Combinatorics of free probability theory

Roland Speicher

Author address:

HEIDELBERG

'Free probability theory and operator spaces', IHP, Paris, 1999.

Contents

Part 1

Basic concepts

CHAPTER 1

Some basic concepts of non-commutative probability theory

Definition 1.1. 1) A **non-commutative probability space** $(\mathcal{A}, \varphi)$ consists of a unital algebra $\mathcal{A}$ and a linear functional

$$\varphi : \mathcal{A} \to \mathbb{C}; \qquad \varphi(1) = 1.$$

2) If in addition $\mathcal{A}$ is a $*$-algebra and φ fulfills $\varphi(a^*) = \overline{\varphi(a)}$ for all $a \in \mathcal{A}$, then we call $(\mathcal{A}, \varphi)$ a **$*$-probability space**.
3) If $\mathcal{A}$ is a unital C^*-algebra and φ a state, then we call $(\mathcal{A}, \varphi)$ a **C^*-probability space**.
4) If $\mathcal{A}$ is a von Neumann algebra and φ a normal state, then we call $(\mathcal{A}, \varphi)$ a **W^*-probability space**.
5) Elements $a \in \mathcal{A}$ are called **non-commutative random variables** in $(\mathcal{A}, \varphi)$.
6) Let $a_1, \dots, a_n$ be random variables in some probability space $(\mathcal{A}, \varphi)$. If we denote by $\mathbb{C}\langle X_1, \dots, X_n\rangle$ the algebra generated freely by n non-commuting indeterminates $X_1, \dots, X_n$, then the **joint distribution** $\mu_{a_1,\dots,a_n}$ of $a_1, \dots, a_n$ is given by the linear functional

$$\mu_{a_1,\dots,a_n} : \mathbb{C}\langle X_1, \dots, X_n\rangle \to \mathbb{C} \tag{1}$$

which is determined by

$$\mu_{a_1,\dots,a_n}(X_{i(1)} \dots X_{i(k)}) := \varphi(a_{i(1)} \dots a_{i(k)})$$

for all $k \in \mathbb{N}$ and all $1 \le i(1), \dots, i(k) \le n$.
7) Let $a_1, \dots, a_n$ be a family of random variables in some $*$-probability space $(\mathcal{A}, \varphi)$. If we denote by $\mathbb{C}\langle X_1, X_1^*, \dots, X_n, X_n^*\rangle$ the algebra generated freely by $2n$ non-commuting indeterminates $X_1, X_1^*, \dots, X_n, X_n^*$, then the **joint $*$-distribution** $\mu_{a_1,a_1^*,\dots,a_n,a_n^*}$ of $a_1, \dots, a_n$ is given by the linear functional

$$\mu_{a_1,a_1^*,\dots,a_n,a_n^*} : \mathbb{C}\langle X_1, X_1^*, \dots, X_n, X_n^*\rangle \to \mathbb{C} \tag{2}$$

which is determined by

$$\mu_{a_1,a_1^*,\dots,a_n,a_n^*}(X_{i(1)}^{r(1)} \dots X_{i(k)}^{r(k)}) := \varphi(a_{i(1)}^{r(1)} \dots a_{i(k)}^{r(k)})$$

for all $k \in \mathbb{N}$, all $1 \leq i(1), \dots, i(k) \leq n$ and all choices of $r(1), \dots, r(k) \in \{1, *\}$.

EXAMPLES 1.2. 1) Classical (=commutative) probability spaces fit into this frame as follows: Let $(\Omega, \mathcal{Q}, P)$ be a probability space in the classical sense, i.e., Ω a set, $\mathcal{Q}$ the σ-fiel of measurable subsets of Ω and P a probability measure, then put $\mathcal{A} = L^\infty(\Omega, P)$ and

$$\varphi(X) = \int X(\omega) dP(\omega)$$

for bounded random variables $X : \Omega \to \mathbb{C}$.
Note that the fact $P(\Omega) = 1$ corresponds to $\varphi(1) = \int 1 dP(\omega) = P(\Omega) = 1$.
2) Typical non-commutative random variables are given as operators on Hilbert spaces: Let $\mathcal{H}$ be a Hilbert space and $\mathcal{A} = B(\mathcal{H})$ (or more generally a C^*- or von Neumann algebra). Then the fundamental states are of the form $\varphi(a) = \langle \eta, a\eta \rangle$, where $\eta \in \mathcal{H}$ is a vector of norm 1. (Note that $\|\eta\| = 1$ corresponds to $\varphi(1) = \langle \eta, 1\eta \rangle = 1$.)
All states on a C^*-algebra can be written in this form, namely one has the following GNS-construction: Let $\mathcal{A}$ be a C^*-algebra and $\varphi : \mathcal{A} \to \mathbb{C}$ a state. Then there exists a representation $\pi : \mathcal{A} \to B(\mathcal{H})$ on some Hilbert space $\mathcal{H}$ and a unit vector $\eta \in \mathcal{H}$ such that: $\varphi(a) = \langle \eta, \pi(a)\eta \rangle$.

REMARKS 1.3. 1) In general, the distribution of non-commutative random variables is just the collection of all possible joint moments of these variables. In the case of one normal variable, however, this collection of moments can be identified with a probability measure on $\mathbb{C}$ according to the following theorem: Let $(\mathcal{A}, \varphi)$ be a C^*-probability space and $a \in \mathcal{A}$ a normal random variable, i.e., $aa^* = a^*a$. Then there exists a uniquely determined probability measure μ on the spectrum $\sigma(a) \subset \mathbb{C}$, such that we have:

$$\varphi(p(a, a^*)) = \int_{\sigma(a)} p(z, \bar{z}) d\mu(z)$$

for all polynomials p in two commuting variables, i.e. in particular

$$\varphi(a^n a^{*m}) = \int_{\sigma(a)} z^n \bar{z}^m d\mu(z)$$

for all $n, m \in \mathbb{N}$.
2) We will usually identify for normal operators a the distribution μ_{a,a^*} with the probability measure μ.
In particular, for a self-adjoint and bounded operator x, $x = x^* \in B(\mathcal{H})$, its distribution μ_x is a probability measure on $\mathbb{R}$ with compact support.

3) Note that above correspondence between moments and probability measures relies (via Stone-Weierstrass) on the boundedness of the operator, i.e., the compactness of the spectrum. One can also assign a probability measure to unbounded operators, but this is not necessarily uniquely determined by the moments.
4) The distribution of a random variable contains also metric and algebraic information, if the considered state is faithful.

Definition 1.4. A state φ on a $*$-algebra is called **faithful** if $\varphi(aa^*) = 0$ implies $a = 0$.

Proposition 1.5. *Let $\mathcal{A}$ be a C^*-algebra and $\varphi : \mathcal{A} \to \mathbb{C}$ a faithful state. Consider a self-adjoint $x = x^* \in \mathcal{A}$. Then we have*

$$\sigma(x) = \mathit{supp}\,\mu_x, \tag{3}$$

and thus also

$$\|x\| = \max\{|z| \mid z \in \mathit{supp}\,\mu_x\} \tag{4}$$

and

$$\|x\| = \lim_{p\to\infty} \big(\varphi(x^p)\big)^{1/p}. \tag{5}$$

Proposition 1.6. *Let $(\mathcal{A}_1, \varphi_1)$ and $(\mathcal{A}_2, \varphi_2)$ be $*$-probability spaces such that φ_1 are φ_2 are faithful. Furthermore, let $a_1, \dots, a_n$ in $(\mathcal{A}_1, \varphi_1)$ and $b_1, \dots, b_n$ in $(\mathcal{A}_2, \varphi_2)$ be random variables with the same $*$-distribution, i.e., $\mu_{a_1,a_1^*,\dots,a_n,a_n^*} = \mu_{b_1,b_1^*,\dots,b_n,b_n^*}$.*
1) If $\mathcal{A}_1$ is generated as a $*$-algebra by $a_1, \dots, a_n$ and $\mathcal{A}_2$ is generated as a $*$-algebra by $b_1, \dots, b_n$, then $1 \mapsto 1$, $a_i \mapsto b_i$ $(i = 1, \dots, n)$ can be extended to a $*$-algebra isomorphism $\mathcal{A}_1 \cong \mathcal{A}_2$.
2) If $(\mathcal{A}_1, \varphi_1)$ and $(\mathcal{A}_2, \varphi_2)$ are C^*-probability spaces such that $\mathcal{A}_1$ is as C^*-algebra generated by $a_1, \dots, a_n$ and $\mathcal{A}_2$ is as C^*-algebra generated by $b_1, \dots, b_n$, then we have $\mathcal{A}_1 \cong \mathcal{A}_2$ via $a_i \mapsto b_i$.
3) If $(\mathcal{A}_1, \varphi_1)$ and $(\mathcal{A}_2, \varphi_2)$ are W^*-probability spaces such that $\mathcal{A}_1$ is as von Neumann algebra generated by $a_1, \dots, a_n$ and $\mathcal{A}_2$ is as von Neumann algebra generated by $b_1, \dots, b_n$, then we have $\mathcal{A}_1 \cong \mathcal{A}_2$ via $a_i \mapsto b_i$.

Examples 1.7. 1) the **two-sided shift** u
Let $\mathcal{H}$ be a Hilbert space with $\dim(\mathcal{H}) = \infty$ and orthonormal basis $(e_i)_{i=-\infty}^{\infty}$. We now define u by $(k \in \mathbb{Z})$

$$ue_k = e_{k+1} \qquad \text{and thus} \qquad u^*e_k = e_{k-1}. \tag{6}$$

We have $uu^* = u^*u = 1$, i.e., u is unitary. Consider now the state

$$\varphi(a) = \langle e_0, ae_0\rangle.$$

The distribution $\mu := \mu_{u,u^*}$ is determined by $(n \geq 0)$

$$\varphi(u^n) = \langle e_0, u^n e_0 \rangle = \langle e_0, e_n \rangle = \delta_{n0}$$

and

$$\varphi(u^{*n}) = \langle e_0, u^{*n} e_0 \rangle = \langle e_0, e_{-n} \rangle = \delta_{n0}.$$

Since u is normal, we can identify μ with a probability measure on

$$\sigma(u) = \mathbb{T} := \{z \in \mathbb{C} \mid |z| = 1\}.$$

This measure is the normalized Haar measure on $\mathbb{T}$, i.e.,

$$d\mu(z) = \frac{1}{2\pi} dz.$$

It is uniquely determined by

$$\int_{\mathbb{T}} z^k d\mu(z) = \delta_{k0} \qquad (k \in \mathbb{Z}).$$

2) the **one-sided shift** l:
Let $\mathcal{H}$ be a Hilbert space with $\dim(\mathcal{H}) = \infty$ and basis $(e_i)_{i=0}^{\infty}$. The one-sided shift l is defined by

$$le_n = e_{n+1} \qquad (n \geq 0).$$

Its adjoint l^* is thus given by

$$l^* e_n = e_{n-1} \qquad (n \geq 1) \tag{7}$$

$$l^* e_0 = 0. \tag{8}$$

We have the relations

$$l^* l = 1 \qquad \text{and} \qquad ll^* = 1 - |e_0\rangle\langle e_0|, \tag{9}$$

i.e., l is a (non-unitary) isometry.
Again we consider the state

$$\varphi(a) := \langle e_0, a e_0 \rangle.$$

The distribution μ_{l,l^*} is now determined by $(n, m \geq 0)$

$$\varphi(l^n l^{*m}) = \langle e_0, l^n l^{*m} e_0 \rangle = \langle l^{*n} e_0, l^{*m} e_0 \rangle = \delta_{n0}\delta_{m0}.$$

Since l is not normal, this cannot be identified with a probability measure in the plane. However, if we consider

$$s := l + l^*,$$

then s is self-adjoint and its distribution corresponds to a probability measure on $\mathbb{R}$. This distribution will be one of the most important distributions in free probability, so we will determine it in the following.

THEOREM 1.8. *The distribution μ_s of $s := l + l^*$ with respect to φ is given by*

$$d\mu_s(t) = \frac{1}{2\pi}\sqrt{4-t^2}dt, \tag{10}$$

i.e., for all $n \geq 0$ we have:

$$\varphi(s^n) = \frac{1}{2\pi}\int_{-2}^{2} t^n\sqrt{4-t^2}dt. \tag{11}$$

The proof of Equation (11) will consist in calculating independently the two sides of that equation and check that they agree. The appearing moments can be determined explicitly and they are given by well-known combinatorial numbers, the so-called Catalan numbers.

DEFINITION 1.9. The numbers

$$C_k := \frac{1}{k}\binom{2k}{k-1} \qquad (k \in \mathbb{N}) \tag{12}$$

are called **Catalan numbers**

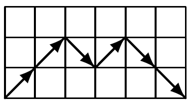
REMARKS 1.10. 1) The first Catalan numbers are

$$C_1 = 1, \qquad C_2 = 2, \qquad C_3 = 5, \qquad C_4 = 14, \qquad C_5 = 42, \qquad \dots .$$

2) The Catalan numbers count the numbers of a lot of different combinatorial objects. One of the most prominent ones are the so-called Catalan paths.

DEFINITION 1.11. A **Catalan path** (of lenght $2k$) is a path in the lattice $\mathbb{Z}^2$ which starts at $(0,0)$, ends at $(2k,0)$, makes steps of the form $(1,1)$ or $(1,-1)$, and never lies below the x-axis, i.e., all its points are of the form (i,j) with $j \geq 0$.

EXAMPLE 1.12. An Example of a Catalan path of length 6 is the following:



REMARKS 1.13. 1) The best way to identify the Catalan numbers is by their recurrence formula

$$C_k = \sum_{i=1}^{k} C_{i-1}C_{k-i}. \tag{13}$$

2) It is easy to see that the number D_k of Catalan paths of length $2k$ is given by the Catalan number C_k: Let Π be a Catalan path from $(0,0)$ to $(2k,0)$ and let $(2i,0)$ $(1 \leq i \leq k)$ be its first intersection with the

x-axis. Then it decomposes as $\Pi = \Pi_1 \cup \Pi_2$, where Π_1 is the part of Π from $(0,0)$ to $(2i,0)$ and Π_2 is the part of Π from $(2i,0)$ to $(2k,0)$. Π_2 is itself a Catalan path of lenght $2(k-i)$, thus there are D_{k-i} possibilities for Π_2. Π_1 is a Catalan path which lies strictly above the x-axis, thus $\Pi_1 = \{(0,0),(1,1),...,(2i-1,1),(2i,0)\} = (0,0) \cup \Pi_3 \cup (2i,0)$ and $\Pi_3 - (0,1)$ is a Catalan path from $(1,0)$ to $(2i-1,0)$, i.e. of length $2(i-1)$ – for which we have D_{i-1} possibilities. Thus we obtain

$$D_k = \sum_{i=1}^{k} D_{i-1} D_{k-i}.$$

Since this recurrence relation and the initial value $D_1 = 1$ determine the series $(D_k)_{k\in\mathbb{N}}$ uniquely it has to coincide with the Catalan numbers $(C_k)_{k\in\mathbb{N}}$.

Proof. We put

$$l^{(1)} := l \qquad \text{and} \qquad l^{(-1)} := l^*.$$

Then we have

$$\begin{aligned}
\varphi\big((l+l^*)^n\big) &= \varphi\big((l^{(1)}+l^{(-1)})^n\big)\\
&= \sum_{i_1,\dots,i_n\in\{-1,1\}} \varphi(l^{(i_n)}\dots l^{(i_1)})\\
&= \sum_{i_1,\dots,i_n\in\{-1,1\}} \langle e_0, l^{(i_n)}\dots l^{(i_1)} e_0\rangle
\end{aligned}$$

Now note that

$$\begin{aligned}
\langle e_0, l^{(i_n)}\dots l^{(i_1)} e_0\rangle &= \langle e_0, l^{(i_n)}\dots l^{(i_2)} e_{i_1}\rangle\\
&\qquad \text{if } i_1 = 1\\
&= \langle e_0, l^{(i_n)}\dots l^{(i_3)} e_{i_1+i_2}\rangle\\
&\qquad \text{if } i_1+i_2 \geq 0\\
&= \dots\\
&= \langle e_0, l^{(i_n)} e_{i_1+i_2+\cdots+i_{n-1}}\rangle\\
&\qquad \text{if } i_1+i_2+\cdots+i_{n-1} \geq 0\\
&= \langle e_0, e_{i_1+i_2+\cdots+i_n}\rangle\\
&\qquad \text{if } i_1+i_2+\cdots+i_n \geq 0\\
&= \delta_{0,i_1+i_2+\cdots+i_n}.
\end{aligned}$$

In all other cases we have $\langle e_0, l^{(i_n)} \dots l^{(i_1)} e_0\rangle = 0$. Thus

$$\begin{aligned}\varphi\big((l+l^*)^n\big) = \#\{(i_1,\dots,i_n) \,|\, & i_m = \pm 1, \\ & i_1+\dots+i_m \geq 0 \text{ for all } 1 \leq m \leq n, \\ & i_1+\dots+i_n = 0\}.\end{aligned}$$

In particular, for n odd we have:

$$\varphi(s^n) = 0 = \frac{1}{2\pi}\int_{-2}^{2} t^n \sqrt{4-t^2}\,dt.$$

Consider now $n = 2k$ even.
A sequence $(i_1,\dots,i_{2k})$ as above corresponds to a Catalan path: $i_m = +1$ corresponds to a step $(1,1)$ and $i_m = -1$ corresponds to a step $(1,-1)$. Thus we have shown that

$$\varphi\big((l+l^*)^{2k}\big) = C_k.$$

So it remains to show that also

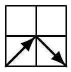
$$\frac{1}{2\pi}\int_{-2}^{2} t^{2k}\sqrt{4-t^2}\,dt = C_k. \tag{14}$$

This will be left to the reader □

EXAMPLES 1.14. Let us write down explicitly the Catalan paths corresponding to the second, fourth and sixth moment.
For $k = 1$ we have

$$\varphi\big((l+l^*)^2\big) = \varphi(l^*l)$$

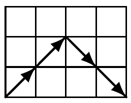
$$\hat{=}\,(+1,-1)$$



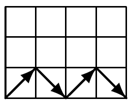
which shows that $C_1 = 1$.
For $k = 2$ we have

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

$$\hat{=}\,(+1,+1,-1,-1)$$



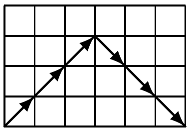
$$(+1,-1,+1,-1)$$



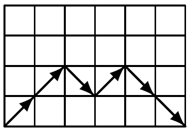
which shows that $C_2 = 2$.
For $k = 3$ we have

$$\begin{aligned}\varphi\big((l+l^*)^6\big) =&\varphi(l^*l^*l^*lll) + \varphi(l^*l^*ll^*ll) + \varphi(l^*ll^*l^*ll)\\ &+ \varphi(l^*l^*ll^*l) + \varphi(l^*ll^*ll^*l)\end{aligned}$$

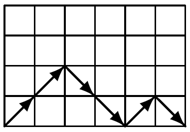
$$\hat{=} (+1, +1, +1, -1, -1, -1)$$



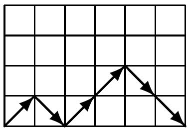
$$(+1, +1, -1, +1, -1, -1)$$



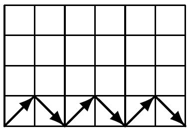
$$(+1, +1, -1, -1, +1, -1)$$



$$(+1, -1, +1, +1, -1, -1)$$



$$(+1, -1, +1, -1, +1, -1)$$



which shows that $C_3 = 5$.

DEFINITIONS 1.15. 1) A unitary element u in a $*$-probability space $(\mathcal{A}, \varphi)$ is called **Haar unitary**, if

$$\varphi(u^k) = \delta_{k0} \qquad \text{for all } k \in \mathbb{Z}. \tag{15}$$

2) A self-adjoint element $s = s^*$ in a $*$-probability space $(\mathcal{A}, \varphi)$ is called **semi-circular with radius** r $(r > 0)$, if its distribution is of the form

$$d\mu_s(t) = \begin{cases} \frac{2}{\pi r^2}\sqrt{r^2 - t^2}dt, & -r \leq t \leq r \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

REMARKS 1.16. 1) Hence $s := l + l^*$ is a semi-circular of radius 2.
2) Let s be a semi-circular of radius r. Then we have

$$\varphi(s^n) = \begin{cases} 0, & n \text{ odd} \\ (r/2)^n \cdot C_{n/2}, & n \text{ even.} \end{cases} \tag{17}$$

CHAPTER 2

Free random variables

Before we introduce our central notion of freeness let us recall the corresponding central notion of independence from classical probability theory – transferred to our algebraic frame.

Definition 2.1. Let $(\mathcal{A}, \varphi)$ be a probability space.
1) Subalgebras $\mathcal{A}_1, \dots, \mathcal{A}_m$ are called **independent**, if the subalgebras $\mathcal{A}_i$ commute – i.e., $ab = ba$ for all $a \in \mathcal{A}_i$ and all $b \in \mathcal{A}_j$ and all i, j with $i \neq j$ – and φ factorizes in the following way:

$$\varphi(a_1 \dots a_m) = \varphi(a_1) \dots \varphi(a_m) \tag{18}$$

for all $a_1 \in \mathcal{A}_1, \dots, a_m \in \mathcal{A}_m$.
2) Independence of random variables is defined by independence of the generated algebras; hence 'a and b independent' means nothing but a and b commute, $ab = ba$, and $\varphi(a^n b^m) = \varphi(a^n)\varphi(b^m)$ for all $n, m \geq 0$.

From a combinatorial point of view one can consider 'independence' as a special rule for calculating mixed moments of independent random variables from the moments of the single variables. 'Freeness' will just be another such specific rule.

Definition 2.2. Let $(\mathcal{A}, \varphi)$ be a probability space.
1) Let $\mathcal{A}_1, \dots, \mathcal{A}_m \subset \mathcal{A}$ be unital subalgebras. The subalgebras $\mathcal{A}_1, \dots, \mathcal{A}_m$ are called **free**, if $\varphi(a_1 \cdots a_k) = 0$ for all $k \in \mathbf{N}$ and $a_i \in \mathcal{A}_{j(i)}$ $(1 \leq j(i) \leq m)$ such that $\varphi(a_i) = 0$ for all $i = 1, \dots, k$ and such that neighbouring elements are from different subalgebras, i.e., $j(1) \neq j(2) \neq \cdots \neq j(k)$.
2) Let $\mathcal{X}_1, \dots, \mathcal{X}_m \subset \mathcal{A}$ be subsets of $\mathcal{A}$. Then $\mathcal{X}_1, \dots, \mathcal{X}_m$ are called free, if $\mathcal{A}_1, \dots, \mathcal{A}_m$ are free, where, for $i = 1, \dots, m$, $\mathcal{A}_i := \mathrm{alg}(1, \mathcal{X}_i)$ is the unital algebra generated by $\mathcal{X}_i$.
3) In particular, if the unital algebras $\mathcal{A}_i := \mathrm{alg}(1, a_i)$ $(i = 1, \dots, m)$ generated by elements $a_i \in \mathcal{A}$ $(i = 1, \dots, m)$ are free, then $a_1, \dots, a_m$ are called **free random variables**. If the $*$-algebras generated by the random variables $a_1, \dots, a_m$ are free, then we call $a_1, \dots, a_m$ **$*$-free**.

Remarks 2.3. 1) Note: the condition on the indices is only on consecutive ones; $j(1) = j(3)$ is allowed

3) Let us state more explicitly the requirement for free random variables: $a_1, \dots, a_m$ are free if we have:

$$\varphi\big(p_1(a_{j(1)}) \dots p_k(a_{j(k)})\big) = 0$$

for all polynomials $p_1, \dots, p_k$ in one variable and all $j(1) \neq j(2) \neq \dots \neq j(k)$, such that $\varphi(p_i(a_{j(i)})) = 0$ for all $i = 1, \dots, k$

4) Freeness of random variables is defined in terms of the generated algebras, but one should note that it extends also to the generated C^*- and von Neumann algebras. One has the following statement: Let $(\mathcal{A}, \varphi)$ be a C^*-probability space and let $a_1, \dots, a_m$ be $*$-free. Let, for each $i = 1, \dots, m$, $\mathcal{B}_i := C^*(1, a_i)$ be the unital C^*-algebra generated by a_i. Then $\mathcal{B}_1, \dots, \mathcal{B}_m$ are also free. (A similar statement holds for von Neumann algebras.)

5) Although not as obvious as in the case of 'independence', freeness is from a combinatorial point of view nothing but a very special rule for calculating joint moments of free variables out of the moments of the single variables. This is made explicit by the following lemma.

LEMMA 2.4. *Let $(\mathcal{A}, \varphi)$ be a probability space and let the unital subalgebras $\mathcal{A}_1, \dots, \mathcal{A}_m$ be free. Denote by $\mathcal{B}$ the algebra which is generated by all $\mathcal{A}_i$, $\mathcal{B} := alg(\mathcal{A}_1, \dots, \mathcal{A}_m)$. Then $\varphi|_{\mathcal{B}}$ is uniquely determined by $\varphi|_{\mathcal{A}_i}$ for all $i = 1, \dots, m$ and by the freeness condition.*

PROOF. Each element of B can be written as a linear combination of elements of the form $a_1 \dots a_k$ where $a_i \in \mathcal{A}_{j(i)}$ $(1 \leq j(i) \leq m)$. We can assume that $j(1) \neq j(2) \neq \dots \neq j(k)$. Let $a_1 \dots a_k \in \mathcal{B}$ be such an element. We have to show that $\varphi(a_1 \dots a_k)$ is uniquely determined by the $\varphi|_{\mathcal{A}_i}$ $(i = 1, \dots, m)$.

We prove this by induction on k. The case $k = 1$ is clear because $a_1 \in \mathcal{A}_{j(1)}$. In the general case we put

$$a_i^0 := a_i - \varphi(a_i)1 \in \mathcal{A}_i.$$

Then we have

$$\begin{aligned}\varphi(a_1 \dots a_k) &= \varphi\big((a_1^0 + \varphi(a_1)1) \dots (a_k^0 + \varphi(a_k)1)\big)\\ &= \varphi(a_1^0 \dots a_k^0) + rest,\end{aligned}$$

where

$$rest = \sum_{\substack{(p(1)<\dots<p(s))\dot\cup \\ \dot\cup(q(1)<\dots<q(k-s))=(1,\dots,k) \\ s<k}} \varphi(a_{p(1)}^0 \dots a_{p(s)}^0) \cdot \varphi(a_{q(1)}) \dots \varphi(a_{q(k-s)}).$$

Since $\varphi(a_i^0) = 0$ it follows $\varphi(a_1^0 \dots a_k^0) = 0$. On the other hand, all terms in $rest$ are of length smaller than k, and thus are uniquely determined by induction hypothesis. $\square$

EXAMPLES 2.5. Let a and b be free.
1) According to the definition of 'freeness' we have directly $\varphi(ab) = 0$ if $\varphi(a) = 0$ and $\varphi(b) = 0$. To calculate $\varphi(ab)$ in general we center our variables as in the proof of the lemma:

$$\begin{aligned}
0 &= \varphi\big((a - \varphi(a)1)(b - \varphi(b)1)\big) \\
&= \varphi(ab) - \varphi(a1)\varphi(b) - \varphi(a)\varphi(1b) + \varphi(a)\varphi(b)\varphi(1) \\
&= \varphi(ab) - \varphi(a)\varphi(b)
\end{aligned}$$

which implies

$$\varphi(ab) = \varphi(a)\varphi(b). \tag{19}$$

2) In the same way we write

$$\varphi\big((a^n - \varphi(a^n)1)(b^m - \varphi(b^m)1)\big) = 0,$$

implying

$$\varphi(a^n b^m) = \varphi(a^n)\varphi(b^m) \tag{20}$$

and

$$\varphi\big((a^{n_1} - \varphi(a^{n_1})1)(b^m - \varphi(b^m)1)(a^{n_2} - \varphi(a^{n_2})1)\big) = 0$$

implying

$$\varphi(a^{n_1} b^m a^{n_2}) = \varphi(a^{n_1+n_2})\varphi(b^m). \tag{21}$$

3) All the examples up to now yielded the same result as we would get for independent random variables. To see the difference between 'freeness' and 'independence' we consider now $\varphi(abab)$. Starting from

$$\varphi\big((a - \varphi(a)1)(b - \varphi(b)1)(a - \varphi(a)1)(b - \varphi(b)1)\big) = 0$$

one arrives finally at

$$\varphi(abab) = \varphi(aa)\varphi(b)\varphi(b) + \varphi(a)\varphi(a)\varphi(bb) - \varphi(a)\varphi(b)\varphi(a)\varphi(b). \tag{22}$$

3) Note that the above examples imply for free commuting variables, $ab = ba$, that at least one of them has vanishing variance, i.e., that $\varphi\big((a - \varphi(a)1)^2\big) = 0$ or $\varphi\big((b - \varphi(b)1)^2\big) = 0$. Indeed, let a and b be free and $ab = ba$. Then we have

$$\varphi(a^2)\varphi(b^2) = \varphi(a^2 b^2) = \varphi(abab) = \varphi(a^2)\varphi(b)^2 + \varphi(a)^2\varphi(b^2) - \varphi(a)^2\varphi(b)^2,$$

and hence

$$0 = \big(\varphi(a^2) - \varphi(a)^2\big)\big(\varphi(b^2) - \varphi(b)^2\big) = \varphi\big((a - \varphi(a)1)^2\big) \cdot \varphi\big((b - \varphi(b)1)^2\big)$$

which implies that at least one of the two factors has to vanish.
4) In particular, if a and b are classical random variables then they can only be free if at least one of them is almost surely constant. This shows that 'freeness' is really a non-commutative concept and cannot

be considered as a special kind of dependence between classical random variables.
5) A special case of the above is the following: If a is free from itself then we have $\varphi(a^2) = \varphi(a)^2$. If, furthermore, $a = a^*$ and φ faithful then this implies that a is a constant: $a = \varphi(a)1$.

REMARKS 2.6. Freeness behaves, similarly to independence, in a good way under special constructions.
1) Functions of free random variables are free: if a and b are free and f and g polynomials, then $f(a)$ and $g(b)$ are free, too. (If a and b are self-adjoint then the same is true for continuous functions in the C^*-case and for measurable functions in the W^*-case.)
3) Freeness behaves well under successive decompositions: Let $\mathcal{A}_1, \dots, \mathcal{A}_m$ be unital subalgebras of $\mathcal{A}$ and, for each $i = 1, \dots, m$, $\mathcal{B}_i^1, \dots, \mathcal{B}_i^{n_i}$ unital subalgebras of $\mathcal{A}_i$. Then we have:
i) If $(\mathcal{A}_i)_{i=1,\dots,m}$ are free in $\mathcal{A}$ and, for each $i = 1, \dots, m$, $(\mathcal{B}_i^j)_{j=1,\dots,n_i}$ are free in $\mathcal{A}_i$, then all $(\mathcal{B}_i^j)_{i=1,\dots,m;j=1,\dots,n_i}$ are free in $\mathcal{A}$.
ii) If all $(\mathcal{B}_i^j)_{i=1,\dots,m;j=1,\dots,n_i}$ are free in $\mathcal{A}$ and if, for each $i = 1, \dots, m$, $\mathcal{A}_i$ is as algebra generated by $\mathcal{B}_i^1, \dots, \mathcal{B}_i^{n_i}$, then $(\mathcal{A}_i)_{i=1,\dots,m}$ are free in $\mathcal{A}$.
4) Part (ii) of the last statement means that we can calculate mixed moments in free variables successively. For example, the statements that $\{a, b\}, c$ are free and that a, b are free is equivalent to a, b, c free. For the calculation of $\varphi(abc)$ this means that we can to it directly, by freeness of a, b, c, as

$$\varphi(abc) = \varphi(a)\varphi(b)\varphi(c)$$

or successively as

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

5) In the following lemma we will state the important fact that constant random variables are free from everything.

LEMMA 2.7. *Let $(\mathcal{A}, \varphi)$ be a probability space and $\mathcal{B} \subset \mathcal{A}$ a unital subalgebra. Then the subalgebras $\mathbb{C}1$ and $\mathcal{B}$ are free.*

PROOF. Consider $a_1 \dots a_k$ as in the definition of freeness and $k \geq 2$. ($k = 1$ is clear.) Then we have at least one $a_i \in \mathbb{C}1$ with $\varphi(a_i) = 0$. But this means $a_i = 0$, hence $a_1 \dots a_k = 0$ and thus $\varphi(a_1 \dots a_k) = 0$. $\square$

Part 2

Cumulants

CHAPTER 3

Motivation: Free central limit theorem

In this section, we want to motivate the appearance of special combinatorial objects in free probability theory.

1. One-dimensional case

DEFINITION 3.1. Let $(\mathcal{A}_N, \varphi_N)$ $(N \in \mathbb{N})$ and $(\mathcal{A}, \varphi)$ be probability spaces and consider

$$a_N \in \mathcal{A}_N \qquad (N \in \mathbb{N}), \qquad a \in \mathcal{A}.$$

We say that a_N **converges in distribution** towards a for $N \to \infty$ and denote this by

$$a_N \xrightarrow{\text{distr}} a,$$

if we have

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

REMARK 3.2. In classical probability theory, the usual form of convergence appearing, e.g., in central limit theorems is 'weak convergence'. If we can identify μ_{a_N} and μ_a with probability measures on $\mathbb{R}$ or $\mathbb{C}$, then weak convergence of μ_{a_N} towards μ_a means

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

If μ_a has compact support – as it is the case for a normal and bounded – then, by Stone-Weierstrass, this is equivalent the the above convergence in distribution.

REMARK 3.3. Let us recall the simplest version of the classical central limit theorem: 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 (i.e., $\mu_{a_r} = \mu_{a_p}$ for all r, p). 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

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

where γ is a normally distributed random variable of variance σ^2.
Explicitly, this means

$$\lim_{N\to\infty} \varphi\big(\big(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\big)^n\big) = \frac{1}{\sqrt{2\pi\sigma^2}} \int t^n e^{-t^2/2\sigma^2} dt \qquad \forall\, n \in \mathbb{N}. \tag{23}$$

THEOREM 3.4. **(free CLT: one-dimensional case)** *Let $(\mathcal{A}, \varphi)$ be a probability space and $a_1, a_2, \dots \in \mathcal{A}$ a sequence of free 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 σ^2.
Explicitly, this means

$$\lim_{N\to\infty} \varphi\big(\big(\frac{a_1 + \cdots + a_N}{\sqrt{N}}\big)^n\big) = \begin{cases} 0, & k \text{ odd} \\ C_k \sigma^n, & n = 2k \text{ even.} \end{cases}$$

PROOF. 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

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

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 $\pi = \{V_1, \dots, V_q\}$ of $\{1, \dots, n\}$ – i.e.,

$$\{1, \dots, n\} = \dot{\cup}_{i=1}^q V_q,$$

– which is determined by: $r(p) = r(q)$ if and only if p and q belong to the same block of π. We will write $(r(1), \dots, r(n)) \hat{=} \pi$ 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 $(r(1), \dots, r(n)) \hat{=} \pi$ by k_π.
For illustration, consider the following example:

$$k_{\{(1,3,4),(2,5),(6)\}} = \varphi(a_1 a_2 a_1 a_1 a_2 a_3) = \varphi(a_7 a_5 a_7 a_7 a_5 a_3).$$

Then we can continue the above calculation with

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

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 $\#\{(r(1),\dots,r(n)) \,\hat{=}\, \pi\}$. 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

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

because $\{a_{r(1)},\dots,\check{a}_r,\dots,a_{r(n)}\}$ is free from a_r. Thus only such partitions π contribute which have no singletons, i.e. $\pi = \{V_1,\dots,V_q\}$ with $\#V_m \geq 2$ for all $m = 1,\dots,q$.

On the other side,

$$\#\{(r(1),\dots,r(n)) \,\hat{=}\, \pi\} = N(N-1)\dots(N-q+1) \approx N^q$$

for $\pi = \{V_1,\dots,V_q\}$. Thus

$$\begin{aligned}
\lim_{N\to\infty} \varphi\big((\frac{a_1 + \cdots + a_N}{\sqrt{N}})^n\big) &= \lim_{N\to\infty} \sum_\pi \frac{\#\{(r(1),\dots,r(n)) \,\hat{=}\, \pi\}}{N^{n/2}} k_\pi \\
&= \lim_{N\to\infty} \sum_\pi N^{q-(n/2)} k_\pi \\
&= \sum_{\substack{\pi = \{V_1,\dots,V_{n/2}\} \\ \#V_m = 2\, \forall m = 1,\dots,n/2}} k_\pi
\end{aligned}$$

This means, that in the limit only pair partitions – i.e., partitions where each block V_m consists of exactly two elements – contribute. In particular, since there are no pair partitions for n odd, we see that the odd moments vanish in the limit:

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

Let now $n = 2k$ 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 π, $(r(1),\dots,r(n)) \,\hat{=}\, \pi$. 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 $m = 1, \ldots, n$, we have by the definition of freeness
$$k_\pi = \varphi(a_{r(1)} \ldots a_{r(n)}) = 0.$$
2. two consecutive indices coincide, i.e., $r(m) = r(m+1) = r$ for some $m = 1, \ldots, n-1$. Then we have
$$\begin{aligned} k_\pi &= \varphi(a_{r(1)} \ldots a_r a_r \ldots a_{r(n)}) \\ &= \varphi(a_{r(1)} \ldots a_{r(m-1)} a_{r(m+2)} \ldots a_{r(n)})\varphi(a_r a_r) \\ &= \varphi(a_{r(1)} \ldots a_{r(m-1)} a_{r(m+2)} \ldots a_{r(n)})\sigma^2 \end{aligned}$$

Iterating of these two possibilities leads finally to the following two possibilities:

1. π has the property that successively we can find consecutive indices which coincide. In this case we have $k_\pi = \sigma^k$.
2. π does not have this property, i.e., 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. We will call such a π **crossing**. In that case we have $k_\pi = 0$.

Thus we have shown

$$\lim_{N\to\infty} \varphi\big((\frac{a_1 + \cdots + a_N}{\sqrt{N}})^n\big) = D_n \sigma^n, \tag{24}$$

where

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

and it remains to show that $D_n = C_{n/2}$. We present two different proofs for this fact:

1. One possibility is to check that the D_n fulfill the recurrence relation of the Catalan numbers: Let $\pi = \{V_1, \ldots, V_k\}$ be a non-crossing pair partition. We denote by V_1 that block of π 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 have either $1 < V_j < m$ or $1 < m < V_j$ (in particular, m has to be even), i.e., π restricted to $\{2, \ldots, m-1\}$ is a non-crossing pair partition of $\{2, \ldots, m-1\}$ and π restricted to $\{m+1, \ldots, n\}$ is a non-crossing pair partition of $\{m+1, \ldots, n\}$. There exist D_{m-2} many non-crossing pair partitions of $\{2, \ldots, m-1\}$ and D_{n-m} many non-crossing pair partitions of $\{m+1, \ldots, n\}$. Both

these possibilities can appear independently from each other and m can run through all even numbers from 2 to n. Hence we get

$$D_n = \sum_{m=2,4,\ldots,n-2,n} D_{m-2} D_{n-m}.$$

This is the recurrence relation for the Catalan numbers. Since also $D_2 = 1 = C_1$, we obtain

$$D_n = C_k \qquad (\text{for } n = 2k \text{ even}).$$

2. Another possibility for proving $D_n = C_k$ is to present a bijection between non-crossing pair partitions and Catalan paths: We map a non-crossing pair partition π to a Catalan path $(i_1, \ldots, i_n)$ by

$$i_m = +1 \qquad \Longleftrightarrow \qquad m \text{ is the first element in some } V_j \in \pi$$
$$i_m = -1 \qquad \Longleftrightarrow \qquad m \text{ is the second element in some } V_j \in \pi$$

Consider the following example:

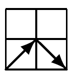
$$\pi = \{(1,4),(2,3),(5,6)\} \mapsto (+1,+1,-1,-1,+1,-1).$$

The proof that this mapping gives really a bijection between Catalan paths and non-crossing pair partitions is left to the reader.

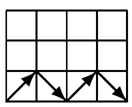
□

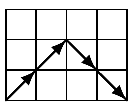
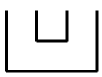
Example 3.5. Consider the following examples for the bijection between Catalan paths and non-crossing pair partitions.

- $n = 2$

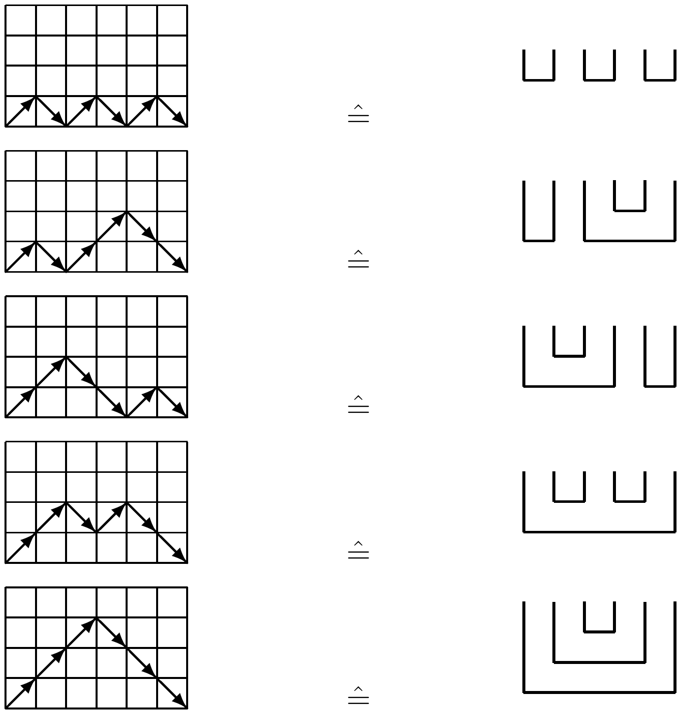
 $\hat{=}$ 

- $n = 4$

 $\hat{=}$ 

 $\hat{=}$ 

- $n = 6$



REMARKS 3.6. 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) As in the classical case the assumptions in the free central limit theorem can be weakened considerably. E.g., the assumption 'identically distributed' is essentially chosen to simplify the writing up; the same proof works also if one replaces this by

$$\sup_{i\in\mathbb{N}} |\varphi(a_i^n)| < \infty \qquad \forall \quad n \in \mathbb{N}$$

and

$$\sigma^2 := \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \varphi(a_i^2).$$

3) The same argumentation works also for a proof of the classical central limit theorem. The only difference appears when one determines the contribution of pair partitions. In contrast to the free case, all partitions give now the same factor. We will see in the forthcoming sections that this is a special case of the following general fact: the transition from classical to free probability theory consists on a combinatorial level in the transition from all partitions to non-crossing partitions.

2. Multi-dimensional case

Remark 3.7. One of the main advantages of our combinatorial approach to free probability theory is the fact that a lot of arguments can be extended from the one-dimensional to the more-dimensional case without great efforts. Let us 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. The calculation of this joint distribution goes in the same way as in the one-dimensional case. Namely, we now have to consider for all $n\in\mathbb{N}$ and all $i(1),\dots,i(n)\in I$

$$\varphi\Big((a_1^{(i(1))}+\cdots+a_N^{(i(1))})\cdots(a_1^{(i(n))}+\cdots+a_N^{(i(n))})\Big) = \sum_{1\leq r(1),\dots,r(n)\leq N} \varphi(a_{r(1)}^{(i(1))}\dots a_{r(n)}^{(i(n))}). \tag{25}$$

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 π 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 also depend on the tuple $(i(1),\dots,i(n))$ and will be denoted 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

$$\lim_{N\to\infty}\varphi\Big(\frac{a_1^{(i(1))}+\cdots+a_N^{(i(1))}}{\sqrt{N}}\cdots\frac{a_1^{(i(n))}+\cdots+a_N^{(i(n))}}{\sqrt{N}}\Big) = \sum_{\substack{\pi \text{ pair partition}\\ \text{of } \{1,\dots,n\}}} k_\pi[i(1),\dots,i(n)]. \tag{26}$$

It only remains to identify the contribution $k_\pi[i(1),\dots,i(n)]$ for a pair partition π. As before, the freeness assumption ensures that $k_\pi[i(1),\dots,i(n)]=0$ for crossing π. So consider finally non-crossing π. Remember that in the case of a non-crossing π 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 π).

DEFINITION 3.8. Let $(\mathcal{A}_N,\varphi_N)$ $(N\in\mathbb{N})$ and $(\mathcal{A},\varphi)$ be probability spaces. Let I be an index set and consider for each $i\in I$ random variables $a_N^{(i)}\in\mathcal{A}_N$ and $a_i\in\mathcal{A}$. We say that $(a_N^{(i)})_{i\in I}$ **converges in distribution** towards $(a_i)_{i\in I}$ and denote this by

$$
(a_N^{(i)})_{i\in I} \xrightarrow{\text{distr}} (a_i)_{i\in I},
$$

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

$$
\lim_{N\to\infty}\varphi_N(a_N^{(i(1))}\dots a_N^{(i(n))}) = \varphi(a_{i(1)}\dots a_{i(n)}). \tag{27}
$$

THEOREM 3.9. **(free CLT: multi-dimensional case)** *Let $(\mathcal{A},\varphi)$ be a probability space and $(a_1^{(i)})_{i\in I},(a_2^{(i)})_{i\in I},\dots\subset\mathcal{A}$ a sequence of free 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

$$
\Big(\frac{a_1^{(i)}+\dots+a_N^{(i)}}{\sqrt{N}}\Big)_{i\in I} \xrightarrow{\textit{distr}} (s_i)_{i\in I}, \tag{28}
$$

where the joint distribution of the family $(s_i)_{i\in I}$ is, for all $n\in\mathbb{N}$ and all $i(1),\dots,i(n)\in I$, given by

$$
\varphi(s_{i(1)}\dots s_{i(n)}) = \sum_{\substack{\pi \text{ non-crossing pair partition}\\ \text{of } \{1,\dots,n\}}} k_\pi[s_{i(1)},\dots,s_{i(n)}], \tag{29}
$$

where

$$k_\pi[s_{i(1)},\dots,s_{i(n)}] = \prod_{(p,q)\in\pi} c_{i(p)i(q)}. \tag{30}$$

NOTATION 3.10. Since $(s_i)_{i\in I}$ gives the more-dimensional generalization of a semi-circular variable, we will call $(s_i)_{i\in I}$ a **semi-circular family of covariance** $(c_{ij})_{i,j\in I}$. It is usual terminology in free probability theory to denote by 'semi-circular family' the more special situation where the covariance is diagonal in i, j, i.e. where all s_i are free (see the following proposition).

PROPOSITION 3.11. *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 families $\{s_i\}_{i\in I_p}$ $(p = 1,\dots,d)$ are free*
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 $\{s_i\}_{i\in I_p}$ $(p = 1,\dots,d)$ are free and consider $i \in I_p$ and $j \in I_q$ with $p \neq q$. Then the freeness 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 they have the same distribution as the $s_r^{(i)}$ 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 free. 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

$$\Big(\frac{a_1^{(i)} + \cdots + a_N^{(i)}}{\sqrt{N}}\Big)_{i\in I} \xrightarrow{\text{distr}} (s_i)_{i\in I}, \tag{31}$$

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 free; since freeness passes over to the limit we get the wanted result that the sets $\{s_i\}_{i\in I_p}$ $(p = 1,\dots,d)$ are free. □

REMARK 3.12. The conclusion we draw from the above considerations is that non-crossing pair partitions appear quite naturally in free probability. If we would consider more general limit theorems (e.g., free Poisson) then all kind of non-crossing partitions would survive in the limit. The main feature of freeness in this combinatorial approach (in particular, in comparision with independence) is reflected by the property 'non-crossing'. In the next sections we will generalize to arbitrary distributions what we have learned from the case of semi-circular families, namely:

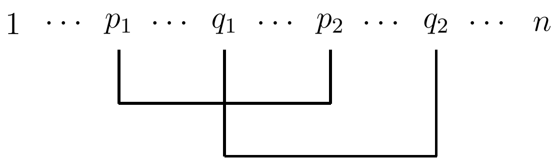
- it seems to be canonical to write moments as sums over non-crossing partitions
- the summands k_π reflect freeness quite clearly, since $k_\pi[s_{i(1)}, \dots, s_{i(n)}]$ vanishes if the blocks of π couple elements which are free.

CHAPTER 4

Free Cumulants

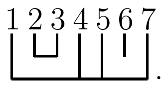
1. Non-crossing partitions

Definitions 4.1. Fix $n \in \mathbb{N}$. We call $\pi = \{V_1, \dots, V_r\}$ a **partition** of $S = (1, \dots, n)$ if and only if the V_i $(1 \le i \le r)$ are pairwisely disjoint, non-void tuples such that $V_1 \cup \dots \cup V_r = S$. We call the tuples $V_1, \dots, V_r$ the **blocks** of π. The number of components of a block V is denoted by $|V|$. Given two elements p and q with $1 \le p, q \le n$, we write $p \sim_\pi q$, if p and q belong to the same block of π.
2) A partition π is called **non-crossing**, if the following situation does not occur: There exist $1 \le p_1 < q_1 < p_2 < q_2 \le n$ such that $p_1 \sim_\pi p_2 \not\sim_\pi q_1 \sim_\pi q_2$:

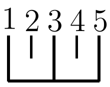


The set of all non-crossing partitions of $(1, \dots, n)$ is denoted by $NC(n)$.

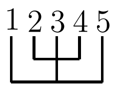
Remarks 4.2. 1) We get a linear graphical representation of a partition π by writing all elements $1, \dots, n$ in a line, supplying each with a vertical line under it and joining the vertical lines of the elements in the same block with a horizontal line.
Example: A partition of the tuple $S = (1, 2, 3, 4, 5, 6, 7)$ is

$\pi_1 = \{(1, 4, 5, 7), (2, 3), (6)\}$ $\hat{=}$ .

The name 'non-crossing' becomes evident in such a representation; e.g., let $S = \{1, 2, 3, 4, 5\}$. Then the partition

$\pi = \{(1, 3, 5), (2), (4)\}$ $\hat{=}$ 

is non-crossing, whereas

$\pi = \{(1, 3, 5), (2, 4)\}$ $\hat{=}$ 

is crossing.
2) By writing a partition π in the form $\pi = \{V_1, \dots, V_r\}$ we will always assume that the elements within each block V_i are ordered in increasing order.
3) In most cases the following recursive description of non-crossing partitions is of great use: a partition π ist non-crossing if and only if at least one block $V \in \pi$ is an interval and $\pi \backslash V$ is non-crossing; i.e. $V \in \pi$ has the form

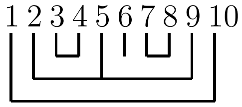
$$V = (k, k+1, \dots, k+p) \qquad \text{for some } 1 \le k \le n \text{ and } p \ge 0,\ k+p \le n$$

and we have

$$\pi \backslash V \in NC(1, \dots, k-1, k+p+1, \dots, n) \cong NC(n-(p+1)).$$

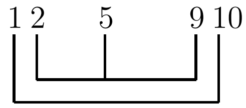
Example: The partition

$$\{(1,10),(2,5,9),(3,4),(6),(7,8)\} \qquad \hat{=}$$



can, by successive removal of intervals, be reduced to

$$\{(1,10),(2,5,9)\} \qquad \hat{=}$$

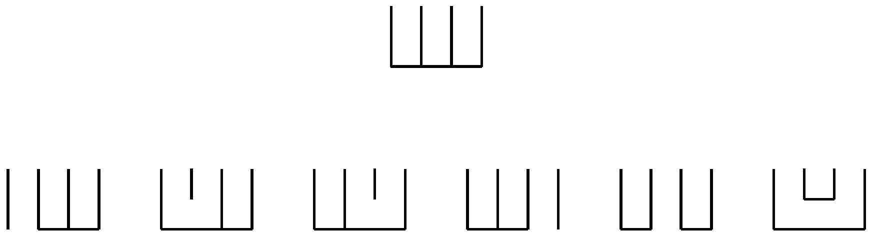


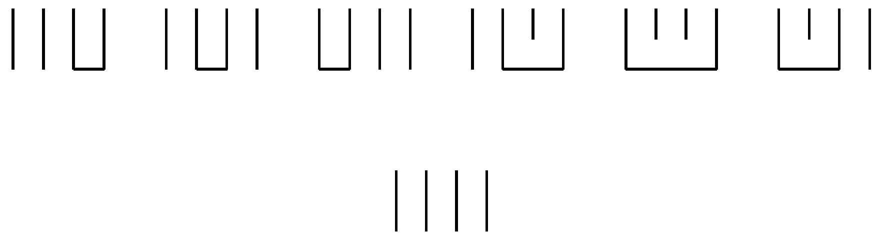
and finally to

$$\{(1,10)\} \qquad \hat{=}$$

4) In the same way as for $(1, \dots, n)$ one can introduce non-crossing partitions $NC(S)$ for each finite linearly ordered set S. Of course, $NC(S)$ depends only on the number of elements in S. In our investigations, non-crossing partitions will appear as partitions of the index set of products of random variables $a_1 \cdots a_n$. In such a case, we will also sometimes use the notation $NC(a_1, \dots, a_n)$. (If some of the a_i are equal, this might make no rigorous sense, but there should arise no problems by this.)

Example 4.3. The following pircture shows $NC(4)$.





NOTATIONS 4.4. 1) If S is the union of two disjoint sets S_1 and S_2 then, for $\pi_1 \in NC(S_1)$ and $\pi_2 \in NC(S_2)$, we let $\pi_1 \cup \pi_2$ be that partition of S which has as blocks the blocks of π_1 and the blocks of π_2. Note that $\pi_1 \cup \pi_2$ is not automatically non-crossing.
2) Let W be a union of some of the blocks of π. Then we denote by $\pi|_W$ the restriction of π to W, i.e.

$$\pi|_W := \{V \in \pi \mid V \subset W\} \in NC(W). \tag{32}$$

3) Let $\pi, \sigma \in NC(n)$ be two non-crossing partitions. We write $\sigma \leq \pi$, if each block of σ is completely contained in one block of π. Hence, we obtain σ out of π by refining the block structure. For example, we have

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

The partial order $\leq$ induces a lattice structure on $NC(n)$. In particular, given two non-crossing partitions $\pi, \sigma \in NC(n)$, we have their **join** $\pi \vee \sigma$ – which is the unique smallest $\tau \in NC(n)$ such that $\tau \geq \pi$ and $\tau \geq \sigma$ – and their **meet** $\pi \wedge \sigma$ – which is the unique biggest $\tau \in NC(n)$ such that $\tau \leq \pi$ and $\tau \leq \sigma$.
4) The maximum of $NC(n)$ – the partition which consists of one block with n components – is denoted by 1_n. The partition consisting of n blocks, each of which has one component, is the minimum of $NC(n)$ and denoted by 0_n.
5) The lattice $NC(n)$ is self-dual and there exists an important anti-isomorphism $K : NC(n) \to NC(n)$ implementing this self-duality. This complementation map K is defined as follows: Let π be a non-crossing partition of the numbers $1, \dots, n$. Furthermore, we consider numbers $\bar{1}, \dots, \bar{n}$ with all numbers ordered in an alternating way

$$1\,\bar{1}\,2\,\bar{2}\dots n\,\bar{n}\,.$$

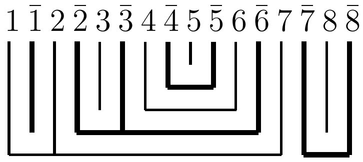
The **complement** $K(\pi)$ of $\pi \in NC(n)$ is defined to be the biggest $\sigma \in NC(\bar{1}, \dots, \bar{n}) \hat{=} NC(n)$ with

$$\pi \cup \sigma \in NC(1, \bar{1}, \dots, n, \bar{n})\,.$$

Example: Consider the partition $\pi := \{(1,2,7),(3),(4,6),(5),(8)\} \in NC(8)$. For the complement $K(\pi)$ we get

$$K(\pi) = \{(1),(2,3,6),(4,5),(7,8)\},$$

as can be seen from the graphical representation:



.

Non-crossing partitions and the complementation map were introduced by Kreweras.

2. Free cumulants

Definition 4.5. Let $(\mathcal{A}, \varphi)$ be a probability space, i.e. $\mathcal{A}$ is a unital algebra and $\varphi : \mathcal{A} \to \mathbb{C}$ a unital linear functional. Then we define corresponding **(free) cumulants** $(k_n)_{n\in\mathbb{N}}$

$$\begin{aligned} k_n : \qquad \mathcal{A}^n &\to \mathbf{C}, \\ (a_1, \dots, a_n) &\mapsto k_n(a_1, \dots, a_n) \end{aligned}$$

indirectly by the following system of equations:

$$\varphi(a_1 \cdots a_n) = \sum_{\pi \in NC(n)} k_\pi [a_1, \dots, a_n] \qquad (a_1, \dots, a_n \in \mathcal{A}), \tag{33}$$

where k_π splits in a multiplicative way in a product of cumulants according to the block structure of π, i.e.

$$k_\pi[a_1, \dots, a_n] := \prod_{i=1}^{r} k_{|V_i|}(a_{i,1}, \dots, a_{i,|V_i|}) \tag{34}$$

for a partition $\pi = \{V_1, \dots, V_r\} \in NC(n)$ consisting of r blocks of the form $V_i = (a_{i,1}, \dots, a_{i,|V_i|})$.

Remark 4.6. The defining relation (2) expresses the moment $\varphi(a_1 \cdots a_n)$ in terms of cumulants, but by induction this can also be resolved for giving the cumulants uniquely in terms of moments:

$$k_n(a_1, \dots, a_n) = \varphi(a_1 \cdots a_n) - \sum_{\substack{\pi \in NC(n) \\ \pi \neq 1_n}} k_\pi [a_1, \dots, a_n] \tag{35}$$

Since, by induction, we know all cumulants of smaller order, i.e., all $k_\pi[a_1, \dots, a_n]$ for $\pi \in NC(n)$ with $\pi \neq 1_n$, this leads to an expression for k_n in terms of moments. Abstractly, this is the so-called Möbius inversion of relation (2) and has the following form

$$k_n(a_1,\dots,a_n)=\sum_{\pi\in NC(n)}\mu(\pi,1_n)\varphi_\pi[a_1,\dots,a_n], \tag{36}$$

where μ is the Möbius function of the lattice of non-crossing partitions and where φ_π is defined in the same multiplicative way as k_π if we put $\varphi_n(a_1,\dots,a_n):=\varphi(a_1\cdots a_n)$. We will first give some examples for this connection between moments and cumulants and then say something more on the general theory of Möbius inversion.

Examples 4.7. Let us give the concrete form of $k_n(a_1,\dots,a_n)$ for small values of n.
1) $n=1$:

$$k_1(a_1)=\varphi(a_1)\,. \tag{37}$$

2) $n=2$: The only partition $\pi\in NC(2),\pi\neq 1_2$ is $\vert\,\vert$. So we get

$$\begin{aligned} k_2(a_1,a_2) &= \varphi(a_1a_2)-k_{\vert\,\vert}[a_1,a_2]\\ &= \varphi(a_1a_2)-k_1(a_1)k_1(a_2)\\ &= \varphi(a_1a_2)-\varphi(a_1)\varphi(a_2). \end{aligned}$$

Using the notation φ_π we can also write this as

$$k_2(a_1,a_2)=\varphi_{\sqcup}[a_1,a_2]-\varphi_{\vert\,\vert}[a_1,a_2]. \tag{38}$$

3) $n=3$: We have to subtract from $\varphi(a_1a_2a_3)$ the terms k_π for all partitions in $NC(3)$ except 1_3, i.e., for the following partitions:

$$\vert\sqcup\,,\qquad \sqcup\vert\,,\qquad \dot{\sqcup}\,,\qquad \vert\,\vert\,\vert\,.$$

With this we obtain:

$$\begin{aligned} k_3(a_1,a_2,a_3) &= \varphi(a_1a_2a_3)-k_{\vert\sqcup}[a_1,a_2,a_3]-k_{\sqcup\vert}[a_1,a_2,a_3]\\ &\quad -k_{\dot{\sqcup}}[a_1,a_2,a_3]-k_{\vert\vert\vert}[a_1,a_2,a_3]\\ &= \varphi(a_1a_2a_3)-k_1(a_1)k_2(a_2,a_3)-k_2(a_1,a_2)k_1(a_3)\\ &\quad -k_2(a_1,a_3)k_1(a_2)-k_1(a_1)k_1(a_2)k_1(a_3)\\ &= \varphi(a_1a_2a_3)-\varphi(a_1)\varphi(a_2a_3)-\varphi(a_1a_2)\varphi(a_3)\\ &\quad -\varphi(a_1a_3)\varphi(a_2)+2\varphi(a_1)\varphi(a_2)\varphi(a_3). \end{aligned}$$

Again we can write this in Möbius inverted form by using the notation φ_π.

$$\begin{aligned} k_3(a_1,a_2,a_3) &= \varphi_{\sqcup\!\sqcup}[a_1,a_2,a_3]-\varphi_{\vert\sqcup}[a_1,a_2,a_3]\\ &\quad -\varphi_{\sqcup\vert}[a_1,a_2,a_3]-\varphi_{\dot{\sqcup}}[a_1,a_2,a_3]\\ &\quad +2\varphi_{\vert\vert\vert}[a_1,a_2,a_3]. \end{aligned} \tag{39}$$

4) For $n = 4$ we consider the special case where all $\varphi(a_i) = 0$. Then we have

$$k_4(a_1, a_2, a_3, a_4) = \varphi(a_1a_2a_3a_4) - \varphi(a_1a_2)\varphi(a_3a_4) - \varphi(a_1a_4)\varphi(a_2a_3). \tag{40}$$

EXAMPLE 4.8. Our results from the last section show that the cumulants of a semi-circular variable of variance σ^2 are given by

$$k_n(s, \dots, s) = \delta_{n2}\sigma^2. \tag{41}$$

More generally, for a semi-circular family $(s_i)_{i\in I}$ of covariance $(c_{ij})_{i\in I}$ we have

$$k_n(s_{i(1)}, \dots, s_{i(n)}) = \delta_{n2}c_{i(1)i(2)}. \tag{42}$$

3. Posets and Möbius inversion

REMARK 4.9. We will now make some general remarks about the concept of Möbius inversion and Möbius function. This is a unifying concept in modern combinatorics because it provides a common frame for quite a lot of situations.

The general frame is a **poset**, i.e. a partially ordered set. We will restrict to the finite case, so we consider a finite set P with a partial order $\leq$. Furthermore let two functions $f, g : P \to \mathbb{C}$ be given which are connected as follows:

$$f(\pi) = \sum_{\substack{\sigma\in P\\ \sigma\leq\pi}} g(\sigma).$$

This is a quite common situation and one would like to invert this relation. This is indeed possible in a universal way, in the sense that one gets for the inversion a formula of a similar kind which involves also another function, the so-called **Möbius function** μ. This μ, however, does not depend on f and g, but only on the poset P.

PROPOSITION 4.10. *Let P be a finite poset. Then there exists a function $\mu : P \times P \to \mathbb{C}$ such that for all functions $f, g : P \to \mathbb{C}$ the following two statements are equivalent:*

$$f(\pi) = \sum_{\substack{\sigma\in P\\ \sigma\leq\pi}} g(\sigma) \qquad \forall \pi \in P \tag{43}$$

$$g(\pi) = \sum_{\substack{\sigma\in P\\ \sigma\leq\pi}} f(\sigma)\mu(\sigma, \pi) \qquad \forall \pi \in P \tag{44}$$

If we also require $\mu(\sigma, \pi) = 0$ if $\sigma \not\leq \pi$, then μ is uniquely determined.

PROOF. We try to define μ inductively by writing our relation $f(\pi) = \sum_{\sigma\leq\pi} g(\sigma)$ in the form

$$f(\pi) = g(\pi) + \sum_{\tau<\pi} g(\tau),$$

which gives

$$g(\pi) = f(\pi) - \sum_{\tau<\pi} g(\tau).$$

If we now assume that we have the inversion formula for $\tau < \pi$ then we can continue with

$$\begin{aligned} g(\pi) &= f(\pi) - \sum_{\tau<\pi}\sum_{\sigma\leq\tau} f(\sigma)\mu(\sigma,\tau) \\ &= f(\pi) - \sum_{\sigma<\pi} f(\sigma)\big(\sum_{\sigma\leq\tau<\pi} \mu(\sigma,\tau)\big). \end{aligned}$$

This shows that μ has to fulfill

$$\mu(\pi,\pi) = 1 \tag{45}$$

$$\mu(\sigma,\pi) = -\sum_{\sigma\leq\tau<\pi} \mu(\sigma,\tau) \qquad (\sigma < \pi). \tag{46}$$

On the other hand this can be used to define $\mu(\cdot,\pi)$ recursively by induction on the length of the interval $[\sigma,\pi] := \{\tau \mid \sigma \leq \tau \leq \pi\}$. This defines $\mu(\sigma,\pi)$ uniquely for all $\sigma \leq \pi$. If we also require μ to vanish in all other cases, then it is uniquely determined. □

EXAMPLES 4.11. 1) The classical example which gave the name to the Möbius inversion is due to Möbius (1832) and comes from number theory: Möbius showed that a relation $f(n) = \sum_{m|n} g(n)$ – where n and m are integer numbers and $m \mid n$ means that m is a divisor of n – can be inverted in the form $g(n) = \sum_{m|n} f(m)\mu(m/n)$ where μ is now the classical Möbius function given by

$$\mu(n) = \begin{cases} (-1)^k, & \text{if } n \text{ is the product of } k \text{ distinct primes} \\ 0, & \text{otherwise.} \end{cases} \tag{47}$$

2) Our free cumulants fit into this frame as follows: We start with a functional φ on our algebra $\mathcal{A}$ and this gives raise to a sequence of multilinear functionals $(\varphi_n)_{n\in\mathbb{N}}$ on $\mathcal{A}$ via

$$\varphi_n(a_1,\dots,a_n) := \varphi(a_1\cdots a_n). \tag{48}$$

We extend these to a corresponding function on non-crossing partitions in a multiplicative way by $(a_1, \dots, a_n \in \mathcal{A})$

$$\varphi(\pi) = \varphi_\pi[a_1, \dots, a_n] := \prod_{V \in \pi} \varphi(\pi|_V), \tag{49}$$

where we used the notation

$$\varphi(\pi|_V) := \varphi_s(a_{i_1}, \dots, a_{i_s}) \qquad \text{for } V = (i_1, \dots, i_s) \in \pi. \tag{50}$$

(To be precise, we have defined a function on $NC(n) \times \mathcal{A}^n$; however, the $a_1, \dots, a_n$ play no role for making the Möbius inversion; they are a kind of dummy variables, which have to be put at the right positions.) Furthermore, we have the Möbius function μ on the lattice of non-crossing partitions. Then we define our free cumulants as follows:

$$k(\pi) = k_\pi[a_1, \dots, a_n] := \sum_{\sigma \leq \pi} \varphi(\sigma)\mu(\sigma, \pi) \qquad (\pi \in NC(N)). \tag{51}$$

In order to make this consistent with our Definition 4.5 it only remains to see that also the $k(\pi)$ defined in this way factorize in a multiplicative way according to the block structure of π. This relies on the fact that also the Möbius function is multiplicative (see next lemma), and can be seen as follows:

$$\begin{aligned}
k(\pi) &= \sum_{\sigma \leq \pi} \varphi(\sigma)\mu(\sigma, \pi) \\
&= \sum_{(\bigcup_{V \in \pi} \sigma|_V) \leq \pi} \prod_{V \in \pi} \varphi(\sigma|_V)\mu(\sigma|_V, \pi|_V) \\
&= \sum_{\sigma|_V \leq \pi|_V \; (V \in \pi)} \prod_{V \in \pi} \varphi(\sigma|_V)\mu(\sigma|_V, \pi|_V) \\
&= \prod_{V \in \pi} \Big(\sum_{\sigma_V \leq \pi|_V} \varphi(\sigma_V)\mu(\sigma_V, \pi|_V) \Big) \\
&= \prod_{V \in \pi} k(\pi|_V)
\end{aligned}$$

3) Our examples from 4.6 give us the following values for the Möbius function on non-crossing partitions:

$$\mu(\vert\,\vert, \sqcup) = -1 \tag{52}$$

$$\mu(\vert\sqcup, \sqcup\!\sqcup) = -1 \tag{53}$$

$$\mu(\sqcup\vert, \sqcup\!\sqcup) = -1 \tag{54}$$

$$\mu(\dot{\sqcup}, \sqcup\!\sqcup) = -1 \tag{55}$$

$$\mu(\vert\vert\vert, \sqcup\!\sqcup) = 2 \tag{56}$$

LEMMA 4.12. *The Möbius function μ on non-crossing partitions is multiplicative, i.e. for $\sigma \leq \pi \leq \omega$ we have*

$$\mu(\sigma, \pi) = \prod_{V \in \omega} \mu(\sigma|_V, \pi|_V). \tag{57}$$

PROOF. We do this by induction on the length of the interval $[\sigma, \pi]$; namely

$$\mu(\pi, \pi) = 1 = \prod_{V \in \omega} 1 = \prod_{V \in \omega} \mu(\pi|_V, \pi|_V)$$

and for $\sigma < \pi$

$$\begin{aligned}
\mu(\sigma, \pi) &= -\sum_{\sigma \leq \tau < \pi} \mu(\sigma, \tau) \\
&= -\sum_{\sigma \leq \tau < \pi} \prod_{V \in \omega} \mu(\sigma|_V, \tau|_V) \\
&= -\sum_{\sigma \leq \tau \leq \pi} \prod_{V \in \omega} \mu(\sigma|_V, \tau|_V) + \prod_{V \in \omega} \mu(\sigma|_V, \pi|_V) \\
&= -\prod_{V \in \omega} \Big(\sum_{\sigma|_V \leq \tau_V \leq \pi|_V} \mu(\sigma|_V, \tau_V) \Big) + \prod_{V \in \omega} \mu(\sigma|_V, \pi|_V) \\
&= 0 + \prod_{V \in \omega} \mu(\sigma|_V, \pi|_V),
\end{aligned}$$

where we used in the last step again the reccurence relation (15) for the Möbius function. Note that at least for one $V \in \omega$ we have $\sigma|_V < \pi|_V$. ☐

Another way to look at the cumulants k_n for $n \geq 2$ is that they organize in a special way the information about how much φ ceases to be a homomorhpism.

PROPOSITION 4.13. *Let $(k_n)_{n \geq 1}$ be the cumulants corresponding to φ. Then φ is a homomorphism if and only if k_n vanishes for all $n \geq 2$.*

PROOF. Let φ be a homomorphism. Note that this means $\varphi_\sigma = \varphi_{0_n}$ for all $\sigma \in NC(n)$. Thus we get

$$k_n = \sum_{\sigma \leq 1_n} \varphi_\sigma \mu(\sigma, 1_n) = \varphi(0_n) \sum_{0_n \leq \sigma \leq 1_n} \mu(\sigma, 1_n),$$

which is, by the recurrence relation (15) for the Möbius function, equal to zero if $0_n \neq 1_n$, i.e. for $n \geq 2$.
The other way around, if k_2 vanishes, then we have for all $a_1, a_2 \in \mathcal{A}$

$$0 = k_2(a_1, a_2) = \varphi(a_1 a_2) - \varphi(a_1)\varphi(a_2).$$

☐

REMARK 4.14. In particular, this means that on constants only the first order cumulants are different from zero:

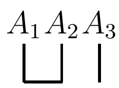
$$k_n(1, \ldots, 1) = \delta_{n1}. \tag{58}$$

CHAPTER 5

Cumulants with products as entries

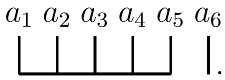
We have defined cumulants as some special polynomials in the moments. Up to now it is neither clear that such objects have any nice properties per se nor that they are of any use for the description of freeness. The latter point will be addressed in the next section, whereas here we focus on properties of our cumulants with respect to the algebraic structure of our underlying algebra $\mathcal{A}$. Of course, the linear structure is clear, because our cumulants are multi-linear functionals. Thus it remains to see whether there is anything to say about the relation of the cumulants with the multiplicative structure of the algebra. The crucial property in a multiplicative context is associativity. On the level of moments this just means that we can put brackets arbitrarily; for example we have $\varphi_2(a_1a_2, a_3) = \varphi((a_1a_2)a_3) = \varphi(a_1(a_2a_3)) = \varphi_2(a_1, a_2a_3)$. But the corresponding statement on the level of cumulants is, of course, not true, i.e. $k_2(a_1a_2, a_3) \neq k_2(a_1, a_2a_3)$ in general. However, there is still a treatable and nice formula which allows to deal with free cumulants whose entries are products of random variables. This formula will be presented in this section and will be fundamental for our forthcoming investigations.

Notation 5.1. The general frame for our theorem is the following: Let an increasing sequence of integers be given, $1 \leq i_1 < i_2 < \cdots < i_m := n$ and let $a_1, \dots, a_n$ be random variables. Then we define new random variables A_j as products of the given a_i according to $A_j := a_{i_{j-1}+1} \cdots a_{i_j}$ (where $i_0 := 0$). We want to express a cumulant $k_\tau[A_1, \dots, A_m]$ in terms of cumulants $k_\pi[a_1, \dots, a_n]$. So let τ be a non-crossing partition of the m-tuple $(A_1, \dots, A_m)$. Then we define $\hat{\tau} \in NC(a_1, \dots, a_n)$ to be that partition which we get from τ by replacing each A_j by $a_{i_{j-1}+1}, \dots, a_{i_j}$, i.e., for a_i being a factor in A_k and a_j being a factor in A_l we have $a_i \sim_{\hat{\tau}} a_j$ if and only if $A_k \sim_\tau A_l$.
For example, for $n = 6$ and $A_1 := a_1a_2$, $A_2 := a_3a_4a_5$, $A_3 := a_6$ and

$\tau = \{(A_1, A_2), (A_3)\}$ $\qquad \hat{=} \qquad$ 

we get

$$\hat\tau = \{(a_1, a_2, a_3, a_4, a_5), (a_6)\} \qquad \hat{=}$$



Note also in particular, that $\hat\tau = 1_n$ if and only if $\tau = 1_m$.

THEOREM 5.2. *Let $m \in \mathbb{N}$ and $1 \le i_1 < i_2 < \cdots < i_m := n$ be given. Consider random variables $a_1, \dots, a_n$ and put $A_j := a_{i_{j-1}+1} \cdots a_{i_j}$ for $j = 1, \dots, m$ (where $i_0 := 0$). Let τ be a partition in $NC(A_1, \dots, A_m)$.*
Then the following equation holds:

$$k_\tau[a_1 \cdots a_{i_1}, \dots, a_{i_{m-1}+1} \cdots a_{i_m}] = \sum_{\substack{\pi \in NC(n) \\ \pi \vee \sigma = \hat\tau}} k_\pi[a_1, \dots, a_n]\,, \tag{59}$$

where $\sigma \in NC(n)$ is the partition $\sigma = \{(a_1, \dots, a_{i_1}), \dots, (a_{i_{m-1}+1}, \dots, a_{i_m})\}$.

REMARKS 5.3. 1) In all our applications we will only use the special case of Theorem 2.2 where $\tau = 1_m$. Then the statement of the theorem is the following: Consider $m \in \mathbb{N}$, an increasing sequence $1 \le i_1 < i_2 < \cdots < i_m := n$ and random variables $a_1, \dots, a_n$. Put $\sigma := \{(a_1, \dots, a_{i_1}), \dots, (a_{i_{m-1}+1}, \dots, a_{i_m})\}$. Then we have:

$$k_m[a_1 \cdots a_{i_1}, \dots, a_{i_{m-1}+1} \cdots a_{i_m}] = \sum_{\substack{\pi \in NC(n) \\ \pi \vee \sigma = 1_n}} k_\pi[a_1, \dots, a_n]\,. \tag{60}$$

2) Of special importance will be the case where only one of the arguments is a product of two elements, i.e. $k_{n-1}(a_1, \dots, a_{m-1}, a_m \cdot a_{m+1}, a_{m+2}, \dots, a_n)$. In that case $\sigma = \{(1), (2), \dots, (m, m+1), \dots, (n)\}$ and it is easily seen that the partitions $\pi \in NC(n)$ with the property $\pi \vee \sigma = 1_n$ are either 1_n or those non-crossing partitions which consist of two blocks such that one of them contains m and the other one contains $m+1$. This leads to the formula

$$\begin{aligned} k_{n-1}(a_1, \dots, a_m \cdot a_{m+1}, \dots, a_n) &= \\ &= k_n(a_1, \dots, a_m, a_{m+1}, \dots, a_n) + \sum_{\substack{\pi \in NC(n) \\ |\pi| = 2,\, m \not\sim_\pi m+1}} k_\pi[a_1, \dots, a_n]. \end{aligned}$$

Note also that the π appearing in the sum are either of the form

$1 \quad \dots \quad m \quad m{+}1 \quad \dots \quad m{+}p \quad m{+}p{+}1 \quad \dots \quad n$

or of the form

$$1 \quad \dots \quad m{-}p \quad m{-}p{+}1 \quad \dots \quad m \quad m{+}1 \quad \dots \quad n$$

Example 5.4. Let us make clear the structure of the assertion with the help of an example: For $A_1 := a_1a_2$ and $A_2 := a_3$ we have $\sigma = \{(a_1, a_2), (a_3)\} \mathrel{\hat{=}} \sqcup|$. Consider now $\tau = 1_2 = \{(A_1, A_2)\}$, implying that $\hat{\tau} = 1_3 = \{(a_1, a_2, a_3)\}$. Then the application of our theorem yields

$$\begin{aligned} k_2(a_1a_2, a_3) &= \sum_{\substack{\pi\in NC(3)\\ \pi\vee\sigma=1_3}} k_\pi[a_1, a_2, a_3] \\ &= k_{\sqcup\!\sqcup}[a_1, a_2, a_3] + k_{|\sqcup}[a_1, a_2, a_3] + k_{\sqcup|}[a_1, a_2, a_3] \\ &= k_3(a_1, a_2, a_3) + k_1(a_1)k_2(a_2, a_3) + k_2(a_1, a_3)k_1(a_2)\ , \end{aligned}$$

which is easily seen to be indeed equal to $k_2(a_1a_2, a_3) = \varphi(a_1a_2a_3) - \varphi(a_1a_2)\varphi(a_3)$.

We will give two proofs of that result in the following. The first one is the original one (due to my student Krawczyk and myself) and quite explicit. The second proof (due to myself) is more conceptual by connecting the theorem with some more general structures. (Similar ideas as in the second proof appeared also in a proof by Cabanal-Duvillard.)

First proof of Theorem 5.2. We show the assertion by induction over the number m of arguments of the cumulant k_τ.
To begin with, let us study the case when $m = 1$. Then we have $\sigma = \{(a_1, \dots, a_n)\} = 1_n = \hat{\tau}$ and by the defining relation (1) for the free cumulants our assertion reduces to

$$\begin{aligned} k_1(a_1 \cdots a_n) &= \sum_{\substack{\pi\in NC(n)\\ \pi\vee 1_n=1_n}} k_\pi[a_1, \dots, a_n] \\ &= \sum_{\pi\in NC(n)} k_\pi[a_1, \dots, a_n] \\ &= \varphi(a_1 \cdots a_n), \end{aligned}$$

which is true since $k_1 = \varphi$.
Let us now make the induction hypothesis that for an integer $m \geq 1$ the theorem is true for all $m' \leq m$.
We want to show that it also holds for $m + 1$. This means that for

$\tau \in NC(m+1)$, a sequence $1 \le i_1 < i_2 < \cdots < i_{m+1} =: n$, and random variables $a_1, \dots, a_n$ we have to prove the validity of the following equation:

$$
\begin{aligned}
k_\tau[A_1, \dots, A_{m+1}] &= k_\tau[a_1 \cdots a_{i_1}, \dots, a_{i_m+1} \cdots a_{i_{m+1}}] \\
&= \sum_{\substack{\pi \in NC(n) \\ \pi \vee \sigma = \hat{\tau}}} k_\pi[a_1, \dots, a_n],
\end{aligned} \tag{61}
$$

where $\sigma = \{(a_1, \dots, a_{i_1}), \dots, (a_{i_m+1}, \dots, a_{i_{m+1}})\}$.
The proof is divided into two steps. The first one discusses the case where $\tau \in NC(m+1)$, $\tau \ne 1_{m+1}$ and the second one treats the case where $\tau = 1_{m+1}$.
Step $1°$: The validity of relation (3) for all $\tau \in NC(m+1)$ except the partition 1_{m+1} is shown as follows: Each such τ has at least two blocks, so it can be written as $\tau = \tau_1 \cup \tau_2$ with τ_1 being a non-crossing partition of an s-tuple $(B_1, \dots, B_s)$ and τ_2 being a non-crossing partition of a t-tuple $(C_1, \dots, C_t)$ where $(B_1, \dots, B_s) \cup (C_1, \dots, C_t) = (A_1, \dots, A_{m+1})$ and $s + t = m + 1$. With these definitions, we have

$$k_\tau[A_1, \dots, A_{m+1}] = k_{\tau_1}[B_1, \dots, B_s]\, k_{\tau_2}[C_1, \dots, C_t]\,.$$

We will apply now the induction hypothesis on $k_{\tau_1}[B_1, \dots, B_s]$ and on $k_{\tau_2}[C_1, \dots, C_t]$. According to the definition of A_j, both $B_k (k = 1, \dots, s)$ and C_l $(l = 1, \dots, t)$ are products with factors from $(a_1, \dots, a_n)$. Put $(b_1, \dots, b_p)$ the tuple containing all factors of $(B_1, \dots, B_s)$ and $(c_1, \dots, c_q)$ the tuple consisting of all factors of $(C_1, \dots, C_t)$; this means $(b_1, \dots, b_p) \cup (c_1, \dots, c_q) = (a_1, \dots, a_n)$ (and $p + q = n$). We put $\sigma_1 := \sigma|_{(b_1, \dots, b_p)}$ and $\sigma_2 := \sigma|_{(c_1, \dots, c_q)}$, i.e., we have $\sigma = \sigma_1 \cup \sigma_2$. Note that $\hat{\tau}$ factorizes in the same way as $\hat{\tau} = \hat{\tau}_1 \cup \hat{\tau}_2$. Then we get with the help of our induction hypothesis:

$$
\begin{aligned}
k_\tau[A_1, \dots, A_{m+1}] &= k_{\tau_1}[B_1, \dots, B_s] \cdot k_{\tau_2}[C_1, \dots, C_t] \\
&= \sum_{\substack{\pi_1 \in NC(p) \\ \pi_1 \vee \sigma_1 = \hat{\tau}_1}} k_{\pi_1}[b_1, \dots, b_p] \cdot \sum_{\substack{\pi_2 \in NC(q) \\ \pi_2 \vee \sigma_2 = \hat{\tau}_2}} k_{\pi_2}[c_1, \dots, c_q] \\
&= \sum_{\substack{\pi_1 \in NC(p) \\ \pi_1 \vee \sigma_1 = \hat{\tau}_1}} \sum_{\substack{\pi_2 \in NC(q) \\ \pi_2 \vee \sigma_2 = \hat{\tau}_2}} k_{\pi_1 \cup \pi_2}[a_1, \dots, a_n] \\
&= \sum_{\substack{\pi \in NC(n) \\ \pi \vee \sigma = \hat{\tau}}} k_\pi[a_1, \dots, a_n]\,.
\end{aligned}
$$

Step 2°: It remains to prove that the equation (3) is also valid for $\tau = 1_{m+1}$. By the definition of the free cumulants we obtain

$$\begin{aligned} k_{1_{m+1}}[A_1, \dots, A_{m+1}] &= k_{m+1}(A_1, \dots, A_{m+1}) \\ &= \varphi(A_1 \cdots A_{m+1}) - \sum_{\substack{\tau \in NC(m+1) \\ \tau \neq 1_{m+1}}} k_\tau[A_1, \dots, A_{m+1}] \,. \end{aligned} \tag{62}$$

First we transform the sum in (4) with the result of step 1°:

$$\begin{aligned} \sum_{\substack{\tau \in NC(m+1) \\ \tau \neq 1_{m+1}}} k_\tau[A_1, \dots, A_{m+1}] &= \sum_{\substack{\tau \in NC(m+1) \\ \tau \neq 1_{m+1}}} \sum_{\substack{\pi \in NC(n) \\ \pi \vee \sigma = \hat{\tau}}} k_\pi[a_1, \dots, a_n] \\ &= \sum_{\substack{\pi \in NC(n) \\ \pi \vee \sigma \neq 1_n}} k_\pi[a_1, \dots, a_n] \,, \end{aligned}$$

where we used the fact that $\tau = 1_{m+1}$ is equivalent to $\hat{\tau} = 1_n$.
The moment in (4) can be written as

$$\varphi(A_1 \cdots A_{m+1}) = \varphi(a_1 \cdots a_n) = \sum_{\pi \in NC(n)} k_\pi[a_1, \dots, a_n] \,.$$

Altogether, we get:

$$\begin{aligned} k_{m+1}[A_1, \dots, A_{m+1}] &= \sum_{\pi \in NC(n)} k_\pi[a_1, \dots, a_n] - \sum_{\substack{\pi \in NC(n) \\ \pi \vee \sigma \neq 1_n}} k_\pi[a_1, \dots, a_n] \\ &= \sum_{\substack{\pi \in NC(n) \\ \pi \vee \sigma = 1_n}} k_\pi[a_1, \dots, a_n] \,. \end{aligned}$$

□

The second proof reduces the statement essentially to some general structure theorem on lattices. This is the well-known Möbius algebra which we present here in a form which looks analogous to the description of convolution via Fourier transform.

NOTATIONS 5.5. Let P be a finite lattice.
1) For two functions $f, g : P \to \mathbb{C}$ on P we denote by $f * g : P \to \mathbb{C}$ the function defined by

$$f * g(\pi) := \sum_{\substack{\sigma_1, \sigma_2 \in P \\ \sigma_1 \vee \sigma_2 = \pi}} f(\sigma_1) g(\sigma_2) \qquad (\pi \in P). \tag{63}$$

2) For a function $f : P \to \mathbb{C}$ we denote by $F(f) : P \to \mathbb{C}$ the function defined by

$$F(f)(\pi) := \sum_{\substack{\sigma\in P\\ \sigma\leq\pi}} f(\sigma) \qquad (\pi \in P). \tag{64}$$

REMARKS 5.6. 1) Note that the operation $(f, g) \mapsto f * g$ is commutative and associative.
2) According to the theory of Möbius inversion F is a bijective mapping whose inverse is given by

$$F^{-1}(f)(\pi) = \sum_{\substack{\sigma\in P\\ \sigma\leq\pi}} f(\sigma)\mu(\sigma, \pi). \tag{65}$$

3) Denote by 1_σ and $1_{\geq\sigma}$ the functions given by

$$1_\sigma(\pi) = \begin{cases} 1, & \text{if } \pi = \sigma \\ 0, & \text{otherwise} \end{cases} \tag{66}$$

and

$$1_{\geq\sigma}(\pi) = \begin{cases} 1, & \text{if } \pi \geq \sigma \\ 0, & \text{otherwise.} \end{cases} \tag{67}$$

Then we have

$$F(1_\sigma)(\pi) = \sum_{\tau\leq\pi} 1_\sigma(\tau) = 1_{\geq\sigma}(\pi),$$

and hence $F(1_\sigma) = 1_{\geq\sigma}$.

PROPOSITION 5.7. *Let P be a finite lattice. Then we have for arbitrary functions $f, g : P \to \mathbb{C}$ that*

$$F(f * g) = F(f) \cdot F(g), \tag{68}$$

where on the right hand side we have the pointwise product of functions.

PROOF. We have

$$\begin{aligned} F(f * g)(\pi) &= \sum_{\sigma\leq\pi} (f * g)(\sigma) \\ &= \sum_{\sigma\leq\pi} \sum_{\sigma_1\vee\sigma_2=\sigma} f(\sigma_1)g(\sigma_2) \\ &= \big(\sum_{\sigma_1\leq\pi} f(\sigma_1)\big)\big(\sum_{\sigma_2\leq\pi} g(\sigma_2)\big) \\ &= F(f)(\pi) \cdot F(g)(\pi), \end{aligned}$$

where we used the fact that $\sigma_1 \vee \sigma_2 \leq \pi$ is equivalent to $\sigma_1 \leq \pi$ and $\sigma_2 \leq \pi$. □

Before starting with the second proof of our Theorem 5.2 we have to show that the embedding of $NC(m)$ into $NC(n)$ by $\tau \mapsto \hat{\tau}$ preserves the Möbius function.

LEMMA 5.8. *Consider $\tau, \pi \in NC(m)$ with $\pi \leq \tau$ and let $\hat{\pi} \leq \hat{\tau}$ be the corresponding partitions in $NC(n)$ according to 5.1. Then we have*

$$\mu(\hat{\pi}, \hat{\tau}) = \mu(\pi, \tau). \tag{69}$$

PROOF. By the reccurence relation (4.15) for the Möbius function we have $\mu(\hat{\pi}, \hat{\tau}) = -\sum_{\hat{\pi} \leq \kappa < \hat{\tau}} \mu(\hat{\pi}, \kappa)$. But the condition $\hat{\pi} \leq \kappa < \hat{\tau}$ enforces that $\kappa \in NC(n)$ is of the form $\kappa = \hat{\omega}$ for a uniquely determined $\omega \in NC(m)$ with $\pi \leq \omega < \tau$. (Note that κ being of the form $\hat{\omega}$ is equivalent to $\kappa \geq \sigma$, where σ is the partition from Theorem 5.2.) Hence we get by induction on the length of $[\hat{\pi}, \hat{\tau}]$.

$$\mu(\hat{\pi}, \hat{\tau}) = -\sum_{\hat{\pi} \leq \hat{\omega} < \hat{\tau}} \mu(\hat{\pi}, \hat{\omega}) = -\sum_{\pi \leq \omega \leq \tau} \mu(\pi, \omega) = \mu(\pi, \tau).$$

□

SECOND PROOF OF THEOREM 5.2. We have

$$\begin{aligned}
k_\tau[A_1, \dots, A_m] &= \sum_{\substack{\pi \in NC(m) \\ \pi \leq \tau}} \varphi_\pi[A_1, \dots, A_m] \mu(\pi, \tau) \\
&= \sum_{\substack{\pi \in NC(m) \\ \pi \leq \tau}} \varphi_{\hat{\pi}}[a_1, \dots, a_n] \mu(\hat{\pi}, \hat{\tau}) \\
&= \sum_{\substack{\pi' \in NC(n) \\ \sigma \leq \pi' \leq \hat{\tau}}} \varphi_{\pi'}[a_1, \dots, a_n] \mu(\pi', \hat{\tau}) \\
&= \sum_{\substack{\pi' \in NC(n) \\ \pi' \leq \hat{\tau}}} \varphi_{\pi'}[a_1, \dots, a_n] \mu(\pi', \hat{\tau}) 1_{\geq \sigma}(\pi') \\
&= F^{-1}(\varphi[a_1, \dots, a_n] \cdot 1_{\geq \sigma})(\hat{\tau}) \\
&= (F^{-1}(\varphi[a_1, \dots, a_n]) * F^{-1}(1_{\geq \sigma}))(\hat{\tau}) \\
&= (k[a_1, \dots, a_n] * 1_\sigma)(\hat{\tau}) \\
&= \sum_{\sigma_1 \vee \sigma_2 = \hat{\tau}} k_{\sigma_1}[a_1, \dots, a_n] 1_\sigma(\sigma_2) \\
&= \sum_{\sigma_1 \vee \sigma = \hat{\tau}} k_{\sigma_1}[a_1, \dots, a_n].
\end{aligned}$$

□

EXAMPLE 5.9. Let s be a semi-circular variable of variance 1 and put $P := s^2$. Then we get the cumulants of P as

$$k_n(P, \dots, P) = k_n(s^2, \dots, s^2) = \sum_{\substack{\pi \in NC(2n) \\ \pi \vee \sigma = 1_{2n}}} k_\pi[s, s, \dots, s, s],$$

where $\sigma = \{(1,2), (3,4), \dots, (2n-1, 2n)\}$. Since s is semi-circular, only pair partitions contribute to the sum. It is easy to see that there is only one pair partition with the property that $\pi \vee \sigma = 1_{2n}$, namely the partition $\pi = \{(1, 2n), (2,3), (4,5), \dots, (2n-2, 2n-1)\}$. Thus we get

$$k_n(P, \dots, P) = 1 \qquad \text{for all } n \geq 1 \tag{70}$$

and by analogy with the classical case we call P a **free Poisson** of parameter $\lambda = 1$. We will later generalize this to more general kind of free Poisson distribution.

CHAPTER 6

Freeness and vanishing of mixed cumulants

The meaning of the concept 'cumulants' for freeness is shown by the following theorem.

THEOREM 6.1. *Let $(\mathcal{A}, \varphi)$ be a probability space and consider unital subalgebras $\mathcal{A}_1, \dots, \mathcal{A}_m \subset \mathcal{A}$. Then the following two statements are equivalent:*

i) *$\mathcal{A}_1, \dots, \mathcal{A}_m$ are free.*

ii) *We have for all $n \geq 2$ and for all $a_i \in \mathcal{A}_{j(i)}$ with $1 \leq j(1), \dots, j(n) \leq m$ that $k_n(a_1, \dots, a_n) = 0$ whenever there exist $1 \leq l, k \leq n$ with $j(l) \neq j(k)$.*

REMARKS 6.2. 1) This characterization of freeness in terms of cumulants is the translation of the definition of freeness in terms of moments – by using the relation between moments and cumulants from Definition 4.5. One should note that in contrast to the characterization in terms of moments we do not require that $j(1) \neq j(2) \neq \cdots \neq j(n)$ nor that $\varphi(a_i) = 0$. Hence the characterization of freeness in terms of cumulants is much easier to use in concrete calculations.
2) Since the unit 1 is free from everything, the above theorem contains as a special case the statement: $k_n(a_1, \dots, a_n) = 0$ if $n \geq 2$ and $a_i = 1$ for at least one i. This special case will also present an important step in the proof of Theorem 6.1 and it will be proved separately as a lemma.
3) Note also: for $n = 1$ we have $k_1(1) = \varphi(1) = 1$.

Before we start the proof of our theorem we will consider, as indicated in the above remark, a special case.

LEMMA 6.3. *Let $n \geq 2$ und $a_1, \dots, a_n \in \mathcal{A}$. Then we have $k_n(a_1, \dots, a_n) = 0$ if there exists a $1 \leq i \leq n$ with $a_i = 1$.*

SECOND PROOF OF THEOREM 5.2. To simplify notation we consider the case $a_n = 1$, i.e. we want to show that $k_n(a_1, \dots, a_{n-1}, 1) = 0$. We will prove this by induction on n.
$n = 2$: the assertion is true, since

$$k_2(a, 1) = \varphi(a1) - \varphi(a)\varphi(1) = 0.$$

$n-1 \to n$: Assume we have proved the assertion for all $k < n$. Then we have

$$\begin{aligned}\varphi(a_1 \cdots a_{n-1} 1) &= \sum_{\pi \in NC(n)} k_\pi[a_1, \dots, a_{n-1}, 1] \\ &= k_n(a_1, \dots, a_{n-1}, 1) + \sum_{\substack{\pi \in NC(n) \\ \pi \neq \mathbf{1}_n}} k_\pi[a_1, \dots, a_{n-1}, 1].\end{aligned}$$

According to our induction hypothesis only such $\pi \neq \mathbf{1}_n$ contribute to the above sum which have the property that (n) is a one-element block of π, i.e. which have the form $\pi = \sigma \cup (n)$ with $\sigma \in NC(n-1)$. Then we have

$$k_\pi[a_1, \dots, a_{n-1}, 1] = k_\sigma[a_1, \dots, a_{n-1}] k_1(1) = k_\sigma[a_1, \dots, a_{n-1}],$$

hence

$$\begin{aligned}\varphi(a_1 \cdots a_{n-1} 1) &= k_n(a_1, \dots, a_{n-1}, 1) + \sum_{\sigma \in NC(n-1)} k_\sigma[a_1, \dots, a_{n-1}] \\ &= k_n(a_1, \dots, a_{n-1}, 1) + \varphi(a_1 \cdots a_{n-1}).\end{aligned}$$

Since $\varphi(a_1 \cdots a_{n-1} 1) = \varphi(a_1 \cdots a_{n-1})$, we obtain $k_n(a_1, \dots, a_{n-1}, 1) = 0$. $\square$

Now we can prove the general version of our theorem.

Proof of Theorem 6.1. (i) $\Longrightarrow$ (ii): If all a_i are centered, i.e. $\varphi(a_i) = 0$, and alternating, i.e. $j(1) \neq j(2) \neq \cdots \neq j(n)$, then the assertion follows directly by the relation

$$k_n(a_1, \dots, a_n) = \sum_{\pi \in NC(n)} \varphi_\pi[a_1, \dots, a_n] \mu(\pi, 1_n),$$

because at least one factor of φ_π is of the form $\varphi(a_l a_{l+1} \cdots a_{l+p})$, which vanishes by the definition of freeness.

The essential part of the proof consists in showing that on the level of cumulants the assumption 'centered' is not needed and 'alternating' can be weakened to 'mixed'.

Let us start with getting rid of the assumption 'centered'. Let $n \geq 2$. Then the above Lemma 6.3 implies that we have for arbitrary $a_1, \dots, a_n \in \mathcal{A}$ the relation

$$k_n(a_1, \dots, a_n) = k_n\big(a_1 - \varphi(a_1)1, \dots, a_n - \varphi(a_n)1\big), \tag{71}$$

i.e. we can center the arguments of our cumulants k_n $(n \geq 2)$ without changing the value of the cumulants.

Thus we have proved the following statement: Consider $n \geq 2$ and $a_i \in \mathcal{A}_{j(i)}$ $(i = 1, \dots, n)$ with $j(1) \neq j(2) \neq \cdots \neq j(n)$. Then we have

$k_n(a_1, \dots, a_n) = 0$.
To prove our theorem in full generality we will now use Theorem 5.2 on the behaviour of free cumulants with products as entries – indeed, we will only need the special form of that theorem as spelled out in Remark 5.3. Consider $n \geq 2$ and $a_i \in \mathcal{A}_{j(i)}$ $(i = 1, \dots, n)$. Assume that there exist k, l with $j(k) \neq j(l)$. We have to show that $k_n(a_1, \dots, a_n) = 0$. This follows so: If $j(1) \neq j(2) \neq \dots \neq j(n)$, then the assertion is already proved. If the elements are not alternating then we multiply neigbouring elements from the same algebra together, i.e. we write $a_1 \dots a_n = A_1 \dots A_m$ such that neighbouring A's come from different subalgebras. Note that $m \geq 2$ because of our assumption $j(k) \neq j(l)$. Then, by Theorem 5.2, we have

$$k_n(a_1, \dots, a_n) = k_m(A_1, \dots, A_m) - \sum_{\substack{\pi \in NC(n), \pi \neq 1_n \\ \pi \vee \sigma = 1_n}} k_\pi[a_1, \dots, a_n],$$

where $\sigma = \{(A_1), \dots, (A_m)\} \in NC(a_1, \dots, a_n)$ is that partition whose blocks encode the information about which elements a_i to multiply in order to get the A_j. Since the A's are alternating we have $k_m(A_1, \dots, A_m) = 0$. Furthermore, by induction, the term $k_\pi[a_1, \dots, a_n]$ can only be different from zero, if each block of π couples only elements from the same subalgebra. So all blocks of σ which are coupled by π must correspond to the same subalgebra. However, we also have that $\pi \vee \sigma = 1_n$, which means that π has to couple all blocks of σ. Hence all appearing elements should be from the same subalgebra, which is in contradiction with $m \geq 2$. Thus there is no non-vanishing contribution in the above sum and we obtain $k_n(a_1, \dots, a_n) = 0$.
(ii) $\implies$ (i): Consider $a_i \in \mathcal{A}_{j(i)}$ $(i = 1, \dots, n)$ with $j(1) \neq j(2) \neq \dots \neq j(n)$ and $\varphi(a_i) = 0$ for all $i = 1, \dots, n$. Then we have to show that $\varphi(a_1 \cdots a_n) = 0$. But this is clear because we have $\varphi(a_1 \cdots a_n) = \sum_{\pi \in NC(n)} k_\pi[a_1, \dots, a_n]$ and each product $k_\pi[a_1, \dots, a_n] = \prod_{V \in \pi} k(\pi|_V)$ contains at least one factor of the form $k_{p+1}(a_l, a_{l+1}, \dots, a_{l+p})$ which vanishes in any case (for $p = 0$ because our variables are centered and for $p \geq 1$ because of our assumption on the vanishing of mixed cumulants). □

Let us also state the freeness criterion in terms of cumulants for the case of random variables.

Theorem 6.4. *Let $(\mathcal{A}, \varphi)$ be a probability space and consider random variables $a_1, \dots, a_m \in \mathcal{A}$. Then the following two statements are equivalent:*

i) *$a_1, \dots, a_m$ are free.*

ii) *We have for all* $n \geq 2$ *and for all* $1 \leq i(1), \dots, i(n) \leq m$ *that* $k_n(a_{i(1)}, \dots, a_{i(n)}) = 0$ *whenever there exist* $1 \leq l, k \leq n$ *with* $i(l) \neq i(k)$.

Proof. The implication i) $\Longrightarrow$ ii) follows of course directly from the foregoing theorem. Only the other way round is not immediately clear, since we have to show that our present assumption ii) implies also the apparently stronger assumption ii) for the case of algebras. Thus let $\mathcal{A}_i$ be the unital algebra generated by the element a_i and consider now elements $b_i \in \mathcal{A}_{j(i)}$ with $1 \leq j(1), \dots, j(n) \leq m$ such that $j(l) \neq j(k)$ for some l, k. Then we have to show that $k_n(b_1, \dots, b_n)$ vanishes. As each b_i is a polynomial in $a_{j(i)}$ and since cumulants with a 1 as entry vanish in any case for $n \geq 2$, it suffices to consider the case where each b_i is some power of $a_{j(i)}$. If we write $b_1 \cdots b_n$ as $a_{i(1)} \cdots a_{i(m)}$ then we have

$$k_n(b_1, \dots, b_n) = \sum_{\substack{\pi \in NC(m) \\ \pi \vee \sigma = 1_m}} k_\pi[a_{i(1)}, \dots, a_{i(m)}],$$

where the blocks of σ denote the neighbouring elements which have to be multiplied to give the b_i. In order that $k_\pi[a_{i(1)}, \dots, a_{i(m)}]$ is different from zero π is only allowed, by our assumption (ii), to couple between the same a_i. So all blocks of σ which are coupled by π must correspond to the same a_i. However, we also have $\pi \vee \sigma = 1_m$, which means that all blocks of σ have to be coupled by π. Thus all a_i should be the same, in contradiction with the fact that we consider a mixed cumulant. Hence there is no non-vanishing contribution in the above sum and we finally get that $k_n(b_1, \dots, b_n) = 0$. $\square$

Definition 6.5. An element c of the form $c = \frac{1}{\sqrt{2}}(s_1 + is_2)$ – where s_1 and s_2 are two free semi-circular elements of variance 1 – is called a **circular elment**.

Examples 6.6. 1) The vanishing of mixed cumulants in free variables gives directly the cumulants of a circular element: Since only second order cumulants of semi-circular elements are different from zero, the only non-vanishing cumulants of a circular element are also of second order and for these we have

$$k_2(c, c) = k_2(c^*, c^*) = \frac{1}{2} - \frac{1}{2} = 0$$
$$k_2(c, c^*) = k_2(c^*, c) = \frac{1}{2} + \frac{1}{2} = 1.$$

2) Let s be a semi-circular of variance 1 and a a random variable which

is free from s. Put $P := sas$. Then, by Theorem 5.2, we get

$$k_n(P, \dots, P) = k_n(sas, \dots, sas) = \sum_{\substack{\pi \in NC(3n) \\ \pi \vee \sigma = 1_{3n}}} k_\pi[s, a, s, s, a, s, \dots, s, a, s],$$

where $\sigma \in NC(3n)$ is the partition $\sigma = \{(1, 2, 3), (4, 5, 6), \dots, (3n - 2, 3n - 1, 3n)\}$. Now Theorem 6.1 tells us that k_π only gives a non-vanishing contribution if all blocks of π do not couple s with a. Furthermore, since s is semi-circular, those blocks which couple within the s have to consist of exactly two elements. But, together with the requirement $\pi \vee \sigma = 1_{3n}$, this means that π must be of the form $\pi_s \cup \pi_a$ where $\pi_s = \{(1, 3n), (3, 4), (6, 7), (9, 10), \dots, (3n - 3, 3n - 2)\} \in NC(1, 3, 4, 6, 7, 9, \dots, 3n - 3, 3n - 2, 3n)$ is that special partition of the positions of the s which glues the blocks of σ together and where $\pi_a \in NC(2, 5, 8, \dots, 3n - 1)$ is just an arbitrary partition for the positions of the a.

$$s \quad a \quad s \quad , \quad s \quad a \quad s \quad , \quad s \quad a \quad s \quad , \quad \dots \quad , \quad s \quad a \quad s$$

Note that $k_\pi[s, a, s, s, a, s, \dots, s, a, s]$ factorizes then into $k_{\pi_s}[s, s, \dots, s] \cdot k_{\pi_a}[a, a, \dots, a] = k_{\pi_a}[a, a, \dots, a]$ and we get

$$k_n(P, \dots, P) = \sum_{\pi_a \in NC(n)} k_{\pi_a}[a, a, \dots, a] = \varphi(a^n).$$

Thus we have the result that the cumulants of P are given by the moments of a:

$$k_n(P, \dots, P) = \varphi(a^n) \qquad \text{for all } n \geq 1. \tag{72}$$

In analogy with the classical case we will call a random variable of such a form a **compound Poisson** with parameter $\lambda = 1$. Note that we recover the usual Poisson element from Example 5.9 by putting $a = 1$.
3) As a generalization of the last example, consider now the following situation: Let s be a semi-circular of variance 1 and $a_1, \dots, a_m$ random variables such that s and $\{a_1, \dots, a_m\}$ are free. Put $P_i := sa_is$. As

above we can calculate the joint distribution of these elements as

$$\begin{aligned} k_n(P_{i(1)},\dots,P_{i(n)}) &= k_n(sa_{i(1)}s,\dots,sa_{i(n)}s) \\ &= \sum_{\substack{\pi\in NC(3n)\\ \pi\vee\sigma=1_{3n}}} k_\pi[s,a_{i(1)},s,s,a_{i(2)},s,\dots,s,a_{i(n)},s] \\ &= \sum_{\pi_a\in NC(n)} k_{\pi_a}[a_{i(1)},a_{i(2)},\dots,a_{i(n)}] \\ &= \varphi(a_{i(1)}a_{i(2)}\cdots a_{i(n)}), \end{aligned}$$

where $\sigma \in NC(3n)$ is as before the partition $\sigma = \{(1,2,3),(4,5,6),\dots,(3n-2,3n-1,3n)\}$. Thus we have again the result that the cumulants of $P_1,\dots,P_m$ are given by the moments of $a_1,\dots,a_m$. This contains of course the statement that each of the P_i is a compound Poisson, but we also get that orthogonaliy between the a_i is translated into freeness between the P_i. Namely, assume that all a_i are orthogonal in the sense $a_ia_j = 0$ for all $i \neq j$. Consider now a mixed cumulant in the P_i, i.e. $k_n(P_{i(1)},\dots,P_{i(n)})$, with $i(l) \neq i(k)$ for some l,k. Of course, then there are also two neighbouring indices which are different, i.e. we can assume that $k = l+1$. But then we have

$$k_n(P_{i(1)},\dots,P_{i(l)},P_{i(l+1)},\dots,P_{i(n)}) = \varphi(a_{i(1)}\dots a_{i(l)}a_{i(l+1)}\dots a_{i(n)}) = 0.$$

Thus mixed cumulants in the P_i vanish and, by our criterion 6.4, $P_1,\dots,P_m$ have to be free.

4) Let us also calculate the cumulants for a Haar unitary element u. Whereas for a semi-circular element the cumulants are simpler than the moments, for a Haar unitary it is the other way around. The $*$-moments of u are described quite easily, whereas the situation for cumulants in u and u^* is more complicated. But nevertheless, there is a structure there, which will be useful and important later. In particular, the formula for the cumulants of a Haar unitary was one of the motivations for the introduction of so-called R-diagonal elements.

So let u be a Haar unitary. We want to calculate $k_n(u_1,\dots,u_n)$, where $u_1,\dots,u_n \in \{u,u^*\}$. First we note that such a cumulant can only be different from zero if the number of u among $u_1,\dots,u_n$ is the same as the number of u^* among $u_1,\dots,u_n$. This follows directly by the formula $k_n(u_1,\dots,u_n) = \sum_{\pi\in NC(n)} \varphi_\pi[u_1,\dots,u_n]\mu(\pi,1_n)$, since, for arbitrary $\pi \in NC(n)$, at least one of the blocks V of π couples different numbers of u and u^*; but then $\varphi(\pi|_V)$, and thus also $\varphi_\pi[u_1,\dots,u_n]$, vanishes. This means in particular, that only cumulants of even length are different from zero.

Consider now a cumulant where the number of u and the number of

u^* are the same. We claim that only such cumulants are different from zero where the entries are alternating in u and u^*. We will prove this by induction on the length of the cumulant. Consider a non-alternating cumulant. Since with u also u^* is a Haar unitary it suffices to consider the cases $k(\dots, u^*, u, u, \dots)$ and $k(\dots, u, u, u^*, \dots)$. Let us treat the first one. Say that the positions of $\dots, u^*, u, u, \dots$ are $\dots, m, m+1, m+2, \dots$. Since 1 is free from everything (or by Lemma 6.3) and by Remark 5.3 we know that

$$\begin{aligned} 0 &= k(\dots, 1, u, \dots) \\ &= k(\dots, u^* \cdot u, u, \dots) \\ &= k(\dots, u^*, u, u, \dots) + \sum_{\substack{|\pi|=2 \\ m \not\sim_\pi m+1}} k_\pi[\dots, u^*, u, u, \dots]. \end{aligned}$$

Let us now consider the terms in the sum. Since each k_π is a product of two lower order cumulants we know, by induction, that each of the two blocks of π must connect the arguments alternatingly in u and u^*. This implies that π cannot connect $m+1$ with $m+2$, and hence it must connect m with $m+2$. But this forces $m+1$ to give rise to a singleton of π, hence one factor of k_π is just $k_1(u) = 0$. This implies that all terms k_π in the above sum vanish and we obtain $k(\dots, u^* \cdot u, u, \dots) = 0$. The other case is analogous, hence we get the statement that only alternating cumulants in u and u^* are different from zero.
Finally, it remains to determine the value of the alternating cumulants. Let us denote by α_n the value of such a cumulant of length $2n$,i.e.

$$\alpha_n := k_{2n}(u, u^*, \dots, u, u^*) = k_{2n}(u^*, u, \dots, u^*, u).$$

The last equality comes from the fact that u^* is also a Haar unitary. We use now again Theorem 5.2 (in the special form presented in part 2) of Remark 5.3) and the fact that 1 is free from everything:

$$\begin{aligned} 0 &= k_{2n-1}(1, u, u^*, \dots, u, u^*) \\ &= k_{2n-1}(u \cdot u^*, u, u^*, \dots, u, u^*) \\ &= k_{2n}(u, u^*, u, u^*, \dots, u, u^*) + \sum_{\substack{\pi \in NC(2n) \\ |\pi|=2, 1 \not\sim_\pi 2}} k_\pi[u, u^*, u, u^*, \dots, u, u^*] \\ &= k_{2n}(u, u^*, u, u^*, \dots, u, u^*) + \sum_{p=1}^{n-1} k_{2n-2p}(u, u^*, \dots, u, u^*) \cdot k_{2p}(u^*, u, \dots, u^*, u) \\ &= \alpha_n + \sum_{p=1}^{n-1} \alpha_{n-p} \alpha_p. \end{aligned}$$

Thus we have the recursion

$$\alpha_n = -\sum_{p=1}^{n-1} \alpha_{n-p}\alpha_p, \tag{73}$$

which is up to the minus-sign and a shift in the indices by 1 the recursion relation for the Catalan numbers. Since also $\alpha_1 = k_2(u, u^*) = 1$ this gives finally

$$k_{2n}(u, u^*, \dots, u, u^*) = (-1)^{n-1}C_{n-1}. \tag{74}$$

3) Let b be a Bernoulli variable, i.e. a selfadjoint random variable whose distribution is the measure $\frac{1}{2}(\delta_{-1} + \delta_1)$. This means nothing but that the moments of b are as follows:

$$\varphi(b^n) = \begin{cases} 1, & \text{if } n \text{ even} \\ 0, & \text{if } n \text{ odd} \end{cases} \tag{75}$$

By the same kind of reasoning as for the Haar unitary one finds the cumulants of b as

$$k_n(b, \dots, b) = \begin{cases} (-1)^{k-1}C_{k-1}, & \text{if } n = 2k \text{ even} \\ 0, & \text{if } n \text{ odd.} \end{cases} \tag{76}$$

CHAPTER 7

Additive free convolution

Our main concern in this section will be the understanding and effective description of the sum of free random variables. On the level of distributions we are addressing the problem of the (additive) free convolution.

DEFINITION 7.1. Let μ and ν be probability measures on $\mathbb{R}$ with compact support. Let x and y be self-adjoint random variables in some C^*-probability space which have the given measures as distribution, i.e. $\mu_x = \mu$ and $\mu_y = \nu$, and which are free. Then the distribution μ_{x+y} of the sum $x + y$ is called the **free convolution** of μ and ν and denoted by $\mu \boxplus \nu$.

REMARKS 7.2. 1) Note that, for given μ and ν as above, one can always find x and y as required. Furthermore, by Lemma 2.4, the distribution of the sum does only depend on the distributions μ_x and μ_y and not on the concrete realisations of x and y. Thus $\mu \boxplus \nu$ is uniquely determined.
2) We defined the free convolution only for measures with compact support. By adequate truncations one can extend the definition (and the main results) also to arbitrary probability measures on $\mathbb{R}$.
3) To be more precise (and in order to distinguish it from the analogous convolution for the product of free variables) one calls $\mu \boxplus \nu$ also the additive free convolution.

In order to recover the results of Voiculescu on the free convolution we only have to specify our combinatorial description to the one-dimensional case

NOTATION 7.3. For a random variable $a \in \mathcal{A}$ we put

$$k_n^a := k_n(a, \dots, a)$$

and call $(k_n^a)_{n\geq 1}$ the **(free) cumulants of** a.

Our main theorem on the vanishing of mixed cumulants in free variables specifies in this one-dimensional case to the linearity of the cumulants.

PROPOSITION 7.4. *Let a and b be free. Then we have*

$$k_n^{a+b} = k_n^a + k_n^b \qquad \text{for all } n \geq 1. \tag{77}$$

PROOF. We have

$$\begin{aligned} k_n^{a+b} &= k_n(a+b, \dots, a+b) \\ &= k_n(a, \dots, a) + k_n(b, \dots, b) \\ &= k_n^a + k_n^b, \end{aligned}$$

because cumulants which have both a and b as arguments vanish by Theorem 6.4. □

Thus, free convolution is easy to describe on the level of cumulants; the cumulants are additive under free convolution. It remains to make the connection between moments and cumulants as explicit as possible. On a combinatorial level, our definition specializes in the one-dimensional case to the following relation.

PROPOSITION 7.5. *Let $(m_n)_{n\geq 1}$ and $(k_n)_{n\geq 1}$ be the moments and free cumulants, respectively, of some random variable. The connection between these two sequences of numbers is given by*

$$m_n = \sum_{\pi \in NC(n)} k_\pi, \tag{78}$$

where

$$k_\pi := k_{\#V_1} \cdots k_{\#V_r} \qquad \text{for} \qquad \pi = \{V_1, \dots, V_r\}.$$

EXAMPLE 7.6. For $n = 3$ we have

$$\begin{aligned} m_3 &= k_{\sqcup\!\sqcup} + k_{|\,\sqcup} + k_{\sqcup\,|} + k_{\sqcup\!\!\!\!\;\cdot\;} + k_{|\,|\,|} \\ &= k_3 + 3k_1k_2 + k_1^3. \end{aligned}$$

For concrete calculations, however, one would prefer to have a more analytical description of the relation between moments and cumulants. This can be achieved by translating the above relation to corresponding formal power series.

THEOREM 7.7. *Let $(m_n)_{n\geq 1}$ and $(k_n)_{n\geq 1}$ be two sequences of complex numbers and consider the corresponding formal power series*

$$M(z) := 1 + \sum_{n=1}^{\infty} m_n z^n, \tag{79}$$

$$C(z) := 1 + \sum_{n=1}^{\infty} k_n z^n. \tag{80}$$

Then the following three statements are equivalent:

(i) *We have for all $n \in \mathbb{N}$*

$$m_n = \sum_{\pi \in NC(n)} k_\pi = \sum_{\pi = \{V_1, \dots, V_r\} \in NC(n)} k_{\#V_1} \dots k_{\#V_r}. \tag{81}$$

(ii) *We have for all $n \in \mathbb{N}$ (where we put $m_0 := 1$)*

$$m_n = \sum_{s=1}^{n} \sum_{\substack{i_1, \dots, i_s \in \{0,1,\dots,n-s\} \\ i_1 + \dots + i_s = n-s}} k_s m_{i_1} \dots m_{i_s}. \tag{82}$$

(iii) *We have*

$$C[zM(z)] = M(z). \tag{83}$$

Proof. We rewrite the sum $m_n = \sum_{\pi \in NC(n)} k_\pi$ in the way that we fix the first block V_1 of π (i.e. that block which contains the element 1) and sum over all possibilities for the other blocks; in the end we sum over V_1:

$$m_n = \sum_{s=1}^{n} \sum_{V_1 \text{ with } \#V_1 = s} \sum_{\substack{\pi \in NC(n) \\ \text{where } \pi = \{V_1, \dots\}}} k_\pi.$$

If $V_1 = (v_1 = 1, v_2, \dots, v_s)$, then $\pi = \{V_1, \dots\} \in NC(n)$ can only connect elements lying between some v_k and v_{k+1}, i.e. $\pi = \{V_1, V_2, \dots, V_r\}$ such that we have for all $j = 2, \dots, r$: there exists a k with $v_k < V_j < v_{k+1}$. There we put $v_{s+1} := n + 1$. Hence such a π decomposes as

$$\pi = V_1 \cup \tilde{\pi}_1 \cup \dots \cup \tilde{\pi}_s,$$

where

$$\tilde{\pi}_j \text{ is a non-crossing partition of } \{v_j + 1, v_j + 2, \dots, v_{j+1} - 1\}.$$

For such π we have

$$k_\pi = k_{\#V_1} k_{\tilde{\pi}_1} \dots k_{\tilde{\pi}_s} = k_s k_{\tilde{\pi}_1} \dots k_{\tilde{\pi}_s},$$

and thus we obtain

$$\begin{aligned}
m_n &= \sum_{s=1}^{n} \sum_{1=v_1<v_2<\dots<v_s\le n} \sum_{\substack{\pi=V_1\cup\tilde\pi_1\cup\dots\cup\tilde\pi_s\\ \tilde\pi_j\in NC(v_j+1,\dots,v_{j+1}-1)}} k_s k_{\tilde\pi_1}\dots k_{\tilde\pi_s}\\
&= \sum_{s=1}^{n} k_s \sum_{1=v_1<v_2<\dots<v_s\le n} \Big(\sum_{\tilde\pi_1\in NC(v_1+1,\dots,v_2-1)} k_{\tilde\pi_1}\Big)\dots\Big(\sum_{\tilde\pi_s\in NC(v_s+1,\dots,n)} k_{\tilde\pi_s}\Big)\\
&= \sum_{s=1}^{n} k_s \sum_{1=v_1<v_2<\dots<v_s\le n} m_{v_2-v_1-1}\dots m_{n-v_s}\\
&= \sum_{s=1}^{n} \sum_{\substack{i_1,\dots,i_s\in\{0,1,\dots,n-s\}\\ i_1+\dots+i_s+s=n}} k_s m_{i_1}\dots m_{i_s} \qquad\qquad (i_k := v_{k+1}-v_k-1).
\end{aligned}$$

This yields the implication (i) $\Longrightarrow$ (ii).
We can now rewrite (ii) in terms of the corresponding formal power series in the following way (where we put $m_0 := k_0 := 1$):

$$\begin{aligned}
M(z) &= 1+\sum_{n=1}^{\infty} z^n m_n\\
&= 1+\sum_{n=1}^{\infty}\sum_{s=1}^{n} \sum_{\substack{i_1,\dots,i_s\in\{0,1,\dots,n-s\}\\ i_1+\dots+i_s=n-s}} k_s z^s m_{i_1} z^{i_1}\dots m_{i_s} z^{i_s}\\
&= 1+\sum_{s=1}^{\infty} k_s z^s \Big(\sum_{i=0}^{\infty} m_i z^i\Big)^s\\
&= C[zM(z)].
\end{aligned}$$

This yields (iii).
Since (iii) describes uniquely a fixed relation between the numbers $(k_n)_{n\ge1}$ and the numbers $(m_n)_{n\ge1}$, this has to be the relation (i). $\square$

If we rewrite the above relation between the formal power series in terms of the Cauchy-transform

$$G(z) := \sum_{n=0}^{\infty} \frac{m_n}{z^{n+1}} \tag{84}$$

and the R-transform

$$R(z) := \sum_{n=0}^{\infty} k_{n+1} z^n, \tag{85}$$

then we obtain Voiculescu's formula.

COROLLARY 7.8. *The relation between the Cauchy-transform $G(z)$ and the R-transform $R(z)$ of a random variable is given by*

$$G[R(z) + \frac{1}{z}] = z. \tag{86}$$

PROOF. We just have to note that the formal power series $M(z)$ and $C(z)$ from Theorem 7.7 and $G(z)$, $R(z)$, and $K(z) = R(z) + \frac{1}{z}$ are related by:

$$G(z) = \frac{1}{z}M(\frac{1}{z})$$

and

$$C(z) = 1 + zR(z) = zK(z), \qquad \text{thus} \qquad K(z) = \frac{C(z)}{z}.$$

This gives

$$K[G(z)] = \frac{1}{G(z)}C[G(z)] = \frac{1}{G(z)}C[\frac{1}{z}M(\frac{1}{z})] = \frac{1}{G(z)}M(\frac{1}{z}) = z,$$

thus $K[G(z)] = z$ and hence also

$$G[R(z) + \frac{1}{z}] = G[K(z)] = z.$$

$\square$

REMARKS 7.9. 1) It is quite easy to check that the cumulants k_n^a of a random variable a are indeed the coefficients of the R-transform of a as introduced by Voiculescu.
2) For a probability measure μ on $\mathbb{R}$ the Cauchy transform

$$G_\mu(z) = \int \frac{1}{z-t}d\mu(t) = \varphi(\frac{1}{z-x})$$

is not only a formal power series but an analytic function on the upper half plane

$$\begin{aligned} G_\mu : \mathbb{C}^+ &\to \mathbb{C}^- \\ z &\mapsto \int \frac{1}{z-t}d\mu(t). \end{aligned}$$

Furthermore it contains the essential information about the distribution μ in an accesible form, namely if μ has a continuous density h with respect to Lebesgue measure, i.e. $d\mu(t) = h(t)dt$, then we can recover this density from G_μ via the formula

$$h(t) = -\frac{1}{\pi}\lim_{\varepsilon\to 0}\Im G_\mu(t + i\varepsilon). \tag{87}$$

This can be seen as follows:

$$G_\mu(t+i\varepsilon) = \int \frac{1}{t+i\varepsilon - s} d\mu(s) = \int \frac{t-s-i\varepsilon}{(t-s)^2+\varepsilon^2} d\mu(s).$$

and thus for $d\mu(s) = h(s)ds$

$$\begin{aligned} -\frac{1}{\pi}\Im G_\mu(t+i\varepsilon) &= \frac{1}{\pi}\int \frac{\varepsilon}{(t-s)^2+\varepsilon^2} d\mu(s) \\ &= \frac{1}{\pi}\int \frac{\varepsilon}{(t-s)^2+\varepsilon^2} h(s)ds. \end{aligned}$$

The sequence of functions

$$f_\varepsilon(s) = \frac{1}{\pi}\frac{\varepsilon}{(t-s)^2+\varepsilon^2}$$

converges (in the sense of distributions) for $\varepsilon \to 0$ towards the 'delta function' δ_t (or to put it more probabilistically: the sequence of Cauchy distributions

$$\frac{1}{\pi}\frac{\varepsilon}{(t-s)^2+\varepsilon^2} ds$$

converges for $\varepsilon \to 0$ weakly towards the delta distribution δ_t). So for h continuous we get

$$\lim_{\varepsilon\to 0}(-\frac{1}{\pi}\Im G_\mu(t+i\varepsilon)) = \int h(s) d\delta_t(s) = h(t).$$

EXAMPLE 7.10. Let us use the Cauchy transform to calculate the density of a semi-circular s with variance 1 out of the knowledge about the moments. We have seen in the proof of Theorem 1.8 that the moments $m_{2n} = C_n$ of a semi circular obey the recursion formula

$$m_{2k} = \sum_{i=1}^{k} m_{2(i-1)} m_{2(k-i)}.$$

In terms of the formal power series

$$A(z) := \sum_{k=0}^{\infty} m_{2k} z^{2k} = 1 + \sum_{k=1}^{\infty} m_{2k} z^{2k} \tag{88}$$

this reads as

$$\begin{aligned}A(z) &= 1+\sum_{k=1}^{\infty} m_{2k}z^{2k}\\ &= 1+\sum_{k=1}^{\infty}\sum_{i=1}^{k}\big(z^2(m_{2(i-1)}z^{2(i-1)})(m_{2(k-i)}z^{2(k-i)})\big)\\ &= 1+z^2(1+\sum_{p=1}^{\infty} m_{2p}z^{2p})(1+\sum_{q=1}^{\infty} m_{2q}z^{2q})\\ &= 1+z^2A(z)A(z).\end{aligned}$$

This implies

$$A(z)=\frac{1-\sqrt{1-4z^2}}{2z^2}. \tag{89}$$

(Note: the other solution $(1+\sqrt{1-4z^2})/(2z^2)$ is ruled out because of the requirement $A(0)=1$.) This gives for the Cauchy tranform

$$\begin{aligned}G(z) &= \frac{A(1/z)}{z}\\ &= \frac{1-\sqrt{1-4/z^2}}{2/z}\\ &= \frac{z-\sqrt{z^2-4}}{2},\end{aligned}$$

and thus for the density

$$h(t)=-\frac{1}{\pi}\lim_{\varepsilon\to 0}\Im\frac{t+i\varepsilon-\sqrt{(t+i\varepsilon)^2-4}}{2}=\frac{1}{2\pi}\Im\sqrt{t^2-4}$$

or

$$h(t)=\begin{cases}\frac{1}{2\pi}\sqrt{4-t^2}, & |t|\le 2\\ 0, & \text{otherwise.}\end{cases} \tag{90}$$

REMARK 7.11. Combining all the above, we have a quite effective machinery for calculating the free convolution. Let μ, ν be probability measures on $\mathbb{R}$ with compact support, then we can calculate $\mu\boxplus\nu$ as follows:

$$\begin{aligned}\mu &\rightsquigarrow G_\mu \rightsquigarrow R_\mu\\ \nu &\rightsquigarrow G_\nu \rightsquigarrow R_\nu\end{aligned}$$

and

$$R_\mu, R_\nu \rightsquigarrow R_\mu+R_\nu=R_{\mu\boxplus\nu}\rightsquigarrow G_{\mu\boxplus\nu}\rightsquigarrow \mu\boxplus\nu.$$

EXAMPLE 7.12. Let
$$\mu = \nu = \frac{1}{2}(\delta_{-1} + \delta_{+1}).$$
Then we have
$$G_\mu(z) = \int \frac{1}{z-t} d\mu(t) = \frac{1}{2}\Big(\frac{1}{z+1} + \frac{1}{z-1}\Big) = \frac{z}{z^2-1}.$$
Put
$$K_\mu(z) = \frac{1}{z} + R_\mu(z).$$
Then $z = G_\mu[K_\mu(z)]$ gives
$$K_\mu(z)^2 - \frac{K_\mu(z)}{z} = 1,$$
which has as solutions
$$K_\mu(z) = \frac{1 \pm \sqrt{1+4z^2}}{2z}.$$
Thus the R-transform of μ is given by
$$R_\mu(z) = K_\mu(z) - \frac{1}{z} = \frac{\sqrt{1+4z^2}-1}{2z}$$
(note: $R_\mu(0) = k_1(\mu) = m_1(\mu) = 0$). Hence we get
$$R_{\mu\boxplus\mu}(z) = 2R_\mu(z) = \frac{\sqrt{1+4z^2}-1}{z},$$
and
$$K(z) := K_{\mu\boxplus\mu}(z) = R_{\mu\boxplus\mu}(z) + \frac{1}{z} = \frac{\sqrt{1+4z^2}}{z},$$
which allows to determine $G := G_{\mu\boxplus\mu}$ via
$$z = K[G(z)] = \frac{\sqrt{1+4G(z)^2}}{G(z)}$$
as
$$G(z) = \frac{1}{\sqrt{z^2-4}}$$
¿From this we can calculate the density
$$\frac{d(\mu \boxplus \mu)(t)}{dt} = -\frac{1}{\pi} \lim_{\varepsilon \to 0} \Im \frac{1}{\sqrt{(t+i\varepsilon)^2-4}} = -\frac{1}{\pi} \Im \frac{1}{\sqrt{t^2-4}},$$
and finally
$$\frac{d(\mu \boxplus \mu)(t)}{dt} = \begin{cases} \frac{1}{\pi\sqrt{4-t^2}}, & |t| \le 2 \\ 0, & \text{otherwise} \end{cases} \tag{91}$$

REMARKS 7.13. 1) The distribution $1/(\pi\sqrt{4-t^2})dt$ is also of importance in classical probability theory. It is called **arc-sine distribution**.
2) The free convolution has the quite surprising (from the probabilistic point of view) property that the convolution of discrete distributions can be an absolutely continuous distribution. (From a functional analytic point of view this is of course not so surprising because it is a well known phenomena that in general the spectrum of the sum of two operators has almost no relation to the spectra of the single operators.
3) In particular, we see that $\boxplus$ is not distributive, as

$$\frac{1}{2}(\delta_{-1}+\delta_{+1})\boxplus\mu \neq \frac{1}{2}(\delta_{-1}\boxplus\mu)+\frac{1}{2}(\delta_{+1}\boxplus\mu)=\frac{1}{2}\mu_{(-1)}+\frac{1}{2}\mu_{(+1)},$$

where $\mu_{(a)}=\mu\boxplus\delta_a$ is the shift of the measure μ by the amount a.

EXAMPLE 7.14. With the help of the R-transform machinery we can now give a more analytic and condensed proof of the central limit theorem: Since free cumulants are polynomials in moments and vice versa the convergence of moments is equivalent to the convergence of cumulants. This means what we have to show for $a_1, a_2, \dots$ being free, identically distributed, centered and with variance σ^2 is

$$R_{(a_1+\cdots+a_n)/\sqrt{n}}(z)\to R_s(z)=\sigma^2 z$$

in the sense of convergence of the coefficients of the formal power series. It is easy to see that

$$R_{\lambda a}(z)=\lambda R_a(\lambda z).$$

Thus we get

$$\begin{aligned}
R_{(a_1+\cdots+a_n)/\sqrt{n}}(z) &= \frac{1}{\sqrt{n}}R_{a_1+\cdots+a_n}(\frac{z}{\sqrt{n}})\\
&= n\frac{1}{\sqrt{n}}R_{a_i}(\frac{z}{\sqrt{n}})\\
&= \sqrt{n}R_{a_i}(\frac{z}{\sqrt{n}})\\
&= \sqrt{n}(k_1+k_2\frac{z}{\sqrt{n}}+k_3\frac{z^2}{n}+\dots)\\
&= \sqrt{n}(\sigma^2\frac{z}{\sqrt{n}}+k_3\frac{z^2}{n}+\dots)\\
&\to \sigma^2 z,
\end{aligned}$$

since $k_1=0$ and $k_2=\sigma^2$.

CHAPTER 8

Products of free variables

1. Description of the product of free variables

REMARKS 8.1. 1) In the last section we treated the sum of free variables, in particular we showed how one can understand and solve from a combinatorial point of view the problem of describing the distribution of $a + b$ in terms of the distributions of a and of b if these variables are free. Now we want to turn to the corresponding problem for the product. Thus we want to understand how we get the distribution of ab out of the distributions of a and of b for a and b free. Note that in the classical case no new considerations are required since this problem can be reduced to the additive problem. Namely we have $ab = \exp(\log a + \log b)$ and thus we only need to apply the additive theory to $\log a$ and $\log b$. In the non-commutative situation, however, the functional equation for the exponential function is not true any more, so there is no clear way to reduce the multiplicative problem to the additive one. Indeed, one needs new considerations. Fortunately, our combinatorial approach allows also such a treatment. It will turn out that the description of the multiplication of free variables is intimately connected with the complementation map K in the lattice of non-crossing partitions. Note that there is no counterpart of this for all partitions. Thus statements around the multiplication of free variables might be quite different from what one expects classically. With respect to additive problems classical and free probability theory go quite parallel (combinatorially this means that one just replaces arguments for all partitions by the corresponding arguments for non-crossing partitions), with respect to multiplicative problems the world of free probability is, however, much richer.
2) As usual, the combinatorial nature for the problem is the same in the one-dimensional and in the multi-dimensional case. Thus we will consider from the beginning the latter case.

THEOREM 8.2. *Let $(\mathcal{A}, \varphi)$ be a probability space and consider random variables $a_1, \dots, a_n, b_1, \dots, b_n \in \mathcal{A}$ such that $\{a_1, \dots, a_n\}$ and*

$\{b_1, \dots, b_n\}$ *are free. Then we have*

(92)

$$\varphi(a_1b_1a_2b_2\dots a_nb_n) = \sum_{\pi\in NC(n)} k_\pi[a_1, a_2, \dots, a_n] \cdot \varphi_{K(\pi)}[b_1, b_2, \dots, b_n]$$

and

(93)

$$k_n(a_1b_1, a_2b_2, \dots, a_nb_n) = \sum_{\pi\in NC(n)} k_\pi[a_1, a_2, \dots, a_n] \cdot k_{K(\pi)}[b_1, b_2, \dots, b_n].$$

PROOF. Although both equations are equivalent and can be transformed into each other we will provide an independent proof for each of them.

i) By using the vanishing of mixed cumulants in free variables we obtain

$$\begin{aligned}\varphi(a_1b_1a_2b_2\dots a_nb_n) &= \sum_{\pi\in NC(2n)} k_\pi[a_1, b_1, a_2, b_2, \dots, a_n, b_n] \\ &= \sum_{\substack{\pi_a\in NC(1,3,\dots,2n-1),\pi_b\in NC(2,4,\dots,2n)\\ \pi_a\cup\pi_b\in NC(2n)}} k_{\pi_a}[a_1, a_2, \dots, a_n] \cdot k_{\pi_b}[b_1, b_2, \dots, b_n] \\ &= \sum_{\pi_a\in NC(1,3,\dots,2n-1)} k_{\pi_a}[a_1, a_2, \dots, a_n] \cdot \Big(\sum_{\substack{\pi_b\in NC(2,4,\dots,2n)\\ \pi_a\cup\pi_b\in NC(2n)}} k_{\pi_b}[b_1, b_2, \dots, b_n]\Big).\end{aligned}$$

Now note that, for fixed $\pi_a \in NC(1, 3, \dots, 2n-1) \hat{=} NC(n)$, the condition $\pi_a \cup \pi_b \in NC(2n)$ for $\pi_b \in NC(2, 4, \dots, 2n) \hat{=} NC(n)$ means nothing but $\pi_b \leq K(\pi_a)$ (since $K(\pi_a)$ is by definition the biggest element with this property). Thus we can continue

$$\begin{aligned}&\varphi(a_1b_1a_2b_2\dots a_nb_n)\\ &\qquad= \sum_{\pi_a\in NC(n)} k_{\pi_a}[a_1, a_2, \dots, a_n] \cdot \Big(\sum_{\pi_b\leq K(\pi_a)} k_{\pi_b}[b_1, b_2, \dots, b_n]\Big) \\ &\qquad= \sum_{\pi_a\in NC(n)} k_{\pi_a}[a_1, a_2, \dots, a_n] \cdot \varphi_{K(\pi_a)}[b_1, b_2, \dots, b_n].\end{aligned}$$

ii) By using Theorem 5.2 for cumulants with products as entries we get

$$k_n(a_1b_1, \dots, a_nb_n) = \sum_{\substack{\pi\in NC(2n)\\ \pi\vee\sigma=1_{2n}}} k_\pi[a_1, b_1, \dots, a_n, b_n],$$

where $\sigma = \{(1, 2), (3, 4), \dots, (2n-1, 2n)\}$. By the vanishing of mixed cumulants only such π contribute in the sum which do not couple a's with b's, thus they are of the form $\pi = \pi_a \cup \pi_b$ with

$\pi_a \in NC(a_1, a_2, \dots, a_n)$ and $\pi_b \in NC(b_1, b_2, \dots, b_n)$. Fix now an arbitrary π_a. Then we claim that there exists exactly one π_b such that $\pi_a \cup \pi_b$ is non-crossing and that $(\pi_a \cup \pi_b) \vee \sigma = 1_{2n}$. This can be seen as follows. π_a must contain a block $(a_l, a_{l+1}, \dots, a_{l+p})$ consisting of consecutive elements, i.e. we have the following situation:

$$\dots\ a_{l-1}\, b_{l-1}\ ,\ a_l\, b_l\ ,\ a_{l+1}\, b_{l+1}\ ,\ \dots\ ,\ a_{l+p-1}\, b_{l+p-1}\ ,\ a_{l+p}\, b_{l+p}\ ,\ \dots$$

But then b_{l-1} must be connected with b_{l+p} via π_b because otherwise $a_l b_l \dots a_{l+p} b_{l+p}$ cannot become connected with $a_{l-1} b_{l-1}$. But if b_{l+p} is connected with b_{l-1} then we can just take away the interval $a_l b_l \dots a_{l+p} b_{l+p}$ and continue to argue for the rest. Thus we see that in order to get $(\pi_a \cup \pi_b) \vee \sigma = 1_{2n}$ we have to make the blocks for π_b as large as possible, i.e. we must take $K(\pi_a)$ for π_b. As is also clear from the foregoing induction argument the complement will really fulfill the condition $(\pi_a \cup K(\pi_a)) \vee \sigma = 1_{2n}$. Hence the above sum reduces to

$$k_n(a_1 b_1, \dots, a_n b_n) = \sum_{\pi_a \in NC(n)} k_{\pi_a}[a_1, \dots, a_n] \cdot k_{K(\pi_a)}[b_1, \dots, b_n].$$

$\square$

REMARK 8.3. Theorem 8.2 can be used as a starting point for a combinatorial treatment of Voiculescu's description of multiplicative free convolution via the so-called S-transform. This is more complicated as in the additive case, but nevertheless one can give again a purely combinatorial and elementary (though not easy) proof of the results of Voiculescu.

2. Compression by a free projection

NOTATION 8.4. If $(\mathcal{A}, \varphi)$ is a probability space and $p \in \mathcal{A}$ a projection (i.e. $p^2 = p$) such that $\varphi(p) \neq 0$, then we can consider the **compression** $(p\mathcal{A}p, \varphi^{p\mathcal{A}p})$, where

$$\varphi^{p\mathcal{A}p}(\cdot) := \frac{1}{\varphi(p)} \varphi(\cdot) \qquad \text{restricted to } p\mathcal{A}p.$$

Of course, $(p\mathcal{A}p, \varphi^{p\mathcal{A}p})$ is also a probability space. We will denote the cumulants corresponding to $\varphi^{p\mathcal{A}p}$ by $k^{p\mathcal{A}p}$.

EXAMPLE 8.5. If $\mathcal{A} = M_4$ are the 4×4-matrices equipped with the normalized trace $\varphi = \mathrm{tr}_4$ and p is the projection

$$p = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

then

$$p \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \alpha_{14} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & \alpha_{24} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & \alpha_{34} \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & \alpha_{44} \end{pmatrix} p = \begin{pmatrix} \alpha_{11} & \alpha_{12} & 0 & 0 \\ \alpha_{21} & \alpha_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

and going over to the compressed space just means that we throw away the zeros and identify pM_4p with the 2×2-matrices M_2. Of course, the renormalized state $\mathrm{tr}_4^{p\mathcal{A}p}$ coincides with the state tr_2 on M_2.

THEOREM 8.6. *Let $(\mathcal{A}, \varphi)$ be a probability space and consider random variables $p, a_1, \dots, a_m \in \mathcal{A}$ such that p is a projection with $\varphi(p) \neq 0$ and such that p is free from $\{a_1, \dots, a_m\}$. Then we have the following relation between the cumulants of $a_1, \dots, a_m \in \mathcal{A}$ and the cumulants of the compressed variables $pa_1p, \dots, pa_mp \in p\mathcal{A}p$: For all $n \geq 1$ and all $1 \leq i(1), \dots, i(n) \leq m$ we have*

$$k_n^{p\mathcal{A}p}(pa_{i(1)}p, \dots, pa_{i(n)}p) = \frac{1}{\lambda} k_n(\lambda a_{i(1)}, \dots, \lambda a_{i(n)}), \tag{94}$$

where $\lambda := \varphi(p)$.

REMARK 8.7. The fact that p is a projection implies

$$\varphi(pa_{i(1)}ppa_{i(2)}p \dots pa_{i(n)}p) = \varphi(pa_{i(1)}pa_{i(2)}p \dots a_{i(n)}p),$$

so that apart from the first p we are in the situation where we have a free product of a's with p. If we would assume a tracial situation, then of course the first p could be absorbed by the last one. However, we want to treat the theorem in full generality. Namely, even without traciality we can arrive at the situation from Theorem 8.2, just by enlarging $\{a_1, \dots, a_m\}$ to $\{1, a_1, \dots, a_m\}$ (which does not interfere with the freeness assumption because 1 is free from everything) and reading $\varphi(pa_{i(1)}pa_{i(2)}p \dots a_{i(n)}p)$ as $\varphi(1pa_{i(1)}pa_{i(2)}p \dots a_{i(n)}p)$.

Proof. We have

$$\begin{aligned}
\varphi_n^{p\mathcal{A}p}(pa_{i(1)}p,\dots,pa_{i(n)}p) &= \frac{1}{\lambda}\varphi_n(pa_{i(1)}p,\dots,pa_{i(n)}p)\\
&= \frac{1}{\lambda}\varphi_{n+1}(1p,a_{i(1)}p,\dots,a_{i(n)}p)\\
&= \frac{1}{\lambda}\sum_{\sigma\in NC(n+1)} k_\sigma[1,a_{i(1)},\dots,a_{i(n)}]\cdot\varphi_{K(\sigma)}[p,p,\dots,p].
\end{aligned}$$

Now we observe that $k_\sigma[1,a_{i(1)},\dots,a_{i(n)}]$ can only be different from zero if σ does not couple the random variable 1 with anything else, i.e. $\sigma\in NC(0,1,\dots,n)$ must be of the form $\sigma=(0)\cup\pi$ with $\pi\in NC(1,\dots,n)$. So in fact the sum runs over $\pi\in NC(n)$ and $k_\sigma[1,a_{i(1)},\dots,a_{i(n)}]$ is nothing but $k_\pi[a_{i(1)},\dots,a_{i(n)}]$. Furthermore, the relation between $K(\sigma)\in NC(0,1,\dots,n)$ and $K(\pi)\in NC(1,\dots,n)$ for $\sigma=(0)\cup\pi$ is given by the observation that on $1,\dots,n$ they coincide and 0 and n are always in the same block of $K(\sigma)$. In particular, $K(\sigma)$ has the same number of blocks as $K(\pi)$. But the fact $p^2=p$ gives

$$\varphi_{K(\sigma)}[p,p,\dots,p]=\lambda^{|K(\sigma)|}=\lambda^{|K(\pi)|}=\lambda^{n+1-|\pi|},$$

where we used for the last equality the easily checked fact that

$$|\pi|+|K(\pi)|=n+1 \qquad \text{for all } \pi\in NC(n). \tag{95}$$

Now we can continue our above calculation.

$$\begin{aligned}
\varphi_n^{p\mathcal{A}p}(pa_{i(1)}p,\dots,pa_{i(n)}p) &= \frac{1}{\lambda}\sum_{\pi\in NC(n)} k_\pi[a_{i(1)},\dots,a_{i(n)}]\lambda^{n+1-|\pi|}\\
&= \sum_{\pi\in NC(n)} \frac{1}{\lambda^{|\pi|}}k_\pi[\lambda a_{i(1)},\dots,\lambda a_{i(n)}].
\end{aligned}$$

Since on the other hand we also know that

$$\varphi_n^{p\mathcal{A}p}(pa_{i(1)}p,\dots,pa_{i(n)}p)=\sum_{\pi\in NC(n)} k_\pi^{p\mathcal{A}p}[pa_{i(1)}p,\dots,pa_{i(n)}p],$$

we get inductively the assertion. Note the compatibility with multiplicative factorization, i.e. if we have

$$k_n^{p\mathcal{A}p}[pa_{i(1)}p,\dots,pa_{i(n)}p]=\frac{1}{\lambda}k_n(\lambda a_{i(1)},\dots,a_{i(n)})$$

then this implies

$$k_\pi^{p\mathcal{A}p}[pa_{i(1)}p,\dots,pa_{i(n)}p]=\frac{1}{\lambda^{|\pi|}}k_\pi(\lambda a_{i(1)},\dots,a_{i(n)}).$$

$\square$

Corollary 8.8. *Let $(\mathcal{A},\varphi)$ be a probability space and $p\in\mathcal{A}$ a projection such that $\varphi(p)\neq 0$. Consider unital subalgebras $\mathcal{A}_1,\dots,\mathcal{A}_m\subset\mathcal{A}$ such that p is free from $\mathcal{A}_1\cup\cdots\cup\mathcal{A}_m$. Then the following two statements are equivalent:*

1. *The subalgebras $\mathcal{A}_1,\dots,\mathcal{A}_m\subset\mathcal{A}$ are free in the original probability space $(\mathcal{A},\varphi)$.*
2. *The compressed subalgebras $p\mathcal{A}_1p,\dots,p\mathcal{A}_mp\subset p\mathcal{A}p$ are free in the compressed probability space $(p\mathcal{A}p,\varphi^{p\mathcal{A}p})$.*

Proof. Since the cumulants of the $\mathcal{A}_i$ coincide with the cumulants of the compressions $p\mathcal{A}_ip$ up to some power of λ the vanishing of mixed cumulants in the $\mathcal{A}_i$ is equivalent to the vanishing of mixed cumulants in the $p\mathcal{A}_ip$. □

Corollary 8.9. *Let μ be a probability measure on $\mathbb{R}$ with compact support. Then, for each real $t\geq 1$, there exists a probability measure μ_t, such that we have $\mu_1=\mu$ and*

$$\mu_{s+t}=\mu_s\boxplus\mu_t \qquad \textit{for all real } s,t\geq 1. \tag{96}$$

Remark 8.10. For $t=n\in\mathbb{N}$, we have of course the convolution powers $\mu_n=\mu^{\boxplus n}$. The corollary states that we can interpolate between them also for non-natural powers. Of course, the crucial fact is that we claim the μ_t to be always probability measures. As linear functionals these objects exist trivially, the non-trivial fact is positivity.

Proof. Let x be a self-adjoint random variable and p be a projection in some C^*-probability space $(\mathcal{A},\varphi)$ such that $\varphi(p)=\frac{1}{t}$, the distribution of x is equal to μ, and x and p are free. (It is no problem to realize such a situation with the usual free product constructions.) Put now $x_t:=p(tx)p$ and consider this as an element in the compressed space $(p\mathcal{A}p,\varphi^{p\mathcal{A}p})$. It is clear that this compressed space is again a C^*-probability space, thus the distribution μ_t of $x_t\in p\mathcal{A}p$ is again a probability measure. Furthermore, by Theorem 8.6, we know that the cumulants of x_t are given by

$$k_n^{\mu_t}=k_n^{p\mathcal{A}p}(x_t,\dots,x_t)=tk_n(\frac{1}{t}tx,\dots,\frac{1}{t}tx)=tk_n(x,\dots,x)=tk_n^{\mu}.$$

This implies that for all $n\geq 1$

$$k_n^{\mu_{s+t}}=(s+t)k_n^{\mu}=sk_n^{\mu}+tk_n^{\mu}=k_n^{\mu_s}+k_n^{\mu_t},$$

which just means that $\mu_{s+t}=\mu_s\boxplus\mu_t$. Since $k_n^{\mu_1}=k_n^{\mu}$, we also have $\mu_1=\mu$. □

Remarks 8.11. 1) Note that the corollary states the existence of μ_t only for $t\geq 1$. For $0<t<1$, μ_t does not exist as a probability

measure in general. In particular, the existence of the semi-group μ_t for all $t > 0$ is equivalent to μ being infinitely divisible (in the free sense).
2) There is no classical analogue of the semi-group μ_t for all $t \geq 1$. In the classical case one can usually not interpolate between the natural convolution powers. E.g., if $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$ is a Bernoulli distribution, we have $\mu * \mu * \mu = \frac{1}{8}\delta_{-3} + \frac{3}{8}\delta_{-1} + \frac{3}{8}\delta_1 + \frac{1}{8}\delta_3$ and it is trivial to check that there is no possibility to write $\mu * \mu * \mu$ as $\nu * \nu$ for some other probability measure $\nu = \mu^{*3/2}$.

Remark 8.12. By using the connection between $N \times N$-random matrices and freeness one can get a more 'concrete' picture of the compression results. The random matrix results are only asymptotically true for $N \to \infty$. However, we will suppress in the following this limit and just talk about $N \times N$ matrices, which give for large N an arbitrary good approximation of the asymptotic results. Let A be a deterministic $N \times N$-matrix with distribution μ and let U be a random $N \times N$-unitary. Then UAU^* has, of course, the same distribution μ as A and furthermore, in the limit $N \to \infty$, it becomes free from the constant matrices. Thus we can take as a canonical choice for our projection p the diagonal matrix which has λN ones and $(1-\lambda)N$ zeros on the diagonal. The compression of the matrices corresponds then to taking the upper left corner of length λN of the matrix UAU^*. Our above results say then in particular that, by varying λ, the upper left corners of the matrix UAU^* give, up to renormalization, the wanted distributions μ_t (with $t = 1/\lambda$). Furthermore if we have two deterministic matrices A_1 and A_2, the joint distribution of A_1, A_2 is not destroyed by conjugating them by the same random unitary U. In this 'randomly rotated' realization of A_1 and A_2, the freeness of UA_1U^* from UA_2U^* is equivalent to the freeness of a corner of UA_1U^* from the corresponding corner of UA_2U^*.

3. Compression by a free family of matrix units

Definition 8.13. Let $(\mathcal{A}, \varphi)$ be a probability space. A **family of matrix units** is a set $\{e_{ij}\}_{i,j=1,\dots,d} \subset \mathcal{A}$ (for some $d \in \mathbb{N}$) with the properties

$$e_{ij}e_{kl} = \delta_{jk}e_{il} \qquad \text{for all } i,j,k,l = 1,\dots,d \tag{97}$$

$$\sum_{i=1}^{d} e_{ii} = 1 \tag{98}$$

$$\varphi(e_{ij}) = \delta_{ij}\frac{1}{d} \qquad \text{for all } i,j = 1,\dots,d. \tag{99}$$

THEOREM 8.14. *Let $(\mathcal{A},\varphi)$ be a probability space and consider random variables $a^{(1)},\dots,a^{(m)} \in \mathcal{A}$. Furthermore, let $\{e_{ij}\}_{i,j=1,\dots,d} \subset \mathcal{A}$ be a family of matrix units such that $\{a^{(1)},\dots,a^{(m)}\}$ is free from $\{e_{ij}\}_{i,j=1,\dots,d}$. Put now $a_{ij}^{(r)} := e_{1i}a^{(r)}e_{j1}$ and $p := e_{11}$, $\lambda := \varphi(p) = 1/d$. Then we have the following relation between the cumulants of $a^{(1)},\dots,a^{(m)} \in \mathcal{A}$ and the cumulants of the compressed variables $a_{ij}^{(r)} \in p\mathcal{A}p$ ($i,j = 1,\dots,d$; $r = 1,\dots,m$) : For all $n \geq 1$ and all $1 \leq r(1),\dots,r(n) \leq m$, $1 \leq i(1),j(1),\dots,i(n),j(n) \leq d$ we have*

$$k_n^{p\mathcal{A}p}(a_{i(1)j(1)}^{(r(1))},\dots,a_{i(n)j(n)}^{(r(n))}) = \frac{1}{\lambda}k_n(\lambda a^{(r(1))},\dots,\lambda a^{(r(n))}) \tag{100}$$

if $j(k) = i(k+1)$ for all $k = 1,\dots,n$ (where we put $i(n+1) := i(1)$) and zero otherwise.

NOTATION 8.15. Let a partition $\pi \in NC(n)$ and an n-tuple of double-indices $(i(1)j(1), i(2)j(2),\dots,i(n)j(n))$ be given. Then we say that π **couples in a cyclic way** (c.c.w., for short) the indices $(i(1)j(1), i(2)j(2),\dots,i(n)j(n))$ if we have for each block $(r_1 < r_2 < \dots < r_s) \in \pi$ that $j(r_k) = i(r_{k+1})$ for all $k = 1,\dots,s$ (where we put $r_{s+1} := r_1$).

PROOF. As in the case of one free projection we calculate

$$\begin{aligned}&\varphi_n^{p\mathcal{A}p}(a_{i(1)j(1)}^{(r(1))},\dots,a_{i(n)j(n)}^{(r(n))})\\ &= \frac{1}{\lambda}\sum_{\sigma\in NC(n+1)} k_\sigma[1,a^{(r(1))},\dots,a^{(r(n))}]\cdot\varphi_{K(\sigma)}[e_{1,i(1)},e_{j(1)i(2)},\dots,e_{j(n)1}].\end{aligned}$$

Again σ has to be of the form $\sigma = (0)\cup\pi$ with $\pi \in NC(n)$. The factor $\varphi_{K(\sigma)}$ gives

$$\begin{aligned}&\varphi_{K(\sigma)}[e_{1,i(1)},e_{j(1)i(2)},\dots,e_{j(n)1}] = \varphi_{K(\pi)}[e_{j(1)i(2)},e_{j(2),j(3)},\dots,e_{j(n)i(1)}]\\ &= \begin{cases}\lambda^{|K(\pi)|}, & \text{if } K(\pi) \text{ c.c.w. } (j(1)i(2),j(2)i(3),\dots,j(n)i(1))\\ 0, & \text{otherwise}\end{cases}\end{aligned}$$

Now one has to observe that cyclicity of $K(\pi)$ in $(j(1)i(2), j(2)i(3),\dots,j(n)i(1))$ is equivalent to cyclicity of π in

$(i(1)j(1), i(2)j(2), \dots, i(n)j(n))$. This claim can be seen as follows: $K(\pi)$ has a block $V = (l, l+1, \dots, l+p)$ consisting of consecutive elements. The relevant part of π is given by the fact that $(l+1), (l+2), \dots, (l+p) \in \pi$ form singletons and that l and $l+p+1$ lie in the same block of π. But cyclicity of $K(\pi)$ in $(j(1)i(2), j(2)i(3), \dots, j(n)i(1))$ evaluated for the block V means that $i(l) = j(l)$, $i(l+1) = j(l+1)$, $\dots$, $i(l+p) = j(l+p)$, and $i(l+p+1) = j(l)$. This, however, are the same requirements as given by cyclicity of π in $(i(1)j(1), i(2)j(2), \dots, i(n)j(n))$ evaluated for the relevant part of π. Now one can take out the points $l, l+1, \dots, l+p$ from $K(\pi)$ and repeat the above argumentation for the rest. This gives a recursive proof of the claim. Having this claim one can continue the above calculation as follows.

$$\begin{aligned}
&\varphi_n^{p\mathcal{A}p}(a_{i(1)j(1)}^{(r(1))}, \dots, a_{i(n)j(n)}^{(r(n))}) \\
&\qquad = \frac{1}{\lambda} \sum_{\substack{\pi \in NC(n) \\ \pi \text{ c.c.w. } (i(1)j(1),\, i(2)j(2),\, \dots,\, i(n)j(n))}} k_\pi[a^{(r(1))}, \dots, a^{(r(n))}] \cdot \lambda^{|K(\pi)|} \\
&\qquad = \sum_{\substack{\pi \in NC(n) \\ \pi \text{ c.c.w. } (i(1)j(1),\, i(2)j(2),\, \dots,\, i(n)j(n))}} \frac{1}{\lambda^{|\pi|}} k_\pi[\lambda a^{(r(1))}, \dots, \lambda a^{(r(n))}],
\end{aligned}$$

where we sum only over such π which couple in a cyclic way $(i(1)j(1), i(2)j(2), \dots, i(n)j(n))$. By comparison with the formula

$$\varphi_n^{p\mathcal{A}p}(a_{i(1)j(1)}^{(r(1))}, \dots, a_{i(n)j(n)}^{(r(n))}) = \sum_{\pi \in NC(n)} k_\pi^{p\mathcal{A}p}[a_{i(1)j(1)}^{(r(1))}, \dots, a_{i(n)j(n)}^{(r(n))}]$$

this gives the statement. $\square$

COROLLARY 8.16. *Let $(\mathcal{A}, \varphi)$ be a probability space. Consider a family of matrix units $\{e_{ij}\}_{i,j=1,\dots,d} \subset \mathcal{A}$ and a subset $\mathcal{X} \subset \mathcal{A}$ such that $\{e_{ij}\}_{i,j=1,\dots,d}$ and $\mathcal{X}$ are free. Consider now, for $i = 1, \dots, d$, the compressed subsets $\mathcal{X}_i := e_{1i}\mathcal{X}e_{i1} \subset e_{11}\mathcal{A}e_{11}$. Then $\mathcal{X}_1, \dots, \mathcal{X}_d$ are free in the compressed probability space $(e_{11}\mathcal{A}e_{11}, \varphi^{e_{11}\mathcal{A}e_{11}})$.*

PROOF. Since there are no $\pi \in NC(n)$ which couple in a cyclic way indices of the form $(i(1)i(1), i(2)i(2), \dots, i(n)i(n))$ if $i(k) \neq i(l)$ for some l and k, Theorem 8.14 implies that mixed cumulants in elements from $\mathcal{X}_1, \dots, \mathcal{X}_d$ vanish. By the criterion for freeness in terms of cumulants, this just means that $\mathcal{X}_1, \dots, \mathcal{X}_d$ are free. $\square$

REMARK 8.17. Again one has a random matrix realization of this statement: Take $\mathcal{X}$ as a subset of deterministic $N \times N$-matrices and consider a randomly rotated version $\tilde{\mathcal{X}} := U\mathcal{X}U^*$ of this. Then, for $N = Md$ we write $M_N = M_M \otimes M_d$ and take as $e_{ij} = 1 \otimes E_{ij}$ the embedding of the canonical matrix units of M_d into M_N. These matrix units are free (for $N \to \infty$, keeping d fixed) from $\tilde{\mathcal{X}}$ and in this picture $e_{1i}\tilde{\mathcal{X}}e_{j1}$ corresponds to $\tilde{\mathcal{X}} \otimes E_{ij}$, i.e. we are just splitting our $N \times N$-matrices into $N/d \times N/d$-sub-matrices. The above corollary states that the sub-matrices on the diagonal are asymptotically free.

THEOREM 8.18. *Consider random variables a_{ij} $(i, j = 1, \dots, d)$ in some probability space $(\mathcal{A}, \varphi)$. Then the following two statements are equivalent.*

1. *The matrix $A := (a_{ij})_{i,j=1}^d$ is free from $M_d(\mathbb{C})$ in the probability space $(M_d(\mathcal{A}), \varphi \otimes tr_d)$.*
2. *The joint cumulants of $\{a_{ij} \mid i, j = 1, \dots, d\}$ in the probability space $(\mathcal{A}, \varphi)$ have the following property: only cyclic cumulants $k_n(a_{i(1)i(2)}, a_{i(2)i(3)}, \dots, a_{i(n)i(1)})$ are different from zero and the value of such a cumulant depends only on n, but not on the tuple $(i(1), \dots, i(n))$.*

PROOF. (1) $\Longrightarrow$ (2) follows directly from Theorem 8.14 (for the case $m = 1$), because we can identify the entries of the matrix A with the compressions by the matrix units.
For the other direction, note that moments (with respect to $\varphi \otimes \mathrm{tr}_d$) in the matrix $A := (a_{ij})_{i,j=1}^d$ and elements from $M_d(\mathbb{C})$ can be expressed in terms of moments of the entries of A. Thus the freeness between A and $M_d(\mathbb{C})$ depends only on the joint distribution of (i.e. on the cumulants in) the a_{ij}. This implies that if we can present a realization of the a_{ij} in which the corresponding matrix $A = (a_{ij})_{i,j=1}^d$ is free from $M_d(\mathbb{C})$, then we are done. But this representation is given by Theorem 8.14. Namely, let a be a random variable whose cumulants are given, up to a factor, by the cyclic cumulants of the a_{ij}, i.e. $k_n^a = d^{n-1}k_n(a_{i(1)i(2)}, a_{i(2)i(3)}, \dots, a_{i(n)i(1)})$. Let furthermore $\{e_{ij}\}_{i,j=1,\dots,d}$ be a family of matrix units which are free from a in some probability space $(\tilde{\mathcal{A}}, \tilde{\varphi})$. Then we compress a with the free matrix units as in Theorem 8.14 and denote the compressions by $\tilde{a}_{ij} := e_{1i}ae_{j1}$. By Theorem 8.14 and the choice of the cumulants for a, we have that the joint distribution of the $\tilde{a}_{ij}$ in $(e_{11}\tilde{\mathcal{A}}e_{11}, \tilde{\varphi}^{e_{11}\tilde{\mathcal{A}}e_{11}})$ coincides with the joint distribution of the a_{ij}. Furthermore, the matrix $\tilde{A} := (\tilde{a}_{ij})_{i,j=1}^d$ is free from $M_d(\mathbb{C})$ in $(M_d(e_{11}\tilde{\mathcal{A}}e_{11}), \tilde{\varphi}^{e_{11}\tilde{\mathcal{A}}e_{11}} \otimes \mathrm{tr}_d)$, because the mapping

$$\tilde{\mathcal{A}} \to M_d(e_{11}\tilde{\mathcal{A}}e_{11})$$
$$y \mapsto (e_{1i}ye_{j1})_{i,j=1}^d$$

is an isomorphism which sends a into $\tilde{A}$ and e_{kl} into the canonical matrix units E_{kl} in $1 \otimes M_d(\mathbb{C})$. □

CHAPTER 9

R-diagonal elements and free anti-commutator

REMARK 9.1. There is a quite substantial difference in our understanding of one self-adjoint (or normal) operator on one side and more non-commuting self-adjoint operators on the other side. Whereas the first case takes place in the classical commutative world, where we have at our hand the sophisticated tools of analytic function theory, the second case is really non-commutative in nature and presents a lot of difficulties. This difference shows also up in our understanding of fundamental concepts in free probability. (An important example of this is free entropy. Whereas the case of one self-adjoint variable is well understood and there are concrete formulas in that case, the multi-dimensional case is the real challenge. In particular, we would need an understanding or more explicit formulas for the general case of two self-adjoint operators, which is the same as the general case of one not necessarily normal operator.) In this section we present a special class of non-normal operators which are of some interest, because they are on one side simple enough to allow concrete calculations, but on the other side this class is also big enough to appear quite canonically in a lot of situations.

NOTATION 9.2. Let a be a random variable in a $*$-probability space. A cumulant $k_{2n}(a_1, \dots, a_{2n})$ with arguments from $\{a, a^*\}$ is said to have **alternating arguments**, if there does not exist any $a_i\,(1 \leq i \leq 2n-1)$ with $a_{i+1} = a_i$. We will also say that the cumulant $k_{2n}(a_1, \dots, a_{2n})$ is **alternating**. Cumulants with an odd number of arguments will always be considered as not alternating.
Example: The cumulant $k_6(a, a^*, a, a^*, a, a^*)$ is alternating, whereas $k_8(a, a^*, a^*, a, a, a^*, a, a^*)$ or $k_5(a, a^*, a, a^*, a)$ are not alternating.

DEFINITION 9.3. A random variable a in a $*$-probability space is called R**-diagonal** if for all $n \in \mathbb{N}$ we have that $k_n(a_1, \dots, a_n) = 0$ whenever the arguments $a_1, \dots, a_n \in \{a, a^*\}$ are not alternating in a and a^*.
If a is R-diagonal we denote the non-vanishing cumulants by $\alpha_n := k_{2n}(a, a^*, a, a^*, \dots, a, a^*)$ and $\beta_n := k_{2n}(a^*, a, a^*, a, \dots, a^*, a)$ $(n \geq 1)$.

The sequences $(\alpha_n)_{n\geq 1}$ and $(\beta_n)_{n\geq 1}$ are called the **determining series** of a.

REMARK 9.4. As in Theorem 8.18, it seems to be the most conceptual point of view to consider the requirement that $a \in \mathcal{A}$ is R-diagonal as a condition on the matrix $A := \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix} \in M_2(\mathcal{A})$. However, this characterization needs the frame of operator-valued freeness and thus we want here only state the result: a is R-diagonal if and only if the matrix A is free from $M_2(\mathbb{C})$ with amalgamation over the scalar diagonal matrices $D := \{\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \mid \alpha, \beta \in \mathbb{C}\}$.

NOTATION 9.5. In the same way as for cumulants and moments we define for the determining series $(\alpha_n)_{n\geq 1}$ of an R-diagonal element (or more general for any sequence of numbers) a function on non-crossing partitions by multiplicative factorization:

$$\alpha_\pi := \prod_{V\in\pi} \alpha_{|V|}. \tag{101}$$

EXAMPLES 9.6. 1) By the first part of Example 6.6, we know that the only non-vanishing cumulants for a circular element are $k_2(c, c^*) = k_2(c^*, c) = 1$. Thus a circular element is R-diagonal with determining series

$$\alpha_n = \beta_n = \begin{cases} 1, & n = 1 \\ 0, & n > 1. \end{cases} \tag{102}$$

2) Let u be a Haar unitary. We calculated its cumulants in part 4 of Example 6.6. In our present language, we showed there that u is R-diagonal with determining series

$$\alpha_n = \beta_n = (-1)^{n-1} C_{n-1}. \tag{103}$$

REMARK 9.7. It is clear that all information on the $*$-distribution of an R-diagonal element a is contained in its determining series. Another useful description of the $*$-distribution of a is given by the distributions of aa^* and a^*a. The next proposition connects these two descriptions of the $*$-distribution of a.

PROPOSITION 9.8. *Let a be an R-diagonal random variable and*

$$\begin{aligned} \alpha_n &: = k_{2n}(a, a^*, a, a^*, \dots, a, a^*), \\ \beta_n &: = k_{2n}(a^*, a, a^*, a, \dots, a^*, a) \end{aligned}$$

the determining series of a.
1) Then we have:

$$k_n(aa^*,\dots,aa^*) = \sum_{\substack{\pi\in NC(n)\\ \pi=\{V_1,\dots,V_r\}}} \alpha_{|V_1|}\beta_{|V_2|}\cdots\beta_{|V_r|} \tag{104}$$

$$k_n(a^*a,\dots,a^*a) = \sum_{\substack{\pi\in NC(n)\\ \pi=\{V_1,\dots,V_r\}}} \beta_{|V_1|}\alpha_{|V_2|}\cdots\beta_{|V_r|} \tag{105}$$

where V_1 denotes that block of $\pi \in NC(n)$ which contains the first element 1. In particular, the $*$-distribution of a is uniquely determined by the distributions of aa^* and of a^*a.
2) In the tracial case (i.e. if $\alpha_n = \beta_n$ for all n) we have

$$k_n(aa^*,\dots,aa^*) = k_n(a^*a,\dots,a^*a) = \sum_{\pi\in NC(n)} \alpha_\pi. \tag{106}$$

PROOF. 1) Applying Theorem 5.2 yields

$$k_n(aa^*,\dots,aa^*) = \sum_{\substack{\pi\in NC(2n)\\ \pi\vee\sigma=1_{2n}}} k_\pi[a,a^*,\dots,a,a^*]$$

with

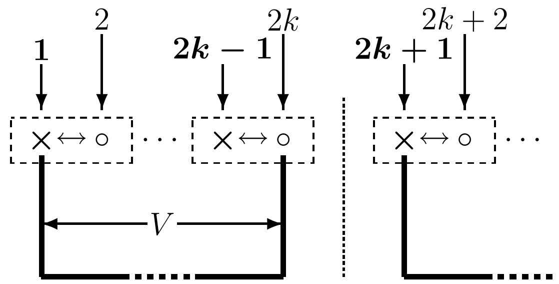
$$\sigma = \{(a,a^*),\dots,(a,a^*)\} \quad \hat{=} \quad \{(\mathbf{1},2),\dots,(\mathbf{2n-1},2n)\}.$$

We claim now the following: The partitions π which fulfill the condition $\pi\vee\sigma = 1_{2n}$ are exactly those which have the following properties: the block of π which contains the element $\mathbf{1}$ contains also the element $2n$, and, for each $k = 1,\dots,n-1$, the block of π which contains the element $2k$ contains also the element $\mathbf{2k+1}$.

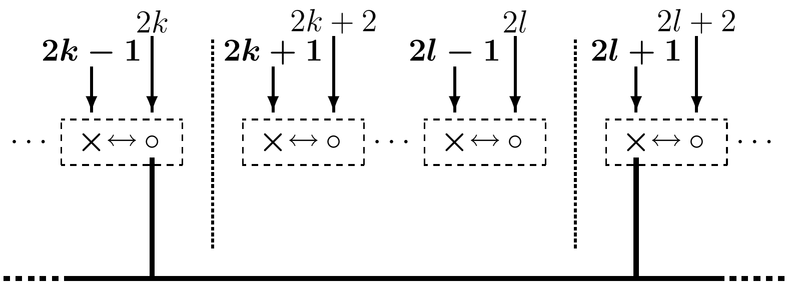
Since the set of those $\pi \in NC(2n)$ fulfilling the claimed condition is in canonical bijection with $NC(n)$ and since $k_\pi[a,a^*,\dots,a,a^*]$ goes under this bijection to the product appearing in our assertion, this gives directly the assertion.

So it remains to prove the claim. It is clear that a partition which has the claimed property does also fulfill $\pi\vee\sigma = 1_{2n}$. So we only have to prove the other direction.

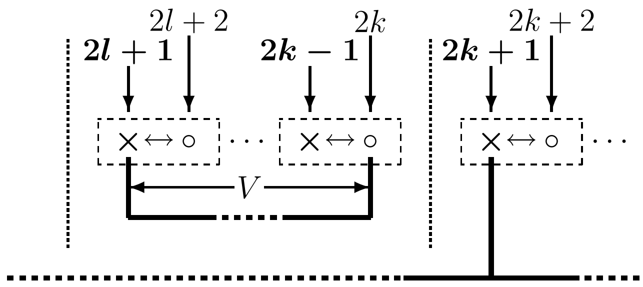
Let V be the block of π which contains the element $\mathbf{1}$. Since a is R-diagonal the last element of this block has to be an a^*, i.e., an even number, let's say $2k$. If this would not be $2n$ then this block V would in $\pi\vee\sigma$ not be connected to the block containing $\mathbf{2k+1}$, thus $\pi\vee\sigma$ would not give 1_{2n}. Hence $\pi\vee\sigma = 1_{2n}$ implies that the block containing the first element $\mathbf{1}$ contains also the last element $2n$.



Now fix a $k = 1, \dots, n-1$ and let V be the block of π containing the element $2k$. Assume that V does not contain the element $\mathbf{2k+1}$. Then there are two possibilities: Either $2k$ is not the last element in V, i.e. there exists a next element in V, which is necessarily of the form $\mathbf{2l+1}$ with $l > k$ ...



... or $2k$ is the last element in V. In this case the first element of V is of the form $\mathbf{2l+1}$ with $0 \le l \le k-1$.



In both cases the block V gets not connected with $\mathbf{2k+1}$ in $\pi \vee \sigma$, thus this cannot give 1_{2n}. Hence the condition $\pi \vee \sigma = 1_{2n}$ forces $2k$ and $\mathbf{2k+1}$ to lie in the same block. This proves our claim and hence the assertion.

2) This is a direct consequence from the first part, if the α's and β's are the same. $\square$

REMARK 9.9. It is also true that, for a R-diagonal, aa^* and a^*a are free. We could also prove this directly in the same spirit as above, but we will defer it to later, because it will follow quite directly from a characterization of R-diagonal elements as those random variables whose $*$-distribution is invariant under multiplication with a free Haar unitary.

0.1. Proposition. Let a and x be elements in a $*$-probability space $(\mathcal{A}, \varphi)$ with a being R-diagonal and such that $\{a, a^*\}$ and $\{x, x^*\}$ are free. Then ax is R-diagonal.

PROOF. We examine a cumulant $k_r(a_1a_2, \dots, a_{2r-1}a_{2r})$ with $a_{2i-1}a_{2i} \in \{ax, x^*a^*\}$ for $i \in \{1, \dots, r\}$.
According to the definition of R-diagonality we have to show that this cumulant vanishes in the following two cases:

(1°) r is odd.
(2°) There exists at least one s $(1 \le s \le r-1)$ such that $a_{2s-1}a_{2s} = a_{2s+1}a_{2s+2}$.

By Theorem 5.2, we have

$$k_r(a_1a_2, \dots, a_{2r-1}a_{2r}) = \sum_{\substack{\pi \in NC(2r) \\ \pi \vee \sigma = 1_{2r}}} k_\pi[a_1, a_2, \dots, a_{2r-1}, a_{2r}] \,,$$

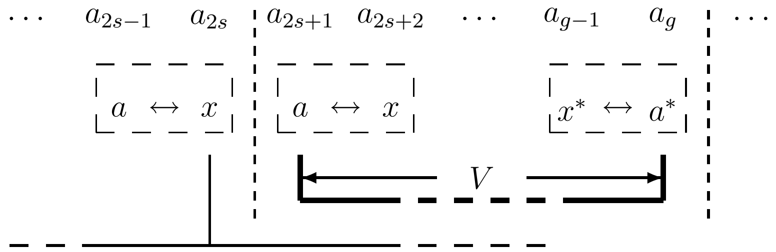
where $\sigma = \{(a_1, a_2), \dots, (a_{2r-1}, a_{2r})\}$.

The fact that a and x are $*$-free implies, by the vanishing of mixed cumulants, that only such partitions $\pi \in NC(2r)$ contribute to the sum each of whose blocks contains elements only from $\{a, a^*\}$ or only from $\{x, x^*\}$.

Case (1°): As there is at least one block of π containing a different number of elements a and a^*, k_π vanishes always. So there are no partitions π contributing to the above sum, which consequently vanishes.

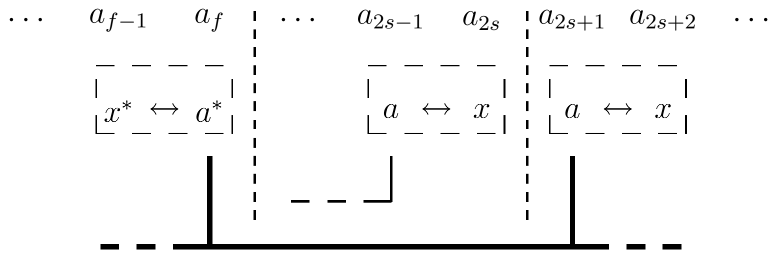
Case (2°): We assume that there exists an $s \in \{1, \dots, r-1\}$ such that $a_{2s-1}a_{2s} = a_{2s+1}a_{2s+2}$. Since with a also a^* is R-diagonal, it suffices to consider the case where $a_{2s-1}a_{2s} = a_{2s+1}a_{2s+2} = ax$, i.e., $a_{2s-1} = a_{2s+1} = a$ and $a_{2s} = a_{2s+2} = x$.
Let V be the block containing a_{2s+1}. We have to examine two situations:

A. On the one hand, it might happen that a_{2s+1} is the first element in the block V. This can be sketched in the following way:



In this case the block V is not connected with a_{2s} in $\pi \vee \sigma$, thus the latter cannot be equal to 1_{2n}.

B. On the other hand, it can happen that a_{2s+1} is not the first element of V. Because a is R-diagonal, the preceding element must be an a^*.



But then V will again not be connected to a_{2s} in $\pi \vee \sigma$. Thus again $\pi \vee \sigma$ cannot be equal to 1_{2n}.

As in both cases we do not find any partition contributing to the investigated sum, this has to vanish. $\square$

THEOREM 9.10. *Let x be an element in a $*$-probability space $(\mathcal{A}, \varphi)$. Furthermore, let u be a Haar unitary in $(\mathcal{A}, \varphi)$ such that $\{u, u^*\}$ and $\{x, x^*\}$ are free. Then x is R-diagonal if and only if (x, x^*) has the same joint distribution as (ux, x^*u^*):*

$$x \ R\text{-diagonal} \quad \Longleftrightarrow \quad \mu_{x,x^*} = \mu_{ux,x^*u^*} \ .$$

PROOF. $\Longrightarrow$: We assume that x is R-diagonal and, by Prop. 9.10, we know that ux is R-diagonal, too. So to see that both have the same $*$-distribution it suffices to see that the respective determining series agree. By Prop. 9.8, this is the case if the distribution of xx^* agrees with the distribution of $ux(ux)^*$ and if the distribution of x^*x agrees with the distribution of $(ux)^*ux$. For the latter case this is directly clear, whereas for the first case one only has to observe that uau^* has the same distribution as a if u is $*$-free from a. (Note that in the non-tracial

case one really needs the freeness assumption in order to get the first u cancel the last u^* via $\varphi((uau^*)^n) = \varphi(ua^n u^*) = \varphi(uu^*)\varphi(a^n) = \varphi(a^n)$.)
$\Longleftarrow$: We assume that the $*$-distribution of x is the same as the $*$-distribution of ux. As, by Prop. 9.10, ux is R-diagonal, x is R-diagonal, too. □

Corollary 9.11. *Let a be R-diagonal. Then aa^* and a^*a are free.*

Proof. Let u be a Haar unitary which is $*$-free from a. Since a has the same $*$-distribution as ua it suffices to prove the statement for ua. But there it just says that uaa^*u^* and $a^*u^*ua = a^*a$ are free, which is clear by the very definition of freeness. □

Corollary 9.12. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with φ a faithful trace and let $a \in \mathcal{A}$ be such that* $\ker(a) = \{0\}$. *Then the following two statements are equivalent.*

1. *a is R-diagonal.*
2. *a has a polar decomposition of the form $a = ub$, where u is a Haar unitary and u, b are $*$-free.*

Proof. The implication 2) $\Longrightarrow$ 1) is clear, by Prop. 9.10, and the fact that a Haar unitary u is R-diagonal.
1) $\Longrightarrow$ 2): Consider $\tilde{u}, \tilde{b}$ in some W^*-probability space $(\tilde{\mathcal{A}}, \tilde{\varphi})$, such that $\tilde{u}$ is Haar unitary, $\tilde{u}$ and $\tilde{b}$ are $*$-free and furthermore $\tilde{b} \geq 0$ has the same distribution as $|a| = \sqrt{a^*a}$ (which is, by traciality, the same as the distribution of $\sqrt{aa^*}$.) But then it follows that $\tilde{a} := \tilde{u}\tilde{b}$ is R-diagonal and the distribution of $\tilde{a}\tilde{a}^* = \tilde{u}\tilde{b}\tilde{b}\tilde{u}^*$ is the same as the distribution of aa^* and the distribution of $\tilde{a}^*\tilde{a} = \tilde{b}\tilde{u}^*\tilde{u}\tilde{b} = \tilde{b}\tilde{b}$ is the same as the distribution of a^*a. Hence the $*$-distributions of a and $\tilde{a}$ coincide. But this means that the von Neumann algebra generated by a is isomorphic to the von Neumann algebra generated by $\tilde{a}$ via the mapping $a \mapsto \tilde{a}$. Since the polar decomposition takes places inside the von Neumann algebras, the polar decompostion of a is mapped to the polar decomposition of $\tilde{a}$ under this isomorphism. But the polar decomposition of $\tilde{a}$ is by construction just $\tilde{a} = \tilde{u}\tilde{b}$ and thus has the stated properties. Hence these properties (which rely only on the $*$-distributions of the elements involved in the polar decomposition) are also true for the elements in the polar decomposition of a. □

Prop. 9.10 implies in particular that the product of two free R-diagonal elements is R-diagonal again. This raises the question how the alternating cumulants of the product are given in terms of the alternating cumulants of the factors. This is answered in the next proposition.

PROPOSITION 9.13. *Let a and b be R-diagonal random variables such that $\{a, a^*\}$ is free from $\{b, b^*\}$. Furthermore, put*

$$\begin{aligned} \alpha_n &:= k_{2n}(a, a^*, a, a^*, \dots, a, a^*) , \\ \beta_n &:= k_{2n}(a^*, a, a^*, a, \dots, a^*, a) , \\ \gamma_n &:= k_{2n}(b, b^*, b, b^*, \dots, b, b^*). \end{aligned}$$

Then ab is R-diagonal and the alternating cumulants of ab are given by

$$\begin{aligned} (107) \quad & k_{2n}(ab, b^*a^*, \dots, ab, b^*a^*) \\ &= \sum_{\substack{\pi=\pi_a\cup\pi_b\in NC(2n)\\ \pi_a=\{V_1,\dots,V_k\}\in NC(1,3,\dots,2n-1)\\ \pi_b=\{V_1',\dots,V_l'\}\in NC(2,4,\dots,2n)}} \alpha_{|V_1|}\beta_{|V_2|}\cdots\beta_{|V_k|}\gamma_{|V_1'|}\cdots\gamma_{|V_l'|} , \end{aligned}$$

where V_1 is that block of π which contains the first element 1.

REMARK 9.14. Note that in the tracial case the statement reduces to

$$(108) \qquad k_{2n}(ab, b^*a^*, \dots, ab, b^*a^*) = \sum_{\substack{\pi_a,\pi_b\in NC(n)\\ \pi_b\leq K(\pi_a)}} \alpha_{\pi_a}\beta_{\pi_b}.$$

PROOF. R-diagonality of ab is clear by Prop. 9.10. So we only have to prove the formula for the alternating cumulants.

By Theorem 5.2, we get

$$k_{2n}(ab, b^*a^*, \dots , ab, b^*a^*) = \sum_{\substack{\pi\in NC(4n)\\ \pi\vee\sigma=1_{4n}}} k_\pi[a, b, b^*, a^*, \dots , a, b, b^*, a^*] ,$$

where $\sigma = \{(a, b), (b^*, a^*), \dots , (a, b), (b^*, a^*)\}$. Since $\{a, a^*\}$ and $\{b, b^*\}$ are assumed to be free, we also know, by the vanishing of mixed cumulants that for a contributing partition π each block has to contain components only from $\{a, a^*\}$ or only from $\{b, b^*\}$.

As in the proof of Prop. 9.8 one can show that the requirement $\pi \vee \sigma = 1_{4n}$ is equivalent to the following properties of π: The block containing 1 must also contain $4n$ and, for each $k = 1, ..., 2n - 1$, the block containing $2k$ must also contain $2k + 1$. (This couples always b with b^* and a^* with a, so it is compatible with the $*$-freeness between a and b.) The set of partitions in $NC(4n)$ fulfilling these properties is in canonical bijection with $NC(2n)$. Furthermore we have to take care of the fact that each block of $\pi \in NC(4n)$ contains either only elements from $\{a, a^*\}$ or only elements from $\{b, b^*\}$. For the image of π in $NC(2n)$ this means that it splits into blocks living on the odd numbers and blocks living on the even numbers. Furthermore, under these identifications the quantity $k_\pi[a, b, b^*, a^*, \dots, a, b, b^*, a^*]$ goes over to the expression as appearing in our assertion. $\square$

An important way how R-diagonal elements can arise is as the product of two free even elements.

NOTATIONS 9.15. We call an element x in a $*$-probability space $(\mathcal{A}, \varphi)$ **even** if it is selfadjoint and if all its odd moments vanish, i.e. if $\varphi(x^{2k+1}) = 0$ for all $k \geq 0$.
In analogy with R-diagonal elements we will call $(\alpha_n)_{n\geq 1}$ with $\alpha_n := k^x_{2n}$ the **determining series** of an even variable x.

REMARKS 9.16. 1) Note that the vanishing of all odd moments is equivalent to the vanishing of all odd cumulants.
2) Exactly the same proof as for Prop. 9.8 shows that for an even variable x we have the following relation between its determining series α_n and the cumulants of x^2:

$$k_n(x^2, \dots, x^2) = \sum_{\pi \in NC(n)} \alpha_\pi. \tag{109}$$

THEOREM 9.17. *Let x, y be two even random variables. If x and y are free then xy is R-diagonal.*

PROOF. Put $a := xy$. We have to see that non alternating cumulants in $a = xy$ and $a^* = yx$ vanish. Since it is clear that cumulants of odd length in xy and yx vanish always it remains to check the vanishing of cumulants of the form $k_n(\dots, xy, xy, \dots)$. (Because of the symmetry of our assumptions in x and y this will also yield the case $k_n(\dots, yx, yx, \dots)$.) By Theorem 5.2, we can write this cumulant as

$$k_n(\dots, xy, xy, \dots) = \sum_{\substack{\pi \in NC(2n) \\ \pi \vee \sigma = 1_{2n}}} k_\pi[\dots, xy, xy, \dots],$$

where $\sigma = \{(1, 2), (3, 4), \dots, (2n-1, 2n)\}$. In order to be able to distinguish y appearing at different possitions we will label them by indices (i.e. $y_i = y$ for all appearing i). Thus we have to look at $k_\pi[\dots, xy_1, xy_2, \dots]$ for $\pi \in NC(2n)$. Because of the freeness of x and y, π only gives a contribution if it does not couple x with y. Furthermore all blocks of π have to be of even length, by our assumption that x and y are even. Let now V be that block of π which contains y_1. Then there are two possibilities.

1. Either y_1 is not the last element in V. Let y_3 be the next element in V, then we must have a situation like this

$$\dots \quad , \quad x \quad y_1 \quad , \quad x \quad y_2 \quad , \quad \dots \quad , \quad y_3 \quad x \quad , \quad \dots$$

Note that y_3 has to belong to a product yx as indicated, because both the number of x and the number of y lying between y_1 and y_3 have to be even. But then everything lying between y_1 and y_3 is not connected to the rest (neither by π nor by σ), and thus the condition $\pi \vee \sigma = 1_{2n}$ cannot be fulfilled.

2. Or y_1 is the last element in the block V. Let y_0 be the first element in V. Then we have a situation as follows

$$\ldots, \quad y_0 \quad x \quad , \quad \ldots \quad , \quad x \quad y_1 \quad , \quad x \quad y_2 \quad , \quad \ldots$$

Again we have that y_0 must come from a product yx, because the number of x and the number of y lying between y_0 and y_1 have both to be even (although now some of the y from that interval might be connected to V, too, but that has also to be an even number). But then everything lying between y_0 and y_1 is separated from the rest and we cannot fulfill the condition $\pi \vee \sigma = 1_{2n}$.

Thus in any case there is no π which fulfills $\pi \vee \sigma = 1_{2n}$ and has also $k_\pi[\ldots, xy_1, xy_2, \ldots]$ different from zero. Hence $k(\ldots, xy, xy, \ldots)$ vanishes. $\square$

COROLLARY 9.18. *Let x and y be two even elements which are free. Consider the free anti-commutator $c := xy + yx$. Then the cumulants of c are given in terms of the determining series $\alpha_n^x = k_{2n}^x$ of x and $\alpha_n^y := k_{2n}^y$ of y by*

$$k_{2n}^c = 2 \sum_{\substack{\pi_1, \pi_2 \in NC(n) \\ \pi_2 \leq K(\pi_1)}} \alpha_{\pi_1}^x \alpha_{\pi_2}^y. \tag{110}$$

PROOF. Since xy is R-diagonal it is clear that cumulants in c of odd length vanish. Now note that φ restricted to the unital $*$-algebra generated by x and y is a trace (because the the unital $*$-algebra generated by x and the unital $*$-algebra generated by y are commutative and the free product preserves traciality), so that the R-diagonality of xy gives

$$\begin{aligned} k_{2n}^c &= k_{2n}(xy + yx, \ldots, xy + yx) \\ &= k_{2n}(xy, yx, \ldots, xy, yx) + k_{2n}(yx, xy, \ldots, yx, xy) \\ &= 2k_{2n}(xy, yx, \ldots, xy, yx). \end{aligned}$$

Furthermore, the same kind of argument as in the proof of Prop. 9.14 allows to write this in the way as claimed in our assertion. □

REMARKS 9.19. 1) The problem of the anti-commutator for the general case is still open. For the problem of the commutator $i(xy-yx)$ one should remark that in the even case this has the same distribution as the anti-commutator and that the general case can, due to cancellations, be reduced to the even case.
2) As it becomes clear from our Prop. 9.14 and our result about the free anti-commutator, the combinatorial formulas are getting more and more involved and one might start to wonder how much insight such formulas provide. What is really needed for presenting these solutions in a useful way is a machinery which allows to formalize the proofs and manipulate the results in an algebraic way without having to spend too much considerations on the actual kind of summations. Such a machinery will be presented later (in the course of A. Nica), and it will be only with the help of that apparatus that one can really formulate the results in a form also suitable for concrete calculations.

REMARK 9.20. According to Prop. 9.10 multiplication preserves R-diagonality if the factors are free. Haagerup and Larsen showed that, in the tracial case, the same statement is also true for the other extreme relation between the factors, namely if they are the same – i.e., powers of R-diagonal elements are also R-diagonal. The proof of Haagerup and Larsen relied on special realizations of R-diagonal elements. Here we will give a combinatorial proof of that statement. In particular, our proof will – in comparison with the proof of Prop. 9.10 – also illuminate the relation between the statements "$a_1, \dots, a_r$ R-diagonal and free implies $a_1 \cdots a_r$ R-diagonal" and "a R-diagonal implies a^r R-diagonal". Furthermore, our proof extends without problems to the non-tracial situation.

PROPOSITION 9.21. *Let a be an R-diagonal element and let r be a positive integer. Then a^r is R-diagonal, too.*

PROOF. For notational convenience we deal with the case $r = 3$. General r can be treated analogously.
The cumulants which we must have a look at are $k_n(b_1, \dots, b_n)$ with arguments b_i from $\{a^3, (a^3)^*\}$ $(i = 1, \dots, n)$. We write $b_i = b_{i,1}b_{i,2}b_{i,3}$ with $b_{i,1} = b_{i,2} = b_{i,3} \in \{a, a^*\}$. According to the definition of R-diagonality we have to show that for any $n \geq 1$ the cumulant $k_n(b_{1,1}b_{1,2}b_{1,3}, \dots, b_{n,1}b_{n,2}b_{n,3})$ vanishes if (at least) one of the following things happens:

(1°) There exists an $s \in \{1, \dots, n-1\}$ with $b_s = b_{s+1}$.

(2°) n is odd.

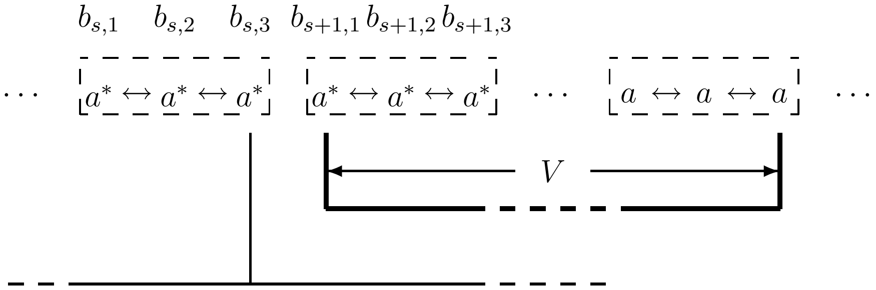
Theorem 5.2 yields

$$k_n(b_{1,1}b_{1,2}b_{1,3},\dots,b_{n,1}b_{n,2}b_{n,3}) \\ = \sum_{\substack{\pi\in NC(3n)\\ \pi\vee\sigma=1_{3n}}} k_\pi[b_{1,1},b_{1,2},b_{1,3},\dots,b_{n,1},b_{n,2},b_{n,3}]\,,$$

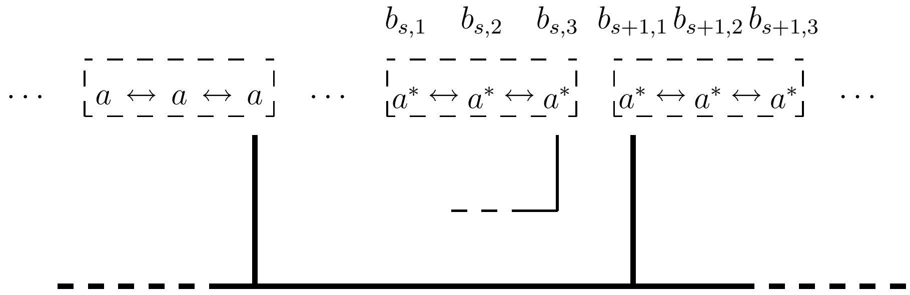
where $\sigma := \{(b_{1,1},b_{1,2},b_{1,3}),\dots,(b_{n,1},b_{n,2},b_{n,3})\}$. The R-diagonality of a implies that a partition π gives a non-vanishing contribution to the sum only if its blocks link the arguments alternatingly in a and a^*.
Case (1°): Without loss of generality, we consider the cumulant $k_n(\dots,b_s,b_{s+1},\dots)$ with $b_s = b_{s+1} = (a^3)^*$ for some s with $1\le s\le n-1$. This means that we have to look at $k_n(\dots,a^*a^*a^*,a^*a^*a^*,\dots)$.
Theorem 5.2 yields in this case

$$k_n(\dots,a^*a^*a^*,a^*a^*a^*,\dots) = \sum_{\substack{\pi\in NC(3n)\\ \pi\vee\sigma=1_{3n}}} k_\pi[\dots,a^*,a^*,a^*,a^*,a^*,a^*,\dots],$$

where $\sigma := \{\dots,(a^*,a^*,a^*),(a^*,a^*,a^*),\dots\}$. In order to find out which partitions $\pi\in NC(3n)$ contribute to the sum we look at the structure of the block containing the element $b_{s+1,1}=a^*$; in the following we will call this block V.
There are two situations which can occur. The first possibility is that $b_{s+1,1}$ is the first component of V; in this case the last component of V must be an a and, since each block has to contain the same number of a and a^*, this a has to be the third a of an argument a^3. But then the block V gets in $\pi\vee\sigma$ not connected with the block containing $b_{s,3}$ and hence the requirement $\pi\vee\sigma=1_{3n}$ cannot be fulfilled in such a situation.



The second situation that might happen is that $b_{s+1,1}$ is not the first component of V. Then the preceding element in this block must be an a and again it must be the third a of an argument a^3. But then the block containing $b_{s,3}$ is again not connected with V in $\pi \vee \sigma$. This possibility can be illustrated as follows:



Thus, in any case there exists no π which fulfills the requirement $\pi \vee \sigma = 1_{3n}$ and hence $k_n(\dots, a^*a^*a^*, a^*a^*a^*, \dots)$ vanishes in this case.

Case (2°): In the case n odd, the cumulant $k_\pi[b_{1,1}, b_{1,2}, b_{1,3}, \dots, b_{n,1}, b_{n,2}, b_{n,3}]$ has a different number of a and a^* as arguments and hence at least one of the blocks of π cannot be alternating in a and a^*. Thus k_π vanishes by the R-diagonality of a.

As in both cases we do not find any partition giving a non-vanishing contribution, the sum vanishes and so do the cumulants $k_n(b_1, \dots, b_n)$. $\square$

REMARK 9.22. We are now left with the problem of describing the alternating cumulants of a^r in terms of the alternating cumulants of a. We will provide the solution to this question by showing that the similarity between $a_1 \cdots a_r$ and a^r goes even further as in the Remark 9.21. Namely, we will show that a^r has the same $*$-distribution as $a_1 \cdots a_r$ if $a_1, \dots, a_r$ are $*$-free and all a_i $(i = 1, \dots, r)$ have the same $*$-distribution as a. The distribution of a^r can then be calculated by an iteration of Prop. 9.14. In the case of a trace this reduces to a result of Haagerup and Larsen. The specical case of powers of a circular element was treated by Oravecz.

PROPOSITION 9.23. *Let a be an R-diagonal element and r a positive integer. Then the $*$-distribution of a^r is the same as the $*$-distribution of $a_1 \cdots a_r$ where each a_i $(i = 1, \dots, r)$ has the same $*$-distribution as a and where $a_1, \dots, a_r$ are $*$-free.*

PROOF. Since we know that both a^r and $a_1 \cdots a_r$ are R-diagonal we only have to see that the respective alternating cumulants coincide.

By Theorem 5.2, we have

$$k_{2n}(a^r, a^{*r}, \dots, a^r, a^{*r}) = \sum_{\substack{\pi \in NC(2nr) \\ \pi \vee \sigma = 1_{2nr}}} k_\pi[a, \dots, a, a^*, \dots, a^*, \dots, a, \dots, a, a^*, \dots, a^*]$$

and

$$k_{2n}(a_1 \cdots a_r, a_r^* \cdots a_1^*, \dots, a_1 \cdots a_r, a_r^* \cdots a_1^*) = \sum_{\substack{\pi \in NC(2nr) \\ \pi \vee \sigma = 1_{2nr}}} k_\pi[a_1, \dots, a_r, a_r^*, \dots, a_1^*, \dots, a_1, \dots, a_r, a_r^*, \dots, a_1^*],$$

where in both cases $\sigma = \{(1, \dots, r), (r+1, \dots, 2r), \dots, (2(n-1)r + 1, \dots, 2nr)\}$. The only difference between both cases is that in the second case we also have to take care of the freeness between the a_i which implies that only such π contribute which do not connect different a_i. But the R-diagonality of a implies that also in the first case only such π give a non-vanishing contribution, i.e. the freeness in the second case does not really give an extra condition. Thus both formulas give the same and the two distributions coincide. $\square$

CHAPTER 10

Free Fisher information

REMARKS 10.1. 1) In classical probability theory there exist two important concepts which measure the 'amount of information' of a given distribution. These are the Fisher information and the entropy (the latter measures the absence of information). There exist various relations between these quantities and they form a cornerstorne of classical probability theory and statistics. Voiculescu introduced free probability analogues of these quantities, called free Fisher information and free entropy, denoted by Φ and χ, respectively. However, the present situation with these quantities is a bit confusing. In particular, there exist two different approaches, each of them yielding a notion of entropy and Fisher information. One hopes that finally one will be able to prove that both approaches give the same, but at the moment this is not clear. Thus for the time being we have to distinguish the entropy χ and the free Fisher information Φ coming from the first approach (via micro-states) and the free entropy χ^* and the free Fisher information Φ^* coming from the second approach (via a non-commutative Hilbert transform). We will in this section only deal with the second approach, which fits quite nicely with our combinatorial theory of freeness. In this approach the Fisher information is the basic quantity (in terms of which the free entropy χ^* is defined), so we will restrict our attention to Φ^*.
2) The concepts of information and entropy are only useful when we consider states (so that we can use the positivity of φ to get estimates for the information or entropy). Thus in this section we will always work in the framework of a von Neumann probability space. Furthermore, it is crucial that we work with a trace. The extension of the present theory to non-tracial situations is unclear.
3) The basic objects in our approach to free Fisher information, the conjugate variables, do in general not live in the algebra $\mathcal{A}$, but in the L^2-space $L^2(\mathcal{A}, \varphi)$ associated to $\mathcal{A}$ with respect to the given trace φ. We will also consider cumulants, where one argument comes from $L^2(\mathcal{A})$ and all other arguments are from $\mathcal{A}$ itself. Since this corresponds to moments where we multiply one element from $L^2(\mathcal{A})$ with elements from $\mathcal{A}$ this is well-defined within $L^2(\mathcal{A})$. (Of course the case where

more than one argument comes from $L^2(\mathcal{A})$ would be problematic.) Thus, by continuity, such cumulants are well-defined and we can work with them in the same way as for the case where all arguments are from $\mathcal{A}$.

NOTATIONS 10.2. Let $(\mathcal{A},\varphi)$ be a W^*-probability space with a faithful trace φ.
1) We denote by $L^2(\mathcal{A},\varphi)$ (or for short $L^2(\mathcal{A})$ if the state is clear) the completion of $\mathcal{A}$ with respect to the norm $\|a\| := \varphi(a^*a)^{1/2}$.
2) For a subset $\mathcal{X}\subset\mathcal{A}$ we denote by

$$L^2(\mathcal{X},\varphi) = L^2(\mathcal{X}) := \overline{\operatorname{alg}(\mathcal{X},\mathcal{X}^*)}^{\|\cdot\|}$$

the closure in $L^2(\mathcal{A})$ of the unital $*$-algebra generated by the set $\mathcal{X}$.
3) A **self-adjoint family of random variables** is a family of random variables $F=\{a_i\}_{i\in I}$ (for some index set I) with the property that with $a_i\in F$ also $a_i^*\in F$ (thus $a_i^* = a_j$ for some $j\in I$).

DEFINITIONS 10.3. Let $(\mathcal{A},\varphi)$ be a W^*-probability space with a faithful trace φ. Let I be a finite index set and consider a self-adjoint family of random variables $\{a_i\}_{i\in I}\subset\mathcal{A}$.
1) We say that a family $\{\xi_i\}_{i\in I}$ of vectors in $L^2(\mathcal{A},\varphi)$ **fulfills the conjugate relations** for $\{a_i\}_{i\in I}$, if

(111)
$$\varphi(\xi_i a_{i(1)}\cdots a_{i(n)}) = \sum_{k=1}^{n} \delta_{ii(k)}\varphi(a_{i(1)}\cdots a_{i(k-1)})\cdot\varphi(a_{i(k+1)}\cdots a_{i(n)})$$

for all $n\geq 0$ and all $i,i(1),\dots,i(n)\in I$.
2) We say that a family $\{\xi_i\}_{i\in I}$ of vectors in $L^2(\mathcal{A},\varphi)$ is a **conjugate system** for $\{a_i\}_{i\in I}$, if it fulfills the conjugate relations (1) and if in addition we have that

$$\xi_i\in L^2(\{a_j\}_{j\in I},\varphi)\qquad\text{for all } i\in I. \tag{112}$$

REMARKS 10.4. 1) A conjugate system for $\{a_i\}_{i\in I}$ is unique, if it exists.
2) If there exists a family $\{\xi_i\}_{i\in I}$ in $L^2(\mathcal{A},\varphi)$ which fulfills the conjugate relations for $\{a_i\}_{i\in I}$, then there exists a conjugate system for $\{a_i\}_{i\in I}$. Namely, let P be the orthogonal projection from $L^2(\mathcal{A},\varphi)$ onto $L^2(\{a_j\}_{j\in I},\varphi)$. Then $\{P\xi_i\}_{i\in I}$ is the conjugate system for $\{a_i\}_{i\in I}$.

PROPOSITION 10.5. *Let $(\mathcal{A},\varphi)$ be a W^*-probability space with a faithful trace φ. Let I be a finite index set and consider a self-adjoint family of random variables $\{a_i\}_{i\in I}\subset\mathcal{A}$. A family $\{\xi_i\}_{i\in I}$ of vectors* *in $L^2(\mathcal{A},\varphi)$ fulfills the conjugate relations for $\{a_i\}_{i\in I}$, if and only if we have for all $n\geq 0$ and all $i, i(1),\dots,i(n)\in I$ that*

$$k_{n+1}(\xi_i, a_{i(1)},\dots,a_{i(n)}) = \begin{cases}\delta_{ii(1)}, & n=1\\ 0, & n\neq 1.\end{cases} \tag{113}$$

Proof. We only show that (1) implies (2). The other direction is similar.
We do this by induction on n. For $n=0$ and $n=1$, this is clear:

$$k_1(\xi_i)=\varphi(\xi_i)=0$$

$$k_2(\xi_i,a_{i(1)}) = \varphi(\xi_i a_{i(1)}) - \varphi(\xi_i)\varphi(a_{i(1)}) = \delta_{ii(1)}.$$

Consider now $n\geq 2$. Then we have

$$\begin{aligned}\varphi(\xi_i a_{i(1)}\cdots a_{i(n)}) &= \sum_{\pi\in NC(n+1)} k_\pi[\xi_i, a_{i(1)},\dots,a_{i(n)}]\\ &= k_n(\xi_i,a_{i(1)},\dots,a_{i(n)}) + \sum_{\pi\neq 1_{n+1}} k_\pi[\xi_i,a_{i(1)},\dots,a_{i(n)}].\end{aligned}$$

By induction assumption, in the second term only such partitions $\pi\in NC(0,1,\dots,n)$ contribute which couple ξ_i with exactly one $a_{i(k)}$, i.e. which are of the form $\pi=(0,k)\cup\pi_1\cup\pi_2$, for some $1\leq k\leq n$ and where $\pi_1\in NC(1,\dots,k-1)$ and $\pi_2\in NC(k+1,\dots,n)$. Thus we get

$$\begin{aligned}\varphi(\xi_i a_{i(1)}\cdots a_{i(n)}) &= k_n(\xi_i,a_{i(1)},\dots,a_{i(n)})\\ &\quad+\sum_{k=1}^n \delta_{ii(k)}\big(\sum_{\pi_1\in NC(1,\dots,k-1)} k_{\pi_1}[a_{i(1)},\dots,a_{i(k-1)}]\big)\\ &\qquad\qquad\cdot\big(\sum_{\pi_2\in NC(k+1,\dots,n)} k_{\pi_2}[a_{i(k+1)},\dots,a_{i(n)}]\big)\\ &= k_n(\xi_i,a_{i(1)},\dots,a_{i(n)}) + \sum_{k=1}^n \delta_{ii(k)}\varphi(a_{i(1)}\cdots a_{i(k-1)})\cdot\varphi(a_{i(k+1)}\cdots a_{i(n)})\\ &= k_n(\xi_i,a_{i(1)},\dots,a_{i(n)}) + \varphi(\xi_i a_{i(1)}\cdots a_{i(n)}).\end{aligned}$$

This gives the assertion. □

Example 10.6. Let $\{s_i\}_{i\in I}$ be a semi-circular family, i.e. s_i $(i\in I)$ are free and each of them is a semi-circular of variance 1. Then the conjugate system $\{\xi_i\}_{i\in I}$ for $\{s_i\}_{i\in I}$ is given by $\xi_i=s_i$ for all $i\in I$. This follows directly from 4.8, which states that

$$k_{n+1}(s_i,s_{i(1)},\dots,s_{i(n)}) = \delta_{n1}\delta_{ii(1)}.$$

DEFINITION 10.7. Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ. Let I be a finite index set and consider a self-adjoint family of random variables $\{a_i\}_{i\in I} \subset \mathcal{A}$. If $\{a_i\}_{i\in I}$ has a conjugate system $\{\xi_i\}_{i\in I}$, then the **free Fisher information** of $\{a_i\}_{i\in I}$ is defined as

$$\Phi^*(\{a_i\}_{i\in I}) := \sum_{i\in I} \|\xi_i\|^2. \tag{114}$$

If $\{a_i\}_{i\in I}$ has no conjugate system, then we put $\Phi^*(\{a_i\}_{i\in I}) := \infty$.

REMARKS 10.8. 1) Note that by considering real and imaginary parts of our operators we could reduce the above frame to the case where all appearing random variables are self-adjoint. E.g. consider the case $\{x = x^*, a, a^*\}$. Then it is easy to see that

$$\Phi^*(x, a, a^*) = \Phi^*\big(x, \frac{a + a^*}{\sqrt{2}}, \frac{a - a^*}{\sqrt{2}i}\big). \tag{115}$$

2) Since an orthogonal projection in a Hilbert space does not increase the length of the vectors, we get from Remark 10.4 the following main tool for deriving estimates on the free Fisher information: If $\{\xi_i\}_{i\in I}$ fulfills the conjugate relations for $\{a_i\}_{i\in I}$, then we have that $\Phi^*(\{a_i\}_{i\in I}) \le \sum_{i\in I} \|\xi_i\|^2$. Note also that we have equality if and only if $\xi_i \in L^2(\{a_j\}_{j\in I}, \varphi)$ for all $i \in I$.
3) In classical probability theory there exists a kind of meta-mathematical principle, the "maximum entropy principle". Consider a classical situation of which only some partial knowledge is available. Then this principle says that a generic description in such a case is given by a probability distribution which is compatible with the partial knowledge and whose classical entropy is maximal among all such distributions. In the same spirit, one has the following non-commutative variant of this: A generic description of a non-commutative situation subject to given constraints is given by a non-commutative distribution which respects the given constraints and whose free entropy is maximal (or whose free Fisher information is minimal) among all such distributions.
4) Thus it is natural to consider minimization problems for the free Fisher information and ask in particular in which case the minimal value is really attained. We will consider four problems of this kind in the following. The first two are due to Voiculescu; in particular the second contains the main idea how to use equality of Fisher informations to deduce something about the involved distributions (this is one of the key observations of Part VI of the series of Voiculescu on the

analogues of entropy and Fisher's information measure in free probability theory). The other two problems are due to Nica, Shlyakhtenko, and Speicher and are connected with R-diagonal distributions and the distributions coming from compression by a free family of matrix units.

PROBLEM 10.9. Let $r > 0$ be given. Minimize $\Phi^*(a, a^*)$ under the constraint that $\varphi(a^*a) = r$.

PROPOSITION 10.10. **(free Cramer-Rao inequality).** *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ. Let $a \in \mathcal{A}$ be a random variable.*
1) Then we have

$$\Phi^*(a, a^*) \geq \frac{2}{\varphi(a^*a)}. \tag{116}$$

2) We have equality in Equation (6) if and only if a is of the form $a = \lambda c$, where $\lambda > 0$ and c is a circular element.

PROOF. 1) If no conjugate system for $\{a, a^*\}$ exists, then the left hand side is infinite, so the statement is trivial. So let $\{\xi, \xi^*\}$ be a conjugate system for $\{a, a^*\}$. Then we have

$$\begin{aligned} 2 &= \varphi(\xi a) + \varphi(\xi^* a^*) \\ &= \langle \xi^*, a \rangle + \langle \xi, a^* \rangle \\ &\leq \|\xi^*\| \cdot \|a\| + \|\xi\| \cdot \|a^*\| \\ &= \big(2\|a\|(\|\xi^*\|^2 + \|\xi\|^2)\big)^{1/2} \\ &= \big(\Phi^*(a, a^*) \cdot 2\varphi(a^*a)\big)^{1/2}, \end{aligned}$$

which gives the assertion.
2) We have equality in the above inequalities exactly if $\xi = ra^*$ for some $r \in \mathbb{C}$. But this means that the only non-vanishing cumulants in the variables a and a^* are

$$\varphi(a^*a) = k_2(a, a^*) = k_2(a^*, a) = \frac{1}{r} k_2(\xi, a) = \frac{1}{r}$$

Hence $r > 0$ and $\sqrt{r}a$ is a circular element, by Example 6.6. ☐

PROBLEM 10.11. Let μ and ν be probability measures on $\mathbb{R}$. Minimize $\Phi^*(x, y)$ for selfadjoint variables x and y under the constraints that the distribution of x is equal to μ and the distribution of y is equal to ν.

PROPOSITION 10.12. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ and consider two selfadjoint random variables $x, y \in \mathcal{A}$.*

Then we have

$$\Phi^*(x,y) \geq \Phi^*(x) + \Phi^*(y). \tag{117}$$

PROOF. We can assume that a conjugate system $\{\xi_x, \xi_y\} \subset L(x,y)$ for $\{x,y\}$ exists (otherwise the assertion is trivial). Then ξ_x fulfills the conjugate relation for x, hence $\Phi^*(x) \leq \|\xi_x\|^2$, and ξ_y fulfills the conjugate relation for y, hence $\Phi^*(y) \leq \|\xi_y\|^2$. Thus we get

$$\Phi^*(x,y) = \|\xi_x\|^2 + \|\xi_y\|^2 \geq \Phi^*(x) + \Phi^*(y).$$

□

PROPOSITION 10.13. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ and consider two selfadjoint random variables $x, y \in \mathcal{A}$. Assume that x and y are free. Then we have*

$$\Phi^*(x,y) = \Phi^*(x) + \Phi^*(y). \tag{118}$$

PROOF. Let $\xi_x \in L^2(x)$ be the conjugate variable for x and let $\xi_y \in L^2(y)$ be the conjugate variable for y. We claim that $\{\xi_x, \xi_y\}$ fulfill the conjugate relations for $\{x,y\}$. Let us only consider the equations involving ξ_x, the others are analogous. We have to show that $k_{n+1}(\xi_x, a_1, \dots, a_n) = 0$ for all $n \neq 1$ and all $a_1, \dots, a_n \in \{x,y\}$ and that $k_2(\xi_x, x) = 1$ and $k_2(\xi_x, y) = 0$. In the case that all $a_i = x$ we get this from the fact that ξ_x is conjugate for x. If we have $a_i = y$ for at least one i, then the corresponding cumulant is always zero because it is a mixed cumulant in the free sets $\{\xi_x, x\}$ and $\{y\}$. Thus $\{\xi_x, \xi_y\}$ fulfills the conjugate relations for $\{x,y\}$ and since clearly $\xi_x, \xi_y \in L^2(x,y)$, this is the conjugate system. So we get

$$\Phi^*(x,y) = \|\xi_x\|^2 + \|\xi_y\|^2 = \Phi^*(x) + \Phi^*(y).$$

□

PROPOSITION 10.14. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ and consider two selfadjoint random variables $x, y \in \mathcal{A}$. Assume that*

$$\Phi^*(x,y) = \Phi^*(x) + \Phi^*(y) < \infty. \tag{119}$$

Then x and y are free.

PROOF. Let $\{\xi_x, \xi_y\} \subset L^2(x,y)$ be the conjugate system for $\{x,y\}$. If we denote by P_x and P_y the orthogonal projections from $L^2(x,y)$ onto $L^2(x)$ and onto $L^2(y)$, respectively, then $P_x\xi_x$ is the conjugate variable for x and $P_y\xi_y$ is the conjugate variable for y. Our assertion on equality in Equation (9) is equivalent to the statements that $P_x\xi_x = \xi_x$

and $P_y\xi_y = \xi_y$, i.e. that $\xi_x \in L^2(x)$ and $\xi_y \in L^2(y)$. This implies in particular $x\xi_x = \xi_x x$, and thus

$$k_{n+1}(x\xi_x, a_1, \dots, a_n) = k_{n+1}(\xi_x x, a_1, \dots, a_n)$$

for all $n \geq 0$ and all $a_1, \dots, a_n \in \{x, y\}$. However, by using our formula for cumulants with products as entries, the left hand side gives

$$LHS = k_{n+2}(x, \xi_x, a_1, \dots, a_n) + \delta_{a_1 x} k_n(x, a_2, \dots, a_n),$$

whereas the right hand side reduces to

$$RHS = k_n(\xi_x, x, a_1, \dots, a_n) + \delta_{x a_n} k_n(x, a_1, \dots, a_{n-1}).$$

For $n \geq 1$, the first term disappears in both cases and we get

$$\delta_{a_1 x} k_n(x, a_2, \dots, a_n) = \delta_{x a_n} k_n(x, a_1, \dots, a_{n-1})$$

for all choices of $a_1, \dots, a_n \in \{x, y\}$. If we put in particular $a_n = x$ and $a_1 = y$ then this reduces to

$$0 = k_n(x, y, a_2, \dots, a_{n-1}) \qquad \text{for all } a_2, \dots, a_{n-1} \in \{x, y\}.$$

But this implies, by traciality, that all mixed cumulants in x and y vanish, hence x and y are free. ☐

Problem 10.15. Let μ be a probability measure on $\mathbb{R}$. Then we want to look for a generic representation of μ by a $d \times d$ matrix according to the following problem: What is the minimal value of $\Phi^*(\{a_{ij}\}_{i,j=1,\dots,d})$ under the constraint that the matrix $A := (a_{ij})_{i,j=1}^d$ is self-adjoint (i.e. we must have $a_{ij} = a_{ji}^*$ for all $i, j = 1, \dots, d$) and that the distribution of A is equal to the given μ? In which cases is this minimal value actually attained?

Let us first derive an lower bound for the considered Fisher informations.

Proposition 10.16. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ. Consider random variables $a_{ij} \in \mathcal{A}$ $(i, j = 1, \dots, d)$ with $a_{ij} = a_{ji}^*$ for all $i, j = 1, \dots, d$. Put $A := (a_{ij})_{i,j=1}^d \in M_d(\mathcal{A})$. Then we have*

$$\Phi^*(\{a_{ij}\}_{i,j=1,\dots,d}) \geq d^3 \Phi^*(A). \tag{120}$$

Proof. If no conjugate system for $\{a_{ij}\}_{i,j=1,\dots,d}$ exists, then the left hand side of the assertion is ∞, thus the assertion is trivial. So let us assume that a conjugate system $\{\xi_{ij}\}_{i,j=1,\dots,d}$ for $\{a_{ij}\}_{i,j=1,\dots,d}$ exists. Then we put

$$\Xi := \frac{1}{d}(\xi_{ji})_{i,j=1}^d \in M_d(L^2(\mathcal{A})) \hat{=} L^2(M_d(\mathcal{A})).$$

We claim that Ξ fulfills the conjugate relations for A. This can be seen as follows:

$$\varphi \otimes \mathrm{tr}_d(XA^n) = \frac{1}{d} \sum_{i(1),\dots,i(n+1)=1}^{d} \varphi\big(\frac{1}{d}\xi_{i(2)i(1)} a_{i(2)i(3)} \dots a_{i(n+1)i(1)}\big)$$

$$= \sum_{k=2}^{n+1} \frac{1}{d^2} \delta_{i(2)i(k)} \delta_{i(1)i(k+1)} \sum_{i(1),\dots,i(n+1)=1}^{d} \varphi\big(a_{i(2)i(3)} \dots a_{i(k-1)i(k)}\big) \cdot \varphi\big(a_{i(k+1)i(k+2)} \dots a_{i(n+1)i(1)}\big)$$

$$= \sum_{k=2}^{n+1} \Big(\frac{1}{d} \sum_{i(2),\dots,i(k-1)=1}^{d} \varphi\big(a_{i(2)i(3)} \dots a_{i(k-1)i(2)}\big)\Big) \cdot \Big(\frac{1}{d} \sum_{i(1),i(k+2),\dots,i(n+1)=1}^{d} \varphi\big(a_{i(1)i(k+2)} \dots a_{i(n+1)i(1)}\big)\Big)$$

$$= \sum_{k=2}^{n+1} \varphi \otimes \mathrm{tr}_d(A^{k-1}) \cdot \varphi \otimes \mathrm{tr}_d(A^{n-k}).$$

Thus we have according to Remark 10.8

$$\Phi^*(A) \le \|\Xi\|^2 = \frac{1}{d} \sum_{i,j=1}^{d} \varphi\big(\frac{1}{d}\xi_{ij}^* \frac{1}{d}\xi_{ij}\big) = \frac{1}{d^3} \sum_{i,j=1}^{d} \|\xi_{ij}\|^2 = \Phi^*(\{a_{ij}\}_{i,j=1,\dots,d}).$$

□

Next we want to show that the lower bound is actually achieved.

PROPOSITION 10.17. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ. Consider random variables $a_{ij} \in \mathcal{A}$ $(i, j = 1, \dots, d)$ with $a_{ij} = a_{ji}^*$ for all $i, j = 1, \dots, d$. Put $A := (a_{ij})_{i,j=1}^d \in M_d(\mathcal{A})$. If A is free from $M_d(\mathbb{C})$, then we have*

$$\Phi^*(\{a_{ij}\}_{i,j=1,\dots,d}) = d^3 \Phi^*(A). \tag{121}$$

PROOF. Let $\{\xi_{ij}\}_{i,j=1,\dots,d}$ be a conjugate system for $\{a_{ij}\}_{i,j=1,\dots,d}$. Then we put as before $\Xi := \frac{1}{d}(\xi_{ji})_{i,j=1}^d$. In order to have equality in Inequality (10), we have to show that $\Xi \in L^2(A, \varphi \otimes \mathrm{tr}_d)$. Since our assertion depends only on the joint distribution of the considered variables $\{a_{ij}\}_{i,j=1,\dots,d}$, we can choose a convenient realization of the assumptions. Such a realization is given by the compression with a free family of matrix units. Namely, let a be a self-adjoint random variable with distribution μ (where μ is the distribution of the matrix A) and let $\{e_{ij}\}_{i,j=1,\dots,d}$ be a family of matrix units which is free from a. Then the compressed variables $\tilde{a}_{ij} := e_{1i} a e_{j1}$ have the same joint distribution as the given variables a_{ij} and by Theorem 8.14 and Theorem 8.18 we know that the matrix $\tilde{A} := (\tilde{a}_{ij}\}_{i,j=1}^d$ is free from $M_d(\mathbb{C})$. Thus we are done if we can show the assertion for the $\tilde{a}_{ij}$ and $\tilde{A}$. Let ξ be the conjugate variable for a, then we get the conjugate variables for $\{\tilde{a}_{ij}\}_{i,j=1,\dots,d}$ also by compression. Namely, let us put $\tilde{\xi}_{ij} := de_{1i}\xi e_{j1}$. Then we have by Theorem 8.14

$$k_n^{e_{11}\tilde{A}e_{11}}(\tilde{\xi}_{i(1)j(1)}, \tilde{a}_{i(2)j(2)}, \dots, \tilde{a}_{i(n)j(n)})$$
$$= \begin{cases} dk_n(\frac{1}{d}d\xi, \frac{1}{d}a, \dots, \frac{1}{d}a), & \text{if } (i(1)j(1), \dots, i(n)j(n)) \text{ cyclic} \\ 0, & \text{otherwise} \end{cases}$$
$$= \begin{cases} \delta_{j(1)i(2)}\delta_{j(2)1(1)}, & n = 2 \\ 0, & n \neq 2. \end{cases}$$

Thus $\{\tilde{\xi}_{ji}\}_{i,j=1,\dots,d}$ fulfills the conjugate relations for $\{\tilde{a}_{ij}\}_{i,j=1,\dots,d}$. Furthermore, it is clear that $\{\tilde{\xi}_{ij}\} \subset L^2(\{\tilde{a}_{kl}\}_{k,l=1,\dots,d})$, hence this gives the conjugate family. Thus the matrix $\tilde{\Xi}$ (of which we have to show that it belongs to $L^2(\tilde{A})$) has the form

$$\tilde{\Xi} = \frac{1}{d}(\tilde{\xi}_{ij})_{i,j=1}^d = (e_{1i}\xi e_{j1})_{i,j=1}^d.$$

However, since the mapping $y \mapsto (e_{1i}ye_{j1})_{i,j=1}^d$ is an isomorphism, which sends a to $\tilde{A}$ and ξ to $\tilde{\Xi}$, the fact that $\xi \in L^2(a)$ implies that $\tilde{\Xi} \in L^2(\tilde{A})$. □

Now we want to show that the minimal value of the Fisher information can be reached (if it is finite) only in the above treated case.

PROPOSITION 10.18. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ. Consider random variables $a_{ij} \in \mathcal{A}$ $(i, j = 1, \dots, d)$ with $a_{ij} = a_{ji}^*$ for all $i, j = 1, \dots, d$. Put $A := (a_{ij})_{i,j=1}^d \in M_d(\mathcal{A})$. Assume that we also have*

$$\Phi^*(\{a_{ij}\}_{i,j=1,\dots,d}) = d^3\Phi^*(A) < \infty. \tag{122}$$

Then A is free from $M_d(\mathbb{C})$.

PROOF. By Theorem 8.18, we know that the freeness of A from $M_d(\mathbb{C})$ is equivalent to a special property of the cumulants of the a_{ij}. We will show that the assumed equality of the Fisher informations implies this property for the cumulants. Let $\{\xi_{ij}\}_{i,j=1,\dots,d}$ be a conjugate system for $\{a_{ij}\}_{i,j=1,\dots,d}$ and put as before $\Xi := \frac{1}{d}(\xi_{ji})_{i,j=1}^d$. Our assumption is equivalent to the statement that $\Xi \in L^2(A)$. But this implies in particular that $XA = AX$ or

$$\sum_{k=1}^{d} \xi_{ki} a_{kj} = \sum_{k=1}^{d} a_{ik} \xi_{jk} \qquad \text{for all } i, j = 1, \dots, d.$$

This gives for arbitrary $n \geq 1$ and $1 \leq i, j, i(1), j(1), \dots, i(n), j(n) \leq d$ the equation

$$\sum_{k=1}^{d} k_{n+1}\big(\xi_{ki} a_{kj}, a_{i(1)j(1)}, \dots, a_{i(n)j(n)}\big) = \sum_{k=1}^{d} k_{n+1}\big(a_{ik} \xi_{jk}, a_{i(1)j(1)}, \dots, a_{i(n)j(n)}\big).$$

Let us calculate both sides of this equation by using our formula for cumulants with products as entries. Since the conjugate variable can, by Prop. 10.5, only be coupled with one of the variables, this formula reduces to just one term. For the right hand side we obtain

$$\begin{aligned}\sum_{k=1}^{d} & k_{n+1}\big(a_{ik} \xi_{jk}, a_{i(1)j(1)}, \dots, a_{i(n)j(n)}\big) \\ &= \sum_{k=1}^{d} \delta_{ji(1)} \delta_{kj(1)} k_n\big(a_{ik}, a_{i(2)j(2)}, \dots, a_{i(n)j(n)}\big) \\ &= \delta_{ji(1)} k_n\big(a_{ij(1)}, a_{i(2)j(2)}, \dots, a_{i(n)j(n)}\big),\end{aligned}$$

whereeas the left hand side gives

$$\begin{aligned}\sum_{k=1}^{d} & k_{n+1}\big(\xi_{ki} a_{kj}, a_{i(1)j(1)}, \dots, a_{i(n)j(n)}\big) \\ &= \sum_{k=1}^{d} \delta_{ki(n)} \delta_{ij(n))} k_n\big(a_{kj}, a_{i(1)j(1)}, \dots, a_{i(n-1)j(n-1)}\big) \\ &= \delta_{ij(n)} k_n\big(a_{i(n)j}, a_{i(1)j(1)}, \dots, a_{i(n-1)j(n-1)}\big).\end{aligned}$$

Thus we have for all $n \geq 1$ and $1 \leq i, j, i(1), j(1), \dots, i(n), j(n) \leq d$ the equality

$$\begin{aligned}\delta_{ji(1)} & k_n\big(a_{ij(1)}, a_{i(2)j(2)}, \dots, a_{i(n)j(n)}\big) \\ &= \delta_{ij(n)} k_n\big(a_{i(n)j}, a_{i(1)j(1)}, \dots, a_{i(n-1)j(n-1)}\big).\end{aligned}$$

If we put now $i = j(n)$ and $j \neq i(1)$ then this gives

$$k_n\big(a_{i(n)j}, a_{i(1)j(1)}, \dots, a_{i(n-1)j(n-1)}\big) = 0.$$

Thus each cumulant which does not couple in a cyclic way the first double index with the second double index vanishes. By traciality this

implies that any non-cyclic cumulant vanishes.
Let us now consider the case $j = i(1)$ and $i = j(n)$. Then we have

$$k_n(a_{j(n)j(1)}, a_{i(2)j(2)}, \dots, a_{i(n)j(n)}) = k_n(a_{i(n)i(1)}, a_{i(1)j(1)}, \dots, a_{i(n-1)j(n-1)}).$$

Let us specialize further to the case $j(1) = i(2)$, $j(2) = i(3)$,..., $j(n-1) = i(n)$, but not necessarily $j(n) = i(1)$. Then we get

$$\begin{aligned} k_n(a_{j(n)i(2)}, a_{i(2)i(3)}, \dots, a_{i(n)j(n)}) &= k_n(a_{i(n)i(1)}, a_{i(1)i(2)}, \dots, a_{i(n-1)i(n)}) \\ &= k_n(a_{i(1)i(2)}, a_{i(2)i(3)}, \dots, a_{i(n)i(1)}). \end{aligned}$$

Thus we see that we have equality between the cyclic cumulants which might differ at one position in the index set $(i(1), \dots, i(n))$ (here at position n). Since by iteration and traciality we can change that index set at any position, this implies that the cyclic cumulants $k_n(a_{i(1)i(2)}, \dots, a_{i(n)i(1)})$ depend only on the value of n. But this gives, according to Theorem 8.18, the assertion. $\square$

PROBLEM 10.19. Let a probability measure ν on $\mathbb{R}_+$ be given. What is the minimal value of $\Phi^*(a, a^*)$ under the constraint that the distribution of aa^* is equal to the given ν? In which cases is this minimal value actually attained?

NOTATION 10.20. Let ν be a probability measure on $\mathbb{R}_+$. Then we call **symmetric square root of** ν the unique probability measure μ on $\mathbb{R}$ which is symmetric (i.e. $\mu(S) = \mu(-S)$ for each Borel set $S \subset \mathbb{R}$) and which is connected with ν via

(123)
$$\mu(S) = \nu(\{s^2 \mid s \in S\}) \qquad \text{for each Borel set } S \text{ such that } S = -S.$$

REMARK 10.21. For $a \in \mathcal{A}$ we put

$$A := \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix} \in M_2(\mathcal{A}). \tag{124}$$

Then we have that $A = A^*$, A is even and the distribution of A^2 is equal to the distribution of a^*a (which is, by traciality, the same as the distribution of aa^*). Thus we can reformulate the above problem also in the following matrix form: For a given ν on $\mathbb{R}_+$ let μ be the symmetric square root of ν. Determine the minimal value of $\Phi^*(a, a^*)$ under the constraint that the matrix

$$A = \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix}$$

has distribution μ. In which cases is this minimal value actually achieved?

Let us start again by deriving a lower bound for the Fisher information $\Phi^*(a, a^*)$.

PROPOSITION 10.22. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ and let $a \in \mathcal{A}$ be a random variable. Put*

$$A := \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix} \in M_2(\mathcal{A}). \tag{125}$$

Then we have

$$\Phi^*(a, a^*) \geq 2\Phi^*(A). \tag{126}$$

PROOF. Again it suffices to consider the case where the left hand side of our assertion is finite. So we can assume that a conjugate system for $\{a, a^*\}$ exists, which is automatically of the form $\{\xi, \xi^*\} \in L^2(a, a^*)$. Let us put

$$\Xi := \begin{pmatrix} 0 & \xi^* \\ \xi & 0 \end{pmatrix}.$$

We claim that Ξ fulfills the conjugate relations for A. This can be seen as follows: $\varphi \otimes \mathrm{tr}_2(\Xi A^n) = 0$ if n is even, and in the case $n = 2m + 1$ we have

$$\begin{aligned}
\varphi \otimes \mathrm{tr}_2(\Xi A^{2m+1}) &= \frac{1}{2}\big(\varphi(\xi^* a^*(aa^*)^m) + \varphi(\xi a(a^*a)^m)\big) \\
&= \frac{1}{2}\sum_{k=1}^{m}\big(\varphi((a^*a)^{k-1}) \cdot \varphi((aa^*)^{m-k}) + \varphi((aa^*)^{k-1}) \cdot \varphi((a^*a)^{m-k})\big) \\
&= \sum_{k=1}^{m} \varphi \otimes \mathrm{tr}_2(A^{2(k-1)}) \cdot \varphi \otimes \mathrm{tr}_2(A^{2(m-k)}).
\end{aligned}$$

Thus we get

$$\Phi^*(A) \leq \|X\|^2 = \frac{1}{2}(\|\xi\|^2 + \|\xi^*\|^2) = \frac{1}{2}\Phi^*(a, a^*).$$

□

Next we want to show that the minimal value is actually achieved for R-diagonal elements.

PROPOSITION 10.23. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ and let $a \in \mathcal{A}$ be a random variable. Put*

$$A := \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix}.$$

Assume that a is R-diagonal. Then we have

$$\Phi^*(a, a^*) = 2\Phi^*(A). \tag{127}$$

PROOF. As before we put

$$\Xi := \begin{pmatrix} 0 & \xi^* \\ \xi & 0 \end{pmatrix}.$$

We have to show that $\Xi \in L^2(A)$. Again we will do this with the help of a special realization of the given situation. Namely, let x be an even random variable with distribution μ (where μ is the symmetric square root of the distribution of aa^*) and let u be a Haar unitary which is $*$-free from x. Then $\tilde{a} := ux$ is R-diagonal with the same $*$-distribution as a. Thus it suffices to show the assertion for

$$\tilde{A} = \begin{pmatrix} 0 & \tilde{a} \\ \tilde{a}^* & 0 \end{pmatrix} \qquad \text{and} \qquad \tilde{\Xi} := \begin{pmatrix} 0 & \tilde{\xi}^* \\ \tilde{\xi} & 0 \end{pmatrix},$$

where $\{\tilde{\xi}, \tilde{\xi}^*\}$ is the conjugate system for $\{\tilde{a}, \tilde{a}^*\}$. Let ζ be the conjugate variable for x. We claim that the conjugate system for $\{ux, xu^*\}$ is given by $\{\zeta u^*, u\zeta\}$. This can be seen as follows: Let us denote $a_1 := ux$ and $a_2 := xu^*$. We will now consider $k_{n+1}(u\zeta, a_{i(1)}, \dots, a_{i(n)})$ for all possible choices of $i(1), \dots, i(n) \in \{1, 2\}$. If we write each a_i as a product of two terms, and use our formula for cumulants with products as entries, we can argue as usual that only such π contribute which connect ζ with the x from $a_{i(1)}$ and that actually $a_{i(1)}$ must be of the form xu^*. But the summation over all possibilities for the remaining blocks of π corresponds exactly to the problem of calculating $k_n(uu^*, a_{i(2)}, \dots, a_{i(n)})$, which is just δ_{n1}. Hence we have

$$k_{n+1}(u\zeta, xu^*, a_{i(2)}, \dots, a_{i(n)}) = k_2(\zeta, x) \cdot k_n(uu^*, a_{i(2)}, \dots, a_{i(n)}) = \delta_{n1}.$$

The conjugate relations for ζu^* can be derived in the same way. Furthermore, one has to observe that x even implies $\varphi(\zeta x^{2m}) = 0$ for all $m \in \mathbb{N}$, and thus ζ lies in the L^2-closure of the linear span of odd powers of x. This, however, implies that ζu^* and $u\zeta$ are elements in $L^2(ux, xu^*) = L^2(a, a^*)$, and thus $\{\zeta u^*, u\zeta\}$ is indeed the conjugate system for $\{a, a^*\}$.

Thus we have

$$\tilde{A} = \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix} \begin{pmatrix} u^* & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\tilde{\Xi} = \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & \zeta \\ \zeta & 0 \end{pmatrix} \begin{pmatrix} u^* & 0 \\ 0 & 1 \end{pmatrix}.$$

But the fact that $\zeta \in L^2(x)$ implies (by using again the fact that only odd powers of x are involved in ζ)

$$\begin{pmatrix} 0 & \zeta \\ \zeta & 0 \end{pmatrix} \in L^2(\begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix})$$

and hence also $\tilde{\Xi} \in L^2(\tilde{A})$. $\square$

Finally, we want to show that equality in Inequality (16) is, in the case of finite Fisher information, indeed equivalent to being in the R-diagonal situation.

PROPOSITION 10.24. *Let $(\mathcal{A}, \varphi)$ be a W^*-probability space with a faithful trace φ and let $a \in \mathcal{A}$ be a random variable. Put*

$$A := \begin{pmatrix} 0 & a \\ a^* & 0 \end{pmatrix}.$$

Assume that we have

$$\Phi^*(a, a^*) = 2\Phi^*(A) < \infty. \tag{128}$$

Then a is R-diagonal.

PROOF. Let $\{\xi, \xi^*\}$ be the conjugate system for $\{a, a^*\}$ and put

$$\Xi := \begin{pmatrix} 0 & \xi^* \\ \xi & 0 \end{pmatrix}.$$

Our assumption on equality in Equation (18) is equivalent to the fact that $\Xi \in L^2(A)$. In particular, this implies that $\Xi A = A\Xi$ or

$$\begin{pmatrix} \xi^* a^* & 0 \\ 0 & \xi a \end{pmatrix} = \begin{pmatrix} a\xi & 0 \\ 0 & a^*\xi^* \end{pmatrix}.$$

We will now show that the equality $\xi^* a^* = a\xi$ implies that a is R-diagonal. We have for all choices of $n \geq 1$ and $a_1, \dots, a_n \in \{a, a^*\}$

$$k_{n+1}(\xi^* a^*, a_1, \dots, a_n) = k_{n+1}(a\xi, a_1, \dots, a_n).$$

Calculating both sides of this equation with our formula for cumulants with products as entries we get

$$\delta_{a_n a^*} k_n(a^*, a_1, \dots, a_{n-1}) = \delta_{a_1 a} k_n(a, a_2, \dots, a_n).$$

Specifying $a_1 = a_n = a$ we get that

$$k_n(a, a_2, \dots, a_{n-1}, a) = 0$$

for arbitrary $a_2, \dots, a_{n-1} \in \{a, a^*\}$. By traciality, this just means that non-alternating cumulants vanish, hence that a is R-diagonal. $\square$

Part 3

Appendix

APPENDIX A

Infinitely divisible distributions

THEOREM A.1. **(general free limit theorem)** *Let, for each $N \in \mathbb{N}$, $(\mathcal{A}_N, \varphi_N)$ be a probability space. Let I be an index set. Consider a triangular field of random variables, i.e. for each $i \in I$, $N \in \mathbb{N}$ and $1 \le r \le N$ we have a random variable $a^{(i)}_{N;r} \in \mathcal{A}_N$. Assume that, for each fixed $N \in \mathbb{N}$, the sets $\{a^{(i)}_{N;1}\}_{i\in I}, \dots, \{a^{(i)}_{N;N}\}_{i\in I}$ are free and identically distributed and that furthermore for all $n \ge 1$ and all $i(1), \dots, i(n) \in I$ the limits*

$$\lim_{N\to\infty} N \cdot \varphi_N(a^{i(1)}_{N;r} \cdots a^{i(n)}_{N;r}) \tag{129}$$

(which are independent of r by the assumption of identical distribution) exist. Then we have

$$\big(a^{(i)}_{N;1} + \dots + a^{(i)}_{N;N}\big)_{i\in I} \xrightarrow{distr} (a_i)_{i\in I}, \tag{130}$$

where the joint distribution of the family $(a_i)_{i\in I}$ is determined by ($n \ge 1$, $i(1), \dots, i(n) \in I$)

$$k_n(a_{i(1)}, \dots, a_{i(n)}) = \lim_{N\to\infty} N \cdot \varphi_N(a^{i(1)}_{N;r} \cdots a^{i(n)}_{N;r}). \tag{131}$$

LEMMA A.2. *Let $(\mathcal{A}_N, \varphi_N)$ be a sequence of probability spaces and let, for each $i \in I$, a random variable $a^{(i)}_N \in \mathcal{A}_N$ be given. Denote by k^N the free cumulants corresponding to φ_N. Then the following two statements are equivalent:*

1. *For each $n \ge 1$ and each $i(1), \dots, i(n) \in I$ the limit*

$$\lim_{N\to\infty} N \cdot \varphi_N(a^{i(1)}_N \cdots a^{i(n)}_N) \tag{132}$$

exists.

2. *For each $n \ge 1$ and each $i(1), \dots, i(n) \in I$ the limit*

$$\lim_{N\to\infty} N \cdot k^N_n(a^{i(1)}_N \cdots a^{i(n)}_N) \tag{133}$$

exists.

Furthermore the corresponding limits agree in both cases.

PROOF. (2) $\Longrightarrow$ (1): We have

$$\lim_{N\to\infty} N\cdot\varphi_N(a_N^{i(1)}\cdots a_N^{i(n)}) = \lim_{N\to\infty}\sum_{\pi\in NC(n)} N\cdot k_\pi[a_N^{i(1)},\dots,a_N^{i(n)}].$$

By assumption (2), all terms for π with more than one block tend to zero and the term for $\pi = 1_n$ tends to a finite limit.
The other direction (1) $\Longrightarrow$ (2) is analogous. □

PROOF. We write

$$(a_{N;1}^{(i(1))}+\cdots+a_{N;N}^{(i(N))})^n = \sum_{r(1),\dots,r(n)=1}^{N} \varphi_N(a_{N;r(1)}^{(i(1))}+\cdots+a_{N;r(n)}^{(i(n))})$$

and observe that for fixed N a lot of terms in the sum give the same contribution. Namely, the tuples $(r(1),\dots,r(n))$ and $(r'(1),\dots,r'(n))$ give the same contribution if the indices agree at the same places. As in the case of the central limit theorem, we encode this relevant information by a partition π (which might apriori be a crossing partition). Let $(r(1),\dots,r(n))$ be an index-tuple corresponding to a fixed π. Then we can write

$$\varphi_N(a_{N;r(1)}^{(i(1))}+\cdots+a_{N;r(n)}^{(i(n))}) = \sum_{\sigma\in NC(n)} k_\sigma^N[a_{N;r}^{(i(1))}+\cdots+a_{N;r}^{(i(n))}]$$

(where the latter expression is independent of r). Note that because elements belonging to different blocks of π are free the sum runs effectively only over such $\sigma\in NC(n)$ with the property $\sigma\le\pi$. The number of tuples $(r(1),\dots,r(n))$ corresponding to π is of order $N^{|\pi|}$, thus we get

$$\lim_{N\to\infty}(a_{N;1}^{(i(1))}+\cdots+a_{N;N}^{(i(n))})^n = \sum_{\pi\in\mathcal{P}(n)}\sum_{\substack{\sigma\in NC(n)\\ \sigma\le\pi}}\lim_{N\to\infty} N^{|\pi|}k_\sigma^N[a_{N;r}^{(i(1))}+\cdots+a_{N;r}^{(i(n))}].$$

By Lemma ..., we get non-vanishing contributions exactly in those cases where the power of N agrees with the number of factors from the cumulants k_σ. This means that $|\pi| = |\sigma|$, which can only be the case if π itself is a non-crossing partition and $\sigma = \pi$. But this gives exactly the assertion. □

DEFINITIONS A.3. Let $\mathcal{H}$ be a Hilbert space.
1) The **full Fock space** over $\mathcal{H}$ is defined as

$$\mathcal{F}(\mathcal{H}) := \bigoplus_{n=0}^{\infty}\mathcal{H}^{\otimes n}, \tag{134}$$

, where $\mathcal{H}^{\otimes 0}$ is a one-dimensional Hilbert, which we write in the form $\mathbb{C}\Omega$ for a distinguished vector of norm one, called **vacuum**.

2) For each $f \in \mathcal{H}$ we define the **creation operator** $l(f)$ by linear extension of

$$l(f)f_1 \otimes \cdots \otimes f_n := f \otimes f_1 \otimes \cdots \otimes f_n. \tag{135}$$

Its adjoint, the **annihilation operator** $l^*(f)$, is given by

$$l(f)f_1 \otimes \cdots \otimes f_n = \langle f, f_1 \rangle f_2 \otimes \cdots \otimes f_n \qquad (n \geq 1) \tag{136}$$

$$l(f)\Omega = 0. \tag{137}$$

3) For each operator $T \in B(\mathcal{H})$ we define the **gauge operator** $\Lambda(T)$ by

$$\Lambda(T)\Omega := 0 \tag{138}$$

$$\Lambda(T)f_1 \otimes f_n := (Tf_1) \otimes f_2 \otimes \cdots \otimes f_n. \tag{139}$$

4) We denote by

$$\mathcal{A}(\mathcal{H}) := C^*(l(f), \Lambda(T) \mid f \in \mathcal{H}, T \in B(\mathcal{H})) \subset B(\mathcal{F}(\mathcal{H})) \tag{140}$$

the unital C^*-algebra generated by all creation, annihilation and gauge operators.
5) The vector state on $\mathcal{A}(\mathcal{H})$ given by the vacuum,

$$\varphi_{\mathcal{H}}(a) := \langle \Omega, a\Omega \rangle \qquad (a \in \mathcal{A}(\mathcal{H})) \tag{141}$$

is called **vacuum expectation state**.

Proposition A.4. *Let a Hilbert space $\mathcal{H}$ decompose as a direct sum $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$. Then we can consider $\mathcal{A}(\mathcal{H}_1)$ and $\mathcal{A}(\mathcal{H}_2)$ in a canonical way as subalgebras of $\mathcal{A}(\mathcal{H})$ and we have that $\mathcal{A}(\mathcal{H}_1)$ and $\mathcal{A}(\mathcal{H}_2)$ are free in $(\mathcal{A}(\mathcal{H}), \varphi_{\mathcal{H}})$.*

Proof. direct verification □

Proposition A.5. *Let $\mathcal{H}$ be a Hilbert space and consider the C^*-probability space $(\mathcal{A}(\mathcal{H}), \varphi_{\mathcal{H}})$. Then the cumulants of the random variables $l(f)$, $l^*(g)$, $\Lambda(T)$ ($f, g \in \mathcal{H}$, $T \in B(\mathcal{H})$) are of the following form: We have ($n \geq 2$, $f, g \in \mathcal{H}$, $T_1, \dots, T_{n-2} \in B(\mathcal{H})$)*

$$k_n(l^*(f), \Lambda(T_1), \dots, \Lambda(T_{n-2}), l(g)) = \langle f, T_1 \dots T_{n-2} g \rangle \tag{142}$$

and all other cumulants with arguments from the set $\{l(f) \mid f \in \mathcal{H}\} \cup \{l^*(g) \mid g \in \mathcal{H}\} \cup \{\Lambda(T) \mid T \in B(\mathcal{H})\}$ vanish.

Proof. For $N \in \mathbb{N}$, put

$$\mathcal{H}_N := \underbrace{\mathcal{H} \oplus \cdots \oplus \mathcal{H}}_{N \text{ times}}$$

and ($f, g \in \mathcal{H}$, $T \in B(\mathcal{H})$)

Then it is easy to see that the random variables $\{l(f), l^*(g), \Lambda(T) \mid f, g \in \mathcal{H}, T \in B(\mathcal{H})\}$ in $(\mathcal{A}(\mathcal{H}), \varphi_{\mathcal{H}})$ have the same joint distribution as the random variables

$$\{l(\frac{f \oplus \cdots \oplus f}{\sqrt{N}}), l^*(\frac{g \oplus \cdots \oplus g}{\sqrt{N}}), \Lambda(T \oplus \cdots \oplus T) \mid f, g \in \mathcal{H}, T \in B(\mathcal{H})\}$$

in $(\mathcal{A}(\mathcal{H}_N), \varphi_{\mathcal{H}_N})$. The latter variables, however, are the sum of N free random variables, the summands having the same joint distribution as $\{l_N(f), l_N^*(g), \Lambda_N(T) \mid f, g \in \mathcal{H}, T \in B(\mathcal{H})\}$ in $(\mathcal{A}(\mathcal{H}), \varphi_{\mathcal{H}})$, where

$$l_N(f) := \frac{1}{\sqrt{N}} l(f)$$
$$l_N^*(g) := \frac{1}{\sqrt{N}} l^*(g)$$
$$\Lambda_N(T) := \Lambda(T).$$

Hence we know from our limit theorem that the cumulants $k_n(a^{(1)}, \dots, a^{(n)})$ for $a^{(i)} \in \{l(f), l^*(g), \Lambda(T) \mid f, g \in \mathcal{H}, T \in B(\mathcal{H})\}$ can also be calculated as

$$k_n(a^{(1)}, \dots, a^{(n)}) = \lim_{N \to \infty} N \cdot \varphi_{\mathcal{H}_N}(a_N^{(1)} \cdots a_N^{(n)}).$$

This yields directly the assertion. □

Definition A.6. Let μ be a probability measure on $\mathbb{R}$ with compact support. We say that μ is **infinitely divisible** (in the free sense) if, for each positive integer n, the convolution power $\mu^{\boxplus 1/n}$ is a probability measure.

Remark A.7. Since $\mu^{\boxplus p/q} = (\mu^{\boxplus 1/q})^{\boxplus p}$ for positive integers p, q, it follows that the rational convolution powers are probability measures. By continuity, we get then also that all convolution powers $\mu^{\boxplus t}$ for real $t > 0$ are probability measure. Thus the property 'infinitely divisible' is equivalent to the existence of the convolution semi-group $\mu^{\boxplus t}$ in the class of probability measures for all $t > 0$.

Corollary A.8. *Let $\mathcal{H}$ be a Hilbert space and consider the C^*-probability space $(\mathcal{A}(\mathcal{H}), \varphi_{\mathcal{H}})$. For $f \in \mathcal{H}$ and $T = T^* \in B(\mathcal{H})$ let a be the self-adjoint operator*

$$a := l(f) + l^*(f) + \Lambda(T). \tag{143}$$

Then the distribution of a is infinitely divisible.

Proof. This was shown in the proof of Prop. ... □

Notation A.9. Let $(t_n)_{n \geq 1}$ be a sequence of complex numbers. We say that $(t_n)_{n \geq 1}$ is **conditionally positive definite** if we have for all $r \in \mathbb{N}$ and all $\alpha_1, \dots, \alpha_r \in \mathbb{C}$ that $\sum_{n,m=1}^{r} \alpha_n \bar{\alpha}_m t_{n+m} \geq 0$.

THEOREM A.10. *Let μ be a probability measure on $\mathbb{R}$ with compact support and let $k_n := k_n^\mu$ be the free cumulants of μ. Then the following two statements are equivalent:*

1. *μ is infinitely divisible.*
2. *The sequence $(k_n)_{n\geq 1}$ of free cumulants of μ is conditionally positive definite.*

PROOF. (1)$\Longrightarrow$(2): Let a_N be a self-adjoint random variable in some C^*-probability space $(\mathcal{A}_N, \varphi_N)$ which has distribution $\mu^{\boxplus 1/N}$. Then our limit theorem tells us that we get the cumulants of μ as

$$k_n = \lim_{N\to\infty} \varphi_N(a_N^n).$$

Consider now $\alpha_1, \dots, \alpha_r \in \mathbb{C}$. Then we have

$$\begin{aligned}\sum_{n,m=1}^{k} \alpha_n \bar{\alpha}_m k_{n+m} &= \lim_{N\to\infty} N \cdot \sum_{n,m=1}^{k} \varphi_N(\alpha_n \bar{\alpha}_m a_N^{n+m}) \\ &= \lim_{N\to\infty} N \cdot \varphi_N\big((\sum_{n=1}^{k} \alpha_n a^n) \cdot (\sum_{m=1}^{k} \alpha_m a^m)^*\big) \\ &\geq 0,\end{aligned}$$

because all φ_N are positive.
(2)$\Longrightarrow$(1): Denote by $\mathbb{C}_0\langle X\rangle$ the polynomials in one variable X without constant term, i.e.

$$\mathbb{C}_0\langle X\rangle := \mathbb{C}X \oplus \mathbb{C}X^2 \oplus \dots.$$

We equip this vector space with an inner product by sesquilinear extension of

$$\langle X^n, X^m\rangle := k_{n+m} \qquad (n, m \geq 1). \tag{144}$$

The assumption (2) on the sequence of cumulants yields that this is indeed a non-negative sesquilinear form. Thus we get a Hilbert space $\mathcal{H}$ after dividing out the kernel and completion. In the following we will identify elements from $\mathbb{C}_0\langle X\rangle$ with their images in $\mathcal{H}$. We consider now in the C^*-probability space $(\mathcal{A}(\mathcal{H}), \varphi_{\mathcal{H}})$ the operator

$$a := l(X) + l^*(X) + \Lambda(X) + k_1 \cdot 1, \tag{145}$$

where X in $\Lambda(X)$ is considered as the multiplication operator with X. By Corollary ..., we know that the distribution of a is infinitely divisible. We claim that this distribution is the given μ. This follows directly from Prop. A5: For $n = 1$, we have

$$k_1(a) = k_1;$$

for $n = 2$, we get

$$k_2(a) = k_2(l^*(X), l(X)) = \langle X, X \rangle = k_2,$$

and for $n > 2$, we have

$$k_n(a) = k_n(l^*(X), \Lambda(X), \dots, \Lambda(X), l(X)\rangle = \langle X, \Lambda(X)^{n-2} X \rangle = \langle X, X^{n-1} \rangle = k_n.$$

Thus all cumulants of a agree with the corresponding cumulants of μ and hence both distributions coincide. □