NOTES ON FREE PROBABILITY

VIJAY KODIYALAM

These are the notes for a set of lectures on Free Probability given as part of the Advanced Instructional School conducted at IMSc, Chennai between Feb. 1 and Feb. 20, 2016. The notes are very closely based on the book of Nica and Speicher and roughly correspond to the first 5 chapters of that book together with a part of Chapter 8. Comments and corrections are very welcome.

1. Basic definitions

1.1. **Some definitions.** A **probability space** consists of a complex, unital algebra A and a linear functional $\phi : A \to \mathbb{C}$ such that $\phi(1_A) = 1$. Elements of A will be called **random variables**. If ϕ is a trace the probability space A will be said to be **tracial**.

If A is equipped with a conjugate linear, product-reversing involution $*$, i.e., A is a **$*$-algebra**, and ϕ is **positive** in the sense that $\phi(a^*a) \geq 0$ for all $a \in A$, then A is said to be a **$*$-probability space**. If A is a $*$-probability space, we may speak of **self-adjoint** ($a = a^*$), **unitary** ($aa^* = 1 = a^*a$) and **normal** ($aa^* = a^*a$) random variables. Our main interest will be in $*$-probability spaces.

1.2. **Adjoint preservation.** If A is a $*$-probability space, the functional ϕ preserves $*$, i.e., $\phi(a^*) = \overline{\phi(a)}$. To see this note first that this implies that if $x \in A$ is self-adjoint, $\phi(x) \in \mathbb{R}$. The converse holds too. For any element $a \in A$ can be expressed uniquely as $x + iy$ where $x, y \in A$ are self-adjoint (and $x = \frac{1}{2}(a + a^*), y = \frac{1}{2i}(a - a^*)$) and so if for self-adjoint x, we have $\phi(x) \in \mathbb{R}$, then $\phi(a) = \phi(x) + i\phi(y)$ is the decomposition of $\phi(a)$ into its real and imaginary parts. Then, $\phi(a^*) = \phi(x) - i\phi(y) = \overline{\phi(a)}$. Finally to see that for self-adjoint x, $\phi(x)$ is real, note that x can be written as $a^*a - b^*b$ where $a = \frac{1}{2}(x + 1)$ and $b = \frac{1}{2}(x - 1)$. Now appeal to the positivity of ϕ.

1.3. **Cauchy-Schwarz.** If A is a $*$-probability space, we also have the following Cauchy-Schwarz inequality for the sesquilinear form defined by ϕ :

$$|\phi(b^*a)|^2 \leq \phi(a^*a)\phi(b^*b).$$

To prove this, note that the quadratic function of the real valued variable t given by $\phi((a - tb)^*(a - tb))$ is always non-negative. Hence its discriminant is negative and this gives:

$$(Re(\phi(b^*a))^2 \leq \phi(a^*a)\phi(b^*b)$$

The RHS is invariant under replacing a with $ae^{i\theta}$ while the maximum of the LHS as θ varies is exactly $|\phi(b^*a)|^2$, yielding the desired inequality.

1.4. **More definitions.** The functional ϕ is said to be **faithful** if $\phi(a^*a) = 0$ implies that $a = 0$. A **morphism** $\lambda : (A, \phi) \to (B, \psi)$ of $*$-probability spaces is a unital $*$-algebra homomorphism such that $\psi \circ \lambda = \phi$. If ϕ is faithful, then λ is necessarily injective.

2. Examples of probability spaces and representations

2.1. **Classical examples.** Let $(X, \mathcal{B}, \mu)$ be a classical probability space. Thus X is a set, $\mathcal{B}$ is a σ-algebra of **measurable subsets** of X and μ is a probability measure on $\mathcal{B}$. Associated to this data is the probability space $A = L^\infty(X, \mathcal{B}, \mu)$ where ϕ is defined by

$$\phi(a) = \int_X a \; d\mu$$

for $a \in A$. This A is then a commutative (hence, of course, tracial) $*$-probability space where $*$ is the usual complex conjugation.

There are other probability spaces that one may associate to this data. For instance, one problem with the above space is the following. Consider a Gaussian random variable. This is one for which the probability density function is given by

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

i.e., the μ-measure of $\{x \in X : a(x) \in I\}$ is given by

$$\int_I \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx,$$

for any interval I. This random variable is certainly not given by an L^∞ function on X since its density is not compactly supported. So the framework needs to be enlarged to be able to treat Gaussian random variables.

Define $L^{\infty-}(X, \mathcal{B}, \mu) = \cap_{1 \leq p < \infty} L^p(X, \mathcal{B}, \mu)$. Note that this is a descending intersection since μ is a finite measure and hence also equals $\cap_{p=1}^\infty L^p(X, \mathcal{B}, \mu)$. Further, this is an algebra. For if $a, b \in L^{\infty-}(X, \mathcal{B}, \mu)$, then for any finite $p \geq 1$, $a, b \in L^{2p}(X, \mathcal{B}, \mu)$. Thus $a^p, b^p \in L^2(X, \mathcal{B}, \mu)$. By Cauchy-Schwarz, $a^p b^p \in L^1(X, \mathcal{B}, \mu)$, or equivalently, $ab \in L^p(X, \mathcal{B}, \mu)$. Since p is arbitrary, $ab \in L^{\infty-}(X, \mathcal{B}, \mu)$. So with $A = L^{\infty-}(X, \mathcal{B}, \mu)$ and ϕ as before we have a more interesting $*$-probability space. Both these spaces are faithful.

2.2. **Matrix examples.** A first non-commutative example of a $*$-probability space is given by $A = M_d(\mathbb{C})$ for $d \geq 2$ with ϕ given by the normalised matrix trace. Note that for $a \in A$, $\phi(a^*a)$ is $\frac{1}{n}$ times the square of the Hilbert-Schmidt norm (sum of the squares of the norms of all entries) of a and thus ϕ is indeed positive and even faithful.

More generally, a method of constructing new $*$-probability spaces from old is the matrix construction. If A is a $*$-probability space with ϕ, then for any $d \in \mathbb{N}$, we have a $*$-probability space $M_d(A)$ with usual multiplication and $*$ and functional $\phi^{(d)}$ given by $\phi^{(d)}(a) = \frac{1}{d}\sum_{i=1}^d \phi(a_{ii})$. It is easily checked that $\phi^{(d)}$ is positive and is tracial if ϕ is so and faithful if ϕ is so.

In particular, we may do the matrix construction with $L^{\infty-}(X, \mathcal{B}, \mu)$ to get the $*$-probability space $A = M_d(L^{\infty-}(X, \mathcal{B}, \mu))$ with linear functional ϕ defined by

$$\phi(a) = \int_X tr(a) d\mu$$

where tr denotes the normalised trace of a matrix. The random variables of this space are called random matrices over $(X, \mathcal{B}, \mu)$, and can also be regarded as matrix valued functions on X.

2.3. **Group algebras.** A more interesting non-commutative example related to the genesis of free probability theory is given by the following. Let G be a countable group and $A = \mathbb{C}G$ - the group algebra of G. Elements of A are finite formal linear combinations of elements of G with the obvious multiplication. Make A a $*$-algebra by conjugate linearly extending $g^* = g^{-1}$. Define ϕ on A by $\phi(g) = \delta_{g,1}$ and extend by linearity. This is a trace on A and is faithful and positive since if $a = \sum_g c_g g$, then $\phi(a^*a) = \sum_g |c_g|^2$.

2.4. **Subalgebras of $\mathcal{L}(\mathcal{H})$.** Finally, an extremely general example is given as follows. Let $\mathcal{H}$ be a Hilbert space and $\mathcal{L}(\mathcal{H})$ be the $*$-algebra of bounded operators on $\mathcal{H}$. Let A be any unital $*$-subalgebra of $\mathcal{L}(\mathcal{H})$ and $\Omega \in \mathcal{H}$ be a unit vector. We refer to Ω as a **vacuum vector**. Define ϕ on A by $\phi(a) = \langle a\Omega|\Omega\rangle$ - the so-called **vector-state** defined by Ω. Then ϕ is positive since $\phi(a^*a) = ||a\Omega||^2 \geq 0$, but is not necessarily faithful or tracial. In particular, $\mathcal{L}(\mathcal{H})$ itself is a $*$-probability space for any choice of unit vector $\Omega \in \mathcal{H}$, the corresponding functional being denoted by ϕ_Ω.

2.5. **Representations.** A **representation** of a $*$-probability space (A, ϕ) is a morphism into $(\mathcal{L}(\mathcal{H}), \phi_\Omega)$ for some Hilbert space $\mathcal{H}$ and unit vector $\Omega \in \mathcal{H}$. Equivalently, it is a unital $*$-homomorphism $\lambda : A \to \mathcal{L}(\mathcal{H})$ such that $\phi(a) = \langle \lambda(a)\Omega|\Omega\rangle$.

All examples discussed so far except for $L^{\infty-}(X, \mathcal{B}, \mu)$ have natural and faithful representations on Hilbert space. For instance, in the group algebra case, the natural Hilbert space that arises is $\ell^2(G)$ - the Hilbert space with orthonormal basis given by $\{\xi_g : g \in G\}$. The vector $\Omega = \xi_1$ and the representation $\lambda : \mathbb{C}G \to \mathcal{L}(\ell^2(G))$ is given by $\lambda(g)(\xi_h) = \xi_{gh}$. Note that $\langle \lambda(g)\Omega|\Omega\rangle = \langle \xi_g|\xi_1\rangle = \delta_{g,1}$, as needed.

The algebra LG is defined to be the double commutant of $\lambda(\mathbb{C}G)$ in $\mathcal{L}(\ell^2(G))$ and is one of the standard examples of a II_1-factor when G has all non-trivial conjugacy classes infinite. The origins of free probability theory lie in the open decision problem of whether $LF_2 \cong LF_3$ where F_k is the free group on k-generators.

3. $*$-DISTRIBUTIONS OF NORMAL ELEMENTS

3.1. **Analytic $*$-distributions.** In general, by the $*$-distribution of a random variable a of a probability space, we will mean some scheme that keeps track of the values of ϕ on the unital $*$-subalgebra generated by a.

If a is normal, the unital $*$-subalgebra generated by a is the span of all $\{a^k(a^*)^l : k, l \geq 0\}$, and so the $*$-distribution must be something that specifies all $\phi(a^k(a^*)^l)$ for $k, l \geq 0$. The best possible such gadget is a compactly supported Borel probability measure μ on $\mathbb{C}$ for which

$$\phi(a^k(a^*)^l) = \int z^k \overline{z}^l d\mu.$$

If such a μ exists, then it is necessarily unique by the Stone-Weierstrass theorem and is called the **analytic $*$-distribution** of the normal random variable a.

While not all normal random variables of $*$-probability spaces need have an analytic $*$-distribution, most interesting ones do. In particular if a $*$-probability space has a representation on Hilbert space, then its normal random variables do have analytic $*$-distributions.

3.2. **$*$-distributions of self-adjoints are $\mathbb{R}$-supported.** Suppose that a is self-adjoint and has an analytic $*$-distribution, say μ. Then μ is supported in $\mathbb{R}$. For consider $\int_{\mathbb{C}} |z-\overline{z}|^2 d\mu = 0$ and since $|z-\overline{z}|^2$ is a non-negative continuous function it must vanish on the support of μ. Hence $supp(\mu) \subseteq \mathbb{R}$. So here, the $*$-distribution of a is a compactly supported measure μ on $\mathbb{R}$ such that

$$\phi(a^p) = \int_{\mathbb{R}} t^p d\mu.$$

4. Examples of $*$-distributions and general $*$-distributions

4.1. **The classical case.** Consider $A = L^\infty(X, \mathcal{B}, \mu)$, and say $a \in A$, so that $a : X \to \mathbb{C}$ is a bounded measurable function. In classical probability theory the distribution of a is the push-forward measure of μ to $\mathbb{C}$, i.e.,

$$\nu(E) = \mu(a^{-1}(E)).$$

The boundedness of a implies that ν is compactly supported.

Consider what we defined as the analytic $*$-distribution of a. Let's check that this is the same as ν. The displayed equation above is the same as:

$$\int_{\mathbb{C}} 1_E d\nu = \int_X 1_E(a(x)) d\mu.$$

By properly approximating we get

$$\int_{\mathbb{C}} f d\nu = \int_X f(a(x)) d\mu.$$

for every bounded measurable function f on $\mathbb{C}$. We now take f to agree with $z^k\overline{z}^l$ on the support of μ and 0 outside to get

$$\int_{\mathbb{C}} z^k\overline{z}^l d\nu = \int_X a(x)^k \overline{a(x)}^l d\mu = \phi(a^k(a^*)^l).$$

4.2. **Matrix algebras.** Consider $A = M_d(\mathbb{C})$ and let $a \in A$ be a normal matrix with eigenvalues $\lambda_1, \cdots, \lambda_d$. We may diagonalize a to get that

$$tr(a^k(a^*)^l) = \frac{1}{d}\sum_{i=1}^{d} \lambda_i^k \overline{\lambda_i}^l = \int_{\mathbb{C}} z^k\overline{z}^l d\mu$$

where $\mu = \frac{1}{d}\sum_{i=1}^d \delta_{\lambda_i}$. This μ is called the **eigenvalue distribution** of a.

4.3. **Haar and p-Haar unitaries.** Let (A, ϕ) be a $*$-probability space. An element $u \in A$ is said to be a **Haar unitary** if it is unitary and $\phi(u^k) = 0$ for $k \in \mathbb{Z} \setminus \{0\}$. For p a positive integer, $u \in A$ is said to be a **p-Haar unitary** if it is unitary, $u^p = 1$, and $\phi(u^k) = 0$ for $k \in \mathbb{Z}$ a non-multiple of p.

If u is a Haar unitary, the Haar measure on $S^1 \subseteq \mathbb{C}$ is its analytic $*$-distribution. For, by definition, $\phi(u^k(u^*)^l) = \delta_{k,l}$ while for the Haar measure μ on S^1, we have

$$\int_{S^1} z^k\overline{z}^l d\mu = \int_0^{2\pi} e^{i(k-l)\theta}\frac{d\theta}{2\pi} = \delta_{k,l}$$

If u is a p-Haar unitary, its analytic $*$-distribution is given by $\mu = \frac{1}{p}\sum_{i=0}^{p-1} \delta_{\omega^i}$ where ω is a primitive p^{th}-root of 1. This is because for any $k \in \mathbb{Z}$,

$$\int_{\mathbb{C}} z^k d\mu = \sum_{i=0}^{p-1} \omega^{ik}$$

which is 1 or 0 according as p divides k or not.

In the probability space $\mathbb{C}G$, elements $g \in G$ of infinite order are Haar unitaries while those of finite order p are p-Haar unitaries.

4.4. **The distribution of $u+u^*$.** Let u be a Haar unitary in a $*$-probability space A. We want to consider whether the self-adjoint element $u+u^*$ has an analytic $*$-distribution and determine its **moments**, which are, by definition, the numbers $\phi((u+u^*)^p)$ for $p \geq 0$.

The moments are easy to determine. We have $(u+u^*)^p = \sum_{k=0}^{p} \binom{p}{k} u^k (u^*)^{p-k}$. So $\phi((u+u^*)^p)$ vanishes unless p is even in which case it is given by $\binom{p}{\frac{p}{2}}$. Now we want a compactly supported measure ν on $\mathbb{R}$ for which

$$\int_{\mathbb{R}} t^p d\nu$$

vanishes for odd p and is given for even p by $\binom{p}{\frac{p}{2}}$.

We know the corresponding measure for u itself which is the Haar measure μ on S^1. Hence

$$\begin{aligned}
\phi((u+u^*)^p) &= \sum_{k=0}^{p} \binom{p}{k} \phi(u^k (u^*)^{p-k}) \\
&= \sum_{k=0}^{p} \binom{p}{k} \int_{S^1} z^k \overline{z}^{p-k} \\
&= \int_{S^1} (z+\overline{z})^p d\mu \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} (2\ cos(\theta))^p\ d\theta \\
&= \frac{1}{\pi} \int_{0}^{\pi} (2\ cos(\theta))^p\ d\theta
\end{aligned}$$

Since we want to write this as a moment, set $t = 2cos(\theta)$, so that $\theta = cos^{-1}(\frac{t}{2})$ and $d\theta = \frac{-dt}{\sqrt{4-t^2}}$. Hence,

$$\phi((u+u^*)^p) = \frac{1}{\pi} \int_{-2}^{2} \frac{t^p dt}{\sqrt{4-t^2}} dt$$

So the sought after compactly supported measure is given by

$$d\nu = \begin{cases} \frac{dt}{\pi\sqrt{4-t^2}} & \text{if } \ |t| < 2 \\ 0 & \text{if } \ |t| \geq 2 \end{cases}$$

Note that when $p = 2k$, knowledge of $\phi((u+u^*)^{2k})$ implies that

$$\int_0^{\pi} (cos(\theta))^{2k} d\theta = \frac{\pi}{4^k} \binom{2k}{k}.$$

We will use this later.

4.5. **Distribution of a polynomial function.** What we did in the previous example can be done more generally. Suppose that a is a normal element in a $*$-probability space which has an analytic $*$-distribution μ. Let P be a polynomial in two (commuting) variables $z, \overline{z}$ and let $b = P(a, a^*)$. Then b has an analytic $*$-distribution ν given by the push-forward of μ by $P : \mathbb{C} \to \mathbb{C}$, i.e., $\nu(E) = \mu(P^{-1}(E))$.

To see this, first note that ν is a Borel measure on $\mathbb{C}$ and further that it is compactly supported since if K is the support of μ, then $P(K)$ is bounded hence contained in some compact L and then the support of ν is contained in L.

Also $\phi(b^k(b^*)^l) = \int_{\mathbb{C}} P(z,\overline{z})^k (P(z,\overline{z})^*)^l d\mu = \int_{\mathbb{C}} w^k \overline{w}^l d\nu$, as needed, just as in an earlier proof.

4.6. **Application to a p-Haar unitary.** We may use this for a p-Haar unitary u. Its analytic $*$-distribution is given by $\mu = \frac{1}{p}\sum_{i=0}^{p-1} \delta_{\omega^i}$ and so the analytic $*$-distribution of $u + u^*$ is given by the push-forward of μ by the map $z + \overline{z}$. Denoting this measure by ν, a little thought shows that ν is also atomic with mass $\frac{1}{p}$ at 1, $\frac{1}{p}$ at -1 if p is even, and $\frac{2}{p}$ at all $2cos(\frac{2k\pi}{p})$ for $0 < k < \frac{p}{2}$.

4.7. **General $*$-distributions.** Let (A, ϕ) be a $*$-probability space and $a \in A$ be some element. The $*$-subalgebra of A generated by a is the linear span of all words in a and a^*. The values of ϕ on such words will be referred to as **$*$-moments** of a.

Let $\mathbb{C}\langle X, X^*\rangle$ be the polynomial algebra in two non-commuting variables X and X^* equipped with its obvious $*$-structure. The **algebraic $*$-distribution** of a is defined to be the ($*$-preserving) linear functional $\mu : \mathbb{C}\langle X, X^*\rangle \to \mathbb{C}$ defined uniquely by:

$$\mu(X^{\epsilon_1} X^{\epsilon_2} \cdots X^{\epsilon_k}) = \phi(a^{\epsilon_1} a^{\epsilon_2} \cdots a^{\epsilon_k})$$

for all $k \geq 0$ and $\epsilon_i \in \{*, 1\}$.

Note that just as the analytic $*$-distribution is a compactly supported measure on $\mathbb{R}$, the algebraic $*$-distribution is a linear functional on $\mathbb{C}\langle X, X^*\rangle$, independent of which probability space the random variable a belongs to.

When $a \in A$ is self-adjoint, its $*$-moments are just its **moments**, i.e., $\phi(a^p)$ for $p \geq 0$, and following classical probability terminology, the first moment $\phi(a)$ is called the **mean** of a and the quantity $\phi(a^2) - \phi(a)^2$ is called the **variance** of a.

5. A non-unitary isometry

5.1. **Definitions.** Let (A, ϕ) be a $*$-probability space. An element $a \in A$ such that a is an **isometry**, i.e., $a^*a = 1$, but is not unitary so that $aa^* \neq 1$ is called a **non-unitary isometry**. Of course, a is not normal and A is necessarily infinite-dimensional. Further assume that a generates A as a $*$-algebra. It is then easy to see that A is the span of all $a^k(a^*)^l$ for $k, l \geq 0$. Assume further that the elements $a^k(a^*)^l$ for $k, l \geq 0$ are all linearly independent and that $\phi(a^k(a^*)^l) = 0$ unless $k = l = 0$. Note that this ϕ is clearly not faithful.

5.2. **Existence.** How are we sure of existence ? Consider the Hilbert space $\mathcal{H} = \ell^2(\mathbb{N} \cup \{0\})$, with orthonormal basis ξ_n for $n \geq 0$. Let $S \in \mathcal{L}(\mathcal{H})$ be the unilateral shift defined by $S\xi_n = \xi_{n+1}$ for $n \geq 0$. One checks that $S^*\xi_n = \xi_{n-1}$ for $n > 0$ and that $S^*\xi_0 = 0$. Hence $SS^* = 1$ but $S^*S = P$ where P is the orthogonal projection onto the closed subspace spanned by all ξ_n for $n > 0$. Thus S is a non-unitary isometry.

Further, the elements $S^k(S^*)^l$ for $k, l \geq 0$ are all linearly independent. For suppose we have a relation of the type $P_0(S) + P_1(S)S^* + P_2(S)(S^*)^2 + \cdots + P_n(S)(S^*)^n = 0$ where the P_i are polynomials in 1 variable that are not all zero. Let m be least so that $P_m \neq 0$ and apply this operator to ξ_m to get $P_m(S)\xi_0 = 0$. This clearly imples that $P_m = 0$.

5.3. **Toeplitz algebra.** Let ψ be the vector state defined on $\mathcal{L}(\mathcal{H})$ by ξ_0, i.e., $\psi(T) = \langle T\xi_0|\xi_0\rangle$. Then the map $\pi : A \to \mathcal{L}(\mathcal{H})$ defined by linearly extending $\pi(a^k(a^*)^l) = S^k(S^*)^l$ is a representation of the $*$-probability space A. The closure of $\pi(A)$ in the norm topolgy on $\mathcal{L}(\mathcal{H})$ is called the **Toeplitz algebra** and is an important example in C^*-algebra theory.

5.4. **Meaning of $*$-distribution.** By the $*$-distribution of a, we meant a gadget that keeps track of the values of ϕ on the $*$-algebra generated by a. In this case, that algebra is spanned by $a^k(a^*)^l$. So is knowing all $\phi(a^k(a^*)^l)$ enough to know its $*$-distribution ? Not really because the corresponding measure for these values is the Dirac mass at 0 which is the same one as for the $0 \in A$. To understand the $*$-distribution of a, we also need to understand the process by which monomials in a and a^* are reduced to the form $a^k(a^*)^l$. This needs a little combinatorial digression.

6. Dyck paths and Catalan numbers

6.1. **Definitions.** By a **NE-SE path** we mean a path in the plane that begins at $(0,0)$ and moves in steps of $(1, \pm 1)$, i.e., NE or SE steps. The length of the path is the number of steps. There is an empty path of length 0. A NE-SE path is called a **Dyck path** if it never goes below the x-axis and ends on the x-axis.

6.2. **Parametrisation of paths.** A NE-SE path of length k is clearly parametrised by a sequence $(\lambda_1, \cdots, \lambda_k)$ where each $\lambda_i \in \{\pm 1\}$ and the i^{th}-step of the path is by $(1, \lambda_i)$. The y-coordinate of the end-point after i-steps is given by $\lambda_1 + \cdots + \lambda_i$. Thus $(\lambda_1, \cdots, \lambda_k)$ parametrises a Dyck path exactly when each $\lambda_1 + \cdots + \lambda_i \geq 0$ and $\lambda_1 + \cdots + \lambda_k = 0$. Since each $\lambda_i = \pm 1$, the length of a Dyck path is necessarily even.

6.3. **The reflection trick.** We want to show that for every integer $p \geq 0$, the number C_p of Dyck paths of length $2p$ is given by

$$C_p = \frac{1}{2p}\binom{2p}{p}.$$

This uses a very pretty and justly famous 'reflection trick' due to Andre.

To apply this, begin by observing that (m, n) is the end-point of a NE-SE path if and only if $m \geq 0$, m and n have the same parity, and $|n| \leq m$. If these conditions hold, any NE-SE path with end-point (m, n) has $\frac{m+n}{2}$ NE steps and $\frac{m-n}{2}$ SE steps. So the number of such is the binomial coefficient $\binom{m}{\frac{m+n}{2}}$.

The reflection trick establishes a bijection between the sets of non-Dyck NE-SE paths ending at $(2p, 0)$ and all NE-SE paths ending at $(2p, -2)$. Given this, the number of Dyck paths ending at $(2p, 0)$ is given by $\binom{2p}{p} - \binom{2p}{p-1} = \frac{1}{2p}\binom{2p}{p}$, as desired.

The bijection is given by reflecting the part of the path that lies to the right of where it first touches $y = -1$ about that line.

6.4. **The Catalan recurrence.** For every positive integer p, the number C_p above is called the p^{th}-**Catalan** number and we will show that it satisfies the recurrence $C_p = \sum_{k=1}^{p} C_{k-1}C_{p-k}$ subject to the initial conditions $C_0 = C_1 = 1$.

To prove this, call a Dyck path irreducible if the only points at which it touches the x-axis are its end-points. It should be clear that the number of irreducible Dyck paths of length $2p$ equals the number C_{p-1} of Dyck paths of length $2p-2$. The reducible Dyck paths of length $2p$ are divided into classes according to the first time they touch the x-axis (leaving the beginning point) which may be any $(2k, 0)$ for $1 \leq k \leq p-1$. The number in the k^{th}-class is clearly $C_{k-1}C_{p-k}$. Thus $C_p = \sum_{k=1}^{p-1} C_{k-1}C_{p-k} + C_{p-1} = \sum_{k=1}^{p} C_{k-1}C_{p-k}$.

7. Distribution of $a + a^*$

7.1. **Connection with Dyck paths.** The first observation is the following. Consider a monomial $a^{\epsilon_1}a^{\epsilon_2}\cdots a^{\epsilon_k}$ where each $\epsilon_i \in \{*, 1\}$. Let $\lambda_i = 1$ or -1 according as ϵ_i is $*$ or 1. Then $\phi(a^{\epsilon_1}a^{\epsilon_2}\cdots a^{\epsilon_k}) = 1$ or 0 according as $(\lambda_1, \cdots, \lambda_k)$ corresponds to a Dyck path or not.

To see this we use the representation π on Hilbert space of the algebra A to conclude that

$$\begin{aligned}
\phi(a^{\epsilon_1}a^{\epsilon_2}\cdots a^{\epsilon_k}) &= \psi(S^{\epsilon_1}S^{\epsilon_2}\cdots S^{\epsilon_k}) \\
&= \langle S^{\epsilon_1}S^{\epsilon_2}\cdots S^{\epsilon_k}\xi_0|\xi_0\rangle \\
&= \langle \xi_0|(S^{\epsilon_k})^*(S^{\epsilon_{k-1}})^*\cdots(S^{\epsilon_1})^*\xi_0\rangle
\end{aligned}$$

Now, a little thought shows that $(S^{\epsilon_k})^*(S^{\epsilon_{k-1}})^*\cdots(S^{\epsilon_1})^*\xi_0$ which is either some ξ_n or vanishes, vanishes exactly when some $\lambda_1 + \cdots + \lambda_j < 0$ (for $1 \leq j \leq k$) and otherwise gives $\xi_{\lambda_1+\cdots+\lambda_k}$. The required equality for $\phi(a^{\epsilon_1}a^{\epsilon_2}\cdots a^{\epsilon_k})$ is now immediate by definition of a Dyck path.

7.2. **Moments.** We now ask, as before, whether the self-adjoint element $a+a^*$ has an analytic $*$-distribution and what its moments are. Again the moments are easy to determine.

For $k \geq 0$, $\phi((a+a^*)^p)$ is the sum of ϕ of all monomials in a and a^* of total degree p and thus vanishes for odd p and for $p = 2k$ gives C_k - the k^{th}-Catalan number. Next, we want a compactly supported probability measure μ on $\mathbb{R}$ whose moments are $\phi((a + a^*)^p)$. Since this vanishes for p odd, it suggests that the measure μ is symmetric about the origin.

7.3. **The measure.** We claim that the measure μ is supported in $[-2, 2]$ and has density given by $\frac{1}{2\pi}\sqrt{4-t^2}$ in that interval. To see this, we need to verify that when $p = 2k$,

$$\phi((a + a^*)^p) = \int_{-2}^{2} \frac{t^p}{2\pi}\sqrt{4-t^2}dt.$$

(That both sides vanish for odd p is clear.)

Set $t = 2cos(\theta), dt = -2sin(\theta)d\theta$ to reduce to calculating

$$\begin{aligned}
\int_0^\pi \frac{4^{k+1}(cos(\theta)^{2k})}{2\pi}(sin(\theta)^2)d\theta &= \frac{4^{k+1}}{2\pi}\left\{\int_0^\pi cos(\theta)^{2k}d\theta - \int_0^\pi cos(\theta)^{2k+2}d\theta\right\} \\
&= \frac{4^{k+1}}{2\pi}\left\{\frac{\pi}{4^k}\binom{2k}{k} - \frac{\pi}{4^{k+1}}\binom{2k+2}{k+1}.\right\}
\end{aligned}$$

A little calculation now shows that this equals C_k, as desired.

7.4. **Semicircular element.** This example motivates the ultra-important definition of a semicircular element. In a $*$-probability space (A, ϕ), for a positive r, a self-adjoint a is said to be (centered) **semi-circular** of radius r if a has analytic $*$-distribution given by $\frac{2}{\pi r^2}\sqrt{r^2 - t^2}dt$ on the interval $[-r, r]$. Equivalently, if its odd moments vanish and $(2k)^{th}$ moment is given by $(\frac{r}{2})^{2k} C_k$ for $k \geq 0$.

A semicircular element of radius 2 is said to be **standard semicircular**. Thus, our calculation has established that $S + S^*$ standard semicircular. However, rather unsatisfactorily, the measure was pulled out of thin air. To remedy this we will study the Cauchy transform.

7.5. **Another example of a standard semicircular element.** In the probability space (A, ϕ) considered earlier the element $\frac{1}{i}(a - a^*)$ is also standard semicircular. To see this it suffices to see that there is an automorphism θ of the $*$-probability space (A, ϕ) that takes a to $-ia$ (so that $a + a^* \mapsto -i(a - a^*)$). Note that A as an algebra is the quotient of $\mathbb{C}\langle A, A^* \rangle$ by the relation $A^* A = 1$. So there is an automorphism of A taking a to $-ia$ (and a^* to ia^*). To verify that this preserves ϕ use the connection with Dyck paths.

Another way to see this is to see that S and $-iS$ are unitarily conjugate in $\mathcal{L}(\ell^2(\mathbb{N} \cup \{0\}))$ by a unitary that fixes ξ_0. The unitary diagonal matrix with (t, t) entry given by $(-i)^t$ (for $t \geq 0$) works.

8. The Cauchy transform

8.1. **Definition.** The Cauchy transform is a method to derive a measure from the knowledge of its moments. Let μ be a probability measure on $\mathbb{R}$. Its **Cauchy transform** is the function G_μ defined on the upper half-plane $\mathbb{H}^+$ by

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z - t} d\mu(t).$$

This is analytic in $\mathbb{H}^+$ and takes values in $\mathbb{H}^-$. Suppose that μ is compactly supported. Say $r = sup\{|t| : t \in Supp(\mu)\}$. Then, for $|z| > r$, we have a power series expansion

$$G_\mu(z) = \sum_{n=0}^{\infty} \frac{\alpha_n}{z^{n+1}},$$

with $\alpha_n = \int_{\mathbb{R}} t^n d\mu$. In particular $lim_{|z| \to \infty} z G_\mu(z) = \alpha_0 = 1$.

8.2. **Inversion formula.** The **Stieltjes inversion formula** states the following. Suppose that G_μ has a continuous extension to $\mathbb{H}^+ \cup \mathbb{R}$ and that on $\mathbb{R}$, this is the function g, so that,

$$g(t) = lim_{\epsilon \to 0} G_\mu(t + i\epsilon).$$

Then,

$$d\mu(t) = -\frac{1}{\pi} Im(g(t))dt.$$

We discuss a couple of applications.

8.3. **Standard semicircular element.** Begin by noting that in this case,

$$G_\mu(z) = \sum_{p\geq 0} \frac{C_p}{z^{2p+1}},$$

and we know that if it comes from a compactly supported measure μ on $\mathbb{R}$, then it converges for $|z|$ large enough which anyway can be directly verified by calculating the radius of convergence.

I find it easier to think in terms of the generating function

$$H(z) = \sum_{p\geq 0} C_p z^p$$

which is analytic in a neighbourhood of $z = 0$. The Catalan recurrence gives

$$H(z)^2 = \frac{H(z) - C_0}{z}.$$

Now $G_\mu(z) = \frac{1}{z}H(\frac{1}{z^2})$, and therefore

$$G_\mu(z)^2 - zG_\mu(z) + 1 = 0.$$

This first holds for $z \in \mathbb{H}^+$ with $|z|$ large enough and then by analyticity for all $z \in \mathbb{H}^+$.

Solve to get

$$G_\mu(z) = \frac{z \pm \sqrt{z^2 - 4}}{2}.$$

To make sense of this, we first note that there is a branch of $\sqrt{z^2 - 4}$ defined on $\mathbb{C} \setminus [-2, 2]$ which gives the positive square-root on the positive x-axis and the negative square-root on the negative x-axis. If we choose this branch then the condition $zG_\mu(z)$ goes to 1 as $|z|$ gets large forces choice of the negative sign in $G_\mu(z)$.

It can now be verified that the function $g(t)$ associated to $G_\mu(z)$ is given by

$$g(t) = \begin{cases} \frac{t+\sqrt{t^2-4}}{2} & \text{if } t < -2 \\ \frac{t-i\sqrt{4-t^2}}{2} & \text{if } |t| \leq 2 \\ \frac{t-\sqrt{t^2-4}}{2} & \text{if } t > 2 \end{cases}$$

Finally, Stieltjes inversion gives the desired measure as was checked earlier.

8.4. **Analytic distribution of $u + u^*$ for a Haar unitary u.** Reconsider the example in §4.4 of Lecture 1. In this case

$$G_\nu(z) = \sum_{p\geq 0} \frac{\binom{2p}{p}}{z^{2p+1}} = \sum_{p\geq 0} \frac{(p+1)C_p}{z^{2p+1}}.$$

As before, consider the generating function

$$K(z) = \sum_{p\geq 0} \binom{2p}{p} z^p = \sum_{p\geq 0} (p+1)C_p z^p = (zH(z))' = H(z) + zH'(z).$$

We have $G_\nu(z) = \frac{1}{z}K(\frac{1}{z^2})$ while $H(z) = \frac{1\pm\sqrt{1-4z}}{2z}$. Hence $K(z) = \frac{\pm 1}{\sqrt{1-4z}}$. Thus $G_\nu(z) = \frac{\pm 1}{\sqrt{z^2-4}}$.

Choosing the same branch of $\sqrt{z^2-4}$ as in the previous example we see that the sign in $G_\nu(z)$ must be positive. Also as before the function $g(t)$ is given by

$$g(t) = \begin{cases} \frac{1}{-\sqrt{t^2-4}} & \text{if } t < -2 \\ \frac{1}{i\sqrt{4-t^2}} & \text{if } |t| \leq 2 \\ \frac{1}{\sqrt{t^2-4}} & \text{if } t > 2 \end{cases}$$

Finally, Stieltjes inversion gives the arcsine law obtained earlier.

Remark: Above is incorrect since g is not cts. on $\mathbb{R}$. Need to use Shaffe's theorem to do it right.

9. C^*-algebras and the continuous functional calculus

9.1. Definitions.

A **normed linear space** is a complex vector space A equipped with a **norm** $||\cdot|| : A \to \mathbb{R}_{\geq 0}$, i.e., a function that satisfies:

- Positive homogeneity: $||\alpha a|| = |\alpha|||a||$ for all $\alpha \in \mathbb{C}, a \in A$.
- Positive definiteness: $||a|| = 0$ if and only if $a = 0$.
- Sub-additivity: $||a+b|| \leq ||a|| + ||b||$ for all $a, b \in A$.

The normed linear space A is said to be a **Banach space** if further, it is

- Complete for the norm induced metric $d(a,b) = ||b-a||$.

A Banach space A is said to be a **Banach algebra** if it is equipped with an associative, bilinear multiplication map $A \times A \to A$ which satisfies:

- Sub-multiplicativity: $||ab|| \leq ||a||||b||$ for all $a, b \in A$.

A Banach algebra is said to be a **C^*-algebra** if it is equipped with a conjugate-linear, product-reversing involution $* : A \to A$ that satisfies:

- The C^*-identity: $||a^*a|| = ||a||^2$ for all $a \in A$.

C^*-algebras need not be unital but we will be interested only in unital C^*-algebras. The two basic examples of these are $\mathcal{C}(X)$ - the algebra of continuous, complex-valued functions on a compact, Hausdorff space X - and $\mathcal{L}(\mathcal{H})$ - the algebra of bounded, linear operators on a Hilbert space $\mathcal{H}$. The Gelfand-Naimark theorems assert that any commutative, unital C^*-algebra is of the form $\mathcal{C}(X)$ while any unital C^*-algebra is isometrically $*$-isomorphic to a norm-closed $*$-subalgebra of $\mathcal{L}(\mathcal{H})$.

If A is a unital C^*-algebra and $a \in A$, the **spectrum** of a, denoted $\sigma(a)$, is defined as

$$\sigma(a) = \{\lambda \in \mathbb{C} : \lambda.1 - a \text{ is not invertible in } A\}.$$

Theorem: $\sigma(a)$ is a non-empty compact subset of $\mathbb{C}$ contained in $\{z : |z| \leq ||a||\}$.

9.2. The continuous functional calculus.

The following result (or sometimes the map Φ featuring in it) is referred to as the **continuous functional calculus** for a.

Theorem: Let A be a unital C^*-algebra and $a \in A$ be a normal element. Then there is a unital $*$-homomorphism $\Phi : \mathcal{C}(\sigma(a)) \to A$ such that $\Phi(id) = a$ and $||\Phi(f)|| = ||f||_\infty$ for all $f \in \mathcal{C}(\sigma(a))$.

Note that the theorem asserts that Φ is an isometry for the $||\cdot||_\infty$ norm on $\mathcal{C}(\sigma(a))$ and the norm on A. Together with linearity, this implies that Φ is 1-1. Consider polynomial functions in z and $\overline{z}$ restricted to $\sigma(a)$, so that the function z corresponds to what is called id in the statement of the theorem. Since Φ is a unital $*$-homomorphism and its value on id is specified (to be a), its value on all polynomials in z and $\overline{z}$ is determined and by Stone-Weierstrass on the whole of $\mathcal{C}(\sigma(a))$. Thus Φ is unique. We will usually denote $\Phi(f)$ by $f(a)$.

The image of Φ is the unital C^*-subalgebra of A generated by a. For clearly, Φ being an isometry, its image is complete and thus norm closed in A and contains a. So it contains the unital C^*-subalgebra generated by a. But this C^*-algebra must contain all polynomials in a and a^* whose norm closure is the image of Φ.

Many simple and useful consequences follow. For instance, let a be a normal element of a unital C^*-algebra A. Then, $||a|| = ||a^*|| = sup\{|z| : z \in \sigma(a)\}$. Note that $||a|| = ||a^*||$ even for non-normal a in a unital C^*-algebra because $||a||^2 = ||a^*a|| \leq ||a^*||.||a||$ giving one inequality and use involutivity of $*$ for the other one. Also, a is self-adjoint $\Leftrightarrow \sigma(a) \subseteq \mathbb{R}$ and unitary $\Leftrightarrow \sigma(a) \subseteq S^1$.

9.3. **Continuous functional calculus for $\mathcal{C}(X)$.** Let $a \in \mathcal{C}(X)$ so that a is a continuous, complex-valued function on X. It is clear that $\sigma(a) = a(X)$. Consider the continuous functional calculus for a which is a unital $*$-homomorphism $\Phi : \mathcal{C}(\sigma(a)) = \mathcal{C}(a(X)) \to \mathcal{C}(X)$. We assert that $\Phi(f) = f \circ a$. For $f \mapsto f \circ a$ is a unital $*$-homomorphism such that $id \mapsto a$ and such that $||\Phi(f)|| = ||f||_\infty$. Now appeal to the uniqueness of Φ.

9.4. **Spectral mapping theorem. Theorem:** Let A be a unital C^*-algebra, $a \in A$ be a normal element and $f \in \mathcal{C}(\sigma(a))$. Then, $\sigma(f(a)) = f(\sigma(a))$.

To prove this, note first that $\lambda \in \sigma(f(a)) \Leftrightarrow 0 \in \sigma((f - \lambda)(a))$, while $\lambda \in f(\sigma(a)) \Leftrightarrow 0 \in (f - \lambda)(\sigma(a))$. So, replacing f by $f - \lambda$, it suffices to see that $f(a)$ non-invertible iff $0 \in f(\sigma(a))$. One direction of this is clear. If $0 \notin f(\sigma(a))$, f is a non-vanishing continuous function on $\sigma(a)$, hence $\frac{1}{f}$ is a continuous function on $\sigma(a)$ and $\frac{1}{f}(a)$ is an inverse of $f(a)$ by the continuous functional calculus, so that $f(a)$ is invertible.

Conversely suppose that $f(a)$ is invertible and that $0 \in f(\sigma(a))$ to derive a contradiction. Say $f(\lambda_0) = 0$ for $\lambda_0 \in \sigma(a)$. Take a continuous function g on $\sigma(a)$ of large norm, say K, such that fg has norm bounded by 1. To construct such a g, apply Urysohn's lemma to the function that is K at λ_0 and 0 outside a neighbourhood of λ_0 where f is less than $\frac{1}{K}$. Then, $K = ||g(a)|| = ||f(a)^{-1}f(a)g(a)|| \leq ||f(a)^{-1}||.||f(a)g(a)|| \leq ||f(a)^{-1}||$. Since K is arbitrary, we're done.

10. Positivity

10.1. **Basics.** An element a of a unital C^*-algebra A is said to be **positive** if it is self-adjoint and $\sigma(a) \subseteq \mathbb{R}_{\geq 0}$. The set of positive elements of A is denoted A_+ and we write $a \geq 0$ for $a \in A_+$. We assert that the set of positive elements in A is a pointed cone, i.e., p, q positive and $\alpha, \beta \geq 0$ implies that $\alpha p + \beta q$ positive and $p, -p$ positive implies $p = 0$.

The pointedness of the cone is clear. For if p and $-p$ are both positive then $\sigma(p) \subseteq \mathbb{R}_+ \cap \mathbb{R}_- = \{0\}$. So $||p|| = 0 \Rightarrow p = 0$.

To prove conicality, since αp is obviously positive for p positive, it suffices to see that the sum $p+q$ of positives p and q is positive. Now, for positive p, $\sigma(p) \subseteq [0, ||p||]$ and so $\sigma(||p|| - p) \subseteq [0, ||p||]$ too. It follows that $||p|| - p$ is (also positive) of norm at most $||p||$.

Hence $||(||p|| - p) + (||q|| - q)|| \leq ||p|| + ||q||$. Hence $\sigma(||p|| - p + ||q|| - q) \subseteq [-||p|| - ||q||, ||p|| + ||q||]$. So $\sigma(p + q) \subseteq [0, 2(||p|| + ||q||)] \subseteq \mathbb{R}_{\geq 0}$. Thus $p + q$ is positive.

The pointed conicality implies that the relation $\geq$ defined on the set of self-adjoint elements of A by $a \geq b$ if $a - b \geq 0$ is a partial order.

The spectral mapping theorem provides lots of positive elements. If A is a unital C^*-algebra, $a \in A$ is normal and $f : \sigma(a) \to \mathbb{R}_{\geq 0}$ is any continuous function, then $f(a)$ is a positive element of A.

10.2. **Another characterization of positivity.** We assert that A_+ consists exactly of those elements of A that are of the form a^*a for some $a \in A$.

One containment is obvious. Suppose that $p \in A_+$, so that p is self-adjoint and $\sigma(p) \subseteq \mathbb{R}_{\geq 0}$. Consider the continuous function $f = \sqrt{\cdot}$ defined on $\sigma(p)$ and let $a = f(p)$. Then $a = a^*$ and $a^2 = p$, by the continuous functional calculus.

As for the other, take a^*a which is clearly self-adjoint (so that $\sigma(a^*a) \subseteq \mathbb{R}$) and we need to see that $\sigma(a^*a) \subseteq \mathbb{R}_{\geq 0}$. Let $f(t) = max\{0, t\}, g(t) = max\{0, -t\}$, defined on $\sigma(a^*a)$, and let $x = f(a^*a), y = g(a^*a)$, both of which are positive elements of A. From the continuous functional calculus, $x - y = a^*a$ and $xy = 0 = yx$.

It now follows that with $b = ay$, $b^*b = ya^*ay = y(x-y)y = -y^3 \in -A_+$. Hence, so does bb^* - since $\sigma(ab) \cup \{0\} = \sigma(ba) \cup \{0\}$ for all $a, b \in A$. Thus $b^*b + bb^* \in -A_+$. But if $b = u + iv$ with self-adjoint u, v, then $b^*b + bb^* = 2(u^2 + v^2) \in A_+$. By pointedness of the cone of positives, $u^2 + v^2 = 0$. So $u^2 = -v^2 \in A_+ \cap -A_+ = \{0\}$. By the C^*-identity, $u = 0 = v$. Thus $b = 0 \Rightarrow y^3 = 0 \Rightarrow y = 0$. So $a^*a = x$ which is positive.

11. C^*-PROBABILITY SPACES

11.1. **Definitions and states.** The $*$-probability space (A, ϕ) is said be a C^*-**probability space** if A is a (unital) C^*-algebra.

For a unital C^*-algebra A, a linear functional $\psi : A \to \mathbb{C}$ is said to be a **state** if it is positive ($\psi(a^*a) \geq 0$ for all $a \in A$) and $\psi(1) = 1$. Thus, the ϕ for a C^*-probability space (A, ϕ) is a state.

We will use the fact that a linear functional ψ on a unital C^*-algebra A is a state iff it is bounded of norm 1 (i.e., $||\psi|| = sup\{\psi(a) : a \in A, ||a|| \leq 1\} = 1$) and $\psi(1) = 1$.

To prove this, suppose first that ψ is a state. It will suffice to see that $|\psi(a)| \leq ||a||$ for all $a \in A$. But by Cauchy-Schwarz, $|\psi(a)| = |\psi(1^*a)| \leq \psi(a^*a)^{\frac{1}{2}}$ - for note that in a C^*-algebra, $1^* = 1$. The element $p = a^*a$ is a positive element of A; if we knew that for positive p, $\psi(p) \leq ||p||$, we would be done by appeal to the C^*-identity. But this holds, since for positive p, $||p|| - p$ is also positive and just apply ψ.

Conversely suppose that ψ is a linear functional of norm 1 that is 1 at 1. We need to see that ψ is positive. Let $a \in A_+$. The basic idea is that $\phi(a)$ is contained in any closed disc in the complex plane that contains $\sigma(a)$. For any $\lambda \in \mathbb{C}$, we have $|\phi(a - \lambda)| \leq ||a - \lambda||$, i.e., $|\phi(a) - \lambda| \leq ||a - \lambda||$. Suppose that $\sigma(a) \subseteq [r, s] \subseteq \mathbb{R}_{\geq 0}$. Also say $[r, s] \subseteq B(\lambda, t)$ for some $\lambda \in \mathbb{C}, t \in \mathbb{R}$. Then $\sigma(a - \lambda) \subseteq B(0, t)$. Since $a - \lambda$ is normal, $||a - \lambda|| \leq t$. Hence $|\phi(a) - \lambda| \leq t \Rightarrow \phi(a) \in B(\lambda, t)$. Since $[r, s]$ is the intersection of all the $B(\lambda, t)$ containing $[r, s]$, it follows that $\phi(a) \in [r, s]$ and is therefore non-negative, as needed.

11.2. C^***-probability spaces associated to** $\mathcal{C}(X)$**.** Let X be a compact Hausdorff space and consider the C^*-algebra $A = \mathcal{C}(X)$ of complex-valued, continuous functions on X. (Or equivalently, by the Gelfand-Naimark theorem, consider a commutative, unital C^*-algebra A.) This can be made a C^*-probability space in a number of ways.

First, recall the Riesz representation theorem which asserts that states ϕ on $\mathcal{C}(X)$ are exactly those of the form

$$\phi(f) = \int_X f d\mu$$

where μ is a uniquely determined outer- and inner-regular Borel probability measure on X.

So take any such ϕ and consider (A, ϕ). This then is a C^*-probability space, and is the most general commutative example of such.

11.3. $C^*_{red}(G)$**.** Recall for a countable group G, the left-regular representation $\lambda : \mathbb{C}G \to \mathcal{L}(\ell^2(G))$ defined by linearly extending $\lambda(g)(\xi_h) = \xi_{gh}$. This is a representation of the $*$-probability space $(\mathbb{C}G, \phi)$ on $\ell^2(G)$ with vacuum vector ξ_1.

Let $C^*_{red}(G)$ be the norm closure of $\lambda(G)$ in $\mathcal{L}(\ell^2(G))$. Then $C^*_{red}(G)$ is a C^*-algebra. The functional ϕ on $\mathbb{C}G$ extends to one on $C^*_{red}(G)$ by the vector state defined by ξ_1 which is clearly a state and thus bounded and hence continuous. Thus it is still a trace on $C^*_{red}(G)$.

It happens to be still faithful on $C^*_{red}(G)$. Seeing this requires a commutant argument. Consider the right-regular representation of $\mathbb{C}G$ on $\mathcal{L}(\ell^2(G))$ defined by linear extension of $\rho(g)(\xi_h) = \xi_{hg^{-1}}$. It is clear that all elements of $\lambda(\mathbb{C}G)$ commute with every $\rho(g)$. It follows by norm approximation that so do all elements of $C^*_{red}(G)$. Now suppose that $\phi(T^*T) = 0$ for some $T \in C^*_{red}(G)$. We need to see that $T = 0$.

Since $\phi(T^*T) = ||T\xi_1||^2$, it follows that $T\xi_1 = 0$. Applying $\rho(g)$ for $g \in G$ and using commutativity with T gives $T\xi_g = 0$ for all $g \in G$. Hence T vanishes.

11.4. **Proof analysis and LG.** Analysing the proof shows why the vector state defined by ξ_1 is faithful even on LG. In fact, take any vector topology (i.e., addition and sclar multiplication being continuous) on $\mathcal{L}(\mathcal{H})$ such that the adjunction operation and multiplication by a fixed element are continuous in this topology. Then, the closure of $\mathbb{C}G$ in this topology is still a $*$-subalgebra of $\mathcal{L}(\mathcal{H})$ and the vector state defined by ξ_1 is faithful on this completion. Taking this topology to be the weak operator topology gives the result for LG. (LG which is defined to be the double commutant of $\lambda(\mathbb{C}G)$ is also the weak, strong or σ-weak operator topology closure of $\mathbb{C}G$.)

12. $*$-distributions, norm and spectrum of normal elements

12.1. **Existence of analytic $*$-distribution.** Let (A, ϕ) be a C^*-probability space and $a \in A$ be normal. Then a has an analytic $*$-distribution μ whose support is contained in $\sigma(a)$ and such that for all $f \in \mathcal{C}(\sigma(a))$ we have

$$\int_{\mathbb{C}} f d\mu = \phi(f(a)).$$

This is easy to see. For consider the continuous functional calculus, say $\Phi : \mathcal{C}(\sigma(a)) \to A$ and the composite map $\phi \circ \Phi : \mathcal{C}(\sigma(a)) \to \mathbb{C}$ which is a positive linear functional. The Riesz representation theorem now yields a Borel probability measure μ on $\sigma(a)$ (hence a compactly supported measure on $\mathbb{C}$ with support contained in $\sigma(a)$) such that the formula above holds. Taking f of the form $z^m\overline{z}^n$, $\Phi(f) = a^m(a^*)^n$ and so this μ is the analytic $*$-distribution of a.

A useful consequence is that any normal element a of any $*$-probability space (A, ϕ) that admits a representation on Hilbert space has an analytic $*$-distribution.

This also follows easily. For suppose that $\theta : A \to \mathcal{L}(\mathcal{H})$ is a representation with vacuum vector $\Omega \in \mathcal{H}$. The element $\theta(a)$ of the C^*-probability space $(\mathcal{L}(\mathcal{H}), \phi_\Omega)$ has an analytic $*$-distribution μ which is also the analytic $*$-distribution of a.

12.2. **The support of the analytic $*$-distribution.** Suppose that (A, ϕ) is a C^*-probability space and $a \in A$ is normal with analytic $*$-distribution μ. If ϕ is faithful, then, $supp(\mu) = \sigma(a)$.

For suppose that $\lambda \in \sigma(a) \setminus supp(\mu)$. Take an neighbourhood U of λ which has 0 μ-measure. Consider a continuous function $f : \sigma(a) \to [0, 1]$ that is 1 at λ and 0 outside U and let $b = f(a)$ so that $||b|| = 1$. However, $\phi(b^*b) = \phi(b^2) = \int_{\mathbb{C}} f^2 d\mu \leq \int_{\mathbb{C}} 1_U d\mu = 0$. Hence $\phi(b^*b) = 0$ and so $b = 0$ by faithfulness of ϕ. The contradiction shows non-existence of λ.

The reason this is useful is to be able to read off properties of a from that of its $*$-distribution. For instance we see that the spectrum of a is determined by its $*$-distribution with respect to a faithful state. Hence so is its norm.

We note that conversely, if $supp(\mu) = \sigma(a)$, then, ϕ is faithful on the C^*-subalgebra of A generated by a. To see this it suffices to see, using the continuous functional calculus, that if $X \subset \mathbb{C}$ is compact and μ is a measure on $\mathbb{C}$ with support X, then, for any continuous function f on X, if $\int_{\mathbb{C}} |f|^2 d\mu = 0$, then, $f = 0$ which is obvious.

12.3. **The norm from the $*$-distribution.** There is a very direct formula that yields the norm of any element of a faithful C^*-probability space (A, ϕ) from its $*$-moments. Explicitly, we assert that

$$||a|| = lim_{n\to\infty} \phi((a^*a)^n)^{\frac{1}{2n}},$$

for any $a \in A$ (not necesarily normal).

Applying the C^*-identity, it suffices to see that

$$||p|| = lim_{n\to\infty} \phi(p^n)^{\frac{1}{n}},$$

for a positive $p \in A$. Positivity of p and submultiplicativity of the norm imply that the sequence on the RHS is a sequence of positive real numbers each less than $||p||$.

So it is enough to see that for any $\alpha \in (0, ||p||)$ the inequality $\phi(p^n)^{\frac{1}{n}} > \alpha$ holds for all sufficiently large n. Note first that $||p|| \in \sigma(p) = supp(\mu)$, where μ is the $*$-distribution of p. So for any $\beta < ||p||$, the interval $[\beta, ||p||]$ has positive μ-measure. Choose any such $\beta > \alpha$ and consider

$$\phi(p^n) = \int_{\sigma(p)} t^n d\mu \geq \beta^n \mu([\beta, ||p||]).$$

So $\phi(p^n)^{\frac{1}{n}} \geq \beta\mu([\beta, ||p||])^{\frac{1}{n}} > \alpha$ if n is large enough.

13. Joint distributions and $*$-distributions

13.1. **Definitions.** Let (A, ϕ) be a probability space and $a_1, \cdots, a_s \in A$. A **joint moment** of $a_1, \cdots, a_s$ is ϕ evaluated on a word in $a_1, \cdots, a_s$. The **joint distribution** of $a_1, \cdots, a_s$ is the linear functional $\mu : \mathbb{C}\langle X_1, \cdots, X_s\rangle \to \mathbb{C}$ defined by

$$\mu(X_{i_1} X_{i_2} \cdots X_{i_n}) = \phi(a_{i_1} a_{i_2} \cdots a_{i_n}),$$

where $\mathbb{C}\langle X_1, \cdots, X_s\rangle$ is the free, associative, unital algebra generated by the indeterminates $X_1, X_2, \cdots, X_s$.

If (A, ϕ) is a $*$-probability space and $a_1, \cdots, a_s \in A$, a **joint $*$-moment** of $a_1, \cdots, a_s$ is ϕ evaluated on a word in $a_1, \cdots, a_s, a_1^*, \cdots a_s^*$. The **joint $*$-distribution** of $a_1, \cdots, a_s$ is a linear functional defined on the free, associative, unital algebra $\mathbb{C}\langle X_1, \cdots, X_s, X_1^*, \cdots, X_s^*\rangle$ defined analogously. Note that $\mathbb{C}\langle X_1, \cdots, X_s, X_1^*, \cdots, X_s^*\rangle$ has a natural $*$-structure defined in the obvious way and that the joint $*$-distribution preserves adjoints.

Both these can be stated in terms of the evaluation homomorphisms $\mathbb{C}\langle X_1, \cdots, X_s\rangle \to A$ in the first case and and $\mathbb{C}\langle X_1, \cdots, X_s, X_1^*, \cdots, X_s^*\rangle \to A$, in the second.

Why we need joint distributions and $*$-distributions is exactly the same reason as in classical probability: knowing individual distributions of random variables does not give information on the distribution of their sum for instance.

14. Examples

14.1. **The Cayley graph.** Let G be a countable group and $g, h \in G$ be elements of infinite order so that they are Haar unitaries in $(\mathbb{C}G, \phi)$ and so the self-adjoint elements $x = g + g^{-1}$ and $y = h + h^{-1}$ of $\mathbb{C}G$ both have the arcsine distribution. Consider the distribution of their sum $\Delta = g + g^{-1} + h + h^{-1}$.

Define the **Cayley graph** of a group G with respect to a set S of generators (S closed under inversion and not containing 1) as the graph with vertex set G and edge between g and h if $g = hs$ for some $s \in S$. Though this is undirected, it is sometimes convenient to regard it as a directed graph where each edge is directed both ways and labelled such that the label from g to gs is s (and from gs to g is s^{-1}, of course).

If G is a finite group and S is the set of non-identity elements of G, the Cayley graph is the complete graph on $|G|$ vertices. If G is $\mathbb{Z}^2$ and $S = \{(\pm 1, 0), (0, \pm 1)\}$, the Cayley graph is the union of the lines $x = n, y = n$ for $n \in \mathbb{Z}$. If G is F_2 - the free group on 2 generators - g, h, say, the Cayley graph with respect to $S = \{g, g^{-1}, h, h^{-1}\}$ is the 4-regular tree. If G is Σ_3 with $S = \{(12), (123), (132)\}$ the Cayley graph is like the edges of a triangular prism.

Suppose that G is generated by g and h. We claim that for any $n \in \mathbb{N}$, $\phi(\Delta^n)$ is given by the number of closed paths in the Cayley graph that begin and end at 1. This is fairly clear. For, $\Delta^n = \sum_{k_1, k_2, \cdots, k_n \in S} k_1 k_2 \cdots k_n$. So $\phi(\Delta^n) = \sum_{k_1, k_2, \cdots, k_n \in S} \delta_{k_1 k_2 \cdots k_n, 1} = |\{(k_1, k_2, \cdots, k_n) \in S^n : k_1 k_2 \cdots k_n = 1\}|$.

Associating to the n-tuple $(k_1, k_2, \cdots, k_n)$ the path which visits in order 1, k_1, $k_1 k_2$, $\cdots$, $k_1 k_2 \cdots k_{n-1}$, 1 gives the desired bijection. In the reverse direction read off the labels of the visited edges in order and keeping track of direction.

14.2. **The $\mathbb{Z}^2$ example.** To count the number of paths on the integer lattice that begin and end at the origin, encode each such by a string of R, L, U, D with the obvious meaning.

A string of R, L, U, D represents such a path exactly when the number of R's equals the number of L's and the number of U's equals the number of D,s. In particular, there are no such paths of odd length.

If the length of the path is $2p$ associate to it the two subsets of $\{1, 2, \cdots, 2p\}$ which correspond to its R or U positions and R or L positions. Each is a p-subset. Conversely given any two p-subsets of $\{1, 2, \cdots, 2p\}$ put R in the positions of their intersection, U in the complement of the R positions in the first and L in the complement of the R positions in the second and D everywhere else. This corresponds to an allowable path. Thus, the number of such is $\binom{2p}{p}^2$.

What is the moment generating function ?

14.3. **The F_2 example.** Consider starting off at 1 on the Cayley graph of F_2, picking an edge incident there at random, going along it and continuing this process. Let u_k be the probability of returning to 1 after k steps and v_k be the probability of returning to 1 for the first time after k steps. Consider the generating functions

$$\begin{aligned} U(z) &= \sum_{k=0}^{\infty} u_k z^k, \text{and} \\ V(z) &= \sum_{k=0}^{\infty} v_k z^k, \end{aligned}$$

where we decree that $u_1 = 1$ while $v_1 = 0$.

Any path that returns to 1 after k steps has a first return after some t steps for $t = 1, 2, \cdots, k$, and is followed by a path that returns to 1 after $k - t$ steps. It follows that $u_k = \sum_{t=1}^{k} v_t u_{k-t}$ for $k \geq 1$ and therefore that $U(z) = U(z)V(z) + 1$. Thus knowing one determines the other.

We will try to determine V for the Cayley graph of F_2 which, we have observed is the 4-regular tree. More generally, we do this for the $d \geq 2$ regular tree. Choose any vertex as a root and draw it growing downward, so that there are d children of the root and $d - 1$ for any other vertex.

Consider paths of length k based at the root which have first return exactly after k steps. The ratio of their number to the total number of paths of length k based at the root gives v_k. To count them, associate two strings of length k to them. One is the sequence of U/D moves. The other is the choice of child at each D move.

The first sequence must begin with a D, end with a U and be a Catalan sequence of length $k - 2$ in between. The other sequence has $1, \cdots, d$ for the first D and $1, \cdots, d - 1$ for any of the other D's. Any such pair of sequences clearly gives an eligible path. Thus the number of such paths is 0 for odd k and for $k = 2p$ is given by:

$$d(d-1)^{p-1} C_{p-1} = d(d-1)^{p-1} \frac{1}{p} \binom{2p-2}{p-1}.$$

Thus

$$\begin{aligned} V(z) &= \sum_{p=1}^{\infty} \frac{d(d-1)^{p-1} C_{p-1}}{d^{2p}} z^{2p} \\ &= \frac{z^2}{d} \sum_{p=1}^{\infty} (\frac{d-1}{d^2})^{p-1} C_{p-1} z^{2p-2} \\ &= \frac{z^2}{d} \sum_{p=0}^{\infty} C_p (\frac{(d-1)z^2}{d^2})^p. \end{aligned}$$

Recalling that the generating function for the Catalans is given by $\frac{1-\sqrt{1-4z}}{2z}$, we see that $V(z)$ is given by

$$\frac{d}{2(d-1)}\left[1-\sqrt{1-4\frac{d-1}{d^2}z^2}\right].$$

It follows that $U(z)$ is given (after a bit of calculation by)

$$\frac{2(d-1)}{d-2+\sqrt{d^2-4(d-1)z^2}} = \frac{3}{1+\sqrt{4-3z^2}},$$

when $d = 4$. Thus the generating function for the number of loops based at 1 is given by $U(dz)$ which is

$$\frac{2(d-1)}{d-2+d\sqrt{1-4(d-1)z^2}} = \frac{3}{1+2\sqrt{1-12z^2}},$$

for $d = 4$.

14.4. **The rotation algebra.** Consider a $*$-algebra A generated by two unitaries x and y subject only to the relation $xy = e^{i\theta}yx$ where $\theta \in \mathbb{R}$ is a fixed parameter. What this means in greater detail is that A is a quotient of the free algebra on 4 variables X, X^*, Y, Y^* by the two sided ideal, say I, generated by $XX^*-1, X^*X-1, YY^*-1, Y^*Y-1$ and $XY-e^{i\theta}YX$. Since I is $*$-stable - each one of its generators is so, upto scaling - the natural $*$-structure on the free algebra on X, X^*, Y, Y^* descends to A, so that A is a $*$-algebra.

The element $XY - e^{i\theta}YX \in I$ may be used to see that $XY^* - e^{-i\theta}Y^*X$ is also in I. It follows that the elements x^my^n span A where $m, n \in \mathbb{Z}$. We will show that these are also linearly independent. Assuming this, define ϕ on A by setting $\phi(x^my^n) = \delta_{m,0}\delta_{n,0}$. We show later that ϕ is positive and faithful.

The multiplication rule $(x^my^n)(x^py^q) = e^{-inp\theta}x^{m+p}y^{n+q}$ is easily seen and seen to imply that ϕ is a trace.

We assert the following combinatorial interpretation of ϕ evaluated on a monomial in x, x^*, y, y^*. To any such monomial, associate a lattice path that begins at the origin and moves one step east, west, north or south according as x, x^*, y, y^* occurs in the monomial. If m is a monomial of degree n, the associated lattice path γ is composed of n linear pieces. We claim that $\phi(m) = 0$ unless γ is a closed path, in which case, $\phi(m) = e^{ik\theta}$ where k is given by

$$k = \int_\gamma x\ dy = -\int_\gamma y\ dx.$$

To prove this, define a linear functional $\Phi : \mathbb{C}\langle X, X^*, Y, Y^*\rangle \to \mathbb{C}$, by the same prescription. On the free algebra, this is certainly well-defined. We claim that this functional descends to A.

To see this it is enough to verify that the functional vanishes on the ideal I. A spanning set for I is given by all binomials of the form w_1bw_2 where w_1, w_2 are arbitrary monomials in X, X^*, Y, Y^* and b is one of the 5 generating binomials of I.

To see that Φ vanishes on, say $w_1(XX^*-1)w_2$, we need to check that $\Phi(w_1XX^*w_2) = \Phi(w_1w_2)$. Note that the finish points of the paths associated to both $w_1XX^*w_2$ and to w_1w_2 are the same so unless this is the origin, both sides vanish. Otherwise, by definition, each of these is an integral over an associated path and a little thought shows that these integrals are indeed equal as also for the other 3 such generators.

The interesting case is for the generator $b = XY - e^{i\theta}YX$. We need to check in this case that $\Phi(w_1XYw_2) = e^{i\theta}\Phi(w_1YXw_2)$. Some thought shows that if (m,n) is the point obtained by traversing either path until w_1, the difference between the integrals for w_1XYw_2 and for w_1YXw_2 is given by

$$\int_\gamma x\ dy,$$

where γ is the unit square with lower left endpoint at (m,n) traversed anticlockwise. Since this clearly gives 1, we're done.

15. The isomorphism theorems

15.1. **The vanilla isomorphism theorem.** The importance of joint $*$-distributions comes from the following theorem and its generalisations to the C^*-algebraic and von Neumann algebraic context.

Theorem: Let (A, ϕ_A) and (B, ϕ_B) be faithful $*$-probability spaces such that there exist $a_1, \cdots, a_s \in A$ and $b_1, \cdots, b_s \in B$ which are generating sets of A and$B$ respectively as unital $*$-algebras and such that the joint $*$-distributions μ_A of $a_1, \cdots, a_s$ and μ_B of $b_1, \cdots, b_s$ are the same. There is then a unique unital $*$-isomorphism $\Phi : A \to B$ such that $\Phi(a_i) = b_i$ for $i = 1, 2, \cdots, s$, which also is an isomorphism of $*$-probability spaces.

To see this, consider the map $\Theta_A = ev_{a_1,\cdots,a_s} : \mathbb{C}\langle X_1, \cdots, X_s, X_1^*, \cdots, X_s^*\rangle \to A$ defined by $\Theta_A(X_i) = a_i, \Theta_A(X_i^*) = a_i^*$. The generation statement says that Θ_A is a surjective map. Similarly, we have surjective Θ_B. The main observation is that $ker(\Theta_A) = ker(\Theta_B)$. This is because, by definition, $\phi_A \circ \Theta_A = \phi_B \circ \Theta_B$. Thus, if $f \in ker(\Theta_A)$, then $f^*f \in ker(\Theta_A) \Rightarrow \phi_A(f^*f(\underline{a})) = 0 \Rightarrow \mu_A(f^*f) = 0 = \mu_B(f^*f) \Rightarrow \phi_B(f^*(\underline{b})f(\underline{b})) = 0 \Rightarrow f(\underline{b}) = 0 \Rightarrow f \in ker(\Theta_B)$, and conversely. So $A \cong B$ by the map taking $f(\underline{a})$ to $f(\underline{b})$. Clearly $\phi_A \circ f(\underline{a}) = \phi_B \circ f(\underline{b})$.

15.2. **On $*$-homomorphisms.** Let $\Phi : A \to B$ be a unital $*$-homomorphism of unital C^*-algebras. Then Φ decreases norm and, in particular, is continuous. For, Φ certainly decreases spectrum. So it decreases norm for normal elements and the C^*-identity implies that this holds also for general elements.

In particular, a unital $*$-isomorphism preserves norm. In fact, even an injective unital $*$-homomorphism preserves norm. To see this, as before, it suffices to see that the norm of positive elements is preserved. So suppose that $||\Phi(p)|| < ||p||$.

Note that for any $f \in \mathcal{C}(\sigma(p))$, $f(\Phi(p))$ makes sense and equals $\Phi(f(p))$. This is clear for polynomial f in z and $\overline{z}$. Now approximate arbitrary f by polynomials on $\sigma(p)$ and use that Φ is continuous.

Choose for f, by Urysohn's lemma, a continuous extension from $\sigma(p)$ to $[0,1]$ of a function that is 0 on $\sigma(\Phi(p))$ and 1 at $||p||$. Then $f(\Phi(p)) = 0$. So $\Phi(f(p)) = 0 \Rightarrow f(p) = 0$ contradicting its norm being 1.

15.3. **The C^*-isomorphism theorem.** The C^*-algebraic version of the previous theorem reads as follows.

Theorem: Let (A, ϕ_A) and (B, ϕ_B) be faithful C^*-probability spaces such that there exist $a_1, \cdots, a_s \in A$ and $b_1, \cdots, b_s \in B$ which are generating sets of A and$B$ respectively as C^*-algebras and such that the joint $*$-distributions μ_A of $a_1, \cdots, a_s$

and μ_B of $b_1, \cdots, b_s$ are the same. There is then a unique unital isometric $*$-isomorphism $\Phi : A \to B$ such that $\Phi(a_i) = b_i$ for $i = 1, 2, \cdots, s$, which also is an isomorphism of C^*-probability spaces.

To prove this, note that by the previous theorem, there is an isomorphism, say Φ_0, of $*$-probability spaces A_0 and B_0 given by taking $f(\underline{a})$ to $f(\underline{b})$, where A_0 and B_0 are the $*$-algebras generated by $a_1, \cdots, a_s$ and $b_1, \cdots, b_s$. The point is that Φ_0 is isometric. To prove this, recall that the norm of any a in A can be computed in terms of its $*$-distribution explicitly. It follows that $f(\underline{a})$ and $f(\underline{b})$ have the same norm.

Now, by norm density of A_0 in A and completeness of B, Φ_0 has a unique continuous extension to $\Phi : A \to B$ which is also isometric. Norm continuity of the multiplication and $*$-operation on A imply that Φ is a unital $*$-homomorphism. Now $\phi_B \circ \Phi = \phi_A$ on A_0. Both sides being continuous, this holds also on A. Uniqueness is clear since A_0 is dense in A and by continuity of $*$-homomorphisms.

15.4. **Representing the rotation algebra.** Consider the group $\mathbb{Z}^2$ and the associated Hilbert space $\ell^2(\mathbb{Z}^2)$ with orthonormal basis $\xi_{(m,n)}$ for $(m,n) \in \mathbb{Z}^2$. Define unitaries $x, y \in \mathcal{L}(\ell^2(\mathbb{Z}^2))$ by

$$\begin{aligned} x\xi_{(m,n)} &= \xi_{(m+1,n)} \\ y\xi_{(m,n)} &= e^{-im\theta}\xi_{(m,n+1)}. \end{aligned}$$

Note that there is a surjective $*$-homomorphism, say Θ, from A to the unital subalgebra of $\mathcal{L}(\ell^2(\mathbb{Z}^2))$ generated by x and y. Since the elements $x^m y^n$ for $m, n \in \mathbb{Z}$ are linearly independent - operating on $\xi_{(0,0)}$, they give $\xi_{(m,n)}$ - it follows that $X^m Y^n \in A$ are also linearly independent.

Let $\tilde{A}$ be the C^*-algebra generated by x and y in $\mathcal{L}(\ell^2(\mathbb{Z}^2))$ which is the norm closure of $\Theta(A)$. Let ψ be the vector state defined on $\tilde{A}$ by $\xi_{(0,0)}$. Since ψ is a trace on A which is norm dense in $\tilde{A}$, ψ is also a trace on A.

Further, ψ is faithful. For suppose there is a $w \in \tilde{A}$ such that $\psi(w^*w) = 0$. Then $w\xi_{(0,0)} = 0$. Now, $\langle w\xi_{(m,n)} | \xi_{(p,q)} \rangle = \langle wx^m y^n \xi_{(0,0)} | x^p y^q \xi_{(0,0)} \rangle = \langle (y^*)^q (x^*)^p w x^m y^n \xi_{(0,0)} | \xi_{(0,0)} \rangle = \psi((y^*)^q (x^*)^p w x^m y^n) = \psi(x^m y^n (y^*)^q (x^*)^p w) = 0$. The map Θ from (A, ϕ) to $(\tilde{A}, \psi)$ is a representation.

The point of all this is that according to the C^*-isomorphism theorem, the C^*-probability space generated by a pair of unitaries x and y satisfying the relation $xy = e^{-i\theta} yx$ an equipped with a faithful, positive ϕ with $\phi(x^m y^n) = \delta_{m,0}\delta_{n,0}$ is uniquely determined.

16. Classical independence and free independence

16.1. **Classical independence.** Let (A, ϕ) be a probability space and $\{A_i : i \in I\}$ be a family of unital subalgebras of A. The family $\{A_i : i \in I\}$ is said to be **classically independent** or **tensor independent** in A if elements of distinct subalgebras A_i and A_j commute and if ϕ admits a factorisation as follows: for any finite subset $J \subseteq I$, $\phi(\prod_{j \in J} a_j) = \prod_{j \in J} \phi(a_j)$.

Random variables $a, b \in A$ are said to be tensor independent if the unital subalgebras of A that they generate are tensor independent. Equivalently, $a, b \in A$ are tensor independent if they commute and $\phi(a^n b^m) = \phi(a^n)\phi(b^m)$ for all $m, n \geq 0$. In particular, when a and b are tensor independent, their joint distribution is completely determined by the individual distributions of a and b.

16.2. **Free independence.** Let (A, ϕ) be a probability space and $\{A_i : i \in I\}$ be a family of unital subalgebras of A. The family $\{A_i : i \in I\}$ is said to be **freely independent** in A if every alternating product of centred elements of A_i is centred, in the following sense. An element $a_i \in A_i$ is said to be **centred** if $\phi(a_i) = 0$. Recall that $\phi(a_i)$ is the mean of a_i. Consider any finite product $a_1 a_2 \cdots a_k$ where, for each $j = 1, 2, \cdots, k$, $a_j \in A_{i(j)}$. It is said to be an **alternating product** if $i(1) \neq i(2) \neq i(3) \neq \cdots \neq i(k)$.

A family $\{X_i : i \in I\}$ of subsets of A is said to be a **freely independent family** if the unital subalgebras they generate are freely independent in A. In particular, random variables $a_i \in A$ for $i \in I$ are said to be **freely independent** if the unital subalgebras they generate are freely independent in A.

If (A, ϕ) is a $*$-probability space, a family $\{X_i : i \in I\}$ of subsets of A is said to be a **$*$-freely independent family** if the unital $*$-subalgebras they generate are freely independent in A. Similarly, random variables $a_i \in A$ for $i \in I$ are said to be **$*$-freely independent** if the unital $*$-subalgebras they generate are freely independent in A.

16.3. **Some remarks.** (1) Free independence depends on the linear functional ϕ. Consider, for instance, the algebra $A = \mathbb{C}\langle X, Y\rangle$ with the functional ϕ being defined by evaluation at $X = 0, Y = 0$. The subalgebras $\mathbb{C}\langle X\rangle$ and $\mathbb{C}\langle Y\rangle$ are then freely independent in (A, ϕ). However, if ψ is defined by arbitrarily extending $X, Y \mapsto 0, XY \mapsto 1$, then clearly $\mathbb{C}\langle X\rangle$ and $\mathbb{C}\langle Y\rangle$ are not freely independent in (A, ψ).
(2) By definition, a family $\{A_i : i \in I\}$ is freely independent if and only if each of its finite subfamilies is freely independent.
(3) If (A, ϕ) is a C^*-probability space, $\{A_i : i \in I\}$ is a family of unital $*$-subalgebras of A, and B_i is the norm closure of A_i in A, then, $\{A_i : i \in I\}$ is a freely independent family if and only if $\{B_i : i \in I\}$ is a freely independent family. One direction is clear since free independence is monotonic in an obvious sense. As for the other, it suffices to observe that any centred element of B_i is a norm limit of centred elements of A_i and that ϕ is norm continuous as is multiplication jointly in its arguments.

17. Free independence for group algebras

17.1. **Freeness for subgroups.** Let G be a group and $\{G_i : i \in I\}$ be a family of subgroups of G. The family $\{G_i : i \in I\}$ is said to be **free in** G if for all $k \geq 1$, $i_i, i_2, \cdots, i_k \in I$ with $i_i \neq i_2 \neq \cdots \neq i_k$ and $g_j \in G_{i_j} \setminus \{1\}$, we have $g_1 g_2 \cdots g_k \neq 1$.

17.2. **The free product construction for groups.** Let $G_i, i \in I$ be a family of groups. There is a unique (upto unique isomorphism) pair of group G and homomorphisms $\phi_i : G_i \to G, i \in I$ that satisfies the following universal property: for any group H and homomorphisms $\psi_i : G_i \to H$, there is a unique homomorphism $\phi : G \to H$ such that $\psi_i = \phi \circ \phi_i$ for all $i \in I$. G is said to be the **free product** of the G_i and denoted by $*_{i\in I} G_i$.

A very standard proof shows that if the pair $(G, \phi_i : G_i \to G, i \in I)$ exists, then it is unique upto unique isomorphism.

To show existence some work is needed. Let $A = \coprod_{i\in I} G_i$, which is a set. Set $W(A)$ to be the word monoid of A, i.e., $W(A)$ consists of all strings of elements of A including the empty string with the operation being concatenation. Define an equivalence relation $\sim$ on $W(A)$ as the one generated by all

$$\begin{aligned} w1_i\tilde{w} &\sim w\tilde{w}, \\ wa_ib_i\tilde{w} &\sim wc_i\tilde{w}, \end{aligned}$$

where $1_i \in G_i$ is the identity element and $a_i, b_i \in G_i$ with $a_ib_i = c_i$ in G_i. Let G be the quotient $W(A)/\sim$. It is easy to check that G is a group. Let $\phi_i : G_i \to G$ be the natural map $\phi_i(a_i) = [a_i]$. This is a group homomorphism. Finally, the universal property also holds as is easily seen.

17.3. **Reduced words.** Any element of G contains a unique **reduced word**. This is a word in $W(A)$ which is either the empty word or of the form $a_1a_2\cdots a_n$ where $n \geq 1$, $a_j \in G_{i_j} \setminus \{1_{i_j}\}$, $i_j \neq i_{j+1}$ for $j = 1, 2, \cdots, n-1$. To see this, fix an $a \in A$ and consider the map $T(a)$ mapping the set of reduced words to itself defined by:

$$T(a)(a_1\cdots a_n) = \begin{cases} a_1\cdots a_n & \text{if } a = 1_i \text{ for some } i \in I \\ aa_1\cdots a_n & \text{if } a \in G_i \setminus \{1_i\} \text{ for some } i \in I \text{ with } i \neq i_1 \\ (aa_1)a_2\cdots a_n & \text{if } a \in G_{i_1} \text{ with } aa_1 \neq 1_{i_1} \\ a_2\cdots a_n & \text{if } a \in G_{i_1} \text{ and } aa_1 = 1_{i_1} \end{cases}$$

Note that $T(a)w \sim aw$ and is reduced, so by induction on the length of w any element of G contains a reduced word. As for uniqueness, for any word $w = b_1b_2\cdots b_n$, not necessarily reduced, define $T(b) = T(b_1)T(b_2)\cdots T(b_n)$, mapping the set of reduced words to itself. Observe that $w \mapsto T(w)$ respects the equivalence relation $\sim$ and so $w_1 \sim w_2 \Rightarrow T(w_1) = T(w_2)$. Also observe that w reduced implies that $T(w)(\epsilon) = w$, where ϵ is the empty word. Hence if w_1 and w_2 are both reduced and equivalent, then $w_1 = T(w_1)(\epsilon) = T(w_2)(\epsilon) = w_2$.

In particular, the canonical maps $\phi_i : G_i \to G$ are injective and we may identify G_i with a subgroup of G. These subgroups are indeed free in G.

17.4. **Motivation for freeness of probability spaces.** Let G be a group and $\{G_i : i \in I\}$ be a family of subgroups of G. Let (A, ϕ) be the probability space $A = \mathbb{C}G, \phi = \delta_1$, and $A_i = \mathbb{C}G_i$ for $i \in I$. Then, the family $\{G_i : i \in I\}$ is free in G exactly when the family $\{A_i : i \in I\}$ is freely independent in A.

First suppose that $\{G_i : i \in I\}$ is free in G. Take an alternating product $a_1a_2\cdots a_k$ of centred elements where $a_j \in A_{i(j)}$. Write $a_j = \sum_{g \in G_{i(j)}} a_j^g g$. That a_j is centred is equivalent to $a_j^1 = 0$. The product $a_1a_2\cdots a_k$ is then equal to

$$\sum_{g_1 \in G_{i(1)}\setminus\{1\}, \cdots, g_k \in G_{i(k)}\setminus\{1\}} a_1^{g_1}\cdots a_k^{g_k} g_1g_2\cdots g_k$$

By freeness of G_i in G, the coefficient of 1 here vanishes. The other direction is even simpler.

18. Free independence and joint moments

18.1. **How freeness determines joint distributions from individual distributions.** Let (A, ϕ) be a probability space and $\{A_i : i \in I\}$ be a family of freely independent subalgebras of A. Let B be the unital subalgebra of A generated by all the A_i. Then, $\phi|_B$ is determined by $\phi|_{A_i}$ for $i \in I$.

To see this, note first that B is spanned by all words $a_1a_2\cdots a_k$ where $a_j \in A_{i_j}$ where we may assume that $i_1 \neq i_2 \neq \cdots \neq i_k$. Now induce on k, the basis case

$k = 1$ being clear. For larger k, set $a_i^0 = a_i - \phi(a_i)$. We then have that

$$\phi(a_1 a_2 \cdots a_k) = \phi(\ (a_1^0 + \phi(a_1))(a_2^0 + \phi(a_2)) \cdots (a_k^0 + \phi(a_k))\ ).$$

The argument of ϕ in the RHS is a sum of 2^k terms each of which except for $a_1^0 a_2^0 \cdots a_k^0$ is a word with smaller k on which ϕ is determined inductively. Since $\phi(a_1^0 a_2^0 \cdots a_k^0) = 0$ by freeness, we are done.

18.2. **Explicit computation.** Say A and B are subalgebras that are freely independent in some larger probability space with functional ϕ. Then, for $a, a_1, a_2 \in A$ and $b, b_1, b_2 \in B$ we have (1) $\phi(ab) = \phi(a)\phi(b)$, (2) $\phi(a_1 b a_2) = \phi(a_1 a_2)\phi(b)$ and (3) $\phi(a_1 b_1 a_2 b_2) = \phi(a_1 a_2)\phi(b_1)\phi(b_2) + \phi(a_1)\phi(a_2)\phi(b_1 b_2) - \phi(a_1)\phi(a_2)\phi(b_1)\phi(b_2)$.

For, $\phi((a - \phi(a))(b - \phi(b))) = 0$ which gives (1). As for (2), consider

$$\phi((a_1 - \phi(a_1))(b - \phi(b))(a_2 - \phi(a_2))) = 0.$$

This implies $\phi(a_1 b a_2) - \phi(a_1 b)\phi(a_2) - \phi(b)\phi(a_1 a_2) + \phi(a_1)\phi(b)\phi(a_2) - \phi(a_1)\phi(b a_2) + \phi(a_1)\phi(b)\phi(a_2) + \phi(a_1)\phi(b)\phi(a_2) - \phi(a_1)\phi(b)\phi(a_2) = 0$, yielding (2). To prove (3), start with

$$\phi((a_1 - \phi(a_1))(b_1 - \phi(b_1))(a_2 - \phi(a_2))(b_2 - \phi(b_2))) = 0$$

and compute.

19. Properties of free independence

19.1. **Simple properties.** (1) Commuting random variables a and b are freely independent only if at least one has vanishing variance. For free independence implies that $\phi(abab) = \phi(a^2)\phi(b)^2 + \phi(a)^2\phi(b^2) - \phi(a)^2\phi(b)^2$ and commutativity with free independence that the LHS is $\phi(a^2)\phi(b^2)$. Taking all terms to the right and factoring yields $(\phi(a^2) - \phi(a)^2)(\phi(b^2) - \phi(b)^2) = 0$ as desired.
(2) In particular, real valued random variables a and b are freely independent only if at least one is constant.
(3) In a $*$-probability space (A, ϕ) with faithful ϕ, for a freely independent family $\{A_1, A_2\}$, $A_1 \cap A_2 = \mathbb{C}$. For if $a \in A_1 \cap A_2$ is self-adjoint, it is free from itself and so $\phi(a^2) - \phi(a)^2 = 0 \Rightarrow \phi((a^* - \phi(a^*))(a - \phi(a))) = 0$. Faithfulness of ϕ gives $a = \phi(a) \in \mathbb{C}$.
(4) In any probability space (A, ϕ), any subalgebra B is freely independent from $\mathbb{C}$.

19.2. **Free independence and traciality.** We first observe the following. Let (A, ϕ) be a probability space and $\{A_i : i \in I\}$ be a freely independent family of subalgebras of A. Let $a_1 a_2 \cdots a_k$ and $b_1 b_2 \cdots b_l$ be centred alternating products such that $a_t \in A_{i_t}$ and $b_t \in A_{j_t}$. Then $\phi(a_1 a_2 \cdots a_k b_l b_{l-1} \cdots b_1) = 0$ unless $k = l$ and $i_t = j_t$ for all t in which case it gives $\phi(a_1 b_1) \cdots \phi(a_k b_k)$.

The proof is by induction on k after observing that the whole product is an alternating product unless $i_k = j_l$ in which case it gives $\phi(a_k b_l)\phi(a_1 a_2 \cdots a_{k-1} b_{l-1} \cdots b_1)$.

Suppose that in the situation above, each $\phi|_{A_i}$ is a trace. Then, restricted to the algebra generated by the A_i, ϕ is a trace. This follows since the centred alternating products together with 1 are a basis of the generated algebra and we may use the observation above.

19.3. **Associativity of free independence.** Suppose that (A, ϕ) is a probability space and $\{A_i : i \in I\}$ be a family of unital subalgebras of A. For each $i \in I$ let $\{B_i^j : j \in J(i)\}$ be a family of unital subalgebras of A_i that generate A_i. The following are then equivalent:
(1) The family $\{A_i : i \in I\}$ is a freely independent family and for each $i \in I$ the family $\{B_i^j : j \in J(i)\}$ is a freely independent family.
(2) The family $\{B_i^j : i \in I, j \in J(i)\}$ is a freely independent family.

To prove this, we first see that (1) $\Rightarrow$ (2) even without the generation hypothesis on the B_i^j. For consider a centred alternating product $b_1 b_2 \cdots b_n$ where $b_t \in B_{i_t}^{j_t}$ with $(i_1, j_1) \neq (i_2, j_2) \neq \cdots \neq (i_t, j_t)$. Break up $\{1, 2, \cdots, n\}$ into maximal intervals on which the i_t are the same. The corresponding product of the b's over these intervals are centred by free independence of $\{B_i^j : j \in J(i)\}$. The final product of these products is also centred by free independence of $\{A_i : i \in I\}$.

As for (2) $\Rightarrow$ (1), the key observation is that the span of alternating products of centred elements of the family $\{B_i^j : j \in J(i)\}$ is the set of centred elements of A_i for each $i \in I$. To see this, first A_i is spanned by alternating products of $\{B_i^j : j \in J(i)\}$. The centred elements of A_i are spanned by the centreings of such. So consider $b_1 b_2 \cdots b_k - \phi(b_1 b_2 \cdots b_k)$. Induce on k to show that such an element is in the span of alternating products of centred elements of $\{B_i^j : j \in J(i)\}$. $b_1 b_2 \cdots b_k - \phi(b_1 b_2 \cdots b_k) = (b_1^0 + \phi(b_1))(b_2^0 + \phi(b_2)) \cdots (b_k^0 + \phi(b_k)) - \phi(b_1 b_2 \cdots b_k)$. The term $b_1^0 b_2^0 \cdots b_k^0$ of this is centred and the other terms may be expressed as linear combinations of term looking like $c_1 c_2 \cdots c_l - \phi(c_1 c_2 \cdots c_l)$ for $l < k$.

19.4. **An exercise.** Suppose that (A, ϕ) is a $*$-probability space and B be a unital $*$-subalgebra of A that is free from $\{u, u^*\}$ for a Haar unitary $u \in A$. Then B and uBu^* are freely independent in A.

To see this, first observe that $\phi(ubu^*) = \phi(b)$ using the freeness of $\{u, u^*\}$ from B and the explicit computation done above. Now since u and u^* are themselves centred, an alternating product $b_1 u b_2 u^* b_3 u b_4 u^* \cdots$ of centred elements $b_1, u b_2 u^*, \cdots$ is also an alternating product of centred elements $b_1, u, b_2, u^*, \cdots$, and is therefore centred.

20. Convergence in distribution

20.1. **Central limit theorem.** Let (A, ϕ) be a $*$-probability space and $a_1, a_2, \cdots \in A$ be a sequence of identically distributed self-adjoint random variables that are independent - either freely or classically. Also suppose that all these random variables are centred and let $\sigma^2 = \phi(a_r^2) \geq 0$ be their common variance. A central limit theorem addresses the limit behaviour of $\frac{a_1 + \cdots + a_N}{\sqrt{N}}$ as $N \to \infty$.

20.2. **Convergence in distribution.** The notion of convergence relevant here is that of convergence in distribution. We say that random variables $a_N \in (A_n, \phi_N)$ converge in distribution to $a \in (A, \phi)$ if each moment of a_N converges to the corresponding moment of a, i.e., for all $n \in \mathbb{N}$, $lim_{N \to \infty} \phi_N(a_N^n) = \phi(a^n)$.

20.3. **Relation to weak convergence.** Recall that a sequence of measures μ_N on $\mathbb{R}$ is said to converge weakly to a measure μ if for every bounded continuous function f on $\mathbb{R}$ the integrals against μ_N converge to the integral against μ.

If the a_N has analytic distribution μ_N and a has analytic distribution μ - all of which, by our definitions, are compactly supported measures, then, if the a_N converge in distribution to a, the μ_N converge weakly to μ.

For consider a bounded countinuous function f on $\mathbb{R}$. Approximate it uniformly by polynomials on the compact set $K = supp(\mu)$. Moment convergence now implies integral convergence.

20.4. **Generalisation.** Even if μ is not compactly supported we may get the same conclusion under some weaker hypotheses.

A probability measure μ on $\mathbb{R}$ is said to be **determined by its moments** if it is the only probability measure on $\mathbb{R}$ with its moments. Carleman's theorem says that $\sum_k (m_{2k})^{-\frac{1}{2k}} = \infty$ suffices for a measure to be determined by its moments.

Two important facts are that the Gaussian measure is determined by its moments and that if μ is determined by its moments and if μ_N converges to μ in distribution (with μ_N having moments of all orders) then μ_N converges weakly to μ.

21. General central limit theorem

21.1. **Finite moment computation.** Suppose that (A, ϕ) is a $*$-probability space and $a_1, a_2, \cdots, a_N$ are centred and identically distributed variables in (A, ϕ) that are either free or tensor independent. Consider computation of the moments of $a_1 + a_2 + \cdots + a_N$.

We have $\phi((a_1 + a_2 + \cdots + a_N)^n) = \sum_{i_1, i_2, \cdots, i_n = 1}^N \phi(a_{i_1} a_{i_2} \cdots a_{i_n})$. The terms are indexed by elements of $[N]^n$. Each element $(i_1, \cdots, i_n) \in [N]^n$ determines an equivalence relation on or partition of $[n]$ determined by $p \sim q$ if $i_p = i_q$. The independence condition implies that $\phi(a_{i_1} a_{i_2} \cdots a_{i_N}) = \phi(a_{j_1} a_{j_2} \cdots a_{j_N})$ whenever the partitions of $\{1, 2, \cdots, n\}$ corresponding to the equality relations among the i_t and j_t are equal. Denote the set of partitions of n by $\mathcal{P}(n)$ and for $\pi \in \mathcal{P}(n)$, let κ_π denote the common value of all $\phi(a_{i_1} a_{i_2} \cdots a_{i_n})$ where $(i_1, \cdots, i_n) \in [N]^n$ determines π.

21.2. **Calculation of A_π^N.** Thus $\phi((a_1 + a_2 + \cdots + a_N)^n) = \sum_{\pi \in \mathcal{P}(n)} \kappa_\pi A_\pi^N$, where A_π^N is the number of elements of $[N]^n$ that determine π. This sum depends on N only through the A_π^N dependence on N. Further, $A_\pi^N = N(N-1)(N-2)\cdots(N - |\pi| + 1)$.

21.3. **Restriction on π.** Only certain π contribute to the sum. If π has any class of cardinality 1, the centering and freeness assumptions imply that $\kappa_\pi = 0$, so we may sum only over those partitions of $[n]$ where each class has cardinality at least 2.

21.4. **Final general form.** For such a π, $A_\pi^N / N^{n/2}$ has a limit as $N \to \infty$, which is 0 or 1 according as the number of classes of π is less than or equals $n/2$. In the latter case, π is a **pair partition** of $[n]$, i.e., one in which each class has exactly 2 elements. Denoting the set of pair partitions of $[n]$ by $\mathcal{PP}(n)$, we have seen that

$$lim_{N \to \infty} \phi\left(\left(\frac{a_1 + \cdots + a_N}{\sqrt{N}} \right)^n \right) = \sum_{\pi \in \mathcal{PP}(n)} \kappa_\pi.$$

In particular, the limit vanishes for odd n since $\mathcal{PP}(n) = \emptyset$.

22. Classical central limit theorem

22.1. **Cardinality of $\mathcal{PP}(n)$.** When the variables $a_1, a_2, \cdots$ are tensor independent with common variance σ^2, the factorization rule for ϕ implies that for each $\pi \in \mathcal{PP}(n)$, $\kappa_\pi = \sigma^n$. The cardinality of $\mathcal{PP}(n)$ is easily seen to be $\frac{(2n)!}{2^n n!}$ which is exactly the $2n^{th}$ moment of the standard normal distribution.

22.2. **Classical CLT.** Thus, if $a_1, a_2, \cdots$ are tensor independent and identically distributed with mean 0 and variance σ^2 then, as $N \to \infty$, $\frac{a_1+\cdots+a_N}{\sqrt{N}}$ converges in distribution to x where x is normally distributed with mean 0 and variance σ^2.

23. Free central limit theorem

23.1. **Further restriction on π.** Next suppose that the variables $a_1, a_2, \cdots$ are freely independent with common variance σ^2. Then even among the elements of $\mathcal{PP}(n)$ only some contribute to the sum.

Consider a $\pi \in \mathcal{PP}(n)$ and any $(i_1, \cdots, i_n)$ that determines π. If no $i_k = i_{k+1}$ then by definition of free independence, $\phi(a_1 a_2 \cdots a_n) = 0$. Else

$$\phi(a_1 a_2 \cdots a_n) = \phi(a_1 \cdots a_{k-1} a_{k+2} \cdots a_n)\phi(a_k a_{k+1}).$$

Proceed to see that κ_π is non-zero exactly when some class of π contains consecutive numbers and the restriction of π to the complement of this class also has non-zero κ. In this case $\kappa_\pi = \sigma^n$.

23.2. **The non-crossing condition.** We claim that κ_π is non-zero exactly when π is a non-crossing pairing in the sense that there do not exist $i < j < k < l$ such that $\{i, k\}$ and $\{j, l\}$ are classes of π. For suppose that this property holds. We will show that some pair of neighbours are a class of π. For consider $min\{|i-j| : \{i, j\} \in \pi\}$. If this is at least 2, then some element strictly between i and j must be related so some element outside. Conversely, if $\kappa_\pi \neq 0$, then the necessity of this condition is clear.

We conclude that

$$lim_{N\to\infty}\phi\left(\left(\frac{a_1+\cdots+a_N}{\sqrt{N}}\right)^n\right) = |\mathcal{NCPP}(n)|\sigma^2.$$

23.3. **Cardinality of $\mathcal{NCPP}(n)$.** Next we claim that $|\mathcal{NCPP}(2p)| = C_p$. This is easily verified for small p. For large p note first that any element of $\mathcal{NCPP}(2p)$ pairs an even and an odd number. So $\mathcal{NCPP}(2p)$ has a decomposition into p sets according to what the partner of 1 is - one of $2, 4, \cdots, 2p$. The set where 1 is paired with $2j$ is of cardinality $|\mathcal{NCPP}(2(j-1))|.|\mathcal{NCPP}(2(p-j))|$. This is the Catalan recurrence.

23.4. **Free central limit theorem.** Thus, if $a_1, a_2, \cdots$ are freely independent and identically distributed with mean 0 and variance σ^2 then, as $N \to \infty$, $\frac{a_1+\cdots+a_N}{\sqrt{N}}$ converges in distribution to s where x is semicircularly distributed with mean 0 and variance σ^2.

The Institute of Mathematical Sciences, Chennai
E-mail address: vijay@imsc.res.in