BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 48, Number 1, January 2011, Pages 147–152
S 0273-0979(2010)01307-0
Article electronically published on October 27, 2010

Random matrix theory: invariant ensembles and universality, by Percy Deift and Dimitri Gioev, Courant Lecture Notes in Mathematics, 18, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, Rhode Island, 2009, x+217 pp., \$33.00, ISBN 978-0-8218-4737-4

While there has been a flurry of recent activity in random matrix theory, many of the ideas go back to the pioneering work of Wigner in the 1950s. He proposed that the local statistical behavior of scattering resonances of neutrons off large nuclei could be modeled by the statistical behavior of the eigenvalues of real symmetric random matrices. The data supported this conjecture and in the subsequent years, many of the conjectures were verified, and the statistical behavior of random matrices was seen to model many problems in physics and pure and applied mathematics.

So what exactly do we mean by a random matrix? At times it means that we are considering matrices whose entries are independent, random variables taken from some distribution. For example, we might consider $n \times n$ symmetric matrices whose diagonal and upper-diagonal elements are independent, identically distributed random variables. Or we might associate a probability distribution to a set of matrices. And no matter which kind of matrices we consider, the central questions are: what statistical properties are shared by them, in particular, what can we say about the eigenvalue distribution, and what can we say universally when n gets large?

The focus of book by Deift and Gioev is on the latter, that is, on classes or ensembles of random matrices with associated distributions and in particular on the following three important, invariant ensembles.

(1) *Unitary Ensembles* (UE) which are $n \times n$ Hermitian matrices together with a distribution that is invariant under unitary conjugation. In other words, if U is any unitary matrix, the measure of a set S of matrices is the same as $US\,U^*$.

(2) *Orthogonal Ensembles* (OE) which are $n \times n$ symmetric matrices together with a distribution that is invariant under conjugation by an orthgonal matrix.

(3) *Symplectic ensembles* (SE) which are $2n \times 2n$ Hermitian self-dual matrices together with a distribution that is invariant under unitary-symplectic conjugation, that is, matrices M that satisfy $M = M^* = JM^tJ^t$ and distributions invariant under the mapping $M \to UMU^*$ with U satisfying $UU^* = I$ and $UJU^* = J$ where

$$J = \mathrm{diag}(J_2, \ldots, J_2), \qquad J_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Each has an associated probability distribution of the form

$$P_n(M)\, dM = \frac{1}{Z_n} e^{-\mathrm{tr}\ Q(M)} dM,$$

2010 *Mathematics Subject Classification.* Primary 60-02; Secondary 47Bxx, 47N50, 60B99, 60E05, 62E99, 82Bxx.

where dM is Lebesgue measure on the algebraically independent entries of M, Q is a real-valued function, generally an even degree polynomial, and Z_n is a normalization constant.

In the case that $Q(x) = x^2$, these measures are equivalent to having, as much as is algebraically possible, matrices whose entries are independent normal or Gaussian random variables. These ensembles are probably the most studied and are indicated with an extra "Gaussian" adjective as the Gaussian Unitary Ensemble (GUE), the Gaussian Orthogonal Ensemble (GOE), and the Gaussian Symplectic Ensemble (GSE).

In each of the above cases one can compute the induced distribution on the space of eigenvalues. For example, in the UE case one can diagonalize a Hermitian matrix as UDU^*, where U is unitary and D is a diagonal matrix and then make a change of variables $M \mapsto (U, D)$. After *integrating out* the unitary part (which is equivalent to computing a Jacobian) one arrives at an induced distribution on the space of eigenvalues. A similar computation can be made for all three ensembles and the resulting probability densities have the form

$$c_n e^{-\frac{\beta}{2}\sum_{i=1}^n Q(x_i)} |\Delta(x_i)|^\beta,$$

where $\Delta(x_i)$ is the Vandermonde determinant

$$\Delta(x_i) = \det(x_j^{i-1})_{1\le i,j\le n} = \prod_{j<i}(x_i - x_j),$$

and c_n is the normalizing constant.

Here, more precisely, we mean that if f is any symmetric function of n real variables, then the expected value of f is

$$c_n \int_{\mathbb{R}^n} f(x_1, x_2, \ldots, x_n) e^{-\frac{\beta}{2}\sum_{i=1}^n Q(x_i)} |\Delta(x_i)|^\beta dx_1 dx_2 \ldots dx_n.$$

For UE $\beta = 2$, for OE $\beta = 1$, and for SE $\beta = 4$, and thus the three ensembles are often referred to as the $\beta = 1, 2$, or 4 ensembles. (It can be shown for SE that eigenvalues occur in pairs and thus all three densities are defined as functions of n variables.)

As soon as the densities are known, one can try to compute some statistical information about the eigenvalues. Since it is easier to describe, we illustrate the ideas with GUE. We can first factor the exponential terms into the Vandermonde determinant so that the entry x_j^{i-1} is replaced by $e^{-x_j^2/2}x_j^{i-1}$, and then using elementary row operations we can replace each row by any polynomial. That is, replace $e^{-x_j^2/2}x_j^{i-1}$ by $e^{-x_j^2/2}p_{i-1}(x_j)$ only changing the determinant by a constant factor. So we choose to replace them by the normalized Hermite polynomials $h_k(x)$ which satisfy

$$\int_{-\infty}^{\infty} h_k(x) h_j(x) e^{-x^2}\, dx = \delta_{jk}.$$

From this it follows, after identifying the constant, that the density on the space of eigenvalues for GUE is

$$c_n e^{-\sum_{i=1}^n x_i^2} |\Delta(x_i)|^2 = \frac{1}{n!} \det K_n(x_i, x_j),$$

where

$$K_n(x_i, x_j) = \sum_{k=0}^{n-1} \varphi_k(x_i)\varphi_k(x_j),$$

and $\varphi_k(x) = h_k(x)e^{-\frac{x^2}{2}}$. The Christoffel-Darboux formula allows one to analyze $K_n(x,y)$ in a more concise form since it says that for $x \neq y$,

$$K_n(x,y) = \sqrt{\frac{n}{2}}\, \frac{\phi_n(x)\phi_{n-1}(y) - \phi_n(y)\phi_{n-1}(x)}{x-y},$$

and for $x = y$, $K_n(x,y)$ is the limit of this expression as $x \to y$.

All the information we need is somehow contained in the function $K_n(x,y)$. It is clear that for large n this information is intimately related to knowledge about the asymptotics of the Hermite polynomials. For example, the density of eigenvalues, $\rho_n(x)$, defined to be the limit of the expected number of eigenvalues in an interval around x divided by its length, as the length tends to zero, is exactly $K_n(x,x)$. Using the known asymptotics for the Hermite polynomials, one can show that

$$\lim_{n\to\infty} \sqrt{\frac{2}{n}}\, \rho_n(\sqrt{2n}\; x\;) = \begin{cases} \frac{2}{\pi}\sqrt{1-x^2} & \text{if } |x| < 1, \\ 0 & \text{if } |x| > 1 \end{cases}$$

holds uniformly on compact sets of $|x| < 1$ and $|x| > 1$. This result, one of the first successes of Wigner's program is called the *Wigner semi-circle law.* Of course, this computation was only for GUE, but at least for UE, the difference is that we would replace the Hermite polynomials with those orthogonal with respect to the weight $e^{-Q(x)}$.

Another important statistic is the gap probability, which is the probability that no eigenvalues are in the interval (a,b). Using algebraic properties of K_n, one can show that the gap probability is given by

$$\sum_{k=0}^{\infty}\left[\frac{(-1)^k}{k!}\int_a^b \cdots \int_a^b \det\left(K_n\left(x_i, x_j\right)\right)\Big|_{i,j=1}^{k}\, dx_1 \cdots dx_k\right].$$

This sum can be reinterpreted as a Fredholm determinant for the operator $I - K_n$, where K_n is the integral operator on $L_2(a,b)$ with kernel $K_n(x,y)$. Now if we wish to think about this operator for n large, it is useful to make a change of variables by replacing $K_n(x,y)$ with

$$\frac{\pi}{\sqrt{2n}} K_n\left(\frac{\pi x}{\sqrt{2n}}, \frac{\pi y}{\sqrt{2n}}\right).$$

Our scaling is determined by the semi-circle law, which tells us that there are on the order of $\sqrt{n}$ eigenvalues in any interval (a,b) contained in the support of the semi-circle, and without scaling the probability of no eigenvalues in the interval tends to zero. So we apply a microscope to the interval and scale its size so that we are able to distinguish the eigenvalues. (Our choice of constants in the scaling is done so that the expected number of eigenvalues per unit interval is one.) A change of variables in the integral shows that this is equivalent to scaling the kernel.

Now once again the asymptotics of the orthogonal polynomials, along with the Christoffel-Darboux formula, become crucial, and one can show that

$$\lim_{n\to\infty} \frac{\pi}{\sqrt{2n}} K_n\left(\frac{\pi x}{\sqrt{2n}}, \frac{\pi y}{\sqrt{2n}}\right) = \frac{\sin \pi(x-y)}{\pi(x-y)}.$$

The end result is that we are now interested in the Fredholm determinant $\det(I-K)$ on $L_2(a,b)$ with kernel $K(x,y) = \frac{\sin \pi(x-y)}{\pi(x-y)}$. This kernel, called the *sine kernel*, is translation invariant and so we can replace (a,b) with an interval of the same length $(0,s)$. In a wonderful connection to differential equations, it is known by the work of Jimbo, Miwa, Môri, and Sato [9] that if we define

$$\sigma(s) = -\frac{d}{ds} \log \det(I-K),$$

then σ satisfies a second-order nonlinear differential equation of Painlevé type. The theory of Painlevé equations and the theory of integrable systems then yield information about the asymptotics of the probability distribution.

The kind of scaling just done is often called "scaling in the bulk", and this idea can be extended to the $\beta = 1$ and 4 cases, although these are considerably more complicated since the kernels are matrix kernels as opposed to scalar kernels.

However, instead of scaling in the bulk one can also scale at the "edge of the spectrum". Here we are really investigating the behavior of the eigenvalues near the largest one, which is around $\sqrt{2n}$. Once again using the asymptotics of the orthogonal polynomials as a guide to capture the correct scaling factors, we let $\lambda_{\max}$ be the largest eigenvalue of our random matrix and define the random variable λ by

$$\lambda_{\max} = \sqrt{2n} + \frac{\lambda}{2^{1/2} n^{1/6}}.$$

Now rescaling our Fredholm determinant kernel leads to a different kernel and letting $(a,b) = (s,\infty)$, we find that in the limit

$$\operatorname{Prob}(\lambda \le s) = \det(I - K_{\text{Airy}}),$$

where K_{Airy} is the integral operator defined on $L^2(s,\infty)$ with kernel

$$K_{\text{Airy}} = \frac{\operatorname{Ai}(x)\operatorname{Ai}'(y) - \operatorname{Ai}'(x)\operatorname{Ai}(y)}{x-y}.$$

Note that all the eigenvalues are less than s if and only if the largest is less than s, so this determinant is the formula for the cumulative distribution of the largest eigenvalue. Note also the similarity between the form of this Airy kernel and the sine kernel.

The determinant $\det(I - K_{\text{Airy}})$, generally denoted by $F_2(s)$, is called a Tracy-Widom distribution. In 1994 it was shown by Tracy and Widom that the distribution could be expressed by an integral involving a Painlevé II transcendent, and thereafter they also discovered the distributions and analogous results for $\beta = 1$ and 4 [14, 15]. These also can be expressed in terms of Painlevé functions:

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

where q satisfies the Painlevé equation

$$q'' = xq + 2q^3$$

with boundary condition $q(x) \sim \mathrm{Ai}(x)$ as $x \to \infty$. (Proofs that such a solution exists, together with the evaluation of its behavior as $x \to -\infty$ was first given by Hastings and McLeod [8]; see also the later works of Clarkson and McLeod [3], Deift and Zhou [7], and also the paper of Ablowitz and Segur [1].)

The discovery of the three distributions had remarkable consequences in analysis, because it soon became apparent that the Tracy-Widom distributions occurred naturally in many other settings. They can be seen in applications related to patience sorting, random permutations, tiling problems, random growth models, and principal component analysis. In some sense, they give rise to a different type of central limit theorem, meaning that if one rescales certain random variables, the limiting distribution is Tracy-Widom.

For an extensive history of random matrix theory, the reader is referred to the classic book by Mehta [11] where much more can be found about other ensembles, universality, and connections to other topics. There are many other facets of random matrix theory and its applications not mentioned in this review. One can consider general β for example, or one can consider Wigner ensembles, those whose entries are independent random variables, or one can perturb UE by adding a fixed Hermitian matrix. There are unexpected applications of random matrix theory to number theory and the spacing of consecutive zeros of L-functions. All of these topics and many more have generated much interesting mathematics in recent years.

Because it is the most transparent case, this review has concentrated on GUE. There are two elements that make things much harder than the case of GUE. One is to extend to more general functions $Q(x)$ and the other is to consider the orthogonal and symplectic ensembles. The book by Deift and Gioev gives a very careful treatment of all three ensembles and as much as is possible and reasonable for general functions $Q(x)$. The goal is to show that the limiting kernels are the same as in the Gaussian cases, and thus the limiting distributions are universal. They precisely define and include all the analytic and algebraic identities necessary to derive the formulas for the basic eigenvalue statistics generally following the approach of Tracy and Widom.

The most important feature of the book is the second half where they give a clear and complete exposition for both OE and SE. They produce the error estimates which yield the limiting kernels for the gap probabilities in the case where $Q(x)$ is a monomial although the methods are applicable for polynomials of even degree. The authors make use an alternative description of the orthogonal and symplectic ensembles which was found in 1999 by Widom [16] combined with a difficult analysis of the asymptotics of the orthogonal polynomials for general $Q(x)$ to obtain the universality results. These estimates are new and more precise than previously derived by the authors themselves [4, 5].

Other relevant results for more general $Q(x)$ and $\beta = 2$ include those found by Bleher and Its [2], Deift, Kricherbauer, K. McLaughlin, Venakides, and Zhou [6], Kuijlaars and McLaughlin [10], Pastur and Shcherbina [12], and for the quartic case and $\beta = 1$ and 4 by Stojanovic, [13].

Much of what is written about random matrix theory comes from a blend of different points of view and different disciplines. The reviewer is grateful that the authors have provided a very nice addition to the literature of random matrix theory that includes a complete, rigorous treatment of the invariant ensembles. It is based on graduate courses given by the authors, but will prove useful as a general reference to the specialist and as a valuable introduction to those curious about random matrices.

References

[1] M.J. Ablowitz and H. Segur, *Asymptotic solutions of the Korteweg-de Vries equation*, Stud. Appl. Math.57 (1977) No 1, 13–44. MR0481656 (58:1757)

[2] P. Bleher and A. Its, *Semiclassical asymptotics of orthogonal polynomials, Riemann-Hilbert problem, and universality in the matrix model*, Ann. of Math. (2) 150 (1999), no. 1, 185–266. MR1715324 (2000k:42033)

[3] P. A. Clarkson and J. B. McLeod, *A connection formula for the second Painlevé transcendent*, Arch. Rational Mech. Anal. 103 (1988), no. 2, 97–138. MR0946971 (89e:34010)

[4] P. Deift and D. Gioev, *Universality at the edge of the spectrum for unitary, orthogonal, and symplectic ensembles of random matrices*, Comm. Pure Appl. Math. 60 (2007), no. 6, 867–910. MR2306224 (2008e:60089)

[5] P. Deift and D. Gioev, *Universality in random matrix theory for orthogonal and symplectic ensembles*, Int. Math. Res. Pap. IMRP 2007, no. 2, Art. ID. MR2335245 (2008e:82026)

[6] P. Deift, T. Kricherbauer, K. T.-R. McLaughlin, S. Venakides, and X. Zhou, *Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory*, Comm. Pure Appl. Math., 52 (1999), 1335–1425. MR1472344 (99g:34038)

[7] P. Deift and X. Zhou, *Asymptotics for the Painlevé II equation*, Comm. Pure Appl. Math., (48) (1995) no 3, 277–337. MR1322812 (96d:34004)

[8] S. P. Hastings and J. B. McLeod, *A boundary value problem associated with the second Painlevé transcendent and the Korteweg-de Vries equation*, Arch. Rational Mech. Anal. 73 (1980), no. 1, 31–51. MR0555581 (81i:34024)

[9] M. Jimbo, T. Miwa, Y. Môri, and M. Sato, *Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent*, Physica D 1 (1980), 80-158. MR0573370 (84k:82037)

[10] A. B. J. Kuijlaars and K. T.-R. McLaughlin, *Generic behavior of the density of states in random matrix theory and equilibrium problems in the presence of real analytic external fields*, Comm. Pure Appl. Math. 53 (2000), no. 6, 736–785. MR1744002 (2001f:31003)

[11] M. Mehta, *Random Matrices*, 3rd ed. Elsevier/Academic Press, Amsterdam, 2004. MR2129906 (2006b:82001)

[12] L. Pastur and M. Shcherbina, *Universality of the local eigenvalue statistics for a class of unitary invariant random matrix ensembles*, J. Statist. Phys. 86 (1997), no. 1-2, 109–147. MR1435193 (98b:82037)

[13] A. Stojanovic, *Universality in orthogonal and symplectic invariant matrix models with quartic potential*, Math. Phys. Anal. Geom. 3 (2000), no. 4, 339–373; MR1845356, errata: Math. Phys. Anal. Geom. 7 (2004), no. 4, 347–349. MR2108624 (2005h:82066)

[14] C. A. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Comm. Math. Phys. 159 (1994), 33-72. MR1257246 (95e:82003)

[15] C. A. Tracy and H. Widom, *On orthogonal and symplectic ensembles*, Comm. Math. Phys. 177 (1996), 727-754. MR1385083 (97a:82055)

[16] H. Widom, *On the relation between orthogonal, symplectic and unitary matrix ensembles.* J. Statist. Phys. 94 (1999), no. 3-4, 347–363. MR1675356 (2000e:82024)

Estelle Basor
American Institute of Mathematics
E-mail address: ebasor@aimath.org