Polynomials in Asymptotically Free Random Matrices

Roland Speicher
Universität des Saarlandes
Saarbrücken, Germany

joint work with Serban Belinschi, Tobias Mai, Piotr Sniady
We are interested in the limiting eigenvalue distribution of an $N \times N$ random matrix for $N \to \infty$.

Typical phenomena for basic random matrix ensembles:

- almost sure convergence to a deterministic limit eigenvalue distribution
- this limit distribution can be effectively calculated
The Cauchy (or Stieltjes) Transform

For any probability measure $\mu$ on $\mathbb{R}$ we define its Cauchy transform

$$G(z) := \int_{\mathbb{R}} \frac{1}{z - t} d\mu(t)$$

This is an analytic function $G : \mathbb{C}^+ \rightarrow \mathbb{C}^-$ and we can recover $\mu$ from $G$ by Stieltjes inversion formula

$$d\mu(t) = -\frac{1}{\pi} \lim_{\varepsilon \to 0} \Im G(t + i\varepsilon) dt$$
Wigner random matrix
and
Wigner's semicircle
\[ G(z) = \frac{z - \sqrt{z^2 - 4}}{2} \]

Marchenko-Pastur distribution
\[ G(z) = \frac{z + 1 - \lambda - \sqrt{(z - (1 + \lambda))^2 - 4 \lambda}}{2z} \]
We are now interested in the limiting eigenvalue distribution of general (selfadjoint) polynomials $p(X_1, \ldots, X_k)$ of several independent $N \times N$ random matrices $X_1, \ldots, X_k$.

Typical phenomena:

- almost sure convergence to a deterministic limit eigenvalue distribution
- this limit distribution can be effectively calculated only in very simple situations
for $X$ Wigner, $Y$ Wishart

\[ p(X, Y) = X + Y \]
\[ G(z) = G_{\text{Wishart}}(z - G(z)) \]

\[ p(X, Y) = XY + YX + X^2 \]
Existing Results for Calculations of the Limit Eigenvalue Distribution

• Marchenko, Pastur 1967: general Wishart matrices $ADA^*$

• Pastur 1972: deterministic $+$ Wigner (deformed semicircle)

• Speicher, Nica 1998; Vasilchuk 2003: commutator or anti-commutator: $X_1X_2 \pm X_2X_1$

• more general models in wireless communications (Tulino, Verdu 2004; Couillet, Debbah, Silverstein 2011):

$$RADA^*R^* \quad \text{or} \quad \sum_i R_iA_iD_iA_i^*R_i^*$$
Asymptotic Freeness of Random Matrices

Basic result of Voiculescu (1991): Large classes of independent random matrices (like Wigner or Wishart matrices) become asymptotically freely independent.

Conclusion: Calculating the asymptotic eigenvalue distribution of polynomials in such matrices is the same as calculating the distribution of polynomials in free variables.
We want to understand distribution of polynomials in free variables.

What we understand quite well is:

*sums of free selfadjoint variables*

So we should reduce:

\[
\text{arbitrary polynomial} \rightarrow \text{sums of selfadjoint variables}
\]

This can be done on the expense of going over to operator-valued frame.
Let $\mathcal{B} \subset \mathcal{A}$. A linear map

$$E : \mathcal{A} \rightarrow \mathcal{B}$$

is a conditional expectation if

$$E[b] = b \quad \forall b \in \mathcal{B}$$

and

$$E[b_1 ab_2] = b_1 E[a] b_2 \quad \forall a \in \mathcal{A}, \quad \forall b_1, b_2 \in \mathcal{B}$$

An operator-valued probability space consists of $\mathcal{B} \subset \mathcal{A}$ and a conditional expectation $E : \mathcal{A} \rightarrow \mathcal{B}$
Consider an operator-valued probability space $E : A \to B$.

Random variables $x_i \in A \ (i \in I)$ are free with respect to $E$ (or free with amalgamation over $B$) if

$$E[a_1 \cdots a_n] = 0$$

whenever $a_i \in B\langle x_{j(i)} \rangle$ are polynomials in some $x_{j(i)}$ with coefficients from $B$ and

$$E[a_i] = 0 \quad \forall i \quad \text{and} \quad j(1) \neq j(2) \neq \cdots \neq j(n).$$
Consider an operator-valued probability space \( E : \mathcal{A} \to \mathcal{B} \).

For a random variable \( x \in \mathcal{A} \), we define the operator-valued Cauchy transform:

\[
G(b) := E[(b - x)^{-1}] \quad (b \in \mathcal{B}).
\]

For \( x = x^* \), this is well-defined and a nice analytic map on the operator-valued upper halfplane:

\[
\mathbb{H}^+(\mathcal{B}) := \{ b \in \mathcal{B} \mid (b - b^*)/(2i) > 0 \}.
\]
**Theorem (Belinschi, Mai, Speicher 2013):** Let \( x \) and \( y \) be selfadjoint operator-valued random variables free over \( \mathcal{B} \). Then there exists a Fréchet analytic map \( \omega : \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B}) \) so that
\[
G_{x+y}(b) = G_x(\omega(b)) \quad \text{for all} \quad b \in \mathbb{H}^+(\mathcal{B}).
\]

Moreover, if \( b \in \mathbb{H}^+(\mathcal{B}) \), then \( \omega(b) \) is the unique fixed point of the map
\[
f_b : \mathbb{H}^+(\mathcal{B}) \to \mathbb{H}^+(\mathcal{B}), \quad f_b(w) = h_y(h_x(w) + b) + b,
\]
and
\[
\omega(b) = \lim_{n \to \infty} f_b^{\circ n}(w) \quad \text{for any} \quad w \in \mathbb{H}^+(\mathcal{B}).
\]

where
\[
\mathbb{H}^+(\mathcal{B}) := \{ b \in \mathcal{B} \mid (b - b^*)/(2i) > 0 \}, \quad h(b) := \frac{1}{G(b)} - b.
\]
The Linearization Philosophy:

In order to understand polynomials $p$ in non-commuting variables, it suffices to understand matrices $\hat{p}$ of linear polynomials in those variables.

- Voiculescu 1987: motivation
- Haagerup, Thorbjørnsen 2005: largest eigenvalue
- Anderson 2012: the selfadjoint version (based on Schur complement)
Theorem (Anderson 2012): One has

- for each $p$ there exists a linearization $\hat{p}$
  (with an explicit algorithm for finding those)

- if $p$ is selfadjoint, then this $\hat{p}$ is also selfadjoint

Note: $\hat{p}$ is the sum of operator-valued free variables!

Conclusion: Combination of linearization and operator-valued subordination allows to deal with case of selfadjoint polynomials.
**Example:** \( p(x, y) = xy + yx + x^2 \)

\( p \) has linearization

\[
\hat{p} = \begin{pmatrix}
0 & x & y + \frac{x}{2} \\
x & 0 & -1 \\
y + \frac{x}{2} & -1 & 0
\end{pmatrix}
\]

and

\[
G_{\hat{p}}(b) = \text{id} \otimes \varphi((b - \hat{p})^{-1}) = \begin{pmatrix}
\varphi((z - p)^{-1}) & \cdots \\
\cdots & \cdots
\end{pmatrix}
\]

\[
b = \begin{pmatrix}
\hat{z} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]
Example: \( p(x, y) = xy +yx + x^2 \)

\[
\hat{p} = \begin{pmatrix}
0 & x & \frac{x}{2} \\
x & 0 & -1 \\
\frac{x}{2} & -1 & 0
\end{pmatrix} + \begin{pmatrix}
0 & 0 & y \\
0 & 0 & 0 \\
y & 0 & 0
\end{pmatrix} = \hat{X} + \hat{Y}
\]

is now the sum of two selfadjoint variables which are free over \( M_3(\mathbb{C}) \), so we can use our subordination result

\[
\begin{pmatrix}
\varphi((z - p)^{-1}) & \cdots \\
\vdots & \ddots
\end{pmatrix} = G_{\hat{p}}(b) = G_{\hat{X} + \hat{Y}}(b) = G_{\hat{X}}(\omega(b))
\]
\[ P(X, Y) = XY + YX + X^2 \]
for independent \( X, Y \); \( X \) is Wigner and \( Y \) is Wishart

\[ p(x, y) = xy + yx + x^2 \]
for free \( x, y \); \( x \) is semicircular and \( y \) is Marchenko-Pastur
What about non-selfadjoint polynomials?

For a measure on $\mathbb{C}$ its Cauchy transform

$$G_\mu(\lambda) = \int_\mathbb{C} \frac{1}{\lambda - z} d\mu(z)$$

is well-defined everywhere outside a set of $\mathbb{R}^2$-Lebesgue measure zero, however, it is analytic only outside the support of $\mu$.

The measure $\mu$ can be extracted from its Cauchy transform by the formula (understood in distributional sense)

$$\mu = \frac{1}{\pi} \frac{\partial}{\partial \bar{\lambda}} G_\mu(\lambda),$$
Better approach by regularization:

\[ G_{\epsilon,\mu}(\lambda) = \int_{\mathbb{C}} \frac{\bar{\lambda} - \bar{z}}{\epsilon^2 + |\lambda - z|^2} d\mu(z) \]

is well–defined for every \( \lambda \in \mathbb{C} \). By sub-harmonicity arguments

\[ \mu_\epsilon = \frac{1}{\pi} \frac{\partial}{\partial \bar{\lambda}} G_{\epsilon,\mu}(\lambda) \]

is a positive measure on the complex plane.

One has: \( \lim_{\epsilon \to 0} \mu_\epsilon = \mu \) weak convergence.
This can be copied for general (not necessarily normal) operators $x$ in a tracial non-commutative probability space $(A, \varphi)$.

Put

$$G_{\epsilon, x}(\lambda) := \varphi \left( (\lambda - x)^* (\lambda - x)^* + \epsilon^2 \right)^{-1}$$

Then

$$\mu_{\epsilon, x} = \frac{1}{\pi} \frac{\partial}{\partial \bar{\lambda}} G_{\epsilon, \mu}(\lambda)$$

is a positive measure on the complex plane, which converges weakly for $\epsilon \to 0$,

$$\mu_x := \lim_{\epsilon \to 0} \mu_{\epsilon, x} \quad \text{Brown measure of } x$$

(L. Brown 1986; Haagerup, Larsen 2000)
Hermitization Method
(Janik, Nowak, Papp, Zahed 1997; Feinberg, Zee 1997)

For given \( x \) we need to calculate

\[ G_{\epsilon, x}(\lambda) = \varphi \left( (\lambda - x)^* \left( (\lambda - x)(\lambda - x)^* + \epsilon^2 \right)^{-1} \right) \]

Let

\[
X = \begin{pmatrix} 0 & x \\ x^* & 0 \end{pmatrix} \in M_2(\mathcal{A}); \quad \text{note: } X = X^* 
\]

Consider \( X \) in the \( M_2(\mathbb{C}) \)-valued probability space with respect to 
\( E = \text{id} \otimes \varphi : M_2(\mathcal{A}) \to M_2(\mathbb{C}) \) given by

\[
E \left[ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right] = \begin{pmatrix} \varphi(a_{11}) & \varphi(a_{12}) \\ \varphi(a_{21}) & \varphi(a_{22}) \end{pmatrix}.
\]
For the argument
\[
\Lambda_\epsilon = \begin{pmatrix} i\epsilon & \lambda \\ \overline{\lambda} & i\epsilon \end{pmatrix} \in M_2(\mathbb{C}) \quad \text{and} \quad X = \begin{pmatrix} 0 & x \\ x^* & 0 \end{pmatrix}
\]
consider now the $M_2(\mathbb{C})$-valued Cauchy transform of $X$

\[
G_X(\Lambda_\epsilon) = E[(\Lambda_\epsilon - X)^{-1}] = \begin{pmatrix} g_{\epsilon,\lambda,11} & g_{\epsilon,\lambda,12} \\ g_{\epsilon,\lambda,21} & g_{\epsilon,\lambda,22} \end{pmatrix}.
\]

One can easily check that

\[
(\Lambda_\epsilon - X)^{-1} = \begin{pmatrix} -i\epsilon((\lambda - x)(\lambda - x)^* + \epsilon^2)^{-1} & (\lambda - x)((\lambda - x)^*(\lambda - x) + \epsilon^2)^{-1} \\ (\lambda - x)^*((\lambda - x)(\lambda - x)^* + \epsilon^2)^{-1} & -i\epsilon((\lambda - x)^*(\lambda - x) + \epsilon^2)^{-1} \end{pmatrix}
\]

thus

\[
g_{\epsilon,\lambda,12} = G_{\epsilon,x}(\lambda).
\]
So for a general polynomial we should

1. hermitize

2. linearise

3. subordinate

But: do (1) and (2) fit together???

We have now to linearize a polynomial in matrices!
Consider $p = xy$ with $x = x^*$, $y = y^*$.

For this we have to calculate the operator-valued Cauchy transform of

$$ P = \begin{pmatrix} 0 & xy \\ yx & 0 \end{pmatrix} $$

Linearization means we should split this in sums of matrices in $x$ and matrices in $y$.

Write

$$ P = \begin{pmatrix} 0 & xy \\ yx & 0 \end{pmatrix} = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & y \\ y & 0 \end{pmatrix} \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} = XYX $$
\[ P = XYX \] is now a selfadjoint polynomial in the selfadjoint variables

\[ X = \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad Y = \begin{pmatrix} 0 & y \\ y & 0 \end{pmatrix} \]

\( XYX \) has linearization

\[
\begin{pmatrix}
0 & 0 & X \\
0 & Y & -1 \\
X & -1 & 0
\end{pmatrix}
\]
thus

\[
P = \begin{pmatrix} 0 & xy \\ yx & 0 \end{pmatrix}
\]

has linearization

\[
\begin{pmatrix}
0 & 0 & 0 & 0 & x & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & y & 0 & 0 & -1 \\
x & 0 & -1 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0
\end{pmatrix}
= \begin{pmatrix}
0 & 0 & 0 & 0 & x & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & -1 \\
x & 0 & -1 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0
\end{pmatrix} + \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\]

and we can now calculate the operator-valued Cauchy transform of this via subordination.
Does eigenvalue distribution of polynomial in independent random matrices converge to Brown measure of corresponding polynomial in free variables?

Conjecture: Consider \( m \) independent selfadjoint Gaussian (or, more general, Wigner) random matrices \( X_N^{(1)}, \ldots, X_N^{(m)} \) and put

\[
A_N := p(X_N^{(1)}, \ldots, X_N^{(m)}), \quad x := p(s_1, \ldots, s_m).
\]

We conjecture that the eigenvalue distribution \( \mu_{A_N} \) of the random matrices \( A_N \) converge to the Brown measure \( \mu_x \) of the limit operator \( x \).

(compare: single ring theorem of Guionnet and Zeitouni 2011)
Brown measure of $xyz - 2yzx + zxy$ with $x, y, z$ free semicircles
Brown measure of $x + iy$ with $x, y$ free Poissons
Brown measure of $x_1 x_2 + x_2 x_3 + x_3 x_4 + x_4 x_1$