CIMAT
Centro de Investigación en Matemáticas, A.C.

NON-COMMUTATIVE PROBABILITY AND RANDOM MATRICES WITH DISCRETE SPECTRUM

T E S I S

Que para obtener el grado de

Maestro en Ciencias

con Especialidad en

Probabilidad y Estadística

Presenta

Adrián de Jesús Celestino Rodríguez

Directores de Tesis:

Prof. Dr. Roland Speicher

Dr. Octavio Arizmendi Echegaray

Autorización de la versión final

Guanajuato, Gto., 19 de octubre de 2018

Centro de Investigación en Matemáticas, A.C.

Thesis manuscript to obtain the degree of

Master of Science in Probability and Statistics

Non-commutative Probability and Random Matrices with Discrete Spectrum

Author:
Adrián de Jesús
Celestino Rodríguez

Thesis Supervisors:
Prof. Dr. Roland Speicher
Dr. Octavio Arizmendi Echegaray

October 19, 2018

CIMAT

To my parents,

Carlos and Verónica.

Acknowledgments

In these lines, I would like to express my gratitude to every person who in some way contributed to the realization of this project. A new stage of two years with a lot of learning has concluded successfully.

First of all, I want to thank to my parents for all the support given to each project that I have started and the new one that I will soon.

I also want to thank to my friends and classmates of the CIMAT-DEMAT community who made my stay in Guanajuato very enjoyable.

Specially, I want to thank to Dr. Octavio Arizmendi Echegaray for his advising, patience, trust and friendship throughout my studies in CIMAT-DEMAT. Thanks to him, I started this journey to beautiful mathematics. I am sure his teaching and advice will help in my doctorate studies. I also thank to Dr. Víctor Pérez-Abreu and Dr. Mario Díaz for their invaluable comments to this manuscript in order to improve it.

I thank to Dr. Roland Speicher and his Research Group in Free Probability in Saarland University. I spent very nice three months as part of the group, knowing a different way of working in Free Probability.

Finally, I want to thank to CIMAT and CONACYT for all their support that I had during my graduate studies. My studies had not been possible without the scholarship granted by CONACYT for two years.

Contents

Introduction

Random Matrix Theory is the branch of mathematics which studies measurable functions with values in matrices. It was initiated through the work of Wishart in 1920's. In this theory, it is of interest to study properties related with the spectrum of random matrices, such as the eigenvalues distribution, singularity of random matrices, and the existence of properties in the limit when the dimension of the random matrices goes to infinity. In this direction, Wigner proved in the 1950's that the averaged spectral distribution of certain random matrices converges to the semicircle distribution.

On the other hand, Free Probability Theory was introduced by Dan Voiculescu in the 1980's aiming to solve the problem of isomorphism between von Neumann algebras generated by free groups. This theory considers analogous aspects with classical probability theory, but in the non-commutative framework of operator algebras, and takes into account non-commutative random variables and a notion of independence called "freeness".

Free Probability Theory has become an increasing and relevant research area in mathematics, and nowadays it is a very active field. Moreover, it has relationships with different branches of mathematics such as combinatorics, operator algebras, probability, representations of symmetric groups, mathematical physics, and applications to wireless communication systems and quantum information theory.

One of the most relevant application of Free Probability is in Random Matrix Theory. Dan Voiculescu showed that independent Guassian matrices of dimension large enough behave as free random variables in a non-commutative probability space. With this connection, Free Probability became a powerful tool to deal problems in random matrices. Since then, it has been studied the asymptotic freeness of several types of random matrices and so, for instance, computing the spectral distribution of limit of sums and products of such matrices.

Even with the success of Free Probability in Random Matrix Theory, there are situations in that the machinery of Free Probability is not enough; its scope is limited and it does not fit correctly and completely in some types of random matrices. Starting from this, variations and extensions of Free Probability have arisen from both theoretical and applied motivations.

One of the settings in which usual Free Probability is not very suitable is when we are considering random matrices having limiting joint distribution with respect non-normalized trace. In particular, this implies that the matrices converge to the random variable zero with respect the normalized trace. Then a different method than usual in Free Probability is required to deal with this kind of random matrices.

An extension of Free Probability known as *Type B Free Probability* appeared in the work [5] of Biane, Goodman and Nica, from a combinatorial motivation of finding an analogous theory to Free Probability where non-crossing partitions are replaced by non-crossing partitions asociated to Coxeter groups of type B. Thereafter, Belinschi and Shlyakhtenko presented in [**?**] an analytic interpretation of the free convolution of type B. In their work, they found out that this convolution can be expressed in terms of some kind of derivative of distributions, leading them to define the *Infinitesimal Free Probability* as a better way to understand Type B Free Probability. With this new notion, an study of infinitesimal free cumulants was presented by Fevrier and Nica in [11]. However, a remaining open question was about finding random matrices models in which Infinitesimal Free Probability could be applied. The first who gave an answer to the above question was Shlyakhtenko in [17]. In his paper, he proved that rotationally invariant random matrices and finite-rank deterministic matrices are asymptotically infinitesimally free and studied the outliers of sums of both types of matrices.

Motivated by the Shlyakhtenko's results, Collins, Hasebe and Sakuma studied in [8] the spectrum of random matrices which are obtained as some selfadjoint polynomials in random matrices with discrete spectrum and rotationally invariant random matrices having a limiting non-commutative distribution, by defining a notion of independence called *Cyclic Monotone Independence*. The objective of this master's thesis is presenting the notion of cyclic monotone independence along with its applications to random matrices, being [8] the basis of this manuscript.

The organization of this work is the following. Aiming to make this text more accessible to read for non-familiarized people in free probability, we present some preliminaries of functional analysis and free probability. More precisely, Chapter 1 exposes some properties of compact operators in order to establish the framework that we shall be using. The final section of the chapter is dedicated to present a notion of convergence of eigenvalues of compact operators which is characterized by convergence of traces. This notion fits in the context of convergence in distribution of non-commutative probability spaces except that the trace of an operator is not unital. This will be fundamental since the random matrices that we shall consider converge to compact operators.

Chapter 2 presents the basics of Free Probability Theory such as the definition of non-commutative probability space, random variables and free independence. We also include a brief presentation on free cumulants with an application in the proof of the Free Central Limit Theorem. Chapter 3 shows the most important relations between Free Probability and Random Matrices. We give some examples of random matrices ensembles and state the celebrated Wigner's semicircle law along a sketch of its proof. We also discuss some of the most important examples of asymptotic freeness of random matrices and some ideas about their proofs.

Chapters 4 and 5 correspond to the work of Collins, Hasebe and Sakuma about Cyclic Monotone Independence and Random Matrix Theory. In Chapter 4, we expose the abstract notion of Cyclic Monotone Independence and compute the eigenvalues of certain polynomials in cyclically monotone independent random variables which are trace class operators. The main contribution of this manuscript is presented in this chapter by providing an alternative proof to some formulas in [8] considering matrices of cyclically monotone independent elements. This allows to show new formulas for another polynomials, which are described in Proposition 4.2.2, Proposition 4.2.5, and Corollary 4.2.8.

Finally in the Chapter 5, we present the fact that random matrices with discrete spectrum and random matrices which are rotationally invariant are asymptotically cyclically monotone independent if they are independent. This is shown in the context of averaged convergence and almost sure convergence, being the Weingarten formula the main tool to achieve this. We also present some generalizations and computations of limiting eigenvalues of random matrices provided by the formulas obtained in Chapter 4 for some polynomials. Numerical simulations are exposed in order to show graphically the agreement of the eigenvalues and the approximation given by the cyclic monotone independence.

Chapter 1

Spectrum of Compact Operators

This chapter is dedicated to present the tools of functional analysis that will be used in this text. In the first section we collect some basic results about compact operators, being the spectral theorem the most important of them. In the Section 2 we present the Hilbert-Schmidt operators and trace class operators, along with some results associated to them. In the Section 3 we expose the notion of convergence in eigenvalues introduced in [8] and a moment method which allows to prove convergence in eigenvalues.

1.1 Preliminaries on Compact Operators

The purpose of this section is to establish the framework of functional analysis that we will need in order to understand the presented results in this text. We give the notion of spectrum of an operator, the definition of compact operator, results about compact operators and the spectral theorem. All the results presented here can be found in standard textbooks such as [10]. Throughout this section, X and Y will denote Hilbert spaces.

The class of objects which we consider in this work will be limit, in some sense, of matrices. We recall that if X is a finite dimensional Hilbert space and $T : X \to X$ is linear operator, then T is $1-1$ if only if T is invertible. However, this is not necessarily the case when X is infinite dimensional. Because of this, when we are considering linear operators on infinite dimensional spaces, it is convenient to consider a larger set than the eigenvalues set of X.

Definition 1.1.1. Let X be a Hilbert space, I be the identity operator on X and T be an linear operator on X.

1. We define the *resolvent set* of T by

$$\rho(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ is invertible}\}.$$

2. We define the *spectrum* of T by

$$\sigma(T) = \{\lambda \in \mathbb{C} \,:\, T - \lambda I \text{ is not invertible}\}.$$

Furthermore, the spectrum of an operator can be partitioned as follows.

- The *point spectrum* $\sigma_p(T)$ of T is formed by the elements $\lambda \in \mathbb{C}$ such that $T - \lambda I$ is not $1-1$, i.e., the eigenvalues of T.

- The *continuous spectrum* $\sigma_c(T)$ of T is formed by the elements $\lambda \in \sigma(T)$ such that $T - \lambda I$ is $1-1$ and $R(T - \lambda I)$ is dense, where $R(S)$ denotes the range of the linear operator S.

- The *residual spectrum* $\sigma_r(T)$ of T is formed by the elements $\lambda \in \mathbb{C}$ such that $T - \lambda I$ is $1-1$ and $R(T - \lambda I) \subset X$ is not dense.

In the framework of Hilbert spaces, we can define the *adjoint* of a linear operator $T : X \to Y$ which is determined by the condition

$$\langle Tx, y\rangle = \langle x, T^*y\rangle, \quad \forall\, x \in X, y \in Y. \tag{1.1}$$

We are interested in *selfadjoint operators*, i.e., linear operators $T : X \to X$ such that $T = T^*$. In particular, we have the following property about the spectrum of a selfadjoint operator.

Proposition 1.1.2. *Let $T : X \to X$ be a selfadjoint operator. Then*

1. *the operator norm satisfies that* $\|T\| = \sup\{|\lambda| \,:\, \lambda \in \sigma(T)\}$*,*

2. $\sigma(T) \subset \mathbb{R}$*.*

The other type of linear operators which we are interested in are the compact operators. This class of linear operators behave in some way as linear operators between finite dimensional spaces. A concrete definition can be written in the following way.

Definition 1.1.3. Let $T : X \to Y$ a linear operator. We say that T is a *compact operator* if the closure of $TB_X = \{Tx \,:\, \|x\|_X \leq 1\} \subset Y$ is compact.

The next lemma give us a characterization of compact operators.

Lemma 1.1.4. *Let $T : X \to Y$ be a linear operator. Then T is compact if only if for any $\{x_n\} \subset X$ bounded sequence we have that $\{Tx_n\} \subset Y$ has a convergent subsequence.*

Readily from the definition, we have that if $T : X \to Y$ is a compact operator, then T is bounded and hence continuous. Indeed, if T is compact then $\overline{TB_X}$ is compact and then TB_X is bounded. As a consequence, the operator norm of T is finite since $\|T\| = \sup_{\|x\|\leq 1} \|Tx\| < \infty$. Even more, we have the following proposition.

Proposition 1.1.5. *The set of compact operators $T : X \to X$, $K(X)$, is a closed subspace of $B(X) = \{T : X \to X : T$ is a bounded linear operator$\}$. Moreover, $K(X)$ is an ideal of $B(X)$, i.e., if $T \in K(X)$ and $S \in B(X)$, then $ST, TS \in K(X)$.*

Example 1.1.6. We say that a linear operator $T : X \to Y$ is a *finite-rank operator* if $\dim(R(T)) < \infty$. We see that if T is a finite-rank operator, then T is compact. Actually, we have that $\overline{TB_X}$ is bounded and closed in $R(T)$ which is finite dimensional. Using the Heine-Borel Theorem, we conclude that $\overline{TB_X}$ is compact.

The converse of the above example is not true in general. However, as we mentioned at the beginning of this section, compact operators behave in a similar way that finite-rank operators. The precise statement of this relation is described in the next result.

Theorem 1.1.7. *Let $T : X \to Y$ be a compact operator. Then, given any $\epsilon > 0$, there exist a finite-dimensional subspace M in $R(T)$ such that, for any $x \in X$,*

$$\inf_{m\in M} \|Tx - m\| \leq \epsilon \|x\|$$

Using the last example and Proposition 1.1.5, we have that the limit (in the operator norm) of finite-rank operators is a compact operator. On the other hand, by Theorem 1.1.7 we also have that the range of compact operators can be approximated by a finite-dimensional subspace. This approximation property implies that every compact operator can be approximated by finite-rank operators.

Theorem 1.1.8. *Let $T : X \to Y$ be a compact operator. Then T is the limit in the operator norm of a sequence of finite-rank operators.*

Now suppose that a compact operator T is invertible in $B(X)$. Then there exists a bounded linear operator S such that $ST = I = TS$ and so I is compact and the closed unit ball in X is also compact. However, this is not possible when X is infinite dimensional. We conclude that if X is infinite dimensional, then every compact operator T is not invertible, and hence $0 \in \sigma(T)$.

The spectrum of compact operators exhibits some interesting properties. Two of these are stated in the next theorem.

Theorem 1.1.9. *Let $T : X \to X$ be a compact operator.*

1. *(Fredholm's alternative) If $\lambda \neq 0$, then $\lambda \in \sigma_p(K)$ or $\lambda \in \rho(K)$, i.e., except possibly by $\lambda = 0$, the spectrum of T consists only of eigenvalues. Moreover, if $\lambda \in \sigma_p(K)$, $\lambda \neq 0$, then its algebraic multiplicity is finite.*

2. *For every $k \in \mathbb{N}$, the set $\{\lambda \in \sigma(T) : |\lambda| \geq \frac{1}{k}\}$ is finite. Then $\sigma(T)$ is countable and its only possible accumulation point is $\lambda = 0$.*

Finally, the following theorem states that compact selfadjoint operators have a spectral decomposition similar to symmetric matrices.

Theorem 1.1.10 (Spectral Theorem for compact selfadjoint operators). *Let $T : X \to X$ be a compact selfadjoint operator. Then X has an orthonormal basis $\{e_n\}_{n=1}^{\infty}$ of eigenvectors of T corresponding to eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$. Moreover, we have that*

1. *The eigenvalues λ_i are real and 0 is the only possible accumulation point.*

2. *The eigenspaces corresponding to distinct eigenvalues are orthogonal.*

3. *The eigenspaces corresponding to non-zero eigenvalues are finite-dimensional.*

Remark 1.1.11. Let $T : X \to X$ be a compact selfadjoint operator. The above theorem establishes that the eigenvalues of T can be ordered as

$$|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \cdots .$$

Furthermore, the previous theorem says that

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n, \qquad \forall\, x \in X. \tag{1.2}$$

If T has finitely many non-zero eigenvalues, then T is a finite-rank operator. Otherwise, it is possible to show that $\lambda_n \to 0$ as $n \to \infty$. The reciprocal is also true. Indeedm assume that T is an operator on a separable Hilbert space X with a countably infinite orthonormal basis $\{e_n\}_{n=1}^{\infty}$ consisting of eigenvectors of T. Moreover, assume that the eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$ corresponding to the eigenvectors $\{e_n\}_{n=1}^{\infty}$ satisfy that $\lambda_n \to 0$ as $n \to \infty$. Then we have that T is a compact operator. This can be seen by considering the sequence of finite-rank operators $\{T_n\}_{n=1}^{\infty}$, where each T_n is defined as $T_n x = \sum_{m=1}^{n} \lambda_m \langle x, e_m \rangle e_m$, for any $x \in X$. Since this sequence converges to T and each T_n is finite-rank, then T_n is compact and its limit is compact. This shows that T is compact.

1.2 Hilbert-Schmidt and Trace Class Operators

For the purpose of this work, we are interested in operators for which a notion of trace can be defined even in the case of infinite dimensional space. Obviously, a general notion of trace should include the well known case of finite dimensional spaces. Before of considering this class of operators, it is convenient to define a larger class of operators.

Definition 1.2.1. Let X be a Hilbert space and $\{e_i\}_{i\in I}$ be an orthonormal basis of X. We say that $T \in B(X)$ is a *Hilbert-Schmidt operator* if

$$\sum_{i\in I} \|Te_i\|^2 < \infty. \tag{1.3}$$

A natural question arises when we want to drop out the choice of an specific orthonormal basis of X. The following results says is that the value of the series in (1.3) does not depend of the basis.

Theorem 1.2.2. *Let X be a Hilbert space. If $\{e_\alpha\}_{\alpha\in I}$ and $\{f_\alpha\}_{\alpha\in J}$ are orthonormal bases for X and $T \in B(H)$ then*

$$\sum_{\alpha\in I} \|Te_\alpha\|^2 = \sum_{\alpha\in J} \|T^* f_\alpha\|^2 = \sum_{\alpha\in J} \|Tf_\alpha\|^2. \tag{1.4}$$

We denote by $S^2(X)$ the set of Hilbert-Schmidt operators in $B(X)$. If $\{e_i\}_{i\in I}$ is an orthonormal basis of X and $T \in B(X)$, we define the *Hilbert-Schmidt norm of T* as

$$\|T\|_2 = \left(\sum_{i\in I} \|Te_i\|^2\right)^{\frac{1}{2}}. \tag{1.5}$$

Then we have that $T \in S^2(X)$ if $T \in B(X)$ and $\|T\|_2 < \infty$. It can be proved that $S^2(X)$ is a vector space, an $*$-ideal, and $\|\cdot\|_2$ is indeed a norm on $S^2(X)$. Moreover, the topology on $S^2(X)$ induced by the Hilbert-Schmidt norm is finer than the subspace topology given by the operator norm on $B(X)$. The later fact is deduced from the next result.

Theorem 1.2.3. *Let X be a Hilbert space. If $T \in S^2(X)$ then $\|T\| \leq \|T\|_2$.*

Remark 1.2.4. A first example of Hilbert-Schmidt operator is given by finite-rank operators. Actually, we have that the space of finite-rank operators on X is a dense subset of $S^2(X)$ with the norm topology given by $\|\cdot\|_2$. But by Theorem 1.2.3, if $\{T_n\}_n$ is a sequence of finite rank operators which converges to a Hilbert-Schmidt operator with respect to $\|\cdot\|_2$, then it does with respect to the operator norm. So, using Theorem 1.1.8, we conclude that if $T \in S^2(X)$, then T is a compact operator.

Using Theorem 1.2.2, if $\{e_\alpha\}_{\alpha\in I}$ and $f_{\alpha\alpha\in J}$ are orthonormal bases for a Hilbert space X and $T \in B(X)$, then we have that

$$\sum_{\alpha\in I}\langle |T|e_\alpha, e_\alpha\rangle = \sum_{\alpha\in J}\langle |T|f_\alpha, f_\alpha\rangle,$$

where $|T| = \sqrt{T^*T}$. This motivate us to give the following definition.

Definition 1.2.5. Let X be a Hilbert space and $\{e_i\}_{i\in I}$ be an orthonormal basis for X. We say that $T \in B(X)$ is *trace class operator* if

$$\sum_{i\in I}\langle |T|e_i, e_i\rangle < \infty.$$

We denote by $S^1(X)$ the set of trace class operators on X. In an analogous way that Hilbert-Schmidt operators, we define

$$\|T\|_1 = \sum_{i\in I}\langle |T|e_i, e_i\rangle.$$

Then $T \in S^1(X)$ if $T \in B(X)$ and $\|T\|_1 < \infty$.

There is a clear relation between trace class operators and Hilbert-Schmidt operators since $\|T\|_1 = \||T|^{1/2}\|_2^2$. Moreover, there is a deeper relationship which is established in the following characterizations of trace class operators. A proof appears in [10] (Proposition 18.8).

Theorem 1.2.6. *Let X be a Hilbert space and $T \in B(X)$. The following statements are equivalent.*

1. *$T \in S^1(X)$.*
2. *$|T|^{1/2} \in S^2(X)$.*
3. *T is product of two elements of $S^2(X)$.*

Remark 1.2.7. From the last theorem, we have that if T is product of two elements of $S^2(X)$, then $T \in S^2(X)$ since $S^2(X)$ is an ideal. In particular we have that $S^1(X) \subset S^2(X)$ and so any trace class operator is compact.

The following result is very useful since it motives the definition of traces in the case of trace class operators. The proof of this result can be also found in Chapter 3 of [10] (Proposition 18.9).

Theorem 1.2.8. *Let X be a Hilbert space. If $T \in S^1(X)$, and $\{e_i\}_{i\in I}$ is an orthonormal basis for X then*

$$\sum_{i\in I} |\langle Te_i, e_i\rangle| < \infty,$$

and if $\{f_j\}_{j\in J}$ is an orthonormal basis for X then

$$\sum_{i\in I}\langle Te_i, e_i\rangle = \sum_{j\in J}\langle Tf_j, f_j\rangle.$$

With the above result, we are able to define the trace of a trace class operator.

Definition 1.2.9. Let X be a Hilbert space and $\{e_i\}_{i\in i}$ be a orthonormal basis for X. If $T \in S^1(X)$, we define the *trace of* T, $\mathrm{Tr}(T)$ as

$$\mathrm{Tr}(T) = \sum_{i\in I}\langle Te_i, e_i\rangle.$$

Remark 1.2.10.

1. We can note that in the case that X is finite dimensional and T is seen as an element of $M_n(\mathbb{C})$, then $\mathrm{Tr}(T)$ corresponds to the usual trace of a matrix which is the sum of the diagonal entries of T.
2. As we expect, the trace can be thought as a function $\mathrm{Tr} : S^1(X) \to \mathbb{C}$ and it satisfies to be a positive linear functional.

As the space $S^2(X)$, the set trace class operators $S^1(X)$ is a complete normed vector space with the norm $\|\cdot\|_1$. On the other hand, it is easy to see that a finite-rank operator is trace class. It can be then show that the space of finite-rank operators is a dense subset of $S^1(X)$ with respect the trace norm.

The next theorem collects some of the basic properties of the trace and the trace norm. Some of them are the same that we have in finite matrices .

Theorem 1.2.11. *Let X be a Hilbert space and $T \in S^1(X)$.*

1. *$T^* \in S^1(X)$ and $\mathrm{Tr}(T^*) = \overline{\mathrm{Tr}(T)}$.*
2. *If $A \in B(X)$ then $AT, TA \in S^1(X)$, $\mathrm{Tr}(AT) = \mathrm{Tr}(TA)$, and $|\mathrm{Tr}(AT)| \leq \|A\|\|T\|_1$.*
3. *$\|T^*\|_1 = \|T\|_1$.*
4. *If $A \in B(X)$ then $\|TA\|_1 = \|T\|_1\|A\|$.*
5. *$\|T\| \leq \|T\|_1$.*

Finally, we present a generalization of the Hilbert-Schmidt and trace class operators called the Schatten class operators.

Definition 1.2.12. Let $1 \leq p < \infty$. An operator $T \in B(X)$ is called a *Schatten operator of class p* if $\|T\|_p := \||T|^p\|_1^{1/p} < \infty$.

We shall denote the set of Schatten operators of class p as $S^p(X)$. This is a Banach space with respect to the norm $\|\cdot\|_p$. For the values $p = 1, 2$ we get the trace class and Hilbert-Schmidt operators, respectively.

The next proposition summarizes some of results related with Schatten class operators.

Proposition 1.2.13. *Let X be a Hilbert space.*

1. *$S^p(X)$ is an $*$-ideal of $B(X)$.*

2. *$S^p(X)$ is a complete normed vector space with respect to $\|\cdot\|_p$.*

3. *If $p \leq q$ then $S^p(X) \subset S^q(X)$ and $\|T\| \leq \|T\|_q \leq \|T\|_p$ for any $T \in S^p(X)$.*

4. *Hölder inequality holds for $\|\cdot\|_p$: if $0 \leq p, q, r < \infty$ and $\frac{1}{p} + \frac{1}{p} = \frac{1}{r}$, $A \in S^p(X)$, $B \in S^q(X)$, then $AB \in S^r(X)$, and $\|AB\|_r \leq \|A\|_p \|B\|_q$.*

According to the above properties, the spaces $S^p(X)$ can be considered as non-commutative analogous spaces to the $L^p(\Omega, \mu)$ spaces.

Finally, as in the cases of trace class and Hilbert-Schmidt operators, it is possible to show that the Schatten operators of class p are compact operators. Hence, if $T \in S^p(X)$ is selfadjoint, by applying the spectral theorem (Theorem 1.1.10) we have that there exists an ortonormal basis $\{e_n\}_{n=1}^{\infty}$ of X consisting of the eigenvectors of T and if λ_n is the eigenvalue associated to e_n, then Equation (1.2) holds. Since $T \in S^p(X)$ then $|T|^p \in S^1(X)$ and computing the trace using the basis $\{e_n\}_n$ we get

$$\|T\|_p^p = \mathrm{Tr}(|T|^p) = \sum_{n=1}^{\infty} \langle |T|^p e_n, e_n \rangle = \sum_{n=1}^{\infty} |\lambda_n|^p,$$

where we use that T is selfadjoint (and then the eigenvalues of $|T|$ are $\{|l\lambda_n|\}_{n=1}^{\infty}$). We can state this result in the following proposition.

Proposition 1.2.14. *Let X be a Hilbert space and $T \in B(X)$ be a compact operator with eigenvalues $\{\lambda_n\}_{n=1}^{\infty}$.*

1. *If $T \in S^p(X)$, then*

$$\|T\|_p = \left(\sum_{n=1}^{\infty} |\lambda_n|^p \right)^{1/p}. \tag{1.6}$$

2. *(Lidskii) If $T \in S^1(X)$, it is satisfied that*

$$\mathrm{Tr}(T) = \sum_{n=1}^{\infty} \lambda_n. \tag{1.7}$$

1.3 Convergence of Eigenvalues of Compact Selfadjoint Operators

In this section, based in Section 2 of [8], we present a notion of convergence of eigenvalues of compact selfadjoint operators. Then we will study a moment method which characterizes the convergence of eigenvalues.

From the above sections, we know that compact selfadjoint operators have discrete spectrum contained in $\mathbb{R}$. In order to define a convergence of eigenvalues of compact operators, it will be necessary to introduce a convenient order of their eigenvalues.

Definition 1.3.1. Let $\{x_i\}_{i=1}^{\infty} \subset \mathbb{R}$ a sequence converging to 0. We will say that $\{x_i\}_{i=1}^{\infty}$ is *properly arranged* if $|x_i| \geq |x_{i+1}|$, $\forall\, i \in \mathbb{N}$.

A proper arrangement of a sequence is not necessarily unique. For this, it is convenient to split the sequence in its non-negative and non-positive part, and then consider the proper arrangements of these two sequences which will be unique.

If a is a compact selfadjoint operator in a separable Hilbert space, we will denote by $\mathrm{EV}(a)$ the multiset of eigenvalues $\lambda_i(a)$ of a. We will assume that the sequence of eigenvalues $\{\lambda_i(a)\}_{i\geq 1}$ is properly arranged. In the same way, we split the sequence into its non-negative and the non-positive part, $\{\lambda_i^+(a)\}_{i\geq 1}$ and $\{\lambda_i^-(a)\}_{i\geq 1}$ respectively, and we will assume that both sequences are properly arranged. The notion of convergence of eigenvalues that we shall use is the following.

Definition 1.3.2. Let a and a_k, $k \in \mathbb{N}$ be compact selfadjoint operators in separable Hilbert spaces H and H_k, respectively. We will say that a_k *converges to a in eigenvalues* if, for all $i \geq 1$,

$$\begin{aligned} \lim_{k\to\infty} \lambda_i^+(a_k) &= \lambda_i^+(a), \\ \lim_{k\to\infty} \lambda_i^-(a_k) &= \lambda_i^-(a). \end{aligned}$$

In case that some sequence of eigenvalues is finite, we will add infinitely many zeros at the end of the sequence. Convergence of eigenvalues will be denoted by $\lim\limits_{k\to\infty} \mathrm{EV}(a_k) = \mathrm{EV}(a)$.

The next result gives us a characterization of the convergence in eigenvalues, which will be useful to establish a kind of moment method respect the trace on Hilbert spaces.

Proposition 1.3.3. *Let a and a_k, $k \in \mathbb{N}$ be compact selfadjoint operators in separable Hilbert spaces H and H_k, respectively. The next statements are equivalent:*

1. *a_k converges to a in eigenvalues.*
2. $\lim\limits_{k\to\infty} \mathrm{Tr}_{H_k}(f(a_k)) = \mathrm{Tr}_H(f(a))$, $\forall\, f \in C_{0,b}(\mathbb{R})$.
3. $\lim\limits_{k\to\infty} \mathrm{Tr}_{H_k}(f(a_k)) = \mathrm{Tr}_H(f(a))$, $\forall\, f \in C^\infty_{0,b}(\mathbb{R})$.

Here $C_{0,b}(\mathbb{R})$ denotes the set of real-valued bounded continuous functions on $\mathbb{R}$ that vanish in a neighborhood of 0, and $C^\infty_{0,b}(\mathbb{R})$ is the set of functions in $C_{0,b}(\mathbb{R})$ that are infinitely many times differentiable.

Proof. We show (1) $\Rightarrow$ (2) and (3) $\Rightarrow$ (1). The implication (2) $\Rightarrow$ (3) is trivial. In the proof, we denote $\lambda_i^\pm(k) := \lambda_i^\pm(a_k)$ for any $i \geq 1$ and $k \in \mathbb{N}$, where we are considering that the non-negative and non-positive parts of the eigenvalues are properly arranged, i.e., $\lambda_1^+(k) \geq \lambda_2^+(k) \geq \cdots \geq 0$ and $\lambda_1^-(k) \leq \lambda_2^-(k) \leq \cdots \leq 0$. In a similar way, we denote $\lambda_i = \lambda_i(a)$ for any $i \geq 1$ and we also consider the properly arranged sequences of non-negative and non-positive eigenvalues.

((1) $\Rightarrow$ (2)) Let $f \in C_{0,b}(\mathbb{R})$. Then there exists $\delta > 0$ such that $f(x) = 0$ for any $x \in (-\delta, \delta)$. By Theorem 1.1.9, there exist $m, \ell \in \mathbb{N}$ such that $0 \leq \lambda^+_{\ell+1} < \delta \leq \lambda^+_\ell$ and $\lambda^-_m \leq -\delta < \lambda^-_{m+1} \leq 0$. Since the sequences of eigenvalues are properly arranged, we have that $0 \leq \lambda^+_i < \delta$ and $-\delta < \lambda^-_j \leq 0$, for any $i > \ell$ and $j > m$. On the other hand, we have that a_k converges to a in eigenvalues. In particular $\lambda^+_{\ell+1}(k) \to \lambda^+_{\ell+1}$ as $k \to \infty$. Since $\lambda^+_{\ell+1} < \delta$, then there exists $k_0^+ \in \mathbb{N}$ such that $\lambda^+_{\ell+1}(k) < \delta$ for any $k \geq k_0^+$. Because of $\{\lambda_i^+(k)\}_{i\geq 1}$ is a non-increasing sequence for any $k \in \mathbb{N}$, we get that $\lambda_i^+(k) < \delta$ for any $k \geq k_0^+$ and $i \geq \ell + 1$. In a similar way, we have that there exists $k_0^- \in \mathbb{N}$ such that $\lambda_i^-(k) > -\delta$ for any $k \geq k_0^-$ and $i \geq m + 1$.

Using the assumption of convergence in eigenvalues and that f is continuous and vanishes in $(-\delta, \delta)$ we have that

$$\begin{aligned}
\lim_{k\to\infty} \mathrm{Tr}_{H_k}(f(a_k)) &= \lim_{k\to\infty} \sum_{i=1}^{\ell} f\left(\lambda_i^+(k)\right) + \lim_{k\to\infty} \sum_{i=1}^{m} f\left(\lambda_i^-(k)\right) \\
&= \sum_{i=1}^{\ell} f\left(\lambda_i^+\right) + \sum_{i=1}^{m} f\left(\lambda_i^-\right) = \mathrm{Tr}_H(f(a)),
\end{aligned}$$

that is what we wanted.

((3) $\Rightarrow$ (1)) We proceed by contradiction, by assuming that convergence in eigenvalues is not satisfied and giving a function $f \in C^\infty_{0,b}(\mathbb{R})$ for which the statement 3 of the proposition is not true. First, we note that the set of eigenvalues of the operators $\{a_k\}_{k\in\mathbb{N}}$ are uniformly bounded. Indeed, consider $f \in C^\infty_{0,b}(\mathbb{R})$ such that f is non-negative, $f(x) = 0$ if $|x| \leq \|a\|$, and $f(x) = 1$ if $|x| > 2\|a\|$. Since a is selfadjoint, all its eigenvalues are in $[-\|a\|, \|a\|]$, and then $\mathrm{Tr}_H(f(a)) = 0$. By assumption, $\mathrm{Tr}_{H_k}(f(a_k))$ converges to $\mathrm{Tr}_H(f(a))$ as $k \to \infty$. Because of f is positive, there exists $k_0 \in \mathbb{N}$ such that $\mathrm{Tr}_{H_k}(f(a_k)) < 1$ for $k \geq k_0$, and hence a_k does not have eigenvalues λ such that $|\lambda| > 2\|a\|$, for each $k \geq k_0$. This allows to conclude that all the eigenvalues of $\{a_k\}_{k\in\mathbb{N}}$ are contained in some interval $[-M, M]$.

We work first with the non-negative eigenvalues. The case of non-positive eigenvalues can be done in the same way. In order to get a contradiction, assume that $\lambda_1^+(k) \not\to \lambda_1^+$ as $k \to \infty$. Then, there exists an increasing sequence of natural numbers $\{k_j\}_{j\geq 1}$ and a real number $\mu_1^+ \in [0, M]$ such that $\lambda_1^+ \neq \mu_1^+$ and $\lambda_1(k_j) \to \mu_1^+$ as $j \to \infty$. Let $\epsilon = |\lambda_1^+ - \mu_1^+|$ which is greater than zero by assumption. Depending of which the sign of $\lambda_1^+ - \mu_1^+$ is, we take the following functions to get a contradiction.

If $\lambda_1^+ - \mu_1^+ < 0$, we take $f \in C^\infty_{0,b}(\mathbb{R})$ such that $f \geq 0$, $f(x) = 1$ if $x \geq \mu_1^+ - \epsilon/4$, and $f(x) = 0$ if $x \leq \lambda_1^+ + \epsilon/4$. Since λ_1^+ is the largest eigenvalue of a, we have that $\mathrm{Tr}_H(f(a)) = 0$. On the other hand, using the sequence $\{k_j\}_{j\in\mathbb{N}}$, we have that $\lambda_1^+(k_j) \geq \mu_1^+ - \epsilon/4$ for large enough j, and then $\mathrm{Tr}_{H_{k_j}}(f(a_{k_j})) \geq f(\lambda_1^+(k_j)) = 1$. Thus we get a contradiction to the fact that $\mathrm{Tr}_{H_k}(f(a_k))$ converges to $\mathrm{Tr}_H(f(a))$ as $k \to \infty$.

The other case is when $\mu_1^+ - \lambda_1^+ < 0$. Consider $f \in C^\infty_{0,b}(\mathbb{R})$ such that $f \geq 0$, $f(x) = 1$ if $x \geq \lambda_1^+ - \epsilon/4$, and $f(x) = 0$ if $x \leq \mu_1^+ + \epsilon/4$. In the same way that the above case, $\mathrm{Tr}_H(f(a)) \geq f(\lambda_1^+) = 1$ and $\mathrm{Tr}_{H_{k_j}}(f(a_{k_j})) = 0$ for large enough j. This again contradicts that $\mathrm{Tr}_{H_k}(f(a_k))$ converges to $\mathrm{Tr}_H(f(a))$ as $k \to \infty$.

We have to prove then that $\lambda_1^+(k)$ converges to λ_1^+ as $k \to \infty$. Now we use induction on i. Assume that the there exists $m \geq 2$ such that $\lambda_i^+(k) \to \lambda_i^+$ as $k \to \infty$, for any $i = 1, \ldots, m-1$. We shall prove that $\lambda_m^+(k) \to \lambda_m^+$ as $k \to \infty$. We proceed again by contradiction, in a similar fashion to the case of λ_1^+. Suppose that there exist μ_m^+ and an increasing sequence $\{k_j\}_{j\in\mathbb{N}}$ such that $\mu_m^+ \neq \lambda_m^+$ and $\lambda_m^+(k_j) \to \mu_m^+$ as $j \to \infty$. Notice that by using a similar argument to the proof of the eigenvalues of $\{a_k\}_{k\in\mathbb{N}}$ are uniformly bounded, we can take $\mu_m^+ \in [0, \lambda_{m-1}^+]$. Let $\epsilon = |\lambda_m^+ - \mu_m^+| > 0$. We have two cases.

If $\lambda_m^+ - \mu_m^+ < 0$, we take $f \in C^\infty_{0,b}(\mathbb{R})$ such that $f \geq 0$, $f(x) = 1$ if $x \geq \mu_m^+ - \epsilon/4$, and $f(x) = 0$ if $x \leq \lambda_m^+ + \epsilon/4$. Since $\lambda_i^+ \geq \mu_m^+ - \epsilon/4$ for $i = 1, \ldots, m-1$ we have that

$$\mathrm{Tr}_H(f(a)) = \sum_{i=1}^{m-1} f(\lambda_i^+) = m - 1.$$

On the other hand, for large enough j, we have that $\lambda_m^+(k_j) \geq \mu_m^+ - \epsilon/4$ and so are $\lambda_i^+(k_j) \geq \mu_m^+ - \epsilon/4$ for $i = 1, \ldots, m$. Hence $\mathrm{Tr}_{H_{k_j}}(f(a_{k_j})) \geq m$ for large enough j, and this contradicts the fact that $\mathrm{Tr}_{H_k}(f(a_k))$ converges to $\mathrm{Tr}_H(f(a))$ as $k \to \infty$.

Finally, if $\mu_m^+ - \lambda_m^+ < 0$, we consider $f \in C_{0,b}^{\infty}(\mathbb{R})$ such that $f \geq 0$, $f(x) = 1$ if $x \geq \lambda_m^+ - \epsilon/4$, and $f(x) = 0$ if $x \leq \mu_m^+ + \epsilon/4$. In the same way that the previous case, we have that $\mathrm{Tr}_H(f(a)) \geq m$ and $\mathrm{Tr}_{H_{k_j}}(f(a_{k_j})) = m - 1$ for large enough j, which lead us again to a contradiction.

Hence we have proved that $\lambda_m^+(k) \to \lambda_m^+$ as $k \to \infty$. Using mathematical induction, we have that $\lambda_i^+(k) \to \lambda_i^+$ as $k \to \infty$, for every $i \in \mathbb{N}$, that is what we wanted to prove. ∎

The above proposition motivates us to define convergence in distribution of compact operators. This notion can be considered as an extension of convergence in distribution in the framework of non-commutative probability spaces. But since we are considering compact selfadjoint operators instead of trace class operators, we have to deal with operators $f(a)$ which are trace class when a is a compact operator and $f \in C_{0,b}(\mathbb{R})$.

Definition 1.3.4. Let $a_i, a_i(n)$, $i = 1, \dots, k$, be compact selfadjoint operators in separable Hilbert spaces H, H_n, respectively. We say that $(a_1(n), \dots, a_k(n))$ *converges in compact distribution to* $(a_1, \dots, a_k)$ with respect to $\mathrm{Tr}_{H_n}, \mathrm{Tr}_H$ when $n \to \infty$ if for each function $f_i \in C_{0,b}(\mathbb{R})$, $i = 1, \dots, k$, $m \in \mathbb{N}$, $(i_1, \dots, i_m) \in \{1, \dots, k\}^m$, we have

$$\lim_{n\to\infty} \mathrm{Tr}_{H_n}\left(f_{i_1}(a_{i_1}(n)) \cdots f_{i_m}(a_{i_m}(n))\right) = \mathrm{Tr}_H\left(f_{i_1}(a_{i_1}) \cdots f_{i_m}(a_{i_m})\right). \tag{1.8}$$

With Proposition 1.3.3, we can prove an analogous of the moment method in the case of convergence of eigenvalues. First, we will need the next approximation lemma, where $\|\cdot\|_{[a,b]}$ denotes the supremum norm of continuous function in $[a, b]$.

Lemma 1.3.5. *Let $0 < \delta < \alpha$ and let $p \in \mathbb{N}$. If $f \in C^p(\mathbb{R})$ and $f(x) = 0$ for all $x \in [-\delta, \delta]$, then for all $\epsilon > 0$, there exists a polynomial P such that $0 = P(0) = p'(0) = \cdots = P^{(p-1)}(0)$, $\|f - P\|_{[-\delta,\delta]} < \epsilon$, and $|P(x)| \leq \epsilon |x|^p$, for all $x \in [-\delta, \delta]$.*

Proof. The lemma can be proved using Weierstrass' approximation theorem. ∎

The statement of the announced moment method respect the trace is the following.

Proposition 1.3.6. *Let a, a_k, $k \geq 1$ be selfadjoint operators on separable Hilbert spaces H, H_k, respectively. We assume that $a \in S^p(H)$ and $a_k \in S^p(H_k)$, for any $k \geq 1$, for some $p \in \mathbb{N}$. Suppose that $\mathrm{Tr}_{H_k}(a_k^n) \to \mathrm{Tr}_H(a^n)$ if $k \to \infty$, for each $n \geq p$. Then a_k converges to a in eigenvalues.*

Proof. We recall the space of p-Schatten class operators $S^p(H)$ is defined in Definition 1.2.12 as the space of operators a such that their p-Schatten norm $\|a\|_p$ is finite. Note that we can take p as even since $S^p(H) \subset S^{p+1}(H)$. In order to prove the proposition, we shall use the characterization of convergence in eigenvalues given in Proposition 1.3.3. Actually, it will be enough to prove that for all $f \in C^p_{0,b}(\mathbb{R})$ it holds that

$$\lim_{k\to\infty} \mathrm{Tr}_{H_k}(f(a_k)) = \mathrm{Tr}_H(f(a)).$$

Let $f \in C^p_{0,b}(\mathbb{R})$ and let $\delta > 0$ be such that $f(x) = 0$, for all $x \in [-\delta, \delta]$. We recall that we can consider a sequence of eigenvalues as an infinite sequence even if the Hilbert space has finite dimension. So, we denote as $\{\lambda_i\}_{i=1}^\infty$ and $\{\lambda_i(k)\}_{i=1}^\infty$ the properly arranged eigenvalues of a, a_k, $k \geq 1$, respectively.

Let $\alpha := \sup\limits_{k\geq 1} \|a_k\|_p$. Notice that because the assumption, we have that

$$\lim_{k\to\infty} \|a_k\|_p = \lim_{k\to\infty} \mathrm{Tr}_{H_k}(a_k^p)^{1/p} = \mathrm{Tr}_H(a^p)^{1/p} = \|a\|_p < \infty,$$

Then $\{\|a_k\|_p\}_{k\in\mathbb{N}}$ is a convergent sequence and so $\alpha < \infty$. We can also assume that $\alpha > 0$ since that in the other case we have $a = 0 = a_k$ for all $k \geq 1$. Observe that the multiset $\mathrm{EV}(a)$ can be seen as a positive measure on $\mathbb{R}$, where the measure of $\{\lambda\}$ is the dimension of the eigenspace of a with eigenvalue λ. So, we can use Chebyshev's inequality to get

$$|\{i \in \mathbb{N} : |\lambda_i(k)| > \delta\}| = |\{i \in \mathbb{N} : \lambda_i(k)^p > \delta^p\}| \leq \frac{\mathrm{Tr}(a_k^p)}{\delta^p} \leq \frac{\alpha^p}{\delta^p}. \tag{1.9}$$

In the same way, we obtain $|\{i \in \mathbb{N} : |\lambda_i| > \delta\}| \leq \frac{\alpha^p}{\delta^p}$. If we define $N = \lceil \frac{\alpha^p}{\delta^p} \rceil$ and from the fact that the eigenvalues are properly arranged, we have that $|\lambda_i(k)|, |\lambda_i| \leq \delta$, for any $k \geq 1$ and $i \geq N$. Using that $\|a\|_p^p, \|a_k\|_p^p < \alpha^p$, we can see that all the eigenvalues of a, a_k are contained in $[-\alpha, \alpha]$. Then, given $\epsilon > 0$, by Lemma 1.3.5, there exists a polynomial $P(x) = b_p x^p + b_{p+1}x^{p+1} + \cdots + b_q x^q$ such that $\|f - P\|_{[-\alpha,\alpha]} < \epsilon/N$ and $|P(x)| \leq \epsilon\alpha^{-p}x^p$ for $x \in [-\delta, \delta]$. The latter is because we can take $\delta > 0$ small enough such that $N \geq \alpha^p$. Separating the sum, using that f vanishes in $[-\delta, \delta]$ and triangle inequality, for all $k \geq 1$ we have that

$$\begin{aligned}
\left|\sum_{i=1}^\infty f(\lambda_i(k)) - \sum_{i=1}^\infty f(\lambda_i)\right| &= \left|\sum_{i=1}^N f(\lambda_i(k)) - \sum_{i=1}^N f(\lambda_i)\right| \\
&= \left|\sum_{i=1}^N f(\lambda_i(k)) - \sum_{i=1}^N f(\lambda_i)\right| + \left|\sum_{i=1}^N P(\lambda_i(k)) - \sum_{i=1}^N P(\lambda_i)\right| \\
&\quad + \left|\sum_{i=1}^N P(\lambda_i) - \sum_{i=1}^N f(\lambda_i)\right|
\end{aligned}$$

$$\leq \quad 2\epsilon + \left| \sum_{i=1}^{N} P(\lambda_i) - \sum_{i=1}^{N} f(\lambda_i) \right|. \tag{1.10}$$

Now, using the bound for P in $[-\delta, \delta]$ we obtain that

$$\sum_{i=N+1}^{\infty} |P(\lambda_i(k))| \leq \epsilon\alpha^{-p} \sum_{i=N+1}^{\infty} |\lambda_i(k)|^p \leq \epsilon\alpha^{-p} \|a_k\|_p^p \leq \epsilon, \tag{1.11}$$

and in the same way we get $\sum_{i=N+1}^{\infty} |P(\lambda_i)| \leq \epsilon$. Hence,

$$\begin{aligned}
\left| \sum_{i=1}^{N} P(\lambda_i(k)) - \sum_{i=1}^{N} P(\lambda_i) \right| \quad &= \quad \left| \sum_{i=1}^{N} P(\lambda_i(k)) + \sum_{i=N+1}^{\infty} P(\lambda_i(k)) - \sum_{i=N+1}^{\infty} P(\lambda_i(k)) \right. \\
&\quad \left. + \sum_{i=N+1}^{\infty} P(\lambda_i) - \sum_{i=N+1}^{\infty} P(\lambda_i) - \sum_{i=1}^{N} P(\lambda_i) \right| \\
&\leq \quad 2\epsilon + \left| \sum_{i=1}^{\infty} P(\lambda_i(k)) - \sum_{i=1}^{\infty} P(\lambda_i) \right| \qquad (1.12)
\end{aligned}$$

On the other hand, by assumption we have that

$$\sum_{i=1}^{\infty} \lambda_i(k)^n = \mathrm{Tr}_{H_k}(a_k^n) \xrightarrow[k\to\infty]{} \mathrm{Tr}_H(a^n) = \sum_{i=1}^{\infty} \lambda_i^n,$$

for any $n \geq p$. Moreover, the coefficients of the terms of degree less than p in P are equal to zero, then we can choose $k_0 \in \mathbb{N}$ such that

$$\left| \sum_{i=1}^{\infty} P(\lambda_i(k)) - \sum_{i=1}^{\infty} P(\lambda_i) \right| \leq \epsilon, \qquad \forall\, k \geq k_0. \tag{1.13}$$

Finally, we can conclude that

$$\left| \sum_{i=1}^{\infty} f(\lambda_i(k)) - \sum_{i=1}^{\infty} f(\lambda_i) \right| \leq 5\epsilon, \qquad \forall\, k \geq k_0 \tag{1.14}$$

using Equations (1.10), (1.12) and (1.13). ■

As a direct consequence of the above proposition, we have the following criterion to prove that the multisets of eigenvalues of two selfadjoint p-Schatten class operators are the same.

Corollary 1.3.7. *Let a, b selfadjoint operators on separable Hilbert spaces H, K, respec-*

tively, such that $a \in S^p(H)$ and $b \in S^p(H)$, for some $p \in \mathbb{N}$. If $\mathrm{Tr}_H(a^n) = \mathrm{Tr}_K(b^n)$ for all $n \geq p$, then $\mathrm{EV}(a) = \mathrm{EV}(b)$.

Proof. Consider the sequences $a_k = a$ and $b_k = b$, for each $k \geq 1$. Then we have that $\mathrm{Tr}_H(a_k^n) = \mathrm{Tr}_K(b_k^n)$, for any $n \geq p$. By Proposition 1.3.6, a_k converges in eigenvalues to b, and, in the same way, b_k converges in eigenvalues to a. Hence $\mathrm{EV}(a) = \mathrm{EV}(b)$. ■

We finish this section by noting that in Proposition 1.3.6, we assumed that there exists a limiting operator a. This assumption may be omitted as we shall see in the next proposition.

Proposition 1.3.8. *Let p be an even positive integer and H_k be a separable Hilbert space, for each $k \geq 1$. Let $a_k \in S^p(H_k)$ be a selfadjoint operator, for each $k \geq 1$. Suppose that $\alpha_n := \lim_{k\to\infty} \mathrm{Tr}_{H_k}(a_k^n) \in \mathbb{R}$ exists for all $n \geq p$. Then there exist a separable Hilbert space H and a selfadjoint operator $a \in S^p(H)$ such that a_k converges to a in eigenvalues and $\alpha_n = \mathrm{Tr}_H(a^n)$, for all $n > p$.*

Proof. As we noted in the proof of Proposition 1.3.6, we have that $\alpha = \sup_{k\in\mathbb{N}} \|a\|_p < \infty$ and then $|\lambda_i(k)| \in [0, \alpha]$, for any $k, i \in \mathbb{N}$. We construct a sequence of real numbers $\{\lambda_i\}_{i\in\mathbb{N}}$ as follows. As a bounded sequence of real numbers, we can take an increasing sequence $\{k_{1,j}\}_{j\geq 1} \subset \mathbb{N}$ such that $\lambda_1(k_{1,j})$ converges to some λ_1 as $j \to \infty$. Then, we take an increasing subsequence $\{k_{2,j}\}_{j\in\mathbb{N}} \subset \{k_{1,j}\}_{j\geq 1}$ such that $\lambda_2(k_{2,j})$ converges to some λ_2 as $j \to \infty$. We proceed inductively in order to find a sequence $\{k_{m,j}\}$ such that $\lambda_m(k_{m,j})$ converges to some λ_m as $j \to \infty$, for every $m \in \mathbb{N}$.

Now consider the subsequence $\{k_{j,j}\}_{j\geq i} \subset \{k_{i,j}\}_{j\in\mathbb{N}}$. Then $\lambda_i(k_{j,j}) \to \lambda_i$ as $j \to \infty$, for each $i \geq 1$. By construction, we have that $\{\lambda_i\}_{i\geq 1}$ is properly arranged. Since p is even, by Fatou's lemma we have that

$$\sum_{i=1}^{p} \lambda_i^p \leq \liminf_{j\to\infty} \sum_{i=1}^{\infty} \lambda_i(k_{j,j})^p = \liminf_{j\to\infty} \mathrm{Tr}_{H_{k_{j,j}}}\left(a_{k_{j,j}}^p\right) \leq \alpha^p.$$

Hence there exist a separable Hilbert space H and a selfadjoint operator $a \in S^p(H)$ such that $\mathrm{EV}(a) = \{\lambda_i\}_{i\in\mathbb{N}}$. We denote $\beta_n = \mathrm{Tr}_H(a^n)$ for $n \geq p$. Since $S^p(H) \subset S^n(H)$ for any $n \geq p$, we have that $\beta_n < \infty$ for any $n \geq p$. We have now to show that $\alpha_n = \beta_n$ for $n \geq p+1$. Let $\epsilon > 0$. By applying Chebyshev's inequality in the same way that we established (1.9), we have that there exists $m \in \mathbb{N}$ such that $|\lambda_i(k)| \leq \epsilon$ for all $k \in \mathbb{N}$ and $i > m$. We have also that $|\lambda_i| \leq \epsilon$ for $i > m$. Recalling that $\beta_n = \sum_{i=1}^{\infty} \lambda_i^n$, for $n \geq p+1$

we have that

$$\begin{aligned} |\alpha_n - \beta_n| &\leq |\alpha_n - \operatorname{Tr}_{H_k}(a_k^n)| + |\operatorname{Tr}_{H_k}(a_k^n) - \beta_n| \\ &\leq |\alpha_n - \operatorname{Tr}_{H_k}(a_k^n)| + \left|\sum_{i=1}^{m}(\lambda_i(k)^n - \lambda_i^n)\right| + \sum_{i=m+1}^{\infty}(|\lambda_i(k)|^n + |\lambda_i|^n) \qquad (1.15) \\ &= |\alpha_n - \operatorname{Tr}_{H_k}(a_k^n)| + \left|\sum_{i=1}^{m}(\lambda_i(k)^n - \lambda_i^n)\right| + \epsilon^{n-p}\left(\sum_{i=m+1}^{\infty}(\lambda_i(k)^p + \lambda_i^p)\right). \end{aligned}$$

Note that the left-hand side does not depend of k. So, taking in particular the sequence $\{k_{j,j}\}_{j\in\mathbb{N}}$, we have have that the first term of the right-hand side of (1.15) converges to zero by hypothesis. Also, the second term converges to zero by the construction of $\{k_{j,j}\}_{j\in\mathbb{N}}$. Finally, the last term is bounded by $2\alpha^p\epsilon^{n-p}$. We can conclude that $\alpha_n = \beta_n$ for each $n \geq p+1$. Notice that if we consider another sequence of limiting eigenvalues $\{\lambda_i'\}_{i\in\mathbb{N}}$ given by another choice of the sequences $\{k_{i,j}\}_{j\in\mathbb{N}}$, we would get that the $\beta_n' = \sum_{i=1}^{\infty}\lambda_i' = \alpha_n = \beta_n$, for $n \geq p+1$. By applying Corollary 1.3.7, we have that $\{\lambda_i'\}_{i\in\mathbb{N}} = \{\lambda_i\}_{i\in\mathbb{N}}$, and hence the limiting eigenvalues do not depend of which sequences $\{k_{i,j}\}_{j\in\mathbb{N}}$ are considered. So, we have that there exists a selfadjoint operator $a \in S^{p+1}(H)$ such that $\operatorname{Tr}_{H_k}(a_k^n) \to \operatorname{Tr}_H(a^n)$ as $k \to \infty$, for any $n \geq p+1$. We finish the proof by invoking Proposition 1.3.6 and therefore a_k converges in eigenvalues to a. ■

Chapter 2

Free Probability Theory

In this chapter, we present the basic definitions and some of the most important results of Free Probability Theory that we use throughout this work. First, we give the definition of non-commutative probability space, and we introduce the notions of random variables and distributions. In the second section, we talk about free independence of non-commutative random variables, which is an analogous concept of independence of classical random variables. In Section 3, we give the definition of free cumulants, state some related results and give an application to the Free Central Limit Theorem.

Mostly of the definition and results appearing here can found in the book of A. Nica and R. Speicher [14]. For an easy reading of this chapter of preliminaries, we omit most of the proofs of the results.

2.1 Non-commutative Probability Spaces

We introduce the notion of non-commutative probability. We recall what a probability space is in the classical sense. Based in concepts like random variables, moments and distributions, we shall find the motivation for their free counterparts.

We review some basics of probability theory.

Definition 2.1.1. We say that a tuple $(\Omega, \mathcal{F}, \mathbb{P})$ is a *classical probability space* if Ω is a set, $\mathcal{F}$ is a σ-algebra of Ω and $\mathbb{P}$ is a probability measure, i.e. $\mathbb{P}$ is a measure $\mathbb{P} : \mathcal{F} \to [0, 1]$ such that $\mathbb{P}(\Omega) = 1$.

It can be said that Ω is the set of possible results of a random experiment, $\mathcal{F}$ is the collection of events, and for any event $A \in \mathcal{F}$, $\mathbb{P}(A)$ is the probability that the event A occurs.

Definition 2.1.2. 1. Given a classical probability space $(\Omega, \mathcal{F}\mathbb{P})$, we say that a function $X : \Omega \to \mathbb{C}$ is a *classical random variable* if $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$, for any B borelian set of $\mathbb{C}$.

2. Given a classical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a classical random variable X, we define the *distribution* of the classical random variable X as the measure μ in $(\mathbb{C}, \mathcal{B}(\mathbb{C}))$ given by

$$\mu(A) = \mathbb{P}(X^{-1}(A)), \quad A \in \mathcal{B}(\mathbb{C}), \tag{2.1}$$

where $\mathcal{B}(\mathbb{C})$ is the Borel σ-algebra of $\mathbb{C}$.

It can be consider that a random variable X is a quantification of the result of an experiment, which means that $X(\omega)$ is assigned to the event $\omega \in \Omega$.

Definition 2.1.3. Given a random variable X in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, for any measurable function $f : \mathbb{C} \to \mathbb{C}$ we define

$$\mathbb{E}(f(X)) := \int_{\Omega} f(X(\omega))\, d\mathbb{P}(\omega) \tag{2.2}$$

provided that the integral exists.

In particular, if μ is the distribution of a *real* random variable X, we define $m_n(\mu) = \mathbb{E}(X^n)$ as the *n-th moment of X*. The latter definition can be extended for the case that μ is any measure by simply writing

$$m_n(\mu) = \int_{\mathbb{R}} t^n\, d\mu(t).$$

The sequence of moments of a measure is very useful since it has a lot of information about μ. Moreover, in several important cases the sequence of moments characterizes the measure.

We know that the sum and product of random variables form new random variables. This means that the set of random variables form an algebra which is commutative. For a while, we only consider random variables with all their moments. So, this algebra can be provided of a linear functional $\mathbb{E}$ which satisfies that $\mathbb{E}(1) = 1$ and $\mathbb{E}(X) \geq 0$ if $X \geq 0$. The notion of non-commutative probability space can be thought as a kind of generalization of the notion of random variable from a merely algebraic point of view.

Definition 2.1.4. We say that a pair $(\mathcal{A}, \varphi)$ is a *non-commutative probability space* if the following conditions are held:

1. $\mathcal{A}$ is a unital algebra over $\mathbb{C}$,

2. φ is a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ such that $\varphi(1_{\mathcal{A}}) = 1$.

An element $a \in \mathcal{A}$ is called *non-commutative random variable.*

An additional property that can be given to φ is that

$$\varphi(ab) = \varphi(ba), \quad \forall\, a, b \in \mathcal{A}. \tag{2.3}$$

In this case, we say that φ is a *trace*.

The previous definition can be extended to the case that $\mathcal{A}$ is an $*$-*algebra* which means that $\mathcal{A}$ has an antilineal operation $\mathcal{A} \ni a \mapsto a^* \in \mathcal{A}$ such that $(a^*)^* = a$ and $(ab)^* = b^*a^*$ for any $a, b \in \mathcal{A}$. Also, if we consider that φ is *positive*, which means that $\varphi(a^*a) \geq 0$ for any $a \in \mathcal{A}$, we say that the pair $(\mathcal{A}, \varphi)$ is a $*$-*probability space.* In this framework, we can distinguish some special random variables $a \in \mathcal{A}$, for instance

1. a is *normal* if $a^*a = aa^*$,
2. a is *selfadjoint* if $a = a^*$,
3. a is *unitary* if $a^*a = aa^* = 1_{\mathcal{A}}$.

In the same framework of $*$-probability spaces, given a family of random variables $\{a_i\}_{i\in I}$, we can define the *set of moments of* $\{a_i\}_{i=1}^{\infty}$ as

$$\left\{\varphi\left(a_{i(1)}^{\epsilon_1} a_{i(2)}^{\epsilon_2} \cdots a_{i(n)}^{\epsilon_n}\right) \;:\; n \geq 1\,,\, (i(1), \ldots, i(n)) \in I^n\,, (\epsilon_1, \ldots, \epsilon_n) \in \{1, *\}^n\right\}. \tag{2.4}$$

We also say that φ is *faithful* if $\varphi(a^*a) = 0$ implies that $a = 0$.

Some basic examples of non-commutative probability spaces are given in the following.

Example 2.1.5. 1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a classical probability space and let $\mathcal{A} = L^{\infty-}(\Omega, \mathbb{P})$ be the algebra of random variables with finite moments of any order. Indeed, $\mathcal{A}$ is a $*$-algebra since we can consider the $*$-operation as the conjugation of complex functions. On the other hand, the integration respect to $\mathbb{P}$, $\mathbb{E}$, is a linear functional on $\mathcal{A}$ such that $\mathbb{E}(1_{\mathcal{A}}) = 1$ since $\mathbb{P}$ is a probability measure. Then $(L^{\infty-}(\Omega, \mathbb{P}), \mathbb{E})$ is a $*$-probability space.

2. For $d \in \mathbb{N}$, we consider $M_d(\mathbb{C})$ the algebra of $d \times d$ complex matrices along with the usual multiplication of matrices. We also have a $*$-operation given by the conjugate transposition. So, $M_d(\mathbb{C})$ is a $*$-algebra. Moreover, if we define $\mathrm{tr} : M_d(\mathbb{C}) \to \mathbb{C}$ by

$$\mathrm{tr}(a) = \frac{1}{d} \sum_{i=1}^{d} a_{ii} \qquad \forall\ a = (a_{i,j})_{i,j=1}^{d} \in M_d(\mathbb{C}), \tag{2.5}$$

then $(M_d(\mathbb{C}), \mathrm{tr})$ is a $*$-probability space.

3. We can combine the latter two examples in the following one. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a classical probability space and consider $\mathcal{A} = L^{\infty-}(\Omega, \mathbb{P})$. For $d \geq 1$, we consider the $*$-algebra of random matrices $M_d(\mathcal{A})$ with the obvious $*$-operation. Then $(M_d(\mathcal{A}), \operatorname{tr} \circ \mathbb{E})$ is a $*$-probability space. Note that we have the algebra isomorphism $M_d(\mathcal{A}) \cong M_d(\mathbb{C}) \otimes \mathcal{A}$. Actually, it is possible to show that if $(\mathcal{A}, \varphi)$ and $(\mathcal{B}, \psi)$ are $*$-probability spaces, then $(\mathcal{A} \otimes \mathcal{B}, \varphi \otimes \psi)$ is again a $*$-probability space.

4. Let $(H, \langle \cdot, \cdot \rangle)$ be a Hilbert space and $B(H)$ the algebra of bounded linear operators on H. In this case $B(H)$ is a $*$-algebra where given $a \in B(H)$, the adjoint a^* is uniquely determined by the relation

$$\langle a\epsilon, \eta \rangle = \langle \epsilon, a^* \eta \rangle, \quad \forall\, \epsilon, \eta \in H. \tag{2.6}$$

We can define a linear functional on $B(H)$ as follows. We take a vector $\epsilon_0 \in H$ such that $\|\epsilon\| = 1$. If we consider the linear functional $\tau : B(H) \to \mathbb{C}$ defined by

$$\tau(a) := \langle a\epsilon_0, \epsilon_0 \rangle, \quad \forall\, a \in B(H), \tag{2.7}$$

then $(B(H), \tau)$ is a $*$-probability space. In general, τ does not hold to be faithful.

Another important concept in classical probability is that of the distribution of a random variable. It is desirable that an analogous concept of distribution in non-commutative probability contains all the information of the moments of a random variable. In order to give the corresponding definition, we denote as $\mathbb{C}\langle X, X^* \rangle$ as the unital algebra freely generated by two non-commutative indeterminate. Actually $\mathbb{C}\langle X, X^* \rangle$ is an $*$-algebra by defining $(X)^* = X^*$.

Definition 2.1.6. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and $a \in \mathcal{A}$ a random variable. The *$*$-distribution of a (in the algebraic sense)* is the linear functional $\mu : \mathbb{C}\langle X, X^* \rangle \to \mathbb{C}$ defined by

$$\mu\left(X^{\epsilon(1)} \cdots X^{\epsilon(k)}\right) = \varphi\left(a^{\epsilon(1)} \cdots a^{\epsilon(k)}\right), \tag{2.8}$$

for any $k \geq 0$ and $\epsilon(1), \ldots, \epsilon(k) \in \{1, *\}$.

This definition can be extended for the case of a family of random variables $\{a_1, \ldots, a_n\}$ just by considering the functional in the algebra of non-commutative polynomials on $2n$ indeterminates given by the mixed moments of $a_1, \ldots, a_n$. This functional is known as the *$*$-joint distribution of $a_1, \ldots, a_n$*.

In the special case of normal random variables, it is possible to give an alternative definition of distribution with an analytic flavor.

Definition 2.1.7. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and $a \in \mathcal{A}$ be a normal random variable. If there exists a probability measure μ with compact support on $\mathbb{C}$ such that

$$\int_{\mathbb{C}} z^k \bar{z}^\ell d\mu(z) = \varphi\left(a^k (a^*)^\ell\right), \quad \forall\, k, \ell \in \mathbb{N} \cup \{0\}, \tag{2.9}$$

we say that μ is the *$*$-distribution (in the analytic sense) of a.*

It is possible to show that if a is selfadjoint, then its possible distribution μ has support contained in $\mathbb{R}$. On the other hand, if a is a random variable, its analytic distribution does not necessarily exist. However, it exists for a good number of important examples. Furthermore, if we add extra analytic structure to the algebra $\mathcal{A}$, we can guarantee the existence of such analytic distribution. The appropriate framework to accomplish this is inside of C^*-algebra theory.

We recall that $(\mathcal{A}, \|\cdot\|)$ is a *C^*-algebra* if $\mathcal{A}$ is a $*$-algebra, $\|\cdot\|$ is a norm such that $(\mathcal{A}, \|\cdot\|)$ is a complete normed vector space, and $\|xx^*\| = \|x\|^2$, for any $x \in \mathcal{A}$. With this, we can define a *C^*-noncommutative probability space* to be a $*$-probability space $(\mathcal{A}, \varphi)$ such that $\mathcal{A}$ is a C^*-algebra.

In the framework of C^*-spaces, we have the next theorem whose prove can be found in Lecture 3 of [14].

Theorem 2.1.8. *Let $(\mathcal{A}, \varphi)$ be a C^*-noncommutative probability space and $a \in \mathcal{A}$ be a normal random variable. Then a has a distribution in the analytic sense.*

Remark 2.1.9. Let $(\mathcal{A}, \varphi)$ be a C^*-noncommutative probability space and $a \in \mathcal{A}$ be a selfadjoint random variable. From the previous theorem, we know that a has analytic distribution μ whose support is compact and contained in $\mathbb{R}$. A very well known result states that a compact supported measure is determined by its moments m_n. Hence, in the framework of C^*-spaces, the analytic distribution and the algebraic distribution have the same information and they determine each other.

2.2 Free Independence

We know present an analogous concept of independence in the framework of non-commutative probability spaces. This is the called *free independence*. It was introduced by Dan Voiculescu in [19] with the objective of solving problems in operator algebras. The adjective *free* is because this independence is close to free products of group. We recall that in classical independence, a family of random variables $\{X_i\}_{i \in I}$ in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ are *independent* if

$$\mathbb{P}\left(\bigcap_{j\in J}\{X_j \in A_j\}\right) = \prod_{j\in J}\mathbb{P}(X_j \in A_j), \qquad \forall\, J \subset I \text{ finite, } A_j \in \mathcal{B}(\mathbb{C}), \forall\, j \in J. \tag{2.10}$$

Then, if X and Y are independent random variables with all their moments, we have that

$$\mathbb{E}(X^n Y^m) = \mathbb{E}(X^n)\mathbb{E}(Y^m), \quad \forall\, m, n \geq 0. \tag{2.11}$$

From an algebraic point of view, independence of two random variables can be seen as a "recipe" for computing moments of $X^n Y^m$ from the moments of X^n and Y^m. Free independence is a recipe for computing mixed moments in a non-commutative framework.

Definition 2.2.1. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and I be a index set. A family of unital subalgebras $\{\mathcal{A}_i\}_{i\in I}$ of $\mathcal{A}$ is called *freely independent* if

$$\varphi(a_1 \cdots a_k) = 0$$

whenever $k \in \mathbb{N}$, $a_j \in \mathcal{A}_{i(j)}$ where $i(j) \in I$ for $j = 1, \ldots, k$, $\phi(a_j) = 0$ for any $j = 1, \ldots, k$, and $i(1) \neq i(2)$, $i(2) \neq i(3), \ldots,$ $i(k-1) \neq i(k)$.

We can define that a family of subsets of $\mathcal{A}$, $\{\mathcal{X}_i\}_{i\in I}$ is called *freely independent* if the family of unital subalgebras $\mathcal{A}_i = \mathrm{Alg}(\mathcal{X}_i)$ is freely independent, where $\mathrm{Alg}(\mathcal{X})$ denotes the unital subalgebra generated by $\mathcal{X}$. In particular, a set of random variables $\{a_i\}_{i\in I}$ is *freely independent* if the family $\mathcal{A}_i = \mathrm{Alg}(a_i)$ is freely independent. In the context of $*$-spaces, we say that the random variables $\{a_i\}_{i\in I}$ are $*$-*freely independent* if the family $\mathcal{A}_i = \mathrm{Alg}(a_i, a_i^*)$ is freely independent. Throughout this work, we shall abbreviate that random variables are freely independent just by saying that the random variables are *free*. In the same way, we refer free independence as *freeness*.

Now we shall see some examples that show how we can use freeness to compute mixed moments.

Example 2.2.2. Let $a, b \in \mathcal{A}$ be two free random variables in a non-commutative probability space $(\mathcal{A}, \varphi)$. In general, we do not have that $\varphi(a) = 0$, but by linearity $\varphi(a - \varphi(a)1_{\mathcal{A}}) = 0 = \varphi(b - \varphi(b)1_{\mathcal{A}})$. By freeness, we get that

$$\begin{aligned} 0 &= \varphi\left((a - \varphi(a)1_{\mathcal{A}})(b - \varphi(b)1_{\mathcal{A}})\right) \\ &= \varphi\left(ab - \varphi(a)b - \varphi(b)a + \varphi(a)\varphi(b)1_{\mathcal{A}}\right) \\ &= \varphi(ab) - \varphi(a)\varphi(b). \end{aligned}$$

It follows that $\varphi(ab) = \varphi(a)\varphi(b)$ whenever a, b are free. Notice that this coincides with the classical case: if X and Y are independent classical random variables, then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Example 2.2.3. Let $a_1, a_2, b_1, b_2 \in \mathcal{A}$ random variables in a non-commutative probability space $(\mathcal{A}, \varphi)$ such that $\{a_1, a_2\}$ and $\{b_1, b_2\}$ are free. Using the idea of the above example, we get that

$$\varphi\left((a_1 - \varphi(a_1)1_{\mathcal{A}})(b_1 - \varphi(b_1)1_{\mathcal{A}})(a_2 - \varphi(a_2)1_{\mathcal{A}})(b_2 - \varphi(b_2)1_{\mathcal{A}})\right) = 0$$

implies that

$$\varphi(a_1 b_1 a_2 b_2) = \varphi(a_1 a_2)\varphi(b_1)\varphi(b_2) + \varphi(a_1)\varphi(a_2)\varphi(b_1 b_2) - \varphi(a_1)\varphi(a_2)\varphi(b_1)\varphi(b_2).$$

If $b_1 = b$ and $b_2 = 1$ then

$$\varphi(a_1 b a_2) = \varphi(a_1 a_2)\varphi(b)\varphi(1_{\mathcal{A}}) + \varphi(a_1)\varphi(a_2)\varphi(b \cdot 1_{\mathcal{A}}) - \varphi(a_1)\varphi(a_2)\varphi(b)\varphi(1_{\mathcal{A}}) = \varphi(a_1 a_2)\varphi(b),$$

which also coincides with the classical case. However, taking $a = a_1 = a_2$ and $b = b_1 = b_2$, if a and b are free then

$$\varphi(abab) = \varphi(a^2)\varphi(b)^2 + \varphi(a)^2\varphi(b^2) - \varphi(a)^2\varphi(b^2).$$

In this case the latter expression differs from the classical one: if X and Y are independent classical random variables, then $\mathbb{E}(XYXY) = \mathbb{E}(X^2)\mathbb{E}(Y^2)$. This shows that free independence is a different rule from classical independence for computing mixed moments.

For the next example, we need the following definition that, as we shall see in the next chapter, it is related with a very important class of random matrices.

Definition 2.2.4. Let $(\mathcal{A}, \varphi)$ a $*$-probability space and $u \in \mathcal{A}$. We say that u is a *Haar unitary* if $u^*u = 1_{\mathcal{A}} = uu^*$ and $\varphi(u^m) = \delta_{0,m}$ for $m \in \mathbb{Z}$.

Example 2.2.5. Let $a, b, u, v \in \mathcal{A}$ random variables in a non-commutative probability space $(\mathcal{A}, \varphi)$ such that $\{u, u^*\}, \{v, v^*\}$ and $\{a, b\}$ are free. We see that uau^* and vbv^* are free. According to the definition, we have to prove that

$$\varphi\left(p_1(uau^*)q_1(vbv^*)\cdots p_r(uau^*)q_r(vbv^*)\right) = 0 \tag{2.12}$$

whenever p_i, q_i are polynomials such that $\varphi(p_i(uau^*)) = 0 = \varphi(q_i(vbv^*))$ for each $1 \leq i \leq r$. Indeed, we have to consider the other three cases when the argument of φ in (2.12) starts with $q_1(vbv^*)$ and finishes with $p_r(uau^*)$ or $q_r(vbv^*)$, but they are treated in an analogous way. Since $uu^* = 1_{\mathcal{A}} = u^*u$ and $vv^* = 1_{\mathcal{A}} = v^*v$ the above example tells us that

$$\varphi(p_i(uau^*)) = \varphi(up_i(a)u^*) = \varphi(uu^*)\varphi(p_i(a)) = \varphi(p_i(a))$$

and in the same way $\varphi(q_i(vbv^*)) = \varphi(q_i(b))$. Hence, if $\varphi(p_i(uau^*)) = 0 = \varphi(q_i(vbv^*))$ then $\varphi(p_i(a)) = 0 = \varphi(q_i(b))$ for each $1 \leq i \leq r$. So we have

$$\varphi\left(p_1(uau^*)q_1(vbv^*)\cdots p_r(uau^*)q_r(vbv^*)\right) = \varphi\left(up_1(a)u^*vq_1(b)q_1(b)v^*\cdots up_r(a)u^*vq_r(b)v^*\right) = 0$$

since we also have that $\varphi(u) = \varphi(u^*) = 0 = \varphi(v) = \varphi(v^*)$ and $\{u, u^*\}, \{v, v^*\}, \{a, b\}$ are free.

With the idea of the previous examples, it can be shown the next proposition via induction. The proposition states that it is possible to compute the joint distribution of free random variables from the moments of each random variable.

Proposition 2.2.6. *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and consider a family of unital subalgebras of $\mathcal{A}$, $\{\mathcal{A}_i\}_{i\in I}$ which are free. If $\mathcal{B} = \mathrm{Alg}(\{\mathcal{A}_i : i \in I\})$, then $\varphi|_{\mathcal{B}}$ is uniquely determined by $\{\varphi|_{\mathcal{A}_i}\}_{i\in I}$.*

Remark 2.2.7.

1. It makes sense to ask for more analogous aspects of classical probability in the framework of non-commutative probability spaces along with the notion of free independence. Free Probability Theory has many interesting objects motivated by looking for analogous objects. For instance, we can have a Free Central Limit Theorem, Free Processes, Free Infinitely Divisible Distributions, Free Brownian Motion, and Free Stochastic Integrals.

2. We can also ask for an analogous theorem of Kolmogorov consistency theorem. More precisely, in the case of classical random variables, we can assure the existence of independent random variables with a given distribution. Can we translate this fact in the framework of free random variables? In other words, we are asking for the existence of the free product of non-commutative probability spaces. The answer to this question is affirmative. The interested reader can check Lectures 6 and 7 of [14] for a proof of this fact.

3. Free Probability Theory has resulted to be a very rich field with many connections with other branches of mathematics. One of its main applications is in Random Matrix Theory. In the next chapter, we shall present some results which free independence appears in several models of random matrices when the dimension of the matrices goes to infinity.

2.3 Free Central Limit Theorem

It is not difficult to note that the definition of free independence does not give a very explicit rule for computing mixed moments compared with the classical rule of independence. For example, we know that if a and b are free random variables then the distribution of $a+b$ is determined by the moments $\{\varphi(a^m), \varphi(b^n) : n, m \geq 0\}$. But computing $\varphi((a+b)^k)$ for each $k \geq 1$ directly by the definition could be tedious. In this section we shall present some quantities that have the same information that the moments but describe the relation of freeness more easily than the moments. These quantities will allow us to give a proof of a version of the Central Limit Theorem where we consider a sequence of non-commutative random variables that are freely independent.

Definition 2.3.1. Let $n \in \mathbb{N}$.

1. A *partition* π of $\{1, \ldots, n\}$ is a collection of subsets $V_1, \ldots, V_k$ of $\{1, \ldots, n\}$ such that $V_i \cap V_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^k V_i = \{1, \ldots, n\}$. The subsets V_i are called the *blocks of* π. If a, b are in the same block of π, then we write $a \sim_\pi b$.

2. We say that π of $\{1, \ldots, n\}$ is a *non-crossing partition* if $\pi = \{V_1, \ldots, V_k\}$ is a partition of $\{1, \ldots, n\}$ such that for any $a < b < c < d$ with $a \sim_\pi c$ and $b \sim_\pi d$, we have that $b \sim_\pi d$. The set of non-crossing partitions of $\{1, \ldots, n\}$ will be denoted as $\mathcal{NC}(n)$.

The set of non-crossing partitions can by provided of a structure of partial ordered set, by defining that for two non-crossing partitions $\pi, \sigma \in \mathcal{NC}(n)$, $\pi \leq \sigma$ if only if the blocks of π are contained in the blocks of σ. Also, it is a well known fact that $|\mathcal{NC}(n)| = \frac{1}{n+1}\binom{2n}{n}$ for $n \geq 1$. This number is known as the *n-th Catalan number*. This quantities appear in many combinatorial situations. We shall see later that they also have an important occurrence in the Free Central Limit Theorem.

Non-crossing partitions have a fundamental role in the combinatorics of Free Probability. The free cumulants are defined through them as follows:

Definition 2.3.2. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. The *free cumulants* of $(\mathcal{A}, \varphi)$ is the family $\{\kappa_n : \mathcal{A}^n \to \mathbb{C}\}_{n\geq 1}$ of multilinear functionals defined by the equation

$$\varphi(a_1 \cdots a_n) = \sum_{\pi \in \mathcal{NC}(n)} \kappa_\pi(a_1, \ldots, a_n) \tag{2.13}$$

where

$$\kappa_\pi(a_1, \ldots, a_n) = \prod_{V \in \pi} \kappa_{|V|}(a_1, \ldots, a_n | V), \quad a_1, \ldots, a_n \in \mathcal{A}, \tag{2.14}$$

and $\kappa_{|V|}(a_1, \ldots, a_n|V) = \kappa_r(a_{i_1}, \ldots, a_{i_r})$ if $V = \{i_1, \ldots, i_r\}$ and $i_1 < \cdots < i_r$.

Remark 2.3.3. 1. According to the definition, for the case $n = 1$ we have that the first cumulant is $\kappa_1(a_1) = \varphi(a_1)$. For the case $n = 2$, we get that

$$\varphi(a_1a_2) = \kappa_{\{\{1,2\}\}}(a_1, a_2) + \kappa_{\{\{1\},\{2\}\}}(a_1, a_2) = \kappa_2(a_1, a_2) + \kappa_1(a_1)\kappa_1(a_2)$$

and so $\kappa_2(a_1, a_2) = \varphi(a_1a_2) - \varphi(a_1)\varphi(a_2)$.

2. The existence of such of family $\{\kappa_n\}_{n\geq 1}$ in the Definition 2.3.2 is provided by the general Möbius inversion theory in the poset of non-crossing partitions. Actually, it turns out that

$$\kappa_n(a_1, \dots, a_n) = \sum_{\pi\in\mathcal{NC}(n)} \varphi_\pi(a_1, \dots, a_n)\mu_n(\pi, 1_n), \tag{2.15}$$

for any $a_1, \dots, a_n \in \mathcal{A}$, where μ_n is the Möbius function on $\mathcal{NC}(n)$ and φ_π is defined in the same fashion that κ_π just by replacing κ by $\varphi_n(a_1, \dots, a_n) = \varphi(a_1 \cdots a_n)$. It can be consulted the Lectures 9 and 10 of [14] for a detailed proof of this fact. The Equations (2.13) and (2.15) are known as *Moment - Cumulant formulas*.

We enunciate the most important theorem in the combinatorics of Free Probability. This theorem states that free independence can be easily described in terms of free cumulants. The proof can be found in Lecture 11 of [14]

Theorem 2.3.4. *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and let $\{\kappa_n\}_{n\geq 1}$ be the corresponding free cumulants. We consider $\{\mathcal{A}_i\}_{i\in I}$ a family of unital subalgebras of $\mathcal{A}$. The following statements are equivalent:*

1. *$\{A_i\}_{i\in I}$ are free.*
2. *For any $n \geq 2$ and any $a_j \in \mathcal{A}_{i(j)}$ with $i(j) \in I$ for $j = 1, \dots, n$, we have that $\kappa_n(a_1, \dots, a_n) = 0$ whenever there exist $1 \leq \ell < k \leq n$ such that $i(\ell) \neq i(k)$.*

The statement 2 in Theorem 2.3.4 is called the *mixed vanishing cumulants condition.*

Theorem 2.3.4 has many implications in Free Probability Theory. The first implication that we can consider is related two sum of free random variables. If a is a random variable in $(\mathcal{A}, \varphi)$, we denote

$$\kappa_n(a) = \kappa_n(a, \dots, a).$$

The sequence of complex numbers $\{\kappa_n(a)\}_{n\geq 1}$ is called *the sequence of free cumulants of a*. According to the moment - cumulant formulas, the sequence of free cumulants of a carries the same information that the moment sequence of a. But unlike it, free cumulants behave in a nicer way when we are dealing with sums of free random variables. More precisely, we have the next proposition which follows from Theorem 2.3.4 and from multilinearity of free cumulants.

Proposition 2.3.5. *Let a and b non-commutative random variables in a non-commutative probability space $(\mathcal{A}, \varphi)$ with corresponding free cumulants $\{\kappa_n\}_{n\geq 1}$. For any $n \geq 1$, it holds that*

$$\kappa_n(a+b) = \kappa_n(a) + \kappa_n(b). \tag{2.16}$$

With the latter proposition, we are able to give a proof of the Free Central Limit Theorem. In order to do that, we introduce the notion of convergence in distribution in the framework of non-commutative probability spaces.

Definition 2.3.6. Let $\{(\mathcal{A}_n, \varphi_n)\}_{n\geq 1}$ and $(\mathcal{A}, \varphi)$ be non-commutative probability spaces. For a index set I, consider the random variables $a_i(n) \in \mathcal{A}_n$ and $a_i \in \mathcal{A}$ for each $i \in I, n \in \mathbb{N}$. We say that the family $\{a_i(n)\}_{i\in I}$ *converges in distribution to* $\{a_i\}_{i\in I}$ if for any $m \in \mathbb{N}$ and $i_1, \ldots, i_m \in I$ we have that

$$\lim_{n\to\infty} \varphi_n\left(a_{i_1}(n)\cdots a_{i_m}(n)\right) = \varphi\left(a_{i_1}\cdots a_{i_m}\right). \tag{2.17}$$

The convergence in distribution is denoted by

$$(a_i(n))_{i\in I} \xrightarrow{d} (a_i)_{i\in I}.$$

In the framework of $*$-probability spaces, we say that $\{a_i(n)\}_{i\in I}$ *converges in $*$-distribution* if $(a_i(n), a_i^*(n))_{i\in I} \xrightarrow{d} (a_i, a_i^*)_{i\in I}$. In general, we can just say *convergence in distribution* instead of $*$-*distribution* when the context is clear.

Now we introduce a special random variable which plays a very important role in Free Probability Theory.

Definition 2.3.7. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space, $x \in \mathcal{A}$ a selfadjoint random variable and r be a positive real number. We say that x is a *semicircular element of radius r* if the moments of x are given by

$$\varphi(x^n) = \begin{cases} (\frac{r}{2})^{2k} C_k & \text{if } n = 2k, \\ 0 & \text{if } n \text{ is odd}, \end{cases} \tag{2.18}$$

where C_k denotes the k-th Catalan number. If $r = 2$, we say that x is a *standard semicircular element*.

It can be showed that the probability measure given by $d\mu_{s,t}(t) = \frac{2}{\pi r^2}\sqrt{r^2 - t^2}1_{[-r,r]}(t)$ has the moments described in (2.18). The distribution $\mu_{s,r}$ is called *semicircle distribution of radius r*.

The Free Central Limit Theorem can be stated in the following way.

Theorem 2.3.8 (Free Central Limit Theorem)**.** *Let $(\mathcal{A},\varphi)$ be a non-commutative probability space, and $\{a_i\}_{i\in\mathbb{N}}$ be a sequence of free selfadjoint random variables with the same distribution. Assume that for every $i\in\mathbb{N}$ we have that $\varphi(a_i)=0$ and $\varphi(a_i^2)=1$. Then*

$$\frac{a_1+\cdots+a_n}{\sqrt{n}} \xrightarrow{d} s, \tag{2.19}$$

where s is a standard semicircular element.

Proof. From the moment - cumulant formulas, we have that the convergence of moments is equivalent to convergence of free cumulants. Then from multilinearity and Theorem 2.3.4 we have that

$$\begin{aligned} \kappa_m\left(\frac{a_1+\cdots+a_n}{\sqrt{n}}\right) &= \frac{1}{n^{\frac{m}{2}}}\sum_{i=1}^{n}\kappa_m(a_i) \\ &= n^{-\frac{m}{2}+1}\kappa_m(a_1) \\ &\xrightarrow[n\to\infty]{} \begin{cases} 0 & \text{if } m\neq 2, \\ 1 & \text{if } m=2, \end{cases} \end{aligned}$$

Let s be the limiting random variable. Hence the only non-zero limiting free cumulant is $\kappa_2(s)=1$. Using the moment - cumulant formula, we get that if n is odd then $m_n(s)=0$. Otherwise, if $n=2k$ then

$$m_n(s)=\sum_{\pi\in\mathcal{NC}_{2k}(k)}1=|\mathcal{NC}_2(2k)|,$$

where $\mathcal{NC}_2(2k)$ denotes the set of non-crossing pairings of $\{1,\ldots,2k\}$. It is possible to establish a bijection between $\mathcal{NC}_2(2k)$ and $\mathcal{NC}(k)$, and hence $m_{2k}(s)=C_k$. We conclude that s is a standard semicircular element. ∎

Remark 2.3.9. Let $(\mathcal{A},\varphi)$ be a C^*-noncommutative probability space and $\{a_i\}_{i\in\mathbb{N}}$ be a sequence of random variables in $\mathcal{A}$ which satisfies the assumptions of Theorem 2.3.8. Consider the selfadjoint element $s_n=(a_1+\cdots+a_n)/\sqrt{n}$ and let μ_n be the distribution of s_n in the analytic sense given by Theorem 2.1.8. By the Free Central Limit Theorem, we have that the moments of μ_n converges to the moments of the standard semicircle distribution $\mu_{s,2}$. Since this distribution is compactly supported in $[-2,2]$, the semicircle distribution is uniquely determined by its moments. For a known result in measure theory, we can conclude then that Theorem 2.3.8 implies that μ_n weakly converges to the standard semicircle distribution.

Chapter 3

Free Probability and Random Matrices

In this chapter, we present some of the most relevant relations between Free Probability and Random Matrix Theory, such as Wigner's semicircle law and asymptotic freeness of Wigner matrices and deterministic matrices. The presentation of this chapter is based in Chapters 1 and 4 of the work of J. Mingo and R. Speicher [12].

3.1 Random Matrices Ensembles

Definition 3.1.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a classical probability space and $n, m \in \mathbb{N}$. A measurable function $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (M_{n\times m}(\mathbb{C}), \mathcal{B}(M_{n\times m}(\mathbb{C})))$ is called a *random matrix*.

Definition 3.1.2. Let X be an $n \times n$ random matrix on a classical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\lambda_1, \ldots, \lambda_n$ its eigenvalues. We called the *averaged spectral distribution* to the probability measure

$$\mu_X = \frac{1}{n} \sum_{i=1}^{n} \int_{\Omega} \delta_{\lambda_i(\omega)} \, d\mathbb{P}(\omega), \tag{3.1}$$

where δ_a denotes the Dirac measure on a.

Remark 3.1.3. We note that in the context of non-commutative probability spaces, the analytic distribution of a normal random variable X in $(M_n(L^{\infty-}(\Omega, \mathbb{P})), \mathbb{E} \otimes \mathrm{tr})$ coincides with its averaged spectral distribution. Indeed, thanks to the finite dimensional spectral theorem, we can write $X = UDU^*$, where U is unitary and D is the diagonal matrix with the eigenvalues of X. Then for any $r, s \geq 0$:

$$\begin{aligned}\mathbb{E}(\mathrm{tr}(X^r(X^*)^s)) &= \mathbb{E}(\mathrm{tr}(D^r(D^*)^s)) \\ &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}\lambda_i^r\overline{\lambda_i}^s\right) \\ &= \int_{\mathbb{C}} z^r\overline{z}^s\, d\mu_X(z).\end{aligned}$$

Hence μ_X is the analytic distribution of X in the sense of the Definition 2.1.7.

A sequence $\{X_n\}_{n=1}^{\infty}$ where for each $n \in \mathbb{N}$, X_n is an $n \times n$ random matrix is called *random matrices ensemble.*

Example 3.1.4 (Wigner Ensemble)**.** Consider $\{X_n\}_n$ a random matrices ensemble where for each $n \geq 1$, $X_n = (x_{ij})_{i,j=1}^n$ is such that X_n is Hermitian and $\{x_{ij}\}_{i\geq j}$ are independent identically distributed random variables with mean zero and variance $1/n$. The ensemble $\{X_n\}_{n=1}^{\infty}$ is called *Wigner ensemble.*

We recall that Z is a standard complex Gaussian random variable if $Z = (X + iY)/\sqrt{2}$, where X and Y are independent real Gaussian random variable with mean 0 and variance 1. The following ensemble is one of the most important in Random Matrix Theory.

Definition 3.1.5 (Gaussian Unitary Ensemble)**.** A *Gaussian Unitary Ensemble (GUE)* $\{Z_n\}_{n=1}^{\infty}$ is a random matrix ensemble such that if $Z_n = (z_{ij})_{i,j=1}^n$, then Z_n is Hermitian, $\{z_{ij}\}_{i\leq j}$ are independent random variables, $\sqrt{n}z_{ij}$ is a standard complex Guassian random variable for $1 \leq i < j \leq n$ and $\sqrt{n}z_{ii}$ is a standard real Guassian random variable for $1 \leq i \leq n$.

The adjective "unitary" in the above definition is because of UZ_nU^* has the same distribution than Z_n, where Z_n is a GUE and U is an $n \times n$ unitary matrix. The latter can be easily proved by computing the Fourier transform of Z_n.

Another very important example of random matrices related with the Gaussian variables is the Wishart ensemble.

Example 3.1.6 (Wishart Ensemble)**.** For each $n \in \mathbb{N}$, let X_n be an $n \times m$ random matrix whose entries are independent standard complex Guassian random variable. Consider the random matrix

$$Y_n = \frac{1}{n}YY^*.$$

The sequence $\{Y_n\}_n$ is called *Wishart ensemble.*

Now we introduce another example of random matrix. For this, we recall that the *unitary group* $\mathcal{U}(n)$ is the group of $n \times n$ matrices U with complex entries such that $U^*U = UU^* = I$. It is known that $\mathcal{U}(n)$ is a compact group, and then there exists a probability measure μ on $\mathcal{U}(n)$ with the property that it is invariant under translations. This measure is called the *Haar measure on* $\mathcal{U}(n)$ and this is unique.

Definition 3.1.7. Let $n \in \mathbb{N}$. We say that an $n \times n$ random matrix U_n is a *Haar unitary random matrix* if the distribution of U_n is the Haar measure on $\mathcal{U}(n)$.

Remark 3.1.8. Since we have not given an explicit formula for the Haar distribution, we would like to obtain a way to simulate Haar unitary random matrices for numerical experiments. A way to obtain it is to consider Z_n an $n \times n$ random matrix whose entries are independent standard complex Gaussian random variables and orthogonalize it by applying the Gram-Schmidt algorithm. This new matrix is distributed according to the Haar measure on $\mathcal{U}(n)$.

Remark 3.1.9. Let $n \in \mathbb{N}$ and let U be an $n \times n$ Haar unitary random matrix. Take r a non-zero integer and $\lambda \in \mathbb{C}$ such that $|\lambda| = 1$. Then U^r and $\lambda^r U^r$ are also Haar unitaries. Then $\mathrm{tr}(U^r) = \mathrm{tr}(\lambda^r U^r) = \lambda^r \,\mathrm{tr}(U^r)$ which implies that $\mathrm{tr}(U^r) = 0$ for $r \neq 0$. In the framework of non-commutative probability spaces, we have that U is a non-commutative random variable in the space $(M_n(L^{\infty-}(\Omega, \mathbb{P})), \mathbb{E} \otimes \mathrm{tr})$ and its algebraic distribution is given by $\mathbb{E}(\mathrm{tr}(U^r)) = \delta_{0,r}$. Hence, U is a Haar unitary in the sense of the Definition 2.2.4.

3.2 Wigner's Semicircle Law

One of the most interesting question in Random Matrix Theory is if we can say something about the limit as $n \to \infty$ of the averaged spectral distribution of a random matrices ensemble $\{A_n\}_{n=1}^{\infty}$. The next result due to Wigner gives an answer in the case of convergence in distribution of GUE random matrices. This theorem was the key in the work of Voiculescu [20] to find the connection between Free Probability Theory and Random Matrix Theory.

Theorem 3.2.1 (Wigner's Semicircle Law)**.** *Let $\{Z_n\}_{n=1}^{\infty}$ be a GUE. Then*

$$\lim_{n\to\infty} \mathbb{E}\left(\mathrm{tr}(Z_n^k)\right) = \begin{cases} C_{k/2} & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases} = \frac{1}{2\pi}\int_{-2}^{2} t^k \sqrt{4-t^2}\, dt \tag{3.2}$$

We shall give a short sketch of the proof. For $n \geq 1$, consider $Z_n = (z_{ij})_{i,j=1}^n$ be a GUE. We want to compute the expected value of the normalized trace of Z_n^k. According to the multiplication of matrices, we have that

$$\mathbb{E}((\mathrm{tr}(Z_n^k)) = \sum_{i_1,\ldots,i_k=1}^{n} \mathbb{E}\left(z_{i_1 i_2} z_{i_2 i_3} \cdots z_{i_k i_1}\right). \tag{3.3}$$

In order to compute the expectation of each term in the above sum, we use the next identity known as the *Wick formula:* if $X = (X_1, \ldots, X_k)$ is a Gaussian vector then

$$\mathbb{E}(X_{i_1} \cdots X_{i_k}) = \sum_{\pi \in \mathcal{P}_2(k)} \mathbb{E}_\pi(X_{i_1}, \ldots, i_k), \tag{3.4}$$

where $\mathcal{P}_2(k)$ denotes the set of pair partitions of $\{1, \ldots, k\}$ and $\mathbb{E}_\pi$ is defined in a similar fashion that (2.14). This formula can be extended for the complex Gaussian case because of multilinearity. According to (3.4), when k is odd every term in the sum of (3.3) is equal to 0 and so $\mathbb{E}(\mathrm{tr}(Z_n^k)) = 0$. In the case that k is even, doing a combinatorial analysis based in the covariances about which pair partitions π have a non-zero contribution to (3.3), it is possible to show that

$$\mathbb{E}(\mathrm{tr}(Z_n^k)) = \sum_{\pi \in \mathcal{P}_2(k)} N^{\ell(\gamma_k \pi) - \frac{k}{2} - 1}, \tag{3.5}$$

where π is consider as a permutation in S_k which disjoint cycles are precisely the blocks of π, $\gamma_k = (1, 2, \ldots, k)$ and $\ell(\sigma)$ denotes the number of disjoint cycles of σ. The next step is to prove that for any pair partition $\pi \in \mathcal{P}_2(2k)$, we have that $\ell(\gamma_{2k}\pi) \leq k + 1$ where the equality holds when π is non-crossing. Finally, taking the limit in (3.5) we get that

$$\lim_{n\to\infty} \mathbb{E}(\mathrm{tr}(Z_n^{2k})) = \sum_{\pi \in \mathcal{NC}_2(2k)} 1 = |\mathcal{NC}_2(2k)| = C_k, \tag{3.6}$$

which finishes the proof.

The previous theorem gives convergence of moments. In the language of non-commutative probability spaces, we have that $\{Z_n\}_{n=1}^{\infty}$ converges in distribution to s, in the sense of Definition 2.3.6. On the other hand, since the semicircle distribution has compact support, it is determined by its moments. This fact along with the convergence of moments allow to conclude that the averaged spectral distribution μ_{X_n} weakly converges to the standard semicircle distribution μ_s.

Remark 3.2.2. The Wigner's semicircle law establish the convergence to the semicircle distribution by proving the convergence of moments respect to $\mathbb{E}\otimes\mathrm{tr}$. Now, we are interested in analyzing the *almost sure convergence* the random variable $\mathrm{tr}(Z_n^k)$. Given that $\mathbb{E}(\mathrm{tr}(Z_n^k))$ converges to $\varphi(s^k)$, it can be proved that $\mathrm{tr}(Z_n^k)$ almost surely converges to $\varphi(s^k)$ by an usual argument of concentration of measures. This arguments can be stated as follows.

Using similar combinatorial ideas as in the proof of Theorem 3.2.1, we can show that $\mathrm{Var}(\mathrm{tr}(Z_n^k)) \leq \frac{C}{n^2}$, for some positive constant C. On the other hand, the almost sure convergence can be proved by establishing that for any $\epsilon > 0$, we have that

$$\mathbb{P}\left(\limsup_{n\to\mathbb{N}} \left\{ |\,\mathrm{tr}(Z_n^k) - \mathbb{E}(\mathrm{tr}(Z_n^k))| \geq \epsilon \right\}\right) = 0.$$

Since we have independence of the random variables, by Borel-Cantelli lemma the above equation can be proved if we show that

$$\sum_{n=1}^{\infty} \mathbb{P}\left(|\,\mathrm{tr}(Z_n^k) - \mathbb{E}(\mathrm{tr}(Z_n^k))| \geq \epsilon \right) < \infty.$$

But using Chevyshev's inequality and the bound for the variance, we have that

$$\mathbb{P}\left(|\,\mathrm{tr}(Z_n^k) - \mathbb{E}(\mathrm{tr}(Z_n^k))| \geq \epsilon \right) \leq \frac{1}{\epsilon^2} \mathrm{Var}(\mathrm{tr}(Z_n^k)) \leq \frac{C}{\epsilon^2 n^2},$$

which allows to prove the almost sure convergence.

The next pictures show how this different two types of convergence behave in the case of Gaussian unitary ensembles.

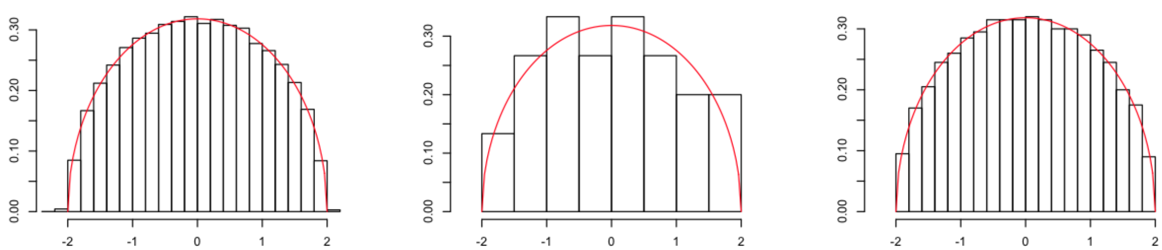

Figure 3.1: Figure from the left is the histogram of the eigenvalues of 1000 realization of 30×30 GUE random matrices. Figure of the middle is the histogram of the eigenvalues of one realization of a 30×30 GUE random matrix. Figure from the right is the histogram of the eigenvalues of a 1000×1000 GUE random matrix. In the three cases, the red line represents the density function of the semicircular distribution.

Remark 3.2.3. With a little bit more effort and following similar ideas of the proof of the Wigner's semicircle law, it is possible to show that if $Z_1, \ldots, Z_m$ are $n \times n$ independent

GUE random matrices, then

$$\lim_{n\to\infty} \mathbb{E}\left(\mathrm{tr}\left((Z_{i_1}^{r_1} - c_{r_1}I)(Z_{i_2}^{r_2} - c_{r_2}I)\cdots(Z_{i_s}^{r_s} - c_{r_s}I)\right)\right) = 0, \tag{3.7}$$

for any $s \geq 1$, $r_1, \ldots, r_s$ are positive integers, $i_1, \ldots, i_s \in \{1, \ldots, m\}$ such that $i_1 \neq i_2, i_2 \neq i_3, \ldots, i_{s-1} \neq i_s$, and $c_m = \lim_{n\to\infty} \mathbb{E}(\mathrm{tr}(Z_n^m))$. We note that the above equation is quite similar to the definition of freeness. We shall address this issue in the following section.

3.3 Asymptotic Freeness of Random Matrices

Definition 3.3.1 (Asymptotic freeness)**.** Let $\{(\mathcal{A}_n, \varphi_n)\}_{n=1}^{\infty}$ be non-commutative probability spaces and let I be an index set. For each $n \in \mathbb{N}$, let $\{a_i(n)\}_{i\in I} \subset \mathcal{A}_n$. We say that the random variables $\{a_i(n)\}_{i\in I}$ are *asymptotically freely independent* if there exist a non-commutative probability space $(\mathcal{A}, \varphi)$ and random variables $\{a_i\}_{i\in I} \subset \mathcal{A}$ such that $\{a_i(n)\}_{i\in I}$ converges in distribution to $\{a_i\}_{i\in I}$ and $\{a_i\}_{i\in I}$ are freely independent.

For the case of asymptotic freeness of random matrices, we just consider the non-commutative probability space $\mathcal{A}_n = (L^{\infty-}(\Omega_n, \mathbb{P}_n), \mathbb{E}\otimes \mathrm{tr})$ in the above definition. The first example of asymptotic freeness of random matrices arises from Remark 3.2.3, which says that independent Gaussian unitary ensembles are asymptotically free.

In the context of random matrices, we can consider a strong notion of asymptotic freeness. This notion arises from studying the almost sure convergence as we see in Remark 3.2.2. The corresponding definition can be stated in the following way.

Definition 3.3.2 (Almost sure asymptotic freeness of random matrices)**.** Let $\{A_n\}_{n=1}^{\infty}$ and $\{B_n\}_{n=1}^{\infty}$ be two sequences such that for each $n \in \mathbb{N}$, A_n and B_n are $n\times n$ random matrices defined in the same classical probability space $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$. Consider $\Omega = \prod_{n=1}^{\infty}\Omega_n$ the product space and $\mathbb{P} = \prod_{n=1}^{\infty}\mathbb{P}_n$ the product measure on Ω. We say that A_n and B_n are *almost surely asymptotically freely independent* if there exist a non-commutative probability space $(\mathcal{A}, \varphi)$, two random variables $a, b \in \mathcal{A}$ which are free and a subset $\Omega' \subset \Omega$ such that $\mathbb{P}(\Omega') = 1$ and $\{A_n(\omega), B_n(\omega)\} \subset (M_n(\mathbb{C}), \mathrm{tr})$ converges in distribution to $\{a, b\}$, for each $\omega \in \Omega'$.

Our first example of asymptotic freeness can be stated as follows.

Theorem 3.3.3 (Asymptotic freeness of GUE)**.** *Let $Z_1(n), \ldots, Z_p(n)$ be independent $n\times n$ GUE random matrices. Then*

$$Z_1(n), \ldots, Z_p(n) \xrightarrow{d} s_1, \ldots, s_p \text{ as } n \to \infty,$$

where s_i is a standard semicircular element for each $1 \leq i \leq p$ and $s_1, \ldots, s_p$ are free. The convergence in distribution also holds almost surely. Then $Z_1(n), \ldots, Z_p(n)$ are almost surely asymptotically free.

It is of interest to analyze if there is an asymptotic relation between GUE random matrices and deterministic matrices. Namely, let $\{D_n\}_{n\geq 1}$ be a sequence of deterministic matrices where $D_n \in M_n(\mathbb{C})$. Assume that there exist a non-commutative probability space $(\mathcal{A}, \varphi)$ and $d \in \mathcal{A}$ such that

$$\lim_{n\to\infty} \operatorname{tr}(D_n^k) = \varphi(d^k), \quad \forall\, k \geq 1,$$

i.e., $D_n \xrightarrow{d} d$ as $n \to \infty$. Consider a sequence $\{Z_n\}_{n=1}^{\infty}$ of GUE random matrices which converges to distribution to a semicircular element s thanks to the Wigner's theorem. We can ask about the convergence in distribution of (Z_n, D_n) as $n \to \infty$, i.e., we have to investigate the convergence of the mixed moments

$$\mathbb{E}(\operatorname{tr}(D_n^{r_1} Z_n D_n^{r_2} \cdots D_n^{r_m} Z_n)) \tag{3.8}$$

for each $m \geq 1$ and $r_1, \ldots, r_m$ non-negative integers. We can use again the Wick formula for the mixed moments of the entries of Z_n. For instance, if $Z_n = (z_{ij})_{i,j=1}^n$, $D_n^{r_k} = (d_{ij}^{(k)})_{i,j=1}^n$ and recalling that $\mathbb{E}(z_{ij} z_{k\ell}) = \delta_{i\ell}\delta_{jk}\frac{1}{n}$ then

$$\begin{aligned}
\mathbb{E}(\operatorname{tr}(D_n^{r_1} Z_n D_n^{r_2} \cdots D_n^{r_m} Z_n)) &= \frac{1}{n} \sum_{i_1, j_1, \ldots, i_m, j_m = 1}^{n} \mathbb{E}\left(d_{j_1 i_1}^{(1)} z_{i_1 j_2} d_{j_2 i_2}^{(2)} z_{i_2 j_3} \cdots d_{j_m i_m}^{(m)} z_{i_m j_1} \right) \\
&= \frac{1}{n} \sum_{i_1, j_1, \ldots, i_m, j_m = 1}^{n} \mathbb{E}\left(z_{i_1 j_2} z_{i_2 j_3} \cdots z_{i_m j_1} \right) d_{j_1 i_1}^{(1)} d_{j_2 i_2}^{(2)} \cdots d_{j_m i_m}^{(m)} \\
&= n^{-\frac{m}{2}-1} \sum_{\pi \in \mathcal{P}_2(m)} \sum_{i_1, j_1, \ldots, i_m, j_m = 1}^{n} \prod_{r=1}^{m} \delta_{i_r j_{\gamma\pi(r)}} d_{j_1 i_1}^{(1)} d_{j_2 i_2}^{(2)} \cdots d_{j_m i_m}^{(m)} \\
&= n^{-\frac{m}{2}-1} \sum_{\pi \in \mathcal{P}_2(m)} \sum_{j_1, \ldots, j_m = 1}^{n} d_{j_1 \gamma\pi(1)}^{(1)} d_{j_2 \gamma\pi(2)}^{(2)} \cdots d_{j_m \gamma\pi(m)}^{(m)}
\end{aligned}$$

where γ is the long cycle in the symmetric group $(1, \ldots, m) \in S_m$ and we are regarding $\pi \in \mathcal{P}_2(m)$ as a permutation in S_m formed as a product of transpositions (the blocks of π). Extending the definition of tr, we can consider tr_σ for $\sigma \in S_m$ as follows: if $\sigma = c_1 \cdots c_{\ell(\sigma)}$ is the decomposition into disjoint cycles of σ, and $c_j = (i_1, \ldots, i_{r(j)})$ then

$\operatorname{tr}_{c_i}(D_{i_1}, D_{i_2}, \ldots, D_{i_{r(j)}}) = \operatorname{tr}(D_{i_1} D_{i_2} \cdots D_{i_{r(j)}})$ and

$$\operatorname{tr}_\sigma(D_1, \ldots, D_m) = \prod_{j=1}^{\ell(\sigma)} \operatorname{tr}_{c_j}(D_{i_1}, \ldots, D_{i_{r(j)}}). \tag{3.9}$$

One can show that for $\sigma \in S_m$ and $A_k = (a_{ij}^{(k)})_{i,j=1}^n$, we have that

$$\operatorname{tr}_\sigma(A_1, \ldots, A_m) = n^{-\ell(\sigma)} \sum_{i_1,\ldots,i_m=1}^{n} a^{(1)}_{i_1 i_{\sigma(1)}} a^{(2)}_{i_2 i_{\sigma(2)}} \cdots a^{(m)}_{i_m i_{\sigma(m)}}. \tag{3.10}$$

Using that $\ell(\gamma\pi) \leq \frac{m}{2} + 1$ where the equality holds when π is non-crossing, we get that

$$\begin{aligned} \mathbb{E}(\operatorname{tr}(D_n^{r_1} Z_n D_n^{r_2} \cdots D_n^{r_m} Z_n)) &= \sum_{\pi \in \mathcal{P}_2(k)} n^{-\ell(\gamma\pi)-1-\frac{m}{2}} \operatorname{tr}_{\gamma\pi}(D_n^{r_1}, D_n^{r_2}, \ldots, D_n^{r_m}) \\ &\xrightarrow[n\to\infty]{} \sum_{\pi \in \mathcal{NC}_2(m)} \varphi_{\gamma\pi}(d^{r_1}, d^{r_2}, \ldots, d^{r_m}). \end{aligned}$$

It is not difficult to see that the last sum correspond to the mixed moments of the product of free random variables (see for instance Lecture 14 of [14]). Hence this allows to conclude that s and d are free. One can consider the case several GUE random matrices and deterministic matrices and work in a similar fashion to get the following theorem due to Voiculescu.

Theorem 3.3.4 (Asymptotic freeness of GUE random matrices and deterministic matrices)**.** *Let $Z_1(n), \ldots, Z_p(n)$ be independent $n \times n$ GUE random matrices and $D_1(n), \ldots, D_q(n)$ be deterministic $n \times n$ matrices such that*

$$D_1(n), \ldots, D_q(n) \xrightarrow{d} d_1, \ldots, d_q \text{ as } n \to \infty.$$

Then it holds that

$$Z_1(n), \ldots, Z_p(n), D_1(n), \ldots, D_q(n) \xrightarrow{d} s_1, \ldots, s_p, d_1, \ldots, d_q \text{ as } n \to \infty,$$

where s_i is a standard semicircular element for each $1 \leq i \leq p$ and $s_1, \ldots, s_p, \{d_1, \ldots, d_q\}$ are free. The convergence in distribution also holds almost surely. Then $Z_1(n), \ldots, Z_p(n)$, $\{D_1(n), \ldots, D_q(n)\}$ are almost surely asymptotically free.

We note that we can also wonder what happens in the case of that the matrices $D_i(n)$ are allowed to be random matrices. We can give an extension of the above theorem in the case that $D_i(n)$ and $Z_i(n)$ are independent and the $D_i(n)$'s have almost sure limit distribution. The proof of the theorem will result by conditioning on the $D_i(n)$'s and using the deterministic version.

Corollary 3.3.5. *Let $Z_1(n), \dots, Z_p(n)$ be independent $n \times n$ GUE random matrices and $D_1(n), \dots, D_q(n)$ be $n \times n$ random matrices such that*

$$D_1(n)(\omega), \dots, D_q(n)(\omega) \xrightarrow{d} d_1, \dots, d_q \text{ as } n \to \infty$$

for almost all ω. Assume that $Z_1(n), \dots, Z_p(n)$, $\{D_1(n), \dots, D_q(n)\}$ are independent. Then it holds that

$$Z_1(n)(\omega), \dots, Z_p(n)(\omega), D_1(n)(\omega), \dots, D_q(n)(\omega) \xrightarrow{d} s_1, \dots, s_p, d_1, \dots, d_q \text{ as } n \to \infty$$

almost surely where s_i is a standard semicircular element for each $1 \leq i \leq p$, $s_1, \dots, s_p$, $\{d_1, \dots, d_q\}$ are free and then $Z_1(n), \dots, Z_p(n)$, $\{D_1(n), \dots, D_q(n)\}$ are almost surely asymptotically free.

Now we consider a different example of asymptotic freeness which concerns to Haar unitary random matrices. For each $n \in \mathbb{N}$, we note that if U_n is a Haar unitary random matrix, then U_n has the same distribution that a Haar unitary element u in a non-commutative probability space $(\mathcal{A}, \varphi)$. In particular we can write $U_n \xrightarrow{d} u$ as $n \to \infty$.

As in the latter example of asymptotic freeness, we are interested in the limiting distribution of Haar unitary random matrices and deterministic matrices $D_1(n), \dots, D_q(n)$ which converges in distribution to $d_1, \dots, d_q$ in some non-commutative probability space. We can start the analysis by considering the term $\mathbb{E}(\mathrm{tr}(D^{r_1}U^{\epsilon_1} \cdots D^{r_m}U^{\epsilon_m}))$ and writing it as a sum due to the trace as we saw in the case of GUE and deterministic matrices. However, the fact of independence and Gaussian distribution of the entries of the GUE led us to consider the Wick formula (3.4) in order to compute the mixed moments in the latter example. For the case of Haar unitary random matrices, we have an analogous, although complicated, formula called the *Weingarten formula* which was introduced by Collins in [7].

Definition 3.3.6. For each $n \in \mathbb{N}$, we called the *Weingarten function* $\mathrm{Wg}(\cdot, n)$ to the linear function on $\mathbb{C}[S_m]$ for $m \leq n$ defined in the canonical basis by

$$\mathrm{Wg}(\sigma, n) = \mathbb{E}\left(u_{11} \cdots u_{mm}\overline{u}_{1\sigma(1)} \cdots \overline{u}_{m\sigma(m)}\right), \quad \forall\, \sigma \in S_m, \tag{3.11}$$

where $U = (u_{ij})_{i,j=1}^n$ is an $n \times n$ Haar unitary random matrix.

With the above definition, we can state the Weingarten formula in the next lemma.

Lemma 3.3.7. *Let $U = (u_{ij})_{i,j=1}^n$ be an $n \times n$ Haar unitary random matrix. Then for*

$n \leq m$ and any $1 \leq i(r), i(r), j(r), j(r) \leq n$ with $1 \leq r \leq m$, we have that

$$\mathbb{E}\left(u_{i(1)j(1)} \cdots u_{i(m)j(m)} \overline{u}_{i'(1)j'(1)} \cdots \overline{u}_{i'(m)j'(m)}\right) = \sum_{\sigma,\tau \in S_m} \prod_{r=1}^{n} \delta_{i(r)i'(\sigma(r))} \delta_{j(r)j'(\tau(r))} \mathrm{Wg}(\tau\sigma^{-1}, n). \tag{3.12}$$

We also have a result about the asymptotic behavior of the Weingarten function as $n \to \infty$. This result can be found in [9].

Lemma 3.3.8. *Let $n \in \mathbb{N}$. For any $\sigma \in S_m$ we have that*

$$\mathrm{Wg}(\sigma, n) = n^{-2n+\ell(\sigma)} \left(\prod_{i=1}^{\ell(\sigma)} (-1)^{m-\ell(c_i)} C_{m-\ell(c_i)} + O(n^{-2}) \right) \tag{3.13}$$

where $\sigma = c_1 \cdots c_{\ell(\sigma)}$ is the decomposition into disjoint cycles of σ, $\ell(\sigma)$ is the number of disjoint cycles which appear in such decomposition, and C_k is the k-th Catalan number.

The Weingarten function will be very useful in Chapter 5 when we are studying the convergence of random matrices which are asymptotic cyclically monotone independent.

The above lemma takes an important role in the proof of the following theorem also due to Voiculescu. We refer to the Lecture 23 [14] and Chapter 4 of [12].

Theorem 3.3.9 (Asymptotic freeness of Haar unitary and deterministic matrices)**.** *Let $U_1(n), \ldots, U_p(n)$ be independent $n \times n$ Haar unitary random matrices and $D_1(n), \ldots, D_q(n)$ be deterministic $n \times n$ matrices such that*

$$D_1(n), \ldots, D_q(n) \xrightarrow{d} d_1, \ldots, d_q \quad \text{as } n \to \infty.$$

Then it holds that

$$U_1(n), U_1^*(n), \ldots, U_p(n), U_p^*(n), D_1(n), \ldots, D_q(n) \xrightarrow{d} u_1, u_1^*, \ldots, u_p, u_p^*, d_1, \ldots, d_q \text{ as } n \to \infty,$$

where u_i is a Haar unitary element for each $1 \leq i \leq p$ and $\{u_1, u_1^*\}, \ldots, \{u_p, u_p^*\}, \{d_1, \ldots, d_q\}$ are free. The convergence in distribution holds also almost surely. Then $\{U_1(n), U_1^*(n)\}, \ldots, \{U_p(n), U_p^*(n)\}$, $\{D_1(n), \ldots, D_q(n)\}$ are almost surely asymptotically free.

Remark 3.3.10. As well as Corollary 3.3.5, we can generalize Theorem 3.3.9 to the case that the matrices D_i's are random matrices independent from the Haar unitariy random matrices and such that the D_i's have an almost sure limit distribution. The proof of this almost sure asymptotic freeness relies to the fact that it is possible to give a bound of order n^{-2} to the covariances of traces. For a discussion of this, the reader can check Chapter 5 of [12].

Combining Example 2.2.5 and Theorem 3.3.9, we get that if $\{A_n\}_{n=1}^{\infty}$ and $\{B_n\}_{n=1}^{\infty}$ are two sequences of deterministic matrices such that $(A_n, B_n) \xrightarrow{d} (a, b)$ and $\{U_n\}_{n=1}^{\infty}$ is a sequence of $n \times n$ Haar unitary random matrices, A_n and $U_n B_n U_n^*$ are asymptotically free. This establishes that two deterministic matrices are asymptotically free when one of them is randomly rotated. Actually, we only need that $A_n \xrightarrow{d} a$ and $B_n \xrightarrow{d} b$ instead of the assumption of joint convergence. A proof of this fact appears in Lecture 23 of [14].

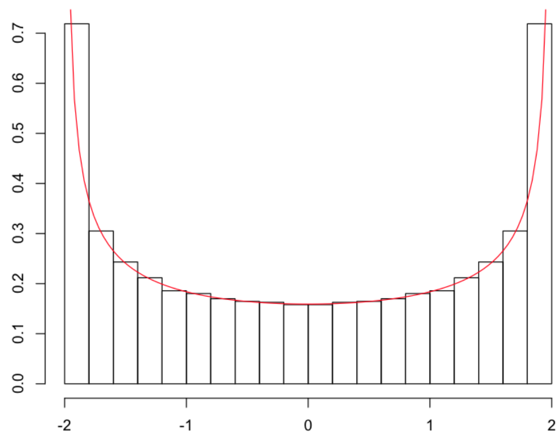

Figure 3.2: Histogram of the eigenvalues of $A_n + U_n B_n U_n^*$, where A_n and B_n are $n \times n$ deterministic diagonal matrices with $n/2$ 1's and $n/2$ -1's in the diagonal, and U_n is a Haar unitary random matrix, with $n = 1000$. The limiting distribution corresponds to the distribution of the sum of two free random variables with the Bernoulli distribution $(\delta_1 + \delta_{-1})/2$. The red line is the density of the limiting distribution which can be effectively computed by analytic methods. We refer to Lecture 12 of [14] for the details.

In particular, by using Example 2.2.5, we have the next corollary which will be useful in the following chapters.

Corollary 3.3.11. *Let $U_1(n), \ldots, U_p(n)$ be independent $n \times n$ Haar unitary random matrices and $D_1(n), \ldots, D_p(n)$ be $n \times n$ deterministic matrices such that $(D_1(n), \ldots, D_q(n))$ converges in distribution. Then $U_1(n)D_1(n)U_1^*(n), \ldots, U_p(n)D_p(n)U_p^*(n)$ are asymptotically free.*

To finish this chapter, we state the general version of asymptotic freeness for Wigner ensembles and deterministic matrices. A combinatorial proof based in sums given in terms of graphs can be studied in detail in Chapter 4 of [12]. Another proof can be found in [1].

Theorem 3.3.12 (Asymptotic freeness of Wigner and deterministic matrices)**.** *Let $\mu_1, \ldots, \mu_p$ be probability measures on $\mathbb{R}$ such that all moments exist and the first moment is equal to* *zero. Let $A_1(n), \ldots, A_p(n)$ be independent $n \times n$ Wigner random matrices where the distribution of the entries of $A_i(n)$ is μ_i for each $i \in \{1, \ldots, p\}$. Let $D_1(n), \ldots, D_q(n)$ be $n \times n$ deterministic matrices such that*

$$D_1(n), \ldots, D_q(n) \xrightarrow{d} d_1, \ldots, d_q \quad \text{as } n \to \infty$$

and

$$\sup_{n \in \mathbb{N},\, r=1,\ldots,q} \|D_r(n)\| < \infty.$$

Then it holds that

$$A_1(n), \ldots, A_p(n), D_1(n), \ldots, D_q(n) \xrightarrow{d} s_1, \ldots, s_p, d_1, \ldots, d_q,$$

where s_i is a standard semicircular element for each $1 \leq i \leq p$ and $s_1, \ldots, s_p$, $\{d_1, \ldots, d_q\}$ are free. The convergence in distribution also holds almost surely. Then $A_1(n), \ldots, A_p(n)$, $\{D_1(n), \ldots, D_q(n)\}$ are almost surely asymptotically free.

In the same way that the other examples of asymptotic freeness, we can allow that the matrices $D_i(n)$'s are random matrices. The conclusion of the above theorem still holds provided that $\{D_1(n), \ldots, D_q(n)\}$ are independent random matrices from the Wigner matrices and have almost sure limit distribution. The proof of this fact follows the same idea of conditioning, using the assumption of independence and the theorem in its deterministic version.

Chapter 4

Cyclic Monotone Independence

In this chapter, we present the abstract notion of cyclic monotone independence. We shall study the basic definitions and compute the eigenvalues of certain polynomials in cyclically monotone elements which can be seen as trace class operators. As we stated in the introduction, this chapter is based in the work of Collins, Hasebe and Sakuma [8].

4.1 The Rule of Cyclic Monotone Independence

In a similar way that we think the space of $n \times n$ matrices with the normalized trace as a non-commutative probability space, we want to put the notion of trace on a Hilbert space in a more general framework. This leads us to define the concept of non-commutative measure space.

Definition 4.1.1. A *non-commutative measure space* is a pair $(\mathcal{A}, \omega)$ where $\mathcal{A}$ is a $*$-algebra over $\mathbb{C}$ and ω is a *tracial weight* which means that:

- ω is defined in a $*$-subalgebra $\mathrm{D}(\omega)$ of $\mathcal{A}$,
- $\omega : \mathrm{D}(\omega) \to \mathbb{C}$ is linear,
- ω is positive, which means that $\omega(a^*a) \geq 0$ for all $a \in \mathrm{D}(\omega)$,
- ω is selfadjoint, i.e. $\omega(a^*) = \overline{\omega(a)}$ for all $a \in \mathrm{D}(\omega)$,
- ω is tracial, i.e. $\omega(ba) = \omega(ab)$ for all $a, b \in \mathrm{D}(\omega)$.

Remark 4.1.2. We notice that in the case that $\mathcal{A}$ is unital, $\mathrm{D}(\omega) = \mathcal{A}$, and $\omega(1_{\mathcal{A}}) = 1$, then (A, ω) is a non-commutative probability space. In this chapter, by a non-commutative probability space, we mean a $*$-probability space $(\mathcal{C}, \phi)$ where the linear functional ϕ is tracial.

We can define some analogous concepts of non-commutative probability in this new setting. For instance:

Definition 4.1.3.
- Let $(\mathcal{A}, \omega)$ be a non-commutative measure space and let $a_1, \ldots, a_k \in \mathrm{D}(\omega)$. We define the *distribution* of $(a_1, \ldots, a_k)$ as the set of mixed moments

$$\{\omega(a_{i_1}^{\epsilon_1} \cdots a_{i_p}^{\epsilon_p}) \,:\, p \geq 1, 1 \leq i_1, \ldots, i_p \leq k, (\epsilon_1, \ldots, \epsilon_k) \in \{1, *\}^p\}. \tag{4.1}$$

- Let $(\mathcal{A}, \omega)$, $(\mathcal{B}, \xi)$ be non-commutative measure spaces and let $a_1, \ldots, a_k \in \mathrm{D}(\omega)$, $b_1, \ldots, b_k \in \mathrm{D}(\xi)$. We say that $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ have *the same distribution* if

$$\omega(a_{i_1}^{\epsilon_1} \cdots a_{i_p}^{\epsilon_p}) = \xi(b_{i_1}^{\epsilon_1} \cdots b_{i_p}^{\epsilon_p}), \qquad \forall\, p \geq 1,\, 1 \leq i_1, \ldots, i_p \leq k,\, (\epsilon_1, \ldots, \epsilon_k) \in \{1, *\}^p. \tag{4.2}$$

The class of elements in the non-commutative measure spaces on which we will work are compact operators. For this reason, it is convenient to define the notion of trace class distribution and convergence in eigenvalues in this new context. More specifically, we have the next definition.

Definition 4.1.4. Let $(\mathcal{A}, \omega)$ be a non-commutative measure space and let $a_1, \ldots, a_k \in \mathrm{D}(\omega)$. We say that $(a_1, \ldots, a_k)$ has *trace class distribution* if there exist a separable Hilbert space H and elements $x_1, \ldots, x_k \in (S^1(H), \mathrm{Tr}_H)$ such that the distribution of $(a_1, \ldots, a_k)$ is the same as that of $(x_1, \ldots, x_k)$. In this case, we define the eigenvalues of a selfadjoint $*$-polynomial $P(a_1, \ldots, a_k)$ to be the eigenvalues of $P(x_1, \ldots, x_k)$.

Definition 4.1.5. Let $(\mathcal{A}, \omega)$, $\{(\mathcal{A}_n, \omega_n)\}_{n=1}^{\infty}$ be non-commutative measure spaces and let $a_1, \ldots, a_k \in \mathrm{D}(\omega)$, $a_1(n), \ldots, a_k(n) \in \mathrm{D}(\omega_n)$, for all $n \geq 1$. We say that $(a_1(n), \ldots, a_k(n))$ *converges in distribution* to $(a_1, \ldots, a_k)$ if

$$\lim_{n\to\infty} \omega_n \left(a_{i_1}(n)^{\epsilon_1} \cdots a_{i_p}(n)^{\epsilon_p}\right) = \omega\left(a_{i_1}^{\epsilon_1} \cdots a_{i_p}^{\epsilon_p}\right), \tag{4.3}$$

for any $p \geq 1$, $1 \leq i_1, \ldots, i_p \leq k$, $(\epsilon_1, \ldots, \epsilon_p) \in \{1, *\}^p$. If the distributions of $a \in \mathrm{D}(\omega)$ and $a_n \in \mathrm{D}(\omega_n)$ are trace class, we define *convergence in eigenvalues* according to Definition 1.3.2.

Translating the results of Section 1.3, we get the next proposition about the convergence in eigenvalues but considering elements in a non-commutative measure space which are trace class.

Proposition 4.1.6. *$(\mathcal{A}, \omega)$, $\{(\mathcal{A}_n, \omega_n)\}_{n=1}^{\infty}$ be non-commutative measure spaces. Suppose that $a_i(n) \in \mathrm{D}(\omega_n)$, $a_i \in \mathrm{D}(\omega)$, for $i = 1, \ldots, k$, have trace class distribution and that $(a_1(n), \ldots, a_k(n))$ converges in distribution to $(a_1, \ldots, a_k)$. Then, for any non-commutative selfadjoint $*$-polynomial P without a constant term, $P(a_1(n), \ldots, a_k(n))$ converges in eigenvalues to $P(a_1, \ldots, a_k)$.*

Proof. This result follows from the above definitions on each monomial of P and the Proposition 1.3.6 with $p = 1$. ∎

The next step is defining a notion of independence in a non-commutative measure space. Motivated for the results of Shlyakhtenko in [17] about the asymptotic behavior of the moments of products of rotationally invariant random matrices and matrices whose all entries are zero except one of them, Collins, Hasebe and Sakuma in [8] defined an abstract notion of independence which fits in the framework of non-commutative probability spaces provided with a tracial weight.

In order to state the notion of independence in which we are interested, we introduce some notation. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be $*$-algebras such that $\mathcal{A}, \mathcal{B} \subset \mathcal{C}$ and $1_{\mathcal{C}} \in \mathcal{B}$. We define the set

$$\mathrm{I}_{\mathcal{B}}(\mathcal{A}) := \operatorname{span}\{b_0 a_1 b_1 \cdots a_n b_n \,:\, n \in \mathbb{N},\, a_1, \ldots, a_n \in \mathcal{A}, b_0, \ldots, b_n \in \mathcal{B}\}. \tag{4.4}$$

We can readily check that $\mathrm{I}_{\mathcal{B}}(\mathcal{A})$ is a $*$-ideal of $\mathcal{C}$ that contains $\mathcal{A}$. The definition of cyclic monotone independence can be written as follows.

Definition 4.1.7. Let $(\mathcal{C}, \tau)$ be a non-commutative probability space with a tracial weight ω. Let $\mathcal{A}, \mathcal{B} \subset \mathcal{C}$ be $*$-subalgebras such that $1_{\mathcal{C}} \in \mathcal{B}$. We say that the pair $(\mathcal{A}, \mathcal{B})$ is *cyclically monotonically independent* with respect to (ω, τ) if $\mathrm{I}_{\mathcal{B}}(\mathcal{A}) \subset \mathrm{D}(\omega)$ and for any $n \in \mathbb{N}$, $a_1, \ldots, a_n \in \mathcal{A}$, $b_1, \ldots, b_n \in \mathcal{B}$, we have that

$$\omega(a_1 b_1 a_2 b_2 \cdots a_n b_n) = \omega(a_1 a_2 \cdots a_n)\tau(b_1)\tau(b_2) \cdots \tau(b_n). \tag{4.5}$$

If $(\mathcal{A}, \mathcal{B})$ is cyclically monotonically independent with respect to (ω, τ), we shall often say that the pair $(\mathcal{A}, \mathcal{B})$ is *cyclically monotone*.

Remark 4.1.8. The above definition can be extended to the case that we are considering subsets instead of subalgebras. More precisely, we say that the pair $(\{a_1, \ldots, a_k\}, \{b_1, \ldots, b_\ell\})$ is *cyclically monotone* if the pair $(\mathrm{alg}\{a_1, \ldots, a_k\}, \mathrm{alg}\{1_{\mathcal{C}}, b_1, \ldots, b_\ell\})$ is cyclically monotone, where $a_1, \ldots, a_k \in \mathrm{D}(\omega)$ and $b_1, \ldots, b_\ell \in \mathcal{C}$, and $\mathrm{alg}(\mathcal{X})$ denotes the (not necessarily unital) $*$-algebra in $\mathcal{C}$ generated by $\mathcal{X} \subset \mathcal{C}$.

An interesting question related with cyclic monotone independence is if given a non-commutative probability space $(\mathcal{B}, \tau)$ and a non-commutative measure space $(\mathcal{A}, \omega)$, there exists an space where the pair $(\mathrm{D}(\omega), \mathcal{B})$ is cyclically monotone. A universal construction is given in the next definition.

Definition 4.1.9. Let $(\mathcal{A}, \omega)$ be a non-commutative measure space and let $(\mathcal{B}, \tau)$ be a non-commutative probability space. We consider $\mathcal{A} * \mathcal{B}$ the algebraic free product of $\mathcal{A}$ and $\mathcal{B}$ where the identity element $1_{\mathcal{B}}$ is identified with the unit elements $\mathbb{C}1_{\mathcal{A}*\mathcal{B}}$. If $\mathcal{B} = \mathcal{B}^0 \oplus \mathbb{C}1_{\mathcal{B}}$ is a direct sum decomposition as a vector space, $\mathcal{X}_1 = \mathcal{B}^0$ and $\mathcal{X}_2 = \mathcal{A}$, then

$$\mathcal{A} * \mathcal{B} = \mathbb{C}1_{\mathcal{A}*\mathcal{B}} \oplus \left(\bigoplus_{n=1}^{\infty} \bigoplus_{\substack{i_1,\ldots,i_n \in \{1,2\} \\ i_1 \neq i_2, \ldots, i_{n-1} \neq i_n}} (\mathcal{X}_1 \otimes \cdots \otimes \mathcal{X}_n) \right).$$

We define the linear functional $\omega \trianglerighteq \tau : \mathrm{D}(\omega \trianglerighteq \tau) \to \mathbb{C}$ by the formula

$$(\omega \trianglerighteq \tau)(b_0 a_1 b_1 a_2 b_2 \cdots a_n b_n) := \omega(a_1 a_2 \cdots a_n)\tau(b_1)\tau(b_2)\cdots\tau(b_{n-1})\tau(b_0 b_n), \tag{4.6}$$

for $n \geq 1, a_1, \ldots, a_n \in \mathrm{D}(\omega), b_0, \ldots, b_n \in \mathcal{B}$, where

$$\mathrm{D}(\omega \trianglerighteq \tau) = \mathrm{D}(\omega) \oplus (\mathrm{D}(\omega) \otimes \mathcal{B}^0) \oplus (\mathcal{B}^0 \otimes \mathrm{D}(\omega)) \oplus (\mathrm{D}(\omega) \otimes \mathcal{B}^0 \otimes \mathrm{D}(\omega)) \oplus \cdots. \tag{4.7}$$

We say that $\omega \trianglerighteq \tau$ is the *cyclic monotone product* of ω and τ.

As we are expecting, it turns out that the new space obtained by the cyclic monotone product is a non-commutative measure space.

Proposition 4.1.10. *Let $(\mathcal{A}, \omega)$ be a non-commutative measure space and let $(\mathcal{B}, \tau)$ be a non-commutative probability space. Then $(\mathcal{A}*\mathcal{B}, \omega \trianglerighteq \tau)$ is a non-commutative measure space.*

Proof. We have to prove that $\omega \trianglerighteq \tau$ is a tracial positive linear functional on $\mathrm{D}(\omega \trianglerighteq \tau)$. The fact that $\omega \trianglerighteq \tau$ is a linear functional is obvious according how it is defined. Now, we shall prove that $\omega \trianglerighteq \tau$ is tracial. Indeed, if $x = b_0 a_1 b_1 \cdots a_n b_n$, $y = b'_0 a'_1 b'_1 \cdots a'_m b'_m$, with $a_i, a'_i \in \mathrm{D}(\omega)$, $b_j, b'_j \in \mathcal{B}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$, then

$$\begin{aligned} (\omega \trianglerighteq \tau)(xy) &= \omega(a_1 \cdots a_n a'_1 \cdots a'_m)\tau(b_0 b'_m)\tau(b_n b'_0) \prod_{i=1}^{n-1} \tau(b_i) \prod_{j=1}^{m-1} \tau(b'_j) \\ &= \omega(a'_1 \cdots a'_m a_1 \cdots a_n)\tau(b'_m b_0)\tau(b'_0 b_n) \prod_{j=1}^{m-1} \tau(b'_j) \prod_{i=1}^{n-1} \tau(b_i) \\ &= (\omega \trianglerighteq \tau)(yx), \end{aligned}$$

where it is used that ω and τ are tracial. It remains to prove that $\omega \trianglerighteq \tau$ is positive. In order to prove it, we take

$$x = \sum_{i=1}^{n} \lambda_i b_{i,0} a_{i,1} b_{i,1} \cdots a_{i,m(i)} b_{i,m(i)},$$

with $\lambda_i \in \mathbb{C}$, $a_{i,j} \in \mathrm{D}(\omega)$ and $b_{i,j} \in \mathcal{B}$. Since ω is positive, we have

$$\begin{aligned} 0 &\leq \omega\left(\left(\sum_{i=1}^{n} \lambda_i a_{i,1} \cdots a_{i,m(i)}\right)^* \left(\sum_{i=1}^{n} \lambda_i a_{i,1} \cdots a_{i,m(i)}\right)\right) \\ &= \sum_{i,j=1}^{n} \overline{\lambda_i} \lambda_j \omega\left(a^*_{i,m(i)} \cdots a^*_{i,1} a_{j,1} \cdots a_{j,m(j)}\right), \end{aligned}$$

for any $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. So, we have that the matrix $A = \left(\omega\left(a^*_{i,m(i)} \cdots a^*_{i,1} a_{j,1} \cdots a_{j,m(j)}\right)\right)_{i,j=1}^{n}$ is positive definite. Since τ is positive, it can also be shown that the matrices

$$\begin{aligned} B' &= \left(\tau(b^*_{i,0} b_{j,0})\right)_{i,j=1}^{n}, \\ B'' &= \left(\tau(b^*_{i,1})\tau(b_{j,1}) \cdots \tau(b^*_{i,m(i)-1})\tau(b_{j,m(j)-1})\tau(b^*_{i,m(i)} b_{j,m(j)})\right)_{i,j=1}^{n} \end{aligned}$$

are positive definite in an analogous way. For instance, we have that

$$\begin{aligned} 0 &\leq \tau\left(\left(\sum_{i=1}^{n} \lambda_i \tau(b_{i,1}) \cdots \tau(b_{i,m(i)-1}) b_{i,m(i)}\right)^* \left(\sum_{i=1}^{n} \lambda_i \tau(b_{i,1}) \cdots \tau(b_{i,m(i)-1}) b_{i,m(i)}\right)\right) \\ &= \sum_{i,j=1}^{n} \overline{\lambda_i} \lambda_j \tau(b^*_{i,1})\tau(b_{j,1}) \cdots \tau(b^*_{i,m(i)-1})\tau(b_{j,m(j)-1})\tau(b^*_{i,m(i)} b_{j,m(j)}) \end{aligned}$$

for any $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$.

Finally, it is known that the Schur product $A \circ B' \circ B''$ is positive definite if A, B' and B'' are. Hence

$$\begin{aligned} (\omega \trianglerighteq \tau)(x^* x) &= (\omega \trianglerighteq \tau)\left(\left(\sum_{i=1}^{n} \lambda_i b_{i.0} a_{i,1} b_{i,1} \cdots a_{i,m(i)} b_{i,m(i)}\right)^* \left(\sum_{i=1}^{n} \lambda_i b_{i.0} a_{i,1} b_{i,1} \cdots a_{i,m(i)} b_{i,m(i)}\right)\right) \\ &= (\lambda_1, \ldots, \lambda_n)^* \left(A \circ B' \circ B''\right) (\lambda_1, \ldots, \lambda_n) \\ &\geq 0, \end{aligned}$$

for any $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, that is what we wanted to prove. ■

Remark 4.1.11. Given $(\mathcal{A}, \omega)$ and $(\mathcal{B}, \tau)$ as in the above proposition, we can construct a non-commutative probability space with a tracial weight provided by their cyclic monotone product and the above proposition. In this space, we have that the pair $(\mathrm{D}(\omega), \mathcal{B})$ is cyclically monotone. For this purpose, consider $\tilde{\tau} : \mathcal{A} * \mathcal{B} \to \mathbb{C}$ the free product of zero function in $\mathcal{A}$ and τ. Then we have that $\tilde{\tau}$ is a tracial state in $\mathcal{A} * \mathcal{B}$. Just by the definition of cyclic monotone product, we have that $(\mathcal{A} * \mathcal{B}, \omega \trianglerighteq \tau, \tilde{\tau})$ is a non-commutative probability space with a tracial weight, such that the pair $(\mathrm{D}(\omega), \mathcal{B})$ is cyclically monotone with respect to $(\omega \trianglerighteq \tau, \tilde{\tau})$.

A final ingredient in relation with the task of looking for analogous definitions of non-commutative probability theory in the case of non-commutative measure spaces with the notion of cyclic monotone independence, is stating an appropriate idea of asymptotic cyclic monotone independence. The main reason of why this is necessary is because the applications studied in this work have relation with random matrices. With this in mind, we have the following definition.

Definition 4.1.12. Let $(\mathcal{A}_n, \omega_n, \tau_n)$ be non-commutative probability spaces with tracial weights ω_n for each $n \geq 1$. Let $a_1(n), \ldots, a_k(n) \in \mathrm{D}(\omega_n)$, $b_1(n), \ldots, b_\ell(n) \in \mathcal{A}_n$. We say that the pair $(\{a_1(n), \ldots, a_k(n)\}, \{b_1(n), \ldots, b_\ell(n)\})$ is *asymptotically cyclically monotone* if there exist a non-commutative probability space $(\mathcal{A}, \omega, \tau)$ with a tracial weight ω, and elements $a_1, \ldots, a_k \in \mathrm{D}(\omega)$, $b_1, \ldots, b_\ell \in \mathcal{A}$ such that

1. the pair $(\{a_1, \ldots, a_k\}, \{b_1, \ldots, b_\ell\})$ is cyclically monotone,
2. for any non-commutative $*$-polynomial $P(x_1, \ldots, x_k, y_1, \ldots, y_\ell)$ such that

$$P(0, \ldots, 0, y_1, \ldots, y_\ell) = 0,$$

we have that $P(a_1(n), \ldots, a_k(n), b_1(n), \ldots, b_\ell(n))$ is an element of the domain of ω_n and the following limit holds:

$$\lim_{n\to\infty} \omega_n \left(P(a_1(n), \ldots, a_k(n), b_1(n), \ldots, b_\ell(n))\right) = \omega \left(P(a_1, \ldots, a_k, b_1, \ldots, b_\ell)\right). \tag{4.8}$$

4.2 Eigenvalues of Polynomials of Cyclically Monotone Elements

An interesting question in relation with the notion of cyclic monotone independence is whether we can give explicit formulas for the eigenvalues of polynomials on cyclically monotone elements in the framework of trace class operators. In the paper [8], the authors establish such formulas for the case of some specific polynomials. The objective of this section is to present the main contribution of this work. In Proposition 4.2.2, we provide a new generalization of the formulas in [8]. This allows to give an alternative proof of the formulas and also sheds light for new ones described in Proposition 4.2.5 and Corollary 4.2.8.

The formulas of the authors of [8] are described in the next theorem.

Theorem 4.2.1. *Let $(\mathcal{A}, \omega, \tau)$ be a non-commutative probability space with tracial weight ω. Consider $a, a_1, \ldots, a_k \in \mathrm{D}(\omega)$ and $b, b_1, \ldots, b_k \in \mathcal{A}$ such that $(a, a_1, \ldots, a_k)$ are trace class with respect to ω and $(\{a_1, \ldots, a_k\}, \{b_1, \ldots, b_k\})$ is cyclically monotone with respect to (ω, τ).*

1. *If $a_1, \ldots, a_k$ are selfadjoint and $B = (\tau(b_i^* b_j))_{i,j=1}^k \in M_k(\mathbb{C})$, then*

$$\mathrm{EV}\left(\sum_{i=1}^k b_i a_i b_i^*\right) = \mathrm{EV}\left(\sqrt{B}\,\mathrm{diag}(a_1, \ldots, a_k)\sqrt{B}\right),$$

where $\sqrt{B}\,\mathrm{diag}(a_1, \ldots, a_k)\sqrt{B} \in (M_k(\mathbb{C}) \otimes \mathcal{A}, \mathrm{Tr}_k \otimes \omega)$.

2. *If $b_1, \ldots, b_k$ are selfadjoint, then*

$$\mathrm{EV}\left(\sum_{i=1}^k a_i b_i a_i^*\right) = \mathrm{EV}\left(\sum_{i=1}^k \tau(b_i) a_i a_i^*\right).$$

3. *If a, b are selfadjoint, $p = \sqrt{\tau(b^2)} + \tau(b)$ and $q = -\sqrt{\tau(b^2)} + \tau(b)$, then*

$$\mathrm{EV}(ab + ba) = (p\,\mathrm{EV}(a)) \sqcup (q\,\mathrm{EV}(a)).$$

4. *If a, b are selfadjoint and $r = \sqrt{\tau(b^2) - \tau(b)^2}$, then*

$$\mathrm{EV}(\mathrm{i}(ab - ba)) = (r\,\mathrm{EV}(a)) \sqcup (-r\,\mathrm{EV}(a)).$$

The above theorem is an example of how cyclic monotone independence can be used to compute the eigenvalue set of some specific selfadjoint polynomials. Noticing that in the above formulas the eigenvalues only depend of the elements b_i's through their first moments, it is natural to ask if in general we can replace the elements b_i's in the polynomials by their moments. In this sense, we have the next result which is the main contribution of this manuscript.

Proposition 4.2.2. *Let $(\mathcal{A}, \omega, \tau)$ be a non-commutative probability space with tracial weight ω. Consider $A_p = \left(a_{ij}^{(p)}\right)_{i,j}^n \in M_n(\mathrm{D}(\omega))$ and $B_q = \left(b_{ij}^{(q)}\right)_{i,j}^n \in M_n(\mathcal{A})$ for $p, q = 1, \ldots, k$.*

Assume that $\left(a_{ij}^{(p)}\right)_{i,j=1,p=1,\ldots,k}^{n}$ *are trace class with respect to* ω *, and that the pair*

$$\left(\left\{a_{i,j}^{(p)}\ :\ p=1,\ldots,k,\ i,j=1,\ldots,n\right\},\left\{b_{i,j}^{(q)}\ :\ q=1,\ldots,k,\ i,j=1,\ldots,n\right\}\right)$$

is cyclically monotone independent with respect to (ω,τ)*. Then*

$$\mathrm{Tr}_n\otimes\omega\left(A_1B_1A_2B_2\cdots A_kB_k\right)=\mathrm{Tr}_n\otimes\omega\left(A_1B_1'A_2B_2'\cdots A_kB_k'\right), \tag{4.9}$$

where for each $p=1,\ldots,k$*, we have that*

$$B_p'=\mathrm{id}_n\otimes\tau(B_p)=\left(\tau\left(b_{ij}^{(p)}\right)\right)_{i,j=1}^{n}\in M_n(\mathbb{C}).$$

Proof. Using linearity of Tr_n and ω, and cyclic monotone independence, it easily follows that

$$\begin{aligned}
\mathrm{Tr}_n\otimes\omega\left(A_1B_1A_2B_2\cdots A_kB_k\right) &= \sum_{\substack{i_1,i_2\ldots,i_k=1\\ j_1,j_2\ldots,j_k=1}}^{n}\omega\left(a_{i_1j_1}^{(1)}b_{j_1i_2}^{(1)}a_{i_2j_2}^{(2)}b_{j_2i_3}^{(2)}\cdots a_{i_kj_k}^{(k)}b_{j_ki_1}^{(k)}\right)\\
&= \sum_{\substack{i_1,i_2\ldots,i_k=1\\ j_1,j_2\ldots,j_k=1}}^{n}\omega\left(a_{i_1j_1}^{(1)}a_{i_2j_2}^{(2)}\cdots a_{i_kj_k}^{(k)}\right)\tau\left(b_{j_1i_2}^{(1)}\right)\tau\left(b_{j_2i_3}^{(2)}\right)\cdots\tau\left(b_{j_ki_1}^{(k)}\right)\\
&= \sum_{\substack{i_1,i_2\ldots,i_k=1\\ j_1,j_2\ldots,j_k=1}}^{n}\omega\left(a_{i_1j_1}^{(1)}\tau\left(b_{j_1i_2}^{(1)}\right)a_{i_2j_2}^{(2)}\tau\left(b_{j_2i_3}^{(2)}\right)\cdots a_{i_kj_k}^{(k)}\tau\left(b_{j_ki_1}^{(k)}\right)\right)\\
&= \mathrm{Tr}_n\otimes\omega\left(A_1B_1'A_2B_2'\cdots A_kB_k'\right).
\end{aligned}$$

■

Since a power of a matrix of the form $A_1B_1\cdots A_kB_k$ has the same form, it follows that:

Corollary 4.2.3. *With the assumptions and notation of Proposition 4.2.2, for any* $m\geq 1$*, we have that*

$$\mathrm{Tr}_n\otimes\omega\left(\left(A_1B_1A_2B_2\cdots A_kB_k\right)^m\right)=\mathrm{Tr}_n\otimes\omega\left(\left(A_1B_1'A_2B_2'\cdots A_kB_k'\right)^m\right), \tag{4.10}$$

i.e., the moments of $A_1B_1A_2B_2\cdots A_kB_k$ *and* $A_1B_1'A_2B_2'\cdots A_kB_k'$ *with respect* $\mathrm{Tr}_n\otimes\omega$ *are the same.*

We recall Corollary 1.3.7. Providing that $A_1B_1\cdots A_kB_k$ and $A_1B_1'\cdots A_kB_k'$ have the same distribution that of selfadjoint trace class operators and the same moments, it follows that

$$\mathrm{EV}(A_1B_1\cdots A_kB_k) = \mathrm{EV}(A_1B_1'\cdots A_kB_k').$$

With this idea, we can give a proof of Theorem 4.2.1 using Proposition 4.2.2.

Proof of Theorem 4.2.1. 1) Assume that $a_1,\ldots,a_k$ are selfadjoint and consider $B = (\tau(b_i^*b_j))_{i,j=1}^k$. Define the matrices

$$A_1 = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_k \end{pmatrix}, \qquad B_0 = \begin{pmatrix} b_1 & b_2 & \cdots & b_k \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

We notice that

$$B_0A_1B_0^* = \begin{pmatrix} \sum_{i=1}^k b_ia_ib_i^* & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

which is a selfadjoint element in $M_k(\mathbb{C})\otimes\mathcal{A}$. By traciality, the moments of $B_0A_1B_0^*$ with respect to $\mathrm{Tr}_k\otimes\omega$ are the same that the moments of $A_1B_0^*B_0$. By Proposition 4.2.2, these moments are the same that of AB. Since B is positive definite, we have that the moments of $B_0A_1B_0^*$ with respect to $\mathrm{Tr}_k\otimes\omega$ are the same that the moments of $\sqrt{B}A_1\sqrt{B}$ which is selfadjoint. Since for any $m\geq 1$ we have that

$$\mathrm{Tr}\otimes\omega((B_0A_1B_0^*)^m) = \omega\left(\left(\sum_{i=1}^k b_ia_ib_i^*\right)^m\right),$$

by Corollary 1.3.7, we conclude that

$$\mathrm{EV}\left(\sum_{i=1}^k b_ia_ib_i^*\right) = \mathrm{EV}\left(\sqrt{B}\,\mathrm{diag}(a_1,\ldots,a_k)\sqrt{B}\right).$$

2) Using the same idea of above, if we define

$$A_1 = \begin{pmatrix} a_1 & a_2 & \cdots & a_k \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \qquad B_1 = \mathrm{diag}(b_1,\ldots,b_k),$$

we have that

$$A_1 B_1 A_1^* = \operatorname{diag}\left(\sum_{i=1}^{k} a_i b_i a_i^*, 0, \ldots, 0\right)$$

which is selfadjoint. By Proposition 4.2.2, if $B_1' = \operatorname{diag}(\tau(b_1), \tau(b_2), \ldots, \tau(b_k))$, it follows that for any $m \geq 1$

$$\begin{aligned}
\omega\left(\left(\sum_{i=1}^{k} a_i b_i a_i^*\right)^m\right) &= \operatorname{Tr}_k \otimes \omega\left((A_1 B_1 A_1^*)^m\right) \\
&= \operatorname{Tr}_k \otimes \omega\left((A_1 B_1' A_1^*)^m\right) \\
&= \omega\left(\left(\sum_{i=1}^{k} a_i \tau(b_i) a_i^*\right)^m\right).
\end{aligned}$$

Hence

$$\operatorname{EV}\left(\sum_{i=1}^{k} a_i b_i a_i^*\right) = \operatorname{EV}\left(\sum_{i=1}^{k} \tau(b_i) a_i a_i^*\right).$$

3) Assume that a, b are selfadjoint. Consider the matrices

$$B_0 = \begin{pmatrix} 1 & b \\ 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}, \quad B_1 = \begin{pmatrix} b & 0 \\ 1 & 0 \end{pmatrix}.$$

It follows that

$$B_0 A_1 B_1 = \begin{pmatrix} ab + ba & 0 \\ 0 & 0 \end{pmatrix}$$

which is selfadjoint because a and b are. By Proposition 4.2.2, for any $m \geq 1$ we have that

$$\begin{aligned}
\omega((ab + ba)^m) &= \operatorname{Tr}_2 \otimes \omega((B_0 A_1 B_1)^m) \\
&= \operatorname{Tr}_2 \otimes \omega((A_1 B_1 B_0)^m) \\
&= \operatorname{Tr}_2 \otimes \omega\left(\left(\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} \tau(b) & \tau(b^2) \\ 1 & \tau(b) \end{pmatrix}\right)^m\right) \\
&= \operatorname{Tr}_2 \otimes \omega\left(a^m \begin{pmatrix} \tau(b) & \tau(b^2) \\ 1 & \tau(b) \end{pmatrix}^m\right),
\end{aligned}$$

If we have a matrix $\begin{pmatrix} x & y \\ 1 & x \end{pmatrix}$, by diagonalizing we have that

$$\operatorname{Tr}_2\left(\begin{pmatrix} x & y \\ 1 & x \end{pmatrix}^m\right) = (x + \sqrt{y})^m + (x - \sqrt{y})^m.$$

Then

$$\begin{aligned}
\omega((ab+ba)^m) &= \mathrm{Tr}_2 \otimes \omega \left(a^m \begin{pmatrix} \tau(b) & \tau(b^2) \\ 1 & \tau(b) \end{pmatrix}^m \right) \\
&= \mathrm{Tr}_2 \otimes \omega \left(a^m \begin{pmatrix} p^m & 0 \\ 0 & q^m \end{pmatrix} \right) \\
&= \omega \left(a^m \left(\tau(b) + \sqrt{\tau(b^2)}\right)^m + a^m \left(\tau(b) - \sqrt{\tau(b)^2}\right)^m \right) \\
&= \omega\left((pa)^m + (qa)^m\right),
\end{aligned}$$

where p and q are defined as in the statement of Theorem 4.2.1. Again by Corollary 1.3.7, we conclude that

$$\mathrm{EV}(ab+ba) = p\mathrm{EV}(a) \sqcup q\mathrm{EV}(a).$$

4) Proceeding in an analogous way of 3), defining the matrices

$$B_0 = \begin{pmatrix} \mathrm{i} & -\mathrm{i}\, b \\ 0 & 0 \end{pmatrix}, \quad A_1 = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}, \quad B_1 = \begin{pmatrix} b & 0 \\ 1 & 0 \end{pmatrix}$$

it follows that

$$B_0 A_1 B_1 = \begin{pmatrix} \mathrm{i}(ab-ba) & 0 \\ 0 & 0 \end{pmatrix}$$

is selfadjoint. For any $m \geq 1$, we have that

$$\begin{aligned}
\omega(((\mathrm{i}(ab+ba))^m) &= \mathrm{Tr}_2 \otimes \omega((B_0 A_1 B_1)^m) \\
&= \mathrm{Tr}_2 \otimes \omega((A_1 B_1 B_0)^m) \\
&= \mathrm{Tr}_2 \otimes \omega \left(\left(\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} \mathrm{i}\,\tau(b) & -\mathrm{i}\,\tau(b^2) \\ \mathrm{i} & -\mathrm{i}\,\tau(b) \end{pmatrix} \right)^m \right) \\
&= \mathrm{Tr}_2 \otimes \omega \left(a^m \begin{pmatrix} \mathrm{i}\,\tau(b) & -\mathrm{i}\,\tau(b^2) \\ \mathrm{i} & -\mathrm{i}\,\tau(b) \end{pmatrix}^m \right),
\end{aligned}$$

By diagonalizing, we have that

$$\mathrm{Tr}_2 \left(\begin{pmatrix} x & y \\ 1 & -x \end{pmatrix}^m \right) = \left(\sqrt{y+x^2}\right)^m + \left(-\sqrt{y+x^2}\right)^m.$$

Hence

$$\omega((\mathrm{i}(ab-ba))^m) = \mathrm{Tr}_2 \otimes \omega \left(a^m \begin{pmatrix} \mathrm{i}\,\tau(b) & -\mathrm{i}\,\tau(b^2) \\ \mathrm{i} & -\mathrm{i}\,\tau(b) \end{pmatrix}^m \right)$$

$$
\begin{aligned}
&= \mathrm{Tr}_2 \otimes \omega \left(a^m \begin{pmatrix} r^m & 0 \\ 0 & (-r)^m \end{pmatrix} \right) \\
&= \omega \left(a^m \left(\sqrt{\tau(b^2) - \tau(b)^2} \right)^m + a^m \left(-\sqrt{\tau(b^2) - \tau(b)^2} \right)^m \right) \\
&= \omega \left((ra)^m + (-ra)^m \right),
\end{aligned}
$$

where r is defined as in the statement of Theorem 4.2.1. We conclude that

$$\mathrm{EV}(\mathrm{i}(ab - ba)) = r\mathrm{EV}(a) \sqcup (-r)\mathrm{EV}(a).$$

■

Remark 4.2.4. We recall the parts 3 and 4 of Theorem 4.2.1 where it is considered only one element a. In these cases, it was possible to obtain explicit formulas from the fact that a commutes with itself and then we can compute the trace of powers of a matrix by adding powers of the eigenvalues of a matrix of some moments of the b's. The result can be expressed as the disjoint union of the eigenvalues of $\lambda_i a$, where the λ_i are the eigenvalues of $\mathrm{id} \otimes \tau\ (B)$. One can generalize this in the following proposition.

Proposition 4.2.5. *Let $(\mathcal{A}, \omega, \tau)$ be a non-commutative probability space with tracial weight ω. Consider $a \in \mathrm{D}(\omega)$ and $b_1, \ldots, b_k, c_1, \ldots, c_k \in \mathcal{A}$ such that a is trace class with respect to ω and $(\{a\}, \{b_1, c_1, \ldots, b_k, c_k\})$ is cyclically monotone with respect to (ω, τ). Assume that $a, b_1, c_1, \ldots, b_k, c_k$ are selfadjoint, $\sum_{i=1}^k b_i a c_i$ is selfadjoint and $B' = (\tau(c_i b_j))_{i,j=1}^k$. If $\lambda_1, \ldots, \lambda_k$ are the k eigenvalues of B' counting multiplicity, then*

$$\mathrm{EV}\left(\sum_{i=1}^k b_i a c_i \right) = \bigsqcup_{i=1}^k \mathrm{EV}(\lambda_i a). \tag{4.11}$$

Proof. As in the proof of Theorem 4.2.1, we define the matrices in $M_k(\mathcal{A})$

$$B = \begin{pmatrix} b_1 & b_2 & \cdots & b_k \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad C = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ c_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_k & 0 & \cdots & 0 \end{pmatrix}, \quad A = \mathrm{diag}(a, \ldots, a).$$

Then $BAC = \mathrm{diag}\left(\sum_{i=1}^k b_i a c_k, 0, \ldots, 0 \right)$ which is selfadjoint. Proceeding as in the latter proof, we have that for $m \geq 1$

$$\omega\left(\left(\sum_{i=1}^k b_i a c_i \right)^m \right) \quad = \quad \mathrm{Tr}_k \otimes \omega((BAC)^m)$$

$$\begin{aligned} &= \mathrm{Tr}_k \otimes \omega((ACB)^m) \\ &= \mathrm{Tr}_k \otimes \omega\left((aB')^m\right) \\ &= \omega\left(\sum_{i=1}^{k} (\lambda_i a)^m\right), \end{aligned}$$

where we apply Proposition 4.2.2 in the third equality and use that $\mathrm{Tr}_k(X^m)$ is the sum of the m-powers of the eigenvalues of X. Hence, it follows that

$$\mathrm{EV}\left(\sum_{i=1}^{k} b_i a c_i\right) = \bigsqcup_{i=1}^{k} \lambda_i \mathrm{EV}(a)$$

by applying Corollary 1.3.7. ∎

Remark 4.2.6. The procedure described in the proofs of the above theorems can be applied to any selfadjoint $*$-polynomial that can be written as the entry (1,1) of a product of matrices $A_1 B_1 \cdots A_k B_k$ as in Proposition 4.2.2, and the rest of the entries are zero. However, the same trick is no longer possible in some polynomials where the number of elements a_i is not the same on each monomial. For instance, consider the polynomial $a + babab$, where $(\{a\},\{b\})$ is cyclically monotone. If we would want to write this polynomial as a product of matrices as in the above proofs, we would have to do the following factorization

$$\begin{pmatrix} 1 & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix} = \begin{pmatrix} a + babab & 0 \\ 0 & 0 \end{pmatrix}.$$

However, a matrix with elements in $\mathrm{alg}(a)$ must not contain any constant. A solution for this issue is considering the decomposition

$$\begin{pmatrix} 1 & b \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & aba \end{pmatrix} \begin{pmatrix} 1 & 0 \\ b & 0 \end{pmatrix} = \begin{pmatrix} a + babab & 0 \\ 0 & 0 \end{pmatrix}.$$

In order to obtain a computation of the moments of the polynomial $a + babab$, we can use Theorem 4.2.1 providing that the new element aba is compatible with the cyclic monotone independence of $(\{a\},\{b\})$. Indeed, we have a more general result.

Proposition 4.2.7. *Let $(\mathcal{A}, \omega, \tau)$ be a non-commutative probability space with tracial weight ω. Let $a_1, \dots, a_k \in \mathrm{D}(\omega)$ and $b_1, \dots, b_k, c_1, \dots, c_k \in \mathcal{A}$. If $(\{a_1, \dots, a_k\}, \{b_1, c_1, \dots, b_k, c_k\})$ is cyclically monotone, then $(\{a_1 c_1 a_1^*, \dots, a_k c_k a_k^*\}, \{b_1, c_1, \dots, b_k, c_k\})$ is cyclically monotone.*

Proof. We have to show that if $x_1, \dots, x_n \in \mathrm{alg}(\{a_1 c_1 a_1^*, \dots, a_k c_k a_k^*\})$ and $y_1, \dots, y_n \in$

$\mathrm{alg}(\{1, b_1, c_1, \ldots, b_k, c_k\})$, then

$$\omega(x_1 y_1 \cdots x_n y_n) = \omega(x_1 \cdots x_n)\tau(y_1)\cdots\tau(y_n).$$

For notational convenience, we shall prove the result for the case $n = 2$. The general case is done in a similar way. Consider the elements

$$\begin{aligned} x_1 &= (a_{i_1} c_{i_1} a_{i_1}^*) \cdots (a_{i_r} c_{i_r} a_{i_r}^*), \\ x_2 &= (a_{j_1} c_{j_1} a_{j_1}^*) \cdots (a_{j_s} c_{j_s} a_{j_s}^*), \end{aligned}$$

for some $1 \leq i_1, \ldots, i_r, j_1, \ldots, j_s \leq k$. If $y_1, y_2 \in \mathrm{alg}(\{1, b_1, c_1, \ldots, b_k, c_k\})$, by cyclic monotone independence, we have that

$$\begin{aligned} \omega(x_i y_1 x_2 y_2) &= \omega\left((a_{i_1} c_{i_1} a_{i_1}^*) \cdots (a_{i_r} c_{i_r} a_{i_r}^*) y_1 (a_{j_1} c_{j_1} a_{j_1}^*) \cdots (a_{j_s} c_{j_s} a_{j_s}^*) y_2\right) \\ &= \omega(a_{i_1} a_{i_1}^* \cdots a_{i_r} a_{i_r}^* a_{j_1} a_{j_1}^* \cdots a_{j_s} a_{j_s}^*)\tau(c_{i_1})\cdots\tau(c_{i_r})\tau(y_1)\tau(c_{j_1})\cdots\tau(c_{j_s})\tau(y_2) \end{aligned}$$

On the other hand

$$\begin{aligned} \omega(x_1 x_2)\tau(y_1)\tau(y_2) &= \omega\left((a_{i_1} c_{i_1} a_{i_1}^*) \cdots (a_{i_r} c_{i_r} a_{i_r}^*)(a_{j_1} c_{j_1} a_{j_1}^*) \cdots (a_{j_s} c_{j_s} a_{j_s}^*)\right)\tau(y_1)\tau(y_2) \\ &= \omega(a_{i_1} a_{i_1}^* \cdots a_{i_r} a_{i_r}^* a_{j_1} a_{j_1}^* \cdots a_{j_s} a_{j_s}^*)\left(\prod_{\ell=1}^{r} \tau(c_{i_\ell})\right)\left(\prod_{\ell=1}^{s} \tau(c_{j_\ell})\right)\tau(y_1)\tau(y_2) \end{aligned}$$

We finish the proof by comparing the above equations. ■

In particular, the above proposition allows to get a formula for another kind of polynomial.

Corollary 4.2.8. *Let $(\mathcal{A}, \omega, \tau)$ be a non-commutative probability space with tracial weight ω. Consider $a_1, \ldots, a_k \in \mathrm{D}(\omega)$ and $b_1, \ldots, b_k, c_1, \ldots, c_k \in \mathcal{A}$ such that $(a_1, \ldots, a_k)$ are trace class with respect to ω and $(\{a_1, \ldots, a_k\}, \{b_1, c_1, \ldots, b_k, c_k\})$ is cyclically monotone with respect to (ω, τ). If $c_1, \ldots, c_k$ are selfadjoint and $B = (\tau(b_i^* b_j))_{i,j=1}^k \in M_k(\mathbb{C})$, then*

$$\mathrm{EV}\left(\sum_{i=1}^{k} b_i a_i c_i a_i^* b_i^*\right) = \mathrm{EV}\left(\sqrt{B}\,\mathrm{diag}(d_1, \ldots, d_k)\sqrt{B}\right),$$

where $d_i = a_i c_i a_i^*$ for $i = 1, \ldots, k$ and $\sqrt{B}\,\mathrm{diag}(d_1, \ldots, d_k)\sqrt{B} \in (M_k(\mathbb{C}) \otimes \mathcal{A}, \mathrm{Tr}_k \otimes \omega)$.

Proof. By Proposition 4.2.7, we have that $(\{d_1, \ldots, d_k\}, \{b_1, \ldots, b_k\})$ is cyclically monotone. Since c_i es selfadjoint, then d_i is also selfadjoint, for $i = 1, \ldots, k$. We conclude by applying Theorem 4.2.1. ■

Chapter 5

Random Matrices and Cyclic Monotone Independence

In this chapter, we shall establish the connection between spectral theory of random matrices and the recently presented theory of cyclic monotone independence. This will be held by proving that Haar invariant random matrices with limiting compact distribution and random matrices which have limiting distribution with respect to the normalized trace tr_n are asymptotically cyclically monotone. The main tool which will be used is the Weingarten calculus which has already been defined in Chapter 3.

5.1 The Weingarten Formula

The goal of this short section is to derive a formula for computing the moments of certain products of deterministic matrices and Haar unitary random matrices via the Weingarten calculus.

First, we introduce some notation. If I is a finite set, let S_I the symmetric group acting on I. In particular, we simply denote S_n as the symmetric group on $\{1, \ldots, n\}$. Let $U = (U_{ij})_{i,j=1}^n$ be a Haar unitary random matrix. We can consider $\mathcal{E}$ a linear transformation on $M_n(\mathbb{C})^{\otimes k}$ that is defined by the equation

$$\mathcal{E}(A) = \mathbb{E}\left(U^{\otimes k} A (U^*)^{\otimes k}\right), \qquad \forall\, A \in M_n(\mathbb{C})^{\otimes k}. \tag{5.1}$$

On the other hand, we can also define $\Phi : M_n(\mathbb{C})^{\otimes k} \to \mathbb{C}[S_k]$ by

$$\Phi(A) = \sum_{\sigma \in S_k} \mathrm{Tr}_{M_n(\mathbb{C})^{\otimes k}} \left(\rho(\sigma)^* A\right) \delta_\sigma, \tag{5.2}$$

where $\{\delta_\sigma\}_{\sigma\in S_k}$ is the canonical basis in $\mathbb{C}[S_k]$ and $\rho : S_k \to M_n(\mathbb{C})^{\otimes k}$ is the representation given by

$$\rho(\sigma)(v_1 \otimes \cdots \otimes v_k) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}, \qquad \text{where } \ v_i \in \mathbb{C}^n, \ i = 1, \dots, k. \tag{5.3}$$

We recall the Weingarten function Wg defined by (3.11). It is proved in [9] that the Weingarten function is related with the map Φ by the equation

$$\Phi(A\mathcal{E}B) = \Phi(A)\Phi(B)\,\mathrm{Wg}, \qquad \forall\, A, B \in M_n(\mathbb{C})^{\otimes k}, \tag{5.4}$$

where

$$\mathrm{Wg} = \sum_{\sigma\in S_k} \mathrm{Wg}(\sigma, n)\delta_\sigma.$$

For our purposes, it will be enough to consider the cases of pure tensors $A = A_1\otimes\cdots\otimes A_k$, $B = B_1 \otimes\cdots\otimes B_k$, where $A_i, B_i \in M_n(\mathbb{C})$, $i = 1, \dots, k$. For this class of elements, we define a kind of multiplicative extension of the trace for every $\sigma \in S_k$. If $c = (i_1 i_2 \cdots i_m)$ is a cycle, let $A_c = A_{i_1} \cdots A_{i_m}$ and if $\sigma = c_1 \cdots c_{\ell(\sigma)}$ is the cycle decomposition of σ, then

$$\mathrm{Tr}_\sigma(A_1, \dots, A_k) = \prod_{i=1}^{\ell(\sigma)} \mathrm{Tr}_n\left(A_{c_i}\right). \tag{5.5}$$

With this notation, it turns out that

$$\mathrm{Tr}_{M_n(\mathbb{C})^{\otimes k}}(\rho(\sigma)^* A) = \mathrm{Tr}_\sigma(A_1, \dots, A_k), \qquad \forall\, \sigma \in S_k. \tag{5.6}$$

We provide a proof of the above formula in the appendix of this work. Finally, combining (5.4) and (5.6), and taking the coefficient of δ_σ, we get the main formula that will be used in the next proofs:

$$\mathbb{E}(\mathrm{Tr}_\sigma(A_1UB_1U^*, \dots, A_kUB_kU^*)) = \sum_{\substack{\sigma_1,\sigma_2,\sigma_3\in S_k\\ \sigma_1\sigma_2\sigma_3=\sigma}} \mathrm{Tr}_{\sigma_1}(A_1, \dots, A_k)\,\mathrm{Tr}_{\sigma_2}(B_1, \dots, B_k)\,\mathrm{Wg}(\sigma_3, n). \tag{5.7}$$

5.2 Asymptotic Cyclic Monotone Independence of Random Matrices

We are ready to state and prove the first main result of the asymptotic cyclic monotone independence of random matrices on average. As we have already mentioned, the main tool for the proof will be the formula (5.7).

Theorem 5.2.1. *Let $n \in \mathbb{N}$. Let $U = U(n)$ be an $n \times n$ Haar unitary random matrix and $A_i = A_i(n)$, $B_j = B_j(n)$, $i = 1, \ldots, k$, $j = 1, \ldots, \ell$ be $n \times n$ random matrices such that*

1. *$((A_1, \ldots, A_k), \mathbb{E} \otimes \mathrm{Tr}_n)$ converges in distribution to a k-tuple of trace class operators as $n \to \infty$.*

2. *$((B_1, \ldots, B_\ell), \mathbb{E} \otimes \mathrm{tr}_n)$ converges in distribution to an ℓ-tuple of elements in a non-commutative probability space as $n \to \infty$.*

3. *$\{A_1, \ldots, A_k\}, \{B_1, \ldots, B_\ell\}, U$ are independent.*

Then the pair $(\{A_1, \ldots, A_k\}, \{UB_1U^*, \ldots, UB_\ell U^*\})$ is asymptotically cyclically monotone with respect to $(\mathbb{E} \otimes \mathrm{Tr}_n, \mathbb{E} \otimes \mathrm{tr}_n)$.

Proof. We shall make some assumptions. First, we can take $k = \ell$ since it is possible to add zero matrices to the A_i's or identity matrices to the B_j's and having the same distribution. In order to simplify the notation, we shall write B_j instead of UB_jU^*. Then it will be enough to prove that the

$$\lim_{n\to\infty} \mathbb{E}\left(\mathrm{Tr}_n(A_1B_1 \cdots A_kB_k)\right) = \lim_{n\to\infty} \mathbb{E}(\mathrm{Tr}_n(A_1 \cdots A_k)) \prod_{i=1}^{k} \lim_{n\to\infty} \mathbb{E}(\mathrm{tr}_n(B_i)). \tag{5.8}$$

This is because of each $*$-polynomial P in $A_1, \ldots, A_k, B_1, \ldots, B_k$ is sum of monomials of the form $B_0'A_1'B_1' \cdots A_m'B_m'$, where $A_i' \in \mathrm{alg}\{A_1, \ldots, A_k\}, B_i' \in \mathrm{alg}\{I_n, B_1, \ldots, B_k\}$. Moreover, since we have traciality, the distribution of $B_0'A_1'B_1' \cdots A_m'B_m'$ respect to Tr_n is the same that of $A_1'B_1' \cdots A_m'B_m'B_0'$. Then, once we have the desired factorization, it is possible to give the non-commutative probability space with a tracial weight according to the limit of this factorization in order to satisfy the rule of cyclic monotone independence of the limiting random variables. This is done via the cyclic monotone product given in the Definition 4.1.9.

Let Z be the cycle $Z = (1 \cdots k) \in S_k$. Recalling that $\{A_i\}_{i=1}^k, \{B_i\}_{i=1}^k, U$ are independent random matrices, we have that:

$$\mathbb{E}(\mathrm{Tr}_n(A_1B_1 \cdots A_kB_k)) = \sum_{\substack{\sigma_1,\sigma_2,\sigma_3 \in S_k \\ \sigma_1\sigma_2\sigma_3 = Z}} \mathbb{E}(\mathrm{Tr}_{\sigma_1}(A_1, \ldots, A_k))\mathbb{E}(\mathrm{Tr}_{\sigma_2}(B_1, \ldots, B_k))\,\mathrm{Wg}(\sigma_3, n). \tag{5.9}$$

We proceed to analyze the later equation in order to find out which terms in the sum have a non-zero contribution in the limit as $n \to \infty$. By assumption, we have that $(A_1, \ldots, A_k)$ converges in distribution, and then $\mathrm{Tr}_{\sigma_1}(A_1, \ldots, A_k)$ is $O(1)$. In the same way, we have that $\mathrm{Tr}_{\sigma_2}(B_1, \ldots, B_k)$ is $O(n^{k-|\sigma_2|})$ since $(B_1, \ldots, B_k)$ converges in distribution with respect to the normalized trace. Here, $|\sigma|$ denotes the minimal number of transpositions to express σ as a product of them, and hence $\ell(\sigma) = k - |\sigma|$. On the other hand, from the asymptotic behavior of the Weingarten function given by (3.13), we also have that $\mathrm{Wg}(\sigma_3, n)$ is $O(n^{-k-|\sigma_3|})$. Hence, each term in the sum (5.9) is $O(n^{-|\sigma_2|-|\sigma_3|})$. From this, we have that the terms in the sum of (5.9) which have a non-zero contribution in the limit are those such that $|\sigma_2| = |\sigma_3| = 0$, which means that $\sigma_2 = \sigma_3 = 1_k$. Recalling that $\sigma_1\sigma_2\sigma_3 = Z$, we conclude that $\sigma_1 = Z$. Finally, using the definition of Tr_σ and the fact that $\mathrm{Wg}(1_k, k) = 1$, we conclude that

$$\mathbb{E}(\mathrm{Tr}_n(A_1B_1\cdots A_kB_k)) = \mathbb{E}(\mathrm{Tr}_n(A_1A_2\cdots A_k))\prod_{i=1}^{k}\mathbb{E}(\mathrm{tr}_n(B_i)) + O(n^{-1}), \tag{5.10}$$

that is what we wanted to prove. ■

Remark 5.2.2. In the previous theorem, it was considered the situation that the random matrices B_j, which converges in distribution in the sense of non-commutative probability, were conjugated by the Haar unitary random matrix U. However, the same proof works if we take the case that the random matrices A_i are conjugated by the Haar unitary. In other words, in the conclusion of the above theorem, we can replace the pair $(\{A_1, \ldots, A_k\}, \{UB_1U^*, \ldots, UB_\ell U^*\})$ with $(\{UA_1U^*, \ldots, UA_kU^*\}, \{B_1, \ldots, B_\ell\})$, and the result is also true because of traciality.

The second main result is about the almost sure convergence of the traces of the random matrix models discussed in the above theorem. It will be necessary to prove the next technical lemma.

Lemma 5.2.3. *Let $n \in \mathbb{N}$. Let $U = U(n)$ be an $n \times n$ Haar unitary random matrix and $A_i = A_i(n)$, $B_j = B_j(n)$, $i = 1, \ldots, k$, $j = 1, \ldots, k$ be $n \times n$ deterministic matrices such that*

1. *$((A_1, \ldots, A_k), \mathrm{Tr}_n)$ converges in distribution to a k-tuple of trace class operators as $n \to \infty$.*

2. *$((B_1, \ldots, B_k), \mathrm{tr}_n)$ converges in distribution to a k-tuple of elements in a non-commutative probability space as $n \to \infty$.*

Then

$$\mathbb{E}\left(|\mathrm{Tr}_n(A_1UB_1U^*\cdots A_kUB_kU^*) - \mathbb{E}\left(\mathrm{Tr}_n(A_1UB_1U^*\cdots A_kUB_kU^*)\right)|^4\right) = O(n^{-2}). \tag{5.11}$$

Proof. As the same way that we did in the last proof, we denote B_i instead of UB_iU^*. In order to prove the lemma, we shall prove the next generalization: if we take $4k$ matrices of each type instead of k, and let

$$\begin{aligned} X_i &= \mathrm{Tr}_n \left(A_{(i-1)k+1} B_{(i-1)k+1} \cdots A_{(i-1)k+k} B_{(i-1)k+k} \right), \quad i = 1,2,3,4, \\ \hat{X}_i &= X_i - \mathbb{E}(X_i), \end{aligned}$$

then $\mathbb{E}\left(\hat{X}_1\hat{X}_2\hat{X}_3\hat{X}_4\right) = O(n^{-2})$. Once we prove this, the lemma follows taking $X_3 = X_1$ and $X_4 = X_2 = \overline{X_1}$ since

$$\overline{X_1} = \mathrm{Tr}_n \left((A_1 B_1 \cdots B_{k-1} A_k B_k)^* \right) = \mathrm{Tr}_n \left(A_k^* B_{k-1}^* \cdots B_1^* A_1^* B_k^* \right).$$

We introduce some notation. For $i = 1,2,3,4$, let

$$\begin{aligned} I_i &= \{(i-1)k+1, (i-1)k+2, \dots, (i-1)k+k\}, \\ Z_i &= ((i-1)k+1(i-1)k+2, \dots, (i-1)k+k) \in S_{I_i}, \\ Y &= Z_1 Z_2 Z_3 Z_4 \in S_{4k}. \end{aligned}$$

Using linearity of expectation, when we expand the moment $\mathbb{E}\left(\hat{X}_1\hat{X}_2\hat{X}_3\hat{X}_4\right)$ we have a sum of 16 terms

$$\mathbb{E}\left(\hat{X}_1\hat{X}_2\hat{X}_3\hat{X}_4\right) = \sum_{A \subset \{1,2,3,4\}} E_A, \tag{5.12}$$

where

$$E_A = (-1)^{|A|} \mathbb{E}\left(\prod_{i \in A} X_i \right) \prod_{i \in A^c} \mathbb{E}(X_i). \tag{5.13}$$

Each of the above expectations can be computed using (5.7) and the definition of Tr_σ. Hence, E_A can be written as

$$E_A = \sum_{\substack{\sigma_1,\sigma_2,\sigma_3 \in S_{4k} \\ \sigma_1\sigma_2\sigma_3 = Y}} \mathrm{Tr}_{\sigma_1}(A_1, \dots, A_{4k}) \, \mathrm{Tr}_{\sigma_2}(B_1, \dots, B_{4k}) f_A(\sigma_1, \sigma_2, \sigma_3, n), \tag{5.14}$$

for some numbers $f_A(\sigma_1, \sigma_2, \sigma_3, n)$, $A \subset \{1,2,3,4\}$, which are equal to zero or are signed products of two, three or four Weingarten functions. We have then

$$\mathbb{E}\left(\hat{X}_1\hat{X}_2\hat{X}_3\hat{X}_4\right) = \sum_{\substack{\sigma_1,\sigma_2,\sigma_3 \in S_{4k} \\ \sigma_1\sigma_2\sigma_3 = Y}} \mathrm{Tr}_{\sigma_1}(A_1, \dots, A_{4k}) \, \mathrm{Tr}_{\sigma_2}(B_1, \dots, B_{4k}) f(\sigma_1, \sigma_2, \sigma_3, n), \tag{5.15}$$

where $f(\sigma_1, \sigma_2, \sigma_3, n) = \sum_{A \subset \{1,2,3,4\}} f_A(\sigma_1, \sigma_2, \sigma_3, n)$.

It is time to use the assumptions. We know that $\mathrm{Tr}_{\sigma_1}(A_1,\ldots,A_{4k}) = O(1)$ and

$$\mathrm{Tr}_{\sigma_2}(B_1,\ldots,B_{4k}) = O(n^{-\ell(\sigma_2)}) = O(n^{-|\sigma_2|+4k}).$$

Since $\mathrm{Wg}(\sigma_3,n) = O(n^{-4k-|\sigma_3|})$, we have that

$$\mathrm{Tr}_{\sigma_2}(B_1,\ldots,B_{4k})f(\sigma_1,\sigma_2,\sigma_3,n) = O(n^{-|\sigma_2|-|\sigma_3|}). \tag{5.16}$$

We have to show now that $\mathrm{Tr}_{\sigma_2}(B_1,\ldots,B_{4k})f(\sigma_1,\sigma_2,\sigma_3,n) = O(n^{-2})$, which is equivalent to show that $f(\sigma_1,\sigma_2,\sigma_3,n) = O(n^{-4k+|\sigma_2|-2})$.

Let $\sigma_1,\sigma_2,\sigma_3$ be fixed permutations in S_{4k} such that $\sigma_1\sigma_2\sigma_3 = Y$. We give an equivalence relation in $\{1,\ldots,4k\}$ as follows: $i \sim j$ if and only if there exists a $\tau \in \langle\sigma_1,\sigma_2,\sigma_3\rangle$, the subgroup of S_{4k} generated by $\sigma_1,\sigma_2,\sigma_3$, such that $\tau(i) = j$. Obviously, $Y \in \langle\sigma_1,\sigma_2,\sigma_3\rangle$, and then every I_i is a subset of some equivalence class of the relation. With this, a partition $\pi(\sigma_1,\sigma_2,\sigma_3) = \{V_1,\ldots,V_m\} \in \mathcal{P}(4)$ can be constructed such that the equivalence classes of the relation described before are exactly $\bigcup_{i\in V_1} I_i, \bigcup_{i\in V_2} I_i, \ldots, \bigcup_{i\in V_m} I_i$. On the other hand, for any $A \subset \{1,2,3,4\}$, let $\pi(A) = \{A, \{b\} : b \in A^c\}$ be a partition in $\mathcal{P}(4)$. Then, the only subsets A such that f_A possibly has a non-zero contribution to the sum in f are those which $\pi(A)$ are coarser than or equal to $\pi(\sigma_1,\sigma_2,\sigma_3)$, and $f_A = 0$ for the other subsets A. Now we have to analyze the order of $f(\sigma_1,\sigma_2,\sigma_3,n)$. We shall do this according the block type of $\pi(\sigma_1,\sigma_2,\sigma_3)$. We have the next cases:

(1) $\pi(\sigma_1,\sigma_2,\sigma_3) = \{1,2,3,4\}$. This means that there is only one equivalence class, and then the group $\langle\sigma_1,\sigma_2,\sigma_3\rangle$ acts transitively on $\{1,\ldots,4k\}$. The only A such that $\pi(A)$ is coarser than or equal to $\{1,2,3,4\}$ is $A = \{1,2,3,4\}$, and then $f_{\{1,2,3,4\}} = \mathrm{Wg}(\sigma_3,n)$. We note that $|\sigma_2| + |\sigma_3| \geq 2$. Indeed, if $|\sigma_2| = 0 = |\sigma_3|$, then $\sigma_2 = 1_{S_{4k}} = \sigma_3$, and since $\sigma_1\sigma_2\sigma_3 = Y$, then $\sigma_1 = Y$. We recall that Y is a product of four disjoint cycles. In particular, there is no $\tau \in \langle Y\rangle$ such that $\tau(1) = k+1$. Then it is not possible that there is only one equivalence class. In the same way, if $|\sigma_2|+|\sigma_3| = 1$, then σ_2 and σ_3 are the identity element and a transposition. Since $\sigma_1\sigma_2\sigma_3 = Y$, then σ_1 has at least three cycles in its cycle decomposition. This is not possible since $\langle\sigma_1,\sigma_2,\sigma_3\rangle$ acts transitively on $\{1,\ldots,4k\}$. We conclude that $|\sigma_2| + |\sigma_3| \geq 2$ and by (5.16), we have that

$$\mathrm{Tr}_{\sigma_2}(B_1,\ldots,B_{4k})f(\sigma_1,\sigma_2,\sigma_3,n) = O(n^{-2}).$$

(2) $\pi(\sigma_1,\sigma_2,\sigma_3)$ is a pair partition. Again, the only A such that f_A is nonzero is the set $A = \{1,2,3,4\}$. We have again that $f = \mathrm{Wg}(\sigma_3,n)$, $|\sigma_2| + |\sigma_3| \geq 2$ and

$$\mathrm{Tr}_{\sigma_2}(B_1,\ldots,B_{4k})f(\sigma_1,\sigma_2,\sigma_3,n) = O(n^{-2}).$$

(3) $\pi(\sigma_1,\sigma_2,\sigma_3)$ has two blocks, one of size 3 and the other of size 1, for instance $\{\{1,2,3\},\{4\}\}$. The subsets A for which f_A is not zero are $\{1,2,3,4\}$ and $\{1,2,3\}$. We know that $f_{\{1,2,3,4\}} = \mathrm{Wg}(\sigma_3,n)$. In order to compute $f_{\{1,2,3\}}$, we use (5.7) as follows

$$\begin{aligned}
E_{\{1,2,3\}} &= (-1)^3\mathbb{E}(X_1X_2X_3)\mathbb{E}(X_4)\\
&= -\left(\sum_{\substack{\sigma_1',\sigma_2',\sigma_3'\in S_{3k}\\ \sigma_1'\sigma_2'\sigma_3'=(1\cdots k)(k+1\cdots 2k)(2k+1\cdots 3k)}} \mathrm{Tr}_{\sigma_1'}(A_1,\ldots,A_{3k})\,\mathrm{Tr}_{\sigma_2'}(B_1,\ldots,B_{3k})\,\mathrm{Wg}(\sigma_3',n)\right)\\
&\quad\times\left(\sum_{\substack{\sigma_1'',\sigma_2'',\sigma_3''\in S_{I_4}\\ \sigma_1''\sigma_2''\sigma_3''=(3k+1\cdots 4k)}} \mathrm{Tr}_{\sigma_1''}(A_{3k+1},\ldots,A_{4k})\,\mathrm{Tr}_{\sigma_2''}(B_{3k+1},\ldots,B_{4k})\,\mathrm{Wg}(\sigma_3'',n)\right)\\
&= -\sum_{\substack{\sigma_1,\sigma_2,\sigma_3\in S_{4k}\\ \sigma_1\sigma_2\sigma_3=Y\\ \sigma_1,\sigma_2 \text{ leave } I_1\cup I_2\cup I_3,\\ \text{and } I_4 \text{ invariant}}} \mathrm{Tr}_{\sigma_1}(A_1,\ldots,A_{4k})\,\mathrm{Tr}_{\sigma_2}(B_1,\ldots,B_{4k})\,\mathrm{Wg}(\sigma_3|_{I_1\cup I_2\cup I_3})\mathrm{Wg}(\sigma_3|_{I_4})\\
&= \sum_{\substack{\sigma_1,\sigma_2,\sigma_3\in S_{4k}\\ \sigma_1\sigma_2\sigma_3=Y}} \mathrm{Tr}_{\sigma_1}(A_1,\ldots,A_{4k})\,\mathrm{Tr}_{\sigma_2}(B_1,\ldots,B_{4k}) f_A(\sigma_1,\sigma_2,\sigma_3,n),
\end{aligned}$$

where the permutations σ_i are formed by the product of σ_i' and σ_i'', $i=1,2,3$. We have that $f_{\{1,2,3\}} = -\mathrm{Wg}(\sigma_3|_{I_1\cup I_2\cup I_3},n)\mathrm{Wg}(\sigma_3|_{I_4},n)$. On the other hand, by Lemma 3.3.8, we can write

$$\mathrm{Wg}(\sigma,n) = n^{-4k-|\sigma_3|}\left(\mu(\sigma_3)+O(n^{-2})\right), \tag{5.17}$$

where $\mu(\sigma_3) = \prod_{i=1}^{\ell(\sigma_3)}(-1)^{|c_i|}C_{|c_i|}$, and $\sigma_3 = c_1\cdots c_{\ell(\sigma_3)}$ is the cycle decomposition of σ_3. We have that μ is multiplicative and hence

$$\begin{aligned}
f &= f_{\{1,2,3\}} + f_{\{1,2,3,4\}}\\
&= -\mathrm{Wg}(\sigma_3|_{I_1\cup I_2\cup I_3},n)\mathrm{Wg}(\sigma_3|_{I_4},n) + \mathrm{Wg}(\sigma_3,n)\\
&= n^{-4k-|\sigma_3|}\left(-\mu(\sigma_3|_{I_1\cup I_2\cup I_3})\mu(\sigma_3|_{I_4}) + \mu(\sigma_3) + O(n^{-2})\right)\\
&= n^{-4k-|\sigma_3|}\left(-\mu(\sigma_3)+\mu(\sigma_3)+O(n^{-2})\right)\\
&= O(n^{-4k-|\sigma_3|-2})
\end{aligned}$$

where in the third equality we use (5.17) and in the fourth equality we use that μ is multiplicative. We can conclude that

$$\mathrm{Tr}_{\sigma_2}(B_1,\ldots,B_{4k}) f(\sigma_1,\sigma_2,\sigma_3,n) = O(n^{-|\sigma_2|-|\sigma_3|-2}).$$

(4) $\pi(\sigma_1, \sigma_2, \sigma_3)$ has three blocks, for instance $\{\{1,2\},\{3\},\{4\}\}$. In this case, the subsets A such that $\pi(A)$ is coarser than or equal to $\pi(\sigma_1,\sigma_2,\sigma_3)$ are $\{1,2,3,4\}, \{1,2,3\}, \{1,2,4\}, \{1,2\}$. Proceeding in a similar way that the latter case we have that

$$\begin{aligned} f &= f_{\{1,2\}} + f_{\{1,2,3\}} + f_{\{1,2,4\}} + f_{\{1,2,3,4\}} \\ &= \mathrm{Wg}(\sigma_3|_{I_1\cup I_2}, n)\mathrm{Wg}(\sigma_3|_{I_3}, n)\mathrm{Wg}(\sigma_3|_{I_4}), n) - \mathrm{Wg}(\sigma_3|_{I_1\cup I_2\cup I_3}, n)\mathrm{Wg}(\sigma_3|_{I_4}, n) \\ &\quad -\mathrm{Wg}(\sigma_3|_{I_1\cup I_2\cup I_4}, n)\mathrm{Wg}(\sigma_3|_{I_3}, n) + \mathrm{Wg}(\sigma_3, n) \\ &= n^{-4k-|\sigma_3|}\,[\mu(\sigma_3|_{I_1\cup I_2})\mu(\sigma_3|_{I_3})\mu(\sigma_3|_{I_4})) - \mu(\sigma_3|_{I_1\cup I_2\cup I_3})\mu(\sigma_3|_{I_4}) \\ &\quad -\mu(\sigma_3|_{I_1\cup I_2\cup I_4})\mu(\sigma_3|_{I_3}) + \mu(\sigma_3) + O(n^{-2})] \\ &= O(n^{-4k-|\sigma_3|-2}), \end{aligned}$$

and again $\mathrm{Tr}_{\sigma_2}(B_1,\ldots,B_{4k})f(\sigma_1,\sigma_2,\sigma_3,n) = O(n^{-|\sigma_2|-|\sigma_3|-2})$.

(5) $\pi(\sigma_1,\sigma_2,\sigma_3) = \{\{1\},\{2\},\{3\},\{4\}\}$. In this final case, the sixteen subsets of $\{1,2,3,4\}$ have a nonzero contribution to the sum in f. Using that μ is multiplicative:

$$f = \sum_{A\subset\{1,2,3,4\}} (-1)^{|A|} n^{-4k-|\sigma_3|} \left(\mu(\sigma_3|_{I_1})\mu(\sigma_3|_{I_2})\mu(\sigma_3|_{I_3})\mu(\sigma_3|_{I_4}) + O(n^{-2})\right) \tag{5.18}$$

We notice that the exactly half of the terms in the sum have positive sign. Hence the dominant terms in the sum get canceled and we conclude that $f = O(n^{-4k+|\sigma_3|-2})$.

In all the cases we proved that $f(\sigma_1,\sigma_2,\sigma_3) = O(n^{-4k+|\sigma_3|-2})$, that it was what we wanted to show and the proof is complete. ∎

The above lemma lead us to a short proof of the almost sure version of Theorem 5.2.1.

Theorem 5.2.4. *Let $n \in \mathbb{N}$. Let $U = U(n)$ be an $n\times n$ Haar unitary random matrix and $A_i = A_i(n)$, $B_j = B_j(n)$, $i = 1,\ldots,k$, $j = 1,\ldots,\ell$ be $n\times n$ deterministic matrices such that*

1. *$((A_1,\ldots,A_k),\mathrm{Tr}_n)$ converges in distribution to a k-tuple of trace class operators as $n\to\infty$.*

2. *$((B_1,\ldots,B_\ell),\mathrm{tr}_n)$ converges in distribution to an ℓ-tuple of elements in a non-commutative probability space as $n\to\infty$.*

Then the pair $(\{A_1,\ldots,A_k\},\{B_1,\ldots,B_\ell\})$ is asymptotically cyclically monotone almost surely with respect to $(\mathrm{Tr}_n,\mathrm{tr}_n)$.

Proof. We write B_i instead of UB_iU^*. Repeating the same arguments in the proof of

Theorem 5.2.1, it will be enough to prove the case $k = \ell$ and that

$$\lim_{n\to\infty} \mathrm{Tr}_n(A_1B_1\cdots A_kB_k) = \lim_{n\to\infty} \mathrm{Tr}_n(A_1\cdots A_k)\prod_{i=1}^{k}\lim_{n\to\infty} \mathrm{tr}_n(B_i) \qquad \text{almost surely.} \tag{5.19}$$

Applying Lemma 5.2.3 and monotone convergence theorem:

$$\mathbb{E}\left(\sum_{n=1}^{\infty} \left|\mathrm{Tr}_n(A_1B_1\cdots A_kUB_k) - \mathbb{E}\left(\mathrm{Tr}_n(A_1B_1\cdots A_kB_k)\right)\right|^4\right) < \infty. \tag{5.20}$$

Since a random variable with finite expectation must be finite almost surely, we have that

$$\sum_{n=1}^{\infty} \left|\mathrm{Tr}_n(A_1B_1\cdots A_kB_k) - \mathbb{E}\left(\mathrm{Tr}_n(A_1B_1\cdots A_kB_k)\right)\right|^4 < \infty \quad \text{a.s.} \tag{5.21}$$

and then $\lim_{n\to\infty} \left|\mathrm{Tr}_n(A_1B_1\cdots A_kB_k) - \mathbb{E}\left(\mathrm{Tr}_n(A_1B_1\cdots A_kB_k)\right)\right| = 0$ almost surely. Finally, by using Theorem 5.2.1, we get that

$$\begin{aligned}\lim_{n\to\infty} \mathrm{Tr}_n(A_1B_1\cdots A_kB_k) &= \lim_{n\to\infty} \mathbb{E}\left(\mathrm{Tr}_n(A_1B_1\cdots A_kB_k)\right) \\ &= \lim_{n\to\infty} \mathrm{Tr}_n(A_1\cdots A_k)\prod_{i=1}^{k}\lim_{n\to\infty} \mathrm{tr}_n(B_i)\end{aligned} \tag{5.22}$$

(we recall that $\mathrm{tr}_n(B_i) = \mathrm{tr}_n(UB_iU^*) = \mathrm{tr}_n(B_i)$ and that B_i's are deterministic matrices, so we can omit the symbol $\mathbb{E}$ in the right-hand side of (5.22)). ■

Remark 5.2.5. In Lemma 5.2.3 and Theorem 5.2.4, it is enough to assume that the matrices A_i, B_j are deterministic. However, this can be extended to the random case as we did in Corollary 3.3.5. This means that if A_i, B_j are random matrices such that $\{A_i\}, \{B_j\}, U$ are independent with U a Haar unitary such that $(A_1, \ldots, A_k)$ and $(B_1, \ldots, B_k)$ almost surely converge in distribution to deterministic elements, we can condition A_i, B_j to be constant and get the same results. For instance, we can take B_i to be the same matrix G in the Gaussian Unitary Ensemble because we know that G almost surely converges in distribution to a semicircular element. Also, we do not need to consider the Haar unitary because the GUE are rotationally invariant.

We can combine the last theorem with the results of convergence of eigenvalues studied in Chapter 1. In particular, we have the next result which will be useful when we are dealing with the convergence in eigenvalues in the compact framework.

Corollary 5.2.6. *We suppose the assumptions of Theorem 5.2.4. Then for any selfadjoint $*$-polynomial $P(x_1, \ldots, x_k, y_1, \ldots, y_\ell)$ such that $P(0, \ldots, 0, y_1, \ldots, y_k) = 0$, the Hermitian* *random matrix $P(A_1, \dots, A_k, UB_1U^*, \dots, UB_\ell U^*)$ converges in eigenvalues to a selfadjoint Hilbert-Schmidt operator almost surely. The limiting eigenvalues can be computed by the rule of cyclic monotone independence.*

Proof. The result follows by Proposition 1.3.8, where we notice that the fact that the limiting operator is Hilbert-Schmidt is because the A_i's are trace class, i.e. they are elements in $S^1 \subset S^2$, and Proposition 1.3.8 works with p an even integer. ∎

5.3 General Compact Case and Several Haar Unitaries Case

This section is devoted to study some generalizations of the results proved in the last section. The first one refers that compact operators with additional assumptions are asymptotic cyclically monotone independent. The proof of this statement relies on the fact that a compact operator can be approximated by trace class operators.

Theorem 5.3.1. *Let $n \in \mathbb{N}$. Let $U = U(n)$ be an $n \times n$ Haar unitary random matrix and $A_i = A_i(n)$, $B_j = B_j(n)$, $i = 1, \dots, k$, $j = 1, \dots, \ell$ be $n \times n$ deterministic matrices such that*

1. *$A_1, \dots, A_k$ are Hermitian,*
2. *$((A_1, \dots, A_k), \mathrm{Tr}_n)$ converges in compact distribution (Definition 1.3.4) to a k-tuple of compact operators $((a_1, \dots, a_k), \mathrm{Tr}_H)$ as $n \to \infty$,*
3. *$((B_1, \dots, B_\ell), \mathrm{tr}_n)$ converges in distribution to an ℓ-tuple of elements in a non-commutative probability space as $n \to \infty$,*
4. *$\sup_{n\in\mathbb{N}} \|B_i(n)\| < \infty$ for every $i = 1, \dots, \ell$, where $\|\cdot\|$ denotes the operator norm on $M_n(\mathbb{C})$.*

Let $P(x_1, \dots, x_k, y_1, \dots, y_\ell)$ be a selfadjoint $*$-polynomial with selfadjoint variables $x_1, \dots, x_k$ such that $P(0, \dots, 0, y_1, \dots, y_\ell) = 0$. Then $P(A_1, \dots, A_k, UB_1U^*, \dots, UB_\ell U^*)$ converges in eigenvalues to a deterministic compact operator almost surely.

Proof. The general idea of the proof is to consider a sequence of truncations of our operators in order to have trace class operators in which we will be able to apply the above results already proved. We shall choose the truncations such that they approximate the polynomial $P(A_1, \dots, A_k, B_1, \dots, B_\ell)$. In this way, the searched limiting eigenvalues will arise by taking the limit when n goes to infinity and also taking the limit in the sequence of truncated operators.

Once again, we can take $k = \ell$ and we simply write B_i instead of UB_iU^*. The assumption (4) is still valid since $\|B_p(n)\| = \|UB_p(n)U^*\|$ when U is unitary. By assumption

(2), A_i converges to a_i in eigenvalues for every $i = 1, \dots, k$. In particular we have that $\sup_n \|A_p(n)\| < \infty$ for any $p = 1, \dots, k$. Then, we can find a sequence $\{\epsilon_j\}_{j=1}^{\infty}$ of real numbers such that it is decreasing, converging to 0 and

$$\{\epsilon_j\}_{j=1}^{\infty} \cap \{\lambda_i^u(A_p(n)), \lim_{N\to\infty} |\lambda_i^u(A_p(N))| \,:\, n, i \geq 1,\, p = 1, \dots, k,\, u \in \{+,-\}\} = \emptyset,$$

i.e., ϵ_j is not any of the absolute value of the eigenvalues of A_i and a_i, for every $i = 1, \dots, k$. For the truncations of the operators, we shall consider continuous functions f_j on $\mathbb{R}$ such that

$$f_j(x) = \begin{cases} 0, & |x| < \epsilon_{j+1}, \\ x, & |x| > \epsilon_j, \end{cases} \tag{5.23}$$

and f_j is non-decreasing, for any $j \in \mathbb{N}$. We denote $A_p^{(j)} = f_j(A_p)$ and $a_p^{(j)} = f_j(a_p)$ for any $j \in \mathbb{N}, p = 1, \dots, k$. By spectral theorem and functional calculus (Theorem 3.4 in [14]), we have that $A_p^{(j)}$ and $a_p^{(j)}$ are trace class operators (they are finite rank operators). Also, by assumption (2) and directly from the definition of convergence in compact distribution, we have that $((A_1^{(j)}, \dots, A_k^{(j)}), \mathrm{Tr}_n)$ converges in compact distribution to $((a_1^{(j)}, \dots, a_k^{(j)}), \mathrm{Tr}_H)$ when $n \to \infty$ for any $j \in \mathbb{N}$. In this frame, we are able to apply Corollary 5.2.6 and deduce that for any $j \in \mathbb{N}$, the random eigenvalues of the polynomial $P(A_1^{(j)}, \dots, A_k^{(j)}, B_1, \dots, B_k)$ converges almost surely to the eigenvalues $\{\lambda_i^{(j)}\}_{i=1}^{\infty}$ of a deterministic selfadjoint Hilbert-Schmidt operator. We denote by $\{\lambda_i^{(j)}(n)\}_{i=1}^{\infty}$ the eigenvalues of $P(A_1^{(j)}, \dots, A_k^{(j)}, B_1, \dots, B_k)$. According to the form of f_j and using functional calculus, we have that

$$\sup_{n\in\mathbb{N}, 1\leq p\leq k} \|A_p^{(j)} - A_p\| \leq \epsilon_j, \qquad \forall\, j \in \mathbb{N}, \tag{5.24}$$

$$\sup_{n\in\mathbb{N}, 1\leq p\leq k} \|A_p^{(j)} - A_p^{(j')}\| \leq \epsilon_j, \quad \forall\, 1 \leq j \leq j'. \tag{5.25}$$

Because of $\sup_{n\in\mathbb{N}, 1\leq k\leq p} \|A_p(n)\| < \infty$ and (5.24), we get that $\sup_{n\in\mathbb{N}, 1\leq k\leq p} \|A_p^{(j)}(n)\| < \infty$ for any $j \geq 1$. The previous bounds and assumption (4) allow to prove that the following random variables converge to zero almost surely as $j \to \infty$:

$$\delta_j = \sup_{n\in\mathbb{N}} \left\| P(A_1, \dots, A_k, B_1, \dots, B_k) - P(A_1^{(j)}, \dots, A_k^{(j)}, B_1, \dots, B_k) \right\|, \tag{5.26}$$

$$\delta_{j,j'} = \sup_{n\in\mathbb{N}} \left\| P(A_1^{(j)}, \dots, A_k^{(j)}, B_1, \dots, B_k) - P(A_1^{(j')}, \dots, A_k^{(j')}, B_1, \dots, B_k) \right\|. \tag{5.27}$$

This finishes the first step of the proof. For the second step, we know that $\lambda_i^{(j)}(n) \to \lambda_i^{(j)}$ as $n \to \infty$ for any $i, j \geq 1$. Since $\delta_j \to 0$, we want to prove now that the eigenvalues of

$P(A_1, \ldots, A_k, B_1, \ldots, B_k)$ converges to the limit of $\lambda_i^{(j)}$ as $j \to \infty$ provided that it exists. Let $\{\lambda_i(n)\}_{i=1}^{\infty}$ be the eigenvalues of $P(A_1, \ldots, A_k, B_1, \ldots, B_k)$. From (5.24), (5.25) and a consequence of Weyl's inequality, it can be followed that

$$\left|\lambda_i^{\pm}(n) - (\lambda_i^{(j)})^{\pm}(n)\right| \leq \delta_j \text{ a.s.}, \quad \left|(\lambda_i^{(j')})^{\pm}(n) - (\lambda_i^{(j)})^{\pm}(n)\right| \leq \delta_{j,j'} \text{ a.s.}, \quad \forall\, i, j, n \in \mathbb{N}. \tag{5.28}$$

The existence of $\lim_{j\to\infty} \lambda_i^{(j)}$ for any $i \geq 1$ is guaranteed by the second inequality in (5.28) because of such inequality implies that $|(\lambda_i^{(j)})^{\pm} - (\lambda_i^{(j')})^{\pm}| \leq \delta_{j,j'}$ for $j \leq j'$. Then $\{\lambda_i^{(j)}\}_{j=1}^{\infty}$ is a Cauchy sequence for any $i \geq 1$ and the limit $\lambda_i^{\pm} = \lim_{j\to\infty}(\lambda_i^{(j)})^{\pm}$ exists for any $i \geq 1$. Hence, the almost sure convergence $\lambda_i^{\pm}(n) \to \lambda_i^{\pm}$ follows from taking $j \to \infty$ in the next inequalities:

$$\begin{aligned}\limsup_{n\to\infty} |\lambda_i^{\pm}(n) - \lambda_i^{\pm}| &\leq \limsup_{n\to\infty}\left|\lambda_i^{\pm}(n) - (\lambda_i^{(j)})^{\pm}(n)\right| + \limsup_{n\to\infty}\left|(\lambda_i^{(j)})^{\pm}(n) - (\lambda_i^{(j)})^{\pm}\right| \\ &\quad + \left|(\lambda_i^{(j)})^{\pm} - \lambda_i^{\pm}\right| \\ &\leq \delta_j + \left|(\lambda_i^{(j)})^{\pm} - \lambda_i^{\pm}\right| \quad \text{almost surely.}\end{aligned}$$

The final step of the proof consists in proving that the sequence $\{\lambda_i\}_{i=1}^{\infty}$ corresponds to the eigenvalues of a selfadjoint compact operator, i.e. we have to prove that $\lambda_i \to 0$ as $i \to \infty$. In order to prove that, we shall use the next result about variational properties of eigenvalues which can be proved using the minimax principle:

- Let $s_1(X) \geq s_2(X) \geq \cdots \geq 0$ be the singular values of a compact operator X. Then $s_i(XYZ) \leq \|X\|\|Z\| s_i(Y)$ and $s_{i+j-1}(X+Y) \leq s_i(X) + s_j(Y)$.

We also recall that if X is a selfadjoint operator, then $s_i(X) = |\lambda_i(X)|$. We know that each monomial in $P(A_1, \ldots, A_k, B_1, \ldots, B_k)$ has a factor A_p for some $1 \leq p \leq k$. We write the monomials as $X_r A_{p(r)} Y_r$, and then $P(A_1, \ldots, A_k, B_1, \ldots, B_k) = \sum_{r=1}^{m} X_r A_{p(r)} Y_r$. Using repeatedly (•), we have a bound for the $(mi - m + 1)$-th singular value of P:

$$\begin{aligned}s_{mi-m+1}(P) &\leq s_i(X_1 A_{p(1)} Y_1) + s_{(m-1)i-m+2}\left(\sum_{r=2}^{m} X_r A_{p(r)} Y_r\right) \\ &\leq \sum_{r=1}^{m} s_i(X_r A_{p(r)} Y_r) \\ &\leq \sum_{r=1}^{m} s_i(A_{p(r)}) \|X_r\| \|Y_r\|.\end{aligned}$$

Let $\epsilon > 0$. We recall that A_p converges in eigenvalues to a compact operator. Hence, there exists $i_0 \in \mathbb{N}$ such that $\sup_{n\in\mathbb{N}, 1\le p\le k} |\lambda_i(A_p(n))| \le \epsilon/m$, for $i \ge i_0$. On the other hand, we have seen that $\sup_{n\in\mathbb{N}} \|A_p(n)\| < \infty$ and $\sup_{n\in\mathbb{N}} \|B_p(n)\| < \infty$ by assumption. Then $\sup_{n\in\mathbb{N}}\{\|X_r\|, \|Y_r\| : 1 \le r \le m\} = K < \infty$ and

$$s_{mi-m+1}(P) \le K^2\epsilon.$$

Since the sequence of singular values is decreasing, we have that $\sup_{n\in\mathbb{N}} \|\lambda_i(n)\| = \sup_{n\in\mathbb{N}} s_i(P)$ converges to 0 a.s. as $i \to \infty$. Finally, we get the conclusion by noticing that

$$\lim_{i\to\infty} |\lambda_i| = \lim_{i\to\infty}\lim_{n\to\infty} |\lambda_i| \le \lim_{i\to\infty}\lim_{n\to\infty} |\lambda_i(n) - \lambda_i| + \lim_{i\to\infty}\sup_{n\in\mathbb{N}} |\lambda_i(n)| = 0. \tag{5.29}$$

■

The second generalization consists in not only to consider one Haar unitary random matrix but several of them. The related results can be proved by combining the previous theorems and the asymptotic freeness of independent Haar unitary random matrices described in Theorem 3.3.9. The first result is a generalization of Theorem 5.2.1 and can be stated in the following way.

Theorem 5.3.2. *Let $n \in \mathbb{N}$. Let $U_i = U_i(n)$ be $n \times n$ independent Haar unitary random matrices for $i = 1, \dots, k$ and $A_i = A_i(n)$, $B_{ij} = B_{ij}(n)$, $i, j = 1, \dots, k$, be $n \times n$ random matrices such that*

1. *$((A_1, \dots, A_k), \mathbb{E} \otimes \mathrm{Tr}_n)$ converges in distribution to a k-tuple of trace class operators as $n \to \infty$.*
2. *$((B_{i1}, \dots, B_{ik}), \mathbb{E} \otimes \mathrm{tr}_n)$ converges in distribution to an k-tuple of elements in a non-commutative probability space as $n \to \infty$, for each $i = 1, \dots, k$,*
3. *$\{A_1, \dots, A_k\}, \{B_{11}, B_{12}, \dots, B_{kk}\}, \{U_1, \dots, U_k\}$ are independent.*

Then the pair $(\{A_1, \dots, A_k\}, \{U_iB_{ij}U_i^*\}_{i,j=1}^k)$ is asymptotically cyclically monotone with respect to $(\mathbb{E} \otimes \mathrm{Tr}_n, \mathbb{E} \otimes \mathrm{tr}_n)$.

Proof. The proof is based in applying Theorem 5.2.1 to a family $\{C_j\}_{j=1}^{k^2}$ obtained by asymptotic freeness. Indeed, let U be a Haar unitary random matrix independent of $\{A_1, \dots, A_k\}$, $\{B_{11}, \dots, B_{kk}\}$ and $\{U_1, \dots, U_k\}$ and let $C_{ij} = U_iB_{ij}U_i^*$. By Theorem 3.3.9, we have that $\{C_{1j}\}_{j=1}^k, \dots, \{C_{kj}\}_{j=1}^k$ are asymptotically free with respect to $\mathbb{E} \otimes \mathrm{tr}_n$, and then we have that $((C_{11}, \dots, C_{kk}), \mathbb{E} \otimes \mathrm{tr}_n)$ converges in distribution to a k^2-tuple of elements in a non-commutative probability space as $n \to \infty$. With this new k^2-tuple, we can now apply

Theorem 5.2.1 in order to get that $(\{A_1, \ldots, A_k\}, \{UC_{11}U^*, \ldots, UC_{kk}U^*\})$ is asymptotically cyclically monotone with respect to $(\mathbb{E} \otimes \mathrm{Tr}_n, \mathbb{E} \otimes \mathrm{tr}_n)$. We finish the proof by noticing that by definition of Haar measure, UU_i has the same distribution as U_i for $1 \leq i \leq k$, and by independence, $(C_{11}, \ldots, C_{kk})$ has the same distribution as $(UC_{11}U^*, \ldots, UC_{kk}U^*)$. Therefore $(\{A_1, \ldots, A_k\}, \{C_{11}, \ldots, C_{kk}\})$ is asymptotically cyclically monotone with respect to $(\mathbb{E} \otimes \mathrm{Tr}_n, \mathbb{E} \otimes \mathrm{tr}_n)$. ■

Using the same ideas, we can get a version of Theorems 5.2.4, 5.3.1 and Corollary 5.2.6 when several Haar unitary random matrices are considered instead of only one. For instance, the statement of the generalized version of Corollary 5.2.6 can be written in the following way.

Corollary 5.3.3. *Let $n \in \mathbb{N}$. Let $U_i = U_i(n)$ be $n \times n$ independent Haar unitary random matrices for $i = 1, \ldots, k$ and $A_i = A_i(n)$, $B_{ij} = B_{ij}(n)$, $i, j = 1, \ldots, k$, be $n \times n$ deterministic matrices such that*

1. *$((A_1, \ldots, A_k), \mathrm{Tr}_n)$ converges in distribution to a k-tuple of trace class operators as $n \to \infty$.*
2. *$((B_{i1}, \ldots, B_{ik}), \mathrm{tr}_n)$ converges in distribution to an k-tuple of elements in a non-commutative probability space as $n \to \infty$, for each $i = 1, \ldots, k$.*

Then for any selfadjoint $*$-polynomial $P(x_1, \ldots, x_k, y_{11}, y_{12}, \ldots, y_{kk})$ such that

$$P(0, \ldots, 0, y_{11}, \ldots, y_{kk}) = 0,$$

the Hermitian random matrix

$$P(A_1, \ldots, A_k, U_1B_{11}U_1^*, U_1B_{12}U_1^*, \ldots, U_1B_{1k}U_1^*, U_2B_{21}U_2^*, \ldots, U_kB_{kk}U_k^*)$$

converges in eigenvalues to a selfadjoint Hilbert-Schmidt operator almost surely. The limiting eigenvalues can be computed by the asymptotic cyclic monotone independence of the pair $(\{A_i\}_{i=1}^k, \{U_iB_{ij}U_i^*\})_{i,j=1}^k)$ and the asymptotic freeness of $\{U_1B_{1j}U_1^*\}_{j=1}^k, \ldots, \{U_kB_{kj}U_k^*\}_{j=1}^k$.

In the context of several Haar unitary random matrices, we can investigate about asymptotic cyclic monotone independence when the A_i matrices are conjugated instead of the B_i matrices in a similar way that we saw in Remark 5.2.2. This problem has an interesting answer which reduces the set of polynomials with a non-zero asymptotic distribution.

Proposition 5.3.4. *Let $n \in \mathbb{N}$. Let $U_i = U_i(n)$ be $n \times n$ independent Haar unitary random matrices for $i = 1, \ldots, k$ and $A_i = A_i(n)$, $B_j = B_j(n)$, $i, j = 1, \ldots, k$, be $n \times n$ deterministic matrices such that*

1. $((A_1, \ldots, A_k), \mathrm{Tr}_n)$ *converges in distribution to a k-tuple of trace class operators as $n \to \infty$.*

2. $((B_1, \ldots, B_k), \mathrm{tr}_n)$ *converges in distribution to a k-tuple of elements in a non-commutative probability space as $n \to \infty$.*

Then for any numbers $1 \leq i_1, \ldots, i_k \leq k$ such that there exist $r, s \in \{1, \ldots, k\}$ with $i_r \neq i_s$, we have that $\mathrm{Tr}_n\left(U_{i_1} A_1 U_{i_1}^* B_1 \cdots U_{i_k} A_k U_{i_k}^* B_k\right)$ converges to zero almost surely.

Proof. Let $1 \leq i_1, \ldots, i_k \leq k$ be positive integers such that there exist r, s with $i_r \neq i_s$. Without loss of generality, we can assume that $i_r = 1$ and $i_s = 2$. It will be enough to show that the conditional expectation

$$\mathbb{E}\left(\left|\mathrm{Tr}_n(U_{i_1} A_1 U_{i_1}^* B_1 \cdots U_{i_k} A_k U_{i_k}^* B_k)\right|^2 |U_2, \ldots, U_k\right) = O(n^{-2}) \qquad \text{almost surely.} \tag{5.30}$$

Indeed if (5.30) holds, then by conditional monotone convergence theorem

$$\sum_{n=1}^{\infty} \left|\mathrm{Tr}_n(U_{i_1} A_1 U_{i_1}^* B_1 \cdots U_{i_k} A_k U_{i_k}^* B_k)\right|^2$$

is finite almost surely and so $\mathrm{Tr}_n(U_{i_1} A_1 U_{i_1}^* B_1 \cdots U_{i_k} A_k U_{i_k}^* B_k)$ converges to zero almost surely. For notation convenience, we shall denote the conditional expectation respect to $\sigma(U_2, \ldots, U_k)$ simply as $\mathbb{E}_{U_1}$.

First we notice that by properties of the trace, we can write

$$\begin{aligned} &\mathbb{E}_{U_1}\left(\left|\mathrm{Tr}_n(U_{i_1} A_1 U_{i_1}^* B_1 \cdots U_{i_k} A_k U_{i_k}^* B_k)\right|^2\right) \\ &= \quad \mathbb{E}_{U_1}\left(\mathrm{Tr}_n(C_1 U_1 D_1 U_1^* \cdots C_\ell U_1 D_\ell U_1^*) \,\mathrm{Tr}_n(C_{\ell+1} U_1 D_{\ell+1} U_1^* \cdots C_{2\ell} U_1 D_{2\ell} U_1^*)\right), \end{aligned}$$

where the D_j's are A_r or A_r^* for some $1 \leq r \leq k$ and the C_i's are products of B_p, B_p^* and the other $U_q A_r U_q^*$ for $q \geq 2$. By the initial assumption, at least one of the $\{C_1, \ldots, C_\ell\}$ and at least one of the $\{C_{\ell+1}, \ldots, C_{2\ell}\}$ have a factor $U_2 A_r U_2^*$ for some $1 \leq r \leq k$. Considering $Z = (1, \ldots, \ell)(\ell+1, \ldots, 2\ell) \in S_{2\ell}$, we can use the Weingarten formula (5.7) in order to get

$$\begin{aligned} &\mathbb{E}_{U_1}\left(\left|\mathrm{Tr}_n(U_{i_1} A_1 U_{i_1}^* B_1 \cdots U_{i_k} A_k U_{i_k}^* B_k)\right|^2\right) \\ &= \sum_{\substack{\sigma_1, \sigma_2, \sigma_3 \in S_{2\ell} \\ \sigma_1\sigma_2\sigma_3 = Z}} \mathrm{Tr}_{\sigma_1}(C_1, \ldots, \ldots, C_{2\ell}) \,\mathrm{Tr}_{\sigma_2}(D_1, \ldots, D_{2\ell}) \mathrm{Wg}(\sigma_3, n). \end{aligned} \tag{5.31}$$

We proceed to compute the order of each factor in the previous sum. Since $(B_1, \ldots, B_k)$ converges in distribution with respect tr_n, we have that $\mathrm{Tr}_{\sigma_1}(C_1, \ldots, C_{2\ell})$ is at least $O(n^{2\ell - |\sigma_1|})$. But we recall that there are at least two factors $U_2 A_r^\epsilon U_2^*$ in two different C_i's. Using again the Weingarten formula, we can show that the trace of a product of C_i's containing at least a factor A_r is $O(1)$. Hence, it is possible to show that for any σ_1, $\mathrm{Tr}_{\sigma_1}(C_1, \dots, C_{2\ell})$ is $O(n^{2\ell-2})$. In fact, if $|\sigma_1| \geq 2$, it is done. Otherwise, we have that σ_1 is equal to the identity element or is equal to a transposition. If σ_1 is the identity element in $S_{2\ell}$, then

$$\mathrm{Tr}_{\sigma_1}(C_1, \dots, C_{2\ell}) = \prod_{i=1}^{2\ell} \mathrm{Tr}_n(C_i).$$

But two of the C_i's contain a factor $U_2 A_r^{\epsilon} U_2^*$, then two of the $\mathrm{Tr}_n(C_i)$'s have order $O(1)$ and so $\mathrm{Tr}_{\sigma_1}(C_1, \dots, C_{2\ell}) = O(n^{2\ell-2})$. A similar reasoning works in the case that σ_2 is a transposition and hence $\mathrm{Tr}_{\sigma_1}(C_1, \dots, C_{2\ell}) = O(n^{2\ell-2})$.

By assumption, we have that $\mathrm{Tr}_{\sigma_2}(D_1, \dots, D_{2\ell}) = O(1)$. Finally, the asymptotic behavior of $\mathrm{Wg}(\sigma_3, n)$ is $O(n^{-2\ell-|\sigma_3|})$ and this allows us to conclude that

$$\mathrm{Tr}_{\sigma_1}(C_1, \dots, \dots, C_{2\ell})\, \mathrm{Tr}_{\sigma_2}(C_1, \dots, C_{2\ell}) \mathrm{Wg}(\sigma_3, n) = O(n^{-2}).$$

■

5.4 Some Computations of Eigenvalues of Polynomials of Random Matrices

In this section, we shall combine the results about asymptotic cyclic monotone independence of random matrices studied in previous sections with the explicit formulas for the eigenvalues of the polynomials of cyclically monotone elements proved in Theorem 4.2.1. We state the corresponding results on random matrices in the framework of compact operators. As we did in the proof of Theorem 5.3.1, firstly we shall prove the theorem for trace class operators and then we will conclude by approximation.

Theorem 5.4.1. *Let $n \in \mathbb{N}$. Let $U = U(n)$ be an $n \times n$ Haar unitary random matrix and $A_i = A_i(n)$, $B_j = B_j(n)$, $i = 1, \dots, k$, $j = 1, \dots, k$ be $n \times n$ deterministic matrices such that*

1. *$A_1, \dots, A_k$ are Hermitian,*

2. *$((A_1, \dots, A_k), \mathrm{Tr}_n)$ converges in compact distribution to a k-tuple of compact operators $((a_1, \dots, a_k), \mathrm{Tr}_H)$ as $n \to \infty$,*

3. *$((B_1, \dots, B_k), \mathrm{tr}_n)$ converges in distribution to a k-tuple of elements in a non-commutative probability space as $n \to \infty$,*

4. $\sup_{n\in\mathbb{N}} \|B_i(n)\| < \infty$ *for every* $i = 1, \ldots, k$.

Under the assumption 3), we can define $\beta_i = \lim_{n\to\infty} \mathrm{tr}_n(B_i)$, $\beta_{ij} = \lim_{n\to\infty} \mathrm{tr}_n(B_i^* B_j)$.

1. *If we denote* $B = (\beta_{ij})_{i,j=1}^k \in M_k(\mathbb{C})$, *then*

$$\lim_{n\to\infty} \mathrm{EV}\left(\sum_{i=1}^k UB_iU^*A_i(UB_iU^*)^*\right) = \mathrm{EV}\left(\sqrt{B}\,\mathrm{diag}(a_1, \ldots, a_k)\sqrt{B}\right) \quad a.s.$$

where $\sqrt{B}\,\mathrm{diag}(a_1, \ldots, a_k)\sqrt{B} \in \big(M_k(\mathbb{C}) \otimes S^1(H), \mathrm{Tr}_k \otimes \mathrm{Tr}_H\big)$.

2. *If* $B_1, \ldots, B_k$ *are Hermitian, then*

$$\lim_{n\to\infty} \mathrm{EV}\left(\sum_{i=1}^k A_iUB_iU^*A_i\right) = \mathrm{EV}\left(\sum_{i=1}^k \beta_i a_i^2\right) \quad a.s.$$

3. *Suppose that* $k = 1$ *and* B_1 *is Hermitian. If we define* $p = \sqrt{\beta_{11}}+\beta_1$, $q = -(\sqrt{\beta_{11}}-\beta_1)$ *and* $r = \sqrt{\beta_{11} - \beta_1^2}$, *then*

$$\begin{aligned}\lim_{n\to\infty} \mathrm{EV}\left(A_1UB_1U^* + UB_1U^*A_1\right) &= p\mathrm{EV}(a_1) \sqcup q\mathrm{EV}(a_1) \quad a.s., \\ \lim_{n\to\infty} \mathrm{EV}\left(\mathrm{i}\left(A_1UB_1U^* - UB_1U^*A_1\right)\right) &= r\mathrm{EV}(a_1) \sqcup (-r)\mathrm{EV}(a_1) \quad a.s.\end{aligned}$$

Proof. Only the proof of (1) is presented here. The other cases can be proved with the same ideas. We recall the notations defined in the proof of Theorem 5.3.1. Once again, we write B_i instead of UB_iU^*. So, if $P(x_1, \ldots, x_k, y_1, \ldots, y_k) = \sum_{i=1}^k y_ix_iy_i^*$, we denote $\{\lambda_i(n)\}_{i=1}^\infty$ as the eigenvalues of the matrix $P(A_1, \ldots, A_k, B_1, \ldots, B_k)$ for each $n \in \mathbb{N}$, $\{\lambda_i\}_{i=1}^\infty$ as the limiting eigenvalues of $P(A_1, \ldots, A_k, B_1, \ldots, B_k)$ (which are defined by Theorem 5.3.1), $\{\lambda_i^{(j)}(n)\}_{i=1}^\infty$ as the eigenvalues of the matrix $P(A_1^{(j)}, \ldots, A_k^{(j)}, B_1, \ldots, B_k)$, and $\{\lambda_i^{(j)}\}_{i=1}^\infty$ as the limiting eigenvalues of $P(A_1^{(j)}, \ldots, A_k^{(j)}, B_1, \ldots, B_k)$. As in the proof of Theorem 5.3.1, we have that $a_p^{(j)}$ are trace class operators. Applying Theorem 4.2.1 along with the almost sure asymptotic cyclically monotone independence provided by Theorem 5.2.4, we get that

$$\left\{\lambda_i^{(j)}\right\}_{i=1}^\infty = \mathrm{EV}\left(\sqrt{B}\,\mathrm{diag}\left(a_1^{(j)}, \ldots, a_k^{(j)}\right)\sqrt{B}\right). \tag{5.32}$$

On the other hand, if we denote the eigenvalues of $\sqrt{B}\,\mathrm{diag}(a_1, \ldots, a_k)\sqrt{B}$ as $\{\lambda_i'\}_{i=1}^\infty$, then we have to show that $\lambda_i = \lambda_i'$ for $i \in \mathbb{N}$. Using triangle's inequality, it can be shown that

$$\left|\lambda_i^\pm - (\lambda_i')^\pm\right| \;\leq\; \left|\lambda_i^\pm - \lambda_i^\pm(n)\right| + \left|\lambda_i^\pm(n) - \left(\lambda_i^{(j)}\right)^\pm(n)\right|$$

$$+\left|\left(\lambda_i^{(j)}\right)^{\pm}(n)-\left(\lambda_i^{(j)}\right)^{\pm}\right|+\left|\left(\lambda_i^{(j)}\right)^{\pm}-(\lambda_i')^{\pm}\right|$$

We have four terms in the right-hand side of the above inequality. The first of them converges to zero as $n\to\infty$ provided by Theorem 5.3.1. The second term is bounded by δ_j for any $n\in\mathbb{N}$. The third term also converges to 0 as $n\to\infty$. Then

$$\left|\lambda_i^{\pm}-(\lambda_i')^{\pm}\right|=\limsup_{j\to\infty}\limsup_{n\to\infty}\left|\lambda_i^{\pm}-(\lambda_i')^{\pm}\right|\leq\limsup_{j\to\infty}\left|\left(\lambda_i^{(j)}\right)^{\pm}-(\lambda_i')^{\pm}\right|.$$

The proof will be complete if we show that $\lambda_i^{(j)}\to\lambda_i'$ as $j\to\infty$, for any $i\geq 1$. This can be done by noticing that, in the same way that we get Equation (5.26), it is satisfied that

$$\lim_{j\to\infty}\left\|\sqrt{B}\,\mathrm{diag}\left(a_1^{(j)},\dots,a_k^{(j)}\right)\sqrt{B}-\sqrt{B}\,\mathrm{diag}(a_1,\dots,a_k)\sqrt{B}\right\|=0$$

along with that $\left|\lambda_i^{(j)}-\lambda_i'\right|\leq\delta_j$, for any $j\geq 1$. Finally, $\limsup_{j\to\infty}\left|\left(\lambda_i^{(j)}\right)^{\pm}-(\lambda_i')^{\pm}\right|=0$ and so $\lambda_i=\lambda_i'$ for any $i\in\mathbb{N}$. Then the proof is now complete. ∎

In a similar way, we can provide a version for random matrices of Propositions 4.2.2 and 4.2.5. For instance, we present the precise statement for Proposition 4.2.5.

Proposition 5.4.2. *Let $n\in\mathbb{N}$. Let $U=U(n)$ be an $n\times n$ Haar unitary random matrix and $A=A(n)$, $B_i=B_i(n)$, $C_j=C_j(n)$, $i,j=1,\dots,k$, be $n\times n$ deterministic matrices such that*

1. *A is Hermitian and (A,Tr_n) converges in compact distribution to a compact operator (a,Tr_H) as $n\to\infty$,*

2. *$((B_1,C_1\dots,B_k,C_k),\mathrm{tr}_n)$ converges in distribution to a $2k$-tuple of elements in a non-commutative probability space as $n\to\infty$,*

3. *$\sup_{n\in\mathbb{N}}\|B_i(n)\|<\infty$, $\sup_{n\in\mathbb{N}}\|C_i(n)\|<\infty$ for every $i=1,\dots,k$.*

Under the assumption 2), let $\beta_{ij}=\lim_{n\to\infty}\mathrm{tr}_n(C_iB_j)$ and $B'=(\beta_{ij})_{i,j=1}^k$. Let $\lambda_1,\dots,\lambda_k$ be the k eigenvalues of B' counting multiplicity. If

$$\sum_{i=1}^k(UB_iU^*)A(UC_iU^*)$$

is Hermitian, then

$$\lim_{n\to\infty} \mathrm{EV}\left(\sum_{i=1}^{k}(UB_iU^*)A(UC_iU^*)\right) = \bigsqcup_{i=1}^{k} \lambda_i \mathrm{EV}(a). \tag{5.33}$$

We also have explicit formulas when we are considering several Haar unitary random matrices instead of only one. However, we have to take into account Proposition 5.3.4 in order to get non-trivial conclusions. The combination of the explicit computation of eigenvalues and the asymptotic cyclic monotone independence in our studied random matrix models can be stated as follows.

Proposition 5.4.3. *Let $n \in \mathbb{N}$. Let $U_i = U_i(n)$ be $n \times n$ independent Haar unitary random matrices for $1 \leq i \leq k$. We consider the assumptions of Theorem 5.4.1. If we define $\beta_i = \lim_{n\to\infty} \mathrm{tr}_n(B_i^* B_i)$, then*

$$\lim_{n\to\infty} \mathrm{EV}\left(\sum_{i=1}^{k} B_iU_iA_iU_i^*B_i^*\right) = \bigsqcup_{i=1}^{k} \mathrm{EV}(\beta_i a_i) \quad a.s.$$

Proof. In the next proof, we only show the case that the limiting random variables $a_1, \ldots, a_k$ are trace class operator. The general case of compact operators can be treated as in the proof of Theorem 5.4.1.

Using Theorem 5.3.4, the limiting trace of monomials with two different U_i's in their factors is equal to zero almost surely. So, we can write for each $m \in \mathbb{N}$

$$\begin{aligned}
\lim_{n\to\infty} \mathrm{Tr}_n\left(\left(\sum_{i=1}^{k} B_iU_iA_iU_i^*B_i^*\right)^m\right) &= \sum_{i=1}^{k} \lim_{n\to\infty} \mathrm{Tr}_n\left((B_iU_iA_iU_i^*B_i^*)^m\right) \\
&= \sum_{i=1}^{k} \lim_{n\to\infty} \mathrm{Tr}_n\left((A_i(U_i^*B_i^*B_iU_i))^m\right) \\
\text{(by a.s. asymptotic cyclic monotonicity)} \quad &= \sum_{i=1}^{k} \lim_{n\to\infty} \mathrm{Tr}_n(A_i^m) \lim_{n\to\infty} \mathrm{tr}_n(B_i^*B_i)^m \\
&= \sum_{i=1}^{k} \mathrm{Tr}_H\left((\beta_i a_i)^m\right).
\end{aligned}$$

We finish the proof by using Proposition 1.3.6. ∎

5.5 Numerical Experiments

The final section of this manuscript is dedicated to present some examples of numerical simulations in order to illustrate the results of Section 5.4. For notational convenience, the index n is omitted in the $n \times n$ matrices.

Example 5.5.1. We consider the case of the commutator and the anti-commutator. We take $n = 300$ for the simulations. Let $A = \mathrm{diag}(2^0, 2^{-1}, 2^{-2}, \ldots, 2^{-n+1})$ and $B = G^2$, where G is a GUE random matrix. Let U be a Haar unitary random matrix. We know that B converges in distribution to s^2, where s is a standard semicircular element. In particular, we have that $\beta_1 = \lim_{n\to\infty} \mathrm{tr}_n(B) = 1$ and $\beta_{11} = \lim_{n\to\infty} \mathrm{tr}_n(B^2) = 2$. Also, it is clear that A converges in compact distribution to a where $\mathrm{EV}(a) = \{2^{-k} : k \in \mathbb{N}_0\}$. Theorem 5.4.1 establishes that

$$\lim_{n\to\infty} \mathrm{EV}(AUBU^* + UBU^*A) = p\mathrm{EV}(a) \sqcup q\mathrm{EV}(a),$$

where in this case, $p = \sqrt{\beta_{11}} + \beta_1 = \sqrt{2} + 1$ and $q = -(\sqrt{\beta_{11}}) - \beta_1) = 1 - \sqrt{2}$. We now present the first ten eigenvalues of a realization of the random matrices.

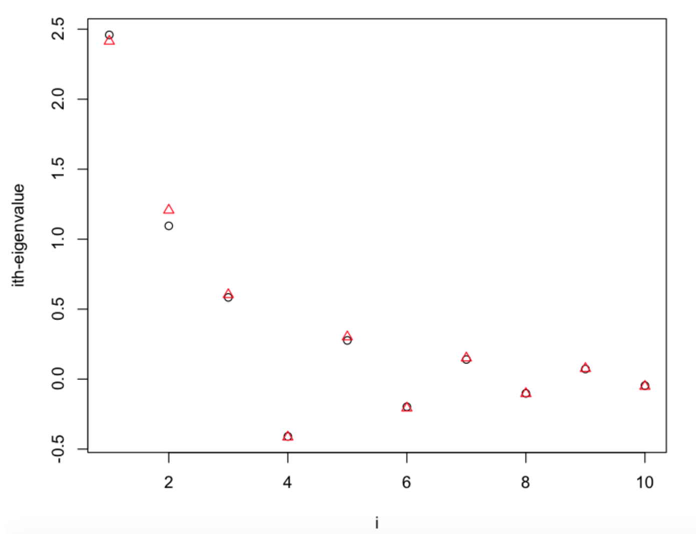

Figure 5.1: Comparison between eigenvalues of $AUBU^* + UBU^*A$ (black circles) and $p\mathrm{EV}(a) \sqcup q\mathrm{EV}(a)$ (red triangles).

We also show a comparison between the first three moments of the commutator with respect Tr_n and the moments provided by the rule of cyclic monotone independence.

k	$\mathrm{Tr}_n((AUBU^*+UBU^*A)^k)$	$\mathrm{Tr}(a^k)(p^k+q^k)$
1	3.894818	4
2	7.910895	8
3	16.33332	16

Table 5.1: First three moments of $AUBU^*+UBU^*A$ compared to the theoretical limiting moments.

The same procedure is done in the case of the anti-commutator. Theorem 5.4.1 establishes that

$$\lim_{n\to\infty} \mathrm{EV}(\mathrm{i}(AUBU^*-UBU^*A)) = r\mathrm{EV}(a) \sqcup -r\mathrm{EV}(a),$$

where $r = \sqrt{\beta_{11}-\beta_1^2} = 1$. The plot of the first ten eigenvalues and the table of the first three moments compared with the moments provided by the cyclic monotone independence rule are the following:

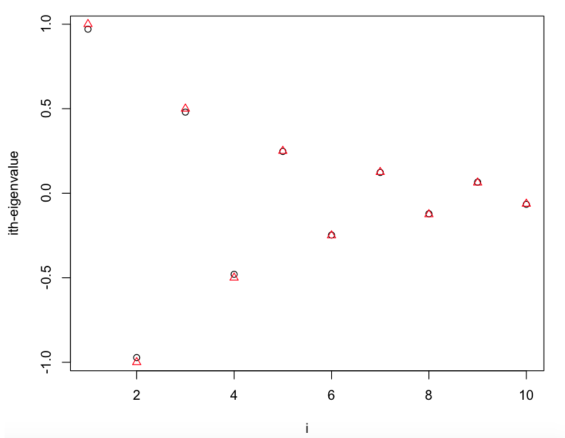

Figure 5.2: Comparison between eigenvalues of $i(AUBU^*-UBU^*A)$ (black circles) and $r\mathrm{EV}(a) \sqcup -r\mathrm{EV}(a)$ (red triangles).

k	$\mathrm{Tr}_n((AUBU^*+UBU^*A)^{2k})$	$2\,\mathrm{Tr}(a^{2k})r^{2k}$
1	2.717913	2.666667
2	2.227108	2.133333
3	2.16769	2.031746

Table 5.2: First three moments of $\mathrm{i}(AUBU^*-UBU^*A)$ compared to the theoretical limiting moments.

Example 5.5.2. In this case, we consider $A_1 = A_2 = A_3 = \mathrm{diag}(2^0, 2^{-1}, \ldots, 2^{-n+1})$ and B_1, B_2, B_3 independent GUE random matrices, with $n = 300$. We have again that (A_1, A_2, A_3) converges in compact distribution to (a_1, a_2, a_3) with respect to Tr_n and (B_1, B_2, B_3) converges in distribution with respect to tr_n. If U_1, U_2, U_3 are independent Haar unitary random matrices, Proposition 5.4.3 states that

$$\lim_{n\to\infty} \mathrm{EV}\left(B_1U_1A_1U_1^*B_1^* + B_2U_2A_2U_2^*B_2^* + B_3U_3A_3U_3^*B_3^*\right) = \mathrm{EV}(\beta_1 a_1) \sqcup \mathrm{EV}(\beta_2 a_2) \sqcup \mathrm{EV}(\beta_3 a_3),$$

where $\beta_i = \lim_{n\to\infty} \mathrm{tr}_n(B_i^* B_i) = 1$, for $i = 1, 2, 3$. Since $\mathrm{EV}(a_i) = \{2^{-k} : k \geq 0\}$, the limiting eigenvalues are

$$\{1, 1, 1, 1/2, 1/2, 1/2, 1/4, 1/4, 1/4, 1/8, 1/8, 1/8, \ldots\}.$$

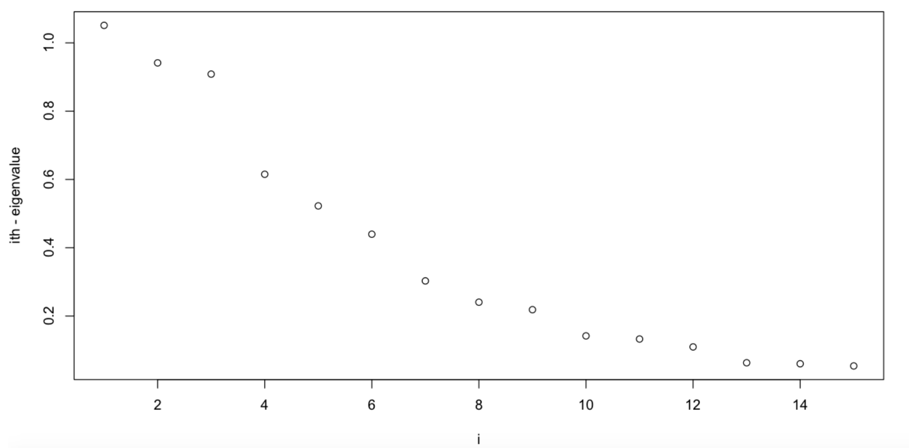

Figure 5.3: First 15 eigenvalues of $\sum_{k=1}^{3} B_iU_iA_iU_i^*B_i^*$.

We also compute the first three moments and the moments given by cyclic monotone independence.

k	$\mathrm{Tr}_n((\sum_{k=1}^{3} B_iU_iA_iU_i^*B_i^*)^k)$	$3\,\mathrm{Tr}(a^k)$
1	5.928757	6
2	3.951528	4
3	3.391974	3.428571

Table 5.3: First three moments of $\sum_{k=1}^{3} B_iU_iA_iU_i^*B_i^*$ compared to the theoretical limiting moments.

Example 5.5.3. We illustrate the asymptotic version in random matrices of Proposition 4.2.2. Let $n = 300$. Let B_1 and B_2 be independent $n \times n$ selfadjoint random matrices whose entries are independent real Gaussian random variables with mean zero and variance 2. Consider $D = \mathrm{diag}(2^0, 2^{-1}, \ldots, 2^{-n+1})$ and U_1, U independent Haar unitary random matrices and independent of B_1 and B_2. We take $A_1 = D$ and $A_2 = U_1 D U_1^*$. Define the block matrices

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}, \quad B = \begin{pmatrix} UB_1^2U^* & 0 \\ 0 & UB_2^2U^* \end{pmatrix}. \tag{5.34}$$

We shall show a realization of the eigenvalues of AB and we shall compare them with the eigenvalues of $A'B'$ where $A' = \lim_{n\to\infty} A$ and

$$B' = \begin{pmatrix} \lim\limits_{n\to\infty} \mathrm{tr}_n(B_1^2) & 0 \\ 0 & \lim\limits_{n\to\infty} \mathrm{tr}_n(B_2^2) \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}.$$

The first 15 eigenvalues of AB and $A'B'$ are plotted in the next graphic:

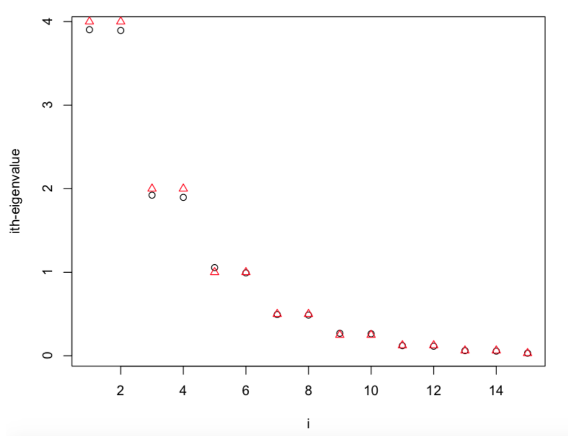

Figure 5.4: Comparison between eigenvalues of AB (black circles) and $A'B'$ (red triangles).

k	$\mathrm{Tr}_{2n}((AB)^k)$	$\mathrm{Tr}_{2n}((A'B')^k)$
1	15.65416	16
2	40.45189	42.66667
3	134.869	146.2857

Table 5.4: First three moments of AB and $A'B'$.

Example 5.5.4. We show another numerical simulation for Proposition 4.2.2. Again, let $n = 300$. Let B_1, B_2 and B_3 be independent $n \times n$ GUE random matrices, $D = \mathrm{diag}(2^0, 2^{-1}, \ldots, 2^{-n+1})$, and U_1, U_2 be independent Haar unitary random matrices. We take $A_1 = D, A_2 = U_1 D U_1^*$ and $A_3 = U_2 D U_2^*$. Define the block matrices

$$A = \begin{pmatrix} A_1 & A_2 \\ A_2^* & A_3 \end{pmatrix}, \quad B = \begin{pmatrix} B_1^2 & B_2^2 \\ B_2^2 & B_3^2 \end{pmatrix}. \tag{5.35}$$

We shall show a realization of the eigenvalues of BAB and we shall compare them with the eigenvalues of $A'B'$ where $A' = \lim_{n\to\infty} A$ and

$$\begin{aligned} B' &= \lim_{n\to\infty} (\mathrm{Id} \otimes \mathrm{tr}_n)(B^2) \\ &= \begin{pmatrix} \lim\limits_{n\to\infty} \mathrm{tr}_n(B_1^4 + B_2^4) & \lim\limits_{n\to\infty} \mathrm{tr}_n(B_1^2 B_2^2 + B_2^2 B_3^2) \\ \lim\limits_{n\to\infty} \mathrm{tr}_n(B_2^2 B_1^2 + B_3^2 B_2^2) & \lim\limits_{n\to\infty} \mathrm{tr}_n(B_2^4 + B_3^4) \end{pmatrix} \\ &= \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}. \end{aligned}$$

In the last equality, we use the fact that the B_1, B_2 and B_3 asymptotically behaves as a free semicircular family. Also again, a graphic of the first 15 eigenvalues of BAB and $A'B'$, and a comparison of the first three moments are provided.

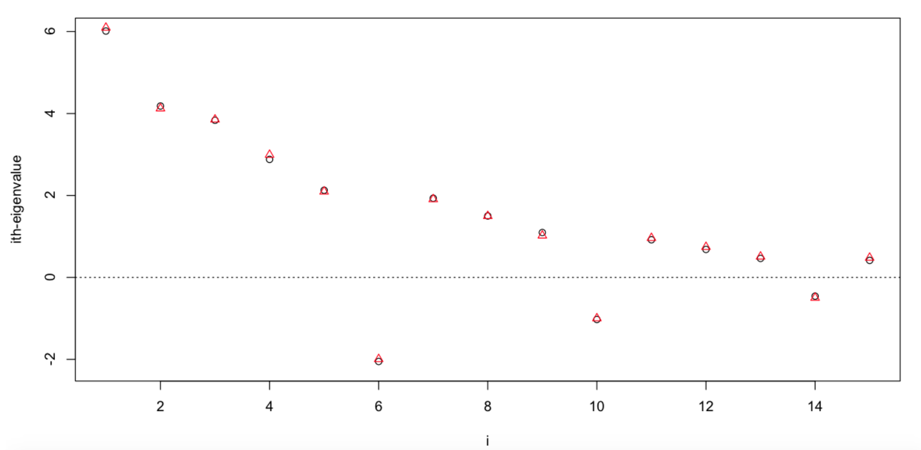

Figure 5.5: Comparison between eigenvalues of BAB (black circles) and $A'B'$ (red triangles).

k	$\mathrm{Tr}_{2n}((BAB)^k)$	$\mathrm{Tr}_{2n}((A'B')^k)$
1	23.70467	24
2	95.90039	96.85024
3	383.8305	393.54777

Table 5.5: First three moments of BAB and $A'B'$.

Example 5.5.5. Finally, we give a numerical example for Proposition 5.4.2. Let $n = 400$. Consider the matrices $A = \mathrm{diag}(2^0, 2^{-1}, 2^{-2}, \dots, 2^{-n+1})$ and B and C independent GUE random matrices, and U be an $n \times n$ Haar unitary random matrix. According to Proposition 5.4.2, the limiting eigenvalues of

$$UBU^*AUCU^* + UCU^*AUBU^* \tag{5.36}$$

are $\lambda_1 \mathrm{EV}(a) \sqcup \lambda_2 \mathrm{EV}(a)$, where $a = \mathrm{diag}(2^0, 2^{-1}, \dots,)$ and λ_1, λ_2 are the eigenvalues of the matrix

$$\lim_{n\to\infty} \mathrm{id} \otimes \mathrm{tr}_n \begin{pmatrix} UCU^*UBU^* & UCU^*UCU^* \\ UBU^*UBU^* & UBU^*UCU^* \end{pmatrix} = \lim_{n\to\infty} \begin{pmatrix} \mathrm{tr}_n(CB) & \mathrm{tr}_n(C^2) \\ \mathrm{tr}_n(B^2) & \mathrm{tr}_n(BC) \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Hence $\lambda_1 = 3$ and $\lambda_2 = -1$. Then, the limiting eigenvalues multiset is

$$\{3 \cdot 2^{-n}, -2^{-n} \ : \ n \geq 0\}.$$

A numerical realization of the matrix (5.36) is done. We provide a comparison of the first three moments and the theoretical limiting moments.

k	$\mathrm{Tr}_n((UBU^*AUCU^* + UCU^*AUBU^*)^k)$	$2\,\mathrm{Tr}(a^k)$
1	4.042675	4
2	14.31484	13.33333
3	32.9917	29.71429

Table 5.6: First three moments of $UBU^*AUCU^* + UCU^*AUBU^*$ and the theoretical limiting moments.

Finally, we compare the first 15 eigenvalues to the theoretical limiting eigenvalues.

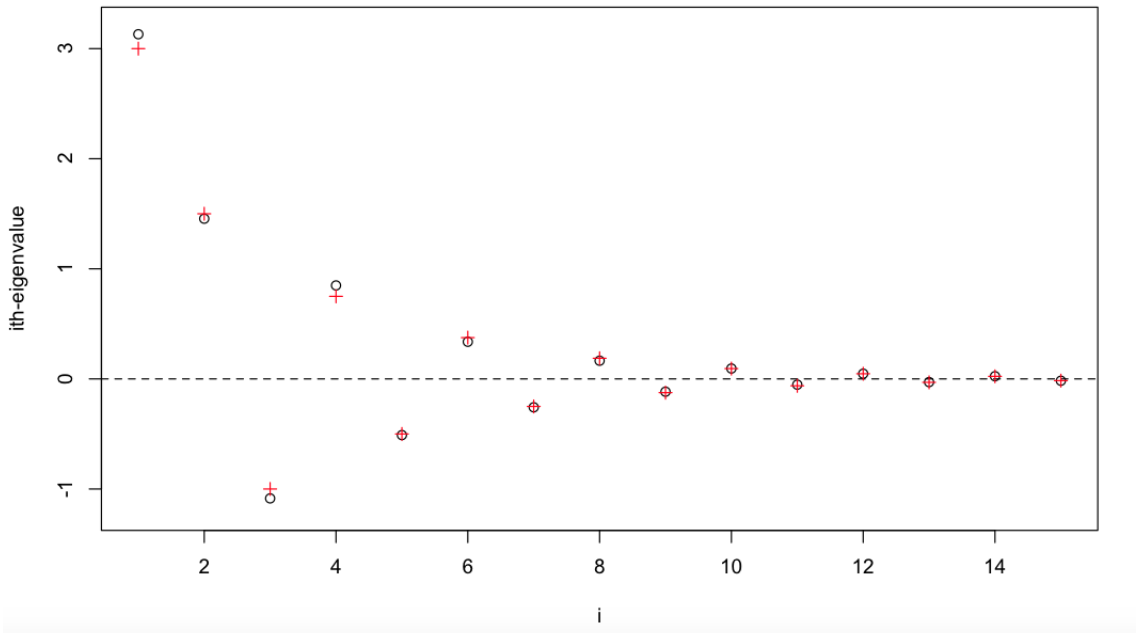

Figure 5.6: Black circles correspond to the eigenvalues a realization of the matrix (5.36). Red crosses correspond to the limiting theoretical eigenvalues.

Appendix A

A characterization of Tr_σ

The objective of this appendix is to give a proof of the Weingarten formula used in Chapter 5. First, we recall the necessary notation to state the result. Let $k \in \mathbb{N}$ and S_k be the symmetric group acting on $\{1, 2, \ldots, k\}$. We consider $\rho : S_k \to M_n(\mathbb{C})^{\otimes k}$ the map where for each $\sigma \in S_k$, $\rho(\sigma) : M_n(\mathbb{C})^{\otimes k} \to M_n(\mathbb{C})^{\otimes k}$ is a linear transformation such that

$$\rho(\sigma)(v_1 \otimes \cdots \otimes v_k) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(k)}, \quad \forall\, v_1, \ldots, v_k \in \mathbb{C}^n. \tag{A.1}$$

We have then that ρ is a unital group homomorphism. Throughout the proof, we shall consider the canonical identification $M_n(\mathbb{C})^{\otimes k} \cong M_{n^k}(\mathbb{C})$ as Hilbert spaces, given by the Kronecker product. In this way, $\rho(\sigma)$ is a permutation matrix for any $\sigma \in S_k$ and so, it is unitary. Since $\mathrm{id} = \rho(\sigma\sigma^{-1}) = \rho(\sigma)\rho(\sigma^{-1})$, we conclude that $\rho(\sigma^{-1}) = \rho(\sigma)^{-1} = \rho(\sigma)^*$.

Now, fix $\sigma \in S_k$. We also fix $A = A_1 \otimes \cdots \otimes A_k \in M_n(\mathbb{C})^{\otimes k}$, where $A_p \in M_n(\mathbb{C})$ for $p = 1, \ldots, k$. For a cycle $c = (j_1, j_2, \ldots, j_m)$ of σ, we define $A_c = A_{j_1} A_{j_2} \cdots A_{j_m}$, and if $\sigma = c_1 c_2 \cdots c_{\ell(\sigma)}$ is the cycle decomposition of σ, we define

$$\mathrm{Tr}_\sigma\left(A_1, \ldots, A_k\right) = \prod_{j=1}^{\ell(\sigma)} \mathrm{Tr}_n\left(A_{c_j}\right). \tag{A.2}$$

The next lemma gives us an alternative definition of Tr_σ. In the rest of the appendix, we write $A_p = \left(a_{ij}^{(p)}\right)_{ij=1}^{n}$ for any $p = 1, \ldots, k$.

Lemma A.1 (Exercise 4.1 in [12])**.** *With the above notation, we have that*

$$\mathrm{Tr}_\sigma\left(A_1, \ldots, A_k\right) = \sum_{i_1, \ldots, i_k = 1}^{n} a^{(1)}_{i_1 i_{\sigma(1)}} a^{(2)}_{i_2 i_{\sigma(2)}} \cdots a^{(k)}_{i_k i_{\sigma(k)}}. \tag{A.3}$$

Proof. We know that (i, j)-entry of $A_p A_q$ is equal to $\sum_{k=1}^{n} a_{ik}^{(p)} a_{kj}^{(q)}$. More generally, the

(i_1, i_1)-entry of $A_{j_1} A_{j_2} \cdots A_{j_m}$ is

$$\sum_{i_2,\ldots,i_m=1}^{n} a_{i_1 i_2}^{(j_1)} a_{i_2 i_3}^{(j_2)} \cdots a_{i_m i_1}^{(j_m)}.$$

Then if $c_k = (j_1, j_2, \ldots, j_m)$ is a cycle of σ, then

$$\mathrm{Tr}_n(A_{c_k}) = \mathrm{Tr}_n(A_{j_1} A_{j_2} \cdots A_{j_m}) = \sum_{i_1,\ldots,i_m=1}^{n} a_{i_1 i_2}^{(j_1)} a_{i_2 i_3}^{(j_2)} \cdots a_{i_m i_1}^{(j_m)},$$

where $j_1, \ldots, j_m$ and the indexes $i_1, \ldots, i_m$ depend of $k \in \{1, \ldots, \ell(\sigma)\}$. Making a change of the indexes, we conclude that

$$\begin{aligned}
\mathrm{Tr}_\sigma(A_1, \ldots, A_k) &= \prod_{i=1}^{\ell(\sigma)} \mathrm{Tr}_n(A_{c_i}) \\
&= \prod_{i=1}^{\ell(\sigma)} \sum_{i_1,\ldots,i_m=1}^{n} a_{i_1 i_2}^{(j_1)} a_{i_2 i_3}^{(j_2)} \cdots a_{i_m i_1}^{(j_m)} \\
&= \sum_{i_1,\ldots,i_k=1}^{n} a_{i_1 i_{\sigma(1)}}^{(1)} a_{i_2 i_{\sigma(2)}}^{(2)} \cdots a_{i_k i_{\sigma(k)}}^{(k)}.
\end{aligned}$$

■

Finally, we state and prove the desired formula which gives us another interesting characterization of Tr_σ in the language of the natural representation ρ.

Theorem A.2. *With the above notation, we have that*

$$\mathrm{Tr}_{M_n(\mathbb{C})^{\otimes k}}(\rho(\sigma)^* A) = \mathrm{Tr}_\sigma(A_1, \ldots, A_k). \tag{A.4}$$

Proof. Recall that by definition, $\rho(\sigma)^*$ is a permutation matrix in $M_{n^k}(\mathbb{C})$. So, $\rho(\sigma)^* A$ is a matrix obtained by exchanging the rows of the matrix A in a specific way given by σ. By the definition of the Kronecker product, a column of A is given by

$$(a_{1i_1}^{(1)} a_{1i_2}^{(2)} \cdots a_{1i_k}^{(k)}, a_{1i_1}^{(1)} a_{1i_2}^{(2)} \cdots a_{2i_k}^{(k)}, \ldots, a_{ni_1}^{(1)} a_{ni_2}^{(2)} \cdots a_{ni_k}^{(k)})$$

$$= (a_{1i_1}^{(1)}, a_{2i_1}^{(1)}, \ldots, a_{ni_1}^{(1)}) \otimes (a_{1i_2}^{(2)}, a_{2i_2}^{(2)}, \ldots, a_{ni_2}^{(2)}) \otimes \cdots \otimes (a_{1i_k}^{(k)}, a_{2i_k}^{(k)}, \ldots, a_{ni_k}^{(k)})$$

for some $i_1, \ldots, i_k \in \{1, \ldots, n\}$. We say that the above column is the column associated to the k-tuple $(i_1, i_2 \ldots, i_k) \in \{1, \ldots, n\}^k$. Again by the rule of the Kronecker product, the entry in the diagonal of A, which is also in the column associated to $(i_1, \ldots, i_k)$, is precisely

$$a^{(1)}_{i_1 i_1} a^{(2)}_{i_2 i_2} \cdots a^{(k)}_{i_k i_k}.$$

On the other hand, since the columns of the matrix $\rho(\sigma)^* A$ are the obtained by applying $\rho(\sigma)^*$ to the columns of A, the column of $\rho(\sigma)^* A$ determined by $(i_1, \ldots, i_k)$ is

$$\rho\left(\sigma^{-1}\right)\left((a^{(1)}_{1i_1}, a^{(1)}_{2i_1}, \ldots, a^{(1)}_{ni_1}) \otimes (a^{(2)}_{1i_2}, a^{(2)}_{2i_2}, \ldots, a^{(2)}_{ni_2}) \otimes \cdots \otimes (a^{(k)}_{1i_k}, a^{(k)}_{2i_k}, \ldots, a^{(k)}_{ni_k})\right)$$

$$= (a^{(\sigma(1))}_{1i_{\sigma(1)}}, a^{(\sigma(1))}_{2i_{\sigma(1)}}, \ldots, a^{(\sigma(1))}_{ni_{\sigma(1)}}) \otimes (a^{(\sigma(2))}_{1i_{\sigma(2)}}, a^{(\sigma(2))}_{2i_{\sigma(2)}}, \ldots, a^{(\sigma(2))}_{ni_{\sigma(2)}}) \otimes \cdots \otimes (a^{(\sigma(k))}_{1i_{\sigma(k)}}, a^{(\sigma(k))}_{2i_{\sigma(k)}}, \ldots, a^{(\sigma(k))}_{ni_{\sigma(k)}}).$$

Note that the entries of $\rho(\sigma)^* A$ are again products of the entries of the matrices $A_1, \ldots, A_k$. Also, the entry in the diagonal of $\rho(\sigma)^* A$ in the column associated to $(i_1, \ldots, i_k)$ is

$$a^{(\sigma(1))}_{i_1 i_{\sigma(1)}} a^{(\sigma(2))}_{i_2 i_{\sigma(2)}} \cdots a^{(\sigma(k))}_{i_k i_{\sigma(k)}}.$$

Summing over all the possible indexes and applying Lemma A.1, we have that

$$\begin{aligned} \mathrm{Tr}_{n^k}(\rho(\sigma)^* A) &= \sum_{i_1, \ldots, i_k = 1}^{n} a^{(\sigma(1))}_{i_1 i_{\sigma(1)}} a^{(\sigma(2))}_{i_2 i_{\sigma(2)}} \cdots a^{(\sigma(k))}_{i_k i_{\sigma(k)}} \\ &= \mathrm{Tr}_\sigma\left(A_{\sigma(1)}, A_{\sigma(2)}, \ldots, A_{\sigma(k)}\right). \end{aligned}$$

Now, if $c = (j_1, \ldots, j_m)$ is a cycle in σ, then $\sigma(j_r) = j_{r+1}$, for $r = 1, \ldots, m$ mod m. By cyclic property of trace, we have that

$$\begin{aligned} \mathrm{Tr}_n\left((A_{\sigma(1)} \otimes \cdots \otimes A_{\sigma(k)})_c\right) &= \mathrm{Tr}_n\left(A_{\sigma(j_1)} A_{\sigma(j_2)} \cdots A_{\sigma(j_{m-1})} A_{\sigma(j_m)}\right) \\ &= \mathrm{Tr}_n\left(A_{j_2} A_{j_3} \cdots A_{j_m} A_{j_1}\right) \\ &= \mathrm{Tr}_n\left(A_{j_1} A_{j_2} A_{j_3} \cdots A_{j_m}\right) \\ &= \mathrm{Tr}_n\left(A_c\right). \end{aligned}$$

Finally, we conclude that

$$\begin{aligned} \mathrm{Tr}_\sigma\left(A_{\sigma(1)}, \ldots, A_{\sigma(k)}\right) &= \prod_{j=1}^{\ell(\sigma)} \mathrm{Tr}_n\left((A_{\sigma(1)} \otimes \cdots \otimes A_{\sigma(k)})_{c_j}\right) \\ &= \prod_{j=1}^{\ell(\sigma)} \mathrm{Tr}_n\left(A_{c_j}\right) \\ &= \mathrm{Tr}_\sigma\left(A_1, \ldots, A_k\right). \end{aligned}$$

Hence

$$\mathrm{Tr}_{M_n(\mathbb{C})^{\otimes k}}\left(\rho(\sigma)^* A\right) = \mathrm{Tr}_\sigma\left(A_{\sigma(1)}, A_{\sigma(2)}, \ldots, A_{\sigma(k)}\right) = \mathrm{Tr}_\sigma\left(A_1, A_2, \ldots, A_k\right),$$

and so (A.4) is proved. ■

Bibliography

[1] G.W Anderson, A. Guionnet, O. Zeitouni. *An Introduction to Random Matrices,* Cambridge Studies in Advanced Mathematics, vol. 118, Cambridge University Press, (2010).

[2] S. Belinschi, H. Bercovici, M. Capitaine. On the outlying eigenvalues of a polynomial in large independent random matrices, preprint (2017). Available online in www.arxiv.org under reference arXiv:1703.08102.

[3] S.T. Belinschi, D. Shlyakhtenko. Free probability of type B: analytic aspects and applications, *American Journal of Mathematics* Vol. 134 (2012), 193-234.

[4] R. Bhatia. *Matrix Analysis,* Graduate Texts in Mathematics 169, Springer-Verlag, New York, (1997).

[5] P. Biane, F. Goodman, A. Nica. Non-crossing cumulants of type B, *Transactions of the American Mathematical Society* 355 (2003), 2263-2303.

[6] M. Capitaine, C. Donati-Martin. Spectrum of deformed random matrices and free probability, To appear in *SMF volume Panoramas et Synthèses*, (2016).

[7] B. Collins. Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and Free Probability, *International Mathematics Research Notices*, no. 17, 953-982, (2003).

[8] B. Collins, T. Hasebe, N. Sakuma. Free probability for purely discrete eigenvalues of random matrices, *Journal of the Mathematical Society of Japan* 70, No. 3 (2018), 1111-1150.

[9] B. Collins, P. Sniady. Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. *Communications in Mathematical Physics* 264 (2006), 773-795.

[10] J. B. Conway. *A Course in Operator Theory,* American Mathematical Society, (2000).

[11] M. Février, A. Nica. Infinitesimal non-crossing cumulants and free probability of type B, *Journal of Functional Analysis.* 258 (2010), No. 9, 2983-3023.

[12] J. Mingo, R. Speicher. *Free Probability and Random Matrices,* Fields Institute Monographs 36, Springer, (2017).

[13] A. Nica, R. Speicher. Commutators of free random variables, *Duke Mathematical Journal* 92 (1998), 553-592.

[14] A. Nica, R. Speicher. *Lectures on the Combinatorics of Free Probability,* London Mathematical Society Lecture Note Series 335, Cambridge University Press, (2006).

[15] A. Nica, R. Speicher. On the multiplication of free n-tuples of non-commutative random variables, *American Journal of Mathematics* Vol. 118 (1996), 799-837.

[16] W. Rudin. *Functional Analysis,* McGraw-Hill, (1974).

[17] D. Shlyakhtenko. Free probability of type B and asymptotics of finite-rank perturbations of random matrices, preprint (2015). Available online in `www.arxiv.org` under reference arXiv:1509.08841.

[18] T. Tao. *Topics in Random Matrix Theory,* American Mathematical Society, (2012).

[19] D. Voiculescu. Addition of certain non-commuting random variables, *Journal of Functional Analysis* 66 (1986), 323-346.

[20] D. Voiculescu. Limit laws for random matrices and free products, *Inventiones Mathematicae* 104 (1991), 201-220.

[21] D. Voiculescu, K. J. Dykema, A. Nica. *Free Random Variables*, CRM Monograph Series 1, American Mathematical Society, (1992).

[22] E. Wigner. On the distribution of the roots of certain symmetric matrices,. *Annals of Mathematics. Second Series* 67 (1958), 325-327.

[23] J. Wishart. The generalized product moment distribution in samples from a normal multivariate population, *Biometrika* 20A.1/2 (July 1928), 32-52.