GOALS Free Probability Lecture 1

1 Noncomutative probability spaces and laws

Definition 1.1. A (noncommutative) $*$–**probability space** consists of a unital $*$-algebra, $\mathcal{A}$ and a state φ on $\mathcal{A}$ (here, state means that $\varphi(1) = 1$ and $\varphi(x^*x) \geq 0$ for all $x \in \mathcal{A}$.). We notate this by the pair $(\mathcal{A}, \varphi)$. Often of interest to operator algebraists are C* or W* probability spaces:

- A **C*–probability space** $(\mathcal{A}, \varphi)$ is a *-probability space where $\mathcal{A}$ is a C*-algebra. As you learned, or will learn soon, φ is automatically norm continuous.
- A **W*–probability space** $(\mathcal{A}, \varphi)$ is a *-probability space where $\mathcal{A}$ is a von Neumann algebra, and φ is normal (ultraweakly continuous)

In any of these circumstances, one often wants φ to be faithful, but that is not a requirement for this definition.

Examples 1.2. (i) If Ω is a measure space, F a σ–algebra of measurable sets, and P is a probability measure defined on F, then $(L^\infty(\Omega, F, P), \int_\Omega \cdot dP)$ is a W*–probability space.

(ii) If one defines $L^{\infty-}(\Omega, F, P) := \bigcap_{p\geq 1} L^p(\Omega, F, P)$, then $(L^{\infty-}(\Omega, F, P), \int_\Omega \cdot dP)$ is a *–probability space. $L^{\infty-}(\Omega, F, P)$ consists of all random variables which have all of their moments. If (Ω, F, P) has no atoms, then this $*$–probability space includes random variables with Gaussian distribution which are necessarily unbounded, thus this is not a C* or W*–probability space.

(iii) $(C([0, 1]), \int \cdot dx)$ is a C*–probability space

(iv) $(M_n(\mathbb{C}), \mathrm{tr})$ is a W*–probability space

(v) $(M_n(L^{\infty-}(\Omega, F, P)), (\int_\Omega \cdot dP) \circ \mathrm{tr})$ is a $*$–probability space. This is the algebra of $n \times n$ random matrices where all entries have all of their moments.

(vi) If Γ is a discrete group, and τ is the canonical trace on $L(\Gamma)$, then $(L(\Gamma), \tau)$ is a W*–probability space.

The notion of a joint distribution in classical probability is essential to the study of the subject. Non-commutative probability has this notion as well, but like many ideas in operator algebras, one must strip away a measure space to make a proper definition. For what follows,

$\mathbb{C}\langle X_1, X_1^*, \cdots, X_n, X_n^*\rangle$ will denote the universal noncommutative $*$-algebra generated by $X_1, X_1^*, \cdots, X_n, X_n^*$. This is often called the algebra of noncommutative polynomials in $X_1, X_1^*, \cdots, X_n, X_n^*$.

Definition 1.3. Let $(\mathcal{A}, \varphi)$ be a $*$–probability space and $a_1, \cdots, a_n \in \mathcal{A}$. The **$*$-distribution** or **law** of $a = (a_1, \cdots, a_n)$ is defined to be the linear functional $\mu_a : \mathbb{C}\langle X_1, X_1^*, \cdots, X_n, X_n^*\rangle \to \mathbb{C}$ given by:

$$\mu_a(P(X_1, X_1^*, \cdots, X_n, X_n^*)) = \varphi(P(a_1, a_1^*, \cdots, a_n, a_n^*))$$

If this above notation is intimidating, note that, for example, $\mu_a(X_1X_2X_1^*) = \varphi(a_1a_2a_1^*)$. In the exercises, you will show that the law of $(a_1, \cdots, a_n)$ uniquely determines the C* or von Neumann algebra that $(a_1, \cdots, a_n)$ generates. Although this definition is a bit abstract, it is the case that for a collection $(Y_1, \cdots, Y_n)$ of classical L^∞ random variables, the values of $\mathbb{E}(P(Y_1, \overline{Y_1}, \cdots, Y_n, \overline{Y_n}))$ determine the joint distribution in the usual probabilistic sense.

Of great importance is the $*$-distribution of a single $a \in (\mathcal{A}, \varphi)$.

Important Example 1.4. Suppose that $\mathcal{A}$ is a C*–probability space and that a is **normal**. We know that $C^*(a) \cong C(\sigma(a))$ and since φ is a state on $C^*(a)$ it follows that under this identification, it given by integration against a probability measure μ supported on $\sigma(a)$. Therefore, if $P \in \mathbb{C}\langle X, X^*\rangle$, it follows that

$$\mu_a(P(X, X^*)) = \varphi(P(a, a^*)) = \int_{\sigma(a)} P(z, \overline{z})d\mu.$$

In this instance, we abuse notation and say that the measure μ is the law of a. Since polynomials in z and $\overline{z}$ are uniformly dense in $C(\sigma(a))$, it follows that μ is uniquely determined by the values of $\varphi(P(a, a^*))$ over all polynomials, P.

In the first example of *-probability spaces given above, it is a good thing to check that the law of a random variable X is nothing other than its distribution in the probabilistic sense.

There are many normal elements that have well known laws. We will discuss a few here:

Example 1.5. Let $A \in (M_n(\mathbb{C}), \mathrm{tr})$ be a normal matrix with eigenvalues $\lambda_1, \cdots, \lambda_n$ (repetition allowed). Since A and A^* are simultaneously diagonalizable,

$$\mathrm{tr}(A^k(A^*)^\ell) = \frac{1}{n}\sum_{j=1}^{n} \lambda_j^k \overline{\lambda_j^\ell}$$

It follows that if one defines μ by $\mu = \frac{1}{n}\sum_{j=1}^{n} \delta_{\lambda_j}$, then for any polynomial P,

$$\mathrm{tr}(P(A, A^*)) = \int_{\sigma(A)} P(z, \overline{z})d\mu$$

meaning that the law of A is simply the normalized counting measure on its eigenvalues

Example 1.6. A unitary $u \in (\mathcal{A}, \varphi)$ is a **Haar unitary** whenever $\varphi(u^k) = \begin{cases} 0 & \text{if } k \in \mathbb{Z} \setminus \{0\} \\ 1 & \text{if } k = 0 \end{cases}$ (note that we have allowed for negative powers of u i.e. positive powers of u^*). As $\sigma(u)$ is a subset of the unit circle, $\mathbb{T}$, it follows that the law of u is a measure μ supported inside of $\mathbb{T}$. We claim that μ is the normalized arc length measure, λ, on the unit-circle. Indeed, note that

$$\int_{\mathbb{T}} z^k d\lambda = \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} d\theta = \begin{cases} 0 & \text{if } k \in \mathbb{Z} \setminus \{0\} \\ 1 & \text{if } k = 0 \end{cases}$$

It follows that $\int_{\mathbb{T}} z^k d\lambda = \varphi(u^k)$ for all $k \in \mathbb{Z}$, giving the desired law. As an added bonus, this example proves that $\sigma(u)$ is all of $\mathbb{T}$ rather than a proper subset of it.

The next example will be arguably the most important self-adjoint distribution in free probability. To define it, we need some setup:

2 The full Fock space and semicircular elements

Let $\mathcal{H}$ be a Hilbert space. We define the **full Fock space** of $\mathcal{H}$, $\mathcal{F}(\mathcal{H})$) to be the Hilbert space:

$$\mathcal{F}(\mathcal{H}) = \mathbb{C}\Omega \oplus \bigoplus_{n \geq 1} \mathcal{H}^{\otimes n}$$

where $\|\Omega\| = 1$. Ω is called the **vacuum vector**. Vectors in $\mathcal{H}$ naturally give rise to **creation and annihilation operators**.

For any $\xi \in \mathcal{H}$, define the **creation operator** $\ell(\xi)$ on $\mathcal{F}(\mathcal{H})$ by the linear extension of:

$$\ell(\xi)\Omega = \xi \text{ and } \ell(\xi)(\xi_1 \otimes \cdots \otimes \xi_n) = \xi \otimes \xi_1 \otimes \cdots \otimes \xi_n$$

One notes that $\ell(\xi)$ extends to a bounded operator on $\mathcal{F}(\mathcal{H})$ with $\|\ell(\xi)\| = \|\xi\|$. One further observes that $\ell(\xi)^*$ satisfies:

$$\ell(\xi)^*\Omega = 0 \text{ and } \ell(\xi)^*(\xi_1 \otimes \cdots \otimes \xi_n) = \langle \xi_1, \xi \rangle \xi_2 \otimes \cdots \otimes \xi_n,$$

consequently we call $\ell(\xi)^*$ an **annihilation operator**. With the definitions of ℓ and ℓ^*, the verification of the following facts is straightforward and left to the reader:

Facts 2.1.

(i) $\ell(\xi)^*\ell(\xi) = \|\xi\|^2 1$ for all $\xi \in \mathcal{H}$. More generally, $\ell(\eta)^*\ell(\xi) = \langle \xi, \eta \rangle 1$ for all $\xi, \eta \in \mathcal{H}$

(ii) If $\xi \in \mathcal{H}$ is a unit vector, then $\ell(\xi)\ell(\xi)^*$ is the orthogonal projection onto $\bigoplus_{n \geq 0} \mathbb{C}\xi \otimes \mathcal{H}^{\otimes n}$

(iii) $\ell(a\xi + b\eta) = a\ell(\xi) + b\ell(\eta)$ for all $a, b \in \mathbb{C}$ and $\xi, \eta \in \mathcal{H}$

(iv) $\ell(a\xi + b\eta)^* = \overline{a}\ell(\xi)^* + \overline{b}\ell(\eta)^*$ for all $a, b \in \mathbb{C}$ and $\xi, \eta \in \mathcal{H}$

Place a normal state φ on $B(\mathcal{F}(\mathcal{H}))$ be setting $\varphi(x) = \langle x\Omega, \Omega\rangle$. We call φ the **vacuum state**. Fix a unit vector $\xi \in \mathcal{H}$. We are interested in computing the law of $s = \ell(\xi) + \ell(\xi)^*$ with respect to φ. Since s is self adjoint, we are after a compactly supported measure μ on $\mathbb{R}$ so that

$$\varphi(s^n) = \int_{\mathbb{R}} x^n d\mu(x)$$

Note that $s^n = \sum_{i_1,\cdots,i_n \in \{1,*\}} \ell(\xi)^{i_1} \cdots \ell(\xi)^{i_n}$, meaning that $\varphi(s^n) = \sum_{i_1,\cdots,i_n \in \{1,*\}} \langle \ell(\xi)^{i_1} \cdots \ell(\xi)^{i_n}\Omega, \Omega\rangle$. We need to examine in which instances $\langle \ell(\xi)^{i_1} \cdots \ell(\xi)^{i_n}\Omega, \Omega\rangle$ is nonzero. Immediately, one sees that the number of indices i_k that are a 1 must match the number of indices that are a $*$ (or else $\ell(\xi)^{i_1} \cdots \ell(\xi)^{i_n}\Omega$ is zero or some $\xi^{\otimes k}$ for $k \geq 1$) thus we can assume $n = 2m$ for some integer m. In addition, it must be the case for each $j \in \{1, n\}$ $\#\{k \geq j \,|\, i_k = 1\} \geq \#\{k \geq j \,|\, i_k = *\}$ or else $\ell(\xi)^{i_1} \cdots \ell(\xi)^{i_n}\Omega = 0$. In this latter event, $\ell(\xi)^{i_1} \cdots \ell(\xi)^{i_n}\Omega = \Omega$ meaning that this inner product is 1. The number of such terms is counted by Dyck paths (to be shown in lecture). The number of Dyck paths of size $2m$ is famously m^{th} Catalan number $C_m = \frac{1}{m+1}\binom{2m}{m}$. This shows that

$$\varphi(s^n) = \begin{cases} 0 & \text{if } n \text{ is odd} \\ C_m & \text{if } n = 2m \end{cases}$$

It is known that

$$\frac{1}{2\pi}\int_{-2}^{2} x^n \sqrt{4 - x^2}dx = \begin{cases} 0 & \text{if } n \text{ is odd} \\ C_m & \text{if } n = 2m \end{cases}$$

Consequently, it follows that the law of s is the **semicircular law** of variance 1: $d\mu = \mathbb{1}_{[-2,2]}\frac{1}{2\pi}\sqrt{4 - x^2}dx$. The term "variance 1" is used because one defines the **variance** of any a as $\varphi(a^*a) - |\varphi(a)|^2$ and it is clear that for s, this quantity is 1.

The semicircular law is the free analogue of the gaussian law in classical probability. This is explored in one of the exercises.

3 Free Independence

In classical probability, one has the notion of independence of events. This notion of independence carries over to random variables: Of one assumes that X and Y are essentially bounded classical random variables, then independence of X and Y is characterized by

$$\mathbb{E}(P(X, \overline{X})Q(Y, \overline{Y})) = \mathbb{E}(P(X, \overline{X})) \cdot \mathbb{E}(Q(Y, \overline{Y}))$$

for any polynomials P and Q. We are after a similar notion of independence in noncommutative probability. As there are more words to consider when the elements do not commute, one needs to take more care:

Definition 3.1. Let $(\mathcal{A}, \varphi)$ be a $*$–probability space and $a \in (\mathcal{A}, \varphi)$. We say that a is **centered** whenever $\varphi(a) = 0$

Centered elements form the backbone of **free independence**

Definition 3.2. Let $(\mathcal{A}, \varphi)$ be a $*$–probability space, I be an index set. and $\{A_i\}_{i\in I}$ be a family of unital $*$-subalgebras of $\mathcal{A}$. We say that the collection $\{A_i\}_{i\in I}$ is **freely independent** if whenever $a_{i_1} \in A_{i_1}, a_{i_2} \in A_{i_2}, \cdots, a_{i_n} \in A_{i_n}$ satisfy the following properties

(i) a_{i_k} is centered for $1 \leq k \leq n$

(ii) $i_k \neq i_{k+1}$ for $1 \leq k \leq n-1$ (note, one is allowed, for instance, $i_1 = i_3$ although this is certainly not the only possibility)

then $a_{i_1} a_{i_2} \cdots a_{i_n}$ is centered.

If J is an index set and $\{S_j\}_{j\in J}$ are subsets of $\mathcal{A}$, and if B_j is the unital *-subalgebra of A generated by S_j, then we say that $\{S_j\}_{j\in J}$ is *****–free if the collection $\{B_j\}_{j\in J}$ is $*$–free

The above definition extends in the natural way to C* and W* probability spaces. We will see that this definition is **not** an extension of classical independence but rather an analogy. One exercise you will do is show that freeness and commutatitivity only coincide when at least one algebra is the scalars.

Let $a \in (\mathcal{A}, \varphi)$, and let $\mathring{a} = a - \varphi(a)1$. Then $\mathring{a}$ is centered, and $a = \mathring{a} + \varphi(a)1$. This trick is very useful in performing computations that involve freeness.

Example 3.3. Suppose a and b are *–free elements of some $(\mathcal{A}, \varphi)$ i.e. $\{a\}$ and $\{b\}$ are free. Let A and B be the unital $*$-algebras generated by A and B respectively. Let's compute $\varphi(ab)$. Since $\mathring{a} \in A$ and $\mathring{b} \in B$, it follows that

$$\varphi(ab) = \varphi((\mathring{a} + \varphi(a)1)(\mathring{b} + \varphi(b)1)) = \varphi(\mathring{a}\mathring{b}) + \varphi(\mathring{a})\varphi(b) + \varphi(a)\varphi(\mathring{b}) + \varphi(a)\varphi(b)$$

where we used linearity. The middle two terms vanish because $\mathring{a}$ and $\mathring{b}$ are centered, and the first term vanishes from the definition of freeness. Therefore, one gets the very pleasing result that

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

One should note that such factorizations do not always happen. For instance, one problem asks you to show that if a and b are $*$–free, then

$$\varphi(a^*b^*ab) = |\varphi(a)|^2\varphi(b^*b) + \varphi(a^*a)|\varphi(b)|^2 - |\varphi(a)|^2 \cdot |\varphi(b)|^2$$

As should be apparent from the above example, the procedure of rewriting elements into their centered from allows one to compute φ on any product of free (not necessarily centered) elements. This lets us deduce the following extremely important fact:

Fact 3.4. Let $(\mathcal{A}, \varphi)$ be a $*$–probability space, $\{A_i\}_{i\in I}$ be a collection of freely independent *–subalgebras of $\mathcal{A}$, and let $\mathcal{B} = \underset{i\in I}{*} A_i$ be the unital algebra generated by the A_i (i.e. the **free product** of the A_i). Then $\varphi(b)$ for any $b \in \mathcal{B}$ is uniquely determined by φ restricted to each A_i.

We conclude this set of notes by examining basic, yet important, examples of freeness.

Example 3.5. Let Γ be a discrete group, and $(L(\Gamma), \tau)$ be the group von Neumann algebra with its canonical trace. Suppose that Γ is the free product of a collection of subgroups $\{\Gamma_i\}_{i\in I}$. We claim that collection of group von Neumann algebras $\{L(\Gamma_i)\}_{i\in I}$ is $*$-free, meaning that

$$(L(\mathop{*}_{i\in I}\Gamma_i), \tau) = \mathop{*}_{i\in I}(L(\Gamma_i), \tau)$$

Let $i_1, \cdots, i_n \in I$ with $i_k \neq i_{k+1}$ for $1 \leq k \leq n-1$ and for each i_k, let x_{i_k} be a centered element in $L(\Gamma_{i_k})$. We need to show that $x_{i_1}x_{i_2}\cdots x_{i_n}$ is centered. Each x_{i_k} is a strong operator limit of a bounded net of elements in the span of $\{u_{g_{i_k}} : g_{i_k} \in \Gamma_{i_k} \setminus \{e\}\}$. As multiplication is strongly continuous on bounded sets, we may approximate each x_{i_k} in this manner. By doing so and expanding out the terms, it suffices to show that $u_{g_{i_1}}\cdots u_{g_{i_n}}$ is centered when each $g_{i_k} \in \Gamma_{i_k} \setminus \{e\}$. Note that $u_{g_{i_1}}\cdots u_{g_{i_n}} = u_{g_{i_1}\cdots g_{i_n}}$. Since $i_k \neq i_{k+1}$ for $1 \leq k \leq n-1$, it follows that the word $g_{i_1}\cdots g_{i_n}$ is reduced in Γ so it can not be the identity so the corresponding unitary is centered. Thus $u_{g_{i_1}}\cdots u_{g_{i_n}}$ is centered which completes the proof.

As a particular sub-example to the above, we see that $L(\mathbb{F}_n) = \mathop{*}_{i=1}^{n} L(\mathbb{Z})$ where $\mathbb{F}_n$ is the free group on n generators and both sides are equipped with their canonical traces.

Our final example concerns the creation and annihilation operators that were defined above.

Example 3.6. Let $\{\xi_i\}_{i\in I}$ be an orthonormal set in a Hilbert space $\mathcal{H}$. We claim that the elements $\ell(\xi_i)_{i\in I}$ are free in $\mathcal{B}(\mathcal{F}(\mathcal{H}))$ under the vacuum state φ. Fixing i for a moment, and noting that $\ell(\xi_i)^*\ell(\xi_i) = 1$, it follows that $\mathrm{W}^*(\ell(\xi_i))$ is the strong closure of the span of $\{\ell(\xi_i)^n(\ell(\xi_i)^*)^m \,|\, n, m \geq 0\}$. The only element in this set that is not centered is when $n = m = 0$, therefore the centered elements in $\mathrm{W}^*(\ell(\xi_i))$ are a limit of a bounded net of elements in the span of $\{\ell(\xi_i)^n(\ell(\xi_i)^*)^m \,|\, n, m \geq 0\} \setminus \{1\}$.

Let $i_1, \cdots, i_n \in I$ with $i_k \neq i_{k+1}$ for $1 \leq k \leq n-1$ and for each i_k, let x_{i_k} be a centered element in $\mathrm{W}^*(\ell(\xi_{i_k}))$. We need to show that $x_{i_1}\cdots x_{i_n}$ is centered. Arguing as in the previous example, it is sufficient to verify that $x_{i_1}\cdots x_{i_n}$ is centered when each $x_{i_k} = \ell(\xi_{i_k})^n(\ell(\xi_{i_k})^*)^m$ with n and m not both zero. By orthogonality, $\ell(\xi_{i_k})^*\ell(\xi_{i_{k+1}}) = 0$. This means that there are only three ways for $x_{i_1}\cdots x_{i_n}$ to be nonzero

(i) Each x_{i_k} is a product of only creation operators

(ii) Each x_{i_k} is a product of only annihilation operators

(iii) There exists a k satisfying $2 \leq k \leq n-1$ so that $x_{i_k} = \ell(\xi_{i_k})^n(\ell(\xi_{i_k})^*)^m$ with n and m nonzero. For all $j > k$, x_{i_j} is a product of only annihilation operators and for all $j < k$, x_{i_j} is a product of only creation operators.

In the last two cases, $x_{i_1}\cdots x_{i_n}\Omega = 0$. In the first case, $x_{i_1}\cdots x_{i_n}\Omega$ is orthogonal to Ω. This establishes the freeness.

The above example implies that the elements $(\ell(\xi_i)+\ell(\xi_i)^*)_{i\in I}$ are free in $\mathcal{B}(\mathcal{F}(\mathcal{H}))$, meaning that we have a concrete spatial representation of a free family of semicircular elements. It turns out φ is faithful on each $W^*(\ell(\xi_i) + \ell(\xi_i)^*)$. From our calculation of the law of

$\ell(\xi_i) + \ell(\xi_i)^*$, it follows that $(W^*(\ell(\xi_i) + \ell(\xi_i)^*), \varphi) \cong L^\infty([-2, 2], d\mu)$ with μ the semicircular law. As μ is diffuse, it follows that $L^\infty([-2, 2], d\mu) \cong (L(\mathbb{Z}), \tau)$. This analysis implies that $(W^*(\{\ell(\xi_i) + \ell(\xi_i)^*\}_{i \in I}), \varphi) \cong (L(\mathbb{F}_I), \tau)$, giving us an alternative spatial representation of a free group factor. This latter representation of a free group factor is often more preferred. The last exercise assigned to you gets at a reason why.