Chapter 22

Free Probability Theory

Roland Speicher

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

Abstract

Free probability theory was created by Dan Voiculescu around 1985, motivated by his efforts to understand special classes of von Neumann algebras. His discovery in 1991 that also random matrices satisfy asymptotically the *freeness* relation transformed the theory dramatically. Not only did this yield spectacular results about the structure of operator algebras, but it also brought new concepts and tools into the realm of random matrix theory. In the following we will give, mostly from the random matrix point of view, a survey on some of the basic ideas and results of free probability theory.

22.1 Introduction

Free probability theory allows one to deal with asymptotic eigenvalue distributions in situations involving several matrices. Let us consider two sequences A_N and B_N of selfadjoint $N \times N$ matrices such that both sequences have an asymptotic eigenvalue distribution for $N \to \infty$. We are interested in the asymptotic eigenvalue distribution of the sequence $f(A_N, B_N)$ for some non-trivial selfadjoint function f. In general, this will depend on the relation between the eigenspaces of A_N and of B_N. However, by the concentration of measure phenomenon, we expect that for large N this relation between the eigenspaces concentrates on *typical* or *generic positions*, and that then the asymptotic eigenvalue distribution of $f(A_N, B_N)$ depends in a deterministic way only on the asymptotic eigenvalue distribution of A_N and on the asymptotic eigenvalue distribution of B_N. Free probability theory replaces this vague notion of *generic position* by the mathematical precise concept of *freeness* and provides general tools for calculating the asymptotic distribution of $f(A_N, B_N)$ out of the asymptotic distribution of A_N and the asymptotic distribution of B_N.

22.2 The Moment Method for Several Random Matrices and the Concept of Freeness

The *empirical eigenvalue distribution* of a selfadjoint $N \times N$ matrix A is the probability measure on $\mathbb{R}$ which puts mass $1/N$ on each of the N eigenvalues λ_i of A, counted with multiplicity. If μ_A is determined by its moments then it can be recovered from the knowledge of all traces of powers of A:

$$\mathrm{tr}(A^k) = \frac{1}{N}\big(\lambda_1^k + \cdots + \lambda_N^k\big) = \int_{\mathbb{R}} t^k d\mu_A(t),$$

where by tr we denote the normalized trace on matrices (so that we have for the identity matrix 1 that $\mathrm{tr}(1) = 1$). This is the basis of the *moment method* which tries to understand the asymptotic eigenvalue distribution of a sequence of matrices by the determination of the asymptotics of traces of powers.

Definition 22.2.1 *We say that a sequence* $(A_N)_{N\in\mathbb{N}}$ *of* $N \times N$ *matrices* has an asymptotic eigenvalue distribution *if the limit* $\lim_{N\to\infty} \mathrm{tr}(A_N^k)$ *exists for all* $k \in \mathbb{N}$.

Consider now our sequences A_N and B_N, each of which is assumed to have an asymptotic eigenvalue distribution. We want to understand, in the limit $N \to \infty$, the eigenvalue distribution of $f(A_N, B_N)$, not just for one f, but for a wide class of different functions. By the moment method, this asks for the investigation of the limit $N \to \infty$ of $\mathrm{tr}\big(f(A_N, B_N)^k\big)$ for all $k \in \mathbb{N}$ and all f in our considered class of functions. If we choose for the latter all polynomials in non-commutative variables, then it is clear that the basic objects which we have to understand in this approach are the asymptotic *mixed moments*

$$\lim_{N\to\infty} \mathrm{tr}(A_N^{n_1} B_N^{m_1} \cdots A_N^{n_k} B_N^{m_k}) \qquad (k \in \mathbb{N};\ n_1, \ldots, n_k, m_1, \ldots, m_k \in \mathbb{N}). \tag{22.2.1}$$

Thus our fundamental problem is the following. If A_N and B_N each have an asymptotic eigenvalue distribution, and if A_N and B_N are in generic position, do the asymptotic mixed moments $\lim_{N\to\infty} \mathrm{tr}(A_N^{n_1} B_N^{m_1} \cdots A_N^{n_k} B_N^{m_k})$ exist? If so, can we express them in a deterministic way in terms of

$$\left(\lim_{N\to\infty} \operatorname{tr}(\varphi(A_N^k))\right)_{k\in\mathbb{N}} \qquad \text{and} \qquad \left(\lim_{N\to\infty} \operatorname{tr}(B_N^k)\right)_{k\in\mathbb{N}}. \tag{22.2.2}$$

Let us start by looking on the second part of the problem, namely by trying to find a possible relation between the mixed moments (22.2.1) and the moments (22.2.2). For this we need a simple example of matrix sequences A_N and B_N which we expect to be in generic position.

Whereas up to now we have only talked about sequences of matrices, we will now go over to random matrices. Namely, it is actually not clear how to produce two sequences of deterministic matrices whose eigenspaces are in generic position. However, it is much easier to produce two such sequences of random matrices for which we have almost surely a generic situation. Indeed, consider two independent random matrix ensembles A_N and B_N, each with almost surely a limiting eigenvalue distribution, and assume that one of them, say B_N, is a *unitarily invariant ensemble*, which means that the joint distribution of its entries does not change under unitary conjugation. This implies that taking $U_N B_N U_N^*$, for any unitary $N \times N$-matrix U_N, instead of B_N does not change anything. But then we can use this U_N to rotate the eigenspaces of B_N against those of A_N into a generic position, thus for typical realizations of A_N and B_N the eigenspaces should be in a generic position.

The simplest example of two such random matrix ensembles are two independent Gaussian random matrices A_N and B_N. In this case one can calculate everything concretely: in the limit $N \to \infty$, $\operatorname{tr}(A_N^{n_1} B_N^{m_1} \cdots A_N^{n_k} B_N^{m_k})$ is almost surely given by the number of non-crossing or planar pairings of the pattern

$$\underbrace{A \cdot A \cdots A}_{n_1\text{-times}} \cdot \underbrace{B \cdot B \cdots B}_{m_1\text{-times}} \cdots \underbrace{A \cdot A \cdots A}_{n_k\text{-times}} \cdot \underbrace{B \cdot B \cdots B}_{m_k\text{-times}},$$

which do not pair A with B. (A pairing is a decomposition of the pattern into pairs of letters; if we connect the two elements from each pair by a line, drawn in the half-plane below the pattern, then non-crossing means that we can do this without getting crossings between lines for different pairs.)

After some contemplation, it becomes obvious that this implies that the trace of a corresponding product of centered powers,

$$\begin{aligned}\lim_{N\to\infty} \operatorname{tr}\Big(&\big(A_N^{n_1} - \lim_{M\to\infty} \operatorname{tr}(A_M^{n_1}) \cdot 1\big) \cdot \big(B_N^{m_1} - \lim_{M\to\infty} \operatorname{tr}(B_M^{m_1}) \cdot 1\big) \cdots \\ &\cdots \big(A_M^{n_k} - \lim_{M\to\infty} \operatorname{tr}(A_M^{n_k}) \cdot 1\big) \cdot \big(B_N^{m_k} - \lim_{M\to\infty} \operatorname{tr}(B_M^{m_k}) \cdot 1\big)\Big)\end{aligned} \tag{22.2.3}$$

is given by the number of non-crossing pairings which do not pair A with B and for which, in addition, each group of A's and each group of B's is connected with some other group. It is clear that if we want to connect the groups in this way we will get some crossing between the pairs, thus there are actually no pairings of the required form and we have that the term (22.2.3) is equal to zero.

One might wonder what advantage is gained by trading the explicit formula for mixed moments of independent Gaussian random matrices for the implicit relation (22.2.3)? The drawback to the explicit formula for mixed moments of independent Gaussian random matrices is that the asymptotic formula for $\mathrm{tr}(A_N^{n_1} B_N^{m_1} \cdots A_N^{n_k} B_N^{m_k})$ will be different for different random matrix ensembles (and in many cases an explicit formula fails to exist). However, the vanishing of (22.2.3) remains valid for many matrix ensembles. The vanishing of (22.2.3) gives a precise meaning to our idea that the random matrices should be in generic position; it constitutes Voiculescu's definition of asymptotic freeness.

Definition 22.2.2 *Two sequences of matrices* $(A_N)_{N\in\mathbb{N}}$ *and* $(B_N)_{N\in\mathbb{N}}$ *are* asymptotically free *if we have the vanishing of* (22.2.3) *for all* $k \geq 1$ *and all* $n_1, m_1, \ldots, n_k, m_k \geq 1$.

Provided with this definition, the intuition that unitarily invariant random matrices should give rise to generic situations becomes now a rigorous theorem. This basic observation was proved by Voiculescu [Voi91] in 1991.

Theorem 22.2.3 *Consider* $N \times N$ *random matrices* A_N *and* B_N *such that: both* A_N *and* B_N *have almost surely an asymptotic eigenvalue distribution for* $N \to \infty$*;* A_N *and* B_N *are independent;* B_N *is a unitarily invariant ensemble. Then,* A_N *and* B_N *are almost surely asymptotically free.*

In order to prove this, one can replace A_N and B_N by A_N and $U_N B_N U_N^*$, where U_N is a Haar unitary random matrix (i.e., from the ensemble of unitary matrices equipped with the normalized Haar measure as probability measure); furthermore, one can then restrict to the case where A_N and B_N are deterministic matrices. In this form it reduces to showing almost sure asymptotic freeness between Haar unitary matrices and deterministic matrices. The proof of that statement proceeds then as follows. First one shows asymptotic freeness in the mean and then one strengthens this to almost sure convergence.

The original proof of Voiculescu [Voi91] for the first step reduced the asymptotic freeness for Haar unitary matrices to a corresponding statement for non-selfadjoint Gaussian random matrices; by realizing the Haar measure on the group of unitary matrices as the pushforward of the Gaussian measure under taking the phase. The asymptotic freeness result for Gaussian random matrices can be derived quite directly by using the genus expansion for their traces. Another more direct way to prove the averaged version of unitary freeness for Haar unitary matrices is due to Xu [Xu97] and relies on *Weingarten type formulas* for integrals over products of entries of Haar unitary matrices.

In the second step, in order to strengthen the above result to almost sure asymptotic freeness one can either [Voi91] invoke concentration of measure results of Gromov and Milman (applied to the unitary group) or [Spe93, Hia00] more specific estimates for the variances of the considered sequence of random variables.

Though unitary invariance is the most intuitive reason for having asymptotic freeness among random matrices, it is not a necessary condition. For example, the above theorem includes the case where B_N are Gaussian random matrices. If we generalize those to Wigner matrices (where the entries above the diagonal are i.i.d, but not necessarily Gaussian), then we loose the unitary invariance, but the conclusion of the above theorem still holds true. More precisely, we have the following theorem.

Theorem 22.2.4 *Let X_N be a selfadjoint Wigner matrix, such that the distribution of the entries is centered and has all moments, and let A_N be a random matrix which is independent from X_N. If A_N has almost surely an asymptotic eigenvalue distribution and if we have*

$$\sup_{N\in\mathbb{N}} \|A_N\| < \infty,$$

then A_N and X_N are almost surely asymptotically free.

The case where A_N consists of block diagonal matrices was treated by Dykema [Dyk93], for the general version see [Min10, And10].

22.3 Basic Definitions

The freeness relation, which holds for many random matrices asymptotically, was actually discovered by Voiculescu in a quite different context; namely canonical generators in operator algebras given in terms of free groups satisfy the same relation with respect to a canonical state, see Section 22.9.1. Free probability theory investigates these freeness relations abstractly, inspired by the philosophy that freeness should be considered and treated as a kind of non-commutative analogue of the classical notion of independence.

Some of the main probabilistic notions used in free probability are the following.

Notation 22.3.1 *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 addressed as* (non-commutative) random variables*, the numbers $\varphi(a_{i(1)} \cdots a_{i(n)})$ for such random variables $a_1, \dots, a_k \in \mathcal{A}$ are called* moments*, the collection of all moments is called the* joint distribution of $a_1, \dots, a_k$.

Definition 22.3.2 *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 $i \in I$, $a_i \in \mathcal{A}$. The elements $(a_i)_{i\in I}$ 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 a_i.*

Freeness, like classical independence, is a rule for calculating mixed moments from knowledge of the moments of individual variables. Indeed, one can easily show by induction that if $(\mathcal{A}_i)_{i\in I}$ are free with respect to φ, then φ restricted to the algebra generated by all $\mathcal{A}_i$, $i \in I$, is uniquely determined by $\varphi|_{\mathcal{A}_i}$ for all $i \in I$ and by the freeness condition. For example, if $\mathcal{A}$ and $\mathcal{B}$ are free, then one has for $a, a_1, a_2 \in \mathcal{A}$ and $b, b_1, b_2 \in \mathcal{B}$ that $\varphi(ab) = \varphi(a)\varphi(b)$, $\varphi(a_1ba_2) = \varphi(a_1a_2)\varphi(b)$, and $\varphi(a_1b_1a_2b_2) = \varphi(a_1a_2)\varphi(b_1)\varphi(b_2) + \varphi(a_1)\varphi(a_2)\varphi(b_1b_2) - \varphi(a_1)\varphi(b_1)\varphi(a_2)\varphi(b_2)$. Whereas the first two factorizations are the same as for the expectation of independent random variables, the last one is different, and more complicated, from the classical situation. It is important to note that freeness plays a similar role in the non-commutative world as independence plays in the classical world, but that freeness is not a generalization of independence: independent random variables can be free only in very trivial situations. Freeness is a theory for non-commuting random variables.

22.4 Combinatorial Theory of Freeness

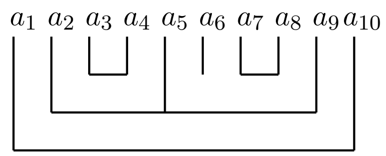
The defining relations for freeness from Def. 22.3.2 are quite implicit and not easy to handle directly. It has turned out that replacing moments by other quantities, so-called *free cumulants*, is advantageous for many questions. In particular, freeness is much easier to describe on the level of free cumulants. The relation between moments and cumulants is given by summing over non-crossing partitions. This combinatorial theory of freeness is due to Speicher [Spe94]; many consequences of this approach were worked out by Nica and Speicher, see [Nic06].

Definition 22.4.1 *For a unital linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ on a unital algebra $\mathcal{A}$ we define* cumulant functionals $\kappa_n : \mathcal{A}^n \to \mathbb{C}$ *(for all $n \geq 1$) by the* moment-cumulant relations

$$\varphi(a_1 \cdots a_n) = \sum_{\pi \in NC(n)} \kappa_\pi[a_1, \ldots, a_n]. \tag{22.4.1}$$

In equation (22.4.1) the summation is running over *non-crossing partitions* of the set $\{a_1, a_2, \ldots, a_n\}$; those are decompositions of that set into disjoint non-empty subsets, called *blocks*, such that there are no crossings between different blocks. In diagrammatic terms this means that if we draw the blocks of such a π below the points $a_1, a_2, \ldots, a_n$, then we can do this without having crossings in our picture. The contribution κ_π in (22.4.1) of such a non-crossing π is a product of cumulants corresponding to the block structure of π. For each block of π we have as a factor a cumulant which contains as arguments those a_i which are connected by that block.

An example of a non-crossing partition π for $n = 10$ is



In this case the blocks are $\{a_1, a_{10}\}$, $\{a_2, a_5, a_9\}$, $\{a_3, a_4\}$, $\{a_6\}$, and $\{a_7, a_8\}$; and the corresponding contribution κ_π in (22.4.1) is given by

$$\kappa_\pi[a_1, \ldots, a_{10}] = \kappa_2(a_1, a_{10}) \cdot \kappa_3\big(a_2, a_5, a_9\big) \cdot \kappa_2(a_3, a_4) \cdot \kappa_1(a_6) \cdot \kappa_2(a_7, a_8).$$

Note that in general there is only one term in (22.4.1) involving the highest cumulant κ_n, thus the moment cumulant formulas can inductively be resolved for the κ_n in terms of the moments. More concretely, the set of non-crossing partitions forms a lattice with respect to refinement order and the κ_n are given by the *Möbius inversion* of the formula (22.4.1) with respect to this order.

For $n = 1$, we get the mean, $\kappa_1(a_1) = \varphi(a_1)$ and for $n = 2$ we have the covariance, $\kappa_2(a_1, a_2) = \varphi(a_1 a_2) - \varphi(a_1)\varphi(a_2)$.

The relevance of the κ_n in our context is given by the following characterization of freeness.

Theorem 22.4.2 *Freeness is equivalent to the vanishing of mixed cumulants. More precisely, the fact that $(a_i)_{i \in I}$ are free is equivalent to: $\kappa_n(a_{i(1)}, \ldots, a_{i(n)}) = 0$ whenever $n \geq 2$ and there are k, l such that $i(k) \neq i(l)$.*

This description of freeness in terms of free cumulants is related to the planar approximations in random matrix theory. In a sense some aspects of this theory of freeness were anticipated (but mostly neglected) in the physics community in the paper [Cvi82].

22.5 Free Harmonic Analysis

For a meaningful harmonic analysis one needs some positivity structure for the non-commutative probability space $(\mathcal{A}, \varphi)$. We will usually consider selfadjoint random variables and φ should be positive. Formally, a good frame for this is a *C^*-probability space*, where $\mathcal{A}$ is a C^*-algebra (i.e., a norm-closed $*$-subalgebra of the algebra of bounded operators on a Hilbert space) and φ is a state, i.e. it is positive in the sense $\varphi(aa^*) \geq 0$ for all $a \in \mathcal{A}$. Concretely this means that our random variables can be realized as bounded operators on a Hilbert space and φ can be written as a vector state $\varphi(a) = \langle a\xi, \xi \rangle$ for some unit vector ξ in the Hilbert space.

In such a situation the distribution of a selfadjoint random variable a can be identified with a compactly supported probability measure μ_a on $\mathbb{R}$, via

$$\varphi(a^n) = \int_{\mathbb{R}} t^n d\mu_a(t) \qquad \text{for all } n \in \mathbb{N}.$$

22.5.1 Sums of free variables: the $\mathcal{R}$-transform

Consider two selfadjoint random variables a and b which are free. Then, by freeness, the moments of $a + b$ are uniquely determined by the moments of a and the moments of b.

Notation 22.5.1 *We say the distribution of $a + b$ is the* free convolution*, denoted by $\boxplus$, of the distribution of a and the distribution of b,*

$$\mu_{a+b} = \mu_a \boxplus \mu_b.$$

Notation 22.5.2 *For a random variable a we define its* Cauchy transform *G and its $\mathcal{R}$*-transform *$\mathcal{R}$ by*

$$G(z) = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{\varphi(a^n)}{z^{n+1}} \qquad \text{and} \qquad \mathcal{R}(z) = \sum_{n=1}^{\infty} \kappa_n(a, \ldots, a) z^{n-1}.$$

One can see quite easily that the moment-cumulant relations (22.4.1) are equivalent to the following functional relation

$$\frac{1}{G(z)} + \mathcal{R}(G(z)) = z. \tag{22.5.1}$$

Combined with the additivity of free cumulants under free convolution, which follows easily by the vanishing of mixed cumulants in free variables, this yields the following basic theorem of Voiculescu.

Theorem 22.5.3 *Let $G(z)$ be the Cauchy-transform of a, as defined in Notation 22.5.2 and define its $\mathcal{R}$-transform by the relation* (22.5.1). *Then we have*

$$\mathcal{R}^{a+b}(z) = \mathcal{R}^{a}(z) + \mathcal{R}^{b}(z)$$

if a and b are free.

We have defined the Cauchy and the $\mathcal{R}$-transform here only as formal power series. Also (22.5.1) is proved first as a relation between formal power series. But if a is a selfadjoint element in a C^*-probability space, then G is also the analytic function

$$G : \mathbb{C}^+ \to \mathbb{C}^-; \qquad G(z) = \varphi\left(\frac{1}{z-a}\right) = \int_{\mathbb{R}} \frac{1}{z-t} d\mu_a(t);$$

and one can also show that (22.5.1) defines then $\mathcal{R}$ as an analytic function on a suitably chosen subset of $\mathbb{C}^+$. In this form Theorem 22.5.3 is amenable to analytic manipulations and so gives an effective algorithm for calculating free convolutions. This can be used to calculate the asymptotic eigenvalue distribution of sums of random matrices which are asymptotically free.

Furthermore, by using analytic tools around the Cauchy transform (which exist for any probability measure on $\mathbb{R}$) one can extend the definition of and most results on free convolution to all probability measures on $\mathbb{R}$. See [Ber93, Voi00] for more details.

We would like to remark that the machinery of free convolution was also found around the same time, independently from Voiculescu and independently from each other, by different researchers in the context of random walks on the free product of groups: by Woess, by Cartwright and Soardi, and by McLaughlin; see, for example, [Woe86].

22.5.2 Products of free variables: the S-transform

Consider a, b free. Then, by freeness, the moments of ab are uniquely determined by the moments of a and the moments of b.

Notation 22.5.4 *We say the distribution of ab is the* free multiplicative convolution, *denoted by $\boxtimes$, of the distribution of a and the distribution of b,*

$$\mu_{ab} = \mu_a \boxtimes \mu_b.$$

Note: even if we start from selfadjoint a and b, their product ab is not selfadjoint, unless a and b commute (which is rarely the case, when a and b are free). Thus the above does not define an operation on probability measures on $\mathbb{R}$ in general. However, if one of the operators, say a, is positive (and thus μ_a

supported on $\mathbb{R}_+$), then $a^{1/2}ba^{1/2}$ makes sense; since it has the same moments as ab (note for this that the relevant state is a trace, as the free product of traces is tracial) we can identify μ_{ab} then with the probability measure $\mu_{a^{1/2}ba^{1/2}}$.

Again, Voiculescu introduced an analytic object which allows to deal effectively with this multiplicative free convolution.

Theorem 22.5.5 *Put $M_a(z) := \sum_{m=0}^{\infty} \varphi(a^m)z^m$ and define the S*-transform *of a by*

$$S_a(z) := \frac{1+z}{z} M_a^{<-1>}(z),$$

where $M^{<-1>}$ denotes the inverse of M under composition. Then we have

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

if a and b are free.

As in the additive case, the moment generating series M and the S-transform are not just formal power series, but analytic functions on suitably chosen domains in the complex plane. For more details, see [Ber93, Hia00].

22.5.3 The free central limit theorem

One of the first theorems in free probability theory, proved by Voiculescu in 1985, was the free analogue of the central limit theorem. Surprisingly, it turned out that the analogue of the Gaussian distribution in free probability theory is the semicircular distribution.

Definition 22.5.6 *Let $(\mathcal{A}, \varphi)$ be a C^*-probability space. A selfadjoint element $s \in \mathcal{A}$ is called* semicircular *(of variance 1) if its distribution μ_s is given by the probability measure with density $\frac{1}{2\pi}\sqrt{4-t^2}$ on the interval $[-2, +2]$. Alternatively, the moments of s are given by the Catalan numbers,*

$$\varphi(s^n) = \begin{cases} \frac{1}{k+1}\binom{2k}{k}, & \text{if } n = 2k \text{ even} \\ 0, & \text{if } n \text{ odd} \end{cases}$$

Theorem 22.5.7 *If ν is a compactly supported probability measure on $\mathbb{R}$ with vanishing mean and variance 1, then*

$$D_{1/\sqrt{N}} \nu^{\boxplus N} \Rightarrow \mu_s,$$

where D_α denotes the dilation of a measure by the factor α, and $\Rightarrow$ means weak convergence.

By using the analytic theory of $\boxplus$ for all, not necessarily compactly supported, probability measures on $\mathbb{R}$, the free central limit theorem can also be extended to this general situation.

The occurrence of the semicircular distribution as limit both in Wigner's semicircle law as well as in the free central limit theorem was the first hint of a relationship between free probability theory and random matrices. The subsequent development of this connection culminated in Voiculescu's discovery of asymptotic freeness between large random matrices, as exemplified in Theorem 22.2.3. When this contact was made between freeness and random matrices, the previously introduced $\mathcal{R}$- and S-transforms gave powerful new techniques for calculating asymptotic eigenvalue distributions of random matrices. For computational aspects of these techniques we refer to [Rao09], for applications in electrical engineering see [Tul04], and also Chapter 40.

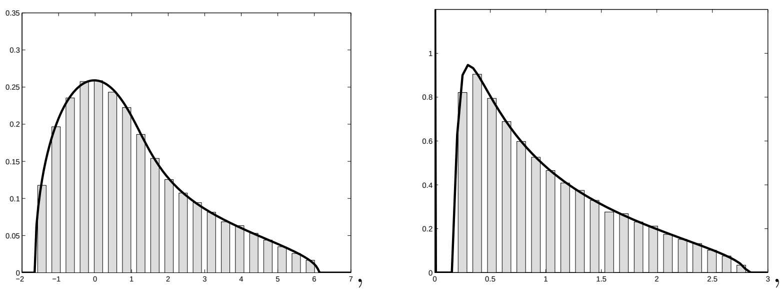


Figure 22.1: Comparison of free probability result with histogram of eigenvalues of an $N \times N$ random matrix, for $N = 2000$: (i) histogram of the sum of independent Gaussian and Wishart matrices, compared with the free convolution of semicircular and free Poisson distribution (rate $\lambda = 1/2$), calculated by using the $\mathcal{R}$-transform; (ii) histogram of the product of two independent Wishart matrices, compared with the free multiplicative convolution of two free Poisson distributions (both with rate $\lambda = 5$), calculated by using the S-transform

22.5.4 Free Poisson distribution and Wishart matrices

There exists a very rich free parallel of classical probability theory, of which the free central limit theorem is just the starting point. In particular, one has the free analogue of infinitely divisible and of stable distributions and corresponding limit theorems. For more details and references see [Ber99, Voi00].

Let us here only present as another instance of this theory the free Poisson distribution. As with the semicircle distribution, the free counterpart of the Poisson law, which is none other than the Marchenko-Pastur distribution, appears very naturally as the asymptotic eigenvalue distribution of an important class of random matrices, namely Wishart matrices.

As in the classical theory the Poisson distribution can be described by a limit theorem. The following statement deals directly with the more general notion of a compound free Poisson distribution.

Proposition 22.5.8 *Let $\lambda \geq 0$ and ν a probability measure on $\mathbb{R}$ with compact support. Then the weak limit for $N \to \infty$ of*

$$\left(\left(1-\frac{\lambda}{N}\right)\delta_0+\frac{\lambda}{N}\nu\right)^{\boxplus N}$$

has free cumulants $(\kappa_n)_{n\geq 1}$ which are given by $\kappa_n = \lambda \cdot m_n(\nu)$ $(n \geq 1)$ (m_n denotes here the n-th moment) and thus an $\mathcal{R}$-transform of the form

$$\mathcal{R}(z)=\lambda\int_{\mathbb{R}}\frac{x}{1-xz}\,d\nu(x).$$

Definition 22.5.9 *The probability measure appearing in the limit of Prop. 22.5.8 is called a* compound free Poisson distribution *with rate λ and jump distribution ν.*

Such compound free Poisson distributions show up in the random matrix context as follows. Consider: rectangular Gaussian $M \times N$ random matrices $X_{M,N}$, where all entries are independent and identically distributed according to a normal distribution with mean zero and variance $1/N$; and a sequence of deterministic $N \times N$ matrices T_N such that the limiting eigenvalue distribution μ_T of T_N exists. Then almost surely, for $M, N \to \infty$ such that $N/M \to \lambda$, the limiting eigenvalue distribution of $X_{M,N}T_N X^*_{M,N}$ exists, too, and it is given by a compound free Poisson distribution with rate λ and jump distribution μ_T.

One notes that the above frame of rectangular matrices does not fit directly into the theory presented up to now (so it is, e.g., not clear what asymptotic freeness between $X_{M,N}$ and T_N should mean). However, rectangular matrices can be treated in free probability by either embedding them into bigger square matrices and applying some compressions at appropriate stages or, more directly, by using a generalization of free probability due to Benaych-Georges [Ben09] which is tailor-made to deal with rectangular random matrices.

22.6 Second Order Freeness

Asymptotic freeness of random matrices shows that the mixed moments of two ensembles in generic position are deterministically calculable from the moments of each individual ensemble. The formula for the calculation of the mixed moments is the essence of the concept of freeness. The same philosophy applies also to finer questions about random matrices, most notably to *global fluctuations of linear statistics*. With this we mean the following: for many examples (like Gaussian or Wishart) of random matrices A_N the magnified fluctuations of traces around the limiting value, $N\big(\mathrm{tr}(A_N^k) - \lim_{M\to\infty} \mathrm{tr}(A_M^k)\big)$, form asymptotically a Gaussian family. If we have two such ensembles in generic position (e.g., if they are independent and one of them is unitarily invariant), then this is also true for mixed traces and the covariance of mixed traces is determined in a deterministic way by the covariances for each of the two ensembles separately. The formula for the calculation of the mixed covariances constitutes the definition of the concept of *second order freeness*. There exist again cumulants and an R-transform on this level, which allow explicit calculations. For more details, see [Col07, Min10].

22.7 Operator-Valued Free Probability Theory

There exists a generalization of free probability theory to an *operator-valued* level, where the complex numbers $\mathbb{C}$ and the expectation state $\varphi : \mathcal{A} \to \mathbb{C}$ are replaced by an arbitrary algebra $\mathcal{B}$ and a conditional expectation $E : \mathcal{A} \to \mathcal{B}$. The formal structure of the theory is very much the same as in the scalar-valued case, one only has to take care of the fact that the "scalars" from $\mathcal{B}$ do not commute with the random variables.

Definition 22.7.1 *1) Let $\mathcal{A}$ be a unital algebra and consider a unital subalgebra $\mathcal{B} \subset \mathcal{A}$. A linear map $E : \mathcal{A} \to \mathcal{B}$ is a* conditional expectation *if $E[b] = b$ for all $b \in \mathcal{B}$ and $E[b_1 a b_2] = b_1 E[a] b_2$ for all $a \in \mathcal{A}$ and all $b_1, b_2 \in \mathcal{B}$. An* operator-valued probability space *consists of $\mathcal{B} \subset \mathcal{A}$ and a conditional expectation $E : \mathcal{A} \to \mathcal{B}$.*

2) Consider an operator-valued probability space $\mathcal{B} \subset \mathcal{A}$, $E : \mathcal{A} \to \mathcal{B}$. Random variables $(x_i)_{i\in I} \subset \mathcal{A}$ are free with respect to E *or* free with amalgamation over $\mathcal{B}$ *if $E[p_1(x_{i(1)}) \cdots p_k(x_{i(k)})] = 0$, whenever $k \in \mathbb{N}$, p_j are elements from the algebra generated by $\mathcal{B}$ and $x_{i(j)}$, neighboring elements are different, i.e., $i(1) \neq i(2) \neq \cdots \neq i(k)$, and we have $E[p_j(x_{i(j)})] = 0$ for all $j = 1, \dots, k$.*

Voiculescu introduced this operator-valued version of free probability theory in [Voi85] and provided in [Voi95] also a corresponding version of free convolution and $\mathcal{R}$-transform. A combinatorial treatment was given by Speicher [Spe98] who showed that the theory of free cumulants has also a nice counterpart in the operator-valued frame.

For $a \in \mathcal{A}$ we define its *(operator-valued) Cauchy transform* $G_a : \mathcal{B} \to \mathcal{B}$ by

$$G_a(b) := E[\frac{1}{b-a}] = \sum_{n\geq 0} E[b^{-1}(ab^{-1})^n].$$

The *operator-valued $\mathcal{R}$-transform* of a, $\mathcal{R}_a : \mathcal{B} \to \mathcal{B}$, can be defined as a power series in operator-valued free cumulants, or equivalently by the relation $bG(b) = 1 + \mathcal{R}(G(b)) \cdot G(b)$ or $G(b) = \big(b - \mathcal{R}(G(b))\big)^{-1}$. One has then as before: If x and y are free over $\mathcal{B}$, then $\mathcal{R}_{x+y}(b) = \mathcal{R}_x(b) + \mathcal{R}_y(b)$. Another form of this is the *subordination property* $G_{x+y}(b) = G_x\big[b - \mathcal{R}_y\big(G_{x+y}(b)\big)\big]$.

There exists also the notion of a *semicircular element* s in the operator-valued world. It is characterized by the fact that only its second order free cumulants are different from zero, or equivalently that its $\mathcal{R}$-transform is of the form $\mathcal{R}_s(b) = \eta(b)$, where $\eta : \mathcal{B} \to \mathcal{B}$ is the linear map given by $\eta(b) = E[sbs]$. Note that in this case the equation for the Cauchy transform reduces to

$$bG(b) = 1 + \eta[G(b)] \cdot G(b); \tag{22.7.1}$$

more generally, if we add an $x \in \mathcal{B}$, for which we have $G_x(b) = E[(b - x)^{-1}] = (b - x)^{-1}$, we have for the Cauchy transform of $x + s$ the implicit equation

$$G_{x+s}(b) = G_x\big[b - \mathcal{R}_s\big(G_{x+s}(b)\big)\big] = \big(b - \eta[G_{x+s}(b)] - x\big)^{-1}. \tag{22.7.2}$$

It was observed by Shlyakhtenko [Shl96] that operator-valued free probability theory provides the right frame for dealing with more general kind of random matrices. In particular, he showed that so-called *band matrices* become asymptotically operator-valued semicircular elements over the limit of the diagonal matrices.

Theorem 22.7.2 *Suppose that $A_N = A_N^*$ is an $N \times N$ random band matrix, i.e., $A_N = (a_{ij})_{i,j=1}^N$, where $\{a_{ij} \mid i \leq j\}$ are centered independent complex Gaussian random variables, with $E[a_{ij}\overline{a}_{ij}] = (1 + \delta_{ij}\sigma^2(i/N, j/N))/N$ for some $\sigma^2 \in L^\infty([0,1]^2)$. Let $\mathcal{B}_N$ be the diagonal $N \times N$ matrices, and embed $\mathcal{B}_N$ into $\mathcal{B} := L^\infty[0,1]$ as step functions. Let $B_N \in \mathcal{B}_N$ be selfadjoint diagonal matrices such that $B_N \to f \in L^\infty[0,1]$ in $\|\cdot\|_\infty$. Then the limit distribution of $B_N + A_N$ exists, and its Cauchy transform G is given by*

$$G(z) = \int_0^1 g(z,x)dx,$$

where $g(z,x)$ is analytic in z and satisfies

$$g(z,x) = \left[z - f(x) - \int_0^1 \sigma^2(y,x)g(z,y)dy\right]^{-1}. \tag{22.7.3}$$

Note that (22.7.3) is nothing but the general equation (22.7.2) specified to the situation $\mathcal{B} = L^\infty[0,1]$ and $\eta : L^\infty[0,1] \to L^\infty[0,1]$ acting as integration operator with kernel σ^2.

Moreover, Gaussian random matrices with a certain degree of correlation between the entries are also asymptotically semicircular elements over an appropriate subalgebra, see [Ras08].

22.8 Further Free-Probabilistic Aspects of Random Matrices

Free probability theory provides also new ideas and techniques for investigating other aspects of random multi-matrix models. In particular, Haagerup and Thorbjornsen [Haa02, Haa05] obtained a generalization to several matrices for a number of results concerning the largest eigenvalue of a Gaussian random matrix.

Much work is also devoted to deriving rigorous results about the large N limit of random multi-matrix models given by densities of the type

$$c_N e^{-N^2 \mathrm{tr} P(A_1,\dots,A_N)} d\lambda(A_1,\dots,A_n),$$

where $d\lambda$ is Lebesgue measure, $A_1,\dots,A_n$ are selfadjoint $N\times N$ matrices, and P a noncommutative selfadjoint polynomial. To prove the existence of that limit in sufficient generality is one of the big problems. For a mathematical rigorous treatment of such questions, see [Gui06].

Free Brownian motion is the large N limit of the Dyson Brownian motion model of random matrices (with independent Brownian motions as entries, compare Chapter 11). Free Brownian motion can be realized concretely in terms of creation and annihilation operators on a full Fock space (see section 22.9.1). There exists also a corresponding *free stochastic calculus* [Bia98]; for applications of this to multi-matrix models, see [Gui07].

There is also a surprising connection with the representation theory of the symmetric groups S_n. For large n, representations of S_n are given by large matrices which behave in some respects like random matrices. This was made precise by Biane who showed that many operations on representations of the symmetric group can asymptotically be described by operations from free probability theory, see [Bia02].

22.9 Operator Algebraic Aspects of Free Probability

A survey on free probability without the mentioning of at least some of its operator algebraic aspects would be quite unbalanced and misleading. We will highlight some of these operator algebraic facets in this last section. For the sake of brevity, we will omit the definitions of standard concepts from operator algebras, since these can be found elsewhere (we refer the reader to [Voi92, Voi05, Hia00] for more information on notions, as well as for references related to the following topics).

22.9.1 Operator Algebraic Models for Freeness

Free group factors

Let $G = \star_{i\in I} G_i$ be the free product of groups G_i. Let $L(G)$ denote the group von Neumann algebra of G, and φ the associated trace state, corresponding to the neutral element of the group. Then $L(G_i)$ can be identified with a subalgebra of $L(G)$ and, with respect to φ, these subalgebras $(L(G_i))_{i\in I}$ are free. This freeness is nothing but the rewriting in terms of φ what it means that the groups G_i are free as subgroups in G. The definition of freeness was modeled according to the situation occurring in this example. The *free* in free probability theory refers to this fact.

A special and most prominent case of these von Neumann algebras are the *free group factors* $L(\mathbb{F}_n)$, where $\mathbb{F}_n$ is the free group on n generators. One hopes to eventually be able to resolve the isomorphism problem: whether the free groups factors $L(\mathbb{F}_n)$ and $L(\mathbb{F}_m)$ are, for $n, m \geq 2$, isomorphic or not.

Creation and annihilation operators on full Fock spaces

Let $\mathcal{H}$ be a Hilbert space. The *full Fock space* over $\mathcal{H}$ is defined as $\mathcal{F}(\mathcal{H}) := \bigoplus_{n=0}^{\infty} \mathcal{H}^{\otimes n}$. The summand $\mathcal{H}^{\otimes 0}$ on the right-hand side of the last equation is a one-dimensional Hilbert space. It is customary to write it in the form $\mathbb{C}\Omega$ for a distinguished vector of norm one, which is called *the vacuum vector*. The vector state $\tau_{\mathcal{H}}$ on $B(\mathcal{F}(\mathcal{H}))$ given by the vacuum vector, $\tau_{\mathcal{H}}(T) := \langle T\Omega, \Omega\rangle$ $(T \in B(\mathcal{F}(\mathcal{H})))$, is called *vacuum expectation state*.

For each $\xi \in \mathcal{H}$, the operator $l(\xi) \in B(\mathcal{F}(\mathcal{H}))$ determined by the formula $l(\xi)\Omega = \xi$ and $l(\xi)\xi_1 \otimes \cdots \otimes \xi_n = \xi \otimes \xi_1 \otimes \cdots \otimes \xi_n$ for all $n \geq 1$, $\xi_1, \ldots, \xi_n \in \mathcal{H}$, is called the *(left) creation operator* given by the vector ξ. As one can easily verify, the adjoint of $l(\xi)$ is described by the formula: $l(\xi)^*\Omega = 0$, $l(\xi)^*\xi_1 = \langle \xi_1, \xi\rangle\Omega$, and $l(\xi)^*\xi_1 \otimes \cdots \otimes \xi_n = \langle \xi_1, \xi\rangle \xi_2 \otimes \cdots \otimes \xi_n$ and is called the *(left) annihilation operator* given by the vector ξ.

The relevance of these operators comes from the fact that orthogonality of vectors translates into free independence of the corresponding creation and annihilation operators.

Proposition 22.9.1 *Let $\mathcal{H}$ be a Hilbert space and consider the probability space $(B(\mathcal{F}(\mathcal{H})), \tau_{\mathcal{H}})$. Let $\mathcal{H}_1, \ldots, \mathcal{H}_k$ be a family of linear subspaces of $\mathcal{H}$, such that $\mathcal{H}_i \perp \mathcal{H}_j$ for $i \neq j$ $(1 \leq i, j \leq k)$. For every $1 \leq i \leq k$ let $\mathcal{A}_i$ be the unital C^*-subalgebra of $B(\mathcal{F}(\mathcal{H}))$ generated by $\{l(\xi) : \xi \in \mathcal{H}_i\}$. Then $\mathcal{A}_1, \ldots, \mathcal{A}_k$ are freely independent in $(B(\mathcal{F}(\mathcal{H})), \tau_{\mathcal{H}})$.*

Also semicircular elements show up very canonically in this frame; namely, if we put $l := l(\xi)$ for a unit vector $\xi \in \mathcal{H}$, then $l + l^*$ is a semicircular element of variance 1. More generally, one has that $l + f(l^*)$ has $\mathcal{R}$-transform $\mathcal{R}(z) = f(z)$

(for f a polynomial, say). This, together with the above proposition, was the basis of Voiculescu's proof of Theorem 22.5.3. Similarly, a canonical realization for the S-transform is $(1 + l)g(l^*)$, for which one has $S(z) = 1/g(z)$. This representation is due to Haagerup who used it for a proof of Theorem 22.5.5.

22.9.2 Free Entropy

Free entropy is, as the name suggests, the counterpart of entropy in free probability theory. The development of this concept is at present far from complete. The current state of affairs is that there are two distinct approaches to free entropy. These should give isomorphic theories, but at present we only know that they coincide in a limited number of situations. The first approach to a theory of free entropy is via *microstates*. This goes back to the statistical mechanics roots of entropy via the Boltzmann formula and is related to the theory of large deviations. The second approach is *microstates free*. This draws its inspiration from the statistical approach to classical entropy via the notion of Fisher information. We will in the following only consider the first approach via microstates, as this relates directly with random matrix questions.

Wigner's semicircle law states that as $N \to \infty$ the empirical eigenvalue distribution μ_{A_N} of an $N \times N$ Gaussian random matrix A_N converges almost surely to the semicircular distribution μ_W, i.e., the probability that μ_{A_N} is in any fixed neighborhood of the semicircle converges to 1. We are now interested in the deviations from this: What is the rate of decay of the probability that μ_{A_N} is close to ν, where ν is an arbitrary probability measure? We expect that this probability behaves like $e^{-N^2 I(\nu)}$, for some *rate function* I vanishing at the semicircle distribution. By analogy with the classical theory of large deviations, I should correspond to a suitable notion of free entropy. This heuristics led Voiculescu to define in [Voi93] the *free entropy* χ in the case of one variable to be

$$\chi(\nu) = \iint \log|s - t| d\nu(s) d\nu(t) + \frac{3}{4} + \frac{1}{2}\log 2\pi.$$

Inspired by this, Ben-Arous and Guionnet proved in [Ben97] a rigorous version of a large deviation for Wigner's semicircle law, where the rate function $I(\nu)$ is, up to a constant, given by $-\chi(\nu) + \frac{1}{2}\int t^2 d\nu(t)$.

Consider now the case of several matrices. By Voiculescu's generalization of Wigner's theorem we know that n independent Gaussian random matrices $A_N^{(1)}, \ldots, A_N^{(n)}$ converge almost surely to a freely independent family $s_1, \ldots, s_n$ of semicircular elements. Similarly as for the case of one matrix, large deviations from this limit should be given by

$$\text{Prob}\left\{(A_N^{(1)}, \ldots, A_N^{(n)}) : \text{distr}((A_N^{(1)}, \ldots, A_N^{(n)}) \approx \text{distr}(a_1, \ldots, a_n)\right\} \sim e^{-N^2 I(a_1, \ldots, a_n)},$$

where $I(a_1, \ldots, a_n)$ should be related to the *free entropy* of the random variables

$a_1, \ldots, a_n$. Since the distribution $\mathrm{distr}(a_1, \ldots, a_n)$ of several non-commuting random variables $a_1, \ldots, a_n$ is a mostly combinatorial object (consisting of the collection of all joint moments of these variables), it is much harder to deal with these questions and, in particular, to get an analytic formula for I. Essentially, the above heuristics led Voiculescu to the following definition [Voi94b] of a free entropy for several variables.

Definition 22.9.2 *Given a tracial W^*-probability space (M, τ) (i.e., M a von Neumann algebra and τ a faithful and normal trace), and an n-tuple $(a_1, \ldots, a_n)$ of selfadjoint elements in M, put*

$$\begin{aligned}\Gamma(a_1, \ldots, a_n; N, r, \epsilon) :=& \\ &\big\{(A_1, \ldots, A_n) \in M_N(\mathbb{C})^n_{sa} : |\mathrm{tr}(A_{i_1} \ldots A_{i_k}) - \tau(a_{i_1} \ldots a_{i_k})| \le \epsilon \\ &\qquad\qquad \textit{for all } 1 \le i_1, \ldots, i_k \le n, 1 \le k \le r\big\}\end{aligned}$$

In words, $\Gamma(a_1, \ldots, a_n; N, r, \epsilon)$ is the set of all n-tuples of $N \times N$ selfadjoint matrices which approximate the mixed moments of the selfadjoint elements $a_1, \ldots, a_n$ of length at most r to within ϵ.

Let Λ denote Lebesgue measure on $M_N(\mathbb{C})^n_{sa}$. Define

$$\chi(a_1, \ldots, a_n; r, \epsilon) := \limsup_{N \to \infty} \frac{1}{N^2} \log \Lambda(\Gamma(a_1, \ldots, a_n; N, r, \epsilon) + \frac{n}{2} \log N,$$

and

$$\chi(a_1, \ldots, a_n) := \lim_{\substack{r \to \infty \\ \epsilon \to 0}} \chi(a_1, \ldots, a_n; r, \epsilon).$$

The function χ is called the free entropy.

Many of the expected properties of this quantity χ have been established (in particular, it behaves additive with respect to free independence), and there have been striking applications to the solution of some old operator algebra problems. A celebrated application of free entropy was Voiculescu's proof of the fact that free group factors do not have Cartan subalgebras (thus settling a longstanding open question). This was followed by several results of the same nature; in particular, Ge showed that $L(\mathbb{F}_n)$ cannot be written as a tensor product of two II_1 factors. The rough idea of proving the absence of some property for the von Neumann algebra $L(\mathbb{F}_n)$ using free entropy is the following string of arguments: finite matrices approximating in distribution any set of generators of $L(\mathbb{F}_n)$ should also show an approximate version of the considered property; one then has to show that there are not many finite matrices with this approximate property; but for $L(\mathbb{F}_n)$ one has many matrices, given by independent Gaussian random matrices, which approximate its canonical generators.

However, many important problems pertaining to free entropy remain open. In particular, we only have partial results concerning the relation to large deviations for several Gaussian random matrices. For more information on those and other aspects of free entropy we refer to [Voi02, Bia03, Gui04].

22.9.3 Other Operator Algebraic Applications of Free Probability 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 that a compression of some $L(\mathbb{F}_n)$ results in another free group factor; more precisely, one has $(L(\mathbb{F}_n))_{1/m} = L(\mathbb{F}_{1+m^2(n-1)})$. By introducing interpolated free group factors $L(\mathbb{F}_t)$ for all real $t > 1$, this formula could be extended by Dykema and Radulescu to any real $n, m > 1$, resulting in the following dichotomy: One has that either all free group factors $L(\mathbb{F}_n)$ $n \geq 2$ are isomorphic or that they are pairwise not isomorphic.

There exist also type III versions of the free group factors; these free analogues of the Araki-Woods factors were introduced and largely classified by Shlyakhtenko.

The study of free group factors via free probability techniques has also had an important application to subfactor theory. Not every set of data for a subfactor inclusion can be realized in the hyperfinite factor; however, work of Shlyakhtenko, Ueda, and Popa has shown that this is possible using free group factors.

By relying on free probability techniques and ideas, Haagerup achieved also a crucial break-through on the famous invariant subspace problem: every operator in a II_1 factor whose Brown measure (which is a generalization of the spectral measure composed with the trace to non-normal operators) is not concentrated in one point has non-trivial closed invariant subspaces affiliated with the factor.

Acknowledgements: This work was supported by a Discovery Grant from NSERC.

References

[And10] G. Anderson, A. Guionnet, and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge University Press (to appear)

[Ben97] G. Ben-Arous and A. Guionnet, Prob. Th. Rel. Fields **108** (1997) 517

[Ben09] F. Benaych-Georges, Prob. Th. Rel. Fields **144** (2009) 471

[Ber93] H. Bercovici and D. Voiculescu, Indiana Univ. Math. J. **42** (1993) 733

[Ber99] H. Bercovici and V. Pata (with an appendix by P. Biane), Ann. of Math. **149** (1999) 1023

[Bia98] P. Biane and R. Speicher, Prob. Th. Relat. Fields **112** (1998) 373

[Bia02] P. Biane, Proceedings of the International Congress of Mathematicians, Beijing 2002, Vol. 2 (2002) 765

[Bia03] P. Biane, M. Capitaine, and A. Guionnet, Invent. Math. **152** (2003) 433

[Cvi82] P. Cvitanovic, P.G. Lauwers, and P.N. Scharbach, Nucl. Phys. B **203** (1982) 385

[Col07] B. Collins, J. Mingo, P. Sniady, and R., Speicher, Documenta Math. **12** (2007) 1

[Dyk93] K. Dykema, J. Funct. Anal. **112** (1993) 31

[Gui04] A. Guionnet, Probab. Surv. **1** (2004) 72

[Gui06] A. Guionnet, Proceedings of the International Congress of Mathematicians, Madrid 2006, Vol. III (2006) 623

[Gui07] A. Guionnet and D. Shlyakhtenko, preprint, arXiv:math/0701787 (2007)

[Haa02] U. Haagerup, Proceedings of the International Congress of Mathematicians, Beijing 2002, Vol 1 (2002) 273

[Haa05] U. Haagerup, S. Thorbjørnsen, Ann. of Math. **162** (2005) 711

[Hia00] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables and Entropy*, Math. Surveys and Monogr. 77, AMS 2000

[Nic06] A. Nica and R. Speicher, *Lectures on the Combinatorics of Free Probability*, London Mathematical Society Lecture Note Series, vol. 335, Cambridge University Press, 2006

[Min10] J. Mingo and R. Speicher, *Free Probability and Random Matrices*, Fields Monograph Series (to appear)

[Rao09] N. R. Rao and A. Edelman, Foundations of Computational Mathematics (to appear)

[Ras08] R. Rashidi Far, T. Oraby, W. Bryc, and R. Speicher, IEEE Trans. Inf. Theory **54** (2008) 544

[Shl96] D. Shlyakhtenko, IMRN 1996 **1996** 1013

[Spe93] R. Speicher, Publ. RIMS **29** (1993) 731

[Spe94] R. Speicher, Math. Ann. **298** (1994) 611

[Spe98] R. Speicher, *Combinatorial theory fo the free product with amalgamtion and operator-valued free probability theory*, Memoirs of the AMS **627** 1998

[Tul04] A. Tulino, S. Verdu, *Random matrix theory and wirless communications*, Foundations and Trends in Communications and Information Theory **1** (2004)

[Voi85] D. Voiculescu, in *Operator Algebras and their Connections with Topology and Ergodic Theory*, Lecture Notes in Math. 1132 (1985), Springer Verlag, 556

[Voi91] D. Voiculescu, Invent. Math. **104** (1991) 201

[Voi92] D. Voiculescu, K. Dykema, and A. Nica, *Free Random Variables*, CRM Monograph Series, Vol. 1, AMS 1992

[Voi93] D. Voiculecu, Comm. Math. Phys. **155** (1993) 71

[Voi94a] D. Voiculescu, Proceedings of the International Congress of Mathematicians, Zürich 1994, 227

[Voi94b] D. Voiculescu, Invent. Math. **118** (1994) 411

[Voi95] D. Voiculescu, Asterisque **223** (1995) 243

[Voi00] D. Voiculescu, in *Lectures on Probabiltiy Theory and Statistics (Saint-Flour, 1998)*, Lecture Notes in Mathematics 1738, Springer, 2000, 279

[Voi02] D. Voiculescu, Bulletin of the London Mathematical Society **34** (2002) 257

[Voi05] D. Voiculescu, Reports on Mathematical Physics **55** (2005) 127

[Woe86] W. Woess, Bollettino Un. Mat. Ital. **5-B** (1986) 961

[Xu97] F. Xu, Commun. Math. Phys. **190** (1997) 287