1 Noncommutative probability

In measure theory, the underlying measure space Ω plays an important role, with the measurable sets and functions viewed as being attached to that space. In contrast, in probability, the role of events and their probabilities are more important, and the sample space Ω itself is usually an abstract space, which has less important role. In particular, one cares more about the random variables and σ-algebras than the sample space itself.

These can be encoded by algebraic structures, which leads to an abstract definition of probability space.

Definition 1.1. *A noncommutative probability space is a pair $(\mathcal{A}, \mathbf{E})$, where A is a complex algebra with a unit, and $\mathbf{E} : \mathcal{A} \to \mathbb{C}$ is a linear map such that $\mathbf{E}(1_A) = 1$.*

Example. 1. *For a given probability space $(\Omega, \Sigma, \mathbf{P})$, $(L^\infty(\Omega), \mathbf{E})$ is a noncommutative probabiltiy space,*

2. *$\mathcal{A} = \bigcap_{1 \le p < \infty} L^p(\Omega)$ with $\mathbf{E}$ being the expected value,*

3. *$\mathcal{A} = M_n(\mathbb{C})$ with $\mathbf{E}(T) = \frac{1}{n}\mathrm{tr}(T)$,*

4. *$\mathcal{A} = M_n(\mathbb{C})$, fix a unit vector $v \in \mathbb{C}^n$ with $\|v\|_2 = 1$ and define $\mathbf{E}(T) = v^* T v$. Then $(\mathcal{A}, \mathbf{E})$ is also a noncommutative probability space.*

Definition 1.2. *Let $(\mathcal{A}, \mathbf{E})$ be a noncommutative probability space. We say $a \in \mathcal{A}$ has a distribution μ, where μ is a probability measure on $\mathbb{R}$, if $\mathbf{E}(a^k) = \int_{-\infty}^{\infty} t^k \, \mathrm{d}\mu(t)$ for all $k \in \mathbb{N}$.*

Note that even in the classical case, μ may not be uniquely determined by the moments, so it might actually be better to view a distribution as a sequence of moments instead of a probability measure in this case.

Example. *Consider $(M_n(\mathbb{C}), \tau_n)$, where $\tau_n(T) = \frac{1}{n}\mathrm{tr}(T)$. Let $T \in M_n(\mathbb{C})$ be self-adjoint. Then there exist an orthonormal basis $e_1, \ldots, e_n$ in $\mathbb{C}^n$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ such that $Te_j = \lambda_j e_j$, $j = 1, \ldots, n$. We have*

$$\tau_n(T^k) = \frac{1}{n}\sum_{j=1}^{n}(T^k)_{j,j} = \frac{1}{n}\sum_{j=1}^{n}\lambda_j^k = \int_{-\infty}^{\infty} \lambda^k \, \mathrm{d}\mu(\lambda),$$

where $\mu = \frac{1}{n}\sum_{j=1}^{n}\delta_{\lambda_j}$. That is, T has distribution μ.

We can construct the space of random matrices as follows. Given a probability space $(\Omega, \Sigma, \mathbf{P})$ and $n \in \mathbb{N}$, we define

$$\mathcal{A} = M_n\left(\bigcap_{1 \le p < \infty} L^p(\Omega)\right).$$

For $X = [x_{i,j}] \in A$, $x_{i,j} : \Omega \to \mathbb{C}$ are random variables with $\mathbf{E}|x_{i,j}|^k < \infty$ for all $k \in \mathbb{N}$. Define $\varphi : \mathcal{A} \to \mathbb{C}$ by $\varphi(X) = \mathbf{E}(\tau_n(X)) = \frac{1}{n}\sum_{j=1}^{n}\int_\Omega x_{j,j} \, \mathrm{d}\mathbf{P}$. Then $(\mathcal{A}, \varphi)$ is a noncommutative probability space.

Suppose that $X(\omega) = X(\omega)^*$ for all ω. Then

$$\varphi(X^k) = \mathbf{E}(\tau_n(X(\omega)^k)) = \mathbf{E}\left(\int_{-\infty}^{\infty} \lambda^k \, \mathrm{d}\mu_\omega(\lambda)\right),$$

where μ_ω is the empirical eigenvalue distribution of $X(\omega)$. Then a distribution of X is given by $\mathbf{E}\mu_\omega$.

2 Group algebras

Now we will consider an important example of a noncommutative probability space. Let G be a group and let $\mathbb{C}[G]$ be the group algebra of G. $\mathbb{C}[G]$ is a vector space with G as a basis. For $x \in \mathbb{C}[G]$, write $x = \sum_{g\in G} x_g g$, where $x_g \in \mathbb{C}$ and only finitely many of them are nonzero. The product of x and y, where $x, y \in \mathbb{C}[G]$, is defined to be

$$\begin{aligned} xy &= \left(\sum_{g\in G} x_g g\right)\left(\sum_{h\in G} y_h h\right) \\ &= \sum_{g,h\in G} x_g y_h gh \\ &= \sum_{k\in G}\left(\sum_{g\in G} x_g y_{g^{-1}k}\right) k. \end{aligned}$$

One can also define a norm $\|x\|_1 = \sum_{g\in G} |x_g|$ and consider the completion of $\mathbb{C}[G]$ under this norm:

$$\ell^1(G) = \left\{\sum_{g\in G} x_g g : x_g \in \mathbb{C}, \sum_{g\in G} |x_g| < \infty\right\}.$$

One can show that for all $x, y \in \ell^1(G)$, $\|xy\|_1 \leq \|x\|_1\|y\|_1$. Note that the unit e of G is also a multiplicative unit in $\mathbb{C}[G]$ and $\ell^1(G)$. We define $\tau : \mathbb{C}[G] \to \mathbb{C}$ (or $\tau : \ell^1(G) \to \mathbb{C}$) by

$$\tau\left(\sum_{g\in G} x_g g\right) = x_e.$$

Let's consider a particular case, where G is isomorphic to $\mathbb{Z}$. Write $G = \{g^n : n \in \mathbb{Z}\}$. Then one can associate $\sum_{n\in\mathbb{Z}} \alpha_n g^n \in \ell^1(G)$ with a function (in fact, Fourier series) $f(t) = \sum_{n\in\mathbb{Z}} \alpha_n e^{2\pi int}$ on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, the dual group of $\mathbb{Z}$. Then the expected value $\tau(\sum_{n\in\mathbb{Z}} \alpha_n g^n)$ in this case can be computed by $\int_0^1 f(t)\, \mathrm{d}t$.

2.1 An application: random walk on G

Let G be a group and let $g_1, g_2, \ldots, g_k \in G$. Define a sequence of random variables $X_n : \Omega \to G$ by $X_0 = e$, and

$$\mathbf{P}(X_{n+1} = X_n g_j) = \frac{1}{2k}, \quad \mathbf{P}(X_{n+1} = X_n g_j^{-1}) = \frac{1}{2k}$$

for all $j = 1, 2, \ldots, k$. Of interest are the numbers $p_n = \mathbf{P}(X_n = e)$, the return probabilities.

Define $a \in \mathbb{C}[G]$ by $a = \frac{1}{2k}\sum_{j=1}^{k}(g_j + g_j^{-1})$. Then we have $p_n = \tau(a^n)$ for all $n = 0, 1, 2, \ldots$, because if we write $S = \{g_1, \ldots, g_k, g_1^{-1}, \ldots, g_k^{-1}\}$, then

$$a^n = \left(\frac{1}{2k}\right)^n \sum_{h_j \in S} h_1 h_2 \cdots h_n,$$

and hence

$$\tau(a^n) = \left(\frac{1}{2k}\right)^n \#\{(h_1, \ldots, h_n) : h_j \in S, h_1 h_2 \cdots h_n = e\}.$$

It is easy to see that the right hand side is exactly p_n.

Let's consider a simpler case, G being isomorphic to $\mathbb{Z}$. Let $S = \{g, g^{-1}\}$ with $g \neq e$. As we have seen before, we can associate $a = \frac{1}{2}(g + g^{-1})$ with $f(t) = \frac{1}{2}(e^{2\pi i t} + e^{-2\pi i t}) = \cos(2\pi t)$. Then $\tau(a^n)$ is given by $\int_0^1 \cos^n(2\pi t)\, \mathrm{d}t$, which implies $p_{2n} = \binom{2n}{n}\frac{1}{2^{2n}}$.

3 Independence

One of the most important notion in probability theory is independence. We will also define independence in noncommutative probability.

Definition 3.1. *Let $(\mathcal{A}, \tau)$ be a noncommutative probability space. Let $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{A}$ be two subalgebras such that $e \in \mathcal{A}_1 \cap \mathcal{A}_2$. We say that $\mathcal{A}_1$ and $\mathcal{A}_2$ are classically independent if*

1. *$a_1 a_2 = a_2 a_1$ if $a_1 \in \mathcal{A}_1, a_2 \in \mathcal{A}_2$,*
2. *if $a_1 \in \mathcal{A}_1$ and $a_2 \in \mathcal{A}_2$ with $\tau(a_1) = \tau(a_2) = 0$, then $\tau(a_1 a_2) = 0$.*

Remark. *Note that if $\mathcal{A}_1$ and $\mathcal{A}_2$ are classically independent, then if $a_1 \in \mathcal{A}_1$ and $a_2 \in \mathcal{A}_2$, one has $\tau(a_1 a_2) = \tau(a_1)\tau(a_2)$ (by applying the definition to $a_j - \tau(a_j)e$). So this is indeed the same as the independence we saw in classical probability.*

Let's see an example of classical independence with a nonclassical example. Let $G = G_1 \times G_2$. We can write

$$G = \{g_1 g_2 : g_1 \in G_1, g_2 \in G_2\},$$

with $g_1 g_2 = g_2 g_1$ if $g_1 \in G_1$ and $g_2 \in G_2$, and $g_1 g_2 = e$ if and only if $g_1 = e$ and $g_2 = e$. In this case, $\mathbb{C}[G_1], \mathbb{C}[G_2] \subseteq \mathbb{C}[G]$, and it is easy to see that $\mathbb{C}[G_1]$ and $\mathbb{C}[G_2]$ are classically independent. In this case, if we fix $g_1 \in G_1$ and $g_2 \in G_2$ and consider a random walk with these elements, then $a = \frac{1}{4}(g_1 + g_1^{-1} + g_2 + g_2^{-1}) = \frac{1}{4}(g_1 + g_1^{-1}) + \frac{1}{4}(g_2 + g_2^{-1})$ is a sum of independent random variables.

In classical independence, some commutativity is required, which might not be very useful in the setting of noncommutative probability. Let's again consider the group algebras as a motivating example. Above we saw that the Cartesian product of groups give some commutativity between the two groups, and hence classical independence between the two group algebras. However, instead of Cartesian product, one can also consider the free product of two groups, which can be seen as the most "general" group that is generated by the elements of the two groups.

More precisely, for groups G_1 and G_2, a word in G_1 and G_2 is a product of the form $s_1 s_2 \cdots s_n$, where each s_j is either an element of G_1 of an element of G_2. Such a word can be reduced by removing the identity elements or reducing a pair of the form $g_1 h_1$ by its product in G_1 or a pair of the form $g_2 h_2$ by its product in G_2. Then every reduced word is an alternating product of elements of G_1 and elements of G_2. The free product $G_1 * G_2$ is the group with elements being the reduced words in G_1 and G_2, under the operation of concatenation of words followed by reduction. In this case, $\mathbb{C}[G_1]$ and $\mathbb{C}[G_2]$ are again subalgebras of $\mathbb{C}[G_1 * G_2]$. This motivates the notion of free independence:

Definition 3.2. *Let $(\mathcal{A}, \tau)$ be a noncommutative probability space. Let $\mathcal{A}_1, \mathcal{A}_2 \subseteq \mathcal{A}$ be two subalgebras such that $e \in \mathcal{A}_1 \cap \mathcal{A}_2$. We say that $\mathcal{A}_1$ and $\mathcal{A}_2$ are freely independent if for any $a_1, a_2, \ldots, a_n$, where $a_i \in \mathcal{A}_{j_i}$, $j_i \neq j_{i+1}$ for $i = 1, 2, \ldots, n-1$ with $\tau(a_i) = 0$ for all i, then $\tau(a_1 a_2 \cdots a_n) = 0$.*

Note that under this definition, if $G = G_1 * G_2$, then $\mathbb{C}[G_1]$ and $\mathbb{C}[G_2]$ are freely independent.

Proof. Write $a_i = \sum_{g \in G_{j_i} \setminus \{e\}} \alpha_g^{(i)} g$. Then

$$a_1 a_2 \cdots a_n = \sum_{g_1, \ldots, g_n} \alpha_{g_1}^{(1)} \alpha_{g_2}^{(2)} \cdots \alpha_{g_n}^{(n)} g_1 g_2 \cdots g_n,$$

where the sum is over $g_i \in G_{j_i} \setminus \{e\}$, $i = 1, \ldots, n$. The term $g_1 g_2 \cdots g_n$ cannot cancel, because each consecutive elements belong to different algebras. Therefore, $\tau(a_1 a_2 \cdots a_n) = 0$. □

3.1 Comparing classical independence and free independence

For $x \in \mathcal{A}$, write $\mathcal{A}_x = \{p(x) : p \text{ is a polynomial}\}$, the algebra generated by x.

First consider the case $\mathcal{A}_x$ and $\mathcal{A}_y$ are classically independent. In this case, the moment of $x + y$ can be computed by

$$\begin{aligned}\tau((x+y)^n) &= \sum_{j=0}^{n} \binom{n}{j} \tau(x^j y^{n-j}) \\ &= \sum_{j=0}^{n} \binom{n}{j} \tau(x^j) \tau(y^{n-j}) \\ &= \tau(x^n) + \tau(y^n) + Q_n(\tau(x), \tau(x^2), \ldots, \tau(x^{n-1}), \tau(y), \tau(y^2), \ldots, \tau(y^{n-1}))\end{aligned}$$

for some polynomial Q_n. For $\alpha = (\alpha_1, \ldots, \alpha_n), \beta = (\beta_1, \ldots, \beta_n) \in \mathbb{C}^n$, define

$$\alpha *_n \beta = (\alpha_1 + \beta_1, \alpha_2 + \beta_2 + Q_2(\alpha_1, \beta_1), \ldots, \alpha_n + \beta_n + Q_n(\alpha_1, \ldots, \alpha_{n-1}, \beta_1, \ldots, \beta_{n-1})).$$

Then $*_n : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}^n$ is commutative, associative, and has a unit (namely the zero vector). Moreover, by solving equations, one can see that $\alpha *_n \beta = 0$ is possible, and so inverse exists. As an algebraic group, there are not too many possibilities, and it turns out that $(\mathbb{C}^n, *_n) \cong (\mathbb{C}^n, +)$. Therefore, there exists a sequence of numbers $\{c_n(x)\}_{n=1}^{\infty}$, called the cumulants, such that

$$(\tau(x), \tau(x^2), \dots, \tau(x^n)) \mapsto (c_1(x), c_2(x), \dots c_n(x))$$

is an isomorphism from $\mathbb{C}^n$ to $\mathbb{C}^n$. In fact, we have

$$\tau(x^n) = \sum_{\pi \text{ partition of } \{1,\dots,n\}} \prod_{B \text{ block of } \pi} c_{|B|}(x).$$

For example, when $n = 4$, there are four ways to partition $\{1, 2, 3, 4\}$ into a 3-element set and a 1-element set, three ways to partition into two 2-element sets, etc., we have

$$\tau(x^4) = c_4(x) + 4c_3(x)c_1(x) + 3c_2(x)^2 + 6c_2(x)c_1(x)^2 + c_1(x)^4.$$

Now, consider the case that $\mathcal{A}_x$ and $\mathcal{A}_y$ are freely independent. Then

$$\tau((x + y)^n) = \sum \tau(a_1 a_2 \cdots a_n), \quad a_1, \dots, a_n \in \{x, y\}.$$

So we would like to know what $\tau(x^{i_1} y^{j_1} x^{i_2} y^{j_2} \cdots x^{i_k} y^{i_k})$ is. Using

$$\tau((x^{i_1} - \tau(x^{i_1})e) \cdots (y^{j_k} - \tau(y^{j_k})e)) = 0$$

and induction, one can show that

$$\tau((x + y)^n) = \tau(x^n) + \tau(y^n) + Q'_n(\tau(x), \tau(x^2), \dots, \tau(x^{n-1}), \tau(y), \tau(y^2), \dots, \tau(y^{n-1}))$$

for some polynomial Q'_n (which is different from Q_n unless $n = 1, 2$). Similarly, one can define

$$\alpha \boxplus_n \beta = (\alpha_1 + \beta_1, \alpha_2 + \beta_2 + Q'_2(\alpha_1, \beta_1), \dots, \alpha_n + \beta_n + Q'_n(\alpha_1, \dots, \alpha_{n-1}, \beta_1, \dots, \beta_{n-1})).$$

This $\boxplus_n$ is commutative, associative (which is not obvious), with unit being the zero vector, and has inverses. Therefore, there is an isomorphism $(\mathbb{C}^n, \boxplus_n) \to (\mathbb{C}^n, +)$, and we write

$$(\tau(x), \tau(x^2), \dots, \tau(x^n)) \mapsto (k_1(x), k_2(x), \dots, k_n(x)).$$

These k_j's are called the free cumulants of x. It is known that

$$\tau(x^n) = \sum_{\pi \text{ non-crossing partition of } \{1,\dots,n\}} \prod_{B \text{ block of } \pi} k_{|B|}(x).$$

For example, when $n = 4$, there are four ways to partition $\{1, 2, 3, 4\}$ into a 3-element set and a 1-element set, three ways to partition into two 2-element sets (but only two of them are non-crossing), etc., we have

$$\tau(x^4) = k_4(x) + 4k_3(x)k_1(x) + 2k_2(x)^2 + 6k_2(x)k_1(x)^2 + k_1(x)^4.$$

Note that the formula the almost the same as that in the classical case except the $k_2(x)^2$ term.

4 Free central limit theorem

Theorem 4.1. *Let $(\mathcal{A}, \tau)$ be a noncommutative probability space, and let $\{x_n\}_{n=1}^{\infty}$ be freely independent identically distributed (f.i.d.) random variables with $\tau(x_n) = 0$, $\tau(x_n^2) = 1$. Define $y_n = (x_1 + \cdots + x_n)/\sqrt{n}$. Then $\tau(y_n^k) \to C_{k/2}$ as $n \to \infty$, where $C_{k/2}$ is the $(k/2)$-th Catalan number (with $C_{k/2} = 0$ if k is odd).*

As we have seen, the moments of the semicircle law are exactly the Catalan numbers. Therefore, it means that sum of f.i.d. random variables converges (after suitable rescaling) in distribution to the semicircle law.

In the classical case, the central limit theorem can be proved by the following fact: if $X, Y \sim N(0,1)$ are independent, then $(X+Y)/\sqrt{2}$ has the same distribution as X. Why would this help? Because in this case, we have $c_n((X+Y)/\sqrt{2}) = c_n(X)$, and so by independence $c_n(X/\sqrt{2}) + c_n(Y/\sqrt{2}) = c_n(X)$, which implies $2(1/\sqrt{2})^n c_n(X) = c_n(X)$. This forces $c_n(X) = 0$ if $n \neq 2$, and it is easy to see that $c_2(X) = \mathrm{Var}(X) = 1$. Now, if $X_1, \dots, X_n$ are i.i.d. with $\mathbf{E}X = 0$ and $\mathrm{Var}(X) = 1$ (and under some suitable moment conditions), we see that $c_m((X_1 + \cdots + X_n)/\sqrt{n}) = n/(\sqrt{n})^m c_m(X_1)$, which is 0 when $m = 1$, 1 when $m = 2$, and goes to 0 as $n \to \infty$ when $m \geq 3$, which are the cumulants of a standard Gaussian, and this proves the central limit theorem.

Therefore, to prove the free central limit theorem, it suffices to construct f.i.d. variables X, Y such that $(X+Y)/\sqrt{2}$ has the same distribution as X with $\tau(X) = 0$, $\tau(X^2) = 1$. The same calculation will apply with cumulants c replaced by the free cumulants k.

4.1 Fock space

Let V be a finite dimensional Euclidean space. We define the Fock space of V by

$$T(V) = \bigoplus_{n=0}^{\infty} V^{\otimes n},$$

where $V^{\otimes 0} = \mathbb{R}\Omega$, and $\Omega \in V$ is a fixed unit vector called the vacuum vector. Let $\mathcal{A} = L(T(V))$, the space of linear operators on $T(V)$, and we define $\tau : \mathcal{A} \to \mathbb{C}$ by $\tau(X) = \langle X\Omega, \Omega \rangle$.

For a unit vector $v \in V$, we define the creation operator $C_v : T(V) \to T(V)$ by $C_v(w) = v \otimes w$. For example, say $w = 4\Omega + 3w_1 + 5w_1 \otimes w_2 + 3w_2 \otimes w_3$, $C_v(w) = 4v + 3v \otimes w_1 + 5v \otimes w_1 \otimes w_2 + 3v \otimes w_2 \otimes w_3$. We also define the destruction operator D_v by $D_v(\Omega) = 0$, and $D_v(w_1 \otimes w_2 \otimes \cdots \otimes w_n) = \langle w_1, v \rangle w_2 \otimes \cdots \otimes w_n$. Define $X_v = C_v + D_v$. Note that

$$\tau(X_v) = \tau(C_v) + \tau(D_v) = \langle C_v\Omega, \Omega \rangle + \langle D_v\Omega, \Omega \rangle = \langle v, \Omega \rangle + 0.$$

Since v is a tensor of order 1 and Ω is a tensor of order 0, $\langle v, \Omega \rangle = 0$, and thus $\tau(X_v) = 0$. Also,

$$\begin{aligned} \tau(X_v^2) &= \langle C_v^2\Omega, \Omega \rangle + \langle C_v D_v \Omega, \Omega \rangle + \langle D_v C_v \Omega, \Omega \rangle + \langle D_v^2 \Omega, \Omega \rangle \\ &= \langle D_v C_v \Omega, \Omega \rangle = \langle \Omega, \Omega \rangle = 1. \end{aligned}$$

For general n, we have

$$\tau(X_v^n) = \sum_{a_j \in \{C_v, D_v\}} \tau(a_1 a_2 \cdots a_n).$$

When n is odd, it is not difficult to see that the summand is always 0. When n is even, the calculation for $n = 2$ suggests that $\tau(a_1 a_2 \cdots a_n)$ is not 0 only when the number of C_v's is the same as the number of D_v's, and there are always more creations than destruction when one reads the string of right to left. If one recalls the definition of Dyck word or Dyck path, then we have $\tau(X_v^n) = C_k$, where $n = 2k$. Therefore, the distribution of X_v is exactly the semicircle law.

Proposition 4.2. *If $v_1, v_2 \in V$ and $\langle v_1, v_2 \rangle = 0$, then $\mathcal{A}_1$ is freely independent of $\mathcal{A}_2$, where $\mathcal{A}_j$ is the algebra generated by C_{v_j} and D_{v_j}, $j = 1, 2$.*

Proof. Note that since $v_1 \perp v_2$, we have $D_{v_1} C_{v_2} = 0$. Also, note that $D_{v_1} C_{v_1}$ is the identity, and so A_1 consists of linear combinations of elements of the form $C_{v_1}^{\alpha} D_{v_1}^{\beta}$, $\alpha, \beta \in \{0, 1, 2 \ldots\}$. Moreover, $\tau(C_{v_1}^{\alpha} D_{v_1}^{\beta}) = \langle C_{v_1}^{\alpha} D_{v_1}^{\beta} \Omega, \Omega \rangle$, and this is nonzero only when $\alpha = \beta = 0$, and in this case $\langle C_{v_1}^{0} D_{v_1}^{0} \Omega, \Omega \rangle = 1$.

To prove free independence, consider a product $Y_1, Y_2, \ldots, Y_n$, $Y_j \in A_{i_j}$, $i_j \in \{1, 2\}$ and $i_j \neq i_{j+1}$ for $j = 1, 2, \ldots, n-1$ and $\tau(Y_j) = 0$. We need to show $\tau(Y_1 Y_2 \cdots Y_n) = 0$. By linearity, we may (and do) restrict to the case in which Y_{i_j} has only one term (that is, $Y_{i_j} = C_{v_{i_j}}^{\alpha_j} D_{v_{i_j}}^{\beta_j}$ with $\alpha_j + \beta_j > 0$). Note that

$$Y_1 Y_2 \cdots Y_n = C_{v_{i_1}}^{\alpha_1} D_{v_{i_1}}^{\beta_1} C_{v_{i_2}}^{\alpha_2} D_{v_{i_2}}^{\beta_2} \cdots C_{v_{i_n}}^{\alpha_n} D_{v_{i_n}}^{\beta_n}.$$

Recall that when $v_1 \perp v_2$, $D_{v_1} C_{v_2} = 0$. Therefore, for the above product being nonzero, it has to be of the form

$$C_{v_{i_1}}^{\alpha_1} \cdots C_{v_{i_j}}^{\alpha_j} D_{v_{i_{j+1}}}^{\beta_{j+1}} \cdots D_{v_{i_n}}^{\beta_n}$$

for some $j \in \{1, \ldots, n\}$. However, in this case we still have $\tau(Y_1 Y_2 \cdots Y_n) = 0$. This proves the proposition. □

The free central limit theorem will be proved by establishing the following proposition.

Proposition 4.3. *If $v_1, v_2 \in V$ and $\langle v_1, v_2 \rangle = 0$, then $(X_{v_1} + X_{v_2})/\sqrt{2}$ and X_{v_1} has the same distribution.*

Proof. Write $v = (v_1 + v_2)/\sqrt{2}$. Since $v_1 \perp v_2$, v is also a unit vector. Moreover,

$$\left(\frac{C_{v_1} + C_{v_2}}{\sqrt{2}} \right)(w) = \frac{v_1 \otimes w + v_2 \otimes w}{\sqrt{2}} = v \otimes w = C_v(w).$$

Similarly, $(D_{v_1} + D_{v_2})/\sqrt{2} = D_v$. Therefore, $(X_{v_1} + X_{v_2})/\sqrt{2} = X_v$. Since we saw that the law of X_v is semicircular, we are done. □

5 More connection to random matrices

From what we did in this semester and what we saw above, we see that Wigner random matrices converges in distribution (in the sense of noncommutative probability) to the semicircle law. It has nothing to do with the free central limit theorem, though.

Yet, random matrices are deeply related to free independence.

Definition 5.1. *A sequence of random variables $X_{n,1}, \ldots, X_{n,k}$ in a noncommutative probability space $(\mathcal{A}_n, \tau_n)$ is asymptotically free independent if*

$$\tau_n((P_1(X_{n,i_1}) - \tau_n((P_1(X_{n,i_1}))e) \cdots (P_m(X_{n,i_m}) - \tau_n((P_m(X_{n,i_m}))e)) \to 0$$

as $n \to \infty$, for any polynomials $P_1, \ldots, P_m$, and $i_1, \ldots, i_m \in \{1, \ldots, k\}$, $i_j \neq i_{j+1}$ for all $j = 1, \ldots, m-1$.

Independent families of random matrices are asymptotically free. A large random matrix will correlate with itself (for instance, $\mathrm{tr}(M^*M)$ is large); however, if we insert an independent random matrix N with zero trace, then the correlation will be largely destroyed (for instance, $\mathrm{tr}(M^*NM)$ is usually small). In fact, we have the following.

Theorem 5.2. *Let $M_{n,1}, \ldots, M_{n,k}$ be independent $n \times n$ Wigner matrices, where the coefficients all have uniformly bounded m-th moments for all m. Then the random matrices $\frac{1}{\sqrt{n}}M_{n,1}, \ldots, \frac{1}{\sqrt{n}}M_{n,k}$ are asymptotically freely independent.*

This can be proved, again, by analyzing the moments carefully.

References

[1] Bercovici, Hari. M564 lectures.

[2] Tao, Terence. Topics in random matrix theory. Graduate Studies in Mathematics, 132. American Mathematical Society, Providence, RI, 2012. x+282 pp. ISBN: 978-0-8218-7430-1