# Non-Commutative Probability Theory

Paul D. Mitchener
e-mail: mitch@uni-math.gwdg.de

November 3, 2005

## Contents

# 1 Classical Probability Theory

The idea of this chapter is to put classical probability theory in a rigorous and abstract setting in such a way that later generalisations will be natural. We will assume that the reader has a fair knowledge of measure theory, as presented, for instance, in the first few chapters of [Rud87].

## 1.1 Probability Spaces

**Definition 1.1** Let $(\Omega, \mu)$ be a measure space. Then the measure $\mu$ is called a *probability measure* if $\mu(\Omega) = 1$. A measure space where the measure is a probability measure is termed a *probability space*.

We think of a measurable set $E \subseteq \Omega$ as a possible *event* and the measure $\mu(E) \in [0, 1]$ as the probability of that event occurring.

Given two events, $E_1$ and $E_2$, we think of the union $E_1 \cup E_2$ as either the event $E_1$ or the event $E_2$ occurring. The intersection $E_1 \cap E_1$ should be thought of as the event $E_1$ and the event $E_2$ both occurring. Viewed in this way, the axioms of measure theory translate into intuitive axioms for a probability space.

## 1.2 Random Variables

**Definition 1.2** Let $\omega$ be a probability space. Then a *complex random variable* if a measurable function $X : \Omega \to \mathbb{C}$. A *real random variable* is a measurable function $X : \Omega \to \mathbb{R}$.

A random variable can be thought of as a function assigning a value depending on which events occur. Thus, given a measurable subset $S \subset \mathbb{C}$, we can define the probability

$$P[X \in S] = \mu(X^{-1}[S])$$

for a complex random variable $X$. The following result is easy to see.

**Proposition 1.3** *Let $X$ be a complex random variable. Then we can define a probability measure on the set $\mathbb{C}$ by writing*

$$\tau_X(S) = P(X \in S)$$

*for each measurable set $S \subseteq \mathbb{C}$.* □

We can rewrite the above formula

$$P(X \in S) = \int_S d\tau_X$$

The corresponding result for real random variables is also true.

**Definition 1.4** The above measure $\tau_X$ is called the *probability law* (or just *law*) of a random variable $X$.

One of the fundamental problems in probability theory is to compute the probability laws of random variables.

**Definition 1.5** Let $X$ be a random variable on a probability space $(X, \mu)$. The *expectation* of $X$ is the integral

$$E(X) = \int_X X d\mu$$

The $k$-th *moment* of $X$ is the expectation

$$m_k(X) = E(X^k)$$

The expectation of a random variable can be thought of as its average value. The following two results follows immediately from properties of the integral.

**Proposition 1.6** *Let* 1 *denote the random variable with constant value* 1*. Then* $E(1) = 1$. □

**Proposition 1.7** *Let $X$ and $Y$ be random variables, and let $\alpha$ and $\beta$ be scalars. Then $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$.* □

The *variance* of a random variable $X$

$$\sigma(X)^2 = E((X - m_1)^2) = m_2(X)^2 - m_1(X)^2$$

can be thought of as a measure of how much and how frequently $X$ differs from its expectation.

The following result follows immediately from our various definitions.

**Proposition 1.8** *Let $X$ be a complex random variable, with law $\tau_X$. Then we have moments*

$$m_k(X) = \int_{\mathbb{C}} t^k d\tau_X(t)$$

□

The moments of a random variable determine its law in the sense of the following result.

**Theorem 1.9** *Two random variables with the same moments have the same probability laws.* □

We will not prove this theorem now since we will look at a similar but more general result later on.

### 1.3 Independence

**Definition 1.10** Let $(\Omega, \mu)$ be a probability space. Then we call two events $E_1, E_2 \subseteq \Omega$ *independent* if $\mu(E_1 \cap E_2) = \mu(E_1)\mu(E_2)$.

When constructing probability spaces, indpendence of two events means the probability of one event occurring is not affected by whether or not the other event takes place. We can make a similar definition for random variables.

**Proposition 1.11** *Let $X_1, X_2, \dots, X_n$ be complex random variables on a probability space $\Omega$. Then we can define a probability measure on the set $\mathbb{C}^n$ by writing*

$$\tau_{X_1,\dots,X_n}(S) = \mu(X_1^{-1}[S] \cap \cdots \cap X_n^{-1}[S])$$

*for each measurable set $S \subseteq \mathbb{C}^n$.* □

The corresponding result for real random variables is of course also true.

**Definition 1.12** The above measure $\tau_X$ is called the *joint law* of the random variable $X_1, \dots, X_n$. The random variables $X_1, \dots, X_n$ are termed *independent* if the laws satisfy the equation

$$\tau_{X_1,\dots,X_n} = \tau_{X_1}\tau_{X_2}\cdots\tau_{X_n}$$

The joint laws of independent random variables are thus determined from their individual laws, which makes them easy to calculate with. We will see some examples of such calculations later on.

## 2 Non-Commutative Probability Spaces

### 2.1 Algebras of Random Variables

The idea behind non-commutative geometry is that we can replace a geometric object by an algebra of functions on that object. This commutative algebra will have certain properties defined by the geometry. We then generalise by looking at non-commutative algebras with the same properties. The book [Con94] looks at this philosophy along with numerous constructions and examples.

This approach to non-commutative geometry also works for probability theory. Let $\Omega$ be a probability space. Then we can form an algebra, $A(\Omega)$, consisting of all complex random variables on $\Omega$. The expectation of a random variable defines a linear functional $E\colon A(\Omega) \to \mathbb{C}$ such that $E(1) = 1$.

The expectation defines the moments of a random variable, and so by theorem 1.9 enables the probability law of a variable, and so all information concerning its behaviour, to be recovered.

**Definition 2.1** A *non-commutative probability space* is a pair $(A, \phi)$, where $A$ is a unital complex algebra, and $\phi$ is a linear functional such that $\phi(1) = 1$. Elements of the algebra $A$ are called *random variables.*

The functional $\phi$ is termed a *trace* if $\phi(XY) = \phi(YX)$ for all random variables $X$ and $Y$.

We usually have slightly more structure to work with, as the following examples show.

**Example 2.2** Let $\Omega$ be a topological space equipped with a probability measure. Then we can form $C(\Omega)$, the commutative $C^*$-algebra of all continuous random variables. The convolution is defined by writing

$$X^\star(x) = \overline{X(x)} \qquad X \in C(\Omega),\ x \in \Omega$$

The expectation is a *state*, that is to say a continuous linear functional $E: C(\Omega) \to \mathbb{C}$ such that $E(1) = 1$ and $E(X^\star X) \geq 0$ for all random variables $X \in C(\Omega)$.

The $C^*$-algebra $C(\Omega)$ acts on the Hilbert space $L^2(\Omega)$ by left-multiplication, and so we have a faithful representation $C(\Omega) \to \mathcal{B}L^2(\Omega)$.

**Example 2.3** The *weak topology* on the space of operators $\mathcal{B}L^2(\Omega)$ is defined by saying that a sequence of operators $(T_n)$ has limit $T$ if and only if the sequence $(\langle X_n f, g\rangle)$ has limit $\langle Tf, g\rangle$ for all functions $f, g \in L^2(\Omega)$.

The $C^*$-algebra $C(\Omega)$ is not closed in the space $\mathcal{B}L^2(\Omega)$ under the weak topology; the closure is the von Neumann algebra $L^\infty(\Omega)$ consisting of all bounded measurable functions on $\Omega$.

The expectation $E: L^\infty(\Omega) \to \mathbb{C}$ is continuous with respect to the weak topology.

**Definition 2.4** Let $(A, \phi)$ be a non-commutative probability space. We call the pair $(A, \phi)$ a *$C^*$-probability space* if the algebra $A$ is a $C^*$-algebra and the functional $\phi$ is a state.

We call the pair $(A, \phi)$ a *$W^*$-probability space* if the algebra $A$ is a von Neumann algebra and the functional $\phi$ is a weakly continuous state.

Von Neumann's double commutant theorem (see for example [Dix81]) is a useful tool when looking at $W^\star$-probability spaces. Given an algebra $A \subseteq \mathcal{B}(H)$, we define the commutant

$$A' = \{T \in \mathcal{B}(H) \mid TX = XT \text{ for all } X \in A\}$$

The *double commutant theorem* then tells us that a $\star$-subalgebra $A \subseteq \mathcal{B}(H)$ is weakly closed (and so a von Neumann algebra) if and only if $A = A''$.

**Example 2.5** Let $\Omega$ be a proability space, and let $\{X_{ij} \mid 1 \leq i, j \leq n\}$ be random variables. Then we can form a *random matrix* $X = (X_{ij}): \Omega \to M_n(\mathbb{C})$.

The algebra of all continuous random matrices defines a $C^*$-algebra, $M_n(C(\Omega))$. Multiplication is of course matrix multiplication. The involution is defined by the formula

$$(X_{ij})^\star = (\overline{X_{ji}})$$

that is to say by taking the transpose of the complex conjugate. The functional $\phi$ is a trace defined by the formula

$$\phi(X) = E(\frac{1}{n}\,tr(X)) = E(\frac{1}{n}(X_{11} + X_{22} + \cdots + X_{nn}))$$

We thus obtain the $C^*$-probability space of continuous random matrices. We can similarly define the $W^*$-probability space of all bounded random matrices.

**Example 2.6** Let $G$ be a discrete group. Let $\mathbb{C}G$ be the complex group algebra, made up of formal sums

$$\sum_{i=1}^{n} \alpha_i g_i \qquad \alpha_i \in \mathbb{C}, g_i \in G$$

with multiplication law

$$\left(\sum_{i=1}^{m} \alpha_i g_i\right)\left(\sum_{j=1}^{n} \beta_j h_j\right) = \sum_{i,j=1}^{m,n} \alpha_i \beta_j g_i \circ h_j$$

and trace

$$\phi\left(\sum_{i=1}^{m} \alpha_i g_i\right) = \sum_{g_i = e} \alpha_i$$

Then the pair $(\mathbb{C}G, \phi)$ is a non-commutative probability space. The algebra $\mathbb{C}G$ can be equipped with an involution

$$\left(\sum_{i=1}^{n} \alpha_i g_i\right)^{\star} = \sum_{i=1}^{n} \overline{\alpha_i} g_i^{-1}$$

The *left-regular representation* of the algebra $\mathbb{C}G$, $\lambda\colon \mathbb{C}G \to \mathcal{B}(l^2G)$, is defined by the formula

$$(\lambda(g)\xi)(h) = \xi(g^{-1}h) \qquad \xi \in l^2G, g, h \in G$$

and extended by linearity.

The left-regular representation can be used to define a $C^*$-norm on the algebra $\mathbb{C}G$ by writing $\|x\| = \|\lambda(x)\|$. The trace on the algebra $\mathbb{C}G$ is then given by the formula

$$\phi(x) = \langle \lambda(x)e, e\rangle$$

We can complete the algebra $\mathbb{C}G$ to a $C^*$-algebra $C_r^*G$, called the *reduced $C^*$-algebra of $G$*. We have a trace, $\phi$, defined on the reduced $C^*$-algebra by the above formula, and so a $C^*$-probability space $(C_r^*G, \phi)$.

Completing this space with respect to the weak topology, we obtain the *group von Neumann algebra $NG$*, and a $W^*$-probability space $(NG, \phi)$. By the double commutant theorem, the von Neumann algebra $NG$ is equal to the double commutant $\lambda[G]'' \subseteq \mathcal{B}(l^2G)$.

## 2.2 Probability Laws

**Definition 2.7** Let $(A, \phi)$ be a non-commutative proability space, and let $X \in A$. Then the $k$-th *moment* of $X$ is the number

$$m_k(X) = \phi(X^n)$$

The *probability law* of $X$ is the linear functional $\tau_X \colon \mathbb{C}[X] \to \mathbb{C}$ defined by the formula

$$\tau_X(P) = \phi(P(X))$$

for each polynomial $P \in \mathbb{C}[X]$.

By linearity, the probability law of a random variable $X$ is determined by its moments.

**Theorem 2.8** *Let $(A, \phi)$ be a $C^*$-probability space, and let $X \in A$ be self-adjoint. Then there is a unique measure, $\mu_X$, on the real line such that*

$$\int_{\mathbb{R}} P(t) d\mu_X(t) = \phi(P(X))$$

*for all polynomials $P \in \mathbb{C}[X]$.*

**Proof:** We define the probability law $\tau_X(P) = \phi(P(X))$ for each polynomial $P$. By the Stone-Weierstrass theorem (see for example [Rud91]), we can extend the functional $\tau_X$ to a unique linear functional

$$\overline{\tau_X} \colon C(\mathbb{R}) \to \mathbb{C}$$

which is continuous under the suprememum norm whenever we restrict to a compact subset of $\mathbb{R}$. By the Riesz representation theorem (see for example [Rud87]), we can define a unique measure $\mu_X$ such that

$$\int_{\mathbb{R}} f \ d\mu_X = \overline{\tau}_X(f)$$

for all $f \in C(\mathbb{R})$, and we are done. □

The real case of theorem 1.9 follows as an immediate corollary. The complex case is then easily deduced.

**Corollary 2.9** *Two classical random variables with the same moments have the same probability laws.* □

Now, let us equip the space of complex continuous functions $C(\mathbb{C})$ with the topology defined by saying that a sequence of functions converges if it converges uniformly on compact subsets. The space of polynomials $\mathbb{C}[X]$ is topologised as a subspace of the space $C(\mathbb{C})$.

**Definition 2.10** Let $(X_n)$ be a sequence of (non-commutative) random variables, with sequence of probability laws $(\tau_n)$. Then we say that the sequence $(X_n)$ *converges in distribution* to a probability law $\tau \colon \mathbb{C}[X] \to \mathbb{C}$ if the sequence $\tau_n(P)$ converges to the functional $\tau(P)$ for each polynomial $P \in \mathbb{C}[X]$.

Less formally, we say that the sequence $(X_n)$ converges in distribution to a random variable $X$ if it converges in distribution to the probability law of $X$.

Note that it is not necessary for the each variable $X_n$ to lie in the same non-commutative probability space for the above definition to make sense.

The following result follows by linearity.

**Proposition 2.11** *Let $(X_n)$ be a sequence of random variables, with sequences of moments $(m_k^{(n)})$. Let $\tau\colon \mathbb{C}[X] \to \mathbb{C}$ be a probability law, with moments $m_k := \tau(X^k)$. Then the sequence $(X_n)$ converges in distribution to the probability law $\tau$ if and only if each sequence of moments $(m_k^{(n)})$ converges to the moment $m_k$.* □

**Proposition 2.12** *Let $(A, \phi)$ be a $C^*$-probability space. Let $(X_n)$ be a sequence in $A$, with norm limit $X$. Then the sequence $(X_n)$ converges in distribution to the random variable $X$.*

**Proof:** Let $P$ be a polynomial. Then the sequence $P(X^n)$ converges to the point $P(X)$. The result now follows by the definition of a probability law and norm-continuity of the functional $\phi$. □

The following result is similar.

**Proposition 2.13** *Let $(A, \phi)$ be a $W^*$-probability space. Let $(X_n)$ be a sequence in $A$, with weak limit $X$. Then the sequence $(X_n)$ converges in distribution to the random variable $X$.* □

## 2.3 Independence

**Definition 2.14** Let $(A, \phi)$ be a non-commutative probability space. Then a family of subalgebras $\{A_\lambda \mid \lambda \in \Lambda\}$ is called *independent* if:

- $[A_{\lambda_1}, A_{\lambda_2}] = 0$ if $\lambda_1 \neq \lambda_2$.
- $\phi(a_1 a_2 \cdots a_n) = \phi(a_1)\phi(a_2)\cdots\phi(a_n)$ whenever $a_k \in A_{\lambda_k}$ and $\lambda_i \neq \lambda_j$ when $i \neq j$.

The following two results follow immediately from the definition of independence.

**Proposition 2.15** *Let $(A, \phi)$ be a non-commutative probability space, where the algebra $A$ is generated by independeny subalgebras $A_1$ and $A_2$. Then the functional $\phi$ is determined by the restrictions $\phi|_{A_1}$ and $\phi|_{A_2}$.* □

**Proposition 2.16** *Let $(A, \phi)$ be a non-commutative probability space, and let $A_1$ and $A_2$ be independent subalgebras. Then*

$$\phi(a_1 a a_2) = \phi(a)\phi(a_1 a_2)$$

*whenever $a_1, a_2 \in A_1$ and $a \in A_2$.* □

We call a set of random variables $\{X_\lambda \mid \lambda \in \Lambda\}$ *independent* if the algebras generated by them form an independent family, that is to say:

- $[X_{\lambda_1}, X_{\lambda_2}] = 0$ if $\lambda_1 \neq \lambda_2$.
- $\phi(X_{\lambda_1} X_{\lambda_2} \cdots X_{\lambda_n}) = \phi(X_{\lambda_1})\phi(X_{\lambda_2}) \cdots \phi(X_{\lambda_n})$ whenever $\lambda_i \neq \lambda_j$ when $i \neq j$.

Similarly, we call a collection of sets independent if the algebras generated by them form an independent family.

**Definition 2.17** Let $(A, \phi)$ be a non-commutative probability space, and let $\{X_\lambda \mid \lambda \in \Lambda\}$ be a family of random variables

Then the *mixed moments* of the family are the various numbers

$$\phi(X_{\lambda_1} \cdots X_{\lambda_n}) \qquad \lambda_i \in \Lambda$$

Let $\mathbb{C}\langle Y_\lambda \rangle$ be the algebra of non-commutative polynomials in variables indexed by the set $\Lambda$. Then we define the *joint law* of the family $\{X_\lambda\}$ by the formula

$$\tau_{X_\lambda}(P) = \phi(P(X_\lambda))$$

whenever $P \in \mathbb{C}\langle Y_\lambda \rangle$ is a polynomial. The above notation implies we write the specific element $X_\lambda$ in place of the abstract variable $Y_\lambda$ in the polynomial $P$.

By linearity, the joint law of a family is determined by its mixed moments. It is clear that independence of a family of random variables depends only on its commutation relations and its joint law.

There is a notion of convergence in distribution of a family of random variables.

**Definition 2.18** Let $\{X_\lambda^{(n)} \mid \lambda \in \Lambda\}$ be a sequence of families of non-commutative random variables, with sequence of joint laws $(\tau_n)$. Then we say that the sequence $\{X_\lambda^{(n)}\}$ *converges in distribution* to a joint law $\tau : \mathbb{C}\langle Y_\lambda \rangle \to \mathbb{C}$ if the sequence $\tau_n(P)$ converges to the functional $\tau(P)$ for each polynomial $P \in \mathbb{C}\langle Y_\lambda \rangle$.

**Proposition 2.19** *Let $\{X_\lambda^{(n)} \mid \lambda \in \Lambda\}$ be a sequence of families of non-commutative random variables. Then the sequence $\{X_\lambda^{(n)}\}$* converges in distribution *to a joint law $\tau : \mathbb{C}\langle Y_\lambda \rangle \to \mathbb{C}$ if and only if for all elements $\lambda_i \in \Lambda$ the sequence of mixed moments $\phi(X_{\lambda_1}^{(n)} \cdots X_{\lambda_k}^{(n)})$ converges to the number $\tau(Y_{\lambda_1} \cdots Y_{\lambda_n})$.* □

# 3 Freeness

## 3.1 Free Products

In this subsection we look at a number of free product constructions without proofs. For further details, see for example chapter 1 of [VDN92].

**Definition 3.1** Let $\{G_\lambda \mid \lambda \in \Lambda\}$ be a family of groups. Then the *free product* $*_{\lambda \in \Lambda} G_\lambda$ is the unique (up to isomorphism) group $G$ equipped with homomorphisms $\psi_\lambda : G_\lambda \to G$ such that for any group $H$ equipped with homomorphisms $\varphi : G_\lambda \to H$ there is a unique homomorphism $\Phi : G \to H$ such that $\Phi \circ \psi_\lambda = \varphi_\lambda$ for all $\lambda \in \Lambda$.

Given a finite family of groups $G_1, \dots, G_n$, we denote the free product by writing $G_1 * G_2 * \cdots * G_n$.

**Definition 3.2** Let $G$ be a group. Then subgroups $G_1$ and $G_2$ are called *free* if there are no relations between them, that is to say, if $w = g_1 g_2 \cdots g_n$, where the elements $g_i$ belong alternately to the subgroups $G_1$ and $G_2$, and $g_i \neq e$ for all $i$, then $w \neq e$.[1]

**Proposition 3.3** *Let $G$ be a group generated by free subgroups $G_1$ and $G_2$. Then $G = G_1 * G_2$.* □

**Definition 3.4** Let $\{A_\lambda \mid \lambda \in \Lambda\}$ be a family of unital algebras. Then the *free product* $*_{\lambda\in\Lambda} A_\lambda$ is the unique (up to isomorphism) unital algebra $A$ equipped with homomorphisms $\psi_\lambda \colon A_\lambda \to A$ such that for any algebra $B$ equipped with homomorphisms $\varphi \colon A_\lambda \to B$ there is a unique homomorphism $\Phi \colon A \to B$ such that $\Phi \circ \psi_\lambda = \varphi_\lambda$ for all $\lambda \in \Lambda$.

The definition of a free product of $C^*$-algebras is exactly analagous to the above.

**Proposition 3.5** *Let $\{G_\lambda \mid \lambda \in \Lambda\}$ be a family of groups. Then looking at free products of algebras and $C^*$-algebras respectively*

$$*_{\lambda\in\Lambda} \mathbb{C} G_\lambda = \mathbb{C}\left(*_{\lambda\in\Lambda} G_\lambda\right) \qquad *_{\lambda\in\Lambda}\, C_r^* G_\lambda = C_r^*\left(*_{\lambda\in\Lambda} G_\lambda\right)$$

□

**Definition 3.6** Let $\{(H_\lambda, \xi_\lambda) \mid \lambda \in \Lambda\}$ be a family of Hilbert spaces, each equipped with a distinguished unit vector $\xi_\lambda \in H_\lambda$. Then the *free product* $*_{\lambda\in\Lambda}(H_\lambda, \xi_\lambda)$ is the Hilbert space

$$H = \mathbb{C}\xi \oplus \left(\oplus_{n\geq 1} \otimes_{\lambda_1 \neq \cdots \neq \lambda_n} \xi^0_{\lambda_1} \otimes \cdots \otimes \xi^0_{\lambda_n}\right)$$

where $\xi^0_{\lambda_i}$ is the orthogonal complement of the vector $\xi_{\lambda_i}$ in the Hilbert space $H_\lambda$.

We consider the Hilbert space $H$ to have the distinguished unit vector $\xi$. Each Hilbert space $H_\lambda$ is embeddded in $H$, and each unit vector $\xi_\lambda$ is identified with the vector $\xi$.

**Definition 3.7** Let $H$ be a Hilbert space. Then we define the *full Fock space*

$$T(H) = \mathbb{C}\xi \oplus \left(\oplus_{n\geq 1} H^{\otimes n}\right)$$

The norm of the vector $\xi$ is equal to one.

**Proposition 3.8** *Let $\{(H_\lambda, \xi_\lambda) \mid \lambda \in \Lambda\}$ be a family of Hilbert spaces with distinguished unit vectors. Then*

$$(T(\oplus_{\lambda\in\Lambda} H_\lambda), \xi) = *_{\lambda\in\Lambda}(T(H_\lambda), \xi)$$

□

[1] Here we use $e$ to denote the identity element of a group.

**Definition 3.9** Let $\{A_\lambda \mid \lambda \in \Lambda\}$ be a family of von Neumann algebras, where $A_\lambda \subseteq \mathcal{B}(H_\lambda)$, and the Hilbert space $H_\lambda$ is equipped with a distinguished unit vector $\xi_\lambda$.

Let us regard $A_\lambda$ as a subalgebra of the operator space $\mathcal{B}(*_{\lambda\in\Lambda}(H_\lambda.\xi_\lambda)$. Then we define the free product as a double commutant

$$*_{\lambda\in\Lambda} A_\lambda = \left(\bigcup_{\lambda\in\Lambda} A_\lambda\right)''$$

**Proposition 3.10** *Let $\{G_\lambda \mid \lambda \in \Lambda\}$ be a family of groups. Then we have group von Neumann algebra*

$$N(*_{\lambda\in\Lambda} G_\lambda) = *_{\lambda\in\Lambda} NG_\lambda$$

□

## 3.2 Free Algebras and Random Variables

*Freeness* is a fundamental concept in non-commutative probability. It is analogous to independence, but only exists as a concept in the non-commutative case. We begin with a result on free products in order to motivate the definition.

**Proposition 3.11** *Let $A_1, \ldots, A_n$ be complex unital algebras. Let $A = A_1 * \cdots * A_n$ be the free product, and suppose that this free product is equipped with some trace, $\phi$, such that $\phi(1) = 1$. Let us identify an algebra $A_i$ with its image in the free product. Then $\phi(a_1 a_2 \cdots a_k) = 0$ whenever:*

- *$a_j \in A_{i(j)}$, where $i(1) \neq i(2), i(2) \neq i(3), \ldots i(k-1) \neq i(k)$.*
- *$\phi(a_i) = 0$ for all $i$.*

□

**Definition 3.12** Let $(A, \phi)$ be a non-commutative probability space. Let $A_1, \ldots, A_n \subseteq A$ be unital subalgebras of $A$. Then the family $\{A_1, \ldots, A_n\}$ is called *free* if $\phi(a_1 a_2 \cdots a_k) = 0$ whenever:

- $a_j \in A_{i(j)}$, where $i(1) \neq i(2), i(2) \neq i(3), \ldots i(k-1) \neq i(k)$.
- $\phi(a_i) = 0$ for all $i$.

We will often talk about freeness of a set of random variables rather than algebras.

**Definition 3.13** Let $(A, \phi)$ be a non-commutative probability space. Random variables $X_1, \ldots, X_n \in A$ are termed free if the family of algebras generated by them is free.

We have a similar notion of a family of sets being free. There are many possible algebraic manipulations involving the definition of freeness.

**Proposition 3.14** *Let $X_1$ and $X_2$ be free random variables. Then*

$$\phi(X_1 X_2) = \phi(X_1)\phi(X_2)$$

**Proof:** By definition of freeness

$$\phi((X_1 - \phi(X_1))(X_2 - \phi(X_2))) = 0$$

The result now follows by linearity of the functional $\phi$. $\square$

**Proposition 3.15** *Let $(A, \phi)$ be a non-commutative probability space, and let $A_1$ and $A_2$ be free subalgebras. Then*

$$\phi(X_1 X X_2) = \phi(X)\phi(X_1 X_2)$$

*whenever $X_1, X_2 \in A_1$ and $X \in A_2$.*

**Proof:** By definition of freeness

$$\phi((X_1 - \phi(X_1))(X - \phi(X))(X_2 - \phi(X_2))) = 0$$

Expanding the above expression, we see that

$$\phi(X_1(X - \phi(X))(X_2 - \phi(X_2)) + \phi(X_1)\phi((X - \phi(X))(X_2 - \phi(X_2))) = 0$$

The second part of the above expression is again zero by freeness. Hence

$$\phi(X_1 X X_2) - \phi(X)\phi(X_1 X_2) = \phi(X_1)\phi(X_2)\phi(X - \phi(X)) = 0$$

and we are done. $\square$

**Proposition 3.16** *Let $(A, \phi)$ be a non-commutative probability space, where the algebra $A$ is generated by free subalgebras $A_1$ and $A_2$. Then the functional $\phi$ is determined by the restrictions $\phi|_{A_1}$ and $\phi|_{A_2}$.*

**Proof:** Let $a \in A$. Then we can write $a = a_1 a_2 \cdots a_n$, where the elements $a_i$ belong alternately to the algebras $A_1$ and $A_2$. Suppose that $n \geq 2$. Observe

$$\phi(a) = \phi((a_1 - \phi(a_1) + \phi(a_1)) \cdots (a_n - \phi(a_n) + \phi(a_n)))$$

so by linearity

$$\phi(a) = \phi((a_1 - \phi(a_1)) \cdots (a_n - \phi(a_n))) + \phi(b)$$

where $b = b_1 b_2 \cdots b_{n-1}$ and the elements $b_i$ belong alternately to the algebras $A_1$ and $A_2$. By definition of freeness, we see that

$$\phi((a_1 - \phi(a_1)) \cdots (a_n - \phi(a_n))) = 0$$

The result now follows by induction. $\square$

## 3.3 Random Walks on Groups

Let $G$ be a discrete group, with finite generating set $\{s_1, \ldots, s_n\}$. Let us consider a random walk on the group $G$, beginning at the point 1, and going in each of the directions represented by elements of the generating set and their inverses, with equal probability $1/2n$.

Each step of the random walk is the random variable

$$X = \frac{1}{2n}(s_1 + \cdots + s_n + s_1^{-1} + \cdots + s_n^{-1})$$

in the space $\mathbb{C}G$.

The position after $k$ steps is given by the random variable $X^k$. The probability at being at position $g \in G$ after $k$ steps is the coefficient of $g$ in the expression for $X^k$, that is to say

$$P(X^k = g) = \langle \lambda(X^k), g \rangle$$

where $\lambda$ is the left regular representation described in example 2.6.

In particular, $P(X^k = e) = \phi(X_k)$.

**Example 3.17** Consider a random walk on $\mathbb{Z}$. Let $s$ be a generator for $\mathbb{Z}$, and let $X$ be the random variable described above. Then

$$X = \frac{1}{2}(s + s^{-1})$$

and

$$X^k = \frac{1}{2^k} \sum_{l=0^k} \binom{k}{l} s^{2l-k}$$

We can thus compute the probability of returning to the starting position after $k$ steps

$$\phi(X^k) = \begin{cases} 0 & k \text{ odd} \\ k!/(2^k(k/2)!) & k \text{ even} \end{cases}$$

These probabilities are the moments of the random variable $X$, and so determine its probability law.

**Example 3.18** Consider a random walk on $F_2$, the free group with two generators, $s_1$ and $s_2$. Then the random variable $X$ is given by the formula

$$X = X_1 + X_2 \qquad X_1 = \frac{1}{4}(s_1 + s_1^{-1}), X_2 = \frac{1}{4}(s_2 + s_2^{-1})$$

Observe $\phi(X_1) = 0, \phi(X_2) = 0$, and

$$X_1X_2 = \frac{1}{16}(s_1s_2 + s_1s_2^{-1} + s_1^{-1}s_2 + s_1^{-1}s_2^{-1}), X_2X_1 = \frac{1}{16}(s_2s_1 + s_2s_1^{-1} + s_2^{-1}s_1 + s_2^{-1}s_1^{-1})$$

The group elements in the above expressions are distinct, so $\phi(X_1X_2) = \phi(X_2X_1) = 0$. It follows that the random variables $X_1$ and $X_2$ are free. Freeness actually gives us the tools to precisely compute the moments $\phi(X^k)$ and the associated probability law. We will return to this subject later on.

# 4 Some Combinatorics

## 4.1 The Catalan Numbers

The Catalan numbers are a sequence that occur naturally in many combinatorial problems. The following definition is probably the one that is easiest to use in combinatorial problems.

**Definition 4.1** The sequence of *Catalan numbers*, $(C_n)$, is defined recursively by writing $c_0 = 0$, $c_1 = 1$, and

$$C_n = \sum_{k=1}^{n} C_{k-1} C_{n-k}$$

when $n > 1$.

**Proposition 4.2** *The Catalan numbers can also be defined by the explicit formula*

$$C_n = \frac{(2n)!}{(n+1)!n!}$$

**Proof:** Let us define the generating function for the Catalan numbers

$$c(x) = \sum_{n=0}^{\infty} C_n x^n$$

We can rewrite our recurrence relation

$$C_n = \sum_{k=0}^{n-1} C_k C_{(n-1)-k}$$

so

$$c(x)^2 = \sum_{n=0}^{\infty} \sum_{k=0}^{n} C_k C_{n-k} x^n = \sum_{n=0}^{\infty} C_{n+1} x^n$$

and

$$c(x) = 1 + xc(x)^2$$

We can solve this equation to obtain the expression

$$c(x) = \frac{1 - \sqrt{1-4x}}{2x}$$

since we know that $c(0) = C_0 = 1$

Taking a binomial expansion of the square root, we obtain the formula we claimed. □

**Corollary 4.3** *The Catalan numbers can be defined recursively by writing*

$$C_0 = 1 \qquad C_n = \frac{2(2n-1)}{n+1} C_{n-1}$$

□

## 4.2 Partitions

**Definition 4.4** A *partition*, $\pi$, of a finite set $S$ is a collection of *blocks* $\{V_1, \ldots, V_s\}$ that partition the set $S$.

If the set $S$ is ordered, we call a partition *crossing* if we can find elements $p_1 \leq q_1 \leq p_2 \leq q_2$ such that the elements $p_1$ and $p_2$ belong to the same block, and the elements $q_1$ and $q_2$ belong to another distinct block.

A partition is called a *pairing* if each block contains exactly two elements.

Let us write $P(n)$ to denote the set of all partitions of the set $\{1, \ldots, n\}$, $NCP(n)$ to denote the set of all non-crossing partitions, $P_2(n)$ to denote the set of pairings, and $NCP_2(n)$ to denote the set of all non-crossing pairings.

**Proposition 4.5** *Let $\pi, \sigma \in NCP(n)$ be non-crossing partitions. Write $\sigma \leq \pi$ whenever each block of the partition $\sigma$ is contained in some block of the partition $\pi$.*

*Then the set $NCP(n)$ is a partially ordered set, with maximal element*

$$1_n = \{(1, \ldots, n)\}$$

*and minimal element*

$$0_n = \{(1), \ldots, (n)\}$$

□

It is quite easy to count the number of pairings.

**Proposition 4.6** *The number of elements in the set $P_2(n)$ is zero if $n$ is odd, and equal to the product*

$$(n-1)(n-3)\cdots 3.1$$

*if $n$ is even.*

**Proof:** Let $\pi = \{V_1, \ldots, V_s\} \in P_2(n)$, where $1 \in V_1$ and the pairs $V_i$ are written in the order of their lower elements. Then $s = n/2$.

Observe $V_1 = \{1, m\}$ so there are $n-1$ choices for the block $V_1$. Having chosen the block $V_1$, there are $n-3$ choices for the block $V_2$, $n-5$ choices for the block $V_3$, and so on. Multiplying these numbers, we arrive at the above formula. □

We can also count the number of non-crossing pairings.

**Proposition 4.7** *The number of elements in the set $NCP_2(n)$ is zero if $n$ is odd, and equal to the Catelan number $C_{n/2}$ if $n$ is even.*

**Proof:** The result when $n$ is odd is obvious. It is similarly obvious that $\natural NCP_2(0) = C_0 = 1$ and $\natural NCP_2(2) = C_1 = 1$.

Let $\pi = \{V_1, \ldots, V_s\} \in NCP_2(n)$ where $1 \in V_1$, so that $V_1 = \{1, m\}$ for some $m > 1$. Then, since the partition $\pi$ is non-crossing, for each block $V_j$, either $1 < k < m$ or $m < k$ for all $k \in V_j$.

It follows that the partition $\pi$ can be restricted to define non-crossing pairings on the sets $\{2,\ldots,m-1\}$ and $\{m+1,\ldots,n\}$. It follows that $m$ must be even. Counting the numbers of such partitions, we obtain the recurrence relation

$$\natural NCP_2(n) = \sum_{l=1}^{n/2} \natural NCP_2(m-2l)\natural NCP_2(n-2l)$$

If we set $n = 2k$, we see that the number $\natural NCP_2(2k)$ satisfies the recurrence relation for the Catalan numbers. The result now follows. □

## 4.3 Complements

The following construction works for non-crossing partitions, but has no analogy in the general case.

**Definition 4.8** Let $NCP(\overline{n})$ be the set of all non-crossing partitions on the ordered set $\{\overline{1},\ldots,\overline{n}\}$. Let $NCP(n,\overline{n})$ be the set of all non-crossing partitions on the ordered set $\{1,\overline{1},\ldots,n,\overline{n}\}$.

Let $\pi \in NCP(n)$ be a partition. Then we define the *complement*, $K(\pi) \in NCP(\overline{n})$, to be the maximal non-crossing partition $\sigma$ such that $\pi \cup \sigma \in NCP(n,\overline{n})$.

Of course, the set $NCP(\overline{n})$ can be cannonically identified with the set $NCP(n)$, so the complement operation can be viewed as a natural map $K\colon NCP(n) \to NCP(n)$.

**Proposition 4.9** *Let $\pi \in NCP_2(n)$ be a non-crossing pairing. Identify $\pi$ with the permutation in which the cycles are the blocks of the partition. Let $\gamma = (1,2\ldots,n)$ be the cyclic permutation. Then $K(\pi) = \gamma^{-1}\pi$.* □

## 4.4 Free Cumulants

**Definition 4.10** Let $(A,\phi)$ be a non-commutative probability space. Then we define the *free cumulants*, $k_n\colon A^n \to \mathbb{C}$, to be the multilinear functionals defined inductively by the *moment-cumulant formula*:

$$k_1(X) = \phi(X) \qquad \phi(X_1\cdots X_n) = \sum_{\pi\in NCP(n)} k_\pi[X_1,\ldots,X_n]$$

where

$$k_\pi[X_1,\ldots,X_n] = \prod_{i=1}^{r} k_{V(i)}[X_1,\ldots,X_n] \qquad \pi = \{V(1),\ldots,V(r)\}$$

and

$$k_V[X_1,\ldots,X_n] = k_s(X_{v(1)},\ldots X_{v(s)}) \qquad V = (v(1),\ldots,v(s))$$

As a special case, for a random variable $X$, we write

$$k_n^X = k_n(X,\ldots,X)$$

**Proposition 4.11** *The free cumulants are well-defined.*

**Proof:** We can write the moment-cumulant formula in the form

$$\phi(X_1 \cdots X_n) = k_n(X_1, \ldots, X_n) + \sum_{\substack{\pi \in NCP(n) \\ \pi \neq 1_n}} k_\pi[X_1, \ldots, X_n]$$

The result now follows by induction. □

We use the version of the moment-cumulant formula in the above proof to work out cumulants explicitly.

The following result indicates why cumulants provide a useful combinatorial way of keeping track of data when looking at free random variables. The proof is purely a matter of combinatorics using the definitions involved, and can be found, for instance, in [Spe94].

**Theorem 4.12** *Let $(A, \phi)$ be a non-commutative probability space, and let $A_1, \ldots, A_m \subseteq A$ be unital subalgebras. Then the family $\{A_1, \ldots, A_n\}$ is free if and only if the free cumulant $k_n(a_1, \ldots, a_n)$ is equal to zero whenever $n \geq 2$, $a_j \in A_{i(j)}$, and there exist $k$ and $l$ such that $i(k) \neq i(l)$.* □

We can generalise the moment-cumulant formula.

**Definition 4.13** Let $\pi \in NCP(n)$. Then we define

$$\phi_\pi[X_1, \ldots, X_n] = \prod_{i=1}^{r} \phi_{V(i)}[X_1, \ldots, X_n] \qquad \pi = \{V(1), \ldots, V(r)\}$$

where

$$\phi_V[X_1, \ldots, X_n] = \phi(X_{v(1)}, \ldots X_{v(s)}) \qquad V = (v(1), \ldots, v(s))$$

**Proposition 4.14** *Let $(A, \phi)$ be a non-commutative probability space. Let $X_1, \ldots, X_n$ be random variables, and let $\pi \in NCP(n)$ be a non-crossing partition. Then*

$$\phi_\pi[X_1, \ldots, X_n] = \sum_{\substack{\sigma \in NCP(n) \\ \sigma \leq \pi}} k_\sigma[X_1, \ldots, X_n]$$

□

# 5 Gaussian and Semicircular Laws

## 5.1 Gaussian Random Variables

Let $X$ be a self-adjoint random variable in a $C^*$-probability space. Then, as we have already seen, there is a unique measure $\mu_X$ such that the law of $X$ is defined by the formula

$$\tau_X(P) = \int_R P(t) d\mu_X(t)$$

Conversely, a measure can be used to define a probability law. Working in this way, we have analogues in the non-commutative world of the classical distributions in probability theory. The following definition is perhaps the most useful example to us.

**Definition 5.1** Let $(A, \phi)$ be a non-commutative probability space. Then a self-adjoint random variable $X \in A$ is called *normal* or *Gaussian*, with expectation $\mu$ and variance $\sigma^2$ if the probability law is defined by the formula

$$\tau_X(P) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} P(t) \exp\left(-\frac{1}{2\sigma^2}(t-\mu)^2\right) \, dt$$

**Proposition 5.2** *Let $X$ be a random variable with Gaussian probability law, with expectation $\mu$ and variance $\sigma^2$.*

*Then $m_1(X) = \mu$. If $\mu = 0$, then we have higher moments*

$$m_k(X) = \begin{cases} 0 & k \text{ odd} \\ \sigma^k (k-1)(k-3)\cdots 3.1 & k \text{ even} \end{cases}$$

**Proof:** The moments of $X$ are defined by the formula

$$m_k(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} t^n \exp\left(-\frac{1}{2\sigma^2}(t-\mu)^2\right) \, dt$$

and we know that

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2\sigma^2}(t-\mu)^2\right) \, dt = 1$$

If we substitute $s = t - \mu$, we see that

$$m_1(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} s \exp\left(-\frac{1}{2\sigma^2}s^2\right) \, dt + \frac{\mu}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2\sigma^2}s^2\right) \, dt = \mu$$

by symmetry of the first integral and the above observation.

Now, let $\mu = 0$. Then, by symmetry, $m_k(X) = 0$ when $k$ is odd. Let $k = 2n$, and write

$$\begin{array}{ll} u = t^{2n-1} & dv = t\exp(-t^2/2\sigma^2) \\ du = (2n-1)t^{2n-2} & v = -\sigma^2 \exp(-t^2/2\sigma^2) \end{array}$$

Using integration by parts

$$m_{2n} = \frac{2n-1}{2\pi} \int_{\mathbb{R}} t^{2n-2} \exp\left(-\frac{1}{2\sigma^2}t^2\right) \, dt$$

It follows that

$$m_{2n} = \sigma^{2n}(2n-1)(2n-3)\cdots 3.1$$

as claimed. □

The following result is easy to check.

**Proposition 5.3** *Let $X$ and $Y$ be independent Gaussian random variables, with expectations $\mu_X$ and $\mu_Y$, and variances $\sigma_X^2$ and $\sigma_Y^2$ respectively. Then $X+Y$ is a Gaussian random variable with expectation $\mu_X + \mu_Y$ and variance $\sigma_X^2 + \sigma_Y^2$.* □

## 5.2 Semicircular Random Variables

As we will see in the next section, the Gaussian law has a fundamental role in the study of independent random variables. The following probability law has a similarly central role when we look at free random variables.

**Definition 5.4** Let $(A, \phi)$ be a non-commutative probability space. Then a self-adjoint random variable $X \in A$ is called *semicircular*, with *centre* $a$, and *radius* $r$, if the probability law is defined by the formula

$$\tau_X(P) = \frac{2}{\pi r^2} \int_{a-r}^{a+r} P(t)\sqrt{r^2 - (t-a)^2}\; dt$$

**Proposition 5.5** *Let $X$ be a random variable with semicircular probability law, centre $a$ and radius $r$.*

*Then $m_1(X) = a$. If $a = 0$, we have higher moments*

$$m_k(X) = \begin{cases} 0 & k \text{ odd} \\ (r/2)^k C_{k/2} & k \text{ even} \end{cases}$$

**Proof:** The moments of $X$ are defined by the formula

$$m_k(X) = \frac{2}{\pi r^2} \int_{a-r}^{a+r} t^k \sqrt{r^2 - (t-a)^2}\; dt$$

and we know that

$$\frac{2}{\pi r^2} \int_{a-r}^{a+r} \sqrt{r^2 - (t-a)^2}\; dt = 1$$

If we substitute $s = t - a$, we see that

$$m_1(X) = \frac{2}{\pi r^2} \int_{-r}^{r} s\sqrt{r^2 - s^2}\; ds + \frac{2}{\pi r^2} \int_{-r}^{r} s\sqrt{r^2 - s^2}\; dt = a$$

by symmetry of the first integral and the above observation.

Now, let $a = 0$. Then, by symmetry, $m_k(X) = 0$ when $k$ is odd. Let $k = 2n$, and write

$$\begin{array}{ll} u = t^{2n-1} & dv = t\sqrt{r^2 - t^2} \\ du = (2n-1)t^{2n-2} & v = -\frac{1}{3}(r^2 - t^2)\sqrt{r^2 - t^2} \end{array}$$

Using integration by parts

$$m_{2n}(X) = \frac{2}{\pi r^2} \int_{-r}^{r} \frac{1}{3} r^2 (2n-1) t^{2n-2} \sqrt{r^2 - t^2}\; dt - \frac{2}{\pi r^2} \int_{-r}^{r} \frac{1}{3}(2n-1) t^{2n} \sqrt{r^2 - t^2}\; dt$$

so

$$m_{2n}(X) = \frac{1}{3}(2n-1)r^2 m_{2n-2}(X) - \frac{1}{3} m_{2n}(X)$$

Rearranging, we see

$$m_{2n}(X) = \frac{2n-1}{2(n+1)} r^2 m_{2n-2}(X)$$

We know that $m_0(X) = 1$. Hence the desired result follows by the characterisation of the Catalan numbers in corollary 4.3. □

## 5.3 Families of Random Variables

In this section we will generalise Gaussian and semicircular random variables by looking at two classes of joint laws for families of random variables. We begin by looking at Gaussian families.

**Definition 5.6** A family of random variables $\{X_\lambda \mid \lambda \in \Lambda\}$ is called a (centred) *Gaussian family*, with *covariance matrix* $C = (C_ij)$ if each variable $X_\lambda$ is Gaussian, and the *Wick formula*

$$\phi(X_{\lambda(1)} \dots X_{\lambda(n)}) = \sum_{\pi \in P_2(n)} \prod_{(p,q) \in \pi} C_{\lambda(p)\lambda(q)}$$

holds.

**Proposition 5.7** . *Let $S$ be a commutative family of Gaussian random variables with centre zero. Then $S$ is a Gaussian family with diagonal covariance matrix if and only if the set $S$ is independent.*

**Proof:** Let $S = \{X_\lambda \mid \lambda \in \Lambda\}$ be a family of Gaussian random variables, each with mean zero.

Then by the Wick formula given in the definition, the mixed moment

$$\phi(X_{\lambda(1)}, \dots, X_{\lambda(n)})$$

is equal to zero whenever $\lambda(i) \neq \lambda(j)$ for $i \neq j$. Since $\phi(X_{\lambda(i)}) = 0$ for all $i$ and the variables commute, it follows that the set $S$ is independent.

Conversely, suppose that the set $S$ is independent. Since the random variables $X_\lambda$ commute, it suffices to look at moments of the form

$$\phi(X_{\lambda(1)}^{r(1)} \cdots X_{\lambda(n)}^{r(n)})$$

where $\lambda(i) \neq \lambda(j)$ whenever $i \neq j$. Independence tells us that

$$\phi(X_{\lambda(1)}^{r(1)} \cdots X_{\lambda(n)}^{r(n)}) = \phi(X_{\lambda(1)}^{r(1)}) \cdots \phi(X_{\lambda(n)}^{r(n)})$$

This value is equal to zero unless the numbers $r(i)$ are all even. Write $C_{\lambda(i)\lambda(j)} = \phi(X_{\lambda(i)} X_{\lambda(j)})$ so that $C_{\lambda(i)\lambda(j)} = 0$ if $i \neq j$. By proposition 4.6 and proposition 5.2, the $n$-th even moment of a Gaussian random variable is equal to the number of pairs in the set $P_2(n)$ multiplied by the 2nd moment. Hence we have the formula

$$\phi(X_{\lambda(1)}^{r(i)} = \sum_{\pi \in P_2(r(i))} C_{\lambda(i)\lambda(i)}$$

and so

$$\phi(X_{\lambda(1)}^{r(1)} \cdots X_{\lambda(n)}^{r(n)}) = \prod_{i=1}^{n} \sum_{\pi \in P_2(r(i))} C_{\lambda(i)\lambda(i)}$$

Write

$$X_{\lambda(1)}^{r(1)} \cdots X_{\lambda(n)}^{r(n)} = X_{\lambda'(1)} \cdots X_{\lambda'(k)}$$

Then, by counting possible pairings, we can write the above formula in the form

$$\phi(X_{\lambda'(1)} \cdots X_{\lambda'(k)}) = \prod_{i=1}^{n} \sum_{\pi \in P_2(k)} C_{\lambda'(i)\lambda(j)}$$

and we are done. □

**Definition 5.8** A family of random variables $\{X_\lambda \mid \lambda \in \Lambda\}$ is called a (centred) *semicircular family*, with *covariance matrix* $C = (C_{ij})$ if each variable $X_\lambda$ is semicircular, and the *Wick formula*

$$\phi(X_{\lambda(1)}, \ldots, X_{\lambda(n)}) = \sum_{\pi \in NCP_2(n)} \prod_{(p,q)\in\pi} C_{\lambda(p)\lambda(q)}$$

holds.

The following result can be proved similarly to proposition 5.7

**Proposition 5.9** *Let $S$ be a family of random variables. Then $S$ is a semicircular family with diagonal covariance matrix if and only if each random $X \in S$ is semicircular with mean zero, and the set $S$ is free.* □

## 5.4 Operators on Fock Space

We will now define an example of an operator with the semicircular probability law. Let $H$ be a Hilbert space. Recall that we define the *full Fock space*

$$T(H) = \mathbb{C}\xi \oplus \left(\oplus_{n\geq 1} H^{\otimes n}\right)$$

The vector $\xi$ is called the *vacuum vector*. We set its norm to be equal to one.

**Definition 5.10** Let $h \in H$. Then we define the *left-creation operator* $l(h)\colon T(H) \to T(H)$ by the formula

$$l(h)(k) = \begin{cases} h & k = \xi \\ h \otimes k & k \in H^{\otimes n},\ n \geq 1 \end{cases}$$

The adjoint, $l(h)^*$, is called the *left-annihilation operator* and is given by the formulae

$$l(h)^*(k_1 \otimes \cdots \otimes k_n) = \langle h, k_1\rangle k_2 \otimes \cdots \otimes k_n \qquad k_i \in H$$

and

$$l(h)^*(\lambda\xi) = 0 \qquad \lambda \in \mathbb{C}$$

**Definition 5.11** Let $H$ be a Hilbert space. Then we define $\Phi(H)$ to be the von Neumann algebra generated by the set of left-creation operators. We define the *vacuum state* on $\Phi(H)$ by the formula

$$v(T) = \langle \xi, T\xi\rangle$$

for any operator $T$.

The algebra $\Phi(H)$ equipped with the state $v$ is a $W^*$-probability space.

**Proposition 5.12** *Let $h \in H$. Then the operator*

$$s(h) = \frac{1}{2}(l(h) + l(h)^*)$$

*has a semicircular probability law, with centre 0 and radius $2\|h\|$.*

**Proof:** We will calculate the moments of the operator $s(h)$. Let us write

$$X_1 = l(h) \qquad X_{-1} = l(h)^*$$

Thenm by linearity of the vacuum state

$$v(s(h)^n = \sum_{k(i)=\pm 1} v(X_{k(n)} \cdots X_{k(1)})$$

By definition of the vacuum state, the number $v(X_{k(n)} \cdots X_{k(1)})$ is zero unless the following conditions all hold:

- Equal numbers of the indices $k(i)$ are equal to 1 and $-1$.
- $k(1) = 1$.
- $\sum_{i=1}^{k} k(i) \geq 0$ whenever $k \leq n$.

If the above conditions all hold, then $v(X_{k(n)} \cdots X_{k(1)}) = \|h\|^n$. The question is thus how many elements there are in the set $S$, consisting of all indices $(k(1), \ldots, k(n))$ there are which satisfy the above conditions.

If $n$ is odd, then the first of the above conditions implies that the set $S$ is empty. Let $n$ be even. We can define a map $\alpha\colon S \to NCP_2(n)$ by writing $\alpha(S) = \{V_1, \ldots V_s\}$ where:

- $V_1$ is the block containing 1 and the lowest number $r$ such that

$$\sum_{i=1}^{r} k(i) = 0$$

- For $p \geq 1$, $V_p = \{s, t\}$ where $s$ is the $p$-th number such that $k(s) = 1$ and $t$ is the lowest number such that

$$\sum_{i=s}^{t} k(i) = 0$$

The map $\alpha$ is clearly a bijection. Hence, by proposition 4.7 the number of elements in the set $S$ is equal to the Catalan number $C_{n/2}$. The result now follows by proposition 5.5. □

There are approaches to free probability theory (see for example [Haa97]) where the operators $s(h)$ are fundamental constructions. Here, we take a different, combinatorial approach, and regard these operators simply as major examples.

A proof of the following result can be found in [VDN92].

**Proposition 5.13** *Let* $\{V_\lambda \mid \lambda \in \Lambda\}$ *be an orthogonal family of subspaces of a Hilbert space* $H$*. Then the family of algebras*

$$\{\Phi(V_\lambda) \mid \lambda \in \Lambda\}$$

*is free in the space* $\Phi(H)$*.* □

**Corollary 5.14** *Let* $\{h_\lambda| \lambda \in \Lambda\}$ *be a family of orthogonal vectors. then the family of random variables* $\{s(h_\lambda) \mid \lambda \in \Lambda\}$ *is a free semicircular family.*

**Proof:** The result follows immediately from the above two propositions and proposition 5.9. □

# 6 The Central Limit Theorem

## 6.1 The Classical Central Limit Theorem

Let $(A, \phi)$ be a non-commutatitve probability space. Let $(X_n)$ be a sequence of either independent or free random variables, all with the same law, such that $\phi(X_n) = 0$ for all $n$. The *central limit theorem* tells us the sequence of random variables

$$\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}}\right)$$

converges in distribution to something definite. To prove this theorem, and say exactly what the limit is, we will look at moments.

There are in fact two central limit theorems; one for independent random variables, and one for free random variables. In this subsection, we will look at the part of the proof of the central limit theorem that is the same in both the independent and free cases, and go on to prove our result in the independent case.

The version of the central limit theorem for free random variables first appeared in [Voi85], and was one of the first major results in free probability. The combinatorial approach to the proof that we take here comes from [Spe98].

**Definition 6.1** Let $\pi$ be a partition of the set $\{1, \dots, n\}$. Then we write $\pi \cong (r(1), \dots, r(n))$ whenever $i$ and $j$ belong to the same block of $\pi$ precisely when $r(i) = r(j)$.

**Lemma 6.2** *Let* $(A, \phi)$ *be a non-commutative probability space. Let* $(X_n)$ *be a sequence of either independent or free random variables, all with the same law, such that* $\phi(X_n) = 0$ *for all* $n$*. Let* $\pi$ *be a partition of the set* $\{1, \dots, n\}$*, and let* $\pi \cong (r(1), \dots, r(n))$*. Then the* mixed moment

$$m_\pi = \phi(X_{r(1)} X_{r(2)} \cdots X_{r(n)})$$

*depends only on the partition* $\pi$*.*

**Proof:** By proposition 2.16 (in the independent case) or 3.16 (in the free case), the expression $\phi(X_{r(1)} \cdots, X_{r(n)})$ can be calculated from the moments of the individual random variables. Since the variables all have the same distribution, it follows that

$$\phi(X_{r(1)} \cdots X_{r(n)}) = \phi(X_{p(1)} \cdots X_{p(n)})$$

whenever

$$r(i) = r(j) \Leftrightarrow p(i) = p(j) \qquad \text{for all } i, j$$

The result now follows. □

**Corollary 6.3** *Let $(X_n)$ be a sequence of random variables as above. Let $A_\pi^N$ be the number of $n$-tuples $(r(1), \ldots, r(n))$ such that $\pi \cong (r(1), \ldots, r(n))$. Then*

$$\phi((X_1 + \cdots + X_N)^n) = \sum_{\pi \in P(n)} m_\pi A_\pi^N$$

**Proof:** By linearity

$$\phi((X_1 + \cdots + X_N)^n) = \sum_{r(1),\ldots,r(n)=1}^{n} \phi(X_{r(1)} \cdots X_{r(n)})$$

The result now follows from the above lemma. □

**Lemma 6.4** *Suppose $\pi = \{V_1, \ldots, V_s\}$ and $V_i = \{r\}$ for some block $V_i$. Then the number $m_\pi$ defined in lemma 6.2 is equal to zero.*

**Proof:** By proposition 2.16 (in the independent case) or proposition 3.15 (in the free case) we can write

$$m_\pi = \phi(X_{r(1)} \cdots X_r \cdots X_{r(n)}) = \phi(X_r)\phi(X_{r(1)} \cdots \tilde{X}_r \cdots X_{r(n)})$$

But $\phi(X_r) = 0$ be definition of the sequence $(X_n)$. □

**Corollary 6.5** *Suppose that $m_\pi \neq 0$ where $\pi = \{V_1, \ldots, V_s\}$. Then $\natural V_i \geq 2$ for all $i$ and $\natural\pi = s \leq n/2$.* □

**Lemma 6.6** *Let $(X_n)$ be a sequence of random variables as above. Then*

$$\lim_{N\to\infty} \phi(\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}}\right)^n) = \sum_{\pi \in P_2(n)} m_\pi$$

**Proof:** By corollary 6.3, we know that

$$\lim_{N\to\infty} \phi(\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}}\right)^n) = \sum_{\pi \in P(n)} \frac{A_\pi^N}{N^{n/2}} m_\pi$$

Observe

$$A_\pi^N = N(N-1)\cdots(N - \natural\pi + 1) = \frac{N!}{(N-\natural\pi)!}$$

so

$$\lim_{N\to\infty} \frac{A_\pi^N}{N^{n/2}} = \lim_{N\to\infty} N^{\natural\pi - n/2} = \begin{cases} 0 & \natural\pi < n/2 \\ 1 & \natural pi = n/2 \end{cases}$$

and

$$\lim_{N\to\infty} \phi(\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}}\right)^n) = \sum_{\pi\in P_2(n)} m_\pi$$

as claimed. □

**Corollary 6.7**

$$\lim_{N\to\infty} \phi(\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}}\right)^n) = 0$$

*if $n$ is odd.*

**Proof:** There are no pairings of the set $\{1, \ldots, n\}$ if $n$ is odd. □

We can now bring our calculations together for independent random variables.

**Theorem 6.8 (Classical Central Limit Theorem)** *Let $(A, \phi)$ be a non-commutative probability space, and let $(X_n)$ be a sequence of independent random variables such that $\phi(X_n) = 0$ for all $n$. Write $\sigma^2 = \phi(X_n^2)$. Then the sequence*

$$\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}}\right)$$

*converges in distribution to a Gaussian law with mean 0 and variance $\sigma^2$.*

**Proof:** Let $\pi \in P_2(n)$. Then, by definition of independence we see that

$$m_\pi = \phi(X_{r(1)} \cdots X_{r(n)}) = \phi(X_1^2)^{n/2} = \sigma^n$$

since the variables are identically distributed.

Hence, by lemma 6.6

$$\lim_{N\to\infty} \phi(\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}}\right)^n) = \sum_{\pi\in P_2(n)} m_\pi = \natural P_2(n)\sigma^n$$

But by proposition 4.6 (for $n$ even)

$$\natural P_2(n) = (n-1)(n-3)\cdots 3.1$$

so

$$\lim_{N\to\infty} \phi(\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}}\right)^n) = \sum_{\pi\in P_2(n)} m_\pi = \begin{cases} 0 & n \text{ odd} \\ \sigma^n (n-1)(n-3)\cdots 3.1 & n \text{ even} \end{cases}$$

But, by proposition 5.2, the above limits are the moments of a Gaussian law with mean $\mu$ and variance $\sigma^2$. The result therefore follows by proposition 2.11. □

## 6.2 The Free Central Limit Theorem

The difference between the independent and the free case of the central limit theorem is the result of the following lemma.

**Lemma 6.9** *Let $(X_i)$ be a free sequence of random variables, identically distributed, such that $\phi(X_i) = 0$ for all $i$. Write $(r/2)^2 = \phi(X_i^2)$. Let $\pi \in P_2(n)$. Then*

$$m_\pi = \begin{cases} (r/2)^n & \pi \in NCP_2(n) \\ 0 & \text{otherwise} \end{cases}$$

**Proof:** Let $\pi \cong (r(1), \ldots, r(n))$. Then there are two possibilities.

- $r(1) \neq r(2) \neq \cdots \neq r(n)$:
  Since $\phi(X_i) = 0$ for all $i$, it follows that

  $$m_\pi = \phi(X_{r(1)} \cdots X_{r(n)})$$

  by definition of freeness.

- There exists $m$ such that $r(m) = r(m+1)$:
  Then the variable $X_{r(m)} X_{r(m+1)} = X_{r(m)}^2$ is free from the set

  $$\{X_{r(1)} \cdots X_{r(m-1)}, X_{r(m+2)} \cdots X_{r(n)}\}$$

  is free since $\pi$ is a pairing. Hence by proposition 3.15

  $$\begin{aligned} m_\pi &= \phi(X_{r(1)} \cdots X_{r(m-1)}, X_{r(m+2)} \cdots X_{r(n)}) \phi(X_{r(m)}^2) \\ &= \phi(X_{r(1)} \cdots X_{r(m-1)}, X_{r(m+2)} \cdots X_{r(n)}) \sigma^2 \end{aligned}$$

The result now follows by induction. □

**Theorem 6.10 (Free Central Limit Theorem)** *Let $(A, \phi)$ be a non-commutative probability space, and let $(X_n)$ be a sequence of free random variables such that $\phi(X_n) = 0$ for all $n$. Write $(r/2)^2 = \phi(X_n^2)$. Then the sequence*

$$\left( \frac{X_1 + \cdots + X_N}{\sqrt{N}} \right)$$

*converges in distribution to a semicircular law with centre* 0 *and radius $r$.*

**Proof:** By lemma 6.6 and lemma 6.9 we see that

$$\lim_{N \to \infty} \phi( \left( \frac{X_1 + \cdots + X_N}{\sqrt{N}} \right)^n ) = \natural NCP_2(n) (r/2)^n$$

But by proposition 4.7 we know that $\natural NCP_2(n) = C_{n/2}$ when $n$ is even. This number is zero when $n$ is odd. Thus, by proposition 5.5 the above limits are the moments of a semicircular law with centre 0 and radius $r$. The result therefore follows by proposition 2.11. □

## 6.3 The Multi-dimensional Case

An elaboration of the calculations in the previous two sections lets us formulate versions of the classical and free central limit theorems for families.

**Theorem 6.11** *Let $(A, \phi)$ be a non-commutative probability space, and let $\{X_\lambda^{(n)} \mid \lambda \in \Lambda\}$ be a sequence of families of random variables such that each family is independent from the other families, $\phi(X_\lambda^{(n)}) = 0$ for all $\lambda$ and all $n$, and there are constants $C_{ij}$ such that $\phi(X_{\lambda(p)}^{(n)} X_{\lambda(q)}^{(n)}) = C_{\lambda(p)\lambda(q)}$ for all $\lambda(p)$, $\lambda(q)$, and $n$.*

*Then the sequence of families of random variables of the form*

$$\left( \frac{X_\lambda^{(1)} + \cdots + X_\lambda^{(N)}}{\sqrt{N}} \right)$$

*converges in distribution as $N \to \infty$ to a Gaussian family with covariance matrix $(C_{ij})$.* □

As in the previous section, the free version of the above result involves replacing pairings by non-crossing pairings.

**Theorem 6.12** *Let $(A, \phi)$ be a non-commutative probability space, and let $\{X_\lambda^{(n)} \mid \lambda \in \Lambda\}$ be a sequence of families of random variables such that each family is free from the other families, $\phi(X_\lambda^{(n)}) = 0$ for all $\lambda$ and all $n$, and there are constants $C_{ij}$ such that $\phi(X_{\lambda(p)}^{(n)} X_{\lambda(q)}^{(n)}) = C_{\lambda(p)\lambda(q)}$ for all $\lambda(p)$, $\lambda(q)$, and $n$.*

*Then the sequence of families of random variables of the form*

$$\left( \frac{X_\lambda^{(1)} + \cdots + X_\lambda^{(N)}}{\sqrt{N}} \right)$$

*converges in distribution as $N \to \infty$ to a semicircular family with covariance matrix $(C_{ij})$.* □

# 7 Limits of Gaussian Random Matrices

## 7.1 Gaussian Matrices

**Definition 7.1** Let $(\Omega, \mu)$ be a (classical) probability space. Then we can form an algebra of functions

$$L = \bigcap_{1 \leq p < \infty} L^p(\Omega)$$

and define a functional $E\colon L \to \mathbb{C}$ by integration with respect to the proabability measure $\mu$. Let $tr\colon M_n(\mathbb{C}) \to \mathbb{C}$ be the trace of a matrix. Then we can define the $W^*$-probability space of $n \times n$ *random matrices*

$$(M_n, \phi_n) = (L \otimes M_n(\mathbb{C}), E \otimes \frac{1}{n}\, tr)$$

In these notes we will restrict our attention to square random matrices, although of course others are possible. There are certain important classes of random matrices. We begin with the most general.

**Definition 7.2** Let $X = (X_{ij}) \in M_n$ be an $n \times n$ real random matrix. Then we call $X$ a *real Gaussian matrix* if each random variable $X_{ij}$ is independent and Gaussian, with expectation zero and identical variance. We call $X$ *standard* if we have variance $\sigma(X_{ij})^2 = 1/n$ for all $i$ and $j$.

A complex random matric $Z$ is called a *complex Gaussian matrix* if it takes the form $Z = X + iY$ where $X$ and $Y$ are independent real Gaussian matrices where each entry has the same variance. The matrix $Z$ if the absolute value of each entry, $|Z_{ij}|$, has variance $1/n$.

A random matrix of the form $X^*X$, where $X$ is a Gaussian matrix, is called a *Wishart matrix*.

**Proposition 7.3** *Every complex Gaussian matrix is an invertible element of the algebra $M_n$.*

**Proof:** An element of the algebra $M_n$ is an equivalence class of measurable functions $\Omega \to M_n(\mathbb{C})$, where we identify functions that agree on a set of measure zero.

Let $N$ be the set of non-invertible matrices, and let $X$ be a complex Gaussian matrix. Then by definition of the Gaussian law, the inverse image $X^{-1}[N]$ has measure zero. Hence, outside of a set of measure zero, the function $X\colon \Omega \to M_n(\mathbb{C})$ agrees with a function where every element of the image is invertible. $\square$

**Definition 7.4** Let $X = (X_{ij}) \in M_n$ be a real $n \times n$ random matrix. Then we call $X$ a *real self-adjoint Gaussian matrix* if the following conditions hold:

- $X_{ij} = X_{ji}$ for all $i, j$.
- Each random variable $X_{ij}$, where $i \leq j$, is independent and Gaussian, with expectation zero and identical variance.

A real random matrix $Y = (Y_{ij}) \in M_n$ is a *real skew-adjoint Gaussian matrix* if the following conditions hold:

- $X_{ij} = X_{ji}$ for all $i, j$.
- Each random variable $X_{ij}$, where $i \leq j$, is independent and Gaussian, with expectation zero and identical variance.

A complex random matrix $Z$ is called a *complex self-adjoint Gaussian random matrix* if it takes the form $Z = X + iY$, where $X$ is a real self-adjoint Gaussian matrix and $Y$ is a real skew-adjoint Gaussian matrix where each entry has the same variance.

**Definition 7.5** A real self-adjoint or skew-adjoint Gaussian matrix $X$ is called *standard* if we have variances $\sigma(X_{ij})^2 = 1/n$ for all $i$ and $j$.

A complex self-adjoint Gaussian matrix $Z$ is called *standard* if the absolute value of each entry, $|Z_{ij}|$, has variance $1/n$.

## 7.2 The Semicircular Law

We now want to look at limits of standard self-adjoint Gaussian matrices as the size increases, initially for single matrices, and then for independent families. We will see that such matrices converge in distribution to semicircular random variables, and independent families converge in distribution to free semicircular families. These results first appeared in [Voi91]. The combinatorial papproach that we take comes from [Spe93].

**Lemma 7.6** *Let $X \in M_n$ be a standard self-adjoint Gaussian matrix. Given a permutation $\sigma \in \Sigma_m$, let us write $Z(\sigma)$ to denote the number of cycles in $\sigma$. Then we have the formula*

$$\phi(X^m) = \sum_{\pi \in P_2(m)} n^{Z(\gamma\pi)-1-m/2}$$

*where we identify a pairing $\pi \in P_2(m)$ with a permutation as in proposition 4.9.*

**Proof:** By definition of the functional $\phi_n$, it is clear that

$$\phi(X^m) = \frac{1}{n} \sum_{i(1),\ldots,i(m)=1}^{n} E(X_{i(1)i(2)} X_{i(2)i(3)} \cdots X_{i(m)i(1)})$$

For convenience, let us count modulo $m$, that is to say set $i(m+1) = i(1)$. Then by proposition 5.7, we see that

$$\phi(X^m) = \frac{1}{n} \sum_{i(1),\ldots,i(m)=1}^{n} \sum_{\pi \in P_2(m)} \prod_{(r,s)\in\pi} E(X_{i(r)i(r+1)} X_{i(s)} X_{i(s+1)})$$

Since the matrix $X$ is a standard self-adjoint Gaussian matrix, we can write, using Kronecker delta notation:[2]

$$E(X_{i(r)i(r+1)} X_{i(s)} X_{i(s+1)}) = \frac{1}{n} \delta_{i(r)i(s+1)} \delta_{i(s)i(r+1)}$$

Thus

$$\phi(X^m) = \frac{1}{n^{1+m/2}} \sum_{\pi \in P_2(m)} \sum_{i(1),\ldots,i(m)=1}^{n} \prod_{(r,s)\in\pi} \delta_{i(r)i(s+1)} \delta_{i(s)i(r+1)}$$

As mentioned in the statement of the result, a pairing $\pi \in P_2(m)$ can be identified with a permutation $\pi \in \Sigma_m$ by defining the cycles of the permutation to be the blocks of the pairing. Thus we can rewrite the above equation

$$\phi(X^m) = \frac{1}{n^{1+m/2}} \sum_{\pi \in P_2(m)} \sum_{i(1),\ldots,i(m)=1}^{n} \prod_{r=1}^{m} \delta_{i(r)i(\pi(r)+1)}$$

[2] That is to say

$$\delta_{ab} = \begin{cases} 1 & a = b \\ 0 & a \neq b \end{cases}$$

or

$$\phi(X^m) = \frac{1}{n^{1+m/2}} \sum_{\pi \in P_2(m)} \sum_{i(1),\ldots,i(m)=1}^{n} \prod_{r=1}^{m} \delta_{i(r)i(\gamma\pi(r))}$$

where

$$\gamma = (1, 2, \ldots, m-1, m) \in \Sigma_m$$

The index function $i$ must be constant on the cycles of the permutation $\gamma\pi$ in order for the product $\prod_{r=1}^{m} \delta_{i(r)i(\gamma\pi(r))}$ to be one; otherwise, the product is zero.

It follows that we have $n^{Z(\gamma\pi)}$ possible index functions which make the above product zero. Hence

$$\sum_{i(1),\ldots,i(m)=1}^{n} \prod_{r=1}^{m} \delta_{i(r)i(\gamma\pi(r))} = n^{Z(\gamma\pi)}$$

and

$$\phi_n(X^m) = \sum_{\pi \in P_2(m)} n^{Z(\gamma\pi)-1-m/2}$$

as claimed. □

**Corollary 7.7** *We have moments $\phi(X^m) = 0$ when $m$ is odd.* □

**Theorem 7.8** *Let $(X_n)$ be a sequence of standard self-adjoint Gaussian matrices, where $X_n \in M_n$. Then the sequence $(X_n)$ converges in distribution to a semi-circular random variable with centre* 0 *and radius* 2.

**Proof:** When $m$ is odd, as we have already remarked, $\phi_n(X_n^m) = 0$ for all $n$. Let $m = 2k$.

Let $\pi \in P_2(2k)$. Then the number of cycles, $Z(\gamma\pi)$, can be at most $1 + k$, and this number is achieved precisely when the pairing is non-crossing. By the above lemma, we have the formula

$$\phi(X_n^{2k})) = \sum_{\pi \in P_2(m)} n^{Z(\gamma\pi)-1-k}$$

Observe

$$\lim_{n\to\infty} n^{Z(\gamma\pi)-1-k} = \begin{cases} 1 & Z(\gamma\pi) = k+1 \\ 0 & Z(\gamma\pi) < k+1 \end{cases}$$

Thus

$$\lim_{n\to\infty} \phi_n(X_n^{2k})) = \natural NCP_2(k)$$

By proposition 4.7, $\natural NCP_2(k)$ is the $k$-th Catalan number. Hence, by proposition 5.5, the sequence $(X_n)$ converges in distribution to a semicircular random variable, as claimed. □

### 7.3 Asymptotic Freeness

We can quite easily extend the main theorem of the previous section to a result on families of Gaussian matrices. As a first step, we have the following result. We omit the proof since the computations needed are almost identical to those in the proof of lemma 7.6.

**Lemma 7.9** *Let $\{X^\lambda \mid \lambda \in \Lambda\}$ be a family of standard self-adjoint $n \times n$ Gaussian matrices. Then*

$$\phi(X^{\lambda(1)} \cdots X^{\lambda(n)}) = \sum_{\pi \in P_2(m)} \prod_{r=1}^{m} \delta_{\lambda(r)\lambda(\pi(r))} n^{Z(\gamma\pi)-1-m/2}$$

□

**Theorem 7.10** *Let $\{X_n^\lambda \mid \lambda \in \Lambda\}$ be a sequence of independent families of standard self-adjoint $n \times n$ Gaussian matrices. Then the sequence of families $(\{X_n^\lambda\})$ converges in distribution to a free semi-circular family.*

**Proof:** When $m$ is odd, the set set $P_2(m)$ is empty, so by the above formula the mixed moments $\phi(X^{\lambda(1)} \cdots X^\lambda(n))$ are all zero. Let $m = 2k$.

Let $\pi \in P_2(2k)$. Then the number of cycles, $Z(\gamma\pi)$, can be at most $1 + k$, and this number is achieved precisely when the pairing is non-crossing. By the above lemma, we have the formula

$$\phi(X^{\lambda(1)} \cdots X^\lambda(n)) = \sum_{\pi \in P_2(m)} \prod_{r=1}^{m} \delta_{\lambda(r)\lambda(\pi(r))} n^{Z(\gamma\pi)-1-m/2}$$

Observe

$$\lim_{n\to\infty} n^{Z(\gamma\pi)-1-k} = \begin{cases} 1 & Z(\gamma\pi) = k+1 \\ 0 & Z(\gamma\pi) < k+1 \end{cases}$$

Thus

$$\lim_{n\to\infty} \phi(X^{\lambda(1)} \cdots X^\lambda(n)) = \sum_{\pi \in P_2(m)} \prod_{r=1}^{m} \delta_{\lambda(r)\lambda(\pi(r))}$$

We can write this formula

$$\lim_{n\to\infty} \phi(X^{\lambda(1)} \cdots X^\lambda(n)) = \sum_{\pi \in NCP_2(m)} \prod_{(r,s)\in\pi} \delta_{\lambda(r)\lambda(s))}$$

which is, by definition, the law of a free semicircular family (with covariance matrix the identity). □

## 8 Sums of Random Variables

### 8.1 Sums of Independent Random Variables

**Definition 8.1** Let $\mu_1$ and $\mu_2$ be two probability measures on the real line, $\mathbb{R}$. Then we define the *convolution*, $d\mu_1 * d\mu_2$, by writing

$$d\mu_1 * d\mu_2(v) = \int_{\mathbb{R}} d\mu_1(u) d\mu_2(u-v)$$

For a Borel set $S \subseteq \mathbb{R}$, we thus define

$$\mu_1 * \mu_2(S) = \int_S \int_{\mathbb{R}} d\mu_1(u) d\mu_2(u - v)$$

It is straightforward to check that the convolution of two probability measures is a well-defined probability measure.

**Proposition 8.2** *Let $X$ and $Y$ be real independent random variables, with laws defined by probability measures $\mu_X$ and $\mu_Y$ respectively. Then the sum $X + Y$ has law defined by the convolution of the measures $\mu_X$ and $\mu_Y$.*

**Proof:** We have moments defined by the formulae

$$m_n(X) = \int_{\mathbb{R}} x^n \, d\mu_X(x) \qquad m_k(Y) = \int_{\mathbb{R}} y^n \, d\mu_Y(y)$$

and

$$m_n(X+Y) = \phi((X+Y)^n) = \sum_{k=0}^{n} \binom{n}{k} \phi(X^k)\phi(Y^{n-k}) = \sum_{k=0}^{n} \int_{\mathbb{R}} \int_{\mathbb{R}} x^k y^{n-k} \, d\mu_X(x) d\mu_Y(y)$$

By linearity of the integral, we have the equation

$$m_n(X + Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} (x + y)^n d\mu_X(x) d\mu_Y(y)$$

Let $z = x + y$. Then

$$m_n(X + Y) = \int_{\mathbb{R}} \int_{\mathbb{R}} z^n d\mu_X(x) d\mu_Y(z - x) = \int_{\mathbb{R}} z^n \, d\mu_X * d\mu_Y(z)$$

Thus the convolution gives the moments of the sum $X + Y$, and the result follows. □

## 8.2 Free Convolution

Let $X$ and $Y$ be free random variables. We would like to define the probability law of the sum $X + Y$ in terms of the probability laws of the random variables $X$ and $Y$. This is in fact possible, using an analogue of the convolution in the independent case.

As elsewhere in these notes, we take a combinatorial approach, following [Spe94]. Alternative, more analytic methods, can be found in [VDN92] and [Haa97].

The first way to look at this convolution involves using cumulants, as introduced in definition 4.10.

**Proposition 8.3** *Let $X$ and $Y$ be free random variables. Then we have free cumulants*

$$k_n^{X+Y} = k_n^X + k_n^Y$$

**Proof:** By definition of the free cumulants

$$\begin{aligned} k_n^{X+Y} &= k_n(X+Y,\ldots,X+Y) \\ &= k_n(X,\ldots,X) + k_n(Y,\ldots,Y) \\ &= k_n^X + k_n^Y \end{aligned}$$

where the second step follows from the fact that the mixed cumulants $k_n\colon A^n \to \mathbb{C}$ are multilinear, and mixed cumulants vanish by freeness of the random variables $X$ and $Y$ and theorem 4.12. □

A simple application of the moment-cumulant formula yields the following result.

**Proposition 8.4** *Let $X$ be a random variable. Let $\pi \in NCP(n)$ be a non-crossing partition, with blocks $\{V_1,\ldots,V_r\}$. Define the cumulant $k_\pi^X$ by writing*

$$k_\pi^X = k_{\natural V_1}^X \cdots k_{\natural V_r}^X$$

*Let $(m_n^X)$ be the sequence of moments of the random variable $X$. Then we have the moment-cumulant formula*

$$m_n^X = \sum_{\pi \in NCP(n)} k_\pi^X$$

□

The above formula can be used to determine the free cumulants in terms of the moments. The following definition therefore makes sense.

**Definition 8.5** Let $\tau$ and $\tau'$ be probability laws, with sequences of moments $(m_k)$ and $(m'_k)$ respectively. Let $\{k_\pi \mid \pi \in NCP(n)\}$ and $\{k'_\pi \mid \pi \in NCP(n)\}$ be the sets of free cumulants arising from these laws.

Then we define the *free convolution*, $\tau \boxplus \tau'$, to be the probability law with squence of free cumulants $(k_n + k'_n)$.

The following result follows immediately from proposition 8.3 and the definition of the free convolution.

**Theorem 8.6** *Let $X$ and $Y$ be free random variables, with probability laws $\tau$ and $\tau'$ respectively. Then the sum $X+Y$ has probability law $\tau \boxplus \tau'$.* □

## 8.3 The Cauchy Transform

Let $X$ be a real random variable. We have already seen, in theorem 2.8, that there is a unique measure, $\mu$, on $\mathbb{R}$, such that the moments of $X$ are given by the formula

$$m_k(X) = \int_{\mathbb{R}} t^k \; d\mu(t)$$

In this section we will see more precisely how to determine the measure $\mu$ from the sequence of moments $(m_k)$.

**Definition 8.7** Let $\mu$ be a probability measure on the real line. Then we define the *Cauchy transform* of $\mu$ by the formula

$$G^{\mu}(z) = \int_{\mathbb{R}} \frac{1}{z-t} d\mu(t)$$

whenever $\Im(z) > 0$.

The Cauchy transform is relevant to us because of the following result, which can be proved by expressing the fraction as a power series and exchanging the integral and summation signs.

**Proposition 8.8** *Let $\mu$ be a probability measure on the real line, with moments*

$$m_k(\mu) = \int_{\mathbb{R}} t^k \ d\mu(t)$$

*Then the moment-generating function*

$$M^{\mu}(z) = \sum_{n=1}^{\infty} m_n z^n$$

*converges to an analytic function in some neighbourhood of* $0$. *Whenever the absolute value* $|z|$ *is sufficiently large, the formula*

$$G^{\mu}(z) = \frac{1}{z} M^{\mu}\left(\frac{1}{z}\right)$$

*holds.* □

In order to recover a measure from its moments, we use the following theorem to invert the Cauchy transformation.

**Theorem 8.9 (The Stieltjes Inversion Formula)** *Let $\mu$ be a probability measure on the real line. Then*

$$d\mu(t) = -\frac{1}{\pi} \lim_{\varepsilon \to 0} \Im(G^{\mu}(t + i\varepsilon)) \ dt$$

*whenever $t \in \mathbb{R}$. Here convergence is with respect to the weak topology on the space of all real probability measures.* □

## 8.4 The $R$-transform

The $R$-transform, along with the Cauchy transformation, gives us an analytic way of computing laws involving sums of free random variables.

**Definition 8.10** Let $\tau$ be a probability law, with sequence of free cumulants $(k_n)$. Then we define the *$R$-transform* by the formula

$$R^{\tau}(z) = \sum_{n=0}^{\infty} k_{n+1} z^n$$

The following result follows immediately from theorem 8.6.

**Theorem 8.11** *Let $X$ and $Y$ be free random variables, then*

$$R^{X+Y}(z) = R^X(z) + R^Y(z)$$

$\square$

There is a close relation between the $R$-transform and the Cauchy transform. In order to establish the relation, we begin with a purely algebraic result.

**Lemma 8.12** *Let $(m_n)_{n\geq 1}$ and $(k_n)_{n\geq 1}$ be sequences of complex numbers, with corresponding formal power series*

$$M(z) = 1 + \sum_{n=1}^{\infty} m_n z^n \qquad C(z) = 1 + \sum_{n=1}^{\infty} k_n z^n$$

*as generating functions. Write*

$$k_\pi = k_{\natural V_1} \cdots k_{\natural V_r} \qquad \pi = \{V_1, \ldots, V_r\} \in NCP(n)$$

*and suppose that*

$$m_n = \sum_{\pi \in NCP(n)} k_\pi$$

*Then*

$$C(zM(z)) = M(z)$$

**Proof:** Given a non-crossing partition $\pi \in NCP(n)$, let $V_1$ denote the block containing the element 1. The given formula for the number $m_n$ tells us that

$$m_n = \sum_{s=1}^{n} \sum_{\substack{V_1 \\ \natural V_1 = s}} \sum_{\substack{\pi \in NCP(n) \\ \pi = \{V_1, \ldots, V_r\}}} k_\pi$$

Let $\pi = \{V_1, \ldots, V_r\} \in NCP(n)$, and

$$V_1 = \{v_1, \ldots, v_s\}$$

Set $v_{s+1} = n+1$. Then, since $\pi$ is a non-crossing partition, for all $j \geq 2$ we can find $k$ such that $v_k < V_j < v_{k+1}$. It follows that

$$\pi = V_1 \cup \overline{\pi_1} \cup \cdots \cup \overline{\pi_s}$$

where $overline\pi_j$ is a non-crossing partition of the set $\{v_j + 1, \ldots, v_{j+1} - 1\}$.

Let $i_j = v_{j+1} - v_j - 1$. Then we can consider the partition $\overline{\pi_j}$ to be an element of the set $NCP(i_j)$. By the definition of the number $k_\pi$ we can write

$$k_\pi = k_s k_{\overline{\pi_1}} \cdots k_{\overline{\pi_s}}$$

since $\natural V_1 = s$. Therefore

$$m_n = \sum_{s=1}^{n} \sum_{\substack{i_1, \ldots, i_s = 0 \\ i_1 + \cdots + i_s + s = n}}^{n-s} \sum_{\substack{\natural V_1 = s \\ \overline{\pi_1}, \ldots, \overline{\pi_s} \in NCP(i_j)}} k_s k_{\overline{\pi_1}} \cdots k_{\overline{\pi_s}}$$

But

$$m_{i_j} = \sum_{\overline{\pi_r} \in NCP(i_j)} k_{\overline{\pi_s}}$$

for all $j$. Hence

$$m_n = \sum_{s=1}^{n} \sum_{\substack{i_1,\dots,i_s=0 \\ i_1+\cdots+i_s+s=n}}^{n-s} k_s m_{i_1} \cdots m_{i_s}$$

Now

$$M(z) = 1 + \sum_{n=1}^{\infty} m_n z^n = 1 + \sum_{n=1}^{\infty} \sum_{s=1}^{n} \sum_{\substack{i_1,\dots,i_s=0 \\ i_1+\cdots+i_s+s=n}}^{n-s} k_s z^s m_{i_1} z^{i_1} \cdots m_{i_s} z^{i_s}$$

We can write this sum

$$M(z) = 1 + \sum_{s=0}^{\infty} k_s z^s \left( 1 + \sum_{i=1}^{\infty} m_i z^i \right)^s = C(zM(z))$$

and we are done. □

**Theorem 8.13** *Let $\tau$ be a probability law, with Cauchy transfrom $G(z)$ and $R$-transform $R(z)$. Then*

$$R(G(z)) + \frac{1}{G(z)} = z \qquad G(R(z) + \frac{1}{z}) = z$$

**Proof:** Let $(m_n)$ be the sequence of moments of the probability law $\tau$, and let $(k_n)$ be the sequence of free cumulants. Write

$$K(z) = R(z) + \frac{1}{z}$$

The $R$-transform is defined by the formula

$$C(z) = 1 + zR(z)$$

where

$$C(z) = \sum_{n=0}^{\infty} k_n z^n$$

By the above lemma and proposition 8.8, we have the relations

$$M(z) = C(zM(z)) \qquad G(z) = \frac{1}{z} M\left(\frac{1}{z}\right)$$

So

$$K(G(z)) = R(G(z)) + \frac{1}{G(z)} = \frac{1}{G(z)} C(G(z)) = \frac{1}{G(z)} C\left(\frac{1}{z} M\left(\frac{1}{z}\right)\right) = \frac{1}{G(z)} M\left(\frac{1}{z}\right) = z$$

Set $w = G(z)$. Then we have the relation $G(K(w)) = w$. Since this formula is purely a formal relation between non-trivial power series, it follows in general that

$$G\left(R(z) + \frac{1}{z}\right) G(K(z)) = z$$

and we are done. □

## 8.5 Random Walks on Free Groups

Let $F_2$ be the free group on 2 generators, $s_1$ and $s_2$. Consider the random variable

$$X = \frac{1}{4}(X_1 + \cdots + X_n) \qquad X_i = s_i + s_i^{-1}$$

representing a step of a random walk going with equal probability in each possible direction on the group $F_2$. It is straightforward to check that the set of random variables $\{X_1, X_2\}$ is free.

The operator $X_1$ has the same law as the operator $s + s^{-1}$ on $l^2(\mathbb{Z})$, where $s$ is a generator of the group $\mathbb{Z}$. Let $S^1 \subseteq \mathbb{C}$ be the unit circle. Then we have an isomorphism of Hilbert spaces

$$l^2(\mathbb{Z}) \to L^2(S^1)$$

defined by mapping the series $(a_k)_{k\in\mathbb{Z}}$ to the Laurent series $\sum_{k=-\infty}^{\infty} a_k z^k$. Under this transformation, the algebra $\mathbb{C}\mathbb{Z}$ is mapped to the algebra of all Laurent polynomials, and the trace is mapped to the function

$$\phi(P) = \int_{S^1} P(z)\ d\mu(z)$$

where $\mu$ is the Haar measure on the unit circle.

The operator $\frac{1}{2}(s + s{-}1)$ is therefore mapped to the Laurent polynomial $z + 1/z$. The law of this operator therefore has Cauchy transformation

$$G_1(z) = \int_{S^1} \frac{1}{z - (w + 1/w)}\ d\mu(w)$$

The Haar measure is defined by writing $d\mu(w) = 1/(2\pi i w)\ dw$ so we can compute the integral as a path-integral

$$G_1(z) = -\frac{1}{2\pi i} \int_{S^1} \frac{1}{w^2 - wz + 1}\ dw$$

Let $p(w) = w^2 - wz - 1$. This polynomial has roots

$$w_1(z) = \frac{1}{2} + \sqrt{\frac{z^2}{4} - 1} \qquad w_2(z) = \frac{1}{2} - \sqrt{\frac{z^2}{4} - 1}$$

For sufficiently large values of $|z|$, the root $w_1(z)$ lies outside of the unit circle, and the root $w_2(z)$ lies inside the unit circle. Hence, by Cauchy's residue formula, we can compute this integral

$$G_1(z) = \frac{1}{w_1(z) - w_2(z)}$$

that is to say

$$G_1(z) = \left(z^2 - 4\right)^{-\frac{1}{2}}$$

Now, the $R$-transformation, $R_1(z)$, is determined by the formula

$$G_1(R_1(z) + 1/z) = z$$

so

$$\left(\left(R_1(z)+\frac{1}{z}\right)^2-4\right)^{-\frac{1}{2}}=z$$

and

$$R_1(z)=-\frac{1}{z}+\sqrt{4-\frac{1}{z^2}}$$

The random variable $X_2$ has the same law as the random variable $X_1$, and therefore $R$-transformation

$$R_2(z)=-\frac{1}{z}+\sqrt{4-\frac{1}{z^2}}$$

By theorem 8.11, the random variable $X=\frac{1}{4}(X_1+X_2)$ has $R$-transformation

$$R(z)=\frac{1}{4}\left(-\frac{1}{z}+\sqrt{4-\frac{1}{z^2}}\right)$$

Let $G$ be the Cauchy transform of the random variable $X$. Then

$$R(G(z))+1/G(z)=z$$

so

$$\frac{3}{4G(z)}+\sqrt{1-\frac{1}{4Gz^2}}=z$$

and we obtain the equation

$$(4z^2-16)G(z)^2-24zG(z)+5=0$$

We can solve this equation to obtain the formula

$$G(z)=\frac{3z+\sqrt{(31/4)z^2+5}}{z^2-4}$$

By proposition 8.8, the moment-generating function

$$M(z)=\sum_{k=1}^{\infty}\phi(X^k)z^k$$

is given by the formula

$$M(z)=\frac{1}{z}G\left(\frac{1}{z}\right)$$

so

$$M(z)=\frac{3+\sqrt{5+(31/4z^2)}}{1-4z^2}$$

This formula enables us to work out the moments by looking at a power series expansion, and therefore the probability law.

# 9 Products of Random Variables

## 9.1 Products of Independent Random Variables

Let $X$ and $Y$ be independent random variables. Then, by the definition of independence, we have moments

$$m_k(XY) = \phi((XY)^k) = \phi(X^k)\phi(Y^k) = m_k(X)m_k(Y)$$

We thus immediately have the law for the product $XY$. Let

$$M(z) = \sum_{n=1}^{\infty} m_n(X)m_n(Y)z^n$$

be the moment-generating function. By proposition 8.8, if the law of the product $XY$ is defined in terms of a probability measure $\mu$, then we have Cauchy transformation

$$G^{\mu}(z) = \frac{1}{z}M\left(\frac{1}{z}\right)$$

whenever the absolute value $|z|$ is sufficiently large.

## 9.2 The Multiplicative Free Convolution

Let $X$ and $Y$ be free random variables. We would like to define the probability law of the product $XY$ in terms of the probability laws of the random variables $X$ and $Y$. This is in fact possible, using calculations involving free cumulants, much as we did in the previous chapter. As we mentioned then, there are also analytic approaches; see [VDN92] or [Haa97] for details.

Our first result is best phrased algebraically.

**Theorem 9.1** *Let $(A, \phi)$ be a non-commutative probability space. Let $A$ and $B$ be free alagebras, and let $a_i \in A$, $b_i \in B$. Then the following formulae hold.*

- $\phi(a_1b_1a_2b_2\cdots a_nb_n) = \sum_{\pi\in NCP(n)} k_\pi[a_1,\ldots,a_n]\phi_{K(\pi)}[b_1,\ldots,b_n]$
- $\phi(a_1b_1a_2b_2\cdots a_nb_n) = \sum_{\pi\in NCP(n)} \phi_{K^{-1}(\pi)}[a_1,\ldots,a_n]k_\pi[b_1,\ldots,b_n]$
- $k_n(a_1b_1, a_2b_2, \cdots, a_nb_n) = \sum_{\pi\in NCP(n)} k_\pi[a_1,\ldots,a_n]k_{K(\pi)}[b_1,\ldots,b_n]$

**Proof:** The moment-cumulant formula tells us that

$$\phi(a_1b_1a_2b_2\cdots a_nb_n) = \sum_{\pi\in NCP(2n)} k_\pi[a_1, b_1, \ldots, a_n, b_n]$$

By freeness and theorem 4.12, the cumulant $k_\pi[a_1, b_1, \ldots, a_n, b_n]$ vanishes when an $a$-variable and a $b$-variable both appear in the same block of the partition $\pi$. By definition of the complement of a partition, it follows that

$$\phi(a_1b_1a_2b_2\cdots a_nb_n) = \sum_{\sigma\in NCP(n)} \left( k_\sigma[a_1,\ldots,a_n] \sum_{\substack{\sigma'\in NCP(n)\\ \sigma'\leq K(\sigma)}} k_{\sigma'}[b_1,\ldots,b_n] \right)$$

By the generalised moment-cumulant formula in proposition 4.14, we have the equation

$$\phi(a_1 b_1 a_2 b_2 \cdots a_n b_n) = \sum_{\sigma \in NCP(n)} \big( k_\sigma[a_1, \ldots, a_n] \phi_{K(\sigma)}[b_1, \ldots, b_n] \big)$$

The other equations follow from further applications of the generalised moment-cumulant formula. □

Since the probability law of a random variable determines and is determined by its free cumulants, the following definition makes sense.

**Definition 9.2** Let $\tau$ and $\tau'$ be probability laws, with sequences of moments $(m_k)$ and $(m'_k)$ respectively. Let $\{k_\pi \mid \pi \in NCP(n)\}$ and $\{k'_\pi \mid \pi \in NCP(n)\}$ be the sets of free cumulants arising from these laws.

Then we define the *multiplicative free convolution*, $\tau \boxtimes \tau'$, to be the probability law with free cumulants

$$k_n^{\tau \boxtimes \tau'} = \sum_{\pi \in NCP(\pi)} k_\pi k'_{K(\pi)}$$

The following result follows immediately from theorem 9.1 and the definition of the multiplicative free convolution.

**Theorem 9.3** *Let $X$ and $Y$ be free random variables, with probability laws $\tau$ and $\tau'$ respectively. Then the product $XY$ has probability law $\tau \boxtimes \tau'$.* □

## 9.3 Composition of Formal Power Series

As in the previous chapter, further calculations involving products of free random variables involve operations on certain formal power series.

**Definition 9.4** Let

$$\psi_1(z) = \sum_{n=1}^{\infty} a_n z^n \qquad \psi_2(z) = \sum_{n=1}^{\infty} b_n z^n$$

be formal power series. Then we define the *composition* $\psi_1 \circ \psi_2(z)$ to be the power series $\sum_{n=1}^{\infty} c_n z^n$ where

$$c_n = \sum_{rs=n} a_r b_s$$

If the functions $\psi_1(z) = \sum_{n=1}^{\infty} a_n z^n$ and $\psi_2(z) = \sum_{n=1}^{\infty} b_n z^n$ are well-defined, then the composition of formal power series is the same as composition of functions.

The following result relating compositions and products is straightforward to check.

**Proposition 9.5** *Let $\psi_1(z)$, $\psi_2(z)$, and $\psi_3(z)$ be formal power series. Then*

$$(\psi_1 . \psi_2) \circ \psi_3 = (\psi_1 \circ \psi_3)(\psi_2 \circ \psi_3)$$

□

**Proposition 9.6** *Let*

$$\psi(z) = \sum_{n=1}^{\infty} a_n z^n$$

*be a formal power series where* $a_1 \neq 0$. *Then there is a unique formal power series*

$$\psi^{-1}(z) = \sum_{n=1}^{\infty} b_n z^n$$

*such that*

$$\psi \circ \psi^{-1}(z) = \psi^{-1} \circ \psi(z) = z$$

**Proof:** We need to define coefficients $b_k$ such that

$$\sum_{rs=n} a_r b_s = \begin{cases} 1 & n = 1 \\ 0 & n \neq 1 \end{cases}$$

We know that $a_1 \neq 0$. We can therefore define $b_1 = 1/a_1$. The remaining coefficients, $b_k$, when $k \geq 2$, are determined inductively by the above formula. □

## 9.4 The $S$-transform

The $S$-transform is an analogue of the $R$-transform when we consider products rather than sums of free random variables. The results of the previous section mean that the following definition makes sense.

**Definition 9.7** Let $\tau$ be a probability law with a sequence of moments $(m_k^\tau)$, where $m_1^\tau \neq 0$. Define the formal power series

$$\psi_\tau(z) = \sum_{n=1}^{\infty} m_n^\tau z^n$$

Then we define the $S$*-transform* of $\tau$ by the formula

$$S^\tau(z) = \psi_\tau^{-1}(z) z^{-1}(1 + z)$$

**Lemma 9.8** *Let* $\tau$ *be a probability law, weith associated sequence of free cumulants* $(k_n)$. *Define a formal power series*

$$\Phi_\tau(z) = \sum_{n=1}^{\infty} k_n z^n$$

*Then we have* $S$*-transformation*

$$S^\tau(z) = \frac{\Phi_\tau^{-1}(z)}{z}$$

**Proof:** Write

$$M(z) = 1 + \psi_\tau(z) \qquad C(z) = 1 + \Phi_\tau(z)$$

Then by lemma 8.12

$$C(zM(z)) = M(z)$$

that is to say

$$\Phi_\tau(z(1 + \psi_\tau(z)) = \psi_\tau(z)$$

Write $w = \psi_\tau^{-1}(z)$. Then

$$\Phi_\tau(\psi_\tau^{-1}(w)(1 + w) = w$$

ie:

$$\Phi_\tau^{-1}(w) = \psi_\tau^{-1}(w)(1 + w)$$

The desired formula now follows by definition of the $S$-transform. □

**Theorem 9.9** *Let $X$ and $Y$ be free random variables. Then*

$$S^{XY}(z) = S^X(z)S^Y(z)$$

**Proof:** By theorem 8.6, we know that we have free cumulants

$$k_n^{XY} = \sum_{\pi \in NCP(n)} k_\pi^X k_{K(\pi)}^Y$$

for all $n$.

Let us define formal powers series $\Phi_X$, $\Phi_Y$, and $\Phi_{XY}$ as in the above lemma. Define

$$NCP'(n) = \{\pi \in NCP(n) \mid (1) \text{ is a block of } \pi\}$$

and write

$$gamma_n = \sum_{\pi \in NCP'(n)} k_\pi^X k_{K(\pi)}^Y \qquad \Phi_\gamma(z) = \sum_{n=1}^{\infty} \gamma_n z^n$$

Then we can check the relation

$$\Phi_{XY}(z) = \Phi_X \circ \Phi_\gamma(z)$$

Define

$$\delta_n = \sum_{\pi \in NCP'(n)} k_\pi^Y k_{K(\pi)}^X \qquad \Phi_\delta(z) = \sum_{n=1}^{\infty} \delta_n z^n$$

The random variables $XY$ and $YX$ have the same probability law. Therefore

$$\Phi_{XY}(z) = \Phi_Y \circ \Phi_\delta(z)$$

The formula

$$z\Phi_{XY}(z) = \Phi_\gamma(z)\Phi_\delta(z)$$

is also easy to verify. It follows that

$$\Phi_\gamma(z)\Phi_X^{-1} \circ \Phi_{XY}(z) \qquad \Phi_\delta(z)\Phi_Y^{-1} \circ \Phi_{XY}(z)$$

so

$$z\Phi_{XY}(z) = \Phi_X^{-1} \circ \Phi_{XY}(z)\Phi_Y^{-1} \circ \Phi_{XY}(z)$$

and by proposiition 9.5

$$\Phi_{XY}^{-1}(z)z = \Phi_X^{-1}(z)\Phi_Y^{-1}(z)$$

that is to say

$$\frac{\Phi_{XY}^{-1}(z)}{z} = \frac{\Phi_X^{-1}(z)}{z}\frac{\Phi_Y^{-1}(z)}{z}$$

and we are done, by the above lemma. □

# 10 More Random Matrices

## 10.1 Constant Matrices

The results we proved earlier about random matrices and freeness can be generalised to include non-Gaussian self-adjoint matrices. The first step in this calculation is to look at constant matrices.

**Definition 10.1** A sequence of matrices, $(D_n)$, where $D_n \in M_n(\mathbb{C})$ is said to *converge in distribution* if the sequence of traces $tr(D_n^m)$ converges for each $m$.

A normal matrix, $D_n$, can be regarded as a constant function $\Omega \to M_n(\mathbb{C})$ on a probability space $\Omega$, and so as a constant random matrix. Taking this point of view, the above definition is the same as the definition of convergence in distribution for a sequence of random variables.

We want to do some calculations involving both constant matrices and Gaussian matrices. Before presenting the results of aour first such calculation, however, it is convenient to introduce some notation.

**Definition 10.2** Let $A^{(1)}, \dots, A^{(m)}$ be $n \times n$ matrices. Write $A^{(k)} = (A_{ij}^{(k)})$. Let $\sigma \in \Sigma_m$ be a permutation. Then we write

$$\underset{\sigma}{tr}(A^{(1)}, \dots, A^{(m)}) = \sum_{j(1),\dots,j(m)=1}^{n} A_{j(1)j(\sigma(1))} \cdots A_{j(m)j(\sigma(m))}$$

Observe (since we divide by $k$ when defining the ordinary trace of an $k \times k$ matrix) that the above trace is a product of ordinary traces multiplied by $n^{Z(\sigma)}$, where $Z(\sigma)$ is the number of cycles of the permutation $\sigma$.

**Lemma 10.3** *Let $X$ be a standard self-adjoint Gaussian $n \times n$ matrix, and let $D$ be a constant $n \times n$ matrix. Then*

$$\phi(D^{q(1)}XD^{q(2)}X \cdots D^{q(m)}X) = \sum_{\pi \in P_2(m)} \underset{\gamma\pi}{tr}[D^{q(1)} \cdots D^{q(m)}]n^{Z(\gamma\pi)-1-m/2}$$

**Proof:** Write $D^k = (D_{ij}^{(k)})$, and $X = (X_{ij})$. Then by definition of the state $\phi$,

$$\phi(D^{q(1)}XD^{q(2)}X\cdots D^{q(m)}X) = \frac{1}{n}\sum_{\substack{i(1),\ldots,i(m)\\ j(1),\ldots,j(m)=1}}^{n} E[D_{j(1)i(1)}^{(q(1))}X_{i(1)j(2)}D_{j(2)i(2)}^{(q(2))}X_{i(2)j(3)}\cdots D_{j(m)i(m)}^{(q(m))}X_{i(m)j(1)}]$$

Since the variables $D_{j(k)i(k)}^{(q(k))}$ are all constant, we can write the above sum as the expression

$$\frac{1}{n}\sum_{\substack{i(1),\ldots,i(m)\\ j(1),\ldots,j(m)=1}}^{n} E[X_{i(1)j(2)}X_{i(2)j(3)}X_{i(m)j(1)}]D_{j(1)i(1)}^{(q(1))}D_{j(2)i(2)}^{(q(2))}\cdots D_{j(m)i(m)}^{(q(m))}$$

which by the Wick formula for moments of Gaussian families can be written

$$\frac{1}{n}\sum_{\substack{i(1),\ldots,i(m)\\ j(1),\ldots,j(m)=1}}^{n}\sum_{\pi\in P_2(m)}\pi_{(r,s)\in\pi}E[X_{i(r)j(r+1)}X_{i(s)j(s+1)}]D_{j(1)i(1)}^{(q(1))}D_{j(2)i(2)}^{(q(2))}\cdots D_{j(m)i(m)}^{(q(m))}$$

The fact that the matrix $(X_{ij})$ is a standard self-adjoint Gaussian random matrix now tells us that $\phi(D^{q(1)}X\cdots D^{q(m)}X) =$

$$\frac{1}{n}\sum_{\substack{i(1),\ldots,i(m)\\ j(1),\ldots,j(m)=1}}^{n}\sum_{\pi\in P_2(m)}\pi_{(r,s)\in\pi}\delta_{i(r)j(r+1)}\delta_{i(s)j(s+1)}\frac{1}{n^{m/2}}D_{j(1)i(1)}^{(q(1))}D_{j(2)i(2)}^{(q(2))}\cdots D_{j(m)i(m)}^{(q(m))}$$

Let us fix $\pi\in P_2(m)$. As usual, we can identify $\pi$ with a permuatation $\pi\in\Sigma_m$. Let $\gamma=(1,2,\ldots,m)\in\Sigma_m$. Then the sum

$$\sum_{\substack{i(1),\ldots,i(m)\\ j(1),\ldots,j(m)=1}}^{n}\pi_{(r,s)\in\pi}\delta_{i(r)j(r+1)}\delta_{i(s)j(s+1)}D_{j(1)i(1)}^{(q(1))}\cdots D_{j(m)i(m)}^{(q(m))}$$

can then be written in the form

$$\sum_{\substack{i(1),\ldots,i(m)\\ j(1),\ldots,j(m)=1}}^{n}\pi_{r=1}^{m}\delta_{i(r)j(\gamma\pi(r))}D_{j(1)i(1)}^{(q(1))}\cdots D_{j(m)i(m)}^{(q(m))} = \sum_{j(1),\ldots,j(m)=1}^{n}D_{j(1)j(\gamma\pi(1))}^{(q(1))}\cdots D_{j(m)j(\gamma(\pi(m))}^{(q(m))}$$

Using definition 10.2, we have the simplification

$$\sum_{\substack{i(1),\ldots,i(m)\\ j(1),\ldots,j(m)=1}}^{n}\pi_{(r,s)\in\pi}\delta_{i(r)j(r+1)}\delta_{i(s)j(s+1)}D_{j(1)i(1)}^{(q(1))}\cdots D_{j(m)i(m)}^{(q(m))} = \frac{1}{n^{Z(\sigma)}}\underset{\sigma}{tr}[D^{(q(1))},\cdots,D^{(q(m))}]$$

and so the formula appearing in the statement of the lemma. □

**Theorem 10.4** *Let $(X_n)$ be a sequence of $n\times n$ standard self-adjoint Gaussian random matrices, and let $(D_n)$ a sequence of $n\times n$ constant matrices. Then the sequence of pairs $\{X_n, D_n\}$ converges in distribution to a pair of free random variables.*

**Proof:** When $m$ is odd, the set set $P_2(m)$ is empty, so by the above formula the moments $\phi(D^{q(1)}XD^{q(2)}X\cdots D^{q(m)}X)$ are all zero. Let $m=2k$.

Let $\pi \in P_2(2k)$. Then the number of cycles, $Z(\gamma\pi)$, can be at most $1+k$, and this number is achieved precisely when the pairing is non-crossing. By the above lemma, we have the formula

$$\phi(D^{q(1)}XD^{q(2)}X\cdots D^{q(m)}X) = \sum_{\pi\in P_2(m)} \underset{\gamma\pi}{tr}[D^{q(1)}\cdots D^{q(m)}]n^{Z(\gamma\pi)-1-m/2}$$

Observe

$$\lim_{n\to\infty} n^{Z(\gamma\pi)-1-k} = \begin{cases} 1 & Z(\gamma\pi)=k+1 \\ 0 & Z(\gamma\pi)<k+1 \end{cases}$$

Thus

$$\lim_{n\to\infty}\phi(D^{q(1)}XD^{q(2)}X\cdots D^{q(m)}X) = \sum_{\pi\in NCP_2(m)} \underset{\gamma\pi}{tr}[D^{q(1)}\cdots D^{q(m)}]$$

We can rewrite this formula

$$\lim_{n\to\infty}\phi(D^{q(1)}XD^{q(2)}X\cdots D^{q(m)}X) = \lim_{n\to\infty}\sum_{\pi\in NCP_2(m)} \phi_{K^{-1}(\pi)}[D^{q(1)},\cdots,D^{q(m)}]$$

By theorem 9.1, the above formula shows that in the limit the mixed moments are those of a pair of free random variables. The result now follows. □

We can generalise the above result to families of Gaussian random variables and constant matrices. The proof is the same as that given above. If we combine this generalisation with theorem 7.10, we see the following result.

**Theorem 10.5** *Let $\{X_n^\lambda \mid \lambda\in\Lambda\}$ be a sequence of independent families of standard self-adjoint $n\times n$ Gaussian matrices. Let $\{D_n^\lambda \mid \lambda\in\Lambda\}$ be a sequence of families of constant random matrices that converge in distribution to random variables $\{d_\lambda \mid \lambda\in\Lambda\}$.*

*Then the sequence of families $(\{X_n^\lambda\})$ converges in distribution to a free semi-circular family $\{s_\lambda \mid \lambda\in\Lambda\}$. Further, the sets*

$$\{s_\lambda \mid \lambda\in\Lambda\} \qquad \{d_\lambda \mid \lambda\in\Lambda\}$$

*are free from each-other.* □

## 10.2 Unitary Random Matrices

In this section we want to imitate our earlier calculations on asymptotic freeness for unitary rather than self-adjoint Gaussian random matrices. We begin by defining the objects we want to examine.

**Definition 10.6** A random matrix $V\in M_n$ is *unitary* if $V^*V=VV^*=I$, where $I=1\otimes I\in L\otimes M_n(\mathbb{C})$ is the identity matrix.

Let $U(n)$ denote the space of complex unitary $n \times n$ matrices. A random unitary matrix is then by definition a map of the form $V \colon \Omega \to U(n)$, where $\Omega$ is a probability space.

However, the space $U(n)$ is a compact topological group, and so can be equipped with the *Haar measure*, $h$, namely the a unique multiplication-invariant probability measure defined on all Borel measurable sets.[3]

**Definition 10.7** Let $\Omega$ be a probability space, with measure $\mu$. Then a random unitary matrix $V \colon \Omega \to U(n)$ is called a *Haar matrix* if it is uniformly distributed on the space of unitary matrices, that is to say the equation

$$h(S) = \mu(V^{-1}[S])$$

holds for any Borel set $S \subseteq U(n)$.

Let $X$ be a standard Gaussian random matrix. By proposition 7.3, the matrix $X$ is invertible (as an element of the von Neumann algebra of random matrices), so we have a polar decomposition

$$X = AV$$

where the random matrix $A$ is positive and the matrix $V$ is unitary.

**Proposition 10.8** *The random unitary matrix $V$ defined above is a Haar matrix.*

**Proof:** Let $W$ be a fixed unitary matrix. Then we want to show that the random matrices $V$ and $WV$ have the same law. The result then follows by theorem 2.8 and the definition of the Haar measure as the unique multiplication-invariant probability measure on the space of unitary matrices.

By functional calculus, we can write

$$V = (X^*X)^{\frac{-1}{2}}X \qquad WV = ((WX^*)(WX))^{\frac{-1}{2}}WX$$

It therefore suffices to show that the random matrix $WX$ is also Gaussian. This follows by the definition of independence, the definition of unitary, and the fact that a sum of independent Gaussian random variables is also Gaussian. □

**Proposition 10.9** *Let $\{X^\lambda \mid \lambda \in \Lambda\}$ be a family of complex Gaussian random matrices. Then the family of unitary matrices*

$$\{V^\lambda \mid \lambda \in \Lambda\}$$

*arising from polar decompositions as above is independent.*

**Proof:** The polar decompositions $X^\lambda = A^\lambda V^\lambda$ are defined by functional calculus:

$$A^\lambda = (X^\lambda(X^\lambda)^*)^{\frac{1}{2}} \qquad V^\lambda = X^\lambda(X^\lambda(X^\lambda)^*)^{-\frac{1}{2}}$$

Since the matrices $\{X^\lambda\}$ are independent, they pairwise commute. By construction, the matrices $\{V^\lambda\}$ also pairwise commute.

[3]A construction of the Haar measure can be found for example in chapter 5 of [Rud91].

Again using independence of the family $\{X^\lambda\}$, we know that

$$\phi(X^{\lambda(1)}X^{\lambda(2)}\cdots X^{\lambda(n)}) = \phi(X^{\lambda(1)})\phi(X^{\lambda(2)})\cdots\phi(X_{\lambda(n)})$$

whenever $\lambda(i) \neq \lambda(j)$ when $i \neq j$

It is easy to check that $\phi(A^\lambda) = |\phi(X^\lambda)|$ for all $\lambda$. Since the state $\phi$ is in fact a trace, we see that

$$\phi(V^{\lambda(1)}V_{\lambda(2)}\cdots V^{\lambda(n)}) = \phi(V^{\lambda(1)})\phi(V^{\lambda(2)})\cdots\phi(V^{\lambda(n)})$$

and we have shown independence of the family $\{V^\lambda\}$.

Let $\lambda \in \Lambda$. Then we want to check that the random unitary matrix $V_\lambda$ is a Haar matrix. For Borel set $S \subseteq U(n)$, define

$$h(S) = \mu((V^\lambda)^{-1}[S])$$

where $\mu$ is a probability measure on the set $\Omega$. Then $h(U(n)) = 1$ so $h$ is a probability measure. □

In order to perform calculations for random unitary matrices, we need to look at some expectations involving the entries.

**Definition 10.10** Let $\pi \in \Sigma_n$ be a permutation, let $N \geq n$, and let $V = (V_{ij})$ be an $N \times N$ Haar matrix. Then we define the *Weingarten function*

$$Wg(N,\pi) = E(V_{11}\ldots V_{nn}\overline{V_{1\pi(1)}}\cdots\overline{V_{n\pi(n)}}$$

The Weingarten function $Wg(N,\pi)$ depends only on the conjugacy class of the permutation $\pi$. The general behaviour of this function is rather complex. The following formula, relating general expectations to the Weingarten function is easily seen, however.

**Proposition 10.11** *Let $V = (V_{ij})$ be an $N \times N$ Haar matrix. Then the expectation*

$$E(V_{i'(1)j'(1)}\ldots V_{i'(n)j'(n)}\overline{V_{i(1)j(1)}}\cdots\overline{V_{i(n)j(n)}}$$

*is equal to the sum*

$$\sum_{\alpha,\beta\in\Sigma_n} \delta_{i(1)i'(\alpha(1))}\cdots\delta_{i(n)i'(\alpha(n))}\delta_{j(1)i'(\beta(1))}\cdots\delta_{j(n)j'(\beta(n))}\,Wg(N,\beta\alpha^{-1})$$

□

We omit the proof of the following (actually quite involved) result on the asymptotic behaviour of the Weingarten function. For details, the article [Nic93], where results involving the asymptotic behaviour of random unitaries were originally examined, can be consulted.

**Lemma 10.12** *Let $\pi \in \Sigma_n$ be a permutation. Then*

$$\lim_{N\to\infty} Wg(N,\pi)N^{2n-Z(\pi)} = 1$$

□

**Definition 10.13** Let $(A, \phi)$ be a $C^*$-probability space. A random variable $V \in A$ is called a *Haar unitary* if $\phi(V^k) = 0$ for all $k \in \mathbb{Z}\backslash\{0\}$.

It is clear than any Haar matrix is a Haar unitary. The following result can be proved using our usual methods on the asymptotic behaviour of random matrices and lemma 10.12.

**Theorem 10.14** *Let $\{U_n^\lambda \mid \lambda \in \Lambda\}$ be a sequence of independent families of $n \times n$ Haar matrices. Let $\{D_n^\lambda \mid \lambda \in \Lambda\}$ be a sequence of families of constant random matrices that converge in distribution to random variables $\{d_\lambda \mid \lambda \in \Lambda\}$.*

*Then the sequence of families $(\{X_n^\lambda\})$ converges in distribution to a free family of Haar unitaries $\{u_\lambda \mid \lambda \in \Lambda\}$. Further, the algebras generated by the sets*

$$\{u_\lambda \mid \lambda \in \Lambda\} \qquad \{d_\lambda \mid \lambda \in \Lambda\}$$

*are free.* □

## 10.3 Random Rotations

To see why we are interested in unitary matrices, note the following proposition, which follows immediately from the definition of freeness.

**Proposition 10.15** *Let $X$ and $Y$ be free random variables. Let $U$ be a Haar unitary that is free from the variables $X$ and $Y$. Then the random variables $X$ and $UYU^*$ are also free.* □

The above proposition along with theorem 10.14 gives us the following result.

**Theorem 10.16** *Let $(X_n)$ and $(Y_n)$ be sequences of constant $n \times n$ matrices which converge in distribution to random variables $X$ and $Y$ respectively. Let $U_n$ be an $n \times n$ Haar matrix. Then the sequences $(X_n)$ and $(U_n Y_n U_n^*)$ converge in distribution to free random variables $X$ and $Y'$ respectively, where $Y'$ has the same distribution as the random variable $Y$.* □

The above theorem tells us that any two free random variables can be realised as a limit (in distribution) of a sequence of constant matrices and a sequence of constant matrices rotated by Haar matrices.

# References


[Con94] A. Connes. *Noncommutative Geometry.* Academic Press, 1994.

[Dix81] J. Dixmier. *Von Neumann Algebras.* North-Holland Publishing Co., 1981.

[Haa97] U. Haagerup. On Voiculescu's $R$- and $S$-transforms for free non-commuting random variables. In *Free probability theory (Waterloo, ON, 1995)*, volume 12 of *Fields Institute Communications*, pages 127–148. American Mathematical Society, 1997.

[Nic93] A. Nica. Asymptotically free families of random unitaries in symmetric groups. *Pacific Journal of Mathematics*, 157:295–310, 1993.

[Rud87] W. Rudin. *Real and Complex Analysis.* McGraw-Hill, 1987.

[Rud91] W. Rudin. *Functional Analysis. Second Edition.* International Series in Pure and Applied Mathematics. McGraw-Hill, 1991.

[Spe93] R. Speicher. Free convolution and the random sum of matrices. *Publications of the RIMS*, 29:731–744, 1993.

[Spe94] R. Speicher. Multiplicative functions on the lattice of non-crossing partitions and free convolution. *Mathematische Annalen*, 298:611–628, 1994.

[Spe98] R. Speicher. *Combinatorial theory of the free product with amalgamation and operator-valued free probability theory*, volume 132 of *Memoirs of the American Mathematical Society.* The American Mathematical Society, 1998.

[VDN92] D.V. Voiculescu, K.J. Dykema, and A. Nica. *Free Random Variables*, volume 1 of *CRM Monograph Series.* American Mathematical Society, 1992.

[Voi85] D.V. Voiculescu. Symmetries of some reduced free product $C^*$-algberas. In *Operator algebras and their connections with topology and ergodic theory (Bucsteni, 1983)*, volume 1132 of *Lecture Notes in Mathematics*, pages 556–588. Springer, 1985.

[Voi91] D.V. Voiculescu. Limit laws for random matrices and free products. *Inventiones Mathematica*, 104:201–220, 1991.