Methods of classical and free probability

Teo Banica

Department of Mathematics, University of Cergy-Pontoise, F-95000 Cergy-Pontoise, France. teo.banica@gmail.com

2010 *Mathematics Subject Classification.* 60B20

Key words and phrases. Free probability, Semicircle law

Abstract. This is a joint introduction to classical and free probability, which are twin sisters. We discuss in detail the foundations and main results of both theories, by insisting on their common features, and by using a light formalism, based on standard calculus. We include as well a brief discussion of more advanced aspects.

Preface

This is an introduction to classical and free probability, by insisting on their common features. The central result in classical probability is the Central Limit Theorem (CLT), the limiting measure being the centered normal law of variance $t > 0$:

$$g_t = \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} dx$$

In free probability, we have a Free Central Limit Theorem (FCLT), the limiting measure being the Wigner semicircle law of parameter $t > 0$:

$$\gamma_t = \frac{1}{2\pi t} \sqrt{4t^2 - x^2} dx$$

In the discrete setting now, the main result in classical probability is the Poisson Limit Theorem (PLT), the limiting measure being the Poisson law of parameter $t > 0$:

$$p_t = e^{-t} \sum_k \frac{t^k}{k!} \delta_k$$

In free probability, we have a Free Poisson Limit Theorem (FPLT), the limiting measure being the Marchenko-Pastur law of parameter $t > 0$:

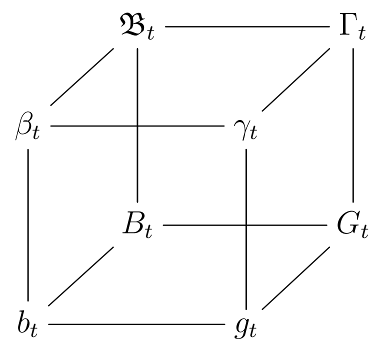
$$\pi_t = \max(1 - t, 0)\delta_0 + \frac{\sqrt{4t - (x - 1 - t)^2}}{2\pi x} dx$$

To summarize, we have discrete and continuous, and classical and free, probabilistic limiting theorems, the limiting measures being as follows:

$$\begin{matrix} \pi_t & - & \gamma_t \\ | & & | \\ p_t & - & g_t \end{matrix}$$

It is possible to further expand this picture, with complexifications, perturbations and so on, and explaining this will be as well a main purpose of the present book.

As a main result regarding the correspondence between classical and free, we have the following cubic diagram, whose faces correspond to the main real/complex, classical/free and discrete/continuous probabilistic limiting theorems:



To be more precise, on the right we have the Gaussian and Wigner laws g_t, γ_t and their complex versions G_t, Γ_t, and on the left we have real/complex, classical/free Bessel laws, replacing the Poisson laws p_t, π_t. We will explain this dagram, in this book.

This book is organized in 4 parts, as follows:

Part I - Here we discuss the main limiting theorems in classical probability, with a quick look into finite groups and random matrices as well.

Part II - Here we discuss the basics of free probability, notably with models for freeness, and with the theory of semicircular and circular variables.

Part III - Here we discuss more advanced aspects of free probability, mostly combinatorial, and we explain the bijection with classical probability.

Part IV - Here we discuss advanced aspects of classical and free probability, in connection with quantum groups, geometry and Wishart matrices.

This book contains, besides the basic of the theory, a few recent contributions as well, in relation with quantum algebra and with the modified Wishart matrices. I am grateful to Mireille Capitaine, Benoît Collins, Steve Curran, Ion Nechita and Roland Speicher, for substantial joint work on the subject. Many thanks go as well to my cats. No serious science can be done without advice from a cat or two.

Contents

Part I

Classical probability

CHAPTER 1

The normal law

1a. Probability theory

In this chapter and in the next one we discuss the foundations of probability theory, with a random matrix and quantum mechanics motivation in mind.

Generally speaking, probability theory is best learned by flipping coins and throwing dices. At a more advanced level, which is playing cards, we have:

THEOREM 1.1. *The probabilities at poker are as follows:*

(1) *One pair:* 0.533.
(2) *Two pairs:* 0.120.
(3) *Three of a kind:* 0.053.
(4) *Full house:* 0.006.
(5) *Straight:* 0.005.
(6) *Four of a kind:* 0.001.
(7) *Flush:* 0.000.
(8) *Straight flush:* 0.000.

PROOF. Let us consider indeed our deck of 32 cards:

$$7, 8, 9, 10, J, Q, K, A$$

The total number of possibilities is:

$$\binom{32}{5} = \frac{32 \cdot 31 \cdot 30 \cdot 29 \cdot 28}{2 \cdot 3 \cdot 4 \cdot 5} = 32 \cdot 31 \cdot 29 \cdot 7$$

(1) For having a pair, the number of possibilities is:

$$N = \binom{8}{1}\binom{4}{2} \times \binom{7}{3}\binom{4}{1}^3 = 8 \cdot 6 \cdot 35 \cdot 64$$

Thus, the probability of having a pair is:

$$P = \frac{8 \cdot 6 \cdot 35 \cdot 64}{32 \cdot 31 \cdot 29 \cdot 7} = \frac{6 \cdot 5 \cdot 16}{31 \cdot 29} = \frac{480}{899} = 0.533$$

(2) For having two pairs, the number of possibilities is:

$$N = \binom{8}{2}\binom{4}{2}^2 \times \binom{24}{1} = 28 \cdot 36 \cdot 24$$

Thus, the probability of having two pairs is:

$$P = \frac{28 \cdot 36 \cdot 24}{32 \cdot 31 \cdot 29 \cdot 7} = \frac{36 \cdot 3}{31 \cdot 29} = \frac{108}{899} = 0.120$$

(3) For having three of a kind, the number of possibilities is:

$$N = \binom{8}{1}\binom{4}{3} \times \binom{7}{2}\binom{4}{1}^2 = 8 \cdot 4 \cdot 21 \cdot 16$$

Thus, the probability of having three of a kind is:

$$P = \frac{8 \cdot 4 \cdot 21 \cdot 16}{32 \cdot 31 \cdot 29 \cdot 7} = \frac{3 \cdot 16}{31 \cdot 29} = \frac{48}{899} = 0.053$$

(4) For having full house, the number of possibilities is:

$$N = \binom{8}{1}\binom{4}{3} \times \binom{7}{1}\binom{4}{2} = 8 \cdot 4 \cdot 7 \cdot 6$$

Thus, the probability of having full house is:

$$P = \frac{8 \cdot 4 \cdot 7 \cdot 6}{32 \cdot 31 \cdot 29 \cdot 7} = \frac{6}{31 \cdot 29} = \frac{6}{899} = 0.006$$

(5) For having a straight, the number of possibilities is:

$$N = 4\left[\binom{4}{1}^4 - 4\right] = 16 \cdot 63$$

Thus, the probability of having a straight is:

$$P = \frac{16 \cdot 63}{32 \cdot 31 \cdot 29 \cdot 7} = \frac{9}{2 \cdot 31 \cdot 29} = \frac{9}{1798} = 0.005$$

(6) For having four of a kind, the number of possibilities is:

$$N = \binom{8}{1}\binom{4}{4} \times \binom{7}{1}\binom{4}{1} = 8 \cdot 7 \cdot 4$$

Thus, the probability of having four of a kind is:

$$P = \frac{8 \cdot 7 \cdot 4}{32 \cdot 31 \cdot 29 \cdot 7} = \frac{1}{31 \cdot 29} = \frac{1}{899} = 0.001$$

(7) For having a flush, the number of possibilities is:

$$N = 4\left[\binom{8}{4} - 4\right] = 4 \cdot 66$$

Thus, the probability of having a flush is:

$$P = \frac{4 \cdot 66}{32 \cdot 31 \cdot 29 \cdot 7} = \frac{33}{4 \cdot 31 \cdot 29 \cdot 7} = \frac{9}{25172} = 0.000$$

(8) For having a straight flush, the number of possibilities is:

$$N = 4 \cdot 4$$

Thus, the probability of having a straight flush is:

$$P = \frac{4 \cdot 4}{32 \cdot 31 \cdot 29 \cdot 7} = \frac{1}{2 \cdot 31 \cdot 29 \cdot 7} = \frac{1}{12586} = 0.000$$

Thus, we have obtained the numbers in the statement. □

Summarizing, probability is basically about binomials and factorials, and ultimately about numbers. We will see later on that, in connection with more advanced questions, of continuous nature, some standard calculus comes into play as well.

Let us discuss now the general theory. The fundamental result in probability is the Central Limit Theorem (CLT), and our first task will be that of explaining this. With the idea in mind of doing things a bit abstractly, our starting point will be:

DEFINITION 1.2. *Let X be a probability space, that is to say, a space with a probability measure, and with the corresponding integration denoted $\mathbb{E}$, and called expectation.*

(1) *The random variables are the real functions $f \in L^\infty(X)$.*
(2) *The moments of such a variable are the numbers $M_k(f) = \mathbb{E}(f^k)$.*
(3) *The law of such a variable is the measure given by $M_k(f) = \int_\mathbb{R} x^k d\mu_f(x)$.*

Here the fact that μ_f exists indeed is not trivial. By linearity, we would like to have a real probability measure making hold the following formula, for any $P \in \mathbb{R}[X]$:

$$\mathbb{E}(P(f)) = \int_\mathbb{R} P(x) d\mu_f(x)$$

By using a continuity argument, it is enough to have this formula for the characteristic functions χ_I of the arbitrary measurable sets of real numbers $I \subset \mathbb{R}$:

$$\mathbb{E}(\chi_I(f)) = \int_\mathbb{R} \chi_I(x) d\mu_f(x)$$

Thus, we would like to have a measure μ_f such that:

$$\mathbb{P}(f \in I) = \mu_f(I)$$

But this latter formula can serve as a definition for μ_f, with the axioms of real probability measures being trivially satisifed, and so we are done.

Next in line, we need to talk about independence. Once again with the idea of doing things a bit abstractly, the definition here is as follows:

DEFINITION 1.3. *Two variables $f,g\in L^\infty(X)$ are called independent when*

$$\mathbb E(f^kg^l)=\mathbb E(f^k)\cdot\mathbb E(g^l)$$

happens, for any $k,l\in\mathbb N$.

Once again, this definition hides some non-trivial things. Indeed, by linearity, we would like to have a formula as follows, valid for any polynomials $P,Q\in\mathbb R[X]$:

$$\mathbb E(P(f)Q(g))=\mathbb E(P(f))\cdot\mathbb E(Q(g))$$

By continuity, it is enough to have this formula for characteristic functions χ_I,χ_J of the arbitrary measurable sets of real numbers $I,J\subset\mathbb R$:

$$\mathbb E(\chi_I(f)\chi_J(g))=\mathbb E(\chi_I(f))\cdot\mathbb E(\chi_J(g))$$

Thus, we are led to the usual definition of independence, namely:

$$\mathbb P(f\in I,g\in J)=\mathbb P(f\in I)\cdot\mathbb P(g\in J)$$

All this might seem a bit abstract, but in practice, the idea is of course that f,g must be independent, in an intuitive, real-life sense. As a first result now, we have:

PROPOSITION 1.4. *Assuming that $f,g\in L^\infty(X)$ are independent, we have*

$$\mu_{f+g}=\mu_f*\mu_g$$

where $*$ is the convolution of real probability measures.

PROOF. We have the following computation, using the independence of f,g:

$$\begin{aligned}M_k(f+g)&=\mathbb E((f+g)^k)\\&=\sum_l\binom{k}{l}\mathbb E(f^lg^{k-l})\\&=\sum_l\binom{k}{l}M_l(f)M_{k-l}(g)\end{aligned}$$

On the other hand, by using the Fubini theorem, we have as well:

$$\begin{aligned}\int_{\mathbb R}x^kd(\mu_f*\mu_g)(x)&=\int_{\mathbb R\times\mathbb R}(x+y)^kd\mu_f(x)d\mu_g(y)\\&=\sum_l\binom{k}{l}\int_{\mathbb R}x^kd\mu_f(x)\int_{\mathbb R}y^ld\mu_g(y)\\&=\sum_l\binom{k}{l}M_l(f)M_{k-l}(g)\end{aligned}$$

Thus μ_{f+g} and $\mu_f*\mu_g$ have the same moments, and so they coincide, as claimed. □

Here is now our second result, which is something more advanced, providing us with some efficient tools for the study of the independence:

THEOREM 1.5. *Assuming that $f,g\in L^\infty(X)$ are independent, we have*

$$F_{f+g}=F_fF_g$$

where $F_f(x)=\mathbb E(e^{ixf})$ is the Fourier transform.

PROOF. We have the following computation, using Proposition 1.4 and Fubini:

$$\begin{aligned}
F_{f+g}(x)&=\int_{\mathbb R}e^{ixy}d\mu_{f+g}(y)\\
&=\int_{\mathbb R}e^{ixy}d(\mu_f*\mu_g)(y)\\
&=\int_{\mathbb R\times\mathbb R}e^{ix(y+z)}d\mu_f(y)d\mu_g(z)\\
&=\int_{\mathbb R}e^{ixy}d\mu_f(y)\int_{\mathbb R}e^{ixz}d\mu_g(z)\\
&=F_f(x)F_g(x)
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

1b. Central limits

Let us discuss now the normal distributions. We will need:

PROPOSITION 1.6. *We have polar coordinates in 2 dimensions,*

$$\begin{cases}x=r\cos t\\ y=r\sin t\end{cases}$$

the corresponding Jacobian being $J=r$.

PROOF. This is elementary, the Jacobian being:

$$\begin{aligned}
J&=\begin{vmatrix}\frac{\partial r\cos t}{\partial r}&\frac{\partial r\cos t}{\partial t}\\ \\ \frac{\partial r\sin t}{\partial r}&\frac{\partial r\sin t}{\partial t}\end{vmatrix}\\
&=\begin{vmatrix}\cos t&-r\sin t\\ \sin t&r\cos t\end{vmatrix}\\
&=r\cos^2t+r\sin^2t\\
&=r
\end{aligned}$$

Thus, we have indeed the formula in the statement. □

As an application, we can compute the Gauss integral:

Theorem 1.7. *We have the following formula,*

$$\int_{\mathbb{R}} e^{-x^2}dx = \sqrt{\pi}$$

called Gauss integral formula.

Proof. Let I be the integral in the statement. By using polar coordinates, we obtain:

$$\begin{aligned}
I^2 &= \int_{\mathbb{R}} e^{-x^2}dx \int_{\mathbb{R}} e^{-y^2}dy \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} e^{-x^2-y^2}dxdy \\
&= \int_0^{2\pi}\int_0^\infty e^{-r^2\cos^2 t - r^2\sin^2 t} rdrdt \\
&= \int_0^{2\pi}\int_0^\infty e^{-r^2} rdrdt \\
&= 2\pi\int_0^\infty \left(-\frac{e^{-r^2}}{2}\right)' dr \\
&= 2\pi\left[0-\left(-\frac{1}{2}\right)\right] \\
&= \pi
\end{aligned}$$

Thus, we are led to the formula in the statement. □

We can now introduce the normal distributions, as follows:

Definition 1.8. *The normal law of parameter* 1 *is the following measure:*

$$g_1 = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$$

More generally, the normal law of parameter $t > 0$ is the following measure:

$$g_t = \frac{1}{\sqrt{2\pi t}}e^{-x^2/2t}dx$$

These are also called Gaussian distributions, with "g" standing for Gauss.

As a first comment, these laws are traditionally denoted $\mathcal{N}(0,1)$ and $\mathcal{N}(0,t)$, but since we will be doing in this book all kinds of probability, namely classical and free, real and complex, discrete and continuous, and so on, we will have to deal with lots of interesting probability measures, and we will be using simplified notations for them.

Let us mention as well that the normal laws traditionally have 2 parameters, the mean and the variance. Here we do not need the mean, all our theory using centered laws.

As a second remark, the above laws have indeed mass 1, as they should. This follows indeed from the Gauss formula, which gives, with $x=\sqrt{2t}\,y$:

$$\begin{aligned}\int_{\mathbb{R}} e^{-x^2/2t}dx &= \int_{\mathbb{R}} e^{-y^2}\sqrt{2t}\,dy\\ &= \sqrt{2t}\int_{\mathbb{R}} e^{-y^2}dy\\ &= \sqrt{2t}\times\sqrt{\pi}\\ &= \sqrt{2\pi t}\end{aligned}$$

Generally speaking, the normal laws appear as bit everywhere, in real life. The reasons behind this phenomenon come from the Central Limit Theorem (CLT), that we will explain in a moment, after developing some theory. As a first result, we have:

PROPOSITION 1.9. *We have the variance formula*

$$V(g_t)=t$$

valid for any $t>0$.

PROOF. We have $M_1=0$, and the second moment M_2 can be computed as follows:

$$\begin{aligned}M_2 &= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} x^2e^{-x^2/2t}dx\\ &= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}}(tx)\left(-e^{-x^2/2t}\right)'dx\\ &= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} te^{-x^2/2t}dx\\ &= \sqrt{\frac{t}{2\pi}}\int_{\mathbb{R}} e^{-x^2/2t}dx\\ &= \sqrt{\frac{t}{2\pi}}\times\sqrt{2\pi t}\\ &= t\end{aligned}$$

We conclude from this that the variance is $V=t$. □

Here is now another result, which is very useful in practice:

THEOREM 1.10. *We have the following formula, valid for any $t>0$:*

$$F_{g_t}(x)=e^{-tx^2/2}$$

In particular, the normal laws satisfy $g_s*g_t=g_{s+t}$, for any $s,t>0$.

PROOF. The Fourier transform formula can be established as follows:

$$\begin{aligned}
F_{g_t}(x) &= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} e^{-y^2/2t+ixy}dy\\
&= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} e^{-(y/\sqrt{2t}-\sqrt{t/2}ix)^2-tx^2/2}dy\\
&= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} e^{-z^2-tx^2/2}\sqrt{2t}dz\\
&= \frac{1}{\sqrt{\pi}}e^{-tx^2/2}\int_{\mathbb{R}} e^{-z^2}dz\\
&= \frac{1}{\sqrt{\pi}}e^{-tx^2/2}\cdot\sqrt{\pi}\\
&= e^{-tx^2/2}
\end{aligned}$$

As for the last assertion, this follows from the fact that $\log F_{g_t}$ is linear in t. $\square$

We are now ready to state and prove the CLT, as follows:

THEOREM 1.11 (CLT). *Given real random variables $f_1, f_2, f_3, \ldots \in L^\infty(X)$ which are i.i.d., centered, and with variance $t > 0$, we have, with $n \to \infty$, in moments,*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n f_i \sim g_t$$

where g_t is the Gaussian law of parameter t, having density $\frac{1}{\sqrt{2\pi t}}e^{-y^2/2t}dy$.

PROOF. In terms of moments, we have the following formula:

$$\begin{aligned}
F_f(x) &= \mathbb{E}\left(\sum_{k=0}^\infty \frac{(ixf)^k}{k!}\right)\\
&= \sum_{k=0}^\infty \frac{(ix)^k\mathbb{E}(f^k)}{k!}\\
&= \sum_{k=0}^\infty \frac{i^kM_k(f)}{k!}\,x^k
\end{aligned}$$

Thus, the Fourier transform of the variable in the statement is:

$$\begin{aligned} F(x) &= \left[F_f\left(\frac{x}{\sqrt{n}}\right)\right]^n \\ &= \left[1-\frac{tx^2}{2n}+O(n^{-2})\right]^n \\ &\simeq \left[1-\frac{tx^2}{2n}\right]^n \\ &\simeq e^{-tx^2/2} \end{aligned}$$

But this latter function being the Fourier transform of g_t, we obtain the result. □

Let us discuss now some further properties of the normal law, which are of more specialized nature. We first have the following result:

Proposition 1.12. *The moments of the normal law are the numbers*

$$M_k(g_t) = t^{k/2} \times k!!$$

where the double factorials are by definition given by

$$k!! = 1 \cdot 3 \cdot 5 \dots (k-1)$$

with the convention $k!! = 0$ when k is odd.

Proof. We have the following computation:

$$\begin{aligned} M_k &= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} x^k e^{-x^2/2t}dx \\ &= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} (tx^{k-1})\left(-e^{-x^2/2t}\right)' dx \\ &= \frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} t(k-1)x^{k-2}e^{-x^2/2t}dx \\ &= t(k-1)\times\frac{1}{\sqrt{2\pi t}}\int_{\mathbb{R}} x^{k-2}e^{-x^2/2t}dx \\ &= t(k-1)M_{k-2} \end{aligned}$$

Now recall from Proposition 1.9 that $M_0 = 1$, $M_1 = 0$. Thus by recurrence, the even moments all vanish, and the odd moments are given by the formula in the statement. □

We can improve the above result, as follows:

Proposition 1.13. *The moments of the normal law are the numbers*

$$M_k(g_t) = t^{k/2}|P_2(k)|$$

where $P_2(k)$ is the set of pairings of $\{1,\dots,k\}$.

PROOF. Let us count the pairings of $\{1,\ldots,k\}$. In order to have such a pairing, we must pair 1 with one of $2,\ldots,k$, and then use a pairing of the remaining $k-2$ points. Thus, we have the following recurrence formula:

$$|P_2(k)|=(k-1)|P_2(k-2)|$$

As for the initial data, this is $P_1=0$, $P_2=1$. We therefore obtain, by recurrence $|P_2(k)|=k!!$, and we are led to the conclusion in the statement. $\square$

We are done done yet, and here is one more improvement:

THEOREM 1.14. *The moments of the normal law are the numbers*

$$M_k(g_t)=\sum_{\pi\in P_2(k)}t^{|\pi|}$$

where $P_2(k)$ is the set of pairings of $\{1,\ldots,k\}$, and $|.|$ is the number of blocks.

PROOF. This follows indeed from Proposition 1.13 above, because the number of blocks of a pairing of $\{1,\ldots,k\}$ is trivially $k/2$, independently of the pairing. $\square$

1c. Spherical integrals

As already mentioned, due to the CLT, the normal laws appear a bit everywhere, in real life. In a purely mathematical context, the simplest way of recovering these laws is by looking at the coordinates over the real spheres $S^{N-1}_\mathbb{R}$, in the $N\to\infty$ limit.

At $N=2$ the sphere is the unit circle $\mathbb{T}$, and with $z=e^{it}$ the coordinates are $\cos t,\sin t$. Let us first integrate powers of these coordinates. We have here:

PROPOSITION 1.15. *We have the following formulae,*

$$\int_0^{\pi/2}\cos^pt\,dt=\int_0^{\pi/2}\sin^pt\,dt=\left(\frac{\pi}{2}\right)^{\varepsilon(p)}\frac{p!!}{(p+1)!!}$$

where $\varepsilon(p)=1$ if p is even, and $\varepsilon(p)=0$ if p is odd, and where

$$m!!=(m-1)(m-3)(m-5)\ldots$$

with the product ending at 2 if m is odd, and ending at 1 if m is even.

PROOF. Let us first compute the integral on the left in the statement:

$$I_p=\int_0^{\pi/2}\cos^pt\,dt$$

We do this by partial integration. We have the following formula:

$$\begin{aligned}(\cos^pt\sin t)'&=p\cos^{p-1}t(-\sin t)\sin t+\cos^pt\cos t\\&=p\cos^{p+1}t-p\cos^{p-1}t+\cos^{p+1}t\\&=(p+1)\cos^{p+1}t-p\cos^{p-1}t\end{aligned}$$

By integrating between 0 and $\pi/2$, we obtain the following formula:

$$(p+1)I_{p+1} = pI_{p-1}$$

Thus we can compute I_p by recurrence, and we obtain:

$$\begin{aligned} I_p &= \frac{p-1}{p}\, I_{p-2} \\ &= \frac{p-1}{p}\cdot\frac{p-3}{p-2}\, I_{p-4} \\ &= \frac{p-1}{p}\cdot\frac{p-3}{p-2}\cdot\frac{p-5}{p-4}\, I_{p-6} \\ &\ \vdots \\ &= \frac{p!!}{(p+1)!!}\, I_{1-\varepsilon(p)} \end{aligned}$$

On the other hand, at $p=0$ we have the following formula:

$$I_0 = \int_0^{\pi/2} 1\, dt = \frac{\pi}{2}$$

Also, at $p=1$ we have the following formula:

$$I_1 = \int_0^{\pi/2} \cos t\, dt = 1$$

Thus, we obtain the result, by recurrence. As for the second formula, regarding the integrals of powers of $\sin t$, this follows from the first formula, with the change of variables $t = \frac{\pi}{2} - s$. Thus, we have proved both formulae in the statement. □

More generally now, we have the following result, which fully computes the integrals of arbitrary polynomials of the coordinates over the unit circle:

THEOREM 1.16. *We have the following formula,*

$$\int_0^{\pi/2} \cos^p t \sin^q t\, dt = \left(\frac{\pi}{2}\right)^{\varepsilon(p)\varepsilon(q)} \frac{p!!q!!}{(p+q+1)!!}$$

where $\varepsilon(p)=1$ if p is even, and $\varepsilon(p)=0$ if p is odd, and where

$$m!! = (m-1)(m-3)(m-5)\ldots$$

with the product ending at 2 if m is odd, and ending at 1 if m is even.

PROOF. This is standard, indeed, by doing a partial integration, and then proving the result by a double recurrence, on both p and q. Let us set indeed:

$$I_{pq} = \int_0^{\pi/2} \cos^p t \sin^q t\, dt$$

In order to do the partial integration, observe that we have:

$$\begin{aligned}(\cos^p t\sin^q t)' &= p\cos^{p-1}t(-\sin t)\sin^q t+\cos^p t\cdot q\sin^{q-1}t\cos t\\ &= -p\cos^{p-1}t\sin^{q+1}t+q\cos^{p+1}t\sin^{q-1}t\end{aligned}$$

By integrating between 0 and $\pi/2$, we obtain, for $p,q>0$:

$$pI_{p-1,q+1}=qI_{p+1,q-1}$$

Thus, we can compute I_{pq} by recurrence. When q is even we have:

$$\begin{aligned}I_{pq} &= \frac{q-1}{p+1}\,I_{p+2,q-2}\\ &= \frac{q-1}{p+1}\cdot\frac{q-3}{p+3}\,I_{p+4,q-4}\\ &= \frac{q-1}{p+1}\cdot\frac{q-3}{p+3}\cdot\frac{q-5}{p+5}\,I_{p+6,q-6}\\ &= \vdots\\ &= \frac{p!!q!!}{(p+q)!!}\,I_{p+q}\end{aligned}$$

But the last term comes from Proposition 1.15, and we obtain the result:

$$\begin{aligned}I_{pq} &= \frac{p!!q!!}{(p+q)!!}\,I_{p+q}\\ &= \frac{p!!q!!}{(p+q)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(p+q)}\frac{(p+q)!!}{(p+q+1)!!}\\ &= \left(\frac{\pi}{2}\right)^{\varepsilon(p)\varepsilon(q)}\frac{p!!q!!}{(p+q+1)!!}\end{aligned}$$

Observe that this gives the result for p even as well, because $I_{pq}=I_{qp}$, by using the change of variables $t=\frac{\pi}{2}-s$. In the remaining case now, where both p,q are odd, we can use once again the formula $pI_{p-1,q+1}=qI_{p+1,q-1}$, and the recurrence goes as follows:

$$\begin{aligned}I_{pq} &= \frac{q-1}{p+1}\,I_{p+2,q-2}\\ &= \frac{q-1}{p+1}\cdot\frac{q-3}{p+3}\,I_{p+4,q-4}\\ &= \frac{q-1}{p+1}\cdot\frac{q-3}{p+3}\cdot\frac{q-5}{p+5}\,I_{p+6,q-6}\\ &= \vdots\\ &= \frac{p!!q!!}{(p+q-1)!!}\,I_{p+q-1,1}\end{aligned}$$

In order to compute the last term, observe that we have:

$$\begin{aligned} I_{p1} &= \int_0^{\pi/2} \cos^p t \sin t\, dt \\ &= -\frac{1}{p+1}\int_0^{\pi/2} (\cos^{p+1} t)'\, dt \\ &= \frac{1}{p+1} \end{aligned}$$

Thus, we can finish our computation in the case p,q odd, as follows:

$$\begin{aligned} I_{pq} &= \frac{p!!q!!}{(p+q-1)!!}\, I_{p+q-1,1} \\ &= \frac{p!!q!!}{(p+q-1)!!}\cdot\frac{1}{p+q} \\ &= \frac{p!!q!!}{(p+q+1)!!} \end{aligned}$$

Thus, we obtain the formula in the statement, the exponent of $\pi/2$ appearing there being $\varepsilon(p)\varepsilon(q) = 0\cdot 0 = 0$ in the present case, and this finishes the proof. □

Let us discuss now the integration over the higher spheres. We will need:

THEOREM 1.17. *We have spherical coordinates in N dimensions,*

$$\begin{cases} x_1 &= r\cos t_1 \\ x_2 &= r\sin t_1\cos t_2 \\ \vdots \\ x_{N-1} &= r\sin t_1\sin t_2\ldots\sin t_{N-2}\cos t_{N-1} \\ x_N &= r\sin t_1\sin t_2\ldots\sin t_{N-2}\sin t_{N-1} \end{cases}$$

the corresponding Jacobian being given by the following formula:

$$J(r,t) = r^{N-1}\sin^{N-2} t_1 \sin^{N-3} t_2\,\ldots\,\sin^2 t_{N-3}\sin t_{N-2}$$

PROOF. The fact that we have indeed spherical coordinates is clear. Regarding the Jacobian, the proof is similar to the one from 2 dimensions, by developing the determinant over the last column, and then by proceeding by recurrence. Indeed, we have:

$$\begin{aligned} J_N &= r\sin t_1\ldots\sin t_{N-2}\sin t_{N-1}\times\sin t_{N-1}J_{N-1} \\ &+ r\sin t_1\ldots\sin t_{N-2}\cos t_{N-1}\times\cos t_{N-1}J_{N-1} \\ &= r\sin t_1\ldots\sin t_{N-2}(\sin^2 t_{N-1}+\cos^2 t_{N-1})J_{N-1} \\ &= r\sin t_1\ldots\sin t_{N-2}J_{N-1} \end{aligned}$$

Thus, we obtain the formula in the statement, by recurrence. □

As a first application, we can compute the volume of the sphere:

THEOREM 1.18. *The volume of the unit sphere in $\mathbb{R}^N$ is given by*

$$\frac{V}{2^N}=\left(\frac{\pi}{2}\right)^{[N/2]}\frac{1}{(N+1)!!}$$

with the convention

$$N!!=(N-1)(N-3)(N-5)\ldots$$

with the product ending at 2 if N is odd, and ending at 1 if N is even.

PROOF. If we denote by Q the positive part of the sphere, obtained by cutting the sphere in 2^N parts, we have:

$$\begin{aligned}
\frac{V}{2^N} &= \int_Q 1\\
&= \int_0^1\int_0^{\pi/2}\ldots\int_0^{\pi/2} r^{N-1}\sin^{N-2}t_1\ldots\sin t_{N-2}\,drdt_1\ldots dt_{N-1}\\
&= \int_0^1 r^{N-1}\,dr\int_0^{\pi/2}\sin^{N-2}t_1\,dt_1\ldots\int_0^{\pi/2}\sin t_{N-2}dt_{N-2}\int_0^{\pi/2}1dt_{N-1}\\
&= \frac{1}{N}\times\left(\frac{\pi}{2}\right)^{[N/2]}\times\frac{(N-2)!!}{(N-1)!!}\cdot\frac{(N-3)!!}{(N-2)!!}\ldots\frac{2!!}{3!!}\cdot\frac{1!!}{2!!}\cdot 1\\
&= \frac{1}{N}\times\left(\frac{\pi}{2}\right)^{[N/2]}\times\frac{1}{(N-1)!!}\\
&= \left(\frac{\pi}{2}\right)^{[N/2]}\frac{1}{(N+1)!!}
\end{aligned}$$

Thus, we obtain the formula in the statement. □

Let us discuss now the computation of the arbitrary polynomial integrals, over the spheres of arbitrary dimension. The result here is as follows:

THEOREM 1.19. *The spherical integral of $x_{i_1}\ldots x_{i_k}$ vanishes, unless each index $a\in\{1,\ldots,N\}$ appears an even number of times in the sequence $i_1,\ldots,i_k$. We have*

$$\int_{S^{N-1}}x_{i_1}\ldots x_{i_k}\,dx=\frac{(N-1)!!l_1!!\ldots l_N!!}{(N+\Sigma l_i-1)!!}$$

with l_a being this number of occurrences.

PROOF. The result holds indeed at $N=2$, due to the formula in Theorem 1.16. In general, we can restrict attention to the case $l_a\in 2\mathbb{N}$, since the other integrals vanish. The integral in the statement can be written in spherical coordinates, as follows:

$$I=\frac{2^N}{V}\int_0^{\pi/2}\ldots\int_0^{\pi/2}x_1^{l_1}\ldots x_N^{l_N}\,J\,dt_1\ldots dt_{N-1}$$

In this formula V is the volume of the sphere, J is the Jacobian, and the 2^N factor comes from the restriction to the $1/2^N$ part of the sphere where all the coordinates are positive. The normalization constant in front of the integral is:

$$\frac{2^N}{V}=\frac{2^N}{N\pi^{N/2}}\cdot\Gamma\left(\frac{N}{2}+1\right)=\left(\frac{2}{\pi}\right)^{[N/2]}(N-1)!!$$

As for the unnormalized integral, this is given by:

$$\begin{aligned}I'=\int_0^{\pi/2}\ldots\int_0^{\pi/2}\quad&(\cos t_1)^{l_1}(\sin t_1\cos t_2)^{l_2}\\&\vdots\\&(\sin t_1\sin t_2\ldots\sin t_{N-2}\cos t_{N-1})^{l_{N-1}}\\&(\sin t_1\sin t_2\ldots\sin t_{N-2}\sin t_{N-1})^{l_N}\\&\sin^{N-2}t_1\sin^{N-3}t_2\ldots\sin^2 t_{N-3}\sin t_{N-2}\\&dt_1\ldots dt_{N-1}\end{aligned}$$

By rearranging the terms, we obtain:

$$\begin{aligned}I' &= \int_0^{\pi/2}\cos^{l_1}t_1\sin^{l_2+\ldots+l_N+N-2}t_1\,dt_1\\&\quad\int_0^{\pi/2}\cos^{l_2}t_2\sin^{l_3+\ldots+l_N+N-3}t_2\,dt_2\\&\quad\vdots\\&\quad\int_0^{\pi/2}\cos^{l_{N-2}}t_{N-2}\sin^{l_{N-1}+l_N+1}t_{N-2}\,dt_{N-2}\\&\quad\int_0^{\pi/2}\cos^{l_{N-1}}t_{N-1}\sin^{l_N}t_{N-1}\,dt_{N-1}\end{aligned}$$

Now by using the above-mentioned formula at $N=2$, this gives:

$$\begin{aligned}I' &= \frac{l_1!!(l_2+\ldots+l_N+N-2)!!}{(l_1+\ldots+l_N+N-1)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(N-2)}\\&\quad\frac{l_2!!(l_3+\ldots+l_N+N-3)!!}{(l_2+\ldots+l_N+N-2)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(N-3)}\\&\quad\vdots\\&\quad\frac{l_{N-2}!!(l_{N-1}+l_N+1)!!}{(l_{N-2}+l_{N-1}+l_N+2)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(1)}\\&\quad\frac{l_{N-1}!!l_N!!}{(l_{N-1}+l_N+1)!!}\left(\frac{\pi}{2}\right)^{\varepsilon(0)}\end{aligned}$$

Now observe that the various double factorials multiply up to quantity in the statement, modulo a $(N-1)!!$ factor, and that the $\frac{\pi}{2}$ factors multiply up to $F = \left(\frac{\pi}{2}\right)^{[N/2]}$. Thus by multiplying with the normalization constant, we obtain the result. □

1d. Hyperspherical laws

We can now recover the normal law, as follows:

THEOREM 1.20. *The moments of the hyperspherical variables are*

$$\int_{S^{N-1}_\mathbb{R}} x_i^k dx = \frac{(N-1)!!k!!}{(N+k-1)!!}$$

and the rescaled variables $y_i = \frac{x_i}{\sqrt{N}}$ become normal with $N \to \infty$.

PROOF. The formula in the statement follows from Theorem 1.19, at $k = 1$. With $N \to \infty$ we have the following estimate:

$$\begin{aligned}\int_{S^{N-1}_\mathbb{R}} x_i^k dx &= \frac{(N-1)}{(N+k-1)!!} \times k!!\\ &\simeq N^{k/2} \times k!!\\ &= N^{k/2} M_k(g_1)\end{aligned}$$

Thus, we are led to the conclusions in the statement. □

It is possible to prove as well that the rescaled variables $x_i/\sqrt{N}$ become independent with $N \to \infty$. We will be back to this. We can talk as well about rotations, as follows:

THEOREM 1.21. *We have the integration formula*

$$\int_{O_N} U_{ij}^k dU = \frac{(N-1)!!k!!}{(N+k-1)!!}$$

and the rescaled variables $V_{ij} = \frac{U_{ij}}{\sqrt{N}}$ become normal with $N \to \infty$.

PROOF. We use the well-known fact that we have an embedding as follows, for any i, which makes correspond the respective integration functionals:

$$C(S^{N-1}_\mathbb{R}) \subset C(O_N)$$

$$x_i \to U_{1i}$$

With this identification made, the result follows from Theorem 1.20. □

1e. Exercises

CHAPTER 2

The Poisson law

2a. Poisson limits

We discuss here the "discrete" counterpart of the results from chapter 1. The mathematics here will involve the Poisson laws p_t, and their versions. Let us start with:

DEFINITION 2.1. *The Poisson law of parameter* 1 *is the following measure,*

$$p_1 = \frac{1}{e}\sum_k \frac{\delta_k}{k!}$$

and the Poisson law of parameter $t > 0$ *is the following measure,*

$$p_t = e^{-t}\sum_k \frac{t^k}{k!}\,\delta_k$$

with the letter "p" standing for Poisson.

Here the notations that we use are in tune with the previous notations g_1, g_t for the normal laws, with "g" standing there for Gauss. As already explained in the previous chapter, the reasons for these notations come from the fact that we will be doing all kinds of probability in this book, and we need simple notations for everything.

As a first observation, the Poisson laws as constructed above have indeed mass 1, as they should, due to the following formula:

$$e^t = \sum_k \frac{t^k}{k!}$$

We will see in the moment why these measures appear a bit everywhere, in discrete contexts, the reasons behind this coming from the Poisson Limit Theorem (PLT).

Let us first develop some general theory. We first have the following result:

THEOREM 2.2. *We have the following formula, for any* $s, t > 0$*,*

$$p_s * p_t = p_{s+t}$$

so the Poisson laws form a convolution semigroup.

PROOF. By using $\delta_k * \delta_l = \delta_{k+l}$ and the binomial formula, we obtain:

$$\begin{aligned}
p_s * p_t &= e^{-s}\sum_k \frac{s^k}{k!}\,\delta_k * e^{-t}\sum_l \frac{t^l}{l!}\,\delta_l\\
&= e^{-s-t}\sum_{kl}\frac{s^kt^l}{k!l!}\delta_{k+l}\\
&= e^{-s-t}\sum_n \delta_n \sum_{k+l=n}\frac{s^kt^l}{k!l!}\\
&= e^{-s-t}\sum_n \frac{\delta_n}{n!}\sum_{k+l=n}\frac{n!}{k!l!}s^kt^l\\
&= e^{-s-t}\sum_n \frac{(s+t)^n}{n!}\,\delta_n\\
&= p_{s+t}
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

Observe the analogy with the formulae $g_s * g_t = g_{s+t}$ and $G_s * G_t = G_{s+t}$ for the Gaussian laws, established in chapter 1. We have as well the following result:

THEOREM 2.3. *The Poisson laws appear as exponentials*

$$p_t = \sum_k \frac{t^k(\delta_1 - \delta_0)^{*k}}{k!}$$

with respect to the convolution of measures $*$.

PROOF. By using the binomial formula, the measure at right is:

$$\begin{aligned}
\mu &= \sum_k \frac{t^k}{k!}\sum_{p+q=k}(-1)^q\frac{k!}{p!q!}\delta_p\\
&= \sum_k t^k \sum_{p+q=k}(-1)^q\frac{\delta_p}{p!q!}\\
&= \sum_p \frac{t^p\delta_p}{p!}\sum_q \frac{(-1)^q}{q!}\\
&= \frac{1}{e}\sum_p \frac{t^p\delta_p}{p!}\\
&= p_t
\end{aligned}$$

Thus, we are led to the conclusion in the statement. □

The Fourier transform computation is as follows:

THEOREM 2.4. *The Fourier transform of p_t is given by*

$$F_{p_t}(x)=\exp\left((e^{ix}-1)t\right)$$

for any $t>0$.

PROOF. We have the following computation:

$$\begin{aligned}
F_{p_t}(x) &= e^{-t}\sum_k\frac{t^k}{k!}F_{\delta_k}(x)\\
&= e^{-t}\sum_k\frac{t^k}{k!}\,e^{ikx}\\
&= e^{-t}\sum_k\frac{(e^{ix}t)^k}{k!}\\
&= \exp(-t)\exp(e^{ix}t)\\
&= \exp\left((e^{ix}-1)t\right)
\end{aligned}$$

Thus, we obtain the formula in the statement. □

We can now establish the Poisson Limit Theorem (PLT), as follows:

THEOREM 2.5. *We have the following convergence, in moments,*

$$\left(\left(1-\frac{t}{n}\right)\delta_0+\frac{t}{n}\delta_1\right)^{*n}\to p_t$$

for any $t>0$.

PROOF. Let us denote by μ_n the measure under the convolution sign:

$$\mu_n=\left(1-\frac{t}{n}\right)\delta_0+\frac{t}{n}\delta_1$$

We have the following computation:

$$\begin{aligned}
F_{\delta_r}(x)=e^{irx} &\implies F_{\mu_n}(x)=\left(1-\frac{t}{n}\right)+\frac{t}{n}e^{ix}\\
&\implies F_{\mu_n^{*n}}(x)=\left(\left(1-\frac{t}{n}\right)+\frac{t}{n}e^{ix}\right)^n\\
&\implies F_{\mu_n^{*n}}(x)=\left(1+\frac{(e^{ix}-1)t}{n}\right)^n\\
&\implies F(x)=\exp\left((e^{ix}-1)t\right)
\end{aligned}$$

Thus, we obtain the Fourier transform of p_t, as desired. □

2b. Bell numbers

At the level of moments now, things are quite subtle, and we first have the folowing result, dealing with the simplest case, namely $t = 1$:

THEOREM 2.6. *The moments of p_1 are the Bell numbers,*

$$M_k(p_1) = |P(k)|$$

where $P(k)$ is the set of partitions of $\{1, \ldots, k\}$.

PROOF. The moments of p_1 are given by the following formula:

$$M_k = \frac{1}{e}\sum_s \frac{s^k}{s!}$$

We have the following recurrence formula for these moments:

$$\begin{aligned}
M_{k+1} &= \frac{1}{e}\sum_s \frac{(s+1)^{k+1}}{(s+1)!}\\
&= \frac{1}{e}\sum_s \frac{(s+1)^k}{s!}\\
&= \frac{1}{e}\sum_s \frac{s^k}{s!}\left(1+\frac{1}{s}\right)^k\\
&= \frac{1}{e}\sum_s \frac{s^k}{s!}\sum_r \binom{k}{r} s^{-r}\\
&= \sum_r \binom{k}{r}\cdot\frac{1}{e}\sum_s \frac{s^{k-r}}{s!}\\
&= \sum_r \binom{k}{r} M_{k-r}
\end{aligned}$$

Let us try now to find a recurrence for the Bell numbers:

$$B_k = |P(k)|$$

A partition of $\{1, \ldots, k+1\}$ appears by choosing r neighbors for 1, among the k numbers available, and then partitioning the $k - r$ elements left. Thus, we have:

$$B_{k+1} = \sum_r \binom{k}{r} B_{k-r}$$

Thus, the numbers M_k satisfy the same recurrence as the numbers B_k. Regarding now the initial values, for the moments of p_1, these are $M_0 = 1$ and:

$$M_1 = \frac{1}{e}\sum_s \frac{s}{s!} = \frac{1}{e}\sum_s \frac{1}{(s-1)!} = \frac{1}{e} \times e = 1$$

Now by using the above recurrence we obtain from this:

$$M_2 = \sum_r \binom{1}{r} M_{k-r} = 1 + 1 = 2$$

Now observe that for the Bell numbers, the initial values are $B_1 = 1$, $B_2 = 2$. Thus the initial values of M_k, B_k coincide, and so these numbers are equal, as stated. □

More generally now, we have the following result:

THEOREM 2.7. *The moments of p_t with $t > 0$ are given by*

$$M_k(p_t) = \sum_{\pi \in P(k)} t^{|\pi|}$$

where $|.|$ is the number of blocks.

PROOF. Observe first that the formula in the statement generalizes the one in Theorem 2.6 above, because at $t = 1$ we obtain, as we should:

$$M_k(p_1) = \sum_{\pi \in P(k)} 1^{|\pi|} = |P(k)| = B_k$$

In general now, the moments of p_t with $t > 0$ are given by:

$$M_k = e^{-t}\sum_s \frac{t^s s^k}{s!}$$

We have the following recurrence formula for these moments:

$$\begin{aligned}
M_{k+1} &= e^{-t}\sum_s \frac{t^{s+1}(s+1)^{k+1}}{(s+1)!}\\
&= e^{-t}\sum_s \frac{t^{s+1}(s+1)^k}{s!}\\
&= e^{-t}\sum_s \frac{t^{s+1}s^k}{s!}\left(1+\frac{1}{s}\right)^k\\
&= e^{-t}\sum_s \frac{t^{s+1}s^k}{s!}\sum_r \binom{k}{r}s^{-r}\\
&= \sum_r \binom{k}{r}\cdot e^{-t}\sum_s \frac{t^{s+1}s^{k-r}}{s!}\\
&= t\sum_r \binom{k}{r}M_{k-r}
\end{aligned}$$

Regarding now the initial values, we have $M_0 = 1$, and:

$$M_1 = e^{-t}\sum_s \frac{t^s s}{s!} = e^{-t}\sum_s \frac{t^s}{(s-1)!} = \frac{1}{e}\times te^t = t$$

Now by using the above recurrence we obtain from this:

$$M_2 = t\sum_r \binom{1}{r}M_{k-r} = t(1+t) = t+t^2$$

Thus, the initial values for our moments are as follows:

$$M_1 = t \quad , \quad M_2 = t+t^2$$

On the other hand, consider the numbers in the statement, namely:

$$S_k = \sum_{\pi\in P(k)} t^{|\pi|}$$

Since a partition of $\{1,\ldots,k+1\}$ appears by choosing r neighbors for 1, among the k numbers available, and then partitioning the $k-r$ elements left, we have:

$$S_{k+1} = t\sum_r \binom{k}{r}S_{k-r}$$

As for the initial values of these numbers, these are:

$$S_1 = t \quad , \quad S_2 = t+t^2$$

Thus the initial values coincide, and so these numbers are the moments, as stated. □

Observe the analogy with the moment formulae for g_t and G_t, from section 1 above. We will be back later with some more conceptual explanations for these results.

2c. Derangements

In relation now with groups, and with applications to mathematics in general, let us first recall the following well-known, fundamental, and very beautiful result:

THEOREM 2.8. *The probability for a random permutation $\sigma \in S_N$ to have no fixed points is*

$$P \simeq \frac{1}{e}$$

in the $N \to \infty$ limit.

PROOF. This is something very classical, which is best viewed by using the inclusion-exclusion principle. Consider indeed the following sets:

$$S_N^i = \left\{\sigma \in S_N \Big| \sigma(i) = i\right\}$$

The set of permutations having no fixed points is then:

$$X_N = \left(\bigcup_i S_N^i\right)^c$$

In order to compute now the cardinality $|X_N|$, consider as well the following sets, depending on indices $i_1 < \ldots < i_k$, obtained by taking intersections:

$$S_N^{i_1 \ldots i_k} = S_N^{i_1} \bigcap \ldots \bigcap S_N^{i_k}$$

Observe that we have the following formula:

$$S_N^{i_1 \ldots i_k} = \left\{\sigma \in S_N \Big| \sigma(i_1) = i_1, \ldots, \sigma(i_k) = i_k\right\}$$

The inclusion-exclusion principle tells us that we have:

$$\begin{aligned}
&\left|\left(\bigcup_i S_N^i\right)^c\right| \\
= \ & |S_N| - \sum_i |S_N^i| + \sum_{i<j} |S_N^i \cap S_N^j| - \ldots + (-1)^N \sum_{i_1 < \ldots < i_N} |S_N^{i_1} \cup \ldots \cup S_N^{i_N}| \\
= \ & |S_N| - \sum_i |S_N^i| + \sum_{i<j} |S_N^{ij}| - \ldots + (-1)^N \sum_{i_1 < \ldots < i_N} |S_N^{i_1 \ldots i_N}|
\end{aligned}$$

Thus, the probability that we are interested in is given by:

$$P = \frac{1}{N!}\left(|S_N| - \sum_i |S_N^i| + \sum_{i<j} |S_N^{ij}| - \ldots + (-1)^N \sum_{i_1 < \ldots < i_N} |S_N^{i_1 \ldots i_N}|\right)$$

Now observe that for any $i_1 < \ldots < i_k$ we have:

$$|S_N^{i_1\ldots i_k}| = (N-k)!$$

We obtain from this:

$$\begin{aligned}
P &= \frac{1}{N!}\sum_{k=0}^N(-1)^k\sum_{i_1<\ldots<i_k}|S_N^{i_1\ldots i_k}|\\
&= \frac{1}{N!}\sum_{k=0}^N(-1)^k\sum_{i_1<\ldots<i_k}(N-k)!\\
&= \frac{1}{N!}\sum_{k=0}^N(-1)^k\binom{N}{k}(N-k)!\\
&= \sum_{k=0}^N\frac{(-1)^k}{k!}\\
&= 1-\frac{1}{1!}+\frac{1}{2!}-\ldots+(-1)^{N-1}\frac{1}{(N-1)!}+(-1)^N\frac{1}{N!}
\end{aligned}$$

Since on the right we have the expansion of $\frac{1}{e}$, this gives the result. □

In order to refine the above result, we will need some notions from group theory:

Definition 2.9. *Given a closed subgroup $G \subset U_N$, the function*

$$\chi(g) = \sum_{i=1}^N g_{ii}$$

is called main character of G. More generally, the function

$$\chi_t(g) = \sum_{i=1}^{[tN]} g_{ii}$$

is called main truncated character of G, of parameter $t \in (0,1]$.

Getting back now to the symmetric group S_N, it is useful in the present context to regard it as being the permutation group of the N coordinate axes of $\mathbb{R}^N$.

Since these permutations are isometries, we have an embedding as follows:

$$S_N \subset O_N$$

In other words, we identify the permutations $\sigma \in S_N$ with the corresponding permutation matrices $M_\sigma \in O_N$.

We can now refine Theorem 2.8, as follows:

THEOREM 2.10. *Consider the symmetric group S_N, regarded as a compact group of matrices, $S_N\subset O_N$, via the standard permutation matrices.*

(1) *The main character $\chi\in C(S_N)$ counts the number of fixed points:*

$$\chi(\sigma)=\#\left\{i\in\{1,\ldots,N\}\Big|\sigma(i)=i\right\}$$

(2) *The law of the main character $\chi\in C(S_N)$ becomes, with $N\to\infty$, a Poisson law of parameter 1, with respect to the counting measure:*

$$\chi\simeq p_1$$

(3) *The truncated character $\chi_t\in C(S_N)$ is given by the following formula:*

$$\chi_t(\sigma)=\#\left\{i\in\{1,\ldots,[tN]\}\Big|\sigma(i)=i\right\}$$

(4) *The law of the truncated character $\chi_t\in C(S_N)$ becomes, with $N\to\infty$, a Poisson law of parameter t, with respect to the counting measure:*

$$\chi_t\simeq p_t$$

PROOF. The idea is that the assertions (1,3) are clear, and that (2,4) basically follow by extending the proof of Theorem 2.8. To be more precise:

(1) We have indeed the following computation:

$$\chi(\sigma)=\sum_i\sigma_{ii}=\sum_i\delta_{\sigma(i)i}=\#\left\{i\in\{1,\ldots,N\}\Big|\sigma(i)=i\right\}$$

(2) With the present formalism, the formula in Theorem 2.8 reads:

$$\lim_{N\to\infty}\mathbb{P}(\chi=0)=\frac{1}{e}$$

A similar application of the inclusion-exclusion principle gives:

$$\lim_{N\to\infty}\mathbb{P}(\chi=k)=\frac{1}{k!e}$$

Thus, we obtain in the limit a Poisson law, as stated.

(3) We have the following computation, generalizing the one in (1) above:

$$\chi_t(\sigma)=\sum_{i=1}^{[tN]}\sigma_{ii}=\sum_{i=1}^{[tN]}\delta_{\sigma(i)i}=\#\left\{i\in\{1,\ldots,[tN]\}\Big|\sigma(i)=i\right\}$$

(4) With the formula (3) in hand, our claim is that the inclusion-exclusion gives:

$$\lim_{N\to\infty}\mathbb{P}(\chi_t=0)=\frac{1}{e^t}$$

Consider indeed the following sets:

$$S_N^i=\left\{\sigma\in S_N\Big|\sigma(i)=i\right\}$$

The set of permutations having no fixed points is then:

$$X_N=\left(\bigcup_iS_N^i\right)^c$$

In order to compute now the cardinality $|X_N|$, consider as well the following sets, depending on indices $i_1<\ldots<i_k$, obtained by taking intersections:

$$S_N^{i_1\ldots i_k}=S_N^{i_1}\bigcap\ldots\bigcap S_N^{i_k}$$

Observe that we have the following formula:

$$S_N^{i_1\ldots i_k}=\left\{\sigma\in S_N\Big|\sigma(i_1)=i_1,\ldots,\sigma(i_k)=i_k\right\}$$

The inclusion-exclusion principle tells us that we have:

$$\begin{aligned}
&\left|\left(\bigcup_iS_N^i\right)^c\right|\\
=&\ |S_N|-\sum_i|S_N^i|+\sum_{i<j}|S_N^i\cap S_N^j|-\ldots+(-1)^N\sum_{i_1<\ldots<i_N}|S_N^{i_1}\cup\ldots\cup S_N^{i_N}|\\
=&\ |S_N|-\sum_i|S_N^i|+\sum_{i<j}|S_N^{ij}|-\ldots+(-1)^N\sum_{i_1<\ldots<i_N}|S_N^{i_1\ldots i_N}|
\end{aligned}$$

Thus, the probability that we are interested in is given by:

$$P=\frac{1}{N!}\left(|S_N|-\sum_i|S_N^i|+\sum_{i<j}|S_N^{ij}|-\ldots+(-1)^N\sum_{i_1<\ldots<i_N}|S_N^{i_1\ldots i_N}|\right)$$

Now observe that for any $i_1<\ldots<i_k$ we have:

$$|S_N^{i_1\ldots i_k}|=(N-k)!$$

We obtain from this:

$$\begin{aligned}
P &= \frac{1}{N!}\sum_{k=0}^N(-1)^k\sum_{i_1<\ldots<i_k}|S_N^{i_1\ldots i_k}| \\
&= \frac{1}{N!}\sum_{k=0}^N(-1)^k\sum_{i_1<\ldots<i_k}(N-k)! \\
&= \frac{1}{N!}\sum_{k=0}^N(-1)^k\binom{N}{k}(N-k)! \\
&= \sum_{k=0}^N\frac{(-1)^k}{k!} \\
&= 1-\frac{1}{1!}+\frac{1}{2!}-\ldots+(-1)^{N-1}\frac{1}{(N-1)!}+(-1)^N\frac{1}{N!}
\end{aligned}$$

Since on the right we have the expansion of $\frac{1}{e}$, this gives the result. More generally, the inclusion-exclusion principle gives:

$$\lim_{N\to\infty}\mathbb{P}(\chi_t=k)=\frac{t^k}{k!e^t}$$

Thus, we obtain in the limit a Poisson law of parameter t, as stated. □

As a conclusion to this, the Poisson laws p_t appear to be quite similar to the real and complex Gaussian laws g_t and G_t, in the sense that:

(1) All these laws appear via basic limiting theorems.
(2) They form semigroups with respect to the convolution.
(3) Their moments can be computed by counting certain partitions.
(4) There is a relation with pure mathematics as well, involving S_N, O_N, U_N.

All this remains of course to be further discussed. One problem for instance comes from the fact that the above computations for S_N, involving truncated characters, are not exactly of the same nature as those for O_N, U_N, involving individual coordinates.

2d. Bessel laws

For the moment, let us keep looking at finite transformation groups of $\mathbb{R}^N$, and do probability theory for the corresponding characters, and truncated characters.

An obvious choice here is the hyperoctahedral group H_N. The definition and basic properties of this remarkable group can be summarized as follows:

THEOREM 2.11. *Consider the hyperoctahedral group $H_N\subset O_N$, consisting of the symmetries of the hypercube in $\mathbb R^N$.*

1. *H_N appears as well as symmetry group of the N coordinate axes of $\mathbb R^N$.*
2. *H_N consists of the permutation-like matrices with entries in $\{-1,0,1\}$.*
3. *We have the cardinality formula $|H_N|=2^NN!$.*
4. *We have a crossed product decomposition $H_N=S_N\rtimes\mathbb Z_2^N$.*
5. *We have a wreath product decomposition $H_N=\mathbb Z_2\wr S_N$.*

PROOF. Consider indeed the standard cube in $\mathbb R^N$, centered at 0, and having as vertices the points having coordinates ±1.

(1) With the above picture in hand, it is clear that the symmetries of the cube coincide with the symmetries of the N coordinate axes of $\mathbb R^N$.

(2) We use here the interpretation in (1). In order to understand the symmetries there, observe first that we have among them the $N!$ permutations of the N coordinate axes.

But each of these permutations $\sigma\in S_N$ can be further "decorated" by a sign vector $\varepsilon\in\{\pm1\}^N$, consisting of the possible ±1 flips which can be applied to each coordinate axis, at the arrival. In matrix terms, this gives the conclusion in the statement.

(3) By using the above interpretation of H_N, we have the following formula:

$$|H_N|=|S_N|\cdot|\mathbb Z_2^N|=N!\cdot2^N$$

(4) We know from (3) that at the level of cardinalities, we have:

$$|H_N|=|S_N\times\mathbb Z_2^N|$$

We can deduce from this that we have a crossed product decomposition, as follows:

$$H_N=S_N\rtimes\mathbb Z_2^N$$

(5) The formula established in (4) means exactly that we have:

$$H_N=\mathbb Z_2\wr S_N$$

Thus, we are led to the conclusion in the statement. □

Getting back now to our character computations, we have:

THEOREM 2.12. *For the hyperoctahedral group $H_N\subset O_N$, the law of the variable*

$$\chi=g_{11}+\ldots+g_{ss}$$

with $s=[tN]$ is, in the $N\to\infty$ limit, the measure

$$b_t=e^{-t}\sum_{k=-\infty}^\infty\delta_k\sum_{p=0}^\infty\frac{(t/2)^{|k|+2p}}{(|k|+p)!p!}$$

where δ_k is the Dirac mass at $k\in\mathbb Z$.

PROOF. We regard H_N as being the symmetry group of the graph $I_N=\{I^1,\ldots,I^N\}$ formed by n segments. The diagonal coefficients are given by:

$$u_{ii}(g)=\begin{cases}0\ \text{ if } g \text{ moves } I^i\\ +1 \text{ if } g \text{ fixes } I^i\\ -1 \text{ if } g \text{ returns } I^i\end{cases}$$

We denote by $\uparrow g,\downarrow g$ the number of segments among $\{I^1,\ldots,I^s\}$ which are fixed, respectively returned by an element $g\in H_N$. With this notation, we have:

$$u_{11}+\ldots+u_{ss}=\uparrow g-\downarrow g$$

We denote by P_N probabilities computed over the group H_N. The density of the law of $u_{11}+\ldots+u_{ss}$ at a point $k\geq 0$ is given by the following formula:

$$\begin{aligned}D(k)&=\mathrm{P}_N(\uparrow g-\downarrow g=k)\\&=\sum_{p=0}^\infty \mathrm{P}_N(\uparrow g=k+p,\downarrow g=p)\end{aligned}$$

Assume first that we have $t=1$. We use the fact that the probability of $\sigma\in S_N$ to have no fixed points is asymptotically $P_0=\frac{1}{e}$. Thus the probability of $\sigma\in S_N$ to have m fixed points is asymptotically $P_m=\frac{1}{em!}$. In terms of probabilities over H_N, we get:

$$\begin{aligned}\lim_{N\to\infty}D(k)&=\lim_{N\to\infty}\sum_{p=0}^\infty(1/2)^{k+2p}\binom{k+2p}{k+p}\mathrm{P}_N(\uparrow g+\downarrow g=k+2p)\\&=\sum_{p=0}^\infty(1/2)^{k+2p}\binom{k+2p}{k+p}\frac{1}{e(k+2p)!}\\&=\frac{1}{e}\sum_{p=0}^\infty\frac{(1/2)^{k+2p}}{(k+p)!p!}\end{aligned}$$

The general case $0<t\leq 1$ follows by performing some modifications in the above computation. The asymptotic density is computed as follows:

$$\begin{aligned}\lim_{N\to\infty}D(k)&=\lim_{N\to\infty}\sum_{p=0}^\infty(1/2)^{k+2p}\binom{k+2p}{k+p}\mathrm{P}_n(\uparrow g+\downarrow g=k+2p)\\&=\sum_{p=0}^\infty(1/2)^{k+2p}\binom{k+2p}{k+p}\frac{t^{k+2p}}{e^t(k+2p)!}\\&=e^{-t}\sum_{p=0}^\infty\frac{(t/2)^{k+2p}}{(k+p)!p!}\end{aligned}$$

Together with $D(-k)=D(k)$, this gives the formula in the statement. □

The above result is quite interesting, because the densities there are the Bessel functions of the first kind. Due to this fact, the limiting measures are called Bessel laws:

DEFINITION 2.13. *The Bessel law of parameter $t > 0$ is the measure*

$$b_t = e^{-t} \sum_{k=-\infty}^{\infty} \delta_k \, f_k(t/2)$$

with the density being the Bessel function of the first kind:

$$f_k(t) = \sum_{p=0}^{\infty} \frac{t^{|k|+2p}}{(|k|+p)!p!}$$

Let us study now these Bessel laws. We first have the following result:

THEOREM 2.14. *The Bessel laws b_t have the property*

$$b_s * b_t = b_{s+t}$$

so they form a truncated one-parameter semigroup with respect to convolution.

PROOF. The Fourier transform of the measure b_t is given by:

$$Fb_t(y) = e^{-t} \sum_{k=-\infty}^{\infty} e^{ky} \, f_k(t/2)$$

We compute now the derivative with respect to t:

$$Fb_t(y)' = -Fb_t(y) + \frac{e^{-t}}{2} \sum_{k=-\infty}^{\infty} e^{ky} \, f_k'(t/2)$$

On the other hand, the derivative of f_k with $k \geq 1$ is given by:

$$\begin{aligned}
f_k'(t) &= \sum_{p=0}^{\infty} \frac{(k+2p)t^{k+2p-1}}{(k+p)!p!} \\
&= \sum_{p=0}^{\infty} \frac{(k+p)t^{k+2p-1}}{(k+p)!p!} + \sum_{p=0}^{\infty} \frac{p\, t^{k+2p-1}}{(k+p)!p!} \\
&= \sum_{p=0}^{\infty} \frac{t^{k+2p-1}}{(k+p-1)!p!} + \sum_{p=1}^{\infty} \frac{t^{k+2p-1}}{(k+p)!(p-1)!} \\
&= \sum_{p=0}^{\infty} \frac{t^{(k-1)+2p}}{((k-1)+p)!p!} + \sum_{p=1}^{\infty} \frac{t^{(k+1)+2(p-1)}}{((k+1)+(p-1))!(p-1)!} \\
&= f_{k-1}(t) + f_{k+1}(t)
\end{aligned}$$

This computation works in fact for any k, so we get:

$$\begin{aligned}
Fb_t(y)' &= -Fb_t(y) + \frac{e^{-t}}{2}\sum_{k=-\infty}^{\infty} e^{ky}(f_{k-1}(t/2) + f_{k+1}(t/2)) \\
&= -Fb_t(y) + \frac{e^{-t}}{2}\sum_{k=-\infty}^{\infty} e^{(k+1)y} f_k(t/2) + e^{(k-1)y} f_k(t/2) \\
&= -Fb_t(y) + \frac{e^y + e^{-y}}{2} Fb_t(y) \\
&= \left(\frac{e^y + e^{-y}}{2} - 1\right) Fb_t(y)
\end{aligned}$$

Thus the log of the Fourier transform is linear in t, and we get the assertion. □

In order to further discuss this, and extend these results, we will need a number of probabilistic preliminaries. We have the following notion, extending Theorem 2.5:

DEFINITION 2.15. *Associated to any compactly supported positive measure ν on $\mathbb{R}$ is the probability measure*

$$p_\nu = \lim_{n\to\infty}\left(\left(1 - \frac{t}{n}\right)\delta_0 + \frac{1}{n}\nu\right)^{*n}$$

where $t = mass(\nu)$, called compound Poisson law.

In what follows we will be interested in the case where ν is discrete, as is for instance the case for $\nu = t\delta_1$ with $t > 0$, which produces the Poisson laws. The following result allows one to detect compound Poisson laws:

PROPOSITION 2.16. *For a discrete measure, written as $\nu = \sum_{i=1}^s t_i\delta_{z_i}$ with $t_i > 0$ and $z_i \in \mathbb{R}$, we have*

$$F_{p_\nu}(y) = \exp\left(\sum_{i=1}^s t_i(e^{iyz_i} - 1)\right)$$

where F denotes the Fourier transform.

PROOF. Let μ_n be the measure appearing in Definition 2.15, under the convolution signs, namely:

$$\mu_n = \left(1 - \frac{t}{n}\right)\delta_0 + \frac{1}{n}\nu$$

We have the following computation:

$$\begin{aligned} F_{\mu_n}(y) &= \left(1-\frac{t}{n}\right)+\frac{1}{n}\sum_{i=1}^s t_ie^{iyz_i}\\ \implies\quad F_{\mu_n^{*n}}(y) &= \left(\left(1-\frac{t}{n}\right)+\frac{1}{n}\sum_{i=1}^s t_ie^{iyz_i}\right)^n\\ \implies\quad F_{p_\nu}(y) &= \exp\left(\sum_{i=1}^s t_i(e^{iyz_i}-1)\right) \end{aligned}$$

Thus, we have obtained the formula in the statement. □

We have as well the following result, providing an alternative to Definition 2.15:

THEOREM 2.17. *For a discrete measure, written as $\nu=\sum_{i=1}^s y_i\delta_{z_i}$ with $y_i>0$ and $z_i\in\mathbb{R}$, we have*

$$p_\nu=\mathrm{law}\left(\sum_{i=1}^s z_i\alpha_i\right)$$

where the variables α_i are Poisson (t_i), independent.

PROOF. Let α be the sum of Poisson variables in the statement:

$$\alpha=\sum_{i=1}^s z_i\alpha_i$$

By using some well-known Fourier transform formulae, we have:

$$\begin{aligned} F_{\alpha_i}(y) &= \exp(t_i(e^{iy}-1))\\ \implies\quad F_{z_i\alpha_i}(y) &= \exp(t_i(e^{iyz_i}-1))\\ \implies\quad F_\alpha(y) &= \exp\left(\sum_{i=1}^s t_i(e^{iyz_i}-1)\right) \end{aligned}$$

Thus we have indeed the same formula as in Proposition 2.16. □

Summarizing, we have a full generalization of the PLT.

Getting back now to the Bessel laws, we have:

THEOREM 2.18. *The Bessel laws b_t are compound Poisson laws, given by*

$$b_t=p_{t\varepsilon}$$

where ε is the centered Bernoulli law, $\varepsilon=\frac{1}{2}(\delta_{-1}+\delta_1)$.

PROOF. This follows indeed by comparing the formula of the Fourier transform of b_t, from the proof of Theorem 2.14 above, with the formula in Proposition 2.16. □

Summarizing, we have a good understanding of H_N, and everything is finally quite similar to what happens for S_N. Our next task will be that of generalizing the results that we have for S_N, H_N. For this purpose, let us consider the following family of groups:

DEFINITION 2.19. *The complex reflection group $H_N^s \subset U_N$, depending on parameters*

$$N \in \mathbb{N}$$

$$s \in \mathbb{N} \cup \{\infty\}$$

are the groups of permutation-type matrices with s-th roots of unity as entries,

$$H_N^s = M_N(\mathbb{Z}_s \cup \{0\}) \cap U_N$$

with the convention $\mathbb{Z}_\infty = \mathbb{T}$, at $s = \infty$.

Observe that at $s = 1, 2$ we obtain in this way the symmetric group S_N and the hyperoctahedral group H_N:

$$H_N^1 = S_N \quad , \quad H_N^2 = H_N$$

Another important particular case is $s = \infty$, where we obtain a group which is actually not finite, denoted as follows:

$$K_N \subset U_N$$

In order to do now the character computations for H_N^s, in general, we need a number of further probabilistic preliminaries. Let us start with:

DEFINITION 2.20. *The Bessel law of level $s \in \mathbb{N} \cup \{\infty\}$ and parameter $t > 0$ is*

$$b_t = p_{t\varepsilon_s}$$

with ε_s being the uniform measure on the s-th roots of unity.

Observe that at $s = 1, 2$ we obtain p_t, b_t. Another important particular case is $s = \infty$, where we obtain a messure which is actually not discrete, denoted as follows:

$$B_t = b_t^\infty$$

Here we use the same convention as in the continuous case, namely capital letters stand for complexifications. As a basic now result on these laws, we have:

THEOREM 2.21. *The generalized Bessel laws b_t^s have the property*

$$b_t^s * b_{t'}^s = b_{t+t'}^s$$

so they form a truncated one-parameter semigroup with respect to convolution.

PROOF. This follows indeed from the Fourier transform formulae from Proposition 2.16, because the log of these Fourier transforms are linear in t. □

Regarding now the moments, the result here is as follows:

THEOREM 2.22. *The moments of the Bessel law b_t^s are the numbers*

$$M_k = |P^s(k)|$$

where $P^s(k)$ is the set of partitions of $\{1, \ldots, k\}$ satisfying

$$\#\circ = \# \bullet (s)$$

as a weighted sum, in each block.

PROOF. Observe first that the formula holds indeed at $s = 1$, where $b_t^1 = p_t$ is the Poisson law of parameter $t > 0$, and where $P^1 = P$ is the set of all partitions. At $s = 2$ we have $P^2 = P_{even}$, and the result is elementary as well. In general, this follows by doing some combinatorics. □

Summarizing, we have full generalizations of our various results regarding p_t, b_t. We can go back now to reflection groups, and we have:

THEOREM 2.23. *For the complex reflection group H_N^s we have, with $N \to \infty$,*

$$\chi_t \sim b_t^s$$

where $b_t^s = p_{t\varepsilon_s}$, with ε_s being the uniform measure on the s-th roots of unity.

PROOF. This follows indeed by doing some computations, by using the inclusion-exclusion principle, generalizing those at $s = 1$ for S_N, and at $s = 2$ for H_N. □

2e. Exercises

CHAPTER 3

Spectral measures

3a. Random matrices

We discuss here more advanced aspects of probability theory, which are of rather "noncommutative" nature, in relation with the random matrices.

The random matrices are simple and fundamental mathematical objects, virtually appearing in all areas of mathematics and physics. They are defined as follows:

Definition 3.1. *A random matrix is a square matrix of type*

$$Z \in M_N(L^\infty(X))$$

with X being a probability space, and $N \in \mathbb{N}$ being an integer.

As basic examples, we have the usual matrices $Z \in M_N(\mathbb{C})$, obtained by taking $X = \{.\}$. Also, we have the usual random variables $Z \in L^\infty(X)$, obtained by taking $N = 1$. In general, what we have is a kind of combination of these 2 situations.

Let us begin with a discussion concerning the usual matrices $Z \in M_N(\mathbb{C})$. These are not exactly random variables in the usual sense, but we can still talk about their moments and their distributions, by being a bit abstract. Let us begin with:

Definition 3.2. *The moments of a usual complex matrix $Z \in M_N(\mathbb{C})$ are the following numbers, indexed by the integers $k \in \mathbb{N}$,*

$$M_k = tr(Z^k)$$

with $tr = N^{-1} \cdot Tr$ being the normalized matrix trace.

As a basic example here, consider the case of a diagonal matrix:

$$Z = \begin{pmatrix} \lambda_1 && \\ & \ddots & \\ && \lambda_N \end{pmatrix}$$

The powers of Z, with respect to integer exponents $k \in \mathbb{N}$, are as follows:

$$Z^k = \begin{pmatrix} \lambda_1^k && \\ & \ddots & \\ && \lambda_N^k \end{pmatrix}$$

Thus the moments are given by the following formula:

$$M_k = \sum_i \lambda_i^k$$

Regarding now the distribution, things are a bit more tricky here. In view of some further generalizations, let us formulate things here as follows:

Definition 3.3. *The distribution of a usual complex matrix $Z \in M_N(\mathbb{C})$ is the following abstract functional $\mu_Z : \mathbb{C}[X] \to \mathbb{C}$, with $tr = N^{-1} \cdot Tr$:*

$$P \to tr(P(Z))$$

In the case where we have a probability measure $\mu_Z \in \mathcal{P}(\mathbb{C})$ such that

$$tr(P(Z)) = \int_{\mathbb{C}} P(x)\, d\mu_Z(x)$$

we identify this measure with the distribution, or law of Z.

Observe that knowing the distribution is the same as knowing the moments. Once again, for illustrating this notion, consider the case of a diagonal matrix:

$$Z = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}$$

We have then the following formula, valid for any polynomial $P \in \mathbb{C}[X]$:

$$P(Z) = \begin{pmatrix} P(\lambda_1) & & \\ & \ddots & \\ & & P(\lambda_N) \end{pmatrix}$$

Now by applying the normalized trace, we obtain from this:

$$\begin{aligned} tr(P(Z)) &= \frac{1}{N}(P(\lambda_1) + \ldots + P(\lambda_N)) \\ &= \frac{1}{N}\int_{\mathbb{C}} P(x) d(\delta_{\lambda_1} + \ldots + \delta_{\lambda_N})(x) \\ &= \int_{\mathbb{C}} P(x) d\left(\frac{1}{N}(\delta_{\lambda_1} + \ldots + \delta_{\lambda_N})\right)(x) \end{aligned}$$

Thus, according to Definition 3.3, the law of Z is the following measure:

$$\mu_Z = \frac{1}{N}(\delta_{\lambda_1} + \ldots + \delta_{\lambda_N})$$

Quite remarkably, the distribution always exists as a probability measure on $\mathbb{C}$:

THEOREM 3.4. *For any matrix $Z \in M_N(\mathbb{C})$ we have the formula*

$$tr(P(Z)) = \frac{1}{N}(P(\lambda_1) + \ldots + P(\lambda_N))$$

where $\lambda_1, \ldots, \lambda_N \in \mathbb{C}$ are the eigenvalues of Z. Thus the complex measure

$$\mu_Z = \frac{1}{N}(\delta_{\lambda_1} + \ldots + \delta_{\lambda_N})$$

is the distribution of Z, in the abstract sense of Definition 3.3.

PROOF. There are several proofs for this fact, and a particularly beautiful and instructive proof, relying on some density arguments which are good to know, is as follows:

(1) Consider first the simplest case, that of a diagonal matrix:

$$Z = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}$$

Here we know from the above discussion that the result holds indeed:

$$\mu_Z = \frac{1}{N}(\delta_{\lambda_1} + \ldots + \delta_{\lambda_N})$$

(2) More generally now, let us discuss the case where our matrix Z is diagonalizable. Here we must have a formula as follows, with D being diagonal:

$$Z = BDB^{-1}$$

Now observe that the moments of Z are given by the following formula:

$$\begin{aligned} tr(Z^k) &= tr(BDB^{-1} \cdot BDB^{-1} \ldots BDB^{-1}) \\ &= tr(BD^kB^{-1}) \\ &= tr(D^k) \end{aligned}$$

We conclude, by linearity, that the matrices Z, D have the same distribution:

$$\mu_Z = \mu_D$$

On the other hand, $Z = BDB^{-1}$ shows that Z, D have the same eigenvalues. Thus, if we denote by $\lambda_1, \ldots, \lambda_N \in \mathbb{C}$ these eigenvalues, we obtain, by using (1):

$$\mu_Z = \frac{1}{N}(\delta_{\lambda_1} + \ldots + \delta_{\lambda_N})$$

(3) Let us discuss now the general case, where $Z \in M_N(\mathbb{C})$ is arbitrary. We will use here a well-known trick, stating that the diagonalizable matrices are dense inside $M_N(\mathbb{C})$. Indeed, consider the set of matrices $Z \in M_N(\mathbb{C})$ having distinct eigenvalues. These latter

matrices are given by the following formula, where P_Z is the characteristic polynomial, and where $\Delta(P)=R(P,P')$ is the discriminant of a polynomial:

$$\Delta(P_Z)\neq 0$$

Thus, what we have is the complement of an algebraic surface, which is dense. Now since the matrices having distinct eigenvalues are diagonalizable, it follows that the matrices which are diagonalizable are dense as well. Now with this in hand, the result simply follows from the result for the diagonalizable matrices, (2) above, by continuity. □

Summarizing, we have a nice theory for the matrices $Z\in M_N(\mathbb{C})$, paralleling that of the random variables $f\in L^\infty(X)$. It is tempting at this point to try to go further, and to unify the matrices and the random variables by talking about random matrices:

$$Z\in M_N(L^\infty(X))$$

However, we will not do so right away, because our matrix theory has a flaw. Indeed, all what has being said above, although being nice and conceptual, does not take into account the adjoint matrix, $Z^*=(\bar{Z}_{ji})$. Thus, before getting any further, we must fix this flaw, and talk about the moments and distribution of the pair (Z,Z^*).

3b. Colored moments

This extension can be done as follows:

Definition 3.5. *The moments of a usual complex matrix $Z\in M_N(\mathbb{C})$ are the following numbers, indexed by the colored integers $k=\circ\bullet\bullet\circ\ldots$*

$$M_k=tr(Z^k)$$

with the powers Z^k being defined by $Z^\circ=Z$, $Z^\bullet=Z^*$ and multiplicativity.

All this might seem a bit complicated, but there is no other way of dealing with such things. Indeed, since the variables Z,Z^* do not commute, unless the matrix is normal, $ZZ^*=Z^*Z$, which is something special, that does not happen in general, we are led to colored exponents $k=\circ\bullet\bullet\circ\ldots$ and to the above definition for the moments.

Regarding the distribution, we can use here a similar idea, as follows:

Definition 3.6. *The distribution of a usual complex matrix $Z\in M_N(\mathbb{C})$ is the abstract functional $\mu_Z:\mathbb{C}<X,X^*>\to\mathbb{C}$ given by:*

$$P\to tr(P(Z))$$

In the case where we have a probability measure $\mu_Z\in\mathcal{P}(\mathbb{C})$ such that

$$tr(P(Z))=\int_{\mathbb{C}}P(x)\,d\mu_Z(x)$$

we identify this measure with the distribution, or law of Z.

Observe that by linearity, the distribution is uniquely determined by the moments. In fact, knowing the distribution is the same thing as knowing the moments.

As a first result now, we can recycle Theorem 3.4 in this setting, as follows:

THEOREM 3.7. *Given a matrix $Z\in M_N(\mathbb C)$ which is self-adjoint, $Z=Z^*$, we have the following formula, valid for any polynomial $P\in\mathbb C<X,X^*>$,*

$$tr(P(Z))=\frac{1}{N}(P(\lambda_1)+\ldots+P(\lambda_N))$$

where $\lambda_1,\ldots,\lambda_N\in\mathbb C$ are the eigenvalues of Z. Thus the complex measure

$$\mu_Z=\frac{1}{N}(\delta_{\lambda_1}+\ldots+\delta_{\lambda_N})$$

is the distribution of Z, in the abstract sense of Definition 3.6.

PROOF. This follows indeed from Theorem 3.4, because due to our self-adjointness assumption $Z=Z^*$, the adjoint matrix plays no role in all this. □

Quite remarkably now, the above result extends to the normal case, $ZZ^*=Z^*Z$. This is something non-trivial, the statement and proof being as follows:

THEOREM 3.8. *Given a matrix $Z\in M_N(\mathbb C)$ which is normal, $ZZ^*=Z^*Z$, we have the following formula, valid for any polynomial $P\in\mathbb C<X,X^*>$,*

$$tr(P(Z))=\frac{1}{N}(P(\lambda_1)+\ldots+P(\lambda_N))$$

where $\lambda_1,\ldots,\lambda_N\in\mathbb C$ are the eigenvalues of Z. Thus the complex measure

$$\mu_Z=\frac{1}{N}(\delta_{\lambda_1}+\ldots+\delta_{\lambda_N})$$

is the distribution of Z, in the abstract sense of Definition 3.6.

PROOF. There are several proofs for this fact, one of them being as follows:

(1) Let us first consider the case where the matrix is diagonal:

$$Z=\begin{pmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_N\end{pmatrix}$$

The powers of Z, with respect to colored integer exponents $k=\circ\bullet\bullet\circ\ldots$ as in Definition 3.5 above, are then given by the following formula, with the convention that the numbers λ^k are given by $\lambda^\circ=\lambda,\lambda^\bullet=\bar\lambda$ and multiplicativity:

$$Z^k=\begin{pmatrix}\lambda_1^k&&\\&\ddots&\\&&\lambda_N^k\end{pmatrix}$$

Thus, the moments of Z are given by the following formula:

$$M_k = \frac{1}{N}(\lambda_1^k + \ldots + \lambda_N^k)$$

Regarding now the distribution, this by definition given by:

$$\mu_Z : \mathbb{C} < X, X^* > \to \mathbb{C}$$

$$P \to tr(P(Z))$$

Now since the matrix is normal, $ZZ^* = Z^*Z$, knowing this distribution is the same as knowing its restriction to the usual polynomials in two variables:

$$\mu_Z : \mathbb{C}[X, X^*] \to \mathbb{C}$$

$$P \to tr(P(Z))$$

By using now the fact that Z is diagonal, we conclude that the distribution is:

$$\mu_Z : \mathbb{C}[X, X^*] \to \mathbb{C}$$

$$P \to \frac{1}{N}(P(\lambda_1) + \ldots + P(\lambda_N))$$

But this functional corresponds to integrating P with respect to the following complex measure, that we agree to still denote by μ_Z, and call distribution of Z:

$$\mu_Z = \frac{1}{N}(\delta_{\lambda_1} + \ldots + \delta_{\lambda_N})$$

Summarizing, modulo a number of standard identifications, the distribution of a diagonal matrix $Z \in M_N(\mathbb{C})$ is a complex probability measure, given by the above formula.

(2) In the general case now, where $Z \in M_N(\mathbb{C})$ is normal and arbitrary, we can use the Spectral Theorem for the normal matrices, which tells us that Z, Z^* are jointly diagonalizable. Thus, after changing the basis, we are done by (1). □

Summarizing, we have now a fix for Definitions 3.2 and 3.3 and Theorem 3.4. Importantly, let us mention that the normality assumption in the above results is really needed. Indeed, we have the following basic counterexample:

PROPOSITION 3.9. *The following matrix, which is not normal,*

$$Z = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

has no distribution $\mu_Z \in \mathcal{P}(\mathbb{C})$ *in the sense of Definition 3.6.*

PROOF. We have the following formulae, which show that Z is not normal:

$$ZZ^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad , \quad Z^*Z = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Now observe that the eigenvalues of Z are 0 and 0. Thus the $*$-law formula in Theorem 3.8 above has no chance to extend to this setting, simply because we have:

$$tr(ZZ^*) = tr(Z^*Z) = \frac{1}{2}$$

Even worse, let us prove now that, as claimed, our matrix Z has no distribution $\mu_Z \in \mathcal{P}(\mathbb{C})$, in the sense of Definition 3.6. For this purpose, observe that we have:

$$tr(ZZ^*ZZ^*) = tr((ZZ^*)^2) = tr\begin{pmatrix}1&0\\0&0\end{pmatrix} = \frac{1}{2}$$

On the other hand, we have as well the following formula:

$$tr(ZZZ^*Z^*) = tr(Z^2(Z^*)^2) = tr\begin{pmatrix}0&0\\0&0\end{pmatrix} = 0$$

Since these numbers as different, we cannot obtain them by integrating with respect to a complex measure $\mu_Z \in \mathcal{P}(\mathbb{C})$, and this gives the conclusion in the statement. □

With these preliminaries in hand, we can go ahead and discuss the random matrices, where things become truly interesting. We can extend Definition 3.5, as follows:

DEFINITION 3.10. *The moments of a random matrix $Z \in M_N(L^\infty(X))$ are the following numbers, indexed by the colored integers $k = \circ\bullet\bullet\circ\ldots$*

$$M_k = \int_X tr(Z^k)$$

with the powers Z^k being defined by $Z^\circ = Z$, $Z^\bullet = Z^*$ and multiplicativity.

Observe that this notion extends Definition 3.5 for the usual matrices $Z \in M_N(\mathbb{C})$, which can be recovered with $X = \{.\}$. Also, in the case $N = 1$, where our matrix is just a random variable $Z \in L^\infty(X)$, we recover in this way the usual moments, or rather the joint moments of $Z, \bar{Z}$. As before, the same philosophical comments apply. Since the variables Z, Z^* do not commute, unless the matrix is normal, $ZZ^* = Z^*Z$, we are led to colored exponents $k = \circ\bullet\bullet\circ\ldots$ and to the above definition for the moments. Regarding now the distribution, we can use here a similar extension, as follows:

DEFINITION 3.11. *The distribution of a random matrix $Z \in M_N(L^\infty(X))$ is the abstract functional $\mu_Z : \mathbb{C} < X, X^* > \to \mathbb{C}$ given by:*

$$P \to \int_X tr(P(Z))$$

In the case where we have a probability measure $\mu_Z \in \mathcal{P}(\mathbb{C})$ such that

$$tr(P(Z)) = \int_{\mathbb{C}} P(x)\, d\mu_Z(x)$$

we identify this measure with the distribution, or law of Z.

Observe that by linearity, the distribution is uniquely determined by the moments. In fact, knowing the distribution is the same thing as knowing the moments. As basic examples, for the usual matrices $Z \in M_N(\mathbb{C})$, obtained by taking $X = \{.\}$, we obtain the previous notion of distribution, from Definition 3.6.

Also, for the usual random variables $Z \in L^\infty(X)$, obtained by taking $N = 1$, we obtain in this way the previous notion of distribution, from chapters 1 and 2 above. Indeed, these variables are normal, $ZZ^* = Z^*Z$, and so the corresponding distributions, in the above abstract sense, can be restricted to usual polynomials $P \in \mathbb{C}[X, X^*]$, and then identified with the usual distributions, in the sense of probability theory.

3c. Spectral theory

In order to further clarify all this, and to discuss as well what happens in the self-adjoint and normal cases, with extensions of Theorem 3.7 and Theorem 3.8 above, to the random matrix setting, we will need some basic functional analysis. In view of some further applications, later on, we will go here slightly further than what is needed for the moment. Let us start with the following standard definition:

Definition 3.12. *A C^*-algebra is a complex algebra with unit A, having:*

1. *A norm $a \to ||a||$, making it a Banach algebra (the Cauchy sequences converge).*
2. *An involution $a \to a^*$, which satisfies $||aa^*|| = ||a||^2$, for any $a \in A$.*

As basic examples, we have the usual matrix algebras $M_N(\mathbb{C})$, with the norm and the involution being the usual matrix norm and involution, given by:

$$||M|| = \sup_{||x||=1} ||Mx|| \quad , \quad (M^*)_{ij} = \overline{M}_{ji}$$

Some other basic examples are the algebras $L^\infty(X)$ of essentially bounded functions $f : X \to \mathbb{C}$ on a measured space X, with usual norm and involution, namely:

$$||f|| = \sup_{x \in X} |f(x)| \quad , \quad f^*(x) = \overline{f(x)}$$

We can put these two basic classes of examples together, as follows:

Proposition 3.13. *The random matrix algebras, namely*

$$A = M_N(L^\infty(X))$$

are C^*-algebras, with the usual norm and involution, given by:

$$||Z|| = \sup_{x \in X} ||Z_x|| \quad , \quad (Z^*)_{ij} = \overline{Z}_{ij}$$

These algebras generalize both the algebras $M_N(\mathbb{C})$, and the algebras $L^\infty(X)$.

Proof. The fact that the C^*-algebra axioms are satisfied is clear from definitions. As for the last assertion, this follows by taking $X = \{.\}$ and $N = 1$, respectively. $\square$

Summarizing, the C^*-algebras are natural generalizations of the random matrix algebras. In what follows we will develop some general "noncommutative probability" theory for the C^*-algebras, then come back to the random matrix algebras later on.

As a first observation, the algebra $B(H)$ of bounded linear operators $T : H \to H$ on a Hilbert space is a C^*-algebra, with usual norm and involution, generalizing those for the algebra $M_N(\mathbb{C})$, which appears when the space is finite dimensional, $H = \mathbb{C}^N$. More generally, any closed $*$-subalgebra $A \subset B(H)$ is a C^*-algebra. It is possible to prove that any C^*-algebra appears in this way, and we will be back to this later. For the moment, let us just record the following basic result, dealing with the random matrix case:

Theorem 3.14. *Any algebra of type $L^\infty(X)$ is an operator algebra, as follows:*

$$L^\infty(X) \subset B(L^2(X))$$

$$f \to (g \to fg)$$

More generally, any random matrix algebra is an operator algebra,

$$M_N(L^\infty(X)) \subset M_N(\mathbb{C}) \otimes B(L^2(X))$$

the embedding being the above one, tensored with the identity.

Proof. Given $f \in L^\infty(X)$, consider the following linear operator, on $H = L^2(X)$:

$$T_f(g) = fg$$

Observe that T_f is indeed well-defined, and bounded as well, because:

$$||fg||_2 = \sqrt{\int_X |f(x)|^2 |g(x)|^2 d\mu(x)} \leq ||f||_\infty ||g||_2$$

The application $f \to T_f$ being linear, involutive, continuous, and injective as well, we obtain in this way a C^*-algebra embedding, as claimed:

$$L^\infty(X) \subset B(H)$$

Finally, the last assertion is clear as well, via some standard identifications. □

In view of the above, the elements of the C^*-algebras $a \in A$ can be thought of as being bounded operators on a Hilbert space $T \in B(H)$. Thus, in order to study them, we can emulate spectral theory, in the abstract C^*-algebra setting. Let us begin with:

Definition 3.15. *The spectrum of an element $a \in A$ is the set*

$$\sigma(a) = \left\{\lambda \in \mathbb{C} \big| a - \lambda \notin A^{-1}\right\}$$

where $A^{-1} \subset A$ is the set of invertible elements.

As a basic example, the spectrum of a usual matrix $M \in M_N(\mathbb{C})$ is the collection of its eigenvalues. Also, the spectrum of a continuous function $f \in C(X)$ is its image. In the case of the trivial algebra $A = \mathbb{C}$, the spectrum of an element is the element itself. As a first, basic result regarding spectra, we have:

PROPOSITION 3.16. *We have the following formula, valid for any $a, b \in A$:*

$$\sigma(ab) \cup \{0\} = \sigma(ba) \cup \{0\}$$

Moreover, there are examples where $\sigma(ab) \neq \sigma(ba)$.

PROOF. We first prove that $1 \notin \sigma(ab) \implies 1 \notin \sigma(ba)$. Assume indeed that $1 - ab$ is invertible, with inverse $c = (1 - ab)^{-1}$. We have then:

$$abc = cab = c - 1$$

By using these formulae, we obtain:

$$\begin{aligned}(1 + bca)(1 - ba) &= 1 + bca - ba - bcaba \\ &= 1 + bca - ba - bca + ba \\ &= 1\end{aligned}$$

A similar computation shows that we have $(1 - ba)(1 + bca) = 1$. We conclude that $1 - ba$ is invertible, with inverse $1 + bca$, which proves our claim.

By multiplying by scalars, we deduce from this that we have, for any $\lambda \in \mathbb{C} - \{0\}$:

$$\lambda \notin \sigma(ab) \implies \lambda \notin \sigma(ba)$$

But this leads to the conclusion in the statement, namely:

$$\sigma(ab) \cup \{0\} = \sigma(ba) \cup \{0\}$$

Regarding now the last claim, let us first recall that for usual matrices $a, b \in M_N(\mathbb{C})$ we have $0 \in \sigma(ab) \iff 0 \in \sigma(ba)$, because ab is invertible if any only if ba is.

However, this latter fact fails for general operators on Hilbert spaces. As a basic example, we can take a, b to be the shift $S(e_i) = e_{i+1}$ on the space $l^2(\mathbb{N})$, and its adjoint. Indeed, we have $S^*S = 1$, and SS^* being the projection onto $e_0^\perp$, it is not invertible. □

Given an element $a \in A$, and a rational function $f = P/Q$ having poles outside $\sigma(a)$, we can construct the element $f(a) = P(a)Q(a)^{-1}$. For simplicity, we write:

$$f(a) = \frac{P(a)}{Q(a)}$$

With this convention, we have the following result:

THEOREM 3.17. *We have the "rational functional calculus" formula*

$$\sigma(f(a)) = f(\sigma(a))$$

valid for any rational function $f \in \mathbb{C}(X)$ having poles outside $\sigma(a)$.

PROOF. In order to prove this result, we can proceed in two steps, as follows:

(1) In the polynomial case, $f\in\mathbb{C}[X]$, we pick $\lambda\in\mathbb{C}$, and we write:

$$f(X)-\lambda=c(X-r_1)\ldots(X-r_n)$$

We have then, as desired:

$$\begin{aligned}\lambda\notin\sigma(f(a))&\iff f(a)-\lambda\in A^{-1}\\&\iff c(a-r_1)\ldots(a-r_n)\in A^{-1}\\&\iff a-r_1,\ldots,a-r_n\in A^{-1}\\&\iff r_1,\ldots,r_n\notin\sigma(a)\\&\iff \lambda\notin f(\sigma(a))\end{aligned}$$

(2) In the general case, $f\in\mathbb{C}(X)$, we pick $\lambda\in\mathbb{C}$, we write $f=P/Q$, and we set $F=P-\lambda Q$. By using (1), we obtain:

$$\begin{aligned}\lambda\in\sigma(f(a))&\iff F(a)\notin A^{-1}\\&\iff 0\in\sigma(F(a))\\&\iff 0\in F(\sigma(a))\\&\iff \exists\mu\in\sigma(a),F(\mu)=0\\&\iff \lambda\in f(\sigma(a))\end{aligned}$$

Thus, we have obtained the formula in the statement. □

Given an element $a\in A$, its spectral radius $\rho(a)$ is the radius of the smallest disk centered at 0 containing $\sigma(a)$. With this convention, we have the following key result:

THEOREM 3.18. *Let A be a C^*-algebra.*

(1) *The spectrum of a norm one element is in the unit disk.*
(2) *The spectrum of a unitary element ($a^*=a^{-1}$) is on the unit circle.*
(3) *The spectrum of a self-adjoint element ($a=a^*$) consists of real numbers.*
(4) *The spectral radius of a normal element ($aa^*=a^*a$) is equal to its norm.*

PROOF. We use the various results established above.

(1) This comes from the following formula, valid when $||a||<1$:

$$\frac{1}{1-a}=1+a+a^2+\ldots$$

(2) Assuming $a^*=a^{-1}$, we have the following norm computations:

$$||a||=\sqrt{||aa^*||}=\sqrt{1}=1$$

$$||a^{-1}||=||a^*||=||a||=1$$

If we denote by D the unit disk, we obtain from this, by using (1):

$$||a|| = 1 \implies \sigma(a) \subset D$$

$$||a^{-1}|| = 1 \implies \sigma(a^{-1}) \subset D$$

On the other hand, by using the rational function $f(z) = z^{-1}$, we have:

$$\sigma(a^{-1}) \subset D \implies \sigma(a) \subset D^{-1}$$

Now by putting everything together we obtain, as desired:

$$\sigma(a) \subset D \cap D^{-1} = \mathbb{T}$$

(3) This follows by using the result (2), just established above, and Theorem 3.18, with the following rational function, depending on $t \in \mathbb{R}$:

$$f(z) = \frac{z + it}{z - it}$$

Indeed, for $t >> 0$ the element $f(a)$ is well-defined, and we have:

$$\left(\frac{a + it}{a - it}\right)^* = \frac{a - it}{a + it} = \left(\frac{a + it}{a - it}\right)^{-1}$$

Thus the element $f(a)$ is a unitary, and by using (2) its spectrum is contained in $\mathbb{T}$. We conclude that we have $f(\sigma(a)) = \sigma(f(a)) \subset \mathbb{T}$. Thus we obtain:

$$\sigma(a) \subset f^{-1}(\mathbb{T}) = \mathbb{R}$$

In other words, we have proved the result.

(4) We already know from (1) that we have the following inequality:

$$\rho(a) \leq ||a||$$

For the converse, we fix a number $\rho > \rho(a)$. We have then:

$$\int_{|z|=\rho} \frac{z^n}{z - a}\, dz = \sum_{k=0}^{\infty} \left(\int_{|z|=\rho} z^{n-k-1} dz\right) a^k = a^{n-1}$$

By applying the norm and taking n-th roots we obtain from this:

$$\rho \geq \lim_{n \to \infty} ||a^n||^{1/n}$$

In the case $a = a^*$ we have $||a^n|| = ||a||^n$ for any exponent of the form $n = 2^k$, and by taking n-th roots we get $\rho \geq ||a||$. But this gives the missing inequality:

$$\rho(a) \geq ||a||$$

In the general case $aa^* = a^*a$ we have $a^n(a^n)^* = (aa^*)^n$. We obtain from this:

$$\rho(a)^2 = \rho(aa^*)$$

Now since aa^* is self-adjoint, we get $\rho(aa^*) = ||a||^2$, and we are done. □

We are now in position of proving a key result, namely:

THEOREM 3.19. *For any normal element $a \in A$ we have an identification as follows:*

$$< a >= C(\sigma(a))$$

In addition, given a function $f \in C(\sigma(a))$, we can apply it to a, and we have

$$\sigma(f(a)) = f(\sigma(a))$$

which generalizes the previous rational calculus formula, in the normal case.

PROOF. It is routine to construct an evaluation map, as follows:

$$ev :< a >\to C(\sigma(a))$$

We first prove that ev is involutive. We use the following formula:

$$b = \frac{b+b^*}{2} - i\cdot\frac{i(b-b^*)}{2}$$

Thus it is enough to prove that for self-adjoint elements b we have:

$$ev_{b^*} = ev_b^*$$

But this is the same as proving that $b = b^*$ implies that ev_b is a real function, which is in turn true, because $ev_b(\chi) = \chi(b)$ is an element of $\sigma(b)$, contained in $\mathbb{R}$.

Since $< a >$ is commutative, each element is normal, so ev is isometric:

$$||ev_b|| = \rho(b) = ||b||$$

It remains to prove that ev is surjective. But this follows from the Stone-Weierstrass theorem, because $ev(< a >)$ is a closed subalgebra of $C(\sigma(a))$, separating the points.

Finally, the last assertion is clear from the first one. □

In order to talk now about noncommutative probability, we first have to develop the theory of positive elements, and positive linear forms. First, we have the following result:

PROPOSITION 3.20. *For an element $a \in A$, the following are equivalent:*

(1) *a is positive, in the sense that $\sigma(a) \subset [0,\infty)$.*
(2) *$a = b^2$, for some $b \in A$ satisfying $b = b^*$.*
(3) *$a = cc^*$, for some $c \in A$.*

PROOF. This is something quite standard, as follows:

(1) $\implies$ (2) Observe that $\sigma(a) \subset \mathbb{R}$ implies $a = a^*$. Thus the algebra $< a >$ is commutative, and by using the Gelfand theorem, we can set:

$$b = \sqrt{a}$$

(2) $\implies$ (3) This is trivial, because we can set $c = b$.

(2) $\implies$ (1) This is clear too, because we have:

$$\sigma(a)=\sigma(b^2)=\sigma(b)^2\subset\mathbb{R}^2=[0,\infty)$$

(3) $\implies$ (1) We proceed by contradiction. By multiplying c by a suitable element of $<cc^*>$, we are led to the existence of an element $d\neq 0$ satisfying:

$$-dd^*\geq 0$$

By writing now $d=x+iy$ with $x=x^*,y=y^*$ we have:

$$dd^*+d^*d=2(x^2+y^2)\geq 0$$

We conclude from this that we have $d^*d\geq 0$. But this contradicts the elementary fact that $\sigma(dd^*),\sigma(d^*d)$ must coincide outside $\{0\}$, coming from Proposition 3.16 above. □

We can talk as well about positive linear forms, as follows:

Definition 3.21. *Consider a linear map $\varphi:A\to\mathbb{C}$.*
1. *φ is called positive when $a\geq 0\implies\varphi(a)\geq 0$.*
2. *φ is called faithful and positive when $a\geq 0,a\neq 0\implies\varphi(a)>0$.*

In the commutative case, $A=C(X)$, the positive linear forms appear as follows, with μ being positive, and strictly positive if we want φ to be faithful and positive:

$$\varphi(f)=\int_X f(x)d\mu(x)$$

In general, the positive linear forms can be thought of as being integration functionals with respect to some underlying "positive measures".

Let us review now the other fundamental result regarding the C^*-algebras, namely the representation theorem of Gelfand, Naimark and Segal, which states that any C^*-algebra appears as an algebra of operators, $A\subset B(H)$, over some Hilbert space H. In the commutative case, the precise statement is as follows:

Proposition 3.22. *Let A be a commutative C^*-algebra, write $A=C(X)$, with X being a compact space, and let μ be a positive measure on X. We have then an embedding*

$$A\subset B(H)$$

where $H=L^2(X)$, with $f\in A$ corresponding to the operator $g\to fg$.

Proof. Given $f\in C(X)$, consider the following operator, on the space $H=L^2(X)$:

$$T_f(g)=fg$$

Observe that T_f is indeed well-defined, and bounded as well, because:

$$||fg||_2=\sqrt{\int_X|f(x)|^2|g(x)|^2d\mu(x)}\leq||f||_\infty||g||_2$$

The application $f \to T_f$ being linear, involutive, continuous, and injective as well, we obtain in this way a C^*-algebra embedding $A \subset B(H)$, as claimed. □

In general, the idea will be that of extending the above construction. In order to do so, we will use a functional analysis trick, coming from the Riesz theorem, which amounts in replacing the positive measures μ with the corresponding integration functionals:

PROPOSITION 3.23. *Let $\varphi : A \to \mathbb{C}$ be a positive linear form.*

(1) *$< a, b >= \varphi(ab^*)$ defines a generalized scalar product on A.*
(2) *By separating and completing we obtain a Hilbert space H.*
(3) *$\pi(a) : b \to ab$ defines a representation $\pi : A \to B(H)$.*
(4) *If φ is faithful in the above sense, then π is faithful.*

PROOF. Almost everything here is straightforward, as follows:

(1) This is clear from definitions, and from Proposition 3.20.

(2) This is a standard procedure, which works for any scalar product.

(3) All the verifications here are standard algebraic computations.

(4) This follows indeed from $a \neq 0 \implies \pi(aa^*) \neq 0 \implies \pi(a) \neq 0$. □

In order to establish the GNS theorem, it remains to prove that any C^*-algebra has a faithful and positive linear form $\varphi : A \to \mathbb{C}$.

This is something more technical:

THEOREM 3.24. *Let A be a C^*-algebra.*

(1) *Any positive linear form $\varphi : A \to \mathbb{C}$ is continuous.*
(2) *A linear form φ is positive iff there is a norm one $h \in A_+$ such that $||\varphi|| = \varphi(h)$.*
(3) *For any $a \in A$ there exists a positive norm one form φ such that $\varphi(aa^*) = ||a||^2$.*
(4) *If A is separable there is a faithful positive form $\varphi : A \to \mathbb{C}$.*

PROOF. The proof here, which is quite technical, inspired from the existence proof of the probability measures on abstract compact spaces, goes as follows:

(1) This follows from Proposition 3.23, via the following inequality:

$$|\varphi(a)| \leq ||\pi(a)||\varphi(1) \leq ||a||\varphi(1)$$

(2) In one sense we can take $h = 1$. Conversely, let $a \in A_+$, $||a|| \leq 1$. We have:

$$|\varphi(h) - \varphi(a)| \leq ||\varphi|| \cdot ||h - a|| \leq \varphi(h)1 = \varphi(h)$$

Thus we have $Re(\varphi(a)) \geq 0$, and it remains to prove that the following holds:

$$a = a^* \implies \varphi(a) \in \mathbb{R}$$

By using $1 - h \geq 0$ we can apply the above to $a = 1 - h$ and we obtain:

$$Re(\varphi(1 - h)) \geq 0$$

We conclude that $Re(\varphi(1)) \geq Re(\varphi(h)) = ||\varphi||$, and so $\varphi(1) = ||\varphi||$. Thus, we can assume $h = 1$. Now observe that for any self-adjoint element a, and $t \in \mathbb{R}$ we have:

$$\begin{aligned} |\varphi(1+ita)|^2 &\leq ||\varphi||^2 \cdot ||1+ita||^2 \\ &= \varphi(1)^2||1+t^2a^2|| \\ &\leq \varphi(1)^2(1+t^2||a||^2) \end{aligned}$$

On the other hand with $\varphi(a) = x + iy$ we have:

$$\begin{aligned} |\varphi(1+ita)| &= |\varphi(1) - ty + itx| \\ &\geq (\varphi(1) - ty)^2 \end{aligned}$$

We therefore obtain that for any $t \in \mathbb{R}$ we have:

$$\varphi(1)^2(1+t^2||a||^2) \geq (\varphi(1) - ty)^2$$

Thus we have $y = 0$, and this finishes the proof of our remaining claim.

(3) Consider the linear subspace of A spanned by the element aa^*. We can define here a linear form by $\varphi(\lambda aa^*) = \lambda||a||^2$. This linear form has norm one, and by Hahn-Banach we get a norm one extension to the whole A. The positivity of φ follows from (2).

(4) Let (a_n) be a dense sequence inside A. For any n we can construct as in (3) a positive form satisfying $\varphi_n(a_na_n^*) = ||a_n||^2$, and then define φ in the following way:

$$\varphi = \sum_{n=1}^\infty \frac{\varphi_n}{2^n}$$

Let $a \in A$ be a nonzero element. Pick a_n close to a and consider the pair (H, π) associated to the pair (A, φ_n), as in Proposition 3.23. We have then:

$$\begin{aligned} \varphi_n(aa^*) &= ||\pi(a)1|| \\ &\geq ||\pi(a_n)1|| - ||a - a_n|| \\ &= ||a_n|| - ||a - a_n|| \\ &> 0 \end{aligned}$$

Thus $\varphi_n(aa^*) > 0$. It follows that we have $\varphi(aa^*) > 0$, and we are done. □

With these ingredients in hand, we can now state and prove:

THEOREM 3.25 (GNS theorem). *Let A be a C^*-algebra.*

1. *A appears as a closed $*$-subalgebra $A \subset B(H)$, for some Hilbert space H.*
2. *When A is separable (usually the case), H can be chosen to be separable.*
3. *When A is finite dimensional, H can be chosen to be finite dimensional.*

PROOF. This result, from [61], follows indeed by combining the construction from Proposition 3.23 with the existence result from Theorem 3.24 above. □

More details on all the above can be found in any operator algebra book.

3d. Spectral measures

Let us discuss now noncommutative probability theory. We first have:

DEFINITION 3.26. *Let A be a C^*-algebra, given with a positive trace $tr : A \to \mathbb{C}$.*

(1) *The elements $a \in A$ are called random variables.*
(2) *The moments of such a variable are the numbers $M_k(a) = tr(a^k)$.*
(3) *The law of such a variable is the functional $\mu_a : P \to tr(P(a))$.*

Here $k = \circ\bullet\bullet\circ\ldots$ is by definition a colored integer, and the powers a^k are defined by the following formulae, and multiplicativity:

$$a^\emptyset = 1 \quad , \quad a^\circ = a \quad , \quad a^\bullet = a^*$$

As for the polynomial P, this is a noncommuting $*$-polynomial in one variable:

$$P \in \mathbb{C} < X, X^* >$$

Observe that the law is uniquely determined by the moments. We have:

THEOREM 3.27. *Let A be a C^*-algebra, with a trace tr, and let $a = a^* \in A$.*

(1) *μ_a is a real probability measure, satisfying $supp(\mu_a) \subset \sigma(a)$.*
(2) *Assuming that tr is faithful, we have $supp(\mu_a) = \sigma(a)$.*

Moreover, all this can be generalized to the normal case, $aa^* = a^*a$.

PROOF. This is standard, coming from Theorem 3.20, and the Riesz theorem. □

Getting back now to the random matrices, we have:

THEOREM 3.28. *In the normal case, $ZZ^* = Z^*Z$, the law,*

$$\mu : \mathbb{C} < X, X^* > \to \mathbb{C}$$

$$P \to \frac{1}{N}\int_X tr(P(Z))$$

when restricted to the usual polynomials in two variables,

$$\mu : \mathbb{C}[X, X^*] \to \mathbb{C}$$

$$P \to \frac{1}{N}\int_X tr(P(Z))$$

must come from a probability measure on the spectrum $\sigma(Z) \subset \mathbb{C}$, as:

$$\mu(P) = \int_{\sigma(T)} P(x)d\mu(x)$$

We agree to use the symbol μ for all these notions.

PROOF. This comes from standard measure theory, as developed above, the only subtle ingredient being the fact that $ZZ^* = Z^*Z$ implies:

$$< Z >= C(\sigma(Z))$$

But this is something that we already know. □

In the self-adjoint case, the statement is as follows:

THEOREM 3.29. *In the self-adjoint case, $Z = Z^*$, the law,*

$$\mu : \mathbb{C} < X, X^* > \to \mathbb{C}$$

$$P \to \frac{1}{N}\int_X tr(P(Z))$$

when restricted to the usual polynomials

$$\mu : \mathbb{C}[X] \to \mathbb{C}$$

$$P \to \frac{1}{N}\int_X tr(P(Z))$$

must come from a probability measure on the spectrum $\sigma(Z) \subset \mathbb{R}$, as:

$$\mu(P) = \int_{\sigma(T)} P(x)d\mu(x)$$

We agree to use the symbol μ for all these notions.

PROOF. This comes once again from standard measure theory and spectral theory, as developed above, the only subtle ingredient being the fact that $Z = Z^*$ implies:

$$\sigma(Z) \subset \mathbb{R}$$

But this is something that we already know. □

3e. Exercises

CHAPTER 4

Random matrices

4a. Normal variables

The random matrices are simple and fundamental mathematical objects, virtually appearing in all areas of mathematics and physics. They are defined as follows:

Definition 4.1. *A random matrix is a square matrix of type*

$$Z \in M_N(L^\infty(X))$$

with X being a probability space, and $N \in \mathbb{N}$ being an integer.

As basic examples, we have the usual matrices $Z \in M_N(\mathbb{C})$, obtained by taking $X = \{.\}$. Also, we have the usual random variables $Z \in L^\infty(X)$, obtained by taking $N = 1$. In general, what we have is a kind of combination of these 2 situations.

We will be mostly interested in what follows in random matrices constructed by using complex Gaussian variables, so let us first discuss the theory here. We first have:

Definition 4.2. *The complex Gaussian law of parameter $t > 0$ is*

$$G_t = law\left(\frac{1}{\sqrt{2}}(a + ib)\right)$$

where a, b are independent, each following the law g_t.

As in the real case, these measures form convolution semigroups:

Theorem 4.3. *The complex Gaussian laws have the property*

$$G_s * G_t = G_{s+t}$$

for any $s, t > 0$, and so they form a convolution semigroup.

Proof. This follows indeed from the real result, established in chapter 1 above, by taking real and imaginary parts. □

We have the following complex analogue of the CLT:

THEOREM 4.4 (CLT). *Given complex random variables $f_1, f_2, f_3, \ldots \in L^\infty(X)$ which are i.i.d., centered, and with variance $t > 0$, we have, with $n \to \infty$, in moments,*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n f_i \sim G_t$$

where G_t is the complex Gaussian law of parameter t.

PROOF. This follows indeed from the real CLT, established in chapter 1 above, by taking real and imaginary parts. □

Regarding now the moments, we first have the following result:

THEOREM 4.5. *The moments of the complex Gaussian law are given by*

$$M_k = |\mathcal{P}_2(k)|$$

for any colored integer $k = \circ\bullet\bullet\circ\ldots$, where $\mathcal{P}_2(k)$ are the matching pairings of $\{1, \ldots, k\}$.

PROOF. This is something well-known, which can be done in several steps, as follows:

(1) We recall from chapter 1 above that the moments of the real Gaussian law g_1, with respect to integer exponents $k \in \mathbb{N}$, are the following numbers:

$$m_k = |P_2(k)|$$

Numerically, we have the following formula, explained as well in section 5:

$$m_k = \begin{cases} k!! & (k \text{ even}) \\ 0 & (k \text{ odd}) \end{cases}$$

(2) We will show here that in what concerns the complex Gaussian law G_1, similar results hold. Numerically, we will prove that we have the following formula, where a colored integer $k = \circ\bullet\bullet\circ\ldots$ is called uniform when it contains the same number of $\circ$ and $\bullet$, and where $|k| \in \mathbb{N}$ is the length of such a colored integer:

$$M_k = \begin{cases} (|k|/2)! & (k \text{ uniform}) \\ 0 & (k \text{ not uniform}) \end{cases}$$

Now since the matching partitions $\pi \in \mathcal{P}_2(k)$ are counted by exactly the same numbers, and this for trivial reasons, we will obtain the formula in the statement, namely:

$$M_k = |\mathcal{P}_2(k)|$$

(3) This was for the plan. In practice now, we must compute the moments, with respect to colored integer exponents $k = \circ\bullet\bullet\circ\ldots$, of the variable in the statement:

$$c = \frac{1}{\sqrt{2}}(a + ib)$$

As a first observation, in the case where such an exponent $k = \circ\bullet\bullet\circ\ldots$ is not uniform in $\circ,\bullet$, a rotation argument shows that the corresponding moment of c vanishes. To be more precise, the variable $c' = wc$ can be shown to be complex Gaussian too, for any $w \in \mathbb{C}$, and from $M_k(c) = M_k(c')$ we obtain $M_k(c) = 0$, in this case.

(4) In the uniform case now, where $k = \circ\bullet\bullet\circ\ldots$ consists of p copies of $\circ$ and p copies of $\bullet$, the corresponding moment can be computed as follows:

$$\begin{aligned}
M_k &= \int (c\bar{c})^p \\
&= \frac{1}{2^p}\int (a^2+b^2)^p \\
&= \frac{1}{2^p}\sum_s \binom{p}{s}\int a^{2s}\int b^{2p-2s} \\
&= \frac{1}{2^p}\sum_s \binom{p}{s}(2s)!!(2p-2s)!! \\
&= \frac{1}{2^p}\sum_s \frac{p!}{s!(p-s)!}\cdot\frac{(2s)!}{2^s s!}\cdot\frac{(2p-2s)!}{2^{p-s}(p-s)!} \\
&= \frac{p!}{4^p}\sum_s \binom{2s}{s}\binom{2p-2s}{p-s}
\end{aligned}$$

(5) In order to finish now the computation, let us recall that we have the following formula, coming from the generalized binomial formula, or from the Taylor formula:

$$\frac{1}{\sqrt{1+t}} = \sum_{k=0}^\infty \binom{2k}{k}\left(\frac{-t}{4}\right)^k$$

By taking the square of this series, we obtain the following formula:

$$\begin{aligned}
\frac{1}{1+t} &= \sum_{ks}\binom{2k}{k}\binom{2s}{s}\left(\frac{-t}{4}\right)^{k+s} \\
&= \sum_p \left(\frac{-t}{4}\right)^p \sum_s \binom{2s}{s}\binom{2p-2s}{p-s}
\end{aligned}$$

Now by looking at the coefficient of t^p on both sides, we conclude that the sum on the right equals 4^p. Thus, we can finish the moment computation in (4), as follows:

$$M_p = \frac{p!}{4^p}\times 4^p = p!$$

(6) As a conclusion, if we denote by $|k|$ the length of a colored integer $k=\circ\bullet\bullet\circ\ldots$, the moments of the variable c in the statement are given by:

$$M_k=\begin{cases}(|k|/2)! & (k\text{ uniform})\\ 0 & (k\text{ not uniform})\end{cases}$$

On the other hand, the numbers $|\mathcal{P}_2(k)|$ are given by exactly the same formula. Indeed, in order to have matching pairings of k, our exponent $k=\circ\bullet\bullet\circ\ldots$ must be uniform, consisting of p copies of $\circ$ and p copies of $\bullet$, with $p=|k|/2$. But then the matching pairings of k correspond to the permutations of the $\bullet$ symbols, as to be matched with $\circ$ symbols, and so we have $p!$ such matching pairings. Thus, we have the same formula as for the moments of c, and we are led to the conclusion in the statement. □

More generally now, we have the following result:

Theorem 4.6. *The moments of the complex normal law are the numbers*

$$M_k(G_t)=\sum_{\pi\in\mathcal{P}_2(k)}t^{|\pi|}$$

where $\mathcal{P}_2(k)$ are the matching pairings of $\{1,\ldots,k\}$, and $|.|$ is the number of blocks.

Proof. This follows indeed from Theorem 4.5 above. □

In relation now with the spheres, and the rotations, we have:

Theorem 4.7. *The rescalings $z_i/\sqrt{N}$ of the complex hyperspherical variables*

$$z_i:S^{N-1}_{\mathbb{C}}\to\mathbb{C}$$

become complex Gaussian with $N\to\infty$. Also, the rescalings $U_{ij}/\sqrt{N}$ of the variables

$$U_{ij}:U_N\to\mathbb{C}$$

become complex Gaussian with $N\to\infty$.

Proof. This follows from the real results, by doing some computations. □

4b. Wigner and Wishart

Now back to the random matrices, with the above ingredients in hand, let us begin by specifying the precise classes of matrices that we are interested in. First we have the complex Gaussian matrices, which are constructed as follows:

Definition 4.8. *A complex Gaussian matrix is a random matrix of type*

$$Z\in M_N(L^\infty(X))$$

which has i.i.d. complex normal entries.

We will be interested as well in the Wigner random matrices, which are the self-adjoint versions of these matrices. These are constructed as follows:

Definition 4.9. *A Wigner matrix is a random matrix of type*

$$Z\in M_N(L^\infty(X))$$

which has i.i.d. complex normal entries, up to the constraint $Z=Z^*$.

In other words, a Wigner matrix must be as follows, with X_i being real normal variables, Z_{ij} being complex normal variables, and all these variables being independent:

$$W=\begin{pmatrix}X_1&Z_{12}&\ldots&\ldots&Z_{1N}\\ \bar{Z}_{12}&X_2&\ddots&&\vdots\\ \vdots&\ddots&\ddots&\ddots&\vdots\\ \vdots&&\ddots&X_{N-1}&Z_{N-1,N}\\ \bar{Z}_{1N}&\ldots&\ldots&\bar{Z}_{N-1,N}&X_N\end{pmatrix}$$

Finally, we will be interested as well in the complex Wishart random matrices, which are the positive versions of these matrices, which are constructed as follows:

Definition 4.10. *A complex Wishart matrix is a random matrix of type*

$$Z=GG^*\in M_N(L^\infty(X))$$

with G being a complex Gaussian matrix.

We discuss now the computation of the law of the Wigner and Wishart matrices. These matrices are self-adjoint, and their law is as follows:

Theorem 4.11. *In the self-adjoint case, $Z=Z^*$, the law,*

$$\mu:\mathbb{C}<X,X^*>\to\mathbb{C}$$
$$P\to\frac{1}{N}\int_X tr(P(Z))$$

when restricted to the usual polynomials, $\mu:\mathbb{C}[X]\to\mathbb{C}$, must come from a probability measure on the spectrum $\sigma(Z)\subset\mathbb{R}$, as:

$$\mu(P)=\int_{\sigma(T)}P(x)d\mu(x)$$

We agree to use the symbol μ for all these notions.

Proof. This is something that we already know, coming from standard measure theory and spectral theory, and based on the fact that $Z=Z^*$ implies $\sigma(Z)\subset\mathbb{R}$. □

In order to compute the laws of the Wigner and Wishart matrices, we will use the moment method. The combinatorics here is that of the Catalan numbers, so let us start with some preliminaries here. First, we have the following well-known result:

THEOREM 4.12. *The Catalan numbers, which are by definition given by*

$$C_k = |NC_2(2k)|$$

satisfy the following recurrence formula,

$$C_{k+1} = \sum_{a+b=k} C_a C_b$$

their generating series $f(z) = \sum_{k\geq 0} C_k z^k$ *satisfies the equation*

$$zf^2 - f + 1 = 0$$

and we have the following explicit formula for these numbers:

$$C_k = \frac{1}{k+1}\binom{2k}{k}$$

PROOF. We must count the noncrossing pairings of $\{1,\ldots,2k\}$. Now observe that such a pairing appears by pairing 1 to an odd number, $2a+1$, and then inserting a noncrossing pairing of $\{2,\ldots,2a\}$, and a noncrossing pairing of $\{2a+2,\ldots,2l\}$. We conclude from this that we have the following recurrence formula for the Catalan numbers:

$$C_k = \sum_{a+b=k-1} C_a C_b$$

Consider now generating series of the Catalan numbers:

$$f(z) = \sum_{k\geq 0} C_k z^k$$

In terms of this generating series, the above recurrence gives:

$$\begin{aligned} zf^2 &= \sum_{a,b\geq 0} C_a C_b z^{a+b+1} \\ &= \sum_{k\geq 1}\sum_{a+b=k-1} C_a C_b z^k \\ &= \sum_{k\geq 1} C_k z^k \\ &= f-1 \end{aligned}$$

Thus f satisfies the following degree 2 equation:

$$zf^2 - f + 1 = 0$$

By solving this equation, and choosing the solution which is bounded at $z=0$, we obtain the following formula:

$$f(z) = \frac{1-\sqrt{1-4z}}{2z}$$

By using now the Taylor formula for $\sqrt{x}$, we obtain the following formula:

$$f(z)=\sum_{k\geq 0}\frac{1}{k+1}\binom{2k}{k}z^k$$

It follows that the Catalan numbers are given by:

$$C_k=\frac{1}{k+1}\binom{2k}{k}$$

Thus, we are led to the conclusion in the statement. □

The Catalan numbers are central objects in probability as well, and we have the following key result here:

THEOREM 4.13. *The normalized Wigner semicircle law,*

$$\gamma_1=\frac{1}{2\pi}\sqrt{4-x^2}dx$$

has the Catalan numbers as even moments. As for the odd moments, these all vanish.

PROOF. The even moments of the Wigner law can be computed with the change of variable $x=2\cos t$, and we are led to the following formula:

$$\frac{1}{2\pi}\int_{-2}^2\sqrt{4-x^2}x^{2k}dx=C_k$$

As for the odd moments, these all vanish, because the density of γ_1 is an even function. Thus, we are led to the conclusion in the statement. □

In order to deal with the Wishart matrices, we will need:

PROPOSITION 4.14. *We have a bijection $NC(k)\simeq NC_2(2k)$, constructed as follows:*

1. *The application $NC(k)\to NC_2(2k)$ is the "fattening" one, obtained by doubling all the legs, and doubling all the strings as well.*
2. *Its inverse $NC_2(2k)\to NC(k)$ is the "shrinking" application, obtained by collapsing pairs of consecutive neighbors.*

PROOF. The fact that the two operations in the statement are indeed inverse to each other is clear, by computing the corresponding two compositions, with the remark that the construction of the fattening operation requires the partitions to be noncrossing. □

As a consequence of the above result, we have the following statement, complementing the various combinatorial results from Theorem 4.12 above:

THEOREM 4.15. *The Catalan numbers appear as well as*

$$C_k=|NC(k)|$$

where $NC(k)$ is the set of all noncrossing partitions of $\{1,\ldots,k\}$.

PROOF. This follows indeed from Proposition 4.14 above. □

We are led to the following result:

THEOREM 4.16. *The moments of the Marchenko-Pastur law,*

$$\pi_1 = \frac{1}{2\pi}\sqrt{4x^{-1}-1}\,dx$$

are the Catalan numbers, $C_k = \frac{1}{k+1}\binom{2k}{k}$.

PROOF. The moments of the Marchenko-Pastur law can be computed with the change of variable $x = 4\cos^2 t$, and we are led to the following formula:

$$\frac{1}{2\pi}\int_0^4 \sqrt{4x^{-1}-1}\,x^k dx = C_k$$

Thus, we are led to the conclusion in the statement. □

4c. Limiting laws

Now back to the random matrices, as a first result, due to Wigner, we have:

THEOREM 4.17. *Given a sequence of Wigner random matrices*

$$Z_N \in M_N(L^\infty(X))$$

which by definition have i.i.d. complex normal entries, up to the constraint $Z_N = Z_N^*$, *we have, in the* $N \to \infty$ *limit,*

$$Z_N \sim \frac{1}{2\pi}\sqrt{4-x^2}dx$$

with the limiting measure being Wigner's semicircle law γ_1.

PROOF. This follows by using the moment method. The matrices being self-adjoint, we are only interested in the usual moments, with respect to exponents $k \in \mathbb{N}$, given by:

$$M_k = \frac{1}{N}\int_X Tr(Z_N^k)$$

But we can compute these moments in a straightforward way, by expanding and using the Wick formula. In the $N \to \infty$ limit the combinatorics simplifies, and we are led, up to some normalizations, to the Catalan numbers:

$$C_k = \frac{1}{k+1}\binom{2k}{k}$$

But these numbers are the even moments of the semicircle law γ_1, as explained in Theorem 4.13 above, and we are done. □

As a second result now, due to Marchenko and Pastur, we have:

THEOREM 4.18. *Given a sequence of complex Wishart random matrices*

$$Z_N = \frac{1}{M}\, Y_N Y_N^* \in M_N(L^\infty(X))$$

with Y_N being $N \times M$ complex Gaussian of parameter 1, we have

$$Z_N \sim \frac{1}{2\pi}\sqrt{4x^{-1} - 1}\, dx$$

with $M = tN \to \infty$, with the limiting measure being the Marchenko-Pastur law π_1.

PROOF. This follows by using the moment method. The matrices being self-adjoint, we are only interested in the usual moments, with respect to exponents $k \in \mathbb{N}$, given by:

$$M_k = \frac{1}{N}\int_X Tr(Z_N^k)$$

But we can compute these moments in a straightforward way, by expanding and using the Wick formula. In the $N \to \infty$ limit the combinatorics simplifies, and we are led, to the Catalan numbers. But these numbers are the moments of the Marchenko-Pastur law π_1, as explained in Theorem 4.16 above, and we are done. □

We will need as well parametric versions of the above measures. The subject here is quite mysterious, at least with our knowledge so far, and we must proceed as follows:

DEFINITION 4.19. *The Wigner and Marchenko-Pastur laws γ_t, π_t of parameter $t > 0$ are constructed as follows:*

(1) *γ_t is the rescaling of the law γ_1 from $[-2, 2]$ to $[-2t, 2t]$.*
(2) *π_t is the law of the square of a variable following the law γ_t.*

We will see later on a more conceptual construction of γ_t, π_t out of their $t = 1$ values γ_1, π_1, by using free convolution. For our purposes here, we will only need the densities and the moments of γ_t, π_t. Regarding the densities, the formulae are as follows:

THEOREM 4.20. *The Wigner law of parameter $t > 0$ is the following measure,*

$$\gamma_t = \frac{1}{2\pi t}\sqrt{4t^2 - x^2}dx$$

and the Marchenko-Pastur law of parameter $t > 0$ is the following measure,

$$\pi_t = \max(1 - t, 0)\delta_0 + \frac{\sqrt{4t - (x - 1 - t)^2}}{2\pi x}\, dx$$

the supports being respectively $[-2t, 2t]$ and $[0, 4t^2]$.

PROOF. This follows indeed by doing some computations. □

Regarding now the moments of γ_t, π_t, the formulae here are quite similar to those for the Gaussian and Poisson laws g_t, p_t, the result being as follows:

THEOREM 4.21. *The moments of the Wigner laws are the numbers*

$$M_k(\gamma_t)=\sum_{\pi\in NC_2(k)}t^{|\pi|}$$

and the moments of the Marchenko-Pastur laws are the numbers

$$M_k(\pi_t)=\sum_{\pi\in NC(k)}t^{|\pi|}$$

where "NC" stands for noncrossing.

PROOF. As a first observation, at $t=1$ the first formula reads:

$$M_k(\gamma_t)=|NC_2(k)|$$

But this is elementary to establish, via partial integration and a recurrence, the numbers in question being the Catalan numbers, shifted via $k\to 2k$:

$$C_k=\frac{1}{k+1}\binom{2k}{k}$$

Regarding now the second formula, at $t=1$ this reads:

$$M_k(\pi_t)=|NC(k)|$$

But this is once again elementary to establish, via partial integration and a recurrence, the numbers in question being the Catalan numbers, unshifted this time.

In the general case, $t>0$, all this follows indeed by doing some computations. □

Getting back now to the random matrices, after all these preliminaries, we can eventually formulate some results. As a first result, due to Wigner, we have:

THEOREM 4.22. *Given a sequence of Wigner random matrices*

$$Z_N\in M_N(L^\infty(X))$$

which by definition have i.i.d. complex normal entries, up to the constraint

$$Z_N=Z_N^*$$

we have, in the $N\to\infty$ limit,

$$Z_N\sim\frac{1}{2\pi t}\sqrt{4t^2-x^2}dx$$

with the limiting measure being Wigner's semicircle law γ_t.

PROOF. This follows by using the moment method. The matrices being self-adjoint, we are only interested in the usual moments, with respect to exponents $k\in\mathbb N$, given by:

$$M_k=\frac{1}{N}\int_X Tr(Z_N^k)$$

But we can compute these moments in a straightforward way, by expanding and using the Wick formula. In the $N\to\infty$ limit the combinatorics simplifies, and we are led, up to some normalizations coming from the parameter $t>0$, to the Catalan numbers:

$$C_k=\frac{1}{k+1}\binom{2k}{k}$$

But these numbers are the even moments of the semicircle laws γ_t, as explained in Theorem 4.21 above, and we are done. □

As a second result now, due to Marchenko and Pastur, we have:

THEOREM 4.23. *Given a sequence of complex Wishart random matrices*

$$Z_N=\frac{1}{M}Y_NY_N^*\in M_N(L^\infty(X))$$

with Y_N being $N\times M$ complex Gaussian of parameter 1, we have

$$Z_N\sim\max(1-t,0)\delta_0+\frac{\sqrt{4t-(x-1-t)^2}}{2\pi x}\,dx$$

with $M=tN\to\infty$, with the limiting measure being the Marchenko-Pastur law π_t.

PROOF. This follows by using the moment method. The matrices being self-adjoint, we are only interested in the usual moments, with respect to exponents $k\in\mathbb N$, given by:

$$M_k=\frac{1}{N}\int_X Tr(Z_N^k)$$

But we can compute these moments in a straightforward way, by expanding and using the Wick formula. In the $N\to\infty$ limit the combinatorics simplifies, and we are led, up to some normalizations coming from the parameter $t>0$, to the Catalan numbers:

$$C_k=\frac{1}{k+1}\binom{2k}{k}$$

But these numbers are the moments of the Marchenko-Pastur laws π_t, as explained in Theorem 4.21 above, and we are done. □

4d. Lie groups

We discuss now some aternative interpretations of the limiting laws that we found, by using Lie groups. Let us start with some basic constructions, as follows:

THEOREM 4.24. *We have the following results:*

(1) *The rotations of $\mathbb R^N$ form the orthogonal group O_N.*
(2) *The rotations of $\mathbb C^N$ form the unitary group U_N.*
(3) $SO_N=\{U\in O_N|\det U=1\}$ *is a group too.*
(4) $SU_N=\{U\in U_N|\det U=1\}$ *is a group too.*

PROOF. This is very standard material, the idea being as follows:

(1) This is clear from definitions.

(2) This is clear from definitions, as well.

(3) Here what we have are the rotations which preserve the orientation.

(4) Here we obtain once again a group, due to $\det(AB) = \det(A)\det(B)$. □

Our claim now is that the Wigner and Marchenko-Pastur laws can be recovered as laws of characters for these groups, at certain small values of $N \in \mathbb{N}$. Let us start our study here with the following well-known result:

THEOREM 4.25. *We have the following formula,*

$$SU_2 = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix} \;\middle|\; |a|^2 + |b|^2 = 1 \right\}$$

which makes SU_2 isomorphic to the unit sphere $S^1_{\mathbb{C}} \subset \mathbb{C}^2$.

PROOF. Consider an arbitrary 2×2 matrix, written as follows:

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Assuming $\det U = 1$, the inverse is then given by:

$$U^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

On the other hand, assuming $U \in U_2$, the inverse must be the adjoint:

$$U^{-1} = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix}$$

Thus our matrix must be of the following special form:

$$U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

Since the determinant is 1, we must have $|a|^2 + |b|^2 = 1$, so we are done with one direction. As for the converse, this is clear, the matrices in the statement being unitaries, and of determinant 1, and so being elements of SU_2. Finally, we have:

$$S^1_{\mathbb{C}} = \left\{ (a,b) \in \mathbb{C}^2 \;\middle|\; |a|^2 + |b|^2 = 1 \right\}$$

Thus, the final assertion in the statement holds as well. □

We have the following reformulation of Theorem 4.25:

THEOREM 4.26. *We have the formula*

$$SU_2=\left\{\begin{pmatrix}x+iy&z+it\\-z+it&x-iy\end{pmatrix}\Big|\,x^2+y^2+z^2+t^2=1\right\}$$

which makes SU_2 isomorphic to the unit real sphere $S^3_\mathbb{R}\subset\mathbb{R}^3$.

PROOF. We recall from Theorem 4.25 above that we have:

$$SU_2=\left\{\begin{pmatrix}a&b\\-\bar{b}&\bar{a}\end{pmatrix}\Big|\,|a|^2+|b|^2=1\right\}$$

With $a=x+iy$, $b=z+it$, we are led to the conclusions in the statement. □

As a philosophical comment, the above parametrization of SU_2 is something very nice, because the parameters (x,y,z,t) range now over the sphere of space-time.

Here is now another reformulation of our main result so far, regarding SU_2, obtained by further building on the parametrization from Theorem 4.26:

THEOREM 4.27. *We have the following formula,*

$$SU_2=\left\{xc_1+yc_2+zc_3+tc_4\,\Big|\,x^2+y^2+z^2+t^2=1\right\}$$

where c_1,c_2,c_3,c_4 are the Pauli matrices, given by:

$$c_1=\begin{pmatrix}1&0\\0&1\end{pmatrix}\quad,\quad c_2=\begin{pmatrix}i&0\\0&-i\end{pmatrix}$$

$$c_3=\begin{pmatrix}0&1\\-1&0\end{pmatrix}\quad,\quad c_4=\begin{pmatrix}0&i\\i&0\end{pmatrix}$$

PROOF. We recall from Theorem 4.26 above that the group SU_2 can be parametrized by the real sphere $S^3_\mathbb{R}\subset\mathbb{R}^4$, in the following way:

$$SU_2=\left\{\begin{pmatrix}x+iy&z+it\\-z+it&x-iy\end{pmatrix}\Big|\,x^2+y^2+z^2+t^2=1\right\}$$

But this gives the formula in the statement, with the Pauli matrices c_1,c_2,c_3,c_4 being the coefficients of x,y,z,t, in this parametrization. □

The above result is often the most convenient one, when dealing with SU_2.

This is because the Pauli matrices have a number of remarkable properties, which are very useful when doing computations. These properties can be summarized as follows:

THEOREM 4.28. *The Pauli matrices multiply according to the following formulae,*

$$c_2^2 = c_3^2 = c_4^2 = -1$$

$$c_2c_3 = -c_3c_2 = c_4$$

$$c_3c_4 = -c_4c_3 = c_2$$

$$c_4c_2 = -c_2c_4 = c_3$$

they conjugate according to the following rules,

$$c_1^* = c_1,\ c_2^* = -c_2,\ c_3^* = -c_3,\ c_4^* = -c_4$$

and they form an orthonormal basis of $M_2(\mathbb{C})$, with respect to the scalar product

$$< a, b >= tr(ab^*)$$

with $tr : M_2(\mathbb{C}) \to \mathbb{C}$ being the normalized trace of 2×2 matrices, $tr = Tr/2$.

PROOF. The first two assertions, regarding the multiplication and conjugation rules for the Pauli matrices, follow from some elementary computations. As for the last assertion, this follows by using these rules. Indeed, the fact that the Pauli matrices are pairwise orthogonal follows from computations of the following type, for $i \neq j$:

$$< c_i, c_j >= tr(c_ic_j^*) = tr(\pm c_ic_j) = tr(\pm c_k) = 0$$

As for the fact that the Pauli matrices have norm 1, this follows from:

$$< c_i, c_i >= tr(c_ic_i^*) = tr(\pm c_i^2) = tr(c_1) = 1$$

Thus, we are led to the conclusion in the statement. □

As a consequence of the above results, we have:

THEOREM 4.29. *We have a double cover map, obtained via the adjoint representation,*

$$SU_2 \to SO_3$$

and this map produces the Euler-Rodrigues formula

$$U = \begin{pmatrix} x^2+y^2-z^2-t^2 & 2(yz-xt) & 2(xz+yt)\\ 2(xt+yz) & x^2+z^2-y^2-t^2 & 2(zt-xy)\\ 2(yt-xz) & 2(xy+zt) & x^2+t^2-y^2-z^2 \end{pmatrix}$$

for the generic elements of SO_3.

PROOF. We recall that the computation for SU_2 leads to the following formula for the generic matrices $V \in SU_2$, with parameters satisfying $x^2 + y^2 + z^2 + t^2 = 1$:

$$V = \begin{pmatrix} x+iy & z+it\\ -z+it & x-iy \end{pmatrix}$$

The point now is that the standard action $SU_2 \curvearrowright \mathbb{C}^2$ produces, via an adjoint action method, an action $SU_2 \curvearrowright \mathbb{R}^4$, which is trivial on one vector, and so comes from an action

$SU_2 \curvearrowright \mathbb{R}^3$. Moreover, this latter action is isometric, and orientation-preserving, and so we have a map as follows, which can be shown to be surjective:

$$SU_2 \to SO_3$$

This quotient map can be explicitely computed, and is given by the following formula, with V being as above, and U being as in the statement:

$$V \to U$$

Thus, the elements of SO_3 are given by the formula in the statement, as claimed. $\square$

Now back to probability, we can now recover our measures, as follows:

THEOREM 4.30. *The main character of SU_2 follows the following law,*

$$\gamma_1 = \frac{1}{2\pi}\sqrt{4-x^2}dx$$

and the modified main character of SO_3 follows the following law,

$$\pi_1 = \frac{1}{2\pi}\sqrt{4x^{-1}-1}\,dx$$

called respectively Wigner and Marchenko-Pastur laws of parameter 1.

PROOF. The results follow by using the above explicit description of SU_2, and SO_3:

(1) The first result follows from Theorem 4.26, by identifying SU_2 with the real sphere $S^3_\mathbb{R}$, the variable $\chi = 2Re(a)$ being semicircular. Indeed, in real notation, we have:

$$\chi\begin{pmatrix}x+iy&z+it\\-z+it&x-iy\end{pmatrix} = 2x$$

We are therefore left with computing the law of the following variable:

$$x \in C(S^3_\mathbb{R})$$

For this purpose, we use the moment method. We have:

$$\begin{aligned}\int_{S^3_\mathbb{R}} x^{2k} &= \frac{3!!(2k)!!}{(2k+3)!!}\\ &= 2\cdot\frac{3\cdot5\cdot7\ldots(2k-1)}{2\cdot4\cdot6\ldots(2k+2)}\\ &= 2\cdot\frac{(2k)!}{2^kk!2^{k+1}(k+1)!}\\ &= \frac{1}{4^k}\cdot\frac{1}{k+1}\binom{2k}{k}\\ &= \frac{C_k}{4^k}\end{aligned}$$

Thus the variable $2x \in C(S^3_{\mathbb{R}})$ has the Catalan numbers as even moments, and so its distribution is the Wigner semicircle law γ_1, as claimed.

(2) This follows by using the quotient map $SU_2 \to SO_3$, and the result for SU_2 from (1). Indeed, via some standard identifications coming from $SU_2 \to SO_3$, in the context of (1), the main character of SO_3 is given by:

$$\chi = 4Re(a)^2$$

Now recall from (1) above that we have:

$$2Re(a) \sim \gamma_1$$

On the other hand, a quick comparison between the moment formulae for the Wigner and Marchenko-Pastur laws, which are very similar, shows that we have:

$$f \sim \gamma_1 \implies f^2 \sim \pi_1$$

Thus, with $f = 2Re(a)$, we obtain the result in the statement. □

We should mention that the above notations γ_1, π_1, which remind the notations g_1, p_1 for the Gaussian and Poisson laws, are not a coincidence. We will see later on that γ_1, π_1 are the "free analogues" of g_1, p_1, appearing via a free CLT, and free PLT.

4e. Exercises

Part II

Free probability

CHAPTER 5

Free probability

5a. Freeness

The common framework for classical and free probability is "noncommutative probability". This is something very general, that we already met in connection with the random matrices, in chapters 3-4. We first recall this material. Let us start with:

Definition 5.1. *A C^*-algebra is a complex algebra A, having a norm $||.||$ making it a Banach algebra, and an involution $*$, related to the norm by the formula*

$$||aa^*|| = ||a||^2$$

which must hold for any $a \in A$.

As a basic example, the algebra $B(H)$ of by the bounded linear operators $T : H \to H$ on a Hilbert space H is a C^*-algebra, with the usual norm and involution:

$$||T|| = \sup_{||x||=1} ||Tx||$$

$$< Tx, y >=< x, T^*y >$$

More generally, any closed $*$-subalgebra of $B(H)$ is a C^*-algebra. It is possible to prove that any C^*-algebra appears in this way, as explained in chapter 3 above:

$$A \subset B(H)$$

In finite dimensions we have $H = \mathbb{C}^N$, and so the operator algebra $B(H)$ is the usual matrix algebra $M_N(\mathbb{C})$, with the usual norm and involution, namely:

$$||M|| = \sup_{||x||=1} ||Mx||$$

$$(M^*)_{ij} = \bar{M}_{ji}$$

As explained in chapter 3 above, elementary algebra shows that the finite dimensional C^*-algebras are exactly the direct sums of matrix algebras:

$$A = M_{n_1}(\mathbb{C}) \oplus \ldots \oplus M_{n_k}(\mathbb{C})$$

Summarizing, the C^*-algebra formalism is something in between the $*$-algebras, which are purely algebraic objects and whose theory basically leads nowhere, and the advanced operator algebras, such as the von Neumann algebras, which are more technical.

As yet another class of examples now, which are of particular importance for us, we have various algebras of functions $f : X \to \mathbb{C}$. The theory here is as follows:

THEOREM 5.2. *The commutative C^*-algebras are exactly the algebras of type $C(X)$, with X being a compact space, the correspondence being as follows:*

(1) *Given a compact space X, the algebra $C(X)$ of continuous functions $f : X \to \mathbb{C}$ is a commutative C^*-algebra, with norm and involution as follows:*

$$||f|| = \sup_{x \in X} |f(x)|$$

$$f^*(x) = \overline{f(x)}$$

(2) *Conversely, any commutative C^*-algebra can be written as $A = C(X)$, with its "spectrum" appearing as the space of Banach algebra characters of A:*

$$X = \left\{\chi : A \to \mathbb{C}\right\}$$

In view of this, given an arbitrary C^*-algebra A, not necessarily commutative, we agree to write $A = C(X)$, and call the abstract space X a compact quantum space.

PROOF. This is something that we know from chapter 3, the idea being as follows:

(1) First of all, the fact that $C(X)$ is indeed a Banach algebra is clear, because a uniform limit of continuous functions must be continuous. As for the formula $||ff^*|| = ||f||^2$, this is something trivial for functions, because on both sides we obtain $\sup_{x \in X} |f(x)|^2$.

(2) Given a commutative C^*-algebra A, the character space $X = \{\chi : A \to \mathbb{C}\}$ is indeed compact, and we have an evaluation morphism $ev : A \to C(X)$. The tricky point, which follows from basic spectral theory, is to prove that ev is indeed isometric. □

In order to do now probability theory, we need one more notion, as follows:

DEFINITION 5.3. *A trace, or expectation, or integration functional, on a C^*-algebra A is a linear form $tr : A \to \mathbb{C}$ having the following properties:*

(1) *tr is unital, and continuous.*
(2) *tr is positive, $a \geq 0 \implies \varphi(a) \geq 0$.*
(3) *tr has the trace property $tr(ab) = tr(ba)$.*

We call tr faithful when $a > 0 \implies \varphi(a) > 0$.

With these notions in hand, we have everything that we need for developing noncommutative probability theory. The basic notions here are as follows:

DEFINITION 5.4. *Let A be a C^*-algebra, given with a trace $tr : A \to \mathbb{C}$.*

(1) *The elements $a \in A$ are called random variables.*
(2) *The moments of such a variable are the numbers $M_k(a) = tr(a^k)$.*
(3) *The law of such a variable is the functional $\mu : P \to tr(P(a))$.*

Here $k=\circ\bullet\bullet\circ\ldots$ is by definition a colored integer, and the corresponding powers a^k are defined by the following formulae, and multiplicativity:

$$a^\emptyset=1\quad,\quad a^\circ=a\quad,\quad a^\bullet=a^*$$

As for the polynomial P, this is a noncommuting $*$-polynomial in one variable:

$$P\in\mathbb C<X,X^*>$$

Observe that the law is uniquely determined by the moments, because we have:

$$P(X)=\sum_k\lambda_kX^k\implies\mu(P)=\sum_k\lambda_kM_k(a)$$

Generally speaking, the above definition is something quite abstract, but there is no other way of doing things, at least at this level of generality. However, in certain special cases, the formalism simplifies, and we recover more familiar objects, as follows:

Theorem 5.5. *Assuming that $a\in A$ is normal, $aa^*=a^*a$, its law corresponds to a probability measure on its spectrum $\sigma(a)\subset\mathbb C$, according to the following formula:*

$$tr(P(a))=\int_{\sigma(a)}P(x)d\mu(x)$$

When the trace is faithful we have $supp(\mu)=\sigma(a)$. Also, in the particular case where the variable is self-adjoint, $a=a^*$, this law is a real probability measure.

Proof. This is something very standard, coming from the continuous functional calculus in C^*-algebras, explained in chapter 3 above. In fact, we can deduce from there that more is true, in the sense that the following formula holds, for any $f\in C(\sigma(a))$:

$$tr(f(a))=\int_{\sigma(a)}f(x)d\mu(x)$$

In addition, assuming that we are in the case $A\subset B(H)$, the measurable functional calculus tells us that the above formula holds for any $f\in L^\infty(\sigma(a))$. $\square$

We have the following notion, generalizing the one in Definition 1.4:

Definition 5.6. *Two subalgebras $B,C\subset A$ are called independent when the following condition is satisfied, for any $b\in B$ and $c\in C$:*

$$tr(bc)=tr(b)tr(c)$$

Equivalently, the following condition must be satisfied, for any $b\in B$ and $c\in C$:

$$tr(b)=tr(c)=0\implies tr(bc)=0$$

Also, two variables $b,c\in A$ are called independent when the algebras that they generate,

$$B=<b>\quad,\quad C=<c>$$

are independent inside A, in the above sense.

Observe that the above two conditions are indeed equivalent. In one sense this is clear, and in the other sense, with $a' = a - tr(a)$, this follows from:

$$\begin{aligned} tr(bc) &= tr[(b' + tr(b))(c' + tr(c))] \\ &= tr(b'c') + t(b')tr(c) + tr(b)tr(c') + tr(b)tr(c) \\ &= tr(b'c') + tr(b)tr(c) \\ &= tr(b)tr(c) \end{aligned}$$

The other remark is that the above notion generalizes indeed the usual notion of independence, from the classical case, the result here being as follows:

THEOREM 5.7. *Given two compact measured spaces Y, Z, the algebras*

$$C(Y) \subset C(Y \times Z)$$

$$C(Z) \subset C(Y \times Z)$$

are independent in the above sense, and a converse of this fact holds too.

PROOF. We have two assertions here, the idea being as follows:

(1) First of all, given two arbitrary compact spaces Y, Z, we have embeddings of algebras as in the statement, defined by the following formulae:

$$f \to [(y, z) \to f(y)]$$

$$g \to [(y, z) \to g(z)]$$

In the measured space case now, the Fubini theorems tells us that:

$$\int_{Y \times Z} f(y)g(z) = \int_Y f(y) \int_Z g(z)$$

Thus, the algebras $C(Y), C(Z)$ are independent in the sense of Definition 5.6.

(2) Conversely now, assume that $B, C \subset A$ are independent, with A being commutative. Let us write our algebras as follows, with X, Y, Z being certain compact spaces:

$$A = C(X) \quad , \quad B = C(Y) \quad , \quad C = C(Z)$$

In this picture, the inclusions $B, C \subset A$ must come from quotient maps, as follows:

$$p : Z \to X \quad , \quad q : Z \to Y$$

Regarding now the independence condition from Definition 5.6, in the above picture, this tells us that the folowing equality must happen:

$$\int_X f(p(x))g(q(x)) = \int_X f(p(x)) \int_X g(q(x))$$

Thus we are in a Fubini type situation, and we obtain from this:

$$Y \times Z \subset X$$

Thus, the independence of $B, C \subset A$ appears as in (1) above. $\square$

It is possible to develop some theory here, but all this is ultimately not very interesting. As a much more interesting notion now, we have the freeness:

Definition 5.8. *Two subalgebras $B, C \subset A$ are called free when the following condition is satisfied, for any $b_i \in B$ and $c_i \in C$:*

$$tr(b_i) = tr(c_i) = 0 \implies tr(b_1 c_1 b_2 c_2 \ldots) = 0$$

Also, two variables $b, c \in A$ are called free when the algebras that they generate,

$$B = < b > \quad , \quad C = < c >$$

are free inside A, in the above sense.

In short, freeness appears by definition as a kind of "free analogue" of independence, taking into account the fact that the variables do not necessarily commute.

As a first observation, of theoretical nature, there is actually a certain lack of symmetry between Definition 5.6 and Definition 5.8, because in contrast to the former, the latter does not include an explicit formula for the quantities of the following type:

$$tr(b_1 c_1 b_2 c_2 \ldots)$$

However, this is not an issue, and is simply due to the fact that the formula in the free case is something more complicated, the result being as follows:

Proposition 5.9. *Assuming that $B, C \subset A$ are free, the restriction of tr to $< B, C >$ can be computed in terms of the restrictions of tr to B, C. To be more precise,*

$$tr(b_1 c_1 b_2 c_2 \ldots) = P\Big(\{tr(b_{i_1} b_{i_2} \ldots)\}_i, \{tr(c_{j_1} c_{j_2} \ldots)\}_j\Big)$$

where P is certain polynomial in several variables, depending on the length of the word $b_1 c_1 b_2 c_2 \ldots$, and having as variables the traces of products of type

$$b_{i_1} b_{i_2} \ldots$$

$$c_{j_1} c_{j_2} \ldots$$

with the indices being chosen increasing, $i_1 < i_2 < \ldots$ and $j_1 < j_2 < \ldots$

Proof. This is something quite theoretical, so let us begin with an example. Our claim is that if b, c are free then, exactly as in the case where we have independence:

$$tr(bc) = tr(b)tr(c)$$

Indeed, let us go back to the computation performed after Definition 5.6, which was as follows, with the convention $a' = a - tr(a)$:

$$\begin{aligned} tr(bc) &= tr[(b' + tr(b))(c' + tr(c))] \\ &= tr(b'c') + t(b')tr(c) + tr(b)tr(c') + tr(b)tr(c) \\ &= tr(b'c') + tr(b)tr(c) \\ &= tr(b)tr(c) \end{aligned}$$

Our claim is that this computation perfectly works under the sole freeness assumption. Indeed, the only non-trivial equality is the last one, which follows from:

$$tr(b')=tr(c')=0\implies tr(b'c')=0$$

In general now, the situation is of course more complicated, but the same trick applies. To be more precise, we can start our computation as follows:

$$\begin{aligned}
&tr(b_1c_1b_2c_2\ldots)\\
=\ &tr\big[(b_1'+tr(b_1))(c_1'+tr(c_1))(b_2'+tr(b_2))(c_2'+tr(c_2))\ldots\ldots\big]\\
=\ &tr(b_1'c_1'b_2'c_2'\ldots)+\text{other terms}\\
=\ &\text{other terms}
\end{aligned}$$

Observe that we have used here the freeness condition, in the following form:

$$tr(b_i')=tr(c_i')=0\implies tr(b_1'c_1'b_2'c_2'\ldots)=0$$

Now regarding the "other terms", those which are left, each of them will consist of a product of traces of type $tr(b_i)$ and $tr(c_i)$, and then a trace of a product still remaining to be computed, which is of the following form, with $\beta_i\in B$ and $\gamma_i\in C$:

$$tr(\beta_1\gamma_1\beta_2\gamma_2\ldots)$$

To be more precise, the variables $\beta_i\in B$ appear as ordered products of those $b_i\in B$ not getting into individual traces $tr(b_i)$, and the variables $\gamma_i\in C$ appear as ordered products of those $c_i\in C$ not getting into individual traces $tr(c_i)$. Now since the length of each such alternating product $\beta_1\gamma_1\beta_2\gamma_2\ldots$ is smaller than the length of the original alternating product $b_1c_1b_2c_2\ldots$, we are led into of recurrence, and this gives the result. □

5b. Free products

Let us discuss now some models for independence and freeness. We first have the following result, which clarifies the analogy between independence and freeness:

THEOREM 5.10. *Given two algebras (B,tr) and (C,tr), the following hold:*

(1) *B,C are independent inside their tensor product $B\otimes C$, endowed with its canonical tensor product trace, given on basic tensors by $tr(b\otimes c)=tr(b)tr(c)$.*
(2) *B,C are free inside their free product $B*C$, endowed with its canonical free product trace, given by the formulae in Proposition 5.9.*

PROOF. Both the assertions are clear from definitions, as follows:

(1) This is clear with either of the definitions of the independence, from Definition 5.6 above, because we have by construction of the trace:

$$\begin{aligned}
tr(bc)&=tr[(b\otimes 1)(1\otimes c)]\\
&=tr(b\otimes c)\\
&=tr(b)tr(c)
\end{aligned}$$

Observe that there is a relation here with Theorem 5.7 as well, due to the following formula for compact spaces, with $\otimes$ being a topological tensor product:

$$C(Y \times Z) = C(Y) \otimes C(Z)$$

To be more precise, the present statement generalizes the first assertion in Theorem 5.7, and the second assertion tells us that this generalization is more or less the same thing as the original statement. All this comes of course from basic measure theory.

(2) This is clear from definitions, the only point being that of showing that the notion of freeness, or the recurrence formulae in Proposition 5.9, can be used in order to construct a canonical free product trace, on the free product of the two algebras involved:

$$tr : B * C \to \mathbb{C}$$

But this can be checked for instance by using a GNS construction. Indeed, consider the GNS constructions for the algebras (B, tr) and (C, tr):

$$B \to B(l^2(B))$$

$$C \to B(l^2(C))$$

By taking the free product of these representations, we obtain a representation as follows, with the $*$ symbol on the right being a free product of pointed Hilbert spaces:

$$B * C \to B(l^2(B) * l^2(C))$$

Now by composing with the linear form $T \to < T\xi, \xi >$, where $\xi = 1_B = 1_C$ is the common distinguished vector of $l^2(B)$ and $l^2(C)$, we obtain a linear form, as follows:

$$tr : B * C \to \mathbb{C}$$

It is routine then to check that tr is indeed a trace, and this is the "canonical free product trace" from the statement. Then, an elementary computation shows that B, C are indeed free inside $B * C$, with respect to this trace, and this finishes the proof. □

5c. Free convolution

As an concrete application of the above results, we have:

THEOREM 5.11. *We have a free convolution operation* $\boxplus$ *for the distributions*

$$\mu : \mathbb{C} < X, X^* > \to \mathbb{C}$$

which is well-defined by the following formula, with b, c *taken to be free:*

$$\mu_b \boxplus \mu_c = \mu_{b+c}$$

This restricts to an operation, still denoted $\boxplus$*, on the real probability measures.*

Proof. We have several verifications to be performed here, as follows:

(1) We first have to check that given two variables b,c which live respectively in certain C^*-algebras B,C, we can recover inside some C^*-algebra A, with exactly the same distributions μ_b,μ_c, as to be able to sum them and then talk about μ_{b+c}. But this comes from Theorem 5.10, because we can set $A=B*C$, as explained there.

(2) The other verification which is needed is that of the fact that if b,c are free, then the distribution μ_{b+c} depends only on the distributions μ_b,μ_c. But for this purpose, we can use the general formula from Proposition 5.9, namely:

$$tr(b_1c_1b_2c_2\ldots)=P\Big(\{tr(b_{i_1}b_{i_2}\ldots)\}_i,\{tr(c_{j_1}c_{j_2}\ldots)\}_j\Big)$$

Here P is certain polynomial, depending on the length of $b_1c_1b_2c_2\ldots$, having as variables the traces of products $b_{i_1}b_{i_2}\ldots$ and $c_{j_1}c_{j_2}\ldots$, with $i_1<i_2<\ldots$ and $j_1<j_2<\ldots$

Now by plugging in arbitrary powers of b,c as variables b_i,c_j, we obtain a family of formulae of the following type, with Q being certain polyomials:

$$tr(b^{k_1}c^{l_1}b^{k_2}c^{l_2}\ldots)=P\Big(\{tr(b^k)\}_k,\{tr(c^l)\}_l\Big)$$

Thus the moments of $b+c$ depend only on the moments of b,c, with of course colored exponents in all this, according to our moment conventions, and this gives the result.

(3) Finally, in what regards the last assertion, regarding the real measures, this is clear from the fact that if b,c are self-adjoint, then so is their sum $b+c$. □

Along the same lines, but with some technical subtleties this time, we can talk as well about multiplicative free convolution, as follows:

Theorem 5.12. *We have a free convolution operation $\boxtimes$ for the distributions*

$$\mu:\mathbb C<X,X^*>\to\mathbb C$$

which is well-defined by the following formula, with b,c taken to be free:

$$\mu_b\boxtimes\mu_c=\mu_{bc}$$

In the case of the self-adjoint variables, we can equally set

$$\mu_b\boxtimes\mu_c=\mu_{\sqrt{b}c\sqrt{b}}$$

and so we have an operation, still denoted $\boxtimes$, on the real probability measures.

Proof. We have two statements here, the idea being as follows:

(1) The verifications for the fact that $\boxtimes$ is indeed well-defined at the general distribution level are identical to those done before for $\boxplus$, with the result basically coming from the formula in Proposition 5.9, and with Theorem 5.10 invoked as well, in order to say that we have a model, and so we can indeed use this formula.

(2) Regarding now the last assertion, regarding the real measures, this was something trivial for $\boxplus$, but is something trickier now for $\boxtimes$, because if we take b, c to be self-adjoint, thier product bc will normally not be self-adjoint, and definitely it will be not if we want b, c to be free, and so the formula $\mu_b \boxtimes \mu_c = \mu_{bc}$ will apparently makes us exit the world of real probability measures. However, this is not exactly the case. Let us set:

$$a = \sqrt{b}c\sqrt{c}$$

This variable is then self-adjoint, and its moments are given by:

$$\begin{aligned} tr(a^k) &= tr[(\sqrt{b}c\sqrt{b})^k] \\ &= tr[\sqrt{b}cb \ldots bc\sqrt{b}] \\ &= tr[\sqrt{b} \cdot \sqrt{b}cb \ldots bc] \\ &= tr[(bc)^k] \end{aligned}$$

Thus, we are led to the conclusion in the statement. □

As a remark here, observe that we have used in the proof of (2) above, and actually for the first time since talking about freeness, the trace property of the trace:

$$tr(ab) = tr(ba)$$

This is quite interesting, philosophically speaking, because in the operator algebra world there are many interesting examples of subalgebras $A \subset B(H)$ coming with natural linear forms $\varphi : A \to \mathbb{C}$ which are continuous and positive, but which are not traces. It is possible to do a bit of free probability on such algebras, but not much.

5d. Group algebras

We would like to have linearization results for $\boxplus$ and $\boxtimes$, in the spirit of the known results for $*$ and $\times$. We will do this slowly, in several steps. As a first objective, we would like to convert our one and only modelling result so far, namely Theorem 5.10, which is a rather abstract result, into something more concrete. And the answer here comes from the group algebras and their versions, that we will show in what follows to model in a very nice and effective way both the independence and the freeness. let us start with:

Theorem 5.13. *Let Γ be a discrete group, and consider the complex group algebra $\mathbb{C}[\Gamma]$, with involution given by the fact that all group elements are unitaries:*

$$g^* = g^{-1} \quad , \quad \forall g \in \Gamma$$

The maximal C^*-seminorm on $\mathbb{C}[\Gamma]$ is then a C^*-norm, and the closure of $\mathbb{C}[\Gamma]$ with respect to this norm is a C^*-algebra, denoted $C^*(\Gamma)$. Moreover,

$$tr(g) = \delta_{g1}$$

defines a positive unital trace $tr : C^*(\Gamma) \to \mathbb{C}$, which is faithful on $\mathbb{C}[\Gamma]$.

Proof. We have two assertions to be proved, the idea being as follows:

(1) In order to prove the first assertion, regarding the maximal seminorm which is a norm, we must find a $*$-algebra embedding as follows, with H being a Hilbert space:

$$\mathbb{C}[\Gamma]\subset B(H)$$

For this purpose, consider the Hilbert space $H=l^2(\Gamma)$, having $\{h\}_{h\in\Gamma}$ as orthonormal basis. Our claim is that we have an embedding, as follows:

$$\pi:\mathbb{C}[\Gamma]\subset B(H)$$

$$\pi(g)(h)=gh$$

Indeed, since $\pi(g)$ maps the basis $\{h\}_{h\in\Gamma}$ into itself, this operator is well-defined, bounded, and is an isometry. It is also clear from the formula $\pi(g)(h)=gh$ that $g\to\pi(g)$ is a morphism of algebras, and since this morphism maps the unitaries $g\in\Gamma$ into isometries, this is a morphism of $*$-algebras. Finally, the faithfulness of π is clear.

(2) Regarding now the second assertion, we can use here once again the above construction. Indeed, we can define a linear form on the image of $C^*(\Gamma)$, as follows:

$$tr(T)=<T\delta_1,\delta_1>$$

This functional is then positive, and is easily seen to be a trace. Moreover, on the group elements $g\in\Gamma$, this functional is given by the following formula:

$$tr(g)=\delta_{g1}$$

Thus, it remains to show that tr is faithful on $\mathbb{C}[\Gamma]$. But this simply follows from the fact that tr is faithful on the image of $C^*(\Gamma)$, which contains $\mathbb{C}[\Gamma]$. □

There are many things that can be said about the group algebras, and we will be back to this later, when doing quantum groups. For the moment, let us just record the following result, which fully clarifies the situation, in the abelian group case:

Proposition 5.14. *Given a discrete abelian group Γ, we have an isomorphism*

$$C^*(\Gamma)\simeq C(G)$$

where $G=\widehat{\Gamma}$ is its Pontrjagin dual, formed by the characters $\chi:\Gamma\to\mathbb{T}$. Moreover,

$$tr(g)=\delta_{g1}$$

corresponds in this way to the Haar integration over G.

Proof. We have two assertions to be proved, the idea being as follows:

(1) Since Γ is abelian, $A=C^*(\Gamma)$ is commutative, so by the Gelfand theorem we have $A=C(X)$. The spectrum $X=Spec(A)$, consisting of the characters $\chi:C^*(\Gamma)\to\mathbb{C}$, can be identified with the Pontrjagin dual $G=\widehat{\Gamma}$, and this gives the result.

(2) Regarding now the last assertion, we must prove here that we have:

$$tr(f) = \int_G f(x)dx$$

But this is clear via the above identifications, for instance because the linear form $tr(g) = \delta_{g1}$, when viewed as a functional on $C(G)$, is left and right invariant. □

Getting back now to our questions, we can now formulate a general modelling result for independence and freeness, providing us with large classes of examples, as follows:

THEOREM 5.15. *We have the following results, valid for group algebras:*

(1) $C^*(\Gamma), C^*(\Lambda)$ *are independent inside* $C^*(\Gamma \times \Lambda)$.

(2) $C^*(\Gamma), C^*(\Lambda)$ *are free inside* $C^*(\Gamma * \Lambda)$.

PROOF. In order to prove these results, we have two possible methods:

(1) We can either use the general results in Theorem 5.10 above, along with the following two isomorphisms, which are both standard:

$$C^*(\Gamma \times \Lambda) = C^*(\Lambda) \otimes C^*(\Gamma)$$

$$C^*(\Gamma * \Lambda) = C^*(\Lambda) * C^*(\Gamma)$$

(2) Or, we can prove this directly, by using the fact that each group algebra is spanned by the corresponding group elements. Indeed, this shows that it is enough to check the independence and freeness formulae on group elements, which is in turn trivial. □

5e. Exercises

CHAPTER 6

Limiting theorems

6a. Fock spaces

We have seen so far the foundations of free probability, in analogy with those of classical probability, regarded with a functional analysis touch. The idea now is that with a bit of luck, the basic theory from the classical case, namely the Fourier transform, and then the CLT, should have free extensions. Let us being our discussion with:

Definition 6.1. *The real probability measures are subject to operations $*$ and $\boxplus$, called classical and free convolution, given by the formulae*

$$\mu_a * \mu_b = \mu_{a+b}$$

$$\mu_\alpha \boxplus \mu_\beta = \mu_{\alpha+\beta}$$

with a, b being independent, and α, β being free, and all variables being self-adjoint.

The problem now is that of linearizing these operations $*$ and $\boxplus$. In what regards $*$, we know from chapter 1 that this operation is linearized by the logarithm $\log F$ of the Fourier transform, which in the present setting, where $\mathbb{E} = tr$, is given by:

$$F_a(x) = tr(e^{ixa})$$

In order to find a similar result for $\boxplus$, we need some good models for pairs of free random variables (b, c). This is a priori not a problem, because once we have $b \in B$ and $c \in C$, we can form the free product $B * C$, which contains b, c as free variables.

However, the initial choice, that of the variables $b \in B$ and $c \in C$ modelling some given laws $\mu, \nu \in \mathcal{P}(\mathbb{R})$, matters a lot. Indeed, any kind of abstract choice here would lead us into an abstract algebra $B * C$, and so into the combinatorics from Proposition 5.9 and its proof, that cannot be solved with bare hands, and that we want to avoid.

In short, we must be tricky, at least in what concerns the beginning of our computation. The idea will be that of temporarily lifting the self-adjointness assumption on our variables b, c, and looking instead for arbitrary random variables β, γ, not necessarily self-adjoint, modelling in integer moments our given laws $\mu, \nu \in \mathcal{P}(\mathbb{R})$, as follows:

$$tr(\beta^k) = M_k(\mu),\ \forall k \in \mathbb{N}$$

$$tr(\gamma^k) = M_k(\nu),\ \forall k \in \mathbb{N}$$

To be more precise here, assuming that β, γ are indeed not self-adjoint, the above formulae are not the general formulae for β, γ, simply because these latter formulae involve colored integers $k = \circ \bullet \bullet \circ \ldots$ as exponents. Thus, in the context of the above formulae, μ, ν are not the distributions of β, γ, but just some "pieces" of these distributions.

Now with this idea in mind, due to Voiculescu and quite tricky, the solution to the law modelling problem comes in a quite straightforward way, involving the good old Hilbert space $H = l^2(\mathbb{N})$ and the good old shift operator $S \in B(H)$, as follows:

Theorem 6.2. *Consider the shift operator on the space $H = l^2(\mathbb{N})$, given by:*

$$S(e_i) = e_{i+1}$$

The variables of the following type, with $f \in \mathbb{C}[X]$ being a polynomial,

$$S^* + f(S)$$

model then in moments, up to finite order, all the distributions $\mu : \mathbb{C}[X] \to \mathbb{C}$.

Proof. We have already met the shift S in chapter 3 above, as the simplest example of an isometry which is not a unitary, $S^*S = 1, SS^* = 1$, with this coming from:

$$S^*(e_i) = \begin{cases} e_{i-1} & (i > 0) \\ 0 & (i = 0) \end{cases}$$

Consider now a variable as in the statement, namely:

$$T = S^* + a_0 + a_1 S + a_2 S^2 + \ldots + a_n S^n$$

We have then $tr(T) = a_0$, then $tr(T^2)$ will involve a_1, then $tr(T^3)$ will involve a_2, and so on. Thus, we are led to a certain recurrence, that we will not attempt to solve now, with bare hands, but which definitely gives the conclusion in the statement. $\square$

Before getting further, with taking free products of such models, let us work out a very basic example, which is something fundamental, that we will need in what follows:

Proposition 6.3. *In the context of the above correspondence, the variable*

$$T = S + S^*$$

follows the Wigner semicircle law on $[-2, 2]$, given by:

$$\gamma_1 = \frac{1}{2\pi}\sqrt{4 - x^2}dx$$

Proof. In order to compute the law of variable T in the statement, we use the moment method. The moments of this variable are the following numbers:

$$\begin{aligned} M_k &= tr(T^k) \\ &= tr((S + S^*)^k) \\ &= \#(1 \in (S + S^*)^k) \end{aligned}$$

Now since S shifts to the right on $\mathbb{N}$, and S^* shifts to the left, while remaining positive, we are left with counting the length k paths on $\mathbb{N}$ starting and ending at 0. Since there are no such paths when $k = 2l+1$ is odd, the odd moments vanish:

$$M_{2l+1} = 0$$

In the case where $k = 2l$ is even, such paths on $\mathbb{N}$ are best represented as paths in the upper half-plane, starting at 0, and then going at each step NE or SE, depending on whether the original path on $\mathbb{N}$ goes at right or left, and finally ending at $k \in \mathbb{N}$. Indeed, with this picture we are led to the following formula for the number of such paths:

$$M_{2l+2} = \sum_s M_{2s} M_{2l-s}$$

But this is exactly the recurrence formula for the Catalan numbers, and so:

$$M_{2l} = \frac{1}{l+1}\binom{2l}{l}$$

Summarizing, the odd moments of T vanish, and the even moments are the Catalan numbers. But these numbers being exactly the moments of the Wigner semicircle law γ_1, as explained in chapter 3 above, we are led to the conclusion in the statement. □

Getting back now to our linearization program for $\boxplus$, the next step is that of taking a free product of the model found in Theorem 6.2 with itself. There are two approaches here, one being a bit abstract, and the other one being more concrete. The abstract approach, which is quite nice, making a link with our main modelling result so far from chapter 5, involving group algebras, is as follows:

Proposition 6.4. *We can talk about semigroup algebras $C^*(\Gamma) \subset B(l^2(\Gamma))$, exactly as we did for the group algebras, and at the level of examples:*

1. *With $\Gamma = \mathbb{N}$ we recover the shift algebra $A = < S >$ on $H = l^2(\mathbb{N})$.*
2. *With $\Gamma = \mathbb{N} * \mathbb{N}$, we obtain the algebra $A = < S_1, S_2 >$ on $H = l^2(\mathbb{N} * \mathbb{N})$.*

Proof. We can talk indeed about semigroup algebras $C^*(\Gamma) \subset B(l^2(\Gamma))$, exactly as we did for the group algebras, the only difference coming from the fact that the semigroup elements $g \in \Gamma$ will now correspond to isometries, which are not necessarily unitaries. Now this construction in hand, both the assertions are clear, as follows:

(1) With $\Gamma = \mathbb{N}$ we recover indeed the shift algebra $A = < S >$ on the Hilbert space $H = l^2(\mathbb{N})$, the shift S itself being the isometry associated to the element $1 \in \mathbb{N}$.

(2) With $\Gamma = \mathbb{N} * \mathbb{N}$ we recover the double shift algebra $A = < S_1, S_2 >$ on the Hilbert space $H = l^2(\mathbb{N} * \mathbb{N})$, the two shifts S_1, S_2 themselves being the isometries associated to two copies of the element $1 \in \mathbb{N}$, one for each of the two copies of $\mathbb{N}$ which are present. □

In what follows we will rather use an equivalent, second approach to our problem, which is exactly the same thing, but formulated in a less abstract way, as follows:

PROPOSITION 6.5. *We can talk about the algebra of creation operators*

$$S_x : v \to x \otimes v$$

on the free Fock space associated to a real Hilbert space H, given by

$$F(H) = \mathbb{C}\Omega \oplus H \oplus H^{\otimes 2} \oplus \ldots$$

and at the level of examples, we have:

(1) *With $H = \mathbb{C}$ we recover the shift algebra $A =< S >$ on $H = l^2(\mathbb{N})$.*
(2) *With $H = \mathbb{C}^2$, we obtain the algebra $A =< S_1, S_2 >$ on $H = l^2(\mathbb{N} * \mathbb{N})$.*

PROOF. We can talk indeed about the algebra $A(H)$ of creation operators on the free Fock space $F(H)$ associated to a real Hilbert space H, with the remark that, in terms of the abstract semigroup notions from Proposition 6.4 above, we have:

$$A(\mathbb{C}^k) = C^*(\mathbb{N}^{*k})$$

$$F(\mathbb{C}^k) = l^2(\mathbb{N}^{*k})$$

As for the assertions (1,2) in the statement, these are both clear, either directly, or by passing via (1,2) from Proposition 6.4, which were clear as well. □

The advantage with this latter model comes from the following result, which has a very simple formulation, without linear combinations or anything:

PROPOSITION 6.6. *Given a real Hilbert space H, and two orthogonal vectors $x, y \in H$,*

$$x \perp y$$

the corresponding creation operators S_x and S_y are free with respect to

$$tr(T) =< T\Omega, \Omega >$$

called trace associated to the vacuum vector.

PROOF. In standard tensor notation for the elements of the free Fock space $F(H)$, the formula of a creation operator associated to a vector $x \in H$ is as follows:

$$S_x(y_1 \otimes \ldots \otimes y_n) = x \otimes y_1 \otimes \ldots \otimes y_n$$

As for the formula of the adjoint of this creation operator, called annihilation operator associated to $x \in H$, this is as follows:

$$S_x^*(y_1 \otimes \ldots \otimes y_n) =< x, y_1 > \otimes y_2 \otimes \ldots \otimes y_n$$

We obtain from this the following formula, valid for any two vectors $x, y \in H$:

$$S_x^* S_y =< x, y > id$$

With these formulae in hand, the result follows by doing some elementary computations, in the spirit of those done for the group algebras, from chapter 5 above. □

6b. R-transform

With this technology in hand, let us go back to our linearization program for $\boxplus$. We know from Theorem 6.2 how to model the individual distributions $\mu \in \mathcal{P}(\mathbb{R})$, and by combining this with Proposition 6.5 and Proposition 6.6, we therefore know how to freely model pairs of distributions $\mu, \nu \in \mathcal{P}(\mathbb{R})$, as required by the convolution problem.

We are therefore left with doing the sum in the model, and then computing its distribution. And the point here is that the following remarkable fact happens:

THEOREM 6.7. *Given two polynomials $f, g \in \mathbb{C}[X]$, consider the variables*

$$R^* + f(R) \quad , \quad S^* + g(S)$$

where R, S are two creation operators, or shifts, associated to a pair of orthogonal norm 1 vectors. These variables are then free, and their sum has the same law as

$$T^* + (f+g)(T)$$

with T being the usual shift on $l^2(\mathbb{N})$.

PROOF. We have two assertions here, the idea being as follows:

(1) The freeness assertion comes from the general freeness result from Proposition 6.6, via the various identifications coming from the previous results.

(2) Regarding now the second assertion, the idea is that this comes from a 45° rotation trick. Let us write indeed the two variables in the statement as follows:

$$X = R^* + a_0 + a_1 R + a_2 R^2 + \ldots$$

$$Y = S^* + b_0 + b_1 S + a_2 S^2 + \ldots$$

Now let us perform the following 45° base change, on the real span of the vectors $r, s \in H$ producing our two shifts R, S:

$$t = \frac{r+s}{\sqrt{2}} \quad , \quad u = \frac{r-s}{\sqrt{2}}$$

The new shifts, associated to these vectors $t, u \in H$, are then given by:

$$T = \frac{R+S}{\sqrt{2}} \quad , \quad U = \frac{R-S}{\sqrt{2}}$$

By using now these new shifts, which are free as well according to Proposition 6.6, we obtain the following equality of distributions:

$$\begin{aligned}
X+Y &= R^*+S^*+\sum_k a_kR^k+b_kS^k\\
&= \sqrt{2}T^*+\sum_k a_k\left(\frac{T+U}{\sqrt{2}}\right)^k+b_k\left(\frac{T-U}{\sqrt{2}}\right)^k\\
&\sim \sqrt{2}T^*+\sum_k a_k\left(\frac{T}{\sqrt{2}}\right)^k+b_k\left(\frac{T}{\sqrt{2}}\right)^k\\
&\sim T^*+\sum_k a_kT^k+b_kT^k
\end{aligned}$$

To be more precise, here in the last two lines we have used the freeness property of T,U in order to cut U from the computation, as it cannot bring anything, and then we did a basic rescaling at the end. Thus, we are led to the conclusion in the statement. □

As a conclusion to all this, the operation $\mu\to f$ from Theorem 6.2 above linearizes the free convolution operation $\boxplus$. In order to reach to something concrete, we are therefore left with a computation inside $C^*(\mathbb{N})$, which is elementary, and whose conclusion is that $R_\mu=f$ can be recaptured from μ via the Cauchy transform G_μ.

The precise result here, due to Voiculescu [90], is as follows:

THEOREM 6.8. *Given a real probability measure μ, define its R-transform as follows:*

$$G_\mu(\xi)=\int_{\mathbb{R}}\frac{d\mu(t)}{\xi-t}\implies G_\mu\left(R_\mu(\xi)+\frac{1}{\xi}\right)=\xi$$

The free convolution operation is then linearized by this R-transform.

PROOF. This can be done by using the above results, in several steps, as follows:

(1) According to Theorem 7.7, the operation $\mu\to f$ from Theorem 6.2 above linearizes the free convolution operation $\boxplus$. We are therefore left with a computation inside $C^*(\mathbb{N})$. To be more precise, consider a variable as in Theorem 6.2 above:

$$X=S^*+f(X)$$

In order to establish the result, we must prove that the R-transform of X, constructed according to the procedure in the statement, is the function f itself.

(2) In order to do so, fix $|z|<1$ in the complex plane, and let us set:

$$w_z=\delta_0+\sum_{k=1}^\infty z_k\delta_k$$

The shift and its adjoint act then as follows, on this vector:

$$Sw_z = z^{-1}(w_z - \delta_0) \quad , \quad S^*w_z = zw_z$$

It follows that the adjoint of our operator X acts as follows on this vector:

$$\begin{aligned} X^*w_z &= (S + f(S^*))w_z \\ &= z^{-1}(w_z - \delta_0) + f(z)w_z \\ &= (z^{-1} + f(z))w_z - z^{-1}\delta_0 \end{aligned}$$

Now observe that this formula can be written as follows:

$$z^{-1}\delta_0 = (z^{-1} + f(z) - X^*)w_z$$

The point now is that when $|z|$ is small, the operator appearing on the right is invertible. Thus, we can rewrite this formula as follows:

$$(z^{-1} + f(z) - X^*)^{-1}\delta_0 = zw_z$$

Now by applying the trace, we are led to the following formula:

$$\begin{aligned} tr\left[(z^{-1} + f(z) - X^*)^{-1}\right] &= \left\langle (z^{-1} + f(z) - X^*)^{-1}\delta_0, \delta_0 \right\rangle \\ &= < zw_z, \delta_0 > \\ &= z \end{aligned}$$

(3) With this formula in hand, we can finish. Let us apply indeed the complex function procedure in the statement to the real probability measure μ modelled by X. The Cauchy transform G_μ, which is a function having real coefficients, is given by:

$$\begin{aligned} G_\mu(\xi) &= tr((\xi - X)^{-1}) \\ &= \overline{tr\Big((\bar{\xi} - X^*)^{-1}\Big)} \\ &= tr((\xi - X^*)^{-1}) \end{aligned}$$

Now observe that, with the choice $\xi = z^{-1} + f(z)$ for our complex variable, the trace formula found in (2) above tells us precisely that we have:

$$G_\mu\Big(z^{-1} + f(z)\Big) = z$$

Thus, by definition of the R-transform, we have the following formula:

$$R_\mu(z) = f(z)$$

But this finishes the proof, as explained in step (1) above. □

Summarizing, the situation in free probability is quite similar to the one in classical probability, the product spaces needed for the basic properties of the Fourier transform being replaced by something "noncommutative", namely the free Fock space models.

6c. Free CLT

With the above linearization technology in hand, we can now establish the following free analogue of the CLT, due to Voiculescu [90]:

THEOREM 6.9 (Free CLT). *Given self-adjoint variables $x_1, x_2, x_3, \ldots$ which are f.i.d., centered, with variance $t > 0$, we have, with $n \to \infty$, in moments,*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n x_i \sim \gamma_t$$

where γ_t is the Wigner semicircle law of parameter t, having density:

$$\gamma_t = \frac{1}{2\pi t}\sqrt{4t^2 - x^2}dx$$

PROOF. We follow the same idea as in the proof of the CLT, from section 1:

(1) At $t = 1$, the R-transform of the variable in the statement on the left can be computed by using the linearization property from Theorem 6.8, and is given by:

$$R(\xi) = nR_x\left(\frac{\xi}{\sqrt{n}}\right) \simeq \xi$$

(2) Regarding now the right term, also at $t = 1$, our first claim here is that the Cauchy transform of the Wigner law γ_1 satisfies the following equation:

$$G_{\gamma_1}\left(\xi + \frac{1}{\xi}\right) = \xi$$

Indeed, we already know from chapter 4 above that the moments of the Wigner law γ_1 are the Catalan numbers:

$$\frac{1}{2\pi}\int_0^4 \sqrt{4 - x^2}x^{2k}dx = C_k$$

Consider now the generating series of the Catalan numbers:

$$f(z) = \sum_{k \geq 0} C_k z^k$$

We also know that f satisfies the following degree 2 equation:

$$zf^2 - f + 1 = 0$$

But this gives the above formula for G_{γ_1}, via a few manipulations.

(3) We conclude from the formula in (2) above that the R-transform of the Wigner semicircle law γ_1 is given by the following formula:

$$R_{\gamma_1}(\xi) = \xi$$

Note that this follows in fact as well from the following formula, coming from Proposition 6.3, and from the technical details of the R-transform:

$$S+S^*\sim\gamma_1$$

Thus, the laws in the statement have the same R-transforms, and so they are equal.

(4) Summarizing, we have proved the free CLT at $t=1$. The passage to the general case, $t>0$, is routine, by some standard dilation computations. □

Regarding the limiting measures γ_t, we have seen most of the theory, namely density and moments, in chapter 4 above, with all this partly coming from SU_2, at $t=1$. One thing which was missing, however, was that of understanding how γ_t exactly appears, out of γ_1. With free probability theory, we can now solve this question, as follows:

THEOREM 6.10. *The Wigner semicircle laws have the property*

$$\gamma_s\boxplus\gamma_t=\gamma_{s+t}$$

so they form a 1*-parameter semigroup with respect to free convolution.*

PROOF. This follows either from Theorem 6.9, or from Theorem 6.8, by using the fact that the R-transform of γ_t, which is given by $R_{\gamma_t}(\xi)=t\xi$, is linear in t. □

As a conclusion to what we have so far, we have:

THEOREM 6.11. *The Gaussian laws g_t and the Wigner laws γ_t, given by*

$$g_t=\frac{1}{\sqrt{2\pi t}}e^{-x^2/2t}dx$$

$$\gamma_t=\frac{1}{2\pi t}\sqrt{4t^2-x^2}dx$$

have the following properties:

(1) *They appear via the CLT, and the free CLT.*
(2) *They form semigroups with respect to $*$ and $\boxplus$.*
(3) *Their transforms are $\log F_{g_t}(x)=-tx^2/2$, $R_{\gamma_t}(x)=tx$.*
(4) *Their moments are $M_k=\sum_{\pi\in D(k)}t^{|\pi|}$, with $D=P_2,NC_2$.*

PROOF. These are all results that we already know, the idea being that (3,4) follow by doing some combinatorics and calculus, and that (1,2) follow from (3,4). □

We will be back later on, with more conceptual explanations for all this.

6d. Free PLT

We have the following free analogue of the PLT, whose statement is identical to that of the PLT, except for the fact that the convolution operation $*$ there is replaced by a free convolution operation $\boxplus$, and that we obtain a different limiting measure:

THEOREM 6.12 (Free PLT). *The following limit converges, for any $t > 0$,*

$$\lim_{n\to\infty}\left(\left(1-\frac{t}{n}\right)\delta_0+\frac{t}{n}\delta_1\right)^{\boxplus n}$$

and we obtain the Marchenko-Pastur law of parameter t,

$$\pi_t = \max(1-t,0)\delta_0 + \frac{\sqrt{4t-(x-1-t)^2}}{2\pi x}\,dx$$

also called free Poisson law of parameter t.

PROOF. Consider the measure in the statement, under the free convolution sign:

$$\mu = \left(1-\frac{t}{n}\right)\delta_0+\frac{t}{n}\delta_1$$

The Cauchy transform of this measure is elementary to compute, given by:

$$G_\mu(\xi) = \left(1-\frac{t}{n}\right)\frac{1}{\xi}+\frac{t}{n}\cdot\frac{1}{\xi-1}$$

We want to compute the following R-transform:

$$R = R_{\mu^{\boxplus n}}(y) = nR_\mu(y)$$

The equation for this function R is as follows:

$$\left(1-\frac{t}{n}\right)\frac{1}{y^{-1}+R/n}+\frac{t}{n}\cdot\frac{1}{y^{-1}+R/n-1} = y$$

By multiplying by n/y, this equation can be written as:

$$\frac{t+yR}{1+yR/n} = \frac{t}{1+yR/n-y}$$

With $n\to\infty$ we obtain the following formula:

$$t+yR = \frac{t}{1-y}$$

Thus we have the following formula:

$$R = \frac{t}{1-y}$$

But this gives the result, since R_{π_t} is elementary to compute from what we have, by "doubling" the results for the Wigner law γ_t, and is given by the same formula. $\square$

As in the continuous case, most of the basic theory of π_t was done in section 4 above, with all this partly coming from the theory SO_3, at $t=1$. One thing which was missing there, however, was that of understanding how the law π_t, with parameter $t>0$, exactly appears, out of π_1.

With free probability theory, we can now solve this question:

THEOREM 6.13. *The Marchenko-Pastur laws have the property*

$$\pi_s \boxplus \pi_t = \pi_{s+t}$$

so they form a 1*-parameter semigroup with respect to free convolution.*

PROOF. This follows either from Theorem 6.12, or from the fact that the R-transform of π_t, computed in the proof of Theorem 6.12, is linear in t. □

In analogy with Theorem 6.11 above, we can summarize the various discrete results that we have, classical and free, as follows:

THEOREM 6.14. *The Poisson laws p_t and the Marchenko-Pastur laws π_t, given by*

$$p_t = e^{-t}\sum_k \frac{t^k}{k!}\,\delta_k$$

$$\pi_t = \max(1-t,0)\delta_0 + \frac{\sqrt{4t-(x-1-t)^2}}{2\pi x}\,dx$$

have the following properties:

(1) *They appear via the PLT, and the free PLT.*
(2) *They form semigroups with respect to $*$ and $\boxplus$.*
(3) *Their transforms are $\log F_{p_t}(x) = t(e^{ix}-1)$, $R_{\pi_t}(x) = t/(1-x)$.*
(4) *Their moments are $M_k = \sum_{\pi\in D(k)} t^{|\pi|}$, with $D = P, NC$.*

PROOF. These are all results that we already know, from the previous chapters, the idea being as follows:

(1,2) follow from (3,4).

(3,4) follow by doing some combinatorics and calculus. □

There is a clear similarity here with Theorem 6.11, especially at the level of the moments, where the global result, including what we have so far, is as follows:

THEOREM 6.15. *The moments of the various central limiting measures, namely*

$$\begin{array}{ccc}\pi_t & - & \gamma_t\\ | & & |\\ p_t & - & g_t\end{array}$$

are always given by the same formula, involving partitions, namely

$$M_k=\sum_{\pi\in D(k)}t^{|\pi|}$$

where the sets of partitions $D(k)$ in question are respectively

$$\begin{array}{ccc}NC & \leftarrow & NC_2\\ \downarrow & & \downarrow\\ P & \leftarrow & P_2\end{array}$$

and where $|.|$ is the number of blocks.

PROOF. This follows by putting together the various moment results that we have, from Theorem 6.11 and Theorem 6.14. □

We will see later on a more conceptual explanation for the above result, in terms of cumulants, and a number of extensions as well, eventually ending up with a cube.

6e. Exercises

CHAPTER 7

Circular variables

7a. Free CCLT

We have seen so far that free probability leads to two key limiting theorems, namely the free analogues of the CLT and PLT. The limiting measures are the Wigner semicircle laws γ_t and the Marchenko-Pastur laws π_t, previously known from SU_2, SO_3 and from random matrices. Together with the Gaussian laws g_t and the Poisson laws p_t, appearing from the classical CLT and PLT, these laws form a square diagram, as follows:

$$\begin{array}{ccc}
\pi_t & - & \gamma_t \\
| & & | \\
p_t & - & g_t
\end{array}$$

Our purpose here will be that of extending this diagram to the right, with a free analogue of the complex central limiting theorem (CCLT), adding to the classical CCLT, and providing us with free analogues Γ_t of the complex Gaussian laws G_t.

As in the classical case, there is actually not so much work to be done here, because we can obtain the free CCLT from the free CLT, simply by taking the real and imaginary parts of our variables. Let us start with the construction of the limiting measure:

Definition 7.1. *The Voiculescu circular law of parameter $t > 0$ is given by*

$$\Gamma_t = law\left(\frac{1}{\sqrt{2}}(a+ib)\right)$$

where a, b are free, each following the Wigner semicircle law γ_t.

In other words, the passage $\gamma_t \to \Gamma_t$ is by definition entirely similar to the passage $g_t \to G_t$ from the classical case, by taking real and imaginary parts. As before in other similar situations, the fact that Γ_t is indeed well-defined is clear from definitions.

Let us start with a number of straightforward results, obtained by complexifying the free probability theory that we have. As a first result, we have, as announced above, the following natural free analogue of the complex central limiting theorem (CCLT):

Theorem 7.2 (Free CCLT). *Given random variables $x_1, x_2, x_3, \ldots$ which are f.i.d., centered, with variance $t > 0$, we have, with $n \to \infty$, in moments,*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n x_i \sim \Gamma_t$$

where Γ_t is the Voiculescu circular law of parameter t.

Proof. This follows indeed from the free CLT, established in section 6 above, by taking real and imaginary parts. Indeed, let us write:

$$x_i = \frac{1}{\sqrt{2}}(y_i + iz_i)$$

The variables y_i satisfy then the assumptions of the free CLT, and so their rescaled averages converge to a semicircle law γ_t, and the same happens for the variables z_i:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n y_i \sim \gamma_t \quad , \quad \frac{1}{\sqrt{n}}\sum_{i=1}^n z_i \sim \gamma_t$$

Now since the two limiting semicircle laws that we obtain in this way are free, their rescaled sum is circular, in the sense of Definition 7.1, and this gives the result. □

The circular laws are very interesting objects, both from a combinatorial and probabilistic point of view, so let us develop now some theory for them. First, we have:

Theorem 7.3. *The Voiculescu circular laws have the property*

$$\Gamma_s \boxplus \Gamma_t = \Gamma_{s+t}$$

so they form a 1-parameter semigroup with respect to free convolution.

Proof. This follows from our previous result stating that the Wigner laws γ_t have the free semigroup convolution property, by taking real and imaginary parts. □

7b. Moments, combinatorics

Less trivial now is the computation of the moments, and we have here:

Theorem 7.4. *The moments of the Voiculescu laws are the numbers*

$$M_k(\Gamma_t) = \sum_{\pi \in \mathcal{NC}_2(k)} t^{|\pi|}$$

with "$\mathcal{NC}_2$" standing for the noncrossing matching pairings.

Proof. This follows indeed by doing some computations. □

All this is quite nice, and we can now formulate a global result, as follows:

THEOREM 7.5. *The complex Gaussian laws G_t and the Voiculescu laws Γ_t, given by*

$$G_t = law\left(\frac{1}{\sqrt{2}}(a+ib)\right)$$

$$\Gamma_t = law\left(\frac{1}{\sqrt{2}}(\alpha+i\beta)\right)$$

where $a,b/\alpha,\beta$ are independent/free, following g_t/γ_t, have the following properties:

(1) *They appear via the complex CLT, and the free complex CLT.*
(2) *They form semigroups with respect to the operations $*$ and $\boxplus$.*
(3) *Their moments are $M_k = \sum_{\pi\in D(k)} t^{|\pi|}$, with $D = \mathcal{P}_2, \mathcal{NC}_2$.*

PROOF. This is a summary of results that we already know, with (1,2) being quite straightforward, and with (3) coming by doing some combinatorics and calculus. □

At the level of moments, we can now put everything together, as follows:

THEOREM 7.6. *The moments of the various central limiting measures, namely*

$$\begin{array}{ccccc}
\pi_t & - & \gamma_t & - & \Gamma_t \\
| & & | & & | \\
p_t & - & g_t & - & G_t
\end{array}$$

are always given by the same formula, involving partitions, namely

$$M_k = \sum_{\pi\in D(k)} t^{|\pi|}$$

where the sets of partitions $D(k)$ in question are respectively

$$\begin{array}{ccccc}
NC & \leftarrow & NC_2 & \leftarrow & \mathcal{NC}_2 \\
\downarrow & & \downarrow & & \downarrow \\
P & \leftarrow & P_2 & \leftarrow & \mathcal{P}_2
\end{array}$$

and where $|.|$ is the number of blocks.

PROOF. This follows by putting together the various moment results that we have, from the previous section, and from Theorem 7.4 above. □

7c. Semigroup models

We present now a number of more advanced results regarding the circular laws Γ_t, which have no classical counterpart. Let us start with:

THEOREM 7.7. *We have the following results:*

1. *Any matrix* $T\in M_N(\mathbb C)$ *has a polar decomposition,* $T=U|T|$.
2. *Assuming* $T\in A\subset M_N(\mathbb C)$*, we have* $U,|T|\in A$.
3. *Any operator* $T\in B(H)$ *has a polar decomposition,* $T=U|T|$.
4. *Assuming* $T\in A\subset B(H)$*, we have* $U,|T|\in\bar{A}$*, weak closure.*

PROOF. All this is standard, the idea being as follows:

(1) In each case under consideration, the first observation is that the matrix or general operator T^*T being positive, it has a square root:

$$|T|=\sqrt{T^*T}$$

(2) With this square root extracted, in the invertible case we can compare the action of T and $|T|$, and we conclude that we have $T=U|T|$, with U being a unitary. In the general, non-invertible case, a similar analysis leads to the conclusion that we have as well $T=U|T|$, but with U being this time a partial isometry.

(3) In what regards now algebraic and topological aspects, in finite dimensions the extraction of the square root, and so the polar decomposition itself, takes place over the matrix blocks of the ambient algebra $A\subset M_N(\mathbb C)$, and so takes place inside A itself.

(4) In infinite dimensions however, we must take the weak closure, an illustrating example here being the functions $f\in A$ belonging to the algebra $A=C(X)$, represented on $H=L^2(X)$, whose polar decomposition leads into the bigger algebra $\bar{A}=L^\infty(X)$. □

As a first illustration for all this, we have:

PROPOSITION 7.8. *The polar decomposition of semicircular variables is*

$$s=eq$$

with the variables e,q *being as follows:*

1. e *has moments* $1,0,1,0,1,\ldots$
2. q *is quarter-circular.*
3. e,q *are independent.*

PROOF. It is enough to prove the result in a model of our choice, and the best choice here is the most straightforward one, namely:

$$s=x\in L^\infty\Big([-2,2],\gamma_1\Big)$$

To be more precise, we endow the interval $[-2,2]$ with the probability measure γ_1, and we consider here the variable $s = x = (x \to x)$, which is trivially semicircular. The polar decomposition of this variable is then $s = eq$, with e, q being as follows:

$$e = sgn(x) \quad , \quad q = |x|$$

Now since e has moments $1, 0, 1, 0, 1, \ldots$, and also q is quarter-circular, and finally e, q are independent, this gives the result in our model, and so in general. □

In general now, in order to deal with such questions, we are in need of results regarding the multiplicative free convolution operation $\boxtimes$. Let us recall that we have:

Definition 7.9. *We have a free convolution operation $\boxtimes$, constructed as follows:*

(1) *For abstract distributions, via $\mu_b \boxtimes \mu_c = \mu_{bc}$, with b, c free.*
(2) *For real measures, via $\mu_b \boxtimes \mu_c = \mu_{\sqrt{b}c\sqrt{b}}$, with b, c self-adjoint and free.*

All this is quite tricky, explained in chapter 5 above, the idea being that, while (1) is straightforward, (2) is not, and comes by considering the variable $a = \sqrt{b}c\sqrt{c}$, which unlike bc is always self-adjoint, and whose moments are given by:

$$\begin{aligned} tr(a^k) &= tr[(\sqrt{b}c\sqrt{b})^k] \\ &= tr[\sqrt{b}cb \ldots bc\sqrt{b}] \\ &= tr[\sqrt{b} \cdot \sqrt{b}cb \ldots bc] \\ &= tr[(bc)^k] \end{aligned}$$

Quite remarkably, the free multiplicative convolution operation $\boxtimes$ can be linearized, in analogy with what happens for the usual multiplicative convolution $\times$, and the additive operations $*, \boxplus$ as well. We have here the following result, due to Voiculescu:

Theorem 7.10. *The free multiplicative convolution operation $\boxtimes$ for the real probability measures $\mu \in \mathcal{P}(\mathbb{R})$ can be linearized as follows:*

(1) *Start with the sequence of moments M_k, and then compute the moment generating function, or Stieltjes transform of the measure:*
$$f(z) = 1 + M_1 z + M_2 z^2 + M_3 z^3 + \ldots$$
(2) *Perform the following operations to the Stieltjes transform:*
$$\psi(z) = f(z) - 1$$
$$\psi(\chi(z)) = z$$
$$S(z) = (1 + z^{-1})\chi(z)$$
(3) *Then $\log S$ linearizes the free multiplicative convolution, in the sense that:*
$$S_{\mu \boxtimes \nu} = S_\mu S_\nu$$

PROOF. There are several proofs here, the idea being as follows:

(1) Voiculescu's original proof is quite similar to the proof of the R-transform theorem, by using some suitable models for the variables involved.

(2) There is another proof by Haagerup, using some alternative arguments, and some other proofs using combinatorics as well, which are all interesting. □

With such kind of technology, we can now prove the following key result, originally found by Voiculescu by using random matrix techniques:

THEOREM 7.11. *The polar decomposition of circular variables is*

$$c = uq$$

with the variables u, q being as follows:

(1) *u is a Haar-unitary.*
(2) *q is quarter-circular.*
(3) *u, q are free.*

PROOF. This is something which looks quite similar to Proposition 7.8, but which is far more difficult, and can be however proved, via some advanced combinatorics. □

In the remainder of this section we discuss yet another approach to Theorem 7.11, this time with algebraic techniques. A bit in analogy with the Voiculescu original proof, these techniques will be once again quite heavy, but they will trivialize the problem.

The idea will be that of using a $l^2(\mathbb{Z} * \mathbb{N})$ trick, with the Haar-unitaries being easy to model on $l^2(\mathbb{Z})$, and with the semicircular and quarter-circulars being easy to model on $l^2(\mathbb{N})$. For this purpose we will use semigroup algebras, jointly generalizing the main models that we have, namely group algebras, and free Fock spaces. Let us start with:

DEFINITION 7.12. *We call "semigroup" a unital semigroup, embeddable into a group:*

$$M \subset G$$

For such a semigroup M, we use the notation

$$M^{-1} = \left\{ m^{-1} \Big| m \in M \right\}$$

as a subset in some group containing M.

Observe that the embeddability assumption $M \subset G$ tells us that the usual cancellation rules hold in M, namely:

$$ab = ac \implies b = c$$
$$ba = ca \implies b = c$$

Regarding the precise relation between M and G, it is possible to talk here about the Grothendieck group G associated to such a semigroup M.

With the above definition in hand, we have the following construction, which unifies the main models that we have, namely the group algebras, and the free Fock spaces:

PROPOSITION 7.13. *Let M be a semigroup. By using the left simplifiability of M one can define, as for discrete groups, an embedding of semigroups, as follows:*

$$(M,\cdot)\to(B(l^2(M)),\mathbb{C})$$

$$\lambda_M(m)\delta_n=\delta_{mn}$$

Let $W^*(M)$ be the von Neumann algebra generated by $\lambda_M(M)$. Together with the following canonical state, it is a noncommutative W^*-probability space:

$$\tau_M(T)=<T\delta_e,\delta_e>$$

The operators in $\lambda_M(M)$ are isometries, but not necessarily unitaries.

PROOF. Indeed, for every $m\in M$, $\lambda_M(m)^*$ is given by:

$$\begin{aligned}\lambda_M(m)^*(\delta_n) &= \sum_{x\in M}<\lambda_M(m)^*\delta_n,\delta_x>\delta_x\\ &= \sum_{x\in M}\delta_{n,mx}\delta_x\end{aligned}$$

Thus we have indeed the isometry property, namely:

$$\lambda_M(m)^*\lambda_M(m)=1$$

We are therefore led to the conclusions in the statement. □

At the level of examples, we have:

PROPOSITION 7.14. *We have the following examples:*

(1) *For groups we obtain the group algebras.*
(2) *$l^2(\mathbb{N}^{*I})$ is the full Fock space over $\mathbb{C}^I$, and via this identification, $(W^*(\mathbb{N}^{*I}),\tau_{\mathbb{N}^{*I}})$ is the algebra of creation operators, with the state associated to the vacuum vector.*

PROOF. This is indeed clear from definitions. □

Now let $M\subset N$ be semigroups, so that $l^2(M)\subset l^2(N)$. For $m,\ m'\in M$ we have:

$$\lambda_M(m)\delta_{m'}=\lambda_N(m)\delta_{m'}$$

Thus if we suppose $M(N-M)=N-M$ then we have:

$$\begin{aligned}\lambda_M(m)^*\delta_{m'} &= \sum_{x\in M}\delta_{m',mx}\delta_x\\ &= \sum_{x\in N}\delta_{m',mx}\delta_x\\ &= \lambda_N(m)^*\delta_{m'}\end{aligned}$$

In particular, if $m_1 \dots m_k \in M$ and $\alpha_1 \dots \alpha_k$ are exponents $\in \{1, *\}$ then:

$$\lambda_M(m_1)^{\alpha_1} \dots \lambda_M(m_k)^{\alpha_k} \delta_e = \lambda_N(m_1)^{\alpha_1} \dots \lambda_N(m_k)^{\alpha_k} \delta_e$$

It follows that we have the following result:

PROPOSITION 7.15. *If $M \subset N$ are semigroups such that $M(N-M) = N - M$ then for every family $\{a_i\}_{i\in I}$ of elements in M, the $*$-distribution joint to*

$$\{\lambda_N(a_i)\}_{i\in I}$$

is equal to the $*$-distribution joint to $\{\lambda_M(a_i)\}_{i\in I}$.

PROOF. This follows indeed from the above discussion. □

Let $\prod_{i\in I} M_i$ be a direct product of semigroups and $a \in M_k$, $k \in I$. Then, by the above the $*$-distribution of $\lambda_{M_k}(a) \in W^*(M_k)$ is equal to the $*$-distribution of:

$$\lambda_{\prod M_i}(a) \in W^*\left(\prod M_i\right)$$

Moreover, let $b \in M_j$ with $j \neq k$. Then $\lambda_{\prod M_i}(a)$ and $\lambda_{\prod M_i}(b)$ are independent. Now let $*_{i\in I} M_i$ be a free product and $a \in M_k$, $k \in I$, and consider the following variable:

$$\lambda_{M_k}(a) \in W^*(M_k)$$

By the above the $*$-distribution of this variable is equal to the $*$-distribution of:

$$\lambda_{*M_i}(a) \in W^*(*M_i)$$

Let us introduce some definitions related to the combinatorics of free semigroups:

DEFINITION 7.16. *Let N be a semigroup. Consider the following order relation on it:*

$$a \preceq_N b \iff b \in aN$$

We say that N is in the class E if it satisfies one of the following equivalent conditions:

(1) *For $\preceq_N$ every bounded subset is totally ordered.*
(2) *$a \preceq c, b \preceq c \implies a \preceq b$ or $b \preceq a$.*
(3) *$aN \cap bN \neq \emptyset \implies aN \subset bN$ or $bN \subset aN$.*
(4) *$NN^{-1} \cap N^{-1}N = N \cup N^{-1}$.*

We have as well:

DEFINITION 7.17. *Let $(a_i)_{i\in I}$ be a family of elements in a semigroup N. We call it "code" if it satisfies the following conditions:*

(1) *The semigroup M generated by the a_i's is isomorphic to $\mathbb{N}^{*I}$ by $a_i \to e_i$.*
(2) *$M(N-M) = N - M$.*

and "prefix" if $a_i \in a_j N \implies i = j$ (i.e. the a_i are not comparable by $\preceq_N$).

Let $(a_i, b_i)_{i\in I}$ be a code, and consider the following family:

$$(\lambda_N(a_i), \lambda_N(b_i))_{i\in I}$$

This family has the same $*$-distribution as a family of creation operators associated to a family of $2I$ orthonormal vectors, acting on the Fock space. In particular the following is a circular family:

$$\left(\frac{1}{2}(\lambda_N(a_i) + \lambda_N(b_i)^*)\right)_{i\in I}$$

Thus the following proposition is a nice criterion for finding circular systems in the algebras $W^*(N)$ of semigroups in the class E:

PROPOSITION 7.18. *For a semigroup $N \in E$, a family*

$$(a_i)_{i\in I} \subset N$$

having at least two elements is a prefix if and only if it is a code.

PROOF. Let $(a_i)_{i\in I}$ be a code which is not a prefix, for instance $a_i = a_j n$ with $i \neq j, n \in N$. Then n is in the semigroup M generated by the a_k and $a_i = a_j n$ with $i \neq j$, so M cannot be free, contradiction. Suppose now that $(a_i)_{i\in I}$ is a prefix and let, with $m \in N$:

$$A = a_{i_1}^{\alpha_1} \ldots a_{i_n}^{\alpha_n} m = a_{j_1}^{\beta_1} \ldots a_{j_s}^{\beta_s}$$

We have $a_{i_1} \preceq A$, $a_{j_1} \preceq A$, and so:

$$i_1 = j_1$$

We simplify A to the left by a_{i_1}. A reccurence on $\Sigma\alpha_i$ shows that:

$$n \leq s$$

$$a_{i_k} = a_{j_k} \quad , \quad \forall k \leq n$$

$$\alpha_k = \beta_k \quad , \quad \forall k < n$$

$$\alpha_n \leq \beta_n$$

$$m = a_{j_n}^{\beta_n - \alpha_n} a_{j_{n+1}}^{\beta_{n+1}} \ldots a_{j_s}^{\beta_s}$$

Finally, m is in the semigroup generated by the a_i, so we have a code. Moreover, for $m = e$ we obtain $n = s$, $a_{j_k} = a_{i_k}$, $\alpha_k = \beta_k$, $(\forall k \leq n)$ so the a_i generate freely M and $(a_i)_{i\in I}$ is a code. □

Let us discuss now the examples. We have here the following result:

PROPOSITION 7.19. *The class E has the following properties:*

1. *All the groups are in E.*
2. *The positive parts of totally ordered abelian groups are in E.*
3. *If G is a group and $M \in E$, then $M \times G \in E$.*
4. *If A_1, A_2 are in E, then the free product $A_1 * A_2$ is in E.*

PROOF. The proof goes as follows:

(1) This is obvious.

(2) This is obvious as well, because M is totally ordered by $\preceq_M$.

(3) Let G be a group and $M \in E$. We have then, as desired:

$$\begin{aligned}
& (M \times G)(M \times G)^{-1} \cap (M \times G)^{-1}(M \times G) \\
= \ & (M \times G)(M^{-1} \times G) \cap (M^{-1} \times G)(M \times G) \\
= \ & (MM^{-1} \times G) \cap (M^{-1}M \times G) \\
= \ & (MM^{-1} \cap M^{-1}M) \times G \\
= \ & (M \cup M^{-1}) \times G \\
= \ & (M \times G) \cup (M^{-1} \times G) \\
= \ & (M \times G) \cup (M \times G)^{-1}
\end{aligned}$$

(4) Let $a, b, c \in A_1 * A_2$ such that $ab = c$. We write, as reduced words:

$$a = x_1 \ldots x_n$$

$$b = y_1 \ldots y_m$$

$$c = z_1 \ldots z_p$$

Now let s be such that:

$$x_n y_1 = 1$$

$$\vdots$$

$$x_{n-s+1} y_s = 1$$

$$x_{n-s} y_{s+1} \neq 1$$

Consider now the following element:

$$u = x_{n-s+1} \ldots x_n = (y_1 \ldots y_s)^{-1}$$

We have then:

$$c = ab = x_1 \ldots x_{n-s} y_{s+1} \ldots y_m$$

Let $i \in \{1, 2\}$ be such that $z_{n-s} \in A_i$. There are two cases:

– If $x_{n-s} \in A_1$ and $y_{s+1} \in A_2$ or if $x_{n-s} \in A_2$ and $y_{s+1} \in A_1$, then $x_1 \ldots x_{n-s} y_{s+1} \ldots y_m$ is a reduced word. In particular, $x_1 = z_1$ and so on, up to $x_{n-s} = z_{n-s}$. Thus $a = z_1 \ldots z_{n-s} u$, with u invertible.

– If $x_{n-s}, y_{s+1} \in A_i$ then $x_1 = z_1$ and so on, up to $x_{n-s-1} = z_{n-s-1}$ and $x_{n-s} y_{s+1} = z_{n-s}$. In this case $a = z_1 \ldots z_{n-s-1} x_{n-s} u$, with u invertible.

Observe that in both cases we obtained that a is of the form $z_1 \ldots z_f x u$ for some f, with u invertible and such that if $z_{f+1} \in A_i$, then there exists $y \in A_i$ with:

$$xy = z_{f+1}$$

Indeed, we can take $f = n - s - 1$ and $x = z_{n-s}, y = 1$ in the first case, $x = x_{n-s}, y = y_{s+1}$ in the second one. Suppose now that $A_1, A_2 \in E$ and let $a, b, a', b' \in A_1 * A_2$ such that $ab = a'b'$. Let $z_1 \ldots z_p$ be the decomposition of $ab = a'b'$ as a reduced word. Then we can decompose, as above:

$$a = z_1 \ldots z_f x u$$

$$a' = z_1 \ldots z_{f'} x' u'$$

We have to show that $a = a'm$ or that $a' = am$ for some $m \in A_1 * A_2$. There are three cases:

– If $f < f'$, then $a' = au^{-1}yz_{f+2} \ldots z_{f'}x'u'$.

– If $f' < f$, then $a = a'u'^{-1}z_{f'+2} \ldots z_f xu$.

– If $f = f'$, then $xy = x'y' = z_{f+1} \in A_i$ for some $i \in \{1, 2\}$. As $A_i \in E$, we have that $x = x'm$ or $x' = xm$ for some $m \in A_i$, so that $a = a'u'^{-1}mu$ or $a' = aumu'$. □

With the above ingredients in hand, we can now formulate our main result about freeness in semigroup algebras, which is as follows:

Proposition 7.20. *The following happen:*

(1) *Let $M \subset N$ be two semigroups in the class E such that*

$$M(N - M) = N - M$$

Let $\lambda = \lambda_N$. Then every element x of the $*$-algebra generated by $\lambda(M)$ can be written as follows, with $p_i, q_i \in M$:

$$x = \Sigma a_i \lambda(p_i)\lambda(q_i)^*$$

(2) *Let $A, B \in E$, $\lambda = \lambda_{A*B}$, $\tau = \tau_{A*B}$, and let x be an element of the $*$-algebra generated by $\lambda(A)$ such that $\tau(x) = 0$. Denote by W_A the set of reduced words beginning by an element of A, and by W_B the set of reduced words beginning by an element of B. Then x maps:*

$$l^2(W_B \cup \{e\}) \to l^2(W_A)$$

(3) *Let $A, B \in E$. Then $\lambda_{A*B}(A)$ and $\lambda_{A*B}(B)$ are $*$-free.*

Proof. The idea here is as follows:

(1) It is enough to prove this for $x = \lambda(m)^*\lambda(n)$ with $m, n \in M$, because the general case will follow easily from this. In order to do so, observe that $x = \lambda(m)^*\lambda(n)$ is different from 0 iff $\exists a, b \in N$ such that:

$$< \lambda(m)^*\lambda(n)\delta_a, \delta_b > \neq 0$$

That is, the following condition must be satisfied:

$$na = mb$$

We know that $\exists c \in N$ with $n = mc$ or with $m = nc$. Moreover, as $M(N-M) = N-M$, it follows that $c \in M$. Thus $x = \lambda(m)^*\lambda(n) \neq 0$ implies that $x = \lambda(c)$ or $x = \lambda(c)^*$ with $c \in M$, and this finishes the proof.

(2) We apply (1) with $M = A$ and $N = A * B$ for writing, with $p_i, q_i \in A$:

$$x = \sum a_i \lambda(p_i)\lambda(q_i)^*$$

Consider now the following element:

$$\tau(\lambda(p_i)\lambda(q_i)^*) = \sum_x \delta_{e,p_i x}\delta_{e,q_i x}$$

This element is nonzero iff $p_i = q_i=$ invertible, and in this case:

$$\lambda(p_i)\lambda(q_i)^* = 1$$

As $\tau(x) = 0$, it follows that we can write:

$$x = \sum a_i \lambda(p_i)\lambda(q_i)^*$$

$$\tau(\lambda(p_i)\lambda(q_i)^*) = 0$$

By linearity, it is enough to prove the result for $x = \lambda(p_i)\lambda(q_i)^*$. Let $m \in W_B \cup \{e\}$ and suppose that $x\delta_m \neq 0$. Then $\lambda(q_i)^*\delta_m \neq 0$ implies that $m = q_i c$ for some word $c \in A * B$. As $q_i \in A$ and $m \in W_B \cup \{e\}$, it follows that q_i is invertible. Now observe that:

$$p_i q_i^{-1} = 1 \implies \tau(x) = 1$$

It follows that we have:

$$x\delta_m = \delta_{p_i q_i^{-1} m} \in l^2(W_A)$$

(3) This follows from (2). Indeed, let $P = x_n \ldots x_1$ be a product of elements in $ker(\tau)$, such that x_{2k} is in the $*$-algebra generated by $\lambda(B)$ and x_{2k+1} is in the $*$-algebra generated by $\lambda(A)$. Then $x_1\delta_e \in l^2(W_A)$. Thus $x_2 x_1 \delta_e \in l^2(W_B)$, and so on. By a reccurence, $P\delta_e$ is in $l^2(W_A)$ or in $l^2(W_B)$. But this implies that $\tau(P) = 0$. □

7d. Polar decomposition

We have the following application of the above:

Theorem 7.21. *Consider a Haar-unitary u, $*$-free from a semicircular s. Then*

$$c = us$$

is a circular variable.

PROOF. Denote by z the image of $1 \in \mathbb{Z}$ and by n the image of $1 \in \mathbb{N}$ by the canonical embeddings into the free product $\mathbb{Z} * \mathbb{N}$. Let $\lambda = \lambda_{\mathbb{Z}*\mathbb{N}}$. We know that $\mathbb{Z} * \mathbb{N} \in E$. Also (zn, nz^{-1}) is obviously a prefix, so it is a code. Thus, the following variable is circular:

$$c = \frac{1}{2}(\lambda(zn) + \lambda(nz^{-1})^*)$$

The point now is that we have:

$$\frac{1}{2}(\lambda(zn) + \lambda(nz^{-1})^*) = us$$

But this gives the result, because $u = \lambda(z)$ is a Haar-unitary, $s = 1/2(\lambda(n) + \lambda(n)^*)$ is semicircular, and u and s are $*$-free. □

We have the following result:

THEOREM 7.22. *Consider the polar decomposition of a circular variable in some W^*-probability space with faithful normal state:*

$$x = vb$$

Then v is Haar-unitary, b is quarter-circular and (v, b) is a $*$-free pair.

PROOF. Consider the following group:

$$G = \mathbb{Z} * (\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z})$$

Let us denote by z, t, a the images of the following elements, by the canonical embeddings into G:

$$1 \in \mathbb{Z}$$
$$(1, \hat{0}) \in \mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})$$
$$(0, \hat{1}) \in \mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})$$

Let $u = \lambda_G(z)$, $d = \lambda_G(a)$ and choose a quarter-circular:

$$q \in W^*(\lambda_G(t))$$

Then (q, d) are independent, and the distribution of d is given by:

$$\mu_d(X^{2k}) = 1 \quad , \quad \mu_d(X^{2k+1}) = 0$$

Then dq is semicircular, so by $c = udq$ is circular, and:

– The module of c is q, which is a quarter-circular.

– The polar part of c is ud, which is obviously a Haar-unitary.

– Consider the automorphism ψ of G which is the identity on $\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ and sends $z \to za$. It extends to a trace-preserving automorphism $\tilde{\psi}$ of $W^*(G)$ which sends:

$$u \to ud \quad , \quad q \to q$$

As u and q are $*$-free, it follows that ud and q are $*$-free. □

Regarding now the moments, the result is as follows:

THEOREM 7.23. *The moments of the Voiculescu laws are the numbers*

$$M_k(\Gamma_t)=\sum_{\pi\in\mathcal{NC}_2(k)}t^{|\pi|}$$

with "$\mathcal{NC}_2$" standing for the noncrossing matching pairings.

PROOF. This is something that we already know, and can be recovered as well by using the polar decomposition result, once again by doing some computations. □

All this is quite nice, and we can now formulate a global result, as follows:

THEOREM 7.24. *The moments of the various central limiting measures, namely*

$$\begin{array}{ccccc}
\pi_t & - & \gamma_t & - & \Gamma_t\\
| & & | & & |\\
p_t & - & g_t & - & G_t
\end{array}$$

are always given by the same formula, involving partitions, namely

$$M_k=\sum_{\pi\in D(k)}t^{|\pi|}$$

where the sets of partitions $D(k)$ in question are respectively

$$\begin{array}{ccccc}
NC & \leftarrow & NC_2 & \leftarrow & \mathcal{NC}_2\\
\downarrow & & \downarrow & & \downarrow\\
P & \leftarrow & P_2 & \leftarrow & \mathcal{P}_2
\end{array}$$

and where $|.|$ is the number of blocks.

PROOF. This is something that we already know, and which follows by putting together the various moment results that we have. □

We will be back to this later, in the next section, after developing some random matrix models as well. We will see there, following the original papers of Voiculescu, that Theorem 7.22 is something very simple, in the random matrix setting.

7e. Exercises

CHAPTER 8

Gaussian matrices

8a. Wigner matrices

Let us begin by specifying the class of matrices that we are interested in. First we have the complex Gaussian matrices and the Wigner matrices, constructed as follows:

Definition 8.1. *A complex Gaussian matrix is a random matrix of type*

$$Z \in M_N(L^\infty(X))$$

having i.i.d. complex normal entries. When imposing the extra condition $Z = Z^*$, the resulting matrix is called a Wigner matrix.

Observe that a Wigner matrix must be as follows, with X_i being real normal variables, Z_{ij} being complex normal variables, and all these variables being independent:

$$W = \begin{pmatrix} X_1 & Z_{12} & \dots & \dots & Z_{1N} \\ \bar{Z}_{12} & X_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & X_{N-1} & Z_{N-1,N} \\ \bar{Z}_{1N} & \dots & \dots & \bar{Z}_{N-1,N} & X_N \end{pmatrix}$$

The relation between the Wigner and Gaussian matrices is via complexification, with the real and imaginary parts of a Gaussian matrix being Wigner matrices, having mutually independent entries.

8b. Asymptotic freeness

We have the following key result, due to Wigner and Voiculescu:

Theorem 8.2. *A family of Wigner matrices, with pairwise independent entries*

$$Z^i \in M_N(L^\infty(X))$$

becomes with $N \to \infty$ semicircular, and free.

Proof. Here the first assertion is due to Wigner, and the second assertion is due to Voiculescu, and both these assertions can be proved by using the moment method. Indeed, in the $N \to \infty$ limit the combinatorics simplifies, and we obtain both the results. □

8c. Complex matrices

Getting now to the complex case, we have here:

THEOREM 8.3. *A family of complex Gaussian matrices, with pairwise independent entries*

$$Z^i \in M_N(L^\infty(X))$$

becomes with $N \to \infty$ circular, and free.

PROOF. This follows from Theorem 8.2 above, which applies to the real and imaginary parts of our complex Gaussian matrices, and gives the result. □

The above results are interesting for both free probability and random matrices. As a illustration here, we have the folowing application to free probability:

THEOREM 8.4. *Polar decomposition of circular variables.*

PROOF. This is indeed quite easy to see in the Gaussian matrix model provided by Theorem 8.3 above. Now since this holds in the model, it must hold in general. □

There are many other applications along this lines, and conversely, free probability can be used as well for the detailed study of the Wigner and Gaussian matrices.

8d. Wishart matrices

We have so far results about the "real" and "complex" cases. In the "real positive" case, the objects of study are the complex Wishart matrices, constructed as follows:

DEFINITION 8.5. *A complex Wishart matrix is a random matrix of type*

$$Z = GG^* \in M_N(L^\infty(X))$$

with G being a complex Gaussian matrix.

We have the following result, improving the Marchenko-Pastur theorem:

THEOREM 8.6. *A family of complex Wishart matrices, with pairwise independent input*

$$Z^i \in M_N(L^\infty(X))$$

becomes with $N \to \infty$ free Poisson, and free.

PROOF. Here the first assertion is the Marchenko-Pastur theorem, and the second assertion follows from the freeness result from Theorem 8.2, or from Theorem 8.3. □

8e. Exercises

Part III

The bijection

CHAPTER 9

Poisson limits

9a. Poisson limits

We have seen so far that free probability leads to three key limiting theorems, namely the free analogues of the PLT, CLT and CCLT. The limiting measures are the Marchenko-Pastur laws π_t, the Wigner semicircle laws γ_t and the Voiculescu circular laws Γ_t. Together with the Poisson laws p_t and the Gaussian laws g_t and G_t appearing from the classical PLT, CLT and CCLT, these laws form a rectangular diagram, as follows:

$$\begin{array}{ccccc}
\pi_t & - & \gamma_t & - & \Gamma_t \\
| & & | & & | \\
p_t & - & g_t & - & G_t
\end{array}$$

In this section we develop some more limiting theorems, by generalizing the free PLT into a free compound Poisson limit theorem (CPLT). At the level of the above diagram, there are no complex analogues of p_t, π_t, but by using measures found via the classical and free CLPT, namely the Bessel laws discussed in section 2 above, and their free analogues to be discussed here, we will be able to modify and then fold the diagram, and eventually complete it into a cube. This is of course quite nice, theoretically speaking, because it leads to a kind of 3D orientation inside the whole subject, which is useful.

Let us start with the following definition:

Definition 9.1. *Associated to any compactly supported positive measure ρ on $\mathbb{R}$ is the probability measure*

$$\pi_\rho = \lim_{n\to\infty}\left(\left(1-\frac{c}{n}\right)\delta_0+\frac{1}{n}\rho\right)^{\boxplus n}$$

where $c = mass(\rho)$, called compound free Poisson law.

In what follows we will be interested in the case where ρ is discrete, as is for instance the case for $\rho = t\delta_1$ with $t > 0$, which produces the free Poisson laws.

The following result allows one to detect compound Poisson laws:

Proposition 9.2. *For a discrete measure, written as $\rho = \sum_{i=1}^s c_i\delta_{z_i}$ with $c_i > 0$ and $z_i \in \mathbb{R}$, we have*

$$R_{\pi_\rho}(y) = \sum_{i=1}^s \frac{c_i z_i}{1 - y z_i}$$

where R denotes the Voiculescu R-transform.

Proof. Let μ_n be the measure appearing in Definition 9.1, under the convolution signs:

$$\mu_n = \left(1 - \frac{c}{n}\right)\delta_0 + \frac{1}{n}\rho$$

The Cauchy transform of μ_n is then given by:

$$G_{\mu_n}(\xi) = \left(1 - \frac{c}{n}\right)\frac{1}{\xi} + \frac{1}{n}\sum_{i=1}^s \frac{c_i}{\xi - z_i}$$

Consider now the R-transform of the measure $\mu_n^{\boxplus n}$, which is given by:

$$R_{\mu_n^{\boxplus n}}(y) = nR_{\mu_n}(y)$$

By using now the general theory of the R-transform, the above formula of G_{μ_n} shows that the equation for $R = R_{\mu_n^{\boxplus n}}$ is as follows:

$$\begin{aligned}
& \left(1 - \frac{c}{n}\right)\frac{1}{y^{-1} + R/n} + \frac{1}{n}\sum_{i=1}^s \frac{c_i}{y^{-1} + R/n - z_i} = y \\
\implies \quad & \left(1 - \frac{c}{n}\right)\frac{1}{1 + yR/n} + \frac{1}{n}\sum_{i=1}^s \frac{c_i}{1 + yR/n - yz_i} = 1
\end{aligned}$$

Now multiplying by n, rearranging the terms, and letting $n \to \infty$, we get:

$$\begin{aligned}
& \frac{c + yR}{1 + yR/n} = \sum_{i=1}^s \frac{c_i}{1 + yR/n - yz_i} \\
\implies \quad & c + yR_{\pi_\rho}(y) = \sum_{i=1}^s \frac{c_i}{1 - yz_i} \\
\implies \quad & R_{\pi_\rho}(y) = \sum_{i=1}^s \frac{c_i z_i}{1 - yz_i}
\end{aligned}$$

This finishes the proof in the free case, and we are done. □

We have as well the following result, providing an alternative to Definition 9.1:

THEOREM 9.3. *For a discrete measure, written as $\rho = \sum_{i=1}^s c_i\delta_{z_i}$ with $c_i > 0$ and $z_i \in \mathbb{R}$, we have*

$$\pi_\rho = \mathrm{law}\left(\sum_{i=1}^s z_i\alpha_i\right)$$

where the variables α_i are free Poisson(c_i), free.

PROOF. Let α be the sum of free Poisson variables in the statement:

$$\alpha = \sum_{i=1}^s z_i\alpha_i$$

We will show that the R-transform of α is given by the formula in Proposition 9.2. We have the following computation:

$$\begin{aligned} & R_{\alpha_i}(y) = \frac{c_i}{1-y} \\ \implies \quad & R_{z_i\alpha_i}(y) = \frac{c_iz_i}{1-yz_i} \\ \implies \quad & R_\alpha(y) = \sum_{i=1}^s \frac{c_iz_i}{1-yz_i} \end{aligned}$$

Thus we have indeed the same formulae as those in Proposition 9.2. □

All the above is quite general, and in practice, in order to obtain concrete results, the simplest measures that we can use as "input" for the CLPT are the measures of type $\rho = t\varepsilon_s$, with $t > 0$, and with ε_s being the uniform measure on the s-th roots of unity.

9b. Bessel laws

We are led in this way the following measures:

DEFINITION 9.4. *Consider the compound Poisson and free Poisson laws*

$$b_t^s = p_{t\varepsilon_s}$$

$$\beta_t^s = \pi_{t\varepsilon_s}$$

with ε_s being the uniform measure on the s-th roots of unity.

(1) *At $s = 1$ we recover the Poisson laws p_t, π_t.*
(2) *At $s = 2$ we have the real Bessel laws b_t, β_t.*
(3) *At $s = \infty$ we have the complex Bessel laws $B_t, \mathfrak{B}_t$.*

The terminology here comes from the fact, that we know from chapter 2 above, that the density of the measure b_t, appearing at $s = 2$, is a Bessel function of the first kind.

Our next task will be that upgrading our results about π_t in this setting, using a parameter $s \in \mathbb{N} \cup \{\infty\}$. First, we have the following result:

THEOREM 9.5. *The free Bessel laws have the property*

$$\beta^s_t \boxplus \beta^s_{t'} = \beta^s_{t+t'}$$

so they form a 1-parameter semigroup with respect to free convolution.

PROOF. This follows from the fact that the R-transform of β^s_t is linear in t, which is something that we already know, from the above. □

Regarding now the moments, we have the following result:

THEOREM 9.6. *The moments of β^s_t are the numbers*

$$M_k = \sum_{\pi \in NC^s(k)} t^{|\pi|}$$

where NC^s are the noncrossing partitions satisfying

$$\# \circ = \# \bullet \, (s)$$

in each block of the partition.

PROOF. This is something quite technical, as follows:

(1) At $t = 1$ the formula is as follows:

$$M_k(\beta^s_1) = |NC^s(k)|$$

But this can be proved by doing some combinatorics.

(2) At $t > 0$ the formula is as follows:

$$M_k = \sum_{\pi \in NC^s(k)} t^{|\pi|}$$

But this is something more technical, requiring substantial computations. □

We have as well the following result:

THEOREM 9.7. *Further properties of the free Bessel laws*

$$\beta^s_t \quad : \quad s \in \mathbb{N} \cup \{\infty\}, \ t > 0$$

involving the free multiplicative convolution operation $\boxtimes$.

PROOF. This is something quite technical as well. □

We will be back to this, in the next chapter.

At the combinatorial level, we can enlarge the previous diagram, as follows:

THEOREM 9.8. *The moments of the various central limiting measures, namely*

$$\begin{matrix}\beta_t^s & - & \gamma_t & - & \Gamma_t\\ | & & | & & |\\ b_t^s & - & g_t & - & G_t\end{matrix}$$

are always given by the same formula, involving partitions, namely

$$M_k=\sum_{\pi\in D(k)}t^{|\pi|}$$

where the sets of partitions $D(k)$ in question are respectively

$$\begin{matrix}NC^s & \leftarrow & NC_2 & \leftarrow & \mathcal{NC}_2\\ \downarrow & & \downarrow & & \downarrow\\ P^s & \leftarrow & P_2 & \leftarrow & \mathcal{P}_2\end{matrix}$$

and where $|.|$ is the number of blocks.

PROOF. This follows by putting together the various moment results that we have. □

9c. The standard cube

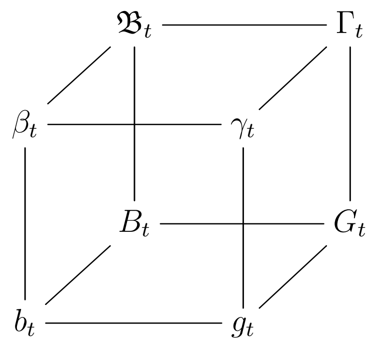
As already mentioned, in what regards the Bessel and free Bessel laws b_t^s,β_t^s, the important particular cases are $s=1,2,\infty$.

It is therefore tempting to leave one of these 3 cases aside, and fold the corresponding diagram, from Theorem 9.8, into a cube.

Quite surprisingly, in order to do so, in a correct way, the case which must be left aside is the most important one, namely $s=1$, corresponding to the Poisson and free Poisson laws p_t,π_t.

We will comment later on this, but let us just start by doing so:

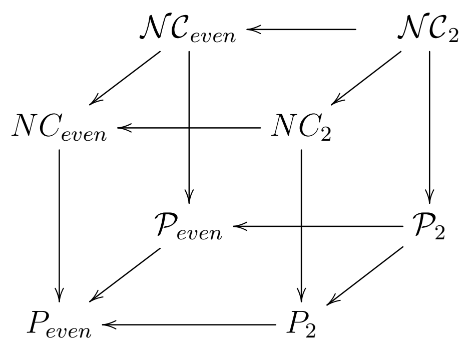
Theorem 9.9. *The moments of the various central limiting measures, namely*



are always given by the same formula, involving partitions, namely

$$M_k = \sum_{\pi \in D(k)} t^{|\pi|}$$

where the sets of partitions $D(k)$ in question are respectively



and where $|.|$ is the number of blocks.

Proof. This follows by putting together the various moment results that we have. □

We will see later more conceptual explanations for all this.

9d. Matrix models

We discuss here the relation with the random matrices.

9e. Exercises

CHAPTER 10

Bessel laws

10a. Basic properties

We denote by $\boxplus$ and $\boxtimes$ the free additive and multiplicative convolutions, and we use Voiculescu's R and S transforms, which linearize them.

Given a real probability measure μ, one can ask whether the convolution powers $\mu^{\boxtimes s}$ and $\mu^{\boxplus t}$ exist, for various values of $s,t>0$. For the free Poisson law, the answer to these questions is well-known. We include here the precise statement, along with a complete proof, which will serve as a model for some subsequent generalizations:

Theorem 10.1. *The measures $\pi^{\boxtimes s}$, $\pi^{\boxplus t}$ exist for any $s,t>0$.*

Proof. The free Poisson law π is the $t=1$ particular case of the free Poisson law of parameter t, given by:

$$\pi_t=\max(1-t,0)\delta_0+\frac{\sqrt{4t-(x-1-t)^2}}{2\pi x}\,dx$$

The Cauchy transform of this measure is given by:

$$G(\xi)=\frac{(\xi+1-t)+\sqrt{(\xi+1-t)^2-4\xi}}{2\xi}$$

We can compute now the R transform, by proceeding as follows:

$$\begin{aligned}
\xi G^2+1=(\xi+1-t)G \quad&\implies\quad Kz^2+1=(K+1-t)z\\
&\implies\quad Rz^2+z+1=(R+1-t)z+1\\
&\implies\quad Rz=R-t\\
&\implies\quad R=t/(1-z)
\end{aligned}$$

This expression being linear in t, the measures π_t form a semigroup with respect to free convolution. Thus we have $\pi_t=\pi^{\boxplus t}$, which proves the second assertion.

Regarding now the measure $\pi^{\boxtimes s}$, there is no explicit formula for its density. However, we can prove that this measure exists, by using some abstract results.

We have the following computation for the S transform of π_t:

$$\begin{aligned}\xi G^2+1=(\xi+1-t)G &\implies zf^2+1=(1+z-zt)f\\ &\implies z(\psi+1)^2+1=(1+z-zt)(\psi+1)\\ &\implies \chi(z+1)^2+1=(1+\chi-\chi t)(z+1)\\ &\implies \chi(z+1)(t+z)=z\\ &\implies S=1/(t+z)\end{aligned}$$

In particular at $t=1$ we have $S(z)=1/(1+z)$, so the Σ transform of π, which is by definition $\Sigma(z)=S(z/(1-z))$, is given by:

$$\Sigma(z)=1-z$$

The Σ transforms of the probability measures which are $\boxtimes$-infinitely divisible are the functions of the form $\Sigma(z)=e^{v(z)}$, where $v:\mathbb{C}-[0,\infty)\to\mathbb{C}$ is analytic, satisfying $v(\bar{z})=\bar{v}(z)$ and $v(\mathbb{C}^+)\subset\mathbb{C}^-$. In the case of the free Poisson law, the function $v(z)=\log(1-z)$ satisfies all the above properties, and this gives the result. □

We have the following remarkable identity:

THEOREM 10.2. *For $s\geq 1$ and $t\in(0,1]$ we have:*

$$\pi^{\boxtimes s-1}\boxtimes\pi^{\boxplus t}=((1-t)\delta_0+t\delta_1)\boxtimes\pi^{\boxtimes s}$$

PROOF. We know from the previous proof that the S transform of π is given by $S(z)=1/(1+z)$, and that the S transform of $\pi^{\boxplus t}$ is given by $S(z)=1/(t+z)$. Thus the measure on the left has the following S transform:

$$S(z)=\frac{1}{(1+z)^{s-1}}\cdot\frac{1}{t+z}$$

The S transform of $\alpha_t=(1-t)\delta_0+t\delta_1$ can be computed as follows:

$$\begin{aligned}f=1+tz/(1-z) &\implies \psi=tz/(1-z)\\ &\implies z=t\chi/(1-\chi)\\ &\implies \chi=z/(t+z)\\ &\implies S=(1+z)/(t+z)\end{aligned}$$

This shows that the measure on the right has the following S transform:

$$S(z)=\frac{1}{(1+z)^s}\cdot\frac{1+z}{t+z}$$

Thus the S transforms of our two measures are the same, and we are done. □

We are now in position of introducing a remarkable two-parameter family of real probability measures. We call them free Bessel laws, because of a certain relationship with the Bessel functions, to be discussed later on:

DEFINITION 10.3. *The free Bessel law is the real probability measure π_{st} with $(s,t)\in(0,\infty)\times(0,\infty)-(0,1)\times(1,\infty)$, defined as follows:*

(1) *For $s\geq 1$ we set $\pi_{st}=\pi^{\boxtimes s-1}\boxtimes\pi^{\boxplus t}$.*
(2) *For $t\leq 1$ we set $\pi_{st}=((1-t)\delta_0+t\delta_1)\boxtimes\pi^{\boxtimes s}$.*

We can regard the free Bessel law π_{st} as being a natural two-parameter generalization of the free Poisson law π, in connection with Voiculescu's free convolution operations $\boxtimes$ and $\boxplus$. Observe that we have the following formulae:

$$\begin{cases}\pi_{s1}=\pi^{\boxtimes s}\\ \pi_{1t}=\pi^{\boxplus t}\end{cases}$$

Concerning the precise range of the parameters (s,t), the above results can be probably improved. The point is that the measure π_{st} still exists for certain points in the critical rectangle $(0,1)\times(1,\infty)$, but not for all of them. The numeric checks show that the critical values of (s,t) seem to form an algebraic curve contained in $(0,1)\times(1,\infty)$, having $s=1$ as an asymptote. However, the case we are the most interested in is $t\in(0,1]$, and here there is no problem: π_{st} exists for any $s>0$.

10b. Diagrams

We have the following result:

PROPOSITION 10.4. *The Stieltjes transform of π_{st} satisfies:*

$$f=1+zf^s(f+t-1)$$

PROOF. We have the following computation:

$$\begin{aligned}S=\frac{1}{(1+z)^{s-1}}\cdot\frac{1}{t+z}\quad&\Longrightarrow\quad \chi=\frac{z}{(1+z)^s}\cdot\frac{1}{t+z}\\ &\Longrightarrow\quad z=\frac{\psi}{(1+\psi)^s}\cdot\frac{1}{t+\psi}\\ &\Longrightarrow\quad z=\frac{f-1}{f^s}\cdot\frac{1}{t+f-1}\end{aligned}$$

This gives the equation in the statement. □

We have the following result:

THEOREM 10.5. *The Stieltjes transform of π_{s1} with $s\in\mathbb{N}$ is given by*

$$f(z)=\sum_{p\in NC_s}z^{k(p)}$$

where $k:NC_s\to\mathbb{N}$ is the normalized length.

PROOF. With the notation $C_k = \#NC_s(k)$, the sum on the right is:

$$f(z) = \sum_k C_k z^k$$

For a given partition $p \in NC_s(k+1)$ we can consider the last s legs of the first block, and make cuts at right of them (see [13] for $s = 2$). This gives a decomposition of p into $s+1$ partitions in NC_s, and we get:

$$C_{k+1} = \sum_{\Sigma k_i = k} C_{k_0} \ldots C_{k_s}$$

By multiplying by z^{k+1} then summing over k we get that the generating series of these numbers satisfies $f - 1 = zf^{s+1}$. But this is the same as the equation $f = 1 + zf^{s+1}$ of the Stieltjes transform of π_{s1}, and we are done. □

Next, we have the following result:

THEOREM 10.6. *The Stieltjes transform of π_{st} with $s \in \mathbb{N}$ is given by:*

$$f(z) = \sum_{p \in NC_s} z^{k(p)} t^{b(p)}$$

where $k, b : NC_s \to \mathbb{N}$ are the normalized length, and the number of blocks.

PROOF. We denote by F_{kb} the number of partitions in $NC_s(k)$ having b blocks, and we set $F_{kb} = 0$ for other integer values of k, b. All sums will be over integer indices ≥ 0. With these notations, the sum on the right in the statement is:

$$f(z) = \sum_{kb} F_{kb} z^k t^b$$

The recurrence formula for the numbers C_k in the previous proof becomes:

$$\sum_b F_{k+1,b} = \sum_{\Sigma k_i = k} \sum_{b_i} F_{k_0 b_0} \ldots F_{k_s b_s}$$

In this formula, each term contributes to $F_{k+1,b}$ with $b = \Sigma b_i$, except for those of the form $F_{00} F_{k_1 b_1} \ldots F_{k_s b_s}$, which contribute to $F_{k+1,b+1}$. We get:

$$\begin{aligned} F_{k+1,b} &= \sum_{\Sigma k_i = k} \sum_{\Sigma b_i = b} F_{k_0 b_0} \ldots F_{k_s b_s} \\ &+ \sum_{\Sigma k_i = k} \sum_{\Sigma b_i = b-1} F_{k_1 b_1} \ldots F_{k_s b_s} \\ &- \sum_{\Sigma k_i = k} \sum_{\Sigma b_i = b} F_{k_1 b_1} \ldots F_{k_s b_s} \end{aligned}$$

This gives the following formula for the polynomials $P_k = \sum_b F_{kb} t^b$:

$$P_{k+1} = \sum_{\Sigma k_i = k} P_{k_0} \dots P_{k_s} + (t-1) \sum_{\Sigma k_i = k} P_{k_1} \dots P_{k_s}$$

In terms of $f = \sum_k P_k z^k$, we get the following equation:

$$f - 1 = zf^{s+1} + (t-1)zf^s$$

But this is the same as the equation $f = 1 + zf^s(f + t - 1)$ of the Stieltjes transform of π_{st}, and we are done. □

10c. Moments

The moments can be expressed in terms of generalized binomial coefficients. We recall that the coefficient corresponding to $\alpha \in \mathbb{R}$, $k \in \mathbb{N}$ is:

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}$$

We denote by $m_1, m_2, m_3, \dots$ the moments of a given probability measure. We have:

Theorem 10.7. *The moments of π_{s1} with $s > 0$ are the Fuss-Catalan numbers:*

$$m_k = \frac{1}{sk+1}\binom{sk+k}{k}$$

Proof. In the case $s \in \mathbb{N}$, we know that $m_k = \#NC_s(k)$. The formula in the statement follows by counting partitions.

In the general case $s > 0$, observe first that the Fuss-Catalan number in the statement is a polynomial in s:

$$\frac{1}{sk+1}\binom{sk+k}{k} = \frac{(sk+2)(sk+3)\dots(sk+k)}{k!}$$

Thus, in order to pass from the case $s \in \mathbb{N}$ to the case $s > 0$, it is enough to check that the k-th moment of π_{s1} is analytic in s. But this is clear from the equation $f = 1 + zf^{s+1}$ of the Stieltjes transform of π_{s1}. □

We have as well the following result:

Theorem 10.8. *The moments of π_{st}, $s > 0$ are the Fuss-Narayana numbers:*

$$m_k = \sum_{b=1}^{k} \frac{1}{b}\binom{k-1}{b-1}\binom{sk}{b-1} t^b$$

Proof. In the case $s \in \mathbb{N}$, we know from Theorem 4.4 that $m_k = \sum_b F_{kb} t^b$, where F_{kb} is the number of partitions in $NC_s(k)$ having b blocks. The formula in the statement follows by counting such partitions. This result can be extended to any $s > 0$, by using a complex variable argument, as in the proof of Theorem 10.7. □

In the case $s \notin \mathbb{N}$, the moments of π_{st} can be further expressed in terms of Gamma functions. In the case $s = 1/2$, the result is as follows:

Proposition 10.9. *The moments of $\pi_{1/2,1}$ are given by:*

$$\begin{aligned} m_{2p} &= \frac{1}{p+1}\binom{3p}{p} \\ m_{2p-1} &= \frac{2^{-4p+3}p}{(6p-1)(2p+1)} \cdot \frac{p!(6p)!}{(2p)!(2p)!(3p)!} \end{aligned}$$

Proof. The even moments of π_{st} with $s = n - 1/2$, $n \in \mathbb{N}$, are given by:

$$\begin{aligned} m_{2p} &= \frac{1}{(n-1/2)(2p)+1}\binom{(n+1/2)(2p)}{2p} \\ &= \frac{1}{(2n-1)p+1}\binom{(2n+1)p}{2p} \end{aligned}$$

With $n = 1$ we get the formula in the statement. Now for the odd moments, we can use here the following well-known identity:

$$\binom{m-1/2}{k} = \frac{4^{-k}}{k!} \cdot \frac{(2m)!}{m!} \cdot \frac{(m-k)!}{(2m-2k)!}$$

With $m = 2np + p - n$ and $k = 2p - 1$ we get:

$$\begin{aligned} m_{2p-1} &= \frac{1}{(n-1/2)(2p-1)+1}\binom{(n+1/2)(2p-1)}{2p-1} \\ &= \frac{2}{(2n-1)(2p-1)+2}\binom{(2np+p-n)-1/2}{2p-1} \\ &= \frac{2^{-4p+3}}{(2p-1)!} \cdot \frac{(4np+2p-2n)!}{(2np+p-n)!} \cdot \frac{(2np-p-n+1)!}{(4np-2p-2n+3)!} \end{aligned}$$

In particular with $n = 1$ we get:

$$\begin{aligned} m_{2p-1} &= \frac{2^{-4p+3}}{(2p-1)!} \cdot \frac{(6p-2)!}{(3p-1)!} \cdot \frac{p!}{(2p+1)!} \\ &= \frac{2^{-4p+3}(2p)}{(2p)!} \cdot \frac{(6p)!(3p)}{(3p)!(6p-1)6p} \cdot \frac{p!}{(2p)!(2p+1)} \end{aligned}$$

This gives the formula in the statement. $\square$

Let us investigate now the free additivity property of π_{st}, in analogy with the well-known free additivity property of the free Poisson laws π_{1t}. We first have:

THEOREM 10.10. *The free Bessel law π_{st} with $s \in \mathbb{N}$ is given by*

$$\pi_{st} = \text{law}\left(\sum_{k=1}^{s} w^k \alpha_k\right)^s$$

where $\alpha_1, \ldots, \alpha_s$ are free random variables, each of them following the free Poisson law of parameter t/s, and $w = e^{2\pi i/s}$.

PROOF. Given a random variable α and a complex number q, we have the following relations between the functional transforms of law(α) and law$(q\alpha)$:

$$\begin{aligned} f_{q\alpha}(z) = f_\alpha(qz) \quad &\Longrightarrow \quad G_{q\alpha}(z) = q^{-1} G_\alpha(q^{-1} z) \\ &\Longrightarrow \quad K_{q\alpha}(z) = q K_\alpha(qz) \\ &\Longrightarrow \quad R_{q\alpha}(z) = q R_\alpha(qz) \end{aligned}$$

Consider now the variable $\alpha = \sum w^k \alpha_k$. We have:

$$\begin{aligned} R_\alpha(z) &= \sum_{k=1}^{s} w^k R_{\alpha_k}(w^k z) \\ &= \sum_{k=1}^{s} w^k \cdot \frac{t}{s} \cdot \frac{1}{1 - w^k z} \end{aligned}$$

This gives the following formula:

$$R_\alpha(z) = t\left(\frac{1}{s} \sum_{k=1}^{s} \frac{w^k}{1 - w^k z}\right) = \frac{t z^{s-1}}{1 - z^s}$$

Consider now the formal measure $\tilde{\pi}_{st}$ having Stieltjes transform $\tilde{f}(z) = f(z^s)$, where f is the Stieltjes transform of π_{st}. The R transform of $\tilde{\pi}_{st}$ can be computed by using the equation of f obtained before, and we have:

$$\begin{aligned} f = 1 + z f^s (f + t - 1) \quad &\Longrightarrow \quad \tilde{f} = 1 + (z\tilde{f})^s (\tilde{f} + t - 1) \\ &\Longrightarrow \quad \xi \tilde{G} = 1 + \tilde{G}^s (\xi \tilde{G} + t - 1) \\ &\Longrightarrow \quad \tilde{K} z = 1 + z^s (\tilde{K} z + t - 1) \\ &\Longrightarrow \quad \tilde{R} z + 1 = 1 + z^s (\tilde{R} z + t) \\ &\Longrightarrow \quad \tilde{R}(z) = t z^{s-1} / (1 - z^s) \end{aligned}$$

Thus we have the equality of R transforms $\tilde{R} = R_\alpha$. In terms of measures we get $\tilde{\pi}_{st} = \text{law}(\alpha)$, hence $\pi_{st} = \text{law}(\alpha^s)$, and we are done. □

It is convenient to introduce the following related measures:

Definition 10.11. *The modified free Bessel laws* $\tilde{\pi}_{st}$ *with* $s\in\mathbb N$ *are given by*

$$\tilde{\pi}_{st}=\mathrm{law}\left(\sum_{k=1}^s w^k\alpha_k\right)$$

where $\alpha_1,\ldots,\alpha_s$ *are free random variables, each of them following the free Poisson law of parameter* t/s, *and* $w=e^{2\pi i/s}$.

We have the following result:

Theorem 10.12. *We have the Poisson limit type convergence*

$$\left(\left(1-\frac{1}{n}\right)\delta_0+\frac{1}{n}\rho\right)^{\boxplus n}\to\tilde{\pi}_{s1}$$

where ρ *is the uniform measure on the* s*-roots of unity.*

Proof. We compute first the R transform of the measure on the left:

$$\begin{aligned}\mu=\left(1-\frac{1}{n}\right)\delta_0+\frac{1}{n}\rho \quad&\Longrightarrow\quad f=\left(1-\frac{1}{n}\right)+\frac{1}{n}\cdot\frac{1}{1-z^s}\\ &\Longrightarrow\quad G(\xi)=\frac{1}{\xi}+\frac{1}{n}\cdot\frac{1}{\xi(\xi^s-1)}\\ &\Longrightarrow\quad (K^s-1)(zK-1)=\frac{1}{n}\\ &\Longrightarrow\quad \left(\left(R+\frac{1}{z}\right)^s-1\right)zR=\frac{1}{n}\end{aligned}$$

This shows that the R transform of $\mu^{\boxplus n}$ satisfies:

$$\left(\left(\frac{R}{n}+\frac{1}{z}\right)^s-1\right)z\,\frac{R}{n}=\frac{1}{n}$$

We multiply by n, then we take the limit $n\to\infty$. We get:

$$\left(\frac{1}{z^s}-1\right)zR=1$$

Thus in the limit $n\to\infty$ we have $R=z^{s-1}/(1-z^s)$, and we are done. □

10d. Product models

We discuss in what follows a random matrix model for the measures π_{st} with $s\in\mathbb N$. We restrict attention to the case $t=1$, since $\pi_{st}=\pi^{\boxtimes s-1}\boxtimes\pi^{\boxplus t}$ and therefore matrix models for π_{st} will follow from matrix models for $\pi^{\boxtimes s}$.

We first recall the definition of a Wishart matrix. Let $Y_1, Y_2, \ldots, Y_p$ be independent vectors in $\mathbb{C}^N$ with identical Gaussian distribution $N(0, \Sigma)$ and set:

$$W = Y_1 Y_1^* + \cdots + Y_p Y_p^*$$

Then the $N \times N$ Hermitian matrix W follows the complex Wishart distribution $W(N, p, \Sigma)$. We also can write $G^* = (Y_1, \ldots, Y_p)$, so that G is a $p \times N$ matrix and:

$$W = G^* G$$

When $\Sigma = \sigma^2 I_{N^2}$, then $G = (g_{ij})_{i=1\ldots p, j=1\ldots N}$ is a Gaussian random matrix with independent entries of variance σ^2, that is such that $\{Re(g_{ij}), Im(g_{ij})\}$ is a family of $2pN$ independent $N(0, \sigma^2/2)$ random variables. When $\Sigma = I_{N^2}/N$ and $\lim_{N\to\infty} p/N = t$, the limiting spectral distribution of W is the free Poisson law of parameter t, i.e. is the measure $\pi_{1t} = \pi^{\boxplus t}$. We have the following result:

Theorem 10.13. *Let $G_1, \ldots, G_s$ be a family of $N \times N$ independent matrices formed by independent centered Gaussian variables, of variance $1/N$. Then with*

$$M = G_1 \ldots G_s$$

the moments of the spectral distribution of (MM^*) converge to the corresponding moments of π_{s1}, as $N \to \infty$.

Proof. We proceed by induction. At $s = 1$ it is well-known that MM^* is a model for π_{11}. So, assume that the result holds for $s - 1 \geq 1$. We have:

$$\begin{aligned} & tr(MM^*)^k \\ = \ & tr(G_1 \ldots G_s G_s^* \ldots G_1^*)^k \\ = \ & tr(G_1 (G_2 \ldots G_s G_s^* \ldots G_1^* G_1)^{k-1} G_2 \ldots G_s G_s^* \ldots G_1^*) \end{aligned}$$

We can pass the first G_1 matrix to the right, we get:

$$\begin{aligned} & tr(MM^*)^k \\ = \ & tr((G_2 \ldots G_s G_s^* \ldots G_1^* G_1)^{k-1} G_2 \ldots G_s G_s^* \ldots G_1^* G_1) \\ = \ & tr(G_2 \ldots G_s G_s^* \ldots G_1^* G_1)^k \\ = \ & tr((G_2 \ldots G_s G_s^* \ldots G_2^*)(G_1^* G_1))^k \end{aligned}$$

We know that $G_1^* G_1$ is a Wishart matrix, hence is a model for π. Also, we know by the induction assumption that $G_2 \ldots G_s G_s^* \ldots G_2^*$ gives a matrix model for $\pi_{s-1,1}$. Since $G_1^* G_1$ and $G_2 \ldots G_s G_s^* \ldots G_2^*$ are asymptotically free, their product gives a matrix model for $\pi_{s-1,1} \boxtimes \pi_{11} = \pi_{s1}$, and we are done. □

We have as well the following result:

Theorem 10.14. *If W is a $W(sN,sN,\frac{1}{sN}I_{(sN)^2})$ complex Wishart matrix and*

$$D=\begin{pmatrix}1_N&0&&0\\0&w1_N&&0\\&&\ddots&\\0&0&&w^{s-1}1_N\end{pmatrix}$$

with $w=e^{2\pi i/s}$ then the moments of the mean empirical distribution of the eigenvalues of $(DW)^s$ converge to the corresponding moments of π_{s1}, as $N\to\infty$.

Proof. We use the following formula of Graczyk, Letac and Massam [63]:

$$E(Tr(DW)^K)=\sum_{\sigma\in S_K}\frac{M^{\gamma(\sigma^{-1}\pi)}}{M^K}\,r_\sigma(D)$$

Here W is a $W(M,M,\frac{1}{M}I_{N^2})$ complex Wishart matrix and D is a deterministic $M\times M$. As for the right term, this is as follows:

(1) π is the cycle $(1,\ldots,K)$.

(2) $\gamma(\sigma)$ is the number of disjoint cycles of σ.

(3) If we denote by $C(\sigma)$ the set of such cycles and for any cycle c, by $|c|$ its length, then:

$$r_\sigma(D)=\prod_{c\in C(\sigma)}Tr(D^{|c|})$$

In our situation we have $K=sk$ and $M=sN$, and we get:

$$E(\mathrm{Tr}(DW)^{sk})=\sum_{\sigma\in S_{sk}}\frac{(sN)^{\gamma(\sigma^{-1}\pi)}}{(sN)^{sk}}\,r_\sigma(D)$$

Now since D is uniformly formed by s-roots of unity, we have:

$$Tr(D^p)=\begin{cases}sN & \text{if } s|p\\ 0 & \text{if } s\not|p\end{cases}$$

Thus if we denote by S_{sk}^s the set of permutations $\sigma\in S_{sk}$ having the property that all the cycles of σ have length multiple of s, the above formula reads:

$$E(Tr(DW)^{sk})=\sum_{\sigma\in S_{sk}^s}\frac{(sN)^{\gamma(\sigma^{-1}\pi)}}{(sN)^{sk}}\,(sN)^{\gamma(\sigma)}$$

In terms of the normalized trace tr, we get:

$$E(Tr(DW)^{sk})=\sum_{\sigma\in S_{sk}^s}(sN)^{\gamma(\sigma^{-1}\pi)+\gamma(\sigma)-sk-1}$$

The exponent on the right, say L_σ, can be estimated by using the distance on the Cayley graph of S_{sk}:

$$\begin{aligned} L_\sigma &= \gamma(\sigma^{-1}\pi)+\gamma(\sigma)-sk-1\\ &= (sk-d(\sigma,\pi))+(sk-d(e,\sigma))-sk-1\\ &= sk-1-(d(e,\sigma)+d(\sigma,\pi))\\ &\leq sk-1-d(e,\pi)\\ &= 0 \end{aligned}$$

Now when taking the limit $N\to\infty$ in the above formula of $E(tr(DW)^{sk})$, the only terms that count are those coming from permutations $\sigma\in S_{sk}^s$ having the property $L_\sigma=0$, which each contribute with a 1 value. We get:

$$\begin{aligned} \lim_{N\to\infty} E(tr(DW)^{sk}) &= \#\{\sigma\in S_{sk}^s \mid L_\sigma=0\}\\ &= \#\{\sigma\in S_{sk}^s \mid d(e,\sigma)+d(\sigma,\pi)=d(e,\pi)\}\\ &= \#\{\sigma\in S_{sk}^s \mid \sigma\in[e,\pi]\} \end{aligned}$$

Now by using Biane's correspondence in [31], this is the same as the number of noncrossing partitions of $\{1,\ldots,sk\}$ having all blocks of size multiple of s. Thus we have reached to the sets $NC_s(k)$ from section 4, and we are done. □

As a consequence of the above random matrix formula, we have the following alternative free probabilistic approach to the free Bessel laws:

THEOREM 10.15. *The moments of the free Bessel law π_{s1} with $s\in\mathbb N$ coincide with those of*

$$\left(\sum_{k=1}^s w^k\alpha_k\right)^s$$

where $\alpha_1,\ldots,\alpha_s$ are free random variables, each of them following the free Poisson law of parameter $1/s$, and $w=e^{2\pi i/s}$.

PROOF. Let $G_1,\ldots,G_s$ be a family of independent $sN\times N$ matrices formed by independent, centered complex Gaussian variables, of variance $1/(sN)$. The following matrices $H_1,\ldots,H_s$ are as well complex Gaussian and independent:

$$H_k=\frac{1}{\sqrt s}\sum_{p=1}^s w^{kp}G_p$$

Thus the following matrix provides a model for $\Sigma w^k\alpha_k$:

$$\begin{aligned} M &= \sum_{k=1}^s w^k H_k H_k^* \\ &= \frac{1}{s}\sum_{k=1}^s\sum_{p=1}^s\sum_{q=1}^s w^{k+kp-kq} G_p G_q^* \\ &= \sum_{p=1}^s\sum_{q=1}^s \left(\frac{1}{s}\sum_{k=1}^s \left(w^{1+p-q}\right)^k\right) G_p G_q^* \\ &= G_1G_2^* + G_2G_3^* + \ldots + G_{s-1}G_s^* + G_sG_1^* \end{aligned}$$

This matrix can be written as:

$$\begin{aligned} M &= \begin{pmatrix} G_1 & G_2 & \ldots & G_{s-1} & G_s \end{pmatrix} \begin{pmatrix} G_2^* \\ G_3^* \\ \ldots \\ G_s^* \\ G_1^* \end{pmatrix} \\ &= \begin{pmatrix} G_1 & G_2 & \ldots & G_{s-1} & G_s \end{pmatrix} \begin{pmatrix} 0 & 1_N & 0 & \ldots & 0 \\ 0 & 0 & 1_N & \ldots & 0 \\ & & & \ddots & \\ 0 & 0 & 0 & \ldots & 1_N \\ 1_N & 0 & 0 & \ldots & 0 \end{pmatrix} \begin{pmatrix} G_1^* \\ G_2^* \\ \ldots \\ G_{s-1}^* \\ G_s^* \end{pmatrix} \\ &= GOG^* \end{aligned}$$

Here $G = (G_1 \ \ldots \ G_s)$ is the $sN \times sN$ Gaussian matrix obtained by concatenating $G_1, \ldots, G_s$, and O is the matrix in the middle. But this latter matrix is of the form $O = UDU^*$ with U unitary, so we have:

$$M = GUDU^*G^*$$

Now since GU is a Gaussian matrix, M has the same law as $M' = GDG^*$, and we get:

$$\begin{aligned} E\left(\left(\sum_{l=1}^s w^l\alpha_l\right)^{sk}\right) &= \lim_{N\to+\infty} E(\mathrm{tr}(M^{sk})) \\ &= \lim_{N\to+\infty} E(\mathrm{tr}(GDG^*)^{sk}) \\ &= \lim_{N\to+\infty} E(\mathrm{tr}(D(G^*G))^{sk}) \end{aligned}$$

Thus with $W = G^*G$ we get the result. □

Summarizing, we have applications to the random matrices, and random matrix models for all the 8 basic probability laws, appearing from limiting theorems.

10e. Exercises

CHAPTER 11

Free cumulants

11a. Partition basics

We will need some combinatorics, in relation with the partitions, including the Möbius inversion formula. Let us start with:

Definition 11.1. *Let $P(k)$ be the set of partitions of $\{1,\ldots,k\}$, and let $\pi,\sigma\in P(k)$.*

(1) *We write $\pi\leq\sigma$ if each block of π is contained in a block of σ.*
(2) *We let $\pi\vee\sigma\in P(k)$ be the partition obtained by superposing π,σ.*

As an illustration here, at $k=2$ we have $P(2)=\{||,\sqcap\}$. The order relation here is very simple, as follows:

$$||\leq\sqcap$$

At $k=3$ now, we have $P(3)=\{|||,\sqcap|,\sqcap\!\!\!\!\!\!|,|\sqcap,\sqcap\hskip-0.7mm\sqcap\}$. The order relation here is as follows:

$$|||\leq\sqcap|,\sqcap\!\!\!\!\!\!|,|\sqcap\leq\sqcap\hskip-0.7mm\sqcap$$

Observe also that we have:

$$\pi,\sigma\leq\pi\vee\sigma$$

In fact, $\pi\vee\sigma$ is the smallest partition with this property. Due to this fact, $\pi\vee\sigma$ is called supremum of π,σ.

We will need as well the following key definition:

Definition 11.2. *The Möbius function of any lattice, and so of P, is given by*

$$\mu(\pi,\sigma)=\begin{cases}1&\text{if }\pi=\sigma\\ -\sum_{\pi\leq\tau<\sigma}\mu(\pi,\tau)&\text{if }\pi<\sigma\\ 0&\text{if }\pi\not\leq\sigma\end{cases}$$

with the construction being performed by recurrence.

As an illustration here, let us go back to the set of 2-point partitions, $P(2)=\{||,\sqcap\}$. We have by definition:

$$\mu(||,||)=\mu(\sqcap,\sqcap)=1$$

Also, we know that we have $||<\sqcap$, with no intermediate partition in between, and so the above recurrence procedure gives:

$$\mu(||,\sqcap)=-\mu(||,||)=-1$$

Finally, we have $\sqcap \not\leq ||$, and so:

$$\mu(\sqcap, ||) = 0$$

Thus, as a conclusion, the Möbius matrix $M_{\pi\sigma} = \mu(\pi, \sigma)$ of the lattice $P(2) = \{||, \sqcap\}$ is as follows:

$$M = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

A similar computation gives the following formula, for $P(3) = \{|||, \sqcap|, \sqcap\!\!\!\!\!\;\;|\,, |\sqcap, \sqcap\!\sqcap\}$:

$$M = \begin{pmatrix} 1 & -1 & -1 & -1 & 2 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In general, the Möbius matrix of $P(k)$ looks a bit like the above matrices at $k = 2, 3$, being upper triangular, with 1 on the diagonal, and so on. We will be back to this.

Back to the general case now, the main interest in the Möbius function comes from the Möbius inversion formula, which states that the following happens:

$$f(\sigma) = \sum_{\pi \leq \sigma} g(\pi) \quad \Longrightarrow \quad g(\sigma) = \sum_{\pi \leq \sigma} \mu(\pi, \sigma) f(\pi)$$

This is something very useful. In linear algebra terms, the statement and proof of this formula are as follows:

Theorem 11.3. *The inverse of the adjacency matrix of P, given by*

$$A_{\pi\sigma} = \begin{cases} 1 & \text{if } \pi \leq \sigma \\ 0 & \text{if } \pi \not\leq \sigma \end{cases}$$

is the Möbius matrix of P, given by $M_{\pi\sigma} = \mu(\pi, \sigma)$.

Proof. This is well-known, coming for the fact that the adjacency matrix A is upper triangular. Indeed, when trying to invert A, we are led to the recurrence in Definition 11.2, and so to the Möbius matrix M, as stated. $\square$

As a first illustration, for $P(2)$ the formula $M = A^{-1}$ appears as follows:

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1}$$

At $k = 3$ now, we have:

$$P(3) = \{|||, \sqcap|, \sqcap\!\!\!\!\!\;\;|\,, |\sqcap, \sqcap\!\sqcap\}$$

Here the formula $M = A^{-1}$ reads:

$$\begin{pmatrix}1&-1&-1&-1&2\\0&1&0&0&-1\\0&0&1&0&-1\\0&0&0&1&-1\\0&0&0&0&1\end{pmatrix}=\begin{pmatrix}1&1&1&1&1\\0&1&0&0&1\\0&0&1&0&1\\0&0&0&1&1\\0&0&0&0&1\end{pmatrix}^{-1}$$

In general, the inversion formula is quite similar.

We refer to the literature for various applications of this formula.

11b. Free cumulants

With these ingredients in hand, let us go back to free probability.

We have the following key definition:

DEFINITION 11.4. *The classical and free cumulants $k_n(a), \kappa_n(a)$ of a noncommutative random variable a are constructed as follows:*

$$\log F_a(\xi) = \sum_n k_n(a)\xi^n$$

$$R_a(\xi) = \sum_n \kappa_n(a)\xi^n$$

More generally, we can define quantities $k_\pi(a), \kappa_\pi(a)$, depending on the partitions

$$\pi \in P(k)$$

by starting with $k_n(a), \kappa_n(a)$, and using multiplicativity over the blocks.

There are many examples here.

11c. Inversion formula

We have the following key result:

THEOREM 11.5. *We have the classical and free moment-cumulant formulae*

$$M_k(a) = \sum_{\pi \in P(k)} k_\pi(a)$$

$$M_k(a) = \sum_{\pi \in NC(k)} \kappa_\pi(a)$$

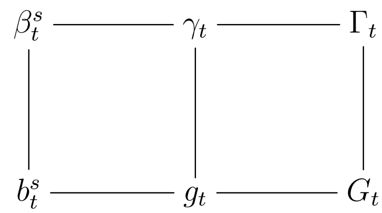
where $k_\pi(a), \kappa_\pi(a)$ are the generalized cumulants and free cumulants of a.

PROOF. This is standard, by using the formulae of F_a, R_a, or by doing some direct combinatorics, based on the Möbius inversion formula, from Theorem 11.3 above. □

11d. Basic examples

There are many examples for all this, and in particular, we have:

THEOREM 11.6. *Cumulants for the main classical and free probability measures:*



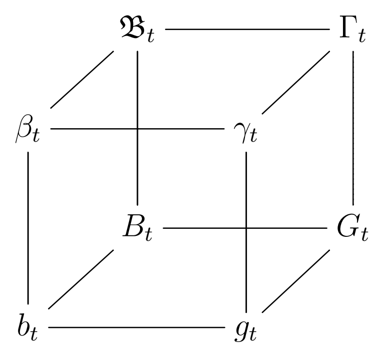
PROOF. This follows from the results that we have. Recall indeed that we have:

$$M_k = \sum_{\pi \in D(k)} t^{|\pi|}$$

We can process this formula by using Theorem 11.5, and we obtain the result. □

As a consequence, we have:

THEOREM 11.7. *Cumulants for the standard cube of probability measures:*



PROOF. This follows from Theorem 11.6 above. □

Summarizing, the cumulant theory provides us with a conceptual explanation for the moment results that we have so far, and for the classical/free correspondence.

11e. Exercises

CHAPTER 12

The bijection

12a. The bijection

We discuss in this section various aspects of the Bercovici-Pata bijection. With the cumulant theory from chapter 11 in hand, we can now formulate the following simple definition, making the connection between classical and free:

DEFINITION 12.1. *We say that a real probability measure*

$$m \in \mathcal{P}(\mathbb{R})$$

is the classical version of another probability measure, called its liberation

$$\mu \in \mathcal{P}(\mathbb{R})$$

when the classical cumulants of m coincide with the free cumulants of μ.

This fits with the main examples from the previous section.

In order to reach to a more advanced theory, let us formulate:

DEFINITION 12.2. *A convolution semigroup of measures*

$$\{m_t\}_{t>0} \quad : \quad m_s * m_t = m_{s+t}$$

is in Bercovici-Pata bijection with a free convolution semigroup of measures

$$\{\mu_t\}_{t>0} \quad : \quad \mu_s \boxplus \mu_t = \mu_{s+t}$$

when the classical cumulants of m_t coincide with the free cumulants of μ_t.

We have the following result:

THEOREM 12.3. *Main examples.*

PROOF. This follows from the results from the previous sections. □

There are many other explicit examples. We will be back to this.

12b. Algebraic results

At the level of the general theory, we first have:

THEOREM 12.4. *Algebraic results.*

PROOF. This is standard. □

12c. Analytic results

At the level of the general theory, we also have:

THEOREM 12.5. *Analytic results.*

PROOF. This is standard too, but more tricky. □

For further general results, we refer to the literature.

Back to the examples now, we have the following surprising result:

THEOREM 12.6. *The normal law is infinitely divisible.*

PROOF. This is tricky. □

The above result raises the following question: what is the classical analogue of the normal law? We refer to the literature for more on this.

12d. Meixner laws

Generally speaking, the BP bijection should be thought of as being something happening in the $N \to \infty$ limit. When $N \in \mathbb{N}$ is fixed the situation is more complicated, and we have here many correspondences, coming from quantum groups, or random matrices, which are not obviously related to the BP bijection. We first have the Meixner laws:

THEOREM 12.7. *Meixner and free Meixner laws.*

PROOF. This is standard. □

Next, we have the hyperspherical laws:

THEOREM 12.8. *Hyperspherical and free hyperspherical laws.*

PROOF. This is standard. □

We have as well the hypergeometric laws:

THEOREM 12.9. *Hypergeometric and free hypergeometric laws.*

PROOF. This is standard. □

We will discuss in the next section a common geometric approach to these laws, with the Meixner laws coming from tori, the hyperspherical laws coming from spheres and unitary groups, and the hypergeometric laws coming from reflection groups.

12e. Exercises

Part IV

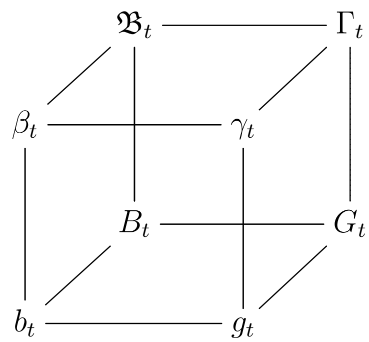
Matrix models

CHAPTER 13

Quantum groups

13a. Quantum groups

We have seen so far the foundations and basic results in free probability, with the onclusion that this is a "twin sister" to the usual classical probability. The relation between classical and free can be understood in many ways, both algebraic and analytic, but most relevant perhaps is the diagram formed by the laws appearing in the main classical and free limiting theorems. This is a familiar cube, as follows:



To be more precise, the x direction corresponds to discrete/continuous, the y direction corresponds to real/complex, and the z direction corresponds to classical/free. Missing from this cube are the Poisson and free Poisson laws p_t, π_t, which do not have complex analogues, and must be replaced by the classical and free real Bessel laws b_t, β_t.

In the remainder of this book we discuss more advanced aspects of free probability, either of philosophical nature, or in relation with physics, or both. The presentation will be often quite brief, but with full references given, to the relevant literature.

Let us start the discussion here with the following basic question, that we have avoided so far, throughout this book: we have been talking all the time about measured spaces X, but in practice, how to find such measured spaces X, in the real life?

The problem is that the measure in question cannot come out of "nowhere", and is usually a Haar measure. Thus, if we really want explicit examples of such spaces X, for our various needs, we must look into Lie groups, and use:

THEOREM 13.1. *The following happen:*

(1) *Any compact Lie group G has a uniform, or Haar measure, meaning a probability measure μ satisfying the following condition, for any $E \subset G$:*

$$\mu(gE) = \mu(Eg) = \mu(E)$$

(2) *More generally, any quotient space G/H of such compact Lie groups has a Haar measure, which must be by definition invariant under the action of G.*

PROOF. This is something well-known, the idea being as follows:

(1) We can indeed pick any probability measure ν on our group G, and then start convolving with itself. The more we convolve, the more ν becomes invariant, and so in the $n \to \infty$ limit we will obtain an invariant measure μ. This is the general idea, but in practice, in order for things to converge, we must use a Cesàro limit, as follows:

$$\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \nu^{*k}$$

The fact that this Cesàro limit converges indeed, and is left and right invariant under the action of G, is standard measure theory and functional analysis. We will be back to this, with some "quantum" generalizations of this, with detailed proofs at that time.

(2) Assume indeed that $X = G/H$ is an homogeneous space, coming from a closed subgroup $H \subset G$, as in the statement. The Haar measure on X appears then as the push-forward of the Haar measure of G, via the canonical quotient map:

$$\pi : G \to G/H$$

We can recover this as well via functional analysis. Indeed, at the level of algebras of functions, the above quotient map π produces an embedding, as follows:

$$i : C(X) \subset C(G)$$

Thus, we can define the Haar integration over X as being the restriction of the Haar integration over G, via this embedding i, and this gives the result. □

The above was quite brief, but we will come up with details in a moment, in a more general setting, that of the compact quantum groups.

Our idea in what follows will be that of finding "noncommutative" analogues of Theorem 13.1, providing us with explicit formulae of type $A = L^\infty(X)$, involving noncommutative von Neumann algebras A, and quantum measured spaces X.

As a first observation, in order to develop such a program, we must use the abstract C^*-algebra formalism explained in chapter 3 above, without any reference to the Hilbert space $H = L^2(X)$ that we want to construct.

Here is a theoretical statement, providing foundations:

THEOREM 13.2. *We can talk about compact quantum measured spaces, as follows:*

1. *The category of compact quantum measured spaces (X,μ) is the category of the C^*-algebras with faithful traces (A,φ), with the arrows reversed.*
2. *In the case where we have a non-faithful trace φ, we can still talk about the corresponding space (X,μ), by performing the GNS construction.*
3. *By taking the weak closure in the GNS representation, we obtain the von Neumann algebra $A''=L^\infty(X)$, in the previous general measured space sense.*
4. *In the particular case of the group algebras, all the group algebras $C^*_\pi(\Gamma)$ give rise to the same space, namely the usual abstract dual $G=\widehat{\Gamma}$.*

PROOF. All this follows from the GNS theorem, and from the other things that we already know, with the whole result itself being something rather philosophical. □

Let us go ahead now with our noncommutative geometry program. Let us start with the following key definition, due to Woronowicz:

DEFINITION 13.3. *A Woronowicz algebra is a C^*-algebra A, given with a unitary matrix $u\in M_N(A)$ whose coefficients generate A, such that the formulae*

$$\Delta(u_{ij})=\sum_k u_{ik}\otimes u_{kj}$$

$$\varepsilon(u_{ij})=\delta_{ij}$$

$$S(u_{ij})=u_{ji}^*$$

define morphisms of C^*-algebras $\Delta:A\to A\otimes A$, $\varepsilon:A\to\mathbb{C}$, $S:A\to A^{opp}$.

We say that A is cocommutative when $\Sigma\Delta=\Delta$, where $\Sigma(a\otimes b)=b\otimes a$ is the flip. We have the following result, which justifies the terminology and axioms:

PROPOSITION 13.4. *The following are Woronowicz algebras:*

1. *$C(G)$, with $G\subset U_N$ compact Lie group. Here the structural maps are:*

$$\begin{aligned}\Delta(\varphi)&=(g,h)\to\varphi(gh)\\ \varepsilon(\varphi)&=\varphi(1)\\ S(\varphi)&=g\to\varphi(g^{-1})\end{aligned}$$

2. *$C^*(\Gamma)$, with $F_N\to\Gamma$ finitely generated group. Here the structural maps are:*

$$\begin{aligned}\Delta(g)&=g\otimes g\\ \varepsilon(g)&=1\\ S(g)&=g^{-1}\end{aligned}$$

Moreover, we obtain in this way all the commutative/cocommutative algebras.

PROOF. In both cases, we have to exhibit a certain matrix u. For the first assertion, we can use the matrix $u=(u_{ij})$ formed by matrix coordinates of G, given by:

$$g=\begin{pmatrix}u_{11}(g)&\ldots&u_{1N}(g)\\ \vdots&&\vdots\\ u_{N1}(g)&\ldots&u_{NN}(g)\end{pmatrix}$$

For the second assertion, we can use the diagonal matrix formed by generators:

$$u=\begin{pmatrix}g_1&&0\\ &\ddots&\\ 0&&g_N\end{pmatrix}$$

Finally, the last assertion follows from the Gelfand theorem, in the commutative case, and in the cocommutative case, we will be back to this later. □

In view of Proposition 13.4, we can now formulate the following definition:

DEFINITION 13.5. *Given a Woronowicz algebra A, we formally write*

$$A=C(G)=C^*(\Gamma)$$

and call G compact quantum group, and Γ discrete quantum group.

When A is both commutative and cocommutative, G is a compact abelian group, Γ is a discrete abelian group, and these groups are dual to each other, $G=\widehat{\Gamma},\Gamma=\widehat{G}$.

In general, we still agree to write $G=\widehat{\Gamma},\Gamma=\widehat{G}$, but in a formal sense.

In general now, the structural maps Δ,ε,S have the following properties:

PROPOSITION 13.6. *Let (A,u) be a Woronowicz algebra.*

(1) *Δ,ε satisfy the usual axioms for a comultiplication and a counit, namely:*

$$\begin{aligned}(\Delta\otimes id)\Delta&=(id\otimes\Delta)\Delta\\ (\varepsilon\otimes id)\Delta&=(id\otimes\varepsilon)\Delta=id\end{aligned}$$

(2) *S satisfies the antipode axiom, on the $*$-subalgebra generated by entries of u:*

$$m(S\otimes id)\Delta=m(id\otimes S)\Delta=\varepsilon(.)1$$

(3) *In addition, the square of the antipode is the identity, $S^2=id$.*

PROOF. The two comultiplication axioms follow from:

$$\begin{aligned}(\Delta\otimes id)\Delta(u_{ij})&=(id\otimes\Delta)\Delta(u_{ij})=\sum_{kl}u_{ik}\otimes u_{kl}\otimes u_{lj}\\ (\varepsilon\otimes id)\Delta(u_{ij})&=(id\otimes\varepsilon)\Delta(u_{ij})=u_{ij}\end{aligned}$$

As for the antipode formulae, the verification here is similar. □

13b. Haar integration

Let us call corepresentation of A any unitary matrix $v\in M_n(A)$ satisfying the same conditions as those satisfied by u, namely:

$$\Delta(v_{ij})=\sum_k v_{ik}\otimes v_{kj}$$

$$\varepsilon(v_{ij})=\delta_{ij}$$

$$S(v_{ij})=v_{ji}^*$$

These corepresentations can be thought of as corresponding to the unitary representations of the underlying compact quantum group G. As main examples, we have $u=(u_{ij})$ itself, its conjugate $\bar{u}=(u_{ij}^*)$, as well as any tensor product between $u,\bar{u}$.

We have the following key result, due to Woronowicz [99]:

THEOREM 13.7. *Any Woronowicz algebra has a unique Haar integration functional,*

$$\left(\int_G\otimes id\right)\Delta=\left(id\otimes\int_G\right)\Delta=\int_G(.)1$$

which can be constructed by starting with any faithful positive form $\varphi\in A^*$, and setting

$$\int_G=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\varphi^{*k}$$

where $\phi*\psi=(\phi\otimes\psi)\Delta$. Moreover, for any corepresentation $v\in M_n(\mathbb{C})\otimes A$ we have

$$\left(id\otimes\int_G\right)v=P$$

where P is the orthogonal projection onto $Fix(v)=\{\xi\in\mathbb{C}^n|v\xi=\xi\}$.

PROOF. Following [99], this can be done in 3 steps, as follows:

(1) Given $\varphi\in A^*$, our claim is that the following limit converges, for any $a\in A$:

$$\int_\varphi a=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\varphi^{*k}(a)$$

Indeed, by linearity we can assume that a is the coefficient of certain corepresentation, $a=(\tau\otimes id)v$. But in this case, an elementary computation gives the following formula, with P_φ being the orthogonal projection onto the 1-eigenspace of $(id\otimes\varphi)v$:

$$\left(id\otimes\int_\varphi\right)v=P_\varphi$$

(2) Since $v\xi=\xi$ implies $[(id\otimes\varphi)v]\xi=\xi$, we have $P_\varphi\geq P$, where P is the orthogonal projection onto the following fixed point space:

$$Fix(v)=\left\{\xi\in\mathbb C^n\Big|v\xi=\xi\right\}$$

The point now is that when $\varphi\in A^*$ is faithful, by using a standard positivity trick, one can prove that we have $P_\varphi=P$. Assume indeed $P_\varphi\xi=\xi$, and let us set:

$$a=\sum_i\left(\sum_jv_{ij}\xi_j-\xi_i\right)\left(\sum_kv_{ik}\xi_k-\xi_i\right)^*$$

We must prove that we have $a=0$. Since v is biunitary, we have:

$$\begin{aligned}
a&=\sum_i\left(\sum_j\left(v_{ij}\xi_j-\frac{1}{N}\xi_i\right)\right)\left(\sum_k\left(v_{ik}^*\bar\xi_k-\frac{1}{N}\bar\xi_i\right)\right)\\
&=\sum_{ijk}v_{ij}v_{ik}^*\xi_j\bar\xi_k-\frac{1}{N}v_{ij}\xi_j\bar\xi_i-\frac{1}{N}v_{ik}^*\xi_i\bar\xi_k+\frac{1}{N^2}\xi_i\bar\xi_i\\
&=\sum_j|\xi_j|^2-\sum_{ij}v_{ij}\xi_j\bar\xi_i-\sum_{ik}v_{ik}^*\xi_i\bar\xi_k+\sum_i|\xi_i|^2\\
&=||\xi||^2-<v\xi,\xi>-\overline{<v\xi,\xi>}+||\xi||^2\\
&=2(||\xi||^2-Re(<v\xi,\xi>))
\end{aligned}$$

By using now our assumption $P_\varphi\xi=\xi$, we obtain from this:

$$\begin{aligned}
\varphi(a)&=2\varphi(||\xi||^2-Re(<v\xi,\xi>))\\
&=2(||\xi||^2-Re(<P_\varphi\xi,\xi>))\\
&=2(||\xi||^2-||\xi||^2)\\
&=0
\end{aligned}$$

Now since φ is faithful, this gives $a=0$, and so $v\xi=\xi$. Thus $\int_\varphi$ is independent of φ, and is given on coefficients $a=(\tau\otimes id)v$ by the following formula:

$$\left(id\otimes\int_\varphi\right)v=P$$

(3) With the above formula in hand, the left and right invariance of $\int_G=\int_\varphi$ is clear on coefficients, and so in general, and this gives all the assertions. See [99]. □

As a consequence of the above result, we can now develop, once again by following [99], the Peter-Weyl theory for the corepresentations of A.

Consider indeed the dense $*$-subalgebra $\mathcal{A} \subset A$ generated by the coefficients of the fundamental corepresentation u, and endow it with the following scalar product:

$$< a, b >= \int_G ab^*$$

We have then the following result:

THEOREM 13.8. *We have the following Peter-Weyl type results:*

1. *Any corepresentation decomposes as a sum of irreducible corepresentations.*
2. *Each irreducible corepresentation appears inside a certain $u^{\otimes k}$.*
3. *$\mathcal{A} = \bigoplus_{v \in Irr(A)} M_{\dim(v)}(\mathbb{C})$, the summands being pairwise orthogonal.*
4. *The characters of irreducible corepresentations form an orthonormal system.*

PROOF. This is something quite routine, and we refer here to [99]. □

We can now solve a problem that we left open before, namely:

PROPOSITION 13.9. *The cocommutative Woronowicz algebras appear as the quotients*

$$C^*(\Gamma) \to A \to C^*_{red}(\Gamma)$$

given by $A = C^*_\pi(\Gamma)$ with $\pi \otimes \pi \subset \pi$, with Γ being a discrete group.

PROOF. This follows from the Peter-Weyl theory. Observe that the assumption $\pi \otimes \pi \subset \pi$, which should be taken in a weak containment sense, is satisfied for the regular representation, as well as the universal representation. □

As another consequence of the above results, once again by basically following [99], we have the following result, dealing with functional analysis aspects:

THEOREM 13.10. *Let A_{full} be the enveloping C^*-algebra of $\mathcal{A}$, and let A_{red} be the quotient of A by the null ideal of the Haar integration. The following are then equivalent:*

1. *The Haar functional of A_{full} is faithful.*
2. *The projection map $A_{full} \to A_{red}$ is an isomorphism.*
3. *The counit map $\varepsilon : A_{full} \to \mathbb{C}$ factorizes through A_{red}.*
4. *We have $N \in \sigma(Re(\chi_u))$, the spectrum being taken inside A_{red}.*

If this is the case, we say that the underlying discrete quantum group Γ is amenable.

PROOF. This is well-known in the group dual case, $A = C^*(\Gamma)$, with Γ being a usual discrete group. In general, the result follows by adapting the group dual case proof:

(1) $\iff$ (2) This simply follows from the fact that the GNS construction for the algebra A_{full} with respect to the Haar functional produces the algebra A_{red}.

(2) $\iff$ (3) Here $\implies$ is trivial, and conversely, a counit map $\varepsilon : A_{red} \to \mathbb{C}$ produces an isomorphism $A_{red} \to A_{full}$, via a formula of type $(\varepsilon \otimes id)\Phi$. See [99].

(3) $\iff$ (4) Here $\implies$ is clear, coming from $\varepsilon(N - Re(\chi(u))) = 0$, and the converse can be proved by doing some functional analysis. Once again, we refer here to [99]. □

13c. Easiness, diagrams

Let us discuss now some "new" examples. Following Wang, we have:

Theorem 13.11. *The following universal algebras are Woronowicz algebras,*

$$\begin{aligned} C(O_N^+) &= C^*\left((u_{ij})_{i,j=1,\ldots,N}\Big|u=\bar{u},u^t=u^{-1}\right) \\ C(U_N^+) &= C^*\left((u_{ij})_{i,j=1,\ldots,N}\Big|u^*=u^{-1},u^t=\bar{u}^{-1}\right) \end{aligned}$$

so the underlying spaces O_N^+,U_N^+ and O_N^*,U_N^* are compact quantum groups.

Proof. The first assertion follows from the elementary fact that if a matrix $u=(u_{ij})$ is orthogonal or biunitary, then so must be the following matrices:

$$u_{ij}^\Delta=\sum_k u_{ik}\otimes u_{kj}$$

$$u_{ij}^\varepsilon=\delta_{ij}$$

$$u_{ij}^S=u_{ji}^*$$

Thus, we can define indeed morphisms Δ,ε,S as in Definition 13.12 above, by using the universality property of the algebras $C(O_N^+)$, $C(U_N^+)$. □

In order to integrate, we will need:

Theorem 13.12. *Assuming that $A=C(G)$ has Tannakian category $C=(C(k,l))$, the Haar integration over G is given by the Weingarten type formula*

$$\int_G u_{i_1j_1}^{e_1}\ldots u_{i_kj_k}^{e_k}=\sum_{\pi,\sigma\in D_k}\delta_\pi(i)\delta_\sigma(j)W_k(\pi,\sigma)$$

for any colored integer $k=e_1\ldots e_k$ and any multi-indices i,j, where D_k is a linear basis of $C(\emptyset,k)$, $\delta_\pi(i)=<\pi,e_{i_1}\otimes\ldots\otimes e_{i_k}>$, and $W_k=G_k^{-1}$, with $G_k(\pi,\sigma)=<\pi,\sigma>$.

Proof. We know from Theorem 13.7 above that the integrals in the statement form altogether the orthogonal projection P^k onto the following space:

$$Fix(u^{\otimes k})=span(D_k)$$

Consider now the following linear map, with $D_k=\{\xi_k\}$ being as in the statement:

$$E(x)=\sum_{\pi\in D_k}<x,\xi_\pi>\xi_\pi$$

By a standard linear algebra computation, it follows that we have $P=WE$, where W is the inverse on $span(T_\pi|\pi\in D_k)$ of the restriction of E. But this restriction is the linear map given by G_k, and so W is the linear map given by W_k, and this gives the result. □

The above result, basically known since Weyl, remains something quite theoretical. In order to apply it in practice, we must combine it with another piece of old and beautiful technology, namely with Brauer type results, of the following type:

$$C = span(D)$$

In order to explain all this, and directly in the quantum group setting, we need some combinatorial preliminaries. Let us begin with the following definition:

DEFINITION 13.13. *Let $P(k,l)$ be the set of partitions between an upper colored integer k, and a lower colored integer l. A collection of subsets*

$$D = \bigsqcup_{k,l} D(k,l)$$

with $D(k,l) \subset P(k,l)$ is called a category of partitions when it has the following properties:

1. *Stability under the horizontal concatenation, $(\pi,\sigma) \to [\pi\sigma]$.*
2. *Stability under vertical concatenation $(\pi,\sigma) \to [^\sigma_\pi]$, with matching middle symbols.*
3. *Stability under the upside-down turning $*$, with switching of colors, $\circ \leftrightarrow \bullet$.*
4. *Each set $P(k,k)$ contains the identity partition $||\ldots||$.*
5. *The sets $P(\emptyset,\circ\bullet)$ and $P(\emptyset,\bullet\circ)$ both contain the semicircle $\cap$.*

The relation with the Tannakian categories coming from:

PROPOSITION 13.14. *Each partition $\pi \in P(k,l)$ produces a linear map*

$$T_\pi : (\mathbb{C}^N)^{\otimes k} \to (\mathbb{C}^N)^{\otimes l}$$

given by the following formula, where $e_1,\ldots,e_N$ is the standard basis of $\mathbb{C}^N$,

$$T_\pi(e_{i_1}\otimes\ldots\otimes e_{i_k}) = \sum_{j_1\ldots j_l} \delta_\pi\begin{pmatrix} i_1 & \ldots & i_k\\ j_1 & \ldots & j_l\end{pmatrix} e_{j_1}\otimes\ldots\otimes e_{j_l}$$

and with the Kronecker type symbols $\delta_\pi \in \{0,1\}$ depending on whether the indices fit or not. The assignement $\pi \to T_\pi$ is categorical, in the sense that we have

$$T_\pi \otimes T_\sigma = T_{[\pi\sigma]}$$

$$T_\pi T_\sigma = N^{c(\pi,\sigma)} T_{[^\sigma_\pi]}$$

$$T_\pi^* = T_{\pi^*}$$

where $c(\pi,\sigma)$ are certain integers, coming from the erased components in the middle.

PROOF. This follows from the elementary computations, as follows:

(1) The concatenation axiom follows from the following computation:

$$\begin{aligned}
&(T_\pi\otimes T_\sigma)(e_{i_1}\otimes\ldots\otimes e_{i_p}\otimes e_{k_1}\otimes\ldots\otimes e_{k_r})\\
=&\sum_{j_1\ldots j_q}\sum_{l_1\ldots l_s}\delta_\pi\begin{pmatrix}i_1&\ldots&i_p\\j_1&\ldots&j_q\end{pmatrix}\delta_\sigma\begin{pmatrix}k_1&\ldots&k_r\\l_1&\ldots&l_s\end{pmatrix}e_{j_1}\otimes\ldots\otimes e_{j_q}\otimes e_{l_1}\otimes\ldots\otimes e_{l_s}\\
=&\sum_{j_1\ldots j_q}\sum_{l_1\ldots l_s}\delta_{[\pi\sigma]}\begin{pmatrix}i_1&\ldots&i_p&k_1&\ldots&k_r\\j_1&\ldots&j_q&l_1&\ldots&l_s\end{pmatrix}e_{j_1}\otimes\ldots\otimes e_{j_q}\otimes e_{l_1}\otimes\ldots\otimes e_{l_s}\\
=&T_{[\pi\sigma]}(e_{i_1}\otimes\ldots\otimes e_{i_p}\otimes e_{k_1}\otimes\ldots\otimes e_{k_r})
\end{aligned}$$

(2) The composition axiom follows from the following computation:

$$\begin{aligned}
&T_\pi T_\sigma(e_{i_1}\otimes\ldots\otimes e_{i_p})\\
=&\sum_{j_1\ldots j_q}\delta_\sigma\begin{pmatrix}i_1&\ldots&i_p\\j_1&\ldots&j_q\end{pmatrix}\sum_{k_1\ldots k_r}\delta_\pi\begin{pmatrix}j_1&\ldots&j_q\\k_1&\ldots&k_r\end{pmatrix}e_{k_1}\otimes\ldots\otimes e_{k_r}\\
=&\sum_{k_1\ldots k_r}N^{c(\pi,\sigma)}\delta_{[^\sigma_\pi]}\begin{pmatrix}i_1&\ldots&i_p\\k_1&\ldots&k_r\end{pmatrix}e_{k_1}\otimes\ldots\otimes e_{k_r}\\
=&N^{c(\pi,\sigma)}T_{[^\sigma_\pi]}(e_{i_1}\otimes\ldots\otimes e_{i_p})
\end{aligned}$$

(3) Finally, the involution axiom follows from the following computation:

$$\begin{aligned}
&T_\pi^*(e_{j_1}\otimes\ldots\otimes e_{j_q})\\
=&\sum_{i_1\ldots i_p}<T_\pi^*(e_{j_1}\otimes\ldots\otimes e_{j_q}),e_{i_1}\otimes\ldots\otimes e_{i_p}>e_{i_1}\otimes\ldots\otimes e_{i_p}\\
=&\sum_{i_1\ldots i_p}\delta_\pi\begin{pmatrix}i_1&\ldots&i_p\\j_1&\ldots&j_q\end{pmatrix}e_{i_1}\otimes\ldots\otimes e_{i_p}\\
=&T_{\pi^*}(e_{j_1}\otimes\ldots\otimes e_{j_q})
\end{aligned}$$

Summarizing, our correspondence is indeed categorical. □

In relation with the quantum groups, we have the following result, from [**27**]:

THEOREM 13.15. *Each category of partitions $D=(D(k,l))$ produces a family of compact quantum groups $G=(G_N)$, one for each $N\in\mathbb N$, via the following formula:*

$$Hom(u^{\otimes k},u^{\otimes l})=span\left(T_\pi\Big|\pi\in D(k,l)\right)$$

To be more precise, the spaces on the right form a Tannakian category, and so produce a certain closed subgroup $G_N\subset U_N^+$, via the Tannakian duality correspondence.

PROOF. This follows indeed from Woronowicz's Tannakian duality, in its "soft" form from [75]. Indeed, let us set:

$$C(k,l)=span\left(T_\pi\Big|\pi\in D(k,l)\right)$$

By using the axioms in Definition 13.13, and the categorical properties of the operation $\pi\to T_\pi$, from Proposition 13.14 above, we deduce that $C=(C(k,l))$ is a Tannakian category. Thus the Tannakian duality applies, and gives the result. □

All the above might seem a bit complicated, but we will see examples in a moment. Philosophically speaking, the quantum groups appearing as in Theorem 13.15 are the simplest from the perspective of Tannakian duality, so let us formulate:

DEFINITION 13.16. *A closed subgroup $G\subset U_N^+$ is called easy when we have*

$$Hom(u^{\otimes k},u^{\otimes l})=span\left(T_\pi\Big|\pi\in D(k,l)\right)$$

for any colored integers k,l, for a certain category of partitions $D\subset P$.

In other words, a compact quantum group is called easy when its Tannakian category appears in the simplest possible way: from a category of partitions. The terminology is quite natural, because Tannakian duality is basically our only non-trivial tool.

Before getting into examples, let us formulate, as a first result based on easiness, the following particularization of Theorem 13.12 above:

THEOREM 13.17. *For an easy quantum group $G\subset U_N^+$, coming from a category of partitions $D=(D(k,l))$, we have the Weingarten integration formula*

$$\int_G u_{i_1j_1}^{e_1}\ldots u_{i_kj_k}^{e_k}=\sum_{\pi,\sigma\in D(k)}\delta_\pi(i)\delta_\sigma(j)W_{kN}(\pi,\sigma)$$

for any $k=e_1\ldots e_k$ and any i,j, where $D(k)=D(\emptyset,k)$, δ are usual Kronecker symbols, and $W_{kN}=G_{kN}^{-1}$, with $G_{kN}(\pi,\sigma)=N^{|\pi\vee\sigma|}$, where $|.|$ is the number of blocks.

PROOF. With notations from Theorem 13.12, the Kronecker symbols are given by:

$$\begin{aligned}\delta_{\xi_\pi}(i)&=<\xi_\pi,e_{i_1}\otimes\ldots\otimes e_{i_k}>\\&=\delta_\pi(i_1,\ldots,i_k)\end{aligned}$$

The Gram matrix being as well the correct one, we obtain the result. □

Getting into examples, we have the following Brauer type result, classical and free:

Theorem 13.18. *The basic unitary quantum groups are all easy,*

$$\begin{array}{ccc} U_N & \longrightarrow & U_N^+ \\ \uparrow & & \uparrow \\ O_N & \longrightarrow & O_N^+ \end{array} \qquad\qquad \begin{array}{ccc} \mathcal{P}_2 & \longleftarrow & \mathcal{NC}_2 \\ \downarrow & & \downarrow \\ P_2 & \longleftarrow & NC_2 \end{array}$$

with on the right being the corresponding categories of partitions.

Proof. This is something crucial, the idea being as follows:

(1) The quantum group U_N^+ is defined via the following relations:

$$u^* = u^{-1} \quad , \quad u^t = \bar{u}^{-1}$$

Thus, the following operators must be in the associated Tannakian category C:

$$T_\pi \quad , \quad \pi = \overset{\cap}{\circ\bullet}$$

$$T_\pi \quad , \quad \pi = \overset{\cap}{\bullet\circ}$$

Thus the associated Tannakian category is $C = span(T_\pi|\pi \in D)$, with:

$$D = < \overset{\cap}{\circ\bullet}\,,\ \overset{\cap}{\bullet\circ} >= \mathcal{N}C_2$$

Thus, we are led to the conclusion in the statement.

(2) The quantum group $O_N^+ \subset U_N^+$ is defined by imposing the following relations:

$$u_{ij} = \bar{u}_{ij}$$

Thus, the following operators must be in the associated Tannakian category C:

$$T_\pi \quad , \quad \pi = \begin{smallmatrix}\circ\\|\\\bullet\end{smallmatrix}$$

$$T_\pi \quad , \quad \pi = \begin{smallmatrix}\bullet\\|\\\circ\end{smallmatrix}$$

Thus the associated Tannakian category is $C = span(T_\pi|\pi \in D)$, with:

$$D = < \mathcal{NC}_2, \begin{smallmatrix}\circ\\|\\\bullet\end{smallmatrix}, \begin{smallmatrix}\bullet\\|\\\circ\end{smallmatrix} >= NC_2$$

Thus, we are led to the conclusion in the statement.

(3) The group $U_N \subset U_N^+$ is defined via the following relations:

$$[u_{ij}, u_{kl}] = 0 \quad , \quad [u_{ij}, \bar{u}_{kl}] = 0$$

Thus, the following operators must be in the associated Tannakian category C:

$$T_\pi \quad , \quad \pi = \begin{smallmatrix}\circ\circ\\ \times\\ \circ\circ\end{smallmatrix}$$

$$T_\pi \quad , \quad \pi = \begin{smallmatrix}\circ\bullet\\ \times\\ \bullet\circ\end{smallmatrix}$$

Thus the associated Tannakian category is $C = span(T_\pi|\pi \in D)$, with:

$$D = < \mathcal{NC}_2, \begin{smallmatrix}\circ\circ\\ \times\\ \circ\circ\end{smallmatrix}, \begin{smallmatrix}\circ\bullet\\ \times\\ \bullet\circ\end{smallmatrix} >= \mathcal{P}_2$$

Thus, we are led to the conclusion in the statement.

(4) In order to deal now with O_N, we can simply use the following formula:

$$O_N = O_N^+ \cap U_N$$

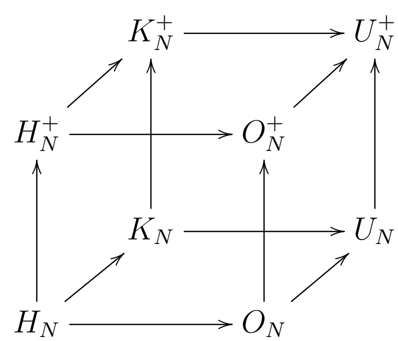
At the categorical level, this tells us that the associated Tannakian category is given by $C = span(T_\pi|\pi \in D)$, with:

$$D =< NC_2, \mathcal{P}_2 >= P_2$$

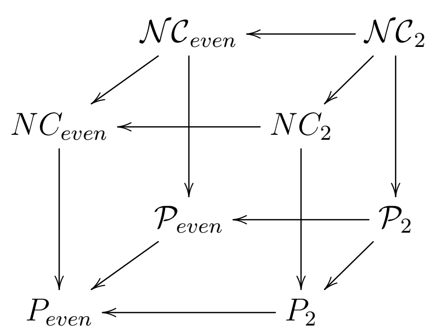
Thus, we are led to the conclusion in the statement. □

We have in fact the following more general result, once again of Brauer type, bringing into picture the corresponding quantum reflection groups as well:

Theorem 13.19. *We have quantum unitary and reflection groups as follows,*



which are all easy, the corresponding categories of partitions being:



Proof. We already have the quantum groups and easiness results on the right, from Theorem 13.18. Regarding the quantum groups on the left, their construction is quite standard, and the proof of their easiness property is standard as well. □

13d. Laws of characters

Let us discuss now probabilistic consequences. We will use:

Definition 13.20. *Associated to any closed subgrpup $G \subset U_N$ is its main character:*

$$\chi = \sum_{i=1}^{N} u_{ii}$$

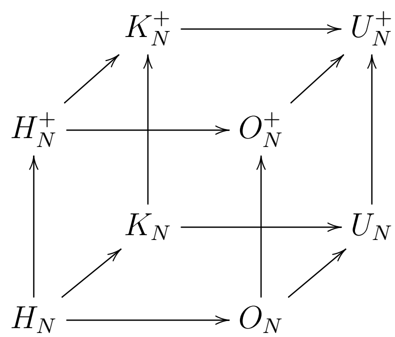
More generally, the truncated character of parameter $t \in (0,1]$ is given by:

$$\chi_t = \sum_{i=1}^{[tN]} u_{ii}$$

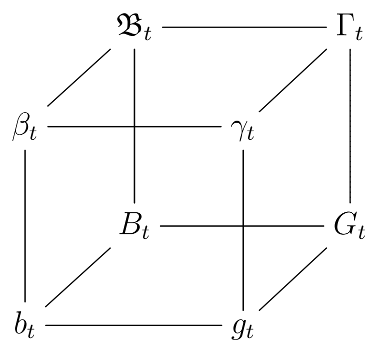
Here the notion of main character is the usual one, that we already met when doing Peter-Weyl theory, in the beginning of this section. As for the notion of truncated character, this is something more technical, inspired from random matrix theory.

We have the following character computations, for the basic quantum groups:

Theorem 13.21. *The truncated characters for the basic quantum groups, namely*



are with $N \to \infty$ the main laws in classical and free probability, namely:



Proof. This is something well-known, which can be done in 3 steps, as follows:

(1) We first need linear independence results for the vectors ξ_π associated to the partitions $\pi \in P(k)$. The simplest method here is that of computing the determinant of

the Gram matrix G_{kN} formed by these vectors, and the formula here, which is well-known and elementary, obtained by writing G_{kN} as a product of two triangular matrices, is:

$$\det(G_{kN}) = \prod_{\pi\in P(k)} \frac{N!}{(N-|\pi|)!}$$

(2) Let us discuss now the case $t = 1$, of the usual characters. For an easy group $G = (G_N)$, coming from a category of partitions $D = (D(k,l))$, it follows from (1) that the asymptotic moments of the main character are given by the following formula:

$$\lim_{N\to\infty}\int_{G_N}\chi^k = \#D(k)$$

But this gives the measures in the statement, at $t = 1$.

(3) In the general $t \in (0,1]$ case, the point is that the Gram and Weingarten matrices are asymptotically diagonal, in all cases under consideration, and this gives:

$$\lim_{N\to\infty}\int_{G_N}\chi_t^k = \sum_{\pi\in D(k)} t^{|\pi|}$$

But this leads to the laws in the statement, via results that we already know. □

13e. Exercises

CHAPTER 14

Free manifolds

14a. Free geometry

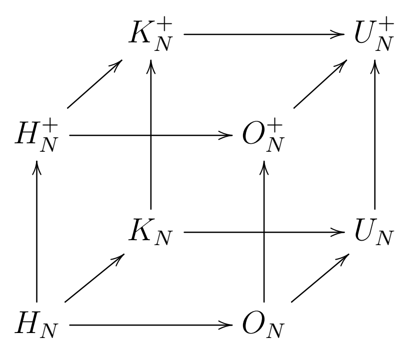
14b. Quotient spaces

14c. Spheres and tori

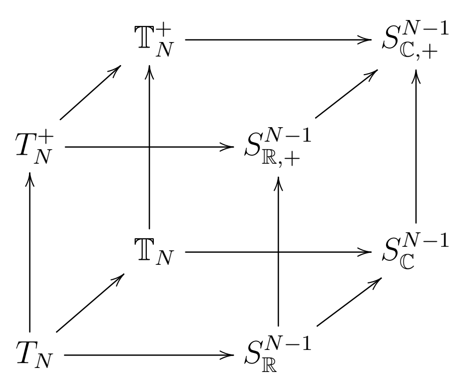
14d. Meixner laws

It is possible to obtain similar results for carefully chosen spaces of type $X = G/H$, and in connection with the questions raised in section 12 above, we have:

THEOREM 14.1. *The spheres and tori associated to the basic quantum groups,*



or rather to the corresponding "quantum geometries" are as follows:



These produce the Meixner, hyperspherical and hypergeometric laws.

Proof. This statement is something rather informal, and we will not attempt to explain everything in detail here. The idea is that it is possible to do geometry and probability on S,T, a bit in a same way as we did on U,K, using Weingarten type formulae. Looking at the laws of various "natural" variables leads then into the classical and free Meixner laws, hyperspherical laws and hypergeometric laws, as stated. □

Summarizing, we have a nice picture for quantum geometry and probability.

14e. Exercises

CHAPTER 15

De Finetti theorems

15a. Invariance questions

We discuss in this chapter probabilistic invariance questions with respect to the basic quantum rotation and reflection groups. Let us start by fixing some notations:

DEFINITION 15.1. *Given an easy quantum group $G_N \subset O_N^+$, we consider the free complex algebra on N variables*

$$\mathcal{P}_N = \mathbb{C} < t_1, \ldots, t_N >$$

and we construct a coaction of $C(G_N)$ on it, as follows:

$$\alpha_N : \mathcal{P}_N \to \mathcal{P}_N \otimes C(G_N)$$

$$t_j \to \sum t_i \otimes u_{ij}$$

Observe that α_N is indeed a coaction, in the sense that we have:

$$(id \otimes \Delta)\alpha_N = (\alpha_N \otimes id)\alpha_N$$

$$(id \otimes \varepsilon)\alpha_N = id$$

With this notion in hand, we can talk about invariant sequences, as follows:

DEFINITION 15.2. *Let $(x_1, \ldots, x_N)$ be a sequence of random variables in a noncommutative probability space (B, φ). We say that the sequence is G_N-invariant if the distribution functional $\varphi_x : \mathcal{P}_N \to \mathbb{C}$ is invariant under the coaction α_N,*

$$(\varphi_x \otimes id)\alpha_N(p) = \varphi_x(p)$$

for all $p \in \mathcal{P}_N$. More explicitly, the sequence $(x_1, \ldots, x_N)$ is G_N-invariant if

$$\varphi(x_{j_1} \ldots x_{j_k}) = \sum_{i_1 \ldots i_k} \varphi(x_{i_1} \ldots x_{i_k}) u_{i_1 j_1} \ldots u_{i_k j_k}$$

as an equality in $C(G_N)$, for any $k \in \mathbb{N}$ and any $1 \leq j_1, \ldots, j_k \leq N$.

In the classical case we recover in this way the usual invariance notion from probability, as shown by the following result:

Proposition 15.3. *In the classical group case, $G_N \subset O_N$, a sequence $(x_1, \ldots, x_N)$ is G_N-invariant in the above sense if and only if*

$$\varphi(x_{j_1} \ldots x_{j_k}) = \sum_{i_1 \ldots i_k} g_{i_1 j_1} \ldots g_{i_k j_k} \varphi(x_{i_1} \ldots x_{i_k})$$

for each $k \in \mathbb{N}$, $1 \leq j_1, \ldots, j_k \leq N$ and $g = (g_{ij}) \in G_N$, and this coincides with the usual notion of G_N-invariance for a sequence of classical random variables.

Proof. This follows indeed by evaluating both sides of the equation in Definition 15.2 at a given group element $g \in G_N$. □

In the classical De Finetti theorem, the independence occurs after conditioning. Likewise the free De Finetti theorem is a statement about freeness with amalgamation. Both these concepts may be expressed in terms of operator-valued probability theory, that we will recall now. First, we have the following definition:

Definition 15.4. *An operator-valued probability space consists of:*

1. *A unital algebra A.*
2. *A unital subalgebra $B \subset A$.*
3. *An expectation $E : A \to B$, satisfying $E(1) = 1$ and $E(b_1 a b_2) = b_1 E(a) b_2$.*

Given such an operator-valued probability space, the joint distribution of a family of variables $(x_i)_{i \in I}$ in the algebra A is by definition the following functional:

$$E_x : B < (t_i)_{i \in I} > \to B$$

$$P \to E(P(x))$$

We refer to [86] and related papers for more on all this, general results and examples, in relation with the operator-valued probability theory.

Next, we have the following key definition:

Definition 15.5. *Let $(x_i)_{i \in I}$ be a family of variables.*

1. *These variables are called independent if the algebra $< B, (x_i)_{i \in I} >$ is commutative, and if $i_1, \ldots, i_k \in I$ are distinct and $p_1, \ldots, p_k \in B < t >$ then:*
$$E(p_1(x_{i_1}) \ldots p_k(x_{i_k})) = E(p_1(x_{i_1})) \ldots E(p_k(x_{i_k}))$$
2. *These variables are called free if for any $i_1, \ldots, i_k \in I$ such that $i_l \neq i_{l+1}$, and any $p_1, \ldots, p_k \in B < t >$ such that $E(p_l(x_{i_l})) = 0$, we have:*
$$E(p_1(x_{i_1}) \ldots p_k(x_{i_k})) = 0$$

In order to deal with invariance questions, we will need the theory of classical and free cumulants, in this setting.

Let us start with the following definition:

Definition 15.6. *Let (A, B, E) be an operator-valued probability space.*

(1) *A B-functional is a N-linear map $\rho : A^N \to B$ such that:*
$$\rho(b_0 a_1 b_1, a_2 b_2 \ldots, a_N b_N) = b_0 \rho(a_1, b_1 a_2, \ldots, b_{N-1} a_N) b_N$$
Equivalently, ρ is a linear map $A^{\otimes_B N} \to B$, where the tensor product is taken with respect to the natural $B - B$ bimodule structure on A.

(2) *Suppose that B is commutative. For $k \in \mathbb{N}$ let $\rho^{(k)}$ be a B-functional. Given $\pi \in P(n)$, we define a B-functional $\rho^{(\pi)} : A^N \to B$ by the formula*
$$\rho^{(\pi)}(a_1, \ldots, a_N) = \prod_{V \in \pi} \rho(V)(a_1, \ldots, a_N)$$
where if $V = (i_1 < \ldots < i_s)$ is a block of π then:
$$\rho(V)(a_1, \ldots, a_N) = \rho_s(a_{i_1}, \ldots, a_{i_s})$$

If B is noncommutative, there is no natural order in which to compute the product appearing in the above formula for $\rho^{(\pi)}$. However, the nesting property of the noncrossing partitions allows for a natural definition of $\rho^{(\pi)}$ for $\pi \in NC(N)$, which we now recall:

Definition 15.7. *For $k \in \mathbb{N}$ let $\rho^{(k)} : A^k \to B$ be a B-functional. Given $\pi \in NC(N)$, define a B-functional $\rho^{(N)} : A^N \to B$ recursively as follows:*

(1) *If $\pi = 1_N$ is the partition having one block, define $\rho^{(\pi)} = \rho^{(N)}$.*

(2) *Otherwise, let $V = \{l + 1, \ldots, l + s\}$ be an interval of π and define:*
$$\rho^{(\pi)}(a_1, \ldots, a_N) = \rho^{(\pi - V)}(a_1, \ldots, a_l \rho^{(s)}(a_{l+1}, \ldots, a_{l+s}), a_{l+s+1}, \ldots, a_N)$$

As before, we refer to [80], [86] for more on all this.

Finally, we have the following definition:

Definition 15.8. *Let $(x_i)_{i \in I}$ be a family of random variables in A.*

(1) *The operator-valued classical cumulants $c_E^{(k)} : A^k \to B$ are the B-functionals defined by the following classical moment-cumulant formula:*
$$E(a_1 \ldots a_N) = \sum_{\pi \in P(N)} c_E^{(\pi)}(a_1, \ldots, a_N)$$

(2) *The operator-valued free cumulants $\kappa_E^{(k)} : A^k \to B$ are the B-functionals defined by the following free moment-cumulant formula:*
$$E(a_1, \ldots, a_N) = \sum_{\pi \in NC(N)} \kappa_E^{(\pi)}(a_1, \ldots, a_N)$$

We refer to [86] for more on the above notions.

We have the following result, which is well-known in the classical case, and which in the free case is due to Speicher [86]:

THEOREM 15.9. *Let $(x_i)_{i\in I}$ a family of random variables in A.*

(1) *If the algebra $< B, (x_i)_{i\in I} >$ is commutative, then $(x_i)_{i\in I}$ are conditionally independent given B if and only if when there are $1 \leq k, l \leq N$ such that $i_k \neq i_l$:*

$$c_E^{(N)}(b_0 x_{i_1} b_1, \ldots, x_{i_N} b_N) = 0$$

(2) *The variables $(x_i)_{i\in I}$ are free with amalgamation over B if and only if when there are $1 \leq k, l \leq N$ such that $i_k \neq i_l$:*

$$\kappa_E^{(N)}(b_0 x_{i_1} b_1, \ldots, x_{i_N} b_N) = 0$$

Note that the condition in (1) is equivalent to the statement that if $\pi \in P(N)$, then the following happens, unless $\pi \leq \ker i$:

$$c_E^{(\pi)}(b_0 x_{i_1} b_1, \ldots, x_{i_N} b_N) = 0$$

Similarly, the condition (2) is equivalent to the statement that if $\pi \in NC(N)$, then the following happens, unless $\pi \leq \ker i$:

$$\kappa_E^{(\pi)}(b_0 x_{i_1} b_1, \ldots, x_{i_N} b_N) = 0$$

Observe also that in the case $B = \mathbb{C}$ we obtain the usual notions of independence and freeness. As before, we refer to [80], [86] for more on all this.

Stronger characterizations of the joint distribution of $(x_i)_{i\in I}$ can be given by specifying what types of partitions may contribute nonzero cumulants.

To be more precise, we have here the following result:

THEOREM 15.10. *Let $(x_i)_{i\in I}$ be a family of random variables in A.*

(1) *Suppose that $< B, (x_i)_{i\in I} >$ is commutative. The B-valued joint distribution of $(x_i)_{i\in I}$ is independent for $D = P$ and independent centered Gaussian for $D = P_2$ if and only if, for any $\pi \in P(N)$, unless $\pi \in D(N)$ and $\pi \leq \ker i$:*

$$c_E^{(\pi)}(b_0 x_{i_1} b_1, \ldots, x_{i_N} b_N) = 0$$

(2) *The B-valued joint distribution of $(x_i)_{i\in I}$ is freely independent for $D = NC$ and freely independent centered semicircular for $D = NC_2$ if and only if, for any $\pi \in NC(N)$, unless $\pi \in D(N)$ and $\pi \leq \ker i$:*

$$\kappa_E^{(\pi)}(b_0 x_{i_1} b_1, \ldots, x_{i_N} b_N) = 0$$

PROOF. These results are well-known, coming from the definition of the classical and free cumulants, in the present setting, via some combinatorics. See [80], [86]. □

Finally, here is one more basic result that we will need:

THEOREM 15.11. *Let $(x_i)_{i\in I}$ be a family of random variables. Define the B-valued moment functionals $E^{(N)}$ by the following formula:*

$$E^{(N)}(a_1,\ldots,a_N)=E(a_1\ldots a_N)$$

(1) *If B is commutative, then for any $\sigma\in P(N)$ and $a_1,\ldots,a_N\in A$ we have:*

$$c_E^{(\sigma)}(a_1,\ldots,a_N)=\sum_{\pi\in P(N),\pi\leq\sigma}\mu_{P(N)}(\pi,\sigma)E^{(\pi)}(a_1,\ldots,a_N)$$

(2) *For any $\sigma\in NC(N)$ and $a_1,\ldots,a_N\in A$ we have:*

$$\kappa_E^{(\sigma)}(a_1,\ldots,a_N)=\sum_{\pi\in NC(N),\pi\leq\sigma}\mu_{NC(N)}(\pi,\sigma)E^{(\pi)}(a_1,\ldots,a_N)$$

PROOF. This follows indeed from the Möbius inversion formula. □

This was the general theory that we will need, in what follows. For further details on operator-valued free probability in general, we refer to [80], [86].

15b. Reverse De Finetti

With the above ingredients in hand, we can now investigate invariance questions for the sequences of random variables. We can first prove a reverse De Finetti theorem:

THEOREM 15.12. *Let $(x_1,\ldots,x_N)$ be a sequence in A.*

(1) *If $x_1,\ldots,x_N$ are freely independent and identically distributed with amalgamation over B, then the sequence is S_N^+-invariant.*
(2) *If $x_1,\ldots,x_N$ are freely independent and identically distributed with amalgamation over B, and have centered semicircular distributions with respect to E, then the sequence is O_N^+-invariant.*
(3) *If $< B,x_1,\ldots,x_N >$ is commutative and $x_1,\ldots,x_N$ are conditionally independent and identically distributed given B, then the sequence is S_N-invariant.*
(4) *If $< x_1,\ldots,x_N >$ is commutative and $x_1,\ldots,x_N$ are conditionally independent and identically distributed given B, and have centered Gaussian distributions with respect to E, then the sequence is O_N-invariant.*

PROOF. Suppose that the joint distribution of $(x_1,\ldots,x_N)$ satisfies one of the conditions in the statement, and let D be the partition family associated to the corresponding easy quantum group. We have then the following computation:

$$\begin{aligned}
&\sum_{i_1\dots i_k}\varphi(x_{i_1}\dots x_{i_k})u_{i_1j_1}\dots u_{i_kj_k}\\
=&\sum_{i_1\dots i_k}\varphi(E(x_{j_1}\dots x_{j_k}))u_{i_1j_1}\dots u_{i_kj_k}\\
=&\sum_{i_1\dots i_k}\sum_{\pi\leq\ker i}\varphi(\xi_E^{(\pi)}(x_1,\dots,x_1))u_{i_1j_1}\dots u_{i_kj_k}\\
=&\sum_{\pi\in D(k)}\varphi(\xi_E^{(\pi)}(x_1,\dots,x_1))\sum_{i,\pi\leq\ker i}u_{i_1j_1}\dots u_{i_kj_k}
\end{aligned}$$

Here ξ denotes the free and classical cumulants in the cases (1,2) and (3,4) respectively. It follows from a direct computation that if $\pi\in D(k)$ then:

$$\sum_{i,\pi\leq\ker i}u_{i_1j_1}\dots u_{i_kj_k}=\begin{cases}1&\text{if }\pi\leq\ker j\\0&\text{otherwise}\end{cases}$$

Applying this above, we find:

$$\begin{aligned}
\sum_{i_1,\dots,i_k}\varphi(x_{i_1}\dots x_{i_k})u_{i_1j_1}\dots u_{i_kj_k}&=\sum_{\pi\leq\ker j}\varphi(\xi_E^{(\pi)}(x_1,\dots,x_1))\\
&=\varphi(x_{j_1}\dots x_{j_k})
\end{aligned}$$

This completes the proof. □

We now begin the technical preparations for our approximation result. We will use the following simple fact:

PROPOSITION 15.13. *Suppose that a sequence $(x_1,\dots,x_N)$ is G_N-invariant. Then there is a right coaction*

$$\widetilde{\alpha}_N:M_N(\mathbb{C})\to M_N(\mathbb{C})\otimes L^\infty(G_N)$$

determined by the following formula:

$$\widetilde{\alpha}_N(p(x))=(ev_x\otimes\pi_N)\alpha_N(p)$$

Moreover, the fixed point algebra of $\widetilde{\alpha}_N$ is the G_N-invariant subalgebra B_N.

PROOF. This follows indeed after identifying the GNS representation of $\mathcal{P}_N$ for the state φ_x with the morphism $ev_x:\mathcal{P}_N\to M_n(\mathbb{C})$. □

There is a natural conditional expectation given by integrating the coaction $\widetilde{\alpha}_n$ with respect to the Haar state:

$$E_N:M_N(\mathbb{C})\to B_N$$

$$E_N(m)=\left(id\otimes\int\right)\widetilde{\alpha}_N(m)$$

The point now is that by using the Weingarten calculus, we can give a simple combinatorial formula for the moment functionals with respect to E_N, in the case where G_N is one of the easy quantum groups under consideration. To be more precise, we have:

PROPOSITION 15.14. *Suppose that $(x_1,\ldots,x_N)$ is G_N-invariant, and that either $G_N=O_N^+,S_N^+$, or that $G_N=O_N,S_N$ and $(x_1,\ldots,x_N)$ commute. We have then*

$$E_N^{(\pi)}(b_0x_1b_1,\ldots,x_1b_k)=\frac{1}{N^{|\pi|}}\sum_{\pi\leq\ker i}b_0x_{i_1}\ldots bx_{i_k}b_k$$

for any π in the partition category $D(k)$ for G_N, and any $b_0,\ldots,b_k\in B_N$.

PROOF. We prove this by induction on the number of blocks of π. First suppose that $\pi=1_k$ is the partition with only one block. Then:

$$\begin{aligned}E_N^{(1_k)}(b_0x_1b_1,\ldots,x_1b_k)&=E_N(b_0x_1\ldots x_1b_k)\\&=\sum_{i_1\ldots i_k}b_0x_{i_1}\ldots x_{i_k}b_k\int u_{i_11}\ldots u_{i_k1}\end{aligned}$$

Here we have used the fact that $b_0,\ldots,b_k$ are fixed by the coaction $\widetilde{\alpha}_n$. Applying the Weingarten integration formula, we have:

$$\begin{aligned}E_N(b_0x_1\ldots x_1b_k)&=\sum_{i_1\ldots i_k}b_0x_{i_1}\ldots x_{i_k}b_k\sum_{\pi\leq\ker i}\sum_{\sigma}W_{kN}(\pi,\sigma)\\&=\sum_{\pi\in D(k)}\left(\sum_{\sigma\in D(k)}W_{kN}(\pi,\sigma)\right)\sum_{i,\pi\leq\ker i}b_0x_{i_1}\ldots x_{i_k}b_k\end{aligned}$$

Now observe that for any $\sigma\in D(k)$ we have:

$$G_{kN}(\sigma,1_k)=N^{|\sigma\vee1_k|}=N$$

It follows that for any $\pi\in D(k)$, we have:

$$\begin{aligned}N\sum_{\sigma\in D(k)}W_{kN}(\pi,\sigma)&=\sum_{\sigma\in D(k)}W_{kN}(\pi,\sigma)G_{kN}(\sigma,1_k)\\&=\delta_{\pi1_k}\end{aligned}$$

Applying this above, we find, as desired:

$$\begin{aligned}E_N(b_0x_1\ldots x_1b_k)&=\sum_{\pi\in D(k)}\frac{1}{N}\delta_{\pi1_k}\sum_{i,\pi\leq\ker i}b_0x_{i_1}\ldots x_{i_k}b_k\\&=\frac{1}{N}\sum_{i=1}^Nb_0x_i\ldots x_ib_k\end{aligned}$$

If the condition (3) or (4) is satisfied, then the general case follows from:

$$E_N^{(\pi)}(b_0x_1b_1,\ldots,x_1b_k)=b_1\ldots b_k\prod_{V\in\pi}E_N(V)(x_1,\ldots,x_1)$$

The one thing we must check here is that if $\pi\in D(k)$ and V is a block of π with s elements, then $1_s\in D(s)$. This is easily verified, in each case.

Suppose now that the condition (1) or (2) is satisfied. Let $\pi\in D(k)$. Since π is noncrossing, π contains an interval $V=\{l+1,\ldots,l+s+1\}$. We then have:

$$E_N^{(\pi)}(b_0x_1b_1,\ldots,x_1b_k)=E_N^{(\pi-V)}(b_0x_1b_1,\ldots,E_n(x_1b_{l+1}\ldots x_1b_{l+s})x_1,\ldots,x_1b_k)$$

To apply induction, we must check that $\pi-V\in D(k-s)$ and $1_s\in D(s)$. Indeed, this is easily verified for NC,NC_2. Applying induction, we have:

$$\begin{aligned}
&E_N^{(\pi)}(b_0x_1b_1,\ldots,x_1b_k)\\
=\ &\frac{1}{N^{|\pi|-1}}\sum_{i,\pi-V\leq\ker i}b_0x_{i_1}\ldots b_l\left(E_n(x_1b_{l+1}\ldots x_1b_{l+s})\right)x_{i_{l+s}}\ldots x_{i_k}b_k\\
=\ &\frac{1}{N^{|\pi|-1}}\sum_{i,\pi-V\leq\ker i}b_0x_{i_1}\ldots b_l\left(\frac{1}{N}\sum_{i=1}^nx_ib_{l+1}\ldots bx_ib_{l+s}\right)x_{i_{l+s}}\ldots x_{i_k}b_k\\
=\ &\frac{1}{N^{|\pi|}}\sum_{i,\pi\leq\ker i}b_0x_{i_1}\ldots x_{i_k}b_k
\end{aligned}$$

This completes the proof. □

15c. Weingarten estimates

In order to advance, we will need some standard Weingarten estimates for our quantum groups, which have their own interest, and that we will discuss now. Regarding the symmetric group S_N, the situation here is very simple, as follows:

Proposition 15.15. *For S_N the Weingarten function is given by*

$$W_{kN}(\pi,\sigma)=\sum_{\tau\leq\pi\wedge\sigma}\mu(\tau,\pi)\mu(\tau,\sigma)\frac{(N-|\tau|)!}{N!}$$

and satisfies the folowing estimate,

$$W_{kN}(\pi,\sigma)=N^{-|\pi\wedge\sigma|}(\mu(\pi\wedge\sigma,\pi)\mu(\pi\wedge\sigma,\sigma)+O(N^{-1}))$$

with μ being the Möbius function of $P(k)$.

Proof. The first assertion follows from the Weingarten formula, namely:

$$\int_{S_N}u_{i_1j_1}\ldots u_{i_kj_k}=\sum_{\pi,\sigma\in P(k)}\delta_\pi(i)\delta_\sigma(j)W_{kN}(\pi,\sigma)$$

Indeed, in this formula the integrals on the left are known, from the explicit integration formula over S_N that we established in chapter 5, namely:

$$\int_{S_N} g_{i_1j_1}\ldots g_{i_kj_k}=\begin{cases}\frac{(N-|\ker i|)!}{N!}&\text{if }\ \ker i=\ker j\\ 0&\text{otherwise}\end{cases}$$

But this allows the computation of the right term, via the Möbius inversion formula, explained in chapter 2. As for the second assertion, this follows from the first one. □

The above result is something special, coming from the fact that the integration over S_N is something very simple. Regarding now the quantum group S_N^+, that we are particularly interested in, let us begin with some explicit computations. We first have the following simple result at $k=2,3$, directly in terms of the quantum group integrals:

Proposition 15.16. *At $k=2,3$ we have the following estimate:*

$$\int_{S_N^+} u_{i_1j_1}\ldots u_{i_kj_k}=\begin{cases}0&(\ker i\neq\ker j)\\ \simeq N^{-|\ker i|}&(\ker i=\ker j)\end{cases}$$

Proof. Since at $k\leq 3$ we have $NC(k)=P(k)$, the Weingarten integration formulae for S_N^+ and S_N coincide, and we obtain, by using the above formula for S_N:

$$\begin{aligned}\int_{S_N^+} u_{i_1j_1}\ldots u_{i_kj_k}&=\int_{S_N} u_{i_1j_1}\ldots u_{i_kj_k}\\ &=\delta_{\ker i,\ker j}\frac{(N-|\ker i|)!}{N!}\end{aligned}$$

But this gives the formula in the statement. □

In general now, the idea will be that of working out a "master estimate" for the Weingarten function, as above. Before starting here, let us record the formulae at $k=2,3$, which will be useful, as illustrations. At $k=2$, with indices $||,\sqcap$, and with the convention that $\approx$ means componentwise dominant term, we have:

$$W_{2N}\approx\begin{pmatrix}N^{-2}&-N^{-2}\\ -N^{-2}&N^{-1}\end{pmatrix}$$

At $k=3$ now, with indices $|||,|\sqcap,\sqcap|,\sqcap\!\!|,\sqcap\!\sqcap$, and same meaning for $\approx$, we have:

$$W_{3N}\approx\begin{pmatrix}N^{-3}&-N^{-3}&-N^{-3}&-N^{-3}&2N^{-3}\\ -N^{-3}&N^{-2}&N^{-3}&N^{-3}&-N^{-2}\\ -N^{-3}&N^{-3}&N^{-2}&N^{-3}&-N^{-2}\\ -N^{-3}&N^{-3}&N^{-3}&N^{-2}&-N^{-2}\\ 2N^{-3}&-N^{-2}&-N^{-2}&-N^{-2}&N^{-1}\end{pmatrix}$$

These formulae follow indeed from the plain formulae for W_{kN} at $k=2,3$ from [15], after rearranging the matrix indices as above. Observe in particular that we have the following formula, which will be of interest in what follows:

$$W_{3N}(|\sqcap,\sqcap|)\simeq N^{-3}$$

In order to deal now with the general case, let us start with:

PROPOSITION 15.17. *The following happen, regarding the partitions in $P(k)$:*

(1) $|\pi|+|\sigma|\leq|\pi\vee\sigma|+|\pi\wedge\sigma|$.
(2) $|\pi\vee\tau|+|\tau\vee\sigma|\leq|\pi\vee\sigma|+|\tau|$.
(3) $d(\pi,\sigma)=\frac{|\pi|+|\sigma|}{2}-|\pi\vee\sigma|$ *is a distance.*

PROOF. All this is well-known, the idea being as follows:

(1)This is well-known, coming from the fact that $P(k)$ is a semi-modular lattice.

(2) This follows from (1), as explained for instance in [22].

(3) This follows from (2), which says that the following holds:

$$\begin{aligned}&\frac{|\pi|+|\tau|}{2}-d(\pi,\tau)+\frac{|\tau|+|\sigma|}{2}-d(\tau,\sigma)\\ \leq\ &\frac{|\pi|+|\sigma|}{2}-d(\pi,\sigma)+|\tau|\end{aligned}$$

Thus, we obtain the triangle inequality:

$$d(\pi,\tau)+d(\tau,\sigma)\geq d(\pi,\sigma)$$

As for the other distance conditions, these are all clear. □

Actually in what follows we will only need (3) in the above statement. As a main result now regarding the Weingarten function, we have:

THEOREM 15.18. *The Weingarten function W_{kN} has a series expansion in N^{-1},*

$$W_{kN}(\pi,\sigma)=N^{|\pi\vee\sigma|-|\pi|-|\sigma|}\sum_{g=0}^\infty K_g(\pi,\sigma)N^{-g}$$

where the objects on the right are defined as follows:

(1) *A path from π to σ is a sequence $p=[\pi=\tau_0\neq\tau_1\neq\ldots\neq\tau_r=\sigma]$.*
(2) *The signature of such a path is $+$ when r is even, and $-$ when r is odd.*
(3) *The geodesicity defect of such a path is $g(p)=\sum_{i=1}^r d(\tau_{i-1},\tau_i)-d(\pi,\sigma)$.*
(4) *K_g counts the signed paths from π to σ, with geodesicity defect g.*

PROOF. The Gram matrix $G_{kN}(\pi,\sigma)=N^{|\pi\vee\sigma|}$ can be written as follows:

$$\begin{aligned}G_{kN}(\pi,\sigma)&=N^{\frac{|\pi|}{2}}N^{|\pi\vee\sigma|-\frac{|\pi|+|\sigma|}{2}}N^{\frac{|\sigma|}{2}}\\&=N^{\frac{|\pi|}{2}}N^{-d(\pi,\sigma)}N^{\frac{|\sigma|}{2}}\end{aligned}$$

Consider now the diagonal matrix $\Delta=diag(N^{\frac{|\pi|}{2}})$, and let us set as well:

$$H(\pi,\sigma)=\begin{cases}0&(\pi=\sigma)\\N^{-d(\pi,\sigma)}&(\pi\neq\sigma)\end{cases}$$

In terms of these two matrices, the above formula simply reads:

$$G_{kN}=\Delta(1+H)\Delta$$

Thus, the Weingarten matrix is given by the following formula:

$$W_{kN}=\Delta^{-1}(1+H)^{-1}\Delta^{-1}$$

In order to compute the inverse of $1+H$, consider the set $P_r(\pi,\sigma)$ of length r paths between π and σ. The powers of H are then given by:

$$\begin{aligned}H^r(\pi,\sigma)&=\sum_{p\in P_r(\pi,\sigma)}H(\tau_0,\tau_1)\ldots H(\tau_{r-1},\tau_r)\\&=\sum_{p\in P_r(\pi,\sigma)}N^{-d(\pi,\sigma)-g(p)}\end{aligned}$$

Thus by using the formula $(1+H)^{-1}=1-H+H^2-H^3+\ldots$ we obtain:

$$\begin{aligned}(1+H)^{-1}(\pi,\sigma)&=\sum_{r=0}^\infty(-1)^rH^r(\pi,\sigma)\\&=N^{-d(\pi,\sigma)}\sum_{r=0}^\infty\sum_{p\in P_r(\pi,\sigma)}(-1)^rN^{-g(p)}\end{aligned}$$

It follows that the Weingarten matrix is given by:

$$\begin{aligned}W_{kN}(\pi,\sigma)&=\Delta^{-1}(\pi)(1+H)^{-1}(\pi,\sigma)\Delta^{-1}(\sigma)\\&=N^{-\frac{|\pi|}{2}-\frac{|\sigma|}{2}-d(\pi,\sigma)}\sum_{r=0}^\infty\sum_{p\in P_r(\pi,\sigma)}(-1)^rN^{-g(p)}\\&=N^{|\pi\vee\sigma|-|\pi|-|\sigma|}\sum_{r=0}^\infty\sum_{p\in P_r(\pi,\sigma)}(-1)^rN^{-g(p)}\end{aligned}$$

Now by rearranging the various terms of the double sum according to their geodesicity defect $g=g(p)$, this gives the formula in the statement. □

As an illustration, we have the following explicit estimates:

THEOREM 15.19. *Consider an easy quantum group $G=(G_N)$, coming from a category of partitions $D=(D(k))$. For any $\pi\leq\sigma$ we have the estimate*

$$W_{kN}(\pi,\sigma)=N^{-|\pi|}(\mu(\pi,\sigma)+O(N^{-1}))$$

and for π,σ arbitrary we have

$$W_{kN}(\pi,\sigma)=O(N^{|\pi\vee\sigma|-|\pi|-|\sigma|})$$

with μ being the Möbius function of $D(k)$.

PROOF. We have two assertions here, the idea being as follows:

(1) The first estimate is clear from Theorem 15.18.

(2) In the case $\pi\leq\sigma$ it is known that K_0 coincides with the Möbius function of $NC(k)$, as explained for instance in [22], so we obtain once again from Theorem 15.18 the fine estimate as well, namely:

$$W_{kN}(\pi,\sigma)=N^{-|\pi|}(\mu(\pi,\sigma)+O(N^{-1}))\qquad\forall\pi\leq\sigma$$

Observe that, by symmetry of W_{kN}, we obtain as well the following estimate:

$$W_{kN}(\pi,\sigma)=N^{-|\sigma|}(\mu(\sigma,\pi)+O(N^{-1}))\qquad\forall\pi\geq\sigma$$

Thus, we are led to the conclusions in the statement. □

When π,σ are not comparable by $\leq$, the situation is quite unclear. The simplest example appears at $k=3$, where we have the following formula, which is elementary:

$$W_{3N}(|\sqcap,\sqcap|)\simeq N^{-3}$$

Observe that the exponent -3 is precisely the dominant one, because:

$$\Big||\sqcap\vee\sqcap|\Big|-\Big||\sqcap\Big|-\Big|\sqcap|\Big|=1-2-2=-3$$

As for the corresponding coefficient, $K_0(|\sqcap,\sqcap|)=1$, this is definitely not the Möbius function, which vanishes for partitions which are not comparable by $\leq$. According to Theorem 15.18, this is rather the number of signed geodesic paths from $|\sqcap$ to $\sqcap|$.

In relation to this, observe that geometrically, $NC(5)$ consists of the partitions $|\sqcap,\sqcap|$, ⊓̩, which form an equilateral triangle with edges worth 1, and then the partitions $|||$, $\sqcap\!\sqcap$, which are at distance 1 apart, and each at distance 1/2 from each of the vertices of the triangle. It is not exactly obvious how to recover the formula $K_0(|\sqcap,\sqcap|)=1$ from this.

Finally, we will need as well the following result:

PROPOSITION 15.20. *We have the following results:*

1. *If $D=NC,NC_2$, then $\mu_{D(k)}(\pi,\sigma)=\mu_{NC(k)}(\pi,\sigma)$.*
2. *If $D=P,P_2$ then $\mu_{D(k)}(\pi,\sigma)=\mu_{P(k)}(\pi,\sigma)$.*

Proof. Let $Q = NC, P$ according to the cases (1,2). It is easy to see in each case that $D(k)$ is closed under taking intervals in $Q(k)$, i.e., if $\pi_1, \pi_2 \in D(k)$, $\sigma \in Q(k)$ and $\pi_1 < \sigma < \pi_2$ then $\sigma \in D(k)$. The result now follows from the definition of the Möbius function. □

15d. De Finetti theorems

With all these combinatorial ingredients in hand, we are now prepared to prove an approximation result for finite sequences, from [22], as follows:

Theorem 15.21. *Suppose that $(x_1, \ldots, x_N)$ is G_N-invariant, and that $G_N = O_N^+, S_N^+$, or that $G_N = O_N, S_N$ and $(x_1, \ldots, x_N)$ commute. Let $(y_1, \ldots, y_N)$ be a sequence of B_N-valued random variables with B_N-valued joint distribution determined as follows:*

1. *$G = O^+$: Free semicircular, centered with same variance as x_1.*
2. *$G = S^+$: Freely independent, y_i has same distribution as x_1.*
3. *$G = O$: Independent Gaussian, centered with same variance as x_1.*
4. *$G = S$: Independent, y_i has same distribution as x_1.*

Then if $1 \leq j_1, \ldots, j_k \leq N$ and $b_0, \ldots, b_k \in B_N$, we have the following estimate,

$$||E_N(b_0 x_{j_1} \ldots x_{j_k} b_k) - E(b_0 y_{j_1} \ldots y_{j_k} b_k)|| \leq \frac{C_k(G)}{N} ||x_1||^k ||b_0|| \ldots ||b_k||$$

with $C_k(G)$ being a constant depending only on k and G.

Proof. First we note that it suffices to prove the result for N sufficiently large. We will assume that N is sufficiently large as for the Gram matrix G_{kN} to be invertible.

Let $1 \leq j_1, \ldots, j_k \leq N$ and $b_0, \ldots, b_k \in B_N$. We have:

$$\begin{aligned}
& E_N(b_0 x_{j_1} \ldots x_{j_k} b_k) \\
= & \sum_{i_1 \ldots i_k} b_0 x_{i_1} \ldots x_{i_k} b_k \int u_{i_1 j_1} \ldots u_{i_k j_k} \\
= & \sum_{i_1 \ldots i_k} b_0 x_{i_1} \ldots x_{i_k} b_k \sum_{\pi \leq \ker i} \sum_{\sigma \leq \ker j} W_{kN}(\pi, \sigma) \\
= & \sum_{\sigma \leq \ker j} \sum_{\pi} W_{kN}(\pi, \sigma) \sum_{i, \pi \leq \ker i} b_0 x_{i_1} \ldots x_{i_k} b_k
\end{aligned}$$

On the other hand, it follows from the assumptions on $(y_1, \ldots, y_N)$ and the various moment-cumulant formulae, that we have:

$$E(b_0 y_{j_1} \ldots y_{j_k} b_k) = \sum_{\sigma \leq \ker j} \xi_{E_N}^{(\sigma)}(b_0 x_1 b_1, \ldots, x_1 b_k)$$

Here, and in what follows, ξ are the relevant free or classical cumulants:

The right hand side can be expanded, via Möbius inversion, in terms of expectation functionals as follows, with π being a partition in NC, P according to the cases (1,2) or (3,4), and with $\pi \leq \sigma$ for some $\sigma \in D(k)$:

$$E_N^{(\pi)}(b_0 x_1 b_1, \ldots, x_1 b_k)$$

Now if $\pi \notin D(k)$ then we claim that this expectation functional is zero.

Indeed this is only possible if $D = NC_2, P_2$ and π has a block with an odd number of legs. But it is easy to see that in these cases x_1 has an even distribution with respect to E_N, and therefore we have, as claimed:

$$E_N^{(\pi)}(b_0 x_1 b_1, \ldots, x_1 b_k) = 0$$

This observation allows to to rewrite the above equation as:

$$E(b_0 y_{j_1} \ldots y_{j_k} b_k) = \sum_{\sigma \leq \ker j} \sum_{\pi \leq \sigma} \mu_{D(k)}(\pi, \sigma) E_N^{(\pi)}(b_0 x_1 b_1, \ldots, x_1 b_k)$$

We therefore obtain the following formula:

$$E(b_0 y_{j_1} \ldots y_{j_k} b_k) = \sum_{\sigma \leq \ker j} \sum_{\pi \leq \sigma} \mu_{D(k)}(\pi, \sigma) N^{-|\pi|} \sum_{i, \pi \leq \ker i} b_0 x_{i_1} \ldots x_{i_k} b_k$$

Comparing these two equations, we find that:

$$\begin{aligned} & E_N(b_0 x_{j_1} \ldots x_{j_k} b_k) - E(b_0 y_{j_1} \ldots y_{j_k} b_k) \\ = & \sum_{\sigma \leq \ker j} \sum_{\pi} \left(W_{kN}(\pi, \sigma) - \mu_{D(k)}(\pi, \sigma) N^{-|\pi|} \right) \sum_{i, \pi \leq \ker i} b_0 x_{i_1} \ldots x_{i_k} b_k \end{aligned}$$

Now since $x_1, \ldots, x_N$ are identically distributed with respect to the faithful state φ, it follows that these variables have the same norm. Therefore, for any $\pi \in D(k)$:

$$\left| \left| \sum_{i, \pi \leq \ker i} b_0 x_{i_1} \ldots x_{i_k} b_k \right| \right| \leq N^{|\pi|} ||x_1||^k ||b_0|| \ldots ||b_k||$$

Combining this with the former equation, we obtain:

$$\begin{aligned} & ||E_N(b_0 x_{j_1} \ldots x_{j_k} b_k) - E(b_0 y_{j_1} \ldots y_{j_k} b_k)|| \\ \leq & \sum_{\sigma \leq \ker j} \sum_{\pi} \left| W_{kN}(\pi, \sigma) N^{|\pi|} - \mu_{D(k)}(\pi, \sigma) \right| ||x_1||^k ||b_0|| \ldots ||b_k|| \end{aligned}$$

Let us set now:

$$C_k(G) = \sup_{N \in \mathbb{N}} N \times \sum_{\sigma, \pi \in D(k)} \left| W_{kN}(\pi, \sigma) N^{|\pi|} - \mu_{D(k)}(\pi, \sigma) \right|$$

But this is finite by our main estimate, which completes the proof. □

We will make use of the inclusions $G_N \subset G_M$ for $N < M$, which correspond to the Hopf algebra morphisms $\omega_{N,M} : C(G_M) \to C(G_N)$ determined by:

$$\omega_{N,M}(u_{ij}) = \begin{cases} u_{ij} & \text{if } 1 \leq i,j \leq N \\ \delta_{ij} & \text{if } \max(i,j) > N \end{cases}$$

We begin by extending the notion of G_N-invariance to infinite sequences:

Definition 15.22. *Let $(x_i)_{i\in\mathbb{N}}$ be a sequence in a noncommutative probability space (A,φ). We say that $(x_i)_{i\in\mathbb{N}}$ is G-invariant if*

$$(x_1,\ldots,x_N)$$

is G_N-invariant for each $N \in \mathbb{N}$.

In other words, the condition is that the joint distribution functional of $(x_1,\ldots,x_N)$ should be invariant under the following action, for each $n \in \mathbb{N}$:

$$\alpha_N : \mathcal{P}_N \to \mathcal{P}_N \otimes C(G_N)$$

It will be convenient to extend these actions to $\mathcal{P}_\infty = \mathbb{C} < t_i | i \in \mathbb{N} >$, by defining $\beta_N : \mathcal{P}_\infty \to \mathcal{P}_\infty \otimes C(G_N)$ to be the unique unital morphism such that:

$$\beta_N(t_j) = \begin{cases} \sum_{i=1}^N t_i \otimes u_{ij} & \text{if } 1 \leq j \leq N \\ t_j \otimes 1 & \text{if } j > N \end{cases}$$

It is clear that β_N is an action of G_N. Also, we have the following relations, where $\iota_N : \mathcal{P}_N \to \mathcal{P}_\infty$ is the natural inclusion:

$$(id \otimes \omega_{N,M})\beta_M = \beta_N$$

$$(\iota_N \otimes id)\alpha_N = \beta_N \iota_N$$

By using these compatibilities, we have the following result:

Proposition 15.23. *A a sequence $(x_i)_{i\in\mathbb{N}}$ is G-invariant if and only if the joint distribution functional $\varphi_x : \mathcal{P}_\infty \to \mathbb{C}$ is invariant under β_N for each $N \in \mathbb{N}$.*

Proof. This is clear indeed from the above discussion. □

In what follows $(x_i)_{i\in\mathbb{N}}$ will be a sequence of self-adjoint random variables in a von Neumann algebra (M,φ). We will assume that M is generated by $(x_i)_{i\in\mathbb{N}}$. We denote by $L^2(M,\varphi)$ the corresponding GNS Hilbert space, with inner product as follows:

$$< m_1, m_2 >= \varphi(m_1 m_2^*)$$

The strong topology on M will be taken by definition with respect to the representation on $L^2(M,\varphi)$. We let $\mathcal{P}_\infty^{\beta_N}$ be the fixed point algebra of the action β_N, and we set:

$$B_N = \left\{ p(x) \middle| p \in \mathcal{P}_\infty^{\beta_N} \right\}''$$

We have then the following formula:

$$(id \otimes \omega_{N,N+1})\beta_{N+1} = \beta_N$$

Thus we have an inclusion as follows, for any $n \geq 1$:

$$B_{N+1} \subset B_N$$

We then define the G-invariant subalgebra by:

$$B = \bigcap_{N \geq 1} B_N$$

With these conventions, we have the following result:

Proposition 15.24. *If an infinite sequence $(x_i)_{i\in\mathbb{N}}$ is G-invariant, then for each $N \in \mathbb{N}$ there is a right coaction*

$$\widetilde{\beta}_N : M \to M \otimes L^\infty(G_N)$$

determined by the following formula, for any $p \in \mathcal{P}_\infty$:

$$\widetilde{\beta}_N(p(x)) = (ev_x \otimes \pi_N)\beta_N(p)$$

The fixed point algebra of $\widetilde{\beta}_N$ is then B_N.

Proof. This is indeed clear from definitions. □

We have as well the following result, which is clear as well:

Proposition 15.25. *In the above context, for each $N \in \mathbb{N}$ there is then a φ-preserving conditional expectation $E_N : M \to B_N$ given by integrating the action $\widetilde{\beta}_N$:*

$$E_N(m) = \left(id \otimes \int\right)\widetilde{\beta}_N(m)$$

By taking the limit as $N \to \infty$, we obtain a φ-preserving conditional expectation onto the G-invariant subalgebra.

Proof. Once again, this is clear from definitions. □

Next, we have the following result:

Proposition 15.26. *Suppose that $(x_i)_{i\in\mathbb{N}}$ is G-invariant. Then:*

1. *For any $m \in M$, the sequence $E_N(m)$ converges in 2-norm and with respect to the strong topology to a limit $E(m) \in B$.*
2. *E is a φ-preserving conditional expectation of M onto B.*
3. *For $\pi \in NC(k)$ and $m_1, \ldots, m_k \in M$ we have, with strong convergence:*

$$E^{(\pi)}(m_1 \otimes \ldots \otimes m_k) = \lim_{n\to\infty} E_N^{(\pi)}(m_1 \otimes \ldots \otimes m_k)$$

Proof. This is again clear from definitions. Note that (1) is just a simple noncommutative reversed martingale convergence theorem. □

We are now prepared to state and prove the main theorem, from [22]:

THEOREM 15.27. *Let $(x_i)_{i\in\mathbb{N}}$ be a G-invariant sequence of self-adjoint random variables in (M,φ), and assume that $M=<(x_i)_{i\in\mathbb{N}}>$. Then there is a subalgebra $B\subset M$ and a φ-preserving conditional expectation $E:M\to B$ such that:*

1. *If $G=(S_N)$, then $(x_i)_{i\in\mathbb{N}}$ are conditionally independent and identically distributed given B.*
2. *If $G=(S_N^+)$, then $(x_i)_{i\in\mathbb{N}}$ are freely independent and identically distributed with amalgamation over B.*
3. *If $G=(O_N)$, then $(x_i)_{i\in\mathbb{N}}$ are conditionally independent, and have Gaussian distributions with mean zero and common variance, given B.*
4. *If $G=(O_N^+)$, then $(x_i)_{i\in\mathbb{N}}$ form a B-valued free semicircular family with mean zero and common variance.*

PROOF. Let $j_1,\ldots,j_k\in\mathbb{N}$ and $b_0,\ldots,b_k\in B$. We have:

$$\begin{aligned}
E(b_0x_{j_1}\ldots x_{j_k}b_k) &= \lim_{N\to\infty}E_n(b_0x_{j_1}\ldots x_{j_k}b_k)\\
&= \lim_{N\to\infty}\sum_{\sigma\leq\ker j}\sum_{\pi}W_{kN}(\pi,\sigma)\sum_{i,\pi\leq\ker i}b_0x_{i_1}\ldots x_{i_k}b_k\\
&= \lim_{N\to\infty}\sum_{\sigma\leq\ker j}\sum_{\pi\leq\sigma}\mu_{D(k)}(\pi,\sigma)N^{-|\pi|}\sum_{i,\pi\leq\ker i}b_0x_{i_1}\ldots x_{i_k}b_k
\end{aligned}$$

Let us recall now from the above that we have the following compatibility formula, where $\widetilde{\iota}_N:W^*(x_1,\ldots,x_N)\to M$ is the obvious inclusion, and $\widetilde{\alpha}_N$ is as before:

$$(\widetilde{\iota}_N\otimes id)\widetilde{\alpha}_N=\widetilde{\beta}_N\widetilde{\iota}_N$$

By using this, and the above cumulant results, we have:

$$E(b_0x_{j_1}\ldots x_{j_k}b_k)=\lim_{N\to\infty}\sum_{\sigma\leq\ker j}\sum_{\pi\leq\sigma}\mu_{D(k)}(\pi,\sigma)E_N^{(\pi)}(b_0x_1b_1,\ldots,x_1b_k)$$

We therefore obtain the following formula:

$$E(b_0x_{j_1}\ldots x_{j_k}b_k)=\sum_{\sigma\leq\ker j}\sum_{\pi\leq\sigma}\mu_{D(k)}(\pi,\sigma)E^{(\pi)}(b_0x_1b_1,\ldots,x_1b_k)$$

We can replace the sum of expectation functionals by cumulants to obtain:

$$E(b_0x_{j_1}\ldots x_{j_k}b_k)=\sum_{\sigma\leq\ker j}\xi_E^{(\sigma)}(b_0x_1b_1,\ldots,x_1b_k)$$

Here and in what follows ξ denotes the relevant free or classical cumulants, depending on the quantum group that we are dealing with, free or classical.

Now since the cumulants are determined by the moment-cumulant formulae, we find that we have the following formula:

$$\xi_E^{(\sigma)}(b_0 x_{j_1} b_1, \ldots, x_{j_k} b_k) = \begin{cases} \xi_E^{(\sigma)}(b_0 x_1 b_1, \ldots, x_1 b_k) & \text{if } \sigma \in D(k) \text{ and } \sigma \leq \ker j \\ 0 & \text{otherwise} \end{cases}$$

The result then follows from the characterizations of these joint distributions in terms of cumulants. □

We refer to [**22**] and related papers for more on the above.

15e. Exercises

CHAPTER 16

Wishart matrices

Bibliography

[1] N. Alexeev, F. Götze and A. Tikhomirov, On the asymptotic distribution of singular values of powers of random matrices, *Lithuanian Math. Journal* **50** (2010), 121–132.

[2] G.W. Anderson, A. Guionnet and O. Zeitouni, An introduction to random matrices, Cambridge Univ. Press (2010).

[3] O. Arizmendi, I. Nechita and C. Vargas, On the asymptotic distribution of block-modified random matrices, *J. Math. Phys.* **57** (2016), 1–27.

[4] G. Aubrun, Partial transposition of random states and non-centered semicircular distributions, *Random Matrices Theory Appl.* **1** (2012), 125–145.

[5] T. Banica, On the polar decomposition of circular variables, *Integral Equations Operator Theory* **24** (1996), 372–377.

[6] T. Banica, The free unitary compact quantum group, *Comm. Math. Phys.* **190** (1997), 143–172.

[7] T. Banica, Symmetries of a generic coaction, *Math. Ann.* **314** (1999), 763–780.

[8] T. Banica, Quantum automorphism groups of homogeneous graphs, *J. Funct. Anal.* **224** (2005), 243–280.

[9] T. Banica, Liberations and twists of real and complex spheres, *J. Geom. Phys.* **96** (2015), 1–25.

[10] T. Banica, Liberation theory for noncommutative homogeneous spaces, *Ann. Fac. Sci. Toulouse Math.* **26** (2017), 127–156.

[11] T. Banica, Block-modified Wishart matrices: the easy case, *Indiana Univ. Math. J.* **69** (2020), 1–34.

[12] T. Banica, S.T. Belinschi, M. Capitaine and B. Collins, Free Bessel laws, *Canad. J. Math.* **63** (2011), 3–37.

[13] T. Banica, J. Bichon and B. Collins, The hyperoctahedral quantum group, *J. Ramanujan Math. Soc.* **22** (2007), 345–384.

[14] T. Banica, J. Bichon and S. Curran, Quantum automorphisms of twisted group algebras and free hypergeometric laws, *Proc. Amer. Math. Soc.* **139** (2011), 3961–3971.

[15] T. Banica and B. Collins, Integration over compact quantum groups, *Publ. Res. Inst. Math. Sci.* **43** (2007), 277–302.

[16] T. Banica and B. Collins, Integration over quantum permutation groups, *J. Funct. Anal.* **242** (2007), 641–657.

[17] T. Banica and B. Collins, Integration over the Pauli quantum group, *J. Geom. Phys.* **58** (2008), 942–961.

[18] T. Banica, B. Collins and P. Zinn-Justin, Spectral analysis of the free orthogonal matrix, *Int. Math. Res. Not.* **17** (2009), 3286–3309.

[19] T. Banica and S. Curran, Decomposition results for Gram matrix determinants, *J. Math. Phys.* **51** (2010), 1–14.

[20] T. Banica, S. Curran and R. Speicher, Classification results for easy quantum groups, *Pacific J. Math.* **247** (2010), 1–26.

[21] T. Banica, S. Curran and R. Speicher, Stochastic aspects of easy quantum groups, *Probab. Theory Related Fields* **149** (2011), 435–462.

[22] T. Banica, S. Curran and R. Speicher, De Finetti theorems for easy quantum groups, *Ann. Probab.* **40** (2012), 401–435.

[23] T. Banica and D. Goswami, Quantum isometries and noncommutative spheres, *Comm. Math. Phys.* **298** (2010), 343–356.

[24] T. Banica and I. Nechita, Asymptotic eigenvalue distributions of block-transposed Wishart matrices, *J. Theoret. Probab.* **26** (2013), 855–869.

[25] T. Banica and I. Nechita, Block-modified Wishart matrices and free Poisson laws, *Houston J. Math.* **41** (2015), 113–134.

[26] T. Banica and I. Nechita, Flat matrix models for quantum permutation groups, *Adv. Appl. Math.* **83** (2017), 24–46.

[27] T. Banica and R. Speicher, Liberation of orthogonal Lie groups, *Adv. Math.* **222** (2009), 1461–1501.

[28] H. Bercovici and V. Pata, Stable laws and domains of attraction in free probability theory, *Ann. of Math.* **149** (1999), 1023–1060.

[29] H. Bercovici and D.V. Voiculescu, Regularity questions for free convolution, *Oper. Theory Adv. Appl.* **104** (1998), 37–47.

[30] J. Bhowmick, F. D'Andrea and L. Dabrowski, Quantum isometries of the finite noncommutative geometry of the standard model, *Comm. Math. Phys.* **307** (2011), 101–131.

[31] P. Biane, Some properties of crossings and partitions, *Discrete Math.* **175** (1997), 41–53.

[32] J. Bichon, Half-liberated real spheres and their subspaces, *Colloq. Math.* **144** (2016), 273–287.

[33] J. Bichon and M. Dubois-Violette, Half-commutative orthogonal Hopf algebras, *Pacific J. Math.* **263** (2013), 13–28.

[34] E. Blanchard, Déformations de C^*-algèbres de Hopf, *Bull. Soc. Math. Fr.* **124** (1996), 141–215.

[35] M. Brannan, Reduced operator algebras of trace-preserving quantum automorphism groups, *Doc. Math,* **18** (2013), 1349–1402.

[36] M. Brannan, A. Chirvasitu and A. Freslon, Topological generation and matrix models for quantum reflection groups, *Adv. Math.* **363** (2020), 1–26.

[37] M. Brannan, B. Collins and R. Vergnioux, The Connes embedding property for quantum group von Neumann algebras, *Trans. Amer. Math. Soc.* **369** (2017), 3799–3819.

[38] M. Brannan and K. Kirkpatrick, Quantum groups and generalized circular elements, *Pacific J. Math.* **282** (2016), 35–61.

[39] R. Brauer, On algebras which are connected with the semisimple continuous groups, *Ann. of Math.* **38** (1937), 857–872.

[40] W. Bryc, Compound real Wishart and q-Wishart matrices, *Int. Math. Res. Not.* **79** (2008), 1–42.

[41] B. Collins, Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability, *Int. Math. Res. Not.* **17** (2003), 953–982.

[42] B. Collins and S. Matsumoto, On some properties of orthogonal Weingarten functions, *J. Math. Phys.* **50** (2009), 1–18.

[43] B. Collins and I. Nechita, Random quantum channels I: graphical calculus and the Bell state phenomenon, *Comm. Math. Phys.* **297** (2010), 345–370.

[44] B. Collins and I. Nechita, Random quantum channels II: entanglement of random subspaces, Rényi entropy estimates and additivity problems, *Adv. Math.* **226** (2011), 1181–1201.

[45] B. Collins and I. Nechita, Gaussianization and eigenvalue statistics for random quantum channels (III), *Ann. Appl. Probab.* **21** (2011), 1136–1179.

[46] B. Collins and P. Śniady, Integration with respect to the Haar measure on unitary, orthogonal and symplectic groups, *Comm. Math. Phys.* **264** (2006), 773–795.

[47] A. Connes, Noncommutative geometry, Academic Press (1994).

[48] S. Curran, Quantum exchangeable sequences of algebras, *Indiana Univ. Math. J.* **58** (2009), 1097–1126.
[49] S. Curran, Quantum rotatability, *Trans. Amer. Math. Soc.* **362** (2010), 4831–4851.
[50] S. Curran, A characterization of freeness by invariance under quantum spreading, *J. Reine Angew. Math.* **659** (2011), 43–65.
[51] S. Curran and R. Speicher, Quantum invariant families of matrices in free probability, *J. Funct. Anal.* **261** (2011), 897–933.
[52] P. Di Francesco, Meander determinants, *Comm. Math. Phys.* **191** (1998), 543–583.
[53] P. Diaconis and M. Shahshahani, On the eigenvalues of random matrices, *J. Applied Probab.* **31** (1994), 49–62.
[54] L. Faddeev, Instructive history of the quantum inverse scattering method, *Acta Appl. Math.* **39** (1995), 69–84.
[55] L. Faddeev, N. Reshetikhin and L. Takhtadzhyan, Quantization of Lie groups and Lie algebras, *Leningrad Math. J.* **1** (1990), 193–225.
[56] A. Freslon, On the partition approach to Schur-Weyl duality and free quantum groups, *Transform. Groups* **22** (2017), 707–751.
[57] A. Freslon, Cut-off phenomenon for random walks on free orthogonal quantum groups, *Probab. Theory Related Fields* **174** (2019), 731–760.
[58] A. Freslon, L. Teyssier and S. Wang, Cutoff profiles for quantum Lévy processes and quantum random transpositions, preprint 2020.
[59] M. Fukuda and P. Śniady, Partial transpose of random quantum states: exact formulas and meanders, *J. Math. Phys.* **54** (2013), 1–31.
[60] I.M. Gelfand, Normierte Ringe, *Mat. Sb.* **9** (1941), 3–24.
[61] I.M. Gelfand and M.A. Naimark, On the imbedding of normed rings into the ring of operators on a Hilbert space, *Mat. Sb.* **12** (1943), 197–217.
[62] D. Goswami, Quantum group of isometries in classical and noncommutative geometry, *Comm. Math. Phys.* **285** (2009), 141–160.
[63] P. Graczyk, G. Letac and H. Massam, The complex Wishart distribution and the symmetric group, *Ann. Statist.* **31** (2003), 287–309.
[64] V.F.R. Jones, Index for subfactors, *Invent. Math.* **72** (1983), 1–25.
[65] V.F.R. Jones, On knot invariants related to some statistical mechanical models, *Pacific J. Math.* **137** (1989), 311–334.
[66] V.F.R. Jones, The Potts model and the symmetric group, in "Subfactors, Kyuzeso 1993" (1994), 259–267.
[67] V.F.R. Jones, Planar algebras I, preprint 1999.
[68] P. Józiak, Remarks on Hopf images and quantum permutation groups, *Canad. Math. Bull.* **61** (2018), 301–317.
[69] P. Józiak, Quantum increasing sequences generate quantum permutation groups, preprint 2019.
[70] C. Köstler, R. Speicher, A noncommutative de Finetti theorem: invariance under quantum permutations is equivalent to freeness with amalgamation, *Comm. Math. Phys.* **291** (2009), 473–490.
[71] F. Lemeux and P. Tarrago, Free wreath product quantum groups: the semigroupal category, approximation properties and free probability, *J. Funct. Anal.* **270** (2016), 3828–3883.
[72] R. Lenczewski, Limit distributions of random matrices, *Adv. Math.* **263** (2014), 253–320.
[73] W. Liu, General de Finetti type theorems in noncommutative probability, *Comm. Math. Phys.* **369** (2019), 837–866.
[74] M. Lupini, L. Mančinska and D.E. Roberson, Nonlocal games and quantum permutation groups, *J. Funct. Anal.* **279** (2020), 1–39.

[75] S. Malacarne, Woronowicz's Tannaka-Krein duality and free orthogonal quantum groups, *Math. Scand.* **122** (2018), 151–160.
[76] V.A. Marchenko and L.A. Pastur, Distribution of eigenvalues in certain sets of random matrices, *Mat. Sb.* **72** (1967), 507–536.
[77] J.A. Mingo and M. Popa, Freeness and the transposes of unitarily invariant random matrices, *J. Funct. Anal.* **271** (2016), 883–921.
[78] J.A. Mingo and M. Popa, Freeness and the partial transposes of Wishart random matrices, *Canad. J. Math.* **71** (2019), 659–681.
[79] J.A. Mingo and R. Speicher, Free probability and random matrices, Springer (2017).
[80] A. Nica and R. Speicher, Lectures on the combinatorics of free probability, Cambridge Univ. Press (2006).
[81] S. O'Rourke and A. Soshnikov, Products of independent non-Hermitian random matrices, *Electron. J. Probab.* **16** (2011), 2219–2245.
[82] S. Raum, Isomorphisms and fusion rules of orthogonal free quantum groups and their complexifications, *Proc. Amer. Math. Soc.* **140** (2012), 3207–3218.
[83] S. Raum and M. Weber, The full classification of orthogonal easy quantum groups, *Comm. Math. Phys.* **341** (2016), 751–779.
[84] G.C. Shephard and J.A. Todd, Finite unitary reflection groups, *Canad. J. Math.* **6** (1954), 274–304.
[85] R. Speicher, Multiplicative functions on the lattice of noncrossing partitions and free convolution, *Math. Ann.* **298** (1994), 611–628.
[86] R. Speicher, Combinatorial theory of the free product with amalgamation and operator-valued free probability theory, *Mem. Amer. Math. Soc.* **132** (1998).
[87] P. Tarrago and J. Wahl, Free wreath product quantum groups and standard invariants of subfactors, *Adv. Math.* **331** (2018), 1–57.
[88] P. Tarrago and M. Weber, Unitary easy quantum groups: the free case and the group case, *Int. Math. Res. Not.* **18** (2017), 5710–5750.
[89] S. Vaes and R. Vergnioux, The boundary of universal discrete quantum groups, exactness and factoriality, *Duke Math. J.* **140** (2007), 35–84.
[90] D.V. Voiculescu, Symmetries of some reduced free product C*-algebras, in "Operator algebras and their connections with topology and ergodic theory", Springer (1985), 556–588.
[91] D.V. Voiculescu, Addition of certain noncommuting random variables, *J. Funct. Anal.* **66** (1986), 323–346.
[92] D.V. Voiculescu, Multiplication of certain noncommuting random variables, *J. Operator Theory* **18** (1987), 223–235.
[93] D.V. Voiculescu, Limit laws for random matrices and free products, *Invent. Math.* **104** (1991), 201–220.
[94] D.V. Voiculescu, K.J. Dykema and A. Nica, Free random variables, AMS (1992).
[95] S. Wang, Free products of compact quantum groups, *Comm. Math. Phys.* **167** (1995), 671–692.
[96] S. Wang, Quantum symmetry groups of finite spaces, *Comm. Math. Phys.* **195** (1998), 195–211.
[97] D. Weingarten, Asymptotic behavior of group integrals in the limit of infinite rank, *J. Math. Phys.* **19** (1978), 999–1001.
[98] E. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, *Ann. of Math.* **62** (1955), 548–564.
[99] S.L. Woronowicz, Compact matrix pseudogroups, *Comm. Math. Phys.* **111** (1987), 613–665.
[100] S.L. Woronowicz, Tannaka-Krein duality for compact matrix pseudogroups. Twisted SU(N) groups, *Invent. Math.* **93** (1988), 35–76.