LECTURE 3: FREE CENTRAL LIMIT THEOREM AND FREE CUMULANTS

Recall from Lecture 2 that if (A, ϕ) is a non-commutative probability space and $A_1, \ldots, A_n$ are subalgebras of A which are free with respect to ϕ, then freeness gives us in principle a rule by which we can evaluate $\phi(a_1 a_2 \ldots a_k)$ for any alternating word in random variables $a_1, a_2, \ldots, a_k$. Thus we can in principle calculate all mixed moments for a system of free random variables. However, we do not yet have any concrete idea of the structure of this factorization rule. This situation will be greatly clarified by the introduction of "free cumulants." Classical cumulants appeared in Lecture 1, where we saw that they are intimately connected with the combinatorial notion of set partitions. Our free cumulants will be linked in a similar way to the lattice of non-crossing set partitions. We will motivate the appearance of free cumulants and non-crossing partition lattices in free probability theory by examining in detail a proof of the central limit theorem by the method of moments.

1. The Classical and Free Central Limit Theorems

Our setting is that of a non-commutative probability space (A, ϕ) and a sequence $(a_i)_{i\in\mathbb{N}} \subset A$ of centred and identically distributed random variables. This means that $\phi(a_i) = 0$ for all $i \geq 1$, and that and $\phi(a_i^n) = \phi(a_j^n)$ for any $i, j, n \geq 1$. We assume that our random variables $a_i,\ i \geq 1$ are either classically independent, or freely independent as defined in Lecture 2. Either form of independence gives us a "factorization rule" for calculating mixed moments in the random variables.

For $k \geq 1$, set

$$S_k := \frac{1}{\sqrt{k}}(a_1 + \cdots + a_k). \tag{1}$$

The Central Limit Theorem is a statement about the limit distribution of the random variable S_k in the large k limit.

Definition 1. *Let $(A_k, \phi_k),\ k \in \mathbb{N}$ and (A, ϕ) be noncommutative probability spaces. Let $(b_k)_{k\in\mathbb{N}}$ be a sequence of random variables with*

Date: Lecture given on October 4, 2007.

$b_k \in A_k$, and let $b \in A$. We say that the sequence $(b_k)_{k\in\mathbb{N}}$ converges in distribution to b, notated by $b_k \xrightarrow{distr} b$, if

$$\lim_{k\to\infty} \phi_k(b_k^n) = \phi(b^n) \tag{2}$$

for any fixed $n \in \mathbb{N}$.

We want to make a statement about convergence in distribution of the random variables $(S_k)_{k\in\mathbb{N}}$ (which all come from the same underlying n.c. probability space). Thus we need to do a moment calculation. Let $[k] = \{1, \dots, k\}$ and $[n] = \{1, \dots, n\}$ and denote by $[k]^{[n]}$ the set of all functions $r : [n] \to [k]$.

$$\phi(S_k^n) = \frac{1}{k^{n/2}} \sum_{r\in[k]^{[n]}} \phi(a_{r(1)} \dots a_{r(n)}).$$

It will be convenient to think of our state ϕ and the first k random variables $a_1, \dots, a_k$ as defining a functional $\hat{\phi} : [k]^{[n]} \to \mathbb{C}$ by

$$\hat{\phi}(r) := \phi(a_{r(1)} \dots a_{r(n)}).$$

Thus what our moment calculation amounts to is summing the possible values of this functional.

$$\phi(S_k^n) = \frac{1}{k^{n/2}} \sum_{r\in[k]^{[n]}} \hat{\phi}(r).$$

It turns out that the fact that the random variables $a_1, \dots, a_k$ are independent and identically distributed makes this task less complex than it initially appears.

The idea of encoding a function $r \in [k]^{[n]}$ by its ordered sequence of fibres was popularized by Gian-Carlo Rota in a series of lectures on the so-called *twelvefold way* in enumerative combinatorics. The simple idea is that a function $r \in [k]^{[n]}$ is equivalent to the data of the ordered k-tuple

$$(F_1, \dots, F_k),$$

where $F_i = r^{-1}(\{i\})$ is the fibre of r over $i \in [k]$. In general, an ordered k-tuple of disjoint subsets of $[n]$ whose union is $[n]$ and some of which might be empty is called a "k-part set composition" of $[n]$. The collection of k-part set compositions of n is denoted $\mathcal{C}^k(n)$.

Example 1.1. *Consider the function $r \in [4]^{[6]}$ defined by*

$$r(1) = 1, r(2) = 2, r(3) = 1, r(4) = 1, r(5) = 2, r(6) = 3. \tag{3}$$

Then r is encoded by the set composition

$$(\{1, 3, 4\}, \{2, 5\}, \{6\}, \emptyset). \tag{4}$$

So, if we regard $\hat{\phi}$ as a functional on $\mathcal{C}^k(n)$, then what we want to do is compute

$$\phi(S_k^n) = \frac{1}{k^{n/2}} \sum_{(F_1,\ldots,F_k)\in\mathcal{C}^k(n)} \hat{\phi}(F_1,\ldots,F_k).$$

From the point of view of set compositions, there is a very natural action of the symmetric group $S(k)$ on $\mathcal{C}^k(n)$; permutations should act simply by permuting parts. That is for $\sigma \in S(k)$ and $(F_1,\ldots,F_k) \in \mathcal{C}^k(n)$ we define

$$\sigma \cdot (F_1,\ldots,F_k) := (F_{\sigma(1)},\ldots,F_{\sigma(k)}).$$

For any $(F_1,\ldots,F_k)$ its **orbit** and **stabilizer** under the action of $S(k)$ are defined to be

$$\begin{aligned}
\mathcal{O}(F_1,\ldots,F_k) &:= \{(F_{\sigma(1)},\ldots,F_{\sigma(k)}) : \sigma \in S(k)\} \\
\mathcal{S}(F_1,\ldots,F_k) &:= \{\sigma \in S(k) : (F_{\sigma(1)},\ldots,F_{\sigma(k)}) = (F_1,\ldots,F_k)\}.
\end{aligned}$$

Orbits are either disjoint or identical. Let $\mathcal{C}^k(n)/S(k)$ denote the set of distinct orbits arising from the group action. The fact that $\hat{\phi}$ is defined in terms of independent and identically distributed random variables $a_1,\ldots,a_k$ allows us to "lift" $\hat{\phi}$ to a function on $\mathcal{C}^k(n)/S(k)$.

Lemma 1. *For any composition $(F_1,F_2,\ldots,F_k) \in \mathcal{C}^{(k)}(n)$, $\hat{\phi}$ is constant on $\mathcal{O}(F_1,F_2,\ldots,F_k)$.*

Proof. The statement is that for any $(F_1,F_2,\ldots,F_k) \in \mathcal{C}^{(k)}(n)$ and any permutation $\sigma \in S(k)$ we have

$$\hat{\phi}(F_1,F_2,\ldots,F_k) = \hat{\phi}(F_{\sigma(1)},F_{\sigma(2)},\ldots,F_{\sigma(k)}).$$

We know that the random variables $a_1,\ldots,a_k$ are independent. This means that we have a factorization rule for calculating mixed moments in $a_1,\ldots,a_k$ in terms of the moments of individual a_i's. In particular this means that $\hat{\phi}(F_1,F_2,\ldots,F_k)$ can be written as some expression in moments $\phi(a_i^j)$, while $\hat{\phi}(F_{\sigma(1)},F_{\sigma(2)},\ldots,F_{\sigma(k)})$ can be written as that same expression except with $\phi(a_i^j)$ replaced by $\phi(a_{\sigma(i)}^j)$. However, since our random variables all have the same distribution, then $\phi(a_i^j) = \phi(a_{\sigma(i)}^j)$ for any i,j. Thus the lemma is proved. $\square$

Consequently, we have

$$\phi(S_k^n) = \frac{1}{k^{n/2}} \sum_{\mathcal{O}\in\mathcal{C}^k(n)/S(k)} \hat{\phi}(\mathcal{O}) \cdot |\mathcal{O}|.$$

It is not difficult to find $|\mathcal{O}|$. For any $(F_1,\dots,F_k) \in \mathcal{C}^k(n)$, we know from the orbit-stabilizer theorem that

$$|\mathcal{O}(F_1,\dots,F_k)| \cdot |\mathcal{S}(F_1,\dots,F_k)| = |S(k)|.$$

Note that the only way that a permutation can stabilize a composition is by permuting its parts which are equal to the emptyset $\emptyset$. Thus, if $p(F_1,\dots,F_k)$ is the number of parts of $(F_1,\dots,F_k)$ which are not equal to the emptyset, we have from the orbit-stabilizer theorem that

$$|\mathcal{O}(F_1,\dots,F_k)| = \frac{k!}{(k-p(F_1,\dots,F_k))!}.$$

Hence

$$\phi(S_k^n) = \frac{1}{k^{n/2}} \sum_{\mathcal{O} \in \mathcal{C}^k(n)/S(k)} \hat{\phi}(\mathcal{O}) \cdot k(k-1)\dots(k-p(\mathcal{O})+1).$$

Now, an orbit $\mathcal{O} \in \mathcal{C}^k(n)/S(k)$ is equivalent to the data of an unordered collection of disjoint non-empty subsets of $[n]$ whose union is $[n]$, in other words a *partition* of $[n]$. This is because if $\{V_1,\dots,V_s\}$ is a partition of $[n]$ with $s \leq k$, we can construct a corresponding orbit by forming the composition

$$(V_1,\dots,V_s,V_{s+1},\dots,V_k),$$

where $V_{s+1},\dots,V_s$ are instances of $\emptyset$, and permuting its parts in all possible ways. Thus there is a one-to-one correspondence between $\mathcal{C}^k(n)/S(k)$ and the set of partitions of $[n]$ which have at most k blocks (the sets that make up a partition $\pi = \{V_1,\dots,V_s\}$ are called blocks). Let $\mathcal{P}(n)$ denote the collection of all partitions π of $[n]$, and define

$$\kappa(\pi) = \kappa(\{V_1,\dots,V_{|\pi|}\}) = \begin{cases} \hat{\phi}(\mathcal{O}(V_1,\dots,V_{|\pi|},\emptyset,\cdots\emptyset)), \text{ if } |\pi| \leq k \\ 0, \text{ if } |\pi| > k \end{cases}$$

where $|\pi|$ denotes the number of blocks of π. Then what we have proved is that

$$\phi(S_k^n) = \frac{1}{k^{n/2}} \sum_{\pi \in \mathcal{P}(n)} \kappa(\pi) \cdot k(k-1)\dots(k-|\pi|+1).$$

The great advantage of this expression over what we started with is that the number of terms does not depend on k. Thus we are in a position to take the $k \to \infty$ limit, provided we can effectively estimate each term of the sum.

Our first observation is the most obvious one, namely we have

$$k(k-1)\dots(k-|\pi|+1) \sim k^{|\pi|} \tag{5}$$

for $k \to \infty$.

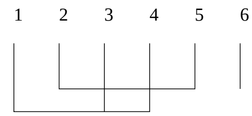

FIGURE 1. *Visualizing the set partition* $\{1,3,4\} \cup \{2,5\} \cup \{6\}$.

Next observe that if π has a block of size 1, then we will have $\kappa(\pi) = 0$. Indeed suppose that $\pi = \{V_1, \ldots, V_m, \ldots, V_s\} \in \mathcal{P}([n])$ with $V_m = \{j\}$ for some $j \in [n]$. Then we will have

$$\kappa(\pi) = \phi(a_1 \ldots a_{j-1} a_j a_{j+1} \ldots a_n) \tag{6}$$

where $a_j \notin \{a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n\}$. Hence we can write

$$\kappa(\pi) = \phi(b a_j c), \tag{7}$$

where $b = a_1 \ldots a_{j-1}$ and $c = a_{j+1} \ldots a_n$ and thus

$$\phi(b a_j c) = \phi(a_j)\phi(bc) = 0, \tag{8}$$

since a_j is (classically or freely) independent of $\{b, c\}$.

Thus the only partitions which contribute to the sum are those with blocks of size at least 2. Note that such a partition can have at most $n/2$ blocks. Now,

$$\lim_{k\to\infty} \frac{k^{|\pi|}}{k^{n/2}} = \begin{cases} 1, & \text{if } |\pi| = n/2 \\ 0, & \text{if } |\pi| < n/2. \end{cases} \tag{9}$$

Hence the only partitions which contribute to the sum in the $k \to \infty$ limit are those with *exactly* $n/2$ blocks, i.e. partitions each of whose blocks has size 2. Such partitions are called "pairings," and the set of pairings is denoted $\mathcal{P}_2(n)$.

Thus we have shown that

$$\lim_{k\to\infty} \phi(S_k^n) = \sum_{\pi \in \mathcal{P}_2(n)} \kappa(\pi). \tag{10}$$

Note that in particular if n is odd then $\mathcal{P}_2(n) = \emptyset$, so that the odd limiting moments vanish. In order to determine the even limiting moments, we must distinguish between the setting of classical independence and free independence.

In the case of classical independence, our random variables commute and factorize completely with respect to ϕ. Thus if we denote by $\phi(a_i^2) = \sigma^2$ the common variance of our random variables, then for any pairing $\pi \in \mathcal{P}_2(n)$ we have $\kappa(\pi) = \sigma^n$. Thus we have

$$\begin{aligned}
\lim_{k\to\infty} \phi(S_k^n) &= \sum_{\pi\in\mathcal{P}_2(n)} \kappa(\pi) \\
&= \sigma^n |\mathcal{P}_2(n)| \\
&= \begin{cases} \sigma^n (n-1)(n-3)\dots 5\cdot 3\cdot 1, & \text{if } n \text{ even} \\ 0, & \text{if } n \text{ odd} \end{cases} .
\end{aligned}$$

From Lecture 1, we recognize these as exactly the moments of a Gaussian random variable of mean 0 and variance σ^2. Hence we get the classical central limit theorem: If $(a_i)_{i\in\mathbb{N}}$ are classically independent random variables which are identically distributed with $\phi(a_i) = 0$ and $\phi(a_i^2) = \sigma^2$, then S_k converges in distribution to a Gaussian random variable with mean 0 and variance σ^2.

We should note that the notion of "convergence in distribution" is *a priori* weaker than the usual notion of "weak convergence" or "convergence in law" considered in probability theory. However, these notions coincide in the case that the distribution of the limit random variable is determined by its moments (i.e. it is the unique distribution with those moments). This is well-known to be the case for the Gaussian distribution.

Now we want to deal with the case where the random variables are freely independent. In this case, $\kappa(\pi)$ will not be the same for all pair partitions $\pi \in \mathcal{P}_2(2n)$ (we focus on the even moments now because we already know that the odd ones are zero). Let's take a look at some examples:

$$\begin{aligned}
\kappa(12|34) &= \phi(a_1a_1a_2a_2) = \phi(a_1)^2\phi(a_2)^2 = \sigma^4 \\
\kappa(14|23) &= \phi(a_1a_2a_2a_1) = \phi(a_1)^2\phi(a_2)^2 = \sigma^4 \\
\kappa(13|24) &= \phi(a_1a_2a_1a_2) = 0.
\end{aligned}$$

(Here we are using a shorthand notation for partitions, e.g. $12|34 = \{\{1,2\},\{3,4\}\}$). In general, we will get $\kappa(\pi) = \sigma^{2n}$ if we can find successively neighbouring pairs of identical random variables in any word in the random variables corresponding to π; if we cannot we will have $\kappa(\pi) = 0$. Geometrically, one sees that the type of pair partitions that give a non-zero contribution are the ones that have the geometric property that they are **non-crossing** (see the next section). Let $NC_2(2n)$ denote the set of non-crossing pair partitions. Then we have as our central limit theorem that

$$\lim_{k\to\infty} \phi(S_k^{2n}) = \sigma^{2n}|NC_2(2n)|.$$

There are many ways to determine the cardinality $C_n := |NC_2(2n)|$; to give it away, the answer turns out to be the ubiquitous **Catalan numbers** which appear in many seemingly unrelated contexts throughout mathematics.

Our first method for counting non-crossing pairings is to find a simple recurrence which they satisfy. The idea is to look at the block of a pairing which contains the number 1. In order for the pairing to be non-crossing, i must be paired with some even number in the set $[2n]$, else we would necessarily have a crossing. Thus 1 must be paired with $2i$ for some $i \in [n]$. Now let i run through all possible values in $[n]$, and count for each the number of non-crossing pairings that contain this block, as in the diagram below.

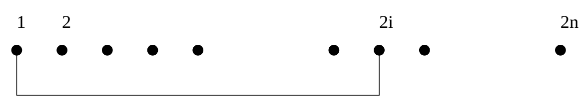

FIGURE 2. *We have C_{i-1} possibilities for inside the block, and C_{n-i} possibilities for outside the block.*

In this way we see that the cardinality C_n of $NC_2(2n)$ must satisfy the recurrence relation

$$C_n = \sum_{i=1}^{n} C_{i-1}C_{n-i}, \tag{11}$$

with initial condition $C_0 = 1$. One can then check directly that the Catalan numbers satisfy this recurrence, hence $C_n = \frac{1}{n+1}\binom{2n}{n}$.

We can also prove directly that $C_n = \frac{1}{n+1}\binom{2n}{n}$ by finding a bijection between $NC_2(2n)$ and some standard set of objects which we can see directly is enumerated by the Catalan numbers. A reasonable choice for this "canonical" set is the collection of $2 \times n$ Standard Young Tableaux. Suppose that we have a $2 \times n$ grid of squares. A Standard Young Tableaux of shape $2 \times n$ is just a filling of the squares of the grid with the numbers $1, \dots, 2n$ which is strictly increasing in each of the 2 rows and each of the n columns. The number of these Standard Young Tableaux is very easy to calculate, using a famous and fundamental result known as the "Hook-Length Formula." The Hook-Length formula tells us that the number of Standard Young Tableaux on the $2 \times n$ rectangle is

$$\frac{(2n)!}{(n+1)!n!} = \frac{1}{n+1}\binom{2n}{n}. \tag{12}$$

Thus we will have proved that $|NC_2(2n)| = \frac{1}{n+1}\binom{2n}{n}$ if we can bijectively associate to each pair partition $\pi \in NC_2(2n)$ a Standard Young

Tableaux on the $2 \times n$ rectangular grid. This is very easy to do. Simply take the "left-halves" of each pair in π and write them in increasing order in the cells of the first row. Then take the "right-halves" of each pair of π and write them in increasing order in the cells of the second row. Figure 2 shows the bijection between $NC_2(6)$ and Standard Young Tableaux on the 2×3 rectangle.

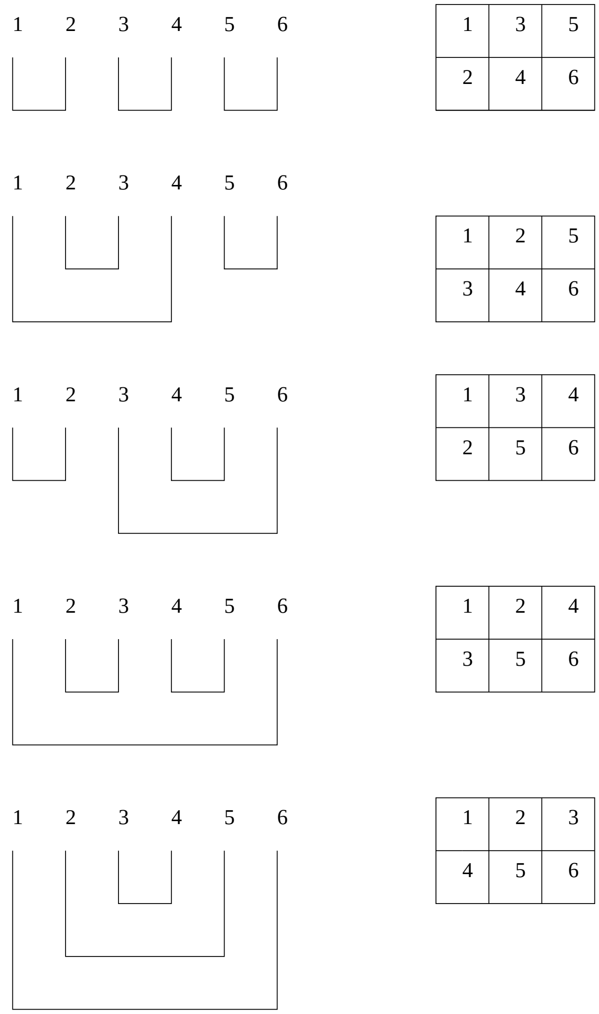

FIGURE 3. *Bijection between $\mathcal{P}_2(n)$ and $2 \times n$ Standard Young Tableaux.*

Definition 2. *A random variable s with odd moments $\phi(s^{2n+1}) = 0$ and even moments $\phi(s^{2n}) = \sigma^{2n} C_n$ where C_n is the n-th Catalan number and $\sigma > 0$ is a constant is called a* **semicircular element** *of variance σ^2.*

Thus we have the “Free Central Limit Theorem:”

Theorem 1. *If $(a_n)_{n\in\mathbb{N}}$ are freely independent and identically distributed with $\phi(a_i) = 0$ and $\phi(a_i^2) = \sigma^2$, then S_k converges in distribution to a semicircular element of variance σ^2 as $k \to \infty$.*

Recall that in Lecture 1 it was shown that for a random matrix X_N chosen from $GUE(N)$ we have that

$$\lim_{N\to\infty} \mathbb{E}[\mathrm{tr}(X^n)] = \begin{cases} 0, & \text{if } n \text{ odd} \\ C_{n/2}, & \text{if } n \text{ even} \end{cases} \tag{13}$$

so that a GUE random matrix is a semicircular element in the limit of large matrix dimension.

We can also define a family of semicircular random variables.

Definition 3. *$(s_i)_{i\in I}$ is called a* **semicircular family** *of covariance $(c_{ij})_{i,j\in I}$ if for any $n \geq 1$ and any n-tuple $i(1), \ldots, i(n) \in I$ we have*

$$\phi(s_{i(1)} \ldots s_{i(n)}) = \sum_{\pi\in NC_2(2n)} \kappa_\pi[s_{i(1)}, \ldots, s_{i(n)}],$$

where

$$\kappa_\pi[s_{i(1)}, \ldots, s_{i(n)}] = \prod_{(p,q)\in\pi} c_{i(p)i(q)}.$$

This is the free analogue of Wick's formula for Gaussian random variables.

In fact, using this language, it was shown in Lecture 1 that if $X_1, \ldots, X_r$ are matrices chosen independently from $GUE(N)$, then they converge in distribution in the large N limit to a semicircular family $s_1, \ldots, s_r$ of covariance $c_{ij} = \delta_{ij}$.

2. Non-Crossing Partitions and Free Cumulants

We begin by repeating the definition of a partition, which was given in Section 1.

Definition 4. *A partition of $[n] = \{1, \ldots, n\}$ is an unordered collection $\pi = \{V_1, \ldots, V_r\}$ of disjoint, non-empty subsets of $[n]$ whose union is $[n]$.*

The V_i's are called the *blocks* of π, and $\mathcal{P}(n) = \mathcal{P}([n])$ denotes the collection of all partitions of $[n]$.

Definition 5. *A partition $\pi \in \mathcal{P}(n)$ is called* **non-crossing** *if there do not exist numbers $p_1, q_1, p_2, q_2 \in [n]$ with $p_1 < q_1 < p_2 < q_2$ such that: p_1 and p_2 are in the same block of π, q_1 and q_2 are in the same block of π, and p_1, q_1 are not in the same block of π.*

The diagram below should make it clear what a "crossing" in a partition is; a non-crossing partition is a partition with no crossings.

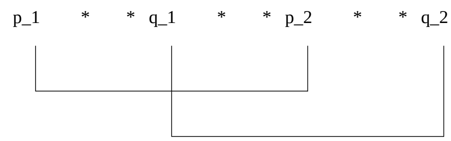

FIGURE 4. *A crossing in a partition*

Definition 6. *The collection of all non-crossing partitions of* $[n]$ *is denoted* $NC(n)$.

Note that $\mathcal{P}(n)$ is partially ordered by

$$\pi_1 \leq \pi_2 :\iff \text{ each block of } \pi_1 \text{ is contained in a block of } \pi_2. \tag{14}$$

We also say that π_1 is a *refinement* of π_2. $NC(n)$ is a subset of $\mathcal{P}(n)$ and inherits this partial order, so $NC(n)$ is an induced sub-poset of $\mathcal{P}(n)$. In fact both are lattices; they have well defined maximum $\vee$ and minimum $\wedge$ operations (though the max of two non-crossing partitions in $NC(n)$ does not necessarily agree with their max when viewed as elements of $\mathcal{P}(n)$).

We now define the important **free cumulants** of a non-commutative probability space (A, ϕ).

Definition 7. *Let* (A, ϕ) *be a noncommutative probability space. The corresponding free cumulants*

$$\kappa_n : A^n \to \mathbb{C}, \ n \geq 1 \tag{15}$$

are defined inductively in terms of mixed moments by

$$\phi(a_1 \ldots a_n) = \sum_{\pi \in NC(n)} \kappa_\pi(a_1, \ldots, a_n), \tag{16}$$

where by definition if $\pi = \{V_1, \ldots, V_r\}$ *then*

$$\kappa_\pi(a_1, \ldots, a_n) = \prod_{i=1}^{r} \kappa_{|V_i|}((a_i)_{i \in V_i}). \tag{17}$$

Example 2.1. *For* $n = 1$, *we have*

$$\phi(a_1) = \kappa_1(a_1) \implies \kappa_1(a_1) = \phi(a_1). \tag{18}$$

Example 2.2. *For $n = 2$, we have*

(19)
$$\phi(a_1a_2) = \kappa_{\{1,2\}}(a_1, a_2) + \kappa_{\{1\}\cup\{2\}}(a_1, a_2) = \kappa_2(a_1, a_2) + \kappa_1(a_1)\kappa_1(a_2).$$

Since we know from the $n = 1$ calculation that $\kappa_1(a_1) = \phi(a_1)$, this yields

(20)
$$\kappa_2(a_1, a_2) = \phi(a_1a_2) - \phi(a_1)\phi(a_2).$$

Example 2.3. *For $n = 3$, we have*

(21)
$$\phi(a_1a_2a_3) = \kappa_{\{1,2,3\}}(a_1, a_2, a_3) + \kappa_{\{1,2\}\cup\{3\}}(a_1, a_2, a_3)$$
(22)
$$+ \kappa_{\{1\}\cup\{2,3\}}(a_1, a_2, a_3) + \kappa_{\{1,3\}\cup\{2\}}(a_1, a_2, a_3) + \kappa_{\{1\}\cup\{2\}\cup\{3\}}(a_1, a_2, a_3)$$
(23)
$$= \kappa_3(a_1, a_2, a_3) + \kappa_2(a_1, a_2)\kappa_1(a_3) + \kappa_2(a_2, a_3)\kappa_1(a_1)$$
(24)
$$+ \kappa_2(a_1, a_3)\kappa_1(a_2) + \kappa_1(a_1)\kappa_1(a_2)\kappa_1(a_3).$$

Thus we find that

(25)
$$\kappa_3(a_1, a_2, a_3) = \phi(a_1a_2a_3) - \phi(a_1)\phi(a_2a_3) - \phi(a_2)\phi(a_1a_3) - \phi(a_3)\phi(a_1a_2) + 2\phi(a_1)\phi(a_2)\phi(a_3).$$

These three examples outline the general procedure of recursively defining κ_n in terms of mixed moments of length n.

Let us also point out how the definition looks for $a_1 = \cdots = a_n = a$, i.e. when all random variables are the same. Then we have

(26)
$$\phi(a^n) = \sum_{\pi \in NC(n)} \kappa_\pi(a, \ldots, a).$$

Thus if we write $\alpha_n := \phi(a^n)$ and $\kappa_\pi^a := \kappa_\pi(a, \ldots, a)$ this reads

$$\alpha_n = \sum_{\pi \in NC(n)} \kappa_\pi^a,$$

which we call the **moment-cumulant formula**.

The following are some important and non-trivial properties of free cumulants:

(1) κ_n is an n-linear function.
(2) There exists a combinatorial formula for dealing with cumulants whose arguments are products of random variables. For example, consider the evaluation of $\kappa_2(a_1a_2, a_3)$. It turns out

that this can be evaluated as

$$\begin{aligned}\kappa_2(a_1a_2, a_3) &= \sum_{\pi\in\{123,1|23,13|2\}} \kappa_\pi \\ &= \kappa_3(a_1,a_2,a_3) + \kappa_1(a_1)\kappa_2(a_2,a_3) + \kappa_2(a_1,a_3)\kappa_1(a_2).\end{aligned}$$

In general, the evaluation of a free cumulant of the form

$$\kappa_n(\prod_{i=1}^{m_1} a_{1i}, \ldots, \prod_{i=1}^{m_n} a_{ni})$$

involves summing $\kappa_\pi(a_{1i},\ldots,a_{1m_1},\ldots,a_{n1},\ldots,a_{nm_n})$ over all $\pi \in NC(m_1+\cdots+m_n)$ which have the property that they connect all different product strings.

Perhaps the most important property of free cumulants is that their vanishing in special situations characterizes free independence:

Theorem 2. *Let (A,ϕ) be a non-commutative probability space and let $\kappa_n, n \geq 1$ be the corresponding free cumulants. Then, subalgebras $A_1,\ldots,A_s \subset A$ are free if and only if all "mixed" cumulants with entries from $A_1,\ldots,A_n$ vanish. That is, $A_1,\ldots,A_n$ are free if and only if: whenever we choose $a_j \in A_{i(j)}$ in such a way that $i(k) \neq i(l)$ for some $k,l \in [n]$, then*

$$\kappa_n(a_1,\ldots,a_n) = 0.$$

The proof of this theorem on freeness and the vanishing of mixed cumulants relies on a key lemma, which we now describe.

Proposition 2.1. *Let (A,ϕ) be a non-commutative probability space and let $(\kappa_n)_{n\in\mathbb{N}}$ be the corresponding free cumulants. For $n \geq 2$, $\kappa_n(a_1,\ldots,a_n) = 0$ if $1 \in \{a_,\ldots,a_n\}$.*

Proof. We consider the case where the last argument a_n is equal to 1, and proceed by induction on n.

For $n = 2$,

$$\kappa_2(a,1) = \phi(a1) - \phi(a)\phi(1) = 0.$$

So the base step is done.

Now assume for the induction hypothesis that the result is true for all $k < n$. We have that

$$\begin{aligned}\phi(a_1\ldots a_{n-1}1) &= \sum_{\pi\in NC(n)} \kappa_\pi(a_1,\ldots,a_{n-1},1) \\ &= \kappa_n(a_1,\ldots,a_{n-1},1) + \sum_{\substack{\pi\in NC(n)\\ \pi\neq[n]}} \kappa_\pi(a_1,\ldots,a_{n-1},1).\end{aligned}$$

According to our induction hypothesis, a partition $\pi \neq [n]$ can have $\kappa_\pi(a_1, \ldots, a_{n-1}, 1)$ different from zero only if $\{n\}$ is a one-element block of π, i.e. $\pi = \sigma \cup \{n\}$ for some $\sigma \in NC(n-1)$. For such a partition we have

$$\kappa_\pi(a_1, \ldots, a_{n-1}, 1) = \kappa_\sigma(a_1, \ldots, a_{n-1})\kappa_1(1) = \kappa_\sigma(a_1, \ldots, a_{n-1}),$$

hence

$$\begin{aligned}\phi(a_1 \ldots a_{n-1}1) &= \kappa_n(a_1, \ldots, a_{n-1}, 1) + \sum_{\sigma \in NC(n-1)} \kappa_\sigma(a_1, \ldots, a_{n-1}) \\ &= \kappa_n(a_1, \ldots, a_{n-1}, 1) + \phi(a_1 \ldots a_{n-1}).\end{aligned}$$

Since $\phi(a_1 \ldots a_{n-1}1) = \phi(a_1 \ldots a_{n-1})$, we have proved that $\kappa_n(a_1, \ldots, a_{n-1}, 1) = 0$. □

Notice how much more efficient the result on the description of freeness in terms of cumulants is in checking freeness of random variables than the original definition of free independence. In the cumulant framework, we can forget about centerdness and weaken "alternating" to "mixed." Also, the problem of adding two freely independent random variables becomes easy on the level of free cumulants. If $a, b \in (A, \phi)$ are free with respect to ϕ, then

$$\begin{aligned}\kappa_n^{a+b} &:= \kappa_n(a+b, \ldots, a+b) \\ &= \kappa_n(a, \ldots, a) + \kappa_n(b, \ldots, b) + (\text{mixed cumulants in } a, b) \\ &= \kappa_n^a + \kappa_n^b.\end{aligned}$$

Thus the problem of calculating moments is shifted to the relation between cumulants and moments. We already know that the moments are polynomials in the cumulants, i.e. we know

$$\alpha_n = \sum_{\pi \in NC(n)} \kappa_\pi,$$

but we want to put this relationship into a framework more amenable to performing calculations.

For any $a \in A$, let us consider formal power series in an indeterminate z defined by

$$M(z) = 1 + \sum_{n=1}^{\infty} \alpha_n z^n, \text{ moment series}$$

$$C(z) = 1 + \sum_{n=1}^{\infty} \kappa_n^a z^n, \text{ cumulant series.}$$

We want to translate the moment-cumulant formula

$$\alpha_n = \sum_{\pi \in NC(n)} \kappa_\pi^a$$

into a statement about the relationship between the moment and cumulant series.

Proposition 2.2.

$$M(z) = C(zM(z)).$$

Proof. The trick is to sum first over the possibilities for the block of π containing 1, as in the derivation of the recurrence for C_n. Suppose that the first block of π looks like

$$V = \{1, v_2, \dots, v_s\}$$

where $1 < v_1 < \cdots < v_s \leq n$. Then we build up the rest of the partition π out of smaller "nested" non-crossing partitions $\pi_1', \dots, \pi_s'$ with $\pi_1' \in NC(\{2, \dots, v_2 - 1\})$, $\pi_2' \in NC(\{v_2 + 1, \dots, v_3 - 1\})$, etc. Hence if we denote $i_1 = |\{2, \dots, v_2 - 1\}|, i_2 = |\{v_2 + 1, \dots, v_3 - 1\}|$, etc., then we have

$$\begin{aligned}
\alpha_n &= \sum_{s=1}^{n} \sum_{\substack{i_1, \dots, i_s \geq 0 \\ s + i_1 + \dots + i_s = n}} \sum_{\pi = V \cup \pi_1' \cup \dots \cup \pi_s'} \kappa_s \kappa_{\pi_1'} \dots \kappa_{\pi_s'} \\
&= \sum_{s=1}^{n} \sum_{\substack{i_1, \dots, i_s \geq 0 \\ s + i_1 + \dots + i_s = n}} \sum_{\pi = V \cup \pi_1' \cup \dots \cup \pi_s'} \kappa_s \Big(\sum_{\pi_1' \in NC(i_1)} \kappa_1' \Big) \dots \Big(\sum_{\pi_s' \in NC(i_s)} \kappa_s' \Big) \\
&= \sum_{s=1}^{n} \sum_{\substack{i_1, \dots, i_s \geq 0 \\ s + i_1 + \dots + i_s = n}} \kappa_s \alpha_{i_1} \dots \alpha_{i_s}.
\end{aligned}$$

Thus we have

$$\begin{aligned}
1 + \sum_{n=1}^{\infty} \alpha_n z^n &= 1 + \sum_{n=1}^{\infty} \sum_{s=1}^{n} \sum_{\substack{i_1, \dots, i_s \geq 0 \\ s + i_1 + \dots + i_s = n}} \kappa_s z^s \alpha_{i_1} z^{i_1} \dots \alpha_{i_s} z^{i_s} \\
&= 1 + \sum_{s=1}^{\infty} \kappa_s z^s \Big(\sum_{i=0}^{\infty} \alpha_i z^i \Big)^s.
\end{aligned}$$

□

Now consider the **Cauchy transform** of a :

$$G(z) := \phi(\frac{1}{z-a}) = \sum_{n=0}^{\infty} \frac{\phi(a)}{z^{n+1}} = \frac{1}{z} M(\frac{1}{z})$$

and the R-**transform** of a defined by

$$R(z) := \frac{C(z)-1}{z} = \sum_{n=0}^{\infty} \kappa_{n+1}^{a} z^n.$$

Also put $K(z) = R(z) + \frac{1}{z} = \frac{C(z)}{z}$. Then we have the relations

$$K(G(z)) = \frac{1}{G(z)} C(G(z)) = \frac{1}{G(z)} C(\frac{1}{z} M(\frac{1}{z})) = \frac{1}{G(z)} z(G(z)) = z.$$

Thus we have the following theorem of Voiculescu on "free convolution:"

Theorem 3. *For a random variable a let $G^a(z)$ be its Cauchy transform and define its R-transform by*

$$G[R(z) + \frac{1}{z}] = z. \tag{27}$$

Then, for a and b freely independent, we have

$$R^{a+b}(z) = R^a(z) + R^b(z). \tag{28}$$

At the moment these are idenitites on the level of formal power series, i.e. identities in the ring $\mathbb{C}[[z]]$. In the next lecture, we will elaborate on their interpretation as identities concerning analytic functions.