Noname manuscript No.
(will be inserted by the editor)

Free Probability Theory

and its avatars in representation theory, random matrices, and operator algebras; also featuring: non-commutative distributions

Roland Speicher

Received: date / Accepted: date

Abstract This article is an invitation to the world of free probability theory. This theory was introduced by Dan Voiculescu at the beginning of the 1980's and has developed since then into a vibrant and very active theory which connects with many different branches of mathematics. We will motivate Voiculescu's basic notion of "freeness", and relate it with problems in representation theory, random matrices, and operator algebras. The notion of "non-commutative distributions" is one of the central objects of the theory and awaits a deeper understanding.

Keywords free probability theory · random matrix · asymptotic representation theory · non-commutative distribution

Mathematics Subject Classification (2010) MSC 46L54

1 Introduction

This article is an invitation to the world of free probability theory. This theory was introduced by Dan Voiculescu at the beginning of the 1980's and has developed since then into a vibrant and very active theory which connects with many different branches of mathematics. However, instead of following the historical development (which started in the subject of operator algebras), we will begin our journey into the free world by looking on some classical type of problems around the asymptotics of the representation theory of the symmetric group S_n or of summing large matrices. We will then show that asymptotically there is some precise and interesting structure present, which will be captured

Supported by the ERC Advanced Grant "Non-commutative Distributions in Free Probability"

Roland Speicher
Fachrichtung Mathematik, Universität des Saarlandes
E-mail: speicher@math.uni-sb.de

by Voiculescu's notion of "freeness". By viewing this notion in paralled to the concept of "independence", but now for non-commuting variables, we will realize that we are actually dealing with some non-commutative version of a probability theory, and questions like a free central limit theorem become meaningful and have a nice answer. We will give the basics of the machinery for calculating "free convolutions", and will then finally give some precise statements for the heuristic observations we made in the beginning about the asymptotics of representation theory and sums of large matrices. We will also outline the original operator algebra context where Voiculescu introduced free probability theory, by saying a few words about the famous "free group factor isomorphism" problem. We will then end with summarizing what we have observed and alert the reader to the fact that free probability should actually be seen as a central piece of a more general program, which tries to develop a non-commutative counterpart of classical mathematics (related with commuting variables) for maximal non-commuting variables. The notion of "non-commutative distributions" is one of the central objects of the theory and awaits a deeper understanding.

2 The asymptotics of representation theory and sums of large matrices

Many problems in mathematics concern properties of a series of objects, like the representation theory of the symmetric groups S_n or the Horn problem about the eigenvalues of sums of $n \times n$-matrices. Often one has a complete description of the solution for each n; however, for large n this solution might become very complicated and unmanageable and hence one is trying to understand the typical or dominant properties of the solution for large n. A typical phenomenon in high dimension is concentration, which says that although there are many possibilities, they will typically concentrate around one specific situation. The challenge in a concrete context is to prove this concentration and, maybe even more interesting, to isolate and understand this typical configuration.

A lot of recent progress in mathematics has been about such large n asymptotics; Voiculescu's free probability theory, in particular, is a subject which allows to deal with typical configurations in an effective way.

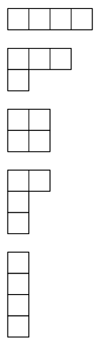
We will first clarify the above statements with two concrete examples of an asymptotic nature and then move on to free probability.

2.1 Two examples for asymptotic statements

Example 1 (Representation theory of S_n)

The irreducible representations of the symmetric group S_n are parametrized by the partitions of n. A partition λ of n (usually denoted by $\lambda \vdash n$), is a decomposition of n as a non-increasing sum of positive integers; partitions are usually depicted in a graphical way by Young diagrams. For example, for $n = 4$, we have five irreducible representations of S_4, corresponding to the following five partitions of 4, respectively the five Young diagrams with 4 boxes:

$4 = 4$	(4)	$=$
$4 = 3 + 1$	$(3, 1)$	$=$
$4 = 2 + 2$	$(2, 2)$	$=$
$4 = 2 + 1 + 1$	$(2, 1, 1)$	$=$
$4 = 1 + 1 + 1 + 1$	$(1, 1, 1, 1)$	$=$

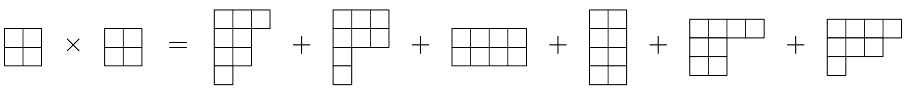


Let us denote the repesentation corresponding to the partition λ by π_λ. We will not give a precise description of this representation, the main question we are interested in is how we can construct new representations out of given ones.

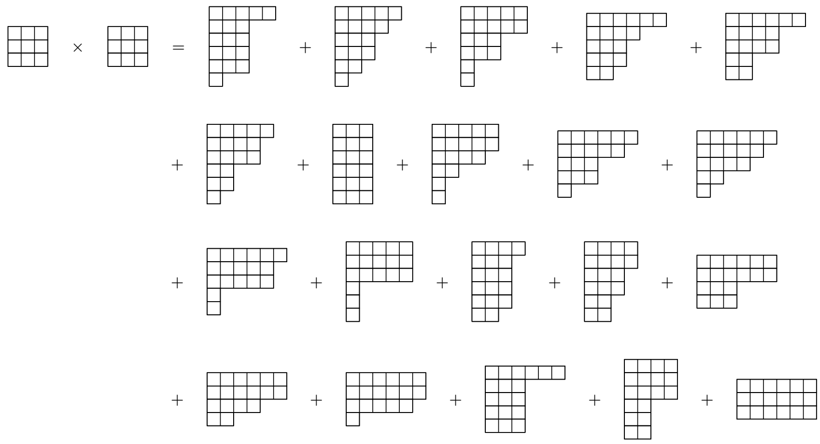
There is a basic construction which makes out of a representation π_1 of S_n and a representation π_2 of S_m a representation of S_{m+n}. This is given by inducing the tensor product representation $\pi_1 \otimes \pi_2$ of $S_m \times S_n \subset S_{m+n}$ from $S_m \times S_n$ to S_{m+n}. However, irreducibility gets lost by doing so. So the induced representation $\mathrm{Ind}_{S_m \times S_n}^{S_{m+n}} \pi_1 \otimes \pi_2$ of the tensor product of an irreducible representation π_1 of S_m and an irreducible representation π_2 of S_n decomposes into the direct sum of several irreducible representations σ_i of S_{n+m}. There is a precise combinatorial rule, the Littlewood-Richardson rule, for determining which Young diagrams appear in this direct sum. This is a bit complicated, so we do not state its precise form; in any case it is algorithmic and can be implemented on a computer easily. In the following we use the Littlewood-Richardson Calculator `lrcalc` by Anders Buch for calculating such decompositions. Here is, for example, the decomposition of the induced representation $\mathrm{Ind}_{S_4 \times S_4}^{S_8} \pi_{(2,2)} \otimes \pi_{(2,2)}$ of S_8:

$$\mathrm{Ind}_{S_4 \times S_4}^{S_8} \pi_{(2,2)} \otimes \pi_{(2,2)} = \pi_{(3,2,2,1)} \oplus \pi_{(3,3,1,1)} \oplus \pi_{(4,4)} \oplus \pi_{(2,2,2,2)} \oplus \pi_{(4,2,2)} \oplus \pi_{(4,3,1)},$$

or in terms of Young diagrams



Let us consider now the same “shape” of the two starting Young diagrams, but in a bigger group, namely $\mathrm{Ind}_{S_9 \times S_9}^{S_{18}} \pi_{(3,3,3)} \otimes \pi_{(3,3,3)}$:



There is no clear structure in this decomposition visible for large n. However, if we give up to understand all the fine details of this decomposition and are only interested in the typical behaviour for large n, then there is some structure appearing. Since we want to compare Young diagrams which correspond to different n, it is advantegeous to renormalize the size of our Young diagrams so that the total area of all boxes of a Young diagram is always equal to 2. (That we choose 2 and not 1 is just convenience to avoid some factors $\sqrt{2}$ in our examples.) We also rotate our Young diagrams by 135 degrees from the original “English” representation to the “Russian” one, see Figure 1.

In this scaling, for large n, most of the diagrams appearing in the decomposition look very similar and there seems to be one dominant limiting shape. In Figure 1 we have randomly chosen two Young diagrams from each decomposition arising from the induced tensor product of the square Young diagram with itself for growing size and one sees a clear limit shape for a typical component. This limit shape, however, is not a Young diagram anymore; it is a kind of continuous version of a Young diagram. It turns out that it is advantageous to encode the information about Young diagrams (and their continuous generalizations) in probability measures on the real line. Such ideas go back to the work of Vershik and Kerov [19,32].

First we associate with our normalized Young diagram λ the interlacing sequences $x_1, \dots, x_m$ and $y_1, \dots, y_{m-1}$ of x-values of the minima and of the maxima, respectively, of the enveloping curve of λ in its Russian representation, see the left diagram in Figure 2. Then we put

$$\mu_\lambda = \sum_{k=1}^{m} \alpha_k \delta_{x_k}, \tag{1}$$

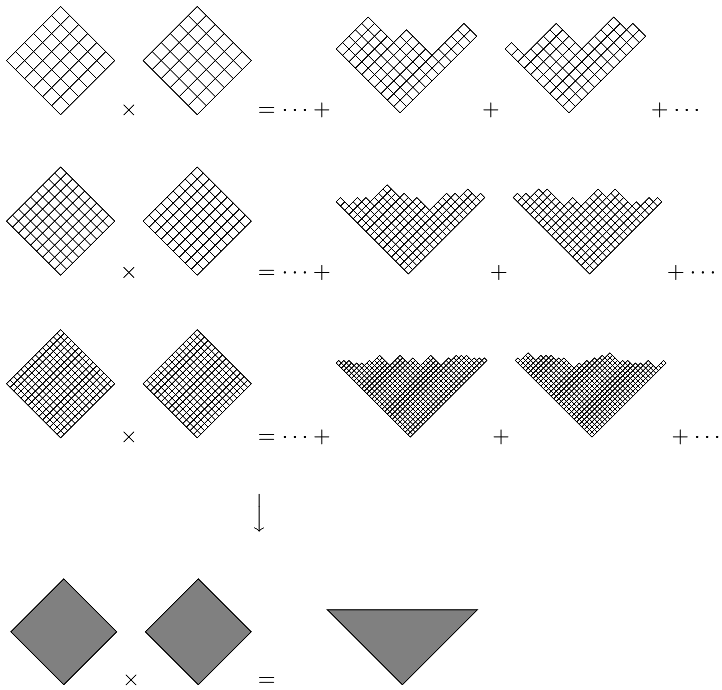


Fig. 1 Two typical Young diagrams appearing in the decomposition of the induced tensor product of a square Young diagram with itself, for the diagrams $(6,6,6,6,6,6)$, $(9,9,9,9,9,9,9,9,9)$, $(15,15,15,15,15,15,15,15,15,15,15,15,15,15,15)$; and the apparent limit shape

where δ_x is the Dirac measure at x, and the weights α_k are given by

$$\alpha_k = \frac{\prod_{i=1}^{m-1}(x_k - y_i)}{\prod_{i\neq k}(x_k - x_i)}. \tag{2}$$

This might look a bit arbitrary, but this measure has some representation theoretic relevance; we will give a precise meaning to this in Section 5. In particular, one might notice that the x_k are exactly the positions where we can add another box in our Young diagram for S_n to get a valid diagram for S_{n+1}.

Let us calculate this measure in the case of our square Young diagrams: we have $m = 2$, $x_1 = -1$, $x_2 = 1$, $y_1 = 0$ (see the right diagram in Figure 2) and thus

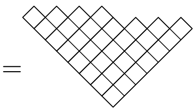
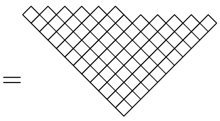
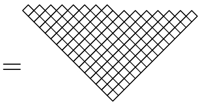
$$\alpha_1 = \frac{x_1 - y_1}{x_1 - x_2} = \frac{-1}{-2} = \frac{1}{2}, \qquad \alpha_2 = \frac{x_2 - y_1}{x_2 - x_1} = \frac{1}{2}.$$

Hence the corresponding measure in this case is given by the symmetric Bernoulli measure with atoms at -1 and $+1$, each with mass $1/2$;

$$\mu_{\blacklozenge} = \mu_{(1)} = \mu_{(2,2)} = \mu_{(3,3,3)} = \cdots = \frac{1}{2}(\delta_{-1} + \delta_{+1}).$$

For the dominating triangular limit shape $\blacktriangledown$ we can calculate a corresponding measure $\mu_{\blacktriangledown}$ as the limit of probability measures for Young diagrams which approximate this limit shape. Let us take for this, for m even, the most "triangular" Young diagram with $m^2/2$ boxes:

$$\lambda_m = (m-1, m-2, \dots, \frac{m}{2}+2, \frac{m}{2}+1, \frac{m}{2}, \frac{m}{2}, \frac{m}{2}-1, \frac{m}{2}-2, \dots, 4, 3, 2, 1)$$

$\lambda_8 =$  $\lambda_{12} =$  $\lambda_{16} =$ 

For $m \to \infty$ this converges to our triangular limit shape, thus we can assign to the latter continuous Young diagram the measure

$$\mu_{\blacktriangledown} = \lim_{m\to\infty} \mu_{\lambda_m}.$$

The interlacing sequences for these λ_m are given by

$$-2 \qquad -2+\frac{4}{m} \qquad \dots \quad -\frac{4}{m} \qquad \frac{2}{m} \qquad \dots \quad 2-\frac{6}{m} \qquad 2-\frac{2}{m}$$

$$-2+\frac{2}{m} \qquad -2+\frac{6}{m} \quad \dots \qquad -\frac{2}{m} \qquad \frac{4}{m} \quad \dots \qquad 2-\frac{4}{m}$$

It is quite straightforward to check that the corresponding measure, given by (1) and (2) converges, for $m \to \infty$, to the arcsine distribution on $[-2, 2]$ which has the density

$$\mu_{\blacktriangledown} = \begin{cases} \frac{1}{\pi\sqrt{4-t^2}}, & |t| \le 2 \\ 0, & |t| > 2. \end{cases} \tag{3}$$

What we have seen here for a special situation remains true in general. If we have two sequences of Young diagrams, whose limit shapes correspond to probability measures μ and ν, respectively, then in the induced tensor representation of these, there dominates asymptotically one shape, which is also associated to a probability measure. Under suitable domains of convergence for the two sequences of Young diagrams, the limit shape depends only on μ and ν. We will denote this dominating measure by $\mu \boxplus \nu$. In our example, we have seen that for the symmetric Bernoulli distribution $\mu_{\blacklozenge}$ and for the arcsine distribution $\mu_{\blacktriangledown}$ we have:

$$\mu_{\blacktriangledown} = \mu_{\blacklozenge} \boxplus \mu_{\blacklozenge}$$

Example 2 (Addition of matrices)

The assignment $(\mu, \nu) \mapsto \mu \boxplus \nu$ in the previous example was quite indirect and the notation as a sum might look strange. We will now present another asymptotic situation where the same operation on probability measures appears much more directly, and corresponds indeed to taking a sum. A discrete probability measure with atoms of equal weights $1/n$ at the positions $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ will now be realized by a symmetric $n \times n$ matrix which has

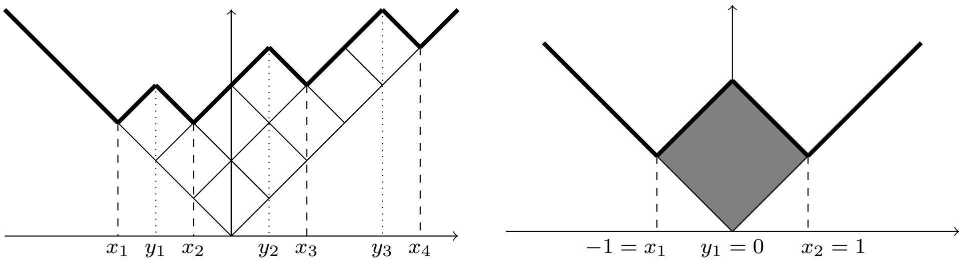


Fig. 2 left: we identify the rescaled Young diagram $(5, 3, 1)$ with the enveloping black curve and encode this via the x-values of its minima, x_1, x_2, x_3, x_4, and the x-values of its maxima, y_1, y_2, y_3; right: the same for the rescaled square Young diagram

these λ_i as eigenvalues. We will try to understand, at least asymptotically, what happens with the eigenvalues if we add two matrices.

Consider two selfadjoint $n \times n$ matrices X and Y. Denote the eigenvalues of X by $\nu = (\nu_1 \leq \cdots \leq \nu_n)$ and the eigenvalues of Y by $\mu = (\mu_1 \leq \cdots \leq \mu_n)$. We ask about the possible eigenvalues $\lambda = (\lambda_1 \leq \cdots \leq \lambda_n)$ of $X + Y$. Again, for given ν and μ there are many possibilities for λ, depending on the relation between the eigenspaces of X and Y. The Horn conjecture (proved by Klyachko, and Knutson and Tao about 15 years ago) gives precise conditions on which λ are possible. As before, for large n this description is not very helpful. However, if we ask for the typical behaviour then we have again a concentration phenomenon: for large n the eigenvalues of the sum concentrate on one distribution. And this dominating distribution is determined by the same operation $\boxplus$ that we observed in the asymptotics of representations of S_n.

Let us look again at the concrete example from before. The distributions $\mu = \nu = \mu_{\blacklozenge}$ will be realized, for even n, by $n \times n$ matrices X and Y, for which half of their eigenvalues are equal to +1 and the other half are equal to -1.

We choose now "randomly" two such matrices and add them. To make this precise we need to put a probability measure on all possibilities for choosing the position between the eigenspaces of X and the eigenspaces of Y. This position is determined by a unitary matrix; since the space of all unitary $n \times n$-matrices is a compact group there exists a Haar measure on this space; this is finite, so we can normalize it to a probability measure. All our statements are now with respect to this measure. So if we want to add our two matrices, we can assume that X is diagonal and the eigenspace of Y is given by a random unitary matrix U. In Figure 3 we show, for $n = 2000$, the histogram of the n eigenvalues of $X + Y$ for two such a random choices of U and compare this with the arcsine distribution. This shows quite clearly that for generic choices of U we will get similar eigenvalue distributions, which are close to the deterministic arcsine distribution $\mu_{\blacktriangledown} = \mu_{\blacklozenge} \boxplus \mu_{\blacklozenge}$.

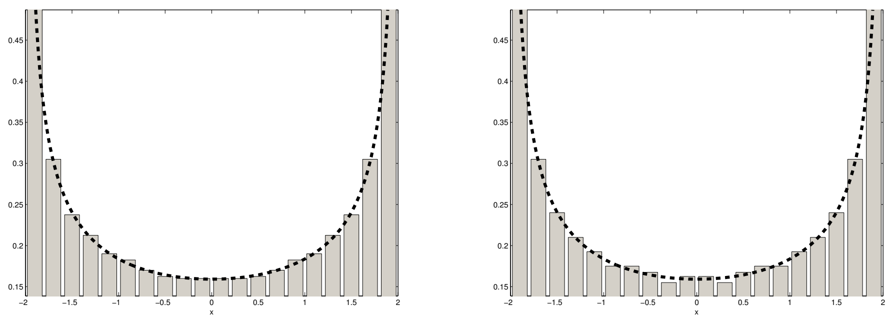


Fig. 3 Histogram of the n eigenvalues of $X+Y$ for X diagonal and two random choices of the eigenspace of Y; both X and Y have eigenvalue distribution $\mu_{\blacklozenge}$; the eigenvalue histogram of $X+Y$ is in both cases compared to $\mu_{\blacktriangledown} = \mu_{\blacklozenge} \boxplus \mu_{\blacklozenge}$; $n = 2000$

2.2 Capturing the structure of the asymptotics

In the previous section we have seen a realization of the asymptotic operation $\mu \boxplus \nu$ for the representations of permutation groups via the sum of two matrices $X+Y$, where X has eigenvalue distribution μ and Y has eigenvalue distribution ν. However, this representation is still an asymptotic one, for $n \to \infty$. We ask now whether we can also find non-asymptotic concrete operators which realize the asymptotic situations which we encountered before? We will try to understand the structure which survives for our $n \times n$-matrices X, Y, when $n \to \infty$. Again we will examine the example $\mu_{\blacktriangledown} = \mu_{\blacklozenge} \boxplus \mu_{\blacklozenge}$.

A selfadjoint matrix X with eigenvalues $+1$ and -1 should clearly correspond to an operator x with $x^2 = 1$. Hence the algebra which describes asymptotically the structure of our matrix X is given by $\mathbb{C}1 + \mathbb{C}x$ with $x^2 = 1$; this is the group algebra $\mathbb{C}\mathbb{Z}_2$, where $\mathbb{Z}_2$ is the group with 2 elements; we identify here the neutral element e in our group with the identity 1 in the algebra. With this algebraic description we can capture the possible eigenvalues of x as $+1$ and -1; for also getting a hold on the multiplicities of the eigenvalues we have to consider a concrete representation of this group algebra; or alternatively, to specify a state (i.e., a positive linear functional) $\tau : \mathbb{C}\mathbb{Z}_2 \to \mathbb{C}$ on the algebra. If we want the eigenvalues to have the same multiplicities we clearly should have a symmetric state, given by $\tau(1) = 1$ and $\tau(x) = 0$. Note that in the group algebra this state τ is nothing but the canonical state which singles out the coefficient of the neutral element, i.e., $\tau(\alpha_1 1 + \alpha_x x) = \alpha_1$.

For the asymptotic description of Y we can take another copy of the group algebra $\mathbb{C}\mathbb{Z}_2$ and the corresponding generator y. However, if we want to build $x+y$ we have to realize the two copies of $\mathbb{C}\mathbb{Z}_2$ in the same algebra. There are two canonical general constructions how we can embed two algebras in a bigger one; one is the tensor product (corresponding to the direct product of the groups) the other one is the free product (corresponding to the free product of the groups). The first possibility is a commutative construction and corresponds in some sense to the notion of independence in classical probability theory; the second possibility on the other hand is a very non-commutative construction and is the appropriate one in our present context.

So let us take the free product $\mathbb{Z}_2 \star \mathbb{Z}_2$ of our two copies of $\mathbb{Z}_2$. This is the (non-commutative!) group which is generated by two generators x and y, subject to the relations $x^2 = 1$ and $y^2 = 1$. This is an infinite group with elements

$$1, x, y, xy, yx, xyx, yxy, xyxy, yxyx, \cdots$$

(Note that the inverses are of course given by $x^{-1} = x$, $y^{-1} = y$ and extension of this, like $(xy)^{-1} = y^{-1}x^{-1} = yx$.) We consider now the corresponding group algebra $\mathbb{C}(\mathbb{Z}_2 \star \mathbb{Z}_2)$, which is given by finite linear combinations of group elements. This contains of course the group algebras of the two copies of $\mathbb{Z}_2$ as the unital subalgebra generated by x and the unital subalgebra generated by y, respectively.

The canonical state τ on $\mathbb{C}(\mathbb{Z}_2 \star \mathbb{Z}_2)$, which picks out the coefficient of the neutral element

$$\tau(\sum_g \alpha_g g) = \alpha_1,$$

is an extension of the corresponding states on the two copies of $\mathbb{C}\mathbb{Z}_2$, hence both x and y have with respect to this still the symmetric distribution on $+1$ and -1. But now x and y live in the same space and we can ask about the distribution of $x + y$.

For this we want to calculate the moments of $x + y$ with respect to our state τ, $\tau((x + y)^k)$. If we draw the Cayley graph of the group $\mathbb{Z}_2 \star \mathbb{Z}_2$ with respect to the generators x and y,

$$\begin{array}{ccccccccccccccccccccc}
 & \cdots & & xyx & & yx & & x & & 1 & & y & & xy & & yxy & & \cdots & \\
\ldots & \circ & \longleftrightarrow & \circ & \longleftrightarrow & \circ & \longleftrightarrow & \circ & \longleftrightarrow & \circ & \longleftrightarrow & \circ & \longleftrightarrow & \circ & \longleftrightarrow & \circ & \longleftrightarrow & \circ & \ldots \\
 & & y & & x & & y & & x & & y & & x & & y & & x & &
\end{array}$$

then the words appearing in the expansion of $(x + y)^k$ correspond to all walks in this graph of length k, starting at 1. Applying τ counts then exactly those walks which end also at 1. These walks, however, are exactly those for which the number of steps to the left is the same as the number of steps to the right (in particular, k must be even, to have such walks), and those closed walks can bijectively be encoded by the position of the $k/2$ steps to the left. Hence the number of such walks, and thus the k-th moment of $x + y$ is given by

$$\varphi\left((x + y)^k\right) = \begin{cases} \binom{k}{k/2}, & k \text{ even} \\ 0, & k \text{ odd} \end{cases}.$$

One might suspect that those numbers are actually the moments of our arcsine distribution (3). In order to see this we can write the central binomial coefficients also in terms of integrating a complex variable z over the unit circle

$\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ as

$$\binom{k}{k/2} = \int_{\mathbb{T}} (z + \bar{z})^k dz.$$

This equation is true, since we need again the same number of z and $\bar{z}$ to make a contribution; note that we have $\int_{\mathbb{T}} z^p dz = \delta_{p0}$ for all $p \in \mathbb{Z}$. But now we can rewrite this complex integral as

$$\begin{aligned}\int_{\mathbb{T}} (z + \bar{z})^k dz &= \frac{1}{2\pi} \int_0^{2\pi} (e^{ix} + e^{-ix})^k dx \\ &= \frac{1}{2\pi} \int_0^{2\pi} (2\cos x)^k dx \\ &= \frac{1}{\pi} \int_0^{\pi} (2\cos x)^k dx \\ &= \frac{1}{\pi} \int_{-2}^{+2} t^k \frac{1}{\sqrt{4 - t^2}} dt.\end{aligned}$$

In the last step we used the substitution $2\cos x = t$, hence $dx = d(\arccos(t/2)) = -dt/\sqrt{4 - t^2}$.

So we see that the moments of our element $x + y$ with respect to the state τ are exactly the moments of the arcsine distribution

$$\tau((x + y)^k) = \int_{-2}^{+2} t^k d\mu_{\blacktriangledown}(t).$$

3 Free Probability Theory

3.1 The notion of freeness

In the previous section we have seen that we can realize our asymptotic situations — showing up in taking induced representations of tensor products or random sums of matrices — by operators in the group algebra of the free product of groups. The main information was actually not given by the algebraic properties of our operators, but was encoded in a probability measure or a state on our algebra. Let us formalize the essentials of this setting.

Definition 1 A pair $(\mathcal{A}, \varphi)$ consisting of a unital algebra $\mathcal{A}$ and a linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ with $\varphi(1) = 1$ is called a *non-commutative probability space*. Often the adjective "non-commutative" is just dropped. Elements from $\mathcal{A}$ are referred to as *(non-commutative) random variables*, the numbers $\varphi(a_{i(1)} \cdots a_{i(k)})$ for such random variables $a_1, \dots, a_m \in \mathcal{A}$ are called *moments*, the collection of all moments is called the *(non-commutative) distribution of* $a_1, \dots, a_m$.

If the algebra $\mathcal{A}$ is a $*$-algebra, then we usually require φ to be a state on $\mathcal{A}$, which means that it is positive in the sense that $\varphi(aa^*) \geq 0$ for all $a \in \mathcal{A}$. The reader might want to check that for a group algebra $\mathbb{C}G$ with the canonical state τ this is indeed the case if the $*$ is given by the inverse on group elements: $(\sum \alpha_g g)^* = \sum \bar{\alpha}_g g^{-1}$. In many cases $\mathcal{A}$ is a $*$-algebra of operators on a Hilbert space $\mathcal{H}$, and the relevant states are vector-states, i.e. of the form $\varphi(a) = \langle a\xi, \xi\rangle$ for a unit vector $\xi \in \mathcal{H}$; for more on this see Section 6.

If the random variables commute, then we are in the classical setting. Usually, one has some positivity structure around — i.e., φ is the expectation with respect to a probability measure and the random variables $x_1, \dots, x_m$ are real-valued — and then the joint distribution in the sense of the collection of all moments can be identified with the analytic distribution of classical random variables as a probability measure $\mu_{x_1,\dots,x_m}$ on $\mathbb{R}^m$ via

$$\varphi(x_{i(1)} \cdots x_{i(k)}) = \int_{\mathbb{R}^k} t_{i(1)} \cdots t_{i(k)} d\mu_{x_1,\dots,x_m}(t_1, \dots, t_m) \tag{4}$$

for all $k \in \mathbb{N}$ and all $1 \leq i(1), \dots, i(k) \leq m$. In particular, in the case of one real-valued random variable x we can identify the distribution of x with the probability measure μ_x via

$$\varphi(x^k) = \int t^k d\mu_x(t).$$

In our algebraic setting we are restricted to situations where all moments exist; and in order that (4) determines $\mu_{x_1,\dots,x_m}$ uniquely we need that the latter is determined by its moments. This is, for example, the case if our random variables are bounded, in which case $\mu_{x_1,\dots,x_m}$ or μ_x are compactly supported.

If we have two representations or two matrices then we assign to them operators x and y, and asymptotically our relevant operations correspond to taking the sum $x+y$. However, in order to make sense out of the sum we must embed x and y in the same space. If we have x and y living in some group algebra $\mathbb{C}G_1$ and $\mathbb{C}G_2$, respectively, then embedding both in the free product $\mathbb{C}(G_1 * G_2)$ seemed to do the job. In general, we can take the free product of the algebras in which x and y live. The crucial point, however, is that we do not only consider this embedding in the free product on the algebraic level, but we also have to extend our states, as the main information is contained in the distributions of our operators. But again we can take our lead for this extension from what we learned in the group case.

For a discrete group G we have the canonical state τ on the group algebra $\mathbb{C}G$, given by singling out the coefficient of the neutral element e,

$$\tau : \mathbb{C}G \to \mathbb{C}, \qquad \sum_{g \in G} \alpha_g g \mapsto \alpha_e.$$

Calculations with respect to this state seem to correspond to the asymptotic operations on representations or matrices. In order to get a better grasp on how we can extend this to more general situations, it is important to anwer the following question: Can we express the fact that G_1 and G_2 are free (in the algebraic sense) in the free product $G_1 * G_2$ also in terms of τ? For this we should note:

- G_1 and G_2 free in $G_1 * G_2$ means, by definition: whenever we have elements $g_1, \dots, g_k$ coming alternatingly from G_1 and G_2 (i.e., $g_j \in G_{i(j)}$ with $i(j) \in \{1, 2\}$ and $i(1) \neq i(2) \neq \cdots \neq g_{k-1} \neq g_k$) and such that each g_j is not the neutral element of $G_{i(j)}$, then $g_1 \cdots g_k$ can also not be the neutral element of $G_1 * G_2$;
- the condition that g_j or $g_1 \cdots g_k$ are not the neutral element can be expressed in terms of τ by $\tau(g_j) = 0$ or $\tau(g_1 \cdots g_k) = 0$, respectively;
- the above statement goes also over to the group algebra: whenever we have elements $a_1, \dots, a_k$ coming alternatingly from $\mathbb{C}G_1$ and $\mathbb{C}G_2$ and such that $\tau(a_j) = 0$ for all $j = 1, \dots, k$, then also $\tau(a_1 \cdots a_k) = 0$.

As it turns out the latter description is a fundamental importance. It was isolated by Voiculescu in the 1980's in the operator algebraic variant of the above situation and termed “freeness”.

Definition 2 Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space and let I be an index set.

1) Let, for each $i \in I$, $\mathcal{A}_i \subset \mathcal{A}$, be a unital subalgebra. The subalgebras $(\mathcal{A}_i)_{i \in I}$ are called *free* or *freely independent*, if $\varphi(a_1 \cdots a_k) = 0$ whenever we have: k is a positive integer; $a_j \in \mathcal{A}_{i(j)}$ (with $i(j) \in I$) for all $j = 1, \dots, k$; $\varphi(a_j) = 0$ for all $j = 1, \dots, k$; and neighboring elements are from different subalgebras, i.e., $i(1) \neq i(2), i(2) \neq i(3), \dots, i(k-1) \neq i(k)$.

2) Let, for each $i \in I$, $x_i \in \mathcal{A}$. The random variables $(x_i)_{i \in I}$ are called *free* or *freely independent*, if their generated unital subalgebras are free, i.e., if $(\mathcal{A}_i)_{i \in I}$ are free, where, for each $i \in I$, $\mathcal{A}_i$ is the unital subalgebra of $\mathcal{A}$ which is generated by x_i. In the same way, subsets $(\mathcal{X}_i)_{i \in I}$ of $\mathcal{A}$ are free, if their generated unital subalgebras are so.

The notion of freeness allows now a precise definition of the operation $\boxplus$ which we encountered in the previous section.

Definition 3 Let μ and ν be probability measures on $\mathbb{R}$ which are determined by their moments. Then the probability measure $\mu \boxplus \nu$ is defined as follows.

- Find a non-commutative probability space $(\mathcal{A}, \varphi)$ and random variables $x, y \in \mathcal{A}$ such that x and y are free and such that $\mu = \mu_x$ and $\nu = \mu_y$, i.e., for all $k \in \mathbb{N}$
$$\varphi(x^k) = \int_{\mathbb{R}} t^k d\mu(t), \qquad \varphi(y^k) = \int_{\mathbb{R}} t^k d\nu(t).$$
- Then $\mu \boxplus \nu$ is the distribution of $x + y$; i.e., it is determined by its moments via
$$\int_{\mathbb{R}} t^k d(\mu \boxplus \nu)(t) = \varphi((x+y)^k) \qquad (t \in \mathbb{N}).$$

For this definition to make sense one has to make the following basic observations about freeness.

- There exists a free product construction which allows to embed any given non-commutative probability spaces $(\mathcal{A}_1, \varphi_1)$, ..., $(\mathcal{A}_m, \varphi_m)$ in a bigger non-commutative probability space $(\mathcal{A}, \varphi)$ such that φ restricted to the image of $\mathcal{A}_i$ coincides with φ_i and such that the images of $\mathcal{A}_1, \dots, \mathcal{A}_m$ are free in $(\mathcal{A}, \varphi)$. This construction preserves also special features of the linear functionals, like positivity or traciality.
- If x and y are free, then the distribution of $x + y$ does only depend on the distribution of x and the distribution of y, but not on the way how x and y are concretely realized.

Example 3 (Calculation of mixed moments for free variables)
In order to see the last point — i.e., how freeness determines mixed moments in x and y out of the moments of x and the moments of y — let us consider the simplest mixed moment, $\varphi(xy)$. If $\varphi(x)$ and $\varphi(y)$ are zero then the definition of freeness says that $\varphi(xy)$ also has to vanish. In the general case we can reduce it to the definition via $\varphi[(x-\varphi(x)1)(y-\varphi(y)1)] = 0$, because the argument of φ is now a product of centered elements which come alternatingly from the unital algebra generated by x and the unital algebra generated by y. Multiplying out terms and using linearity of φ and $\varphi(1) = 1$ gives then $\varphi(xy) = \varphi(x)\varphi(y)$ in general. In the same way one can deal with any mixed moment in x and y. Whereas for moments of the form $\varphi(x^n y^m) = \varphi(x^n)\varphi(y^m)$ one gets the same result as for classical independent random variables, genuine non-commutative moments are quite different (and more complicated); e.g.,

$$\varphi[(x - \varphi(x)1)(y - \varphi(y)1)(x - \varphi(x)1)(y - \varphi(y)1)] = 0$$

results in

$$\varphi(xyxy) = \varphi(x^2)\varphi(y)^2 + \varphi(x)^2\varphi(y^2) - \varphi(x)^2\varphi(y)^2.$$

The above conveys to the reader hopefully the feeling that we have here a kind of non-commutative version of classical probability theory in the following respects:

- Even if we are dealing with abstract elements in algebras or with operators on Hilbert spaces, the concrete realization of these elements is not important, the main quantities of interest are their distributions;
- the notion of “freeness” is very much in analogy to the notion of “independence” in classical probability theory; thus “freeness” is also referred to as “free independence”.

Investigating the concept of freeness, its realizations and implications in various contextes constitutes the subject of *free probability theory.* We are dealing with non-commuting operators in the spirit of classical probability theory, in a context where the notion of freeness is, explicitly or implicitly, of relevance.

3.2 The calculation of the free convolution $\boxplus$

In the last section we gave a precise definition of the free convolution $\boxplus$. However, it remains to develop a machinery which allows to calculate the free convolution. Up to now we have defined $\mu_\diamondsuit \boxplus \mu_\diamondsuit$, but can we also calculate somehow directly that this is actually the arcsine distribution $\mu_\heartsuit$ from (3)?

In the case of the classical convolution (which describes the sum of independent random variables) the Fourier transform provides a powerful analytic tool for doing calculations (and also for proving theorems). It turns out that the role of the Fourier transform is taken over by the Cauchy transform in free probability theory.

Definition 4 For any probability measure μ on $\mathbb{R}$ we define its *Cauchy transform G* by

$$G(z) := \int_{\mathbb{R}} \frac{1}{z-t} d\mu(t). \tag{5}$$

This is an analytic function $G : \mathbb{C}^+ \to \mathbb{C}^-$. In particular in the random matrix context, one prefers to work with $-G : \mathbb{C}^+ \to \mathbb{C}^+$, which is called the *Stieltjes transform.*

Formally, if one expands the geometric series in the definition of G, then one gets the power series expansion

$$G(z) = \sum_{k\geq 0} \frac{\int_{\mathbb{R}} t^k d\mu(t)}{z^{k+1}};$$

if μ is compactly supported, this converges for sufficiently large $|z|$. Hence the Cauchy transform is, like the Fourier transform, a generating function for the moments of the considered distribution. As in the case of the Fourier transform one can also invert the mapping $\mu \mapsto G$; one can recover μ from G by the *Stieltjes inversion formula*

$$d\mu(s) = -\frac{1}{\pi} \lim_{\varepsilon\to 0} \Im G(s+i\varepsilon)dt. \tag{6}$$

This follows formally by observing that

$$-\frac{1}{\pi}\Im G(s+i\varepsilon) = -\frac{1}{\pi}\int_{\mathbb{R}} \Im \frac{1}{(s-t)+i\varepsilon} d\mu(t) = \int_{\mathbb{R}} \frac{1}{\pi}\frac{\varepsilon}{(s-t)^2+\varepsilon^2} d\mu(t)$$

and that the latter integral is the (classical) convolution of μ with the Poisson kernel (or Cauchy distribution), which converges weakly for $\varepsilon \to 0$ to δ_s.

A crucial role for dealing with free convolution will now be played by the functional inverse $K(z)$ of the Cauchy transform $G(z)$. Observe that $K(z)$ has a simple pole at zero. Accordingly, let us split $K(z)$ into this singular part $1/z$ and the remaining analytic part $\mathcal{R}(z)$, which we call the *$\mathcal{R}$-transform.* So we write $K(z) = 1/z + \mathcal{R}(z)$ and have $G[1/z + \mathcal{R}(z)] = z$.

We will now rewrite this in the language of a random variable x in some probability space $(\mathcal{A}, \varphi)$ with distribution $\mu_x = \mu$, hence $G(z) = \varphi\left((z-x)^{-1}\right)$. (We will in the following encounter not only polynomials in x, but also power series. At least if x is bounded, i.e., if μ is compactly supported, these make sense for $|z|$ sufficiently large and the following formal calculations can be rigorously justified.) In terms of x the equation $G[1/z + \mathcal{R}(z)] = z$ reads as

$$\varphi\left(\frac{1}{\frac{1}{z}(1 + z\mathcal{R}(z)) - x}\right) = z, \qquad \text{hence} \qquad \varphi\left(\frac{1}{1 - (zx - z\mathcal{R}(z))}\right) = 1.$$

Let us put $r(z) = z(x - \mathcal{R}(z))$ and expand the resolvent as a Neumann series

$$q(z) = \sum_{k \geq 1} r(z)^k = \frac{1}{1 - r(z)} - 1 = \frac{1}{1 - (zx - z\mathcal{R}(z))} - 1.$$

Then the above tells us that $\varphi(q(z)) = 0$.

Consider now $x_1, \dots, x_n$ which are free in $(\mathcal{A}, \varphi)$. For each x_i we have the corresponding $K_i(z)$, $\mathcal{R}_i(z)$, $r_i(z)$, and $q_i(z)$ as above. We can now calculate

$$\begin{aligned}
\varphi\left(\frac{1}{1 - (r_1(z) + \cdots r_n(z))}\right) &= 1 + \sum_{k \geq 1} \sum_{1 \leq i(1), \dots, i(k) \leq n} \varphi\left(r_{i(1)}(z) \cdots r_{i(k)}(z)\right) \\
&= 1 + \sum_{k \geq 1} \sum_{\substack{1 \leq i(1) \neq i(2) \neq \\ i(3) \neq \cdots \neq i(k) \leq n}} \varphi\left(q_{i(1)}(z) \cdots q_{i(k)}(z)\right) \\
&= 1.
\end{aligned}$$

For the last equation we used that all terms of the form $\varphi(q_{i(1)}(z) \cdots q_{i(k)}(z))$ vanish, because the argument of φ is an alternating product in centered free variables. Thus, with

$$G(z) = \varphi\left(\frac{1}{z - (x_1 + \cdots + x_n)}\right)$$

being the Cauchy transform of $x_1 + \cdots + x_n$, we have seen that

$$\begin{aligned}
&G\left[\frac{1}{z} + \mathcal{R}_1(z) + \cdots + \mathcal{R}_n(z)\right] \\
&\qquad = \varphi\left(\frac{z}{(1 + z\mathcal{R}_1(z) + \cdots + z\mathcal{R}_n(z)) - z(x_1 + \cdots + x_n)}\right) \\
&\qquad = z\varphi\left(\frac{1}{1 - (r_1(z) + \cdots + r_n(z))}\right) \\
&\qquad = z,
\end{aligned}$$

i.e., the $\mathcal{R}$-transform of $x_1 + \cdots + x_n$ is given by $\mathcal{R}(z) = \mathcal{R}_1(z) + \cdots + \mathcal{R}_n(z)$.

Let us phrase our conclusion formally in the following theorem. This is due to Voiculescu [34]; his original proof was much more operator theoretic, the quite elementary and direct proof we presented above follows Lehner [20] (see also [14,41]). Other proofs were given by Haagerup [16] (using operator models) and Speicher [23,29] (relying on the combinatorial approach to freeness, which will be outlined in the next section).

Theorem 1 *For a compactly supported probability measure we define its $\mathcal{R}$-transform $\mathcal{R}_\mu$ as the analytic function in some disk about 0, which is uniquely determined by*

$$G_\mu[1/z + \mathcal{R}(z)] = z, \tag{7}$$

where G_μ is the Cauchy transform of μ.

Then we have the additivity of the $\mathcal{R}$-transform under free convolution*, i.e., for compactly supported probability measures μ and ν, we have in the intersection of their domain*

$$\mathcal{R}_{\mu \boxplus \nu}(z) = \mathcal{R}_\mu(z) + \mathcal{R}_\nu(z). \tag{8}$$

In [6], the free convolution as well as this theorem was extended to all probability measures on $\mathbb{R}$. Whereas the Cauchy transform is always defined on all of the upper complex half-plane, the domain of the $\mathcal{R}$-transform needs some more careful analysis. An alternative approach to this is via subordination functions, which writes $G_{\mu\boxplus\nu}$ in the form $G_{\mu\boxplus\nu}(z) = G_\mu(\omega(z))$, where the subordination function ω is always analytic on all of the upper half plane, and depends both on μ and ν via some fixed point equation in terms of G_μ and G_ν. We will not follow up more on this, but want to point out that this subordination approach, though equivalent to the $\mathcal{R}$-transform approach, is much better behaved from an analytic point of view and is the modern state of the art for dealing with questions on free convolutions. For more information on this, see [4,10,21,30].

Example 4 (Calculation of $\mu_\blacklozenge \boxplus \mu_\blacklozenge$ via $\mathcal{R}$-transform)
Recall that $\mu := \mu_\blacklozenge = 1/2(\delta_{-1} + \delta_{+1})$. Hence its Cauchy transform is given by

$$G_\mu(z) = \int \frac{1}{z-t} d\mu(t) = \frac{1}{2}\Big(\frac{1}{z+1} + \frac{1}{z-1}\Big) = \frac{z}{z^2-1}.$$

Now we apply the relation (7) between Cauchy and $\mathcal{R}$-transform to get the following algebraic equation for the $\mathcal{R}$-transform of μ

$$z = G_\mu[\mathcal{R}_\mu(z) + 1/z] = \frac{\mathcal{R}_\mu(z) + 1/z}{(\mathcal{R}_\mu(z) + 1/z)^2 - 1}.$$

This quadratic equation can easily be solved and we get as a solution $\mathcal{R}_\mu(z) = (\sqrt{1+4z^2} - 1)/(2z)$. (The other solution can be excluded by looking at the asymptotic behavior for $z \to \infty$.) By the additivity (8) of the $\mathcal{R}$-transform, we get for the $\mathcal{R}$-transform of $\nu := \mu \boxplus \mu$ that $\mathcal{R}_\nu(z) = 2\mathcal{R}_\mu(z) = (\sqrt{1+4z^2} - 1)/z$. Using again the relation (7) gives now a quadratic equation for the Cauchy transform of ν, which can be solved to yield $G_\nu(z) = 1/\sqrt{z^2-4}$. Applying the Stieltjes inversion formula to this gives then finally

$$d\nu(t) = -\frac{1}{\pi}\Im\frac{1}{\sqrt{t^2-4}}dt = \begin{cases} \frac{1}{\pi\sqrt{4-t^2}}, & |t| \leq 2 \\ 0, & \text{otherwise} \end{cases},$$

i.e., we have indeed that $\mu_\blacklozenge \boxplus \mu_\blacklozenge = \nu$ is the arcsine distribution $\mu_\blacktriangledown$.

Example 5 (Free central limit theorem)
If one has the notion of a convolution, then one of the first questions which come to mind is about the analogue of a central limit theorem: what can one say about the limit of $(D_{1/\sqrt{n}}\mu)^{\boxplus n}$ if n goes to ∞; $D_r\nu$ denotes here the dilation of the measure ν by the factor r. To put the same question in terms of the random variables: assume $x_1, x_2, \dots$ are free and identically distributed; what can we say about the limit distribution of $(x_1 + \cdots + x_n)/\sqrt{n}$? With the $\mathcal{R}$-transform machinery this is easy to answer. Let us denote by $\mathcal{R}_n$ the $\mathcal{R}$-transform of $(D_{1/\sqrt{n}}\mu)^{\boxplus n}$. By the additivity of the $\mathcal{R}$-transform for free variables and by the scaling behaviour of the $\mathcal{R}$-transform $\mathcal{R}_{D_r\nu}(z) = r\mathcal{R}_\nu(rz)$, we get

$$\mathcal{R}_n(z) = n\frac{1}{\sqrt{n}}\mathcal{R}_1(z/\sqrt{n}),$$

where $\mathcal{R}_1$ is the $\mathcal{R}$-transform of μ. If we expand the $\mathcal{R}$-transforms into a power series about 0

$$\mathcal{R}_n(z) = \kappa_1^{(n)} + \kappa_2^{(n)}z + \kappa_3^{(n)}z^2 + \cdots$$

(the $\mathcal{R}$-transform is an analytic function in a sufficiently small ball about zero, if the measure has compact support), then we see that those coefficients scale as $\kappa_i^{(n)} = (\sqrt{n})^{i-2}\kappa_i^{(1)}$; this means $\kappa_1^{(1)}$ should better be zero (this corresponds to the assumption that μ has mean zero, which always is assumed in a central limit theorem) and, if this is the case, the only surviving coefficient in the limit is $\kappa_2^{(1)}$ (which is actually the variance of μ). Hence, under the usual assumptions $\kappa_1^{(1)} = 0$ and $\kappa_2^{(1)} = 1$, we get $\lim_{n\to\infty}\mathcal{R}_n(z) = z$. It is easy to see that pointwise convergence of the $\mathcal{R}$-transform is the same as weak convergence of the corresponding measures. So what we get in the central limit is a measure, whose $\mathcal{R}$-transform is given by $\mathcal{R}(z) = z$. Plugging this into (7) gives for the corresponding Cauchy transform the equation $zG(z) = 1 + G(z)^2$, which has the solution $G(z) = (z - \sqrt{z^2-4})/2$. By the Stieltjes inversion formula this gives a measure μ_s with support $[-2, 2]$ and density $d\mu_s(t) = \sqrt{4-t^2}/(2\pi)dt$. According to the form of its density, this is usually referred to as a *semicircular distribution.*

This free central limit theorem was one of the first theorems in free probability theory, proved by Voiculescu [33] around 1983. The appearance of the semicircular distribution in this limit pointed to connections with random matrix theory, we will come back to this in Section 4.

3.3 The combinatorics of freeness

The coefficients in the power series expansion of the Cauchy transform of a probability measure μ are essentially the moments of μ. The coefficients of the $\mathcal{R}$-transform (which already played a role in our proof of the free central limit theorem) are, via the relation (7) between the Cauchy and the $\mathcal{R}$-transform, some polynomials in the moments and are of quite some importance for the development of free probability theory. We denote them by $(\kappa_i)_{i\geq 1}$ and call them *free cumulants* (in analogy to the classical cumulants, which are essentially the coefficients in the expansion of the logarithm of the Fourier transform). The relation between moments and cumulants has a beautiful combinatorial structure, which was discovered in [29] (see also [25]) and is governed by the lattice of non-crossing partitions. Here we mean partitions of sets, not of numbers as we encountered them in Section 1 in the context of Young diagrams.

Definition 5 A *partition* of $\{1,\dots,k\}$ is a decomposition $\pi = \{V_1,\dots,V_r\}$ with $V_i \neq \emptyset$, $V_i \cap V_j = \emptyset$ $(i \neq j)$, and $\bigcup_i V_i = \{1,\dots,k\}$. The V_i are the *blocks* of π. If each block of π has exactly two elements, then we call π a *pairing.* A partition π is *non-crossing* if we do not have $p_1 < q_1 < p_2 < q_2$, such that p_1, p_2 are in the same block, and q_1, q_2 in another, different block. By $NC(k)$ we will denote the set of all non-crossing partitions of $\{1,\dots,k\}$.

If we replace the Cauchy transform and the $\mathcal{R}$-transform by the more natural generating power series in the moments and the free cumulants,

$$M(z) := 1 + \sum_{i\geq 1} m_i z^i, \qquad C(z) := 1 + \sum_{i\geq 1} \kappa_i z^i,$$

then we have the relations $G(z) = M(1/z)/z$ and $C(z) = 1 + z\mathcal{R}(z)$, and the relation (7) between G and $\mathcal{R}$ reads now $C[zM(z)] = M(z)$. In terms of the coefficients this is equivalent to the recursion

$$m_k = \sum_{s=1}^{k} \sum_{\substack{i_1,\dots,i_s\in\{0,1,\dots,k-s\}\\ i_1+\cdots+i_s+s=k}} \kappa_s m_{i_1}\cdots m_{i_s}.$$

Iterating this yields finally the following formula for writing the moments in terms of the free cumulants,

$$m_k = \sum_{\pi\in NC(k)} \kappa_\pi. \tag{9}$$

Here κ_π is a product of cumulants, one factor κ_r for each block of π of size r. The relation between moments and classical cumulants — which is the relation between the coefficients of an exponential generating series and the logarithm of this seris — is exactly the same, with the only difference that the summation runs there over all partitions of k instead over the non-crossing ones.

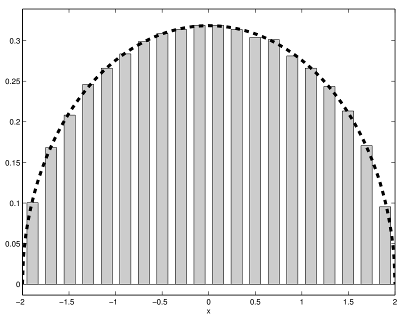


Fig. 4 Histogram of the n eigenvalues of one realization of an $n \times n$ Gaussian random matrix for $n = 2000$; compared to the semicircle distribution. A Gaussian random matrix is selfadjoint, and apart from this all its entries are independent, and distributed according the Gaussian distribution; in order to get a limit for $n \to \infty$ one has to scale the matrix with $1/\sqrt{n}$. This convergence of the eigenvalue distribution of a Gaussian random matrix to the semicircle distribution was proved by Wigner in 1955 and is usually referred to as *Wigner's semicircle law.*

Since for the semicircle distribution only the second cumulant is different from zero, the formula (9) gives in this case that the $2k$-th moment of the normalized semicircular distribution is given by the number of non-crossing pairings of $2k$ elements. This is analogous to the classical statement that the $2k$-th moment of the normalized Gaussian distribution is given by the number of all pairings of $2k$ elements, which is just $(2k-1)!! := (2k-1)\cdot(2k-3)\cdots 5\cdot 3\cdot 1$.

This combinatorial description of freeness in terms of non-crossing partitions goes much further and a lot of the progress in free probability theory relied on this description; more information on this can be found in [23]; see also [24].

4 Free Probability and the Asymptotic Eigenvalue Distribution of Random Matrices

The limit distribution in the free central limit theorem, the semicircle distribution μ_s, had appeared before in the literature, namely in the random matrix context as one of the basic limiting eigenvalue distributions; see Figure 4. This connection motivated Voiculescu to look for some deeper relation between free probability theory and random matrices, resulting in many asymptotic freeness results for random matrices. A glimpse of this we saw in Section 1. Here we want to be a bit more precise on this connection.

One should note that our presentation of the notion of freeness is quite orthogonal to its historical development. Voiculescu introduced this concept in an operator algebraic context (we will say a few words on this in Section 6); at this point, when he introduced the $\mathcal{R}$-transform and proved the free central limit theorem around 1983, there was no relation at all with random matrices (or the asymptotics of representations); this connection was only revealed later in [36].

Random $n \times n$-matrices for finite n have a quite complicated distribution and a nice and controllable structure shows up only in the limit $n \to \infty$; in particular, freeness cannot be found among finite matrices, one can see it in the random matrix world only asymptotically for $n \to \infty$. Hence we need some notion of convergence and, in particular, the concept of "asymptotic freeness".

Definition 6 Let $(\mathcal{A}_n, \varphi_n)$, for $n \in \mathbb{N}$, and $(\mathcal{A}, \varphi)$ be non-commutative probability spaces. Let I be an index set. For each $i \in I$, let $a_i^{(n)} \in \mathcal{A}_n$ $(n \in \mathbb{N})$ and $a_i \in \mathcal{A}$. We say that $(a_i^{(n)})_{i\in I}$ *converges in distribution* to $(a_i)_{i\in I}$, denoted by $(a_i^{(n)})_{i\in I} \stackrel{\text{distr}}{\longrightarrow} (a_i)_{i\in I}$, if we have that each joint moment of $(a_i^{(n)})_{i\in I}$ converges to the corresponding joint moment of $(a_i)_{i\in I}$, i.e., if we have

$$\lim_{n\to\infty} \varphi_n(a_{i_1}^{(n)} \cdots a_{i_k}^{(n)}) = \varphi(a_{i_1} \cdots a_{i_k}) \tag{10}$$

for all $k \in \mathbb{N}$ and all $i_1, \dots, i_k \in I$.

Definition 7 Let, for each $n \in \mathbb{N}$, $(\mathcal{A}_n, \varphi_n)$ be a non-commutative probability space. Let I be an index set and consider for each $i \in I$ and each $n \in \mathbb{N}$ random variables $a_i^{(n)} \in \mathcal{A}_n$. Let $I = I_1 \cup \cdots \cup I_m$ be a decomposition of I into m disjoint subsets. We say that $\{a_i^{(n)} \mid i \in I_1\}, \dots, \{a_i^{(n)} \mid i \in I_m\}$ are *asymptotically free* (for $n \to \infty$), if $(a_i^{(n)})_{i\in I}$ converges in distribution towards $(a_i)_{i\in I}$ for some random variables $a_i \in \mathcal{A}$ $(i \in I)$ in some non-commutative probability space $(\mathcal{A}, \varphi)$ and if the limits $\{a_i \mid i \in I_1\}, \dots, \{a_i \mid i \in I_m\}$ are free in $(\mathcal{A}, \varphi)$.

In [36] Voiculescu could show that not only one of the most basic distributions in free probability theory, the semicircular distribution, shows up as the limit of random matrices, but that also the whole concept of freeness can be found in the random matrix world, at least asymptotically. The following is the most basic such result [38], for independent Gaussian random matrices and can be seen as a vast generalization of Wigner's semicircle law to the multivariate setting. Note that whereas Wigner's theorem can be formulated within classical probability theory, the multivariate situation is genuinely non-commutative and one needs a non-commutative version of probability theory even for its formulation.

Theorem 2 *Let $A_1^{(n)}, \dots, A_p^{(n)}$ be p independent $n \times n$ Gaussian random matrices and let $D_1^{(n)}, \dots, D_q^{(n)}$ be q random $n \times n$ matrices such that almost surely*

$$D_1^{(n)}(\omega), \dots, D_q^{(n)}(\omega) \stackrel{\text{distr}}{\longrightarrow} d_1, \dots, d_q \qquad \text{as } n \to \infty.$$

Furthermore, assume that, for each n, $A_1^{(n)}, \dots, A_p^{(n)}, \{D_1^{(n)}, \dots, D_q^{(n)}\}$ are independent. Then we have almost surely for $n \to \infty$ that

$$A_1^{(n)}(\omega), \dots, A_p^{(n)}(\omega), D_1^{(n)}(\omega), \dots, D_q^{(n)}(\omega) \stackrel{\text{distr}}{\longrightarrow} s_1, \dots, s_p, d_1, \dots, d_q,$$

where each s_i is semicircular and $s_1, \dots, s_p, \{d_1, \dots, d_q\}$ are free. So in particular, we have that $A_1^{(n)}, \dots, A_p^{(n)}, \{D_1^{(n)}, \dots, D_q^{(n)}\}$ are almost surely asymptotically free.

The independence between two matrices means that all entries of the first matrix are independent from all entries of the second matrix; and similar for sets of matrices.

Another important variant of this theorem is for unitary random matrices. Let us first recall what we mean with a Haar unitary random matrix. Let $\mathcal{U}(n)$ denote the group of unitary $n\times n$ matrices, i.e. $n\times n$ complex matrices U which satisfy $U^*U = UU^* = I_n$. Since $\mathcal{U}(n)$ is a compact group, one can take dU to be Haar measure on $\mathcal{U}(n)$ normalized so that $\int_{\mathcal{U}(n)} dU = 1$, which gives a probability measure on $\mathcal{U}(n)$. A *Haar distributed unitary random matrix* is a matrix U_n chosen at random from $\mathcal{U}(n)$ with respect to Haar measure. There is a useful theoretical and practical way to construct such Haar unitaries: take an $n \times n$ (non-selfadjoint!) random matrix whose entries are independent standard complex Gaussians and apply the Gram-Schmidt orthogonalization procedure; the resulting matrix is then a Haar unitary random matrix.

The distribution of each such Haar unitary random matrix is, for each n and thus also in the limit $n \to \infty$, a Haar unitary in the sense of the following definition.

Definition 8 Let $(\mathcal{A}, \varphi)$ be a non-commutative probability space such that $\mathcal{A}$ is a $*$-algebra. An element $u \in \mathcal{A}$ is called a *Haar unitary* if u is unitary, i.e. $u^*u = 1_{\mathcal{A}} = uu^*$ and if $\varphi(u^m) = \delta_{0,m}$ for all $m \in \mathbb{Z}$.

Here is now the version of the previous theorem for the unitary case. This is again due to Voiculescu [36, 38].

Theorem 3 *Let $U_1^{(n)}, \dots, U_p^{(n)}$ be p independent $n \times n$ Haar unitary random matrices, and let $D_1^{(n)}, \dots, D_q^{(n)}$ be q random $n \times n$ matrices such that*

$$D_1^{(n)}(\omega), \dots, D_q^{(n)}(\omega) \stackrel{\text{distr}}{\longrightarrow} d_1, \dots, d_q \qquad as \qquad n \to \infty.$$

Furthermore, we assume that, for each n, $\{U_1^{(n)}, U_1^{(n)*}\}$, $\dots$, $\{U_p^{(n)}, U_p^{(n)*}\}$, $\{D_1^{(n)}, \dots, D_q^{(n)}\}$ are independent. Then we have almost surely for $n \to \infty$ that

$$U_N^{(1)}(\omega), U_N^{(1)*}(\omega) \dots, U_N^{(p)}(\omega), U_N^{(p)*(\omega)}, D_N^{(1)}(\omega), \dots, D_N^{(q)(\omega)} \stackrel{\text{distr}}{\longrightarrow} u_1, u_1^*, \dots, u_p, u_p^*, d_1, \dots, d_q,$$

where each u_i is a Haar unitary and $\{u_1, u_1^*\}, \dots, \{u_p, u_p^*\}, \{d_1, \dots, d_q\}$ are free. So, in particular, $\{U_1^{(n)}, U_1^{(n)*}\}$, $\dots$, $\{U_p^{(n)}, U_p^{(n)*}\}$, $\{D_1^{(n)}, \dots, D_q^{(n)}\}$ are almost surely asymptotically free.

This theorem explains now the findings of Figure 3. There we looked on two matrices of the form $X = D_1$ and $Y = UD_2U^*$, where D_1 and D_2 are diagonal matrices (each with asymptotic eigenvalue distribution $\mu_{\diamond}$) and U is a random Haar unitary (corresponding to a random choice of the eigenspaces of Y). Hence by the above result $\{U, U^*\}$ and $\{D_1, D_2\}$ are asymptotically free. It is an easy exercise (just using the definition of freeness) to see that this implies that also D_1 and UD_2U^* are asymptotically free. Hence the asymptotic eigenvalue distribution of $X + Y$ is given by $\mu_{◆} \boxplus \mu_{◆}$, which we calculated as $\mu_{▼}$ in Example 4.

There are many more matrix examples which show asymptotic freeness (like general Wigner matrices, see [2,22]) or some more general notions of freeness (like “traffic independence”, see [9]).

Random matrix theory is at the moment a very active and fascinating subject in mathematics; to get an idea of its diversity, beauty, and depth one should have a look on the collection of survey articles on various aspects of random matrices in [1]. Free probability has brought to this theory, with its concept of freeness and its quite developed tool box, a very new perspective on random matrices; resulting, among others, in new and very general calculation techniques for the asymptotic eigenvalue distribution of many classes of random matrices; see, for example, [5,30].

Since random matrices are used in many applied fields as models for basic scenarios, methods from random matrix theory, in particular also free probability theory, have become an indispensible tool for calculations in such subjects. A prominent example of this type are wireless networks, where the channels connecting transmitter and receiver antenna are modelled by a random matrix and where the capacity (i.e., the amount of information which can be transmitted through such channels) depends on the eigenvalue distribution of the random matrix. For more information on the use of random matrices and free probability in such a context one should see [31,11].

5 Free Probability and the Asymptotics of Representations of the Symmetric Groups

On an intuitive level typical representations of S_n for large n are given by large matrices, which behave in many respects like random matrices; hence the asymptotics of operations on representations should have some relation to asymptotic operations on random matrices, and the latter is described by free probability theory. This relation between the asymptotics of representations and free probability was discovered and made precise by Biane in [7], see also [8]. In the following we will present some of his results.

In Example 1 we encoded the information about a representation corresponding to a Young diagram λ by a probability measure μ_λ on $\mathbb{R}$ according to equations (1) and (2). The relevance of this μ_λ comes from the fact that it is the eigenvalue distribution of a matrix $\Gamma(\lambda)$, which contains the information about the representation π_λ:

$$\Gamma(\lambda) = \frac{1}{\sqrt{n}} \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & \pi_\lambda(12) & \pi_\lambda(13) & \pi_\lambda(14) & \dots & \pi_\lambda(1n) \\ 1 & \pi_\lambda(12) & 0 & \pi_\lambda(23) & \pi_\lambda(24) & \dots & \pi_\lambda(2n) \\ 1 & \pi_\lambda(13) & \pi_\lambda(23) & 0 & \pi_\lambda(34) & \dots & \pi_\lambda(3n) \\ 1 & \pi_\lambda(14) & \pi_\lambda(24) & \pi_\lambda(34) & 0 & \dots & \pi_\lambda(4n) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \pi_\lambda(1n) & \pi_\lambda(2n) & \pi_\lambda(3n) & \pi_\lambda(4n) & \dots & 0 \end{pmatrix}$$

Note that the entries in this matrix, in particular also 0 and 1, are themselves $d \times d$-matrices (where d is the dimension of the representation π_λ); $\pi_\lambda(ij)$ denotes the matrix representing the transposition $(ij) \in S_n$.

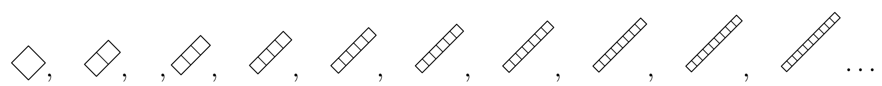
Taking traces of powers of this matrix gives

$$\mathrm{tr}((\Gamma(\lambda)^k) = \frac{1}{n^{k/2}} \frac{1}{n+1} \sum_{0 \le i_1 \ne i_2 \ne \dots \ne i_k \ne i_1 \le n} \chi_\lambda((i_1 i_2)(i_2 i_3) \cdots (i_k i_1)),$$

where $\chi_\lambda(\sigma) = \mathrm{tr}(\pi_\lambda(\sigma))$ $(\sigma \in S_n)$ is the normalized character corresponding to the representation π_λ and where (ij) denotes the identity permutation if either i or j is equal to 0. Working out the asymptotics of the conjugation classes which appear in the form $(i_1 i_2)(i_2 i_3) \cdots (i_k i_1)$, gives in the end for $\mathrm{tr}((\Gamma(\lambda)^k))$ a formula which is in leading order of the form (9), relating moments and free cumulants. This shows that the characters applied to the conjugation classes are in leading order given by the free cumulants of the distribution μ_λ. For making this more precise, we should note that asymptotically we are dealing with Young diagrams which are converging to some limiting Young curve after our rescaling with $\sqrt{n}$ in both linear directions. Hence we are dealing with "balanced" diagrams whose linear dimensions are of order $\sqrt{n}$.

Definition 9 Let $A > 1$. We say that a Young diagram $\lambda \vdash n$ (recall that this means that λ has n boxes) is *A-balanced* if its width (i.e., the maximum of its numbers of rows and columns) is at most $A\sqrt{n}$.

Very thin diagrams, like $\lambda_n = (n)$,



are not balanced and do not have a limit for $n \to \infty$ in this scaling, and we have nothing to say about the asymptotic behaviour of the representations corresponding to them.

For balanced diagrams however we can say a lot. A rigorous treatment of the above arguments yields, for example, the following theorem of Biane [7] about the asymptotic behaviour of the characters of balanced representations.

Theorem 4 *For all $A > 1$ and m positive integer, there exists a constant $K > 0$ such that, for all A-balanced Young diagrams $\lambda \vdash n$, and all permutations $\sigma \in S_n$ satisfying $|\sigma| \leq m$, one has*

$$\left| \chi_\lambda(\sigma) - |\lambda|^{-|\sigma|/2} \prod_{j \geq 2} \kappa_{j+1}(\lambda)^{l_j} \right| \leq K |\lambda|^{-1-|\sigma|/2},$$

if σ has l_2 cycles of length 2, l_3 cycles of length 3, etc., and $|\sigma| = \sum_{j \geq 2}(j-1)l_j$ is the minimal number to write σ as a product of transpositions; $\kappa_j(\lambda)$ $(j \geq 1)$ denote the free cumulants of μ_λ.

For some further results on bounds for characters see [13].

Also the asymptotic behaviour of operations on representations (like inductions, tensor products, etc) corresponding to balanced Young diagrams can be described via an analysis of the matrix $\Gamma(\lambda)$ in terms of operations coming from free probability theory on the corresponding measures μ_λ. Here is the precise statement, again due to Biane [7], concerning the induction of the tensor product.

Theorem 5 *For every $A > 1$ and p positive integer, there exists q_0 and K, $C_0 > 0$ such that for all $q \geq q_0$, all $C > C_0$, and all Young diagrams $\lambda_1 \vdash m$, $\lambda_2 \vdash n$ satisfying $m, n \geq q_0$ and width(λ_1), width$(\lambda_2) \leq A\sqrt{q}$, the subspace of all irreducible representations λ appearing in $\mathrm{Ind}_{S_m \times S_n}^{S_{m+n}} \pi_{\lambda_1} \otimes \pi_{\lambda_2}$ satisfying*

$$\left| \int_{\mathbb{R}} t^k d\mu_\lambda(t) - \int_{\mathbb{R}} t^k d(\mu_{\lambda_1} \boxplus \mu_{\lambda_2})(t) \right| \leq C q^{k/2-1/4} \qquad \textit{for all } k \leq n,$$

has dimension larger than $(1 - K/C^2) \cdot \dim \left(\mathrm{Ind}_{S_m \times S_n}^{S_{m+n}} \pi_{\lambda_1} \otimes \pi_{\lambda_2} \right)$.

This theorem makes our observation from Example 1 precise that asymptotically most irreducible components in ◆ × ◆ look like ▾, because we have seen in Example 4 that $\mu_{◆} \boxplus \mu_{◆} = \mu_{▾}$.

This theorem also shows that a diagram contributes with the dimension of its corresponding irreducible representation. Hence, when we chose a diagram at random in Example 1 in the decomposition, this actually means that we have to choose it with a probability which is proportional to the dimension of its irreducible representation; there is a combinatorial formula, the hook formula, which allows to calculate this dimension from the diagram.

6 Free Probability and Operator Algebras

In Section 3.1 we motivated and introduced the notion of freeness by rewriting the algebraic freeness of groups in a free product of groups in terms of the canonical state τ on the group algebra. When Voiculescu created free probability theory in [33] he was essentially led by this idea, however in an operator algebraic context. Operator algebras are $*$- algebras of bounded operators on a Hilbert space which are closed in some canonical topologies. (C^*-algebras are closed in the operator norm, von Neumann algebras are closed in the weak operator topology; the first topology is the operator version of uniform convergence, the latter of pointwise convergence.) Since the group algebra of a group can be represented on itself by bounded operators given by left multiplication (this is the regular representation of a group), one can take the closure in the appropriate topology of the group algebra and get thus C^*-algebras and von Neumann algebras corresponding to the group. The *group von Neumann algebra* arising from a group G in this way is usually denoted by $L(G)$. This construction, which goes back to the foundational papers of Murray and von Neumann in the 1930's, is, for G an infinite discrete group, a source of important examples in von Neumann algebra theory, and much of the progress in von Neumann algebra theory was driven by the desire to understand the relation between groups and their von Neumann algebras better. The group algebra consists of finite sums over group elements; going over to a closure means that we allow also some infinite sums. One should note that the weak closure, in the case of infinite groups, is usually much larger than the group algebra and it is very hard to control which infinite sums are added. Von Neumann algebras are quite large objects and their classification is notoriously difficult. Deep results of Connes show that amenable groups satisfying an additional technical (ICC) condition give always the same von Neumann algebra, the so-called *hyperfinite factor*. Going beyond amenable groups is a challenge.

A special and most prominent case for this are the free (non-commutative!) groups $\mathbf{F}_n$ on n generators, leading to the so-called *free group factors* $L(\mathbf{F}_n)$. Already Murray and von Neumann showed that the free group factors are not isomorphic to the hyperfinite factor; however, whether $L(\mathbf{F}_n)$ and $L(\mathbf{F}_m)$ are, for $n, m \geq 2$, isomorphic or not, is still one of the big open questions in the subject. It was the context of this problem, in which Voiculescu introduced the concept of "freeness"; led by the idea that $\mathbf{F}_n$ is the free product of n copies of $\mathbf{F}_1 = \mathbb{Z}$. Hence if one could make sense out of the corresponding phrase "$L(\mathbf{F}_n)$ is the free product of n copies of $L(\mathbb{Z})$" ($L(\mathbb{Z})$ is a commutative, and thus well understood, von Neumann algebra) then one might hope for a better understanding of $L(\mathbf{F}_n)$. The first step, to give rigorous meaning to a free product of von Neumann algebras, was achieved by the definition of freeness. Note that the canonical state τ on the group algebra extends continously to the group von Neumann algebra; and in the setting of infinite sums algebraic notions are usually not very helpful and we are really in need of a characterizaton of freeness in terms of the state τ.

That this notion of freeness does then also show up for random matrices, as shown later by Voiculescu in [36], was quite unsuspected and had an tremendous impact, both on operator algebras as well as on random matrix theory. The fact that freeness occurs for von Neumann algebras as well as for random matrices means that the former can be modeled asymptotically by the latter and this insight resulted in the first progress on the free group factors since Murray and von Neumann. In particular, Voiculescu showed in

[35] that the compression $L(\mathbf{F}_n)_r := pL(\mathbf{F}_n)p$ of $L(\mathbf{F}_n)$ by a projection p of trace $\mathrm{tr}(p) = r$ results in another free group factor; more precisely, one has $(L(\mathbf{F}_n))_{1/m} = L(\mathbf{F}_{1+m^2(n-1)})$. By introducing interpolated free group factors $L(\mathbf{F}_t)$ for all real $t > 1$, this formula could be extended by Dykema [12] and Radulescu [26] to any real $n, m > 1$, resulting in the following dichotomy. (The fundamental group of a von Neumann algebra M is the multiplicative subgroup of $\mathbb{R}^+$ consisting of all r for which $L(M)_r$ is isomorphic to M.)

Theorem 6 *We have exactly one of the following two possibilities.*
(i) All interpolating free group factors are isomorphic:

$$L(\mathbf{F}_s) \simeq L(\mathbf{F}_t) \qquad \text{for all } 1 < s, t \leq \infty.$$

In this case the fundamental group of each $L(\mathbf{F}_t)$ *is equal to* $\mathbb{R}^+$.
(ii) The interpolating free group factors are pairwise non-isomorphic:

$$L(\mathbf{F}_s) \not\simeq L(\mathbf{F}_t) \qquad \text{for all } 1 < s \neq t \leq \infty.$$

In this case the fundamental group of each $L(\mathbf{F}_t)$*, for* $t \neq \infty$*, is equal to* $\{1\}$.

Another spectacular result of Voiculescu about the structure of the free group factors, building on the free probability concept of "free entropy", was that free group factors do not have Cartan subalgebras [37]; thus settling a longstanding open question. There have been many more consequences of free probability arguments around subfactors [28], quantum groups [3, 27], invariant subspaces [17], or q-deformed von Neumann algebras [15], and much more.

7 Non-Commutative Distributions

During our journey through the typical asymptotic behaviour of representations or random sums of matrices we have come accross the structure of operator algebras, and we have encountered the distribution of non-commuting operators as one of the central objects.

In operator theory the shift from the concrete action of operators in the Hilbert space or from their algebraic structure to their distributions is quite some change of paradigm. Actually, the real part of the one sided shift has with respect to the canonical vacuum expectation a semicircular distribution. The one sided shift is arguably the most important operator in single operator theory, but apparently nobody before Voiculescu ever looked on its distribution.

We have seen various incarnations of non-commutative distributions, like:

- quite combinatorially, the collection of all mixed moments of the operators;
- more analytically, the restriction of a state to the algebra generated by our operators;

– in the one dimensional case we have also used analytic functions, like the Cauchy transform, for dealing with distributions; there exists also a beginning of a non-commutative function theory (aka free analysis) [18, 39, 40]), which is intended to provide us with analytic tools for dealing with distributions of non-commuting operators.

However, if one compares the situation with the classical, commutative setting then one still has the feeling that an important point of view is missing. Classically, the distribution of a k-tuple of random variables is a probability measure on $\mathbb{R}^k$. Though a positive linear functional on continuous functions is, via the Riesz representation theorem, the same as a probability measure, one surely has the feeling that for many (in particular, probabilistic) purposes having events and probabilities of those is maybe a better description than just being able to average over functions. In the non-commutative world we can average, but we do not (yet) have a good mental picture of non-commutative events and probabilities. Still we are dreaming of substitutes . . . and we would like to answer questions like: what is a good notion of density for non-commutative distributions (see [15]) or what are non-commutative zero-one laws or . . .

The domain of free probability, free analysis, and the other theories we are still hoping for consists in studying operators which do not commute. But this non-commutativity is of a special type. Usually there are no algebraic relations between our operators; the commutativity relations from the classical situation are not replaced by some deformations of commutativity, but they are just dropped. We like to address this as a "maximal non-commutative" situation. It seems that in this regime there exists a maximal non-commutative world in parallel to the classical commutative world.

We have only scratched the surface of such a theory. Starting from our commutative world we have sailed ahead into the non-commutative, passing some islands which show milder forms of non-commutativity, and finally we have landed at the shore of a new maximal non-commutative continent. We are looking forward to exploring it further.

Acknowledgements I would like to thank Philippe Biane, Franz Lehner, Hannah Markwig and Jonathan Novak, as well as an anonymous referee, for suggestions to improve the presentation.

References

1. Akemann, G., Baik, J., Di Francesco, P.: The Oxford handbook of random matrix theory. Oxford University Press (2011)
2. Anderson, G., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Cambridge Studies in Advanced Mathematics 118. Cambridge University Press (2010)
3. Banica, T., Speicher, R.: Liberation of orthogonal Lie groups. Advances in Mathematics **222**(4), 1461–1501 (2009)
4. Belinschi, S.T., Bercovici, H.: A new approach to subordination results in free probability. Journal d'Analyse Mathématique **101**(1), 357–365 (2007)
5. Belinschi, S.T., Mai, T., Speicher, R.: Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem. Journal für die reine und angewandte Mathematik (Crelles Journal) (2013)

6. Bercovici, H., Voiculescu, D.: Free convolution of measures with unbounded support. Indiana University Mathematics Journal **42**(3), 733–774 (1993)
7. Biane, P.: Representations of symmetric groups and free probability. Advances in Mathematics **138**(1), 126–181 (1998)
8. Biane, P.: Free probability and combinatorics. Proc. ICM 2002, Beijing **Vol. II**, 765–774 (2002)
9. Cébron, G., Dahlqvist, A., Male, C.: Universal constructions for spaces of traffics. ArXiv e-prints (2016)
10. Chistyakov, G., Götze, F.: The arithmetic of distributions in free probability theory. Open Mathematics **9**(5), 997–1050 (2011)
11. Couillet, R., Debbah, M.: Random matrix methods for wireless communications. Cambridge University Press (2011)
12. Dykema, K.: Interpolated free group factors. Pacific Journal of Mathematics **163**(1), 123–135 (1994)
13. Féray, V., Śniady, P.: Asymptotics of characters of symmetric groups related to Stanley character formula. Annals of Mathematics **173**(2), 887–906 (2011)
14. Figa-Talamanca, A., Steger, T.: Harmonic analysis for anisotropic random walks on homogeneous trees, 1994. Memoirs of the American Mathematical Society **110**(531), 531 (1994)
15. Guionnet, A., Shlyakhtenko, D.: Free monotone transport. Inventiones mathematicae **197**(3), 613–661 (2014)
16. Haagerup, U.: On Voiculescu's R- and S-transforms for free non-commuting random variables. In: Free probability theory, vol. 12, pp. 127–148. Fields Institute Communications (1997)
17. Haagerup, U.: Random matrices, free probability and the invariant subspace problem relative to a von Neumann algebra. Proc. ICM 2002, Beijing **Vol. I**, 273–290 (2002)
18. Kaliuzhnyi-Verbovetskyi, D.S., Vinnikov, V.: Foundations of free noncommutative function theory, vol. 199. American Mathematical Society (2014)
19. Kerov, S.V.: Transition probabilities for continual young diagrams and the markov moment problem. Functional Analysis and its Applications **27**(2), 104–117 (1993)
20. Lehner, F.: On the computation of spectra in free probability. Journal of Functional Analysis **183**(2), 451–471 (2001)
21. Maassen, H.: Addition of freely independent random variables. Journal of functional analysis **106**(2), 409–438 (1992)
22. Mingo, J., Speicher, R.: Free probability and Random matrices. Fields Institute Monographs. Springer (to appear)
23. Nica, A., Speicher, R.: Lectures on the combinatorics of free probability, *London Mathematical Society Lecture Note Series*, vol. 335. Cambridge University Press, Cambridge (2006). DOI 10.1017/CBO9780511735127. URL http://dx.doi.org/10.1017/CBO9780511735127
24. Novak, J., LaCroix, M.: Three lectures on free probability. arXiv preprint arXiv:1205.2097 (2012)
25. Novak, J., Sniady, P.: What is ... a free cumulant? Notices of the AMS **58**(2), 300–301 (2011)
26. Radulescu, F.: Random matrices, amalgamated free products and subfactors of the von Neumann algebra of a free group, of noninteger index. Inventiones mathematicae **115**(1), 347–389 (1994)
27. Raum, S., Weber, M.: The full classification of orthogonal easy quantum groups. Communications in Mathematical Physics **341**(3), 751–779 (2016)
28. Shlyakhtenko, D.: Free probability, planar algebras, subfactors and random matrices. Proc. ICM 2010 Hyderabad **Vol. III**, 1603–1623 (2010)
29. Speicher, R.: Multiplicative functions on the lattice of non-crossing partitions and free convolution. Mathematische Annalen **298**(1), 611–628 (1994)
30. Speicher, R.: Free probability and random matrices. Proc. ICM 2014 **Vol. III**, 477–501 (2014)
31. Tulino, A.M., Verdú, S.: Random matrix theory and wireless communications, vol. 1. Now Publishers Inc (2004)

32. Vershik, A.M., Kerov, S.V.: Asymptotic theory of characters of the symmetric group. Functional analysis and its applications **15**(4), 246–255 (1981)
33. Voiculescu, D.: Symmetries of some reduced free product C*-algebras. Operator Algebras and their Connections with Topology and Ergodic Theory pp. 556–588 (1985)
34. Voiculescu, D.: Addition of certain non-commuting random variables. Journal of functional analysis **66**(3), 323–346 (1986)
35. Voiculescu, D.: Circular and semicircular systems and free product factors. Operator algebras, unitary representations, enveloping algebras, and invariant theory (Paris, 1989) **92**, 45–60 (1990)
36. Voiculescu, D.: Limit laws for random matrices and free products. Inventiones mathematicae **104**(1), 201–220 (1991)
37. Voiculescu, D.: The analogues of entropy and of Fisher's information measure in free probability theory III: The absence of Cartan subalgebras. Geometric and Functional Analysis **6**(1), 172–199 (1996)
38. Voiculescu, D.: A strengthened asymptotic freeness result for random matrices with applications to free entropy. International Mathematics Research Notices **1998**, 41–63 (1998)
39. Voiculescu, D.V.: Aspects of free analysis. Japanese Journal of Mathematics **3**(2), 163–183 (2008)
40. Williams, J.D.: Analytic function theory for operator-valued free probability. Journal für die reine und angewandte Mathematik (Crelles Journal) (2013)
41. Woess, W.: Random walks on infinite graphs and groups, vol. 138. Cambridge university press (2000)