Near-extreme eigenvalues of random matrices and systems of coupled Painlevé II equations

Tom Claeys

UCLouvain (Belgium)

Based on joint work with Antoine Doeraene

CIRM Luminy

March 2, 2017

The Airy ensemble

- Consider the **Airy ensemble** or Airy point process characterized by correlation functions

$$\rho_k(x_1, \dots, x_k) = \det\left(K^{\mathrm{Ai}}(x_i, x_j)\right)_{i,j=1,\dots,k},$$

where K^{Ai} is the Airy kernel

$$K^{\mathrm{Ai}}(x, y) = \frac{\mathrm{Ai}\,(x)\mathrm{Ai}\,'(y) - \mathrm{Ai}\,'(x)\mathrm{Ai}\,(y)}{x - y}$$

  - distributions of individual eigenvalues are well understood
  - what about distributions of quantities involving more than 1 eigenvalue, such as spacing distributions?

The Airy ensemble

- This process appears as limiting process for
  - **largest eigenvalues** in many random matrix ensembles as the size goes to infinity
  - lengths of first rows of **random partitions** of n following the Plancherel measure
  - random **tilings**
  - non-intersecting **Brownian paths**
- We order the random points as

$$\zeta_1 > \zeta_2 > \zeta_3 > \dots$$

(the largest particle exists almost surely)

Distribution of largest values

- The distribution of the **largest particle** is given by the Fredholm determinant
$$\mathbb{P}\left(\zeta_1 < x\right) = \det\left(1 - \left.K^{\mathrm{Ai}}\right|_{(x,+\infty)}\right)$$
- The distribution of the k**-th largest particle** ζ_k is generated by the Fredholm determinant
$$F(x;s) := \det\left(1 - (1-s)\left.K^{\mathrm{Ai}}\right|_{(x,+\infty)}\right) = \mathbb{E}\left(s^{n_{(x,+\infty)}}\right)$$
- It is given by
$$\mathbb{P}\left(\zeta_k < x\right) = \sum_{j=0}^{k-1} \frac{1}{j!}\frac{d^j}{ds^j}\left.F(x;s)\right|_{s=0}$$

Tracy-Widom formula's

- The Fredholm determinants $F(x;s) = \det\left(1 - (1-s)\, K^{\mathrm{Ai}}\big|_{(x,+\infty)}\right)$ can be expressed in terms of solutions to the **Painlevé II equation** *(Tracy-Widom '93)*:

$$F(x;s) = \exp\left(-\int_x^{+\infty} (\xi - x) q^2(\xi;s) d\xi\right)$$

where q is the **Ablowitz-Segur** (for $s \neq 0$) or **Hastings-McLeod** (for $s = 0$) solution of Painlevé II characterized by

$$q_{xx} = xq + 2q^3, \qquad q(x;s) \sim \sqrt{1-s}\,\mathrm{Ai}(x),\ x \to +\infty$$

Joint distributions

- What can we say about the distribution of **quantities involving more than one particle**? For instance
  - distribution $\mathbb{P}\left(\zeta_k - \zeta_\ell < x\right)$ of the **spacing** between ζ_k and ζ_ℓ, $k < \ell$,
    - $k = 1, \ell = 2$: see *(Bornemann-Forrester-Witte '12, Perret-Schehr '14, Deift-Trogdon '16)*
  - distribution $\mathbb{P}\left(\zeta_1 + \zeta_2 + \ldots + \zeta_k < x\right)$ of the **sum** of the k largest particles,
  - **joint distribution** of k particles

$$\zeta_{i_1} > \ldots > \zeta_{i_k}$$

  - distribution of **truncated linear statistics** *(Grabsch-Majumdar-Texier '16)*

$$\sum_{j=1}^{k} f(\zeta_j)$$

Generating function

- Distributions of quantities involving k particles $\zeta_{i_1}, \dots, \zeta_{i_k}$ are generated by **Fredholm determinants with k discontinuities** of the form

$$F(x_1, \dots, x_k; s_1, \dots, s_k) = \det \left(1 - \sum_{j=1}^{k} (1 - s_j) \chi_{(x_j, x_{j-1})} K^{\mathrm{Ai}} \chi_{(x_j, x_{j-1})} \right)$$

  where

  - $+\infty =: x_0 > x_1 > x_2 > \dots > x_k$
  - χ_A is the characteristic function of the set A
  - $s_1, \dots, s_k \in [0, 1]$

- Our goal: find a **Tracy-Widom type expression** for this Fredholm determinant $F(\vec{x}; \vec{s})$ for general k, in terms of Painlevé type ODEs

Main result

Theorem (Claeys-Doeraene, in progress)

$$F(\vec{x};\vec{s}) = \prod_{j=1}^{k} \exp\left(-\int_0^{+\infty} \xi u_j^2(\xi;\vec{x},\vec{s})d\xi\right)$$

and $u_1, \ldots, u_k$ solve the system of k ODEs

$$u_j''(x) = (x + x_j)u_j(x) + 2u_j(x)\sum_{i=1}^{k} u_i^2(x), \quad j = 1, \ldots, k$$

with asymptotic behaviour

$$u_j(x) \sim \sqrt{s_{j+1} - s_j}\ \mathrm{Ai}(x + x_j), \quad x \to +\infty$$

where we write $s_{k+1} = 1$.

Coupled Painlevé II equations

- Some remarks
  - system of equations depends on $x_1, \ldots, x_k$, relevant solutions depend on $s_1, \ldots, s_k$
  - for $k = 1$, the system reduces to the (shifted) Painlevé II equation

$$u_1'' = (x + x_1)u_1 + 2u_1^3, \quad u_1(x) \sim \sqrt{1 - s_1}\,\mathrm{Ai}\,(x + x_1), \quad x \to +\infty$$

$q(x; s) = u_1(x - x_1; s)$ is the Ablowitz-Segur/Hastings-McLeod solution, and we recover the standard Tracy-Widom formula

Coupled Painlevé II equations

- If $s_{j-1} < s_j$, $u_j(x)$ is real-valued for real x; if $s_{j-1} > s_j$, $u_j(x)$ is purely imaginary for real x
- if $s_1 < s_2 < \ldots < s_k$, all u_j's are real, and then

$$F(\vec{x}; \vec{s}) = \exp\left(- \int_0^{+\infty} \xi \|\vec{u}(\xi)\|^2 d\xi \right),$$

with

$$u_j''(x) = (x + x_j)u_j(x) + 2u_j(x)\|\vec{u}(x)\|^2, \quad j = 1, \ldots, k$$

with $\|.\|$ the 2-norm of the vector $\vec{u} = (u_1, \ldots, u_k)$.

Coupled Painlevé II equations

- If the s_j's are increasing, the system of coupled Painlevé II equations is a **traveling wave reduction of the defocusing vector nonlinear Schrödinger equation**

$$i\vec{q}_t = \vec{q}_{xx} - x\vec{q} - 2\vec{q}\|\vec{q}\|^2, \qquad \vec{q} = (q_1, \ldots, q_k)$$

  - for $k = 2$, this is called the Manakov system
  - If a solution $\vec{q}(x,t)$ has the form $q_j(x,t) = u_j(x)e^{-ix_j t}$, then $\vec{u}$ has to satisfy the system of coupled Painlevé II equations

$$u_j''(x) = (x + x_j)u_j(x) + 2u_j(x)\|\vec{u}(x)\|^2, \quad j = 1, \ldots, k$$

Probabilistic interpretation

- We have

$$F(x_1,\ldots,x_k;s_1,\ldots,s_k)=\mathbb{E}\left(\prod_{j=1}^{k}s_j^{\,n_{(x_j,x_{j-1})}}\right)$$

- **Joint distribution of k particles** $\zeta_{i_1}>\ldots>\zeta_{i_k}$:

$$\mathbb{P}\left(\zeta_{i_1}<x_1,\ldots,\zeta_{i_k}<x_k\right)$$
$$=\sum\frac{1}{j_1!\ldots j_k!}\,\frac{\partial^{j_1+\ldots+j_k}}{\partial s_1^{j_1}\ldots\partial s_k^{j_k}}F(x_1,\ldots,x_k;s_1,\ldots,s_k)\Big|_{\vec{s}=0}$$

where the sum is taken over all $j_1,\ldots,j_k$ such that

$$j_1<i_1,\quad j_1+j_2<i_2,\quad\ldots,\quad j_1+\ldots+j_k<i_k$$

Example 1: gap probability

- The probability of having **no particles in a finite interval** (x_2, x_1) is given by

$$F(x_1, x_2; 1, 0) = \exp\left(\int_0^{+\infty} \xi |u_1(\xi)|^2 d\xi\right) \exp\left(-\int_0^{+\infty} \xi |u_2(\xi)|^2 d\xi\right)$$

with

$$u_j''(x) = (x + x_j)u_j(x) + 2u_j(x) \sum_{i=1}^{k} u_i(x)^2, \quad j = 1, 2$$

and

$$u_{1,2}(x)^2 \sim \mp \mathrm{Ai}(x + x_{1,2})^2, \quad x \to +\infty$$

Example 2: thinning

- **Thinned Airy ensemble** is the process obtained by **removing each particle independently** with probability $s \in (0,1)$ *(Bohigas-de Carvalho-Pato '09, Bothner-Buckingham '17)*
  - interpret remaining particles as **observed**, removed as **unobserved**
  - interpolates between Airy ($s = 0$) and Poisson process ($s \to 1$)
- Distribution of the largest particle ζ_1, **conditioned on the position of the largest observed particle** ξ_1 being less than y can be expressed in terms of $F(x, y; 0, s)$

Example 2: thinning

- Distribution of the largest particle ζ_1, **conditioned on the position of the largest observed particle** ξ_1 being less than y can be expressed in terms of $F(x,y;0,s)$: if $y<x$,

$$\mathbb{P}\left(\zeta_1 < x | \xi_1 < y\right) = \frac{F(x,y;0,s)}{F^{\mathrm{TW}}(y;s)}$$
$$= \frac{1}{F^{\mathrm{TW}}(y;s)} \exp\left(-\int_0^{+\infty} \xi u_1(\xi)^2 d\xi\right) \exp\left(-\int_0^{+\infty} \xi u_2(\xi)^2 d\xi\right)$$

with

$$u_1(\xi)^2 \sim s\mathrm{Ai}(\xi+x)^2, \qquad \xi \to +\infty$$
$$u_2(\xi)^2 \sim (1-s)\mathrm{Ai}(\xi+y)^2, \qquad \xi \to +\infty$$

Example 3: first spacing

- Distribution of the **first spacing** $\zeta_1 - \zeta_2$:

$$\mathbb{P}\left(\zeta_1 - \zeta_2 > \sigma\right) = \int_{\mathbb{R}} v(\zeta + \sigma, \zeta) F^{\mathrm{TW}}(\zeta; 0) d\zeta$$

where

$$v(x_1, x_2) = \int_{\mathbb{R}} \xi \, \frac{-\partial^2}{\partial s_2 \partial x_1} \left(u_1^2(\xi; x_1, x_2; 0, s_2) + u_2^2(\xi; x_1, x_2; 0, s_2)\right)\Big|_{s_2 = 0} d\xi$$

  - different expressions in terms of Hastings-McLeod solution to Painlevé II equation and associated ψ-functions (fundamental solutions to the Lax pair for Painlevé II), obtained by *Bornemann-Forrester-Witte '12, Perret-Schehr '14*
  - related to limit distribution of first halting time for the Toda algorithm applied to random matrices (*Deift-Trogdon '16*)

Example 4: sum of k largest particles

- Distribution of the **sum** $\zeta_1 + \ldots + \zeta_k$: for $k = 2$,

$$\mathbb{P}\left(\zeta_1 + \zeta_2 < \sigma\right) = \int_{\mathbb{R}} v(\sigma - \zeta, \zeta) F^{\mathrm{TW}}(\zeta; 0) d\zeta$$

  - limit distribution of sum of k first components of a random partition w.r.t. **Plancherel measure**
  - limit distribution of maximal sum of lengths of k **disjoint increasing subsequences** of a random permutation *(Baik-Deift-Johansson '99, Borodin-Okounkov-Olshanski '00, Okounkov '00, Johansson '01)*

Sketch of proof

- Step 1: approximate the Airy ensemble by the **GUE for large size** n
  - Fredholm determinant as limit of **Hankel determinants**

$$F(\vec{x};\vec{s}) = \lim_{n\to+\infty} \frac{H_n(w_{n,\vec{x},\vec{s}})}{H_n(e^{-\frac{n}{2}x^2})}$$

where

$$H_n(w) = \det\left(\int_{\mathbb{R}} \xi^{j+k} w(\xi) d\xi\right)_{j,k=0}^{n-1}$$

and the weights are given by

$$w_{n,\vec{x},\vec{s}}(\xi) = s_j e^{-\frac{n}{2}\xi^2}, \quad \xi \in \left(2 + x_j n^{-2/3}, 2 + x_{j-1} n^{-2/3}\right), \; j = 1, \ldots, k+1$$

with

$$x_0 = +\infty, \quad x_{k+1} = -\infty, \quad s_{k+1} = 1$$

Sketch of proof

- Step 2: **differential identities** for $H_n(w_{n,\vec{x},\vec{s}})$ with respect to $\lambda_j = 2 + x_j n^{-2/3}$

$$\frac{\partial}{\partial \lambda_j} \ln H_n(w_{n,\vec{x},\vec{s}})$$
$$= (s_j - s_{j-1}) e^{-\frac{n}{2}\lambda_j^2} \frac{\kappa_{n-1}}{\kappa_n} \left(p_n'(\lambda_j) p_{n-1}(\lambda_j) - p_n(\lambda_j) p_{n-1}'(\lambda_j) \right)$$

  in terms of orthogonal polynomials p_k with respect to $w_{n,\vec{x},\vec{s}}$ and leading coefficients κ_k

Sketch of proof

- Step 3: **Riemann-Hilbert analysis** to obtain large n asymptotics for orthogonal polynomials
  - requires local approximation in the vicinity of the edge 2, where all discontinuities λ_j lie
  - local approximation requires a model Riemann-Hilbert problem (generalization of *Xu-Zhao '11*)
  - Lax pair associated to RH problem $\longrightarrow$ compatibility conditions lead to system of coupled Painlevé II equations
  - asymptotics for $p_n(\lambda_j)$ involve $u_j(0; \vec{x}, \vec{s})$

Sketch of proof

- Step 4: Substitute asymptotics for orthogonal polynomials in the differential identity and **integrate the differential identity**
  - starting point of integration $\lambda_1 = \lambda_2 = \ldots = \lambda_k = +\infty$ (explicit formula for Hankel determinant: Selberg integral)
  - first decrease λ_k, then λ_{k-1}, and so on
  - this explains appearance of integrals

$$\int_0^{+\infty} \xi u_k^2(\xi; \vec{x}, \vec{s}) d\xi + \ldots + \int_0^{+\infty} \xi u_1^2(\xi; \vec{x}, \vec{s}) d\xi$$

Outlook

- Various asymptotic regimes
  - large gap asymptotics for $F(\vec{x};\vec{s})$ where some of the x_j's tend to $-\infty$
- Other point processes
  - Bessel $\longrightarrow$ systems of coupled Painlevé III equations?
  - sine $\longrightarrow$ systems of coupled Painlevé V equations?