Non-Commutative Probability for the Spectral Analysis of Simplicial Complexes

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for the degree of Master of Science
in the Department of Mathematics and Statistics
University of Saskatchewan
Saskatoon

By

Diego Barba de la Mora

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department Mathematics and Statistics
142 McLean Hall
106 Wiggins Road Saskatoon
University of Saskatchewan
Saskatoon, Saskatchewan
S7N 5E6
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan
S7N 5C9
Canada

Abstract

Free probability theory, invented by Voiculescu, and greatly expanded by Speicher, is a young and active area of research with numerous applications in pure and applied mathematics.

This Master thesis is a comprehensive study of a specific result in the recent preprint by C. Vargas, in which Vargas presents a survey of applications of non-commutative and free probability to topological data analysis. The relevant result from the preprint reveals a new interpretation of Betti numbers for simplicial complexes in terms of distributions in an operator-valued probability space.

This thesis is mostly an exposition of the areas of free probability and algebraic topology; here, we do not present cutting-edge research in either free probability or algebraic topology. The author did a literature review for both fields and presents here the results in a comprehensive way along with detailed proofs and motivating examples that one may not find in a research paper. We believe that this thesis would help researchers to quickly grasp the main ideas and tools in both fields, and we hope it will help to advance the research in both areas and to develop applications in related areas.

Acknowledgements

I wish to sincerely thank my supervisor, JC, for all his guidance and personalized training in free probability, his help in organizing and redacting the thesis, and his help writing and editing mathematics documents; I especially thank him for giving me the opportunity and financial support to study in Saskatoon. I also wish to thank Christopher, for his help with everything that involved algebraic topology, for joining me in the free probability seminar, and for his valuable input in writing this thesis. I also thank Professor Steve Rayan, for being always available for help and advice in many aspects of my master's program.

I express thanks to the various professors who received me in CIMAT for two events I attended, especially Dr. Carlos Vargas and Dr. Octavio Arizmendi. I specifically appreciate very much Carlos' generous help as I was learning the contents I now present. He patiently answered my questions, gave me examples, and directed me to useful resources, both in person and by email. I am also grateful to Dr. Mihai Popa, who strongly encouraged me to come to Saskatchewan in the first place, and introduced me to JC. And I thank both Dr. Popa and Dr. Patterson, from UTSA, for their support at the end of my undergraduate years.

I thank my friends in Saskatchewan, especially Cloud, Simon, and Penny, for all the ways in which they were there when I needed them. I thank Justin, my friend and landlord, for offering me such a peaceful place to stay in Saskatoon. Also, I want to thank my office-mate Evan, for answering all my quick questions and difficulties in LaTeX, and for being there when I really needed to take a break from the math and talk to anyone about anything.

I thank my parents who, while being far away from me, were always supportive through their prayers, their good advice, and their financial support. Finally, and above all, I give thanks to God Almighty, "Him without whom nothing can be done", for giving me physical, emotional, mental, and spiritual health.

Dedicated to Mary, our blessed mother

Contents

Introduction

Non-commutative probability is the systematic study of non-commutative random variables, defined as abstract elements of non-commutative algebras. Free probability is the main branch of non-commutative probability, and it studies joint random variables under the view of free independence, or freeness. The language used in free probability is analogous to the language of classical probability. Namely, we will speak of expectations, moments, distributions, independence, there will be a central limit theorem, etc.

The theories of free probability and non-commutative probability have been very fruitful recently in its applications to numerous areas of pure and applied mathematics. This Master's thesis attempts to introduce the reader to one new application of non-commutative probability presented in a recent preprint by C. Vargas [15], in which the framework of non-commutative probability is used in the study of algebraic topology and topological data analysis. Thus this thesis is roughly written in two parts. The first part (chapters 1 and 2) introduces the reader to non-commutative and free probability. Our presentation here closely follows that of A. Nica and R. Speicher in [12]. The second part of the thesis (chapters 3-6) focuses on the algebraic topology of simplicial complexes, introducing the reader to all the concepts used, and presenting the main result from non-commutative probability in algebraic topology. This part is based roughly on [15].

Chapter 1 gives the definitions of non-commutative probability spaces, $*$-probability spaces, moments and distributions, with their basic properties and with plenty of examples. The semicircular distribution, which is particularly important, is introduced in section 1.3. Finally, in section 1.4 we define the Cauchy transform of a compactly supported probability measure and briefly discuss techniques for computing analytic distributions.

Chapter 2 gives an overall view of free probability. The definition of free independence is provided after a discussion of joint moments and joint distributions. A few examples and multiple computations are provided. The combinatorics of free probability are then introduced, namely crossing and non-crossing partitions, and their relation to Catalan numbers. Then, the free central limit theorem is stated in section 2.4 and proved via combinatorics. Although this result is not used elsewhere in this thesis, we choose to include it for two reasons. First, to illustrate the main way in which free probability is analogous to classical probability theory, and second, to help the reader see how non-crossing partitions appear naturally in free probability. Then, without going too far into the combinatorics, free cumulants are defined by the moment-cumulant formulas, and presented with examples and computations. Then we discuss free convolution and the additivity of cumulants. The $\mathcal{R}$-transform is only briefly introduced, and its properties only stated with no proof.

Chapter 3 introduces simplicial complexes. Since textbooks rarely agree on their definition and notations of simplicial complexes, most of this chapter is devoted to clarifying the conventions used in the thesis.

Specifically, we use finite simplicial complexes in $\mathbb{R}^n$, with ordered and oriented simplices whose order and orientations are predetermined by a fixed order of the vertices. As defined here, a simplicial complex is a collection of simplices satisfying some combinatorial properties, a simplex is the convex hull of any given set of points in general position in $\mathbb{R}^n$, and to each simplicial complex is associated a fixed geometric realization in $\mathbb{R}^n$.

Once simplicial complexes are defined and our notation established, we introduce the homology of simplicial complexes by defining boundary operators on simplices, then boundary operators on simplicial complexes, and finally its corresponding boundary matrix, denoted here for a given simplicial complex, by J.

Chapter 4 is very short and simply provides a proof of the finite-dimensional version of the Hodge theorem. The proof of this this finite-dimensional version of the theorem requires only linear algebra, and so it is very accessible. The proof of the general Hodge theorem, on the other hand, requires familiarity with differential geometry that would be too advanced for non-experts. Furthermore the author found no textbooks or papers providing this simple proof, and thus hopes it will be helpful to some readers. A similar, finite-dimensional version of this theorem was proved and posted online in [10].

Chapter 5 begins by introducing Operator-Valued Probability Spaces, a generalization of non-commutative probability that will be used for the spectral analysis of simplicial complexes. Immediately, we move on to prove the main result as presented in Vargas' preprint [15], which is simply a probabilistic interpretation of the Hodge theorem in the framework of operator-valued non-commutative probability.

We conclude, in Chapter 6, by discussing future developments that, in the author's opinion, can either benefit from the results of this thesis, or help advance them further. Some of these future developments have been previously explored to some extent.

1 Non-Commutative Probability Theory

The theory of free probability was invented some thirty-five years ago by Dan-Virgil Voiculescu. It began as a tool for approaching certain problems in operator algebras [16]. Eventually, free probability became its own research area, with manifestations in many branches of pure and applied mathematics. A key result in the historical development of free probability was the *free central limit theorem,* proved analytically by Voiculescu in [18]. The framework on which free probability is developed is non-commutative probability. Non-commutative probability is a field of study analogous to classical probability, but with random variables defined on abstract algebraic structures; in particular, random variables are allowed to be non-commutative. Many different branches of mathematics that are somehow "probabilistic" are then unified in non-commutative probability. For example, this framework allows us to see classical probability, matrix algebra and random matrix theory as different examples of the same kind of object. The distinction between what is called free probability and what is called non-commutative probability is the kind of independence adopted: Free probability is the branch of non-commutative probability that takes freeness as its notion of independence. Other kinds of non-commutative probability are given by classical independence, Boolean independence, monotone independence and others, however these alternative theories (except for classical probability) are less interesting and useful than free probability; furthermore, freeness is the only one that satisfies certain properties that are very desirable.

Many notions of "non-commutative analogs" were being used for a long time, for example in the study of operator algebras and in quantum physics. The first systematic development of non-commutative probability, however, was done by Voiculescu in his treatment, [16], of free probability. He gave us the notion of freeness, or free independence (Chapter 2), analogous to independence in the sense of classical probability. Roland Speicher is a key player in the development of free probability; he revealed a vast combinatorial side of the theory by defining the free cumulants via non-crossing partitions. This allows us to give a proof of the free central limit theorem (section 2.4), which is far simpler than Voiculescu's analytic approach based on the $\mathcal{R}$-transform (section 2.6). Free probability is now a growing and quite accessible research field of its own. The usefulness of the theory lies mostly in the interplay between the analytical and combinatorial sides of the theory.

In this chapter, we introduce the basics of non-commutative probability. Free probability and its combinatorics will be the subject of chapter 2.

1.1 Non-Commutative Probability Spaces

Recall that in classical probability theory we have a sample space Ω, with a collection $\mathcal{F}$ of measurable sets, or events, and a probability measure $\mathbf{P}$. Classical random variables are then, by definition, measurable functions $X : \Omega \to \mathbb{C}$. Since multiplication of random variables is defined pointwise, classical random variables are always commutative: $X(\omega)Y(\omega) = Y(\omega)X(\omega), \quad \forall \omega \in \Omega$. The idea behind of non-commutative probability is to redefine random variables in an abstract way so that they are allowed to be non-commutative.

Recall that an *algebra over a field*, or an *algebra,* is a vector space equipped with a bilinear multiplication operation, that is, a multiplication which is linear on both arguments. More precisely, a vector space $\mathcal{A}$ over a field $\mathbb{F}$ is an algebra, if it has a binary multiplication which gives, for each $a, b, c \in \mathcal{A}$, and $\alpha \in \mathbb{F}$

$$\alpha(ab) = (\alpha a)b = a(\alpha b) \qquad a(b+c) = ab + ac \qquad (a+b)c = ac + bc.$$

We say that an algebra is *unital* if it contains a multiplicative identity, or unit. That is, an algebra $\mathcal{A}$ is unital if there exists an element $1_{\mathcal{A}} \in \mathcal{A}$ such that for all $a \in \mathcal{A}$, $1_{\mathcal{A}}a = a1_{\mathcal{A}} = a$. The unit of $\mathcal{A}$ will also be denoted 1, when there is no danger of ambiguity.

1.1 Definition. A *non-commutative probability space* is a pair $(\mathcal{A}, \varphi)$, consisting of a unital algebra $\mathcal{A}$ over $\mathbb{C}$ and a *unital* linear functional $\varphi : \mathcal{A} \to \mathbb{C}$, where *unital* means $\varphi(1_{\mathcal{A}}) = 1 \in \mathbb{C}$.

Elements of $\mathcal{A}$ are called *random variables*, and φ is called an *expectation functional*. Given an element a in the non-commutative probability space $(\mathcal{A}, \varphi)$, the values $\varphi(a^n)$, for $n \geq 1$ are called the *moments* of a.

Throughout this work, the integrals we use are written in the notation of Lebesgue integration and measure theory. See [5].

1.2 Example. Consider classical probability theory, where random variables are, by definition, measurable functions $X : \Omega \to \mathbb{C}$ on a space Ω with a probability measure $\mathbf{P}$, and we have the usual expectation function $\mathbf{E}$ defined by

$$\mathbf{E}[X] = \int_{\Omega} X d\mathbf{P}.$$

In order for $\mathbf{E}$ to be an expectation functional in the sense of non-commutative probability, we require that $\mathbf{E}[X]$ be finite for all random variables X. The algebra of measurable functions that we will use is $L^{\infty}(\Omega, \mathbf{P})$, the set of all *bounded* $\mathbf{P}$-measurable functions on Ω. (One could rather take the space $L^{\infty-}(\Omega, \mathbf{P})$ of $\mathbf{P}$-measurable functions on Ω, not necessarily bounded, but with *finite moments of all orders.*) Then

$$\Big(L^{\infty}(\Omega, \mathbf{P}),\ \mathbf{E}[\cdot]\Big)$$

is a non-commutative probability space with unit being the constant-1 function. Notice, however, that all random variables in this space are necessarily commutative, which makes example somewhat uninteresting.

1.3 Example. Consider now the algebra $\mathcal{M}_n(\mathbb{C})$ of all $n \times n$ matrices with complex coefficients. We will take for expectation functional the *normalized trace,*

$$\mathrm{tr}_n := \frac{1}{n}\,\mathrm{Tr},$$

where Tr is the trace of a matrix, so that tr_n is unital, $\mathrm{tr}_n(I) = 1$, the unit I being the identity matrix. So $(\mathcal{M}_n(\mathbb{C}), \mathrm{tr}_n)$ is a non-commutative probability space. Note the special property of traces, namely $\mathrm{tr}_n(ab) = \mathrm{tr}_n(ba)$ for any $a, b \in \mathcal{M}_n(\mathbb{C})$.

By diagonalizing a matrix $a \in \mathcal{M}_n(\mathbb{C})$, (or using its Jordan canonical form,) we can see that the expectation is given by the normalized sum of its eigenvalues,

$$\mathrm{tr}_n(a) = \frac{1}{n}\sum_{i=1}^{n} a_{ii} = \frac{1}{n}\sum_{i=1}^{n} \lambda_i,$$

or we could say in other words, the *average* eigenvalue.

1.4 Definition. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. Then φ is said to be a *trace* (or to be *tracial*) if and only if

$$\varphi(ab) = \varphi(ba) \quad \text{for all } a, b \in \mathcal{A}.$$

The two above examples are often combined in the theory of *random matrices*, as in the following example.

1.5 Example. Given a non-commutative probability space $(\mathcal{A}, \varphi)$, we may take

$$\Big(\mathcal{M}_n(\mathcal{A}),\ \varphi \circ \mathrm{tr}_n\Big),$$

the $n \times n$ matrices with entries in $\mathcal{A}$, and the expectation, as above would be given by the expectation of the eigenvalues. (Some authors write $\varphi \otimes \mathrm{tr}_n$, since we have an isomorphism $\mathcal{M}_n(\mathcal{A}) \cong \mathcal{A} \otimes \mathcal{M}_n(\mathbb{C})$.) So for a random matrix $[a_{ij}]_{i,j} \in \mathcal{M}_n(\mathcal{A})$, with entries $a_{ij} \in \mathcal{A}$, we would have

$$(\varphi \circ \mathrm{tr}_n)([a_{ij}]_{i,j}) = \varphi\left(\frac{1}{n}\sum_{i=1}^{n} a_{ii}\right) = \frac{1}{n}\sum_{i=1}^{n} \varphi(a_{ii}).$$

1.6 Definition. A $*$-*probability space* is a non-commutative probability space, $(\mathcal{A}, \varphi)$, in which $\mathcal{A}$ is a $*$-*algebra*, (i.e., there is a linear anti-involution $a \mapsto a^*$ on $\mathcal{A}$, $(a^*)^* = a$, $(ab)^* = b^*a^*$,) and φ is a *positive* functional in the sense that $\varphi(a^*a) \geq 0$ for all $a \in \mathcal{A}$.

All of the examples above are also $*$-algebras, with the $*$-action given, respectively, by complex conjugation of a complex-valued function, and by conjugate-transpose of a (deterministic) matrix or a random matrix. In the spirit of matrix theory, we define the following vocabulary for general $*$-random variables.

1.7 Definition. Let $(\mathcal{A}, \varphi)$ be a $*$-probability space and $a \in \mathcal{A}$.

(1) We say a is *normal* if $a^*a = aa^*$.

(2) We say a is *unitary* if $a^*a = aa^* = 1_{\mathcal{A}}$ (i.e. $a^* = a^{-1}$).

(3) And we say a is *selfadjoint* if $a^* = a$.

In a $*$-probability space $(\mathcal{A}, \varphi)$, we define $*$-*moments* of a random variable $a \in \mathcal{A}$ to be all expressions of the form

$$\varphi\left(a^{\varepsilon(1)}a^{\varepsilon(2)}\ldots a^{\varepsilon(m)}\right),$$

for $m \geq 0$, where for each $k \in [m]$, $\varepsilon(k) \in \{1, *\}$, and the case $m = 0$ is taken to be $\varphi(1)$. (Throughout this thesis we use the notation $[n] = \{1, 2, \cdots, n\}$, for any $n \in \mathbb{N}$.) Then we can see that every $*$-moment can be simplified to the form

$$\varphi\left(\left(a^{\varepsilon(1)}\right)^{i(1)}\left(a^{\varepsilon(2)}\right)^{i(2)}\cdots\left(a^{\varepsilon(n)}\right)^{i(n)}\right),$$

with $n \geq 0$, $(a^{\varepsilon(k)})^{i(k)} \neq 1$, and $\varepsilon(1) \neq \varepsilon(2) \neq \cdots \neq \varepsilon(n)$.

Furthermore, if a is a normal random variable it has moments of the form

$$\varphi\left((a^*)^m a^n\right), \quad \text{for } m, n \geq 0,$$

if a is unitary the moments simplify to

$$\varphi\left(a^n\right), \quad \text{for } n \in \mathbb{Z},$$

and if a is selfadjoint its moments have the form

$$\varphi\left(a^n\right), \quad \text{for } n \geq 0.$$

We follow the vocabulary of classical probability in speaking of the *mean* and *variance* of a random variable. Given any random variable $x \in (\mathcal{A}, \varphi)$ we call $\varphi(x)$ the *mean* and $\varphi(x^2) - \varphi(x)^2$ the *variance* of x.

1.8 Definition. Let $(\mathcal{A}, \varphi)$ be a $*$-probability space. Then φ is *faithful* if and only if for $a \in \mathcal{A}$,

$$\varphi(a^*a) = 0 \quad \text{implies} \quad a = 0.$$

The following example, on group algebras, is important for free probability in general. This is because group algebras are the main motivation for the definition of free independence, to be given in section 2.2. In fact, it could be said that free probability was invented by Voiculescu to study Von-Neumann algebras on free groups, as he explains in [19].

1.9 Example. Let G be any group. We define the *group algebra* $\mathbb{C}G$ to be the algebra of all formal finite linear combinations of elements in G,

$$\sum_{g \in G} \alpha_g g \quad (\alpha_g \in \mathbb{C}),$$

with the ordinary addition of linear combinations, and with a distributive multiplication defined in the natural way:

$$\left(\sum_{g \in G} \alpha_g g\right)\left(\sum_{h \in G} \beta_h h\right) = \sum_{f \in G}\left(\sum_{gh=f} \alpha_g \beta_h\right) f.$$

It is instructive to see this as a generalization of the algebra of polynomials by seeing G as, in a sense, the abstract group of "monomials", in which case we could take the expectation τ_G to be seen as "evaluation at 0", or the coefficient of the identity element $e \in G$. This gives a $*$-probability space

$$\left(\mathbb{C}G, \tau_G\right),$$

with the expectation functional given by

$$\tau_G\Big(\sum \alpha_g g\Big) = \alpha_e,$$

and with the $*$-action given by

$$\Big(\sum \alpha_g g\Big)^* = \Big(\sum \overline{\alpha_g} g^{-1}\Big).$$

It is then easy to see that τ_G is indeed a positive functional. Furthermore, τ_G is a faithful trace. This is because for a random variable $a = \sum \alpha_g g$ in $\mathbb{C}G$, we have

$$a^* a = \sum_{f \in G} \Big(\sum_{g^{-1}h = f} \overline{\alpha_g} \alpha_h \Big) f,$$

and so

$$\tau_G(a^* a) = \Big(\sum_{g^{-1}h = e} \overline{\alpha_g} \alpha_h \Big) = \Big(\sum_{g = h} \overline{\alpha_g} \alpha_h \Big) = \Big(\sum_{g \in G} |\alpha_g|^2 \Big) \geq 0$$

(hence φ is positive), with equality only if $a = 0$ (hence φ faithful), and to show that φ is tracial we note that $\tau_G(ab) = \tau_G(ba)$ for all $a, b \in \mathbb{C}G$, because for group elements $g, h \in G$, we have $gh = e$ if and only if $hg = e$.

In light of the above example, we may define polynomials, commutative and non-commutative, as elements of a group algebra as follows.

1.10 Definition.

(1) The algebra of (complex) polynomials in one variable, denoted $\mathbb{C}[X]$ is defined to be the group algebra of the multiplicative infinite cyclic group generated by an element X.

(2) More generally, $\mathbb{C}[X_1, X_2, \ldots, X_n]$ denotes the algebra of polynomials in n variables, defined to be the group algebra of the (multiplicative) free Abelian group generated by the elements $X_1, X_2, \ldots, X_n$.

(3) $\mathbb{C}\langle X_1, \ldots, X_n \rangle$ is the algebra of non-commutative (or free) polynomials, defined as the group algebra of the free group generated by $X_1, \ldots, X_n$.

1.11 Remark. Let $\mathbb{F}_n$ be the free group generated by n non-commuting elements. We have the isomorphisms

$$\begin{aligned} \mathbb{C}[X_1, \ldots, X_n] &\cong \mathbb{C}\mathbb{Z}^n, \\ \mathbb{C}\langle X_1, \ldots, X_n \rangle &\cong \mathbb{C}\mathbb{F}_n. \end{aligned}$$

1.12 Example. Recall that a vector space $\mathcal{H}$ with an inner product $\langle \cdot, \cdot \rangle$ is called a *Hilbert Space* if $\mathcal{H}$ forms a *complete metric space*, with the norm defined for $v \in \mathcal{H}$ by $\|v\| = \sqrt{\langle v, v \rangle}$. Let $\mathcal{H}$ be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and $B(\mathcal{H})$ the algebra of bounded linear operators on $\mathcal{H}$. If we take a vector $\xi \in \mathcal{H}$ with $\|\xi\| = 1$, we get a functional φ_ξ given by $\varphi_\xi(T) = \langle T\xi, \xi \rangle$ for all $T \in B(\mathcal{H})$. Then we have a $*$-probability space

$$\Big(B(\mathcal{H}), \varphi_\xi\Big),$$

where for each $T \in B(\mathcal{H})$, T^* is the unique operator determined by the property that $\langle u, T^* v \rangle = \langle Tu, v \rangle$ for all $u, v \in \mathcal{H}$. A concrete example of this kind is given in example 1.3.

1.2 Distribution of Random Variables

The central idea of classical probability theory is reducing random variables to their distributions and joint distributions. In a similar way, the central idea in non-commutative probability is the reduction of random variables to a suitable notion of distribution and joint distribution. Joint distributions will be discussed in section 2.1; here, we give the definitions and examples concerning the notion of a $*$-distribution for a single random variable. There are two such notions, both essential to the theory: *distribution in the analytical sense*, for the special case of normal random variables, and *distribution in the algebraic sense*, for the general case. Intuitively, the distribution of a random variable should be an object that captures the information of all its moments while ignoring the specific nature of the random variable or the space where it lives. In other words a "standardized" way of keeping track of $*$-moments. This is done via free non-commutative polynomials.

1.13 Definition. If $(\mathcal{A}, \varphi)$ is a $*$-probability space and $a \in \mathcal{A}$, then the *$*$-distribution* of a (in the algebraic sense) is the function $\mu : \mathbb{C}\langle X, X^*\rangle \to \mathbb{C}$ defined for all polynomials $P \in \mathbb{C}\langle X, X^*\rangle$ by

$$\mu(P) = \varphi\left(P\left(a, a^*\right)\right).$$

Two random variables are said to be *identically distributed,* or to have *the same distribution,* if there is a function $\mu : \mathbb{C}\langle X, X^*\rangle \to \mathbb{C}$ such that μ is the distribution of both random variables.

Here a distribution (i.e., $*$-distribution, the terms will be used interchangeably) is simply the collection of all moments (and linear combinations of moments), but parametrized by non-commutative polynomials. In this way two corresponding moments can be compared from different random variables. It should be noted that X and X^* are just two different symbols, or more precisely, they are two free, non-commutative indeterminate variables, so that we might have written $\mathbb{C}\langle X, Y\rangle$ instead of $\mathbb{C}\langle X, X^*\rangle$.

For example, let $P \in \mathbb{C}\langle X, X^*\rangle$ be the free polynomial $P(X, X^*) = XX^* - X^*X$. If $a \in (\mathcal{A}, \varphi)$ has $*$-distribution $\mu : \mathbb{C}\langle X, X^*\rangle \to \mathbb{C}$, then we would have that

$$\mu(P) = \varphi(aa^*) - \varphi(a^*a).$$

And of course, if in this case a happens to be a normal random variable, or if φ happens to be a trace, then this would simplify to zero.

Analytic Distributions

If a is a normal random variable (i.e. $a^*a = aa^*$), then all the moments of a simplify to the form $\varphi(a^k a^{*\ell})$. In this case we can take "distribution" to mean essentially the same as in classical probability: a measure that gives moments by way of integration:

1.14 Definition. Let $(\mathcal{A}, \varphi)$ be a $*$-probability space and $a \in \mathcal{A}$. If μ is a probability measure on $\mathbb{C}$ with compact support, then μ is said to be the $*$-distribution of a (in the analytic sense) if and only if

$$\varphi(a^k a^{*\ell}) = \int_{\mathbb{C}} z^k \bar{z}^\ell d\mu(z) \quad \text{for all } k, \ell \in \mathbb{N}.$$

1.15 Remark. For an arbitrary random variable a, there may exist no analytic distribution, however if a distribution exists, then it is unique. This follows from the Stone-Weierstrass theorem on $\mathbb{C}$.

1.16 Remark. If a is a self-adjoint random variable with distribution μ, then $\text{supp}(\mu) \subseteq \mathbb{R}$. This can be seen by the following computation.

$$\begin{aligned}\int_{\mathbb{C}} |z - \bar{z}|^2 \, d\mu &= \int_{\mathbb{C}} (z - \bar{z})(\bar{z} - z) \, d\mu \\ &= \int_{\mathbb{C}} 2z\bar{z} - zz - \bar{z}\bar{z} \, d\mu \\ &= 2\varphi(aa^*) - \varphi(a^*a^*) - \varphi(aa) = 0.\end{aligned}$$

Since $z \mapsto |z - \bar{z}|$ is a continuous non-negative function, it must vanish on the support of μ. But $z - \bar{z} = 0$ if and only if $z \in \mathbb{R}$, so we have that

$$\text{supp}(\mu) \subseteq \mathbb{R}.$$

Then, for selfadjoint random variables the condition that makes μ a $*$-distribution of a is

$$\varphi(a^k) = \int_{\mathbb{R}} z^k d\mu(z) \quad \text{for all } k \in \mathbb{N}.$$

Examples of Analytic Distributions

The reader should observe that an algebraic distribution provides no additional information or tools for studying random variables; it only gives us a standard vocabulary, as it were, for talking about its moments. An analytic distribution, on the other hand, provides new analytical tools for the computation of its moments, and of the moments of other related variables, as will be seen below. When we study normal random variables, therefore, we always prefer to have a measure μ to serve as the distribution in the analytical sense. Therefore, an important question to answer about normal random variables is whether a distribution exists. This question only makes sense when we are talking about distributions in the analytic sense, (since distributions in the algebraic sense always exist,) so in the rest of this work we will often refer to analytic distributions simply as distributions, unless there is danger of ambiguity. The following are examples of random variables for which we compute analytic distributions.

1.17 Example. Consider the non-commutative probability space $(L^\infty(\Omega, \mathbf{P}), \mathbf{E})$ of (bounded) classical random variables. Then, for any random variable $a : \Omega \to \mathbb{C}$ in $L^\infty(\Omega, \mathbf{P})$, there is a $*$-distribution of a, namely the probability measure ν on $\mathbb{C}$ called "the distribution of a" in classical probability. Recall that ν is defined on any Borel-measurable set $B \subseteq \mathbb{C}$ by

$$\nu(B) = \mathbf{P}(a^{-1}(B)) = \mathbf{P}\{\omega \in \Omega : a(\omega) \in B\}. \tag{1.1}$$

Since $a \in L^\infty(\Omega, \mathbf{P})$ is bounded by definition, we must have that ν has compact support. Indeed, if a is bounded, then for some $r \in \mathbb{R}$, $\{a(\omega) : \omega \in \Omega\}$ is contained in $D_r(0)$, the closed disc centered at 0 with radius r; then, any open set disjoint from $D_r(0)$ must have measure zero (by the definition of ν), which implies $\text{supp}(\nu) \subseteq D_r(0)$. Since $\text{supp}(\nu)$ is closed and bounded in $\mathbb{C}$, it is therefore compact. If we denote by f the characteristic function of $B \subseteq \mathbb{C}$ (i.e. $f(z)$ is 1 if $x \in B$ and 0 if $x \notin B$), then equation 1.1 can be rewritten as an integral as

$$\int_{\mathbb{C}} f(z)\ d\nu(z) = \int_{\Omega} f(a(\omega))\ d\mathbf{P}(\omega). \tag{1.2}$$

By using a standard argument from real analysis and Lebesgue integration, (see [5],) we may extend equation (1.2) to be valid for any bounded measurable function f. Namely, one first takes linear combinations of characteristic functions, and then (1.2) is valid with f being any step function, and by taking limits we may conclude (1.2) for any bounded measurable function f. In particular, if $k, \ell \in \mathbb{N}$, and we take f to be

$$f(z) = \begin{cases} z^k \bar{z}^\ell, & |z| \leq r \\ 0, & \text{else,} \end{cases}$$

(so that f is bounded,) then we have

$$\int_{\mathbb{C}} z^k \bar{z}^\ell\ d\nu(z) = \int_{\mathbb{C}} f(z)\ d\nu(z) = \int_{\Omega} f(a(\omega))\ d\mathbf{P}(\omega) = \int_{\Omega} a^k a^{*\ell}\ d\mathbf{P} = \mathbf{E}[a^k a^{*\ell}].$$

(The first equality holds because $f(z)$ and the polynomial $P(z) = z^k \bar{z}^\ell$ are equal "almost everywhere" with respect to ν, the second equality by definition of ν, the third because a is bounded by r, and the last one by definition of $\mathbf{E}$.) Hence we have proved that ν is the analytic distribution of a in the framework of non-commutative probability.

1.18 Example. We now take a normal matrix $a \in \mathcal{M}_n(\mathbb{C})$, i.e. $aa^* = a^*a$, and we compute its $*$-moments in $(\mathcal{M}_n(\mathbb{C}), \text{tr}_n)$,

$$\text{tr}_n(a^k a^{*\ell}).$$

By simultaneously diagonalizing a and a^* as $a = pdp^{-1}$ and $a^* = pd^*d^{-1}$, (note that diagonalizable matrices can be simultaneously diagonalized if the commute with one another,) we have

$$\text{tr}(a^k a^{*\ell}) = \text{tr}(d^k d^{*\ell}) = \frac{1}{n} \sum_{i=1}^{n} \lambda_i^k \bar{\lambda}_i^\ell = \int_{\mathbb{C}} z^k \bar{z}^\ell\ d\mu,$$

where $\mu = \sum_{i=1}^n \frac{1}{n} \delta_{\lambda_i}$, and δ_c is the Dirac point-mass measure at $c \in \mathbb{C}$. Then the probability measure μ is the distribution of a and is called the *eigenvalue distribution*, where $\lambda_1, \cdots, \lambda_n$ are the eigenvalues of a (with multiplicities).

1.19 Example. Let G be a group and $g \in G$ be an element of infinite order. If we take g to be a random variable in $(\mathbb{C}G, \tau_G)$, then it is a unitary random variable (i.e. $g^* = g^{-1}$) and its moments are given by

$$\tau_G(g^k) = \begin{cases} 1, & k = 0 \\ 0, & k \in \mathbb{Z} \setminus 0. \end{cases}$$

Let μ be the uniform probability measure supported on the unit circle $\mathbb{T} \subseteq \mathbb{C}$ (the so-called "Haar measure"). That is, the measure μ on $\mathbb{C}$, is supported on $\mathbb{T} = \{e^{it} : t \in [0, 2\pi)\}$, and determined by

$$e^{it} d\mu(t) = \frac{1}{2\pi} e^{it} dt \quad \text{for} \quad t \in [0, 2\pi).$$

Then, we claim μ is the distribution for the element g of infinite order. Indeed, we may parametrize the unit circle as $z = e^{it},\ 0 \leq t < 2\pi$, so we have, for $k \in \mathbb{Z}$

$$\int_{\mathbb{C}} z^k \ d\mu = \frac{1}{2\pi} \int_0^{2\pi} e^{kit} \ dt = \begin{cases} 1, & k = 0 \\ 0, & k \in \mathbb{Z} \setminus \{0\}. \end{cases}$$

Now, similarly, consider the case of an element g of order p. The moments of g are given by

$$\tau_G(g^k) = \begin{cases} 1, & k \equiv 0 \mod p \\ 0, & k \in \mathbb{Z} \setminus p\mathbb{Z}. \end{cases}$$

Then, the measure of g is $\mu = \frac{1}{p} \sum_{n=0}^{p-1} \delta_{\omega_n}$, where $\omega_0, \cdots, \omega_{p-1}$ are the distinct p-th roots of unity $\omega_n = e^{n\frac{2\pi i}{p}}$. We have, when k is not a multiple of p

$$\int_{\mathbb{C}} z^k \ d\mu = \frac{1}{p} \sum_{n=0}^{p-1} e^{nk\frac{2\pi i}{p}} = \frac{1}{p} \sum_{n=0}^{p-1} (e^{\frac{2k\pi i}{p}})^n = \frac{1}{p} \left(\frac{1 - \left(e^{\frac{2k\pi i}{p}}\right)^p}{1 - e^{\frac{2k\pi i}{p}}} \right) = \frac{1 - 1}{p\left(1 - e^{\frac{2k\pi i}{p}}\right)} = 0$$

And clearly, we have, when k is a multiple of p, that $e^{\frac{2\pi i}{p}k} = 1$ and therefore we get $\int z^k \ d\mu = 1$, as expected. This is an important distribution in free probability and in some applications. We will see it again in example (1.21) below. Random variables with these distributions receive the name of Haar unitary and p-Haar unitary, due to Voiculescu in [17].

1.20 Definition. Let $(\mathcal{A}, \varphi)$ be a $*$-probability space.

(1) A random variable $u \in \mathcal{A}$ is called a *Haar unitary* if it is unitary and its moments are given by

$$\varphi(u^k) = 0, \quad \text{for } k \in \mathbb{Z} \setminus \{0\}.$$

(2) A random variable $u \in \mathcal{A}$ is called a *p-Haar unitary* if it is unitary and its moments are given by

$$\varphi(u^k) = \begin{cases} 1, & k \equiv 0 \mod p \\ 0, & k \in \mathbb{Z} \setminus p\mathbb{Z}. \end{cases}$$

One important task in free probability is the following. Given a random variable a, consider the self-adjoint random variable $a + a^*$. We would like to answer two questions:

(1) What are the moments of $a + a^*$? More specifically, is there a "nice" formula for computing the $*$-moments?

(2) Is there a distribution μ of $a + a^*$?

The following two examples introduce some of the computation used often in non-commutative probability. The first involves a unitary (normal) random variable, and the second is similar, but involves a non-normal random variable.

1.21 Example. Let $(\mathcal{A}, \varphi)$ be a $*$-probability space and $u \in \mathcal{A}$ be a Haar unitary. We will consider the self-adjoint random variable $u + u^*$ and answer the two questions above. We begin by finding a formula for the moments $\varphi((u+u^*)^n)$. Since u commutes with u^*, we simplify $(u+u^*)^n$ using the binomial expansion.

$$(u+u^*)^n = \sum_{k=0}^{n} \binom{n}{k} u^k (u^*)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} u^{2k-n}.$$

But since $\varphi(u^{2k-n})$ vanishes except in the case $2k-n=0$, then we have

$$\varphi\big((u+u^*)^k\big) = \sum_{k=0}^{n} \binom{n}{k} \varphi\big(u^{2k-n}\big) = \begin{cases} \binom{n}{n/2}, & n \text{ even} \\ 0, & n \text{ odd.} \end{cases} \tag{1.3}$$

The next question is whether $u+u^*$ has a distribution. The answer is yes, and we compute it as follows. We need a measure μ such that $\int_{\mathbb{R}} t^n d\mu = \varphi((u+u^*)^n)$. There are techniques to determine μ from the moments computed in equation (1.3), namely by using the Cauchy transform, (see section 1.4,) but this is more complicated. Instead, we use the known distribution of u to derive the distribution of $u+u^*$. Since $\varphi(u^n) = \int_0^{2\pi} e^{int} dt$, then it follows that $\varphi((u+u^*)^n) = \frac{1}{2\pi}\int_0^{2\pi} (e^{it}+e^{it})^n dt$. This is because

$$\begin{aligned} \varphi\Big((u+u^*)^n\Big) &= \varphi\left(\sum_{k=0}^{n} \binom{n}{k} u^k (u^*)^{n-k}\right) \\ &= \sum_{k=0}^{n} \binom{n}{k} \varphi\left(u^k (u^*)^{n-k}\right) \\ &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{2\pi} \int_0^{2\pi} (e^{it})^k (e^{-it})^{n-k} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{k=0}^{n} \binom{n}{k} (e^{it})^k (e^{-it})^{n-k}\right) dt = \frac{1}{2\pi} \int_0^{2\pi} (e^{it}+e^{-it})^n dt. \end{aligned}$$

Now, we simplify,

$$\frac{1}{2\pi} \int_0^{2\pi} (e^{it}+e^{-it})^n dt = \frac{1}{2\pi} \int_0^{2\pi} (2\cos(t))^n dt = \frac{1}{2\pi} 2 \int_0^{\pi} (2\cos(t))^n dt = \frac{1}{\pi} \int_0^{\pi} (2\cos(t))^n dt.$$

Notice we are not trying to evaluate the integral, we already did that by computing the moments in (1.3). We are trying to determine the distribution μ by rewriting the integral as $\int_{\mathbb{R}} t^n K(t) dt$, (the function $K : \mathbb{R} \to \mathbb{R}$ is often referred to as a kernel of integration, or as the density of μ) in which case μ will be given by $d\mu = K(t)dt$. We now use the substitution $v = 2\cos t$, i.e.

$$t = \arccos v, \qquad\qquad dt = \frac{dv}{-\sqrt{2^2 - v^2}},$$

and the integral becomes

$$\frac{1}{\pi}\int_0^{\pi}(2\cos(t))^n dt = \frac{1}{\pi}\int_2^{-2} v^n \frac{dv}{-\sqrt{4-v^2}} = \frac{1}{\pi}\int_{-2}^{2} v^n \frac{dv}{\sqrt{4-v^2}}.$$

So in conclusion we have the equality of the four expressions:

$$\varphi((u+u^*)^n) = \frac{1}{\pi}\int_0^{\pi}(2\cos(t))^n dt = \int_{\mathbb{R}} t^n K(t)dt = \begin{cases} \binom{n}{n/2}, & n \text{ even} \\ 0, & n \text{ odd.} \end{cases} \tag{1.4}$$

where

$$K(t) = \begin{cases} \frac{1}{\pi\sqrt{4-t^2}}, & |t|<2 \\ 0, & |t|\geq 2 \end{cases}$$

is the so-called "arcsine density". Therefore μ, given by $d\mu = K(t)dt$, is the measure with compact support $\text{supp}(\mu) = [-2,2]$ and it is the distribution of $u+u^*$.

The following example, due to Voiculescu [16], is analogous to the previous one, but it involves a non-normal random variable a, instead of the normal random variable u studied above.

1.3 The Semicircular Distribution

Consider a $*$-probability space $(\mathcal{A},\varphi)$ and a random variable $a\in\mathcal{A}$ with the following properties

(i) $a^*a = 1$, but $aa^* \neq 1$.

(ii) $\varphi(a^k a^{*\ell}) = 1$ if $k=\ell=0$, and 0 otherwise.

(iii) The set $\{a^k a^{*\ell} | k,\ell \geq 0\}$ is linearly independent.

(iv) $\mathcal{A}$ is generated (as a unital $*$-algebra) by a. I.e., $\mathcal{A} = \text{alg}(1,a,a^*)$.

The $*$-moments of a are indeed determined by (i) and (ii). To show this, we need to show that all $*$-moments can be written in the form described in (ii). This can be seen by noting that all elements of $\mathcal{A}$ can be written in the form $a^k a^{*\ell}$, because by using property (i), we have

$$a^k a^{*\ell} a^i a^{*j} = \begin{cases} a^{k+i-\ell}a^{*\ell}, & \ell < i \\ a^k a^{*j}, & \ell = i \\ a^k a^{*(j+\ell-j)}, & \ell < i. \end{cases}$$

Thus we may conclude that $\mathcal{A} = \text{span}\{a^k a^{*\ell} | k,\ell\geq 0\}$.

Before continuing, we provide a concrete realization of a random variable with these properties. Take the Hilbert space $\mathbf{c}_c$ of infinite sequences with finite support, (also denoted by $\ell(\mathbb{N}\cup\{0\})$,) i.e.

$$\mathbf{c}_c = \left\{ (x_0, x_1, x_2, x_3, \cdots) \;\middle|\; \begin{array}{lll} \forall n \in \mathbb{N}\cup\{0\} & & x_n \in \mathbb{C}, \\ \exists N \in \mathbb{N}\cup\{0\} & \text{s.t.} \quad x_n = 0 \quad \text{for} & n \geq N \end{array} \right\},$$

equipped with the inner product defined for all $x = (x_0, x_1, x_2, \cdots)$, and $y = (y_0, y_1, y_2, \cdots)$ by

$$\langle x, y\rangle = \sum_{n=0}^{\infty} x_n \bar{y}_n.$$

For any $n \in \mathbb{N}\cup\{0\}$, let $e_n = (0, 0, \cdots, 0, 1, 0, \cdots)$, having $x_n = 1$ and zeroes elsewhere. The set $\{e_n | n \in \mathbb{N}\cup\{0\}\}$ is an orthonormal basis. Then, as in example, 1.12 we have the $*$-probability space $(B(\mathbf{c}_c), \varphi_{e_0})$ of bounded linear operators with the expectation function defined on any operator $T : \mathbf{c}_c \to \mathbf{c}_c$ by

$$\varphi_{e_0}(T) = \langle Te_0, e_0\rangle.$$

More precisely, if $T(e_0) = (y_0, y_1, y_2, \cdots)$, then $\varphi_{e_0}(T) = y_0$. Now, consider the *one-sided shift operator* $S : \mathbf{c}_c \to \mathbf{c}_c$ determined by

$$S(e_n) = e_{n+1}$$

It follows that its adjoint is given by

$$S^*(e_0) = 0, \qquad\qquad S^*(e_n) = e_{n-1}, \quad n > 0.$$

Then the reader can check that, if $\mathcal{A} = \text{alg}(S, S^*)$, then the $*$-probability space $(\mathcal{A}, \varphi_{e_0})$ and the random variable S satisfy the three properties listed above, namely

(i) $S^*S = 1$, but $SS^* \neq 1$,

(ii) $\varphi_{e_0}(S^k S^{*\ell}) = 1$ if $k = \ell = 0$, and 0 otherwise, and

(iii) the set $\{S^k S^{*\ell} | k, \ell \geq 0\}$ is linearly independent.

Hence the discussion that follows applies to this concrete random variable, as well as to any random variable with the same distribution. We now continue to work in the general framework of $a \in (\mathcal{A}, \varphi)$ with the properties stated above. Since a is non-normal, we do not attempt to find a distribution for it. Instead, as in example 1.21, we will now focus on the selfadjoint random variable $a + a^*$ and attempt to answer the same two questions: Is there a "nice" formula for the moments? And is there a compactly-supported distribution of $a + a^*$?

First we compute the moments of $a + a^*$.

$$\begin{aligned} \varphi\left((a+a^*)^n\right) &= \varphi\left(\sum_{\varepsilon(1),\cdots,\varepsilon(n)\in\{1,*\}} a^{\varepsilon(1)} a^{\varepsilon(2)} \cdots a^{\varepsilon(n)}\right) \\ &= \sum_{\varepsilon(1),\cdots,\varepsilon(n)\in\{1,*\}} \varphi\left(a^{\varepsilon(1)} a^{\varepsilon(2)} \cdots a^{\varepsilon(n)}\right) \end{aligned}$$

Most of the moments $\varphi\left(a^{\varepsilon(1)}a^{\varepsilon(2)}\cdots a^{\varepsilon(n)}\right)$ will vanish, and those remaining will be equal to 1. So, each moment $\varphi\left((a+a^*)^n\right)$ will count the number of possible choices of a function $\varepsilon : [n] \to \{1, *\}$ so that $a^{\varepsilon(1)}a^{\varepsilon(2)}\cdots a^{\varepsilon(n)}$ reduces to 1. That is,

$$\varphi\left((a+a^*)^n\right) = \#\left\{\varepsilon : [n] \to \{1,*\} \quad \middle| \quad \varphi\left(a^{\varepsilon(1)}a^{\varepsilon(2)}\cdots a^{\varepsilon(n)}\right) = 1\right\}$$

Here, for a set A, $\#A$ denotes the number of elements in A. Let us now solve this counting problem. The criterion for the moment reducing to 1 and not to 0 is, roughly, that before every factor of a, there is one factor of a^*, which will cancel it, and that all factors will end up cancelled in this way. To express this more precisely, we look at functions $\gamma : [n] \to \{-1, 1\}$ (referred to as *paths with n steps*) instead of the functions $\varepsilon : [n] \to \{1, *\}$. with a one to one correspondence between them given by replacing each $*$ in ε by a 1 in γ, and each 1 in ε by a -1 in γ. The characterization of the nonvanishing terms is whether their corresponding path γ satisfies the following conditions:

(1) For each $k \in [n]$, $\sum_{i=1}^{k} \gamma(i) \geq 0$.

(2) $\sum_{i\in[n]} \gamma(i) = 0$.

(Condition (1) ensures all a^*'s come before an a to cancel it, and (2) ensures there is at the end the same number of terms a as a^*.) If a function $\gamma : [n] \to \{-1, 1\}$ satisfies these two conditions it is called, by definition, a *Dick path with n steps.*

The Catalan numbers C_k are defined for all $k \in \mathbb{N} \cup \{0\}$ by

$$C_k = \frac{1}{k+1}\binom{2k}{k},$$

or equivalently by the inductive equations

$$\begin{cases} C_0 = C_1 = 1, \\ C_k = \sum_{i=1}^{k} C_{i-1}C_{k-i}. \end{cases}$$

1.22 Proposition. *The number of all possible Dick paths $\gamma : [n] \to \{-1, 1\}$ with n steps is given by the Catalan numbers as follows:*

$$\#\{\gamma : [n] \to \{-1, 1\} \mid \gamma \text{ is a Dick path}\} = \begin{cases} 0, & n \text{ odd}, \\ C_k, & n = 2k. \end{cases}$$

And therefore, the even moments of $a + a^*$ are given by the Catalan numbers:

$$\varphi\left((a+a^*)^n\right) = \begin{cases} 0, & n \text{ odd}, \\ C_k, & n = 2k. \end{cases}$$

For a full treatment of the combinatorics involved and a proof of this fact, see lecture 2 of [12].

Next we are interested in finding a distribution for $a + a^*$. The desired distribution will be given by the *semicircular density of radius 2*:

$$d\mu = \begin{cases} \frac{1}{2\pi}\sqrt{4-t^2}dt, & |t| \leq 2 \\ 0, & |t| > 2. \end{cases}$$

1.23 Definition (Voiculescu [17])**.** Let $(\mathcal{A}, \varphi)$ be a $*$-probability space and $x \in \mathcal{A}$ be a selfadjoint element. If x has the analytic distribution $\frac{2}{\pi r^2}\sqrt{r^2-t^2}dt$, supported on $[-r, r]$ then x is called a *semicircular random variable with radius* r. By computing an integral, one can show that the variance of x is $r^2/4$. Since the variance is an important quantity in non-commutative probability, we often refer to x as a *semicircular variable of variance* σ, where $\sigma = r^2/4$. If a semicircular variable has radius 2, (i.e. variance 1,) we call it a *standard semicircular random variable.*

1.24 Proposition. *The following equality holds for all* $n \in \mathbb{N} \cup \{0\}$.

$$\frac{1}{2\pi}\int_{-2}^{2} t^n\sqrt{4-t^2}dt = \begin{cases} 0, & n \text{ odd}, \\ C_k, & n = 2k. \end{cases}$$

Proof. We use the trigonometric substitution $t = 2\cos\theta, dt = -2\sin\theta d\theta$, so that $\sqrt{4-t^2} = 2\sin\theta$.

$$\begin{aligned} \frac{1}{2\pi}\int_{-2}^{2} t^n\sqrt{4-t^2}dt &= \frac{1}{2\pi}\int_{\pi}^{0} (2\cos\theta)^n(2\sin\theta)(-2\sin\theta\ d\theta) \\ &= \frac{1}{2\pi}\int_{0}^{\pi} (2\cos\theta)^n(2\sin\theta)^2)\ d\theta \\ &= \frac{1}{2\pi}\int_{0}^{\pi} (2\cos\theta)^n(4-(2\cos\theta)^2)\ d\theta \\ &= \frac{4}{2}\left(\frac{1}{\pi}\int_{0}^{\pi}(2\cos\theta)^n\ d\theta\right) - \frac{1}{2}\left(\frac{1}{\pi}\int_{0}^{\pi}(2\cos\theta)^{n+2}\ d\theta\right). \end{aligned}$$

We already computed in equation 1.4 of example 1.21 that

$$\frac{1}{\pi}\int_{0}^{\pi}(2\cos\theta)^n\ d\theta = \begin{cases} \binom{n}{n/2}, & n \text{ even} \\ 0, & n \text{ odd}. \end{cases}$$

So the integral vanishes for odd n, and for even n we have

$$\begin{aligned} \frac{1}{2\pi}\int_{-2}^{2} t^n\sqrt{4-t^2}dt &= 2\left(\frac{1}{\pi}\int_{0}^{\pi}(2\cos\theta)^n\ d\theta\right) - \frac{1}{2}\left(\frac{1}{\pi}\int_{0}^{\pi}(2\cos\theta)^{n+2}\ d\theta\right) \\ &= 2\binom{n}{n/2} - \frac{1}{2}\binom{n+2}{(n+2)/2}. \qquad \text{Let} \quad \frac{n}{2} = p. \\ &= 2\binom{2p}{p} - \frac{1}{2}\binom{2p+2}{p+1} \\ &= 2\left(\frac{(2p)!}{p!\ p!}\right) - \frac{1}{2}\left(\frac{(2p+2)!}{(p+1)!(p+1)!}\right) \\ &= \left(1 - \frac{1}{2}\frac{(2p+2)(2p+1)}{(p+1)(p+1)}\right)\binom{2p}{p} \\ &= \frac{1}{p+1}\binom{2p}{p} = C_p. \end{aligned}$$

□

1.4 The Cauchy Transform

The proof above only verified that the measure μ gives the right moments via integration; it gives no indication of how one might have come to find μ in the first place; note, for example, that we were able to find the measure of $u + u^*$ in example 1.21 because we had a distribution of u to begin with, which we can not have for a because it is non-normal. In this section we study the main method by which we obtain the analytic distribution of a random variable given knowledge of its moments. This method is the Cauchy transform. Let us denote by $\mathbb{C}^+$ the *upper half-plane,* and by $\mathbb{C}^-$ the *lower half-plane*, thus we have

$$\mathbb{C}^+ = \{x + iy \in \mathbb{C} : x \in \mathbb{R}, y > 0\}, \qquad \mathbb{C}^- = \{x + iy \in \mathbb{C} : x \in \mathbb{R}, y < 0\}.$$

1.25 Definition. Let μ be a compactly supported measure on $\mathbb{R}$. Then the *Cauchy transform* of μ is the function $G_\mu : \mathbb{C}^+ \to \mathbb{C}^-$ given by

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z - t} d\mu(t), \quad z \in \mathbb{C}^+.$$

Note that the integral $\int_{\mathbb{R}} \frac{1}{z-t} d\mu(t)$ converges absolutely for any $z \in \mathbb{C}^+$, and it takes values in $\mathbb{C}^-$. (See [5] for clarification.) Indeed, if $t \in \mathbb{R}$ and $z \in \mathbb{C}^+$, then $z - t \in \mathbb{C}^+$, and so $\frac{1}{z-t} \in \mathbb{C}^-$. Finally we just need to show that the integral converges; thus (by the way Lebesgue integration is defined) we will have that $G_\mu(z) \in \mathbb{C}^-$. To do this it suffices to show that $\int_{\mathbb{R}} |\frac{1}{z-t}| d\mu(t)$ is finite. And indeed, if we write $z = x + iy$, we have

$$\int_{\mathbb{R}} \left| \frac{1}{z - t} \right| d\mu(t) = \int_{\mathbb{R}} \frac{1}{|x - t + iy|} d\mu(t) = \int_{\mathbb{R}} \frac{1}{\sqrt{(x-t)^2 + y^2}} d\mu(t) \leq \int_{\mathbb{R}} \frac{1}{\sqrt{y^2}} d\mu(t) = \int_{\mathbb{R}} \frac{1}{y} d\mu(t) = \frac{1}{y},$$

and since $y > 0$, the integral is finite for all $z \in \mathbb{C}^+$.

The important property of the Cauchy transform is that $G_\mu(z)$ can be expressed as an infinite sum using only information about the moments of μ (i.e. the moments of a random variable with distribution μ, or in other words the values of $\int_{\mathbb{R}} t^n d\mu(t)$ for $n \in \mathbb{N} \cup \{0\}$) in the following way. Let $a \in (\mathcal{A}, \varphi)$ be a selfadjoint random variable, and denote its moments by m_n, i.e.

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

Then we can write

$$\int \frac{1}{z - t} d\mu(t) = \frac{1}{z} \int \frac{1}{1 - \frac{t}{z}} d\mu(t) = \frac{1}{z} \int \sum_{n=0}^{\infty} \left(\frac{t}{z} \right)^n d\mu(t).$$

Of course, this power series expansion is only valid where it converges, which is for all values z (z fixed) such that $\left|\frac{t}{z}\right| < 1$. This is satisfied by taking $t \in \text{supp}(\mu)$ and $|z| > \sup\{|t| : t \in \text{supp}(\mu)\}$. In that case, convergence is absolute, and we may integrate term-by-term:

$$\frac{1}{z} \int \sum_{n=0}^{\infty} \left(\frac{t}{z} \right)^n d\mu(t) = \sum_{n=0}^{\infty} \frac{\int t^n d\mu(t)}{z^{n+1}} \qquad |z| > \sup\{|t| : t \in \text{supp}(\mu)\}.$$

So we have concluded that for z large enough, we have

$$G_\mu(z) = \sum_{n=0}^{\infty} \frac{m_n}{z^{n+1}}. \tag{1.5}$$

Hence, given the selfadjoint random variable a and its moments m_n, our technique for finding its distribution μ (when it exists) is to compute its Cauchy transform using equation (1.5) and then to recover the measure μ by applying the following result, known as the *Stieltjes inversion formula.* We denote by $\mathrm{Im}(z)$ the imaginary part of the complex number z, i.e. if $x, y \in \mathbb{R}$ then $\mathrm{Im}(x + iy) = y$.

1.26 Proposition. *Let $t \in \mathbb{R}$ and $y > 0$. If ϕ is any continuous function, then*

$$\lim_{y \to 0^+} \int_{\mathbb{R}} \phi(t) \left[-\frac{1}{\pi} \mathrm{Im} \Big(G_\mu(t + iy) \Big) \right] dt = \int_{\mathbb{R}} \phi \, d\mu. \tag{1.6}$$

Hence the measure μ is recovered from the Cauchy transform by choosing suitable test functions ϕ and evaluating equation (1.6). The proof is a standard computation in real analysis and measure theory. See [14], or [2].

2 Free Probability

In chapter 1 we introduced the framework on which free probability is built. The cornerstone of the theory will be the definition of free independence, or freeness (analogous to independence of classical random variables). As we will see, it is freeness that gives us all the non-trivial structure of free probability that makes it interesting and essentially different from classical probability.

2.1 Joint Moments and Joint Distributions

To begin our discussion about free probability, it will be useful to first define joint moments and joint distributions in a way similar to the definition of distributions and $*$-distributions in the analytical sense.

Instead of just defining the joint distributions and moments of a finite family $a_1, a_2, \cdots, a_s$ we will give the most general statement by taking an arbitrary family of random variables $(a_i)_{i\in I}$, where I is an arbitrary indexing set (possibly infinite or even uncountable). Similarly we use the algebra of non-commutative polynomials in an arbitrary collection of variables,

$$\mathbb{C}\langle X_i : i \in I\rangle,$$

whose elements are (finite) linear combinations of non-commutative monomials of the form $X_{i_1}X_{i_2}\dots X_{i_n}$, for $n \in \mathbb{N}\cup\{0\}$ and $i_1, \cdots, i_n \in I$.

2.1 Definition. Let $(a_i)_{i\in I}$be random variables in a non-commutative probability space $(\mathcal{A}, \varphi)$.

(1) The family of *joint moments* of $(a_i)_{i\in I}$ is the collection of values

$$\big\{\varphi(a_{i_1}a_{i_2}\dots a_{i_n}) : n \in \mathbb{N}, i_1, \cdots, i_n \in I\big\}.$$

(2) The *joint distribution* of the random variables $(a_i)_{i\in I}$ is the functional $\mu : \mathbb{C}\langle X_i : i \in I\rangle \to \mathbb{C}$ given by

$$\mu(P) = \varphi\big(P(a_i : i \in I)\big),$$

i.e. it is a parametrization of the joint moments via free polynomials.

2.2 Definition. Let $(a_i)_{i\in I}$ be random variables in a $*$-probability space $(\mathcal{A}, \varphi)$.

(1) The family of *joint $*$-moments* of $a_1, \cdots, a_s$ is the collection of values

$$\big\{\varphi(a_{i_1}^{\varepsilon_1}a_{i_2}^{\varepsilon_2}\dots a_{i_n}^{\varepsilon_n}) : n \in \mathbb{N}, \quad i_r \in I, \quad \varepsilon_r \in \{1, *\} \quad \forall r \in [n]\big\}.$$

(2) The *joint distribution* of the random variables $(a_i)_{i\in I}$ is the functional $\mu : \mathbb{C}\langle X_i, X_i^* : i \in I\rangle \to \mathbb{C}$ given by

$$\mu(P) = \varphi\big(P(a_i, a_i^* : i \in I)\big).$$

2.2 Free Independence

The intuition behind independence, at least in the classical setting, is that two random variables are independent if the outcome of one does not affect the other. This intuition will not be of much use for us. Rather, free independence of two non-commutative random variables indicates that they satisfy no algebraic relationships whatsoever, not even commutativity. If they commute, then they are too dependent from a free probability point of view.

It is more practical to see independence of random variables as a way of indicating the behaviour of their mixed moments. Recall that in classical probability, if two random variables $X, Y : \Omega \to \mathbb{C}$ are independent, then we compute their mixed moments by

$$\mathbf{E}[X^m Y^n] = \mathbf{E}[X^m]\mathbf{E}[Y^n]. \tag{2.1}$$

This property actually characterizes classical independence, and the definition of free independence given will be analogous to it. It is stated in terms of random variables of expectation zero, namely by indicating how to evaluate their mixed moments. Notice in the classical setting, as in equation (2.1), we could actually speak of independence of the algebras $\mathrm{alg}(1, X)$ and $\mathrm{alg}(1, Y)$ instead of just independence of X and Y.. Something similar is done below.

2.3 Definition. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and let I be an indexing set.

(1) Subalgebras $\{\mathcal{A}_i\}_{i\in I}$ of $\mathcal{A}$ are said to be *freely independent*, or *free*, if and only if for any $n \in \mathbb{N}$

$$\left.\begin{array}{rl} a_r \in \mathcal{A}_{i(r)} & \forall r \in [n] \\ \varphi(a_r) = 0 & \forall r \in [n] \\ i(1) \neq i(2) \neq i(3) \neq \cdots \neq i(n) & \end{array}\right\} \implies \varphi\big(a_1 a_2 a_3 \cdots a_n\big) = 0. \tag{2.2}$$

(2) Random variables $\{a_i\}_{i\in I}$ in $\mathcal{A}$ are said to be freely independent, or free, if and only if the unital algebras generated by them form a freely independent family.

We say of a random variable that it is *centered* if it has expectation zero. Equation 2.2 may be summarized as saying that a product of centered random variables is also centered, given that adjacent variables belong to distinct subalgebras $\mathcal{A}_i$. Notice that the condition $i(1) \neq \cdots \neq i(n)$ only requires that adjacent indices be distinct; so, for example, we could have that $i(1) = i(3)$, but never that $i(r) = i(r+1)$.

For comparison, we give the definition of classical independence. We will use it in the central limit theorem in section 2.4.

2.4 Definition. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and let I be an indexing set.

(1) Subalgebras $\{\mathcal{A}_i\}_{i\in I}$ of $\mathcal{A}$ are said to be *classically independent*, or *tensor independent,* if and only if the algebras $\mathcal{A}_i$ commute with one another, (in the sense that for all $i \neq j$ in I, and any $a \in \mathcal{A}_i, b \in \mathcal{A}_j$, the random variables a and b commute: $ab = ba$,) and for any random variables $a_1 \in \mathcal{A}_{i_1}, a_2 \in \mathcal{A}_{i_2}, \cdots, a_n \in \mathcal{A}_{i_n}$, with all $i_1, \cdots, i_n$ distinct, we have

$$\varphi\big(a_1 a_2 a_3 \cdots a_n\big) = \varphi(a_1)\varphi(a_2)\cdots\varphi(a_n).$$

(2) Random variables $\{a_i\}_{i\in I}$ in $\mathcal{A}$ are said to be classically independent, or tensor independent, if and only if the unital algebras generated by them form a classically independent family.

At first sight the definition of freeness may seem odd. As in the notion of classical independence, we expect free independence to give us a way of computing mixed moments from individual moments. For example, if a and b are free in $(\mathcal{A}, \varphi)$, and we know all the moments of a and b separately, then what are the values of $\varphi(ab)$, $\varphi(aba)$, or $\varphi(a^2b^2)$? If a and b have expectation zero, then the first two are zero, but what about the third? And what in the case when a and b are not centered? The solution is to proceed by "centering" the random variables. As we see in the examples below, we will apply freeness on the variables $a - \varphi(a)1_{\mathcal{A}}$ and $b - \varphi(b)1_{\mathcal{A}}$, instead of doing it on a and b. In the future we will omit the unit $1 \in \mathcal{A}$ when it is being scaled by a constant, i.e. for a scalar $\alpha \in \mathbb{C}$, we write $\alpha = \alpha 1 \in \mathcal{A}$, when there is no danger of ambiguity.

2.5 Example. Let $a, b \in (\mathcal{A}, \varphi)$ be free. Let us compute the mixed moment $\varphi(ab)$. We do this indirectly by computing $\varphi\big((a - \varphi(a))(b - \varphi(b))\big)$ and symplifying. Since by linearity we have $\varphi(a - \varphi(a)1) = \varphi(a) - \varphi(a)\varphi(1) = 0$, and similarly for b, then we have by freeness that

$$\varphi\Big(\big(a - \varphi(a)\big)\big(b - \varphi(b)\big)\Big) = 0.$$

Then we simplify.

$$\begin{aligned}
0 &= \varphi\Big(\big(a - \varphi(a)\big)\big(b - \varphi(b)\big)\Big) \\
&= \varphi\big(ab - a\varphi(b) - \varphi(a)b + \varphi(a)\varphi(b)\big) \\
&= \varphi(ab) - \varphi(a)\varphi(b) - \varphi(a)\varphi(b) + \varphi(a)\varphi(b) \\
&= \varphi(ab) - \varphi(a)\varphi(b).
\end{aligned}$$

And by rearranging we get

$$\varphi(ab) = \varphi(a)\varphi(b).$$

By substituting a by a^m and b by b^n, and applying the same logic, we also have that

$$\varphi(a^m b^n) = \varphi(a^m)\varphi(b^n). \tag{2.3}$$

This result is the same as the classical case seen in equation 2.1. Something similar happens in the mixed moments of order three:

$$\varphi(aba) = \varphi(a^2)\varphi(b), \qquad \varphi(a^i b^j a^k) = \varphi(a^{i+k})\varphi(b^j). \tag{2.4}$$

The first instance in which we see a result that is unique to free probability is in the order-four moments.

2.6 Example. Let a and b be free random variables in $(\mathcal{A}, \varphi)$. We now compute the value of $\varphi(abab)$ by the same method as above.

$$\begin{aligned} 0 &= \varphi\Big(\big(a-\varphi(a)\big)\big(b-\varphi(b)\big)\big(a-\varphi(a)\big)\big(b-\varphi(b)\big)\Big) \qquad \text{(by freeness)} \\ &= \varphi\Big(\quad abab \quad -aba\varphi(b) \quad -ab\varphi(a)b \quad +ab\varphi(a)\varphi(b) \\ &\qquad -a\varphi(b)ab \quad +a\varphi(b)a\varphi(b) \quad +a\varphi(b)\varphi(a)b \quad -a\varphi(b)\varphi(a)\varphi(b) \\ &\qquad -\varphi(a)bab \quad +\varphi(a)ba\varphi(b) \quad +\varphi(a)b\varphi(a)b \quad -\varphi(a)b\varphi(a)\varphi(b) \\ &\qquad +\varphi(a)\varphi(b)ab \quad -\varphi(a)\varphi(b)a\varphi(b) \quad -\varphi(a)\varphi(b)\varphi(a)b \quad +\varphi(a)\varphi(b)\varphi(a)\varphi(b) \quad \Big). \end{aligned}$$

By linearity (noting $\varphi(a)$ and $\varphi(b)$ are just scalars) this simplifies to

$$\begin{aligned} = \quad & \varphi(abab) \quad -\varphi(aba)\varphi(b) \quad -\varphi(abb)\varphi(a) \quad +\varphi(ab)\varphi(a)\varphi(b) \\ & -\varphi(aab)\varphi(b) \quad +\varphi(aa)\varphi(b)\varphi(b) \quad +\varphi(ab)\varphi(b)\varphi(a) \quad -\varphi(a)\varphi(b)\varphi(a)\varphi(b) \\ & -\varphi(a)\varphi(bab) \quad +\varphi(a)\varphi(b)\varphi(ba) \quad +\varphi(a)\varphi(a)\varphi(bb) \quad -\varphi(a)\varphi(a)\varphi(b)\varphi(b) \\ & +\varphi(a)\varphi(ab)\varphi(b) \quad -\varphi(a)\varphi(a)\varphi(b)\varphi(b) \quad -\varphi(a)\varphi(b)\varphi(a)\varphi(b) \quad +\varphi(a)\varphi(b)\varphi(a)\varphi(b). \end{aligned}$$

Now all mixed moments must be simplified by using the previous results 2.3 and 2.4. Then everything will be expressed only in terms of the values $\varphi(a^n)$ and $\varphi(b^n)$ for $n \in \mathbb{N}$. The simplified equation is

$$0 = \varphi(abab) - \varphi(aa)\varphi(b)^2 - \varphi(a)^2\varphi(bb) + \varphi(a)^2\varphi(b)^2,$$

Which gives the value of the mixed moment, as desired.

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

Free Groups and Free independence

Let us now illustrate the historical origins of the notion of free independence by exploring its relation with freeness of subgroups in a group. Recall the following definition from group theory and compare it to the definition of free independence.

2.7 Definition. Let $\{G_i\}_{i\in I}$ be a family of subgroups of a group G with identity e. We say that the subgroups G_i are *free* in G if and only if

$$\left.\begin{aligned} g_r \in G_{i(r)} \qquad & \forall r \in [n] \\ g_r \neq e \qquad & \forall r \in [n] \\ i(1) \neq i(2) \neq \cdots \neq i(n) & \end{aligned}\right\} \implies g_1 g_2 g_3 \cdots g_n \neq e. \tag{2.5}$$

And we say a family of elements $\{g_i\}_{i\in I}$ are free in G if and only if the subgroups $\{\langle g_i \rangle\}_{i\in I}$ generated by the g_i's are free.

This definition essentially says that subgroups in a group are free if they do not satisfy any algebraic relationships with one another. As we might expect, the notions of freeness "coincide" when we look at group algebras of free groups as non-commutative probability spaces.

2.8 Proposition. *For each $i \in I$, let G_i be a subgroup of a group G. Then the following are equivalent:*

(i) $\{G_i\}_{i\in I}$ are free subgroups of G.

(ii) The group algebras $\{\mathbb{C}G_i\}_{i\in I}$ are freely independent in the $*$-probability space $(\mathbb{C}G, \tau_G)$.

Proof. Suppose first that the groups G_i are free in G, and let us show that the subalgebras $\{\mathbb{C}G_i\}$ are free in $(\mathbb{C}G, \tau_G)$. Let $a_1, a_2, \cdots, a_n$ be elements of the group algebra $\mathbb{C}G$ with expectation zero, i.e. $\tau_G(a_r) = 0$ for $r \in [n]$, and let $i : [n] \to I$ be a function such that for every $r \in [n]$, $a_r \in \mathbb{C}G_{i(r)}$, with the property that $i(r) \neq i(r+1)$, for $1 \leq r \leq n-1$.

We now show that $\tau_G(a_1 \cdots a_n) = 0$. Recall that τ_G is defined by $\tau_G\left(\sum_{g\in G} \alpha_g g\right) = \alpha_e$, so what we have to show is that the coefficient of e in the product $a_1 a_2 \cdots a_n \in \mathbb{C}G$ is zero. Since $\tau_G(a_r) = 0$, then for each $r \in [n]$, a_r has the form

$$a_r = \sum_{g\neq e} \alpha_g^{(r)} g$$

. Therefore we have the following computation

$$\begin{aligned} a_1 a_2 \cdots a_n &= \left(\sum_{g\neq e} \alpha_g^{(1)} g\right)\left(\sum_{g\neq e} \alpha_g^{(2)} g\right) \cdots \left(\sum_{g\neq e} \alpha_g^{(n)} g\right) \\ &= \sum_{g_1,\cdots g_n \neq e} \left(\alpha_{g_1}^{(1)} \alpha_{g_2}^{(2)} \cdots \alpha_{g_n}^{(n)}\right) g_1 g_2 \cdots g_n. \end{aligned}$$

Since the elements g_r $(1 \leq r \leq n)$ satisfy the properties

$$g_r \in G_{i(r)}, \qquad g_r \neq e, \qquad i(1) \neq i(2) \neq \cdots \neq i(n),$$

then by freeness of groups in G, it follows that the product $g_1 g_2 \cdots g_n$ is not equal to e. This in turns proves that $\tau_G(a_1 \cdots a_n) = 0$, as desired.

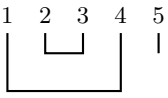
Conversely, suppose that the subalgebras $\{\mathbb{C}G_i\}$ are free in the probability space $(\mathbb{C}G, \tau_G)$. We show that $\{G_i\}$ are free groups in G. This is easy. Take elements $g_1, \cdots g_n$, with $g_r \in G_{i(r)} \setminus \{e\}$ and $i(1) \neq \cdots \neq i(n)$. Then we can consider the elements g_r as being random variables in $\mathbb{C}G_{i(r)}$. We clearly have $\tau_G(g_r) = 0$ because $g_r \neq e$. Then from freeness of the subalgebras $\mathbb{C}G_i$ we immediately conclude that the product $g_i g_2 \cdots g_n \in G$ has expectation zero as a random variable in $(\mathbb{C}G, \tau_G)$, hence it is not equal to e, which proves freeness of the groups $\{G_i\}_{i\in I}$. □

2.3 The Combinatorics of Free Probability

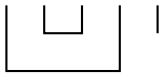
Many of the similarities and distinctions between classical probability and free probability can be better seen from a combinatorial point of view. In fact, it is the relationship between the analysis and the combinatorics

that makes free probability powerful in tackling many kinds of problems. This discussion will culminate in section 2.5, when we introduce the free cumulants. For now, we give a basic overview of the prerequisites in combinatorics and some interesting examples that illustrate how combinatorics will play a role. The results and definitions in the rest of this chapter are mainly due to Speicher [13].

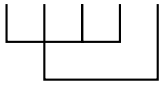
Now, for each n, consider finite ordered sets $[n] = \{1, 2, \cdots, n\}$. A partition of the set $\{1, \cdots, n\}$ is a family $\pi = \{V_1, V_2, \cdots, V_s\}$ of pairwise disjoint nonempty sets whose union is $\{1, \cdots, n\}$; we call the sets V_i *blocks* of π, and we denote $\#\pi = s$, the number of blocks in π. We denote by $P(n)$ the set of all partitions of $[n]$. We will denote each partition in $P(n)$ with a diagram consisting of the ordered elements $1, \cdots, n$, and where for each block $V \in \pi$, we connect the elements of V with lines underneath the numbers. For example, for the partition $\big\{\{1, 4\}, \{2, 3\}, \{5\}\big\}$ we have the following diagram.



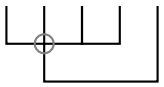
This partition has two blocks of size 2 and one block of size 1. We will always drop the numbers and draw only the lines of the diagram.



As long as we are dealing with an ordered set $\{1 < 2 < \cdots < n\}$, there will be no ambiguity in dropping the numbers. In a similar way, the diagram


$\in P(5)$

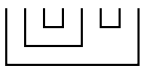
is identified with the partition $\pi = \Big\{\{1, 3, 4\}, \{2, 5\}\Big\}$. We observe, however that two lines representing different blocks "cross" each other:



When this happens we call π a *crossing partition*. (See formal definition below.) If there is no "crossing" in the diagram of a partition, then we call it a *non-crossing partition*, and we denote the set of non-crossing partitions of $[n]$ as $NC(n)$. So for example we have


$\in NC(5)$.

The last two examples above have both one block of size 2 and one block of size 3. A block of size 2 is called a *pairing* or *pair*. If a partition in $P(n)$ has only blocks of size two we call it a *pair partition*, and we denote the sets of pair partitions and non-crossing pair partitions by $P_2(n)$ or $NC_2(n)$, respectively. For example,

$=$ 
$\in NC_2(8) \subset P_2(8)$.

We summarize everything in the following definition.

2.9 Definition. (1) Let A be any set. A collection $\pi = \{V_i : i \in I\}$ of subsets of A is called a *partition of* A if and only if $V_i \neq \varnothing$ for all $i \in I$, $V_i \cap V_j = \varnothing$ for $i \neq j$, and $\bigcup_{i\in I} V_i = A$.

(2) The set of all partitions of $[n]$ is denoted $P(n)$.

(3) A partition $\pi \in P(n)$ is called *crossing* if and only if there exist $i, j, p, q \in [n]$, $i < p < j < q$, such that i, j belong to a block $U \in \pi$, and $p, q \in V \in \pi$, with $U \neq V$. Otherwise, π is called *non-crossing*. The set of non-crossing partitions in $P(n)$ is denoted $NC(n)$.

(4) A partition in $P(n)$ is called a *pairing* if and only if all its blocks are of size 2; and we denote $P_2(n)$ and $NC_2(n)$ the sets of all pair partitions and non-crossing pair partitions of $[n]$, respectively.

The first illustration of how crossing and non-crossing partitions come into play in free probability will be illustrated by the following interesting fact.

2.10 Lemma. (1) *If X is a classical random variable with the standard normal distribution $N(0,1)$, i.e.*

$$\mathbf{E}[X^n] = \frac{1}{2\pi}\int_{\mathbb{R}} t^n e^{\frac{-t^2}{2}}\, dt,$$

then its moments are given by

$$\mathbf{E}[X^n] = \begin{cases} 0, & n \text{ odd} \\ \#P_2(n), & n \text{ even.} \end{cases}.$$

Furthermore, we have that for n even,

$$\#P_2(n) = (n-1)(n-3)(n-5)\cdots(5)(3)(1).$$

(2) *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. If $s \in \mathcal{A}$ is a semicircular random variable, i.e.*

$$\varphi(s^n) = \frac{1}{2\pi}\int_{-2}^{2} t^n \sqrt{4-t^2}\, dt,$$

then its moments are given by

$$\varphi(s^n) = \begin{cases} 0, & n \text{ odd} \\ .\#NC_2(n), & n \text{ even.} \end{cases}$$

Furthermore, we have that for $n = 2k$, n even,

$$\#NC_2(2k) = C_k,$$

where C_k is the k-th Catalan number introduced in example 1.3.

The proof of this lemma is just a series of computations and is left to the reader. To prove that $\#NC_2(2k) = C_k$, one must provide a one-to-one correspondence between partitions $\pi \in NC_2(2k)$ and Dyck paths γ with $2k$ steps, (see section 1.3,) by matching a $+1$ in γ with the "opening" of a pair in π, and a -1 in γ to the "closing" of a pair in π.

The fact expressed in proposition 2.10 suggests that these two distributions play the same role in their respective "parallel worlds" of classical and free probability. This is indeed the case, in a certain sense. As we will see in the central limit theorem, the transition from classical probability to free probability amounts precisely to exchanging the role of partitions by the role of non-crossing partitions.

2.11 Proposition. *For all $k \in \mathbb{N}$, the number of non-crossing pair partitions of $2k$ elements is equal to the number of all non-crossing partitions of k elements. I.e.,*

$$\#NC_2(2k) = \#NC(k).$$

In summary, we have that

$$C_k = \frac{1}{2\pi}\int_{-2}^{2} t^{2k}\sqrt{4-t^2}dt = \#NC_2(2k) = \#NC(k).$$

2.4 Central Limit Theorem

Central to classical probability and statistics is the *Gaussian* or *normal distribution* (denoted $N(0,1)$ when it has mean 0 and variance 1). The reason it is particularly important is that it is the asymptotic distribution appearing in the classical central limit theorem. (See the statement of the theorem below.) In free probability this important role is played by the semicircular distribution, introduced in definition 1.23.

2.12 Definition. If for each $k \in \mathbb{N}$, we have a non-commutative probability space $(\mathcal{A}_k, \varphi_k)$ and a random variable a_k in $\mathcal{A}_k$, and a is a random variable in a non-commutative probability space $(\mathcal{A}, \varphi)$, then we say a_k *converges in distribution* to a as $k \to \infty$ if and only if for every fixed $n \in \mathbb{N}$, the sequence $\left(\varphi_k(a_k^n)\right)_{k\in\mathbb{N}}$ in $\mathbb{C}$, consisting of the n-th moments of a_k in $(\mathcal{A}_k, \varphi_k)$, converges to $\varphi(a^n)$ as $k \to \infty$. In this case, we write

$$a_k \xrightarrow{\text{dist.}} a.$$

2.13 Theorem. *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space, let $(a_i)_{i\in\mathbb{N}}$ be a collection of selfadjoint, centered, and identically distributed random variables in $(\mathcal{A}, \varphi)$ with variance 1. I.e,*

$$\varphi(a_i) = 0, \qquad \varphi(a_i^2) = 1, \qquad \varphi(a_i^n) = \varphi(a_j^n) \qquad \forall i, j, n \in \mathbb{N}.$$

We consider the sequence of random variables

$$\left(\frac{a_1 + a_2 + \cdots a_k}{\sqrt{k}}\right)_{k\in\mathbb{N}}. \tag{2.6}$$

(1) *If the variables $\{a_i\}_{i\in I}$ are classically independent, then the sequence (2.6) converges in distribution to a random variable s with the normal distribution $N(0,1)$.*

(2) *If the variables $\{a_i\}_{i\in I}$ are free, then the sequence (2.6) converges in distribution to a random variable s with the semicircular distribution.*

Proof. In accordance with the definition of convergence in distribution, all we need to do is fix an arbitrary $n \in \mathbb{N}$, compute the n-th moment

$$\varphi\left(\left(\frac{a_1 + \cdots + a_k}{\sqrt{k}}\right)^n\right)$$

and then take the limit as $k \to \infty$.

Let $\{a_i\}_i$ be independent random variables (either classical or free independence; the distinction will be made at the end of the proof). We proceed with the following computations.

$$\begin{aligned}
&\lim_{k\to\infty} \varphi\left(\left(\frac{a_1 + \cdots + a_k}{\sqrt{k}}\right)^n\right) \\
&= \lim_{k\to\infty} \frac{1}{k^{\frac{n}{2}}} \varphi\Big((a_1 + \cdots + a_k)^n\Big) \\
&= \lim_{k\to\infty} \frac{1}{k^{\frac{n}{2}}} \sum_{i:[n]\to[k]} \varphi\left(a_{i(1)} a_{i(2)} \cdots a_{i(n)}\right)
\end{aligned}$$

In the last line above we simply expanded the n-th power of the sum to terms of the form $a_{i(1)}a_{i(2)}\cdots a_{i(n)}$, where the sum runs through all arbitrary choices of subscripts with $1 \leq i(1), i(2), \cdots, i(n) \leq k$. We now observe that the concrete choice of indices $i : [n] \to [k]$ does not affect the value of the mixed moment,

$$\varphi\left(a_{i(1)} a_{i(2)} \cdots a_{i(n)}\right),$$

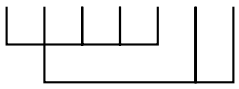
as long as the partition *induced* by i on $[n]$ is the same in the following sense. Given a function $f : A \to S$, the *partition induced on A by f* is defined to be the partition

$$\Big\{f^{-1}\{s\} \subseteq A \ : \ s \in S, \ f^{-1}\{s\} \neq \varnothing\Big\}.$$

Therefore the partition induced by i on $[n]$ is an element $\pi_i \in P(n)$. Since the variables $\{a_i\}_{i\in\mathbb{N}}$ are independent and identically distributed, then two joint moments $\varphi\left(a_{i(1)} a_{i(2)} \cdots a_{i(n)}\right)$ and $\varphi\left(a_{j(1)} a_{j(2)} \cdots a_{j(n)}\right)$ are equal whenever i and j induce the same partition in $P(n)$. For example, if $\{a, b, c, d\}$ are independent and identically distributed random variables, then

$$\varphi\left(abaaabb\right) = \varphi\left(cdcccdd\right) = \varphi\left(dadddaa\right),$$

because all these mixed moments induce the partition



in $P(7)$. Indeed, this is true in general because the joint distribution $\mu : \mathbb{C}\langle X_1, \cdots, X_k\rangle \to \mathbb{C}$ of the independent variables $(a_1, \cdots, a_k)$ is determined by the single-variable distributions of the a_i's, according to the mixed-moment computations given by classical and free independence. And since the variables $(a_1, \cdots, a_k)$ are identically distributed, then the joint distribution $\mu : \mathbb{C}\langle X_1, \cdots, X_k\rangle \to \mathbb{C}$ remains unchanged after permuting the variables $(a_{\sigma(1)}, \cdots, a_{\sigma(k)})$ with some $\sigma \in S_k$.

Then the computation above becomes

$$\lim_{k\to\infty} \frac{1}{k^{\frac{n}{2}}} \sum_{i:[n]\to[k]} \varphi\left(a_{i(1)}a_{i(2)}\cdots a_{i(n)}\right) = \lim_{k\to\infty} \frac{1}{k^{\frac{n}{2}}} \sum_{\pi\in P(n)} N_\pi \kappa_\pi,$$

where κ_π denotes the value of $\varphi\left(a_{i(1)}a_{i(2)}\cdots a_{i(n)}\right)$ if $i:[n]\to[k]$ is any function that induces the partition π, and N_π is the number of functions $i:[n]\to[k]$ which give the partition π. As one might expect, many terms of this sum will vanish, and others will go to zero as $k\to\infty$. We will now prove that all the non-zero terms (after taking the limit) correspond precisely to pair partitions in $P_2(n)$. First, we show that if π has a block of size 1, then $\kappa_\pi = 0$. Indeed, if $\pi = \{V_1, V_2, \cdots, \{t\}, \cdots, V_s\}$ for some $t\in[k]$, then

$$\kappa_\pi = \varphi\left(a_{i(1)}\cdots a_t \cdots a_{i(n)}\right) \tag{2.7}$$

$$= \varphi\left(c_1 a_t c_2\right) \tag{2.8}$$

$$= \varphi(a_t)\varphi(c_1c_2) = 0 \tag{2.9}$$

Here we are taking c_1, c_2 to be in the algebra generated by $\{a_1,\cdots,a_k\}\setminus\{a_t\}$, which is free from the algebra generated by a_t because $\{a_i\}_{i\in\mathbb{N}}$ is free. Equation (2.9) follows by the computations of mixed moments of independent variables, as computed in 2.4 for free variables. (The same holds for classically independent variables.) And the moment vanishes because $\varphi(a_t)=0$. Therefore all the remaining nonzero terms will correspond to partitions π with blocks of size 2 or more. This implies (by the pigeonhole principle) $\#\pi \le n/2$.

Next we show that non-vanishing terms, after taking the limit, will satisfy $\#\pi \ge n/2$. We write the expression above as

$$\sum_{\substack{\pi\in P(n)\\ \#\pi\le\frac{n}{2}}} \left(\lim_{k\to\infty}\frac{N_\pi}{k^{\frac{n}{2}}}\right)\kappa_\pi,$$

and we will show that the limit in parentheses is zero whenever $\#\pi < \frac{n}{2}$.

First, given a partition π, we compute the value of N_π. This is a standard problem in enumerative combinatorics. N_π is by definition the number of functions $i:[n]\to[k]$ that induce the partition π. Let $\pi = \{V_1, V_2, \cdots, V_{\#\pi}\}$, then to construct all possible functions $i:[n]\to[k]$ that induce π, we just need to choose for each $1\le j\le \#\pi$, without repetition, the value of i on each block V_j. Since there are k possible values for i, the number of functions i that induce π is $k(k-1)(k-2)\cdots(k-\#\pi+1)$.

Now we compute the limit

$$\begin{aligned}
&\lim_{k\to\infty} \frac{N_\pi}{k^{\frac{n}{2}}} \\
&= \lim_{k\to\infty} \frac{k(k-1)(k-2)\cdots(k-\#\pi+1)}{k^{\frac{n}{2}}} \\
&= \lim_{k\to\infty} \frac{k}{k}\frac{(k-1)}{k}\frac{(k-2)}{k}\cdots\frac{(k-\#\pi+1)}{k}k^{\#\pi-\frac{n}{2}} \\
&= \begin{cases} 1, & \frac{n}{2} = \#\pi \\ 0, & \frac{n}{2} > \#\pi. \end{cases}
\end{aligned}$$

Thus we have shown that in the limit all terms of the summation vanish unless they correspond to a partition π with $\#\pi = n/2$, and with no blocks of size 1. It follows that each remaining term will correspond to a pair partition. Thus we have so far that

$$\lim_{k\to\infty} \varphi\left(\left(\frac{a_1 + \cdots + a_k}{\sqrt{k}}\right)^n\right) = \sum_{\pi\in P_2(n)} \kappa_\pi.$$

To conclude the proof we compute the value of κ_π. This will require us now to discern between the cases of classical independence and free independence. Recall that κ_π is by definition the moment $\varphi(a_{i(1)}a_{i(2)}\cdots a_{i(n)})$, where $i : [n] \to [k]$ is any function inducing the partition π.

If $\{a_i\}_{i\in\mathbb{N}}$ is *classically independent*, then $\varphi(a_{i(1)}a_{i(2)}\cdots a_{i(n)})$ will factor into a product of second moments, and since the variance of the random variables is 1, then $\kappa_\pi = 1$ for each $\pi \in P_2(n)$. Therefore in the classically independent case we have

$$\lim_{k\to\infty} \varphi\left(\left(\frac{a_1 + \cdots + a_k}{\sqrt{k}}\right)^n\right) = \sum_{\pi\in P_2(n)} 1 = \#P_2(n)$$

As we saw in lemma 2.10, these are precisely the moments of the standard normal distribution. And therefore that random variable converges to a normal distribution of variance 1.

In the case of *free independence,* We have two cases for the moment $\varphi(a_{i(1)}a_{i(2)}\cdots a_{i(n)})$.

Case 1: All pairs of neighbouring indices are different:

$$i(1) \neq i(2) \neq \cdots \neq i(n).$$

In this case, by freeness, $\varphi(a_{i(1)}a_{i(2)}\cdots a_{i(n)}) = 0$.

Case 2: There exists $r \in [n]$ such that $i(r) = i(r+1)$. In this case we may factor the mixed moment by using freeness as we did in equation (2.9). This gives

$$\begin{aligned}&\varphi\left(a_{i(1)}\cdots a_{i(r)}a_{i(r+1)}\cdots a_{i(n)}\right)\\ =&\varphi\left(a_{i(r)}a_{i(r+1)}\right)\varphi\left(a_{i(1)}\cdots a_{i(r-1)}a_{i(r+2)}\cdots a_{i(n)}\right).\end{aligned}$$

We can evaluate $\varphi\left(a_{i(r)}a_{i(r+1)}\right) = 1$, and so we are left with a moment $\varphi\left(a_{i(1)}\cdots a_{i(r-1)}a_{i(r+2)}\cdots a_{i(n)}\right)$ of a lesser order.

This new moment then falls on one of the two cases above, and so we repeat. Eventually this process ends either with a zero moment, if at some point we get an expression of case 1, or with a moment of 1, if all variables are eliminated two at a time. The latter will happen precisely if and only if the indices i induce a *non-crossing* partition. To see that this is the case , consider a crossing pair partition $\pi \in P_2(n) \setminus NC_2(n)$ induced by an index-function $i : [n] \to [k]$. By definition, then, and because π is a pairing, we must have that there are blocks $\{h, j\}, \{\ell, m\} \in \pi$ such that $h < \ell < j < m$. There are exclusively two possibilities for the moment $\varphi(a_{i(1)}a_{i(2)}\cdots a_{i(n)})$; either all neighbouring pairs of variables are cancelled by repeatedly applying case 1 above, in which case the expected value is 1, or at some point we fall into case 2 above, i.e.

neighbouring indices (after possibly applying case 1 a number of times) are all different, in which case this moment is zero. Since the factors $a_{i(h)}$ and $a_{i(j)}$ will never be adjacent because $a_{i(\ell)}$ is intermediate, the factors $a_{i(h)}$ and $a_{i(j)}$ can not be cancelled before the factors $a_{i(\ell)}$ and $a_{i(m)}$. But since a similar argument can be made for $a_{i(\ell)}$ and $a_{i(m)}$, then we can not have cancellation in the reverse order either. Thus these four terms can never be canceled by repeated application of case 1, and therefore case 2 applies eventually to the moment $\varphi\left(a_{i(1)}a_{i(2)}\cdots a_{i(n)}\right)$ which must then be zero.

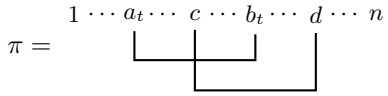
Conversely, consider an index-function $i:[n]\to[k]$ with induced partition $\pi\in P_2(n)$ such that

$$\varphi\left(a_{i(1)}a_{i(2)}\cdots a_{i(n)}\right)=0.$$

This can only happen if case 2 applies eventually. Without loss of generality, suppose that adjacent indices are distinct,

$$i(1)\neq i(2)\neq\cdots\neq i(n)$$

(otherwise apply case 1 as many times as possible, reset the value of n, and re-label the indices). We will now show that π is a crossing partition. Take $1\leq a_1<b_1\leq n$ such that a and b are paired by π. If there exists a',b' such that $a_1<a'<b'<b_1$ and $\{a',b'\}$ is a pair in π, then let $a_2=a'$, and $b_2=b'$. Continue in a similar way: for each a_m,b_m paired in π, if there exist a',b' paired in π such that $a_m<a'<b'<b_m$ then we define $a_{m+1}=a',b_{m+1}=b'$. Eventually this will terminate with a pair $\{a_t,b_t\}\in\pi$ such that there is no pair $\{a',b'\}$ with $a_t<a'<b'<b_t$. There must be an element $c\in[n]$, between a and b, because consecutive indices $i(r),i(r+1)$ are not equal. So c must be paired in π with some other element d such that either $a_t<c<b_t<d$ or $d<a_t<c<b_t$.

$\pi=$ 

In both cases this proves, by definition, that π is a crossing partition.

Thus we have shown, as desired, that $\varphi\left(a_{i(1)}a_{i(2)}\cdots a_{i(n)}\right)=0$ if and only if $i:[n]\to[k]$ induces a crossing partition, and it is 1 otherwise. Therefore,

$$\lim_{k\to\infty}\varphi\left(\left(\frac{a_1+\cdots+a_k}{\sqrt{k}}\right)^n\right)=\sum_{\pi\in NC_2(n)}1=\#NC_2(n),$$

which gives precisely the moments of a semicircular distribution, by lemma 2.10. This concludes the proof. □

2.5 Free Cumulants

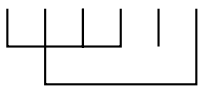
The values κ_π from the proof in the previous section serve as motivation for the study of free cumulants. The values labelled there as κ_π were indeed cumulants, but they are specific to the semicircular distribution. In order to introduce cumulants in general we will need the notion of multiplicative family of functions in NC. From now on, we will use NC to denote the set of all non-crossing partitions of any set $[n]$, i.e. $NC=\bigcup_{n\in\mathbb{N}}NC(n)$.

2.14 Definition. A *multiplicative family of functions in NC* (or respectively, in P) is a collection $(f_\pi)_{\pi\in NC}$ such that for each $\pi \in NC(n)$ (resp. $\pi \in P(n)$), $f_\pi : A^n \to \mathbb{C}$ is a multilinear function $(a_1, \cdots, a_n) \mapsto f_\pi[a_1, \cdots a_n]$ with the property that its values factor in accordance with the blocks of π as follows:

$$f_\pi[a_1, \cdots, a_n] = \prod_{\substack{V\in\pi \\ V=\{i_1<\cdots<i_s\}}} f_s(a_{i_1}, \cdots a_{i_s}).$$

Where $f_n(a_1, \cdots, a_n) = f_{\{[n]\}}[a_1, \cdots, a_n]$, $\{[n]\} \in NC(n)$ the partition with just one block. So that the functions $(f_n)_{n\in\mathbb{N}}$ determine all others.

For example, if

$\pi =$  $= \Big\{\{1,3,4\},\{2,6\},\{5\}\Big\} \in P(6),$

then

$$f_\pi[a_1, a_2, a_3, a_4, a_5, a_6] = f_3(a_1, a_3, a_4) f_2(a_2, a_6) f_1(a_5).$$

2.15 Definition. Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space. We define free cumulants $(\kappa_\pi)_{\pi\in NC}$ to be the multiplicative family of functions in NC determined inductively by the following relation to moments in $\mathcal{A}$.

$$\varphi\left(a_1 \cdots a_n\right) = \sum_{\pi\in NC(n)} \kappa_\pi[a_1, \cdots, a_n]$$

And we define also the sequence of functions $\kappa_n : \mathcal{A}^n \to \mathbb{C}$, $n \in \mathbb{N}$, by

$$\kappa_n = \kappa_{\{[n]\}}.$$

To see how moments in $\mathcal{A}$ determine the values κ_π, we begin by determining κ_1. Take $n = 1$, so $NC(1) = \{\{1\}\}$. We have, for any $a \in \mathcal{A}$,

$$\varphi(a) = \kappa_{\{\{1\}\}}[a] = \kappa_1(a).$$

So $\kappa_1(a) = \varphi(a)$ for all $a \in \mathcal{A}$. Next, to determine κ_2, we fix $n = 2$ and use the value of κ_1 obtained above. First, we have that

$NC(2) =$ 

So the definition gives, for arbitrary $a, b \in \mathcal{A}$,

$$\begin{aligned}\varphi(a,b) &= \kappa_{\text{I I}}[a,b] + \kappa_{\sqcup}[a,b] \\ &= \kappa_1(a)\kappa_1(b) + \kappa_2(a,b) \\ &= \varphi(a)\varphi(b) + \kappa(a,b).\end{aligned}$$

So for any $a, b \in \mathcal{A}$, the second cumulant is given by

$$\kappa_2(a,b) = \varphi(ab) - \varphi(a)\varphi(b). \tag{2.10}$$

Now, once the first two cumulants are known, they are used to compute the third free cumulant as follows. There are five non-crossing partitions of three ordered elements:

$$NC(3) = \Big\{ \quad \text{  } \quad \Big\}.$$

Therefore, we compute the third cumulant $\kappa_3 = \kappa_{\sqcup\!\sqcup}$ is computed, according to the definition, from the moment-cumulant formula:

$$\begin{aligned}
\varphi(abc) =& \kappa_{\sqcup\!\sqcup} + \kappa_{\sqcup\,|} + \kappa_{|\,\sqcup} + \kappa_{\sqcup\!\sqcup} + \kappa_{|\,|\,|} \\
=& \kappa_3(a,b,c) + \kappa_2(a,b)\kappa_1(c) + \kappa_1(a)\kappa_2(b,c) + \kappa_2(a,c)\kappa_1(b) + \kappa_1(a)\kappa_1(b)\kappa_1(c) \\
=& \kappa_3(a,b,c) + \big(\varphi(a,b) - \varphi(a)\varphi(b)\big)\varphi(c) \\
&+ \varphi(a)\big(\varphi(bc) - \varphi(b)\varphi(c)\big) + \big(\varphi(a,c) - \varphi(a)\varphi(c)\big)\varphi(b) + \varphi(a)\varphi(b)\varphi(c) \\
=& \kappa_3(abc) + \varphi(a)\varphi(bc) + \varphi(b)\varphi(ac) + \varphi(c)\varphi(ab) - 2\varphi(a)\varphi(b)\varphi(c).
\end{aligned}$$

And therefore for all $a, b, c \in \mathcal{A}$,

$$\kappa_3(a,b,c) = \varphi(abc) - \varphi(a)\varphi(bc) - \varphi(b)\varphi(ac) - \varphi(c)\varphi(ab) + 2\varphi(a)\varphi(b)\varphi(c) \tag{2.11}$$

The fourth cumulant $\kappa_4(a,b,c,d) = \kappa_{\sqcup\sqcup\sqcup}[a,b,c,d]$ would be computed in the same way, by writing the moment-cumulant formula over $NC(4)$, simplifying according to the known values of the first three cumulants, and solving for κ_4. It is too long to write for arbitrary $a, b, c, d \in \mathcal{A}$, (it has 14 terms, one for each non-crossing partition in $NC(4)$,) but for the special case where all four random variables are centered, i.e. $\varphi(a) = \varphi(b) = \varphi(c) = \varphi(d) = 0$, most of the terms vanish and the cumulant simplifies to

$$\kappa_4(a,b,c,d) = \varphi(abcd) - \varphi(ab)\varphi(cd) - \varphi(ad)\varphi(cd). \tag{2.12}$$

Similarly, in the special case $a = b = c = d$, the cumulant simplifies to

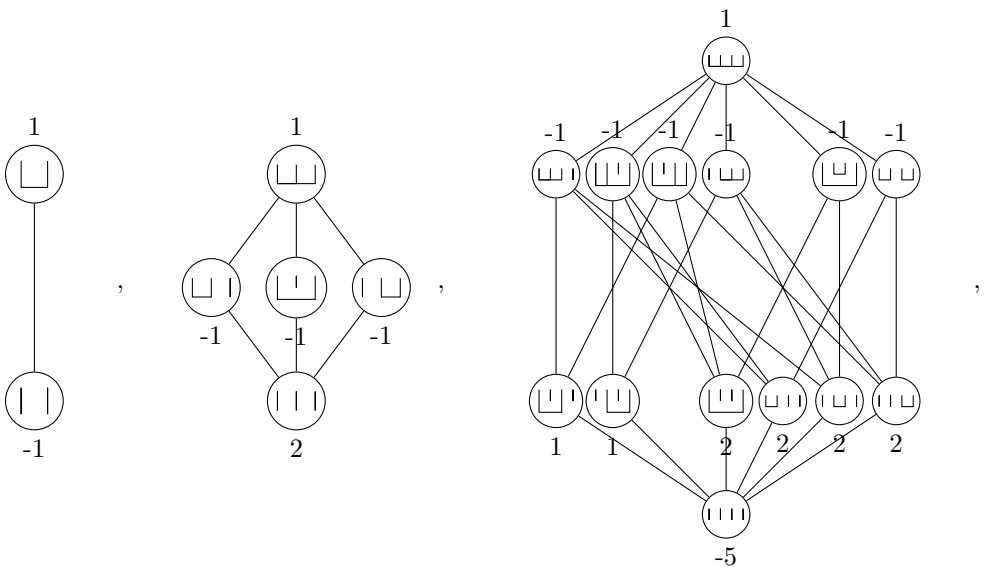
$$\kappa_4(a,a,a,a) = \varphi(aaaa) - 4\varphi(aaa)\varphi(a) - 2\varphi(aa)^2 + 10\varphi(aa)\varphi(a)^2 - 5\varphi(a)^4. \tag{2.13}$$

Admittedly, the method used above for computing the n-th cumulant is tedious, it gets very long quickly as n increases, and it has the disadvantage that, in order to compute κ_n, all previous cumulants must have been computed in advance. In order to overcome the difficulty, techniques from advanced combinatorics are essential, namely *lattice theory and the Möbius inversion formula on a lattice.* Unfortunately this subject is too deep to be covered in this thesis. For reference, we direct the reader to lectures 9 and 10 of [12]. The rough idea is that there exists a function $\mu_n : NC(n) \to \mathbb{Z}$ such that the cumulants may be evaluated directly by

$$\kappa_n(a_1, \cdots a_n) = \sum_{\pi \in NC(n)} \mu_n(\pi)\varphi_\pi[a_1, \cdots a_n].$$

Where φ is treated as a multiplicative family in NC by writing $\varphi_n(a_1, a_2, \cdots, a_n) = \varphi(a_1 a_2 \cdots a_n)$. For example, we would write $\varphi_{\sqcup\,|\,|}[a,b,c,d] = \varphi(abd)\varphi(c)$. In turn, one computes the function μ_n combinatorially

from the lattice structure of $NC(n)$, namely by the lattice diagram of $NC(n)$. For example, the lattice diagrams of $NC(2)$, $NC(3)$ and $NC(4)$ are



where each partition $\pi \in NC(n)$ above is labeled with its corresponding value $\mu_n(\pi)$. Notice, for example that the coefficients of φ_π that determine the second and third cumulants in equations (2.10) and (2.11) precisely match the respective values $\mu_n(\pi)$, as indicated in the diagram. In fact, the author wrote equations (2.12) and (2.13) on fourth cumulants by reading the diagram. Details on the construction and interpretation of these diagrams may be found in [12].

The following theorem is a central result of free probability. It gives a characterization of free independence in terms of cumulants. Roughly, it states that a collection of random variables is free if and only if their mixed cumulants vanish.

2.16 Theorem. *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and let $(\kappa_n)_{n\in\mathbb{N}}$ be the corresponding free cumulants. Let $\mathcal{A}_i$ for $i \in I$, be subalgebras of $\mathcal{A}$. Then the following are equivalent.*

(*i*) $\{\mathcal{A}_i\}_{i\in I}$ *are freely independent.*

(*ii*) *If $n \geq 2$, and $a_1, \cdots, a_n$ are random variables with $a_r \in \mathcal{A}_{i(r)}$, for $i(1), \cdots i(n) \in I$, and if there exist $p, q \in [n]$, $p < q$ with $i(p) \neq i(q)$, then $\kappa_n(a_1, a_2, \cdots, a_n) = 0$.*

As this theorem illustrates, the main use of free cumulants is the simplification of definitions, computations and proofs. For example, recall the definition of freeness given in section 2.2. In order to show random variables are free, according the definition, one needs to show that mixed moments vanish under certain conditions; namely, they have to be moments of *centered* random variables, and neighbouring random variables in the mixed moment must be generated by *different* random variables from the free family. According to the criterion given by this theorem, on the other hand, if we look at the same random variables from the point of view of their joint free cumulants, all we need to show is that mixed cumulants vanish; no other conditions

need be satisfied. The proof of this theorem is not particularly hard, but it uses nontrivial combinatorial properties of NC, and it is a rather long proof, so it is omitted. See [12].

2.6 The $\mathcal{R}$-Transform and Free Convolution

In this section, we will study random variables by focusing on their sequence of moments and their sequence of free cumulants. For a fixed random variable a in a non-commutative probability space $(\mathcal{A}, \varphi)$, he sequence of moments, as seen in the previous chapter, is simply $\big(\varphi(a^n)\big)_{n\in\mathbb{N}}$. Similarly the sequence

$$\Big(\kappa_n(a, \cdots, a)\Big)_{n\in\mathbb{N}}$$

will be referred to as the *sequence of cumulants of* a.

2.17 Notation. Given $a \in (\mathcal{A}, \varphi)$, for simplicity we will write

$$m_n^a = \varphi(a^n) \qquad\qquad \kappa_n^a = \kappa_n(a, \cdots, a).$$

If μ is the $*$-distribution of a in some $*$-probability space, then we may also write m_n^μ for m_n^a and κ_n^μ for κ_n^a. And, when there is no risk of ambiguity, we drop the superscript, and we just write m_n and κ_n.

2.18 Remark. In a non-commutative probability space, the sequence of moments contains all the information about the distribution of a. Since the the sequence of moments is related to the sequence of cumulants by the moment-cumulant formula,

$$\varphi\left(a_1 \cdots a_n\right) = \sum_{\pi\in NC(n)} \kappa_\pi[a_1, \cdots, a_n],$$

then it follows that both the sequence of moments can be recovered from the sequence of cumulants, and vice versa. In other words, the sequences of moments and cumulants contain exactly the same information about the distribution of a random variable.

2.19 Remark. In the framework of a non-commutative probability space, the sequence of moments captures all the information about the single random variable. In the framework of a $*$-probability space, the $*$-distribution of a single random variable is far more complicated, as seen in chapter 1. But in the case of a selfadjoint random variable, we do have that the sequence of moments captures all of its probabilistic information. The same can be said about the sequence of cumulants, by remark 2.18.

In non-commutative probability in general, we are often interested in studying sums of random variables. See, for instance, what we did in example 1.21, where we defined the normal random variables known as Haar unitaries u, whose distribution is quite straight forward. Then we found analytically the distribution of the sum $u + u^*$, a selfadjoint random variable. It was found to have the arcsine distribution and a sequence of moments given by $n \mapsto \binom{n}{n/2}$ if n is even, and $n \mapsto 0$ if n is odd. Similarly, we studied the example of section 1.3, in which we studied a non-normal random variable a with a very simple distribution. The sum $a+a^*$ was

considered, as a selfadjoint random variable, and it was found to have the semicircular distribution, which gives the moment sequence $n \mapsto C_{n/2}$ if n is even, and zero otherwise.

It is especially important in free probability to compute the sum of free, identically distributed random variables, as we saw, for example, in the central limit theorem (section 2.4). We, therefore, define free convolution to capture the essence of addition of freely independent random variables.

2.20 Definition. (1) Let $(\mathcal{A}, \varphi)$, and $(\mathcal{B}, \psi)$ be $*$-probability spaces and $a \in \mathcal{A},\ b \in \mathcal{B}$. If there exists a $*$-probability space $(\mathcal{C}, \tau)$ and random variables a', b' such that a, a' are identically distributed, b, b' are identically distributed, and a', b' are free, then we say that $a' + b'$ are the free convolution of a and b. In this case, we write

$$a \boxplus b = a' + b'.$$

(This is a slight abuse of notation, since a', b', and $(\mathcal{C}, \tau)$ are not necessarily unique, but since the freeness of a' and b' uniquely determines mixed $*$-moments in terms of the $*$-moments of a and of b, then the $*$-distribution of $a' + b'$ is uniquely determined by the $*$-distributions of a and b.)

(2) If μ and ν are $*$-distributions $\mathbb{C}\langle X, X^*\rangle \to \mathbb{C}$, then the $*$-distribution $\eta : \mathbb{C}\langle X, X^*\rangle \to \mathbb{C}$ is said to be the free convolution of μ and ν, if there exist random variables a and b in some $*$-probability space such that η is the $*$-distribution of $a \boxplus b$. In this case we write

$$\mu \boxplus \nu = \eta.$$

As the reader has seen up to know, the moments of a sum of random variables are not predictable in any obvious way. This is the case both in sums of non-free random variables, as in the computations for $u + u^*$ and $a + a^*$, mentioned above, and also in sums of free random variables, as is seen in the computations of the mixed moments of free random variables, as in the central limit theorem. In other words, computing free convolutions is non-trivial from the point of view of moments. However, from the point of view of cumulants, this is not the case. The single most important property of cumulants is that they are additive on free random variables.

2.21 Proposition ([12])**.** *Let a, and b be freely independent random variables in a non-commutative probability space. Then we have, for all $n \in \mathbb{N}$,*

$$\kappa_n^{a+b} = \kappa_n^a + \kappa_n^b.$$

Proof. This is a direct consequence of the vanishing of mixed cumulants, and the properties of κ_n as an

n-linear function:

$$\begin{aligned}
\kappa_n^{a+b} &= \kappa_n(a+b, \cdots, a+b) \\
&= \sum_{(x_1,\cdots,x_n)\in\{a,b\}^n} \kappa(x_1, \cdots, x_n) \qquad \text{(by multilinear property)} \\
&= \kappa_n(a, \cdots, a) + \kappa_n(b, \cdots, b) \qquad \text{(by vanishing of mixed cumulants)} \\
&= \kappa_n^a + \kappa_n^b.
\end{aligned}$$

□

An equivalent way of stating the theorem above would be: *For any random variables a, and b in any pair of non-commutative probability spaces, we have*

$$\kappa_n^{a\boxplus b} = \kappa_n^a + \kappa_n^b.$$

In light of this result, it is easy to see that the computation of free convolution is better done in terms of cumulants than in terms of moments. Therefore, the main technique used in free probability to compute free convolutions is, in summary, passing from knowledge of distributions to knowledge of cumulants, perform direct addition to obtain the cumulants of the convolution, and finally return to knowledge of distributions. The transition between moments and cumulants is done either combinatorially, by directly applying the moment-cumulant formula, algebraically, as in the kinds of computations done in the central limit theorem, or analytically, which leads us to define Voiculescu's "$\mathcal{R}$-transform."

It is common practice in basic combinatorics to record the information of a sequence as the coefficients of a formal power series and attempt using different analytic tools. We do exactly this (with a shift on the indices,) on the sequence of free cumulants.

2.22 Definition. Let a be a random variable in a non-commutative probability space. Then the $\mathcal{R}$-transform $\mathcal{R}_a$ of a (also denoted $\mathcal{R}_\mu$, and referred to as the $\mathcal{R}$-transform of μ, if μ is the distribution of a) is the formal power series

$$\mathcal{R}_a(z) = \sum_{n=0}^{\infty} \kappa_{n+1} z^n.$$

Therefore we have the desired additivity of the $\mathcal{R}$-transform: *If a and b are free random variables in a non-commutative probability space, then*

$$\mathcal{R}_{a+b}(z) = \mathcal{R}_a(z) + \mathcal{R}_b(z).$$

The use of the $\mathcal{R}$-transform is to take advantage of selfadjoint random variables with an analytic distribution. It is related to the Cauchy transform in the following way.

2.23 Theorem. *Let μ be a compactly supported probability measure on $\mathbb{R}$. Then the $\mathcal{R}$-transform of μ is related to the Cauchy transform of μ by*

$$G_\mu\Big(\mathcal{R}_\mu(z) + \frac{1}{z}\Big) = z.$$

See [12] for a proof involving the combinatorics of NC and other analytic computations.

Observe that the $\mathcal{R}$-transform can be computed by finding the inverse function of G_μ:

$$G_\mu^{-1}(z) = \mathcal{R}_\mu(z) + \frac{1}{z}.$$

Recall also that the Cauchy transform can be written in two forms, one in terms of the sequence of moments and the other in terms of an analytic distribution. Namely, we have

$$G_\mu(z) = \int \frac{1}{z-t} d\mu(t) = \sum_{n=0}^{\infty} \frac{m_n^\mu}{z^{n+1}}.$$

And so a standard procedure to transition from distributions to cumulants is to compute the Cauchy transform of a selfadjoint random variable, (whether this is done by its sequence of moments or by its analytic distribution), compute its inverse, and solve for $\mathcal{R}_a(z)$. A similar procedure is used to return from cumulants to moments.

3 Simplicial Complexes

Chapters 1 and 2 were an introduction to free probability and its techniques. The rest of this thesis will focus specifically on a recent application of non-commutative probability to algebraic topology. This research area is in its very early stages and, as the reader will notice, is significantly unfinished. This chapter introduces the main topological objects to be studied, simplicial complexes. We will introduce definitions, notations and properties concerning simplicial complexes and basic homology theory; then, we will record all the information that determines a simplicial complex in the form of a block-matrix. This matrix will be studied in subsequent chapters as a non-commutative random variable. All of our notation in this and subsequent chapters is intended to match the notations used by Vargas in [15].

3.1 Simplices and Complexes

Simplicial complexes are a kind of topological space made up of "triangulations" of different dimensions. First, we define these k-dimensional "triangles", or properly called, *simplices*.

3.1 Definition. Let $v_0, v_1, \ldots, v_k$ be points in $\mathbb{R}^n$, for some $n \geq k$. Then $v_0, v_1, \ldots, v_k$ are said to be in *general position* (or *affinely independent*) if and only if the vectors $v_1 - v_0, v_2 - v_0, \ldots, v_k - v_0$ are linearly independent in $\mathbb{R}^n$.

3.2 Definition. Let $v_0, v_1, \ldots, v_k$ be in general position in $\mathbb{R}^n$ for some $n \geq k$. Then the *k-dimensional simplex*, or *k-simplex* on the vertices $v_0, v_1, \ldots, v_k$, denoted $[v_0, v_1, \ldots, v_k]$, is defined to be the convex hull of its vertices, i.e. the set

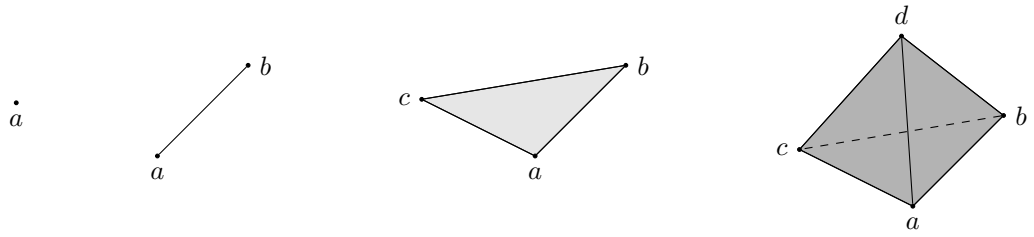
$$[v_0, v_1, \ldots, v_k] = \left\{ x \in \mathbb{R}^n : x = \sum_{j=0}^{k} \alpha_j v_j,\ \alpha_j \geq 0,\ \sum_{j=0}^{k} \alpha_j = 1 \right\}.$$

Given simplices σ and τ, we say τ is a *subsimplex* of σ if and only if every vertex of τ is a vertex of σ. If τ is a subsimplex of σ, and the dimension of τ is one less than the dimension of σ, then we say τ is a *face of* σ. Thus faces of σ are obtained by removing exactly one vertex from those of σ. Given a k-simplex $[v_0, v_1, \ldots, v_k]$, we will denote the $(k-1)$-face on all but one of the vertices by the notation

$$[v_0, \ldots, \hat{v}_i \ldots, v_k] = [v_0, \ldots, v_{i-1}, v_{i+1} \ldots, v_k],$$

so that the "hat" in $\hat{v}_i$ indicates that the vertex v_i is omitted.

For example, given points a, b, c and d in standard position in $\mathbb{R}^n$, the simlpex $[a] = \{a\}$ is simply a point, $[a, b]$ is a line, $[a, b, c]$ is a triangle, and $[a, b, c, d]$ is a tetrahedron.

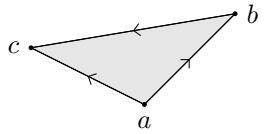


3.3 Notation. Throughout this work, we will assume that the vertices of a simplex are linearly ordered, and this will be expressed by indices. A simplex $[v_{j_0}, v_{j_1}, \ldots, v_{j_k}]$ will always be expressed with the vertices in increasing order. In other words, we use the convention that

$$\textit{A simplex may be expressed as } [v_{j_0}, v_{j_1}, \ldots, v_{j_k}], \textit{ if and only if } \; j_0 < j_1 < \cdots < j_k.$$

Furthermore, when drawing diagrams, arrows may be added to the edges (i.e. 1-dimensional subsimplices), pointing from lesser to greater, in the order of the vertices.

So for example, the simplex $[a, b, c]$ above, with the order $a < b < c$, (i.e., $v_1 = a, v_2 = b, v_3 = c$,) may be represented pictorially by



3.4 Definition. A *simplicial complex* X in $\mathbb{R}^n$ is a collection of simplices in $\mathbb{R}^n$ such that:

1. If $\sigma, \tau \in X$, then the set $\sigma \cap \tau$ is itself a simplex, and it is a subsimplex of both σ and τ.
2. If $\sigma \in X$ and τ is a subsimplex of σ, then $\tau \in X$.

Furthermore, we will denote simplicial complexes as disjoint unions

$$X = \bigcup_{i=0}^{d} X_i,$$

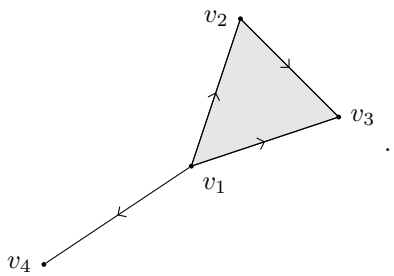
where X_i is the set of i-simplices in X, and $d = \max\{i : X_i \neq \varnothing\}$. d is called the *dimension* of X. We define also n_i to be the number of i-simplices in X, and N to be the total number of simplices in X, i.e.

$$n_i = |X_i| \quad \text{and} \quad N = n_0 + n_1 + \cdots + n_d.$$

0-dimensional simplices in a simplicial complex are often referred to as *vertices*, and 1-dimensional simplices are referred to as *edges*.

It should also be noted that by the definition of a simplex, the vertices of each simplex are in general position, but this needs not be the case for all vertices in the complex.

3.5 Example. Consider the points $v_1 = (0,0),\ v_2 = (1,3),\ v_3 = (3,1),\ v_4 = (-3,-2)$ in $\mathbb{R}^2$ and the 2-dimensional simplicial complex K in $\mathbb{R}^2$ represented pictorially by



For future reference, we will call the simplicial complex K the *kite*. Note that v_1, v_2, v_3 are in general position because they are not colinear, and v_1, v_4 are in general position because they are distinct. Also, note that v_1, v_2, v_3, v_4 are *not* in general position, because they are four co-planar points (obviously, since they are 4 points in $\mathbb{R}^2$) but this is no issue for us; it just means that the 3-simplex $[v_1, v_2, v_3, v_4]$ cannot be defined.

The simplicial-complex structure of X is

$$\begin{aligned} X = \Big\{ & [v_1], [v_2], [v_3], [v_4], \\ & [v_1, v_2], [v_1, v_3], [v_1, v_4], [v_2, v_3], \\ & [v_1, v_2, v_3] \Big\}. \end{aligned}$$

And the elements of X_1, X_2, X_3 are written in different lines, respectively.

3.6 Notation. As in the notation above, we will always assume the collection of all vertices (or 0-simplices) will be ordered by indices so that we will always write indices as in

$$X_0 = \big\{ [v_1], [v_2], \ldots, [v_{n_0}] \big\},$$

and therefore an arbitrary simplex in X will have the form

$$\sigma = [v_{i_0}, v_{i_1}, \ldots, v_{i_k}], \quad \text{for} \quad 1 \le i_0 < i_1 < \cdots < i_k \le n_0.$$

We will fix an order on X as follows. First, on each X_j we will use the lexicographic (or dictionary) order: for any $\sigma = [v_{i(0)}, \ldots, v_{i(j)}], \tau = [v_{k(0)}, \ldots, v_{k(j)}]$ in X_j,

$$\sigma < \tau \quad \text{if and only if} \quad \begin{cases} & i(0) < k(0) & & & & \\ \text{or} & i(0) = k(0), & i(1) < k(1) & & & \\ \text{or} & i(0) = k(0), & i(1) = k(1), & i(2) < k(2) & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ \text{or} & i(0) = k(0), & i(1) = k(1), & \cdots & , & i(j) < k(j). \end{cases}$$

And thus we write the j-simplices in order as

$$X_j = \left\{\sigma_1^{(j)}, \sigma_2^{(j)}, \cdots, \sigma_{n_j}^{(j)}\right\}.$$

Then we declare that for $j < \ell$, and $\sigma_i^{(j)} \in X_j, \sigma_k^{(\ell)} \in X_\ell$, we have

$$\sigma_i^{(j)} < \sigma_k^{(\ell)}$$

Thus we have a linear order in X and we write its N elements it in increasing order:

$$\begin{aligned} X = \quad & \left\{\sigma_1^{(0)}, \sigma_2^{(0)}, \cdots, \sigma_{n_0}^{(0)}, \right. \\ & \sigma_1^{(1)}, \sigma_2^{(1)}, \cdots, \sigma_{n_1}^{(1)}, \\ & \vdots \\ & \left.\sigma_1^{(d)}, \sigma_2^{(d)}, \cdots, \sigma_{n_d}^{(d)}\right\} \end{aligned}$$

For example, the simplices of X from example 3.5 were written in increasing order.

The main use of simplicial complexes is in reducing topological objects to a combinatorial structure; one then uses combinatorial methods to express its topological properties. Since a simplicial complex is simply a set of simplices, we will also talk of its *geometric realization* in $\mathbb{R}^n$

$$|X| = \bigcup_{i=0}^{d} \bigcup_{j=1}^{n_i} \sigma_j^{(i)},$$

And when we talk about the topology of a simplicial complex, we are referring to the topology of its geometric realization.

3.2 Chain Complexes and Homology

We now define chain groups on a simplicial complex. For our purposes, the chain groups are vector spaces over $\mathbb{C}$.

3.7 Definition. Let X be a simplicial complex of dimension d. Then for any $0 \leq i \leq d$, the i-th chain group of X is the formal vector space over $\mathbb{C}$, denoted C_i (or $C_i(X)$), with basis X_i, i.e.

$$C_i = \mathbb{C}\left\{\sigma_1^{(i)}, \sigma_2^{(i)}, \cdots, \sigma_{n_i}^{(i)}\right\}.$$

The motivation for defining the chain groups is to define the homology groups, algebraic structures which capture some of the topological data of the simplicial complex. To this end we define the boundary operator $\partial_i : C_i \to C_{i-1}$

3.8 Definition. Let X be a simplicial complex as above with vertices $v_1, v_2, \cdots, v_{n_0}$. Then for $1 \leq i \leq d$, we define the boundary operator $\partial_i : C_i \to C_{i-1}$ (or $\partial_i(X)$,) to be the linear map determined as follows.

Let $1 \leq j(0) < j(1) < \cdots < j(i) \leq n_0$, then

$$\partial_i \left(\left[v_{j(0)}, v_{j(1)}, \cdots, v_{j(i)} \right] \right) = \sum_{k=0}^{i} (-1)^k \left[v_{j(0)}, v_{j(1)}, \cdots, \hat{v}_{j(k)}, \cdots, v_{j(i)} \right],$$

where the "hat" in $\hat{v}_{j(k)}$ indicates $v_{j(k)}$ has been deleted from the list, i.e.

$$\left[v_{j(0)}, v_{j(1)}, \cdots, v_{j(k-1)}, \hat{v}_{j(k)}, v_{j(k+1)}, \cdots, v_{j(i)} \right] = \left[v_{j(0)}, v_{j(1)}, \cdots, v_{j(k-1)}, v_{j(k+1)}, \cdots, v_{j(i)} \right] \in C_{i-1}.$$

We define $\partial_0 : C_0 \to 0$ to be the zero map.

The chain complex of X is then the set $(C_\bullet, \partial_\bullet) = \{(C_i, \partial_i) : 0 \leq i \leq d\}$, and is summarized in the following diagram.

$$C_d \xrightarrow{\partial_d} C_{d-1} \xrightarrow{\partial_{d-1}} \cdots \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0$$

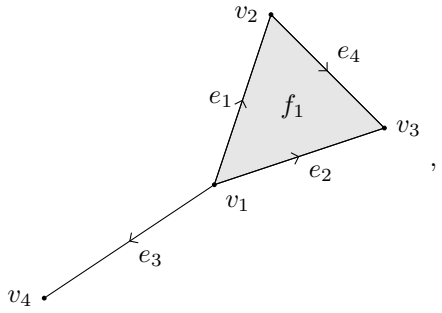
In light of boundary operators, we denote faces of a simplex as either even or odd. Recall that a face of an i-dimensional simplex is by definition an $(i-1)$-dimensional subsimplex.

3.9 Definition. A face $\left[v_{j(0)} \cdots, \hat{v}_{j(k)}, \cdots, v_{j(i)} \right]$ of a simplex $\left[v_{j(0)}, \cdots, v_{j(i)} \right]$ is said to be an *even face* if k is even, and an *odd face* if k is odd.

In other words, even faces of a simplex are the ones that get a coefficient of 1 in the boundary of the simplex, and odd simplices are the ones given a coefficient of -1. Note that a face is only even or odd *relative to the simplex that contains it.* So for example one simplex could be an even face in one simplex and an odd face in another.

In practice, we usually do not write out simplices as the complete list of vertices. Instead, we only have a list of simplices written in order and an indication of which are subsimplices of which, see the example below. A way to remember in a simplex which faces are even and which are odd, we begin by writing the faces in increasing order. Then the last of the faces (the greatest in the linear order) is even, and then we proceed from last to first, alternating between even and odd.

3.10 Example. Let us take the kite, K, introduced in example 3.5. I.e., K is the 2-dimensional simplicial complex



,

whose simplices we denote as

$$X = \{\ v_1, v_2, v_3, v_4,\ e_1, e_2, e_3, e_4,\ f_1\ \}$$

(v_1 is being identified with $[v_1]$, the edges e_i are written in order, and $f_1 = [v_1, v_2, v_3]$). Then the chain groups of X are

$$C_0 = \mathbb{C}\{v_1, v_2, v_3, v_4\}, \qquad C_1 = \mathbb{C}\{e_1, e_2, e_3, e_4\}, \qquad C_2 = \mathbb{C}\{f_1\},$$

and the boundary operators are determined on each simplex as follows.

$$\begin{array}{lll} \partial_0 : C_0 \to 0 & \partial_1 : C_1 \to C_0 & \partial_2 : C_2 \to C_1 \\ v_1 \mapsto 0 & e_1 \mapsto v_2 - v_1 & f_1 \mapsto e_4 - e_2 + e_1 \\ v_2 \mapsto 0 & e_2 \mapsto v_3 - v_1 & \\ v_3 \mapsto 0 & e_3 \mapsto v_4 - v_1 & \\ v_4 \mapsto 0 & e_4 \mapsto v_3 - v_2 & \end{array}$$

So, for instance, e_3 and e_1 are even faces of f_1, and e_2 is an odd face of f_1; in every edge the final endpoint is an even face and the initial enpoint is odd; and, we have that v_2 is an even face of e_1 but an odd face of e_4.

To motivate the coming theorem, let us compute $(\partial_1 \circ \partial_2)(f_1)$.

$$\begin{aligned} \partial_1\Big(\partial_2(f)\Big) &= \partial_1(e_3 - e_2 + e_1) \\ &= \partial_1(e_4) - \partial_1(e_2) + \partial_1(e_1) \\ &= (v_3 - v_2) - (v_3 - v_1) + (v_2 - v_1) \\ &= 0. \end{aligned}$$

As we see in the following theorem, this is no coincicdence.

An important invariant in algebraic topology is the homology group of a simplicial complex, which will be defined as a quotient of vector spaces in C_i. The following theorem is necessary for homology to make sense.

3.11 Theorem. *If $(C_\bullet, \partial_\bullet)$ is the chain complex of a simplicial complex, then we have, for each $1 \leq i \leq d$,*

$$\partial_{i-1} \circ \partial_i = 0.$$

Proof. Take a basis element $\sigma_r^{(i)}$ in C_i. We have the following computations.

$$\begin{aligned} (\partial_{i-1} \circ \partial_i)\left(\sigma_r^{(i)}\right) =& \partial_{i-1}\left(\partial_i\left(\sigma_r^{(i)}\right)\right) \\ =& \partial_{i-1}\left(\partial_i\left(\left[v_{j(0)}, v_{j(1)}, \cdots, v_{j(i)}\right]\right)\right) \qquad \text{(for some } j : [i] \to [n_0]\text{)} \\ =& \partial_{i-1}\left(\sum_{k=0}^{i} (-1)^k \left[v_{j(0)}, \cdots, \hat{v}_{j(k)}, \cdots, v_{j(i)}\right]\right) \\ =& \sum_{k=0}^{i} (-1)^k \partial_{i-1}\left(\left[v_{j(0)}, \cdots, \hat{v}_{j(k)}, \cdots, v_{j(i)}\right]\right) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{i} (-1)^k \partial_{i-1} \left(\left[v_{j(0)}, \cdots, v_{j(k-1)}, v_{j(k+1)}, \cdots, v_{j(i)} \right] \right) \\
&= \sum_{k=0}^{i} (-1)^k \left(\left(\sum_{\ell=0}^{k-1} (-1)^\ell \left[v_{j(0)}, \cdots, \hat{v}_{j(\ell)}, \cdots, v_{j(k-1)}, v_{j(k+1)}, \cdots, v_{j(i)} \right] \right) \right. \\
&\qquad \left. + \left(\sum_{\ell=k+1}^{i} (-1)^{\ell-1} \left[v_{j(0)}, \cdots, v_{j(k-1)}, v_{j(k+1)}, \cdots, \hat{v}_{j(\ell)}, \cdots, v_{j(i)} \right] \right) \right) \\
&= \sum_{k=0}^{i} (-1)^k \left(\left(\sum_{\ell=0}^{k-1} (-1)^\ell \left[v_{j(0)}, \cdots, \hat{v}_{j(\ell)}, \cdots, \hat{v}_{j(k)}, \cdots, v_{j(i)} \right] \right) \right. \\
&\qquad \left. + \left(\sum_{\ell=k+1}^{i} (-1)^\ell (-1) \left[v_{j(0)}, \cdots, \hat{v}_{j(k)}, \cdots, \hat{v}_{j(\ell)}, \cdots, v_{j(i)} \right] \right) \right) \\
&= \sum_{k=0}^{i} \sum_{\ell=0}^{k-1} (-1)^{k+\ell} \left[v_{j(0)}, \cdots, \hat{v}_{j(\ell)}, \cdots, \hat{v}_{j(k)}, \cdots, v_{j(i)} \right] \\
&\quad - \sum_{k=0}^{i} \sum_{\ell=k+1}^{i} (-1)^{k+\ell} \left[v_{j(0)}, \cdots, \hat{v}_{j(k)}, \cdots, \hat{v}_{j(\ell)}, \cdots, v_{j(i)} \right] \\
&= \sum_{0 \le \ell < k \le i} (-1)^{k+\ell} \left[v_{j(0)}, \cdots, \hat{v}_{j(\ell)}, \cdots, \hat{v}_{j(k)}, \cdots, v_{j(i)} \right] \\
&\quad - \sum_{0 \le k < \ell \le i} (-1)^{k+\ell} \left[v_{j(0)}, \cdots, \hat{v}_{j(k)}, \cdots, \hat{v}_{j(\ell)}, \cdots, v_{j(i)} \right] \\
&= \sum_{0 \le \ell < k \le i} (-1)^{k+\ell} \left[v_{j(0)}, \cdots, \hat{v}_{j(\ell)}, \cdots, \hat{v}_{j(k)}, \cdots, v_{j(i)} \right] \\
&\quad - \sum_{0 \le \ell < k \le i} (-1)^{k+\ell} \left[v_{j(0)}, \cdots, \hat{v}_{j(\ell)}, \cdots, \hat{v}_{j(k)}, \cdots, v_{j(i)} \right] \quad = 0
\end{aligned}
$$

□

3.12 Corollary. *Let $(C_\bullet, \partial_\bullet)$ be the chain complex of a simplicial complex. Then we have, for each $0 \le i < d$, that $\mathrm{im}(\partial_{i+1})$ is a subspace of $\ker(\partial_i)$ (both as subspaces of C_i).*

The proof is immediate from theorem 3.11.

3.13 Definition. Let X be a simplicial complex as above, with chain complex $(C_\bullet, \partial_\bullet)$ and $0 \le i < d$. Then the i-th homology group of X is the complex vector space given by the quotient

$$H_i = \frac{\ker(\partial_i)}{\mathrm{im}(\partial_{i+1})}$$

(We may also write $H_i(X)$ when X needs to be specified.) Furthermore, the i-th Betti number β_i (or $\beta_i(X)$) is defined as the dimension of the i-th homology group.

$$\beta_i = \dim(H_i)$$

There would be no point for us in diverting too much on homology theory. All we need to know for this thesis is the following facts:

1. Homology groups (and hence Betti numbers) are topological invariants, and even homotopy-type invariants. Hence it is heavily used theoretically to distinguish non-homeomorphic topological spaces. For example, to prove that $\mathbb{R}^m$ and $\mathbb{R}^n$ are not homeomorphic for $m \neq n$, one shows that upon removing a single point from each, the remaining spaces have different homologies unless $m = n$.

2. Betti numbers have a nice, intuitive geometric interpretation. The Betti number $\beta_0(X)$ is exactly the number of path-connected components of X as a topological space; $\beta_1(X)$ is equal to the number of "holes" in the space X (for example, if X has the shape of a solid coffee mug, then it has $\beta_1(X) = 1$, if X is has the shape of two circles with a single point in common, then $\beta_1(X) = 2$); $\beta_2(X)$ is the number of "hollow spaces" enclosed by X, for example both a sphere and a torus have $\beta_2 = 1$. And so on in higher dimensions.

3. The sequence of Betti numbers is used often in pure and applied mathematics to encode many properties of a topological space. In particular, it is often possible to obtain the sequence of Betti numbers for an unknown space either theoretically, e.g. via cohomology, or experimentally, e.g. via topological data analysis (see section 6.2). In both cases, knowledge of the homology or Betti numbers of a space reveals important information.

For the precise statements of these properties, see any standard textbook on algebraic topology, such as [11, 7].

An equivalent way to define a chain complex is by writing

$$C_\bullet = \bigoplus_{i=0}^{d} C_i \qquad \text{and} \qquad \partial_\bullet : C_\bullet \to C_\bullet,$$

with $\partial_\bullet$ a linear map defined by the linear property and its value on basis vectors

$$\partial_\bullet\left(\sigma_j^{(i)}\right) = \partial_i\left(\sigma_j^{(i)}\right).$$

In this notation, theorem 3.11 simplifies to

$$\partial_\bullet^2 = 0$$

We will, from now on, use the isomorphism $C_\bullet \cong \mathbb{C}^{n_0+n_1+\cdots+n_d} = \mathbb{C}^N$, with the standard basis on $\mathbb{C}^N$ written as

$$\begin{Bmatrix} e_1^{(0)}, e_2^{(0)}, \cdots, e_{n_0}^{(0)}, \\ e_1^{(1)}, e_2^{(1)}, \cdots, e_{n_1}^{(1)}, \\ \vdots \\ e_1^{(d)}, e_2^{(d)}, \cdots, e_{n_d}^{(d)} \end{Bmatrix}$$

and the isomorphism given by $\sigma_j^{(i)} \mapsto e_j^{(i)}$. We will use the standard inner product $\langle a, b\rangle$ on $\mathbb{C}^N$. (That is, $\langle a, b\rangle = b^* a$, where A^* is the conjugate transpose of a matrix A, $a, b \in \mathbb{C}^N$ are seen as column vectors, and $(b^*)(a)$ is seen as matrix multiplication.)

3.3 The Boundary Operator

We denote by $J(X)$, or just J, the *boundary matrix* of X, which is by definition the natural map $J : \mathbb{C}^N \to \mathbb{C}^N$ corresponding to $\partial_\bullet$ under the isomorphism above. J is simply the map such that for each $i = 0, \cdots, d$ and $j \in [n_i]$, we have

$$J\big(e_j^{(i)}\big) = \sum_{\substack{0\le k\le d\\ 1\le \ell\le n_k}} \alpha_\ell^{(k)} e_\ell^{(k)}, \quad \text{for some constants } \alpha_\ell^{(k)} \in \mathbb{C}, \text{ if and only if } \quad \partial_\bullet\big(\sigma_j^{(i)}\big) = \sum_{\substack{0\le k\le d\\ 1\le \ell\le n_k}} \alpha_\ell^{(k)} \sigma_\ell^{(k)}.$$

In other words, J is the unique linear map that makes the following diagram commute.

$$\begin{array}{ccc} C_\bullet & \xrightarrow{\cong} & \mathbb{C}^N \\ {\scriptstyle\partial_\bullet}\downarrow & & \downarrow{\scriptstyle J} \\ C_\bullet & \xrightarrow{\cong} & \mathbb{C}^N. \end{array}$$

More concretely, if $\psi : C_\bullet \to \mathbb{C}^N$ is the isomorphism given by $\psi\big(\sigma^{(i)} j\big) = e^{(i)} j$, then we have

$$J = \psi \partial_\bullet \psi^{-1}.$$

Similarly, we define $d_i : \mathbb{C}^{n_i} \to \mathbb{C}^{n_{i-1}}$ to be the natural map corresponding to $\partial_i : C_i \to C_{i-1}$. In other words, the following diagram commutes:

$$\begin{array}{ccc} C_i & \xrightarrow{\cong} & \mathbb{C}^{n_i} \\ {\scriptstyle\partial_i}\downarrow & & \downarrow{\scriptstyle d_i} \\ C_{i-1} & \xrightarrow{\cong} & \mathbb{C}^{n_{i-1}}, \end{array}$$

and we have

$$d_i = \psi_{|C_i} \partial_i \psi^{-1}_{|C_i}.$$

The direct-sum decomposition $\mathbb{C}^N = \mathbb{C}^{n_0} \oplus \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_d}$ gives J the structure of a block matrix, where the sizes of the block are given by the numbers $(n_0, n_1, \cdots, n_d)$.

We will use the following notation for block matrices of this form.

3.14 Notation. If A is an $N \times N$ matrix with the block structure given by $N = n_0 + n_1 + \cdots + n_d$, then we will say A is an $(n_0, n_1, \cdots, n_d)$-block matrix, and write the blocks of A as $\big[A^{(i,k)}\big]_{i,k=0}^{d}$, that is

$$A = \left(\begin{array}{c|c|c|c} A^{(0,0)} & A^{(0,1)} & \cdots & A^{(0,d)} \\ \hline A^{(1,0)} & A^{(1,1)} & \cdots & A^{(1,d)} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline A^{(d,0)} & A^{(d,1)} & \cdots & A^{(d,d)} \end{array}\right),$$

where $A^{(i,k)}$ is an $n_i \times n_k$ matrix whose elements are given by

$$A^{(i,k)} = \Big[a_{j,\ell}^{(i,k)}\Big]_{j,\ell} \quad (1 \le j \le n_i, \quad 1 \le \ell \le n_k).$$

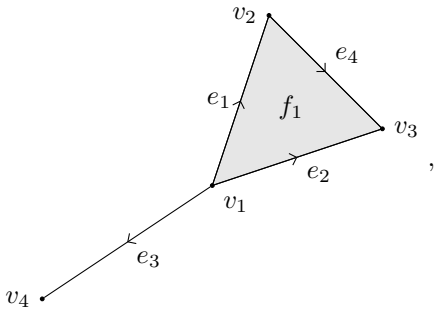
Since $J : \mathbb{C}^N \to \mathbb{C}^N$ restricts on each of the subspaces $\mathbb{C}^{n_0} \oplus \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_d}$ to a map $\mathbb{C}^{n_i} \to \mathbb{C}^{n_{i-1}}$, then $J = \left[J^{(i,k)}\right]_{i,k=0}^{d}$ is an $(n_0, n_1, \cdots, n_d)$-block matrix whose only nonzero blocks are $J^{(i,i-1)}$ for $i = 1, 2, \cdots, d$. And we get precisely that $J^{(i,i-1)} = d_i$.

$$\left(\begin{array}{c|c|c|c|c|c} 0 & d_1 & 0 & \cdots & 0 & 0 \\ \hline 0 & 0 & d_2 & \cdots & 0 & 0 \\ \hline 0 & 0 & 0 & \ddots & 0 & 0 \\ \hline \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hline 0 & 0 & 0 & \cdots & 0 & d_d \\ \hline 0 & 0 & 0 & \cdots & 0 & 0 \end{array}\right).$$

And finally, we observe that the entries of d_i are given by

$$(d_i)_{j,\ell} = \begin{cases} 1, & \text{if } \sigma_\ell^{(i-1)} \text{ is an even face of } \sigma_j^{(i)}, \\ -1, & \text{if } \sigma_\ell^{(i-1)} \text{ is an odd face of } \sigma_j^{(i)}, \\ 0, & \text{else.} \end{cases}$$

3.15 Example. Let us see what the boundary matrix $J(K)$ is, for the kite complex K in example 3.5



We have, for this complex, $N = 9$, so

$$C_\bullet \cong \mathbb{C}^9 = \mathbb{C}^4 \oplus \mathbb{C}^4 \oplus \mathbb{C}^1 = \mathbb{C}\{v_1, v_2, v_3, v_4\} \oplus \mathbb{C}\{e_1, e_2, e_3, e_4\} \oplus \mathbb{C}\{f_1\}.$$

Therefore, $J(K)$ is a $(4, 4, 1)$-block matrix

$$J(K) = \left(\begin{array}{c|c|c} 0 & d_1 & 0 \\ \hline 0 & 0 & d_2 \\ \hline 0 & 0 & 0 \end{array}\right) = \left(\begin{array}{cccc|cccc|c} 0 & 0 & 0 & 0 & -1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

4 Finite-Dimensional Hodge Theorem

In this chapter we continue using the definitions and notation of simplicial complexes, their homology and their corresponding boundary matrices to provide a proof of the finite-dimensional version of the Hodge theorem. For this chapter, let us take a simplicial complex $X = \bigcup_{i=0}^{d} X_i$ with $|X_i| = n_i$, $N = n_0 + \cdots + n_d$, with boundary matrix J. As in the previous chapter, $J = \left[J^{(i,k)}\right]_{i,k=0}^{d}$ is an $(n_0, \cdots, n_d)$-block matrix whose only nonzero blocks are $J^{(i-1,i)} = d_i$.

4.1 Definition. The *combinatorial Laplacian* of X is defined to be the $N \times N$ matrix $\Delta = JJ^* + J^*J$.

4.2 Remark. Δ can be expressed as a block diagonal matrix of the operators Δ_i, which we define by $\Delta_i = d_i^* d_i + d_{i+1} d_{i+1}^*$.

$$\Delta = \left(\begin{array}{c|c|c|c|c} d_1 d_1^* & 0 & \cdots & & 0 \\ \hline 0 & d_1^* d_1 + d_2 d_2^* & & & \\ \hline \vdots & & \ddots & & \vdots \\ \hline & & & d_{d-1}^* d_{d-1} + d_d d_d^* & 0 \\ \hline 0 & & \cdots & 0 & d_d^* d_d \end{array}\right)$$

$$= \begin{pmatrix} \Delta_0 & & & & 0 \\ & \Delta_1 & & & \\ & & \ddots & & \\ & & & \Delta_{d-1} & \\ 0 & & & & \Delta_d \end{pmatrix}.$$

The Hodge theorem is a key result in differential geometry and algebraic geometry. However, the standard statement and proof of the theorem is in a general setting that requires much knowledge not relevant to this thesis. Here, we only deal with a finite-dimensional, linear-algebraic version of the theorem. A similar finite-dimensional version of this theorem was posted online and proved in a similar way in [10].

Recall that the i-th homology group of X is defined by

$$H_i(X) = \frac{\ker(\partial_i)}{\mathrm{im}(\partial_{i+1})}.$$

The following theorem is a mere reformulation of the results of Eckmann in [4].

4.3 Theorem. *We have the following vector space isomorphism.*

$$\ker(\Delta_i) \cong H_i(X)$$

Proof. First, we will show that $\ker(\Delta_i) = \ker(d_i) \cap \ker(d_{i-1}^*)$. It is clear that $\ker(d_i) \cap \ker(d_{i-1}^*) \subset \ker(\Delta_i)$. To show the other containment take $x \in \ker(\Delta_i)$. Since $\Delta_i x = 0$, we also have that $\langle \Delta_i x, x\rangle = 0$. On the other hand, we compute

$$\begin{aligned}\langle \Delta_i x, x\rangle &= \langle (d_i^* d_i + d_{i+1} d_{i+1}^*)x, x\rangle \\ &= \langle d_i^* d_i x + d_{i+1} d_{i+1}^* x, x\rangle \\ &= \langle d_i^* d_i x, x\rangle + \langle d_{i+1} d_{i+1}^* x, x\rangle \\ &= \langle d_i x, d_i x\rangle + \langle d_{i+1}^* x, d_{i+1}^* x\rangle.\end{aligned}$$

Since $\langle u, u\rangle \geq 0$ for any $u \in \mathbb{R}^{n_i} \cong C_i$, we must have

$$\langle d_i x, d_i x\rangle + \langle d_{i+1}^* x, d_{i+1}^* x\rangle. = 0,$$

which implies that $d_i x = 0$ and $d_{i+1}^* x = 0$. Thus we conclude $x \in \ker(d_i) \cap \ker(d_{i-1}^*)$, and therefore $\ker(\Delta_i) = \ker(d_i) \cap \ker(d_{i-1}^*)$.

Now, we will show that

$$\mathbb{R}^{n_i} \cong H_i \oplus \ker(d_{i+1}^*)^\perp \oplus \ker(d_i)^\perp. \tag{4.1}$$

This will prove the theorem because we also have the orthogonal decomposition

$$\begin{aligned}\mathbb{R}^{n_i} &= \big(\ker(d_{i+1}^*) \cap \ker(d_i)\big) \oplus \big(\ker(d_{i+1}^*) \cap \ker(d_i)\big)^\perp \\ &= \big(\ker(d_{i+1}^*) \cap \ker(d_i)\big) \oplus \ker(d_{i+1}^*)^\perp \oplus \ker(d_i)^\perp.\end{aligned}$$

So by uniqueness of direct-sum decompositions, we would conclude

$$H_i \cong \ker(d_i) \cap \ker(d_{i-1}^*),$$

as desired.

We decompose the space $\ker(d_i)$ as a direct sum

$$\ker(d_i) \cong \mathrm{im}(d_{i+1}) \oplus H_i.$$

This decomposition is immediate from the definition of H_i as the quotient

$$H_i = \frac{\ker(d_i)}{\mathrm{im}(d_{i+1})}.$$

Now we decompose C_i (or $\mathbb{R}^{n_i}$).

$$\begin{aligned}\mathbb{R}^{n_i} &= \ker(d_i) \oplus \ker(d_i)^\perp \\ &\cong (\mathrm{im}(d_{i+1}) \oplus H_i) \oplus \ker(d_i)^\perp \\ &\cong H_i \oplus \mathrm{im}(d_{i+1}) \oplus \ker(d_i)^\perp \\ &= H_i \oplus \ker(d_{i+1}^*)^\perp \oplus \ker(d_i)^\perp\end{aligned}$$

This is precisely as in equation (4.1), and thus concludes the proof. (In the last line, we used the result from finite-dimensional linear algebra that $\mathrm{im}(T) = \ker(T^*)^\perp$ for any linear map T.) □

5 Spectral Analysis of Simplicial Complexes

This chapter is based on the results of Vargas in [15], and it gives the main results of this thesis. They are expressed in a slightly more general version of non-commutative probability called operator-valued non-commutative probability.

5.1 Operator-Valued Probability Spaces

5.1 Definition. A triple $(\mathcal{A}, \mathcal{B}, \mathbf{E})$ is an *Operator-Valued Probability Space* (or *OVPS*) if and only if $\mathcal{A}$ is a unital algebra over $\mathbb{C}$, $\mathcal{B} \subset \mathcal{A}$ is a unital subalgebra, and $\mathbf{E} : \mathcal{A} \to \mathcal{B}$ is a *conditional expectation*, which is by definition a linear map satisfying the following conditions.

(i) $\mathbf{E}(b) = b$, for all $b \in \mathcal{B}$.

(ii) If $a \in \mathcal{A}$, and $b_1, b_2 \in \mathcal{B}$, then $\mathbf{E}(b_1 a b_2) = b_1 \mathbf{E}(a) b_2$

5.2 Remark. Note that condition (ii) implies $\mathbf{E}(1) = 1$.

5.3 Example. Every non-commutative probability space $(\mathcal{A}, \varphi)$ is also an OVPS $\Big(\mathcal{A}, \mathbb{C}\{1_{\mathcal{A}}\}, \varphi'\Big)$, where for $a \in \mathcal{A}$, $\varphi'(a) = \varphi(a)1$.

5.4 Example. Given a non-commutative probability space $(\mathcal{A}, \varphi)$, we can consider the space of random matrices as an OVPS $(M_n(\mathcal{A}), M_n(\mathbb{C}), \mathbf{E})$, where $\mathbf{E}$ is given by mapping a random matrix into the deterministic matrix given by "entry-wise expectation":

$$\mathbf{E}\Big(\left[a_{ij}\right]_{i,j} \Big) = \Big[\varphi\left(a_{ij}\right) 1 \Big]_{i,j}.$$

5.5 Example. Given a measure space $(\Omega, \mathcal{F}, \mu)$, and partition $\Omega = \bigcup_{i=1}^{n} S_i$ into measurable sets S_i, we take χ_i to be the characteristic function of S_i, $\text{alg}(1, \chi_1, \cdots, \chi_n) = \text{span}(\chi_1, \cdots, \chi_n)$ we have

$$\Big(\ L^{\infty}(\Omega, \mu), \ \text{alg}(\chi_1, \cdots, \chi_n), \ \mathbf{F} \Big),$$

where $\mathbf{F} = \sum_{i=1}^{n} \mathbf{E}_i$, and

$$\mathbf{E}_i(f) = \frac{\mathbf{E}(f \cdot \chi_i)}{\mathbf{E}(\chi_i)} \chi_i.$$

5.6 Example. Take $N = n_0 + n_1 + \cdots + n_d$, and the unital algebra $\mathcal{A} = \mathcal{M}_N(\mathbb{C})$. We will look at elements in $\mathcal{A}$ as $(n_0, n_1, \cdots, n_d)$-block matrices

$$A = \Big[A^{(i,k)} \Big]_{i,k=0}^{d},$$

where $A^{(i,k)}$ is any $n_i \times n_k$ matrix with complex coefficients,

$$A^{(i,k)} = [a_{j,\ell}^{(i,k)}]_{j,\ell}.$$

Next, we develop an OVPS structure on $\mathcal{A}$. For each $0 \leq i \leq d$, let $P_i \in \mathcal{A}$ be the block matrix

$$P_i = [A^{(j,\ell)}] \quad \text{such that} \quad A^{(j,\ell)} = \begin{cases} I_{n_i}, & \text{if } j = \ell = i, \\ 0, & \text{else.} \end{cases}$$

That is,

$$P_i = \begin{pmatrix} 0_{n_0} & & & & & & 0 \\ & 0_{n_1} & & & & & \\ & & \ddots & & & & \\ & & & I_{n_i} & & & \\ & & & & \ddots & & \\ 0 & & & & & & 0_{n_d} \end{pmatrix},$$

where, for all $n \in \mathbb{N}$, 0_n and I_n are respectively the $n \times n$ zero matrix and identity matrix. In other words, using the basis $\{e_j^{(i)} | 0 \leq i \leq d, 1 \leq j \leq n_i\}$ for $\mathbb{R}^N = \mathbb{R}^{n_0} \oplus \cdots \oplus \mathbb{R}^{n_d}$ used above, $P_i : \mathbb{R}^N \to \mathbb{R}^N$ is the operator determined by

$$P_i\left(e_\ell^{(k)}\right) = \begin{cases} e_\ell^{(k)}, & i = k \\ 0, & i \neq k, \end{cases}$$

for all $e_\ell^{(k)}$ $0 \leq k \leq d$ and $\ell \in [n_k]$, and its matrix is as described.

Then take $\mathcal{B} = \text{alg}(P_0, \cdots, P_d)$, the subalgebra generated by the matrices P_i. Note $\sum P_i = I_N \in \mathcal{B}$, so $\mathcal{B}$ is unital. We will now define a conditional expectation $\mathbf{E}$ as follows. For all $0 \leq i \leq d$ let $\mathbf{E}_i$ be defined for all$a \in \mathcal{A}$ by

$$\mathbf{E}_i(a) = \frac{\text{Tr}(P_i a P_i)}{\text{Tr}(P_i)} P_i.$$

Then we define

$$\mathbf{E} = \sum_{i=0}^{d} \mathbf{E}_i,$$

and indeed we have that

$$(\mathcal{A}, \mathcal{B}, \mathbf{E})$$

is an operator-valued probability space.

This last example is the one we will use throughout the rest of this chapter.

In previous chapters we saw how a topological space X with the structure of a simplicial complex can be encoded by its boundary matrix J. We also saw how the Hodge theorem gives a method for computing information about the homology of the topological space; namely, computing the kernels of the blocks of $\Delta = J^*J + JJ^*$ gives the Betti numbers $\beta_0, \cdots, \beta_d$ of X. Now, in the notation from the last example, we

can practically treat the topological space of a simplicial complex as a random variable in an operator-valued probability space, namely by taking its boundary matrix as a random variable $J \in \mathcal{M}_N(\mathbb{C})$, where N is the number of simplices in X. Now, the result we present in this thesis is one example of the ways in which it may be advantageous to study topological spaces from the point of view of non-commutative probability.

5.7 Definition. Let A be a random variable in the OVPS $\big(\mathcal{M}_N(\mathbb{C}),\, \mathrm{alg}(P_0, \cdots, P_n),\, \mathbf{E}\big)$, i.e. A is the block matrix $A = \big[A^{(i,k)}\big]_{i,k=0}^{d}$. Consider each of the $n_i \times n_i$ blocks $A^{(i,i)}$ in the diagonal as random variables in the $*$-probability space $\big(\mathcal{M}_{n_i}(\mathbb{C}), \mathrm{tr}_{n_i}\big)$. Then $\mu = (\mu_0, \mu_1, \cdots, \mu_d)$ is said to be the *multivariate distribution* of A in $\big(\mathcal{M}_N(\mathbb{C}),\, \mathrm{alg}(P_0, \cdots, P_n),\, \mathbf{E}\big)$, if and only if, for each $i = 0, \cdots, d$, μ_i is a compactly supported measure on $\mathbb{C}$ and $A^{(i,i)}$ has distribution μ_i, that is,

$$\mathrm{tr}_{n_i}\big((A^{(i,i)})^k\big) = \int_{\mathbb{C}} t^k d\mu_i(t), \qquad \forall k \in \mathbb{N}.$$

Now, take a simplicial complex X, with boundary matrix J, and combinatorial Laplacian Δ, defined in chapter 4 as $\Delta = J^*J + JJ^*$. Recall that Δ is a block-diagonal matrix

$$\Delta = \begin{pmatrix} \Delta_0 & & 0 \\ & \ddots & \\ 0 & & \Delta_d \end{pmatrix},$$

where, for each i, we have $\Delta_i = d_i^* d_i + d_{i+1} d_{i+1}^*$, with Δ. As we saw in the Hodge theorem from chapter 4, each Δ_i has a kernel isomorphic to the i-th homology group of X, i.e.

$$\ker(\Delta_i) \cong H_i. \tag{5.1}$$

In particular, we obtain the Betti numbers (defined as $\beta_i = \dim(H_i)$) by

$$\dim\ker(\Delta_i) = \beta_i \tag{5.2}$$

Now, we want to interpret equation (5.2) as a property of operator-valued random variables. We then observe that the kernel of a matrix A equals the eigenspace of A with respect to the eigenvalue 0. Eigenvalues, in turn, are recorded, with multiplicity, in the analytic distribution of A, as a random variable. We now present the main result of this thesis.

5.8 Theorem (Vargas [15])**.** *Let $X = \bigcup_{i=0}^{d} X_i$ be a d-dimensional simplicial complex with i-dimensional simplices $X_i = \{\sigma_1^{(i)}, \cdots, \sigma_{n_i}^{(i)}\}$, and $N = n_0 + \cdots + n_d$. Let J be the boundary matrix of X. Consider the OVPS*

$$\Big(\mathcal{M}_N(\mathbb{C}),\, \mathrm{alg}(P_0, \cdots, P_n),\, \mathbf{E}\Big),$$

introduced in the example above, and the combinatorial Laplacian $\Delta = J^*J + JJ^*$, as a random variable in $\mathcal{M}_N(\mathbb{C})$. Let $\mu = (\mu_0, \cdots, \mu_d)$ be the multivariate distribution of Δ. Then the Betti numbers of X can be recovered as the (scaled) weights of the point 0 in each of the distributions in μ, namely

$$\mu_i\big(\{0\}\big) = \frac{1}{n_i}\beta_i.$$

Proof. This theorem is a direct consequence of the Hodge theorem. From the Hodge theorem we have that $\dim\ker(\Delta_i) = \beta_i$. or equivalently, that the multiplicity of the eigenvalue 0 in Δ_i is equal to β_i. But as we saw in example 1.18, the distribution of Δ is given by

$$\mu_i = \sum_{j=1}^{n_i} \frac{1}{n_i} \delta_{\lambda_j},$$

where $\lambda_1, \cdots, \lambda_{n_i}$ are the eigenvalues of Δ_i, and thus we conclude that $\mu_i\big(\{0\}\big) = \frac{1}{n_i}\beta_i$, as desired. □

This result gives a new way of computing Betti numbers from probability distributions. The computation of Betti numbers is very important in some aspects of applied topology. For example, the field of Topological Data Analysis (TDA), studies data in a topological, rather than a statistical way. The idea behind TDA is to construct a simplicial complex of a suitable dimension to resemble a set of data points, and then being able to study the shape of the data in a rather qualitative way. The computations of Betti numbers is an essential part of the process. Specifically, bar-codes and histograms are specific ways in which TDA records topological information about data points, and non-commutative probability has the potential of being a great computational tool for this. For an introduction to TDA see [3], and for the uses of Non-commutative probability in constructing histograms see section 5.1 of [15]

As the avid reader will notice, this chapter has made no mention of freeness. Indeed, the topics studied in this chapter are in a very early stage of development, and at this point there is very little understanding of what role free independence might play in the non-commutative-probability view of topological spaces. In other words, the understanding of simplicial complexes, or topological spaces in general, as non-commutative random variables, is currently very superficial and involves only the abstract framework of non-commutative probability spaces. The results of this chapter are only a taste of what this theory could be developed into. Once free probability comes into the picture, we will then be able to apply the powerful tools we saw in chapter 2, such as free independence, free products, the combinatorics of freeness, free cumulants, and the $\mathcal{R}$-transform. Until we acquire a conceptual understanding of freeness for simplicial complexes, this theory will remain very limited. In the next chapter we explore some specific directions on which this theory may be developed.

6 Directions for Further Research

The main purpose of this thesis so far has been on one hand to introduce the reader to free probability and its powerful techniques, and on the other hand to expose the recently developed, and vastly unexplored, connections between non-commutative probability and algebraic topology. It is currently unclear how far research could advance in this direction, but there are various possible starting points and open questions which might indicate a way forward.

6.1 Topological Interpretations of Freeness

In chapter 3 we reduced simplicial complexes to their corresponding boundary matrices, and in chapter 5 we developed a way of seeing them as operator-valued non-commutative random variables. As we briefly mentioned earlier, the concept of free independence never came into the picture. In the opinion of the author, the next most important work to do in this research area is understanding the role of free independence of simplicial complexes from a topological point of view. Freeness is important because it is the key concept that allows us to use all the tools introduced in chapter 2, such as the moment-cumulant formulas, the vanishing of free cumulants and the additivity of the $\mathcal{R}$-transform.

One first step towards understanding freeness of simplicial complexes , that is, freeness of their boundary matrices, or of their chain complexes, is understanding their moments. Let X be a simplicial complex with chain complex $(C_\bullet, \partial_\bullet)$, and boundary matrix J. Consider J to be a random variable in the OVPS $\big(\mathcal{M}_N(\mathbb{C}), \text{alg}(P_0, \cdots, P_n), \mathbf{E}\big)$. Then one important open question is: *What is the meaning of the $*$-moments of J?* We have seen in one special case, how of the moments $\mathbf{E}\big((J^*J + JJ^*)^n\big)$ give the Betti numbers. Next we would like to understand what other topological information might be encoded in the distribution or the multi-variate distribution of J.

In order to understand moments better, it may be useful to look at the moments of other block matrices associated to a simplicial complex; for example, some generalizations of adjacency and incidence matrices to simplicial complexes.

Some progress has been done in understanding the moments of certain random variables related to graphs and simplicial complexes. Given a graph G we have upper adjacency matrices A_0^1, lower adjacency matrices A_1^0, incidence matrices I_G, and, if the graph is oriented, boundary matrices J_G. Taking these matrices as elements in the $*$-probability space $\mathcal{M}_n(\mathbb{C})$, (where the meaning of n varies depending on what matrix we are using,) some moments may be interpreted as counting certain paths of points along edges, paths of edges

along points, or other interpretations. Adjacency matrices may be generalized in simplicial complexes to the matrices A_i^j (there is a 1 if two i-dimensional simplices are adjacent through a j-dimentional simplex). This might be a good place to start studying the topological or combinatorial meanings of moments. We could also take the $*$-probability space $(\mathcal{M}_n(\mathbb{C}), \pi)$, with the non-tracial expectation $\pi : \mathcal{M}_n(\mathbb{C}) \to \mathbb{C}$, given by $\pi\Big([a_{i,j}]\Big) = a_{1,1}$. This space has the advantage that it is considerably easier to keep track the combinatorial or topological meaning of the moments of a matrix.

One additional way in which freeness could be studied further is through certain products of graphs, and by extension, products of simplicial complexes. In particular we are interested in the free product of graphs, see for example [1]. Although it results in an infinite graph, work could be done to transfer all of its information to a suitable kind of random variable, e.g some bounded linear map on a Hilbert space. In [8], the authors study how certain algebraic-topological operations on simplicial complexes affect the combinatorial Laplacian.

6.2 Non-Commutative Probability for Topological Data Analysis

Topological Data Analysis, or TDA, is a modern technique for reading large data sets in a qualitative way. The data comes in the form of points (or point clouds) from a metric space, and the idea is that, instead of studying the data with the standard techniques of statistics, one studies the "shape of the data". Namely, points are joined to one another to form graphs or simplicial complexes, and the topological properties of the resulting structure are taken to be properties of the data set. The way one constructs a simplicial complex out of data points in a metric space is by specifying criteria under which two points are joined by a line, or, more generally, $n+1$ pints are joined by an n-simplex. This will in turn create a simplicial complex with the data points as the vertices; this complex is the object to be studied. Of course, the resulting space depends of the "criteria" used for joining points with simplices. Two examples of simplicial complex costructions are the Čech complex and the Vietoris-Rips Complex. These constructions, however, vary according to a parameter. Namely, the parameter that indicates "how close together" a cluster of points needs to be for the points to be connected by a simplex. Finally, instead of fixing this parameter to obtain a fixed complex, one allows the parameter to vary, and the collection of varying complexes is taken as a whole. Then the properties of this varying simplicial complex serve to describe the data set.

This is done via persistence diagrams. As the parameter changes, the homology of the resulting complex changes, either by the addition or the removal of basis vectors. Persistence diagrams keep track of how long each basis vector of the homology stays. The basis vectors that stay longer are said to be persistent, and are considered as more pronounced properties of the data set. Homology vectors that only appear for a short period of the parameter are neglected as noise. For a full treatment of topological data analysis, see [3]. The book [6] has an introduction to topological data analysis as well as many other topics in applied topology.

In TDA, Betti numbers are one of the essential properties that tell us about the "shape of data". So, from

a purely computational point of view, the connection between Non-commutative probability and algebraic topology might be very fruitful. The task of computing the dimension of homology is reduced to the task of computing eigenvalues, or equivalently, eigenvalue distributions.

6.3 On Non-Commutative Mathematics

Beginning with Voiculescu's work on free probability, there has been a recent movement to rewrite the many branches of mathematics in a non-commutative setting. A great part of this movement is illustrated by what is called free non-commutative function theory [9], which is a development of analysis in the non-commutative setting, along with its notion of integration, etc. This theory is currently in early stages.

The topics covered in this work, therefore may contribute to the future developments of a non-commutative framework for topology and geometry.

References

[1] Luigi Accardi, Romuald Lenczewski, and Rafał Sałapata. Decompositions of the free product of graphs. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 10(03):303–334, 2007.

[2] Naum Akhiezer. *The Classical Moment Problem*. Oliver & Boyd, 1965.

[3] Gunnar Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

[4] Beno Eckmann. Harmonische funktionen und randwertaufgaben in einem komplex. *Commentarii Mathematici Helvetici*, 17(1):240–255, 1944.

[5] Gerald B Folland. *Real analysis: modern techniques and their applications*. John Wiley & Sons, 1999.

[6] Robert W Ghrist. *Elementary applied topology*. 2014.

[7] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2001.

[8] Danijela Horak and Jürgen Jost. Spectra of combinatorial laplace operators on simplicial complexes. *Advances in Mathematics*, 244:303–336, 2013.

[9] Dmitry S Kaliuzhnyi-Verbovetskyi and Victor Vinnikov. *Foundations of free noncommutative function theory*, volume 199. American Mathematical Soc., 2014.

[10] Dan Lee. Baby hodge theorem. `http://qcpages.qc.cuny.edu/~dlee2/babyhodge.pdf`. Accessed: 2018-10-15.

[11] James R Munkres. *Elements of Algebraic Topology*. CRC Press, 1993.

[12] Alexandru Nica and Roland Speicher. *Lectures on the Combinatorics of Free Probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge Uniersity Press, 2006.

[13] Roland Speicher. Multiplicative functions on the lattice of non-crossing partitions and free convolution. *Mathematische Annalen*, 298(1):611–628, 1994.

[14] Elias M. Stein and Guido Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton Mathematics Series. Princeton University Press, 1971.

[15] Carlos Vargas. Non-commutative probability theory for topological data analysis. *preprint*, 2017.

[16] Dan-Virgil Voiculescu. Symmetries of some reduced free product C^*-algebras. In *Operator algebras and their connections with topology and ergodic theory*, pages 556–588. Springer, 1985.

[17] Dan-Virgil Voiculescu. Circular and semicircular systems and free product factors. In *Operator Algebras, Unitary Representations, Enveloping Algebras, and Invariant Theory*, volume 92 of *Progress in Mathematics*, pages 45–60. Birkhäuser, 1990.

[18] Dan-Virgil Voiculescu. Limit laws for random matrices and free products. *Inventiones mathematicae*, 104(1):201–220, 1991.

[19] Dan-Virgil Voiculescu. Free probability and the von neumann algebras of free groups. *Reports on Mathematical Physics*, 55(1):127–133, 2005.