PIERRE VAN MOERBEKE

Random matrices and permutations, matrix integrals and integrable systems

Séminaire N. Bourbaki, 1999-2000, exp. n° 879, p. 411-433.
<http://www.numdam.org/item?id=SB_1999-2000__42__411_0>

Séminaire BOURBAKI Juin 2000
52^{e} année, 1999-2000, n^{o} 879, p. 411 à 433

RANDOM MATRICES AND PERMUTATIONS, MATRIX INTEGRALS AND INTEGRABLE SYSTEMS

by **Pierre VAN MOERBEKE**

Contents

The support of a National Science Foundation grant # DMS-98-4-50790, a Nato, a FNRS and a Francqui Foundation grant is gratefully acknowledged.

0. INTRODUCTION

The purpose of this lecture is to give a survey of recent interactions between the theory of random matrices, the theory of random permutations and so-called integrable models. The latter are described either by integrals over tangent spaces to symmetric spaces, or by integrals over classical groups.

Going back in time, t'Hooft was led, in the 70's, to so-called matrix models in order to understand the behavior of quantum gauge theories with gauge group $SU(N)$ when N gets large. In their pioneering work, Bessis, Itzykson and Zuber [8] considered the integral $Z_N^{(\varepsilon)}$ taken over the space $\mathcal{H}_N$ of Hermitian matrices of size N, whose log has the following expansion:

$$\log Z_N^{(\varepsilon)} = \log \frac{\int_{\mathcal{H}_N} dM\, e^{-\frac{1}{2}trM^2 - \frac{\varepsilon}{N}trM^4}}{\int_{\mathcal{H}_N} dM\, e^{-\frac{1}{2}trM^2}} = \sum_{g=0}^{\infty} N^{2-2g} \sum_{k=0}^{\infty} (-\varepsilon)^k W(g,k)$$

where $W(g,n)$ is the number of graphs with k vertices drawn on a surface of genus g, or in a dual language, the number of ways to cover a surface of genus g with k squares. Replacing $\frac{\varepsilon}{N}$ tr M^4 with $\sum_{i\geq 1} t_i \operatorname{tr} M^i$ leads to coverings of surfaces of genus g by n-gons of various n. The sequence of such integrals over N is a solution of the standard Toda lattice equations.

Two-matrix integrals (Chadha-Mahoux-Mehta [11], Itzykson-Zuber [20])

$$\int\int_{\mathcal{H}_N \times \mathcal{H}_N} dM_1 dM_2 e^{-tr(M_1^2 + M_2^2 - cM_1M_2 + \alpha M_1^4 + \beta M_2^4)}$$

provide a solution to the 2d-Toda lattice. They describe an Ising model on a random lattice, with spin $\sigma = \pm 1$ at each site with nearest neighbor interaction; the interaction between two sites depends on whether the spins are equal or not.

Kontsevich [25] proves that $Z_N(t)$ is a τ-function for the KdV equation for N large, with expansion (set $t_i = -c_i \operatorname{tr} Z^{-i}$ for appropriate integer c_i's)

$$\begin{aligned}
\log Z_N(t) &= \log \frac{\int_{\mathcal{H}_N} dX e^{-\frac{1}{2}tr\, X^2 Z + \frac{1}{6}tr\, X^3}}{\int_{\mathcal{H}_N} dX\, e^{-\frac{1}{2}tr\, X^2 Z}} \\
&= \sum_{\Gamma \in G_N} \frac{(\frac{1}{2})\#\text{vertices in }\Gamma}{\#\operatorname{Aut}\Gamma} \prod_{\text{ribbons }\alpha,\beta} \frac{2}{z_\alpha + z_\beta} \\
&= \sum_{n\geq 1, d_1,\ldots,d_n \geq 0} \frac{t_{2d_1+1}\ldots t_{2d_n+1}}{n!} \int_{\bar{\mathcal{M}}_{g,n}} \prod_{i=1}^{n} c_1(\mathcal{L}_i)^{d_i},
\end{aligned}$$

where G_N are all non-equivalent ribbon graphs Γ with N distinct loops obtained by picking vertices from which emerge 3 ribbons, interconnecting them and associating a z_α with each edge $(1 \leq \alpha \leq N)$. Each ribbon has two edges, one going with some z_α and another going with some z_β $(1 \leq \alpha, \beta \leq N)$. The product in the formula above is taken over all such ribbons of the graph. The second formula involves Chern classes $c_1(\mathcal{L}_i)$ of line bundles on the Deligne-Mumford smooth compactification $\overline{\mathcal{M}}_{g,n}$ of the moduli space $\mathcal{M}_{g,n}$ of smooth Riemann surfaces of genus g with n distinct marked points $x_1, \ldots, x_n$; the fibers of the line bundle are given by $T^*_{x_i}\mathcal{C}$ at each point $(\mathcal{C}, x_1, \ldots, x_n) \in \overline{\mathcal{M}}_{g,n}$.

Recently, other integrals have arisen, like

$$\int_{O(n)} e^{xtrM} dM, \quad \int_{U(n)} e^{\sqrt{x}\,\mathrm{tr}(M+\bar{M})} dM, \quad \int_{U(n)} \det(I+M)^k e^{-x\ tr\bar{M}} dM,$$

whose expansion in x relate to problems in combinatorics. The integrals over the space of Hermitian, symmetric and symplectic matrices, namely

$$\int_{\mathcal{H}_n,\ \mathcal{S}_n \text{ or } \mathcal{T}_{2n}} e^{-trV(M)} dM,$$

play a prominent role in theory of random matrices. These two sets of integrals provide the main thrust of this lecture. Sequences of such integrals (in n) provide solutions to integrable lattices. For a more detailed overview on these questions, see the MSRI lectures [39]. The last section contains a table listing these connections between matrix integrals, moment matrices and integrable lattices.

1. LARGEST INCREASING SEQUENCES IN RANDOM PERMUTATIONS

Let S_N be the group of permutations π_N, equipped with the uniform probability distribution:

(1.0.1) $$P(\pi_N) = 1/N!.$$

An *increasing subsequence* of $\pi_N \in S_N$ is a sequence $1 \le j_1 < \ldots < j_k \le N$, such that $\pi(j_1) < \ldots < \pi(j_k)$. Define

(1.0.2) $$L_N(\pi_N) = \text{ length of the longest increasing subsequence of } \pi_N\ .$$

Problem. — Find the probability $P(L_N \le n)$.

Examples. — For $\pi_7 = (\underline{3}, 1, \underline{4}, 2, \underline{6}, \underline{7}, 5)$, we have $L(\pi_7) = 4$. For $\pi_5 = (5, \underline{1}, \underline{4}, 3, 2)$, we have $L(\pi_5) = 2$.

1.1. Robinson-Schensted-Knuth correspondence and symmetric functions

I give here a very brief summary of well-known, but necessary facts to approach the problem:

– A *Young diagram* λ is a finite sequence of non-increasing, non-negative integers $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_\ell \geq 0$; also called a *partition* of $n = |\lambda| := \lambda_1 + \ldots + \lambda_\ell$, with $|\lambda|$ being the weight. It can be represented by a diagram, having λ_1 boxes in the first row, λ_2 boxes in the second row, etc., all aligned to the left. A *dual Young diagram* $\hat{\lambda} = (\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots)$ is the diagram obtained by flipping the diagram λ about its diagonal; thus $\hat{\lambda}_1 = \ell$ =length of first column of λ.

– A *Young tableau* of shape λ is an array of positive integers a_{ij} (at place (i,j) in the Young diagram) placed in the Young diagram λ, which are non-decreasing from left to right *and* strictly increasing from top to bottom.

– A *standard Young tableau* of shape λ is an array of integers $1, \ldots, n$ placed in the Young diagram, which are strictly increasing from left to right *and* from top to bottom. The number of Young tableaux of a given shape $\lambda = (\lambda_1 \geq \ldots \geq \lambda_m)$ is given by a number of formulae (for the Schur polynomial s_λ, see below)[(1)]

$$\begin{aligned}
f^\lambda &= \#\{\text{standard tableaux of shape } \lambda\} \\
&= \text{coefficient of } x_1 x_2 \ldots x_n \text{ in } s_\lambda(x) \\
&= \frac{|\lambda|!}{\prod_{\text{all } i,j} h_{ij}^\lambda} = |\lambda|! \det\left(\frac{1}{(\lambda_i - i + j)!}\right) \\
&= |\lambda|! \prod_{1 \leq i < j \leq m} (h_i - h_j) \prod_1^m \frac{1}{h_i!}, \quad \text{with } h_i := \lambda_i - i + m,\ m := \hat{\lambda}_1. \qquad (1.1.1)
\end{aligned}$$

– The *Schur polynomial* s_λ associated with a Young diagram λ is a symmetric function in the variables $x_1, x_2, \ldots$ (finite or infinite), defined by

$$s_\lambda(x_1, x_2, \ldots) := \sum_{\{a_{ij}\} \text{ tableaux of } \lambda} \prod_{ij} x_{a_{ij}}. \tag{1.1.2}$$

– The linear *space* Λ_n *of symmetric polynomials* in $x_1, \ldots, x_n$ with rational coefficients comes equipped with an inner product, which can also be expressed as an integral over the unitary group $U(n)$ for Haar measure dM:

$$\begin{aligned}
\langle f, g \rangle &= \frac{1}{n!} \int_{(S_1)^n} f(z_1, \ldots, z_n) g(\bar{z}_1, \ldots, \bar{z}_n) \prod_{1 \leq k < \ell \leq n} |z_k - z_\ell|^2 \prod_1^n \frac{dz_k}{2\pi i z_k} \\
&= \int_{U(n)} f(M) g(\bar{M}) dM. \qquad (1.1.3)
\end{aligned}$$

– An *orthonormal basis of the space* Λ_n is given by the Schur polynomials above $s_\lambda(x_1, \ldots, x_n)$, in which the numbers a_{ij} are restricted to $1, \ldots, n$; therefore we have the following "*Fourier series*":

$$f(x_1, \ldots, x_n) = \sum_{\substack{\lambda \text{ with} \\ \hat{\lambda}_1 \leq n}} \langle f, s_\lambda \rangle\, s_\lambda(x_1, \ldots, x_n), \text{ with } \langle s_\lambda, s_{\lambda'} \rangle = \delta_{\lambda\lambda'}. \tag{1.1.4}$$

[(1)] $h_{ij}^\lambda := \lambda_i + \hat{\lambda}_j - i - j + 1$ is the *hook length* of the i,jth box in the Young diagram.

In particular, one computes the following Fourier series:

$$(x_1 + \cdots + x_n)^k = \sum_{\substack{|\lambda|=k \\ \hat{\lambda}_1 \leq n}} f^\lambda s_\lambda. \tag{1.1.5}$$

If $\lambda = (\lambda_1 \geq ... \geq \lambda_\ell)$, with[2] $\hat{\lambda}_1 = \ell > n$, then obviously $s_\lambda = 0$.
– *Robinson-Schensted-Knuth correspondence*: There is a 1-1 correspondence

$$S_n \longrightarrow \left\{ \begin{array}{l} (P,Q), \text{two standard Young} \\ \text{tableaux from } 1,\ldots,n, \text{ where} \\ P \text{ and } Q \text{ have the same shape} \end{array} \right\}. \tag{1.1.6}$$

Given a permutation $i_1, ..., i_n$, the correspondence constructs two standard Young tableaux P, Q having the same shape λ. This construction is inductive. Namely, having obtained two equally shaped Young diagrams P_k, Q_k from $i_1, ..., i_k$, with the numbers $(i_1, ..., i_k)$ in the boxes of P_k and the numbers $(1, ..., k)$ in the boxes of Q_k, one creates a new diagram Q_{k+1}, by putting the *next number* i_{k+1} *in the first row of* P, according to the following rule:

(i) if $i_{k+1} \geq$ all numbers appearing in the first row of P_k, then one creates a new box with i_{k+1} in that box to the right of the first column,

(ii) if not, place i_{k+1} in the box (of the first row) with the smallest higher number. That number then gets pushed down to the second row of P_k according to the rule (i) or (ii), as if the first row had been removed.

The diagram Q is a bookkeeping device; namely, add a box (with the number $k+1$ in it) to Q_k exactly at the place, where the new box has been added to P_k. This produces a new diagram Q_{k+1} of same shape as P_{k+1}.

The inverse of this map is constructed essentially by reversing the steps above.

Example. — $\pi = (5, 1, 4, 3, 2) \in S_5$,

$$\begin{array}{l} 5 \end{array} \qquad \begin{array}{l} 1 \\ 5 \end{array} \qquad \begin{array}{ll} 1 & 4 \\ 5 & \end{array} \qquad \begin{array}{ll} 1 & 3 \\ 4 & \\ 5 & \end{array} \qquad \begin{array}{ll} 1 & 2 \\ 3 & \\ 4 & \\ 5 & \end{array}$$

$$\begin{array}{l} 1 \end{array} \qquad \begin{array}{l} 1 \\ 2 \end{array} \qquad \begin{array}{ll} 1 & 3 \\ 2 & \end{array} \qquad \begin{array}{ll} 1 & 3 \\ 2 & \\ 4 & \end{array} \qquad \begin{array}{ll} 1 & 3 \\ 2 & \\ 4 & \\ 5 & \end{array}$$

[2] Remember, from the definition of the dual Young diagram, that $\hat{\lambda}_1$ is the length of the first column of λ.

Hence

$$\pi \longrightarrow (P(\pi), Q(\pi)) = \left(\begin{pmatrix} 1 & 2 \\ 3 & \\ 4 & \\ 5 & \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ 2 & \\ 4 & \\ 5 & \end{pmatrix} \right)$$

and $L_5(\pi) = 2 = \#$columns of P or Q.

The Robinson-Schensted-Knuth correspondence has the following properties
- $\pi \mapsto (P, Q)$, then $\pi^{-1} \mapsto (Q, P)$
- length (longest increasing subsequence of π) $= \#$ (columns in P)
- length (longest decreasing subsequence of π) $= \#$ (rows in P)
- $\pi^2 = I$, then $\pi \mapsto (P, P)$
- $\pi^2 = I$, with k fixed points, then P has exactly k columns of odd length.

1.2. Plancherel measure, integrals over $U(n)$ and Toeplitz matrices

A next set of ideas is due to Vershik & Kerov [40], Diaconis & Shashahani [13], Biane [9], Rains [33, 34], Baik & Rains [7]. For a nice state-of-the-art account, see Aldous & Diaconis [5]. The uniform probability measure (see (1.1.1) for f^λ) (1.0.1) on S_N induces on Young diagrams, via the RSK bijection (1.1.6), the "Plancherel" probability measure

$$P_N(\lambda) = \frac{(f^\lambda)^2}{N!}, \quad |\lambda| = N, \tag{1.2.1}$$

and so from (1.2.1) and the RSK bijection, one deduces a number of formulae below; notice, equality (ii) follows from (1.1.1) and (iii) follows from (i) and (1.1.5):

$$\begin{aligned}
P_N(L(\pi_N) \leq n) &= \frac{1}{N!} \# \left\{ \begin{array}{c} (P, Q), \text{ standard Young tableaux each of} \\ \text{arbitrary shape } \lambda, \text{ with } |\lambda| = N, \ \lambda_1 \leq n \end{array} \right\} \\
&\overset{\text{(i)}}{=} \frac{1}{N!} \sum_{\substack{|\lambda|=N \\ \lambda_1 \leq n}} (f^\lambda)^2 = \frac{1}{N!} \sum_{\substack{|\lambda|=N \\ \hat{\lambda}_1 \leq n}} (f^\lambda)^2, \quad \text{by duality of Young diagrams} \\
&\overset{\text{(ii)}}{=} N! \sum_{m=1}^{n} \frac{1}{m!} \sum_{\substack{h \in \mathbb{Z}^m, h_i \geq 1 \\ \sum h_j = N + \frac{m(m-1)}{2}}} \prod_{1 \leq i < j \leq m} (h_i - h_j)^2 \prod_1^m \frac{1}{h_j!^2} \\
&\overset{\text{(iii)}}{=} \frac{1}{N!} \sum_{\substack{|\lambda|=|\mu|=N \\ \hat{\lambda}_1, \hat{\mu}_1 \leq n}} f^\lambda f^\mu \langle s_\lambda, s_\mu \rangle \\
&= \frac{1}{N!} \langle (x_1 + \cdots + x_n)^N, (x_1 + \cdots + x_n)^N \rangle \\
&\overset{\text{(iv)}}{=} \frac{1}{N!} \int_{U(n)} |\operatorname{tr} M|^{2N} dM
\end{aligned} \tag{1.2.2}$$

$$= \frac{1}{N!} \binom{2N}{N}^{-1} \int_{U(n)} (\operatorname{tr}(M + \bar{M}))^{2N} dM,$$

using in the last equality the fact that $\int_{U(n)} ((\operatorname{tr} M)^j (\overline{\operatorname{tr} M})^j) dM = 0$ for $i \neq j$. In 1990, Gessel [17] considered the generating function below and showed that it equals a Toeplitz determinant (determinant of a matrix, whose (i, j)th entry depends on $i - j$ only).

$$\begin{aligned}
\tau_n(x) &:= \sum_{N=0}^{\infty} \frac{x^N}{N!} P(L(\pi_N) \leq n) \\
&\overset{\text{(i)}}{=} \int_{U(n)} e^{\sqrt{x}\, tr(M+\bar{M})} dM \\
&\overset{\text{(ii)}}{=} \frac{1}{n!} \int_{(S^1)^n} \Delta_n(z) \Delta_n(\bar{z}) \prod_{k=1}^{n} \left(e^{\sqrt{x}(z_k + \bar{z}_k)} \frac{dz_k}{2\pi i z_k} \right) \\
&\overset{\text{(iii)}}{=} \det \left(\int_0^{2\pi} e^{2\sqrt{x}\cos\theta} e^{i(k-\ell)\theta} d\theta \right)_{1 \leq k, \ell \leq n} \quad \text{(Toeplitz determinant)} \qquad (1.2.3) \\
&\overset{\text{(iv)}}{=} \prod_{k=0}^{n-1} h_k(x)^{-1}, \quad \text{where } h_k(x) := \frac{\tau_{k+1}}{\tau_k} = \int_{S^1} |p_k(z)|^2 e^{\sqrt{x}(z+\bar{z})} dz, \\
&\overset{\text{(v)}}{=} e^x \prod_{k=n}^{\infty} h_k(x)^{-1},
\end{aligned}$$

where $p_k(z)$ are monic orthogonal polynomials on the circle S^1 for the weight $e^{\sqrt{x}(z+\bar{z})} dz$:

$$\int_{S^1} p_n(z) \overline{p_m(z)} e^{\sqrt{x}(z+\bar{z})} \frac{dz}{2\pi i z} = h_n(x) \delta_{mn}. \qquad (1.2.4)$$

Equality (1.2.3)(i) follows from (1.2.2)(iv); moreover (1.2.3)(ii) uses Haar measure on $U(n)$. Identity (1.2.3) (iii) follows from the fact that the product of the two Vandermonde appearing in the integral (ii) can be expressed as sum of determinants:

$$\begin{aligned}
\Delta_n(u) \Delta_n(v) &= \sum_{\sigma \in S_n} \det \left(u_{\sigma(k)}^{\ell-1} v_{\sigma(k)}^{k-1} \right)_{1 \leq \ell, k \leq n} \\
&= \det \left(p_{\ell-1}^{(1)}(u_k) \right)_{1 \leq \ell, k \leq n} \det \left(p_{\ell-1}^{(2)}(v_k) \right)_{1 \leq \ell, k \leq n}.
\end{aligned}$$

The fact above that $\Delta_n(u)\Delta_n(v)$ can also be expressed as a product of two determinants, involving monic polynomials, implies (iv). Finally, by Szegö's strong limit theorem for Toeplitz determinants, we have $\lim_{n \to \infty} \tau_n = e^x$, thus leading to (v).

1.3. Virasoro constraints, integrable systems and Painlevé V equation

THEOREM 1.1. — *For every $\ell \geq 0$, the Gessel generating function*

$$\sum_{N=0}^{\infty} \frac{x^N}{N!} P(L(\pi_N) \leq \ell) = \int_{U(\ell)} e^{\sqrt{x}\,\mathrm{tr}(M+\bar{M})} dM \tag{1.3.1}$$

$$= \exp \int_0^x \log\left(\frac{x}{u}\right) g_\ell(u)du,$$

where g_ℓ is the unique solution to the initial value problem:

$$\left\{\begin{array}{l} g'' - \dfrac{g'^2}{2}\left(\dfrac{1}{g-1} + \dfrac{1}{g}\right) + \dfrac{g'}{u} + \dfrac{2}{u} g(g-1) - \dfrac{\ell^2}{2u^2}\dfrac{g-1}{g} = 0 \\ \\ g_\ell(u) = 1 - \dfrac{u^\ell}{(\ell!)^2} + O(u^{\ell+1}), \ \text{near } u = 0. \quad \textbf{(Painlevé V)}. \end{array}\right. \tag{1.3.2}$$

Theorem 1.1 is due to Hisakado [19], Tracy-Widom [36], by methods of functional analysis and Adler-van Moerbeke [3], by integrable methods. A similar statement holds for the set S_{2n}^0 of fixed-point free involutions π^0(i.e., $(\pi^0)^2 = I$ and $\pi^0(k) \neq k$ for $1 \leq k \leq 2n$). Put the uniform distribution on S_{2n}^0 :

$$P(\pi_{2n}^0) = \frac{1}{(2n-1)!!} = \frac{2^n n!}{(2n)!}. \tag{1.3.3}$$

Then, in the following statement, due to Adler-van Moerbeke [3] and Baik-Rains [7], $O(n)_\pm$ refers to orthogonal matrices, with determinant $= \pm 1$.

THEOREM 1.2. — *The generating function*

$$2\sum_{k=0}^{\infty} \frac{(x^2/2)^k}{k!} P(L(\pi_{2k}^0) \leq \ell + 1) = E_{O(\ell+1)_-} e^{x\ trM} + E_{O(\ell+1)_+} e^{x\ trM}$$

$$= \exp\left(\int_0^x \frac{f_\ell^-(u)}{u} du\right) + \exp\left(\int_0^x \frac{f_\ell^+(u)}{u} du\right), \tag{1.3.4}$$

where $f = f_\ell^\pm$ is the unique solution to the initial value problem:

$$\left\{\begin{array}{l} f''' + \dfrac{1}{u}f'' + \dfrac{6}{u}f'^2 - \dfrac{4}{u^2}ff' - \dfrac{16u^2+\ell^2}{u^2}f' + \dfrac{16}{u}f + \dfrac{2(\ell^2-1)}{u} = 0 \\ \\ f_\ell^\pm(u) = u^2 \pm \dfrac{u^{\ell+1}}{\ell!} + O(u^{\ell+2}), \ \text{near } u = 0. \quad \textbf{(Painlevé V)} \end{array}\right. \tag{1.3.5}$$

Proof of Theorem 1.1. — A brief outline will be given here, because of its interesting connection with an integrable system, called the Toeplitz lattice (see [3]). Inserting times t_i, with $i = ..., -2, -1, 0, 1, 2, ...,$ and $t_0 = 0$, in the integrals (1.2.3)(ii) over $(S^1)^n$, one obtains, setting $t := (..., t_{-1}, t_0, t_1, ...)$:

$$I_n(t) = \int_{(S^1)^n} |\Delta_n(z)|^2 \prod_{k=1}^{n} e^{\sum_{i\neq 0} t_i z_k^i} \frac{dz_k}{2\pi i z_k}, \quad n \geq 0. \tag{1.3.6}$$

Virasoro constraints [3]. — The I_n's satisfy a set of three linear partial differential equations, forming an sl(2,$\mathbb{Z}$)-algebra (Virasoro constraints), explicitly given by the three equations:

$$
\begin{aligned}
&\left(\sum_{i\neq\mp1,0}(i\pm1)t_{i\pm1}\frac{\partial}{\partial t_i}+n\left(t_{\pm1}+\frac{\partial}{\partial t_{\mp1}}\right)\right)I_n(t)=0\\
\text{(1.3.7)}\qquad \text{and}\quad &\left(\sum_{i\neq0}it_i\frac{\partial}{\partial t_i}\right)I_n(t)=0.
\end{aligned}
$$

Toeplitz lattice. — The t-dependent semi-infinite moment matrix,

$$
\text{(1.3.8)}\qquad m_n(t):=\left(\langle z^k,z^\ell\rangle_t\right)_{0\le k,\ell\le n-1}=\left(\oint_{S^1}\frac{\rho(z)dz}{2\pi iz}z^{k-\ell}e^{\sum_{i\neq0}(t_iz^i)}\right)_{0\le k,\ell\le n-1}
$$

is a Toeplitz matrix and satisfies the simple differential equations:

$$
\text{(1.3.9)}\qquad \frac{\partial\mu_{ij}}{\partial t_k}=\mu_{i+k,j},\qquad \frac{\partial\mu_{ij}}{\partial t_{-k}}=\mu_{i,j+k},\quad \text{for } k\geq1.
$$

In analogy with (1.2.3)(iii), define $\tau_n(t)$, instead of $\tau_n(x)$:

$$
\tau_n(t):=\det m_n(t).
$$

The Borel factorization $m_\infty(t)=S_1^{-1}(t)S_2(t)$ with a lower-triangular matrix S_1 (with 1's along the diagonal) and an upper-triangular matrix S_2, enables one to define(3) $L_1:=S_1\Lambda S_1^{-1}$ and $L_2:=S_2\Lambda^\top S_2^{-1}$, which combined with the Toeplitz nature of m_∞ has a peculiar "rank 2"-structure, where x_i and y_i are certain integrals over $U(n)$ and where $h_i=\tau_{i+1}/\tau_i$, $h_i/h_{i-1}=1-x_iy_i$, $h:=\operatorname{diag}(h_0,h_1,\ldots)$ and $x_0=y_0=1$:

$$
L_1=h\begin{pmatrix}
-x_1y_0 & 1-x_1y_1 & 0 & 0 & \\
-x_2y_0 & -x_2y_1 & 1-x_2y_2 & 0 & \\
-x_3y_0 & -x_3y_1 & -x_3y_2 & 1-x_3y_3 & \\
-x_4y_0 & -x_4y_1 & -x_4y_2 & -x_4y_3 & \\
 & & & & \ddots
\end{pmatrix}h^{-1}
$$

and

$$
\text{(1.3.10)}\qquad L_2=\begin{pmatrix}
-x_0y_1 & -x_0y_2 & -x_0y_3 & -x_0y_4 & \\
1-x_1y_1 & -x_1y_2 & -x_1y_3 & -x_1y_4 & \\
0 & 1-x_2y_2 & -x_2y_3 & -x_2y_4 & \\
0 & 0 & 1-x_3y_3 & -x_3y_4 & \\
 & & & & \ddots
\end{pmatrix}.
$$

(3) where Λ is the shift matrix $(\Lambda v)_n=v_{n+1}$, $n\geq0$, i.e., Λ is the semi-infinite matrix with all 0 entries, except for 1's just above the diagonal.

The **2d-Toda Lattice** hierarchy[(4)]

$$\frac{\partial L_i}{\partial t_n} = [(L_1^n)_+ \, , L_i] \quad \text{and} \quad \frac{\partial L_i}{\partial t_{-n}} = [(L_2^n)_- \, , L_i] \, , \quad i = 1, 2, n \geq 1, \tag{1.3.11}$$

maintains the "rank 2" nature of the semi-infinite matrices L_1 and L_2. The first equation in this hierarchy is equivalent to the **discrete sinh-Gordon equation**

$$\frac{\partial^2 q_n}{\partial t_1 \partial t_{-1}} = -e^{q_n - q_{n-1}} + e^{q_{n+1} - q_n}, \quad \text{where} \quad q_n = \log \frac{\tau_{n+1}}{\tau_n}.$$

The equations (1.3.11) induce on the x_i and y_i the **Toeplitz lattice** equations; see [3].

Using the three equations (1.3.7) and $\partial/\partial t_1$ of the first equation of (1.3.7), and setting all $t_i = 0$, enable one to extract various t-partials, like $\partial^2 \log \tau(t)/\partial t_{-1}\partial t_1$ and $\partial^2 \log \tau(t)/\partial t_{-2}\partial t_1$, in terms of pure partials in $\partial^k/\partial x^k$, where $x = t_1 t_{-1}$. Substituting these expressions in the integrable Toeplitz lattice equations leads to Painlevé V (1.3.2). The integrable system associated with the integral (1.3.4) in Theorem 1.2 (with t_i's inserted) is the standard Toda lattice, instead of the Toeplitz lattice; the integral satisfies the Virasoro constraints associated with the Toda lattice; see [3].

1.4. Random permutations π_N for large N

Around 1960 and based on Monte-Carlo methods, Ulam [38] conjectured that

$$\lim_{n \to \infty} \frac{E(L_n)}{\sqrt{n}} = c \ \text{ exists.}$$

An argument of Erdös & Szekeres [14], dating back from 1935 showed that $E(L_n) \geq \frac{1}{2}\sqrt{n-1}$, and thus $c \geq 1/2$. In '72, Hammersley [18] showed rigorously that the limit exists. Logan and Shepp [23] showed the limit $c \geq 2$, and finally Vershik and Kerov [40] that $c = 2$. The next major contribution was due to Johansson [22] and Baik-Deift-Johansson [6]:

THEOREM 1.3 ("Law of large numbers" and "Central limit theorem")

One has:

(1.4.1)

$$\lim_{n \to \infty} \frac{L_n}{2\sqrt{n}} = 1, \quad \textit{and} \quad \lim_{n \to \infty} P\left(\frac{L_n - 2\sqrt{n}}{n^{1/6}} \leq x\right) = \exp\left(-\int_x^\infty (u - x) g^2(u)\, du\right),$$

where $g(x)$ is a solution of (a version of) Painlevé II,

$$\begin{cases} g'' = xg + 2g^3 \quad (\textbf{Painlevé II}) \\ g(x) \cong A(x) := \int_{-\infty}^{\infty} e^{ixy - y^3/3} dy \cong \frac{e^{-\frac{2}{3}x^{\frac{3}{2}}}}{2\sqrt{\pi}x^{1/4}} \ \textit{for } x \nearrow \infty. \end{cases} \tag{1.4.2}$$

[(4)] $(\ )_+$ denotes the upper-triangular part of $(\ )$, including the diagonal, whereas $(\ )_-$ denotes the strictly lower-triangular part of $(\ )$.

The point $\gamma = \frac{2\sqrt{x}}{n} = 1$ corresponds to a "*phase transition*" for the partition function

$$e^{-x}\tau_n(x) = e^{-x}\int_{U(n)} e^{\sqrt{x}\; tr(M+\bar{M})}dM = \prod_{k=n}^{\infty} h_k(x)^{-1} = e^{-x}\sum_{n=0}^{\infty}\frac{x^N}{N!}P(L(\pi_N)\le n)$$

On the one hand, Johansson shows, for all $\varepsilon > 0$, that there exist C and $\delta > 0$ such that
(1.4.3)

$$e^{-x}\tau_n(x) \le Ce^{-\delta x}, \text{ for } \frac{2\sqrt{x}}{n} > 1+\varepsilon \text{ and } e^{-x}\tau_n(x) \ge 1-\frac{C}{n}, \text{ for } \frac{2\sqrt{x}}{n} < 1-\varepsilon.$$

On the other hand, setting $P_{n,N} := P(L(\pi_N) \le n)$, Johansson's de-Poissonization Lemma goes as follows: given any $\alpha > 3$, there are constants $C = C(\alpha)$ and $N_0 = N_0(\alpha)$ such that the following holds for $N \ge N_0$ and all $0 \le n \le N$, with $m(\alpha) = (\alpha^2 - 2\alpha - 1)/2$:

$$(1.4.4)\quad e^{-x}\sum_{n=0}^{\infty}\frac{x^N}{N!}P_{n,N}\big|_{x=N+\alpha\sqrt{N\log N}} - \frac{C}{N^{m(\alpha)}} \le P_{n,N} \le e^{-x}\sum_{n=0}^{\infty}\frac{x^N}{N!}P_{n,N}\big|_{x=N-\alpha\sqrt{N\log N}} + \frac{C}{N^{m(\alpha)}}.$$

The two relations (1.4.3) and (1.4.4) combined lead to the law of large numbers in (1.4.1). To prove the "central limit theorem" (1.4.1), one uses the Riemann-Hilbert approach to obtain asymptotics. Indeed, as a first observation, setting $\rho_x(z) := e^{\sqrt{x}(z+z^{-1})}$, the only solution Y_{n+1} (2×2 matrix) to the following Riemann-Hilbert problem *à la* Fokas, Its and Kitaev [15], but adapted to S^1:

(1) $Y(z)$ holomorphic in $\mathbb{C}\backslash S^1$

(2) $Y_-(z) = Y_+(z)\begin{pmatrix} 1 & \frac{\rho_x(z)}{z^{n+1}} \\ 0 & 1 \end{pmatrix}$

(3) $Y(z)\begin{pmatrix} z^{-(n+1)} & 0 \\ 0 & z^{n+1}\end{pmatrix} = \big(I + O(z^{-1}\big)$, when $z\to\infty$ is given by the matrix

$$(1.4.5)\qquad Y_{n+1}(z) = \begin{pmatrix} p_{n+1}(z) & \int_{S^1} p_{n+1}(u)\frac{\rho_x(u)}{u-z}\frac{du}{2\pi i u^{n+1}} \\ -h_n^{-1}z^n\overline{p_n(1/\bar{z})} & -h_n^{-1}\int_{\mathbb{R}}\overline{p_n(1/\bar{u})}\frac{\rho_x(u)}{u-z}\frac{du}{2\pi i u^{n+1}} \end{pmatrix},$$

where $p_k(z)$ are the monic orthogonal polynomials (1.2.4) on the circle S^1 for the weight $\rho_x(z)dz$.

The Riemann-Hilbert formulation is an efficient tool to find the asymptotics of $(Y_{k+1}(0))_{21} = -h_k^{-1}(x)$. Then, summing up, $\sum_n^{\infty}\log h_k^{-1}(x) = \log(e^{-x}\tau_n(x))$, Baik-Deift-Johansson [6] show the following result: Define u such that $\frac{2\sqrt{x}}{n+1} =$

$1-\frac{u}{2^{1/3}(n+1)^{2/3}}$, with $-M<u<M$ for a given $M>0$. Then the estimate below holds, where g is the solution to (1.4.2):

$$\left|\log(e^{-x}\tau_n(x))-\int_u^\infty 2if(y)dy\right|=\frac{C(M)}{n^{1/3}}+Ce^{-1/4M^{3/2}}$$

with $2if(y) = -\int_y^\infty g^2(\alpha)d\alpha$. This estimate combined with Johansson's de-Poissonization Lemma leads to (1.4.1) and (1.4.2).

2. THE SPECTRUM OF RANDOM MATRICES

Random matrix theory deals with an ensemble of matrices M, having some symmetry condition to guarantee the reality of the spectrum, e.g., the Hermitian ensemble $\mathcal{H}_n$, the symmetric ensemble $\mathcal{S}_n$ or the symplectic ensemble $\mathcal{T}_{2n}$. Define a probability measure, based on a Haar measure, which in "polar coordinates" is expressed in terms of the Vandermonde with a power β:

$$c_n e^{-trV(M)}dM = c_n|\Delta_n(z)|^\beta \prod_1^n e^{-V(z_k)}dz_k dU,$$

with $V' = \frac{g}{f}$ rational, dM = Haar measure. The three ensembles $\mathcal{H}_n$, $\mathcal{S}_n$, $\mathcal{T}_{2n}$ appear very naturally as the tangent spaces (at the identity) to the simplest symmetric spaces:

non-compact symmetric space G/K	$\mathfrak{g}=\mathfrak{k}\oplus\mathfrak{p}$ with $\mathfrak{p}$=tangent space to G/K at I	space $\mathfrak{p}=\{M\in\mathfrak{g}, \text{with...}\}$	Haar measure on $\mathfrak{p}$
$SL(n,\mathbb{C})/SU(n)$	$sl(n,\mathbb{C})=su(n)\oplus\mathcal{H}_n$	$M=\bar{M}^\top$	$\beta=2$
$SL(n,\mathbb{R})/SO(n)$	$sl(n,\mathbb{R})=so(n)\oplus\mathcal{S}_n$	$M=M^\top$	$\beta=1$
$SU^*(2n)/USp(n)$	$su^*(2n)=(sp(n,\mathbb{C})\cap u(2n))\oplus\mathcal{T}_{2n}$	$M=\bar{M}^\top,\ M=J\bar{M}J^{-1}$	$\beta=4$

Question. — What is the statistics of the spectrum of M ?

For $\mathcal{H}_n$ and $\mathcal{S}_n$, if the probability $P(M\in dM)$ satisfies the following two requirements: (i) invariance under conjugation by unitary transformations $M\mapsto UMU^{-1}$, (ii) the random variables M_{ii}, $\Re M_{ij}$, $\Im M_{ij}$, $1\le i<j\le n$ are independent, then $V(z)$ is quadratic (Gaussian ensemble) ([28]).

2.1. Virasoro constraints, Toda and Pfaff lattices and KP equations

Consider weights of the form $\rho(z)dz := e^{-V(z)}dz$ on an interval $F=[A,B]\subseteq\mathbb{R}$, with rational logarithmic derivative and subjected to the following boundary conditions:

$$(2.1.1)\qquad -\frac{\rho'}{\rho}=\frac{g}{f}=\frac{\sum_0^\infty b_i z^i}{\sum_0^\infty a_i z^i},\qquad \lim_{z\to A,B} f(z)\rho(z)z^k=0 \text{ for all } k\ge 0,$$

and a disjoint union of intervals $E=\bigcup_1^{2r}[x_{2i-1},x_{2i}]\subset F\subset\mathbb{R}$.

THEOREM 2.1 (Adler-van Moerbeke [1])

The vector of integrals $I(t,x;\beta)=(I_0=1, I_1(t,x;\beta),\dots)$, *with* $t:=(t_1,t_2,\dots)$ *and* $x:=(x_1,\dots,x_{2r})$ *and*

$$I_n(t,x;\beta):=\int_{E^n}|\Delta_n(z)|^\beta\prod_{k=1}^n\left(e^{\sum_1^\infty t_iz_k^i}\rho(z_k)dz_k\right)\ \textit{for}\ \ n>0 \tag{2.1.2}$$

satisfies the following Virasoro constraints[5] *for all* $k\geq -1$*:*

$$\left(-\sum_1^{2r}x_i^{k+1}f(x_i)\frac{\partial}{\partial x_i}+\sum_{i\geq 0}\left(a_i\ {}^\beta\mathbb{J}^{(2)}_{k+i}(t,n)-b_i\ {}^\beta\mathbb{J}^{(1)}_{k+i+1}(t,n)\right)\right)I(t,x;\beta)=0,$$

in terms of the coefficients a_i, b_i *of the rational function* $(-\log\rho)'$ *and the end points* x_i *of the subset* E, *as in (2.1.1). The* ${}^\beta\mathbb{J}^{(2)}_k$ *and* ${}^\beta\mathbb{J}^{(1)}_k$ *form a Virasoro and a Heisenberg algebra respectively, interacting as follows*

$$\begin{aligned}\left[{}^\beta\mathbb{J}^{(2)}_k,\ {}^\beta\mathbb{J}^{(2)}_\ell\right] &= (k-\ell)\ {}^\beta\mathbb{J}^{(2)}_{k+\ell}+c\left(\frac{k^3-k}{12}\right)\delta_{k,-\ell}\\ \left[{}^\beta\mathbb{J}^{(2)}_k,\ {}^\beta\mathbb{J}^{(1)}_\ell\right] &= -\ell\ \ {}^\beta\mathbb{J}^{(1)}_{k+\ell}+c'k(k+1)\delta_{k,-\ell}.\\ \left[{}^\beta\mathbb{J}^{(1)}_k,\ {}^\beta\mathbb{J}^{(1)}_\ell\right] &= \frac{k}{\beta}\delta_{k,-\ell},\end{aligned} \tag{2.1.3}$$

with central charge $c=1-6\left(\left(\frac{\beta}{2}\right)^{1/2}-\left(\frac{\beta}{2}\right)^{-1/2}\right)^2$ *and* $c'=\left(\frac{1}{\beta}-\frac{1}{2}\right)$.

Moreover:

(i) $\tau_n(t):=\frac{1}{n!}I_n(t,x,\beta)|_{\beta=2}$ *form the* τ*-functions of the standard Toda lattice; in particular, each* τ_n *satisfies the KP equation:*

$$\left(\left(\frac{\partial}{\partial t_1}\right)^4+3\left(\frac{\partial}{\partial t_2}\right)^2-4\frac{\partial^2}{\partial t_1\partial t_3}\right)\log\tau_n+6\left(\frac{\partial^2}{\partial t_1^2}\log\tau_n\right)^2=0.$$

(ii) $\tau_{2n}(t):=\begin{cases}\frac{1}{(2n)!}I_{2n}(t,x;\beta)|_{\beta=1}\\ \frac{1}{n!}I_n(2t,x;\beta)|_{\beta=4}\end{cases}$ *form the* τ*-functions of the Pfaff lattice (i.e., they are Pfaffians, rather than determinants); in particular, they satisfy the Pfaff-KP equation:*

$$\left(\left(\frac{\partial}{\partial t_1}\right)^4+3\left(\frac{\partial}{\partial t_2}\right)^2-4\frac{\partial^2}{\partial t_1\partial t_3}\right)\log\tau_{2n}+6\left(\frac{\partial^2}{\partial t_1^2}\log\tau_{2n}\right)^2=12\frac{\tau_{2n-2}\tau_{2n+2}}{\tau_{2n}^2}.$$

[5] When E equals the whole range F, then the $\partial/\partial x_i$'s are absent in the formulae (2.1.7).

2.2. The Gaussian ensemble: PDE's for the statistics of the spectrum

THEOREM 2.2. — *For the Gaussian ensemble, the probabilities*[(6)]*: ($\beta = 2, 1, 4$)*

$$
\begin{aligned}
P_n(E) &:= P_n(\text{ all spectral points of } M \in E) \\
&= \frac{\int_{\mathcal{H}_n(E),\ \mathcal{S}_n(E) \text{ or } \mathcal{T}_n(E)} e^{-tr\ M^2} dM}{\int_{\mathcal{H}_n(\mathbb{R}),\ \mathcal{S}_n(\mathbb{R}) \text{ or } \mathcal{T}_n(\mathbb{R})} e^{-tr\ M^2} dM} \qquad (2.2.1)\\
&= \frac{\int_{E^n} |\Delta_n(z)|^\beta \prod_{k=1}^n e^{-z_k^2} dz_k}{\int_{\mathbb{R}^n} |\Delta_n(z)|^\beta \prod_{k=1}^n e^{-z_k^2} dz_k}
\end{aligned}
$$

satisfy the following PDE's in the x_i's ($\mathcal{B}_k = \sum_1^{2r} x_i^{k+1} \frac{\partial}{\partial x_i}$, $\delta_{1,4}^\beta = 1$ for $\beta = 1$, 4 and 0 otherwise, $f_n(x) = \frac{d}{dx} \log P_n$):[(7)]

$$
\delta_{1,4}^\beta Q \left(\frac{P_{n-\frac{2}{1}} P_{n+\frac{2}{1}}}{P_n^2} - 1 \right) \qquad \begin{cases} 2 & \text{and } n \text{ even, for } \beta = 1 \\ 1 & \text{and } n \text{ arbitrary, for } \beta = 4 \end{cases}
$$

$$
= \begin{cases} \left(\mathcal{B}_{-1}^4 + (Q_2 + 6\mathcal{B}_{-1}^2 F)\mathcal{B}_{-1}^2 + (2 - \delta_{1,4}^\beta) \dfrac{4}{\beta} (3\mathcal{B}_0^2 - 4\mathcal{B}_{-1}\mathcal{B}_1 + 6\mathcal{B}_0) \right) \log P_n \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{for } E = \bigcup_1^{2r} [x_{2i-1}, x_{2i}] \\ f_n''' + 6f_n'^2 + \left(\dfrac{4x^2}{\beta} (\delta_{1,4}^\beta - 2) + Q_2 \right) f_n' - \dfrac{4x}{\beta} (\delta_{1,4}^\beta - 2) f_n, \quad \text{for } E = [-\infty, x]. \end{cases}
$$

Note that for $\beta = 2$ and for $E = [-\infty, x]$, the equation above takes on the simple form:

$$
f_n''' + 6f_n'^2 + 8nf_n' + 4xf_n = 0. \qquad \textbf{(Painlevé IV)} \tag{2.2.2}
$$

Equation (2.2.2) was obtained by Tracy-Widom [35] and the rest of Theorem 2.2 is due to M. Adler, T. Shiota and P. van Moerbeke ([4] and [1]).

2.3. Infinite Hermitian matrix ensembles

Consider probability (2.2.1) for $\beta = 2$ and let the size n of the matrices go to ∞. To perform this limit, one uses a different representation; namely,

$$
P_n(E) := \frac{\int_{\mathcal{H}_n(E)} e^{-tr\ M^2} dM}{\int_{\mathcal{H}_n} e^{-tr\ M^2} dM} = \frac{1}{n!} \int_{E^n} \det(K_n(z_k, z_\ell))_{1 \le k, \ell \le n} \prod_1^n \rho(dz_i), \tag{2.3.1}
$$

(6) $\mathcal{H}_n(E)$, $\mathcal{S}_n(E)$ or $\mathcal{T}_n(E)$ denotes the subset of matrices with spectrum in $E \subseteq \mathbb{R}$.

(7) Also, define the *invariant* polynomials

$$
Q = 12n \left(n + 1 - \frac{2}{\beta} \right) \quad \text{and} \quad Q_2 = 4(1 + \delta_{1,4}^\beta) \left(2n + \delta_{1,4}^\beta (1 - \frac{2}{\beta}) \right).
$$

can be represented in terms of a reproducing kernel
(2.3.2)
$$K_n(y,z) := \sum_{j=1}^{n} \frac{p_{j-1}(y)}{\sqrt{h_{j-1}}}\frac{p_{j-1}(z)}{\sqrt{h_{j-1}}}, \text{ where } \int_{\mathbb{R}} K_n(y,z)K_n(z,u)\rho(dz) = K_n(y,u).$$

The reproducing property follows from the orthogonality of the *monic orthogonal polynomials* $p_k = p_k(z)$ for the Gaussian weight $e^{-z^2}dz$ on $\mathbb{R}$, and the L^2-norms $h_k = \int_{\mathbb{R}} p_k^2(z)e^{-z^2}dz$ of the p_k's. We also have the following Fredholm determinant formula

$$P(\text{exactly } k \text{ eigenvalues } \in [x_1,x_2]) = \frac{(-1)^k}{k!}\left(\frac{\partial}{\partial\lambda}\right)^k \det(I - \lambda K_n(y,z)I_{[x_1,x_2]}(z))\Big|_{\lambda=1}.$$

When the size n of the Hermitian matrices tends to ∞, the following holds:

– *Wigner's semi-circle law*: For this ensemble and for very large n, the density of eigenvalues tends to the semi-circle distribution on the interval $[-\sqrt{2n}, \sqrt{2n}]$.

– *Bulk scaling limit*: From the formula above, it follows that the average number of eigenvalues per unit length near $z = 0$ ("the bulk") is given by $\sqrt{2n}/\pi$ and thus the average distance between two consecutive eigenvalues is given by $\pi/\sqrt{2n}$. Upon using this rescaling, one shows ([24, 27, 29, 32, 21])

$$\lim_{n\nearrow\infty} \frac{\pi}{\sqrt{2n}}K_n\left(\frac{\pi y}{\sqrt{2n}}, \frac{\pi z}{\sqrt{2n}}\right) = \frac{\sin \pi(y-z)}{\pi(y-z)} =: K(y,z) \qquad \text{(Sine kernel)}$$

with

$$P(\text{no eigenvalues } \in [0,x]) = \det(I - K(y,z)I_{[0,x](z)}) = \exp \int_0^{\pi x} \frac{f(x)}{x}dx, \tag{2.3.3}$$

where $f(x,\lambda)$ is a solution to the **Painlevé V** differential equation, (see Jimbo, Miwa, Mori, Sato [21]), ($' = \partial/\partial x$)

$$(xf'')^2 - 4(xf' - f)(-f'^2 - xf' + f) = 0 \text{ with } f(x;\lambda) \cong -\frac{x}{\pi}, \text{ for } x \simeq 0.$$

– *Edge scaling limit*: I will particularly concentrate on the "edge" $\sqrt{2n}$ of the Wigner semi-circle. There the scaling is $\sqrt{2}n^{1/6}$, which makes the problems considerably subtler; this will be used in (2.3.7) (see [10, 16, 28, 35]). The main result can be stated as follows:

THEOREM 2.3. — *Given the spectrum $z_1 \geq z_2 \geq \ldots$ of the large random Hermitian matrix M, define the "eigenvalues" in a new scale:*

$$u_i = 2n^{\frac{2}{3}}\left(\frac{z_i}{\sqrt{2n}} - 1\right) \quad \textit{for } n \nearrow \infty \tag{2.3.4}$$

then the statistics of the largest "eigenvalue" u_1 (in the new scale) is given by the same probability distribution as the length of the longest increasing sequence:

$$P(u_1 \le x) = \exp\left(-\int_x^\infty (\alpha - x) g^2(\alpha) d\alpha \right),$$

(2.3.5)
$$\textit{with} \begin{cases} g'' = xg + 2g^3 \quad \textbf{(Painlevé II)} \\ g(x) \cong A(x) \cong -\frac{e^{-\frac{2}{3}x^{\frac{3}{2}}}}{2\sqrt{\pi}x^{1/4}} \ \textit{for } x \nearrow \infty. \end{cases}$$

More generally, $P(\text{no points } u_i \in [x_1, x_2])$ satisfies the partial differential equation (2.3.6)

$$\left(\mathcal{B}_{-1}^3 - 4(\mathcal{B}_0 - \frac{1}{2})\right) f + 6(\mathcal{B}_{-1} f)^2 = 0, \ \textit{with } f := \sum_1^2 \frac{\partial \log P}{\partial x_i}, \quad \mathcal{B}_k = \sum_1^2 x_i^{k+1} \frac{\partial}{\partial x_i}.$$

□

Tracy-Widom [35] have first obtained result (2.3.5) by methods of functional analysis. In [4], equations (2.3.5) and (2.3.6) were derived using integrable systems and Virasoro constraints.

Proof. — Setting $z = \sqrt{2n} + \frac{u}{\sqrt{2}n^{1/6}}$, in the kernel $K_n(x, y)$ of (2.3.2), one proves

(2.3.7)
$$\lim_{n \nearrow \infty} \frac{1}{\sqrt{2}n^{1/6}} K_n \left(\sqrt{2n} + \frac{u}{\sqrt{2}n^{1/6}}, \sqrt{2n} + \frac{v}{\sqrt{2}n^{1/6}} \right) = K(u, v),$$

where the Airy kernel K is defined in terms of the Airy function:

(2.3.8)
$$K(u, v) = \int_0^\infty A(x + u) A(x + v) dx,$$
$$A(u) = \int_{-\infty}^\infty e^{iux - x^3/3} dx \cong \frac{e^{-(2/3)u^{3/2}}}{2\sqrt{\pi}u^{1/4}}, \quad \text{for } u \nearrow \infty.$$

To find the differential equations (2.3.6), one proceeds (sketchily) as follows: as a first step, notice that, setting $L^2 := (\frac{\partial}{\partial t_1})^2 + q(t)$, $t := (t_1, t_2, ...)$, the solution $q(t)$ and $\Psi(t; z)$ (with asymptotics $\Psi(t; z) = e^{\sum_1^\infty t_i z^i}(1 + O(z^{-1}))$, $z \in \mathbb{C}$, $z \to \infty$) to the initial value problem:

(2.3.9)
$$\begin{cases} (L^2 - z^2)\Psi(t; z) = 0, \quad \frac{\partial L^2}{\partial t_n} = [(L^n)_+, L^2], \quad \frac{\partial \Psi}{\partial t_n} = (L^n)_+ \Psi, \\ \\ q(t^{(0)}) = -x, \quad \text{at } t = t^{(0)} := (x, 0, \frac{2}{3}, 0, 0, ...), \end{cases}$$

is given by Kontsevich's integral (Z diagonal and N arbitrary), which itself is intimately related to the Airy function $A(x)$,

(2.3.10)
$$\begin{cases} q(t) &= 2\frac{\partial^2}{\partial t_1^2} \log \tau(t), \\ \Psi(t; z) &= e^{\sum_1^\infty t_i z^i} \tau(t - [z^{-1}])/\tau(t), \\ \Psi(t^{(0)}; z) &= z^{1/2} A(x + z^2) = e^{xz + \frac{2}{3}z^3}(1 + O(z^{-1})), \quad z \to \infty. \end{cases}$$

with

$$\tau(t) = \frac{\int_{\mathcal{H}_N} dX e^{-tr(X^3/3+X^2Z)}}{\int_{\mathcal{H}_N} dX e^{-tr(X^2Z)}}, \quad t_n := -\frac{1}{n} trZ^{-n} + \frac{2}{3}\delta_{n3}$$

and

$$[z^{-1}] = (z^{-1}, z^{-2}/2, z^{-3}/3, ...) \in \mathbb{C}^\infty$$

These data define a kernel

$$K_t(z^2, z'^2) := \frac{1}{z^{\frac{1}{2}} z'^{\frac{1}{2}}} \int_0^\infty \Psi(t;z)\Psi(t;z')dt_1, \text{ where } t = (t_1, t_2, ...) \tag{2.3.11}$$

which flows off the Airy kernel $K(z^2, z'^2)$, defined in (2.3.8), upon using (2.3.10).

Then, one shows that both, Kontsevich's integral $\tau(t)$ and the product $\tau(t, E) :=$ $\tau(t)$.
$\det(I - K_t(\lambda, \lambda')I_E(\lambda'))$, with $E = \bigcup_1^{2r} [x_{2i-1}, x_{2i}] \subset \mathbb{R}$, satisfy Virasoro constraints, of which the two first ones read (after the time shift $t_3 \mapsto t_3 + 2/3$, in view of (2.3.9)):

$$\begin{aligned}
\sum_{i=1}^{2r} \frac{\partial}{\partial x_i} \log \tau(t, E) &= \left(\frac{\partial}{\partial t_1} + \frac{1}{2}\sum_{i\geq 3} it_i \frac{\partial}{\partial t_{i-2}}\right) \log \tau(t, E) + \frac{t_1^2}{4} \\
\sum_{i=1}^{2r} x_i \frac{\partial}{\partial x_i} \log \tau(t, E) &= \left(\frac{\partial}{\partial t_3} + \frac{1}{2}\sum_{i\geq 1} it_i \frac{\partial}{\partial t_i}\right) \log \tau(t, E) + \frac{1}{16},
\end{aligned} \tag{2.3.12}$$

Both, $\tau(t)$ and $\tau(t, E)$, also satisfy the Korteweg-de Vries equation (KP equation, depending on odd t_i's only)

$$\left(\left(\frac{\partial}{\partial t_1}\right)^4 - 4\frac{\partial^2}{\partial t_1 \partial t_3}\right) \log \tau + 6\left(\frac{\partial^2}{\partial t_1^2} \log \tau\right)^2 = 0. \quad \textbf{(KdV)} \tag{2.3.13}$$

Differentiating (2.3.12) in t_1 and t_3, and setting all $t_i = 0$ enable one to express the t-partials $\frac{\partial^2}{\partial t_1^2} \log \tau\Big|_{t=0}$, $\frac{\partial^4}{\partial t_1^4} \log \tau\Big|_{t=0}$, $\frac{\partial^2}{\partial t_1 \partial t_3} \log \tau\Big|_{t=0}$ appearing in the KdV equation in terms of partials $\partial/\partial x_i$.

Setting these expressions in the the KdV-equation at $t = 0$, implies that the statistics of the scaled eigenvalues u_i (for the Airy kernel (2.3.8))

$$P(E^c) := P(\text{all "eigenvalues"}, u_i \in E^c) = \det(I - KI_E) = \left.\frac{\tau(t, E)}{\tau(t)}\right|_{t=0},$$

satisfies the partial differential equation (2.3.6) with

$$f := \sum_1^{2r} \frac{\partial \log P}{\partial x_i}, \quad \mathcal{B}_k = \sum_1^{2r} x^{k+1} \frac{\partial}{\partial x_i}.$$

When $E = (x, \infty)$, the equation (2.3.6) for $f = \partial \log P((-\infty, x])/\partial x$ becomes a 3rd order ODE (Chazy-type equation)

$$f''' - 4xf' + 2f + 6f'^2 = 0. \tag{2.3.14}$$

According to [12] , this equation has a first integral, which is a Painlevé II equation

$$f''^2 + 4f'(f'^2 - xf' + f) = 0. \tag{2.3.15}$$

and this equation has a solution given by $f := g'^2 - xg^2 - g^4$ and $f' = -g^2$, which establishes (2.3.5) and (2.3.6).

3. LARGE RANDOM MATRICES AND PERMUTATIONS: A DIRECT CONNECTION VIA ENUMERATIVE GEOMETRY

For large random Hermitian matrices, whose entries have Gaussian real and imaginary part of variance = 1/2, the Wigner semi-circle has support on $(-2\sqrt{n}, 2\sqrt{n})$, instead of $[-\sqrt{2n}, \sqrt{2n}]$. Then the (slightly modified) scaling (2.3.4) is given by

$$u_i = n^{2/3}\left(\frac{z_i}{2n^{1/2}} - 1\right) \quad \text{for } n \nearrow \infty. \tag{3.1}$$

Given the set of partitions $\lambda = (\lambda_1 \geq \lambda_2 \geq ...)$ of n, in view of (1.4.1), consider

$$v_i = n^{1/3}\left(\frac{\lambda_i}{2n^{1/2}} - 1\right) \quad \text{for } n \nearrow \infty. \tag{3.2}$$

Theorems 1.3 and 2.3 show that the variables u_1 and v_1 have the same probability distribution. Okounkov [31] shows *in a direct way* that the joint distribution of all u_i and v_i coincide; more specifically:

THEOREM 3.1 (Okounkov, [31]). — *For any $m \geq 1$, any $\xi_1, ..., \xi_m > 0$, the following holds:*

$$E\Big(\sum_{i=1}^{\infty} e^{\xi_1 u_i} ... \sum_{i=1}^{\infty} e^{\xi_m u_i}\Big) = E\Big(\sum_{i=1}^{\infty} e^{\xi_1 v_i} ... \sum_{i=1}^{\infty} e^{\xi_m v_i}\Big), \quad \textit{when} \quad n \nearrow \infty.$$

The left hand side, which relates to random permutations, hinges on the following expressions:

$$\left(\frac{1}{2\sqrt{n}}\right)^{k_1+...+k_m} \frac{n^{m/2}}{n!} tr(X_1^{k_1})...X_m^{k_m}))$$
$$= \left(\frac{1}{2}\right)^{k_1+...+k_m} \sum_{\mathbf{S}} n^{\frac{1}{2}(\chi(\mathbf{S})-m)} \left|\mathrm{Cov}_{\mathbf{S}}(k_1, ..., k_m)\right|,$$

In this expression, the X_i are certain matrices, such that e.g. $\lambda_j - j$ is an eigenvalue of X_1, if (j, λ_j) is a corner of the Young diagram (i.e., $\lambda_j > \lambda_{j+1}$). Also $\mathrm{Cov}_{\mathbf{S}}(k_1, ..., k_m)$ are coverings of S^2 ramified according to a rule, given by the numbers $k_1, ..., k_m$.

The right hand side relates to random matrices and has an interpretation as a "map" on a surface. It hinges on the following formula for $M \in U(n)$:

$$\left(\frac{1}{2\sqrt{n}}\right)^{k_1+\ldots+k_m} E(tr(M^{k_1})\ldots tr(M^{k_m})) \\ = \left(\frac{1}{2}\right)^{k_1+\ldots+k_m} \sum_{\mathbf{S}} n^{\chi(\mathbf{S})-m} \left|\mathrm{Map}_{\mathbf{S}}(k_1, \ldots, k_m)\right|,$$

where the sum is taken over all the homeomorphism classes of orientable, not necessarily connected surfaces $\mathbf{S}$; $\chi(\mathbf{S})$ is its Euler number and $\mathrm{Map}_{\mathbf{S}}(k_1, \ldots, k_m)$ denotes the ways to cover the surface $\mathbf{S}$ with m polygons (a k_1-gon, a k_2-gon,..., a k_m-gon) by pairwise gluing the sides of different or the same k_i-gons. Each polygon has a marked vertex, to distinguish a k-gon from its $k-1$ rotations.

The main proposition, established by Okounkov and leading to Theorem 3.1, is the following: for connected surfaces, as $k_i \to \infty$,

$$\mathrm{Cov}_{\mathbf{S}}(k_1, \ldots, k_m) \cong \mathrm{Map}_{\mathbf{S}}(k_1, \ldots, k_m).$$

4. INTEGRALS, MOMENT MATRICES AND INTEGRABLE SYSTEMS

The first column contains a list of matrix integrals and a Fredholm determinant. After perturbing the integral by inserting times t_i's and possibly s_i's, the new integral thus obtained satisfies linear PDE's (Virasoro constraints). The integral can be represented as a determinant of a "moment" matrix, (defined by an appropriate inner-product) or a Pfaffian, if the moment matrix is skew. Performing an appropriate "Borel factorization" of this associated moment matrix, one shows that each of these integrals, as a function of the t_i and s_i's, satisfies an integrable lattice or an integrable PDE. In all the cases, combining both systems of equations leads to ODE's or PDE's for the corresponding integral, in x or in the boundary of E. The last column lists the connection with probability. For more information on the 6th integral and the double matrix integral, see [37, 3, 2].

τ-functions satisfying Virasoro constraints, after inserting t_i's	= determinant or Pfaffian of m_ℓ of the form:	underlying integrable lattice or PDE	connecting with
$\int_{\mathcal{H}_\ell(E)} e^{-trV(M)}dM$	Hänkel	Toda lattice	spectrum of random Hermitian matrices
$\int_{S_\ell(E)} e^{-trV(M)}dM$	skew-symmetric	Pfaff lattice	spectrum of random symmetric matrices
$\int_{\mathcal{T}_\ell(E)} e^{-trV(M)}dM$	skew-symmetric	Pfaff lattice	spectrum of random symplectic matrices
$\int_{U(\ell)} e^{\sqrt{x}\ tr(M+\bar{M})}dM$	Toeplitz	Toeplitz lattice	longest increasing sequence in random permutations
$\int_{O(\ell)_\pm} e^{x\ trM}dM$	Hänkel	Toda lattice	longest increasing sequence in random involutions
$\int_{U(\ell)} \det(I+M)^k e^{-xtr\bar{M}}dM$	Toeplitz	Toeplitz lattice	longest increasing sequence in random words
$\det\left(I-K(y,z)I_{(x,\infty)}(z)\right)$ (K sine or Airy kernel)	Fredholm determinant	KdV equation	spectrum of infinite random Hermitian matrices (bulk or edge scaling limit)
$\displaystyle\int\int_{\mathcal{H}_\ell\times\mathcal{H}_\ell(E)} dM_1dM_2 \times e^{-tr(M_1^2+M_2^2-cM_1M_2)}$	general	2d-Toda lattice	coupled random matrices

REFERENCES

[1] M. ADLER and P. VAN MOERBEKE: *Symmetric random matrices and the Pfaff Lattice*, to appear in: Annals of Mathematics, sept 2000 (`solv-int/9903009`) and *The Hermitian, symmetric and symplectic ensembles and PDE's* (`math.-phys./009001`).

[2] M. ADLER and P. VAN MOERBEKE: *The spectrum of coupled random matrices*, Annals of Mathematics, **149**, 921–976 (1999).

[3] M. ADLER and P. VAN MOERBEKE: *Integrals over classical groups, random permutations, Toda and Toeplitz lattices*, Comm. Pure Appl. Math. (2000) (`math.CO/9912143`).

[4] M. ADLER, T. SHIOTA and P. VAN MOERBEKE: *Random matrices, vertex operators and the Virasoro algebra*, Phys. Lett. **A 208**, 101-112, (1995). And *Random matrices, Virasoro algebras and non-commutative KP*, Duke Math. J. **94**, 379-431 (1998).

[5] D. ALDOUS and P. DIACONIS: *Longest increasing subsequences: From patience sorting to the Baik-Deift-Johansson theorem*, Bull. Am. Math. Soc. (new series) **36** (4), 413–432 (1999).

[6] J. BAIK, P. DEIFT and K. JOHANSSON: *On the distribution of the length of the longest increasing subsequence of random permutations*, Math. Archive, Journal Amer. Math. Soc. **12**, 1119-1178 (1999) (`math.CO/9810105`).

[7] J. BAIK and E. RAINS: *Algebraic aspects of increasing subsequences*, (1999), `math.CO/9905083`

[8] D. BESSIS, Cl. ITZYKSON and J.-B. ZUBER : *Quantum field theory techniques in graphical enumeration*, Adv. Appl. Math. **1**, 109-157 (1980).

[9] P. BIANE: *Representations of symmetric groups and free probability*, preprint (1998).

[10] M. BOWICK and E. BRÉZIN: *Universal scaling of the tail of the density of eigenvalues in random matrix models*, Phys. Letters **B 268**, 21–28 (1991).

[11] S. CHADHA, G. MAHOUX, M.L. MEHTA : *A method of integration over matrix variables* : II, J. Phys. A: Math. Gen. 14, 579-586 (1981).

[12] C.M. COSGROVE: *Chazy classes IX-XII of third-order differential equations*, Stud. Appl. Math. **104**, 3, 171-228 (2000).

[13] P. DIACONIS, M. SHASHAHANI: *On the eigenvalues of random matrices* J. Appl. Prob., suppl. in honour of Takàcs **31**A, 49-61 (1994).

[14] P. ERDÖS and G. SZEKERES: *A combinatorial theorem in geometry*, Compositio Math., **2**, 463–470 (1935).

[15] A.S. FOKAS, A.R. ITS, A.V. KITAEV: *The isomonodromy approach to matrix models in 2d quantum gravity*, Comm. Math. Phys., **147**, 395–430 (1992).

[16] P.J. FORRESTER: *The spectrum edge of random matrix ensembles*, Nucl. Phys. B, (1993).

[17] I. M. GESSEL: *Symmetric functions and P-recursiveness*, J. of Comb. Theory, Ser A, **53**, 257–285 (1990).

[18] J.M. HAMMERSLEY: *A few seedlings of research*, Proc. Sixth. Berkeley Symp. Math. Statist. and Probability, Vol. 1, 345–394, University of California Press (1972).

[19] M. HISAKADO: *Unitary matrix models and Painlevé III*, Mod. Phys. Letters, **A 11** 3001–3010 (1996).

[20] Cl. ITZYKSON, J.-B. ZUBER: *The planar approximation*, J. Math. Phys. **21**, 411–421 (1980).

[21] M. JIMBO, T. MIWA, Y. MORI and M. SATO: *Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent*, Physica 1D, 80–158 (1980).

[22] K. JOHANSSON: *The Longest Increasing Subsequence in a Random Permutation and a Unitary Random Matrix Model*, Math. Res. Lett., **5**, no. 1-2, 63–82 (1998).

[23] B.F. LOGAN and L.A. SHEPP: *A variational problem for random Young tableaux*, Advances in Math., **26**, 206–222 (1977).

[24] R.D. KAMIEN, H.D. POLITZER, M.B. WISE: *Universality of random-matrix predictions for the statistics of energy levels* Phys. rev. letters **60**, 1995–1998 (1988).

[25] M. KONTSEVICH: *Intersection theory on the moduli space of curves and the matrix Airy function*, Comm. Math. Phys. **147**, 1-23 (1992).

[26] D. KNUTH: "The Art of Computer programming, Vol III: Searching and Sorting", Addison-Wesley, Reading, MA, 1973.

[27] G. MAHOUX, M.L. MEHTA: *A method of integration over matrix variables: IV*, J. Phys. I (France) **1**, 1093–1108 (1991).

[28] M.L. MEHTA: Random matrices, 2nd ed. Boston: Acad. Press, 1991.

[29] T. NAGAO, M. WADATI: *Correlation functions of random matrix ensembles related to classical orthogonal polynomials*, J. Phys. Soc of Japan, **60** 3298–3322 (1991).

[30] A.M. ODLYZKO: *On the distribution of spacings between the zeros of the zeta function* **48**, 273-308 (1987).

[31] A. OKOUNKOV: *Random matrices and random permutations* (1999), `math.CO/9903176`

[32] L.A. PASTUR: *On the universality of the level spacing distribution for some ensembles of random matrices*, Letters Math. Phys., **25** 259–265 (1992).

[33] E. M. RAINS: *Topics in Probability on compact Lie groups*, Harvard University doctoral dissertation, (1995).
5B

[34] E. M. RAINS: *Increasing subsequences and the classical groups*, Elect. J. of Combinatorics, **5**, R12 (1998).

[35] C.A. TRACY and H. WIDOM: *Level-Spacings distribution and the Airy kernel*, Commun. Math. Phys., **159**, 151–174 (1994).

[36] C.A. TRACY and H. WIDOM: *Random unitary matrices, permutations and Painlevé* (1999), `math.CO/9811154`

[37] C.A. TRACY and H. WIDOM: *On the distribution of the lengths of the longest monotone subsequences in random words* (1999), `math.CO/9904042`

[38] S.M. ULAM: *Monte Carlo calculations in problems of mathematical physics*, in *Modern Mathematics for the Engineers*, E.F. Beckenbach ed., McGraw-Hill, 261–281 (1961).

[39] P. VAN MOERBEKE: *Integrable lattices: random matrices and permutations*, MSRI-volume on Random matrices and exactly solvable models, Eds.: P. Bleher, A. Its, Oxford University press, 2000 (`math.CO/00-10`).

[40] A.M. VERSHIK and S.V. KEROV: *Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux*, Soviet Math. Dokl., **18**, 527–531 (1977).

Pierre VAN MOERBEKE

Département de Mathématiques
Université de Louvain
B–1348 Louvain-la-Neuve (Belgium)
E-mail : `vanmoerbeke@geom.ucl.ac.be`

and

Brandeis University
Waltham, Mass 02454, USA
E-mail : `vanmoerbeke@math.brandeis.edu`