Airy and Pearcey Processes

Craig A. Tracy
UC Davis

Probability, Geometry and Integrable Systems
MSRI
December 2005

Random Matrix Models

Probability Space: $(\Omega, \Pr, \mathcal{F})$:

- Gaussian Orthogonal Ensemble (GOE, $\beta = 1$):
  - $\Omega = N \times N$ real symmetric matrices
  - $\Pr =$ "unique" measure that is invariant under orthogonal transformations and matrix elements are iid random variables. Explicitly,

$$\Pr(A \in \mathcal{B}) = \int_{\mathcal{B}} \mathrm{e}^{-\mathrm{tr}(A^2)}\, dA$$

- Gaussian Unitary Ensemble (GUE, $\beta = 2$)
  - $\Omega = N \times N$ (complex) hermitian matrices
  - $\Pr =$ "unique" measure that is invariant under unitary transformations and the independent real and imaginary matrix elements are iid random variables
- Gaussian Symplectic Ensemble (GSE, $\beta = 4$)

Limit Laws: $N \to \infty$

Eigenvalues, which are random variables, are real and with probability one they are distinct.

If $\lambda_{\mathbf{max}}(\mathbf{A})$ denotes the **largest eigenvalue** of the random matrix A, then for each of the **three Gaussian ensembles** we introduce the corresponding distribution function

$$F_{N,\beta}(t) := \mathrm{Pr}_{\beta}\left(\lambda_{\max} < t\right), \beta = 1, 2, 4.$$

The basic limit laws (Tracy-Widom) state that[a]

$$F_{\beta}(s) := \lim_{N\to\infty} F_{N,\beta}\left(2\sigma\sqrt{N} + \frac{\sigma s}{N^{1/6}}\right), \ \beta = 1, 2, 4,$$

exist and are given explicitly by

[a] Here σ is the standard deviation of the Gaussian distribution on the off-diagonal matrix elements.

$$\begin{aligned} F_2(s) &= \det\left(I - K_{\text{Airy}}\right) \\ &= \exp\left(-\int_s^\infty (x-s)q^2(x)\,dx\right) \end{aligned}$$

where

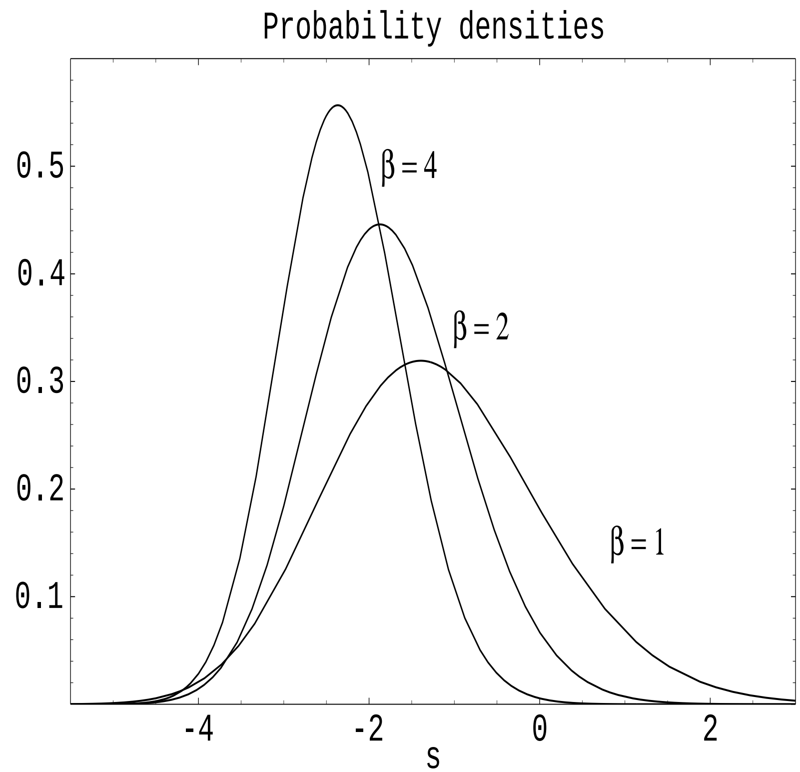
$$K_{\text{Airy}} \doteq \frac{\text{Ai}(x)\text{Ai}'(y) - \text{Ai}'(x)\text{Ai}(y)}{x-y}$$

acting on $L^2(s,\infty)$ (**Airy kernel**)

and q is the unique solution to the **Painlevé II equation**

$$q'' = sq + 2q^3, \quad q(s) \sim \text{Ai}(s) \text{ as } s \to \infty.$$

(Called the Hastings-McLeod solution.)



$$
\begin{aligned}
F_1(s) &= \exp\left(-\frac{1}{2}\int_s^\infty q(x)\,dx\right)\,(F_2(s))^{1/2}\,, \\
F_4(s/\sqrt{2}) &= \cosh\left(\frac{1}{2}\int_s^\infty q(x)\,dx\right)\,(F_2(s))^{1/2}\;.
\end{aligned}
$$

RMT Universality Theorems

Do limit laws depend upon the underlying Gaussian assumption on the probability measure?

To investigate this for unitarily invariant measures ($\beta = 2$):

$$\exp\left(-\mathrm{tr}(A^2)\right) \to \exp\left(-\mathrm{tr}(V(A))\right).$$

Bleher & Its chose

$$V(A) = gA^4 - A^2, g > 0,$$

and subsequently a large class of potentials V was analyzed by Deift/Kriecherbauer/McLaughlin/Venakides/Zhou.

Requires proving new **Plancherel-Rotac**h type formulas for **nonclassical** orthogonal polynomials. The proofs use **Riemann-Hilbert methods**. Generic behavior is GUE. However, by tuning V new universality classes will emerge.

Universality theorems for **orthogonal & symplectic** invariant measures:

- Stojanovic analyzed the quartic potential.
- Deift & Gioev considered a class of polynomial potentials whose equilibrium measure is supported on a single interval. Their starting point is Widom's representation of the correlation kernels for the $\beta = 1, 4$ cases in terms of the unitary ($\beta = 2$) correlation kernel plus a correction.

All these results can be summarized by

Generic edge behavior is described by Airy kernel

Noninvariant RMT Measures

Soshnikov proved that for real symmetric Wigner matrices[a] (complex hermitian Wigner matrices) the limiting distribution of the largest eigenvalue is F_1 (respectively, F_2). The **significance** of this result is that nongaussian Wigner measures lie **outside** the **"integrable class"** (e.g. there are no Fredholm determinant representations for the distribution functions) yet the limit laws are the same as in the integrable cases.

[a] A symmetric Wigner matrix is a random matrix whose entries on and above the main diagonal are independent and identically distributed random variables with distribution function F. Soshnikov assumes all odd moments vanish and even moments are finite satisfying a Gaussian type growth condition.

Next Largest, Next-next Largest, etc. Eigenvalue Distributions

Let $\hat{\lambda}_k^{(n)}$ denote the rescaled k^{th} eigenvalue measured from the EDGE OF THE SPECTRUM. We are interested in

$$F_\beta(s,k) = \lim_{n\to\infty} \mathrm{Pr}_{n\beta}\left(\hat{\lambda}_k^{(n)} \le s\right), \ \beta = 1,2,4.$$

($k = 1$ is the case of largest eigenvalue.) Define

$$D_2(s,\lambda) = \det\left(I - \lambda K_{\text{Airy}}\right), \ 0 \le \lambda \le 1,$$

then

$$F_2(s,k+1) - F_2(s,k) = \frac{(-1)^k}{k\,!}\,\frac{\partial^k}{\partial\lambda^k}\, D_2(s,\lambda)\Big|_{\lambda=1^-} \ k \ge 0, F_2(s,0) := 0$$

We have a Painlevé representation for $D(s,1)$.

What is the Painlevé representation for $D(s,\lambda)$?

The answer (TW) is remarkably simple:

$$D_2(s,\lambda) = \exp\left[-\int_s^\infty (x-s)\, q^2(x,\lambda) d\, x\right]$$

where $q(x,\lambda)$ satisfies the same Painlevé II equation but with boundary condition

$$q(x,\lambda) \sim \sqrt{\lambda}\,\mathrm{Ai}(x),\ x \to \infty.$$

Thus $F_2(s,k)$ are expressible in terms of

$$q(s,1),\ \frac{\partial q}{\partial \lambda}(s,1),\ \ldots,\ \frac{\partial^k q}{\partial \lambda^k}(s,1)$$

Will same hold for orthogonal and symplectic ensembles?

i.e. Take $\lambda = 1$ results and simply make replacement

$$q(x) = q(x,1) \to q(x,\lambda)\,?$$

Let

$$D_1(s,\lambda) := \lim_{\text{Edge Scaling}} \det\left(I - \lambda K_{n,\text{GOE}}\right) = \det{}_2\left(I - \lambda K_{1,\text{Airy}}\right)$$

$$D_4(s,\lambda) := \lim_{\text{Edge Scaling}} \det\left(I - \lambda K_{n,\text{GSE}}\right) = \det\left(I - \lambda K_{4,\text{Airy}}\right)$$

REMARKS:

1. Convergence for $\beta = 4$ is in trace-class norm. For $\beta = 1$ convergence is to the regularized determinant, $\det_2$, in the Hilbert-Schmidt norm (TW).

2.

$$F_\beta(s,k+1) = F_\beta(s,k) + \frac{(-1)^k}{k!}\,\frac{\partial^k}{\partial\lambda^k}\,D_\beta^{1/2}(s,\lambda)\Big|_{\lambda=1},\ \beta = 1,4,$$

with $F_\beta(s,0) := 0$.

Painlevé Representations for D_1 and D_4

Momar Dieng proved the following:

$$D_4(s,\lambda) = D_2(s,\lambda)\cosh^2\left(\frac{\mu(s,\lambda)}{2}\right)$$

$$D_1(s,\lambda) = D_2(s,\tilde{\lambda})\,\frac{\lambda - 1 - \cosh\mu(s,\tilde{\lambda}) + \sqrt{\tilde{\lambda}}\,\sinh\mu(s,\tilde{\lambda})}{\lambda - 2}$$

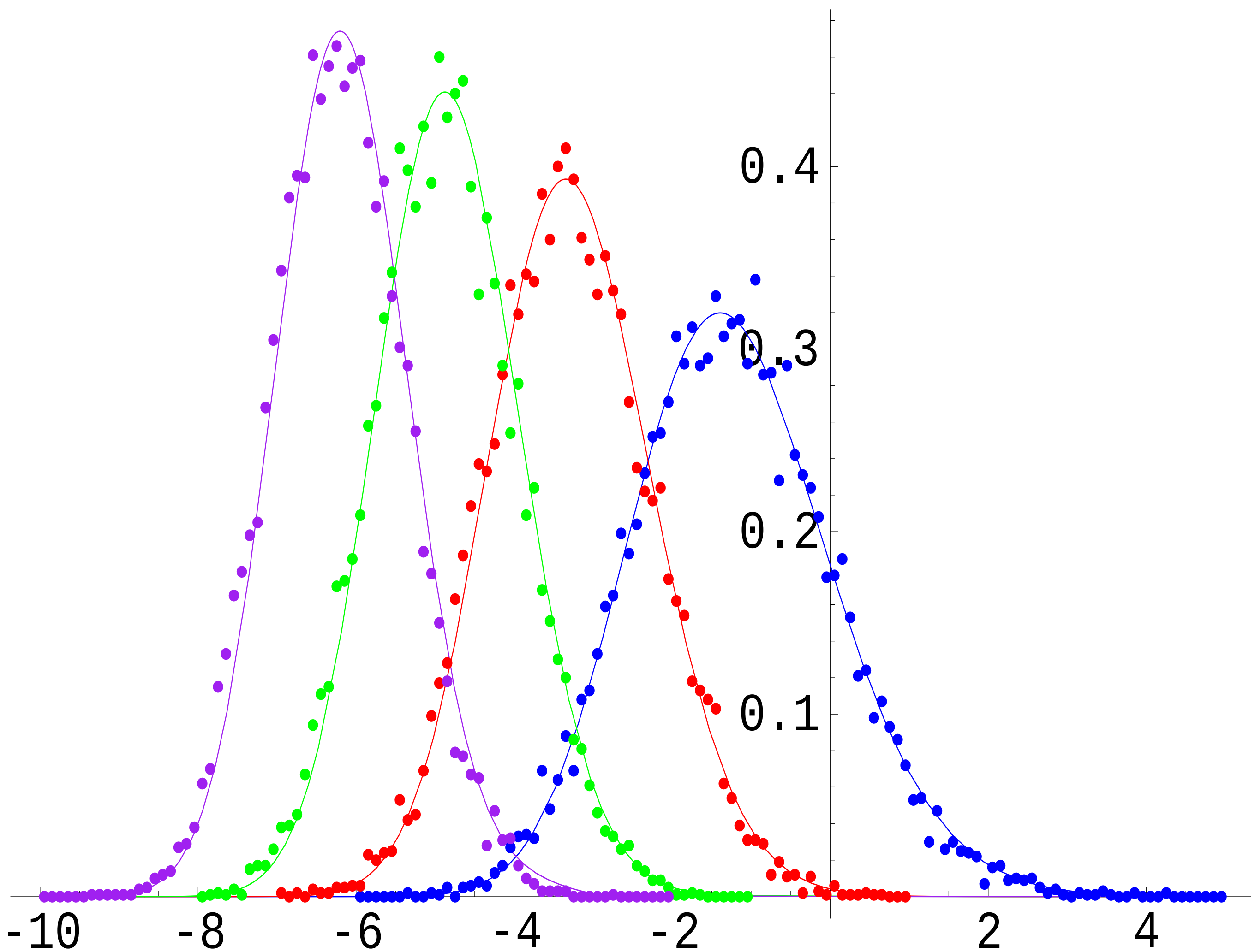
with

$$\mu(s,\lambda) := \int_s^\infty q(x,\lambda)\,dx \qquad \text{and} \qquad \tilde{\lambda} := 2\lambda - \lambda^2$$

In the symplectic case the prescription $q(x,1) \to q(x,\lambda)$ is valid; whereas for the orthogonal case, a NEW FORMULA appears.

Note, in the orthogonal case, that D_2 and q are evaluated at $\tilde{\lambda}$.

10^4 REALIZATIONS OF $10^3 \times 10^3$ GOE MATRICES



Appearance of Limit Laws Outside of RMT

Major breakthrough when Baik, Deift, Johansson proved that the limiting distribution of the **length of the longest increasing subsequence** in a random permutation is F_2.

Random permutation of $\{1, 2, \ldots, 10\}$:

$$\sigma = \{\mathbf{3}, 7, 10, \mathbf{5}, 9, \mathbf{6}, \mathbf{8}, 1, 4, 2\}, \;\; \ell_{10}(\sigma) = 4$$

Patience Sorting Algorithm (Aldous, Diaconis)

	2		
	4	6	
1	5	9	
3	7	10	8

BDJ Theorem:

$$\lim_{n\to\infty} \Pr\left(\frac{\ell_n - 2\sqrt{n}}{n^{1/6}} \le x\right) = F_2(x)$$

and with convergence of moments, e.g.

$$\begin{aligned} E(\ell_n) &= 2\sqrt{n} + \int_{-\infty}^{\infty} x f_2(x)\, dx \; n^{1/6} + \mathrm{o}(n^{1/6}) \\ &= 2\sqrt{n} - 1.7710868074\, n^{1/6} + \mathrm{o}(n^{1/6}) \end{aligned}$$

A simulation with 100 trials for $n = 10^5$ gives an average number of piles per trial

621.96

which should be compared with the asymptotic expected value

620.389

The $2\sqrt{n}$ term alone gives **632.456**.

Key Points in the Proof of the BDJ Theorem

- Gessel proved (uses RSK and Jacobi-Trudi identity for s_λ)

$$\sum_{n\geq 0} \Pr(\ell_n \leq k)\, \frac{t^{2n}}{n!} = \det(T_k(\varphi))$$

  where $T_k(\varphi)$ is a $k \times k$ Toeplitz matrix with symbol $\varphi(z) = e^{t(z+1/z)}$.

- Use Case/Geronimo-Borodin/Okounkov identity that relates a Toeplitz determinant to a Fredholm determinant of an operator on $\ell^2(\{0,1,\ldots\})$

$$\sum_{n\geq 0} \Pr(\ell_n \leq k)\, \frac{t^{2n}}{n!} = \det(I - K_k)$$

Specifically, $\varphi = \varphi_+ \varphi_-$, then

$$K_k(i,j) = \sum_{\ell \geq 0} (\varphi_-/\varphi_+)_{k+i+\ell+1} \, (\varphi_+/\varphi_-)_{-k-j-\ell-1}$$

- Show $K_k \to K_{\text{Airy}}$ in trace class norm: Use saddle point method on Fourier coefficients appearing in CGBO identity. Find

Nontrivial limit only when two saddle points coalesce

Airy function generic behavior

$$K_{\text{Airy}}(x,y) = \int_0^\infty \text{Ai}(x+z)\text{Ai}(y+z)\, dz$$

The one-third scaling is a direct result of this coalescence—viz. the cubic power in the Airy function integral:

$$\text{Ai}(z) = \frac{1}{2\pi i} \int_C e^{\xi^3/3 - z\xi} \, d\xi$$

- A dePoissonization lemma to get limit theorem.

The BDJ Theorem resulted in a burst of activity relating the distribution functions of RMT to problems in combinatorics, representation theory of the symmetric group, growth processes and determinantal random fields

Cast of Players

M. Adler, D. Aldous, J. Baik, P. Bleher, T. Bodineau, A. Borodin, P. Deift, P. Diaconis, P. Ferrari, P. Forrester, J. Gravner, T. Imamura, A. Its, K. Johannson, J. Martin, K. McLaughlin, N. O'Connell, A. Okounkov, G. Olshanski, M. Prähoffer, E. Rains, N. Reshetikhin, T. Sasamoto, A.Soshnikov, H. Spohn, C. Tracy, P. van Moerbeke, H. Widom, ...

From Brownian Motion to the Airy Process

$$t \longrightarrow B_t$$

is a Gaussian process: Fix $t_1 < t_2 < \cdots < t_m$,

$$(B_{t_1}, B_{t_2}, \ldots, B_{t_m})$$

is a multivariate Gaussian, e.g.

$$\Pr\left(B_t \leq x\right) = \Phi(x)$$

The **Airy Process** (Prähoffer & Spohn, Johansson)

$$t \longrightarrow A_t$$

is the process underlying F_2, e.g.

$$\Pr\left(A_t \leq x\right) = F_2(x)$$

Dyson BM

GUE initial conditions and independent matrix elements of a Hermitian matrix H independently undergo Ornstein-Uhlenbeck diffusion

$$t \to H_t.$$

Transition density

$$p(H, H'; t_2 - t_1) := \exp\left(-\frac{\mathrm{tr}(H - qH')^2}{1 - q^2}\right)/Z$$

$q = \mathrm{e}^{t_1 - t_2} < 1.$

As $t_2 \to \infty$, measure approaches GUE measure.

Each eigenvalue feels an electric field

$$E(x_i) = \sum_{i \neq j} \frac{1}{x_i - x_j} - x_i$$

Many times: $t_1 < t_2 < \cdots < t_m$

With GUE initial conditions the density for H_t in neighborhood of H_k at time $t = t_k$ is

$$e^{-\mathrm{tr}(H_1^2)} \prod_{j=2}^{m} p(H_j, H_{j-1}, t_j - t_{j-1})$$

Use HCIZ integral to integrate out unitary parts to obtain determinantal measure on eigenvalues $x_j(t)$

Focus on the **largest eigenvalue**

$$t \longrightarrow x_{\mathrm{max}}(t)$$

In **edge scaling limit** obtain

$$t \longrightarrow A_t$$

Airy Process

Defined by the distribution functions

$$\Pr(A_{t_1} \le \xi_1, \ldots, A_{t_m} \le \xi_m)$$

Probability expressed as a Fredholm determinant of extended Airy kernel, an $m \times m$ matrix kernel. Entries $L_{ij}(x, y)$ given by

$$\int_0^\infty e^{-z\,(t_i - t_j)}\,\mathrm{Ai}(x+z)\,\mathrm{Ai}(y+z)\,dz, \quad i \ge j,$$

$$-\int_{-\infty}^0 e^{-z\,(t_i - t_j)}\,\mathrm{Ai}(x+z)\,\mathrm{Ai}(y+z)\,dz, \; i < j$$

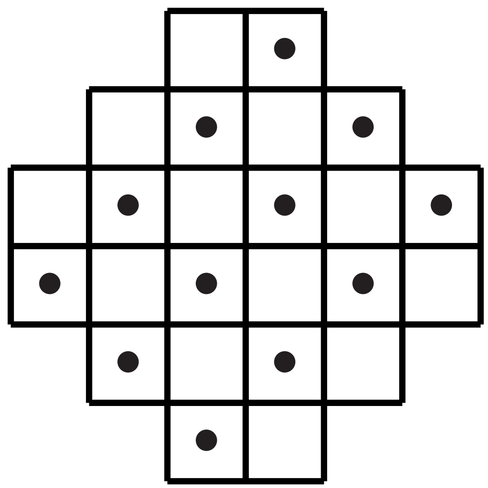
$$K_{ij}(x, y) = L_{ij}(x, y)\,\chi_{(\xi_j, \infty)}(y).$$

Probability equals $\det\,(I - K)$.

Aztec Diamond A_n

Elkies, Kuperberg, Larsen, Propp, . . .

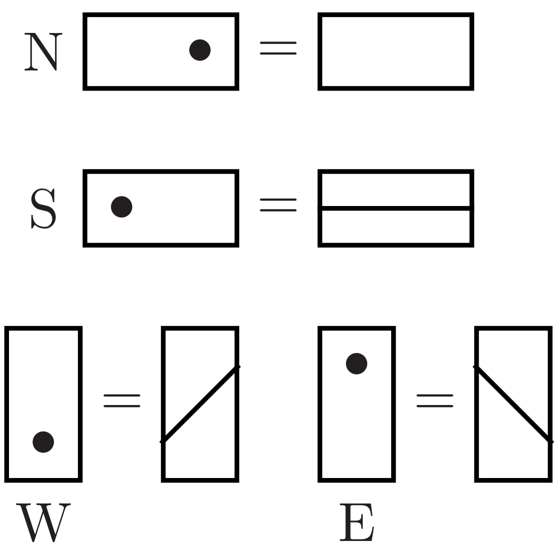
A_n: Union of all lattice squares that lie inside $\{|x| + |y| \le n + 1\}$.



A_3 with checkerboarding

‡ Tile with 2×1 and 1×2 dominoes.

‡ Checkerboard lattice. Four types of tiles: N, S, E, W.



‡ $X_n(t)$ is top line.

‡ The Northern Polar Region (NPR) is exactly the part of the domino tiling that lies above $X_n(t)$, and consists only of N-dominoes.

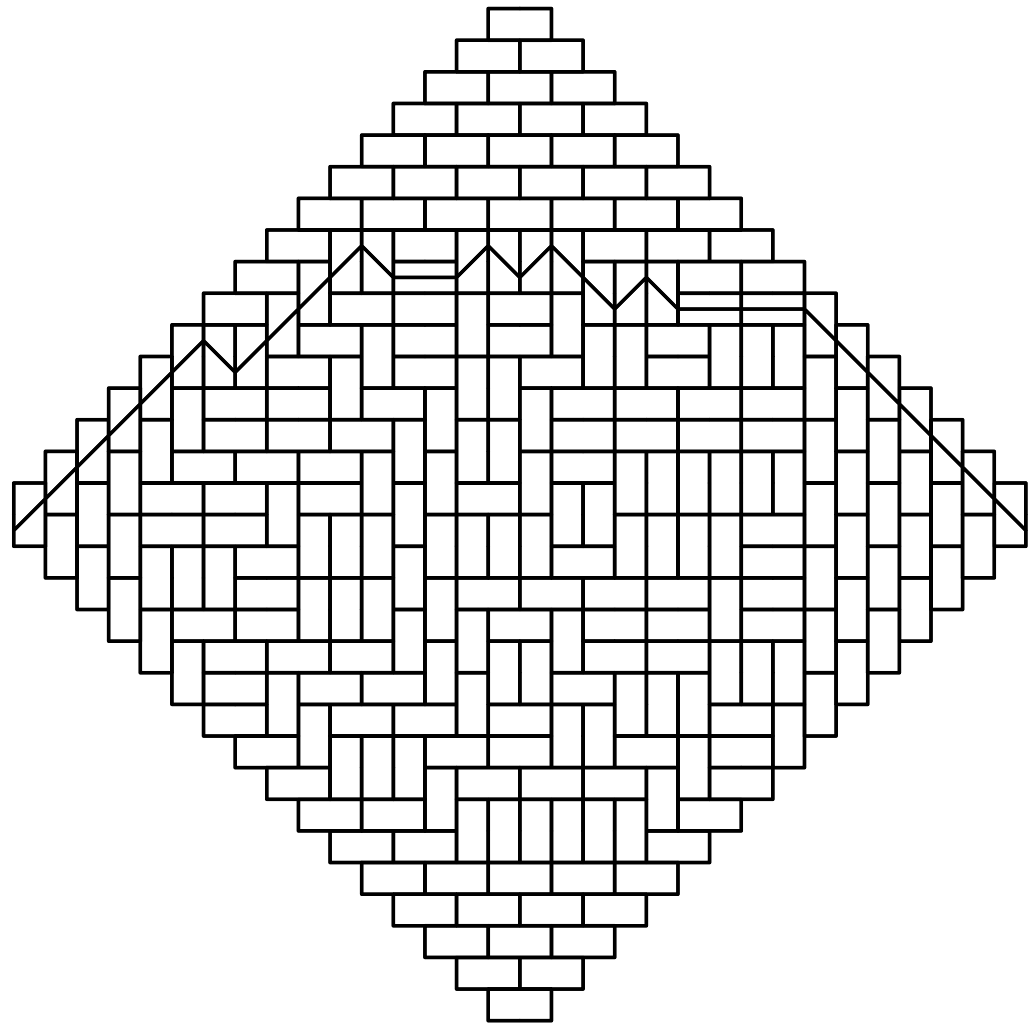
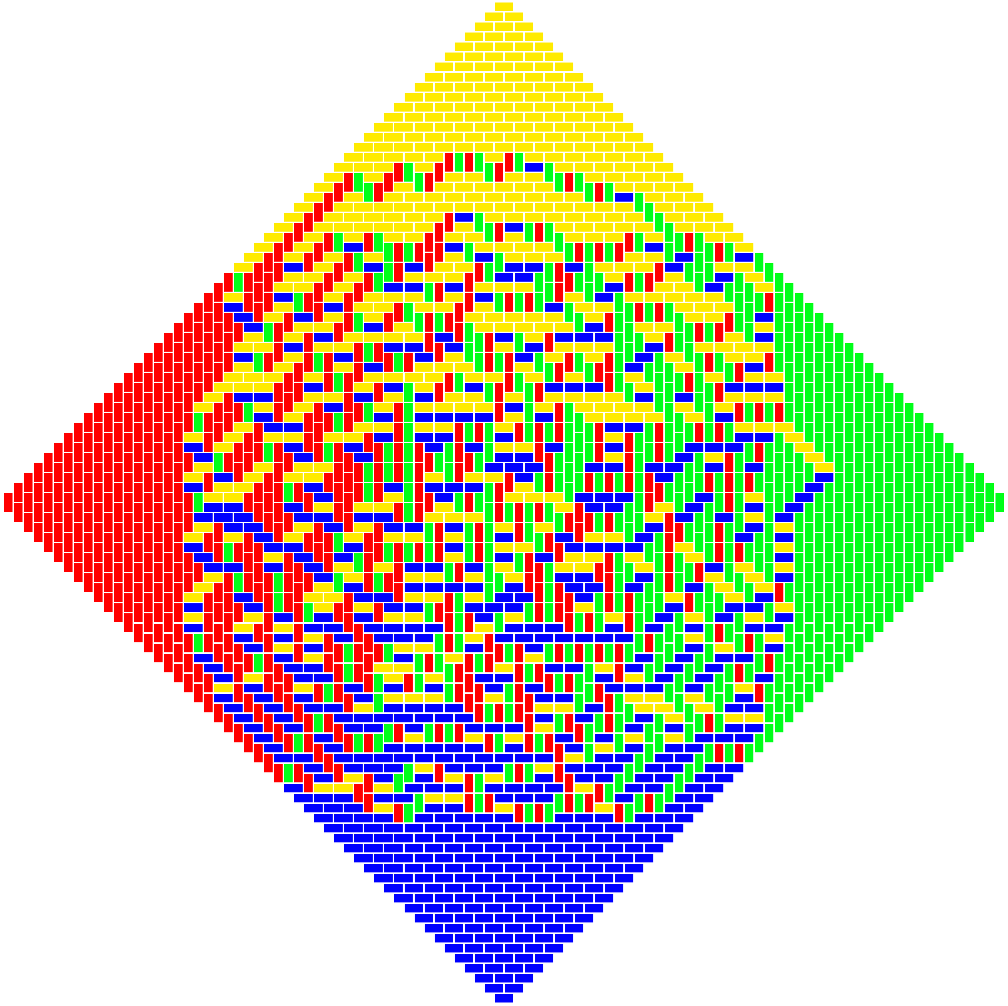


Figure 1: Top Curve $X_n(t)$ [Johansson]



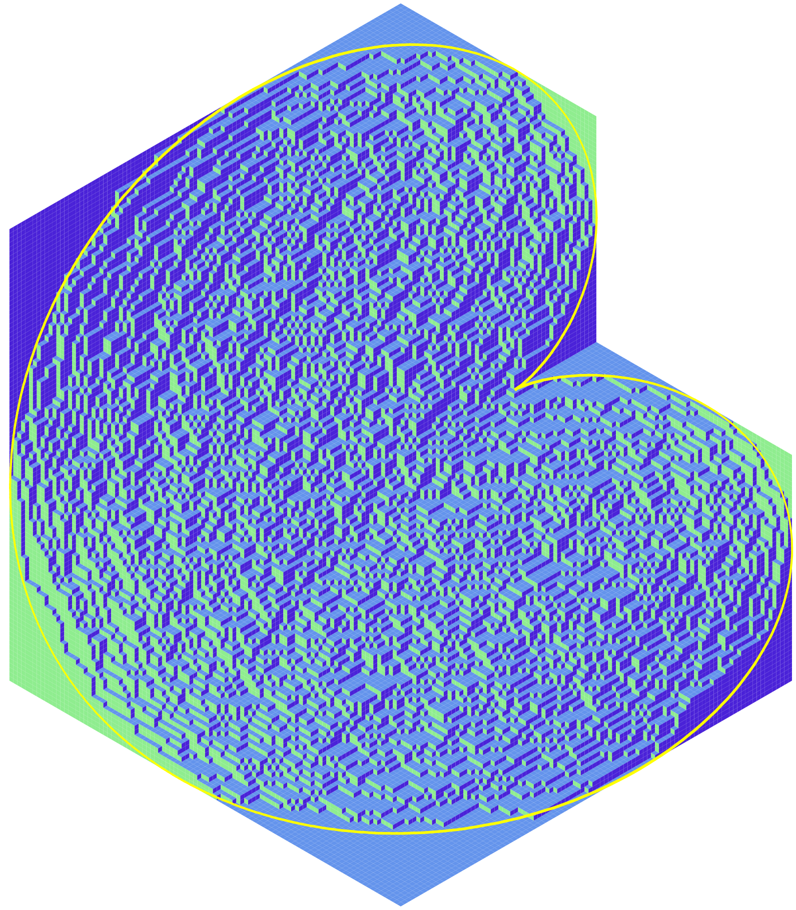
Random Tilings Research Group

Theorem (Johansson) Let $X_n(t)$ be the NPR-boundary process and A_t the Airy process, then

$$\frac{X_n(2^{-1/6}n^{2/3}t) - n/\sqrt{2}}{2^{-5/6}n^{1/3}} \to A_t - t^2,$$

as $n \to \infty$, in the sense of convergence of finite-dimensional distributions.

Remark: Johansson derives an extended kernel for the distribution functions and shows convergence to the extended Airy kernel.



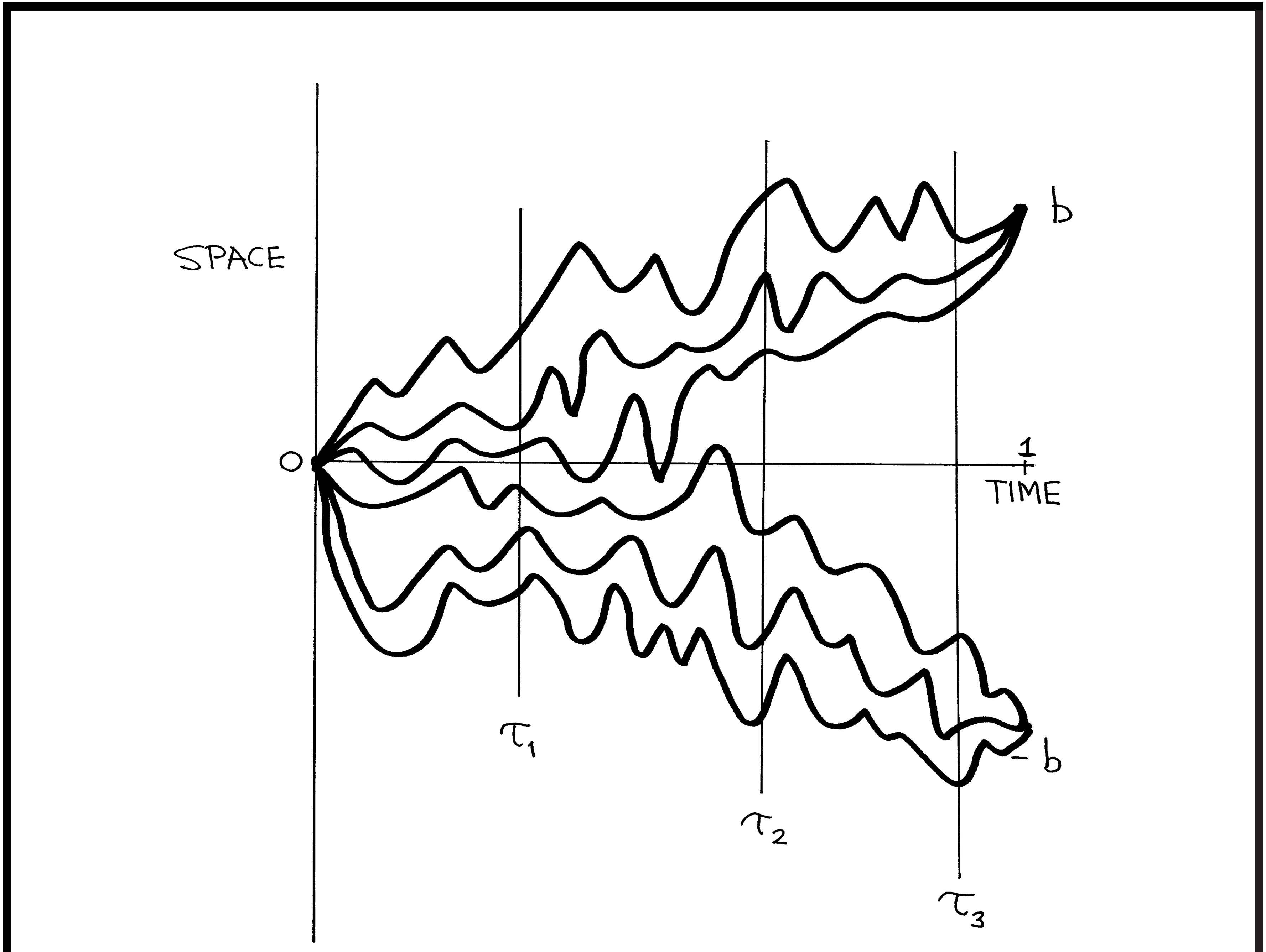
Random 3D Young Diagram
Okounkov and Reshetikhin

Pearcey Process

Brézin & Hikami (1998), Aptekarev, Bleher, Kuijlaars (2004–05), Okounkov & Reshetikhin (2003–05), Tracy & Widom (2004–05), Adler & van Moerbeke(2005)

Airy functions (fold singularity) $\longrightarrow$
Pearcey functions (cusp singularity)

Saddle point analysis: Airy is coalescence of two saddle points whereas Pearcey arises from the coalescence of three saddle points


SPACE
b
0
1
TIME
τ_1
-b
τ_2
τ_3

Take $b = \sqrt{n}$, $\tau_k \to \tau_c + \tau_k/\sqrt{n}$, then in the limit $n \to \infty$, the operator K converges to the to K^{Pearcey} whose kernel, **extended Pearcey kernel**, has i, j entry

$$-\frac{1}{4\pi^2} \int_{\mathcal{C}} \int_{-i\infty}^{i\infty} \mathrm{e}^{-s^4/4+\tau_j s^2/2-ys+t^4/4-\tau_i t^2/2+xt} \, \frac{dsdt}{s-t}$$

The t contour $\mathcal{C}$ consists of the rays from $\pm\infty e^{i\pi/4}$ to 0 and the rays from 0 to $\pm e^{-i\pi/4}$. For $m = 1$ and $\tau_1 = 0$ this reduces to the Pearcey kernel of Brézin & Hikami.

Open Problem: Prove the existence of an actual limiting process consisting of infinitely many paths, with correlation functions and spacing distributions described by the extended Pearcey kernel. For each fixed time that there is a limiting random point field follows from a theorem of Lenard. But the construction of the time-dependent random point field is still open.

HENRY