# COMBINATORIAL ASPECTS OF FREE PROBABILITY THEORY

ROLAND SPEICHER*

## 1. Free central limit theorem and non-crossing pairings

One of the main ideas in free probability theory is to consider the notion of free independence in analogy with the notion of classical independence. In this spirit, the first investigations of Voiculescu in free probability theory centered around free analogues of some of the most fundamental statements from classical probability theory. In particular, he proved a free analogue of a central limit theorem and introduced and described a free analogue of "convolution". His investigations were quite analytical and centered around the concept of the "R-transform", an analytic function which plays in free probability theory the same role as the logarithm of the Fourier transform in classical probability theory. However, in this analytic approach it does not become obvious at all why the R-transform and the logarithm of the Fourier transform should be of an analogue nature.

Our approach to free probability theory is much more combinatorial in nature and will reveal in a much clearer way the parallelism between classical and free probability theory.

In order to see what kind of combinatorial objects are relevant for free probability theory we will now start with giving an algebraic proof of the free central limit theorem. This approach will show the similar nature of classical and free probability theory very clearly, because the same kind of proof can be given for the classical central limit theorem. Most of the arguments will be the same, only in the very end one has to distinguish whether one is in the classical or in the free situation. For convenience, we will restrict to the simplest case where we have identically distributed variables.

*Research supported by a Discovery Grant and a Leadership Support Initiative Award from the Natural Sciences and Engineering Research Council of Canada and by a Premier's Research Excellence Award from the Province of Ontario.

Lectures at the Summer School "Freie Wahrscheinlichkeitstheorie", Göttingen, August 2005.

**General setting for our central limit theorem.** Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and $a_1, a_2, \dots \in \mathcal{A}$ a sequence of identically distributed random variables which are either classically independent or freely independent. Furthermore, assume that the variables are centered, $\varphi(a_r) = 0$ $(r \in \mathbb{N})$, and denote by $\sigma^2 := \varphi(a_r^2)$ the common variance of the variables. A central limit theorem asks about the limit behaviour of

$$\frac{a_1 + \cdots + a_N}{\sqrt{N}}$$

if $N$ tends to infinity.

Of course, one has to specify the kind of convergence, and the only meaningful way one can expect to have convergence is in distribution. So let us first define this concept.

**Definition 1.1.** Let $(\mathcal{A}_N, \varphi_N)$ $(N \in \mathbb{N})$ and $(\mathcal{A}, \varphi)$ be non-commutative probability spaces and consider random variables $b_N \in \mathcal{A}_N$ for each $N \in \mathbb{N}$, and $b \in \mathcal{A}$. We say that $b_N$ **converges in distribution** towards $b$ for $N \to \infty$, and denote this by

$$b_N \overset{\text{distr}}{\longrightarrow} b,$$

if we have

$$\lim_{N \to \infty} \varphi_N(b_N^n) = \varphi(b^n) \qquad \text{for all } n \in \mathbb{N}.$$

*Remark* 1.2. This form of convergence seems to be weaker than the usual form of convergence appearing in the classical central limit theorem. There statements are usually in terms of "weak convergence". If we can identify the distribution $\mu_{b_N}$ of $b_N$ and the distribution $\mu_b$ of $b$ with probability measures on $\mathbb{R}$, then weak convergence of $\mu_{b_N}$ towards $\mu_b$ means

$$\lim_{N \to \infty} \int f(t) d\mu_{b_N}(t) = \int f(t) d\mu_b(t) \qquad \text{for all continuous } f.$$

However, in many cases this is equivalent to our definition above. Namely, if the limit measure is determined by its moments, then it suffices to check the convergence of all moments to conclude weak convergence. In full generality, this is a not totally obvious statement, but in the case when $\mu_b$ has compact support this is just an easy application of the Stone-Weierstrass theorem, which allows to approximate continuous functions on compact sets by polynomials. Usually, our random variables are selfadjoint bounded operators on a Hilbert space, so that their distribution is a compactly supported probability measure on $\mathbb{R}$.

However, one should note that also in the case of the classical central limit theorem, where the limit distribution is a normal distribution (and thus not compactly supported), it is enough to check convergence of moments, because a normal distribution is determined by its moments.

**General central limit theorem.** In order to see that we have convergence in distribution of $(a_1 + \cdots + a_N)/\sqrt{N}$ we should calculate the limit $N \to \infty$ of all moments of $(a_1 + \cdots + a_N)/\sqrt{N}$. Let us first see how much we can say about such moments for finite $N$. In the following we fix a positive integer $n$. Then we have

$$\varphi\big((a_1 + \cdots + a_N)^n\big) = \sum_{1 \le r(1),\dots,r(n) \le N} \varphi(a_{r(1)} \dots a_{r(n)}).$$

Since all $a_r$ have the same distribution we have

$$\varphi(a_{r(1)} \dots a_{r(n)}) = \varphi(a_{p(1)} \dots a_{p(n)})$$

whenever

$$r(i) = r(j) \quad \iff \quad p(i) = p(j) \qquad \forall \quad 1 \le i, j \le n.$$

(This is a consequence of the fact that both classical independence and free independence give a rule for calculating mixed moments out of the values of the moments of the variables.) Thus the value of $\varphi(a_{r(1)} \dots a_{r(n)})$ depends on the tuple $(r(1), \dots, r(n))$ only through the information which of the indices are the same and which are different. We will encode this information by a partition (i.e., a decomposition into disjoint subsets) $\pi = \{V_1, \dots, V_s\}$ of the set $\{1, \dots, n\}$. This partition $\pi$ is determined as follows: Two numbers $p$ and $q$ belong to the same block $V_m$ of $\pi$ (for some $m = 1, \dots, s$) if and only if $r(p) = r(q)$. We will write $(r(1), \dots, r(n)) \mathrel{\hat{=}} \pi$ in this case. Furthermore we denote the common value of $\varphi(a_{r(1)} \dots a_{r(n)})$ for all tuples $(r(1), \dots, r(n))$ with $(r(1), \dots, r(n)) \mathrel{\hat{=}} \pi$ by $k_\pi$.

For illustration, consider the following example: Since $a_1$ has the same moments as $a_7$, since $a_2$ has the same moments as $a_5$, and since $a_3$ has the same moments as $a_8$, the free independence of $a_1, a_2, a_3$ produces for $\varphi(a_1a_2a_1a_1a_2a_3)$ the same result as the free independence of $a_7, a_5, a_8$ for $\varphi(a_7a_5a_7a_7a_5a_8)$, and we denote the common value of both expressions by

$$k_{\{(1,3,4),(2,5),(6)\}} = \varphi(a_1a_2a_1a_1a_2a_3) = \varphi(a_7a_5a_7a_7a_5a_8).$$

With this notations we can continue the above calculation with

$$\varphi\big((a_1 + \cdots + a_N)^n\big) = \sum_{\pi \text{ partition of } \{1,\dots,n\}} k_\pi \cdot A_\pi^N,$$

where $A_\pi^N$ is the number of tuples corresponding to $\pi$, i.e.,

$$A_\pi^N := \#\{(r(1), \dots, r(n)) \mid 1 \le r(1), \dots, r(n) \le N, (r(1), \dots, r(n)) \mathrel{\hat{=}} \pi\}.$$

Note that there are only finitely many partitions of $\{1, \dots, n\}$ so that the above is a finite sum. The only dependence on $N$ is now via the numbers $A_\pi^N$. It remains to examine the contribution of the different partitions. We will see that most of them will give no contribution in the normalized limit, only very special ones survive.

Firstly, we will argue that partitions with singletons do not contribute: Consider a partition $\pi = \{V_1, \dots, V_q\}$ with a singleton, i.e., we have $V_m = \{r\}$ for some $m$ and some $r$. Then we have

$$k_\pi = \varphi(a_{r(1)} \dots a_r \dots a_{r(n)}) = \varphi(a_r) \cdot \varphi(a_{r(1)} \dots \check{a}_r \dots a_{r(n)}) = 0,$$

because $\{a_{r(1)}, \dots, \check{a}_r, \dots, a_{r(n)}\}$ is either classically independent or freely independent from $a_r$ and because our variables are centered. Thus only such partitions $\pi$ contribute which have no singletons, i.e. only $\pi = \{V_1, \dots, V_q\}$ with $\#V_m \geq 2$ for all $m = 1, \dots, q$. Note that this implies in particular that in our sum we can restrict to $\pi$ for which $q \leq n/2$.

Consider now a $\pi = \{V_1, \dots, V_q\}$; then we have $N$ possibilities for the common index corresponding to the first block $V_1$, $N-1$ possibilities for the common index corresponding to the second block $V_2$ (since this index has to be different from the one of the first block), and so on. Thus we have that

$$A_\pi^N = N(N-1)\dots(N-|\pi|+1),$$

which grows asymptocially like $N^{|\pi|}$ for large $N$. Thus

$$\begin{aligned}\lim_{N\to\infty} \varphi\big(\big(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\big)^n\big) &= \lim_{N\to\infty} \sum_\pi \frac{A_\pi^N}{N^{n/2}} k_\pi \\ &= \lim_{N\to\infty} \sum_\pi N^{|\pi|-(n/2)} k_\pi.\end{aligned}$$

Now note that for each appearing $\pi$ the factor $N^{|\pi|-(n/2)}$ has a limit (because only $|\pi| \leq n/2$ appear in our sum), and that this limit is either 1 or 0, depending on whether $|\pi| = n/2$ or $|\pi| < n/2$. This means, in the limit $N \to \infty$ all partitions with $|\pi| < n/2$ are supressed and we get exactly a contribution $k_\pi$ for each $\pi$ which has the property that it has no singleton and that its number of blocks is equal to $n/2$. This means of course that $\pi$ has to be a **pair partition** or **pairing**, i.e., a partition where each block $V_m$ consists of exactly two elements.

Thus, we have now arrived at the following result:

$$\lim_{N\to\infty} \varphi\big(\big(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\big)^n\big) = \sum_{\substack{\pi \text{ pair partition} \\ \text{of } \{1,\dots,n\}}} k_\pi \tag{1.1}$$

Up to now, there has been no difference between the case of classical independence and the case of free independence. The structure of the limiting moments is in both cases the same, namely they are calculated by summing over pair partitions. However, we still have not determined the weighting factors $k_\pi$ for these pair partitions. That is the point where we have to distinguish the two cases.

However, before we do this, let us note that the general formula (1.1) is enough to conclude that odd moments vanish in both cases. This conclusion comes from the simple observation that there are no pair partitions of a set with an odd number of elements. Thus:

$$\lim_{N\to\infty} \varphi\big((\frac{a_1+\cdots+a_N}{\sqrt{N}})^n\big) = 0 \qquad \text{for } n \text{ odd}. \tag{1.2}$$

**Classical central limit theorem.** The actual calculation of the limit distribution will now depend on whether we have classically independent or freely independent variables. Let us first consider the classical case. The usual factorization rule for classically independent random variables gives directly that for any pair partition $\pi$, the corresponding $k_\pi$ factorizes into a product of second moments, thus we have $k_\pi = \sigma^n$. So we get in this case

$$\lim_{N\to\infty} \varphi\big((\frac{a_1+\cdots+a_N}{\sqrt{N}})^n\big) = \sigma^n \cdot \Big(\#\text{pair partitions of } \{1,\dots,n\}\Big). \tag{1.3}$$

It is easy to see that the number of pair partitions of the set $\{1,\dots,n\}$ is, in the case $n$ even, given by $(n-1)\cdot(n-3)\cdots 5\cdot 3\cdot 1$.

On the other hand, one can also check that these numbers are exactly the moments of a centered normal distribution, i.e., that

$$\frac{1}{\sqrt{2\pi\sigma^2}}\int t^n e^{-t^2/2\sigma^2}dt = \begin{cases} 0, & \text{if } n \text{ odd} \\ \sigma^n\cdot(n-1)\cdot(n-3)\cdots 5\cdot 3\cdot 1, & \text{if } n \text{ even.}\end{cases} \tag{1.4}$$

Putting all this together, we have thus proved the following version of the classical central limit theorem.

**Theorem 1.3. (Classical CLT)** *Let $(\mathcal{A},\varphi)$ be a probability space and $a_1, a_2, \dots \in \mathcal{A}$ a sequence of independent and identically distributed random variables. Furthermore, assume that all variables are centered, $\varphi(a_r) = 0$ ($r \in \mathbb{N}$), and denote by $\sigma^2 := \varphi(a_r^2)$ the common variance of the variables.*
*Then we have*

$$\frac{a_1+\cdots+a_N}{\sqrt{N}} \xrightarrow{distr} \gamma,$$

*where $\gamma$ is a normally distributed random variable of variance $\sigma^2$.*

*Remark* 1.4. Let us recall that this statement means explicitely

$$(1.5)\quad \lim_{N\to\infty} \varphi\big(\big(\frac{a_1+\cdots+a_N}{\sqrt{N}}\big)^n\big) = \frac{1}{\sqrt{2\pi\sigma^2}}\int t^n e^{-t^2/2\sigma^2}dt \qquad \forall n\in\mathbb{N}.$$

As we mentioned before, the normal distribution is determined by its moments, so this algebraic form of statement is equivalent to the usual formulation in terms of weak convergence.

**Free central limit theorem.** Now we want to switch to the free case. So we start off again at the general formula (1.1) and it remains to specify what the weighting factors $k_\pi$ are in the case of freely independent random variables.

Since we know that the odd moments vanish in this case, too, it suffices to consider even moments. So let $n = 2k$ be even and consider a pair partition $\pi = \{V_1, \dots, V_k\}$. Let $(r(1), \dots, r(n))$ be an index-tuple corresponding to this $\pi$, $(r(1), \dots, r(n)) \hat{=} \pi$. Then there exist the following two possibilities:

(1) all consecutive indices are different:

$$r(1) \neq r(2) \neq \cdots \neq r(n).$$

Since $\varphi(a_{r(m)}) = 0$ for all $m = 1, \dots, n$, we have by the definition of free independence

$$k_\pi = \varphi(a_{r(1)} \dots a_{r(n)}) = 0.$$

(2) two consecutive indices coincide, i.e., $r(m) = r(m+1) = r$ for some $m = 1, \dots, n-1$. Because the index $r$ does not appear any more among the other indices we have that $\{a_{r(1)}, \dots, a_{r(m-1)}, a_{r(m+2)}, \dots, a_{r(n)}\}$ is freely independent from $a_{r(m)}a_{r(m+1)} = a_r a_r$ and we get by the factorization property

$$\varphi(b_1 b b_2) = \varphi(b_1 b_2)\varphi(b)$$

for $b$ and $\{b_1, b_2\}$ free that

$$\begin{aligned} k_\pi &= \varphi(a_{r(1)} \dots a_r a_r \dots a_{r(n)}) \\ &= \varphi(a_{r(1)} \dots a_{r(m-1)} a_{r(m+2)} \dots a_{r(n)})\varphi(a_r a_r) \\ &= \varphi(a_{r(1)} \dots a_{r(m-1)} a_{r(m+2)} \dots a_{r(n)})\sigma^2 \end{aligned}$$

It is clear that we can use the above argumentation in the second case again and either get zero for $k_\pi$ or reduce the length of the considered word further. So we see that the contribution of a pairing $\pi$ can now be either $\sigma^n$ or 0, i.e. only special pairings will make a contribution. These special pairing are exactly those for which in each recursion step we are in the first case, i.e., we successively can find consecutive indices which coincide. If we have a pairing for which this is not possible, then this means that there exist $p_1 < q_1 < p_2 < q_2$ such that $p_1$ is paired with $p_2$ and $q_1$ is paired with $q_2$. Such a pairing is obviously called "crossing".

So we have arrived at the conclusion that in the free case not all pairings contribute, but only very special ones, namely those which are not crossing in the above sense. This are now exactly the fundamental combinatorial objects, on which our description of free probability theory will rest.

**Notation 1.5.** A pairing of $\{1, \dots, n\}$ is called **non-crossing** if there does not exist $1 \leq p_1 < q_1 < p_2 < q_2 \leq n$ such that $p_1$ is paired with $p_2$ and $q_1$ is paired with $q_2$.

Thus we have shown

$$\lim_{N\to\infty} \varphi\big(\big(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\big)^n\big) = D_n \sigma^n, \tag{1.6}$$

where

$$D_n := \#\{\pi \mid \pi \text{ non-crossing pair partition of } \{1, \dots, n\}\}.$$

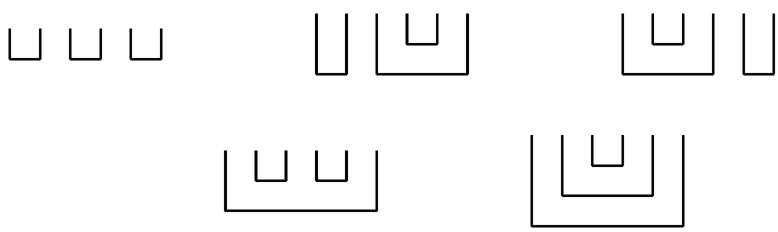
Clearly, for $n$ odd, $D_n$ vanishes, since there is no pairing of an odd number of elements.

Here are the first three values of $D_n$ for even $n$.

- $D_2 = 1$, since there is only one non-crossing pairing of 2 elements:

- $D_4 = 2$; here are the two non-crossing pairings of 4 elements (note that there is one additional, crossing, pairing for $n = 4$)

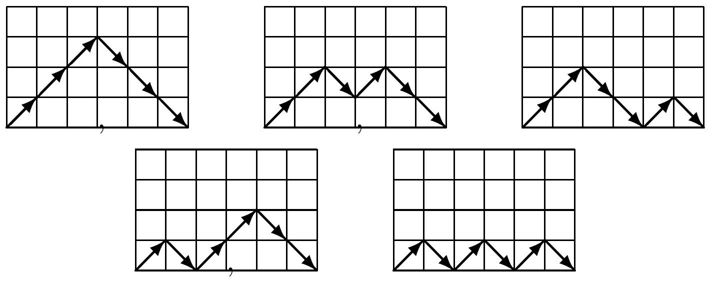
- $D_6 = 5$; here are the five non-crossing pairings of 6 elements (there are 15 pairings of 6 elements, the other 10 are crossing)



It is quite easy to calculate the number of non-crossing pair partitions for general $n$. It will turn out that these are given by the so-called Catalan numbers.

**Notation 1.6.** A Catalan path of lenght $2n$ is a path in the integer lattice $\mathbb{Z}^2$ which starts at $(0,0)$, ends at $(2n,0)$, always makes steps of the form $(1,1)$ or $(1,-1)$ and is not allowed to fall under the $x$-axis).

As example, here are the 5 Catalan paths with 6 steps .



*Remark* 1.7. Catalan paths of length $2n$ are counted by the famous Catalan numbers

$$c_n = \frac{1}{n+1}\binom{2n}{n}$$

These are determined by $c_0 = c_1 = 1$ and the recurrence relation $(n \geq 2)$

$$c_n = \sum_{k=1}^{n} c_{k-1}c_{n-k}.$$

The first fews Catalan numbers are

$$c_1 = 1, \qquad c_2 = 2, \qquad , c_3 = 5, \qquad c_4 = 14, \qquad , c_5 = 42, \cdots$$

**Lemma 1.8.** *The number $D_{2k}$ of non-crossing pair-partitions of the set $\{1,\dots,2k\}$ is given by the Catalan number $c_k$.*

*Proof.* Since clearly $D_2 = 1 = c_1$, it suffices to check that the $D_{2k}$ fulfill the recurrence relation of the Catalan numbers: Let $\pi = \{V_1,\dots,V_k\}$ be a non-crossing pair partition. We denote by $V_1$ that block of $\pi$ which contains the element 1, i.e., it has to be of the form $V_1 = (1,m)$. Then the property "non-crossing" enforces that, for each $V_j$ $(j \neq 1)$, we can not have a crossing between $V_1$ and $V_j$, i.e, we have either $1 < V_j < m$ or $1 < m < V_j$. (In particular, this implies that $m$ has to be even, $m = 2l$.) This means that $\pi$ restricted to $\{2,\dots,m-1\}$ is a non-crossing pair partition of $\{2,\dots,m-1\}$ and $\pi$ restricted to $\{m+1,\dots,n\}$ is a non-crossing pair partition of $\{m+1,\dots,n\}$. There exist $D_{m-2}$ many non-crossing pair partitions of $\{2,\dots,m-1\}$ and $D_{n-m}$ many non-crossing pair partitions of $\{m+1,\dots,n\}$, where we put consistently

$D_0 := 1$. Both these possibilities can appear independently from each other and $m = 2l$ can run through all even numbers from 2 to $n$. Hence we get

$$D_{2k} = \sum_{l=1}^{k} D_{2(l-1)} D_{2(k-l)}.$$

But this is the recurrence relation for the Catalan numbers, so the assertion follows. □

We will call a random variable with vanishing odd moments and even moments given by the Catalan numbers a semi-circular element. The reason for this terminology will become clear when we present the corresponding probability density in the next lecture.

**Definition 1.9.** A **semi-circular element of variance** $\sigma^2$ is a random variable $s$ whose moments are given by

$$\varphi(s^k) = \begin{cases} c_n \sigma^{2n}, & \text{if } k = 2n \\ 0, & \text{if } k \text{ odd.} \end{cases}$$

Putting all this together, we have proved the free central limit theorem in the following form.

**Theorem 1.10. (Free CLT)** *Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and $a_1, a_2, \dots \in \mathcal{A}$ a sequence of freely independent and identically distributed random variables. Assume furthermore $\varphi(a_r) = 0$ $(r \in \mathbb{N})$ and denote by $\sigma^2 := \varphi(a_r^2)$ the common variance of the variables.*
*Then we have*

$$\frac{a_1 + \cdots + a_N}{\sqrt{N}} \xrightarrow{distr} s,$$

*where $s$ is a semi-circular element of variance $\sigma^2$.*

*Remarks* 1.11. 1) According to the free central limit theorem the semi-circular distribution has to be considered as the free analogue of the normal distribution and is thus one of the basic distributions in free probability theory.
2) Semi-circular elements appear also canonically as concrete operators on the full Fock space. (For completeness, the main definitions around the full Fock space are given in an appendix.) I will in my lectures put the emphasis on combinatorial considerations, however, in many cases there are alternate approaches relying on Fock space constructions. I will not elaborate on these, but I want to indicate at least the relation

between these two approaches. Very roughly, the combinatorics of non-crossing partitions is hidden implicitely in calculations with creation and annihilation operators on full Fock spaces.

To see the simplest manifestation of this let us check that the sum of creation and annihilation operator (which is, up to normalization, the real part of the unilateral shift) is a semi-circular element. So we want to calculate the distribution of the operator $\omega := \omega(f)$. Let $f$ be a unit vector, then we can describe the situation as follows: We have $\omega = l + l^*$, where $l^*l = 1$ and $l^*\Omega = 0$. Our state is given by

$$\varphi(\cdot) = \langle \Omega, \cdot\Omega \rangle.$$

This information is enough to calculate all moments of $\omega$ with respect to $\varphi$. Clearly all odd moments are zero, and we have

$$\varphi(\omega^2) = \varphi(l^*l) = 1$$

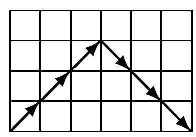
$$\varphi(\omega^4) = \varphi(l^*l^*ll) + \varphi(l^*ll^*l) = 2$$

$$\begin{aligned}\varphi(\omega^6) &= \varphi(l^*l^*l^*lll) + \varphi(l^*l^*lll^*l) \\ &\quad + \varphi(l^*l^*ll^*ll) + \varphi(l^*ll^*l^*ll) + \varphi(l^*ll^*ll^*l) \\ &= 5.\end{aligned}$$

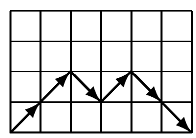
A closer look on those examples reveals that the sequences in $l^*$'s and $l$'s which contribute to the calculation of the $2n$-th moment of $\omega$ can be identified with Catalan paths of lenght $2n$.

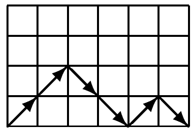
As example let us consider the 5 Catalan paths with 6 steps . We draw them in the pictures below, and for each of them we indicate the corresponding sequence in the calculation of $\varphi(\omega^6)$. (Note that you have to read the sequence in $l, l^*$ backwards to match it with the path.)

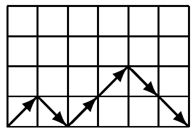
$l^*\, l^*\, l^*\, l\, l\, l$

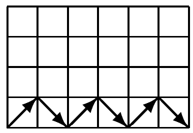


$l^*\, l^*\, l\, l^*\, l\, l$



$l^* \, l \, l^* \, l^* \, l \, l$ 

$l^* \, l^* \, l \, l \, l^* \, l$ 

$l^* \, l \, l^* \, l \, l^* \, l$ 

Thus we have finally

$$\varphi(\omega^k) = \begin{cases} c_n, & \text{if } k = 2n \\ 0, & \text{if } k \text{ odd.} \end{cases}$$

and hence $\omega$ is a semi-circular element of variance 1.

Our parallel treatment of the classical and the free central limit theorem shows the similarity between these two theorems very clearly. In particular, we have learned the simplest manifestation of the following basic observation: the transition from classical to free probability theory consists on a combinatorial level in the transition from all partitions to non-crossing partitions.

**The multi-dimensional case.** One of the main advantages of our combinatorial approach to free probability theory is the fact that, in contrast to an analytical treatment, a lot of arguments can be extended from the one-dimensional to the more-dimensional case without any problems. In the following we want to demonstrate this for the free central limit theorem.

We replace now each $a_r$ by a family of random variables $(a_r^{(i)})_{i\in I}$ and assume that all these families are free and each of them has the same joint distribution and that all appearing random variables are centered. We ask for the convergence of the joint distribution of the random variables $\big((a_1^{(i)}+\cdots+a_N^{(i)})/\sqrt{N}\big)_{i\in I}$ when $N$ tends to infinity. Let us first define the obvious generalization of our notion of convergence to this multi-dimensional setting.

**Definition 1.12.** Let $(\mathcal{A}_N,\varphi_N)$ $(N\in\mathbb{N})$ and $(\mathcal{A},\varphi)$ be non-commutative probability spaces. Let $I$ be an index set and consider for each $i\in I$ random variables $b_N^{(i)}\in\mathcal{A}_N$ and $b_i\in\mathcal{A}$. We say that $(b_N^{(i)})_{i\in I}$ **converges in distribution** towards $(b_i)_{i\in I}$ and denote this by

$$(b_N^{(i)})_{i\in I}\xrightarrow{\text{distr}}(b_i)_{i\in I},$$

if we have that each joint moment of $(b_N^{(i)})_{i\in I}$ converges towards the corresponding joint moment of $(b_i)_{i\in I}$, i.e., if we have for all $n\in\mathbb{N}$ and all $i(1),\dots,i(n)\in I$

$$\lim_{N\to\infty}\varphi_N(b_N^{(i(1))}\dots b_N^{(i(n))})=\varphi(b_{i(1)}\dots b_{i(n)}). \tag{1.7}$$

In the free central limit theorem, the calculation of the joint distribution of $(a_1^{(i)}+\cdots+a_N^{(i)})/\sqrt{N}$ works in the same way as in the one-dimensional case. Namely, we now have to consider for all $n\in\mathbb{N}$ and all $i(1),\dots,i(n)\in I$

$$\varphi\Big((a_1^{(i(1))}+\cdots+a_N^{(i(1))})\cdots(a_1^{(i(n))}+\cdots+a_N^{(i(n))})\Big) = \sum_{1\le r(1),\dots,r(n)\le N}\varphi(a_{r(1)}^{(i(1))}\dots a_{r(n)}^{(i(n))}). \tag{1.8}$$

Again, we have that the value of $\varphi(a_{r(1)}^{(i(1))}\dots a_{r(n)}^{(i(n))})$ depends on the tuple $(r(1),\dots,r(n))$ only through the information which of the indices are the same and which are different, which we will encode as before by a partition $\pi$ of $\{1,\dots,n\}$. The common value of $\varphi(a_{r(1)}^{(i(1))}\dots a_{r(n)}^{(i(n))})$ for all tuples $(r(1),\dots,r(n))\,\hat{=}\,\pi$ will now, in addition, also depend on the tuple $(i(1),\dots,i(n))$ and we will denote it by $k_\pi[i(1),\dots,i(n)]$. The next steps are the same as before. Singletons do not contribute because of the centeredness assumption and only pair partitions survive in the limit. Thus we arrive at

$$\lim_{N\to\infty}\varphi\Big(\frac{a_1^{(i(1))}+\cdots+a_N^{(i(1))}}{\sqrt{N}}\cdots\frac{a_1^{(i(n))}+\cdots+a_N^{(i(n))}}{\sqrt{N}}\Big) = \sum_{\substack{\pi\text{ pair partition}\\ \text{of }\{1,\dots,n\}}}k_\pi[i(1),\dots,i(n)]. \tag{1.9}$$

It only remains to identify the contribution $k_\pi[i(1),\dots,i(n)]$ for a pair partition $\pi$. As before, the free independence assumption ensures that

$k_\pi[i(1),\dots,i(n)] = 0$ for crossing $\pi$. So consider finally non-crossing $\pi$. Remember that in the case of a non-crossing $\pi$ we can find two consecutive indices which coincide, i.e., $r(m) = r(m+1) = r$ for some $m$. Then we have

$$\begin{aligned} k_\pi[i(1),\dots,i(n)] &= \varphi(a_{r(1)}^{(i(1))} \dots a_r^{(i(m))} a_r^{(i(m+1))} \dots a_{r(n)}^{(i(n))}) \\ &= \varphi(a_{r(1)}^{(i(1))} \dots a_r^{(i(m-1))} a_r^{(i(m+2))} \dots a_{r(n)}^{(i(n))}) \varphi(a_r^{(i(m))} a_r^{(i(m+1))}) \\ &= \varphi(a_{r(1)}^{(i(1))} \dots a_{r(m-1)}^{(i(m-1))} a_{r(m+2)}^{(i(m+2))} \dots a_{r(n)}^{(i(n))}) c_{i(m)i(m+1)}, \end{aligned}$$

where $c_{ij} := \varphi(a_r^{(i)} a_r^{(j)})$ is the covariance of $(a_r^{(i)})_{i\in I}$.

Iterating of this will lead to the final result that $k_\pi[i(1),\dots,i(n)]$ is given by the product of covariances $\prod_{(p,q)\in\pi} c_{i(p)i(q)}$ (one factor for each block $(p,q)$ of $\pi$).

This form of the limiting moments motivates the following generalization of the notion of a semi-circular element to the multi-dimensional case.

**Definition 1.13.** A family $(s_i)_{i\in I}$ of random variables in some non-commutative probability space is called a **semi-circular family of covariance** $(c_{ij})_{i,j\in I}$, if its distribution is of the following form: for all $n \in \mathbb{N}$ and all $i(1),\dots,i(n) \in I$ we have

$$\varphi(s_{i(1)} \dots s_{i(n)}) = \sum_{\substack{\pi \text{ non-crossing pair partition} \\ \text{of } \{1,\dots,n\}}} k_\pi[s_{i(1)},\dots,s_{i(n)}], \tag{1.10}$$

where

$$k_\pi[s_{i(1)},\dots,s_{i(n)}] = \prod_{(p,q)\in\pi} c_{i(p)i(q)}. \tag{1.11}$$

*Examples* 1.14. For illustration, let us write down these formulas explicitely for small moments:

$$\varphi(s_1 s_2) = c_{12}$$

$$\varphi(s_1 s_2 s_3 s_4) = c_{12}c_{34} + c_{14}c_{23}$$

$$\varphi(s_1 s_2 s_3 s_4 s_5 s_6) = c_{12}c_{34}c_{56} + c_{12}c_{36}c_{45} + c_{14}c_{23}c_{56} + c_{16}c_{23}c_{45} + c_{16}c_{25}c_{34}.$$

With this notation of a semi-circular family we can summarize our calculations in the following multi-dimensional version of a free central limit theorem.

**Theorem 1.15. (Free CLT - multi-dimensional version)** *Let* $(\mathcal{A},\varphi)$ *be a non-commutative probability space and*

$(a_1^{(i)})_{i\in I}, (a_2^{(i)})_{i\in I}, \dots \subset \mathcal{A}$ *a sequence of freely independent families with the same joint distribution of* $(a_r^{(i)})_{i\in I}$ *for all* $r \in \mathbb{N}$. *Assume furthermore that all variables are centered*

$$\varphi(a_r^{(i)}) = 0 \qquad (r \in \mathbb{N}, i \in I)$$

*and denote the covariance of the variables – which is independent of r – by*

$$c_{ij} := \varphi(a_r^{(i)} a_r^{(j)}) \qquad (i, j \in I).$$

*Then we have*

$$\Big(\frac{a_1^{(i)} + \cdots + a_N^{(i)}}{\sqrt{N}}\Big)_{i\in I} \overset{distr}{\longrightarrow} (s_i)_{i\in I}, \tag{1.12}$$

*where* $(s_i)_{i\in I}$ *is a semi-circular family of covariance* $(c_{ij})_{i,j\in I}$.

*Remark* 1.16. Clearly, we can also do a multi-dimensional version of the classical central limit theorem in the same way. Then the limit is a Gaussian family, whose joint moments are given by a similar formula as for semi-circular families, the only difference is again that the summation runs over all pairings instead of non-crossing pairings. So for a Gaussian family $(\gamma_i)$ of covariance $(c_{ij})$ we would have

$$\varphi(\gamma_1\gamma_2\gamma_3\gamma_4) = c_{12}c_{34} + c_{14}c_{23} + c_{13}c_{24}$$

and the moment $\varphi(\gamma_1\gamma_2\gamma_3\gamma_4\gamma_5\gamma_6)$ would be given by a sum over the 15 pairings of 6 elements.

This combinatorial description of the joint moments of Gaussian families goes usually (in particular, in the physical community) under the name of "Wick formula".

A simple special case of a semi-circular family is given if the covariance is a diagonal matrix. We can use our free central limit theorem to conclude quite easily that this is equivalent to having freely independent semi-circular elements. This is a special case of the following proposition.

**Proposition 1.17.** *Let* $(s_i)_{i\in I}$ *be a semi-circular family of covariance* $(c_{ij})_{i,j\in I}$ *and consider a disjoint decomposition* $I = \dot\bigcup_{p=1}^d I_p$. *Then the following two statements are equivalent:*

(1) *The sets* $\{s_i \mid i \in I_1\}, \dots, \{s_i \mid i \in I_p\}$ *are freely independent.*
(2) *We have* $c_{ij} = 0$ *whenever* $i \in I_p$ *and* $j \in I_q$ *with* $p \neq q$.

*Proof.* Assume first that the families $\big(\{s_i \mid i \in I_p\}\big)_{p=1,\dots,d}$ are freely independent and consider $i \in I_p$ and $j \in I_q$ with $p \neq q$. Then the free independence of $s_i$ and $s_j$ implies in particular

$$c_{ij} = \varphi(s_i s_j) = \varphi(s_i)\varphi(s_j) = 0.$$

If on the other side we have $c_{ij} = 0$ whenever $i \in I_p$ and $j \in I_q$ with $p \neq q$, then we can use our free central limit theorem in the following way: Choose $a_r^{(i)}$ such that (for each $r$) they have the same distribution as the $s_i$ up to second order (i.e. first moments vanish and covariance is given by $(c_{ij})_{i,j\in I}$) and such that the sets $\{a_r^{(i)}\}_{i\in I_p}$ $(p = 1, \dots, d)$ are freely independent. Note that these requirements are compatible because of our assumption on the covariances $c_{ij}$. Of course, we can also assume that the joint distribution of $(a_r^{(i)})_{i\in I}$ is independent of $r$. But then our free central limit theorem tells us that

$$\Big(\frac{a_1^{(i)} + \cdots + a_N^{(i)}}{\sqrt{N}}\Big)_{i\in I} \xrightarrow{\text{distr}} (s_i)_{i\in I}, \tag{1.13}$$

where the limit is given exactly by the semi-circular family with which we started (because this has the right covariance). But we also have that the sets $\{(a_1^{(i)} + \cdots + a_N^{(i)})/\sqrt{N}\}_{i\in I_p}$ $(p = 1, \dots, d)$ are freely independent; since free independence passes over to the limit we get the wanted result that the sets $\{s_i\}_{i\in I_p}$ $(p = 1, \dots, d)$ are freely independent. □

Note that this means in particular that $k_\pi[s_{i(1)}, \dots, s_{i(n)}]$ vanishes if $\pi$ couples semi-circular elements which are free.

*Example* 1.18. Assume $s_1$ and $s_2$ are two semi-circular elements which are free. Let us also assume that both have variance 1. Then the above tells us that their mixed moments are given by counting the non-crossing pairings which connect $s_1$ with $s_1$ and $s_2$ with $s_2$ (no blocks connecting $s_1$ with $s_2$ are allowed). For example, we have in such a situation

$$\varphi(s_1s_1s_2s_2s_1s_2s_2s_1) = 2,$$

because there are two contributing non-crossing pairings, namely

$$\{(1,2),(3,4),(5,8),(6,7)\} \qquad \text{and} \qquad \{(1,8),(2,5),(3,4),(6,7)\}.$$

**Conclusion.** The general conclusion which we draw from this lecture is that non-crossing partitions appear quite naturally in free probability. From a combinatorial point of view, the transition from classical probability theory to free probability theory consists in replacing all partitions by non-crossing partitions.

But there are also more specific features shown by our treatment of the free central limit theorem. In the next lectures we will generalize to arbitrary distributions what we have learned from the case of semi-circular families, namely:

(1) it seems to be canonical to write moments as sums over non-crossing partitions
(2) the appearing summands $k_\pi$ are multiplicative in the sense that they factorize in a product according to the block structure of $\pi$
(3) the summands $k_\pi$ reflect free independence quite clearly, since $k_\pi[s_{i(1)}, \dots, s_{i(n)}]$ vanishes if the blocks of $\pi$ couple elements which are freely independent

## Appendix: Full Fock Space

**Definition 1.19.** Let $\mathcal{H}$ be a Hilbert space.
1) The *full Fock space over* $\mathcal{H}$ is the Hilbert space

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

where $\Omega$ is a distinguished unit vector, called *vacuum.*
2) The *vacuum expectation* is the state

$$\begin{aligned} B(\mathcal{F}(\mathcal{H})) &\to \mathbb{C} \\ a &\mapsto \langle \Omega, a\Omega \rangle. \end{aligned}$$

3) For each $f \in \mathcal{H}$ we define the *(left) annihilation operator* $l^*(f)$ and the *(left) creation operator* $l(f)$ by

$$l^*(f)\Omega = 0$$
$$l^*(f) f_1 \otimes \dots \otimes f_n = \langle f, f_1 \rangle f_2 \otimes \dots \otimes f_n$$

and

$$l(f) f_1 \otimes \dots \otimes f_n = f \otimes f_1 \otimes \dots \otimes f_n.$$

Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada
*E-mail address*: speicher@mast.queensu.ca