ANNALES
DE L'INSTITUT
HENRI
POINCARÉ

PROBABILITÉS
ET STATISTIQUES

Free probability for pairs of faces II: 2-variables bi-free partial R-transform and systems with rank ≤ 1 commutation[1]

Dan-Virgil Voiculescu

Department of Mathematics, University of California at Berkeley, Berkeley, CA 94720-3840, USA. E-mail: dvv@math.berkeley.edu

Abstract. We compute the generating series for the simplest class of bi-free cumulants, beyond the free cumulants, the two-bands bi-free cumulants of a pair of a left and a right variable. We also consider two-faced systems with a commutation condition implying that two-bands moments, that is expectation values of the product of a monomial of left and a monomial of right variables determine the other moments. Examples include hyponormal operators, dual systems in free entropy theory and bi-partite systems.

Résumé. Pour une paire de variables, une gauche et une droite, nous calculons la série génératrice des plus simples cumulants bi-libres au-delà des cumulants libres. On considère aussi des systèmes à deux faces, satisfaisant une condition de commutation qui assure que les moments à deux bandes, c'est-à-dire les espérances des produits d'un monôme de variables gauches et d'un monôme de variables droites, déterminent tous les autres moments non-commutatifs. Les opérateurs hyponormaux, les systèmes duaux de la théorie de l'entropie libre, ainsi que les systèmes bipartites, fournissent des exemples.

MSC: Primary 46L54; secondary 47B20

Keywords: Bi-free cumulant; Partial bi-free R-transform; Two-bands moment; Hyponormal operator

0. Introduction

In Part I [10] we introduced bi-freeness, which underlies the extension of free probability to two-faced systems of non-commutative random variables, that is, to systems with "left" and "right" variables. We proved on the basis of general Lie-theory considerations the existence of bi-free cumulants and found the analogue of Gaussian variables and distributions in this theory.

Here in Part II our focus will be on the bi-freeness machinery to handle a wider class of distributions of left and right variables, which had been an important motivation for us to study bi-freeness and which goes beyond bi-free Gaussians. These are distributions where algebraic relations between left and right variables insure that all moments can be computed from the "two-bands" moments $\varphi(LR)$ where L and R are monomials in left and, respectively, right variables. The systems satisfy commutation relations between left and right variables, a feature which is preserved under bi-free addition. The simplest example is provided by the bi-partite systems where left and right variables commute with each other. Other examples like those arising from dual systems in free entropy theory [9] and from hyponormal operators [1,3] are characterized by commutators of rank ≤ 1 between left and right variables. To get a hold on additive bi-free convolution for such systems it suffices to compute their "two-bands" bi-free cumulants. In the simplest case of one left variable and one right variable we find a formula connecting the two-variables generating series of two-bands moments and the partial R-transform which is the two-variables generating series of the two-bands

cumulants. This solves the problem of computing bi-free additive convolutions of the distributions of hermitian and antihermitian parts of normal operators or more generally of hyponormal operators with rank ≤ 1 self-commutator. It may be interesting to note that two-bands moments from pairs of variables arising from normal or hyponormal operators appear also in other questions in mathematical physics [4].

Not counting the Introduction, this paper has four sections. Section 1 contains preliminaries about bi-freeness and bi-free cumulants from Part I [10] and about notation we will use. Section 2 gives our main result about the partial R-transform for a pair consisting of one left variable and one right variable. This is analogous to our result for the usual one-variable R-transform [8]. Since neither our original approach via a canonical Toeplitz operator form for the variable [8] nor a combinatorial machinery based on partitions [5] were at hand, it was very convenient that one of the proofs of Haagerup [2] could be used as a starting point for the bi-free result. Then in Section 3 we discuss systems of left and right variables satisfying a rank ≤ 1 commutation condition, their distributions and behavior under bi-free addition.

Section 4 deals with the characterization of classical independence of left and right variables by the vanishing of the bi-free cumulants which mix left and right indices.

1. Preliminaries

1.1. *Bi-freeness*

Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and $((a_i^{(k)})_{i \in I}, (b_j^{(k)})_{j \in J})$, $k = 1, 2$ be two-faced families of non-commutative random variables in $(\mathcal{A}, \varphi)$. Then the definition in Section 2.4. Part I of bi-freeness of the two-faced families for $k = 1$ and $k = 2$ amounts in view of the results in Section 2 of Part I to the existence of two vector spaces with specified state vector $(\mathcal{X}_k, \mathring{\mathcal{X}}_k, \xi_k)$, $k = 1, 2$ with the following properties. There are $\tilde{a}_i^{(k)} \in \mathcal{L}(\mathcal{X}_k)$, $\tilde{b}_j^{(k)} \in \mathcal{L}(\mathcal{X}_k)$, for $i \in I$, $j \in J$, $k = 1, 2$ so that if $(\mathcal{X}, \mathring{\mathcal{X}}, \xi) = (\mathcal{X}_1, \mathring{\mathcal{X}}_1, \xi_1) * (\mathcal{X}_2, \mathring{\mathcal{X}}_2, \xi_2)$ and λ_k, ρ_k are the left and right representations of $\mathcal{L}(\mathcal{X}_k)$ on $\mathcal{L}(\mathcal{X})$ [7,11], then the joint distribution of the variables $a_i^{(k)}$, $b_j^{(k)}$, $i \in I$, $j \in J$, $k \in \{1, 2\}$ in $(\mathcal{A}, \varphi)$ is the same as the joint distribution in $(\mathcal{L}(\mathcal{X}), \varphi_\xi)$ of the variables $\lambda_k(\tilde{a}_i^{(k)})$, $\rho_k(\tilde{b}_j^{(k)})$, $i \in I$, $j \in J$, $k = \{1, 2\}$. Moreover, one can always choose $\mathcal{X}_k = \mathcal{A}$, $\mathring{\mathcal{X}}_k = \ker \varphi$, $\xi_k = 1$, $k = 1, 2$ and $\tilde{a}_i^{(k)}$, $\tilde{b}_j^{(k)}$ to be the left multiplication operators on $\mathcal{A}$ by $\tilde{a}_i^{(k)}$ and, respectively, $\tilde{b}_j^{(k)}$. This is due to the fact that in the definition any $(\mathcal{X}_k, \mathring{\mathcal{X}}_k, \xi_k)$ and $\tilde{a}_i^{(k)}$, $\tilde{b}_j^{(k)}$ yield the same joint distribution of the family as long as the joint distribution of each $(\tilde{a}_i^{(k)})_{i \in I} \cup (\tilde{b}_j^{(k)})_{j \in J}$ in $(\mathcal{L}(\mathcal{X}_k), \varphi_{\xi_k})$ equals that of $(a_i^{(k)})_{i \in I} \cup (b_j^{(k)})_{j \in J}$ in $(\mathcal{A}, \varphi)$.

1.2. *Bi-free cumulants*

If I and J are index sets for left and, respectively, right variables of a two-faced family of non-commutative random variables, we obtained in Theorem 5.7 of Part I a general result about existence and uniqueness of bi-free cumulants R_α indexed by maps $\alpha : \{1, \ldots, n\} \to I \amalg J$, $n \geq 1$. Here R_α is a homogeneous polynomial of degree n in commutative variables $X_{\alpha K}$ of degree $|K|$, where the index αK denotes the map $\alpha \circ b_K : \{1, \ldots, |K|\} \to I \amalg J$, where $\varnothing \neq K \subset \{1, \ldots, n\}$ is a subset and $b_K : \{1, \ldots, |K|\} \to K$ the increasing bijection. Equivalently the R_α's are indexed by sequences $\alpha = (h_1, h_2, \ldots, h_n) \in (I \amalg J)^n$ and are homogeneous polynomials of degree n in the commuting variables $X_{h_{k_1}, \ldots, h_{k_r}}$ of degree r, where $(h_{k_1}, \ldots, h_{k_r})$, $1 \leq k_1 < \cdots < k_r \leq n$ runs over the subsequences of $h_1, \ldots, h_n$. If $z = ((z_i)_{i \in I}, (z_j)_{j \in J})$ is a two-faced family in $(\mathcal{A}, \varphi)$, then $R_\alpha(z)$ or $R_\alpha(\mu_z)$ where μ_z is the distribution of z, will denote the value of R_α when we give the values $\varphi(z_{\alpha(k_1)} \cdots z_{\alpha(k_r)})$ to the variables $X_{\alpha(k_1), \ldots, \alpha(k_r)}$ where $1 \leq k_1 < \cdots < k_r \leq n$. The polynomial R_α is characterized by the requirements that the coefficient of the leading term in $X_{\alpha(1), \ldots, \alpha(n)}$ equals 1 and that the cumulant additivity property

$$R_\alpha(z' + z'') = R_\alpha(z') + R_\alpha(z'')$$

is satisfied if z' and z'' are bi-free.

In the present paper we will be interested in perhaps the simplest bi-free cumulants which are not free cumulants. These are those where $n = p + q$, $p > 0$, $q > 0$ and there are $\beta : \{1, \ldots, p\} \to I$, $\gamma : \{1, \ldots, q\} \to J$ so that

$\alpha : \{1, \ldots, p+q\} \to I \sqcup J$ is given by $\alpha(k) = \beta(k)$ if $1 \leq k \leq p$ and $\alpha(p+l) = \gamma(l)$ if $1 \leq l \leq q$. We shall denote R_α by $R_{\beta\gamma}$ in this particular case.

Actually, we will be studying in detail the case of $R_{\beta\gamma}$, when $|I| = |J| = 1$. In this situation it will be sufficient to specify p and q and we will use the notation R_{pq}. This is the case of one left and one right variable a and b. Then R_{pq} is a polynomial in variables X_{rs}, $0 \leq r \leq p$, $0 \leq s \leq q$, $r + s > 0$ and $R_{pq}(a, b)$ corresponds to setting X_{rs} equal to $\varphi(a^r b^s)$. Here the coefficient of X_{pq} is 1 and R_{pq} is completely determined by the additivity property for the addition of bi-free pairs.

2. The partial bi-free R-transform $R_{(a,b)}(z, w)$ for a pair of variables

2.1.

Let (a, b) in $(\mathcal{A}, \varphi)$ be a two-faced pair of non-commutative random variables, that is a will be a left variable and b will be a right variable. We will prove a formula for the generating series

$$R_{(a,b)}(z, w) = \sum_{\substack{m \geq 0 \\ n \geq 0 \\ m+n \geq 1}} R_{m,n}(a, b) z^m w^n,$$

where $R_{m,n}(a, b)$ are bi-free cumulants for (a, b) of the type discussed in the section of preliminaries. In particular, this means $R_{m,n}(a, b)$ will be a polynomial in moments $\varphi(a^p b^q)$ where $0 \leq p \leq m$, $0 \leq q \leq n$, $p + q \geq 1$. We will show how to derive $R_{(a,b)}(z, w)$ from a two-variables Green's function

$$\varphi\big((z1 - a)^{-1}(w1 - b)^{-1}\big) = G_{(a,b)}(z, w).$$

The result will also involve the one-variable Green's functions

$$\varphi\big((z1 - a)^{-1}\big) = G_a(z), \qquad \varphi\big((z1 - b)^{-1}\big) = G_b(z)$$

and the formula for one-variable R-transforms $R_a(z) = K_a(z) - z^{-1}$, $R_b(z) = K_b(z) - z^{-1}$ we found in [8], where $K_a(z), K_b(z)$ are the inverses of $G_a(z), G_b(z)$ at ∞. In establishing a formula for $R_{(a,b)}(z, w)$ we will use in an essential way insights from one of the proofs given by U. Haagerup [2] for our formula for the one-variable R-transform. It may be interesting to note that certain two-variable Green's functions play an important role in the theory of hyponormal operators [1,3].

Throughout Section 2 $(\mathcal{M}, \varphi)$ will denote a non-commutative probability space.

Lemma 2.1. *Let* $((\mathcal{A}_i, \mathcal{B}_i))_{i \in I}$ *be a bi-free family of included pairs of faces in* $(\mathcal{M}, \varphi)$. *Let* $a_k \in \mathcal{A}_{\alpha(k)}$, $b_l \in \mathcal{B}_{\beta(l)}$, $1 \leq k \leq m$, $1 \leq l \leq n$ *be so that* $\varphi(a_k) = \varphi(b_l) = 0$ *and* $\alpha(1) \neq \alpha(2) \neq \cdots \neq \alpha(m)$, $\beta(1) \neq \beta(2) \neq \cdots \neq \beta(n)$. *Then we have*

$$\varphi(a_m \cdots a_1 b_n \cdots b_1) = \delta_{m,n} \prod_{1 \leq k \leq \min(m,n)} \delta_{\alpha(k), \beta(k)} \varphi(a_k b_k).$$

Proof. Let $(\mathcal{X}_i, \overset{\circ}{\mathcal{X}}_i, \xi_i)$, $i \in I$ be vector spaces with specified state vector and let $\mu_i : \mathcal{A}_i \to \mathcal{L}(\mathcal{X}_i)$, $\nu_i : \mathcal{B}_i \to \mathcal{L}(\mathcal{X})$ be homomorphisms so that

$$\varphi(a_m \cdots a_1 b_n \cdots b_1) = p\big(\lambda_{\alpha(m)}\big(\mu_{\alpha(m)}(a_m)\big)$$
$$\cdots \lambda_{\alpha(1)}\big(\mu_{\alpha(1)}(a_1)\big)$$
$$\cdots \rho_{\beta(n)}\big(\nu_{\beta(n)}(b_n)\big)$$
$$\cdots \rho_{\beta(1)}\big(\nu_{\beta(1)}(b_1)\big)\big)$$

as required by the bi-freeness assumption.

We shall proceed by induction over $\min(m, n)$. The case of $\min(m, n) = 0$ and $\delta_{m,n} = 0$ is the case of $m > 0 = n$ or $m = 0 < n$, that is there are either only left or only right variables in the moment and the moment is then 0 by the definition of freeness. The case of $m = 0 = n$ is the case of the empty word, which should be viewed as being $\varphi(1) = 1$.

Assume now $\min(m, n) > 0$. Then

$$\rho_{\beta(n)}\big(v_{\beta(n)}(b_n)\big) \cdots \rho_{\beta(1)}\big(v_{\beta(1)}(b_1)\big)\xi$$
$$= v_{\beta(1)}(b_1)\xi_{\beta(1)} \otimes \cdots \otimes v_{\beta(n)}(b_n)\xi_{\beta(n)} \in \mathring{\mathcal{X}}_{\beta(1)} \otimes \cdots \otimes \mathring{\mathcal{X}}_{\beta(n)}.$$

If $\alpha(1) \neq \beta(1)$, then

$$\lambda_{\alpha(m)}\big(\mu_{\alpha(m)}(a_m)\big) \cdots \lambda_{\alpha(1)}\big(\mu_{\alpha(1)}(a_1)\big)\rho_{\beta(n)}\big(v_{\beta(n)}(b_n)\big) \cdots \rho_{\beta(1)}\big(v_{\beta(1)}(b_1)\big)\xi$$
$$= \mu_{\alpha(m)}(a_m)\xi_{\alpha(m)} \otimes \cdots \otimes \mu_{\alpha(1)}(a_1)\xi_{\alpha(1)}$$
$$\otimes v_{\beta(1)}(b_1)\xi_{\beta(1)} \otimes \cdots \otimes v_{\beta(n)}(b_n)\xi_{\beta(n)}$$
$$\in \mathring{\mathcal{X}}_{\alpha(m)} \otimes \cdots \otimes \mathring{\mathcal{X}}_{\alpha(1)} \otimes \mathring{\mathcal{X}}_{\beta(1)} \otimes \cdots \otimes \mathring{\mathcal{X}}_{\beta(n)}$$

and hence $\varphi(a_m \cdots a_1 b_n \cdots b_1) = 0$. This proves the assertion in case $\alpha(1) \neq \beta(1)$, that is $\delta_{\alpha(1),\beta(1)} = 0$.

If $\alpha(1) = \beta(1)$, then

$$\lambda_{\alpha(m)}\big(\mu_{\alpha(m)}(a_m)\big) \cdots \lambda_{\alpha(1)}\big(\mu_{\alpha(1)}(a_1)\big)\big(v_{\beta(1)}(b_1)\xi_{\beta(1)} \otimes \cdots \otimes v_{\beta(n)}(b_n)\xi_{\beta(n)}$$
$$= \lambda_{\alpha(m)}\big(\mu_{\alpha(m)}(a_m)\big) \cdots \lambda_{\alpha(2)}\big(\mu_{\alpha(2)}(a_2)\big)\big(\mu_{\alpha(1)}(a_1)v_{\beta(1)}(b_1)\xi_{\beta(1)} - \varphi(a_1 b_1)\xi_{\beta(1)}\big)$$
$$\otimes v_{\beta(2)}(b_2)\xi_{\beta(2)} \otimes \cdots \otimes v_{\beta(n)}\xi_{\beta(n)}\big)$$
$$+ \varphi(a_1 b_1)\lambda_{\alpha(m)}\big(\mu_{\alpha(m)}(a_m)\big) \cdots \lambda_{\alpha(2)}\big(\mu_{\alpha(2)}(a_2)\big)\rho_{\beta(n)}\big(v_{\beta(n)}(b_n)\big) \cdots \rho_{\beta(1)}\big(v_{\beta(1)}(b_1)\big)$$
$$\in \mathring{\mathcal{X}}_{\alpha(m)} \otimes \cdots \otimes \mathring{\mathcal{X}}_{\alpha(2)} \otimes \mathring{\mathcal{X}}_{\beta(1)} \otimes \cdots \otimes \mathring{\mathcal{X}}_{\beta(n)}$$
$$+ \varphi(a_1 b_1)\lambda_{\alpha(m)}\big(\mu_{\alpha(m)}(a_m)\big) \cdots \lambda_{\alpha(2)}\big(\mu_{\alpha(2)}(a_2)\big)\rho_{\beta(n)}\big(v_{\beta(n)}(b_n)\big) \cdots \rho_{\beta(2)}\big(v_{\beta(2)}(b_2)\big)\xi$$

and hence

$$\varphi(a_m \cdots a_1 b_n \cdots b_1) = \varphi(a_1 b_1)\varphi(a_m \cdots a_2 b_n \cdots b_2).$$

The assertion then follows from the induction hypothesis applied to $\varphi(a_m \cdots a_2 b_n \cdots b_2)$. $\square$

2.2.

We introduce in this subsection some of the assumptions and notation we shall use to derive the main result of the section. To make use of some of the technique in [2], we will bring as far as possible, our notations close to those in [2]. We will, however, avoid the precise inequalities for radii of convergence of power-series, by using the language of germs of holomorphic functions. These will be often germs of holomorphic functions near 0 or ∞ taking values in $\mathbb{C}$ or in some Banach space.

We will assume $(\mathcal{A}, \varphi)$, our non-commutative probability space, is a unital Banach algebra $\mathcal{A}$ and that the expectation functional φ is continuous. We will deal with two two-faced pairs of variables (a_1, b_1) and (a_2, b_2) in $(\mathcal{A}, \varphi)$ which will be added (note that here our notation differs from [2] where the variables to be added are a and b).

We define $h_{a_k}(t_k) = \varphi((1 - t_k a_k)^{-1})$, $h_{b_k}(s_k) = \varphi((1 - s_k b_k)^{-1})$, $k = 1, 2$, are germs of holomorphic functions near 0 with values in $\mathbb{C}$. Further we shall also consider $a_k(t_k) = (1 - t_k a_k)^{-1} - h_{a_k}(t_k)1$, $b_k(s_k) = (1 - s_k b_k)^{-1} - h_{b_k}(s_k)1$, $k = 1, 2$, which are germs of holomorphic functions near 0 taking values in $\mathcal{A}$.

We shall also use some of our usual notation [8,11] for the one-variable R-transform. So $G_{a_k}(z_k) = \varphi((z_k 1 - a_k)^{-1})$, $G_{b_k}(z_k) = \varphi((z_k 1 - b_k)^{-1})$ will be germs of holomorphic functions near ∞ while $G_a(t^{-1}) = \varphi((t^{-1} 1 - a)^{-1})$, etc., will be germs of holomorphic functions near 0. Then $K_{a_k}(z_k)$ will be a germ of meromorphic function

near 0, taking the value ∞ at 0, which is the inverse of $G_{a_k}(z_k)$, etc. The R-transform $R_{a_k}(z_k) = K_{a_k}(z_k) - z_k^{-1}$ is then the germ of a holomorphic function near 0.

We will also need two-variable analogs of $G_a(z)$, namely the

$$G_{(a,b)}(z, w) = \varphi\big((z1 - a)^{-1}(w1 - b)^{-1}\big)$$

which is a germ of holomorphic function near (∞, ∞) in $(\mathbb{C} \cup \{\infty\}) \times (\mathbb{C} \cup \{\infty\})$. We will also use the related

$$H_{(a,b)}(t, s) = \varphi\big((1 - ta)^{-1}(1 - sb)^{-1}\big)$$

which is a germ of holomorphic function near $(0, 0) \in \mathbb{C} \times \mathbb{C}$.

Theorem 2.1. *We have the following equality of germs of holomorphic functions near* $(0, 0) \in \mathbb{C}^2$

$$R_{(a,b)}(z, w) = 1 + z R_a(z) + w R_b(w) - \frac{zw}{G_{a,b}(K_a(z), K_b(w))}$$

if (a, b) *is a two-faced pair in a Banach-algebra non-commutative probability space.*

2.3.

The proof of the theorem will occupy the rest of Section 2 and will consist of several steps to establish two facts. The first will be, that if (a_1, b_1) and (a_2, b_2) are bi-free in $(\mathcal{A}, \varphi)$ then

$$\left(\frac{zw}{G_{a_1,b_1}(K_{a_1}(z), K_{b_1}(w))} - 1\right)$$

$$+ \left(\frac{zw}{G_{a_2,b_2}(K_{a_2}(z), K_{b_2}(w))} - 1\right)$$

$$= \frac{zw}{G_{a_1+a_2,b_1+b_2}(K_{a_1+a_2}(z), K_{b_1+b_2}(w))} - 1$$

as germs of holomorphic functions near $(0, 0)$. The second will be, that the coefficient of $z^m w^n$, $m \geq 0$, $n \geq 0$, $m + n > 0$ of the series expansion of

$$1 + z R_a(z) + w R_b(w) - \frac{zw}{G_{a,b}(K_a(z), K_b(w))}$$

is a polynomial of bi-degree (m, n) in the $\varphi(a^k b^l)$, where $\varphi(a^k b^l)$ is given bi-degree (k, l), and that the coefficient of the leading term $\varphi(a^m b^n)$ is 1. Clearly the additivity part will be the more difficult part of the proof, the checking of the bi-degree properties being routine.

2.4.

To compute $G_{(a_1+a_2,b_1+b_2)}(z, w)$ when (a_1, b_1) and (a_2, b_2) are bi-free we shall first compute $H_{(a_1+a_2,b_1+b_2)}(t, s)$. Here (z, w) will be near (∞, ∞), while (t, s) will be near $(0, 0)$. An argument in [2] can then be used to transform $H_{(a_1+a_2,b_1+b_2)}(t, s) = \varphi((1 - t(a_1 + a_2))^{-1}(1 - s(b_1 + b_2))^{-1})$ into an expectation value of the form

$$\varphi\big((1 - t_2 a_2)^{-1}\big(1 - \rho a_1(t_1)a_2(t_2)\big)^{-1}(1 - t_1 a_1)^{-1}(1 - s_2 b_2)^{-1}$$

$$\times \big(1 - \sigma b_1(s_1)b_2(s_2)\big)^{-1}(1 - s_1 b_1)^{-1}\big)$$

for t_1, t_2, s_1, s_2 close to 0 and ρ, σ sufficiently close to 1. The last quantity can be further transformed expanding $(1 - \rho a_1(t_1)a_2(t_2))^{-1}$ and $(1 - \sigma b_1(s_1)b_2(s_2))^{-1}$ in geometric series and expressing the $(1 - ta)^{-1}$, $(1 - sb)^{-1}$ as $a(t) + h_a(t)1$ and $b(s) + h_b(s)1$. Thus we get to a computation which will be recorded as the next lemma.

Lemma 2.2. *Assuming (a_1, b_1) and (a_2, b_2) are bi-free, we have*

$$\varphi\big(\big(a_2(t_2) + h_{a_2}(t_2)1\big)\big(a_1(t_1)a_2(t_2)\big)^k\big(a_1(t_1) + h_{a_1}(t_1)1\big)\big(b_2(s_2) + h_{b_2}(s_2)1\big)$$
$$\times \big(b_1(s_1)b_2(s_2)\big)^l\big(b_1(s_1) + h_{b_1}(s_1)1\big)\big)$$
$$= \delta_{k,l}\big(\varphi\big(a_1(t_1)b_1(s_1)\big)\big)^k\big(\varphi\big(a_2(t_2)b_2(s_2)\big)\big)^l\big(\varphi\big(a_1(t_1)b_1(s_1)\big)$$
$$+ h_{a_1}(t_1)h_{b_1}(s_1)\big)\big(\varphi\big(a_2(t_2)b_2(s_2)\big) + h_{a_2}(t_2)h_{b_2}(s_2)\big)$$

for s_1, s_2, t_1, t_2 near 0 and $k \geq 0, l \geq 0$.

Proof. We start by noticing that among the 16 terms which arise from expanding the left-hand side, only the following four may be non-zero.

$$\varphi\big(\big(a_2(t_2)a_1(t_1)\big)^{k+1}\big(b_2(s_2)b_1(s_1)\big)^{l+1}\big)$$
$$+ \varphi\big(a_2(t_2)\big(a_1(t_1)a_2(t_2)\big)^k b_2(s_2)\big(b_1(s_1)b_2(s_2)\big)^l\big)h_{a_1}(t_1)h_{b_1}(s_1)$$
$$+ \varphi\big(\big(a_1(t_1)a_2(t_2)\big)^k a_1(t_1)\big(b_1(s_1)b_2(s_2)\big)^l b_1(s_1)\big)h_{a_2}(t_2)h_{b_2}(s_2)$$
$$+ \varphi\big(\big(a_1(t_1)a_2(t_2)\big)^k\big(b_1(s_1)b_2(s_2)\big)^l\big)h_{a_1}(t_1)h_{a_2}(t_2)h_{b_2}(s_2).$$

Indeed since $\varphi(a_j(t_j)) = \varphi(b_i(s_i)) = 0$ and the expectation values are such that Lemma 2.1 applies. First, the number of left and right factors, that is of $a_j(t_j)$'s and $b_i(s_i)$'s must be equal, so this accounts for the vanishing of 8 terms in the expansion in which there is an even number of factors of one kind and an odd number of the other kind. Then there are four more terms which vanish because the sequence of indices 1 and 2 for the a's must be the same as for the b's. The application of Lemma 2.1 then gives that the four terms, which may not vanish, are actually equal to

$$\delta_{k,l}\big(\varphi\big(a_2(t_2)b_2(s_2)\big)^{k+1}\big(\varphi\big(a_1(t_1)b_1(s_1)\big)\big)^{k+1}$$
$$+ \big(\varphi\big(a_2(t_2)b_2(s_2)\big)\big)^{k+1}\big(\varphi\big(a_1(t_1)b_1(s_1)\big)\big)^k h_{a_1}(t_1)h_{b_1}(s_1)$$
$$+ \big(\varphi\big(a_1(t_1)b_1(s_1)\big)\big)^{k+1}\big(\varphi\big(a_2(t_2)b_2(s_2)\big)\big)^k h_{a_2}(t_2)h_{b_2}(s_2)$$
$$+ \big(\varphi\big(a_1(t_1)b_1(s_1)\big)\big)^k\big(\varphi\big(a_2(t_2)b_2(s_2)\big)\big)^k h_{a_1}(t_1)h_{a_2}(t_2)h_{b_1}(s_1)h_{b_2}(s_2),$$

which is just the right-hand side of the equality we wanted to prove. $\square$

2.5.

We will use the previous result in conjunction with expansions of $(1 - \rho a_1(t_1)a_2(t_2))^{-1}$ and $(1 - \sigma b_1(s_1)b_2(s_2))^{-1}$. The reader should keep in mind for this that as t and s tend to zero also $a_k(t)$, $b_k(s)$ tend to zero. Thus, if ρ and σ are bounded, the expansions converge for t_k, s_k near zero.

Lemma 2.3. *Assume (a_1, b_1) and (a_2, b_2) are bi-free and $|\rho| \leq C$, $|\sigma| \leq C$ for some constant $C > 0$. Then we have*

$$\varphi\big(\big((1 - t_1a_1)\big(1 - \rho a_1(t_1)a_2(t_2)\big)(1 - t_2a_2)\big)^{-1}$$
$$\times \big((1 - s_1b_1)\big(1 - \sigma b_1(s_1)b_2(s_2)\big)(1 - s_2b_2)\big)^{-1}\big)$$
$$= \frac{(\varphi(a_1(t_1)b_1(s_1)) + h_{a_1}(t_1)h_{b_1}(s_1))(\varphi(a_2(t_2)b_2(s_2)) + h_{a_2}(t_2)h_{b_2}(s_2))}{1 - \rho\sigma\varphi(a_1(t_1)b_1(s_1))\varphi(a_2(t_2)b_2(s_2))}$$

for t_1, t_2, s_1, s_2 near 0.

Proof. The left-hand side equals

$$\varphi\big((a_2(t_2) + h_{a_2}(t_2)1)(1 - \rho a_1(t_1)a_2(t_2))^{-1}(a_1(t_1) + h_{a_1}(t_1)1)$$
$$\times (b_2(s_2) + h_{b_2}(s_1)1)(1 - \sigma b_1(s_1)b_2(s_2))^{-1}(b_1(s_1) + h_{b_1}(s_1)1)\big)$$

which in view of Section 2.5 and Lemma 2.2 is further equal to

$$\sum_{k\geq 0, l\geq 0} \rho^k \sigma^l \varphi\big((a_2(t_2) + h_{a_2}(t_2)1)(a_1(t_1)a_2(t_2))^k(a_1(t_1) + h_{a_1}(t_1)1)$$
$$\times (b_2(s_1) + h_{b_2}(s_2)1)(b_1(s_1)b_2(s_2))^l(b_1(s_1) + h_{b_2}(s_1)1)\big)$$
$$= \sum_{k\geq 0} \rho^k \sigma^k \big(\varphi(a_1(t_1)b_1(s_1))\big)^k \big(\varphi(a_2(t_2)b_2(s_2))\big)^k$$
$$\cdot \big(\varphi(a_1(t_1)b_1(s_1) + h_{a_1}(t_1)h_{b_1}(s_1))(\varphi(a_2(t_2)b_2(s_2)) + h_{a_2}(t_2)h_{b_2}(s_2))\big).$$

The lemma then follows from

$$\sum_{k\geq 0} \rho^k \sigma^k \big(\varphi(a_1(t_1)b_1(s_1))\big)^k \big(\varphi(a_2(t_2)b_2(s_2))\big)^k$$
$$= \big(1 - \rho\sigma\varphi(a_1(t_1)b_1(s_1))\varphi(a_2(t_2)b_2(s_2))\big)^{-1}. \qquad \square$$

2.6.

The next lemma is a result from [2] (part of Lemma 3.3 and Theorem 3.4, see pages 139–141 in [2]). For the reader's convenience a proof will be sketched.

Lemma 2.4. (a) *If t_1, t_2 are near zero and $t_1 h_{a_1}(t_1) = t_2 h_{a_2}(t_2)$, then $\rho = (h_{a_1}(t_1)h_{a_2}(t_2))^{-1}$,*

$$t = \frac{t_1 h_{a_1}(t_1)}{h_{a_1}(t_1) + h_{a_2}(t_2) - 1}$$

are such that

$$(1 - t_1 a_1)(1 - \rho a_1(t_1)a_2(t_2))(1 - t_2 a_2) = \frac{h_{a_1}(t_1) + h_{a_2}(t_2) - 1}{h_{a_1}(t_1)h_{a_2}(t_2)}\big(1 - t(a_1 + a_2)\big).$$

(b) *If t sufficiently close to zero is given then t_1, t_2 sufficiently close to zero exist so that $t_1 h_{a_1}(t_1) = t_2 h_{a_2}(t_2)$ and*

$$t = \frac{t_1 h_{a_1}(t_1)}{h_{a_1}(t_1) + h_{a_2}(t_2) - 1}$$

and $|\rho| < 2$ where $\rho = (h_{a_1}(t_1)h_{a_2}(t_2))^{-1}$.

Proof. (a) We have

$$(1 - t_1 a_1)(1 - \rho a_1(t_1))(a_2(t_2))(1 - t_2 a_2)$$
$$= (1 - t_1 a_1)(1 - t_2 a_2) - \rho\big(1 - h_{a_1}(t_1)1 + t_1 h_{a_1}(t_1)a_1\big)$$
$$\times \big(1 - h_{a_2}(t_2)1 + t_2 h_{a_2}(t_2)a_2\big)$$
$$= c_0 + c_1 a_1 + c_2 a_2 + c_3 a_1 a_2,$$

where

$$c_3 = t_1 t_2 \big(1 - \rho h_{a_1}(t_1) h_{a_2}(t_2)\big) = 0,$$

$$c_0 = 1 - \rho\big(h_{a_1}(t_1) - 1\big)\big(h_{a_2}(t_2) - 1\big) = \frac{h_{a_1}(t_1) + h_{a_2}(t_2) - 1}{h_{a_1}(t_1) h_{a_2}(t_2)},$$

$$c_1 = -t_1\big(1 + \rho h_{a_1}(t_1) - \rho h_{a_1}(t_1) h_{a_2}(t_2)\big) = -t_1\big(h_{a_2}(t_2)\big)^{-1},$$

$$c_2 = -t_2\big(1 + \rho h_{a_2}(t_2) - \rho h_{a_1}(t_1) h_{a_2}(t_2)\big) = -t_2\big(h_{a_1}(t_1)\big)^{-1}.$$

In view of the assumption that $t_1 h_{a_1}(t_1) = t_2 h_{a_2}(t_2)$, we have that $c_1 = c_2$ and, moreover,

$$t = \frac{t_1 h_{a_1}(t_1)}{h_{a_1}(t_1) + h_{a_2}(t_2) - 1} = c_0^{-1} \cdot t_1 \big(h_{a_2}(t_2)\big)^{-1} = -c_1 c_0^{-1} = -c_2 c_0^{-1}.$$

Thus we have

$$(1 - t_1 a_1)\big(1 - \rho a_1(t_1) a_2(t_2)\big)(1 - t_2 a_2)$$

$$= c_0(1 - ta_1 - ta_2) = \frac{h_{a_1}(t_1) + h_{a_2}(t_2) - 1}{h_{a_2}(t_2) h_{a_2}(t_2)}\big(1 - t(a_1 + a_2)\big).$$

(b) Since $t_k \to t_k h_{a_k}(t_k)$ is a holomorphic map near zero taking value 0 at 0 and with first order derivative equal 1 at 0, there is $p(z)$ holomorphic near 0, $p(0) = 0$, $p'(0) = 1$ so that $t_1 h_{a_1}(t_1) = p(t_1) h_{a_2}(p(t_1))$ for t_1 near zero. Thus, given t close to zero we must show there is t_1 so that

$$t = \frac{t_1 h_{a_1}(t_1)}{h_{a_1}(t_1) + (t_1/p(t_1)) h_{a_1}(t_1) - 1}.$$

The right-hand side of this equation is a holomorphic function of t_1 near zero, with first-order jet $(0, 1)$, hence there is a solution t_1 which is a holomorphic function of t near zero with first-order jet $(0, 1)$ at 0. Also taking $t_2 = p(t_1)$, ρ becomes a holomorphic function of t_1 near zero, which takes the value 1 at 0. In particular, we have $|\rho| < 2$ if t_1 is close enough to 0. $\square$

2.7.

Clearly a statement similar to the previous lemma also holds with b_k, s_k, s replacing a_k, t_k, t. The same remark should also be made about the next lemma, which provides further details from [2] to the previous lemma.

Lemma 2.5. *If a_1, a_2 are free in $(\mathcal{A}, \varphi)$ and t, t_1, t_2 are near zero and satisfy the relations in Lemma* 2.4*, then we have*: $h_{a_1 + a_2}(t) = h_{a_1}(t_1) + h_{a_2}(t_2) - 1$ *and* $t h_{a_1 + a_2}(t) = t_1 h_{a_1}(t_1) = t_2 h_{a_2}(t_2)$.

Proof. Passing to inverses and applying φ, the relation in Lemma 2.4(a) gives

$$\varphi\big((1 - t_2 a_2)^{-1}\big(1 - \rho a_1(t_1) a_2(t_2)\big)^{-1}(1 - t_1 a_1)^{-1}\big)$$

$$= \frac{h_{a_1}(t_1) h_{a_2}(t_2)}{h_{a_1}(t_1) + h_{a_2}(t_2) - 1} h_{a_1 + a_2}(t).$$

On the other hand, taking for instance $b_1 = b_2 = 0$, $s_1 = s_2 = 0$, we are in the bi-free setting in which Lemma 2.3 holds, and we get

$$\varphi\big((1 - t_2 a_2)^{-1}\big(1 - \rho a_1(t_1) a_2(t_2)\big)^{-1}(1 - t_1 a_1)^{-1}\big) = h_{a_1}(t_1) h_{a_2}(t_2)$$

since we will have $b_1(s_1) = b_2(s_2) = 0$ and $h_{b_1}(s_1) = h_{b_2}(s_2) = 1$. This gives then $h_{a_1 + a_2}(t) = h_{a_1}(t_1) + h_{a_2}(t_2) - 1$. Multiplying this relation with the equality for t in Lemma 2.4(b), we get $t h_{a_1 + a_2}(t) = t_1 h_{a_1}(t_1)$. The remaining equality $t_1 h_{a_1}(t_1) = t_2 h_{a_2}(t_2)$ is among the assumptions in Lemma 2.4(a). $\square$

Lemma 2.6. *Assume* (a_1, b_1) *and* (a_2, b_2) *are bi-free in* $(\mathcal{A}, \varphi)$ *and assume* t, s *are near zero in* $\mathbb{C}$. *Let further* t_1, t_2, s_1, s_2 *be the numbers near zero provided by Lemma* 2.4 *for* a_1, a_2, t *and, respectively,* b_1, b_2, s. *Then we have*

$$\frac{h_{a_1}(t_1)h_{b_1}(s_1)}{H_{a_1,b_1}(t_1,s_1)} + \frac{h_{a_2}(t_2)h_{b_2}(s_2)}{H_{a_2,b_2}(t_2,s_2)} - 1 = \frac{h_{a_1+a_2}(t)h_{b_1+b_2}(s)}{H_{a_1+a_2,b_1+b_2}(t,s)}.$$

Proof. We shall get the above result by transforming the equality in Lemma 2.3, when the conditions in Lemma 2.4 are satisfied by t, t_1, t_2 and ρ and also the analogous conditions are satisfied by s, s_1, s_2 and σ with respect to b_1, b_2. The left-hand side of the equality in Lemma 2.3 becomes

$$\frac{h_{a_1}(t_1)h_{a_2}(t_2)}{h_{a_1}(t_1) + h_{a_2}(t_2) - 1} \cdot \frac{h_{b_1}(s_1)h_{b_2}(s_2)}{h_{b_1}(s_1) + h_{b_2}(s_2) - 1} \cdot H_{a_1+a_2,b_1+b_2}(t,s)$$

$$= \frac{h_{a_1}(t_1)h_{a_2}(t_2)h_{b_1}(s_1)h_{b_2}(s_2)}{h_{a_1+a_2}(t)h_{b_1+b_2}(s)} H_{a_1+a_2,b_1+b_2}(t,s)$$

in view of Lemma 2.5.

On the other hand, the right-hand side in Lemma 2.3 equals

$$\frac{H_{a_1,b_1}(t_1,s_1)H_{a_2,b_2}(t_2,s_2)}{1 - (h_{a_1}(t_1)h_{a_2}(t_2)h_{b_1}(s_1)h_{b_2}(s_2))^{-1}(H_{a_1,b_1}(t_1,s_1) - h_{a_1}(t_1)h_{b_1}(s_1))(H_{a_2,b_2}(t_2,s_2) - h_{a_2}(t_2)h_{b_2}(s_2))}.$$

Dividing by $h_{a_1}(t_1)h_{a_2}(t_2)h_{b_1}(s_1)h_{b_2}(s_2)$ then gives

$$\frac{H_{a_1+a_2,b_1+b_2}(t,s)}{h_{a_1+a_2}(t)h_{b_1+b_2}(s)}$$

$$= \frac{H_{a_1,b_1}(t_1,s_1)H_{a_2,b_2}(t_2,s_2)}{H_{a_1,b_1}(t_1,s_1)h_{a_2}(t_2)h_{b_2}(s_2) + H_{a_2,b_2}(t_2,s_2)h_{a_1}(t_1)h_{b_1}(s_1) - H_{a_1,b_1}(t_1,s_1)H_{a_2,b_2}(t_2,s_2)}.$$

Taking inverses we get

$$\frac{h_{a_1+a_2}(t)h_{b_1+b_2}(s)}{H_{a_1+a_2,b_1+b_2}(t,s)} = \frac{h_{a_2}(t_2)h_{b_2}(s_2)}{H_{a_2,b_2}(t_2,s_2)} + \frac{h_{a_1}(t_1)h_{b_1}(s_1)}{H_{a_1,b_1}(t_1,s_1)} - 1$$

which is the equity we wanted to prove. $\square$

With the lemmas we have proved, we will be now able to complete the two steps of the proof of Theorem 2.1, which we outlined in Section 2.3.

2.8. *Step* 1 *of the proof of Theorem* 2.1

This step will be derived from Lemma 2.6. Remark that $th_a(t) = G_a(t^{-1})$ if t is near 0, since t^{-1} is then near ∞ on the Riemann sphere and G_a is defined near ∞. Similarly $ts H_{a,b}(t,s) = G_{a,b}(t^{-1}, s^{-1})$ if t and s are near 0.

Next, for (z, w) near $(0, 0)$ let $t_j = (K_{a_j}(z))^{-1}$, $s_j = (K_{b_j}(z))^{-1}$, $t = (K_{a_1+a_2}(w))^{-1}$, $s = (K_{b_1+b_2}(w))^{-1}$. Then t_1, t_2, t satisfy the assumptions in Lemma 2.4. Indeed we have $t_j h_{a_j}(t_j) = G_{a_j}(t_j^{-1}) = G_{a_j}(K_j(z)) = z$ for $j = 1, 2$ and also $th_{a_1+a_2}(t) = z$. Further, we also have

$$\frac{t_1 h_{a_1}(t_1)}{h_{a_1}(t_1) + h_{a_2}(t_2) - 1} = \frac{z}{z/t_1 + z/t_2 - 1}$$

$$= \frac{1}{K_{a_1}(z) + K_{a_2}(z) - z^{-1}} = \frac{1}{K_{a_1+a_2}(z)} = t.$$

The relations for s_1, s_2, s are checked similarly. With these choices of t_j, s_j, t, s, the equality in Lemma 2.6 is an addition property for

$$
\begin{aligned}
\frac{h_{a_j}(t_j)h_{b_j}(s_j)}{H_{a_j,b_j}(t_j,s_j)} - 1 &= \frac{t_j h_{a_j}(t_j)s_j h_{b_j}(s_j)}{t_j s_j H_{a_j,b_j}(t_j,s_j)} - 1 \\
&= \frac{G_{a_j}(K_{a_j}(z))G_{b_j}(K_{b_j}(w))}{G_{a_j,b_j}(K_{a_j}(z),K_{b_j}(w))^{-1}} \\
&= \frac{zw}{G_{a_j,b_j}(K_{a_j}(z),K_{b_j}(w))} - 1
\end{aligned}
$$

which is the desired result.

2.9. *Step* 2 *of the proof of Theorem* 2.1

It is easy to see that since

$$
\begin{aligned}
\big(G_{a,b}(z,w)\big)^{-1} &= \big(z^{-1}w^{-1}H_{a,b}\big(z^{-1},w^{-1}\big)\big)^{-1} \\
&= zw\bigg(\sum_{k\geq 0, l\geq 0} z^{-k}w^{-l}\varphi\big(a^k b^l\big)\bigg)^{-1}
\end{aligned}
$$

the coefficients of $z^p w^q$ in the expansion of $(G_{a,b}(z,w))^{-1}$ near (∞,∞) are polynomials in the moments $\varphi(a^k b^l)$. Since $K_a(z) = z^{-1} + R_a(z)$ has also an expansion with coefficients polynomials in $\varphi(a^k) = \varphi(a^k b^0)$ and the same also holds for $K_b(w)$, we infer that the expansion of

$$
zw\big(G_{a,b}\big(K_a(z), K_b(w)\big)\big)^{-1} - 1
$$

has as coefficients polynomials in the moments $\varphi(a^k b^l)$.

To prove that the coefficient of $z^k w^l$ is a polynomial of bi-degree (k,l) when $\varphi(a^p b^q)$ is assigned bi-degree (p,q), amounts then to proving that

$$
\begin{aligned}
&\lambda^{-1}z\mu^{-1}w\big(G_{\lambda a,\mu b}\big(K_{\lambda a}\big(\lambda^{-1}z\big), K_{\mu b}\big(\mu^{-1}w\big)\big)\big)^{-1} \\
&= zw\big(G_{a,b}\big(K_a(z), K_b(w)\big)\big)^{-1}.
\end{aligned}
$$

Indeed, we have obviously $h_{\lambda a}(\lambda^{-1}t) = h_a(t)$, $H_{\lambda a,\mu b}(\lambda^{-1}t,\mu^{-1}s) = H_{a,b}(t,s)$, which then give for $G_a(z) = z^{-1}h_a(z^{-1})$ and $G_{a,b}(z,w) = z^{-1}w^{-1}H_{a,b}(z^{-1},w^{-1})$, and that $G_{\lambda a}(\lambda z) = \lambda^{-1}G_a(z)$ and $G_{\lambda a,\mu b}(\lambda z,\mu w) = \lambda^{-1}\mu^{-1}G_{a,b}(z,w)$. Since K_a is the inverse of G_a we have $G_{\lambda a}(\lambda K_a(\lambda z)) = \lambda^{-1}G_a(K_a(\lambda z)) = \lambda^{-1}(\lambda z) = z$ so that $\lambda K_a(\lambda z) = K_{\lambda a}(z)$. This gives $K_{\lambda a}(\lambda^{-1}z) = \lambda K_a(z)$ and hence

$$
\begin{aligned}
&\big(G_{\lambda a,\mu b}\big(K_{\lambda a}\big(\lambda^{-1}z\big), K_{\mu b}\big(\mu^{-1}w\big)\big)\big)^{-1} \\
&= \big(G_{\lambda a,\mu b}\big(\lambda K_a(z), \mu K_b(w)\big)\big)^{-1} \\
&= \lambda\mu\big(G_{a,b}\big(K_a(z), K_b(w)\big)\big)^{-1}
\end{aligned}
$$

which gives the desired result when we multiply by $\lambda^{-1}z\mu^{-1}w$.

We still must show that in the polynomial in moments $\varphi(a^p b^q)$ which is the coefficient of $z^k w^l$, in the expansion of $zw(G_{a,b}(K_a(z), K_b(w)))^{-1} - 1$ the coefficient of the leading term $\varphi(a^k b^l)$ is 1. We have

$$
\begin{aligned}
&zw\big(G_{a,b}\big(K_a(z), K_b(w)\big)\big)^{-1} - 1 \\
&= zw\big(\big(K_a(z)\big)^{-1}\big(K_b(w)\big)^{-1}H_{a,b}\big(\big(K_a(z)\big)^{-1}, \big(K_b(w)\big)^{-1}\big)\big)^{-1} - 1.
\end{aligned}
$$

Let

$$\tilde{H}_{a,b}(s,t) = 1 - H_{a,b}(s,t) = -\sum_{\substack{m \geq 0 \\ n \geq 0 \\ m+n \geq 1}} s^k t^l \varphi(a^k b^l).$$

It is easily seen that in the sense of convergence of formal power series, we have

$$zw\big(G_{a,b}\big(K_a(z), K_b(w)\big)\big)^{-1} - 1$$
$$= -1 + \big(1 + zR_a(z)\big)\big(1 + wR_b(w)\big)$$
$$\times \left(1 + \sum_{m \geq 0} \tilde{H}_{a,b}\big(z\big(1 + R_a(z)\big)^{-1}, w\big(1 + R_b(w)\big)^{-1}\big)^m\right).$$

To check on the coefficients of leading terms, i.e., those terms which involve just one moment, we may leave out in the previous expansion any terms which contain products of ≥ 2 moments. Thus we arrive at

$$-1 + \big(1 + zR_a(z) + wR_b(w)\big)\big(1 + \tilde{H}_{a,b}\big(z\big(1 - R_a(z)\big), w\big(1 - R_b(w)\big)\big)\big)$$

which can then be further reduced to

$$-1 + \left(1 + \sum_{p \geq 1} z^p \varphi(a^p) + \sum_{q \geq 1} w^q \varphi(b^q)\right)\left(1 - \sum_{\substack{k \geq 0, l \geq 0 \\ k+l > 0}} z^k w^l \varphi(a^k b^l)\right)$$

and then even further reduced to

$$\sum_{p \geq 1} z^p \varphi(a^p) + \sum_{q \geq 1} w^q \varphi(b^q) - \sum_{\substack{k \geq 0, l \geq 0 \\ k+l > 0}} z^k w^l \varphi(a^k b^l).$$

This shows that in order to get the right coefficient of the leading terms, the partial two-variables bi-free R-transform is

$$R_{a,b}(z,w) = 1 - zw\big(G_{a,b}\big(K_a(z), K_b(w)\big)\big)^{-1} + zR_a(z) + wR_b(w). \qquad \square$$

Remark 2.1. *The bi-free cumulants $R_{m,n}(a,b)$ have integer coefficients. This can be seen easily from the fact that the one variable cumulants $R_n(a)$ have integer coefficients and also the expansion we obtained for $R_{a,b}(z,w)$ which involves $G_{a,b}(z,w)$.*

Remark 2.2. *In addition to the relation to the usual one-variable R-transform [8], note also a certain resemblance in the formula for $R_{a,b}(z,w)$ to formulae for Boolean cumulants [6].*

Remark 2.3. *The reader familiar with the role of analytic subordination in free convolution may be interested to note that both the proof in [2] as well as in our present Lemmata 2.4, 2.5, 2.6 there are traces of this role. In particular, in the latter, this was noted in the reviewer's comments.*

3. Systems with rank ≤ 1 commutation

Definition 3.1. *An* implemented non-commutative probability space *is a triple $(\mathcal{A}, \varphi, \mathcal{P})$, where $(\mathcal{A}, \varphi)$ is a non-commutative probability space and $\mathcal{P} = \mathcal{P}^2 \in \mathcal{A}$ is an idempotent so that $\mathcal{P}a\mathcal{P} = \varphi(a)\mathcal{P}$. An* implemented C^*-probability space $(\mathcal{A}, \varphi, \mathcal{P})$ *will satisfy additional requirements that $(\mathcal{A}, \varphi)$ be a C^*-probability space and that $\mathcal{P} = \mathcal{P}^*$.*

Example 3.1. *If $(\mathcal{A}, \varphi)$ is a C^*-probability space, the GNS construction gives a $*$-representation $\pi : \mathcal{A} \to \mathcal{B}(\mathcal{H})$ and a unit vector $\xi \in \mathcal{H}$ so that $\langle \pi(a)\xi, \xi \rangle = \varphi(a)$ if $a \in \mathcal{A}$. Let $\mathcal{P} \in \mathcal{B}(\mathcal{H})$ be the orthogonal projector onto $\mathbb{C}\xi$ and let $\tilde{\mathcal{A}}$ be the C^*-algebra generated by $\pi(\mathcal{A})$ and $\mathcal{P}$ and $\tilde{\varphi}$ the state $\langle \cdot \xi, \xi \rangle$ on $\tilde{\mathcal{A}}$. Then $(\tilde{\mathcal{A}}, \tilde{\varphi}, \mathcal{P})$ is an implemented C^*-probability space and π gives a morphism of $(\mathcal{A}, \varphi)$ into $(\tilde{\mathcal{A}}, \tilde{\varphi})$ so that $\tilde{\varphi}(\pi(a)) = \varphi(a)$. Remark also that the GNS construction can be replaced by a construction in which π is a faithful representation, so that $(\mathcal{A}, \varphi)$ embeds into $(\tilde{\mathcal{A}}, \tilde{\varphi}, \mathcal{P})$.*

Definition 3.2. *A system with rank ≤ 1 commutation in an implemented non-commutative probability space $(\mathcal{A}, \varphi, \mathcal{P})$ is a two-faced family $((a_i)_{i \in I}, (b_j)_{j \in J})$ in $(\mathcal{A}, \varphi)$, so that $[a_i, b_j] = \lambda_{i,j} \mathcal{P}$ for some $\lambda_{i,j} \in \mathbb{C}$, $i \in I$, $j \in J$. We will call $(\lambda_{ij})_{i \in I, j \in J}$ the coefficients matrix of the system.*

Remark 3.1. *A bipartite system in an implemented non-commutative probability space is the same thing as a system with rank ≤ 1 commutation with coefficients matrix $(\lambda_{ij})_{i \in I, j \in J} = 0$.*

Proposition 3.1. *Let μ and ν be the distributions of two systems with rank ≤ 1 commutation, with index sets (I, J) and, respectively, (K, L) for the variables.*

(a) Let π be the distribution of a two-faced system of variables indexed by $(I \sqcup K, J \sqcup L)$ so that if π', π'' are the restrictions of π to the index pairs (I, J) and (K, L), then $\pi' = \mu$, $\pi'' = \nu$ and π', π'' are bi-free when seen as parts of the distribution π. Then π is the distribution of a system with rank ≤ 1 commutation and coefficients matrix $(\lambda_{ij})_{i \in I, j \in J} \oplus (\lambda_{kl})_{k \in K, l \in L}$ the direct sum of the coefficients matrices for the systems with distributions μ and ν.

(b) If $I = K$, $J = L$ then the bi-free convolution $\mu \boxplus\boxplus \nu$ is the distribution of a system with rank ≤ 1 commutation and coefficients matrix the sum of the coefficients matrices for the systems with distributions μ and ν.

Proof. (a) Let $((a_i')_{i \in I}, (b_j')_{j \in J})$ in $(\mathcal{A}', \varphi', \mathcal{P}')$ and $((a_k'')_{k \in K}, (b_l'')_{l \in L})$ in $(\mathcal{A}'', \varphi'', \mathcal{P}'')$ be the systems with distributions μ, ν and let $(\lambda_{ij}')_{i \in I, j \in J}$, $(\lambda_{kl}'')_{k \in K, l \in L}$ be the corresponding coefficients matrices (a slight deviation from the notation without primes and double primes in the statement of proposition). Let further $(\mathcal{X}', \mathring{\mathcal{X}}', \xi')$ and $(\mathcal{X}'', \mathring{\mathcal{X}}'', \xi'')$ denote the vector spaces with specified state-vectors $(\mathcal{A}'\mathcal{P}', (I - \mathcal{P}')\mathcal{A}'\mathcal{P}', \mathcal{P}')$ and $(\mathcal{A}''\mathcal{P}'', (I - \mathcal{P}'')\mathcal{A}''\mathcal{P}'', \mathcal{P}'')$ endowed with the natural left actions of $\mathcal{A}'$ and $\mathcal{A}''$. Consider $(\mathcal{X}, \mathring{\mathcal{X}}, \xi) = (\mathcal{X}', \mathring{\mathcal{X}}', \xi') * (\mathcal{X}'', \mathring{\mathcal{X}}'', \xi'')$ and $\lambda', \rho', \lambda'', \rho''$ the left and right actions of $\mathcal{L}(\mathcal{X}')$ and $\mathcal{L}(\mathcal{X}'')$ on $\mathcal{X}$ viewed as homomorphisms into $\mathcal{L}(\mathcal{X})$. The actions of $\mathcal{A}'$ and $\mathcal{A}''$ on $\mathcal{X}'$ and $\mathcal{X}''$, respectively, yield homomorphisms $\eta' : \mathcal{A}' \to \mathcal{L}(\mathcal{X}')$, $\eta'' : \mathcal{A}'' \to \mathcal{L}(\mathcal{X}'')$. It is easily seen that denoting by $\mathcal{P}$ the idempotent in $\mathcal{L}(\mathcal{X})$ so that $\mathcal{P}\mathcal{X} = \mathbb{C}\xi$ and $\ker \mathcal{P} = \mathring{\mathcal{X}}$, we have

$$[\lambda'(\eta'(a_i')), \rho'(\eta'(b_j'))] = \lambda_{ij}'\mathcal{P},$$
$$[\lambda''(\eta''(a_k'')), \rho''(\eta''(b_l''))] = \lambda_{kl}''\mathcal{P},$$
$$[\lambda'(\eta'(a_i')), \rho''(\eta''(b_l''))] = 0,$$
$$[\lambda''(\eta''(a_k'')), \rho'(\eta'(b_j'))] = 0.$$

Then $((\lambda'(\eta'(a_i')))_{i \in I} \sqcup (\lambda''(\eta''(a_k'')))_{k \in K}, (\lambda'(\eta'(b_j')))_{j \in J} \sqcup (\lambda''(\eta''(b_l'')))_{l \in L})$ is a system with rank ≤ 1 commutation in $(\mathcal{L}(\mathcal{X}), \varphi_\xi, \mathcal{P})$, which from the distributions point of view is made of bi-free copies of μ and ν and has coefficients matrix $(\lambda_{ij}')_{i \in I, j \in J} \oplus (\lambda_{kl}'')_{k \in K, l \in L}$.

(b) We can use the construction in the proof of (a). Then the system $((\lambda'(\eta'(a_i')) + \lambda''(\eta''(a_i'')))_{i \in I}, (\rho'(\eta'(b_j')) + \rho''(\eta''(b_j'')))_{j \in J})$ in $(\mathcal{L}(\mathcal{X}), \varphi_\xi, \mathcal{P})$ is a system with rank ≤ 1 commutation with distribution $\mu \boxplus\boxplus \nu$. Since

$$[\lambda'(\eta'(a_i')) + \lambda''(\eta''(a_i'')), \rho'(\eta'(b_j')) + \rho''(\eta''(b_j''))] = \lambda_{ij}' + \lambda_j''$$

the coefficients matrix is clearly $(\lambda_{ij}' + \lambda_{ij}'')_{i \in I, j \in J}$. $\qquad\qquad\qquad\qquad\square$

3.1.

We will introduce in this subsection the terminology of *bands in index sequences*. For systems with rank ≤ 1 commutations, we will have to deal with up to two bands, but there is no extra cost to give general definitions, which may provide a better perspective.

If (I, J) is a pair of index sets for a two-faced system of non-commutative random variables, a map $\alpha : \{1, \ldots, m\} \to I \sqcup J$ will be called an *index sequence of length m*. A *band in the index sequence* α is a maximal interval $[p, q]$, $1 \leq p \leq q \leq m$, $p, q \in \mathbb{N}$, so that all $\alpha(r)$, $p \leq r \leq q$ are in only one of the two sets I and J, that is the corresponding variables will be either all left or all right variables. An index sequence α which has n bands will also be called an *n-band index sequence*. It may also be useful to say α *starts left* if $\alpha(1) \in I$ and, hence, also its first band is left or α *starts right* if $\alpha(1) \in J$. Similarly we may say α *ends left* or α *ends right* and clearly α starts and ends on the same side if the band number is odd and starts and ends on different sides if the and number is even. If α has n bands the corresponding non-commutative moment will be called *n bands moment* and the corresponding cumulant will also be called *n bands cumulant*. Also the part of the distribution of a two-faced system of non-commutative random variables involving moments with $\leq n$ bands, will be called its *n bands distribution*. Clearly the n bands distribution can also be given by the cumulants with $\leq n$ bands. In particular, *the bi-free additive convolution gives rise to operations on the corresponding n bands distributions*. It may also be useful at times to consider a subset of the moments or cumulants with $\leq n$ bands by requiring that those with precisely n bands start left or right, etc. We may thus speak of the *n bands distribution starting left* (we repeat that the starting condition applies only to the moments or cumulants with maximum number of bands, that is with n bands).

Proposition 3.2. *The distribution of a system with rank ≤ 1 commutation is completely determined by its 2-bands distribution starting left and its coefficients matrix.*

Proof. Let $(\mathcal{A}, \varphi, \mathcal{P})$ be an implemented non-commutative probability space and let $((z_i)_{i \in I}, (z_j)_{j \in J})$ be a system with rank ≤ 1 commutation in $(\mathcal{A}, \varphi, \mathcal{P})$. Let further $I J = \{(i_1, \ldots, i_p, j_1, \ldots, j_q) \mid p \geq 0, q \geq 0, i_1, \ldots, i_p \in I, j_1, \ldots, j_q \in J\}$ which contains a special "empty" element with $p = q = 0$. If $k \in I \sqcup J$ we will denote by T_k the linear operator of right multiplication by z_k on $\mathcal{A}$, that is $T_k a = a z_k$ if $a \in \mathcal{A}$. By $\mathcal{V} \subset \mathcal{A}$ we will denote the subspace spanned by the $\mathcal{P} z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_q}$ where $(i_1, \ldots, i_p, j_1, \ldots, j_q) \in I J$ which means that we assume $\mathcal{P} \in \mathcal{V}$ corresponding to the empty element of $I J$. We claim that $T_k \mathcal{V} \subset \mathcal{V}$. If $k \in J$ this is obvious since $(i_1, \ldots, i_p, j_1, \ldots, j_q, k) \in I J$ in this case and hence $T_k \mathcal{P} z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_q} = \mathcal{P} z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_q} z_k \in \mathcal{V}$. If $k \in I$ we have

$$T_k \mathcal{P} z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_q} = \mathcal{P} z_{i_1} \cdots z_{i_p} z_k z_{j_1} \cdots z_{j_q}$$

$$+ \sum_{t=1}^{q} \mathcal{P} z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_{t-1}} [z_{j_t}, z_k] z_{j_{t+1}} \cdots z_{j_q}$$

$$= \mathcal{P} z_{i_1} \cdots z_{i_p} z_k z_{j_1} \cdots z_{j_q}$$

$$- \sum_{t=1}^{q} \varphi(z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_{t-1}}) \lambda_{k, j_t} \mathcal{P} z_{j_{t+1}} \cdots z_{j_q}$$

since $[z_{j_t}, z_k] = -\lambda_{k, j_t} \mathcal{P}$ and

$$\mathcal{P} z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_{t-1}} \mathcal{P} = \varphi(z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j \in 1}) \mathcal{P}.$$

This shows that $T_k \mathcal{V} \subset \mathcal{V}$ and, moreover, the formulae for computing $T_k \mathcal{P} z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_q}$ are determined by the 2-bands distribution starting left and the coefficients matrix of the system with rank ≤ 1 commutation. The proposition then follows by applying several times this fact to compute $\mathcal{P} z_{k_1} \cdots z_{k_r} = T_{k_r} \cdots T_{k_1} \mathcal{P}$ where $k_1, \ldots, k_r \in I \sqcup J$ and then passing to $\varphi(z_{k_1} \cdots z_{k_r})$ which is given by

$$\varphi(z_{k_1} \cdots z_{k_r}) \mathcal{P} = (T_{k_r} \cdots T_{k_1} \mathcal{P}) \mathcal{P}$$

and since $T_{k_1} \cdots T_{k_r} \mathcal{P} \in \mathcal{V}$ the computation of $(T_{k_1} \cdots T_{k_r} \mathcal{P}) \mathcal{P}$ gives a polynomial in

$$\mathcal{P} z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_q} \mathcal{P} = \varphi(z_{i_1} \cdots z_{i_p} z_{j_1} \cdots z_{j_q}) \mathcal{P}.$$

Clearly at each step the computations involve only the 2-bands distribution starting left and the coefficients matrix, which proves the proposition.   □

Example 3.2. *The* dual systems *in Section 5 of* [9] *provide examples of systems with rank ≤ 1 commutation. Given $(\mathcal{M}, \tau)$ a von Neumann algebra with a normal faithful trace-state, $I \in \mathcal{B} \subset \mathcal{M}$ a $*$-subalgebra and $X_j = X_j^* \in \mathcal{M}$, $1 \leq j \leq n$, a dual system to $(X_1, \ldots, X_n; \mathcal{B})$ in $L^2(\mathcal{M}, \tau)$ is an n-tuple $(Y_1, \ldots, Y_n)$ of operators $Y_j = Y_j^* \in \mathcal{B}(L^2(\mathcal{M}, \tau))$, $1 \leq j \leq n$ such that $[\mathcal{B}, Y_j] = 0$ and $[X_j, Y_k] = i \delta_{jk} \mathcal{P}$ where $\mathcal{P}$ is the orthogonal projection onto $\mathbb{C}1$. Clearly $((X_1, \ldots, X_n) \amalg (b_k)_{k \in K}), (Y_1, \ldots, Y_n)$ where $(b_k)_{k \in K}$ is any family of elements in $\mathcal{B}$, is a system with rank ≤ 1 commutation in $(\mathcal{B}(L^2(\mathcal{M}, \tau)), \omega, \mathcal{P})$ where ω is the state $\langle \cdot 1, 1 \rangle$ on $\mathcal{B}(L^2(\mathcal{M}, \tau))$. The coefficient matrix in this case is $\lambda_{pq} = i \delta_{pq}$ if $1 \leq p, q \leq n$ and $\lambda_{pk} = 0$ if $1 \leq p \leq n$, $k \in K$. There are additional features in this example, such as the framework and the distribution being C^*, the X_p's and Y_q's being hermitian and the joint distribution of X_p's and $(b_k)_{k \in K}$ being a tracial $*$-distribution. Proposition* 3.1 *can be viewed as an algebraic version and generalization of results for dual systems in* [9].

Example 3.3. *Let $\mathcal{H}$ be a Hilbert space and let $T \in \mathcal{B}(\mathcal{H})$ be a hyponormal operator with rank one self-commutator, that is $T^*T - TT^* = \lambda \mathcal{P}$, $\lambda > 0$ where $\mathcal{P}$ is the orthogonal projection onto an one-dimensional subspace $\mathbb{C}\xi \in \mathcal{H}$, $\|\xi\| = 1$. In $(\mathcal{B}(\mathcal{H}), \omega, \mathcal{P})$ where $\omega(S) = \langle S\xi, \xi \rangle$ for $S \in \mathcal{B}(\mathcal{H})$, one can derive from T examples of pairs with rank 1 self-commutator. Indeed if $\Omega = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2; \mathbb{C})$ we can consider $(aT + bT^*, cT + dT^*)$ which has commutator*

$$\left[aT + bT^*, cT + dT^*\right] = -\det(\Omega) \lambda \mathcal{P}.$$

In case $\Omega = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix}$ we get a pair of hermitian operators, like in the simplest case of the dual systems in Example 3.2. *Clearly, the bi-free additive convolutions of the distributions of such pairs are distributions of the same kind, that is arising from a hyponormal operator. On the other hand, taking for instance $\Omega = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ we get the pair (T, T^*) and bi-free addition of such pairs leads to a pair which does not correspond in the same straightforward way to a hyponormal operator with rank one self-commutator.*

Example 3.4. *The bi-free Gaussian two-faced system in Theorem 7.4 of Part I* [10] *is also an example with rank ≤ 1 commutation. Indeed, if $\mathcal{H}$ is a Hilbert space, $\mathcal{T}(\mathcal{H})$ the full Fock space and $h, h^*: I \amalg J \to \mathcal{H}$ then $z_i = l(h(i)) + l^*(h^*(i))$, $i \in I$ and $z_j = r(h(j)) + r^*(h^*(j))$, $j \in J$ has rank ≤ 1 commutation in $(\mathcal{B}(\mathcal{T}(\mathcal{H})), \omega, \mathcal{P})$ where $\omega(T) = \langle T1, 1 \rangle$, $\mathcal{P}\eta = \langle \eta, 1 \rangle 1$. We have $[z_i, z_j] = (\langle h(j), h^*(i) \rangle - \langle h(i), h^*(j) \rangle) \mathcal{P}$.*

4. Classically independent left and right faces

Theorem 4.1. *Let $\pi = ((a_i)_{i \in I}, (b_j)_{j \in J})$ be a two-faced system of non-commutative random variables in $(\mathcal{A}, \varphi)$.*

(a) *The distribution of π is equal to the distribution of $\pi' = ((a_i')_{i \in I}, (b_j')_{j \in J})$ where $a_i' = a_i$, $b_j' = 0$ iff the only non-vanishing bi-free cumulants of π correspond to index sequences α the range of which is in I. The obvious analogue for index sequences with range in J also holds.*

(b) *The distribution of π has the property that the only non-vanishing bi-free cumulants correspond to index sequences with ranges either in I or in J iff the families of non-commutative random variables $(a_i)_{i \in I}$ and $(b_j)_{j \in J}$ are classically independent in $(\mathcal{A}, \varphi)$.*

Proof. (a) Proposition 5.6 of Part I [10] implies a multi-homogeneity indexed by $I \amalg J$ for bi-free addition, which implies that the bi-free cumulants in view of their uniqueness are also multi-homogeneous. More precisely, if $\alpha: \{1, \ldots, n\} \to I \amalg J$, for every $k \in I \amalg J$ assign the degree $\#\alpha^{-1}(k)$ and for every moment of π assign the corresponding multi degree $I \amalg J \ni k \to \#\alpha^{-1}(k) \in \mathbb{Z}_{\leq 0}$ where α is the index sequence of the monomial the expectation of which is the moment. Then the bi-free cumulant is multi-homogeneous with degrees map corresponding to its

index sequence. In view of this, π' is obtained from π by multiplying the b_i's by 0, we infer that the corresponding cumulants are obtained from those of π by multiplying the ones involving right indices by 0. This proves the only if part and also the if part since the cumulants completely determine the distribution.

(b) Like in (a) consider besides π' also π'' where $a_i'' = 0$ and $b_j'' = b_j$. In view of Proposition 2.16 of Part I [10] the bi-free sum of π' and π'' consists of classically independent copies of $(a_i)_{i \in I}$ and $(b_j)_{j \in J}$ and in view of the addition property of cumulants its bi-free cumulants are either zero or free cumulants of $(a_i)_{i \in I}$ or $(b_j)_{j \in J}$. This proves the if part of (b) and the only if actually also follows since cumulants completely determine the distribution. $\qquad\square$

Corollary 4.1. *If μ and ν are distributions for two-faced families with classically independent left and right variables and same index sets, then $\mu \boxplus\boxplus \nu$ is also a distribution with classically independent left and right variables.*

References

[1] K. Clancey. *Seminormal Operators. Lecture Notes in Math.* **742**. Springer, Berlin, 1979. MR0548170

[2] U. Haagerup. On Voiculescu's R- and S-transforms for free non-commuting variables. In *Free Probability Theory* 127–148. *Fields Institute Communications* **12**. D.-V. Voiculescu (Ed.). Amer. Math. Soc., Providence, RI, 1997. MR1426838

[3] M. Martin and M. Putinar. *Lectures on Hyponormal Operators. Operator Theory: Advances and Applications* **39**. Birkhäuser, Basel, 1989. MR1028066

[4] M. Mineev-Weinstein, M. Putinar and R. Teodorescu. Random matrices in $2D$ Laplacian growth and operator theory. *J. Phys. A* **41** (2008) 26. MR2455097

[5] R. Speicher. Combinatorial theory of the free product and operator-valued free probability theory. *Mem. Amer. Math. Soc.* **132** (627) (1998) x+88. MR1407898

[6] R. Speicher and R. Woroudi. Boolean convolution. In *Free Probability Theory* 267–279. *Fields Institute Communications* **12**. D.-V. Voiculescu (Ed.). Amer. Math. Soc., Providence, RI, 1997. MR1426845

[7] D.-V. Voiculescu. Symmetries of some reduced free product C^*-algebras. In *Operator Algebras and Their Connections with Topology and Ergodic Theory* 556–588. *Lecture Notes in Math.* **1132**. Springer, Berlin, 1985. MR0799593

[8] D.-V. Voiculescu. Addition of certain non-commuting random variables. *J. Funct. Anal.* **66** (1986) 323–346. MR0839105

[9] D.-V. Voiculescu. The analogues of entropy and of Fisher's information measure in free probability theory V. Non-commutative Hilbert transforms. *Invent. Math.* **132** (1998) 189–227. MR1618636

[10] D.-V. Voiculescu. Free probability for pairs of faces I. *Comm. Math. Phys.* **332** (3) (2014) 955–980.

[11] D.-V. Voiculescu, K. J. Dykema and A. Nica. *Free Random Variables. CRM Monograph Series* **1**. Amer. Math. Soc., Providence, RI, 1992. MR1217253