The Boolean quadratic forms and tangent law

Wiktor Ejsmont join work Patrycja Hęćka

Probablistic Operator Algebra Seminar, Berkeley

Plan of the talk

1. Free probability
2. Free central limit theorems for quadratic forms
3. Boolean probability
4. Boolean central limit theorems for quadratic forms

Basic notation

We consider a non commutative probability space $(\mathcal{A}, \tau)$. $\mathcal{A}$ is an algebra τ is a linear functional such that

1. τ is linear, weak*-continuous,
2. $\tau(\mathbb{I}) = 1$ - normal,
3. $\tau(XX^*) \geq 0$ - positive,
4. $\tau(XX^*) = 0$ implies $X = 0$ - faithful.

Distribution

A (noncommutative) random variable X is a self-adjoint element of $\mathcal{A}$ with a probability measure μ on $\mathbb{R}$ such as

$$\tau(X^n) = \int_{\mathbb{R}} x^n \mu(dx).$$

Free independence

We say that subalgebras $\mathcal{A}_i \subset \mathcal{A}$ are free independent if for every choice of $i_1 \neq i_2 \ldots \neq i_n$ and every choice of $X_i \in \mathcal{A}_i$ such that $\tau(X_i) = 0$ we have

$$\tau(X_1 X_2 \ldots X_n) = 0,$$

Example

Let G be the free product of groups $G_i, i \in I$, let $A = \mathbb{C}G$ be the group algebra,

$\phi(g) = \delta_{g=e}$ be the group trace, and set $A_i = \mathbb{C}G_i \subset A$.

Then $A_i : i \in I$ are free independent.

Let μ and ν be probability measures on $\mathbb{R}$, and X, Y self-adjoint free random variables with respective distributions μ and ν.

The distribution of $X + Y$ is called the free additive convolution of μ and ν and is denoted by $\mu \boxplus \nu$.

The analytic approach to free convolution is based on the Cauchy transform

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

For measures with compact support the Cauchy transform is analytic at infinity and related to the moment generating function M_X as follows:

$$M_X(z) = \sum_{n=0}^{\infty} \tau(X^n)\, z^n = \frac{1}{z} G_X(1/z).$$

The Cauchy transform has an inverse in some neighbourhood of infinity which has the form

$$G_{\mu}^{-1}(z) = \frac{1}{z} + R_{\mu}(z),$$

where $R_{\mu}(z)$ is analytic in a neighbourhood of zero and is called *R-transform*.

The coefficients of its series expansion

$$R_X(z) = \sum_{n=0}^{\infty} K_{n+1}(X)\, z^n.$$

are called *free cumulants* of the random variable X.

The free convolution can now be computed via the identity

$$R_{\mu \boxplus \nu}(z) = R_\mu(z) + R_\nu(z).$$

The Wigner semicircle law has density

$$d\mu(x) = \frac{1}{2\pi}\sqrt{4 - x^2}\, dx$$

on $-2 \leq x \leq 2$. Its Cauchy-Stieltjes transform is given by the formula

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

and the R-transform is

$$R_X(z) = z.$$

We say that a sequence X_n of random variables *converges in distribution* towards X as $n \to \infty$, denoted by

$$X_n \xrightarrow{distr} X$$

if we have for all $m \in \mathbb{N}$

$$\lim_{n\to\infty} \tau(X_n^m) = \tau(X^m).$$

Let X_i be free copies of a centered random variable X of variance 1, then the sequence of quadratic forms

$$Q_n = \frac{1}{n}\sum_{i,j=1}^{n} a_{i,j}^{(n)} X_i X_j$$

converges in distribution to a compound free Poisson distribution with jump distribution μ.

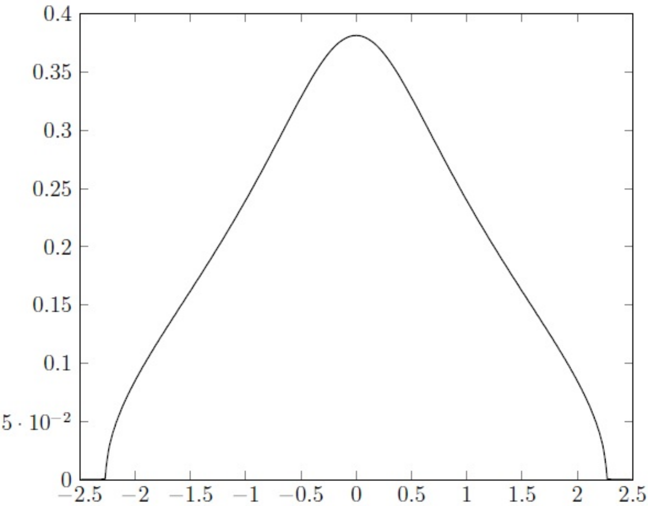
The free tangent law, W. Ejsmont and F. Lehner 2000

Let $X_1, X_2, \ldots, X_n \in \mathcal{A}_{sa}$ be free centred identically distributed random variables of variance 1, then

$$Q_n = \frac{1}{n} \sum_{k,j=1}^{n} i(X_k X_j - X_j X_k) \xrightarrow{distr} Y,$$

where $R_Y(z) = \tan(z)$.

It is worth mentioning that one of the fundamental examples of Nevanlinna functions is the tangent function, see Donoghue, 1974.



Density of the free tangent law

Proof 1.

Let

$$Q_n = \frac{1}{n} \sum_{i,j=1}^{n} a_{i,j}^{(n)} X_i X_j$$

where X_i are free centred identically distributed random variables of variance 1.
Let

$$Q'_n = \frac{1}{n} \sum_{i,j=1}^{n} a_{i,j}^{(n)} X_i X_j$$

where X_i sequence of standard free semicircular variables.
Then

$$lim_{n\to\infty} K_r(Q_n) = lim_{n\to\infty} K_r(Q'_n)$$

Proof 1.

We may assume without loss of generality that X_i are standard free semicircular variables.

The cumulants can be computed using formula of Leonov/Shiryaev and evaluate to

$$K_r\Big(\sum_{\substack{k,j=1\\k<j}}^{n} i(X_kX_j - X_jX_k)\Big) = \mathrm{Tr}(A_n^r) \text{ where } A_n = \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}_n,$$

hence the odd cumulants vanish.

The eigenvalues of the matrix A_n are $\lambda_k = \cot\left(\frac{\pi}{2n} + \frac{k}{n}\pi\right)$ for $k \in \{0, \dots, n-1\}$ hence the even cumulants evaluate to

$$K_{2m}\Big(\sum_{\substack{k,j=1\\k<j}}^{n} i(X_kX_j - X_jX_k)\Big) = \sum_{k=0}^{n-1} \cot^{2m}\left(\frac{\pi}{2n} + \frac{k}{n}\pi\right)$$

$$= n^{2m}\frac{T_{2m-1}}{(2m-1)!} + \mathcal{O}(n^{2m-2})$$

where T_{2m-1} are the tangent numbers

$$\tan z = \sum_{n=1}^{\infty} T_n \frac{z^n}{n!} = z + \frac{2}{3!}z^3 + \frac{16}{5!}z^5 + \frac{272}{7!}z^7 + \cdots,$$

Hence

$$\lim_{n\to\infty} K_{2m}(Q_n) = \frac{T_{2m-1}}{(2m-1)!}$$

and we conclude that

$$\lim_{n\to\infty} R_{Q_n}(z) = \tan(z).$$

Proof 2.

The matrix $\frac{1}{n}A_n = \frac{1}{n}\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}_n$ has characteristic polynomial

$$\chi_n(\lambda) = \frac{i(\lambda - \frac{i}{n})^n + i(\lambda + \frac{i}{n})^n}{2i}.$$

The cumulant generating function

$$R_{Q_n}(z) = \sum_{k=1}^{\infty} \frac{\mathrm{Tr}(A_n^k)}{n^k} z^{k-1},$$

can be obtained from the logarithmic derivative of the characteristic polynomial.

Indeed if we factorize the characteristic polynomial $\chi_n(\lambda) = \prod_{i=1}^{n}(\lambda - \lambda_i)$ then

$$\frac{\chi_n'(\lambda)}{\chi_n(\lambda)} = \sum_{i=1}^{n} \frac{1}{\lambda - \lambda_i}$$

and

$$\frac{1}{z}\frac{\chi_n'(1/z)}{\chi_n(1/z)} = \sum_{k=0}^{\infty}\sum_{i=1}^{n} \lambda_i^k z^k = n + zR_{Q_n}(z).$$

In our case

$$\lim_{n\to\infty} R_{Q_n}(z) = \lim_{n\to\infty} \frac{1}{z}\left(\frac{1}{z}\frac{\chi_n'(1/z)}{\chi_n(1/z)} - n\right)$$
$$= \tan(z).$$

The free zigzag law

Let $X_1, X_2, \ldots, X_n \in \mathcal{A}_{sa}$ be free centred identically distributed random variables of variance 1,, then

$$Q_n = \frac{1}{2n} \sum_{\substack{k,l=1 \\ k<l}}^{n} \left(X_k X_l + X_l X_k + i(X_k X_l - X_l X_k) \right) \xrightarrow{distr} Y,$$

where $R_Y(z) = \frac{1}{2}(\tan(z) + \sec(z) - 1)$.

An alternating permutation (or zigzag permutation) of the set $\{1, 2, 3, ..., n\}$ is an arrangement of those numbers so that each entry is alternately greater or less than the preceding entry.

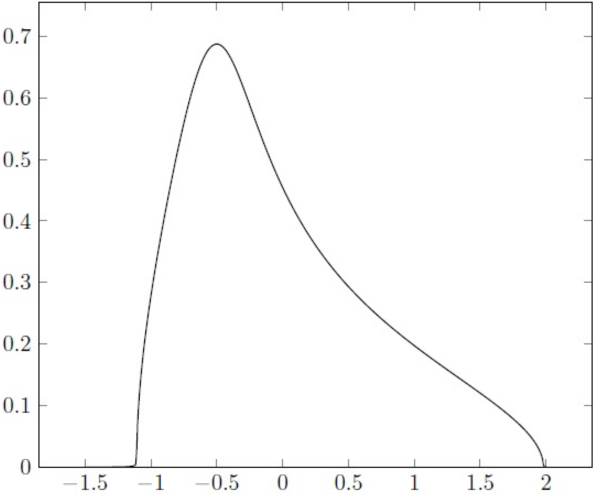
For example, the five alternating permutations of 1, 2, 3, 4 are:

- 1, 3, 2, 4 because $1 < 3 > 2 < 4$,
- 1, 4, 2, 3 because $1 < 4 > 2 < 3$,
- 2, 3, 1, 4 because $2 < 3 > 1 < 4$,
- 2, 4, 1, 3 because $2 < 4 > 1 < 3$,
- 3, 4, 1, 2 because $3 < 4 > 1 < 2$.

The determination of the number E_n of alternating permutations of the set $\{1, ..., n\}$ is called André's problem.

Theorem (Désiré André, 1879)

$$\sum_{n=0}^{\infty} \frac{E_n}{n!} = \tan(z) + \sec(z).$$



Density of the free zigzag law

Free generalized tangent law

Let $X_1, X_2, \dots, X_n \in \mathcal{A}_{sa}$ be free centred identically distributed random variables of variance 1,, then for any $a, b \in \mathbb{R}$ with $a^2 + b^2 = 1$ and $b \neq 0$, the limit law

$$Q_n = \frac{1}{n} \sum_{\substack{k,j=1 \\ k<j}}^{n} \left(a(X_k X_j + X_j X_k) + ib(X_k X_j - X_j X_k)\right) \xrightarrow{distr} Y,$$

has R-transform

$$R_Y(z) = \frac{\tan(bz)}{b - a\tan(bz)}.$$

A tangent number $T_n^{(k+1)}$ of order $k+1$ defined by

$$\tan^{k+1}(z) = \sum_{n=k+1}^{\infty} \frac{T_n^{(k+1)} z^n}{n!}$$

Carlitz and Scoville's [Duke Math. J. 39 (1972)]

$$\frac{\tan(z)}{1 - x\tan(z)} = \sum_{n=1}^{\infty} P_n(x)\frac{z^n}{n!}$$

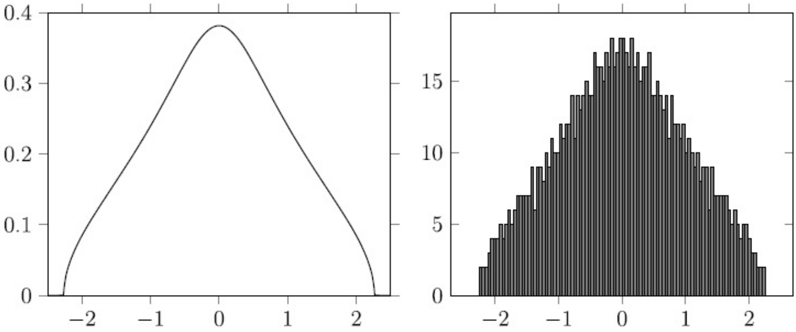
where $P_n(x) = \sum_{k=0}^{n-1} T_n^{(k+1)} x^k$

Free generalized tangent law

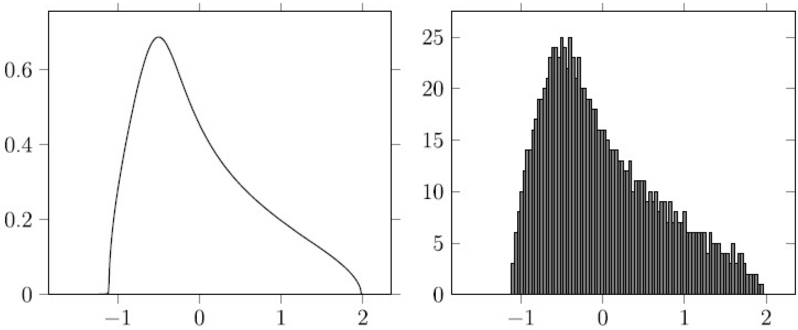
$$R_Y(z) = \frac{\tan(bz)}{b - a\tan(bz)} = \sum_{n=1}^{\infty} \left(\sum_{k=0}^{n-1} T_n^{(k+1)} a^k b^{n-1-k} \right) \frac{z^n}{n!}$$

$$K_n = \frac{\sum_{k=0}^{n-1} T_n^{(k+1)} a^k b^{n-1-k}}{n!}$$

Random matrix model



Random matrix model



Random matrix model

Let $X_{N\times NM}$ be a complex Gaussian random matrix of size $N \times NM$
and let

$$A_M = \begin{bmatrix} 0 & a+ib & a+ib & \dots & a+ib \\ a-ib & 0 & a+ib & \dots & a+ib \\ a-ib & a-ib & 0 & \ddots & a+ib \\ \vdots & & \ddots & \ddots & \\ a-ib & a-ib & a-ib & \dots & 0 \end{bmatrix}.$$

Random matrix model

Let P_N be a sequence of $N \times N$ deterministic matrices all of whose moments with respect to the normalized trace converge to 1, e.g., the identity matrices $P_N = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]_N$ or any projection matrix of large rank like $P_N = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]_N - \frac{1}{N}\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]_N$, then the spectral measures of

$$\frac{1}{M} X_{N \times NM}[A_M \otimes P_N] X^*_{N \times NM}$$

converge in distribution to the free generalized tangent law

$$R(z) = \frac{\tan(bz)}{b - a\tan(bz)}.$$

Boolean Independence.

We say that subalgebras $\mathcal{A}_i \subset \mathcal{A}$ are Boolean independent if for every choice of $i_1 \neq i_2 \ldots \neq i_n$ and every choice of $X_i \in \mathcal{A}_i$, we have

$$\tau(X_1 X_2 \ldots X_n) = \tau(X_1)\tau(X_2)\ldots\tau(X_n)$$

Example

The most natural state on the group algebra of the free group $\mathrm{F_N}$ with the free generators x_1 , $x_2, \ldots,$ x_N is the Haagerup state

$$H_q(g) = q^{l(g)}, \quad g \in \mathrm{F_N}, q \in (0,1)$$

where $g = x_{i_1}^{n_1} x_{i_2}^{n_2} \ldots x_{i_k}^{n_k}$, $l(g) = \sum_i |n_i|$ and $l(e) = 0$.

Then group subalgebras $C[G_i]$ are Boolean independent where G_i are cyclic group generated by x_i.

Let μ and ν be probability measures on $\mathbb{R}$, and X, Y self-adjoint Boolean independent random variables with respective distributions μ and ν.

The distribution of $X + Y$ is called the Boolean additive convolution of μ and ν and is denoted by $\mu \uplus \nu$.

The analytic approach to Boolean convolution is based on the Cauchy transform

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

For measures with compact support the Cauchy transform is analytic at infinity and related to the moment generating function M_X as follows:

$$M_X(z) = \sum_{n=0}^{\infty} \tau(X^n)\, z^n = \frac{1}{z} G_X(1/z).$$

Moreover, the moment generating transform can be written as

$$M_{\mu}(z) = \frac{1}{1 - H_{\mu}(z)},$$

where $H_{\mu}(z)$ is analytic in a neighbourhood of zero. The coefficients of its series expansion

$$H_X(z) = \sum_{n=1}^{\infty} K_n(X) z^n$$

are called *Boolean cumulants* of the random variable X.

Boolean Gaussian distribution

The Boolean Gaussian law with mean zero and variance a^2 has distribution

$$\frac{1}{2}\delta_{-a} + \frac{1}{2}\delta_{a}.$$

Boolean Gaussian distribution

In contrast to the classical convolution, it is not true that, for arbitrary $a \in \mathbb{R}$, the convolution $\mu \uplus \delta_c$ is equal to the shift of measure μ by the amount c.

For example, $(\frac{1}{2}\delta_{-a} + \frac{1}{2}\delta_a) \uplus \delta_c$ is equal to

$$\frac{\left(1 + \frac{c}{\sqrt{4a^2+c^2}}\right)\delta_{(c+\sqrt{4a^2+c^2})/2} + \left(1 - \frac{c}{\sqrt{4a^2+c^2}}\right)\delta_{(c-\sqrt{4a^2+c^2})/2}}{2}.$$

Boolean Gaussian distribution

We say that a family X_i $i \in \{1, \ldots, n\}\}$ is a *Boolean standard normal family* if $K_1(X_i) = K_2(X_i) = 1$, $K_r(X_i) = 0$ for $r > 2$ and X_i are Boolean independent

Boolean generalized tangent law, Ejsmont Hęćka

Let $X_1, \ldots, X_n$ be identically distributed Boolean independent random variables with mean $\tau(X_i) = \frac{1}{\sqrt{n}}$ and variance 1. Then the sequence of quadratic forms

$$Q_n = \frac{1}{n} \sum_{\substack{k,j=1 \\ k<j}}^{n} \left(a(X_k X_j + X_j X_k) + ib(X_k X_j - X_j X_k)\right) \xrightarrow{d} Y, \quad a, b \in \mathbb{R}$$

where the H-transform of the limit distribution has the form

$$H_Y(z) = \frac{1}{z} \frac{\tan(bz)}{b - a\tan(bz)} - 1.$$

The limit Boolean and free cumulants are not exactly the same

$$K_r^{\text{Boolean}} = K_{r+1}^{\text{free}}$$

The main problem

Let $X_1, \ldots, X_n$ be identically distributed Boolean independent normal random variables and

$$Q_n = \frac{1}{n} \sum_{\substack{k,j=1 \\ k<j}}^{n} \left(a(X_k X_j + X_j X_k) + ib(X_k X_j - X_j X_k) \right)$$

In the case of the Boolean normal distribution, with mean c, variance 1 takes the particular form

$$K_r(Q_n) = c^2 \operatorname{Tr}(JA_n^r)$$

where

$$J := \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \dots & 1 \\ 1 & 1 & 1 & \ddots & 1 \\ \vdots & & \ddots & \ddots & \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}$$

$$A_n := \begin{bmatrix} 0 & a+ib & a+ib & \dots & a+ib \\ a-ib & 0 & a+ib & \dots & a+ib \\ a-ib & a-ib & 0 & \ddots & a+ib \\ \vdots & & \ddots & \ddots & \\ a-ib & a-ib & a-ib & \dots & 0 \end{bmatrix}.$$

Now we focus on the special case $a = 0$ and $b = 1$ because in this situation we are able to determine the corresponding measures.

We call the limit law μ_Y the Boolean tangent law if

$$H_Y(z) = \frac{\tan z}{z} - 1.$$

In this case the corresponding transforms are given by

$$M_\mu(z) = \frac{z}{2z - \tan(z)}, \quad G_\mu(z) = \frac{1}{2z - z^2 \tan(1/z)}.$$

If we use the Stieltjes inversion formula, namely

$$d\mu(x) = -\frac{1}{\pi} \lim_{\epsilon \to 0^+} \operatorname{Im} G_\mu(x+i\epsilon) = 0 \quad \text{for } 2/x \neq \tan(1/x) \text{ and } x \neq 0.$$

Thus the measure μ has no absolutely continuous part.

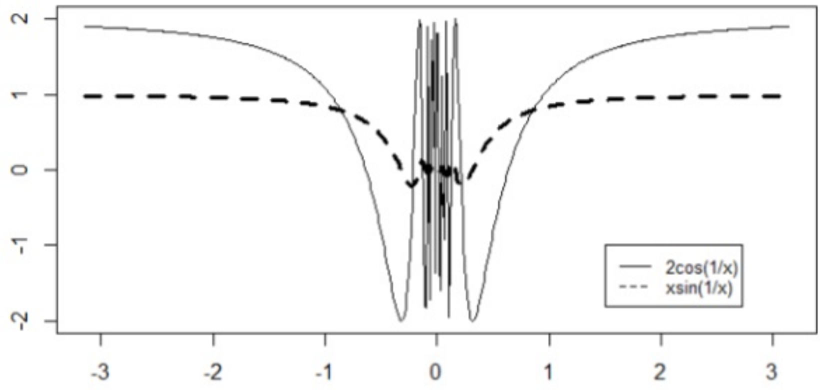
In order to determine the atoms, we compute the limits

$$\lim_{\epsilon \to 0^+} i\epsilon G_\mu (x + i\epsilon) \quad \text{for } \ 2/x = \mathtt{tan}(1/x) \text{ and } x = 0.$$

Finally, we get

$$\mu(\{x\}) = \begin{cases} \frac{x^2}{4-x^2} & \text{for } x \in \{x \mid 2/x = \tan(1/x)\}, \\ \frac{1}{2} & \text{for } x = 0. \end{cases}$$

$$2/x = \tan(1/x) \iff 2\cos(1/x) = x\sin(1/x)$$



Thank you for your attention