An Introduction to Free Probability

GOALS Summer School 2022
Expository Talk
Prepared by: Jennifer Pi

1 History and Motivation

- Operator algebras are very suitable for noncommutative mathematics
  - A general trend of the time, and our time, is to make noncommutative analogs of existing mathematics; this has some basis in quantum physics
- In the 1980's, Voiculescu wanted to study the free group factors $L(\mathbb{F}_n)$.
  - If the theory of von Neumann algebras is like "noncommutative measure theory", then the theory of II_1 factors is like non-commutative probability theory (since we have a trace τ such that $\tau(1) = 1$).

we'll see in talk 2 that this yielded some results for $L(\mathbb{F}_n)$

$\implies$ Place free products of groups into a framework of noncommutative probability

- Free probability theory = noncommutative probability theory + free independence.

Voiculescu likes to say...

$\boxed{\text{Random Mxs.}} \xleftrightarrow{\text{F.P.}} \boxed{\text{Opr Algs.}}$

Why do we keep studying it?

- Analogs of classical probability (central limit theorem, Brownian motion, entropy) have developed that can be applied to study operator algebras. A bit more on this in the second talk.
- In the 1990's, Voiculescu discovered that freeness occurs asymptotically for many random matrices. This allows for operator algebras to be modeled asymptotically by random matrices, and helps us understand some objects e.g. the free group factors. Conversely, free probability brought a conceptual approach to understanding the asymptotic eigenvalue distribution of random matrices.
- The subject itself is beautiful and multi-faceted: Nica & Speicher have a combinatorial approach to freeness, while Voiculescu's original approach is analytic.

Classical Probability (Brief Review/Definitions)

Briefly, a **probability space** is a measure space (Ω, P) such that $P(\Omega) = 1$.

A **random variable** X is a measurable function $X : \Omega \to S$, where S is another measurable space (usually $\mathbb{R}$ or $\mathbb{C}$).

Several key ideas: The **expectation** of X is $\mathbb{E}[X] = \int_\Omega X(\omega)\, dP(\omega)$, and measures the mean of X.

The **distribution** of X describes the probability of certain events involving X, i.e. it details all quantities of the form

$$P(X \in A) \text{ for all Borel } A \subseteq \Omega.$$

Usually see: $P(X \in A) = \int_A X(\omega)\, dF(\omega)$

& $F : \Omega \to \mathbb{R}$ is the distribution fxn.

Independence is a property which intuitively tells us that "two random variables X and Y are unrelated" in a specific sense.
(More on this later).

Non-Commutative Probability Spaces, Random Variables, and NC Laws

We can view classical probability spaces as an algebra of random variables along with a functional given by the expectation of these random variables.

Definition 1.1. A **noncommutative probability space** is a unital algebra A over $\mathbb{C}$ together with a linear functional $\phi : A \to \mathbb{C}$ such that $\phi(1) = 1$.
If A is a C^*-algebra and ϕ is a state, we call (A, ϕ) a C^*-probability space. If A is a W^*-algebra and ϕ is a trace, we call (A, ϕ) a W^*-probability space.

We should *think of* the linear functional ϕ as an analog to the expectation functional:

$$\phi(x) = \int_A x \qquad (= \mathbb{E}[x]).$$

Exps. 1.) $(L^\infty(\Omega, P), \mathbb{E})$ is a W^*-p.s.

2.) Set $L = \bigcap_{p>1} L^p(\Omega, P)$. Then $(M_n(L), \mathbb{E}\circ \mathrm{Tr}_n)$ is a ncps.

3.) For Γ a discrete grp, $(C^*_r(\Gamma), \mathrm{tr})$ is a C^*-ps, & $(L(\Gamma), \mathrm{tr})$ is a W^*-p.s.

Definition 1.2. A **random variable** in (A, ϕ) is an element $x \in A$. The **distribution** or **noncommutative law** of x is the linear functional $\lambda_x : \mathbb{C}[x] \to \mathbb{C}$ defined by $p(x) \mapsto \phi(p(x))$.

So, the distribution is determined by the moments $\varphi(x^k)$ $\forall k \in \mathbb{N}$. Why is this the right analog?

Fact: For classical random variables, the collection of all moments $\mathbb{E}[X^k], k \geq 1$, completely determines the probability distribution of X.

if we know $\varphi(x^k)$ $\forall k \in \mathbb{N}$, then for any poly $p = a_n x^n + \ldots + a_1 x + a_0$,

$$\varphi(p(x)) = a_n \varphi(x^n) + \ldots + a_1 \varphi(x) + a_0$$

2 Free Independence

Classical Independence: Why is it important?

Definition 2.1. Two events A and B are independent if $P(A \cap B) = P(A)P(B)$.
We call two random variables X and Y **independent** if for all $A \in \sigma(X)$ and $B \in \sigma(Y)$, A and B are independent.

- Knowing individual distributions of X and Y completely determines the joint distribution of (X, Y).

we want an analog of this notion for n.c. random variables.

Free Independence

The definition of free independence can be summed up as: subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_s$ are freely independent if "the alternating product of centered elements is centered". More formally:

Definition 2.2. Let $(\mathcal{A}, \phi)$ be a noncommutative probability space, with unital subalgebras $\mathcal{A}_1, \ldots, \mathcal{A}_s$. We say that $\mathcal{A}_1, \ldots, \mathcal{A}_s$ are **freely independent** if whenever $r \geq 2$ and $a_1, \ldots, a_r \in \mathcal{A}$ satisfy:

- $\phi(a_i) = 0$ for $i \in [r]$,
- $a_i \in \mathcal{A}_{j_i}$ with $1 \leq j_i \leq s$ for $i \in [r]$
- $j_1 \neq j_2, j_2 \neq j_3, \ldots, j_{r-1} \neq j_r$,

then we must have $\phi(a_1 \cdots a_r) = 0$.

What is "free" about this?
Recall from group theory: we call a family $(G_i)_{i \in I}$ of subgroups of a group G *free* if there are no non-trivial algebraic relations among the G_i's,
i.e. $g_1 \ldots g_n \neq e$ whenever $g_j \neq e$ for all $1 \leq j \leq n$ and $g_j \in G_{i(j)}$ with $i(j) \neq i(j+1)$ for $1 \leq j \leq n-1$.

For group von Neumann algebras: the free independence of the subalgebras generated by $(\lambda(G_i))_{i\in I}$ in $(L(G),\tau)$ is equivalent to the family of subgroups $(G_i)_{i\in I}$ being algebraically free.
Indeed, if we suppose that $(\lambda(G_i))_{i\in I}$ are freely independent, then whenever $g_1\dots g_n$ is a word satisfying the alternating condition with $g_j \neq e$ for all $1\le j\le n$, then

$$\tau(g_1\dots g_n) = 0;$$

but recall that $\tau(e) = 1$, so we conclude that $g_1\dots g_n \neq e$. On the other hand, let $(G_i)_{i\in I}$ be an algebraically free family of subgroups. Then recall that τ is defined by δ_e, so for any alternating product of elements $g_1\dots g_n$ satisfying $\tau(g_j)=0$ (equivalently $g_j\neq e$), we have

$$\tau(g_1\dots g_n) = \langle g_1\dots g_n\delta_e, \delta_e\rangle = 0 \text{ since } g_1\dots g_n \neq e.$$

Finally, note that the free independence of the sets $(\lambda(G_i))_{i\in I}$ is equivalent to the free independence of the von Neumann algebras generated by them, by definition.

Why is this the right analogue?
Knowing the individual distributions on subalgebras completely determines joint distributions:

Theorem 2.3. *Let $(\mathcal{B},\phi)$ be a noncommutative probability space. Consider unital subalgebras $\mathcal{A}_1,\dots,\mathcal{A}_2\subseteq\mathcal{B}$ which are freely independent. Let $\mathcal{A}$ be the algebra generated by $\mathcal{A}_1,\dots,\mathcal{A}_s$.*
Then $\phi|_{\mathcal{A}}$ is determined by $\phi|_{\mathcal{A}_1},\dots,\phi|_{\mathcal{A}_s}$ along with the freeness condition.

Proof/Example: Freeness Determines Joint Distribution

Theorem 2.4 (Abbreviated Version). *If unital subalgebras $\mathcal{A}_1,\dots,\mathcal{A}_2\subseteq\mathcal{B}$ are freely independent, and $\mathcal{A}=alg(\cup_{i=1}^s\mathcal{A}_i)$, then $\phi|_{\mathcal{A}}$ is determined by $\phi|_{\mathcal{A}_1},\dots,\phi|_{\mathcal{A}_s}$.*

Proof goes by induction on r, the length of words $a_1\dots a_r$.
For $r=2$, suppose $a_1\in\mathcal{A}_{i_1}$ and $a_2\in\mathcal{A}_{i_2}$ with $i_1\neq i_2$. Since the subalgebras are free,

$$\phi[(a_1-\phi(a_1)1)(a_2-\phi(a_2)1] = 0.$$

Expanding the term in brackets, we have

$$(a_1-\phi(a_1)1)(a_2-\phi(a_2)1) = a_1a_2-\phi(a_2)a_1-\phi(a_1)a_2+\phi(a_1)\phi(a_2)1.$$

Hence,

$$\phi(a_1a_2) = \phi[\phi(a_2)a_1+\phi(a_1)a_2-\phi(a_1)\phi(a_2)1] = \phi(a_1)\phi(a_2).$$

Connection to Random Matrices: Asymptotic Freeness

Let $X_1^{(N)},\dots,X_s^{(N)}$ be independent GUE random matrices in $M_N(\mathbb{C})_{\text{sa}}$.

Set an r-tuple of positive integers $m_1,\dots,m_r$ and alternating indices $i_1,\dots,i_r\in[s]$.

Consider $Y_N := ((X_{i_1}^{(N)})^{m_1} - c_{m_1} I) \cdots ((X_{i_r}^{(N)})^{m_r} - c_{m_r} I)$, where c_m is the asymptotic value of X_i^m. Then $\mathbb{E}(\mathrm{tr}(Y_N)) \to 0$.

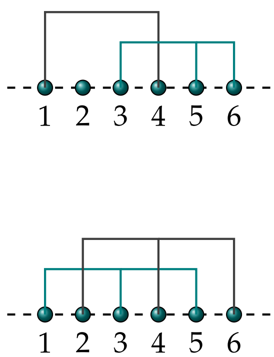
not obvious! see Ch. 2 of Mingo-Speicher if interested.

Basically, as $N \to \infty$, the matrices $X_1^{(N)}, \ldots X_s^{(N)}$ satisfy the freeness condition.

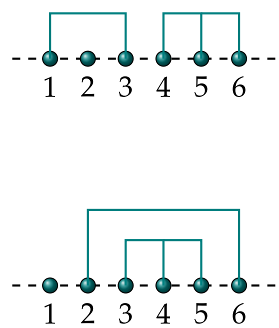
3 Free Cumulants

The given definition of freeness is sometimes hard to check, and though it does give a way to find all mixed moments, this is computationally annoying to do. We now want to give an equivalent formulation of freeness that is easier to check in practice.

Non-Crossing Partitions

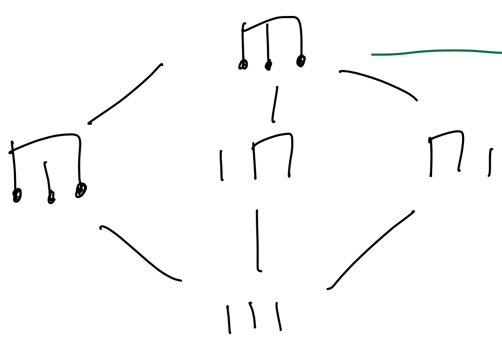


Crossing Partitions



Non-Crossing Partitions

Lattice of Non-Crossing Partitions, Example: NC(3)



Notation: 1_n for the maximal ptn

Ordered by reverse refinement!

Free Cumulants & Moment-Cumulant Formula

The **free cumulants** are n-linear functionals $\kappa_n : \mathcal{A}^n \to \mathbb{C}$ defined inductively via the **moment-cumulant formula**:

$$\phi(a_1 \cdots a_n) = \sum_{\pi \in \mathrm{NC}(n)} \kappa_\pi(a_1, \ldots, a_n),$$

where if $\pi = \{V_1, \ldots, V_r\}$, then

$$\kappa_\pi(a_1, \ldots, a_n) = \prod_{\substack{V \in \pi \\ V = (i_1, \ldots, i_\ell)}} \kappa_\ell(a_{i_1}, \ldots, a_{i_\ell}).$$

Example. In the $n = 2$ case,

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

Since $\kappa_1(a_i) = \phi(a_i)$, we get:

$$\kappa_2(a_1, a_2) = \phi(a_1 a_2) - \phi(a_1)\phi(a_2).$$

Now you try this with $n = 3$:

$$\phi(a_1a_2a_3) = \kappa_{\{(1,2,3)\}}(a_1,a_2,a_3) + \kappa_{\{(1,2),(3)\}}(a_1,a_2,a_3) + \kappa_{\{(1),(2,3)\}}(a_1,a_2,a_3) + \kappa_{\{(1,3),(2)\}}(a_1,a_2,a_3) + \kappa_{\{(1),(2),(3)\}}(a_1,a_2,a_3)$$

$$= \kappa_3(a_1,a_2,a_3) + \kappa_2(a_1,a_2)\kappa_1(a_3) + \kappa_1(a_1)\kappa_2(a_2,a_3) + \kappa_1(a_2)\kappa_2(a_1,a_3) + \kappa_1(a_1)\kappa_1(a_2)\kappa_1(a_3)$$

$$\Rightarrow \kappa_3(a_1,a_2,a_3) = \phi(a_1a_2a_3) + \phi(a_3)\big(\underbrace{\kappa_2(a_1,a_2)}_{\phi(a_1a_2)-\phi(a_1)\phi(a_2)}\big) + \phi(a_1)\underbrace{\kappa_2(a_2,a_3)}_{\phi(a_2a_3)-\phi(a_2)\phi(a_3)} + \phi(a_2)\underbrace{\kappa_2(a_1,a_3)}_{\phi(a_1a_3)-\phi(a_1)\phi(a_3)} + \phi(a_1)\phi(a_2)\phi(a_3)$$

$$\boxed{\kappa_3(a_1,a_2,a_3) = \phi(a_1a_2a_3) + \phi(a_3)\phi(a_1a_2) + \phi(a_1)\phi(a_2a_3) + \phi(a_2)\phi(a_1a_3) - 2\phi(a_1)\phi(a_2)\phi(a_3)}$$

Rephrasing Free Independence

Point: can always write $\kappa_n(\cdots)$ as a sum of $\alpha_\pi \phi_\pi(\cdots)$, $\alpha_\pi \in \mathbb{Z}$.

Definition 3.1. Let $(\mathcal{A}, \phi)$ be a non-commutative probability space. Elements $a_1, \ldots, a_s \in \mathcal{A}$ are **free** or **freely independent** if the generated unital subalgebras $\mathcal{A}_i = \mathrm{alg}(1, a_i)$ are free in $\mathcal{A}$ with respect to ϕ.

Theorem 3.2. *The random variables $a_1, \ldots, a_s \in \mathcal{A}$ are free iff all mixed cumulants of the $a_1, \ldots, a_s$ vanish. More explicitly, $a_1, \ldots, a_s$ are free iff whenever we choose $i_1, \ldots, i_n \in \{1, \ldots, s\}$ in such a way that $i_k \neq i_\ell$ for some $k, \ell \in [n]$, then $\kappa_n(a_{i_1}, \ldots, a_{i_n}) = 0$.*

Cor Constants $\mathbb{C}1$ are always free from any subalg. $\mathcal{B} \subseteq \mathcal{A}$
(if $\mathcal{A}$ is unital & has $\mathbb{C}1$!)

↳ Pf Hint: Induct on # of elts in cumulant. Use moment-cumulant formula.

4 Exploration: Free Central Limit Theorem

(Classical) Central Limit Theorem

Theorem 4.1. *Let $(a_i)_{i\in\mathbb{N}}$ be a sequence of independent, identically distributed classical random variables, with $\phi(a_i) = 0$ and $\phi(a_i^2) = \sigma^2$, and moments of all orders existing. Then,*

$$\frac{1}{\sqrt{k}}(a_1 + \cdots + a_k) =: S_k \to \mathcal{N}(0, \sigma^2) \text{ in distribution,}$$

where $\mathcal{N}(0, \sigma^2)$ is the normal distribution with mean 0 and variance σ^2.

What do we mean by "convergence in distribution"?

Definition 4.2. For random variables $(a_i)_{i\in\mathbb{N}}$, we say that $a_i \to X$ in distribution if the corresponding probability measures $\mu_{a_i} \to \mu_X$ weakly, i.e.

$$\int f(a_i)d\mu_{a_i} \to \int f(X)d\mu_X \quad \text{for all } f \in C_b(\mathbb{R}).$$

Distributions and Moments

Central to the proof is the fact that distributions are **determined by their moments**, i.e.
If μ_X has moments $\alpha_k = \int X^k d\mu_X$ for all $k \in \mathbb{N}$, and if ν has the same moments $\{\alpha_k\}_{k\in\mathbb{N}}$, then $\nu = \mu_X$.

Note: this holds when μ_X has moments of all orders!

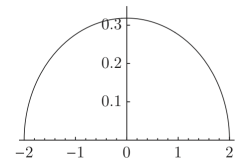
Recall that: for a noncommutative random variable $a \in \mathcal{A}$, we defined the **distribution of** a by the collection:

$\phi(p(a))$ for all $p \in \mathbb{C}[x]$. and this is completely determined by $\varphi(a^n)$ $\forall n \in \mathbb{N}$.

Definition 4.3. If $(a_k)_{k\in\mathbb{N}}$ is a sequence of noncommutative random variables with $a_k \in \mathcal{A}_k$, we say $a_k \to a \in \mathcal{A}$ in distribution if

for all $n \in \mathbb{N}$, $\phi(a_k^n) \to \phi(a^n)$ as $k \to \infty$.

Fig. 1.2 The graph of $(2\pi)^{-1}\sqrt{4-t^2}$. The $2k^{th}$ moment of the semi-circle law is the Catalan number $C_k = (2\pi)^{-1}\int_{-2}^{2} t^{2k}\sqrt{4-t^2}\,dt$.



Free Central Limit Theorem

Definition 4.4. A self-adjoint nc random variable s with odd moments $\phi(s^{2n+1}) = 0$ and even moments $\phi(s^{2n}) = \sigma^{2n} \cdot C_n$, where C_n is the nth Catalan number, is a **semi-circular element of variance** σ^2. If $\sigma = 1$, we call s a **standard semi-circular element**.

Theorem 4.5. *If $(a_i)_{i\in\mathbb{N}}$ are self-adjoint, freely independent, identically distributed nc random variables with $\phi(a_i) = 0$ and $\phi(a_i^2) = \sigma^2$, then*

$$\frac{1}{\sqrt{k}}(a_1 + \cdots + a_k) = S_k \to \mathcal{S}(\sigma^2) \text{ in distribution}$$

Proof of Free Central Limit Theorem

First, we unpack the definition of convergence in distribution:
All we have to do is compute asymptotic moments, and check that $\lim_{k\to\infty}\phi(S_k^{2n+1}) = 0$, while $\lim_{k\to\infty}\phi(S_k^{2n}) = \sigma^{2n}\cdot C_n$.

We start by expanding $\phi(S_k^n)$:

$$\phi(S_k^n) = \phi\left(\left(\frac{1}{\sqrt{k}}(a_1 + \cdots + a_k)\right)^n\right) = k^{-n/2}\sum_{i:[n]\to[k]}\phi(a_{i_1}\cdots a_{i_n}).$$

There are many possible functions $i : [n] \to [k]$ appearing in the sum above, but since the a_i's are identically distributed and (freely) independent, $\phi(a_{i_1}\cdots a_{i_n})$ only depends on the number of different indices and the number of each.

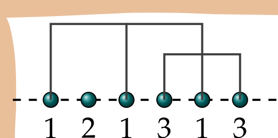
Example: $\phi(\cdot)$ only depends on shape of $\ker i$

In the classical case,

$$\begin{aligned}\phi(a_1a_2a_1a_3a_1) &= \phi(a_1^3)\phi(a_2)\phi(a_3)\\ &= \phi(a_3^3)\phi(a_1)\phi(a_2)\\ &= \phi(a_3a_1a_3a_2a_3).\end{aligned}$$

In the free and more general cases, we still have $\phi(a_{i_1}\cdots a_{i_n})$ depends only on the indices appearing and their configuration, since freeness gives us some rule for calculating the mixed moments of $a_{i_1}\cdots a_{i_n}$ in terms of the individual moments of the a_i's, which are identical. you will get some practice w/ doing mixed moments in terms of pure ones in exercises.

Notation: For a multi-index $i = (i_1, \ldots, i_n)$, we define the kernel, denoted $\ker i$ to be the partition whose blocks correspond to the different values of the indices, e.g.

the multi-index $i = (1,2,1,3,1,3)$ has kernel: 

Lemma 4.6. *When keri =kerj $= \pi$, then we have*

$$\phi(a_{i_1} \cdots a_{i_n}) = \phi(a_{j_1} \cdots a_{j_n}) =: \phi(\pi).$$

Continuing computing $\phi(S_k^n)$, we left off with

$$\phi(S_k^n) = k^{-n/2} \sum_{i:[n]\to[k]} \phi(a_{i_1} \cdots a_{i_n})$$
$$= k^{-n/2} \sum_{\pi \in \mathcal{P}(n)} \phi(\pi) \cdot |\{i : [n] \to [k] \mid \ker i = \pi\}|$$

To count the last thing, suppose π has ℓ blocks. Then we can label each block with a distinct index in $[k]$; so we have k choices for block 1, $k-1$ choices for block 2, and so on...$k - \ell + 1$ choices for block ℓ. Thus,

$$\phi(S_k^n) = k^{-n/2} \sum_{\pi \in \mathcal{P}(n)} \phi(\pi) \cdot k(k-1) \cdots (k - |\pi| + 1).$$

Note that now the number of terms in the sum does not depend on k!

We now see that many of the terms of the sum above vanish, i.e. $\phi(\pi) = 0$ for many π.
First, if π has a singleton, then $\phi(\pi) = 0$ since $\phi(a_i) = 0$.
So we only consider π with block size ≥ 2. This means $|\pi| \leq n/2$. On the other hand, note that $k \cdot (k-1) \cdots (k - |\pi| + 1)$ is asymptotically like $k^{|\pi|}$, and

$$\boxed{\lim_{k\to\infty} \frac{k^{|\pi|}}{k^{n/2}} = \begin{cases} 1 & \text{if } |\pi| = n/2 \\ 0 & \text{if } |\pi| < n/2. \end{cases}}$$

$\implies$ Asymptotically, any term $\phi(\pi)$ with $|\pi| < n/2$ vanishes!
We only need to consider pairings $\pi \in \mathcal{P}_2(n)$.

So far, we've shown

$$\lim_{k\to\infty} \phi(S_k^n) = \sum_{\pi \in \mathcal{P}_2(n)} \phi(\pi).$$

As a direct corollary, we have that all odd asymptotic moments are zero!

$$lim_{k\to\infty} \phi(S_k^{2n+1}) = 0.$$

Now we only have to compute $\lim_{k\to\infty} S_k^{2n}$**.**
We'll return to this after y'all do some exercises!