Support properties of spectra of polynomials in Wigner matrices

Greg W. Anderson
University of Minnesota
www.math.umn.edu/~gwanders
gwanders@umn.edu

Advances in Random Matrix Theory

Institute for Mathematics and its Applications
University of Minnesota
June 18–29, 2012
The spectrum and the support of the law of a noncommutative random variable

Let \((A, \phi)\) be a noncommutative probability space.

Let \(\phi\) be \textit{faithful}, i.e., suppose that for all \(A \in A\) if \(A \geq 0\) and \(A \neq 0\), then \(\phi(A) > 0\).

**Proposition (Support equals spectrum in presence of faithfulness)**

\textit{Notation and assumptions as are above. For every noncommutative random variable \(A \in \text{Mat}_n(A)_{\text{sa}}\) the support of the law of \(A\) and the spectrum of \(A\) are equal.}

We present a proof in the next couple of frames.
Proof of “support equals spectrum in presence of faithfulness”

Lemma

The linear functional \( \phi_n = (A \mapsto \frac{1}{n} \sum_{i=1}^{n} \phi(A[i, i])) \) on \( \text{Mat}_n(A) \) is a faithful state.

Proof Fix \( A \in \text{Mat}_n(A) \) such that \( A \geq 0 \) and \( A \neq 0 \).

Then

\[
\phi_n(A) = \frac{1}{n} \sum_{i=1}^{n} \phi(A^{1/2}[i, j]A^{1/2}[j, i]) \geq 0
\]

and moreover at least one term on the right does not vanish by faithfulness of \( \phi \).
Proof of “support equals spectrum in presence of faithfulness” (continued)

Lemma

Fix $A \in \mathcal{A}_{sa}$ and let $A_0 \subset A$ be the $C^*$-subalgebra generated by $A$. Then every $B \in A_0$ is of the form $B = f(A)$ for some continuous function $f : \text{Spec}(A) \to \mathbb{C}$ and moreover $\|B\| = \sup_{x \in \text{Spec}(A)} |f(x)|$. (In other words, the functional calculus is an isomorphism from the $C^*$-algebra of continuous $\mathbb{C}$-valued functions on $\text{Spec}(A)$ to $A_0$.)

Proof This is a basic consequence of the theory of the Gelfand transform.
The preceding two lemmas granted, we may assume that $A$ is the space of continuous $\mathbb{C}$-valued functions defined on on compact subset $K \subset \mathbb{R}$ and that $A$ is the identity map $K \to K$. Then $\phi$ is represented by a probability measure on $\text{Spec}(A)$. Faithfulness of $\phi$ implies via Urysohn’s lemma that $\text{supp } \mu = \text{Spec}(A)$. \qed
Theorem (Voiculescu)

For each $f \in \text{Mat}_n(\mathbb{C}\langle X_1, \ldots, X_m \rangle)_{\text{sa}}$ the empirical distribution $\mu_f^{(N)}$ converges momentwise to $\mu_f$.

Equivalently (in the more or less the format in which Voiculescu’s theorem was originally stated), for any finite sequence $\ell_1, \ldots, \ell_k \in \{1, \ldots, m\}$,

$$
\lim_{N \to \infty} \frac{1}{N} \mathbb{E} \text{tr} \left( \frac{\Xi^{(N)}_{\ell_1}}{\sqrt{N}} \cdots \frac{\Xi^{(N)}_{\ell_k}}{\sqrt{N}} \right) = (1_{\mathcal{H}}, \Xi_{\ell_1} \cdots \Xi_{i_k} 1_{\mathcal{H}}).
$$

This is the foundational example for free probability theory. It is a far-reaching generalization of Wigner’s semicircle law.
More probabilistic versions of Voiculescu’s result can be proved without a great deal of trouble. In particular, in the setup for HST one can prove the following statement.

**Theorem (Amplification of Voiculescu’s theorem)**

\[ \mu_f^{(N)} \text{ converges weakly to } \mu_f, \text{ almost surely.} \]

The paper [Male 2010] contains a proof of this, and of a more general result. The paper [Meckes-Szarek 2011] gives an interesting treatment of concentration phenomena related to this result.
Proposition (Folklore about faithfulness)

The restriction of the state

\[(A \mapsto (1_{\mathcal{H}}, A1_{\mathcal{H}})) : B(\mathcal{H}) \to \mathbb{C}\]

to the \(C^*\)-subalgebra generated by the free semicircular family \(\{\Xi_i\}_{i=1}^m\) is faithful and thus for every \(f \in \text{Mat}_n(\mathbb{C}\langle X_{1\ldots m}\rangle)_{\text{sa}}\) we have \(\text{supp } \mu_f = \text{Spec}(f(\Xi))\).

I thank K. Dykema for explaining this important point to me. It is quite elementary but hard to pin down in the literature.

This is advantageous because the spectrum of \(f(\Xi)\) is a more tractable object to study than the support of \(\mu_f\).

We give the following proof as an excuse simply to better acquaint the audience with the structure of Boltzmann-Fock space.
Proof

Put

\[ w(\ell_1 \cdots \ell_k) = \theta_{\ell_1} \cdots \theta_{\ell_k} v(0) \]

for all finite sequences \( \ell_1 \cdots \ell_k \) in \( \{1, \ldots, m\} \), including the empty sequence.

Then, as previously noted, \( \{w(\ell_1 \cdots \ell_k)\} \) is just a relabeling of the canonical orthonormal basis \( \{v(i)\} \).

For example,

\[ v(0) = w(\emptyset), \quad w(1) = v(1), \ldots, \quad w(m) = v(m), \quad w(11) = v(m+1), \ldots, \]
Thus we have

\[ \theta_\ell \omega(\ell_1 \cdots \ell_k) = \omega(\ell \ell_1 \cdots \ell_k), \quad \theta^*_\ell \omega(\ell_1 \cdots \ell_k) = \mathbf{1}_{\ell_1=\ell} \omega(\ell_2 \cdots \ell_k). \]

Recall that

\[ \Xi_\ell = \theta_\ell + \theta^*_\ell. \]

Analogously, we now define \( \hat{\theta}_\ell, \hat{\Xi}_\ell \in B(\mathcal{H}) \) “on the right side” by

\[ \hat{\theta}_\ell \omega(\ell_1 \cdots \ell_k) = \omega(\ell_1 \cdots \ell_k \ell), \quad \hat{\theta}^*_\ell \omega(\ell_1 \cdots \ell_k) = \mathbf{1}_{\ell=\ell_k} \omega(\ell_1 \cdots \ell_{k-1}) \]

and we put

\[ \hat{\Xi}_\ell = \hat{\theta}_\ell + \hat{\theta}^*_\ell. \]
We have commutation relations

\[ \theta_i \hat{\theta}_j = \hat{\theta}_j \theta_i, \quad \theta_i^{*} \hat{\theta}_j^{*} = \hat{\theta}_j^{*} \theta_i^{*}, \]

\[ \theta_i \hat{\theta}_j^{*} - \hat{\theta}_j^{*} \theta_i = -\delta_{ij} p_{\mathcal{H}}, \quad \theta_i^{*} \hat{\theta}_j - \hat{\theta}_j \theta_i^{*} = \delta_{ij} p_{\mathcal{H}}, \]

where \( p_{\mathcal{H}} \) denotes orthogonal projection to the groundstate \( 1_{\mathcal{H}} = \nu(0) \)

and thus

\[ \Xi_i \Xi_j = \Xi_j \Xi_i. \]
Proof (continued)

Let $\mathcal{A} \subset B(\mathcal{H})$ (resp., $\hat{\mathcal{A}} \subset B(\mathcal{H})$) be the $C^*$-subalgebra generated by the $\Xi_i$ (resp., by the $\hat{\Xi}_i$).

Clearly $\mathcal{A}$ and $\hat{\mathcal{A}}$ commute with each other.

Furthermore, it is not difficult to verify that $\hat{\mathcal{A}}1_\mathcal{H}$ is dense in $\mathcal{H}$. 
Now let $A \in A$ satisfy $A \geq 0$ and $A \neq 0$.

Clearly there exists $h \in \mathcal{H}$ such that $(h, A^{1/2}h) > 0$.

Since $\hat{A}1_\mathcal{H}$ is dense in $\mathcal{H}$, we may assume that $h = \hat{A}1_\mathcal{H}$ for some $\hat{A} \in \hat{A}$.

Then temporarily writing $\phi = (A \mapsto (1_\mathcal{H}, A1_\mathcal{H}))$ to give the state of $B(\mathcal{H})$ a name, we have

$$0 < (\hat{A}1_\mathcal{H}, A^{1/2}\hat{A}1_\mathcal{H}) = (1_\mathcal{H}, \hat{A}^* A^{1/2}\hat{A}1_\mathcal{H}) = \phi(\hat{A}^* A^{1/2}\hat{A}) = \phi(\hat{A}^* \hat{A}A^{1/2}).$$
Making further use of the hypothesis that operators in $\mathcal{A}$ commute with operators in $\hat{\mathcal{A}}$, we have

$$0 \leq \phi\left(\left(\sqrt{t}\hat{A}^*\hat{A} - A^{1/2}/\sqrt{t}\right)^2\right) = t\phi((\hat{A}^*\hat{A})^2) + \phi(A)/t - 2\phi(\hat{A}^*\hat{A}A^{1/2})$$

for $t > 0$.

The last inequality forces $\phi(A) > 0$.

Thus $\phi|_{\mathcal{A}}$ is indeed faithful.
The precursor to HST was the following result, which we are now in a position to state.

**Theorem ([Haagerup-Thorbjørnsen 2005])**

*Notation and assumptions are as for HST. For every* \( f \in \text{Mat}_n(\mathbb{C}\langle X_1, \ldots, X_m \rangle) \) (not necessarily self-adjoint)

\[
\lim_{N \to \infty} \left[ f \left( \frac{\Xi(N)}{\sqrt{N}} \right) \right] = \lfloor f(\Xi) \rfloor \text{ a.s.}
\]

**Recovery of HT from HST and amplification of Voiculescu’s theorem** After replacing \( f \) by \( ff^* \) we may assume that \( f \) is self-adjoint. The upper bound follows from HST and \( \text{supp}(f(\Xi)) = \text{supp}(\mu_f) \). The lower bound follows from the amplified version of Voiculescu’s theorem.
In the cited paper Schultz proves the analogue of HT with GUE replaced by GOE or GSE. The new difficulty emerging in the proof is that there are correction terms which must be dealt with.

Later when more technical details of the proof of HST have been discussed, we can say more precisely what this new difficulty is.
In the cited paper Capitaine and Donati-Martin prove the analogue of HT with the independent GUE matrices replaced by independent matrices of the following class.

Fix a random variable $Z$ with a symmetric distribution satisfying a Poincaré inequality, i.e., $\text{Var}(f(Z)) \leq cE|f'(Z)|^2$ for all nice enough functions $f : \mathbb{R} \to \mathbb{C}$, for a constant $c$ depending only on $Z$. Let $\mu$ be the law of $Z$. Consider random matrices $X \in \text{Herm}_N$ such that the family of random variables $(X, \hat{e}_{ij})$ is i.i.d., each with law $\mu$.

The algebra no longer works out quite so nicely so a new tool is needed. Capitaine and Donati-Martin use a certain form of “integration by parts” introduced in the influential paper [Khorunzhy-Khoruzhenko-Pastur]. We will discuss the latter circle of ideas at a later point in the course.
A natural question raised by HST is to consider what happens for polynomials in independent GUE matrices and some other independent matrices where the other matrices already are known to have a convergent joint distribution in the sense of noncommutative probability. This was studied in [Male 2010], and the exact analogue of HST was obtained in this more general setting, as well as the exact analogue of the amplification of Voiculescu’s theorem. The exact formulation Male’s result is too difficult to attempt here since it requires more background on free probability than we have supplied. It remains an open question to generalize Male’s results by replacing GUE matrices by Wigner matrices. Doing so, one would to a large extent succeed in recovering and generalizing results of Bai-Silverstein type.
A natural question raised by Male’s result is to consider what happens if one replaces the GUE matrices by unitary matrices. This was studied in [Collins-Male 2011] and the exact analogue of HT was recovered in this case. The method of proof is very clever, efficiently reducing the proof to the previous results in [Male 2010].
In [Anderson, Ann. Probab., to appear] the GUE matrices $\Xi^{(N)}_\ell$ in HST are generalized to the following collection of matrices. Let

$$\{\{\xi_\ell(i,j)\}_{1\leq i\leq j\leq m}\}_{\ell=1}^\infty$$

be an array of independent $\mathbb{C}$-valued random variables with finite absolute fourth moments and zero means such that the law of $\xi_\ell(i,j)$ depends only on $\ell$ and $1_{i<j}$. Assume furthermore that $\xi_\ell(1,1)$ is real-valued almost surely, that real and imaginary parts of $\xi_\ell(1,2)$ are independent and that $\mathbb{E}|\xi_\ell(1,2)|^2 = 1$ for all $\ell$. In the setting of HST consider random $\Xi^{(N)} \in \text{Herm}_N$ with entries $\Xi^{(N)}_\ell[i,j] = \xi_\ell(i,j)$ for $\ell = 1, \ldots, m$ and $i, j = 1, \ldots, N$. Then with these matrices $\Xi^{(N)}_\ell$ instead of GUE matrices, one can still draw the same conclusion as in HST. One also has the analogue of the amplification of Voiculescu’s theorem.
Our next goal is... to fill in background concerning largest eigenvalues of single Wigner-type matrices. Our main concern is to gather robust rough methods for capturing eigenvalues in compact sets.

This is the first stage of the long battle to drag the eigenvalues all the way into the support.
Given a $\mathbb{C}$-valued random variable $Z$, let $\|Z\|_\infty$ denote the essential supremum of $|Z|$ and let $\|Z\|_p = (E|Z|^p)^{1/p}$ for $p \in [1, \infty)$.

*Smooth* means infinitely differentiable.

Occasionally when brevity demands it we write $\land$ and $\lor$ for min and max, respectively.
The matrix norms $\| \cdot \|_p$

- Given $A \in \text{Mat}_{k \times \ell}$, let $\mu_1 \geq \cdots \geq \mu_{k \land \ell} \geq 0$ be the singular values of $A$, let $\|A\| = \|A\|_\infty = \mu_1$ and more generally for $p \in [1, \infty)$ let $\|A\|_p = \left( \sum_{i=1}^{k \land \ell} \mu_i^p \right)^{1/p}$. 

- $\| \cdot \|_p$ is a norm on $\text{Mat}_{k \times \ell}$.

- $\|UXV\|_p = \|X\|_p$ for unitary $U \in \text{Mat}_k$, unitary $V \in \text{Mat}_\ell$ and any $X \in \text{Mat}_{k \times \ell}$.

- $\sum_{i=1}^{k \land \ell} |A[i, i]|^p \leq \|A\|_p^p$ for $p \in [1, \infty)$.

- $\lim_{p \to \infty} \|A\|_p = \|A\|$.

- $\|AB\|_r \leq \|A\|_p \|B\|_q$ whenever $AB$ is defined and $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$.

- For the canonical basis $\{\hat{e}_{ij}\}_{i,j=1}^N$ of $\text{Herm}_N$ one has $\max_{i,j=1}^N \|\hat{e}_{ij}\|_1 \leq 1$.

See [Simon] for a good introduction to this material as well as applications in mathematical physics. Alternatively see the standard text [Horn-Johnson].
The following result is a key component of the proof of HST.

**Proposition (Rough control of eigenvalues)**

On a common probability space, for each positive integer $N$, let $\Xi^{(N)}$ be an $N$-by-$N$ GUE matrix. Then we have

$$\limsup_{N \to \infty} \| \Xi^{(N)} \|_{\infty} \leq \infty, \quad \sup_{p \in [1, \infty)} \limsup_{N \to \infty} \left\| \frac{\Xi^{(N)}}{\sqrt{N}} \right\|_p < \infty.$$

This is very old news and far weaker than what is known. In fact both quantities in question equal 2.

It is striking that the HST method does not require knowledge of the exact constant in order to work.
We reduce the proposition below to a well-known combinatorial estimate of Füredi-Komlós. We will do so by general methods so that the proposition can be extended to handle a wide class of sequences of Wigner-type matrices.
We say that random $X \in \text{Herm}_N$ is of Wigner type if the real random variables $(\hat{e}_{ij}, X)$ are independent, have absolute moments of all orders, and have mean zero.

Equivalently: random $X \in \text{Herm}_N$ is of Wigner type if the entries $X[i,j]$ for $1 \leq i \leq j \leq N$ are independent, have absolute moments of all orders, have mean zero, and have independent real and imaginary parts.

If $X \in \text{Herm}_N$ is of Wigner type we write $X \in \text{Wig}_N$.

To maintain flexibility we do not insist on normalizing the variance of entries of a Wigner-type matrix.
Let $\mathcal{W}_N(k)$ denote the set consisting of $(2k + 1)$-tuples

$$(i_0, \ldots, i_{2k}) = (i_\alpha)_{\alpha=0}^{2k} \in \{1, \ldots, N\}^{2k+1}$$

with the following two properties:

- $i_0 = i_{2k}$.
- Every set on the list $\{i_0, i_1\}, \ldots, \{i_{2k-1}, i_{2k}\}$ appears there at least twice.

We call such $(2k + 1)$-tuples weak Wigner words of length $2k + 1$. 
For $X \in \text{Wig}_N$ and positive integers $k$ one has

$$
\mathbb{E}[X]^{2k} \leq \mathbb{E}\text{tr}X^{2k} = \sum_{(i_\alpha)_{\alpha=0}^{2k} \in \mathcal{W}_N(k)} \mathbb{E} \prod_{\alpha=1}^{2k} X[i_{\alpha-1}, i_\alpha].
$$

The omitted terms vanish since they are of the form

$$
\mathbb{E}(UV), \quad \mathbb{E}U = 0, \quad U \text{ and } V \text{ independent}.
$$

This is the standard first move in the combinatorial analysis of largest eigenvalues of Wigner matrices.
Let $k$ be a positive integer. For $X \in \text{Herm}_N$ (yes, deterministic, not random) we define

$$\|X\|_{2k} = \left( \sum_{(i_\alpha)_{\alpha=0}^{2k} \in \mathcal{W}_N(k)} \prod_{\alpha=1}^{2k} |X[i_{\alpha-1}, i_{\alpha}]| \right)^{\frac{1}{2k}}.$$

For $X \in \text{Wig}_N$ the relation

$$\|X\|_{2k} \leq \|X\|_{2k} \leq (\text{Etr} X^{2k})^{\frac{1}{2k}} \leq \|X\|_{2k}$$

is just a rewrite (in admittedly eccentric notation) of the “standard first move” recalled in the previous frame.

We use the scare-quotes because $\|\cdot\|_{2k}$ is only approximately a norm.

The main reason for introducing $\|\cdot\|_{2k}$ is that it has an obvious monotonicity property which $\|\cdot\|_{2k}$ does not.
Lemma

Given $X_1, \ldots, X_m \in \text{Herm}_N$ and a positive integer $k$ we have

$$\left[\left[ \sum_{\ell=1}^{m} X_\ell \right]\right]_{2k} \leq \frac{1 + \sqrt{5}}{2} \sum_{\ell=1}^{m} \|X_\ell\|_{2k}.$$

The lemma is perhaps not so important but the proof is entertaining.
Proof

We say that a real random variable $\chi$ has the *Fibonacci law* if it is bounded, of mean zero and satisfies

$$\chi^2 = \chi + 1,$$

in which case

$$E\chi^n = f_{n-1} \geq 1$$

for integers $n \geq 2$, where $f_n$ is the $n^{th}$ Fibonacci number and

$$\|\chi\|_\infty = \frac{1 + \sqrt{5}}{2}.$$
Now let \( \{\chi_{ij}\}_{1 \leq i \leq j \leq N} \) be an i.i.d. family of Fibonacci random variables. Let \( \{e_{ij}\}_{i,j=1}^{N} \) be the family of elementary \( N \)-by-\( N \) matrices. Put

\[
\tilde{X}_\ell = \sum_{i=1}^{N} \chi_{ii} |X_\ell[i, i]| + \sum_{1 \leq i < j \leq N} \chi_{ij} |X_\ell[i, j]| (e_{ij} + e_{ji}),
\]

\[
Z = \sum_{\ell=1}^{m} X_\ell, \quad \tilde{Z} = \sum_{\ell=1}^{N} \tilde{X}_\ell,
\]

noting that \( \tilde{X}_\ell, \tilde{Z} \in \text{Wig}_N \) and \( Z \in \text{Herm}_N \).
Clearly we have

\[ \left\| \left[ X_\ell \right]_{2k} \right\|_{2k} \leq \frac{1 + \sqrt{5}}{2} \left\| X_\ell \right\|_{2k}. \]

It is easy to see that for \((i_\alpha)_{\alpha=0}^{2k} \in \mathcal{W}_N(k)\) we have

\[ \prod_{\alpha=1}^{2k} |Z[i_{\alpha-1}, i_\alpha]| \leq E \prod_{\alpha=1}^{2k} \tilde{Z}[i_{\alpha-1}, i_\alpha] \]

and thus

\[ \| Z \|_{2k} \leq \left\| \left[ \tilde{Z} \right]_{2k} \right\|_{2k}. \]

Since \(\| \cdot \|_{2k} \) is truly a norm, the result follows. \( \square \)
Lemma (Rough control via $\| \cdot \|_{2k}$)

For each sufficiently large positive integer $N$ let there be given $X^{(N)} \in \text{Wig}_N$ and an even positive integer $q_N$. Fix constants $C \in [0, \infty)$ and $\epsilon \in (0, \infty]$. Assume that

$$\limsup_{N \to \infty} \left\| \left[ \left[ X^{(N)} \right] \right]_{q_N} \right\|_{q_N} \leq C \quad \text{and} \quad \liminf_{N \to \infty} \frac{q_N}{\log N} \geq \epsilon.$$ 

Then

$$\sup_{q \in [1, \infty)} \limsup_{N \to \infty} \left\| \left[ X^{(N)} \right] \right\|_q \leq K \quad \text{and} \quad \limsup_{N \to \infty} \left\| \left[ X^{(N)} \right] \right\|_{\infty} \leq K$$

for a constant $K \in [0, \infty)$ depending only on $C$ and $\epsilon$. 
Rough control of eigenvalues via $\mathbb{I} \cdot \mathbb{I}_{2k}$ (proof)

**Proof** Pick $0 < \eta < \epsilon$ and $C < D < \infty$ arbitrarily. For some positive integer $N_0$ and for all $N \geq N_0$ the random matrix $X^{(N)}$ is defined and we have

$$e^{q_N/\eta} \geq N \quad \text{and} \quad \| [X^{(N)}] \|_{q_N} \leq \| [ [X^{(N)}] ]_{q_N} \|_{q_N} \leq D,$$

which already proves the first claim with $K = D$, and we have

$$\sum_{N \geq N_0} \Pr \left( [X^{(N)}] > De^{2/\eta} \right) \leq \sum_{N \geq N_0} E \left( \frac{X^{(N)}}{De^{2/\eta}} \right)^{q_N} \leq \sum_{N \geq N_0} \frac{1}{N^2} < \infty.$$

The latter statement proves the second claim with $K = De^{2/\eta}$ by the Borel-Cantelli lemma.

After optimizing over $D$ and $\eta$ we can take, say, $K = Ce^{2/\epsilon}$. \qed
Let

\[ \mathcal{W}_N(k, w) \subset \mathcal{W}_N(k) \]

denote the subset consisting of \((2k + 1)\)-tuples in which exactly \(w\) distinct integers appear.

**Theorem ([Füredi-Komlós 1981])**

*For positive integers \(k, w\) and \(N\) we have*

\[ |\mathcal{W}_N(k, w)| \leq 2^{2k}(2k)^{6(k-w+1)} N^w 1_{w-1 \leq k}. \]

See for example [Anderson-Guionnet-Zeitouni] for treatment and application of this theorem at length.
It is well-known that

\[ |\mathcal{W}_N(k, k+1)| = \frac{1}{k+1} \binom{2k}{k} N(N-1) \cdots (N-w+1) \leq 2^{2k-1} N^w. \]

This follows, say, by interpreting elements of the set in question as rooted planar trees with \( k + 1 \) vertices each bearing a distinct label chosen from \( \{1, \ldots, N\} \). These are the least “singular” of the elements of \( \mathcal{W}_N(k) \).

This is the main point in the combinatorial moment-method proof of the semicircle law.
Idea of proof of FK estimate

We carry the explanation just far enough to see how the form of the estimate comes about.

Fix

\[ i = (i_\alpha)_{\alpha=0}^{2k} \in \mathcal{W}_N(k, w). \]

Put

\[ I_1 = \{ \alpha \in \{1, \ldots, 2k\} \mid i_\alpha \not\in \{i_0, \ldots, i_{\alpha-1}\} \}. \]

Consider the tree \( T = (V, E) \) where

\[ V = \{ i_\alpha \in \{1, \ldots, N\} \mid \alpha = 0, \ldots, 2k \} \quad \text{and} \quad E = \{ \{i_{\alpha-1}, i_\alpha\} \mid \alpha \in I_1 \}. \]

Note that

\[ |V| = w, \quad |I_1| = w - 1. \]

The sequence \( i = (i_\alpha)_{\alpha=0}^{2k} \) is only approximately a walk on \( T \).

Sometimes it takes steps along edges of \( T \) but other times it takes a gratuitous leap.
Let

$$I_2 \subset \{1, \ldots, 2k\}$$

be the set of “times” $\alpha$ at which the walk-with-jumps $i$ completes a visit to an edge of $T$ it has visited exactly once before.

Note that

$$|I_2| = |I_1| = w - 1 \leq k.$$ 

by the definition of a weak Wigner word.
Let $A$ be the partition of the set $\{0, \ldots, 2k\}$ into the vertex sets of the connected components of the graph

$$(\{0, \ldots, 2k\}, \{\{\alpha - 1, \alpha\} | \alpha \in l_1 \cup l_2\}).$$

Each block of the partition $A$ is a set of consecutive integers.

Note that

$$|A| = 2k - 2w + 3.$$ 

The point of this “parsing” operation is that for each $A \in A$ the subsequence

$$i_A = \{i_\alpha\}_{\alpha \in A}$$

is an honest walk on $T$. 
It can be shown that $i$ is uniquely determined from knowledge of

- $I_1$;
- $I_2$;
- $i_\alpha$ for each $\alpha \in I_1 \cup \{0\}$;
- $i_{\min A}$ for each $0 \not\in A \in \mathcal{A}$, and
- $i_{\text{crit} A}$ for each $0 \not\in A \in \mathcal{A}$ where $\text{crit} A$ is the time of the last visit of the walk $i_A$ to the set $\{i_\alpha \mid \alpha \leq \min A\}$.

Counting these choices somewhat crudely we get the bound

$$\frac{(2k!)(2w-2)!}{(2w-2)!(2k-2w+2)!} \cdot \frac{(2w-2)!}{(w-1)!(w-1)!} \cdot N^w w^{4(k-w+1)}$$

whence after further simplification the announced bound.
Lemma

Fix $X \in \operatorname{Wig}_N$ and $(i_\alpha)_{\alpha=0}^{2k} \in \mathcal{W}_N(k)$. If

$$\sup_{q \in [2, \infty)} \max_{i,j=1}^N q^{-c} \|X[i,j]\|_q \leq 1.$$ 

then

$$\mathbb{E} \prod_{\alpha=1}^{2k} |X[i_{\alpha-1}, i_\alpha]| \leq 2^{2ck} (2k)^{2c(k-w+1)}.$$ 

Roughly, in words, the hypothesis says that you have polynomial control of $L^q$-norms of entries.
Consider the (unoriented) graph $G = (V, E)$ where

$$V = \{i_\alpha \mid \alpha = 0, \ldots, 2k\} \quad \text{and} \quad E = \{\{i_{\alpha-1}, i_\alpha\} \mid \alpha = 1, \ldots, 2k\}.$$ 

Since this time, in contrast to the procedure used in the proof of the FK bound, no edges were discarded,

$$i = (i_\alpha)_{\alpha=0}^{2k}$$

as an honest walk on $G$. 
Proof of technical lemma (concluded)

Let \( p_n \) be the number of edges of \( G \) visited \( n \) or more times by \( i \). Clearly \( p_1 = p_2 \) (this is just a re-iteration of the definition of a weak Wigner word) and \( \sum_n p_n = \sum_{n=1}^{2k} p_n = 2k \).

Furthermore since \( G \) is connected \( p_1 + 1 \leq w \). It follows that \( \sum_{n \geq 3} p_n \leq 2(k - w + 1) \).

Finally we have

\[
\mathbb{E} \prod_{\alpha=1}^{2k} |X[i_{\alpha-1}, i_\alpha]| \leq \prod_{n=2}^{\infty} e^{cn \log(n)(p_n - p_{n+1})} \\
= 2^{2cp_2} \prod_{n=3}^{2k} e^{c(n \log(n) - (n-1) \log(n-1)) p_n} \leq 2^{2ck} \prod_{n=3}^{2k} e^{c \log(n) p_n} \\
\leq 2^{2ck} \prod_{n=3}^{2k} (2k)^{cp_n} \leq 2^{2ck} (2k)^{2c(k-w+1)}. \tag*{□}
\]
Corollary

Let $X \in \text{Wig}_N$ and a constant $c \in [0, \infty)$ be given such that

$$\sup_{q \in [2, \infty)} \max_{i,j=1}^N q^{-c} \|X[i,j]\|_q \leq 1.$$ 

Then for positive integers $k$ we have

$$2(2k)^{6+2c} \leq N \Rightarrow \left\|\left\|\left[\left[\frac{X}{2^{c+1} \sqrt{N}}\right]\right]\right\|_{2k} \right\|_{2k} \leq (2N)^{\frac{1}{2k}}.$$
\begin{align*}
\left\| \left[ \left[ \frac{X}{\sqrt{N}} \right] \right] \right\|_{2k}^{2k} & \leq \sum_{w=1}^{k+1} |\mathcal{W}_N(k, w)| \frac{2^{2ck}(2k)^{2c(k-w+1)}}{N^k} \\
& \leq \sum_{w=1}^{k+1} 2^{2k}(2k)^6(k-w+1) N^2 \frac{2^{2ck}(2k)^{2c(k-w+1)}}{N^k} \\
& \leq 2^{(1+c)(2k)} N \sum_{w=1}^{k+1} \frac{(2k)(6+2c)(k-w+1)}{N^{k-w+1}} \\
& \leq 2^{1+(1+c)(2k)} N.
\end{align*}
Recall that to get rough control of eigenvalues of GUE matrices we needed to prove the following statement.

**Proposition**

On a common probability space, for each positive integer $N$, let $\Xi^{(N)}$ be an $N$-by-$N$ GUE matrix. Then we have

$$\limsup_{N \to \infty} \left\| \sqrt{N} \Xi^{(N)} \right\|_{\infty} < \infty,$$

$$\sup_{p \in [1, \infty)} \limsup_{N \to \infty} \left\| \sqrt{N} \Xi^{(N)} \right\|_{p} < \infty.$$
Rough control of eigenvalues of more general Wigner matrices

The next result vastly generalizes the previous one since for a standard normal random variable $Z$ we have $\|Z\|_q = O(\sqrt{q})$. It also handles fake GUE matrices.

**Proposition**

Fix a constant $c \geq 0$. On a common probability space, for each positive integer $N$, let there be given $\Xi^{(N)} \in \text{Wig}_N$ such that

$$
\sup_{q \in [2, \infty)} \sup_{N=1}^\infty \max_{i,j=1} q^{-c} \left\| \Xi^{(N)}[i,j] \right\|_q \leq 1.
$$

Then for an absolute constant $K$ depending only on $c$ we have

$$
\left\| \limsup_{N \to \infty} \left[ \frac{\Xi^{(N)}}{\sqrt{N}} \right] \right\|_{\infty} \leq K, \quad \sup_{p \in [1, \infty)} \limsup_{N \to \infty} \left\| \left[ \frac{\Xi^{(N)}}{\sqrt{N}} \right] \right\|_p \leq K.
$$

**Proof** The preceding corollary along with the rough control in terms of $\|\mathbb{I} \cdot \mathbb{I}_{2k}\|_{2k}$ together prove this.
We have analyzed the proof of Bai-Yin and extracted a combinatorial estimate from it. Derivation of the estimate is too technical to go into. But the estimate itself is relatively simple to state and it interesting to compare it to the FK estimate.
The Bai-Yin parameter

Let

\[ \mathbf{i} = (i_{\alpha})_{\alpha=0}^{2k} \in \mathcal{W}_N(k, w) \]

be a weak Wigner word. Consider again the graph

\[ G = (\{i_{\alpha} \mid \alpha = 0, \ldots, 2k\}, \{\{i_{\alpha-1}, i_{\alpha}\} \mid \alpha = 1, \ldots, 2k\}) \].

As before, let \( p_n \) denote the number of edges of \( G \) visited at least \( n \) times by \( \mathbf{i} \).

We define the Bai-Yin parameter of \( \mathbf{i} \) to be the quantity

\[ p_1 + 1 - w + p_3. \]

In words this parameter is the number of edges of \( G \) not needed to make \( G \) connected plus the number of edges visited three or more times.
The Bai-Yin parameter is a rough measure of the “badness” of $G$ and the walk $i$ on $G$ which is sensitive to third moments.

The importance of this parameter is something one can learn by attending to the proofs in Bai-Yin.
Let $\mathcal{W}_N(k, w, t)$ denote the subset of $\mathcal{W}_N(k, w)$ consisting of weak Wigner words with Bai-Yin parameter equal to $t$.

**Theorem (FKBY estimate)**

For integers $k, w, t, N > 0$ we have

$$|\mathcal{W}_N(k, w, t)| \leq 2^{6k} w^{3t} t^{2(k-w+1)} N^w 1_{w \leq k} 1_{t \leq 2(k-w+1)}.$$

For comparison recall that the FK estimate is

$$|\mathcal{W}_N(k, w)| \leq 2^{2k} (2k)^{6(k-w+1)} N^w 1_{w \leq k+1}.$$

The FKBY estimate can be proved by only a light modification of FK arguments. It does not require some alternative development of graph-theoretical tools.
The FKBY estimate is good enough to prove the following estimate which cannot so far as we know be proved using the FK estimate.

**Proposition**

*On a common probability space, for each positive integer $N$, let there be given $\Xi^{(N)} \in \text{Wig}_N$ such that*

$$\sup_{N=1}^{\infty} \max_{i,j=1}^{N} \left\| \Xi^{(N)}[i,j] \right\|_3 \leq 1 \quad \text{and} \quad \sup_{N=1}^{\infty} \max_{i,j=1}^{N} \left\| \Xi^{(N)}[i,j] \right\|_{\infty} \leq \sqrt{N}.$$

*Then for some absolute constant $K$ we have*

$$\left\| \limsup_{N \to \infty} \left[ \frac{\Xi^{(N)}}{\sqrt{N}} \right] \right\|_{\infty} < K \quad \text{and} \quad \sup_{p \in [1,\infty)} \limsup_{N \to \infty} \left\| \left[ \frac{\Xi^{(N)}}{\sqrt{N}} \right] \right\|_p < K.$$

This result is strong enough to permit deduction of the theorem of Bai-Yin from the preceding results on rough control derived from the FK bound.
Our next goal is... is to briefly explain by example the role of truncation techniques.
In order to show how one can exploit truncation techniques, we carry out an exercise. Recall the theorem of Bai-Yin.

**Theorem (Bai-Yin [Bai-Yin 1988])**

Let \(\{\xi_{ij}\}_{1 \leq i \leq j < \infty}\) be an independent family of \(\mathbb{C}\)-valued random variables such that the law of \(\xi_{ij}\) depends only on \(1_{i < j}\), \(\xi_{11}\) is real almost surely, \(\|\xi_{11}\|_4 < \infty\) and \(\|\xi_{12}\|_4 < \infty\), and \(\mathbb{E}\xi_{11} = 0 = \mathbb{E}\xi_{12}\). Let \(\Xi^{(N)}\) be the random \(N\)-by-\(N\) hermitian matrix with entries \(\Xi^{(N)}[i, j] = \xi_{ij}\) for \(i \leq j\). Then

\[
\lim_{N \to \infty} \left[ \Xi^{(N)} \sqrt{\frac{2}{N}} \right] = 2\|\xi_{12}\|_2 \text{ almost surely.}
\]
In order to separate important issues from less important ones, consider the following three statements.

**Proposition (A)**

Hypotheses are those of the theorem strengthened so that \( \| \xi_{12} \|_2 = 0 \). Then \( \lim_{N \to \infty} \left[ \frac{\Xi(N)}{\sqrt{N}} \right] = 0 \) almost surely.

**Proposition (B)**

Hypotheses are those of the theorem strengthened so that \( \max(\| \xi_{11} \|_{\infty}, \| \xi_{12} \|_{\infty}) < \infty \). Then \( \lim_{N \to \infty} \left[ \frac{\Xi(N)}{\sqrt{N}} \right] = 2 \| \xi_{12} \|_2 \) almost surely.

**Proposition (C)**

Hypotheses are exactly as in the theorem but one concludes only that \( \limsup_{N \to \infty} \left[ \frac{\Xi(N)}{\sqrt{N}} \right] \leq K \| \xi_{12} \|_4 \) for an absolute constant \( K \).
In order to complete our separation of issues, we will prove the following logical relationship.

**Proposition ("A,B,C implies Bai-Yin")**

*Propositions A,B,C together imply the theorem.*

Before starting the proof we comment on the proofs of statements A, B and C.
Proposition $A$ asserts that the theorem of Bai-Yin holds when the Wigner matrices in question are diagonal. To prove this is an exercise in the Borel-Cantelli lemma. To get the necessary convergence of a series one exploits the integration identity

$$
\mathbb{E}Z^4 = \int_0^\infty \Pr(Z^4 > u) \, du = 2 \int_0^\infty t \Pr(Z > \sqrt{t}) \, dt.
$$
Proposition $B$ asserts the weaker version of the theorem of Bai-Yin that was proved earlier by Füredi-Komlós. Most of the arguments needed for that proof have been rehearsed above. For the remaining details see, for example, [Anderson-Guionnet-Zeitouni].
The FKBY criterion stated above along with the application following it are strong enough to prove statement C but do not make the constant $K$ explicit.
Proof of “A,B,C implies Bai-Yin”

By Proposition A we may assume that diagonal entries of \( \Xi^{(N)} \) vanish identically. For any constant \( c > 0 \), let \( \Xi^{(N)}_{\leq c} \) be the \( N \)-by-\( N \) random hermitian matrix with entries 

\[
\Xi^{(N)}_{\leq c}[i,j] = \xi_{ij} \mathbf{1}_{|\xi_{ij}| \leq c}
\]

for \( i \leq j \) and put \( \hat{\Xi}^{(N)}_{\leq c} = \Xi^{(N)}_{\leq c} - E \Xi^{(N)}_{\leq c} \). Then for any \( c > 0 \) we have almost surely

\[
\lim_{N \to \infty} \left[ \left[ \frac{\Xi^{(N)}_{\leq c}}{\sqrt{N}} \right] \right] = 2 \left\| \xi_{12} \mathbf{1}_{|\xi_{12}| \leq c} - E \xi_{12} \mathbf{1}_{|\xi_{12}| \leq c} \right\|_2
\]

and

\[
\limsup_{N \to \infty} \left[ \left[ \frac{\Xi^{(N)} - \Xi^{(N)}_{\leq c}}{\sqrt{N}} \right] \right] \leq K \left\| \xi_{12} \mathbf{1}_{|\xi_{12}| > c} - E \xi_{12} \mathbf{1}_{|\xi_{12}| > c} \right\|_4
\]

by Propositions B and C, respectively. By dominated convergence the preceding statements imply the desired result.

\[\square\]
Bai-Silverstein in [Bai-Silverstein 1998] start their proof with a truncation step similar in form to the example considered above.


It seems that quite generally support questions involving Wigner matrices satisfying parsimonious moment conditions can be reduced to the study of Wigner matrices with bounded entries.
Our next goal is...

...to recall the Poincaré inequality for Gaussian random variables along with a carefully organized proof of it designed to expedite a certain generalization to be considered later.
Let $Z \in \mathbb{R}^n$ be a random vector with i.i.d. centered Gaussian entries, all of variance $\sigma^2$.

Let $f : \mathbb{R}^n \to \mathbb{C}$ be infinitely differentiable and suppose that partial derivatives of all orders have polynomial growth; hereafter we call such functions *tame*.

Let $D_i$ denote the operation of differentiation of functions on $\mathbb{R}^n$ with respect to the $i^{th}$ coordinate.
Statement of the Poincaré inequality

**Theorem (The Poincaré inequality for Gaussian random variables)**

Notation and assumptions are as above. We have

\[
E|f(Z)|^2 - |Ef(Z)|^2 \leq \sigma^2 E \sum_{i=1}^{n} |D_i[f](Z)|^2.
\]

For simplicity we have given ourselves assumptions concerning the regularity of \( f \) that are much stronger than necessary but which are no trouble to verify in the applications we have in mind.

In the next several frames we give a proof by way of a lemma which we will use later to generalize the Poincaré inequality in a useful way.
Setup for a technical lemma

For the moment we assume of the random vector $Z \in \mathbb{R}^n$ only that its entries have absolute moments of all orders so as to rule out integrability issues, and we let $\sigma^2$ denote an arbitrary positive constant.

Let $X, Y \in \mathbb{R}^n$ be i.i.d. copies of $Z$ and for $t \in (0, 1)$ put

$$X^{(t)} = \sqrt{t}X + \sqrt{1-t}Y.$$  

We introduce the auxiliary functions

$$h_i^{(y,t)} : \mathbb{R}^n \rightarrow \mathbb{C} \text{ for } y \in \mathbb{R}^n, 0 < t < 1 \text{ and } i = 1, \ldots, n$$

defined by

$$h_i^{(y,t)}(x) = (f^*(x) - f^*(y))f_i(\sqrt{t \cdot x} + \sqrt{1-t \cdot y}).$$

Finally, let $M_i$ denote the operation of multiplication of functions on $\mathbb{R}^n$ by the $i^{th}$ coordinate.
Statement of the technical lemma

Lemma (The pre-Poincaré identity)

Assumptions and notation are as above. We have an integration identity

\[ \mathbb{E}|f(Z)|^2 - |\mathbb{E}f(Z)|^2 - \sigma^2 \mathbb{E} \int_0^1 \sum_{i=1}^n D_i[f^*](X)D_i[f](X^{(t)}) \, d\sqrt{t} \]

\[ = \sum_{i=1}^n \mathbb{E} \int_0^1 (M_i - \sigma^2 D_i)[h_i^{(Y,t)}](X) \, d\sqrt{t}. \]

Later we will present a nice method for bounding the right side.
Proof of the technical lemma

We start by using the fundamental theorem of calculus.

\[ E|f(Z)|^2 - |Ef(Z)|^2 = Ef^*(X)(f(X) - f(Y)) \]
\[ = \int_0^1 \frac{d}{dt}Ef^*(X)f(X(t))
\]
\[ = E\int_0^1 \sum_{i=1}^n f^*(X) \left( \frac{X_i}{2\sqrt{t}} - \frac{Y_i}{2\sqrt{1-t}} \right) D_i[f](X^{(t)})
\]
Proof of the technical lemma (continuation)

Then we observe a symmetry.

\[
E \int_0^1 \sum_{i=1}^n f^*(X) \frac{Y_i}{2\sqrt{1-t}} D_i[f](\sqrt{t} X + \sqrt{1-t} Y) \, dt
= E \int_0^1 \sum_{i=1}^n f^*(Y) \frac{X_i}{2\sqrt{t}} D_i[f](\sqrt{1-t} Y + \sqrt{t} X) \, dt.
\]

Thus we have

\[
E|f(Z)|^2 - |E f(Z)|^2
= E \int_0^1 \sum_{i=1}^n X_i(f^*(X) - f^*(Y)) D_i[f](X(t)) \, d\sqrt{t}
= \sum_{i=1}^n E \int_0^1 M_i[h_i^{(Y,t)}](X) \, d\sqrt{t}.
\]
Proof of the technical lemma (conclusion)

For $i = 1, \ldots, n$ we have

$$D_i[h_i^{(y,t)}](x) = D_i[f^*](x)D_i[f](\sqrt{t}x + \sqrt{1 - ty})$$

$$+ (f^*(x) - f^*(y))D_i^2[f](\sqrt{t}x + \sqrt{1 - ty})\sqrt{t}$$

and thus

$$E \int_0^1 D_i[h_i^{(y,t)}](X)d\sqrt{t} = E \int_0^1 D_i[f^*](X)D_i[f](X^{(t)}) d\sqrt{t}.$$

The right side has only the one term because the other is killed by odd symmetry. The proof of the lemma is complete.
Proof of the Poincaré inequality (conclusion)

For a standard normal random variable $\zeta$ and tame function $g : \mathbb{R} \to \mathbb{C}$ we have

$$E\zeta g(\zeta) = \sigma^2 E g'(\zeta).$$

Let $X, Y \in \mathbb{R}^n$ be i.i.d. copies of a Gaussian random vector $Z \in \mathbb{R}^n$ with i.i.d. centered entries of variance $\sigma^2$.

Plugging into the technical lemma we can now kill the error on the right side, thus obtaining the (very well-known) identity

$$E|f(Z)|^2 - |Ef(Z)|^2 = \sigma^2 E \int_0^1 \sum_{i=1}^n D_i[f^*](X)D_i[f](tX + \sqrt{1 - t^2} Y) \, dt$$

whence the result via Cauchy-Schwarz. The proof of the theorem is complete.

\[\square\]
Our next goal... 

...is to introduce a technique for approximately integrating by parts and then by means of this technique to prove a variant of the Poincaré inequality.

The papers [Khorunzhy-Khoruzhenko-Pastur] and [Lytova-Pastur 2009 A] are the sources of this excellent idea.

It has been taken up elsewhere in RMT, e.g., it is used in [O’Rourke-Renfrew-Soshnikov 2011 A] and [O’Rourke-Renfrew-Soshnikov 2011 B] and called there the decoupling formula. See also [Lytova-Pastur 2009 B].

We reconfigure the idea a bit to handle the new situation of polynomials in Wigner matrices. The idea is quite robust so this turns out not to be especially difficult.
Let $Z$ be a real random variable with absolute moments of all orders.

Let $f : \mathbb{R} \to \mathbb{C}$ be an infinitely differentiable function such that derivatives of all orders have polynomial growth. (Hereafter such is called tame.)

Let $U_1, \ldots, U_k$ be real random variables independent of $Z$ such that $U_i$ has the beta distribution of parameters 1 and $i$.

Recall that this means $E\varphi(U_k) = k \int_0^1 (1 - t)^{k-1} \varphi(t) \, dt$.

Let $\kappa_i(Z)$ denote the $i^{th}$ cumulant of $Z$.

Recall that $\sum_{n=1}^{\infty} \frac{\kappa_n(Z)}{n!} t^n = \log \left( \sum_{n=0}^{\infty} \frac{EZ^n}{n!} t^n \right)$.
Integration by parts identity

Proposition (Integration by parts identity)

Assumptions and notation are as above. We have

\[
E \left( Zf(Z) - \sum_{i=0}^{k-1} \frac{\kappa_i+1(Z)}{i!} f^{(i)}(Z) \right) = 
E \left( \frac{Z^{k+1}}{k!} f^{(k)}(U_k Z) - \sum_{i=0}^{k-1} \frac{\kappa_i+1(Z)}{i!} \frac{Z^{k-i}}{(k-i)!} f^{(k)}(U_{k-i} Z) \right).
\]

Things are set up so that there are no issues of integrability to contend with.

We review relevant definitions and then prove the proposition in the next several slides.
By definition we have a formal power series identity

\[ \sum_{n=1}^{\infty} \frac{\kappa_n(Z)}{n!} t^n = \log \left( \sum_{n=0}^{\infty} \frac{E_{Z^n}}{n!} t^n \right). \]

Applying \( \frac{d}{dt} \) on both sides and rearranging, we get a factorization

\[
\left( \sum_{n=0}^{\infty} \frac{E_{Z^{n+1}}}{n!} t^n \right) = \left( \sum_{n=0}^{\infty} \frac{\kappa_{n+1}(Z)}{n!} t^n \right) \left( \sum_{n=0}^{\infty} \frac{E_{Z^n}}{n!} t^n \right)
\]
Cumulant moment relations

... and thus relations

$$\frac{\mathbf{E}Z^{n+1}}{n!} = \sum_{i=0}^{n} \frac{\kappa_{i+1}(Z)}{i!} \frac{\mathbf{E}Z^{n-i}}{(n-i)!}$$

for integers $n \geq 0$. 
Recall that Taylor’s formula says that

\[
f(x) = \sum_{n=0}^{k-1} \frac{x^n}{n!} f^{(n)}(0) + \frac{x^k}{(k-1)!} \int_0^1 f^{(k)}(tx)(1 - t)^{k-1} \, dt
\]

\[
= \sum_{n=0}^{k-1} \frac{x^n}{n!} f^{(n)}(0) + \frac{x^k}{k!} \mathbf{E} f^{(k)}(U_k x).
\]
Simple consequences of the preceding identities (toward proof of IBPI)

We have

\[ \sum_{n=0}^{k-1} \frac{EZ^{n+1}}{n!} f^{(n)}(0) = \sum_{i=0}^{k-1} \sum_{n=i}^{k-1} \frac{\kappa_{i+1}(Z)}{i!} \frac{EZ^{n-i}}{(n-i)!} f^{(n)}(0), \]

and

\[
EZf(Z) = \sum_{n=0}^{k-1} \frac{EZ^{n+1}}{n!} f^{(n)}(0) + E \frac{Z^{k+1}}{k!} f^{(k)}(U_k Z),
\]

\[
Ef^{(i)}(Z) = \sum_{n=i}^{k-1} \frac{EZ^{n-i}}{(n-i)!} f^{(n)}(0) + E \frac{Z^{k-i}}{(k-i)!} Ef^{(k)}(U_{k-i} Z)
\]

for \( i = 0, \ldots, k - 1. \)
One simply combines the identities on the preceding frame appropriately to get the claimed formula.
Going forward, we will make use of the preceding theory exclusively through use of a drastically simplified version of the preceding identity. Needless to say it does not exhaust the possibilities inherent in IBPI.
As before let $Z$ be a real random variable with absolute moments of all orders.

Fix an integer $k \geq 2$ and assume that $Z$ has the same first $k$ moments as a centered Gaussian random variable. (Hereafter such is called Gaussian through the $k^{th}$ order.)

As before let $f : \mathbb{R} \to \mathbb{C}$ be tame. Put

$$\mathcal{M}_k[f](x) = \sup_{t \in [0,1]} |f^{(k)}(tx)|.$$

**Proposition (Drastically simplified IBPI)**

Assumptions and notation are as above. We have

$$|\mathbb{E}(Zf(Z) - \text{Var}(Z)f'(Z))| \leq 2\|Z\|^{k+1}_{2(k+1)}\left\|\mathcal{M}_k[f](Z)\right\|_2.$$
By hypothesis

\[ \kappa_j(Z) = \delta_j \text{Var}(Z) \quad \text{for } j = 1, \ldots, k. \]

Thus IBPI simplifies in the present case to

\[
\mathbb{E}(Zf(Z) - f'(Z)) = \mathbb{E}\left( \frac{Z^{k+1}}{k!} f^{(k)}(U_k Z) - \text{Var}(Z) \frac{Z^{k-1}}{(k-1)!} f^{(k)}(U_{k-1} Z) \right).
\]

Now make the obvious application of the Hölder inequality and the definitions.


Benoit Collins, Camille Male *The strong asymptotic freeness of Haar and deterministic matrices* arXiv:1105.4345

A. Deya, I. Nourdin, *Convergence of Wigner integrals to the tetilla law* arXiv:1107.3538


