

# A permutation model for free random variables

Ion Nechita

Institut Camille Jordan, Université Lyon 1

*joint work with Florent Benaych-Georges (Paris 6)*

Ottawa, September 3, 2008

Freeness and random matrices

The model of permutations

A classical probability analogue

Some combinatorics

Conclusion & perspectives

# Voiculescu's definition

Free probability theory

=

Non-commutative probability theory

+

Free independence

# Non-commutative probability spaces

## Definition

A non-commutative probability space (*ncps*) is a couple  $(\mathcal{A}, \varphi)$  where  $\mathcal{A}$  is a unital algebra and  $\varphi : \mathcal{A} \rightarrow \mathbb{C}$  is a linear functional with  $\varphi(1) = 1$ .

Elements  $x \in \mathcal{A}$  are called random variables.

## Examples

- ▶  $(L^\infty(\Omega, \mathbb{P}), \mathbb{E})$ ;
- ▶  $L^{\infty-}(\Omega, \mathbb{P}) = \bigcap_{1 \leq p < \infty} L^p(\Omega, \mathbb{P})$ ;
- ▶  $\mathcal{M}_n(\mathbb{C})$  with the normalized trace  $\text{tr}_n(A) = \frac{1}{n} \text{Tr}(A)$ ;
- ▶ The group algebra  $\mathbb{C}[G]$  with the state

$$\varphi \left( \sum_{g \in G} x_g g \right) = x_e,$$

where  $e$  is the neutral element of  $G$ .

# Free independence

## Definition

Let  $(\mathcal{A}, \varphi)$  be a ncps. Unital subalgebras  $\mathcal{A}_1, \dots, \mathcal{A}_n, \dots$  are called **freely independent** (or **free**), if

$$\varphi(a_1 a_2 \cdots a_k) = 0$$

whenever we have

- ▶  $a_j \in \mathcal{A}_{i(j)}$  for all  $j = 1, \dots, k$ ;
- ▶  $\varphi(a_j) = 0$  for all  $j = 1, \dots, k$ ;
- ▶  $i(1) \neq i(2), i(2) \neq i(3), \dots, i(k-1) \neq i(k)$ .

Random variables  $(x_j)$  in  $\mathcal{A}$  are called free if their generated unital algebras are free.

# Free independence

Free independence is a rule for computing **mixed moments**.

## Examples

- ▶ If  $a$  and  $b$  are free random variables, then

$$\varphi [(a - \varphi(a)1)(b - \varphi(b)1)] = 0,$$

which implies

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

- ▶ Similarly, if the families  $\{a_1, a_2\}$  and  $\{b_1, b_2\}$  are free, then

$$\begin{aligned}\varphi(a_1 b_1 a_2 b_2) &= \varphi(a_1 a_2) \varphi(b_1) \varphi(b_2) + \varphi(a_1) \varphi(a_2) \varphi(b_1 b_2) \\ &\quad - \varphi(a_1) \varphi(a_2) \varphi(b_1) \varphi(b_2).\end{aligned}$$

# Convergence in distribution and the free CLT

## Definition

Let  $(\mathcal{A}_n, \varphi_n)_{n \in \mathbb{N}}$  and  $(\mathcal{A}, \varphi)$  be ncps and consider random variables  $a_n \in \mathcal{A}_n$  and  $a \in \mathcal{A}$ . We say that  $a_n$  **converges in distribution** towards  $a$ , and we write  $a_n \xrightarrow{d} a$ , if  $\lim_{n \rightarrow \infty} \varphi_n(a_n^k) = \varphi(a^k)$  for all  $k \geq 1$ .

## Theorem (free Central Limit Theorem)

Let  $(a_n)_n$  be a sequence of free, identically distributed random variables such that  $\varphi(a_n) = 0$  and  $\varphi(a_n^2) = 1$ . Then

$$\frac{a_1 + \cdots + a_n}{\sqrt{n}} \xrightarrow{d} s,$$

where  $s$  is a standard **semicircular** random variable.  $s$  has distribution

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

# Gaussian random matrices

## Definition

A **Gaussian random matrix** is a matrix  $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C})$  with random elements such that

- ▶  $A$  is self-adjoint:  $a_{ij} = \bar{a}_{ji}$  for all  $i, j$ ;
- ▶  $\{a_{ii}\}_{1 \leq i \leq n}$  are independent real Gaussian random variables with mean 0 and variance  $1/n$ ;
- ▶  $\{\Re a_{ij}\}_{1 \leq i < j \leq n}$  and  $\{\Im a_{ij}\}_{1 \leq i < j \leq n}$  are independent real Gaussian random variables with mean 0 and variance  $1/2n$ .

Gaussian random matrices can be seen as random variables in the ncps  $(\mathcal{M}_n(L^{\infty-}(\Omega, \mathbb{P})), \text{tr} \otimes \mathbb{E})$ , where

$$\text{tr} \otimes \mathbb{E}(A) = \mathbb{E}[\text{tr}(A)] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[a_{ii}].$$



# Voiculescu's theorem

## Theorem

Let  $(A_n)$  and  $(B_n)$  be independent sequences of Gaussian random matrices. Then

$$(A_n, B_n) \xrightarrow{d} (s_1, s_2),$$

where  $s_1$  and  $s_2$  are free standard semicircular random variables. In particular, we say that independent Gaussian random matrices are *asymptotically free*.

How about non-random (**deterministic**) matrices ?

Can one construct deterministic matrices  $(A_n, B_n, C_n, \dots)$  which are asymptotically free ?

P. Biane ('95): Yes, on the group algebra of the symmetric group.

## The set-up

Consider the group  $\mathcal{S}$  of finitely supported permutations on the set of nonnegative integers  $\mathbb{N} = \{0, 1, \dots\}$ .

The ncps we shall work with is the group algebra  $\mathbb{C}[\mathcal{S}]$  together with its canonical trace

$$\varphi \left( \sum_{\sigma} x_{\sigma} \sigma \right) = x_e,$$

where  $e$  is the identity permutation.

For all  $r, n \geq 1$  and  $t \in [0, \infty)$ , define the random variables

$$M_r(n, t) = \frac{1}{n^{r/2}} \sum_{\substack{\underbrace{(0a_1a_2 \cdots a_r)}_{\text{designs the cycle}} \\ 0 \rightarrow a_1 \rightarrow a_2 \rightarrow \cdots \rightarrow a_r \rightarrow 0}} ,$$

where the sum runs over all  $r$ -uplets  $(a_1, \dots, a_r)$  of pairwise distinct integers of  $[1, nt]$ .

# The result

## Theorem

The non-commutative distribution of the family  $(M_r(n, t))_{r \geq 1, t \in [0, +\infty)}$  converges, as  $n$  goes to infinity, to the one of a family  $(M_r(t))_{r \geq 1, t \in [0, +\infty)}$  such that

- ▶  $(M_1(t))_{t \in [0, +\infty)}$  is a *free Brownian motion* (Biane's result);
- ▶ for all  $r, t$ , one has

$$M_r(t) = t^{\frac{r}{2}} U_r(t^{-1/2} M_1(t)),$$

where the  $U_r$ 's are the *Chebyshev polynomials of second kind*.

# The free Brownian motion

## Definition

Let  $(\mathcal{A}, \varphi)$  be a ncps. A family of random variables  $(B_t)_{t \geq 0}$  is called a **free Brownian motion** if

- ▶  $B_0 = 0$ ;
- ▶ For all  $s \leq t$ ,  $B_t - B_s$  is free with the unital algebra generated by  $\{B_u, u \leq s\}$ ;
- ▶ For all  $s \leq t$ ,  $B_t - B_s$  has a semicircular distribution with mean 0 and variance  $t - s$ .

## Chebyshev polynomials of second kind

- ▶  $U_0(x) = 1$ ,  $U_1(x) = x$ ,  $U_2(x) = x^2 - 1$ ,  $U_3(x) = x^3 - 2x$ , etc.
- ▶  $U_n$  is a degree  $n$  polynomial defined by

$$U_n(2 \cos \theta) = \frac{\sin(n+1)\theta}{\sin \theta}, \quad \forall n \geq 0.$$

- ▶ They satisfy the recurrence relation

$$U_1(x)U_n(x) = U_{n-1}(x) + U_{n+1}(x), \quad \forall n \geq 1.$$

- ▶ They are orthogonal on  $[-2, 2]$  with respect to the semicircular (!) weight

$$w(x) = \frac{1}{2\pi} \sqrt{4 - x^2}.$$

## Asymptotically free deterministic matrices

Fix  $r = 1$  and consider the elements

$$A(n) = \frac{1}{\sqrt{n}} \sum_{a=1}^n (0a) \quad \text{and} \quad B(n) = \frac{1}{\sqrt{n}} \sum_{b=n+1}^{2n} (0b).$$

By the main theorem, the family  $(A(n), B(n))$  converges in distribution to a free family  $(s_1 = B_1, s_2 = B_2 - B_1)$  of standard semicircular elements. Hence,  $A(n)$  and  $B(n)$  are **asymptotically free**.

$A(n)$  and  $B(n)$  act by right multiplication on the group algebra  $\mathbb{C}[S]$ . The (finite-dimensional) matrices of these operators have the same joint distribution as  $(A(n), B(n))$ .

## Classical vs. Free probability

Free probability has been constructed in a deep analogy with classical probability theory. There is a “dictionary” between the objects of the two theories:

Classical Probability	Free Probability
classical (or tensor) independence	free independence
Gaussian distribution	semicircular distribution
(general) partitions	non-crossing partitions
classical cumulants	free cumulants, <i>etc.</i>

Does our model of permutations have a **classical-probability** analogue ?



# The classical model

Recall the definition of the random variables used in the free setting

$$M_r(n, t) = \frac{1}{n^{r/2}} \sum (0a_1 a_2 \cdots a_r)$$

**Idea:** replace permutations by sets:

$$L_r(n, t) = \frac{1}{n^{r/2}} \sum \{a_1, a_2, \dots, a_r\}$$

## The classical model

Let  $\mathcal{G}$  be the group of finite sets of positive integers endowed with the symmetric difference operation  $\Delta$ . Consider the **commutative** ncps  $(\mathbb{C}[\mathcal{G}], \psi)$ , where  $\psi$  is the canonical trace defined by

$$\psi \left( \sum_A x_A A \right) = x_\emptyset.$$

For all  $r, n \geq 1$  and  $t \in [0, \infty)$ , define the random variables

$$L_r(n, t) = \frac{1}{n^{r/2}} \sum \{a_1, a_2, \dots, a_r\},$$

where the sum runs over all  **$r$ -uplets**  $(a_1, \dots, a_r)$  of pairwise distinct integers of  $[1, nt]$ .

# The main theorem in the classical case

## Theorem

The (non-commutative) distribution of the family  $(L_r(n, t))_{r \geq 1, t \in [0, +\infty)}$  converges, as  $n$  goes to infinity, to the one of a family  $(L_r(t))_{r \geq 1, t \in [0, +\infty)}$  such that

- ▶  $(L_1(t))_{t \in [0, +\infty)}$  is a (classical) **Brownian motion**;
- ▶ for all  $r, t$ , one has

$$L_r(t) = t^{\frac{r}{2}} H_r(t^{-1/2} L_1(t)),$$

where the  $H_r$ 's are the **Hermite polynomials**.

**Remark:** The Hermite polynomials are orthogonal with respect to the Gaussian distribution.

## Combinatorics and free probability

- ▶  $M_{r=1}(t=1)$  is a standard semicircular random variable;  $L_{r=1}(t=1)$  is a standard Gaussian. Their moments are given by

$$\varphi(M_1(1)^{2n}) = C_n = \frac{1}{n+1} \binom{2n}{n}$$

and,

$$\psi(L_1(1)^{2n}) = (2n)!! = 1 \cdot 3 \cdot 5 \cdots (2n-1).$$

- ▶ What about the moments of  $M_r(1)$  and  $L_r(1)$  for  $r \geq 2$  ?
- ▶ Combinatorial approach to free probability: free cumulants, non-crossing partitions, etc.

## Some insight from classical probability

Let  $X$  and  $Y$  be two independent classical random variables. What are the moments of  $X + Y$  ?

$$\mathbb{E}[(X + Y)^n] = \sum_{k=0}^n \binom{n}{k} \mathbb{E}[X^k Y^{n-k}] = \sum_{k=0}^n \binom{n}{k} \mathbb{E}[X^k] \mathbb{E}[Y^{n-k}].$$

**Idea:** use Fourier transform !

$$\mathcal{F}_{X+Y}(t) = \mathcal{F}_X(t) \cdot \mathcal{F}_Y(t).$$

Write  $\log \mathcal{F}_Z(t) = \sum_{n=0}^{\infty} \frac{(-it)^n}{n!} c_n(Z)$ . If  $X$  and  $Y$  are independent, then

$$c_n(X + Y) = c_n(X) + c_n(Y).$$

The quantities  $c_n$  are called **classical cumulants**.

## Free cumulants

The free cumulants  $\kappa_n$  are multilinear functionals  $\kappa_n : \mathcal{A}^n \rightarrow \mathbb{C}$  defined by

$$\varphi(a_1 a_2 \cdots a_n) = \sum_{\sigma \in NC(n)} \kappa_{\sigma}(a_1, a_2, \dots, a_n),$$

where

- ▶  $NC(n)$  is the lattice of non-crossing partitions of  $\{1, \dots, n\}$ . A partition  $\pi$  is **non-crossing** if there are no  $i < j < k < l$  such that  $i \overset{\pi}{\sim} k$  and  $j \overset{\pi}{\sim} l$ .
- ▶  $\kappa_{\sigma}$  is defined as a product over the blocks of  $\sigma$ .

### Theorem

*If  $a$  and  $b$  are free random variables, then*

$$\kappa_n(a + b, \dots, a + b) = \kappa_n(a, \dots, a) + \kappa_n(b, \dots, b) \quad \forall n \geq 1$$

**Notation:**  $NC_2(n)$  is the set of non-crossing **pairings** of  $n$ .

## Moments and free cumulants of the family $M_r$

Fix  $t = 1$  and let  $M_r = M_r(1)$ . Let  $\vec{r} = (r_1, \dots, r_p)$  be a vector of positive integers and put  $|\vec{r}| = r_1 + \dots + r_p$ . Consider the following sets of non-crossing pairings:

$$NC_2(\vec{r}) = \{\pi \in NC_2(|\vec{r}|) \mid \pi \wedge \hat{1}_{\vec{r}} = \hat{0}_{|\vec{r}|}\} \text{ and}$$

$$NC_2^*(\vec{r}) = \{\pi \in NC_2(\vec{r}) \mid \pi \vee \hat{1}_{\vec{r}} = \hat{1}_{|\vec{r}|}\}.$$

### Theorem

*The distribution of the family  $(M_r)_{r \geq 1}$  is characterized by the fact that its mixed moments are given by*

$$\varphi(M_{r_1} M_{r_2} \cdots M_{r_p}) = \#NC_2(\vec{r})$$

*and its free cumulants are given by*

$$\kappa_p(M_{r_1}, M_{r_2}, \dots, M_{r_p}) = \#NC_2^*(\vec{r}).$$

## Moments and classical cumulants of the family $L_r$

Fix  $t = 1$  and let  $L_r = L_r(1)$ . Let  $\Pi_2(\vec{r})$  be the set of general (i.e. possibly crossing) pairings  $\pi$  of  $\{1, \dots, |\vec{r}|\}$  such that  $\pi \wedge \hat{1}_{\vec{r}} = \hat{0}_{|\vec{r}|}$  and  $\Pi_2^*(\vec{r}) = \{\pi \in \Pi_2(\vec{r}) \mid \pi \vee \hat{1}_{\vec{r}} = \hat{1}_{|\vec{r}|}\}$ .

### Theorem

*The distribution of the family  $(L_r)_{r \geq 1}$  is characterized by the fact that its mixed moments are given by*

$$\psi(L_{r_1} L_{r_2} \cdots L_{r_p}) = \#\Pi_2(\vec{r})$$

*and its classical cumulants are given by*

$$c_p(L_{r_1}, L_{r_2}, \dots, L_{r_p}) = \#\Pi_2^*(\vec{r}).$$



## Conclusion

- ▶ We generalized Biane's result beyond transpositions
- ▶ Constructed a classical-probability analogue of the model
- ▶ More lines added to the "dictionary":

Classical Probability	Free Probability
...	...
subsets, $\Delta$	symmetric group
$L_r(t)$	$M_r(t)$
Hermite polynomials	Chebyshev 2 <sup>nd</sup> kind

- ▶ Some interesting combinatorics

## Perspectives

- ▶ Recently, we noticed that  $L_r(t)$  is the  $r$ -th multiple stochastic integral of a Brownian motion. Michael Anshelevich studied similar questions in the free case (*work in progress*).
- ▶ Are there analogous models for other types of independence in non-commutative probability theory (boolean, monotone) ?
- ▶ The moments of the random variables  $M_r$  count some particular semi-standard Young tableaux. Is there a connection with representation theory ?

Thank you !

<http://arxiv.org/abs/0801.4229>