BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 57, Number 1, January 2020, Pages 161–169
https://doi.org/10.1090/bull/1675
Article electronically published on June 25, 2019

Combinatorics and random matrix theory, by Jinho Baik, Percy Deift, and Toufic Suidan, Graduate Studies in Mathematics, Vol. 172, American Mathematical Society, Providence, RI, 2016, xi+461 pp., ISBN 978-0-8218-4841-8

Dynamical approach to random matrix theory, by L. Erdős and H. T. Yau, Courant Lecture Notes in Mathematics, Vol. 28, American Mathematical Society, Providence, RI, 2016, ix+226 pp., ISBN 978-1-4704-3648-3

A fundamental phenomenon in the statistical study of large complex systems is that of *universality*: in the limit where the number N of components of the system goes to infinity, the distribution of various natural statistics of that system, after suitable normalisation, will often converge (in a suitable probabilistic sense) to a universal limiting distribution, the nature of which is largely independent of the microscopic features of the individual components of such a system. For instance, the two most fundamental theorems in probability can both be interpreted as basic examples of universality:

- *Law of large numbers.* If $X_1, X_2, \ldots$ are independent and identically distributed (or *iid* for short) real random variables of finite first moment $\mathbf{E}|X_i| < \infty$ with the normalisation $\mathbf{E}X_i = 0$, then the normalised averages $\frac{X_1+\cdots+X_N}{N}$ converge in probability to the deterministic constant 0.
- *Central limit theorem.* If $X_1, X_2, \ldots$ are iid real random variables of finite second moment $\mathbf{E}|X_i|^2 < \infty$ with the normalisation $\mathbf{E}X_i = 0, \mathbf{E}X_i^2 = 1$, then the normalised averages $\frac{X_1+\cdots+X_N}{\sqrt{N}}$ converge in distribution to the normal distribution $N(0,1)$ of mean zero and variance once.

Note here how the distribution of the individual random variables X_i has almost no bearing on the universal limiting distribution (which is 0 in the law of large numbers, or $N(0,1)$ in the central limit theorem), so long as they are suitably normalised and obey some finite moment condition. In particular, these laws are equally valid for continuous random variables (such as Gaussian variables with law $N(0,1)$) as they are for discrete random variables (such as the Bernoulli distribution that takes values in $\{-1,+1\}$, with an equal probability of each).

These basic examples of universality are now extremely well understood, with many proofs, generalisations, refinements, and other variants. However, there is another instance of the universality phenomenon, originating from the study of random matrices, for which a satisfactory understanding has only begun to emerge in the last few decades, and even then only for certain classes of random models of interest. Roughly speaking, the book of Baik, Deift, and Suidan [3] describes the progress on understanding such universality for "integrable" models which admit exact determinantal formulae for their statistics, while the book of Erős and Yau [8] describes the progress for "Wigner" type models which admit a useful heat flow dynamics that allows such models to approach a "Gaussian" type equilibrium state. It is interesting that while these two sources of universality (determinantal structure and heat flow structure) seem to operate in completely different ways (and to an

2010 *Mathematics Subject Classification.* 05-XX, 15-XX, 33-XX, 35-XX, 41-XX, 47-XX, 52-XX, 60-XX, 82-XX.

almost disjoint class of random models), the final results of the methods developed to exploit these two sources are remarkably similar, suggesting that the universality phenomenon is far broader than what any single technique can establish on its own.

Let us now give a more precise example of a universality result. Let N be a large integer, and consider a *Wigner random matrix ensemble* given by a random Hermitian $N \times N$ matrix $H = (h_{ij})_{1 \leq i,j \leq N}$, $h_{ji} = \overline{h_{ji}}$, where the upper triangular entries $h_{ij}, 1 \leq i, j \leq N$ are jointly independent, and all have mean zero and variance $1/N$:

$$\mathbf{E}h_{ij} = 0; \quad \mathbf{E}|h_{ij}|^2 = \frac{1}{N}.$$

The $\frac{1}{N}$ normalisation is convenient in order to keep the eigenvalues of H of bounded size. For technical reasons one often imposes an additional decay hypothesis on the distribution of these variables; a typical one is the subexponential decay hypothesis

$$\mathbf{P}(|h_{ij}| \geq N^{-1/2}\lambda) \leq C \exp(-\lambda^\theta)$$

for all $\lambda > 0$ and i, j, and some constants $C, \theta > 0$ independent of N.

The most important example of a Wigner ensemble is that of the *Gaussian Unitary Ensemble* (GUE), in which h_{ij} are distributed according to a complex Gaussian of mean zero and variance $1/N$ when $i \neq j$, and a real Gaussian of mean zero and variance $1/N$ when $i = j$. The terminology can be explained by observing that if H is drawn from the GUE, then for any unitary matrix $U \in U(N)$, the random matrix UHU^{-1} has the same distribution as H. At the opposite extreme, another example of a Wigner ensemble is a (normalised) *symmetric Bernoulli random matrix*, in which all the entries h_{ij} take values in $\{-1/\sqrt{N}, +1/\sqrt{N}\}$ with an equal probability of $1/2$ of each.

The famous *Wigner semicircle law* [20] asserts that in the limit $N \to \infty$, the spectrum of H is asymptotically distributed according to the semicircular measure $\rho_{sc}(x)\, dx$, where $\rho_{sc} : \mathbf{R} \to \mathbf{R}^+$ is the function

$$\rho_{sc}(x) := \frac{1}{2\pi} \max(4 - x^2, 0)^{1/2}.$$

More precisely, if one lets $\lambda_1 \geq \cdots \geq \lambda_N$ denote the N (necessarily real) eigenvalues of H in descending order, then for any interval $[a, b]$, the random variable

$$\frac{1}{N} \#\{1 \leq i \leq n : \lambda_i \in [a, b]\}$$

converges in probability to $\int_a^b \rho(x)\, dx$. One can interpret this universal semicircle law as a "noncommutative" version of the central limit theorem; this can be made precise using the formalism of free probability.

Among other things, the semicircle law suggests that the largest eigenvalue λ_1 should converge in probability to 2, which is indeed the case. Given this "law of large numbers", it is then natural to inquire as to a "central limit theorem" that measures how this random variable λ_1 fluctuates around this universal limiting value of 2. Such a central limit theorem exists, but the universal limit is somewhat of a surprise: the normalised[1] random variable $N^{2/3}(\lambda_1 - 2)$ converges in distribution

[1] In [3] a slightly different normalisation for H_N is used, so that λ_1 fluctuates around $\sqrt{2}$ rather than 2.

to the *Tracy–Widom distribution*

$$\lim_{N\to\infty} \mathbf{P}(N^{2/3}(\lambda_1 - 2) \leq t) = \exp(-\int_t^\infty (s-t)u(s)^2\ ds) \tag{1}$$

where $u : s \mapsto u(s)$ is the (unique, global) solution of the *Painlevé II equation*

$$u'' = su + 2u^3$$

with boundary condition $u(s) \sim \mathrm{Ai}(s)$ as $s \to +\infty$ (where $\mathrm{Ai}(s)$ is the Airy function).

This remarkable law for the normalised largest eigenvalue $N^{2/3}(\lambda_1 - 2)$—resembling a distorted version of the normal distribution—was first established in the case of GUE by Tracy and Widom [17]. A derivation can be found in [3, Chapter 6.5] and is a good illustration of several of the techniques covered in that text. Firstly, one views the probability density function of the GUE as a constant multiple of $e^{-N\mathrm{tr}(H^2)/2}\ dH$, where dH is Lebesgue measure on the space of Hermitian matrices. A relatively straightforward calculation then lets one describe the distribution $d\mu_{\mathrm{GUE}} = \rho_N(\lambda_1, \ldots, \lambda_N)d\lambda_1 \cdots d\lambda_N$ of the eigenvalues $\lambda_1, \ldots, \lambda_N$ of H as a *determinantal point process*, in which the joint probability density function $\rho_N(\lambda_1, \ldots, \lambda_N)$ of the eigenvalues is shown to be proportional to a determinant $\det(K(\lambda_i, \lambda_j))_{1\leq i,j\leq N}$, where the kernel K can be described using a formula of Christoffel and Darboux in terms of orthogonal polynomials with respect to the one-dimensional measure $e^{-Nx^2/2}\ dx$ associated to the GUE measure—that is to say, the classical Hermite polynomials. The probability $\mathbf{P}(N^{2/3}(\lambda_1 - 2) \leq t)$ can then be expressed as a Fredholm determinant $\det(1 - \mathbb{K}_{t,N})$ for an explicit integral operator $\mathbb{K}_{t,N}$ with kernel a truncated and rescaled version of K. Near the edge $\lambda_1 \approx 2$ of the spectrum, the Hermite polynomials have asymptotics controlled by the Airy function $\mathrm{Ai}(s)$, and so K and $K_{t,N}$ do so as well. So it remains to compute the Fredholm determinant $\det(1 - \mathbb{A}_t)$ corresponding to a limiting operator $\mathbb{A}_t$ of $\mathbb{K}_{t,N}$. Taking logarithmic determinants, it suffices to calculate the resolvent $(1 - \mathbb{A}_t)^{-1}$. This operator can be expressed via some algebraic manipulations in terms of the resolvent of a certain matrix Cauchy integral operator, which can in turn be interpreted in terms of a solution to a *Riemann–Hilbert problem*—the problem of locating a matrix-valued function m on the exterior of a contour in the complex plane, which obeys a specified jump condition across that contour, as well as some other technical regularity properties. One can then differentiate this function m with respect to the parameter t and use the uniqueness theory for Riemann–Hilbert problems to obtain a differential equation for m which can eventually be transformed to the Painléve II equation.

Variants of this computation can also be established in the bulk of the spectrum, for instance giving a limiting distribution (known as the *Gaudin–Mehta distribution*) for the normalised eigenvalue gaps $\gamma_c N(\lambda_{i+1} - \lambda_i)$, where $i = (c + o(1))N$ for some $0 < c < 1$, where $0 < \gamma_c < \infty$ is a normalisation factor depending on c that can be expressed in terms of the semicircular density ρ_{sc}. One can also compute the asymptotics of other bulk quantities relating to eigenvalue gaps in the bulk, such as normalised k-point correlation functions. See [8] for details. For the sake of exposition we shall restrict our discussion here mostly to the edge case.

The above computations were specific to the GUE model, but the Tracy–Widom distribution is highly universal, appearing as the limiting distribution in many other contexts. For instance, the methods above largely extend to other unitarily

invariant models, in which the density $e^{-N\mathrm{tr}(H^2)/2}\,dH$ is replaced by more general densities $e^{-N\mathrm{tr}V(H)}\,dH$ for various potential functions V, with the main difference being that the orthogonal polynomials are now with respect to a more general measure, but with the same Tracy–Widom law emerging in the (suitably normalised) limit under reasonable hypotheses on V (as well as Gaudin–Mehta laws emerging in the bulk); see for instance the text [5] for details. Such ensembles are not Wigner ensembles in general, because the upper diagonal entries will be correlated with each other rather than being jointly independent due to couplings in the Hamiltonian $\mathrm{tr}V(H)$. Nevertheless, as discussed extensively in [8], it is now known that the Tracy–Widom law is also the universal limiting distribution for any other Wigner ensemble than GUE, despite the fact that for general ensembles of Wigner type the determinantal structure is not directly present. Initially, this particular universality result at the edge was proven (under an additional symmetry hypothesis) by a sophisticated application of the moment method [15], but the method did not extend to the bulk and it was difficult to relax the additional hypotheses on the matrix ensemble. In recent years, a more flexible "three-stage" procedure has been developed for demonstrating this and many further universality results in Wigner-type models (although some of the more advanced results still need to deviate from this procedure). This procedure is the main focus of the text [8].

The three-stage procedure exploits a heat flow structure within the class of Wigner ensembles that connects an arbitrary such ensemble to the well-understood GUE ensemble. Indeed, if $H = H_0$ is a Wigner random matrix, one can consider the random matrices H_t for $t > 0$ from H_0 by the *matrix Ornstein–Uhlenbeck process*

$$dH_t = \frac{1}{\sqrt{N}} d\mathbf{B}_t - \frac{1}{2} H_t\, dt,$$

where $\mathbf{B}_t$ is a Hermitian matrix Brownian motion process indepedent of H_0. Then for each $t > 0$, H_t has the same distribution as an average

$$e^{-t/2} H_0 + \sqrt{1 - e^{-t}} H^G, \tag{2}$$

where H^G is drawn from GUE independently of H_0. In particular, if H_0 is a Wigner random matrix, then so is each H_t, and H_t converges in distribution to GUE as $t \to +\infty$.

This heat flow was first studied by Dyson [6], who observed that the eigenvalues $\lambda_i = \lambda_i(t)$ of H_t evolve by what is now known as *Dyson Brownian motion*:

$$d\lambda_i = \frac{1}{\sqrt{N}} dB_i - \frac{\lambda_i}{2}\, dt + \frac{1}{N} \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j}\, dt,$$

where B_i are independent Brownian motions. Formally, this flow is the gradient flow associated to the GUE eigenvalue measure $d\mu_{\mathrm{GUE}}$ mentioned previously. Using informal arguments, Dyson conjectured that this flow should relax to global equilibrium in unit time scales $t \sim 1$, but should relax to a local equilibrium (at scales comparable to the average eigenvalue spacing) at much smaller time scales, such as $t \sim N^{-1+\varepsilon}$ for any $\varepsilon > 0$. In his words,

> The picture of the gas coming into equilibrium in two well-separated stages, with microscopic and macroscopic time scales, is suggested with the help of physical intuition. A rigorous proof that this picture is accurate would require a much deeper mathematical analysis.

The first step in this direction was taken by Johansson [10], who studied Wigner matrices which were *Gaussian divisible* at some unit time scale $t \sim 1$, which meant that they were distributed according to a law of the form (2) for that value of t. He found that for such class of matrices, the joint distribution function still enjoyed a (more complicated) determinantal structure, and was able to exploit this to obtain some universal limiting behaviour in the bulk.

A different approach of this, closer to the spirit of Dyson's analysis, was subsequently worked out by Erdős, Schlein, and Yau [7] and presented (in streamlined form) in [8]. Their methods are an extension of the classical theory of Bakry and Émery [1] that studied multidimensional diffusive processes with respect to an equilibrium measure $d\mu = \frac{1}{Z} e^{-H(x)}\ dx$ on a Euclidean space $\mathbf{R}^N$. Roughly speaking, their finding was that if the underlying Hamiltonian H obeyed good convexity properties, then the convergence to equilibrium was quite fast, and furthermore the quantitative estimates relating to this convergence could be used to obtain good *log-Sobolev* estimates relating such quantites as the entropy

$$S(f) := \int_{\mathbf{R}^N} f \log f\ d\mu$$

and Dirichlet energy

$$D(\sqrt{f}) := \int_{\mathbf{R}^N} |\nabla \sqrt{f}|^2\ d\mu$$

for any reasonable test function $f : \mathbf{R}^N \to \mathbf{R}^+$. In the case of Dyson Brownian motion, the associated Hamiltonian H takes the form

$$H(\lambda_1, \ldots, \lambda_N) = \frac{1}{2} \sum_{i=1}^N \lambda_i^2 - \frac{2}{N} \sum_{1 \leq i < j \leq N} \log(\lambda_i - \lambda_j).$$

This Hamiltonian is indeed convex, and Bakry–Émery theory can be used to largely recover the results of Johansson at unit time scales $t \sim 1$ by showing that Gaussian divisible measures and GUE measures are asymptotically indistinguishable with respect to various classes of test functions f. However, in [7] it was also shown that one can obtain results at shorter time scales $t \sim N^{-1+\varepsilon}$ by exploiting the stronger convexity properties of H in certain directions, as anticipated by Dyson. To adapt the Bakry–Émery theory to this nonisotropic context, Erdős, Schlein, and Yau introduced a (slightly artificial) modified Hamiltonian,

$$\tilde{H}(\lambda) = H(\lambda) + \frac{1}{2\tau} \sum_{i=1}^N (\lambda_j - \gamma_j)^2,$$

where $0 < \tau \ll 1$ is a timescale parameter and γ_j is the "classical location" of λ_j—the deterministic location of λ_j predicted by the semicircle law via the formula $\int_{-\infty}^{\gamma_j} \rho_{sc}(x)\ dx = \frac{j}{N}$. This Hamiltonian replaces the Dyson Brownian motion with a somewhat different flow, now known as the *local relaxation flow*, but because the new Hamiltonian $\tilde{H}$ now enjoys strong convexity in all directions, the local relaxation flow can be used to establish good log-Sobolev estimates (with $\mu = \frac{1}{Z} e^{-H(x)}\ dx$ replaced by a modified measure $\frac{1}{\tilde{Z}} e^{-\tilde{H}(x)}\ dx$). To use these estimates, it is first necessary to show that the eigenvalues λ_j of the random matrix H are close enough to their classical locations γ_j that the effect of the correction term $\frac{1}{2\tau} \sum_{i=1}^N (\lambda_j - \gamma_j)^2$ on the equilibrium measure μ is manageable. To establish this fact (known as *eigenvalue rigidity*) requires a refinement of the Wigner semicircle

law that is now known as the *local semicircle law*. This can in turn be established by a detailed analysis of the resolvents $(H-z)^{-1}$ for various complex numbers z, in particular utilising a *self-consistent equation* that relates these $N \times N$ resolvents in several ways to analogous $N-1 \times N-1$ or $N-2 \times N-2$ resolvents arising from minors of H. Again, we refer the reader to [8] for details.

By combining the first stage of local semicircular laws with the second stage of local relaxation flow analysis, one can extend the universality results of Johansson to a significantly larger class of Wigner type ensembles, which have the Gaussian divisible form (2) for a very small value $t = N^{-1+\varepsilon}$ of t. To remove the requirement of Gaussian divisibility completely and obtain a fully satisfactory universality result, one uses comparison methods, which form the third and final stage of the Erdős–Schlein–Yau paradigm. The phenomenon being exploited here is that if Wigner matrix ensembles $H = (h_{ij})_{1 \leq i,j \leq N}, H' = (h'_{ij})_{1 \leq i,j \leq N}$ have very similar moments (for instance, if $\mathbf{E}h_{ij}^a \approx \mathbf{E}(h'_{ij})^a$ for $a = 0, 1, 2, 3$), then the spectral statistics of H will also be very close in distribution to the corresponding statistics of H'. The simplest way to demonstrate this phenomenon follows a classical argument of Lindeberg [11] which he used to prove the central limit theorem, and it is based on exchanging the entries of H with the entries of H' one pair at a time (keeping the matrix self-adjoint throughout) and using perturbative methods to measure how the spectral statistics vary with each such replacement. This is easiest to establish for the resolvents $(H-z)^{-1}$ which obey a number of very useful identities regarding their behaviour under perturbation, but the method can also be adapted to other statistics, such as individual eigenvalues. As with the second stage, the local semicircle law from the first stage plays a key role in the estimates needed to make the Lindeberg exchange method work. See for instance [16] for an early instance of this technique (which was in turn inspired by previous use of the Lindeberg method in random matrix theory in [4]). Because any Wigner matrix H_0 is close in moments to the slightly Gaussian divisible counterpart (2) when $t = N^{-1+\varepsilon}$ is small, this three-stage argument now completes the demonstration of universality of the normalised largest eigenvalue $N^{2/3}(\lambda_1 - 2)$, and similar arguments also apply in the bulk.

Laws such as the Tracy–Widom law have also arisen in many contexts that do not initially seem to bear any relationship to random matrix theory. A simple combinatorial example comes from Ulam's problem [18] on the longest increasing sequence that one can find inside a random permutation $\pi(1), \ldots, \pi(N)$ of the first N natural numbers $\{1, \ldots, N\}$. If one denotes this length by ℓ_N, then it can be shown that the normalised expectation $\frac{1}{\sqrt{N}}\mathbf{E}\ell_N$ converges to 2, a result first established in [19] and [12]. In later work culminating in [2], it was found that the normalised fluctuation $\frac{\ell_N - 2\sqrt{N}}{N^{1/6}}$ converged in distribution as $N \to \infty$ to precisely the same Tracy–Widom law (1) that appeared in the random matrix theory setting! One can begin to partially explain this seemingly amazing coincidence by using the famous *Robinson–Schensted correspondence* ([13], [14]) between permutations and pairs of standard Young tableaux of equal shape to re-interpret ℓ_N as the longest length of a Young diagram of size N drawn using the natural Plancherel measure arising from the representation theory of S_N. The well-known correspondence principle between quantum mechanics and classical mechanics in physics suggests that representation theory can be viewed as a "discretised" or "quantised" version of

continuous branches of mathematics such as symplectic geometry or random matrix theory. As discussed in [3], these statistics are also related to the statistics of discrete nonintersecting random walks, while the eigenvalues of GUE can similarly be related to statistics of nonintersecting Brownian motions, giving a more tangible link between the two random variables.

The next step is to try to express the distribution of ℓ_N in a determinantal form. As it turns out, this is not easy to achieve directly, but one must first pass to a "Poissonised" version of ℓ_N in which the random permutation on N elements (which can be described in terms of a random N-element subset of the unit square) is replaced instead by a random Poisson process on the unit square of total intensity N. The two statistics are closely related to each other, as discussed in [3]. A formula of Gessel [9] then describes the statistics of the Poisson analogue L_N of ℓ_N as the determinant of a certain Toeplitz matrix $(\phi_{j-k})_{0\leq j,k\leq n-1}$ for some explicit coefficients ϕ_k. It is then possible to use the theory of Riemann–Hilbert problems to write this determinant in terms of a Fredholm determinant $\det(1-\mathrm{K}_n)$ that closely resembles the Fredholm determinant appearing in the random matrix setting. To compute the limiting statistics as N (and n) go to infinity, one has to compute the asymptotics of various contour integrals, which can be done by the classical saddle point (or "steepest descent") method.

The book [3] under review treats all of these calculations (and many related calculations for other statistics of these types) in great detail and in a largely self-contained fashion. A particular strength of the text is the detailed treatment of the combinatorics of Young tableaux and related objects such as Schur polynomials and the Robinson–Schensted and Robinson–Schensted–Knuth correspondences; the constructions and derivations here do not require any representation-theoretic background and are well suited for students coming from a probability or random matrix theory background. The book also covers the basics of the theory of orthogonal polynomials, determinantal processes, Riemann–Hilbert problems and Dyson Brownian motion, and the correlation functions of unitarily invariant ensembles, although these topics are also covered well in many other modern random matrix theory texts. Many further examples of both discrete and continuous random models exhibiting the same universal laws, such as the fluctuations of the Aztec diamond or asymptotics of direct last passage percolation, are also discussed in some depth. This book is thus a valuable resource for students and researchers in these fields. My only criticism would be that while the book does an excellent job of presenting all the necessary theory and calculations to compute the asymptotics of all the random systems studied, it does not dwell much on explaining *why* these particular universal laws arise (for instance, why one would expect Painléve equations to play any role in the subject), and what the most essential feature of the random systems is that is responsible for that universality. But that is more a failing of our current understanding of the subject than of the book and its authors in particular.

Similarly, the book [8] gives a fully self-contained and careful treatment of all aspects of the three-stage procedure needed to establish universality for Wigner type models. The text wisely does not aim to cover the strongest and most general results in the literature, instead focusing on the most representative results to convey the key ideas, although the last chapter does discuss several further generalisations, variants, and historical notes. The methods here give a satisfying explanation of the universality phenomenon in the case of Wigner models as primarily being a consequence of the rapid convergence to local equilibrium of Dyson Brownian

motion, although (as with [3]) we still do not have a fully intuitive explanation as to why it is specific laws such as the Tracy–Widom law that end up being the universal limiting distribution.

Both texts cover material that is already contained in several papers scattered across the literature, but they give a streamlined and coherent presentation of their respective topics. Together, the two books, while very orthogonal in content, represent the two most fruitful methods we currently have to understand the universality phenomenon for random matrix models and their discrete counterparts, and they will serve as an excellent starting point for students in this area.

References

[1] D. Bakry and M. Émery, *Diffusions hypercontractives* (French), Séminaire de probabilités, XIX, 1983/84, Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206, DOI 10.1007/BFb0075847. MR889476

[2] J. Baik, P. Deift, and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, J. Amer. Math. Soc. **12** (1999), no. 4, 1119–1178, DOI 10.1090/S0894-0347-99-00307-0. MR1682248

[3] J. Baik, P. Deift, and T. Suidan, *Combinatorics and random matrix theory*, Graduate Studies in Mathematics, vol. 172, American Mathematical Society, Providence, RI, 2016. MR3468920

[4] S. Chatterjee, *A generalization of the Lindeberg principle*, Ann. Probab. **34** (2006), no. 6, 2061–2076, DOI 10.1214/009117906000000575. MR2294976

[5] P. A. Deift, *Orthogonal polynomials and random matrices: a Riemann-Hilbert approach*, Courant Lecture Notes in Mathematics, vol. 3, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1999. MR1677884

[6] F. J. Dyson, *A Brownian-motion model for the eigenvalues of a random matrix*, J. Mathematical Phys. **3** (1962), 1191–1198, DOI 10.1063/1.1703862. MR148397

[7] L. Erdős, B. Schlein, and H.-T. Yau, *Universality of random matrices and local relaxation flow*, Invent. Math. **185** (2011), no. 1, 75–119, DOI 10.1007/s00222-010-0302-7. MR2810797

[8] L. Erdős and H.-T. Yau, *A dynamical approach to random matrix theory*, Courant Lecture Notes in Mathematics, vol. 28, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2017. MR3699468

[9] I. M. Gessel, *Symmetric functions and P-recursiveness*, J. Combin. Theory Ser. A **53** (1990), no. 2, 257–285, DOI 10.1016/0097-3165(90)90060-A. MR1041448

[10] K. Johansson, *Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices*, Comm. Math. Phys. **215** (2001), no. 3, 683–705, DOI 10.1007/s002200000328. MR1810949

[11] J. W. Lindeberg, *Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung* (German), Math. Z. **15** (1922), no. 1, 211–225, DOI 10.1007/BF01494395. MR1544569

[12] B. F. Logan and L. A. Shepp, *A variational problem for random Young tableaux*, Advances in Math. **26** (1977), no. 2, 206–222, DOI 10.1016/0001-8708(77)90030-5. MR1417317

[13] G. d. B. Robinson, *On the representations of the symmetric group*, Amer. J. Math. **60** (1938), no. 3, 745–760, DOI 10.2307/2371609. MR1507943

[14] C. Schensted, *Longest increasing and decreasing subsequences*, Canad. J. Math. **13** (1961), 179–191, DOI 10.4153/CJM-1961-015-3. MR0121305

[15] A. Soshnikov, *Universality at the edge of the spectrum in Wigner random matrices*, Comm. Math. Phys. **207** (1999), no. 3, 697–733, DOI 10.1007/s002200050743. MR1727234

[16] T. Tao and V. Vu, *Random matrices: universality of local eigenvalue statistics*, Acta Math. **206** (2011), no. 1, 127–204, DOI 10.1007/s11511-011-0061-3. MR2784665

[17] C. A. Tracy and H. Widom, *Level-spacing distributions and the Airy kernel*, Comm. Math. Phys. **159** (1994), no. 1, 151–174. MR1257246

[18] S. M. Ulam, *Monte Carlo calculations in problems of mathematical physics*, Modern Mathematics for the Engineer: Second Series, McGraw-Hill, New York, 1961, pp. 261–281. MR0129165

[19] A. M. Veršik and S. V. Kerov, *Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux* (Russian), Dokl. Akad. Nauk SSSR **233** (1977), no. 6, 1024–1027. MR0480398

[20] E. P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Ann. of Math. (2) **62** (1955), 548–564, DOI 10.2307/1970079. MR0077805

Terence Tao

Department of Mathematics
University of California at Los Angeles
Los Angeles, California
Email address: tao@math.ucla.edu