

Lecture 8: April 13, 2011

As we saw last time, if $C_n = \frac{1}{n+1} \binom{2n}{n}$ are the Catalan #s and $\mathcal{C}(z) = \sum_n C_n z^n$, then for $|z| < 1/4$,

$$\mathcal{C}(z) = \frac{1}{2z} (1 - \sqrt{1-4z}).$$

Now, since C_n counts the # of Dyck paths of length $2n$, if S is the unilateral shift and $X = S + S^*$, then

$$\varphi(X^n) = \begin{cases} C_{n/2}, & n \text{ even} \\ 0, & n \text{ odd} \end{cases}$$

So, if we set $M(z) = \sum_n \varphi(X^n) z^n$, we get

$$M(z) = \mathcal{C}(z^2) = \frac{1}{2z^2} (1 - \sqrt{1-4z^2}), \quad |z| < \frac{1}{2}.$$

Note: the expression on the RHS is a well-defined function for larger $|z|$ too — the only question is making sense of $\sqrt{}$; we can certainly make sense of this function on \mathbb{C}_+ or \mathbb{C}_- (upper- or lower-half-plane). The power series only converges for $|z| < \frac{1}{2}$ since this function is not analytic @ $z = \pm \frac{1}{2}$.

But the larger function does have (critical!) meaning to the distribution μ_X .

Recall: in Lecture 6, we computed that (at least in the ball of convergence) the moment power-series of a self-adjoint r.v. x with law μ_x is actually an integral transform of μ_x :

$$M(z) = \sum_{n=0}^{\infty} \varphi(x^n) z^n = \int_{\mathbb{R}} \frac{1}{1-tz} \mu_x(dt).$$

So, we see that, at least for $|z| < \frac{1}{2}$,

$$\frac{1}{2z^2} (1 - \sqrt{1-4z^2}) = M(z) = \int_{\mathbb{R}} \frac{1}{1-tz} \mu_x(dt).$$

Key observation: both $\frac{1}{2z^2} (1 - \sqrt{1-4z^2})$ and $\int_{\mathbb{R}} \frac{1}{1-tz} \mu_x(dt)$ are analytic functions in \mathbb{C}_{\pm} . Since they agree on the ball $B_{1/2}(0)$, they must agree everywhere on \mathbb{C}_{\pm} .

To match up with historical convention (and for convenience as we will soon see), we will customarily work with a simple change of variables from $M(z)$.

Def: Let μ be a (compactly-supported) probability measure on \mathbb{R} . The Stieltjes transform aka Cauchy transform of μ is the analytic function

$$G_\mu: \mathbb{C}_+ \rightarrow \mathbb{C}_-$$

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

Remarks: (1) $\operatorname{Im} \frac{1}{z-t} = \operatorname{Im} \frac{\bar{z}-t}{|z-t|^2} = \frac{1}{|z-t|^2} \operatorname{Im} \bar{z} < 0$ if $\operatorname{Im} z > 0$, so since μ is supported in \mathbb{R} , indeed $G_\mu: \mathbb{C}_+ \rightarrow \mathbb{C}_-$.

(2) To check that G_μ is analytic everywhere in \mathbb{C}_+ , differentiate $\frac{\partial}{\partial \bar{z}}$ under the integral — this is justified easily using the DCT and the real support of μ . Since $\frac{1}{z-t}$ is analytic on \mathbb{C}_+ , $\frac{\partial}{\partial \bar{z}} = 0$. This argument fails for any $z \in \operatorname{supp} \mu$, since $\frac{1}{z-t}$ blows up here.

(3) The relation between G_{μ_X} and the moment series M_X of X is simply

$$G_{\mu_X}(z) = \int_{\mathbb{R}} \frac{1}{z-t} \mu_X(dt) = \frac{1}{z} \int_{\mathbb{R}} \frac{1}{1-t/z} \mu_X(dt) = \frac{1}{z} M_X\left(\frac{1}{z}\right),$$

for $|z|$ sufficiently large that $\frac{1}{z} \in$ ball of convergence of M_X .

In particular, we find that for $X=S+S^*$,

$$G_{\mu_X}(z) = \frac{1}{z} M_X\left(\frac{1}{z}\right) = \frac{1}{z} \cdot \frac{z^2}{2} \left(1 - \sqrt{1 - \frac{4}{z^2}}\right) = \frac{1}{2} (z - \sqrt{z^2 - 4}),$$

for $|z| > 2$. Note, however, that G_{μ_X} and $\frac{1}{2}(z - \sqrt{z^2 - 4})$ are both analytic in \mathbb{C}_+ ; since they agree on the large set $\{z \in \mathbb{C}_+ : |z| > 2\}$, they therefore agree on their whole domain. I.e.

$$G_{\mu_X}(z) = \frac{1}{2} (z - \sqrt{z^2 - 4}), \quad z \in \mathbb{C}_+.$$

(4) The normalization we used to pick the correct sign of $M_z(z)$ is $M_z(0) = 1$ (because $M_z(0) = \varphi(z^*) = \varphi(1) = 1 - i\varepsilon$. φ is a state, so μ_x is a probability measure). This translates to the property

$$\lim_{|z| \rightarrow \infty} z \cdot G_\mu(z) = 1$$

for any Stieltjes transform.

Why work with G_μ instead of M_z ? It turns out that the distribution μ_z can be picked off the formula for G_μ easily.

Theorem: (Stieltjes inversion formula) Let μ be a probability measure on \mathbb{R} . Then

$$\mu(dt) = -\frac{1}{\pi} w\text{-}\lim_{\varepsilon \downarrow 0} G_\mu(t+i\varepsilon) dt.$$

Remarks: (1) What this weak limit statement says literally is that, for $f \in C_c(\mathbb{R})$,

$$\int_{\mathbb{R}} f(t) \mu(dt) = -\frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \int_{\mathbb{R}} f(t) G_\mu(t+i\varepsilon) dt.$$

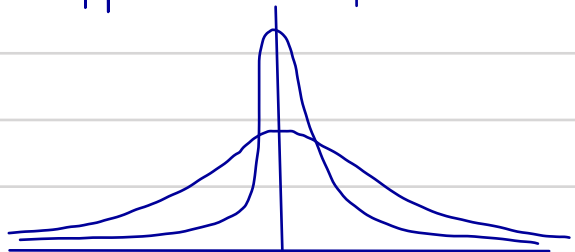
So, although μ may not have a density, the limiting values of $-\frac{1}{\pi} \text{Im} G$ on \mathbb{R} give an approximate density for μ .

(2) In particular, if μ does have a density $\mu(dt) = g(t)dt$, then

$$g(t) = -\frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \text{Im} G(t+i\varepsilon).$$

Pf. This is really just the Poisson integral formula for the unit disk conformally mapped to the upper half-plane. For $\varepsilon > 0$ and $t \in \mathbb{R}$, define

$$P_\varepsilon(t) = \frac{1}{\pi} \frac{\varepsilon}{t^2 + \varepsilon^2}.$$



Note: $\int_{\mathbb{R}} P_\varepsilon(t) dt = 1$, $P_\varepsilon > 0$, and P_ε concentrates around 0 as $\varepsilon \downarrow 0$.

Now, calculate that

$$\begin{aligned} -\frac{1}{\pi} \operatorname{Im} G_{\mu}(t+i\varepsilon) &= -\frac{1}{\pi} \operatorname{Im} \int_{\mathbb{R}} \frac{1}{t+i\varepsilon-s} \mu(ds) \\ &= -\frac{1}{\pi} \operatorname{Im} \int_{\mathbb{R}} \frac{t-s-i\varepsilon}{(t-s)^2+\varepsilon^2} \mu(ds) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{\varepsilon}{(t-s)^2+\varepsilon^2} \mu(ds) \\ &= \int_{\mathbb{R}} P_{\varepsilon}(t-s) \mu(ds). \end{aligned}$$

That is: $-\frac{1}{\pi} \operatorname{Im} G_{\mu}(t+i\varepsilon) = P_{\varepsilon} * \mu(t)$. Since P_{ε} is an approx. identity, the result follows. ///

In particular, this means we can easily find the distribution of $X = S+S^*$.

$$G_{\mu_X}(z) = \frac{1}{2} (z - \sqrt{z^2 - 4}).$$

$$-\frac{1}{\pi} G_{\mu_X}(t+i\varepsilon) = -\frac{1}{2\pi} (t+i\varepsilon) + \frac{1}{2\pi} \sqrt{(t+i\varepsilon)^2 - 4}.$$

Letting $\varepsilon \downarrow 0$, we get $-\frac{t}{2\pi} + \frac{1}{2\pi} \sqrt{t^2 - 4}$.

- real whenever $|t| \geq 2$.
- when $|t| < 2$, $\operatorname{Im} = \frac{1}{2\pi} \sqrt{4 - t^2}$.

Conclusion: μ_X has a density:

$$\frac{d\mu_X}{dt} = \frac{1}{2\pi} \sqrt{4 - t^2} \mathbb{1}_{|t| \leq 2}.$$

This is the famous Semicircle Law that first appeared (famously) in Wigner's seminal work in random matrix theory. This is no accident — as we will soon see, the semicircle law is the most important distribution in free prob., and the operator $X = S+S^*$ is its most important "carrier".

To review:

- * Given a self-adjoint element $x \in (\mathcal{O}, \varphi)$, its moments are the real numbers $\{\varphi(x^n); n \in \mathbb{N}\}$. There is at most one probability measure μ_x on \mathbb{R} that has the same moments:
compactly-supported
$$\int_{\mathbb{R}} t^n \mu_x(dt) = \varphi(x^n), \quad n = 1, 2, 3, \dots$$

(If (\mathcal{O}, φ) arises from an operator algebra and φ is sufficiently continuous, existence of μ_x is guaranteed.)

- * To find μ_x from the moments of x , calculate the Stieltjes transform

$$G_{\mu_x}(z) = \sum_{n=0}^{\infty} \frac{\varphi(x^n)}{z^{n+1}}, \quad |z| \text{ sufficiently large.}$$

This only makes sense if $\{\varphi(x^n)\}$ has at-most-exponential growth.

Exercise: If $\mathcal{O} \subset \mathcal{B}(H)$ and φ is a vector state on \mathcal{O} , show that for $a \in \mathcal{O}$, $|\varphi(a^n)| \leq \|a\|^n$.

- * If μ_x exists, the power-series will have an analytic continuation to all of \mathbb{C}_+ . So we compute this extension (hopefully, symmetries in the sequence $\{\varphi(x^n)\}$ will yield an algebraic or differential equation that G_{μ_x} solves).

- * Once G_{μ_x} is known, calculate μ_x via the Stieltjes inversion formula

$$\mu(dt) = -\frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \operatorname{Im} G_{\mu_x}(t+i\varepsilon) dt.$$

Now, what does any of this have to do with freeness?

Free Convolution

Freeness actually yields a new, important operation on probability measures. To understand this, we first need the concept of a joint distribution.

Def: Let $x, y \in (\mathcal{O}, \varphi)$. The joint distribution of x, y , $\mu_{x,y}$, is the collection of all joint moments $\varphi(x^{n_1} y^{m_1} \dots x^{n_k} y^{m_k})$.

We have already seen one example: the $*$ -distribution μ_{a,a^*} is really the joint distribution of a and a^* .

As in that example, we can embody $\mu_{x,y}$ as a linear functional on non-commutative polynomials in 2 variables

$$\begin{aligned}\mu_{x,y}: \mathbb{C}[X,Y] &\rightarrow \mathbb{C} \\ \mu_{x,y}(P) &= \varphi(P(x,y)).\end{aligned}$$

Remark: we might also want to talk about joint $*$ -distributions - the collection of all joint moments in x, x^*, y, y^* . For the time-being, we are primarily interested in the case x, y are self-adjoint.

Now, freeness is an independence rule: i.e. it is an algorithm for computing joint moments from individual ones. I.e.

If x, y are free, then (μ_x, μ_y) determines $\mu_{x,y}$.

On the other hand, here is a typical sort of statement:

Prop: If $x, y \in (\mathcal{A}, \varphi)$ are self-adjoint, then $\mu_{x,y}$ determines μ_{x+y} .

Pf. μ_{x+y} is determined by the moments

$$\varphi((x+y)^n) = \sum_{\eta \in \{1,2\}^n} \varphi(\underbrace{x\eta_1 \cdots x\eta_n}_{\text{these are exactly the data contained in } \mu_{x,y}}), \quad x_1 = x, x_2 = y.$$

these are exactly the data contained in $\mu_{x,y}$. ///

So, putting these things together, we have:

$$x, y \text{ self adjoint \& free} \implies x+y$$

$$(\mu_x, \mu_y) \implies \mu_{x,y} \implies \mu_{x+y} \equiv \mu_x \boxplus \mu_y.$$

I.e. freeness gives us a new way to combine measures corresponding to "independent" random variables: free convolution

Our goal for much of this course is to understand the operation \boxplus . It is in this understanding that we will develop and use much more sophisticated combinatorial tools than we've seen so far.

Before we start down that path, we must first answer a basic question.

Question: \boxplus is nominally an operation on prob. measures, but to define it, we need to realize a given pair of distributions as free random variables. So:

Given two compactly-supported probability measures μ, ν , does there exist a NCPs with elements x, y that are free such that $\mu_x = \mu$ and $\mu_y = \nu$?

Next time, we will see how to construct such an NCPs. In fact, we'll see we can do it not only with two distributions, but in fact with an infinite sequence of them.