Internship report:
The S-transform and the boxed convolution in free probability

Timothé Ringeard

February-July 2023

Norwegian University of Science and Technology (NTNU), Trondheim,
under the supervision of Kurusch Ebrahimi-Fard

Contents

1 Free probability **3**
1.1 Motivation and first definitions 3
1.2 Cumulants 4

2 The operator-valued case **6**
2.1 Combinatorics of non-crossing partitions and Catalan pairs 6
2.2 Series of multilinear functions 9
2.3 The operator-valued case of free probability 12

3 Twisted multiplicativity of the S-transform **13**
3.1 The boxed convolutions 13
3.2 The twisted multiplicativity 14

Introduction

The goal of this report is to present the mathematical research that I have done in Trondheim between February and July 2023, under the supervision of Kurusch Ebrahimi-Fard, that I thank a lot. We met weekly, he was always very respondant to all my requests, and often sent me links to online-talks. In particular, I followed the ACPMS talk every two weeks. In addition to the life in Trondheim, I had the chance to visit the mathematics department at the University of Tromsø, Norway, where I was welcomed by Cordian Riener. I also met with a few researchers and PhD students, and I made a short talk.

My internship was mathematically divided into two parts. I first learned about the general theory of Hopf algebras, and in particular the Faà di Bruno Hopf algebra, the Connes-Moscovici Hopf algebra, and the Chen-Fliess Hopf algebra. I tried to make parallels between the two latters, because the last one is still not completely understood. I won't talk more about this subject in this report. The second part of my internship was the study of free probabilities, and is the object of this report.

Free probability can be considered as a subdomain of probability, but it has historically arised in the study of operator algebras. Most of the objects in this field were invented by one person, Dan Voiculescu, allowing him to solve several problems in operator algebra around the year 1983. However, a connection was rapidly made between his concepts and probability, around 1990, and the field of free probability was born. See [13] for one of the early papers.

The defining feature of free probability is that there is a alternative notion of indenpendance between random variable, that is called freeness. Even if this freeness is quite complicated to define, it has a lot of properties in common with the independance, and also has a lot of combinatorial properties in itself. For example, there is also a Central Limit Theorem in free probability, where the gaussian distribution is replaced with the semi-circle distribution. As it is presented in this report, freeness is very algebraic: The fact that two random vairables a and b are free is is a statement about the expectation of polynomials in a and b.

Apart from problems in operator algebras, free probability was proven useful in the field of random matrices. Indeed, for several models of random matrices (for example GUE), the asymptotic distribution of eigenvalues are free for independant sequences of matrices. More precisely, if $A_n, B_n, n \geq 1$ are independant GUE, then at the limit $n \to \infty$, it is known that the distribution of eigenvalues are almost surely two both semi-circle laws, one for (A_n), and one for (B_n). And now, the freeness rule gives the distribution of eigenvalues of the sequence of matrices $(A_n B_n)$, for example. The reader is refered to [7] for more details.

This report focuses on a generalisation of free probability spaces, called operator-valued, where the expectation takes values in a arbitrary, possibly non-commutative algebra instead of a base field. Even if this domain has, to my knowledge, no application, it is still very useful to understand the combinatorics behind the scalar-valued case.

To be more precise, this report focused on the folliowing problem: Given two elements a and b in a operator-valued non-commutative probability space, how to find the distribution of the element ab ? This problem has a well-known answer in the scalar-valued case, using the so-called S-transform and the boxed convolution operation. In the operator-valued case, a generalisation of the S-transform was given (see [10] and [4]) and the problem was essentially solved. My contribution was to give a generalisation of the boxed convolution operation, giving a second answer to the problem.

Let me finally give a short outline of the report.

Section 1 is an introduction to the field of free probabilities. It tries to motivate it from classical probability, and introduces the main objects. It is slightly biased towards the operator-valued case, which in my opinion captures best the combinatorial structures behind it, and is expanded in the rest of the report. The central object in this section is probably the cumulant map, which is very useful to describe free variables thanks to Theroem 1.10.

Section 2 describes a generalisation of non-commutative probability spaces, where the expectation takes values in a general algebra. In order to do so, Catalans pairs are introduced to study general bijections between Catalan-type objects, in this paper between binary planar trees and non-crossing partitions. Multilinear function series are also introduced, and finally the operator-valued probability spaces are defined.

Section 3 aims to give a (new) proof of Theorem 3.6 regarding the twisted multiplicativity of the S-tranform, based on the combinatorial objects constructed in the previous section, and on four boxed convolution operations.

In the whole report, let us fix a field $\mathbb{K}$ of zero characteristic. It is not restrictive to think $\mathbb{K} = \mathbb{R}$.

1 Free probability

1.1 Motivation and first definitions

In this section, we try to motivate free probability from known concepts of classical probability. Let us consider the two following principles:

1. Random variables are elements in a large space, that will be called $\mathcal{A}$.
2. Topology is not significant: the functions of a random variable $a \in \mathcal{A}$ are only polynomials a.

Given the second principle, it is natural to assume that the space $\mathcal{A}$ is an algebra. With a few additionnal choices, we arrive to the following definition for the space of random variables $\mathcal{A}$.

Definition 1.1.
1. A *non-commutative probability space* is a pair $(\mathcal{A}, \mathbb{E})$, where $\mathcal{A}$ is an unital algebra over $\mathbb{K}$ and $\mathbb{E} : \mathcal{A} \to \mathbb{K}$ is a linear map such that $\mathbb{E}(1) = 1$. Elements $a \in \mathcal{A}$ are called *(non-commutative) random variables.*
2. For an element $a \in \mathcal{A}$, its *distribution* is the collection of its moments $(\mathbb{E}(a^n))_{n \in \mathbb{N}}$, or alternatively, the formal series, called the *moment series* of a:

$$M^a(x) := \sum_{n \geq 0} \mathbb{E}(a^n) x^n$$

Example 1.2.
1. Let us consider a classical probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then the space $\mathcal{A} = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, together with the usual expectation map $\mathbb{E}(a) = \int_\Omega a \, \mathrm{d}\mathbb{P}$, form a non-commutative probability space.
2. In the same setup, one could also take $\mathcal{A} = \bigcap_{p>1} L^p(\Omega, \mathcal{F}, \mathbb{P})$.

Note that in the usual definition, the distribution of a random variable a is determined by the values taken by $\mathbb{E}(f(a))$ for all f in a certain set of function, for example continuous bounded functions. It is still the case here, and since we don't have topology, the only choice of functions is polynomials. And indeed, the distribution of a is determined by

$$\mathbb{E}(P(a)), \quad P \in \mathbb{K}[X]$$

The name "non-commutative" is given because the algebra $\mathcal{A}$ can be non-commutative. However, if it is commutative, as in two examples above, one can define the notion of independance. Note that it will correspond to the usual notion of independance in the two examples above.

Definition 1.3. (Independance) Let $(\mathcal{A}, \mathbb{E})$ be a non-commutative probability space, and assume that $\mathcal{A}$ is commutative.

1. A collection of subalgebras $(\mathcal{A}_i)_{i \in I}$ of $\mathcal{A}$ is said to be *independant* if for any $i_1, \ldots, i_n \in I$ that are all different, and any $a_{i_k} \in \mathcal{A}_{i_k}$ for $1 \leq k \leq n$,

$$\mathbb{E}(a_{i_1} \cdots a_{i_n}) = \mathbb{E}(a_{i_1}) \cdots \mathbb{E}(a_{i_n}).$$

2. A collection of random variables $(a_i)_{i \in I} \subseteq \mathcal{A}$ is said to be *independant* if the collection of generated algebras $(\langle a_i \rangle)_{i \in I}$ is independant.

Note that if $\mathbb{E}$ is given on independant subalgebras $(\mathcal{A}_i)_{i\in I}$, then $\mathbb{E}$ is completely determined by the independance rule on the subalgebra generated by all the $\mathcal{A}_i$ for $i \in I$. In general the non-comutativity makes independance not very meaningful, but one can define a meaningful alternative: freeness.

Definition 1.4. (Freeness) Let $(\mathcal{A}, \mathbb{E})$ be a non-commutative probability space.

1. A collection of subalgebras $(\mathcal{A}_i)_{i\in I}$ of $\mathcal{A}$ is said to be *free* if for any $i_1, \ldots, i_n \in I$ such that $i_{k-1} \neq i_k$, and any $a_{i_k} \in \mathcal{A}_{i_k}$ such that $\mathbb{E}(a_{i_k}) = 0$,
$$\mathbb{E}(a_{i_1} \cdots a_{i_n}) = 0.$$
2. A collection of random variables $(a_i)_{i\in I} \subseteq \mathcal{A}$ is said to be *free* if the collection of generated algebras $(\langle a_i \rangle)_{i\in I}$ is free.

Note that $\mathbb{E}$ is again completely determined by the freeness rule on the subalgebra generated by $(\mathcal{A}_i)_{i\in I}$, but it is a bit harder to see.

Example 1.5. Assume a, b are free random variables. Define $\overline{a} := a - \mathbb{E}(a)$ and $\overline{b} = b - \mathbb{E}(b)$, so that $\mathbb{E}(\overline{a}) = \mathbb{E}(\overline{b}) = 0$, and $\overline{a} \in \langle a \rangle$, $\overline{b} \in \langle b \rangle$. Then

$$\begin{aligned}\mathbb{E}(ab) &= \mathbb{E}((\overline{a} + \mathbb{E}(a))(\overline{b} + \mathbb{E}(b))) \\ &= \mathbb{E}(\overline{a}\overline{b}) + \mathbb{E}(a)\mathbb{E}(\overline{b}) + \mathbb{E}(b)\mathbb{E}(\overline{a}) + \mathbb{E}(a)\mathbb{E}(b) \\ &= \mathbb{E}(\overline{a}\overline{b}) + \mathbb{E}(a)\mathbb{E}(b) \\ &= \mathbb{E}(a)\mathbb{E}(b)\end{aligned}$$

since $\mathbb{E}(\overline{a}\overline{b}) = 0$ by the freeness rule. Similarly,

$$\begin{aligned}\mathbb{E}(aba) &= \mathbb{E}(b)\mathbb{E}(a^2), \\ \mathbb{E}(abab) &= \mathbb{E}(a^2)\mathbb{E}(b)^2 + \mathbb{E}(b^2)\mathbb{E}(a)^2 - \mathbb{E}(a^2)\mathbb{E}(b^2).\end{aligned}$$

1.2 Cumulants

Let us fix a non-commutative probability space $(\mathcal{A}, \mathbb{E})$ in this section.

The distribution of an element a in a non-commutative probability space can be given by its sequence of moments, but we will see that it can equivalently be given by another sequence, its *cumulants*. Moreover, the latter behave much better with respect to freeness. The construction of cumulants requires non-crossing partitions.

Definition 1.6. (Non-crossing partitions)

1. For $n \geq 0$, let SP_n denote the set of partitions on $[n] = \{1, \ldots, n\}$, and $\mathrm{NCP}_n \subseteq \mathrm{SP}_n$ is the set of non-crossing partitions of $[n]$. The latter is a set partition $P = \{V_1, \ldots, V_k\}$ of $[n]$ without elements $a < b < c < d$ in $[n]$ such that $a, c \in V_i$ and $b, d \in V_j$, and $i \neq j$. Let $\mathrm{SP} := \bigcup_{n\geq 0} \mathrm{SP}_n$ and $\mathrm{NCP} := \bigcup_{n\geq 0} \mathrm{NCP}_n$. By convention $\mathrm{NCP}_0 = \{\emptyset\} = \mathrm{SP}_0$ consists of the empty partition. If $P \in \mathrm{SP}_n$, we write $|P| = n$.
2. Given $P = \{V_1, \ldots, V_k\} \in \mathrm{SP}_n$, and a sequence of scalars $(a_n)_{n\geq 0} \subseteq \mathbb{K}$, one defines
$$a_P := \prod_{1\leq i\leq k} a_{|V_i|}$$

3. Similarly, if for all $m \geq 0, f_m : \mathcal{A}^m \to \mathbb{K}$ is a multilinear function, and $P = \{V_1, \dots, V_k\} \in \mathrm{SP}_n$, define
$$f_P(x_1, \dots, x_n) := \prod_{1 \leq i \leq k} f_{|V_i|}(x_{c^i_1}, \dots, x_{c^i_{j_i}}),$$
where $V_i = \{c^i_1, \dots, c^i_{j_i}\}$ and $c^i_1 < \dots < c^i_{j_i}$. The function f_P is then a multilinear function $\mathcal{A}^n \to \mathbb{K}$.

Example 1.7. The partition $P = \{\{1,5\},\{2\},\{3,4\}\}$ is diagramatically represented as , and it is non-crossing. For a sequence (a_n), $a_P = a_2^2 a_1$, and for a mulitilinear function f, $f_P(x_1, \dots, x_5) = f(x_1 x_5) f(x_2) f(x_3 x_4)$. Note that everything is well defined thanks to the commutativity of $\mathbb{K}$.

Let us now define the cumulant map and the cumulants of an random variable.

Definition 1.8. (Cumulants)

1. Let $(\mathcal{A}, \mathbb{E})$ be a non-commutative probability space. Define $\kappa = (\kappa_n)_{n \geq 0}$ to be a the sequence of multilinear functions, where $\kappa_n : \mathcal{A}^n \to \mathbb{K}$, such that for all $n \geq 0, x_1, \dots, x_n \in \mathcal{A}$,
$$\mathbb{E}(x_1 \cdots x_n) = \sum_{P \in \mathrm{NCP}_n} \kappa_P(x_1, \dots, x_n).$$
The maps κ_n are called the *cumulant maps.*

2. Given a random variable $a \in \mathcal{A}$, define its *cumulants* to be the scalars
$$\kappa^a_n := \kappa_n(a, \dots, a).$$
Define finally its cumulant series to be
$$K^a(x) = \sum_{n \geq 1} \kappa^a_n x^n.$$
This series is sometimes called the *R-transform* of a.

It turns out that for a fixed random variable a, the moment series and the cumulant series are related through Lagrange inversion, which is very famous in combinatorics (see [12]):

Proposition 1.9. *Let $a \in \mathcal{A}$. Then it holds that*
$$M^a(x) = (1 + K^a)(x M^a(x)).$$

The cumulants are in fact a central piece in the study of free probability, mostly because of the following theorem, which is orally refered to as "freeness $\iff$ mixed cumulants vanish".

Theorem 1.10. *Let $a, b \in \mathcal{A}$. Then a and b are free if and only if, for any $n \geq 1$, $c_1, \dots, c_n \in \{a, b\}$ that are not all equal,*
$$\kappa(c_1, \dots, c_n) = 0$$

Thanks to these cancelations of cumulants, the distribution of a sum of free elements is very simple to compute:

Proposition 1.11. *Let $a, b \in \mathcal{A}$ be two free random variables. Then*
$$K^{a+b} = K^a + K^b$$

Proof. Let $n \geq 1$. As κ_n is a n-linear map,

$$\begin{aligned}\kappa_n^{a+b} &= \kappa_n(a+b, a+b, \ldots, a+b)\\ &= \sum_{(c_1,\ldots,c_n)\in\{a,b\}^n} \kappa_n(c_1,\ldots,c_n)\\ &= \kappa_n(a,a,\ldots,a) + \kappa_n(b,b,\ldots,b)\\ &= \kappa_n^a + \kappa_n^b.\end{aligned}$$

□

The R-transform is, however, not the best object to study the distribution of the element ab, where a and b are free elements. Instead, one uses the S-transform:

Definition 1.12. Let $a \in \mathcal{A}$, and assume that $\mathbb{E}(a) \neq 0$. The one can define its S-transform S^a to be

$$S^a := \frac{1}{x}(xK^a)^{\circ -1},$$

where $\cdot^{\circ -1}$ is the inverse map for the composition of series. Note that xK^a is invertible since $K^a = \mathbb{E}(a) + xQ$ for some series Q, and $\mathbb{E}(a) \neq 0$ by assumption.

Remark 1.13. The series S^a still characterises the distribution of a. Indeed, one can find back its cumulant series by $K^a = \frac{1}{x}(xS^a)^{\circ -1}$.

Proposition 1.14. *Let $a, b \in \mathcal{A}$ be two free random variables such that $\mathbb{E}(a) \neq 0$, $\mathbb{E}(b) \neq 0$. Then the S-transform of ab factorises:*

$$S^{ab} = S^a S^b.$$

Note that $\mathbb{E}(ab) = \mathbb{E}(a)\mathbb{E}(b) \neq 0$ so S^{ab} is well-defined.

The proof is a bit too long to fit in this report, but can be found in [8]. The main goal of this report is to extend this result to a more general setting, that is, to the operator-valued case, defined in the next section.

2 The operator-valued case

Before defining operator-valued (non-commutative) probability spaces, and especially the important cumulant map, one first need some results about the combinatorics of non-crossing partitions and binary trees. Moreover, the formalism given by multilinear functions series fits the setting very well. The reader will therefore find two preliminary sections before the actual construction of operator-valued probability spaces.

2.1 Combinatorics of non-crossing partitions and Catalan pairs

Definition 2.1. Let $C = (C_n)_{n\geq 0}$ be a collection of finite sets, and assume that $C_0 = \{1_C\}$. A *Catalan pair* (C, f) over C is defined in terms of a *Catalan map* $f : C \times C \to C$ inducing bijections $\bigsqcup_{k=0}^{n} (C_k \times C_{n-k}) \to C_{n+1}$ for all $n \geq 0$.

Remark 2.2. 1. If (C, f) is a Catalan pair, then $|C_n| = \frac{1}{n+1}\binom{2n}{n}$ is the n-th Catalan number.

2. A Catalan map f gives a way to construct all elements of C from the element of $C_0 = \{1_C\}$. Indeed, from Definition 2.1 it is clear that any element in $\bar{C} := (C_n)_{n>0}$ can be written as iterated expressions of the map f and 1_C such as, for example, $f(1_C, 1_C)$ or $f(f(1_C, f(1_C, 1_C)),\ f(1_C, 1_C))$. Moreover, these expressions are unique. Both of these facts follow from simple inductions.
3. Inductions are particularly efficient in the context of Catalan pairs: if a property is true for the element 1_C and preserved by the Catalan map f (in the sense that if a, b both satisfy the property so does $f(a,b)$), then the property is true for all elements.

Lemma 2.3. *Let (C,f) and (C',f') be Catalan pairs. Then there exists a unique isomorphism of Catalan pairs, i.e., there exists a unique bijection $\varphi : C \to C'$ such that $\varphi(1_C) = 1_{C'}$ and for all $x, y \in C$, $\varphi(f(x,y)) = f'(\varphi(x), \varphi(y))$.*

Proof. The map φ can be constructed on C_n by induction on n, as does unicity. ☐

We are going to define Catalan structures on both planar binary trees and non-crossing partitions. The motivation is that using the former – more obvious and visual – Catalan structure can be helpful in computations involving the latter together with its more hidden Catalan structure. Let us first recall some notations for planar binary trees.

Definition 2.4. 1. A *planar binary tree* is a rooted tree where each internal vertex has exactly two children, a right one and a left one. More formally, for $n \geq 0$, the set of planar binary trees with n internal vertices is

$$\begin{aligned} Y_0 &= \{|\} \\ Y_n &= \bigcup_{\substack{k+l=n-1 \\ k,l\geq 0}} \{\ {}^{\sigma}\curlyvee^{\tau}\ \mid \sigma \in Y_k,\ \tau \in Y_l\} \qquad n > 0, \end{aligned}$$

where $|$ is the so-called empty tree, i.e., with no internal vertex. Let $Y := \bigcup_{n\geq 0} Y_n$ and $\overline{Y} := \bigcup_{n\geq 1} Y_n$. Denoting for a tree $\tau \in Y_n$ the set of internal vertices by $V(\tau)$, we write $|\tau| := |V(\tau)| = n$. We will abusively call vertex an internal vertex.

2. For a tree $\tau \in \overline{Y}_n$, the "left-to-right" ordering of its vertices is the linear order $o(\tau) = \{v_1 < \cdots < v_n\}$ on its vertices $v_i \in V(\tau)$, defined recursively by

$$o\left({}^{\sigma}\curlyvee^{\tau}_{x} \right) = \{o(\sigma) < x < o(\tau)\}.$$

We can therefore identify the vertices of τ with the set $\{1, \ldots, |\tau|\}$.

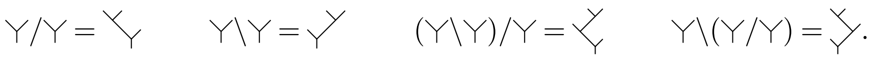
As an example, we consider the planar binary tree on the lefthand side of Figure 2.1, where only internal vertices are shown. The "left-to-right" ordering $o(\tau)$ is indicated by the labelling of the vertices of τ.

Definition 2.5. We define two grafting operations, called over and under, on trees

$$/, \backslash : Y \times Y \to Y,$$

where τ/σ (read τ over σ) is the tree obtained by adding τ as the left-child of the left-most internal vertex of σ, and $\tau\backslash\sigma$ (read τ under σ) is the tree obtained by adding σ as the right-child of the right-most internal vertex of τ.

For example:



Remark 2.6. For $\tau \in \overline{Y}$, we define the substitution $\mu_\tau : \overline{Y}^{|\tau|} \to \overline{Y}$, where the tree $\mu_\tau(\sigma_1, \ldots, \sigma_{|\tau|})$ is obtained from the tree τ by replacing each vertex $v_i \in V(\tau)$ with the tree σ_i. Note that the order on the vertices is $o(\mu_\tau(\sigma_1, \ldots, \sigma_{|\tau|})) = \{o(\sigma_1) < \cdots < o(\sigma_{|\tau|})\}$, where we identify the vertices of σ_i with their image in $\mu_\tau(\sigma_1, \ldots, \sigma_{|\tau|})$. We remark that this operation defines a set-operadic structure on $\overline{Y}$, which allows to write $\mu_\tau(\sigma_1, \ldots, \sigma_{|\tau|})$ in operadic notation

$$\tau \circ (\sigma_1, \ldots, \sigma_{|\tau|}) \in \overline{Y}. \tag{1}$$

The operation μ_τ is not used in the rest of the report, but still is important in omitted proofs. We refer the reader to Frabetti's work [6] for more details, and also to [4].

The definition of planar binary trees implies the following statement.

Lemma 2.7. *Define the map* $\mathrm{Cat}_Y : Y \times Y \to Y$, $(\sigma, \tau) \mapsto {}^{\sigma}\mathsf{Y}^{\tau}$. *Then* (Y, Cat_Y) *is a Catalan pair.*

Let us now add some notations for non-crossing partitions.

Definition 2.8.

1. For $P, Q \in \mathrm{NCP}$, define $P * Q$ to be the non-crossing partition resulting from concatenation of the two partitions.
2. For $P, Q \in \mathrm{NCP}$, define the right-merging $P \underline{*} Q$ to be the non-crossing partition obtained from $P * Q$ by merging the last block of P and the last block of Q into a single block.

Observe that both concatenation and right-merging of partitions are associative but non-commutative operations. As an example for $P * Q$ and $P \underline{*} Q$, consider the two partitions  and  , then

Lemma 2.9. *Define the map* $\mathrm{Cat}_{\mathrm{NCP}} : \mathrm{NCP} \times \mathrm{NCP} \to \mathrm{NCP}$,

$$(P, Q) \mapsto P * (|\underline{*} Q). \tag{2}$$

Then $(NCP, \mathrm{Cat}_{\mathrm{NCP}})$ *is a Catalan pair.*

Combining this with Lemma 2.7 and Lemma 2.3 gives:

Theorem 2.10. *There exists a unique isomorphism* $\varphi : Y \to \mathrm{NCP}$ *of Catalan pairs.*

Remark 2.11. We give an explicit description of the bijection $\varphi : Y \to \mathrm{NCP}$, which is graded in the sense that $\varphi(Y_n) = \mathrm{NCP}_n$, and was already considered in [9]. Let us construct $\varphi(\tau)$ for $\tau \in Y_n$, $n \geq 1$. First, number the vertices of τ by elements from $[n]$ using the left-to-right order. Then $\varphi(\tau) \in \mathrm{NCP}(n)$ is the non-crossing partition on $[n]$ where the block of $i \in [n]$ is the set of all vertices connected to i by "right arms". More precisely, define $R(x,y)$ to be the binary relation on $[n]$: "x is the right child of y". Then $\varphi(\tau)$, seen as an equivalence relation, is the reflexive-symmetric-transitive closure of R. Figure 1 shows an example, and Example 2.12 shows the first few values of φ on small trees.

To verify that the constructed map φ is the same as the one in Definition 2.10, it suffices to notice that $\varphi(|) = \emptyset$, and that for all trees σ, τ in Y, the inductive formula holds

$$\varphi({}^{\sigma}\mathsf{Y}^{\tau}) = \varphi(\sigma) * (|\underline{*}\varphi(\tau)). \tag{3}$$

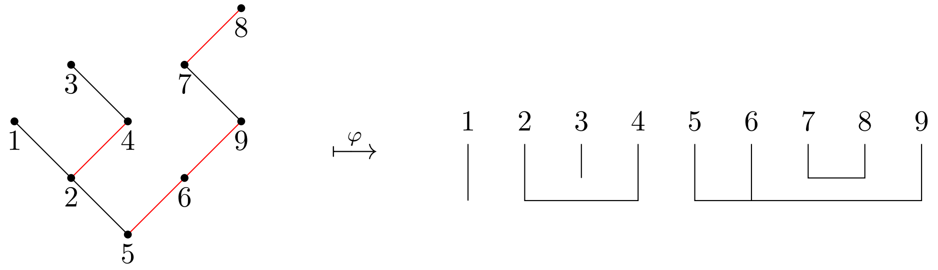


Figure 1: An example value of the function φ. Only internal vertices are shown in the tree on the left.

Example 2.12. Here are the first values of φ:

The following lemma shows the relation between the inductive formula (3) and the two grafting operations on trees, over and under, given in Definition 2.5.

Lemma 2.13. *For* $\sigma, \tau \in Y$,

$$\varphi(\sigma/\tau) = \varphi(\sigma) * \varphi(\tau) \tag{4}$$
$$\varphi(\sigma\backslash\tau) = \varphi(\sigma)\underline{*}\varphi(\tau). \tag{5}$$

Proof. The first equality can be shown by induction on the size of τ. If $\tau = |$, the equality is clear. Otherwise, let us write $\tau = {}^{\tau_1}\curlyvee^{\tau_2}$ with $|\tau_1| < |\tau|$. Then $\sigma/\tau = {}^{\sigma/\tau_1}\curlyvee^{\tau_2}$ so

$$\begin{aligned}
\varphi(\sigma/\tau) &= \varphi(\ {}^{\sigma/\tau_1}\curlyvee^{\tau_2}\ ) \\
&\overset{(3)}{=} \varphi(\sigma/\tau_1) * (|\underline{*}\varphi(\tau_2)) \\
&\overset{\text{ind.}}{=} \varphi(\sigma) * \varphi(\tau_1) * (|\underline{*}\varphi(\tau_2)) \\
&= \varphi(\sigma) * \varphi(\ {}^{\tau_1}\curlyvee^{\tau_2}\ ) \\
&= \varphi(\sigma) * \varphi(\tau).
\end{aligned}$$

The second equality follows by a symmetry argument. □

2.2 Series of multilinear functions

We follow the terminology of Dykema in [1]. Recall that $\mathbb{K}$ denotes the base field of characteristic zero over which all algebraic structures are defined.

Definition 2.14. Let B be a (non necessarily commutative) $\mathbb{K}$-algebra with unit 1. A sequence of multilinear functions $f = (f_n)_{n\geq 0}$, $f_n : B^{\otimes n} \to B$ (with $B^{\otimes 0} = \mathbb{K}$) will be called a multilinear series. The set of these multilinear series will be called $\mathrm{Mult}[[B]]$.

1. We can multiply such series by defining for $f, g \in \mathrm{Mult}[[B]]$:
$$(f \cdot g)_n(x_1, \dots, x_n) := \sum_{k=0}^{n} f_k(x_1, \dots, x_k) g_{n-k}(x_{k+1}, \dots, x_n). \tag{6}$$
$(\mathrm{Mult}[[B]], \cdot)$ is a monoid, with unit $1 := (\delta_{n,0})$. We will often write fg for $f \cdot g \in \mathrm{Mult}[[B]]$.
2. We can compose such series by defining for $f, g \in \mathrm{Mult}[[B]]$:
$$(f \circ g)_n(x_1, \dots, x_n) := \sum_{\substack{n=k_1+\cdots+k_l \\ k_i \geq 1}} f_l(g_{k_1}(x_1, \dots, x_{k_1}), \dots, g_{k_l}(x_{n-k_l+1}, \dots, x_n)). \tag{7}$$
$(\mathrm{Mult}[[B]], \circ)$ is a monoid, with unit $I := (\delta_{n,1}\,\mathrm{Id})$.

Note that if $f \cdot g = 0$, then either $f = 0$ or $g = 0$. This is, however, not the case for the composition $\circ$.

We define now groups with respect to multiplication (6) and composition (7) of elements in $\mathrm{Mult}[[B]]$:
$$G_B^{\mathrm{inv}} := \{f \in \mathrm{Mult}[[B]] \mid f_0 \in B^{\times}\} \subset \mathrm{Mult}[[B]]$$
$$G_B^{\mathrm{dif}} := \{f \in \mathrm{Mult}[[B]] \mid f_0 = 0, f_1 \in \mathrm{GL}(B)\} \subset \mathrm{Mult}[[B]],$$
where $B^{\times}$ is the set of invertible elements of B. We note that replacing $f_0 \in B^{\times}$ by $f_0 = 1$ and $f_1 \in \mathrm{GL}(B)$ by $f_1 = \mathrm{Id}$ would not change much. We also define
$$G_B^I := I \cdot G_B^{\mathrm{inv}}.$$
Alternatively,
$$G_B^I := \{f \in G_B^{\mathrm{dif}} \mid \forall n, f_n(x_1, x_2, \dots, x_n) = x_1 f_n(1, x_2, \dots, x_n)\}.$$
For $f \in G_B^{\mathrm{inv}}$, we will denote by f^{-1} the multiplicative inverse of f. The compositional inverse of $f \in G_B^{\mathrm{dif}}$ will be denoted by $f^{\circ -1}$. Regarding expressions involving both composition and multiplication, $\circ$ and $\cdot$, the former operation will be given precedence, i.e., $f \cdot g \circ h \cdot k = f \cdot (g \circ h) \cdot k$.

Lemma 2.15. *Composition and multiplication, $\circ$ and $\cdot$, are both associative and non-commutative. Moreover, composition is right-distributive over multiplication, that is, for any $f, g, h \in \mathrm{Mult}[[B]]$*
$$(f \cdot g) \circ h = (f \circ h) \cdot (g \circ h).$$

Proof. See Dykema [1] for details. □

Lemma 2.16. *The set G_B^I forms a group with respect to composition $\circ$.*

Proof. Let $f, g \in G_B^I \subseteq G_B^{\mathrm{dif}}$. As $I = I \cdot 1 \in G_B^I$, it suffices to show that $f \circ g$ and $f^{\circ -1}$ (calculated in G_B^{dif}) are elements of G_B^I. Let us write $f = IF$ and $g = IG$, for $F, G \in G_B^{\mathrm{inv}}$. Then
$$\begin{aligned} f \circ g = (IF) \circ (IG) &= IG \cdot (F \circ (IG)) \\ &= I \cdot (G \cdot (F \circ (IG))) \in G_B^I. \end{aligned}$$

Moreover,

$$I = f \circ f^{\circ -1} = (IF) \circ f^{\circ -1} = f^{\circ -1} \cdot F \circ f^{\circ -1}.$$

But $F \circ f^{\circ -1} \in G_B^{\mathrm{inv}}$ since $F \in G_B^{\mathrm{inv}}$, so

$$f^{\circ -1} = I(F \circ f^{\circ -1})^{-1} \in G_B^I.$$

□

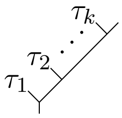
Definition 2.17. (S-transform) Let $f \in G_B^I$. Then, as $f^{\circ -1} \in G_B^I$, we can write

$$f^{\circ -1} = I \cdot S_f.$$

$S_f \in G_B^{\mathrm{inv}}$ is called the S-transform of f.

From one or several multilinear series, one can construct other multilinear maps:

Definition 2.18. For $f, g \in \mathrm{Mult}[[B]]$ and a planar binary tree $\tau \in Y_n$, which we write as

$\tau =$ 

we define:

1. $f_\tau : B^{\otimes n} \to B$ inductively as follows:

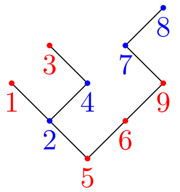
$$\begin{cases} f_| &= 1 \in B \\ f_\tau(x_1, \ldots, x_n) &= f_k\Big(f_{\tau_1}(x_1, \ldots, x_{j_1 - 1})x_{j_1}, \ \ldots \ , \ f_{\tau_k}(x_{j_{k-1}+1}, \ldots, x_{n-1})x_n\Big) \in B \end{cases}$$

2. $(f \cup g)_\tau : B^{\otimes n} \to B$ inductively as follows:

$$\begin{cases} (f \cup g)_| &= 1 \in B \\ (f \cup g)_\tau(x_1, \ldots, x_n) &= g_k\Big((g \cup f)_{\tau_1}(x_1, \ldots, x_{j_1 - 1})x_{j_1}, \ \ldots \ , \ (g \cup f)_{\tau_k}(x_{j_{k-1}+1}, \ldots, x_{n-1})x_n\Big) \in B \end{cases}$$

   We remark that $(f \cup g)_\tau$ is defined by alternating between f and g. Note that $(f \cup g)$ as a whole (without indices) is not defined, in particular, it is not an element of $\mathrm{Mult}[[B]]$.

Example 2.19. Suppose

$\tau =$  ,

where only internal vertices are shown. For $f, g \in \mathrm{Mult}[[B]]$, we have

$$f_\tau(x_1, \ldots, x_9) = f_3\Big(f_2(f_1(x_1)x_2, f_1(x_3)x_4)x_5, \ x_6, \ f_2(x_7, x_8)x_9\Big)$$

$$(f \cup g)_\tau(x_1, \ldots, x_9) = g_3\Big(f_2(g_1(x_1)x_2, g_1(x_3)x_4)x_5, \ x_6, \ f_2(x_7, x_8)x_9\Big)$$

$$(g \cup f)_\tau(x_1, \ldots, x_9) = f_3\Big(g_2(f_1(x_1)x_2, f_1(x_3)x_4)x_5, \ x_6, \ g_2(x_7, x_8)x_9\Big).$$

2.3 The operator-valued case of free probability

Suppose B a unital algebra over the field $\mathbb{K}$ of zero characteristic. One can define a more general version of non-commutative probability space, where the expectation $\mathbb{E}$ takes values in B instead of K. These space will be called operator-valued, or B-valued. We give here the basics definitions, the reader can see [7] for more details and background.

Definition 2.20. 1. A *B-valued non-commutative probability space* is a pair $(\mathcal{A}, \mathbb{E})$, where $\mathcal{A}$ is a unital algebra over B and $\mathbb{E} : \mathcal{A} \to B$ is a unital linear map, $\mathbb{E}(1) = 1$, such that for all $x, y \in B$ and for all $a \in \mathcal{A}$,

$$\mathbb{E}(xay) = x\mathbb{E}(a)y. \tag{8}$$

We say that $\mathbb{E}$ is B-balanced.

Elements $a \in \mathcal{A}$ are called *(non-commutative) random variables.* Their distribution is usually given by their *moment series* $M^a \in \mathrm{Mult}[[B]]$ defined by

$$M^a(x_1, \ldots, x_n) = \mathbb{E}(ax_1ax_2a \cdots ax_na).$$

2. (Freeness) A collection of B-subalgebras $(\mathcal{A}_i)_{i \in I}$ of $\mathcal{A}$ is said to be *free* if $\mathbb{E}(a_1 \cdots a_n) = 0$ when the following conditions holds:

   - $n \geq 1$,
   - $\mathbb{E}(a_i) = 0$ for $1 \leq i \leq n$,
   - $a_i \in \mathcal{A}_{j_i}$ for $1 \leq i \leq n$,
   - $j_i \neq j_{i+1}$ for $1 \leq i < n$.

   A collection of elements $(a_i)_{i \in I} \subseteq \mathcal{A}$ is said to be free if their respective generated B-subalgebras are free. In this paper, for simplicity we shall focus on the case of two free elements $a, b \in \mathcal{A}$.

Next we define we operator-valued cumulants in terms of operator-valued moment-cumulant relations using tree-indexed series of multilinear functions (Definition 2.18).

Definition 2.21. 1. (Cumulants) The series of operator-valued cumulants $\kappa \in \mathrm{Mult}[[\mathcal{A}]]$ is defined by requiring that for all $n \geq 0$, $a_1, \ldots, a_n \in \mathcal{A}$,

$$\mathbb{E}(a_1 \cdots a_n) = \sum_{\tau \in Y_n} \kappa_\tau(a_1, \ldots, a_n).$$

If $a \in \mathcal{A}$, define its corresponding cumulant series $K^a \in \mathrm{Mult}[[B]]$ by

$$K^a_n(x_1, \ldots, x_n) := \kappa(a, x_1a, \ldots, x_na).$$

We also define $k^a := IK^a \in G^I_B$:

$$k^a_n(x_1, \ldots, x_n) := x_1\kappa(a, x_2a, \ldots, x_na) = \kappa(x_1a, x_2a, \ldots, x_na).$$

Finally, when $\mathbb{E}(a) \in B^\times$, we define the S-transform of the operator-valued random variable a by

$$S^a = S_{k^a},$$

which is well-defined since $\kappa(a) = \mathbb{E}(a) \in B^\times$ implying $K^a \in G^{\mathrm{inv}}_B$ and $k^a \in G^I_B$.

From the definition of κ and the B-balancedness of $\mathbb{E}$, the following can be noticed:

Lemma 2.22. *Let $n \geq 1$, $a_1,\dots,a_n \in \mathcal{A}$, $x_0,\dots,x_n \in B$. Then*

$$\kappa(x_0a_1x_1, a_2x_2, \dots, a_nx_n) = x_0\kappa(a_1, x_1a_2, \dots, x_{n-1}a_n)x_n$$

In the next sections, we fix a B-valued probability space $(\mathcal{A},\mathbb{E})$. The following theorem is the operator-valued version of the very important Theroem 1.10

Theorem 2.23. [11, Thm. 9.4] *Elements $a,b \in \mathcal{A}$ are free if and only if for all $n \geq 1$, $c_1,\dots,c_n \in \{a,b\}$ such that the c_i are not all equal, and for all $x_1,\dots,x_{n-1} \in B$, we have $\kappa_n(c_1x_1, c_2x_2, \dots, c_{n-1}x_{n-1}, c_n) = 0$.*

We are now interested in products $ab \in \mathcal{A}$, with fixed free elements $a,b \in \mathcal{A}$. Let us first give a formula for the cumulants of the $\mathcal{A}$-product ab as a function of the cumulants of a and of b. This formula motivates the definition of one of the boxed convolution operation given in Definition 3.3.

Proposition 2.24. *If $a,b \in \mathcal{A}$ are free, then*

$$k_n^{ab}(x_1,\dots,x_n) = \sum_{\tau \in Y_n} (k^a \cup k^b)_{R(\tau)}(x_1, 1, x_2, 1, \dots, 1, x_n, 1).$$

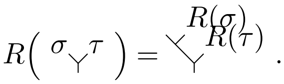
Using definition 3.3, this will be written more concisely as $k^{ab} = k^a \boxed{*}\| \; k^b$.

3 Twisted multiplicativity of the S-transform

3.1 The boxed convolutions

Definition 3.1. Define a map $R : Y \to Y$ recursively by

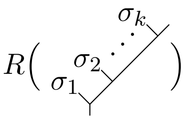
$$R(|) = |$$



Denote

$$Y_{2n}^b := R(Y_n) \quad \text{and} \quad Y^b := R(Y).$$

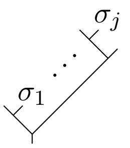
Observe that $Y_{2n}^b \subseteq Y_{2n}$, and that R is injective. Moreover, note that

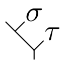


$$= \mathrm{RC}(R(\sigma_1),\dots,R(\sigma_k)).$$

where RC is the rotated comb tree defined by

$$\mathrm{RC}(\sigma_1,\dots,\sigma_j) := \tag{9}$$



Remark 3.2. The set Y^b_{2n}, together with the map $(\sigma,\tau)\mapsto$ , form a Catalan pair.

The following definition proposes several operator-valued boxed-convolution operations, which are meant to generalise both the classical boxed-convolution and the reduced boxed-convolution defined in the scalar-valued case. See Section 18 of the standard reference [8] by A. Nica and R. Speicher. The first boxed convolution operation is also motivated by Proposition 2.24.

Definition 3.3. Let $f,g\in \mathrm{Mult}[[B]]$. Define

$$(f \boxed{*}\!\| \; g)_n(x_1,\dots,x_n) := \sum_{\tau\in Y_n} (f\cup g)_{R(\tau)}(x_1,1,x_2,1,\dots,1,x_n,1) \tag{10}$$

$$(f \;\|\!\boxed{*}\; g)_n(x_1,\dots,x_n) := \sum_{\tau\in Y_n} (f\cup g)_{R(\tau)}(1,x_1,1,x_2,\dots,1,x_n) \tag{11}$$

$$(f \;\overline{\boxed{*}}\; g)_n(x_1,\dots,x_n) := \sum_{\tau\in Y_{n-1}} (g\cup f)_{\curlyvee R(\tau)}(x_1,1,x_2,1,\dots,1,x_n) \tag{12}$$

$$(f \;\|\!\boxed{*}\!\|\; g)_n(x_1,\dots,x_n) := \sum_{\tau\in Y_n} (f\cup g)_{\curlyvee R(\tau)}(1,x_1,1,x_2,1,\dots,1,x_n,1) \tag{13}$$

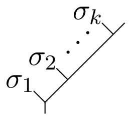
Note that if $f,g\in G^{\mathrm{dif}}_B$, then $f \boxed{*}\!\| \; g$, $f \;\|\!\boxed{*}\; g$ as well as $f \;\overline{\boxed{*}}\; g$ are in G^{dif}_B, but $f \;\|\!\boxed{*}\!\|\; g\in G^{\mathrm{inv}}_B$. Also, if $f,g\in G^I_B$, then $f \boxed{*}\!\| \; g$ and $f \;\overline{\boxed{*}}\; g$ are in G^I_B.

3.2 The twisted multiplicativity

Lemma 3.4. *For $f,g\in\mathrm{Mult}[[B]]$,*

$$(f \boxed{*}\!\| \; g) = g\circ (f \;\overline{\boxed{*}}\; g) \tag{14}$$

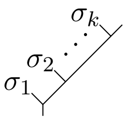
$$(g \;\|\!\boxed{*}\; f) = f\circ ((f \;\|\!\boxed{*}\!\|\; g)I). \tag{15}$$

Proof. These equalities result from the decomposition $\tau =$  for $\tau\in Y$, which implies the unique decomposition $\tau = \mathrm{RC}(\sigma_1,\dots,\sigma_k)$ for $\tau\in Y^b$, thanks to equality 9.

Let us derive (14), the one one being done similarly. Suppose $n\geq 0$. Then

$$(f \boxed{*}\!\| \; g)_n(x_1,\dots,x_n) = \sum_{\tau\in Y_n} (f\cup g)_{R(\tau)}(x_1,1,x_2,1,\dots,1,x_n,1)$$

$$= \sum_{\substack{n=k_1+\cdots+k_j\\ k_i\geq 1}} \sum_{\sigma_i\in Y_{k_i-1}} (f\cup g)_{\mathrm{RC}(R(\sigma_1),\dots,R(\sigma_j))}(x_1,1,\dots,1,x_n,1) \qquad \text{with } \tau =$$



$$= \sum_{\substack{n=k_1+\cdots+k_j\\ k_i\geq 1}} \sum_{\sigma_i\in Y_{k_i-1}} g_j\Big((g\cup f)_{\curlyvee R(\sigma_1)}(x_1,1,\dots,1,x_{k_1}),\ \dots,\ (g\cup f)_{\curlyvee R(\sigma_j)}(x_{n-k_j+1},1,\dots,1,x_n)\Big)$$

$$= \sum_{\substack{n=k_1+\cdots+k_j\\ k_i\geq 1}} g_j\Big((f \;\overline{\boxed{*}}\; g)_{k_1}(x_1,\dots,x_{k_1}),\ \dots,\ (f \;\overline{\boxed{*}}\; g)_{k_j}(x_{n-k_j+1},\dots,x_n)\Big)$$

$$= \Big(g\circ (f \;\overline{\boxed{*}}\; g)\Big)_n(x_1,\dots,x_n).$$

□

For multilinear series in $I\cdot\mathrm{Mult}[[B]]$, the same boxed convolutions $f \boxed{*}\!\| \; g$ and $f \;\|\!\boxed{*}\; g$ have another factorisation:

Lemma 3.5. *For $f,g \in I\cdot\mathrm{Mult}[[B]]$,*

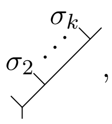
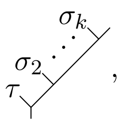
$$(f \boxed{*}\!\| \; g) = (f \overset{=}{\boxed{*}} g)\cdot(f \;\|\!\boxed{*}\!\|\; g) \tag{16}$$

$$(g \;\|\!\boxed{*}\; f) = (f \;\|\!\boxed{*}\!\|\; g)\cdot(f \overset{=}{\boxed{*}} g) \tag{17}$$

Note that it is note necessarily true if we only assume $f,g \in \mathrm{Mult}[[B]]$

Proof. As $f,g \in I\cdot\mathrm{Mult}[[B]]$, the following fact is true: If $\sigma,\tau \in Y$,

$$(f\cup g)_{{}^{\tau}\curlyvee^{\sigma}}(x,\dots,x) = (g\cup f)_{\tau}(x,\dots,x)(f\cup g)_{\curlyvee^{\sigma}}(x,\dots,x). \tag{18}$$

Indeed, let us write $\curlyvee^{\sigma} =$  , so that ${}^{\tau}\curlyvee^{\sigma} =$  , thus

$$\begin{aligned}(f\cup g)_{{}^{\tau}\curlyvee^{\sigma}}(x,\dots,x) &= g_k((g\cup f)_{\tau}(x,\dots,x)x,(g\cup f)_{\sigma_2}(x,\dots,x)x,\dots,(g\cup f)_{\sigma_k}(x,\dots,x)x)\\ &= (g\cup f)_{\tau}(x,\dots,x)g_k(x,(g\cup f)_{\sigma_2}(x,\dots,x)x,\dots,(g\cup f)_{\sigma_k}(x,\dots,x)x)\\ &= (g\cup f)_{\tau}(x,\dots,x)(f\cup g)_{\curlyvee^{\sigma}}(x,\dots,x).\end{aligned}$$

The equalities from the Lemma then follow from the unique decomposition $\tau = {}^{\sigma}\!\curlyvee^{\rho}$ for $\tau \in Y^b$. Let us derive (16), the other one is done similarly. Suppose that $n \geq 0$. Then

$$\begin{aligned}(f \boxed{*}\!\| \; g)_n(x_1,\dots,x_n) &= \sum_{\tau\in Y_n}(f\cup g)_{R(\tau)}(x_1,1,x_2,1,\dots,1,x_n,1)\\ &= \sum_{\substack{k+l=n-1\\ k,l\geq 0}}\sum_{\substack{\sigma\in Y_k\\ \rho\in Y_l}}(f\cup g)_{{}^{R(\sigma)}\curlyvee^{R(\rho)}}(x_1,1,x_2,1,\dots,1,x_n,1) \qquad \text{with } \tau = {}^{\sigma}\curlyvee^{\rho}\\ &\overset{(18)}{=} \sum_{\substack{k+l=n-1\\ k,l\geq 0}}\sum_{\substack{\sigma\in Y_k\\ \rho\in Y_l}}(g\cup f)_{\curlyvee^{R(\sigma)}}(x_1,1,x_2,1,\dots,1,x_{k+1})(f\cup g)_{\curlyvee^{R(\rho)}}(1,x_{k+2},1,\dots,1,x_n,1)\\ &= \sum_{\substack{k+l=n-1\\ k,l\geq 0}}(f \overset{=}{\boxed{*}} g)_{k+1}(x_1,x_2,\dots,x_{k+1})(f \;\|\!\boxed{*}\!\|\; g)_l(x_{k+2},\dots,x_n)\\ &= \Big((f \overset{=}{\boxed{*}} g)(f \;\|\!\boxed{*}\!\|\; g)\Big)_n(x_1,\dots,x_n).\end{aligned}$$

□

We consider now the properties of the S-transform (Definition 2.17) with respect to the first convolution product in Definition 3.3.

Theorem 3.6. *Let $f,g \in G_B^I$. Then the S-transform satisfies the following identity with respect to the boxed convolution product* (10)

$$S_{f \boxed{*}\| \, g} = S_g\cdot(S_f\circ(S_g^{-1}IS_g)).$$

Proof. Suppose that $f = IF$ and $g = IG$, for $F, G \in G_B^{\mathrm{inv}}$. Let us write the following products

$$h := f \boxed{*}\| \, g, \quad \overline{h} := f \, \|\boxed{*} \, g, \quad h_1 := f \, \overline{\boxed{*}} \, g, \quad H_2 := f \, \|\boxed{*}\| \, g.$$

We already noticed that the products $f \boxed{*}\| \, g$ and $f \, \overline{\boxed{*}} \, g$ are in G_B^I. Hence, we can write $h = IH$ and $h_1 = IH_1$. Lemma 3.4 then tells us that:

$$h = g \circ h_1 \tag{19}$$

$$\overline{h} = f \circ (H_2 I). \tag{20}$$

As $f, g \in I \cdot \mathrm{Mult}[[B]]$, h and $\overline{h}$ also factorise by Lemma 3.5:

$$h = h_1 H_2 \tag{21}$$

$$\overline{h} = H_2 h_1. \tag{22}$$

First observe that inverting (19) yields

$$IS_h = (IS_{h_1}) \circ (IS_g) = IS_g \cdot S_{h_1} \circ (IS_g)$$

and thus

$$S_h = S_g \cdot S_{h_1} \circ (IS_g). \qquad (\star)$$

Then, the same equality (19) can be written as

$$IH = (IG) \circ (IH_1) = IH_1(G \circ (IH_1)).$$

But $IH = IH_1 H_2$ by (21) so

$$H_2 = G \circ (IH_1).$$

Therefore, by (22), $\overline{h} = H_2 I H_1 = (G \circ (IH_1)) \cdot IH_1$ so

$$\overline{h} = (GI) \circ (IH_1).$$

Now (20) can be turned into $(GI) \circ (IH_1) = f \circ (G \circ (IH_1) \cdot I)$, and then, composing on the left by IS_f,

$$(IS_f) \circ (GI) \circ (IH_1) = G \circ (IH_1) \cdot I.$$

Composing on the right by (IS_{h_1}) gives

$$(IS_f) \circ (GI) = GIS_{h_1}.$$

But the left-hand side equals $GI \cdot (S_f \circ (GI))$, so

$$S_f \circ (GI) = S_{h_1}.$$

Going back to $(\star)$, we can now say that

$$S_h = S_g \cdot S_f \circ (GI) \circ (IS_g),$$

and we get the desired equality since

$$(GI) \circ (IS_g) = G \circ (IS_g) IS_g = S_g^{-1} IS_g.$$

□

Let us finally examine the consequence of this Lemma in operator-valued free probability. Given Proposition 2.24, applying it to K^a and K^b for a and b free elements, we get

Theorem 3.7. *Let $a, b \in \mathcal{A}$ be free elements in an operator-valued free probability space, such that $\mathbb{E}(a) \neq 0$ and $\mathbb{E}(b) \neq 0$. Then*

$$S^{ab} = S^b \cdot S^a \circ ((S^a)^{-1} \cdot I \cdot S^a).$$

References

[1] Ken Dykema. *Multilinear function series and transforms in free probability theory.* Advances in Mathematics 208 (2007), 351-407. (online).

[2] Kurusch Ebrahimi-Fard, Frédéric Patras. *Monotone, free, and boolean cumulants: a shuffle algebra approach.* Advances in Mathematics 328 (2018), 112-132. (arXiv).

[3] Kurusch Ebrahimi-Fard, Loic Foissy, Joachim Kock, Frédéric Patras, *Operads of (noncrossing) partitions, interacting bialgebras, and moment-cumulant relations*, Advances in Mathematics 369 (2020), 107170. (arXiv).

[4] Kurusch Ebrahimi-Fard, Nicolas Gilliers. *On the twisted factorisation of the T-transform.* (arXiv)

[5] Paul Edelman. *Multichains, non-crossing partitions and trees.* Discrete Mathematics 40 (1982), 171-179. (online).

[6] Alessandra Frabetti. *Groups of tree-expanded series.* Journal of Algebra 319 (2008), 377-413. (arXiv).

[7] James Mingo and Roland Speicher. *Free Probability and Random Matrices.* Springer New York, NY, 2017.

[8] Alexandru Nica, Roland Speicher. Lecture on the Combinatorics of free probability. Cambridge University Press, 2006. (online).

[9] Jean-Christophe Novelli, Jean-Yves Thibon. *Noncommutative Symmetric Functions and Lagrange Inversion II: Noncrossing partitions and the Farahat-Higman algebra.* Advances in Applied Mathematics, Volume 140, 2022. (arXiv).

[10] Roland Speicher. *A Short Proof for the Twisted Multiplicativity Property of the Operator-Valued S-transform.* Complex Analysis and Operator Theory (2022) 16: 81. (arXiv).

[11] Roland Speicher. *Non-Commutative Distributions.* 2019. (Lecture notes).

[12] Richard P. Stanley. *Enumerative combinatorics Vol. 2.* 1999. Cambridge University Press. (online)

[13] Dan Voiculescu, Kenneth Dykema and Alexandru Nica. *Free Random Variables.* Centre de Recherches Mathématiques Montréal, 1992.

[14] Dan Voiculescu. *Operations on certain non-commutative operator-valued random variables.* Astérisque 232 (1995), 243-275.