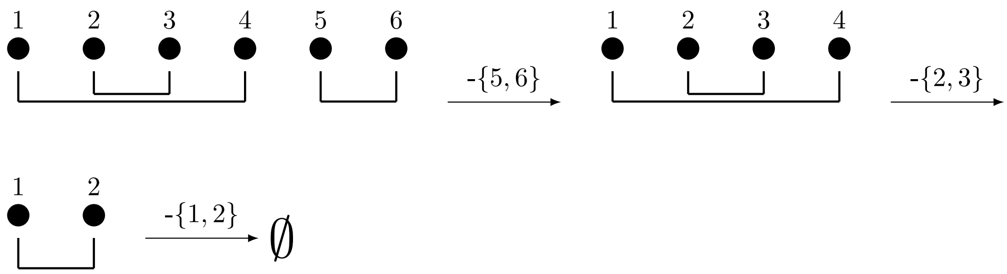
Definition 3.5 (Non-crossing pairing)**.** Let k be even. A pairing $\Pi \in \mathcal{P}_2(k)$ is non-crossing if there are no pairs $\{a,c\},\{b,d\} \in \Pi$ where $1 \le a < b < c < d \le k$. Denote:

$$\mathcal{NC}_2(k) := \{\Pi \in \mathcal{P}_2(k) \ : \ \Pi \text{ is non-crossing}\} \subseteq \mathcal{P}_2(k).$$

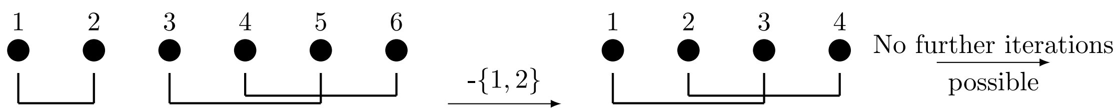
A pairing $\Pi \in \mathcal{P}_2(k)$ such that $\Pi \notin \mathcal{NC}_2(k)$ is called a crossing pairing.

Note that for even k, all non-crossing pairings in $\mathcal{NC}_2(k)$ have the property that the pair containing element 1 is of the form $\{1, 2s\}$, where s is even. If this was not the case and the pair $\{1, 2s\}$ exists for an odd s, the number of elements between 1 and $2s$ would also be odd, which would force one of the elements between 1 and $2s$ to have to pair outside the elements from 1 to $2s$, resulting in a crossing. So the pair containing element 1 must be of the form $\{1, 2s\}$ with s even. The remaining pairs can then only pair within $\{2, \ldots, 2s-1\}$ or $\{2s+1, k\}$. For the those remaining pairs the same fact holds. This yields an iterative structure for the build up of non-crossing pairings, resulting that at least one pair in a non-crossing pairing must be of the form $\{i, i+1\}$ for $1 \leq i \leq k-1$. Now let a non-crossing pairing being broken down in the following way. There has been concluded that any non-crossing pairing must at least have one pair $\{i, i+1\}$ consisting of neighboring elements. If this pair is removed, a non-crossing pairing of $k-2$ elements remain. The same can be said for the remaining non-crossing pairing. Pairs containing neighboring elements can keep being removed from the remaining non-crossing pairing, until the empty set remains. This is not the case for crossing pairings. Below follows an example.

Example 3.9. Let $k = 6$ and consider the following non-crossing pairing $\Pi = \{\{1,4\},\{2,3\},\{5,6\}\} \in \mathcal{NC}_2(6)$. In this case there are two neighboring pairs in the initial pairing. Remove pair $\{5,6\}$. The remaining non-crossing pairing becomes $\Pi = \{\{1,4\},\{2,3\}\} \in \mathcal{NC}_2(4)$. Remove again a neighboring pair. This time, the only option is pair $\{2,3\}$. The remaining non-crossing pairing becomes $\Pi = \{\{1,2\}\} \in \mathcal{NC}_2(2)$. Note that by removing a neighboring pair, the nodes get re-ordered from 1, such that neighboring pairs keep appearing. The final pair $\{1,2\}$ can now be removed, resulting in the empty set. See the schematic representation below.



Now consider a crossing pairing $\Pi = \{\{1,2\},\{3,5\},\{4,6\}\} \in \mathcal{NC}_2(6)$. Remove the pair $\{1,2\}$ resulting in $\Pi = \{\{1,3\},\{2,4\}\} \in \mathcal{NC}_2(4)$. Now, a pairing without any pairs that consists of neighboring elements is left, and there can not be iterated further to get to the empty set. See the schematic representation below.



In example 3.8 a small glimpse of the fact that only non-crossing pairings yield non N-terms in the sum of Theorem 3.8. These are the only terms of importance, because when N is send to infinity, only these terms remain. If this fact is proven, and it can also be shown that the number of non-crossing pairings of a set $[2k]$ are the same as the Catalan numbers, $\#\mathcal{NC}_2(2k) = C_k$, the proof of Wigner's Semicircle Law in Theorem 2.1 is concluded.

Proposition 3.9. *Let k be even. Let the permutation $\gamma \in S_k$ be $\gamma = (12\ldots k)$ and let the permutation $\Pi \in S_k$ consist of 2-cycles containing the elements of pairs of pairings $\Pi \in \mathcal{P}_2(k)$. Then:*

i) $\#\gamma\Pi - \frac{k}{2} - 1 \leq 0$ *for all* $\Pi \in \mathcal{P}_2(k)$,

ii) $\#\gamma\Pi - \frac{k}{2} - 1 = 0 \iff \Pi \in \mathcal{NC}_2(k)$.

Proof. First the following claim is proven.
Claim: For all $\Pi \in \mathcal{P}_2(k) : \{i, i+1\} \in \Pi \iff \gamma\Pi$ contains the cycles $(i+1)$ and $(i, i+2, \ldots)$.
Proof: Let $\Pi \in \mathcal{P}_2(k)$ and $\{i, i+1\} \in \Pi$. Then Π is a permutation containing the cycle $(i\ i+1)$ and

$i+1 \xrightarrow{\Pi} i \xrightarrow{\gamma} i+1$. Indeed $\gamma\Pi(i+1) = (i+1)$ and therefore $(i+1)$ is a cycle in $\gamma\Pi$. Furthermore, $i \xrightarrow{\Pi} i+1 \xrightarrow{\gamma} i+2$, so $\gamma\Pi(i) = (i+2)$ and therefore $(i, i+2, \ldots)$ is a cycle in $\gamma\Pi$. Now for the reverse. Let $\gamma\Pi$ contain the cycles $(i+1)$ and $(i, i+2, \ldots)$. Then $\gamma\Pi(i+1) = (i+1)$ and $(i+1)$ is a fixed point of $\gamma\Pi$. So $\gamma\Pi(i+1) = (i+1) \xrightarrow{\gamma^{-1}} \Pi(i+1) = \gamma^{-1}(i+1) \implies \Pi(i+1) = i$. Indeed $\{i, i+1\} \in \Pi$, which concludes the proof of the claim.
First 'ii) $\Longleftarrow$ ' is shown. Let $\Pi \in \mathcal{NC}_2(k)$ for even k. Then there exists a pair $\{i, i+1\} \in \Pi$, which can be removed. Doing so, by the previous claim, yields in removal of the cycle $(i+1)$ in $\gamma\Pi$ and in removal of the element i in the cycle $(i, i+2, \ldots)$. The number of elements k is reduced by 2 and the result is a non-crossing pairing in $\mathcal{NC}_2(k-2)$ which again contains a pair of the kind $\{i, i+1\}$. After $\frac{k}{2} - 1$ iterations of removal of pairs of neighboring element, there are $k - 2(\frac{k}{2} - 1) = 2$ elements left. Since the remaining 2 elements are neighbors, by the claim, there are 2 cycles left after $\frac{k}{2} - 1$ iterations, namely (i) and $(i+1)$. Again by the claim, per iteration one cycle was removed from the cycles of $\gamma\Pi$. So $\#\gamma\Pi = 2 + 1(\frac{k}{2} - 1) = \frac{k}{2} + 1$.
Secondly 'ii) $\Longrightarrow$ ' is shown by proof by contradiction. Let $\#\gamma\Pi - \frac{k}{2} - 1 = 0$. Suppose $\Pi \notin \mathcal{NC}_2(k)$ is a crossing pairing. Remove all pairs of neighbors until it cannot be done anymore. Since Π is crossing, this can only be done $\frac{k}{2} - x$ amount of times for $x > 1$, since removing pairs of neighbors does not result in the empty set. After removal of pairs of neighbors, by the claim, for the resulting pairing $\tilde{\Pi}$ the permutation $\gamma\tilde{\Pi}$ does not have any fixed points. So each cycle in $\gamma\tilde{\Pi}$ has at least 2 elements. So $\#\gamma\tilde{\Pi} \leq x$, which means $\#\gamma\Pi = \#\gamma\tilde{\Pi} + 1(\frac{k}{2} - x) \leq x + \frac{k}{2} - x = \frac{k}{2}$. However, $\#\gamma\Pi - \frac{k}{2} - 1 = 0$ so there has been reached a contraction. Therefore $\Pi \in \mathcal{NC}_2(k)$.
Now for the proof of i). By 'ii) $\Longleftarrow$ ' for $\Pi \in \mathcal{NC}_2(k) \subseteq \mathcal{P}_2(k)$, $\#\gamma\Pi - \frac{k}{2} - 1 = 0$ and by 'ii) $\Longrightarrow$ ' for $\Pi \notin \mathcal{NC}_2(k)$, $\#\gamma\Pi \leq \frac{k}{2} \leq \frac{k}{2} + 1$. So for all $\Pi \in \mathcal{P}_2(k)$: $\#\gamma\Pi - \frac{k}{2} - 1 \leq 0$. □

Remark 3.5.1. Because of the results of Proposition 3.9 and Theorem 3.8, it becomes clear why the Wigner matrices need to be scaled by $\frac{1}{\sqrt{N}}$. Without the scaling $\frac{1}{\sqrt{N}}$, there would not be any N-terms with a power 0 in the sum and only terms with a power bigger than 0. Therefore, the limit would blow up if N is send to infinity in Theorem 2.1. So scaling the Wigner matrices by $\frac{1}{\sqrt{N}}$ is of great importance in order for the the convergence in Wigner's Semicircle Law to happen.

3.7 Proof of Wigner's Semicircle Law for GUE matrices

Now, almost all the tools are acquired to prove the main Theorem 2.1. Recall what needed to be proven in Wigner's Semicircle Law:

$$\lim_{N\to\infty} \mathbb{E}(m_{\mu_{\bar{W}_N},k}) = m_{\mu_{sc},k}.$$

By Equation 6 and Theorem 3.8, the following expression for the averaged k-th order moments the ESD of $\bar{W}_N$ was obtained:

$$\mathbb{E}(m_{\mu_{\bar{W}_N},k}) = \sum_{\Pi\in\mathcal{P}_2(k)} N^{\#(\gamma\pi)-\frac{k}{2}-1}.$$

By Proposition 3.2, the following expressions for the k-th order moments of the semicircle was proven:

$$m_{\mu_{sc},2k+1} = 0,$$
$$m_{\mu_{sc},2k} = C_k.$$

Here C_k denotes the k-th Catalan number. Let k be odd. In Proposition 3.5 it was proven that $\#\mathcal{P}_2(2k+1) = 0$. Consequently, $\mathbb{E}(m_{\mu_{\bar{W}_N},2k+1}) = 0$ and for odd k Theorem 2.1 is true:

$$\lim_{N\to\infty} \mathbb{E}(m_{\mu_{\bar{W}_N},2k+1}) = 0 = m_{\mu_{sc},2k+1}.$$

Let k be even. In Proposition 3.9 it was shown that for a non-crossing pairing $\Pi \in \mathcal{NC}_2(2k)$, $\#\gamma\Pi - \frac{2k}{2} - 1 = 0$ and for all other pairings $\Pi \notin \mathcal{NC}_2(2k)$, $\#\gamma\Pi - \frac{2k}{2} - 1 \leq 0$. Therefore:

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}(m_{\mu_{\bar{W}_N},2k}) &= \lim_{N\to\infty} \sum_{\Pi\in\mathcal{P}_2(2k)} N^{\#(\gamma\pi)-\frac{2k}{2}-1} = \sum_{\Pi\in\mathcal{P}_2(2k)} \lim_{N\to\infty} N^{\#(\gamma\pi)-\frac{2k}{2}-1} \\
&= \sum_{\Pi\in\mathcal{NC}_2(2k)} \lim_{N\to\infty} N^{\#(\gamma\pi)-\frac{2k}{2}-1} + \sum_{\Pi\notin\mathcal{NC}_2(2k)} \lim_{N\to\infty} N^{\#(\gamma\pi)-\frac{2k}{2}-1} \\
&= \sum_{\Pi\in\mathcal{NC}_2(2k)} 1 + \sum_{\Pi\notin\mathcal{NC}_2(2k)} 0 \\
&= \sum_{\Pi\in\mathcal{NC}_2(2k)} 1 \\
&= \#\mathcal{NC}_2(2k).
\end{aligned}$$

It remains to be proven that $\#\mathcal{NC}_2(2k) = C_k$. It suffices to show that $\#\mathcal{NC}_2(2k)$ and C_k share the same recurrence relation and the same starting value. The sequences $\#\mathcal{NC}_2(2k)$ and C_k indeed share the same starting value for $k = 1$: $\#\mathcal{NC}_2(2) = 1 = C_1$. Furthermore, by Proposition 3.1[2], C_k satisfies the following recurrence relation:

$$C_k = \sum_{l=1}^{k} C_{l-1}C_{k-l}.$$

To show this, let $\Pi \in \mathcal{NC}_2(2k)$. Remember that for a non-crossing pairing Π the pair containing element 1 must be of the form $\{1, 2l\}$. Let $\Pi^{'} \in \mathcal{NC}_2(2l-2)$ and $\Pi^{''} \in \mathcal{NC}_2(2k-2l)$. Then for every $l = 1, 2, \ldots, k$, Π can be split in the following way:

$$\Pi = \{1, 2l\} \cup \Pi^{'} \cup \Pi^{''}.$$

So $\#\mathcal{NC}_2(2k)$ satisfies the following recurrence relation:

$$\begin{aligned}
\#\mathcal{NC}_2(2k) &= \sum_{l=1}^{k} \#\mathcal{NC}_2(2l-2))\#\mathcal{NC}_2(2k-2l)) \\
&= \sum_{l=1}^{k} \#\mathcal{NC}_2(2(l-1))\#\mathcal{NC}_2(2(k-l)).
\end{aligned}$$

It follows that $\#\mathcal{NC}_2(2k) = C_k$. As a result, also the averaged even moments of the ESD of $\bar{W}_N$ converge to the even moments of the semicircle distribution:

$$\lim_{N\to\infty} \mathbb{E}(m_{\mu_{\bar{W}_N},2k}) = C_k = m_{\mu_{sc},2k}.$$

Hence, for all k:

$$\lim_{N\to\infty} \mathbb{E}(m_{\mu_{\bar{W}_N},k}) = m_{\mu_{sc},k}.$$

This concludes the proof of Wigner's Semicircle Law for GUE matrices.

References

[Anderson et al., 2009] Anderson, G. W., Guionnet, A., Lyon, E., and Zeitouni, O. (2009). *An Introduction to Random Matrices.*

[Fraleigh et al., 1990] Fraleigh, J., Beauregard, R., and Katz, V. (1990). *Linear Algebra.* Addison-Wesley Publishing Company.

[Speicher, 2020] Speicher, R. (2020). Lecture notes on "random matrices".

[Z.,] Z., R. Catalan numbers - algebraic proof of the recurrence relation. Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/3304420 (version: 2019-07-26).

[Wolf, 2021] Wolf, Vanessa, "Random Matrix Theory: A Combinatorial Proof of Wigner's Semicircle Law" (2021). Scripps Senior Theses. 1683. https://scholarship.claremont.edu/scripps$_theses/1683$

A Appendix Python Code

```
##################################################################################################

import math as m
import numpy as np
import matplotlib.pyplot as plt

##################################################################################################

## Generating the the ESD of W_N with a variety of distributions against the Semicircle distribution

##################################################################################################

## Generating the NxN Wigner Matrix
N=500 #matrix size
W = np.zeros(shape=(N,N), dtype=np.complex128)
eigval_arr = np.array([])

###Place desired distribution here (see below for the different distributions)###
##Normal distribution
mu = 0 #mean
std_dev = 1 #variable s.t. variance=1 for complex random variables on the off-diagonal

for i in range(N):
    for j in range(i,N): #elements of W on and above diagonal are being generated
        if i==j:
            W[i,j] = np.random.normal(mu,std_dev,1) ###Place distribution for diagonal elements here###
        else:
            real_part = np.random.normal(mu,std_dev,1) ###Place distribution for non-diagonal elements here###
            imag_part = np.random.normal(mu,std_dev,1) ###Place distribution for non-diagonal elements here###
            W[i,j]= (complex(real_part, imag_part))/np.sqrt(2)
            W[j,i] = np.conj(W[i, j]) #W is Hermitian to ensure real eigenvalues

W_bar = W/np.sqrt(N) #W is scaled


## Generating the eigenvalues of the Wigner Matrix
eigval_W_bar = np.linalg.eig(W_bar)[0] #N eigenvalues of W_bar
                                       #np.linalg.eig()[0] -> eigenvalues (the one used)
                                       #np.linalg.eig()[1] -> normalized eigenvectors (not needed)
eigval_arr = np.append(eigval_arr, eigval_W_bar)

## Generating the density function of sigma_sc of the Semicircle distribution
x = np.linspace(-2.0, 2.0, num=250)
sigma_sc = (1/(2*np.pi))*np.sqrt(4 - x**2)


## Plotting the density function of the Semicircle distribution
plt.plot(x, sigma_sc, 'r', label='$'r'\sigma_{sc}$', linewidth=4, color='#EE4B2B')


## Plotting the Empircal Spectral Distribution of the Wigner Matrix
plt.hist(eigval_arr, bins=30, range=(-3,3),normed=True, histtype='bar', ec='black') #normed=True -> normalized


## Plot specifications
plt.legend(fontsize=30)
plt.xticks(fontsize=30)
plt.yticks(fontsize=30)
plt.show()

##################################################################################################

###Different distributions to copy into the above code

##Normal distribution
#mu = 0 #mean
#std_dev = 1 #variable s.t. variance=1 for complex random variables on the off-diagonal

#np.random.normal(mu,size=1) #for diagonal elements
#np.random.normal(mu,std_dev,1) #for non-diagonal elements

##Laplace distribution
#mu = 0 #mean
#std_dev = 1/np.sqrt(2) #variable s.t. variance=1 for complex random variables on the off-diagonal

#np.random.laplace(mu,size=1) #for diagonal elements
#np.random.laplace(mu,std_dev,1) #for non-diagonal elements

#Logistic Distribution
#mu = 0 #mean
#std_dev = np.sqrt(3)/np.pi #variable s.t. variance=1 for random variables on the off-diagonal

#np.random.logistic(mu,size=1) #for diagonal elements
#np.random.logistic(mu,std_dev,1) #for non-diagonal elements
```

```python
###########################################################################################

import math as m
import numpy as np
import matplotlib.pyplot as plt

###########################################################################################

## Generating the ESD of W_N and the Semicirle distribution with hatched area between interval [s,t]

###########################################################################################

# Generating the NxN Wigner Matrix
N=500 #matrix size
W = np.zeros(shape=(N,N), dtype=np.complex128)
eigval_arr = np.array([])

mu = 0 #mean for real standard Gaussian random variables
std_dev = 1 #variable s.t. variance=1 for real random Gaussian variables

for i in range(N):
    for j in range(i,N): #elements of W on and above diagonal are being generated
        if i==j:
            W[i,j] = np.random.normal(mu,std_dev,1)
        else:
            real_part = np.random.normal(mu,std_dev,1)
            imag_part = np.random.normal(mu,std_dev,1)
            W[i,j]= (complex(real_part, imag_part))/np.sqrt(2) #ensure variance=1 for complex standard Gaussian random variables
            W[j,i] = np.conj(W[i, j]) #W is Hermitian to ensure real eigenvalues

W_bar = W/np.sqrt(N) #W is scaled


## Generating the eigenvalues of the Wigner Matrix
eigval_W_bar = np.linalg.eig(W_bar)[0] #N eigenvalues of W_bar
                                       #np.linalg.eig()[0] -> eigenvalues (the one used)
                                       #np.linalg.eig()[1] -> normalized eigenvectors (not needed)
eigval_arr = np.append(eigval_arr, eigval_W_bar)

## Generating the density function of sigma_sc of the Semicircle distribution
fig, ax = plt.subplots()
x = np.linspace(-2.0, 2.0, num=250)
sigma_sc = (1/(2*np.pi))*np.sqrt(4 - x**2)



## Plotting the density function of the semicircle distribution with hatched area between interval [s,t]
#plt.plot(x, sigma_sc, 'r', label='$'r'\sigma_{sc}$', linewidth=4, color='#EE4B2B')
#plt.fill_between(x, sigma_sc, where= (0.5<x)&(x<1.5), hatch='/')


## Plotting the Empircal Spectral Distribution of the Wigner Matrix with hatched area between interval [s,t]
n, bins, patches = ax.hist(eigval_arr, bins=30, range=(-3,3),normed=True, histtype='bar', ec='black') #normed=True -> normalized count
for i in range(len(patches)):
    if bins[i] >= 0.4 and bins[i+1] <= 1.7:
        patches[i].set_hatch('/')
    else:
       patches[i].set_facecolor('#98F5FF')



## Plot specifications
plt.xticks([0.5,1,1.5], ['[s',',','t]'])
plt.legend(fontsize=30)
#plt.title('Histogram of the ESD of eigenvalues of $\overline{W_{'+str(N)+'}}$',fontsize=35)
plt.xticks(fontsize=50)
plt.yticks(fontsize=30)
#plt.yticks(np.array([0,1/(np.pi)]),['0',r'$\frac{1}{\pi}$'])
plt.show()
```