COMBINATORIAL ASPECTS OF FREE PROBABILITY THEORY

ROLAND SPEICHER*

3. Multiplicaton of free variables and S-transform

3.1. Multiplication of free random variables. To show that our description of freeness in terms of free cumulants has also a significance apart from dealing with additive free convolution, we will apply it to the problem of the product of free random variables: Consider $a_1, \dots, a_n, b_1, \dots, b_n$ such that $\{a_1, \dots, a_n\}$ and $\{b_1, \dots, b_n\}$ are free. We want to express the distribution of the product random variables $a_1 b_1, \dots, a_n b_n$ in terms of the distribution of the a's and of the b's.

Notation 3.1. 1) Analogously to k_π we define for

$$\pi = \{V_1, \dots, V_r\} \in NC(n)$$

the expression

$$\varphi_\pi[a_1, \dots, a_n] := \varphi_{V_1}[a_1, \dots, a_n] \dots \varphi_{V_r}[a_1, \dots, a_n],$$

where

$$\varphi_V[a_1, \dots, a_n] := \varphi(a_{v_1} \cdots a_{v_l}) \qquad \text{for} \qquad V = (v_1, \dots, v_l).$$

Examples:

$$\begin{aligned}
\varphi_{\sqcup\!\sqcup}[a_1, a_2, a_3] &= \varphi(a_1 a_2 a_3) \\
\varphi_{\vert\,\sqcup}[a_1, a_2, a_3] &= \varphi(a_1)\varphi(a_2 a_3) \\
\varphi_{\sqcup\,\vert}[a_1, a_2, a_3] &= \varphi(a_1 a_2)\varphi(a_3) \\
\varphi_{\sqcup\!\!\!\vert\,}[a_1, a_2, a_3] &= \varphi(a_1 a_3)\varphi(a_2) \\
\varphi_{\vert\,\vert\,\vert}[a_1, a_2, a_3] &= \varphi(a_1)\varphi(a_2)\varphi(a_3)
\end{aligned}$$

2) Let $\sigma, \pi \in NC(n)$. Then we write

$$\sigma \leq \pi$$

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

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

if each block of σ is contained as a whole in some block of π, i.e. σ can be obtained out of π by refinement of the block structure.
Example:

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

It is now straighforwared to generalize our moment-cumulant formula

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

in the following way.

Proposition 3.2. *Consider $n \in \mathbb{N}$, $\sigma \in NC(n)$ and $a_1, \dots, a_n \in \mathcal{A}$. Then we have*

$$\varphi_\sigma[a_1, \dots, a_n] = \sum_{\substack{\pi \in NC(n) \\ \pi \leq \sigma}} k_\pi[a_1, \dots, a_n]. \tag{3.1}$$

Consider now

$$\{a_1, \dots, a_n\}, \{b_1, \dots, b_n\} \qquad \text{free.}$$

We want to express alternating moments in a and b in terms of moments of a and moments of b. We have

$$\varphi(a_1 b_1 a_2 b_2 \cdots a_n b_n) = \sum_{\pi \in NC(2n)} k_\pi[a_1, b_1, a_2, b_2, \dots, a_n, b_n].$$

Since the a's are free from the b's, the vanishing of mixed cumulants in free variables tells us that only such π contribute to the sum whose blocks do not connect a's with b's. But this means that such a π has to decompose as

$$\pi = \pi_1 \cup \pi_2 \qquad \text{where} \quad \begin{aligned} &\pi_1 \in NC(1,3,5,\dots,2n-1) \\ &\pi_2 \in NC(2,4,6,\dots,2n). \end{aligned}$$

Thus we have

$$\varphi(a_1 b_1 a_2 b_2 \cdots a_n b_n) = \sum_{\substack{\pi_1 \in NC(odd), \pi_2 \in NC(even) \\ \pi_1 \cup \pi_2 \in NC(2n)}} k_{\pi_1}[a_1, a_2, \dots, a_n] \cdot k_{\pi_2}[b_1, b_2, \dots, b_n]$$

$$= \sum_{\pi_1 \in NC(odd)} \Big(k_{\pi_1}[a_1, a_2, \dots, a_n] \cdot \sum_{\substack{\pi_2 \in NC(even) \\ \pi_1 \cup \pi_2 \in NC(2n)}} k_{\pi_2}[b_1, b_2, \dots, b_n] \Big).$$

Note now that for a fixed π_1 there exists a maximal element σ with the property $\pi_1 \cup \sigma \in NC(2n)$ and that the second sum is running over all $\pi_2 \leq \sigma$.

Definition 3.3. Let $\pi \in NC(n)$ be a non-crossing partition of the numbers $1, \dots, n$. Introduce additional numbers $\bar{1}, \dots, \bar{n}$, with alternating order between the old and the new ones, i.e. we order them in the way

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

We define the *complement* $K(\pi)$ of π as the maximal $\sigma \in NC(\bar{1}, \dots, \bar{n})$ with the property

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

If we present the partition π graphically by connecting the blocks in $1, \dots, n$, then σ is given by connecting as much as possible the numbers $\bar{1}, \dots, \bar{n}$ without getting crossings among themselves and with π. Of course, we identify $NC(\bar{1}, \dots, \bar{n})$ in the end with $NC(n)$, so that we can consider the complement as a mapping on $NC(n)$,

$$K : NC(n) \to NC(n).$$

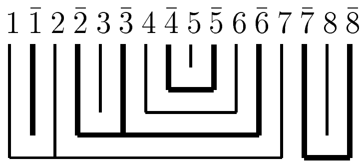
Here is an example for a complement: Consider the partition

$$\pi := \{(1, 2, 7), (3), (4, 6), (5), (8)\} \in NC(8).$$

Then

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

as can be seen from the graphical representation:



.

This natural notation of the complement of a non-crossing partition is also due to Kreweras. Note that there is no analogue of this for the case of all partitions.

With this definition we can continue our above calculation as follows:

$$\begin{aligned}\varphi(a_1 b_1 a_2 b_2 \cdots a_n b_n) &= \sum_{\pi_1 \in NC(n)} \Big(k_{\pi_1}[a_1, a_2, \dots, a_n] \cdot \sum_{\substack{\pi_2 \in NC(n) \\ \pi_2 \leq K(\pi_1)}} k_{\pi_2}[b_1, b_2, \dots, b_n] \Big) \\ &= \sum_{\pi_1 \in NC(n)} k_{\pi_1}[a_1, a_2, \dots, a_n] \cdot \varphi_{K(\pi_1)}[b_1, b_2, \dots, b_n].\end{aligned}$$

This looks a bit unsymmetric in the role of cumulants and moments. By invoking the moment-cumulant formula one can bring this into a much more symmetric form on the level of cumulants.

Theorem 3.4. *Consider*

$$\{a_1, \dots, a_n\}, \{b_1, \dots, b_n\} \qquad \textit{free.}$$

Then we have

(3.2)
$$\varphi(a_1 b_1 a_2 b_2 \cdots a_n b_n) = \sum_{\pi \in NC(n)} k_\pi[a_1, a_2, \dots, a_n] \cdot \varphi_{K(\pi)}[b_1, b_2, \dots, b_n],$$

(3.3)
$$\varphi(a_1 b_1 a_2 b_2 \cdots a_n b_n) = \sum_{\pi \in NC(n)} \varphi_{K^{-1}(\pi)}[a_1, a_2, \dots, a_n] \cdot k_\pi[b_1, b_2, \dots, b_n],$$

and

(3.4)
$$k_n(a_1 b_1, a_2 b_2, \dots, a_n b_n) = \sum_{\pi \in NC(n)} k_\pi[a_1, a_2, \dots, a_n] \cdot k_{K(\pi)}[b_1, b_2, \dots, b_n].$$

Examples 3.5. 1) Let us write down explicitly the first and third formula for small n.
For $n = 1$ we get

$$\varphi(ab) = k_1(a)\varphi(b)$$

and

$$k_1(ab) = k_1(a)k_1(b),$$

both of which are just versions of the factorization rule

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

For $n = 2$ we get

$$\varphi(a_1 b_1 a_2 b_2) = k_1(a_1)k_1(a_2)\varphi(b_1 b_2) + k_2(a_1, a_2)\varphi(b_1)\varphi(b_2)$$

and

$$k_2(a_1 b_1, a_2 b_2) = k_1(a_1)k_1(a_2)k_2(b_1, b_2) + k_2(a_1, a_2)k_1(b_1)k_1(b_2)$$

which are both reformulations of the formula

$$\varphi(a_1 b_1 a_2 b_2) = \varphi(a_1)\varphi(a_2)\varphi(b_1 b_2) + \varphi(a_1 a_2)\varphi(b_1)\varphi(b_2) - \varphi(a_1)\varphi(a_2)\varphi(b_1)\varphi(b_2).$$

2) Let us specialize the second formula of our theorem to the case where $b_1, \dots, b_n$ are elements choosen from free semi-circular elements s_i. The only non-trivial cumulants are then

$$k_2(s_i, s_j) = \delta_{ij},$$

and we have

$$\varphi(a_1 s_{p(1)} \cdots a_n s_{p(n)}) = \sum_{\pi \in NC_2^p(n)} \varphi_{K^{-1}(\pi)}[a_1, \dots, a_n], \tag{3.5}$$

where $NC_2^p(n)$ denotes those non-crossing pairings of n elements whose blocks connect only the same p-indices, i.e., only the same semicirculars. An example is

$$\varphi(a_1 s_1 a_2 s_1 a_3 s_2 a_4 s_2) = \varphi(a_2)\varphi(a_4)\varphi(a_1 a_3).$$

3.2. **Multiplicative free convolution and S-transform.** Restricted to the case

$$a_1 = \cdots = a_n = a, \qquad b_1 = \cdots = b_n = b,$$

the above theorem tells us, on a combinatorial level, how to get, if a and b are free, the moments of ab out of the moments of a and the moments of b. As for the additive problem, we might want to introduce a *multiplicative free convolution* $\boxtimes$ on real probability measures by the prescription

$$\mu_{ab} = \mu_a \boxtimes \mu_b \qquad \text{if } a \text{ and } b \text{ are free.}$$

However, there is a problem with this. Namely, if a and b do not commute, then ab is not selfadjoint, even if a and b are so. So, it is not clear why ab should have a corresponding real probability measure as distribution.

However, in the case that a is a positive operator, then $a^{1/2}$ makes sense and ab has the same moments as $a^{1/2}ba^{1/2}$. (φ restricted to the algebra generated by a and b is necessarily a trace.) The latter, however, is a selfadjoint operator, and has a real probability measure as corresponding distribution. Then the above definition of $\boxtimes$ has to be understood as

$$\mu_a \boxtimes \mu_b = \mu_{a^{1/2}ba^{1/2}}.$$

This allows to define $\boxtimes$ if at least one of the involved real probability measures has support on the positive real axis $\mathbb{R}_+$. If we want to consider $\boxtimes$ as a binary operation, then it acts on probability measures on $\mathbb{R}_+$.

(Again, the above arguments are only for measures which are determined by their moments. However, as in the additive case, truncation and approximation arguments allow to extend the definition of $\boxtimes$ to a binary operation on all probability measures on $\mathbb{R}_+$.)

In order to deal with this free product of random variables in a more analytic way Voiculescu introduced the so-called *S-transform.*

Theorem 3.6. *1) The S-transform of a random variable a is determined as follows. Let χ denote the inverse under composition of the series*

$$\psi(z) := \sum_{n=1}^{\infty} \varphi(a^n) z^n, \tag{3.6}$$

then

$$S_a(z) = \chi(z) z^{-1}(1+z). \tag{3.7}$$

2) If a and b are free, then we have

$$S_{ab}(z) = S_a(z) \cdot S_b(z). \tag{3.8}$$

Note that in this theorem we do not require a or b to be self-adjoint or positive operators. In this formulation, the theorem is true in full generality, since it is essentially a statement about mixed moments of free variables. However, if one wants to extract a probability distribution out of the S-transform or extend to situations where no moments exist (and consider the S-transform not just as a formal power series, but as an analytic function on an appropriate domain), then one has to make special requirements concerning a and b. (Apart from choosing a and b as positive operators, there is another important special situation, namely when they are unitary operators. Note that the product of unitary operators is always unitary. This allows to define $\boxtimes$ also for probability measures on the unit circle, correspoding to the multiplication of free unitary operators.)

The original proof of Voiculescu for this theorem was not easy and relied on studying the exponential map of $\boxtimes$. Apart from the original approach there exist also a proof of Haagerup using creation and annihilation operators and a combinatorial proof (due to Nica and myself) relying on the considerations from the last section. I will say a bit more about this later.

According to our remarks above the S-transform should allow to calculate the multiplicative free convolution between a probability measure μ_a on $\mathbb{R}$ and a probability measure μ_b on $\mathbb{R}_+$. However, one might notice that in the case that $\varphi(a) = 0$, the series ψ has no inverse χ as power series in z, and thus the S-transform seems to be not defined in that case. Thus, in the usual formulation one has to restrict to situations where the first moment does not vanish. However, since in the case $\varphi(a) = 0$ the second moment of a must, apart from the trivial case $a = 0$, be different from zero, χ makes sense as a power series in $\sqrt{z}$, and in this way one can remove the restriction $\varphi(a) \neq 0$ and use the above theorem to calculate the multiplicative convolution between two real probability measures, of which at least one has support on the positive half-axis.

3.3. **Compression by free projection.** Before I say a bit more about a combinatorial proof of the S-transform formula I want to indicate that even without using the S-transform one can get interesting properties about free multiplicative convolution from our combinatorial description.

Definition 3.7. Let $(\mathcal{A},\varphi)$ be a non-commutative probability space and $p\in\mathcal{A}$ a projection with $\varphi(p)\neq 0$. Then we define the *compression* $(p\mathcal{A}p,\tilde{\varphi})$ by

$$p\mathcal{A}p := \{pap \mid a\in\mathcal{A}\}$$

(which is a unital algebra with p as unit) and

$$\tilde{\varphi}(pap) := \frac{1}{\varphi(p)}\varphi(pap).$$

If p is free from $a\in\mathcal{A}$ then we can calculate the distribution of pap in the compressed space in the following way by using our combinatorial formula from Theorem 3.4.

Using $p^k=p$ for all $k\geq 1$ and $\varphi(p)=:\alpha$ gives

$$\varphi_{K(\pi)}[p,p,\dots,p]=\varphi(p\dots p)\varphi(p\dots p)\cdots=\alpha^{|K(\pi)|},$$

where $|K(\pi)|$ denotes the number of blocks of $K(\pi)$. We can express this number of blocks also in terms of π, since we always have the relation

$$|\pi|+|K(\pi)|=n+1. \tag{3.9}$$

Thus we can continue our calculation of Theorem 3.4 in this case as

$$\begin{aligned}\frac{1}{\alpha}\varphi[(ap)^n] &= \frac{1}{\alpha}\sum_{\pi\in NC(n)} k_\pi[a,\dots,a]\alpha^{n+1-|\pi|}\\ &= \sum_{\pi\in NC(n)} \frac{1}{\alpha^{|\pi|}}k_\pi[\alpha a,\dots,\alpha a],\end{aligned}$$

which shows that

$$k_n^{p\mathcal{A}p}(pap,\dots,pap)=\frac{1}{\alpha}k_n(\alpha a,\dots,\alpha a) \tag{3.10}$$

for all n. By our results on the additive free convolution, this gives the surprising result that the renormalized distribution of pap is given by

$$\mu_{pap}^{p\mathcal{A}p}=\mu_{\alpha a}^{\boxplus 1/\alpha}. \tag{3.11}$$

For example, for $\alpha=1/2$, we have

$$\mu_{pap}^{p\mathcal{A}p}=\mu_{a/2}^{\boxplus 2}=\mu_{a/2}\boxplus\mu_{a/2}.$$

Note that given a compactly supported probability measure μ we can always find a selfadjoint operator a having this μ as its distribution and, by a free product construction, a projection p with prescribed trace $\varphi(p)=:\alpha$ between 0 and 1, such that a and p are free. Since $\frac{1}{\alpha}pap$ is again a selfadjoint operator (in the compressed space), its distribution must be a probability measure. However, as we just saw its distribution is $\mu^{\boxplus 1/\alpha}$. As usual one can extend this statement by approximation to all probability measures on $\mathbb{R}$. Thus we have derived the following result of Nica and myself.

Theorem 3.8. *Let μ be a probability measure on $\mathbb{R}$. Then there exists a semi-group (μ_α) ($\alpha \in \mathbb{R}$, $\alpha \geq 1$) such that*

$$\mu_1 = \mu,$$

$$\mu_{\alpha+\beta} = \mu_\alpha \boxplus \mu_\beta \qquad (\alpha, \beta \geq 1)$$

and the mapping

$$\alpha \mapsto \mu_\alpha$$

is continuous with respect to the weak topology on probability measures.

This theorem has no classical counter part. Even though there exist many parallel statements in classical and free probability theory, there is no bijection (neither in statements nor in proofs) between both theories.

Let us state this compression result also in the original probability space $(\mathcal{A}, \varphi)$.

Corollary 3.9. *We have for all real probability measures and all t with $0 < t < 1$ that*

$$\mu \boxtimes \big((1-t)\delta_0 + t\delta_{1/t}\big) = (1-t)\delta_0 + t\mu^{\boxplus 1/t}. \tag{3.12}$$

Of course, one might check this statement also by invoking the S-transform, but the above derivation shows more clearly the reason for the discrepancy between classical and free probability theory in this case. The theorem relies crucially on the existence and the properties (in particular, eq. (3.9)) of the complementation map for non-crossing partitions, and there is no analogue for this in the classical world of all partitions.

3.4. **Combinatorial proof for multiplicativity of S-transform.** Let us now see how one can prove the multiplicativity of the S-transform for the product $c = ab$ of free variables a and b, starting from our combinatorial description in Theorem 3.4

$$k_n^c = \sum_{\pi \in NC(n)} k_\pi^a \cdot k_{K(\pi)}^b.$$

(We have used here the notation $k_\pi^a := k_\pi[a, a, \dots, a]$.)
We expect that the translation of this into generating power series

should result in the multiplicativity for the S-transform. So let us define the generating power series

$$\Phi_a(z) := \sum_{n=1}^{\infty} k_n^a z^n$$

$$\Phi_b(z) := \sum_{n=1}^{\infty} k_n^b z^n$$

$$\Phi_c(z) := \sum_{n=1}^{\infty} k_n^c z^n$$

As before, in deriving the relation between the moment and the cumulant series, we do the summation in

$$k_n^c = \sum_{\pi \in NC(n)} k_\pi^a \cdot k_{K(\pi)}^b$$

in such a way that we fix the first block of π and then sum over all remaining possibilities. This results in the relation

$$\Phi_c = \Phi_a \circ \Phi_\gamma$$

where

$$\Phi_\gamma(z) := \sum_{n=1}^{\infty} \gamma_n z^n$$

with

$$\gamma_n := \sum_{\pi \in NC'(n)} k_\pi^a \cdot k_{K(\pi)}^b;$$

the summation in the latter expression does not run over all of $NC(n)$, but only over the subset

$$NC'(n) := \{\pi \in NC(n) \mid (1) \text{ is a singleton of } \pi\}.$$

The relation $\Phi_c = \Phi_a \circ \Phi_\gamma$ reflects the fact that if we fix the first block V of π, we essentially have to deal with the partitions between consecutive points of V, however to each of them we have to add one point from the block V as a singleton.

If we exchange the roles of a and b and introduce the corresponding quantities

$$\Phi_\delta(z) := \sum_{n=1}^{\infty} \delta_n z^n$$

with

$$\delta_n := \sum_{\pi \in NC'(n)} k_\pi^b \cdot k_{K(\pi)}^a,$$

then we have also

$$\Phi_c = \Phi_b \circ \Phi_\delta.$$

(Note that ab and ba have the same distribution, thus $\Phi_{ba} = \Phi_c$.) In order to proceed one has to observe the following symmetrization lemma

$$z \cdot \Phi_c(z) = \Phi_\gamma(z) \cdot \Phi_\delta(z), \tag{3.13}$$

which encodes a nice bijection between

$$NC(n) \longleftrightarrow \bigcup_{1 \le j \le n} NC'(j) \times NC'(n+1-j).$$

Simple algebraic manipulations with these relations leads to

$$\frac{\Phi_c^{<-1>}(z)}{z} = \frac{\Phi_a^{<-1>}(z)}{z} \cdot \frac{\Phi_b^{<-1>}(z)}{z}.$$

This is exactly the multiplicativity of the S-transform, provided we have

$$S_a(z) = \frac{\Phi_a^{<-1>}(z)}{z}. \tag{3.14}$$

That this is indeed the case follows from our relation between the moment and the cumulant generating power series from the last lecture. There we showed that for

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

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

we have

$$C[zM(z)] = M(z). \tag{3.15}$$

Note that in our present notation we have

$$M(z) = 1 + \psi(z)$$

and

$$C(z) = 1 + \Phi(z),$$

so that (3.15) reads as

$$\Phi[z(1 + \psi(z)] = \psi(z),$$

or, equivalently,

$$\psi^{<-1>}(z) \cdot (1 + z) = \Phi^{<-1>}(z).$$

But this gives for the S-transform

$$S(z) = \psi^{<-1>}(z) \cdot \frac{1+z}{z} = \frac{1}{z} \cdot \Phi^{<-1>}(z).$$

Department of Mathematics and Statistics, Queen's University, Kingston, ON K7L 3N6, Canada

E-mail address: speicher@mast.queensu.ca