Free Probability and Free Entropy
(A High-Level Introduction)

Jennifer Pi

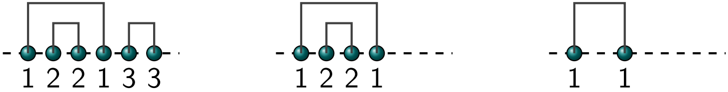
University of California, Irvine

GOALS Summer School 2022

What did you discover from Exercise 1?

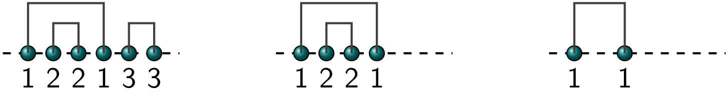
What did you discover from Exercise 1?

We see that for $\pi \in \mathcal{P}_2(2n)$, we have $\phi(\pi) = \sigma^{2n}$ if and only if we can successively remove pairs of matching random variables until we end with a single pair.



What did you discover from Exercise 1?

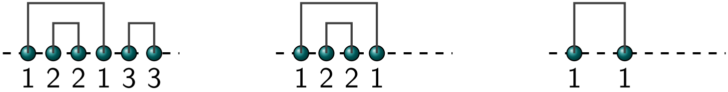
We see that for $\pi \in \mathcal{P}_2(2n)$, we have $\phi(\pi) = \sigma^{2n}$ if and only if we can successively remove pairs of matching random variables until we end with a single pair.



Actually, this occurs if and only if π is non-crossing.

What did you discover from Exercise 1?

We see that for $\pi \in \mathcal{P}_2(2n)$, we have $\phi(\pi) = \sigma^{2n}$ if and only if we can successively remove pairs of matching random variables until we end with a single pair.



Actually, this occurs if and only if π is non-crossing.

Otherwise, $\phi(\pi) = 0$.

Finishing Up the Proof

At the end of the first lecture, we knew:

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

For even moments, since only non-crossing partitions π give a non-zero contribution, we have

$$\lim_{k\to\infty} \phi(S_k^{2n}) = \sum_{\pi\in NC_2(2n)} \phi(\pi) = \sigma^{2n} \cdot |NC_2(2n)|.$$

Since $|NC_2(2n)| = C_n$, the nth Catalan number, we are done!

Theorem (Free Central Limit Theorem)

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

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

Aside: Interpolated Free Group Factors

Definition (Interpolated Free Group Factors)

Let $(\mathcal{M}, \tau)$ be a tracial von Neumann algebra and let $\mathcal{R}$ be a copy of the hyperfinite II_1 factor in $\mathcal{M}$. Also let $\omega = \{X^t \mid t \in T\}$ be a semicircular family such that $\mathcal{R}$ and ω are free.

Then for $1 < r \leq \infty$, we define $L(\mathbb{F}_r)$ as the factor $(\mathcal{R} \cup \{p_t X^t p_t \mid t \in T\})''$, where $p_t \in \mathcal{R}$ are projections satisfying $r = 1 + \sum_{t \in T} \tau(p_t)^2$.

Aside: Compressions/Amplifications

For a II_1 factor $\mathcal{M}$ and $0 < t < 1$, we define the *compression* of $\mathcal{M}$ as

$$\mathcal{M}_t \cong p\mathcal{M}p, \qquad \text{for any } p \in \mathcal{P}(\mathcal{M}),\ \tau(p) = t.$$

We extend this notion to *amplifications of* $\mathcal{M}$ by taking tensors with matrix algebras; for $1 < t < \infty$, if we write $t = n \cdot \ell$, where $0 < \ell < 1$ and $n \in \mathbb{N}$, then we define

$$\mathcal{M}_t \cong p\mathcal{M}p \otimes M_n(\mathbb{C}) \cong M_n(p\mathcal{M}p), \qquad \text{for any } p \in \mathcal{P}(\mathcal{M}),\ \tau(p) = \ell.$$

Definition (Fundamental Group)

The **fundamental group** of a II_1 factor $\mathcal{M}$ is $\{t \in \mathbb{R}_+ \ : \ \mathcal{M}_t \cong \mathcal{M}\}$. It is a multiplicative subgroup of $\mathbb{R}_+$.

Aside: Free Group Factor (FGF) Alternative

Theorem (Dykema '92; Radulescu '94)

One of the following two statements must be true:

1. $L(\mathbb{F}_r) \cong L(\mathbb{F}_s)$ *for all* $1 < r, s \leq \infty$, *and the fundamental group of* $L(\mathbb{F}_r)$ *is* $\mathbb{R}_+$ *for all* $1 < r \leq \infty$.
2. $L(\mathbb{F}_r) \ncong L(\mathbb{F}_s)$ *for all* $1 < r, s \leq \infty$, *and the fundamental group of* $L(\mathbb{F}_r)$ *is* $\{1\}$ *for all* $1 < r \leq \infty$.

Addition Formula

$L(\mathbb{F}_r) * L(\mathbb{F}_s) \cong L(\mathbb{F}_{r+s})$, for $1 < r, s \leq \infty$.

Compression Formula

$L(\mathbb{F}_r)_t \cong L\left(\mathbb{F}\left(1 + \frac{r-1}{t^2}\right)\right)$, for $1 < r \leq \infty$, $0 < t < \infty$.

Microstates Free Entropy

Motivation

What does the word "entropy" bring to mind?

Motivation

What does the word “entropy” bring to mind?

In classical information theory, entropy measures the amount of “information” or “uncertainty” in a random variable.

Boltzmann’s formula from physics says the entropy of a “macrostate” is obtained by counting how many “microstates” correspond to that “macrostate”.

Motivation

What does the word "entropy" bring to mind?

In classical information theory, entropy measures the amount of "information" or "uncertainty" in a random variable.

Boltzmann's formula from physics says the entropy of a "macrostate" is obtained by counting how many "microstates" correspond to that "macrostate".

Classical entropy of a distribution μ on $\mathbb{R}^n$ with density ρ is given by

$$H(\mu) := -\int_{\mathbb{R}} \rho(x_1, \ldots, x_n) \log \rho(x_1, \ldots, x_n)\ dx_1 \cdots dx_n.$$

Idea of $\chi(X)$

Let $X = (x_1, \ldots, x_n)$ be a tuple of self-adjoint elements of (M, τ).

Our "microstates" will be given by tuples of matrices in $M_N(\mathbb{C})_{sa}$ that approximate the mixed moments of X.

"Counting the number of microstates" will be taking the Lebesgue measure in $M_N(\mathbb{C})_{sa} \cong \mathbb{C}^{N^2} \cong \mathbb{R}^{2N^2}$.

We then do some appropriate normalization and want to take limits as $N \to \infty$, and as the level of approximation improves.

Definition of $\chi(X)$

Let (M, τ) be a tracial W^*-probability space and let $x_1, \dots, x_n$ be an n-tuple of self-adjoint elements in M. The *set of approximating microstates* is:

$$\begin{aligned}&\Gamma(x_1, \dots, x_n; N, r, \epsilon)\\ &:= \{(A_1, \dots, A_n) \in M_N(\mathbb{C})^n_{sa} : |\mathrm{tr}(A_{i_1} \cdots A_{i_k}) - \tau(x_{i_1} \cdots x_{i_k})| \leq \epsilon\\ &\text{for all } 1 \leq i_1, \dots, i_k \leq n, 1 \leq k \leq r\}.\end{aligned}$$

Further define:

$$\chi(x_1, \dots, x_n; r, \epsilon) := \limsup_{N \to \infty} \left(\frac{1}{N^2} \log(\Lambda(\Gamma(x_1, \dots x_n; N, r, \epsilon))) + \frac{n}{2} \log(N) \right).$$

The *(microstates) free entropy* $\chi(x_1, \dots, x_n)$ is

$$\chi(x_1, \dots, x_n) := \lim_{\substack{r \to \infty \\ \epsilon \to 0}} \chi(x_1, \dots, x_n; r, \epsilon).$$

Applications to Operator Algebras

Theorem

Let (M, τ) be a tracial von Neumann algebra generated by self-adjoint $x_1, \ldots, x_n$. Assume that $\chi(x_1, \ldots, x_n) > -\infty$. Then

1. *(Voiculescu '96) M does not have property Γ. In particular, M is a factor.*
   M does not have property Γ if for any bounded sequence (t_k) s.t. $\|[x, t_k]\|_2 \to 0$ for all $x \in M$, we have $\|t_k - \tau(t_k)1\|_2 \to 0$ (every central sequence is trivial).

Applications to Operator Algebras

Theorem

Let (M, τ) be a tracial von Neumann algebra generated by self-adjoint $x_1, \ldots, x_n$. Assume that $\chi(x_1, \ldots, x_n) > -\infty$. Then

1. *(Voiculescu '96) M does not have property Γ. In particular, M is a factor.*
   M does not have property Γ if for any bounded sequence (t_k) s.t. $\|[x, t_k]\|_2 \to 0$ for all $x \in M$, we have $\|t_k - \tau(t_k)1\|_2 \to 0$ (every central sequence is trivial).
2. *(Voiculescu '96) M does not have a Cartan subalgebra.*
   A Cartan subalgebra is a maximal abelian subalgebra whose normalizer generates M.

Applications to Operator Algebras

Theorem

Let (M, τ) be a tracial von Neumann algebra generated by self-adjoint $x_1, \ldots, x_n$. Assume that $\chi(x_1, \ldots, x_n) > -\infty$. Then

1. *(Voiculescu '96) M does not have property Γ. In particular, M is a factor.*
   M does not have property Γ if for any bounded sequence (t_k) s.t. $\|[x, t_k]\|_2 \to 0$ for all $x \in M$, we have $\|t_k - \tau(t_k)1\|_2 \to 0$ (every central sequence is trivial).
2. *(Voiculescu '96) M does not have a Cartan subalgebra.*
   A Cartan subalgebra is a maximal abelian subalgebra whose normalizer generates M.
3. *(Ge '98) M is prime.*
   M is prime if it cannot be decomposed as $M = M_1 \bar{\otimes} M_2$ for II_1 factors M_1, M_2.

Non-Microstates Free Entropy

Motivation

Another measure of the amount of "information" a random variable carries is the classical Fisher information.

Classically, entropy can be recovered through an appropriate integral of Fisher information.

This approach provides a more algebraic flavor and avoids the difficulty of finding the size of microstate spaces.

We will skip the classical formulation on this one!

Non-Commutative Derivatives

Definition

Define the *partial non-commutative derivatives* ∂_i as linear mappings

$$\partial_i : \mathbb{C}\langle X_1, \ldots, X_n\rangle \to \mathbb{C}\langle X_1, \ldots, X_n\rangle \otimes \mathbb{C}\langle X_1, \ldots X_n\rangle \text{ by}$$

$$\partial_i 1 = 0, \qquad \partial_i X_j = \delta_{ij} 1 \otimes 1 \quad \text{for } j = 1, \ldots, n,$$

and by the Leibniz rule:

$$\partial_i(P_1 P_2) = \partial_i(P_1) \cdot 1 \otimes P_2 + P_1 \otimes 1 \cdot \partial_i(P_2) \qquad \text{for } P_1, P_2 \in \mathbb{C}\langle X_1, \ldots, X_n\rangle .$$

Exercise: What is ∂_i on monomials? Compute

$$\partial_i(X_{i(1)} \cdots X_{i(m)}) =$$

Non-Commutative Derivatives

Definition

Define the *partial non-commutative derivatives* ∂_i as linear mappings

$$\partial_i : \mathbb{C}\langle X_1, \dots, X_n \rangle \to \mathbb{C}\langle X_1, \dots, X_n \rangle \otimes \mathbb{C}\langle X_1, \dots X_n \rangle \text{ by}$$

$$\partial_i 1 = 0, \qquad \partial_i X_j = \delta_{ij} 1 \otimes 1 \quad \text{for } j = 1, \dots, n,$$

and by the Leibniz rule:

$$\partial_i(P_1 P_2) = \partial_i(P_1) \cdot 1 \otimes P_2 + P_1 \otimes 1 \cdot \partial_i(P_2) \qquad \text{for } P_1, P_2 \in \mathbb{C}\langle X_1, \dots, X_n \rangle .$$

Exercise: What is ∂_i on monomials? Compute

$$\partial_i(X_{i(1)} \cdots X_{i(m)}) = \sum_{k=1}^{m} \delta_{i,i(k)} X_{i(1)} \cdots X_{i(k-1)} \otimes X_{i(k+1)} \cdots X_{i(m)}.$$

∂_i as operators on $L^2(M)$

Recall for $x \in M$ we denote $\|x\|_2^2 = \tau(x^*x)$, and $L^2(M)$ is the closure of M under this 2-norm.

If $x_1, \ldots, x_n \in M_{\text{sa}}$, we can consider the operators ∂_i as derivatives on $\mathbb{C}\langle x_1, \ldots, x_n\rangle \subseteq M$ according to

$$\begin{array}{ccc} \mathbb{C}\langle X_1, \ldots X_n\rangle & \xrightarrow{\partial_i} & \mathbb{C}\langle X_1, \ldots, X_n\rangle \otimes \mathbb{C}\langle X_1, \ldots, X_n\rangle \\ \downarrow{\scriptstyle eval} & & \downarrow{\scriptstyle eval} \\ \mathbb{C}\langle x_1, \ldots x_n\rangle & \longrightarrow & \mathbb{C}\langle x_1, \ldots, x_n\rangle \otimes \mathbb{C}\langle x_1, \ldots, x_n\rangle \end{array}$$

(if the evaluation map is an algebra isomorphism).

Motivation for ξ_i, the conjugate variables

As an operator on $L^2(x_1, \ldots, x_n)$, ∂_i is unbounded, but we want them to be "nice", i.e. closable.
$\implies \partial_i^*$ should be densely defined, i.e. $1 \otimes 1 \in D(\partial_i^*)$.

If $1 \otimes 1 \in D(\partial_i^*)$, then set $\xi_i := \partial_i^*(1 \otimes 1)$. Then we have the following relation:

$$\begin{aligned} \tau(\xi_i P(x_1, \ldots, x_n)) &= \langle \partial_i^*(1 \otimes 1) P(\overline{x}), 1 \rangle \\ &= \langle 1 \otimes 1, \partial_i P(\overline{x}) \rangle \\ &= \tau \otimes \tau(\partial_i P(\overline{x})). \end{aligned}$$

Free Fisher Information

Let $x_1, \ldots, x_n$ be self-adjoint elements of (M, τ).

1. We say $\xi_1, \ldots, \xi_n$ form a **conjugate system** for $x_1, \ldots, x_n$ if for all i, $\xi_i \in L^2(x_1, \ldots, x_n)$, and they satisfy the **conjugate relations:** for all $P \in \mathbb{C}\langle X_1, \ldots, X_n \rangle$,

$$\tau(\xi_i P(x_1, \ldots, x_n)) = \tau \otimes \tau((\partial_i P)(x_1, \ldots, x_n)).$$

2. The **free Fisher information** of $x_1, \ldots, x_n$ is defined by

$$\Phi^*(x_1, \ldots, x_n) = \begin{cases} \sum_{i=1}^n \|\xi_i\|_2^2, & \text{if } \xi_1, \ldots, \xi_n \text{ is a conjugate system for } x_1, \ldots, x_n \\ +\infty, & \text{if no conjugate system exists.} \end{cases}$$

Additivity of Φ^* is Equivalent to Freeness

Theorem

Consider $x_1, \ldots, x_n, y_1, \ldots, y_m$ self-adjoint elements of the same tracial von Neumann algebra (M, τ).
Then $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_m\}$ are free if and only if

$$\Phi^*(x_1, \ldots, x_n, y_1, \ldots, y_m) = \Phi^*(x_1, \ldots, x_n) + \Phi^*(y_1, \ldots, y_m).$$

Proof uses a formulation in terms of cumulants characterizing when $\{\xi_i\}$ is a conjugate system for $\{x_i\}$.

Unification Problem

It was an open question for quite some time whether $\chi(X) = \chi^*(X)$. Since the result $\mathrm{MIP}^* = \mathrm{RE}$ gave a negative answer to the Connes Embedding Problem (this means there exist II_1 factors that do not embed into any ultrapower of $\mathcal{R}$), we now know there are cases where $\chi^*(X) < \chi(X)$.

It is still open whether or not $\chi(X) = \chi^*(X)$ on all $\mathcal{R}^{\mathcal{U}}$-embeddable tracial von Neumann algebras.

It is also open whether or not $\chi(X)$ is a von Neumann algebra invariant, i.e. if $W^*(X) = W^*(Y)$, then is it true that $\chi(X) = \chi(Y)$?

Brief Mention: Other Notions of Dimension

- Hayes' 1-bounded entropy
  - defined via covering numbers of the microstate spaces.
  - known to be a von Neumann algebra invariant.
- Charlesworth and Nelson's free Stein dimension
  - algebraic flavor; defined by taking dimension of an appropriate space of derivations.
  - known to be a $*$-algebra invariant.
- Jekel's free entropy for types in model theory
  - generalizes Hayes' 1-bounded entropy to expressions that involve sup and inf in addition to the trace polynomials.
  - known to be a von Neumann algebra invariant.

Takeaways

What can you take away from this expository talk?

What can you take away from this expository talk?

- What is free independence?

- (Vaguely) What is free entropy, and why do people try to study it?

- Free Probability, Operator Algebras, Random Matrices

What can you take away from this expository talk?

- What is free independence?
  - Definition: "alternating centered moments vanish", or "all mixed cumulants vanish".
  - A rule for determining joint moments from pure ones.
- (Vaguely) What is free entropy, and why do people try to study it?
- Free Probability, Operator Algebras, Random Matrices

What can you take away from this expository talk?

- What is free independence?
  - Definition: “alternating centered moments vanish”, or “all mixed cumulants vanish”.
  - A rule for determining joint moments from pure ones.
- (Vaguely) What is free entropy, and why do people try to study it?
  - 2 main versions: microstates and non-microstates.
  - Hope to define basic invariants that can tell von Neumann algebras apart.
- Free Probability, Operator Algebras, Random Matrices

What can you take away from this expository talk?

- What is free independence?
  - Definition: “alternating centered moments vanish”, or “all mixed cumulants vanish”.
  - A rule for determining joint moments from pure ones.
- (Vaguely) What is free entropy, and why do people try to study it?
  - 2 main versions: microstates and non-microstates.
  - Hope to define basic invariants that can tell von Neumann algebras apart.
- Free Probability, Operator Algebras, Random Matrices
  - Many applications to uncovering the structure of tracial von Neumann algebras. One famous example is the role of free probability in proving the free group factor alternative.
  - Many “nice” classes of random matrices are asymptotically free.

What can you take away from this expository talk?

- What is free independence?
  - Definition: "alternating centered moments vanish", or "all mixed cumulants vanish".
  - A rule for determining joint moments from pure ones.
- (Vaguely) What is free entropy, and why do people try to study it?
  - 2 main versions: microstates and non-microstates.
  - Hope to define basic invariants that can tell von Neumann algebras apart.
- Free Probability, Operator Algebras, Random Matrices
  - Many applications to uncovering the structure of tracial von Neumann algebras. One famous example is the role of free probability in proving the free group factor alternative.
  - Many "nice" classes of random matrices are asymptotically free.
- There is more to explore!

Thank you!

Hope you learned something from this! Ask me questions anytime!