Painlevé Equations in Connection with Random Matrix Theory

Eli Vigneron

April 2020

Abstract. The focus of this talk is to examine the connection between the Painlevé equations and random matrix theory. Solutions to a Painlevé type equation can give us insight into the distributions of the eigenvalues of the classical matrix ensembles. More precisely, integral operator discriminants related to eigenvalue distributions can be shown to satisfy differential equations involving the Painlevé equations in the large n limit. Equations II, III, and V have been studied in detail, but the others are known to arise as well. We will show that the Tracy-Widom distributions can arise from a Painlevé II equation.

1 Introduction

The Painlevé equations have historically found many applications, one such example being in the numerical study of solitons. Painlevé equations are typically introduced when considering questions regarding second-order differential equations that satisfy what is called the **Painlevé property** — the only movable singularities are poles, and these poles move with the initial conditions. To illustrate by way of example, we have the following

$$y' + y^2 = 0, \qquad y(0) = \alpha$$

has solution

$$y(x) = \frac{\alpha}{\alpha x + 1},$$

with a movable pole at $x = -1/\alpha$.

In particular the Painlevé equations take the form $y'' = R(x, y, y')$ where R is analytic in x and rational in y and y'. Painlevé showed that equations satisfying the Painlevé property can be transformed into one of either a linear equation, an elliptic equation, a Riccati equation, or one of the following types:

(I) $y'' = 6y^2 + t$,

(II) $y'' = 2y^3 + ty + \alpha$,

(III) $y'' = \frac{1}{y}y'^2 - \frac{y'}{t} + \frac{\alpha y^2+\beta}{t} + \gamma y^3 + \frac{\delta}{y}$,

(IV) $y'' = \frac{1}{2y}y'^2 - \frac{3}{2}y^3 + 4ty^2 + 2(t^2 - \alpha)y + \frac{\beta}{y}$,

(V) $y'' = \left(\frac{1}{2y} + \frac{1}{y-1}\right)y'^2 - \frac{1}{t}y' + \frac{(y-2)^2}{t}\left(\alpha y + \frac{\beta}{y}\right) + \gamma\frac{y}{t} + \delta\frac{y(y+1)}{y-1}$,

(VI) $y'' = \frac{1}{2}\left(\frac{1}{y} + \frac{1}{y-1} + \frac{1}{y-t}\right)y'^2 - \left(\frac{1}{t} + \frac{1}{t-1} + \frac{1}{y-t}\right)y' + \frac{y(y-1)(y-t)}{t^2(t-1)^2}\Big[\alpha - \beta\frac{t}{y^2} + \gamma\frac{t-1}{(y-2)^2} + \left(\frac{1}{2} - \delta\right)\frac{t(t-1)}{(y-2)^2}\Big]$.

Solutions to a Painlevé type equation can give us insight into the distributions of the eigenvalues of the classical matrix ensembles. More precisely, integral operator discriminants related to eigenvalue distributions can be shown to satisfy differential equations involving the Painlevé equations in the large n limit. Equations II, III, and V have been studied in detail, but the others are known to arise as well.

2 Eigenvalue Distributions

We will make use of the terms *bulk* and *edge* to refer to properties of the eigenvalue distributions, where as suggested *bulk* refers to properties that obtain for all of the eigenvalues and *edge* to those eigenvalues which are either the largest or the smallest.

If we pick $K(x, y)$ as in the below Table 1, and set

$$K[f](x) = \int K(x, y)f(y)dy$$

then, under appropriate integration limits, this operator will be well defined. It happens that, in the limit of large random matrices, the determinant becomes a 'Fredholm determinant'. A summary of some of the various eigenvalue distributions is as follows

Table 1: [ER05]

Painlevé	Statistic	Interval	Kernel	$K(x,y)$
V	'bulk'	$[-s,s]$	sine	$\frac{sin(\pi(x-y)}{\pi(x-y)}$
III	'hard edge'	$(0,s]$	Bessel	$\frac{\sqrt{y}J_\alpha(\sqrt{x})J'_\alpha(\sqrt{y})-\sqrt{x}J_\alpha(\sqrt{y})J'_\alpha(\sqrt{x})}{2(x-y)}$
II	'soft edge'	$[s,\infty)$	Airy	$\frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(y)-\mathrm{Ai}'(x)\,\mathrm{Ai}(y)}{x-y}$

These distributions have arisen in many places and are becoming increasingly more and more important. There are some interesting connections with these distributions and the zeros of the Riemann zeta function. It has been found that the zeros of the Riemann zeta function along the critical line $Re(z) = \frac{1}{2}$ when z is large, are distributed similarly to eigenvalues of random matrices in the GUE.

3 Painlevé II

In this section we will examine the connection between random matrix theory and Painlevé II. Our main reference for this section will be the Tracy-Widhom paper [TW70].

As alluded to previously, interest in limit laws is typically found as $n \to \infty$ and obtaining nontrivial limits requires centering and normalizing the random variables. We will be chiefly concerned here with the limit law associated to the largest eigenvalue. Letting

$$F_{n,\beta} \coloneqq \mathbb{P}_{\beta,n}(\lambda_{\max}), \quad \beta = 1,2,4$$

denote denote the distribution function of the largest eigenvalue, then

$$F_\beta(x) \coloneqq \lim_{n\to\infty} F_{n,\beta}\left(2\sigma\sqrt{n} + \frac{\sigma x}{n^{1/6}}\right), \quad \beta = 1,2,4$$

exists by the basic limit laws as in the Tracy-Widom papers [TW93b, TW94b, TW96]. Further they are explicitly found to be

$$F_2(x) \coloneqq \exp\left(-\int_x^\infty (y-x)q^2(y)dy\right) \tag{i}$$

with q being the unique solution to the Painlevé II equation

$$q'' = xq + 2q^3 \tag{ii}$$

under the boundary condition

$$q(x) \sim \mathrm{Ai}(x), \qquad \text{as } x \to \infty. \tag{iii}$$

In passing we note that it has be shown [HM80] that

$$q(x) = \sqrt{-\frac{x}{2}}\Big(1 + \frac{1}{8x^3} + O\frac{1}{x^6}\Big), \quad \text{as } x \to -\infty.$$

$F_1(x)$ and $F_4(x)$ are given as

$$F_1(x)^2 := F_2(x) \exp\Big(- \int_x^\infty q(y)dy\Big) \tag{iv}$$

and

$$F_4\Big(\frac{x}{2^{2/3}}\Big)^2 := F_2(x)\Big(\cosh\Big(- \int_x^\infty q(y)dy\Big)\Big)^2 \tag{v}$$

The distributions $f_1(x), f_2(x)$ and $f_4(x)$ are the derivatives of these and they can be computed numerically as we will demonstrate via MATLAB. Having numerical evaluations is particularly useful for applications to data analysis.

Before making use of MATLAB, let us rewrite equation ii as a system of first-order ODEs:

$$\frac{d}{dx}\begin{pmatrix} q \\ q' \end{pmatrix} = \begin{pmatrix} q' \\ xq + 2q^3 \end{pmatrix} \tag{vi}$$

We can solve this as an initial value problem where $x = x_0 =$ a large enough positive value, and we integrate back along the x-axis.

Our boundary condition iii, is now the initial values

$$\begin{cases} q(x_0) = \mathrm{Ai}(x_0) \\ q'(x_0) = \mathrm{Ai}'(x_0). \end{cases}$$

Finally, here is the MATLAB code

Listing 1: Computing Tracy-Widom Distributions in MATLAB

```
% define the system
deq = @(x,y) [y(2); x*y(1)+2*y(1)^3];
```

```

% set the integration bounds and output times
x0 = 5;
xn = -8;
sspan = linspace(x0,xn,1000);

% compute the initial values
y0 = [airy(x0); airy(1,x0)];

% set the tolerance and integrate the system
opts = odeset('RelTol',1e-13,'AbsTol',1e-15);
[x,y] = ode45(deq,sspan,y0,opts);

% the function q(x) is the initial entry of y
q=y(:,1);

% initialize
dI0=0;
I0=0;
J0=0;

% numerically integrate
dI=-[0;cumsum((q(1:end-1).^2+q(2:end).^2)/2.*diff(s))]+dI0;
I=-[0;cumsum((dI(1:end-1)+dI(2:end))/2.*diff(s))]+I0;
J=-[0;cumsum((q(1:end-1)+q(2:end))/2.*diff(s))]+J0;

% here we have the distributions
F2=exp(-I);
F1=sqrt(F2.*exp(-J));
F4=sqrt(F2).*(exp(J/2)+exp(-J/2))/2;
x4=x/2^(2/3);

% and here are the probability distributions
f2=gradient(F2,x);
f1=gradient(F1,x);
f4=gradient(F4,x4);

% plotting the probability distributions
plot(x,f1,x,f2,x4,f4)
ylabel('f_\beta(x)',Rotation,0)
```

```
xlabel('x')
legend('\beta = 1','\beta = 2','\beta = 4')
```

Here is the plot:

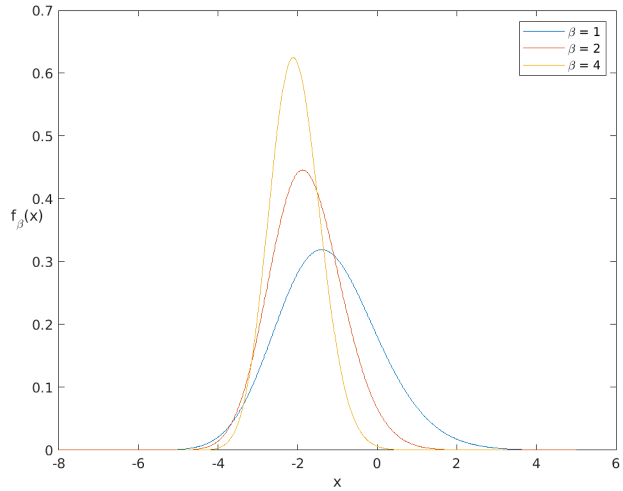

Figure 1: Tracy-Widom distributions

The connection between these Tracy-Widom distributions and random matrix theory is expressed via the following three theorems.

Theorem 1. *Denoting the largest eigenvalue of the β-Hermite ensemble $G_\beta(n,n)$ (for $\beta = 1,2,4$) by λ_{max}, the normalized largest eigenvalue is, in turn, found to be*

$$\lambda'_{max} = n^{1/6}(\lambda_{max} - 2\sqrt{n})$$

And as $n \to \infty$,

$$\lambda'_{max} \xrightarrow{\mathcal{D}} F_\beta(x).$$

Theorem 2. *[Joh01, Theorem 1.1] Denoting the largest eigenvalue of the real Laguerre ensemble $W_1(m,n)$ ($\beta = 1$) by λ_{max}, the normalized largest eigenvalue is, in turn, found to be*

$$\lambda'_{max} = \frac{\lambda_{max} - \mu_{mn}}{\sigma_{mn}}$$

where

$$\mu_{mn} = (\sqrt{m-1} + \sqrt{n})^2, \qquad \sigma_{mn} = (\sqrt{m-1} + \sqrt{n})\left(\frac{1}{\sqrt{m-1}} + \frac{1}{n}\right)^{\frac{1}{3}}.$$

And if $m/n \to \gamma \geq 1$ *as* $n \to \infty$,

$$\lambda'_{max} \xrightarrow{\mathcal{D}} F_1(x).$$

Theorem 3. *Denoting the largest eigenvalue of the real Laguerre ensemble* $W_2(m,n)$ $(\beta = 2)$ *by* λ_{max}*, the normalized largest eigenvalue is, in turn, found to be*

$$\lambda'_{max} = \frac{\lambda_{max} - \mu_{mn}}{\sigma_{mn}}$$

where

$$\mu_{mn} = (\sqrt{m} + \sqrt{n})^2, \qquad \sigma_{mn} = (\sqrt{m} + \sqrt{n})\left(\frac{1}{\sqrt{m}} + \frac{1}{n}\right)^{\frac{1}{3}}.$$

And if $m/n \to \gamma \geq 1$ *as* $n \to \infty$,

$$\lambda'_{max} \xrightarrow{\mathcal{D}} F_2(x).$$

We will provide a sketch of the proof of Theorem 3, i.e. the Tracy-Widom limit for the largest eigenvalue of a Wishart matrix. First recall that

$$F_2(s) = \exp\left(- \int_s^\infty (x-s) q^2(x) dx \right)$$

Let S be an operator on $L^2(t, \infty)$. Then we have the fixed n formula written as $\det(I - S)$. We make use of the scaling $t = \tau(s) = \mu_n + \sigma_n s$ and define

$$S_\tau(x,y) = \sigma_n S(\mu_n + \sigma_n x, \mu_n + \sigma_n y)$$

to be an operator on $L^2(s, \infty)$ so that it hase the same eigenvalues as S does on $L^2(t, \infty)$. Then

$$F_n(\mu_n + \sigma_n s) = \det(I - S_\tau).$$

Define $\bar{S}$ to be the Airy operator on $L^2(s, \infty)$ with kernel

$$\bar{S}(x,y) = \frac{\mathrm{Ai}(x)\,\mathrm{Ai}'(x) - \mathrm{Ai}(y)\,\mathrm{Ai}'(x)}{x - y}.$$

If we can show that $det(I - S_\tau) \to det(I - \bar{S})$ we are done, since it was shown by Tracy and Widom (1994) that F_2 satisfies $F_2(s) = \det(I - \bar{S})$. It is, in turn, sufficient to show that $S_\tau \to \bar{S}$ in trace class norm, since the Fredholm determinant $\det(I - A)$ is continuous function of A in the trace class norm on operators. There is a formula for the commutator $[D, S] = DS - SD$

(where D is the operator $(Df)(x) = f'(x)$) derived by Widom (1999) for the class of unitary ensembles. In particular, for the Laguerre ensemble this operator will have kernel

$$\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) S(x, y) = -\phi(x)\psi(y) - \psi(x)\phi(y) \tag{vii}$$

where $\psi \in \mathcal{H}_n = \text{span}\{\phi_0, \phi_1, \ldots \phi_{n-1}\}$ and ψ is orthogonal to $\mathcal{H}_n$ such that $\int_0^\infty \psi = 0$. Putting $\xi_n(x) = \phi_n(x)/x$ and $a_n = \sqrt{n(n + \alpha_n)}$, for the Laguerre ensemble

$$\phi(x) = (-1)^n \sqrt{\frac{a_n}{2}} \{\sqrt{n + \alpha_n} \xi_n(x) - \sqrt{n} \xi_{n-1}(x)\}$$
$$\psi(x) = (-1)^n \sqrt{\frac{a_n}{2}} \{\sqrt{n} \xi_n(x) - \sqrt{n + \alpha_n} \xi_{n-1}(x)\}.$$

Then vii yields

$$S(x, y) = \int_0^\infty \phi(x + z)\psi(x + z) + \psi(x + z)\phi(y + z) dz. \tag{viii}$$

If we let $\phi_\tau(s) = \sigma_n \phi(\mu_n + s\sigma_n)$, it can be shown that for each s,

$$\phi_\tau(s), \quad \psi_\tau(s) \to \frac{1}{\sqrt{2}} \operatorname{Ai}(s) \tag{ix}$$

as $n \to \infty$, and further

$$\phi_\tau(s), \quad \psi_\tau(s) = O(e^{-s}). \tag{x}$$

Taken together, results viii, ix, x imply that

$$S(x, y) = \int_0^\infty \phi_\tau(x + u)\psi_\tau(y + u) + \psi_\tau(x + u)\phi_\tau(y + u) du \tag{xi}$$
$$\to \int_0^\infty \operatorname{Ai}(x + u) \operatorname{Ai}(y + u) du = \bar{S}(x, y). \tag{xii}$$

where $\bar{S}$ is the Airy kernel.

We can write these integral formulas in terms of operators on $L^2(s, \infty)$ as

$$S_\tau = G_\tau H_\tau + H_\tau G_\tau, \qquad \bar{S} = 2G^2$$

with kernels

$$
\begin{aligned}
G_\tau(x,y) &= \phi_\tau(x+y-s), \qquad G(x,y) = 2^{-1/2}\,\mathrm{Ai}(x+y)\\
H_\tau(x,y) &= \psi_\tau(x+y-s)
\end{aligned}
$$

and we have the following inequality

$$|S_\tau - \bar{S}|_1 \leq 2|G_\tau|_2|H_\tau - G|_2 + 2|G_\tau - G|_2|G|_2 \to 0$$

where the subscript 1 denotes the trace class norm on operators and the ℓ_1 norm on singular values, 2 denotes the Hilbert-Schmidt norm (ℓ_2 norm on singular values). Indeed this inequality follows from ix and x which entail that both G_τ and $H_\tau \to G$ in Hilbert-Schmidt norm on $L^2(s,\infty)$.

Let us discuss a few takeaways from these theorems. The main takeaway from the second theorem can be summarized by noting that even for moderate values of m and n, the Tracy-Widom distribution F_1 can serve as a usable numerical approximation of the null distribution of the largest principal component from Gaussian data. Johnstone [Joh01] notes some heuristics,

1. around 83% if the distribution is less than $\mu_{mn} = (\sqrt{m-1} + \sqrt{n})^2$;
2. around 95% and 99% of the distribution is below $\mu_{mn} + \sigma_{mn}$ and $\mu_{mn} + 2\sigma_{mn}$ respectively.

When the population covariance has exactly r eigenvalues larger than 1, the distribution of the $r+1$ sample eigenvalue can be approximately bounded above by the Tracy-Widom law for an $m \times (n-r)$ matrix, resulting in approximately conservative P-values.

4 Asymptotics for F_β

We consider the tail behaviour of $F_\beta(x)$ as $x \to +\infty$. To do so let us introduce

$$
\begin{aligned}
F(x) &= \exp\Big(-\frac{1}{2}\int_x^\infty (y-x)q(y)^2 dy\Big)\\
E(x) &= \exp\Big(-\frac{1}{2}\int_x^\infty q(y) dy\Big).
\end{aligned}
$$

Then we can write

$$F_1(x) = E(x)F(x), \;\; F_2(x) = F(x)^2, \;\text{ and }\; F_4(x/\sqrt{2}) = \frac{1}{2}\Big(E(x) + \frac{1}{E(x)}\Big)F(x).$$

and letting $x \to +\infty$ we find that

$$\begin{aligned} F(x) &= 1 - \frac{e^{-\frac{4}{3}x^{3/2}}}{32\pi x^{3/2}}\Big(1 + O\frac{1}{x^{3/2}}\Big) \\ E(x) &= 1 - \frac{e^{-\frac{2}{3}x^{3/2}}}{4\sqrt{\pi}x^{3/2}}\Big(1 + O\frac{1}{x^{3/2}}\Big). \end{aligned}$$

The case when $x \to -\infty$ is considerably more difficult, the complete solution for $\beta = 1, 2, 4$ being found only recently [BBD08]. We will not go into detail here aside from citing the relevant results. As $x \to -\infty$

$$\begin{aligned} F_1(x) &= \tau_1 \frac{e^{-\frac{1}{24}|x|^3 - \frac{1}{3\sqrt{2}}|x|^{3/2}}}{|x|^{1/16}}\Big(1 - \frac{1}{24\sqrt{2}|x|^{3/2}} + O(|x|^{-3})\Big), \\ F_2(x) &= \tau_2 \frac{e^{-\frac{1}{12}|x|^3}}{|x|^{1/8}}\Big(1 + \frac{1}{2^6|x|^{3/2}} + O(|x|^{-6})\Big), \\ F_4(x/\sqrt{2}) &= \tau_4 \frac{e^{-\frac{1}{24}|x|^3 - \frac{1}{3\sqrt{2}}|x|^{3/2}}}{|x|^{1/16}}\Big(1 + \frac{1}{24\sqrt{2}|x|^{3/2}} + O(|x|^{-3})\Big), \end{aligned}$$

where

$$\tau_1 = 2^{-11/48}e^{\frac{1}{2}\zeta'(-1)}, \quad \tau_2 = 2^{1/24}e^{\zeta'(-1)}, \quad \tau_4 = 2^{-35/48}e^{\frac{1}{2}\zeta'(-1)}$$

and ζ is the Riemann zeta function.

5 Painlevé V

We turn our attention now to another probability distribution, namely the spacings of the eigenvalues of the Gaussian unitary ensemble, $G_2(n, n)$. The normalized spacing of the eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_m$ is computed by way of

$$\delta_k' = \frac{\lambda_{k-1} - \lambda_k}{\pi\beta}\sqrt{2\beta n - \lambda_k^2}, \qquad k \approx n/2.$$

When $\beta = 2$, the probability distribution $p(x)$ for the eigenvalue spacing's can be found via the Painlevé V equation:

$$(t\sigma'')^2 + 4(t\sigma' - \sigma)(t\sigma' - \sigma + (\sigma')^2) = 0$$

with boundary condition

$$\sigma(t) \approx -\frac{t}{\pi} - \left(\frac{t}{\pi}\right)^2, \qquad \text{as } t \to 0^+.$$

$p(x)$ is then given by

$$p(x) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}E(x)$$

where

$$E(x) = \exp\left(\int_0^{\pi x} \frac{\sigma(t)}{t}\mathrm{d}t.\right)$$

Differentiation then yields

$$p(x) = \frac{1}{x^2}(\pi x\sigma'(\pi x) - \sigma(\pi x) + \sigma(\pi x)^2)E(x).$$

We have only given a brief overview of some of the connections between Random Matrix theory and the Painlevé equations. For a complete picture of these connections, we refer the reader to the Tracy-Widom papers [TW93b, TW94c, TW94a] as well as a paper by Forrester [For00].

References

[BBD08] Jinho Baik, Robert Buckingham, and Jeffery DiFranco. Asymptotics of tracy-widom distributions and the total integral of a painlevé ii function. *Communications in Mathematical Physics*, 280(2):463–497, Feb 2008.

[ER05] Alan Edelman and N. Raj Rao. Random matrix theory. *Acta Numerica*, 14:233–297, 2005.

[For00] P. J. Forrester. Painlevé transcendent evaluation of the scaled distribution of the smallest eigenvalue in the laguerre orthogonal and symplectic ensembles, 2000.

[HM80] S. Hastings and J. McLeod. A boundary value problem associated with the second painlevé transcendent and the korteweg-de vries equation. *Archive for Rational Mechanics and Analysis*, 73:31–51, 03 1980.

[Joh01] Iain M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *The Annals of Statistics*, 29(2):295–327, 2001.

[TW70] Craig Tracy and Harold Widom. *The Distributions of Random Matrix Theory and their Applications*, pages 753–765. 01 1970.

[TW93a] Craig A. Tracy and Harold Widom. Introduction to random matrices. *Lecture Notes in Physics*, page 103–130, 1993.

[TW93b] Craig A. Tracy and Harold Widom. Level-spacing distributions and the airy kernel. *Communications in Mathematical Physics*, 305(1):115–118, Jan 1993.

[TW94a] Craig A. Tracy and Harold Widom. Fredholm determinants, differential equations and matrix models. *Communications in Mathematical Physics*, 163(1):33–72, Jun 1994.

[TW94b] Craig A. Tracy and Harold Widom. Level-spacing distributions and the airy kernel. *Communications in Mathematical Physics*, 159(1):151–174, Jan 1994.

[TW94c] Craig A. Tracy and Harold Widom. Level spacing distributions and the bessel kernel. *Communications in Mathematical Physics*, 161(2):289–309, Mar 1994.

[TW96] Craig A. Tracy and Harold Widom. On orthogonal and symplectic matrix ensembles. *Communications in Mathematical Physics*, 177(3):727–754, Apr 1996.