Lecture 6: April 8, 2011

Let's compute an example of a (non-commutative) distribution.

Eg. Let G be a group, and suppose $u \in G$ has ∞ order. In the NCPS $(\mathbb{C}G, \varphi_G)$, we then have for $n \in \mathbb{Z}$

$$u^n = 1 \iff n = 0, \text{ so } \varphi_G(u^n) = \delta_{n0}.$$

Now, in $\mathbb{C}G$, $u^* = u^{-1}$; so the self-adjoint element $x = u + u^*$ is $x = u + u^{-1}$. Let's compute its moments. This is easy, since u, u^{-1} commute:

$$(u+u^{-1})^n = \sum_{k=0}^{n} \binom{n}{k} u^k (u^{-1})^{n-k} = \sum_{k=0}^{n} u^{2k-n}.$$

Thus $\varphi(x^n) = \sum_{k=0}^{n} \varphi(u^{2k-n}) = \sum_{k=0}^{n} \binom{n}{k} \delta_{2k-n,0} = \begin{cases} \binom{n}{n/2}, & n \text{ even} \\ 0, & n \text{ odd} \end{cases}$

So, we've computed the moments. How do we get from here to μ_x?

<u>General approach</u>: look at the (ordinary) generating function

$$M(z) = \sum_{n=0}^{\infty} \varphi(x^n) z^n = \sum_{m=0}^{\infty} \binom{2m}{m} z^{2m}.$$

[We might just think of this as a formal power series; later, we'll discuss convergence sets.]

We need a tool to relate this "moment series" to the measure. At least formally,

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

So, we need 2 things:

(1) to be able to sum the series $M(z)$ and work with this analytic function (a la generatingfunctionology) $\leftarrow$ combinatorics $\Leftarrow$ A complex analytic transform of the measure μ_x; in spirit, like the Fourier transform.

(2) Come up with an inversion formula for the transform $\leftarrow$ Complex analysis.

We'll get to both. Meanwhile, we can hit this example directly – if we know "μ_u"...

Distributions of non-self-adjoint elements

If $x \in (\mathcal{A}, \varphi)$ is self-adjoint, we now know (in principle) how to calculate its distribution. But what if x is not self-adjoint?

First, let's see what happens in the classical context.

Eg. If $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ is $\mathbb{C}$-valued, its distribution μ_X is a prob. measure on $\mathbb{C}$. It is still true that

$$\mathbb{E}(f(X)) = \int_{\mathbb{C}} f(z)\, \mu_X(|dz|) \quad \forall f \in C(\mathbb{C})$$

but single-variable polynomials are no longer dense in the complex space. Looking back to the Weierstrass approximation theorem, we will need polynomials in z and $\bar{z}$.

Prop: $\mathbb{E}(X^n \bar{X}^m) = \int_{\mathbb{C}} z^n \bar{z}^m \mu_X(|dz|)$
and μ_X is the unique compactly-supported prob. measure on $\mathbb{C}$ with these (joint) moments.

So, this suggests that, in the non-self-adjoint case, we need to look at joint moment in a and a^*.

Caution: Unless a happens to commute with a^*, monomials of the form $a^n a^{*m}$ only tell a small piece of the story.

Def: An element $a \in \mathcal{A}$ is called normal if $aa^* = a^*a$.

In this case (in the operator algebra context), the spectral theorem can be used to prove that

Thm: For a normal element a in an operator algebra NCPS $(\mathcal{A}, \varphi)$, $\exists!$ compactly-supported prob. measure μ_a on $\mathbb{C}$ s.t.

$$\varphi(a^n a^{*m}) = \int_{\mathbb{C}} z^n \bar{z}^m \mu_a(|dz|), \quad n, m \in \mathbb{N}.$$

When a is not normal there may be no such measure; even if there is, it may not say much about a.

Def: If $a \in (\mathcal{A}, \varphi)$, the <u>$*$-distribution</u> μ_{a,a^*} of a is the collection of all joint moments in a and a^* (e.g. $\varphi(a^7 a^* a^2 a^{*4} a^3)$).
We can think of μ_{a,a^*} as a linear functional on $\mathbb{C}[X,Y]$ – the space of non-comm. polynomials in 2 variables:

$$\mu_{a,a^*}(P(X,Y)) = \varphi(P(a,a^*)).$$

This puts it on the same kind of footing as a measure on $\mathbb{C}$, which can be thought of (a la the preceding discussion) as a linear functional on the space $\mathbb{C}\langle X,Y\rangle$ of commutative polynomials in 2 variables.
If a is normal, then μ_{a,a^*} "collapses" to a functional on $\mathbb{C}\langle X,Y\rangle$, and we recover the usual measure on $\mathbb{C}$. Let's look at the $\mathbb{C}G$ example again.

Eg. In $(\mathbb{C}G, \varphi_G)$ with u of ∞ order, since $uu^* = u^*u = 1$, the $*$-distr. is just:

$$\varphi(u^{n_1} u^{*m_1} \cdots u^{n_k} u^{*m_k}) = \varphi(u^{n_1 + \cdots + n_k - m_1 - \cdots - m_k}) = \delta_{n_1 + \cdots - m_k, 0}.$$

We can identify this with a measure on $\mathbb{C}$: its moments are

$$\varphi(u^n u^{*m}) = \varphi(u^{n-m}) = \delta_{n-m,0}.$$

<u>Ansatz</u>: these are exactly the moments of the Haar measure μ_{S^1} on S^1:

$$\int_{S^1} z^n \bar{z}^m \mu_{S^1}(|dz|) = \int_0^{2\pi} (e^{it})^n (e^{-it})^m \frac{1}{2\pi} dt$$
$$= \int_0^{2\pi} e^{i(n-m)t} \frac{dt}{2\pi} = \delta_{n-m,0}$$

So $\mu_{u,u^*} = \mu_u = \mu_{S^1}$. For this reason, we call such an element u in $\mathbb{C}G$ a <u>Haar unitary</u>. (In general, any el't $u \in (\mathcal{A}, \varphi)$ with $uu^* = u^*u = 1$ and $\varphi(u^n) = \delta_{n,0}$ for $n \in \mathbb{Z}$ is called a Haar unitary, because its distribution is Haar measure on S^1, and it is unitary.)

In this case, we can easily use this $*$-distribution (of u) to compute the distribution of $u + u^*$, as the next calculation shows.

Eg. Let $u \in (\mathcal{A}, \varphi)$ be a Haar unitary. Set $x = u + u^*$. For $n \in \mathbb{N}$,

$$\varphi(x^n) = \varphi((u+u^*)^n) = \sum_{k=0}^{n} \binom{n}{k} \varphi(u^k u^{*(n-k)})$$

But these #s are recorded in μ_{u,u^*}, which (in this case) is described by a measure on $\mathbb{C}$.

As established, $\varphi(u^k u^{*(n-k)}) = \int_{\mathbb{C}} z^k \bar{z}^{n-k} \mu_{u,u^*}(|dz|)$

$$= \int_0^{2\pi} (e^{it})^k (e^{-it})^{n-k} \frac{dt}{2\pi}$$

Thus

$$\varphi(x^n) = \sum_{k=0}^{n} \binom{n}{k} \int_0^{2\pi} (e^{it})^k (e^{-it})^{n-k} \frac{dt}{2\pi} = \int_0^{2\pi} (e^{it} + e^{-it})^n \frac{dt}{2\pi}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} (2\cos t)^n \, dt$$

by symmetry of cos

$$= \frac{1}{\pi} \int_0^{\pi} (2\cos t)^n \, dt$$

Now, we make the following change of variables: $r = 2\cos t$. On $(0,\pi)$ this is invertible, $t = \cos^{-1}(r/2)$, and so

$$dt = \frac{-1}{\sqrt{1-(r/2)^2}} \cdot \frac{1}{2} dr = \frac{-dr}{\sqrt{4-r^2}} \quad \text{on } -2 \le r \le 2$$

The minus sign is because of the orientation reversal $\to$ $0 \mapsto 2$, $\pi \mapsto -2$; thus, making the substitution we get

$$\varphi(x^n) = \frac{1}{\pi} \int_2^{-2} r^n \frac{-dr}{\sqrt{4-r^2}} = \frac{1}{\pi} \int_{-2}^{2} r^n \frac{dr}{\sqrt{4-r^2}}.$$

$\Uparrow$

So, we have shown that $x = u + u^*$ has a density: $\mu_x(dr) = \frac{1}{\sqrt{4-r^2}} \mathbb{1}_{[-2,2]}(r) dr$.

"arcsine law"

Remark: Of course, the $*$-distribution μ_{a,a^*} of an element $a \in \mathcal{A}$ determines the law μ_{a+a^*}. The above example shows how this (easily) works in the case that a is normal. In general, however, it is a lot trickier.

<u>Exercise</u>: Let G be a group, and $u \in G$ have finite order $p \in \mathbb{N}$: $u^p = 1$, $u^n \neq 1$ for $1 < n < p$. In $(\mathbb{C}G, \varphi_G)$, compute

(a) The $*$-distribution of u.
(b) The distribution of $u + u^*$.
(c) The distribution of $i(u - u^*)$.

(Such an element u is called a p-Haar unitary.)

A case study of a non-normal element

Consider the Hilbert space $\ell^2(\mathbb{N})$. It has, as a Hilbert-space basis, the "standard basis vectors"

$$e_n = (0, 0, \ldots, 0, \underset{n^{th}\text{ position}}{1}, 0, \ldots), \quad n = 0, 1, 2, \ldots$$

We can use them to define a bounded operator S on $\ell^2(\mathbb{N})$, called the <u>unilateral shift</u>:

$$S(e_n) = e_{n+1}.$$

This operator is bounded – in fact, it is an isometry: for $h \in \ell^2(\mathbb{N})$, expand $h = \sum h_n e_n$; then

$$\|Sh\|_2^2 = \|\sum_n h_n Se_n\|_2^2 = \|\sum_n h_n \underset{\text{orthonormal}}{e_{n+1}}\|_2^2 = \sum_n |h_n|^2 = \|h\|_2^2.$$

Now, in finite dimensions, isometry $\Rightarrow$ unitary, but not here: S is not invertible, since e_0 is not in its range.

Let's compute the adjoint S^* (which exists, since S is bounded).

$$\langle S^* e_n, e_m \rangle \equiv \langle e_n, Se_m \rangle = \langle e_n, e_{m+1} \rangle = \delta_{n,m+1}.$$

Now, if $n \geq 1$, we can write this as $\langle e_n, e_{m+1} \rangle = \delta_{n-1,m} = \langle e_{n-1}, e_m \rangle$. This is true for all m, so for $n \geq 1$ $S^* e_n = e_{n-1}$. But

$$\langle S^* e_0, e_m \rangle = \langle e_0, Se_m \rangle = \langle e_0, e_{m+1} \rangle = 0 \quad \text{b/c } m+1 \geq 1.$$

Thus, $S^*e_0 = 0$. So S^* is a lowering operator, but it kills e_0.

Now, we computed $\|Sh\|_2 = \|h\|_2$; in fact, by polarization, this gives $\langle Sh, Sk\rangle = \langle h,k\rangle$, and so

$$\langle h,k\rangle = \langle Sh,Sk\rangle = \langle S^*Sh, k\rangle \quad \forall h,k$$

$$\implies S^*S = I.$$

But: $SS^*e_0 = S(0) = 0 \neq e_0 \to$ so $SS^* \neq I$. In particular,

S <u>is not normal</u>.

Now, let us consider the algebra $\mathfrak{A} = \langle S\rangle_*$: the smallest $*$-subalgebra of $\mathcal{B}(\ell^2(\mathbb{N}))$ that contains S. Here is a (somewhat) surprising fact.

<u>Prop</u>: $\mathfrak{A} = \mathrm{span}_{\mathbb{C}}\{S^n S^{*m} : n,m \in \mathbb{N}\}$. Moreover, all these monomials are linearly independent.

<u>Pf</u>. First, let's see how to multiply such monomials: $S^{n_1}S^{*m_1} \cdot S^{n_2}S^{*m_2}$.

There are two cases:

- if $m_1 \le n_2$, then $S^{*m_1}S^{n_2} = S^{*m_1}S^{m_1}S^{n_2-m_1} = S^{n_2-m_1}$,
  so $S^{n_1}S^{*m_1} \cdot S^{n_2}S^{*m_2} = S^{n_1+n_2-m_1}S^{*m_2}$
- if $m_1 > n_2$, then $S^{*m_1}S^{n_2} = S^{*m_1-n_2}S^{*n_2}S^{n_2} = S^{*m_1-n_2}$,
  so $S^{n_1}S^{*m_1} \cdot S^{n_2}S^{*m_2} = S^{n_1}S^{*m_1-n_1+m_2}$

In either case, the product is another monomial of the form S^nS^{*m}. This shows the above span is actually an algebra. Moreover,

$$(S^nS^{*m})^* = S^mS^{*n}$$

and so this span is a $*$-subalgebra of $\mathcal{B}(\ell^2(\mathbb{N}))$, which contains S. Since <u>any</u> $*$-algebra containing S must contain all such monomials, it follows that this is the minimal $*$-subalgebra, as claimed.

That they are linearly independent is left as an <u>Exercise</u>.
[Hint: use the fact that $\langle S^nS^{*m}e_m, e_n\rangle = 1$.]

So, even though S is not normal, all words in S and S^* reduce to monomials of the form $S^n S^{*m}$. But the way they reduce is very different from the situation for normal operators.

One way to see this is in the framework of a NCPS. Consider the vector state

$$\varphi_0(A) = \langle Ae_0, e_0 \rangle, \quad A \in \mathfrak{A}.$$

As we know, all vector states are positive. In fact, this one is faithful. The reason is simple:

$$\varphi_0(S^n S^{*m}) = \langle S^n S^{*m} e_0, e_0 \rangle = \langle S^{*m} e_0, S^{*n} e_0 \rangle = \begin{cases} 0 & \text{if } n \text{ or } m \neq 0 \\ 1 & \text{if } n = m = 0. \end{cases}$$

<u>Exercise</u>: use this to show φ_0 is faithful on $\mathfrak{A}$.

<u>Remark</u>: if S were normal, the above joint moments of S would prompt us to find a measure on $\mathbb{C}$ with those moments. In fact, the unique such measure is δ_0 — so S would be distributed like the 0 element. This is unconvincing — and shows tangibly why non-normal elements don't have measures as distributions.

To better understand what's going on, let us forget for the moment that all monomials in $\mathfrak{A}$ reduce to $S^n S^{*m}$. Let us look at an arbitrary word

$$S^{\varepsilon_1} S^{\varepsilon_2} \cdots S^{\varepsilon_n}, \quad \varepsilon_1, \dots, \varepsilon_n \in \{1, *\}$$

and compute φ_0 of it. This will really show how different S is from normal.

$$\varphi_0(S^{\varepsilon_1} \cdots S^{\varepsilon_n}) = \langle \underbrace{S^{\varepsilon_1} \cdots S^{\varepsilon_n} e_0}, e_0 \rangle$$

what vector is this? By induction, easy to see that it is either e_k for some $k \leq n$ or is 0.

When is the state 0? Well, S is a raising operator and S^* is a lowering operator, with the proviso that dropping below e_0 means killing the vector — it is just 0 from then on.

Ie. identifying $\varepsilon = *$ with $\varepsilon = -1$, we get the following.

$S^{\varepsilon_1} \cdots S^{\varepsilon_n} e_0 \quad \varepsilon_n \varepsilon_{n-1} \cdots \varepsilon_2 \varepsilon_1 \in \{1,-1\}^n$

In this string, we keep a running total
$\varepsilon_n,\ \varepsilon_n + \varepsilon_{n-1},\ \varepsilon_n + \varepsilon_{n-1} + \varepsilon_{n-2}, \ldots$
If this running total ever drops below 0,
then $S^{\varepsilon_1} \cdots S^{\varepsilon_n} e_0 = 0$.

Eg. $S^*S^*SSS^*SS e_0 = e_{1+1-1+1+1-1-1} = e_1$.
$S^*SSSS^*S^*S e_0 = 0$.

Now, in the end, we take $\langle S^{\varepsilon_1} \cdots S^{\varepsilon_n} e_0, e_0 \rangle$

I.e. even when $S^{\varepsilon_1} \cdots S^{\varepsilon_n} e_0 \neq 0$, the state is 0 unless it equals e_0.

Conclusion:

$$\varphi_0(S^{\varepsilon_1} \cdots S^{\varepsilon_n}) = \begin{cases} 1, & \text{if the reversed sequence } \varepsilon_n \cdots \varepsilon_1 \text{ has all } \geq 0 \\ & \text{running totals and } \varepsilon_n + \cdots + \varepsilon_n = 0. \\ 0, & \text{otherwise.} \end{cases}$$

The set of such binary strings is a well studied combinatorial object. They can be viewed as lattice paths, called <u>Dyck paths</u>. We will take a close look at them next day.