Symmetry, Integrability and Geometry: Methods and Applications SIGMA **12** (2016), 083, 20 pages

On the Scaling Limits of Determinantal Point Processes with Kernels Induced by Sturm–Liouville Operators⋆

Folkmar BORNEMANN

Zentrum Mathematik – M3, Technische Universität München, 80290 München, Germany
E-mail: *bornemann@tum.de*
URL: http://www-m3.ma.tum.de/bornemann

Received April 15, 2016, in final form August 16, 2016; Published online August 19, 2016
http://dx.doi.org/10.3842/SIGMA.2016.083

Abstract. By applying an idea of Borodin and Olshanski [*J. Algebra* **313** (2007), 40–60], we study various scaling limits of determinantal point processes with trace class projection kernels given by spectral projections of selfadjoint Sturm–Liouville operators. Instead of studying the convergence of the kernels as functions, the method directly addresses the strong convergence of the induced integral operators. We show that, for this notion of convergence, the Dyson, Airy, and Bessel kernels are universal in the bulk, soft-edge, and hard-edge scaling limits. This result allows us to give a short and unified derivation of the known formulae for the scaling limits of the classical random matrix ensembles with unitary invariance, that is, the Gaussian unitary ensemble (GUE), the Wishart or Laguerre unitary ensemble (LUE), and the MANOVA (multivariate analysis of variance) or Jacobi unitary ensemble (JUE).

Key words: determinantal point processes; Sturm–Liouville operators; scaling limits; strong operator convergence; classical random matrix ensembles; GUE; LUE; JUE; MANOVA

2010 Mathematics Subject Classification: 15B52; 34B24; 33C45

Dedicated to Percy Deift at the occasion of his 70th birthday.

1 Introduction

We consider determinantal point processes on a (not necessarily bounded) interval $\Lambda = (a, b)$ with a correlation kernel given by a trace class projection kernel,

$$K_n(x,y) = \sum_{j=0}^{n-1} \phi_j(x)\phi_j(y), \tag{1.1}$$

where $\phi_0, \phi_1, \dots, \phi_{n-1}$ are orthonormal in $L^2(\Lambda)$; each ϕ_j may have some dependence on n that we suppress from the notation. We recall (see, e.g., [2, Section 4.2]) that for such processes the joint probability density of the n points is given by

$$p_n(x_1, \dots, x_n) = \frac{1}{n!} \det_{i,j=1}^{n} K_n(x_i, x_j),$$

the mean counting probability is given by the density (note that $\operatorname{tr} K_n = n$)

$$\rho_n(x) = n^{-1} K_n(x,x),$$

⋆This paper is a contribution to the Special Issue on Asymptotics and Universality in Random Matrices, Random Growth Processes, Integrable Systems and Statistical Physics in honor of Percy Deift and Craig Tracy. The full collection is available at http://www.emis.de/journals/SIGMA/Deift-Tracy.html

and the gap probabilities are given, by the inclusion-exclusion principle, in terms of a Fredholm determinant, namely

$$E_n(J) = \mathbb{P}(\{x_1, \ldots, x_n\} \cap J = \varnothing) = \det(I - \chi_J K_n \chi_J).$$

The various scaling limits are usually derived from an appropriate convergence of the kernel $K_n(x, y)$ by considering the large n asymptotic of the eigenfunctions ϕ_j, which can be technically quite involved[1].

Borodin and Olshanski [4] suggested, for discrete point processes, a different, conceptually and technically much simpler approach based on selfadjoint difference operators. We will show that their method, generalized to selfadjoint Sturm–Liouville operators, allows us to give a short and unified derivation of the various scaling limits for the random matrix ensembles with unitary invariance that are based on the classical orthogonal polynomials (Hermite, Laguerre, Jacobi).

The Borodin–Olshanski method

The method proceeds along three steps: First, we identify the induced integral operator K_n as the spectral projection (where we denote by χ_A the characteristic function of a Borel subset $A \subset \mathbb{R}$ and by $\chi_A(L_n)$ the application of that function to the selfadjoint operator L_N in the sense of measurable functional calculus [17, Theorem VIII.6])

$$K_n = \chi_{(-\infty,0)}(L_n)$$

of some selfadjoint ordinary differential operator L_n on $L^2(\Lambda)$. Any scaling of the point process by $x = \sigma_n \xi + \mu_n$ ($\sigma_n \neq 0$) yields, in turn, the induced rescaled operator

$$\tilde{K}_n = \chi_{(-\infty,0)}(\tilde{L}_n),$$

where $\tilde{L}_n$ is a selfadjoint differential operator on $L^2(\tilde{\Lambda}_n)$, $\tilde{\Lambda}_n = (\tilde{a}_n, \tilde{b}_n)$.

Second, if $\tilde{\Lambda}_n \subset \tilde{\Lambda} = (\tilde{a}, \tilde{b})$ with $\tilde{a}_n \to \tilde{a}$, $\tilde{b}_n \to \tilde{b}$, we aim for a selfadjoint operator $\tilde{L}$ on $L^2(\tilde{\Lambda})$ with a core C such that eventually $C \subset D(\tilde{L}_n)$ and

$$\tilde{L}_n u \to \tilde{L} u, \qquad u \in C. \tag{1.2}$$

The point is that, if the test functions from C are particularly nice, such a convergence is just a simple consequence of the *locally uniform convergence of the coefficients* of the differential operators $\tilde{L}_n$ – a convergence that is, typically, an easy calculus exercise. Now, given (1.2), the concept of *strong resolvent convergence* (see Theorem A.1) immediately yields[2], if $0 \notin \sigma_{\mathrm{pp}}(\tilde{L})$,

$$\tilde{K}_n \chi_{\tilde{\Lambda}_n} = \chi_{(-\infty,0)}(\tilde{L}_n)\chi_{\tilde{\Lambda}_n} \xrightarrow{s} \chi_{(-\infty,0)}(\tilde{L}).$$

Third, we take an interval $J \subset \tilde{\Lambda}$, eventually satisfying $J \subset \tilde{\Lambda}_n$, such that the operator $\chi_{(-\infty,0)}(\tilde{L})\chi_J$ is trace class with kernel $\tilde{K}(x, y)$ (which can be obtained from the generalized eigenfunction expansion of $\tilde{L}$, see Section A.2). Then, we immediately get the strong convergence

$$\tilde{K}_n \chi_J \xrightarrow{s} \tilde{K}\chi_J.$$

Remark 1.1. Tao [20, Section 3.3] sketches the Borodin–Olshanski method, applied to the bulk and edge scaling of GUE, as a heuristic device. Because of the microlocal methods that he uses to calculate the projection $\chi_{(-\infty,0)}(\tilde{L})$, he puts his sketch under the headline "The Dyson and Airy kernels of GUE via semiclassical analysis".

---

[1] Based on the two-scale Plancherel–Rotach asymptotic of classical orthogonal polynomials or, methodologically more general, on the asymptotic of Riemann–Hilbert problems; see, e.g., Tracy and Widom [21, 22], Deift [6], Lubinsky [16], Johnstone [12, 13], Collins [5], Forrester [8], Anderson et al. [2], and Kuijlaars [14].

[2] By "$\xrightarrow{s}$" we denote the strong convergence of operators acting on L^2.

Scaling limits and other modes of convergence

Given that one just has to establish the convergence of the coefficients of a differential operator (instead of an asymptotic of its eigenfunctions), the Borodin–Olshanski method is an extremely simple device to determine all the scalings $x = \sigma_n \xi + \mu_n$ that would yield some *meaningful* limit $\tilde{K}_n \chi_J \to \tilde{K} \chi_J$, namely in the strong operator topology. Other modes of convergence have been studied in the literature, ranging from some weak convergence of k-point correlation functions over convergence of the kernel functions to the convergence of gap probabilities, that is,

$$\tilde{E}_n(J) = \det\big(I - \chi_J \tilde{K}_n \chi_J\big) \to \det\big(I - \chi_J \tilde{K} \chi_J\big) = \tilde{E}(J).$$

From a probabilistic point of view, the latter convergence is of particular interest and has been shown in at least three ways:

1. By Hadamard's inequality, convergence of the determinants follows directly from the locally uniform convergence of the kernels K_n [2, Lemma 3.4.5] and, for unbounded J, from additional large deviation estimates [2, Lemma 3.3.2]. This way, the limit gap probabilities in the bulk and soft edge scaling limit of GUE can rigorously be established (see, e.g., Anderson et al. [2, Sections 3.5 and 3.7]). Johansson [11, Lemma 3.1] gives some general conditions on a scaling of the K_n such that the determinant converges to the soft edge of GUE.

2. Since $A \mapsto \det(I - A)$ is continuous with respect to the trace class norm [18, Theorem 3.4], $\tilde{K}_n \chi_J \to \tilde{K} \chi_J$ in trace class norm would generally suffice. Such a convergence can be proved by factorizing the trace class operators into Hilbert–Schmidt operators and obtaining the L^2-convergence of the factorized kernels once more from locally uniform convergence, see the work of Johnstone [12, 13] on the scaling limits of the LUE/Wishart ensembles and on the limits of the JUE/MANOVA ensembles.

3. Since $\chi_J \tilde{K}_n \chi_J$ and $\chi_J \tilde{K} \chi_J$ are selfadjoint and positive semi-definite, yet another way is by observing that the convergence $\tilde{K}_n \chi_J \to \tilde{K} \chi_J$ in trace class norm is, for continuous kernels, equivalent [18, Theorem 2.20] to the combination of both, the convergence $\tilde{K}_n \chi_J \to \tilde{K} \chi_J$ in the *weak* operator topology and the convergence of the traces

$$\int_J \tilde{K}_n(\xi, \xi)\, d\xi \to \int_J \tilde{K}(\xi, \xi)\, d\xi. \tag{1.3}$$

Once again, these convergences follow from locally uniform convergence of the kernels; see Deift [6, Section 8.1] for an application of this method to the bulk scaling limit of GUE.

Since convergence in the strong operator topology implies convergence in the weak one, the Borodin–Olshanski method would thus establish the convergence of gap probabilities if we were only able to show condition (1.3) by some additional, similarly short and simply argument. Note that, by the ideal property of the trace class, condition (1.3) implies the same condition for all $J' \subset J$. We fall, however, short of conceiving a proof strategy for condition (1.3) that would be *independent* of all the laborious proofs of locally uniform convergence of the kernels.

Remark 1.2. Contrary to the discrete case considered by Borodin and Olshanski, it is also not immediate to infer from the strong convergence of the induced integral operators the *pointwise* convergence of the kernels. In Section 2 we will need only a single such instance, namely

$$\tilde{K}_n(0, 0) \to \tilde{K}(0, 0), \tag{1.4}$$

to prove a limit law $\tilde{\rho}_n(t)dt \xrightarrow{w} \tilde{\rho}(t)dt$ for the mean counting probability. Using mollified Dirac deltas, pointwise convergence would generally follow, for continuously differentiable $\tilde{K}_n(\xi, \eta)$, if we were able to bound, locally uniform, the gradient of $\tilde{K}_n(\xi,\eta)$. Then, by dominated convergence, criterion (1.3) would already be satisfied if we established an integrable bound of $\tilde{K}_n(\xi,\xi)$ on J. Since the scalings laws are, however, maneuvering just at the edge between trivial cases (i.e., zero limits) and divergent cases, it is conceivable that a proof of such bounds might not be significantly simpler than a proof of convergence of the gap probabilities itself.

The main result

To prepare we recall how an integral kernel $K_n(x,y)$ is getting covariantly transformed in the presence of an affine coordinate change $x = \sigma_n\xi + \mu_n$, $y = \sigma_n\eta + \mu_n$: by invariance of the 1-form

$$K_n(x,y)dy = \tilde{K}_n(\xi,\eta)d\eta$$

the transformed kernel $\tilde{K}$ is given by

$$\tilde{K}_n(\xi,\eta) = \sigma_n K_n(\sigma_n\xi + \mu_n, \sigma_n\eta + \mu_n). \tag{1.5}$$

Using the Borodin–Olshanski method, we will prove the following general result for selfadjoint Sturm–Liouville operators; a result that adds a further class of problems to the *universality* [14] of the Dyson, Airy, and Bessel kernel[3] in the bulk, soft-edge, and hard-edge scaling limits.

Theorem 1.3. *Let Λ be one of the three domains $\Lambda = (-\infty,\infty)$, $\Lambda = (0,\infty)$, or $\Lambda = (0,1)$, and let L_n be a selfadjoint realization on $L^2(\Lambda)$ of the formally selfadjoint Sturm–Liouville operator*[4]

$$-\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q_n(x) - \lambda_n$$

with coefficients $p, q_n \in C^\infty(\Lambda)$ such that $p(x) > 0$ for all $x \in \Lambda$. Assume that, for $t \in \Lambda$ and $n \to \infty$, there are asymptotic expansions

$$n^{-2\kappa'}\lambda_n \sim \omega, \qquad n^{-2\kappa'}q_n(n^\kappa t) \sim q_*(t), \qquad n^{2\kappa''}p(n^\kappa t) \sim p_*(t) > 0, \tag{1.6}$$

with a remainder that is of order $O(n^{-1})$ locally uniform in t, and exponents normalized by

$$\kappa + \kappa' + \kappa'' = 1, \qquad \kappa \geqslant 0, \tag{1.7}$$

where $\kappa < \frac{2}{3}$ if $\Lambda = (0,1)$. Further assume that these expansions can be differentiated[5]*, that the roots of $q_*(t) - \omega$ are simple, and that the spectral projection $K_n = \chi_{(-\infty,0)}(L_n)$ is normalized by*

$$\operatorname{tr} K_n = n.$$

Let a scaling by $x = \sigma_n\xi + \mu_n$ induce the transformed projection kernel $\tilde{K}_n$ according to (1.5).

Then, depending on particular choices of σ_n and μ_n, the following three scaling limits hold.

- *Bulk scaling limit: given $t \in \Lambda$ with $q_*(t) < \omega$, the scaling parameters*

$$\sigma_n = \frac{n^{\kappa-1}}{\tilde{\rho}(t)}, \qquad \mu_n = n^\kappa t,$$

[3]For the definitions of the kernels K_{Dyson}, K_{Airy}, K_{Bessel} see (A.3), (A.4) and (A.5).

[4]Since, in this paper, we consider always a particular selfadjoint realization of a formal differential operator, we will use the same letter to denote both.

[5]We say that an expansion $f_n(t) - f(t) = O(1/n)$ can be differentiated if $f_n'(t) - f'(t) = O(1/n)$.

where

$$\tilde\rho(t) = \frac{1}{\pi}\sqrt{\frac{(\omega - q_*(t))_+}{p_*(t)}}, \tag{1.8}$$

yield, for a bounded interval J, the strong limit

$$\tilde K_n \chi_J \xrightarrow{s} K_{\text{Dyson}}\chi_J.$$

At $\xi = 0$, the mean counting probability density $\rho_n(x) = n^{-1}K_n(x,x)$ transforms to the new variable t as

$$\tilde\rho_n(t) = n^\kappa \rho(n^\kappa t).$$

Under condition (1.4), and if $\tilde\rho$ as defined in (1.8) has unit mass on Λ, there is the limit law

$$\tilde\rho_n(t)dt \xrightarrow{w} \tilde\rho(t)dt.$$

- *Soft-edge scaling limit: given $t_* \in \Lambda$ with $q_*(t_*) = \omega$, the scaling parameters*

$$\sigma_n = n^{\kappa - \frac{2}{3}}\left(\frac{p_*(t_*)}{q_*'(t_*)}\right)^{1/3}, \qquad \mu_n = n^\kappa t_*,$$

yield, for $s \in \mathbb{R}$ and a (not necessarily bounded) interval $J \subset (s,\infty)$, the strong limit

$$\tilde K_n \chi_J \xrightarrow{s} K_{\text{Airy}}\chi_J.$$

- *Hard-edge scaling limit: given that $\Lambda = (0,\infty)$ or $\Lambda = (0,1)$ with*

$$p(0) = 0, \qquad p'(0) > 0, \qquad q_n(x) = q(x) = \gamma^2 x^{-1} + O(1), \qquad x \to 0, \tag{1.9}$$

the scaling parameters

$$\sigma_n = \frac{p'(0)}{4\omega n^{2\kappa'}}, \qquad \mu_n = 0,$$

yield, for a bounded interval $J \subset (0,\infty)$, the strong limit[6]

$$\tilde K_n \chi_J \xrightarrow{s} K^{(\alpha)}_{\text{Bessel}}\chi_J\Big|_{\alpha = 2\gamma/\sqrt{p'(0)}}. \tag{1.11}$$

Remark 1.4. Whether the interval J in the strong operator limit $\tilde K_n\chi_J \xrightarrow{s} K\chi_J$ can be chosen unbounded or not depends on whether the limit operator $K\chi_J$ is trace class or not (see the explicit formulae of the traces given in the appendix for each of the three limits): only in the former case we get a representation of the scaling limit in terms of a particular integral kernel, cf. Theorem A.3. Note that it is impossible to use $J = \Lambda$ since $\operatorname{tr} K_n = n \to \infty$.

[6] Here, if $0 \leqslant \alpha < 1$, the selfadjoint realization L_n is defined by means of the boundary condition

$$2xu'(x) - \alpha u(x) = o(x^{-\alpha/2}), \qquad x \to 0. \tag{1.10}$$

Outline of the paper

The proof of Theorem 1.3 is subject of Section 2. In Section 3 we apply it to the classical orthogonal polynomials, which yields a short and unified derivation of the known formulae for the scaling limits for the classical random matrix ensembles with unitary invariance (GUE, LUE/Wishart, JUE/MANOVA). In fact, by a result of Tricomi, the only input needed is the weight function w of the orthogonal polynomials; from there one gets in a purely formula based fashion (by simple manipulations which can easily be coded in any computer algebra system), first, to the coefficients p and q_n as well as to the eigenvalues λ_n of the Sturm–Liouville operator L_n and next, by applying Theorem 1.3, to the particular scaling limits.

To emphasize that our main result and its application is largely independent of concretely identifying the limit projection kernel $\tilde{K}$, we postpone this identification to Lemmas A.5, A.7 and A.9: there, using generalized eigenfunction expansions, we calculate the Dyson, Airy, and Bessel kernels directly from the limit differential operator $\tilde{L}$.

2 Proof of the main result for Sturm–Liouville operators

We start the proof of Theorem 1.3 with some preparatory steps before we deal with the particular scaling limits. Since L_n is a selfadjoint realization on $L^2(\Lambda)$ of the Sturm–Liouville operator

$$L_n = -\frac{d}{dx}\left(p(x)\frac{d}{dx}\right) + q_n(x) - \lambda_n$$

with $p, q_n \in C^\infty(\Lambda)$ and $p(x) > 0$ for $x \in \Lambda$, we have $C_0^\infty(\Lambda) \subset D(L_n)$.

Preparatory Step 1: transformation

The scaling

$$x = \sigma_n \xi + \mu_n, \qquad \sigma_n \neq 0,$$

maps $x \in \Lambda$ bijectively to $\xi \in \tilde{\Lambda}_n$. Since such an affine coordinate transform just induces a *unitary equivalence* of integral and differential operators, the spectral projection relation

$$K_n = \chi_{(-\infty,0)}(L_n)$$

is left invariant if the kernel $K_n(x,y)$ is transformed according to (1.5) and the differential operator L_n is transformed using $d/dx = \sigma_n^{-1} d/d\xi$ as

$$-\frac{1}{\sigma_n^2}\frac{d}{d\xi}\left(p(\sigma_n\xi+\mu_n)\frac{d}{d\xi}\right) + q_n(\sigma_n\xi+\mu_n) - \lambda_n.$$

Since the spectral projection to the negative part of the spectrum of a differential operator is left invariant if we multiply that operator by some *positive* constant $\tau_n\sigma_n^2$, $\tau_n > 0$, we see that

$$\tilde{K}_n = \chi_{(-\infty,0)}(\tilde{L}_n),$$

where the transformed differential operator is given finally by

$$\tilde{L}_n = -\frac{d}{d\xi}\left(\tilde{p}_n(\xi)\frac{d}{d\xi}\right) + \tilde{q}_n(\xi)$$

with coefficients

$$\tilde{p}_n(\xi) = \tau_n p(\sigma_n\xi+\mu_n), \qquad \tilde{q}_n(\xi) = \tau_n\sigma_n^2\left(q_n(\sigma_n\xi+\mu_n) - \lambda_n\right). \tag{2.1}$$

Preparatory Step 2: strong operator limit

Suppose the transformed domain $\tilde{\Lambda}_n = (a_n, b_n)$ satisfies $a_n \to a$, $b_n \to b$. Then, with $\tilde{\Lambda} = (a, b)$ we have that, eventually, $C_0^\infty(\tilde{\Lambda}) \subset D(\tilde{L}_n)$. Further, suppose that the coefficients of $\tilde{L}_n$ converge locally uniform in $\tilde{\Lambda}$ as (where the limit of $\tilde{p}_n(\xi)$ can be differentiated)

$$\tilde{p}_n(\xi) \to \tilde{p}(\xi), \qquad \tilde{q}_n(\xi) \to \tilde{q}(\xi),$$

such that the limit coefficients $\tilde{p} > 0$ and $\tilde{q}$ are smooth functions and

$$\tilde{L} = -\frac{d}{d\xi}\left(\tilde{p}(\xi)\frac{d}{d\xi}\right) + \tilde{q}(\xi) \tag{2.2}$$

defines a Sturm–Liouville operator that is essentially selfadjoint on $C_0^\infty(\tilde{\Lambda}) \subset L^2(\tilde{\Lambda})$. Then, by dominated convergence, we get the convergence $\tilde{L}_n u \to \tilde{L}u$ in $L^2(\tilde{\Lambda})$ for each test function u in the core $C_0^\infty(\tilde{\Lambda})$. Hence, by Theorem A.1 we have the strong operator convergence

$$\tilde{K}_n \chi_J \xrightarrow{s} \chi_{(-\infty,0)}(\tilde{L})\chi_J$$

if $0 \notin \sigma_{\text{pp}}(L)$ and, eventually, $J \subset \tilde{\Lambda}_n$. In the particular cases considered in the following limit steps of the proof, the spectrum of $\tilde{L}$ is always absolutely continuous, that is, $\sigma_{\text{pp}}(L) = \varnothing$. Finally, by Theorem A.3, under the finite trace condition mentioned already in Remark 1.4, there is an integral kernel $\tilde{K}$ such that

$$\chi_{(-\infty,0)}(\tilde{L})\chi_J = \tilde{K}\chi_J,$$

which finishes the proof of a strong operator convergence in general.

Preparatory Step 3: Taylor expansions of the coefficients

The case $\mu_n = n^\kappa t$

Suppose that $t \in \Lambda$ is fixed. The choice $\tau_n = 1/p(\mu_n) > 0$ is then admissible and we get, if

$$\sigma_n = o\big(n^{\kappa-1/2}\big),$$

from (1.6), (1.7), and (2.1) by a Taylor expansion

$$\tilde{p}_n(\xi) = 1 + o(1), \qquad \tilde{q}_n(\xi) = \frac{\sigma_n^2 n^{2-2\kappa}}{p_*(t)}\big(q_*(t) - \omega + \sigma_n n^{-\kappa} q_*'(t)\cdot\xi\big) + o(1), \tag{2.3}$$

which holds locally uniform in $\xi \in \tilde{\Lambda}$ (where the expansion of $\tilde{p}_n(\xi)$ can be differentiated).

The case $\mu_n = 0$

Suppose that the assumptions in (1.9) are met. If $\sigma_n \to 0^+$, the choice $\tau_n = 4\sigma_n/p'(0) > 0$ is admissible and we get from (2.1) by a Taylor expansion

$$\tilde{p}_n(\xi) = 4\xi + o(1), \qquad \tilde{q}_n(\xi) = \frac{4\gamma^2}{p'(0)\xi} - \frac{4\sigma_n\lambda_n}{p'(0)} + o(1), \tag{2.4}$$

which holds locally uniform in $\xi \in \tilde{\Lambda}$ (where the expansion of $\tilde{p}_n(\xi)$ can be differentiated).

Limit Step 1: bulk scaling limit

If $q_*(t) \neq \omega$, by inserting

$$\sigma_n = \sigma_n(t) = \pi n^{\kappa-1}\sqrt{\frac{p_*(t)}{|\omega - q_*(t)|}}$$

we read off from (2.3) the limit coefficients $\tilde{p}(\xi) = 1$ and $\tilde{q}(\xi) = -s\pi^2$, where $s = \operatorname{sign}(\omega - q_*(t))$; that is, the limit differential operator (2.2) is given by

$$\tilde{L} = -\frac{d^2}{d\xi^2} - s\pi^2.$$

Note that, for the domains Λ and the values of κ considered, we have $\tilde{\Lambda} = (-\infty, \infty)$.

Lemma A.5 states that $\tilde{L}$ is essentially selfadjoint on $C_0^\infty(\tilde{\Lambda})$ and that its unique selfadjoint extension has absolutely continuous spectrum: $\sigma(\tilde{L}) = \sigma_{\text{ac}}(\tilde{L}) = [-s\pi^2, \infty)$. Thus, for $s = -1$, the spectral projection $\chi_{(-\infty,0)}(\tilde{L})$ is zero. For $s = 1$, the spectral projection can be calculated by a generalized eigenfunction expansion, yielding the *Dyson* kernel (A.3).

We will see in the next step that the dichotomy between $s = \pm 1$ is also reflected in the structure of the support of the limit law $\tilde{\rho}$.

Limit Step 2: limit law

The result for the bulk scaling limit allows, in passing, to calculate a limit law of the mean counting probability density $\rho_n(x) = n^{-1}K_n(x,x)$: we observe that $x = n^\kappa t$ transforms the density $\rho_n(x)$ into

$$\tilde{\rho}_n(t) = n^{\kappa-1}K_n(n^\kappa t, n^\kappa t) = \frac{n^{\kappa-1}}{\sigma_n(t)}\tilde{K}_n(0,0) = \frac{1}{\pi}\sqrt{\frac{|\omega - q_*(t)|}{p_*(t)}}\tilde{K}_n(0,0).$$

Thus, to get to a limit, we have to *assume* condition (1.4), so that a pointwise rendering of the bulk scaling limit just considered yields[7]

$$\tilde{K}_n(0,0) \to [q_*(t) < \omega]K_{\text{Dyson}}(0,0) = [q_*(t) < \omega].$$

This way we get

$$\tilde{\rho}_n(t) \to \tilde{\rho}(t) = \frac{1}{\pi}\sqrt{\frac{(\omega - q_*(t))_+}{p_*(t)}}.$$

Hence, by Helly's selection theorem, the probability measure $\tilde{\rho}_n(t)dt$ converges vaguely to $\tilde{\rho}(t)dt$, which is, in general, just a sub-probability measure. If, however, it is checked that $\tilde{\rho}(t)dt$ has unit mass, the convergence is weak.

Limit Step 3: soft-edge scaling limit

If $q_*(t_*) = \omega$, by inserting[8]

$$\sigma_n = \sigma_n(t_*) = n^{\kappa-2/3}\left(\frac{p_*(t_*)}{q_*'(t_*)}\right)^{1/3}$$

[7] The Iverson bracket $[S]$ stands for 1 if the statement S is true, 0 otherwise.

[8] Note that, by the assumption made on the simplicity of the roots of $q_*(t) - \omega$, we have $q_*'(t_*) \neq 0$.

we read off from (2.3) the limit coefficients $\tilde{p}(\xi) = 1$ and $\tilde{q}(\xi) = \xi$; that is, the limit differential operator (2.2) is

$$\tilde{L} = -\frac{d^2}{d\xi^2} + \xi.$$

Note that, for the domains Λ and the values of κ considered, we have $\tilde{\Lambda} = (-\infty, \infty)$.

Lemma A.7 states that $\tilde{L}$ is essentially selfadjoint on $C_0^\infty(\tilde{\Lambda})$ and that its unique selfadjoint extension has absolutely continuous spectrum: $\sigma(\tilde{L}) = \sigma_{\mathrm{ac}}(\tilde{L}) = (-\infty, \infty)$. The spectral projection can be calculated by a generalized eigenfunction expansion, yielding the *Airy* kernel (A.4).

Limit Step 4: hard-edge scaling limit

For $\Lambda = (0, \infty)$ or $\Lambda = (0, 1)$, we take a scaling

$$x = \sigma_n \xi,$$

with $\sigma_n \to 0^+$ appropriately chosen, to explore the vicinity of the "hard edge" $x = 0$; note that such a scaling yields $\tilde{\Lambda} = (0, \infty)$. We make the assumptions stated in (1.9). By inserting

$$\sigma_n = n^{-2\kappa'} \frac{p'(0)}{4\omega}$$

we read off from (2.4), using (1.6), the limit coefficients $\tilde{p}(\xi) = 4\xi$ and $\tilde{q}(\xi) = \alpha^2\xi^{-1} - 1$, where α is defined as in (1.11); that is, the limit differential operator (2.2) is given by

$$\tilde{L} = -4\frac{d}{d\xi}\left(\xi \frac{d}{d\xi}\right) + \alpha^2\xi^{-1} - 1\bigg|_{\alpha = 2\gamma/\sqrt{p'(0)}}.$$

If $\alpha \geqslant 1$, Lemma A.9 states that the limit $\tilde{L}$ is essentially selfadjoint on $C_0^\infty(\tilde{\Lambda})$ and that the spectrum of its unique selfadjoint extension is absolutely continuous: $\sigma(\tilde{L}) = \sigma_{\mathrm{ac}}(\tilde{L}) = [-1, \infty)$. The spectral projection can be calculated by a generalized eigenfunction expansion, yielding the *Bessel* kernel (A.5).

Remark 2.1. The theorem also holds in the case $0 \leqslant \alpha < 1$ if the particular selfadjoint realization L_n is defined by the boundary condition (1.10), see Remark A.10.

3 Application to classical orthogonal polynomials

In this section we apply Theorem 1.3 to the kernels associated with the classical orthogonal polynomials, that is, the Hermite, Laguerre, and Jacobi polynomials. In random matrix theory, the thus induced determinantal processes are modeled by the spectra of the Gaussian unitary ensemble (GUE), the Wishart or Laguerre unitary ensemble (LUE), and the MANOVA (multivariate analysis of variance) or Jacobi unitary ensemble (JUE).

To prepare the study of the individual cases, we first discuss their common structure. Let $P_n(x)$ be the sequence of classical orthogonal polynomials belonging to the weight function $w(x)$ on the (not necessarily bounded) interval (a, b). We normalize $P_n(x)$ such that $\langle \phi_n, \phi_n \rangle = 1$, where $\phi_n(x) = w(x)^{1/2} P_n(x)$. The functions ϕ_n form a complete orthogonal set in $L^2(a, b)$; conceptual proofs of the completeness can be found, e.g., in Andrews, Askey and Roy [3] (Section 5.7 for the Jacobi polynomials, Section 6.5 for the Hermite and Laguerre polynomials).

By a result of Tricomi [7, Section 10.7], the $P_n(x)$ satisfy the eigenvalue problem

$$-\frac{1}{w(x)}\frac{d}{dx}\left(p(x)w(x)\frac{d}{dx}P_n(x)\right) = \lambda_n P_n(x), \qquad \lambda_n = -n\big(r' + \tfrac{1}{2}(n+1)p''\big),$$

where $p(x)$ is a *quadratic* polynomial[9] and $r(x)$ a *linear* polynomial such that

$$\frac{w'(x)}{w(x)} = \frac{r(x)}{p(x)}.$$

In terms of ϕ_n, a brief calculation shows that

$$-\frac{d}{dx}\left(p(x)\frac{d}{dx}\phi_n(x)\right) + q(x)\phi_n(x) = \lambda_n\phi_n(x), \qquad q(x) = \frac{r(x)^2}{4p(x)} + \frac{r'(x)}{2}.$$

Therefore, by the completeness of the ϕ_n, the formally selfadjoint Sturm–Liouville operator $L = -\frac{d}{dx}p(x)\frac{d}{dx} + q(x)$ has a particular selfadjoint realization on $L^2(a,b)$ (which we continue to denote by the letter L) with spectrum

$$\sigma(L) = \{\lambda_0, \lambda_1, \lambda_2, \ldots\}$$

and corresponding eigenfunctions ϕ_n. Hence, if the eigenvalues are, eventually, strictly increasing, the projection kernel (1.1) defines an integral operator K_n with $\operatorname{tr} K_n = n$ such that, eventually,

$$K_n = \chi_{(-\infty,0)}(L_n), \qquad L_n = L - \lambda_n.$$

Note that this relation remains true if we choose to make some parameters of the weight w (and, therefore, of the functions ϕ_j) to depend on n. For the scaling limits of K_n, we are now in the realm of Theorem 1.3: given the weight $w(x)$ as the only input all the other quantities can now be obtained simply by routine calculations.

Hermite polynomials

The weight is $w(x) = e^{-x^2}$ on $\Lambda = (-\infty,\infty)$; hence

$$p(x) = 1, \qquad r(x) = -2x, \qquad q(x) = x^2 - 1, \qquad \lambda_n = 2n,$$

and, therefore,

$$\kappa = \kappa' = \tfrac{1}{2}, \qquad \kappa'' = 0, \qquad p_*(t) = 1, \qquad q_*(t) = t^2, \qquad \omega = 2.$$

Theorem 1.3 is applicable and we directly read off the following well-known scaling limits of the GUE (see, e.g., [2, Chapter 3]):

- bulk scaling limit: if $-\sqrt{2} < t < \sqrt{2}$, the transformation

  $$x = \frac{\pi\xi}{n^{1/2}\sqrt{2-t^2}} + n^{1/2}t$$

  induces $\tilde{K}_n$ with a strong limit given by the *Dyson kernel*;
- limit law: the transformation $x = n^{1/2}t$ induces the mean counting probability density $\tilde{\rho}_n$ with a weak limit given by the *Wigner semicircle law*

  $$\tilde{\rho}(t) = \frac{1}{\pi}\sqrt{(2-t^2)_+};$$

- soft-edge scaling limit: the transformation

  $$x = \pm\big(2^{-1/2}n^{-1/6}\xi + \sqrt{2n}\big)$$

  induces $\tilde{K}_n$ with a strong limit given by the *Airy kernel*.

[9] With the sign chosen such that $p(x) > 0$ for $x \in (a,b)$.

Laguerre polynomials

The weight is $w(x) = x^\alpha e^{-x}$ on $\Lambda = (0,\infty)$; hence

$$p(x) = x, \qquad r(x) = \alpha - x, \qquad q(x) = \frac{(\alpha - x)^2}{4x} - \frac{1}{2}, \qquad \lambda_n = n.$$

In random matrix theory, the corresponding determinantal point process is modeled by the spectra of complex $n \times n$ Wishart matrices with a dimension parameter $m \geqslant n$; the Laguerre parameter α is then given by $\alpha = m - n \geqslant 0$. Of particular interest in statistics [12] is the simultaneous limit $m, n \to \infty$ with

$$\frac{m}{n} \to \theta \geqslant 1,$$

for which we get

$$\kappa = 1, \qquad \kappa' = \frac{1}{2}, \qquad \kappa'' = -\frac{1}{2}, \qquad p_*(t) = t, \qquad q_*(t) = \frac{(\theta - 1 - t)^2}{4t}, \qquad \omega = 1.$$

Note that

$$\omega - q_*(t) = \frac{(t_+ - t)(t - t_-)}{4t}, \qquad t_\pm = \left(\sqrt{\theta} \pm 1\right)^2.$$

Theorem 1.3 is applicable and we directly read off the following well-known scaling limits of the Wishart ensemble [12]:

- bulk scaling limit: if $t_- < t < t_+$,

  $$x = \frac{2\pi t \xi}{\sqrt{(t_+ - t)(t - t_-)}} + nt$$

  induces $\tilde{K}_n$ with a strong limit given by the *Dyson kernel*;
- limit law: the scaling $x = nt$ induces the mean counting probability density $\tilde{\rho}_n$ with a weak limit given by the *Marchenko–Pastur law*

  $$\tilde{\rho}(t) = \frac{1}{2\pi t}\sqrt{((t_+ - t)(t - t_-))_+};$$
- soft-edge scaling limit: with signs chosen consistently as either $+$ or $-$,

  $$x = \pm n^{1/3} \theta^{-1/6} t_\pm^{2/3} \xi + n t_\pm \tag{3.1}$$

  induces $\tilde{K}_n$ with a strong limit given by the *Airy kernel*.

Remark 3.1. The scaling (3.1) is better known in the asymptotically equivalent form

$$x = \sigma \xi + \mu, \qquad \mu = \left(\sqrt{m} \pm \sqrt{n}\right)^2, \qquad \sigma = \left(\sqrt{m} \pm \sqrt{n}\right)\left(\frac{1}{\sqrt{m}} \pm \frac{1}{\sqrt{n}}\right)^{1/3},$$

which is obtained from (3.1) by replacing θ with m/n, see [12, p. 305].

In the case $\theta = 1$, which implies $t_- = 0$, the lower soft-edge scaling (3.1) breaks down and has to be replaced by a scaling at the *hard* edge:

- hard-edge scaling limit: if $\alpha = m - n$ is a constant[10], $x = \xi/(4n)$ induces $\tilde{K}_n$ with a strong limit given by the *Bessel kernel* $K_{\text{Bessel}}^{(\alpha)}$.

[10] By Remark 2.1, there is no need to restrict ourselves to $\alpha \geqslant 1$: since $\phi_n(x) = x^\alpha \tilde{\phi}_n(x)$ with $\tilde{\phi}_n(x)$ extending smoothly to $x = 0$, we have, for $\alpha \geqslant 0$,

$$x^{\alpha/2}(2x\phi_n'(x) - \alpha\phi_n(x)) = 2x^{1+\alpha}\tilde{\phi}_n'(x) = O(x), \qquad x \to 0.$$

Hence, the selfadjoint realization L_n is compatible with the boundary condition (1.10).

Jacobi polynomials

The weight is $w(x) = x^\alpha(1-x)^\beta$ on $\Lambda = (0,1)$; hence

$$p(x) = x(1-x), \qquad r(x) = \alpha - (\alpha+\beta)x, \qquad q(x) = \frac{(\alpha-(\alpha+\beta)x)^2}{4x(1-x)} - \frac{\alpha+\beta}{2},$$

and

$$\lambda_n = n(n+\alpha+\beta+1).$$

In random matrix theory, the corresponding determinantal point process is modeled by the spectra of complex $n\times n$ MANOVA matrices with dimension parameters $m_1, m_2 \geqslant n$; the Jacobi parameters α, β are then given by $\alpha = m_1 - n \geqslant 0$ and $\beta = m_2 - n \geqslant 0$. Of particular interest in statistics [13] is the simultaneous limit $m_1, m_2, n \to \infty$ with

$$\frac{m_1}{m_1+m_2} \to \theta \in (0,1), \qquad \frac{n}{m_1+m_2} \to \tau \in (0, 1/2],$$

for which we get

$$\kappa = \kappa'' = 0, \qquad \kappa' = 1, \qquad p_*(t) = t(1-t), \qquad q_*(t) = \frac{(\theta-\tau-(1-2\tau)t)^2}{4\tau^2 t(1-t)}, \qquad \omega = \frac{1-\tau}{\tau}.$$

Note that

$$\omega - q_*(t) = \frac{(t_+-t)(t-t_-)}{4\tau^2 t(1-t)}, \qquad t_\pm = \left(\sqrt{\theta(1-\tau)} \pm \sqrt{\tau(1-\theta)}\right)^2.$$

Theorem 1.3 is applicable and we directly read off the following (less well-known) scaling limits of the MANOVA ensemble [5, 13]:

- bulk scaling limit: if $t_- < t < t_+$,

$$x = \frac{2\pi\tau t(1-t)\xi}{n\sqrt{(t_+-t)(t-t_-)}} + t$$

  induces $\tilde{K}_n$ with a strong limit given by the *Dyson kernel*;

- limit law: (because of $\kappa = 0$ there is no transformation here) the mean counting probability density ρ_n has a weak limit given by the law [23]

$$\rho(t) = \frac{1}{2\pi\tau t(1-t)}\sqrt{((t_+-t)(t-t_-))_+};$$

- soft-edge scaling limit: with signs chosen consistently as either $+$ or $-$,

$$x = \pm n^{-2/3}\frac{(\tau t_\pm(1-t_\pm))^{2/3}}{(\tau\theta(1-\tau)(1-\theta))^{1/6}}\xi + t_\pm \tag{3.2}$$

  induces $\tilde{K}_n$ with a strong limit given by the *Airy kernel*.

Remark 3.2. Johnstone [13, p. 2651] gives the soft-edge scaling in terms of a trigonometric parametrization of θ and τ. By putting

$$\theta = \sin^2\frac{\phi}{2}, \qquad \tau = \sin^2\frac{\psi}{2},$$

we immediately get

$$t_\pm = \sin^2 \frac{\phi \pm \psi}{2}$$

and (3.2) becomes

$$x = \pm\sigma_\pm \xi + t_\pm, \qquad \sigma_\pm = n^{-2/3}\left(\frac{\tau^2 \sin^4(\phi \pm \psi)}{4 \sin\phi \sin\psi}\right)^{1/3}.$$

In the case $\theta = \tau = 1/2$, which is equivalent to $m_1/n, m_2/n \to 1$, we have $t_- = 0$ and $t_+ = 1$. Hence, the lower and the upper soft-edge scaling (3.2) break down and have to be replaced by a scaling at the *hard* edges:

- hard-edge scaling limit: if $\alpha = m_1 - n$, $\beta = m_2 - n$ are constants[11], $x = \xi/(4n^2)$ induces $\tilde{K}_n$ with a strong limit given by the *Bessel kernel* $K^{(\alpha)}_{\text{Bessel}}$; by symmetry, the *Bessel kernel* $K^{(\beta)}_{\text{Bessel}}$ is obtained for $x = 1 - \xi/(4n^2)$.

A Appendices

A.1 Generalized strong convergence

The notion of *strong resolvent convergence* [24, Section 9.3] links the convergence of differential operators, tested for an appropriate class of smooth functions, to the strong convergence of their spectral projections. We recall a slight generalization of that concept, which allows the underlying Hilbert space to vary.

Specifically we consider, on an interval (a, b) (not necessarily bounded) and on a sequence of subintervals $(a_n, b_n) \subset (a, b)$ with $a_n \to a$ and $b_n \to b$, selfadjoint operators

$$L\colon\ D(L) \subset L^2(a, b) \to L^2(a, b), \qquad L_n\colon\ D(L_n) \subset L^2(a_n, b_n) \to L^2(a_n, b_n).$$

By means of the natural embedding (that is, extension by zero) we take $L^2(a_n, b_n) \subset L^2(a, b)$; the multiplication operator induced by the characteristic function $\chi_{(a_n,b_n)}$, which we will denote by the same symbol, constitutes the orthogonal projection of $L^2(a, b)$ onto $L^2(a_n, b_n)$. Following Stolz and Weidmann [19, Section 2], we say that L_n converges to L in the sense of *generalized strong convergence* (gsc), if for some $z \in \mathbb{C} \setminus \mathbb{R}$, and hence, a forteriori, for all such z,

$$R_z(L_n)\chi_{(a_n,b_n)} \xrightarrow{s} R_z(L), \qquad n \to \infty,$$

in the strong operator topology of $L^2(a, b)$.[12]

Theorem A.1 (Stolz and Weidmann [19, Theorem 4/5])**.** *Let the selfadjoint operators L_n and L satisfy the assumptions stated above and let C be a core of L such that, eventually, $C \subset D(L_n)$.*

(i) If $L_n u \to Lu$ for all $u \in C$, then $L_n \xrightarrow{gsc} L$.

(ii) If $L_n \xrightarrow{gsc} L$ and if the endpoints of the interval $\Delta \subset \mathbb{R}$ do not belong to the pure point spectrum $\sigma_{\text{pp}}(L)$ of L, the spectral projections to Δ converge as

$$\chi_\Delta(L_n)\chi_{(a_n,b_n)} \xrightarrow{s} \chi_\Delta(L).$$

[11] For the cases $0 \leqslant \alpha < 1$ and $0 \leqslant \beta < 1$, see the justification of the limit given in footnote 10.

[12] We denote by $R_z(L) = (L - z)^{-1}$ the resolvent of an operator L.

A.2 Generalized eigenfunction expansion of Sturm–Liouville operators

Let L be a formally selfadjoint Sturm–Liouville operator on the interval (a,b),

$$Lu = -(pu')' + qu,$$

with smooth coefficient functions $p>0$ and q. We have the *limit point case* (LP) at the boundary point a if there is some $c\in(a,b)$ and some $z\in\mathbb{C}$ such that there exists at least one solution of $(L-z)u=0$ in (a,b) for which $u\notin L^2(a,c)$; otherwise, we have the *limit circle case* (LC) at a. According to the Weyl alternative [24, Theorem 8.27], in the LP case there exists actually for *all* $c\in(a,b)$ and *all* $z\in\mathbb{C}$ at least one solution of $(L-z)u=0$ in (a,b) for which $u\notin L^2(a,c)$; yet, if $z\in\mathbb{C}\setminus\mathbb{R}$, there is a one-dimensional space of solutions u of $(L-z)u=0$ for which there is nevertheless $u\in L^2(a,c)$. The same structure and notion applies to the boundary point b.

Theorem A.2. *Let L be a formally selfadjoint Sturm–Liouville operator on the interval (a,b) as defined above. If there is the LP case at a and b, then L is essentially self-adjoint on the domain $C_0^\infty(a,b)$ and, for $z\in\mathbb{C}\setminus\mathbb{R}$, the resolvent $R_z(L)=(L-z)^{-1}$ of its unique selfadjoint extension (which we continue to denote by the letter L) is of the form*

$$R_z(L)\phi(x) = \frac{1}{W(u_a,u_b)}\left(u_b(x)\int_a^x u_a(y)\phi(y)\,dy + u_a(x)\int_x^b u_b(y)\phi(y)\,dy\right). \tag{A.1}$$

Here u_a and u_b are the non-vanishing solutions of the equation $(L-z)u=0$, uniquely determined up to a factor by the conditions $u_a\in L^2(a,c)$ and $u_b\in L^2(c,b)$ for some $c\in(a,b)$, and W denotes the Wronskian

$$W(u_a,u_b) = p(x)(u_a'(x)u_b(x) - u_a(x)u_b'(x)),$$

which is a constant for $x\in(a,b)$.

A more general formulation of this theorem, which includes also the LC case, can be found, e.g., in [24, Theorem 8.26/8.29]; see [25, pp. 41–42] for a proof that $C_0^\infty(a,b)$ is a core of L if the coefficients are smooth. In the following, we write (A.1) briefly in the form

$$R_z(L)\phi(x) = \int_a^b G_z(x,y)\phi(y)\,dy$$

with the *Green's kernel*

$$G_z(x,y) = \frac{1}{W(u_a,u_b)}\begin{cases} u_b(x)u_a(y), & x>y,\\ u_a(x)u_b(y), & \text{otherwise.}\end{cases}$$

If the imaginary part of $G_z(x,y)$ has finite boundary values as z approaches the real line from above, there is a simple formula for the spectral projection associated with L that often applies if the spectrum of L is absolutely continuous.

Theorem A.3.

(i) *Assume that there exits, as $\epsilon\to 0^+$, the limit*

$$\pi^{-1}\,\mathrm{Im}\,G_{\lambda+i\epsilon}(x,y)\to K_\lambda(x,y),$$

locally uniform in $x,y\in(a,b)$ for each $\lambda\in\mathbb{R}$ except for some isolated points λ for which the limit is replaced by

$$\epsilon\,\mathrm{Im}\,G_{\lambda+i\epsilon}(x,y)\to 0.$$

Then the spectrum is absolutely continuous, $\sigma(L) = \sigma_{ac}(L)$*, and, for a Borel set* Δ*,*

$$\langle \chi_\Delta(L)\phi, \psi\rangle = \int_\Delta \langle K_\lambda \phi, \psi\rangle d\lambda, \qquad \phi, \psi \in C_0^\infty(a,b). \tag{A.2}$$

(*ii*) *Assume further, for some* $(a',b') \subset (a,b)$*, that*

$$\int_{a'}^{b'}\int_{a'}^{b'}\left(\int_\Delta |K_\lambda(x,y)|d\lambda\right)^2 dxdy < \infty.$$

Then $\chi_\Delta(L)\chi_{(a',b')}$ *is a Hilbert–Schmidt operator on* $L^2(a,b)$ *with kernel*

$$\chi_{(a',b')}(y)\int_\Delta K_\lambda(x,y)d\lambda.$$

If $\int_\Delta K_\lambda(x,y)d\lambda$ *is a continuous function of* $x, y \in (a',b')$*,* $\chi_\Delta(L)\chi_{(a',b')}$ *is a trace class operator with trace*

$$\operatorname{tr}\chi_\Delta(L)\chi_{(a',b')} = \int_{a'}^{b'}\int_\Delta K_\lambda(x,x)d\lambda dx.$$

Proof. With E denoting the spectral resolution of the selfadjoint operator L, we observe that, for a given $\phi \in C_0^\infty(a,b)$, the Borel–Stieltjes transform of the positive measure $\mu_\phi(\lambda) = \langle E(\lambda)\phi, \phi\rangle$ can be simply expressed in terms of the resolvent as follows, see [15, Section 32.1]:

$$\int_{-\infty}^{\infty}\frac{d\mu_\phi(\lambda)}{\lambda - z} = \langle R_z(L)\phi, \phi\rangle.$$

If we take $z = \lambda + i\epsilon$ and let $\epsilon \to 0^+$, we obtain by the locally uniform convergence of the integral kernel of R_z that there exits either the limit

$$\pi^{-1}\operatorname{Im}\langle R_{\lambda+i\epsilon}(L)\phi, \phi\rangle \to \langle K_\lambda\phi, \phi\rangle$$

or, at isolated points λ,

$$\epsilon \operatorname{Im}\langle R_{\lambda+i\epsilon}(L)\phi, \phi\rangle \to 0.$$

By a theorem of de la Vallée–Poussin [18, Theorem 11.6(ii/iii)], the singular part of μ_ϕ vanishes, $\mu_{\phi,\text{sing}} = 0$; by Plemelj's reconstruction the absolutely continuous part satisfies [18, Theorem 11.6(iv)]

$$d\mu_{\phi,\text{ac}}(\lambda) = \langle K_\lambda\phi, \phi\rangle d\lambda.$$

Since $C_0^\infty(a,b)$ is dense in $L^2(a,b)$, approximation shows that $E_{\text{sing}} = 0$, that is, $\sigma(L) = \sigma_{\text{ac}}(L)$. Since $\langle\chi_\Delta(L)\phi, \phi\rangle = \int_\Delta d\mu_\phi(\lambda)$, we thus get, by the symmetry of the bilinear expressions, the representation (A.2), which finishes the proof of (i). The Hilbert–Schmidt part of part (ii) follows using the Cauchy–Schwarz inequality and Fubini's theorem and yet another density argument; the trace class part follows from [9, Theorem IV.8.3] since $\chi_{(a',b')}\chi_\Delta(L)\chi_{(a',b')}$ is a selfadjoint, positive-semidefinite operator. ■

We apply this theorem to the spectral projections used in the proof of Theorem 1.3. The first two examples could have been dealt with by Fourier techniques [20, Section 3.3]; applying, however, the same method in all the examples renders the approach more systematic.

Example A.4 (Dyson kernel)**.** Consider $Lu = -u''$ on $(-\infty, \infty)$. Since $u \equiv 1$ is a solution of $Lu = 0$, both endpoints are LP; for a given $\operatorname{Im} z > 0$ the solutions u_a (u_b) of $(L - z)u = 0$ being L^2 at $-\infty$ (∞) are spanned by

$$u_a(x) = e^{-ix\sqrt{z}}, \qquad u_b(x) = e^{ix\sqrt{z}}.$$

Thus, Theorem A.2 applies: L is essentially selfadjoint on $C_0^\infty(-\infty, \infty)$, the resolvent of its unique selfadjoint extension is represented, for $\operatorname{Im} z > 0$, by the Green's kernel

$$G_z(x, y) = \frac{i}{2\sqrt{z}} \begin{cases} e^{i(x-y)\sqrt{z}}, & x > y, \\ e^{-i(x-y)\sqrt{z}}, & \text{otherwise.} \end{cases}$$

For $\lambda > 0$ there is the limit

$$\pi^{-1} \operatorname{Im} G_{\lambda+i0}(x, y) = K_\lambda(x, y) = \frac{\cos\left((x-y)\sqrt{\lambda}\right)}{2\pi\sqrt{\lambda}},$$

for $\lambda < 0$ the limit is zero; both limits are locally uniform in $x, y \in \mathbb{R}$. For $\lambda = 0$ there would be divergence, but we obviously have

$$\epsilon \operatorname{Im} G_{i\epsilon}(x, y) \to 0, \qquad \epsilon \to 0^+,$$

locally uniform in $x, y \in \mathbb{R}$. Hence, Theorem A.3 applies: $\sigma(L) = \sigma_{\text{ac}}(L) = [0, \infty)$ and (A.2) holds for each Borel set $\Delta \subset \mathbb{R}$. Given a bounded interval (a, b), we may estimate for the specific choice $\Delta = (-\infty, \pi^2)$ that

$$\int_a^b \int_a^b \left(\int_{-\infty}^{\pi^2} |K_\lambda(x, y)| \, d\lambda \right)^2 dx dy$$
$$= \int_a^b \int_a^b \left(\int_0^{\pi^2} \left| \frac{\cos\left((x-y)\sqrt{\lambda}\right)}{2\pi\sqrt{\lambda}} \right| d\lambda \right)^2 dx dy \leqslant \left(\int_a^b \int_0^{\pi^2} \frac{d\lambda}{2\pi\sqrt{\lambda}} \right)^2 = (b-a)^2.$$

Therefore, Theorem A.3 yields that $\chi_{(-\infty,\pi^2)}(L)\chi_{(a,b)}$ is Hilbert–Schmidt with the *Dyson kernel*

$$\int_{-\infty}^{\pi^2} K_\lambda(x, y) \, d\lambda = \int_0^{\pi^2} \frac{\cos\left((x-y)\sqrt{\lambda}\right)}{2\pi\sqrt{\lambda}} d\lambda = \frac{\sin(\pi(x-y))}{\pi(x-y)},$$

restricted to $x, y \in (a, b)$. Here, the last equality is simply obtained from

$$(x-y) \int_0^{\pi^2} \frac{\cos\left((x-y)\sqrt{\lambda}\right)}{2\sqrt{\lambda}} d\lambda = \int_0^{\pi^2} \frac{d}{d\lambda} \sin\left((x-y)\sqrt{\lambda}\right) d\lambda = \sin(\pi(x-y)).$$

Since the resulting kernel is continuous for $x, y \in (a, b)$, Theorem A.3 gives that $\chi_{(-\infty,\pi^2)}(L)\chi_{(a,b)}$ is a trace class operator with trace

$$\operatorname{tr} \chi_{(-\infty,\pi^2)}(L)\chi_{(a,b)} = b - a.$$

To summarize, we have thus obtained the following lemma.

Lemma A.5. *The operator $Lu = -u''$ is essentially selfadjoint on $C_0^\infty(-\infty, \infty)$. The spectrum of its unique selfadjoint extension is*

$$\sigma(L) = \sigma_{\text{ac}}(L) = [0, \infty).$$

Given (a, b) bounded, $\chi_{(-\infty,\pi^2)}(L)\chi_{(a,b)}$ is trace class with trace $b - a$ and kernel

$$K_{\text{Dyson}}(x, y) = \int_0^{\pi^2} \frac{\cos\left((x-y)\sqrt{\lambda}\right)}{2\pi\sqrt{\lambda}} d\lambda = \frac{\sin(\pi(x-y))}{\pi(x-y)}. \tag{A.3}$$

Example A.6 (Airy kernel)**.** Consider the differential operator $Lu = -u'' + xu$ on $(-\infty,\infty)$. Since the specific solution $u(x) = \operatorname{Bi}(x)$ of $Lu = 0$ is not locally L^2 at each of the endpoints, both endpoints are LP. For a given $\operatorname{Im} z > 0$ the solutions u_a (u_b) of $(L-z)u = 0$ being L^2 at $-\infty$ (∞) are spanned by [1, equation (10.4.59-64)]

$$u_a(x) = \operatorname{Ai}(x-z) - i\operatorname{Bi}(x-z), \qquad u_b(x) = \operatorname{Ai}(x-z).$$

Thus, Theorem A.2 applies: L is essentially selfadjoint on $C_0^\infty(-\infty,\infty)$, the resolvent of its unique selfadjoint extension is represented, for $\operatorname{Im} z > 0$, by the Green's kernel

$$G_z(x,y) = i\pi \begin{cases} \operatorname{Ai}(x-z)\left(\operatorname{Ai}(y-z) - i\operatorname{Bi}(y-z)\right), & x > y, \\ \operatorname{Ai}(y-z)\left(\operatorname{Ai}(x-z) - i\operatorname{Bi}(x-z)\right), & \text{otherwise}. \end{cases}$$

For $\lambda \in \mathbb{R}$ there is thus the limit

$$\pi^{-1}\operatorname{Im} G_{\lambda+i0}(x,y) = K_\lambda(x,y) = \operatorname{Ai}(x-\lambda)\operatorname{Ai}(y-\lambda),$$

locally uniform in $x, y \in \mathbb{R}$. Hence, Theorem A.3 applies: $\sigma(L) = \sigma_{\text{ac}}(L) = \mathbb{R}$ and (A.2) holds for each Borel set $\Delta \subset \mathbb{R}$. Given $s > -\infty$, we may estimate for the specific choice $\Delta = (-\infty, 0)$ that

$$\left(\int_s^\infty \int_s^\infty \left(\int_\Delta |K_\lambda(x,y)|\,d\lambda\right)^2 dx dy\right)^{1/2} \leqslant \int_s^\infty \int_0^\infty \operatorname{Ai}(x+\lambda)^2 d\lambda dx = \tau(s)$$

with

$$\tau(s) = \frac{1}{3}\left(2s^2\operatorname{Ai}(s)^2 - 2s\operatorname{Ai}'(s)^2 - \operatorname{Ai}(s)\operatorname{Ai}'(s)\right).$$

Therefore, Theorem A.3 yields that $\chi_{(-\infty,0)}(L)\chi_{(s,\infty)}$ is Hilbert–Schmidt with the *Airy kernel*

$$\int_{-\infty}^0 K_\lambda(x,y)\,d\lambda = \int_0^\infty \operatorname{Ai}(x+\lambda)\operatorname{Ai}(y+\lambda)\,d\lambda = \frac{\operatorname{Ai}(x)\operatorname{Ai}'(y) - \operatorname{Ai}'(x)\operatorname{Ai}(y)}{x-y},$$

restricted to $x, y \in (s,\infty)$. Here, the last equality is obtained from a Christoffel–Darboux type of argument: First, we use the underlying differential equation,

$$x\operatorname{Ai}(x+\lambda) = \operatorname{Ai}''(x+\lambda) - \lambda\operatorname{Ai}(x+\lambda),$$

and partial integration to obtain

$$\begin{aligned} x\int_0^\infty \operatorname{Ai}(x+\lambda)\operatorname{Ai}(y+\lambda)\,d\lambda &= \int_0^\infty \operatorname{Ai}''(x+\lambda)\operatorname{Ai}(y+\lambda)\,d\lambda - \int_0^\infty \lambda\operatorname{Ai}(x+\lambda)\operatorname{Ai}(y+\lambda)\,d\lambda \\ &= -\operatorname{Ai}'(x)\operatorname{Ai}(y) - \int_0^\infty \operatorname{Ai}'(x+\lambda)\operatorname{Ai}'(y+\lambda)\,d\lambda - \int_0^\infty \lambda\operatorname{Ai}(x+\lambda)\operatorname{Ai}(y+\lambda)\,d\lambda. \end{aligned}$$

Next, we exchange the roles of x and y and substract to get the assertion. Since the resulting kernel is continuous, Theorem A.3 gives that $\chi_{(-\infty,0)}(L)\chi_{(s,\infty)}$ is a trace class operator with trace

$$\operatorname{tr}\chi_{(-\infty,0)}(L)\chi_{(s,\infty)} = \tau(s) \to \infty, \qquad s \to -\infty.$$

To summarize, we have thus obtained the following lemma.

Lemma A.7. *The differential operator* $Lu = -u'' + xu$ *is essentially selfadjoint on* $C_0^\infty(-\infty, \infty)$. *The spectrum of its unique selfadjoint extension is*

$$\sigma(L) = \sigma_{\text{ac}}(L) = (-\infty, \infty).$$

Given $s > -\infty$, *the operator* $\chi_{(-\infty,0)}(L)\chi_{(s,\infty)}$ *is trace class with kernel*

$$K_{\text{Airy}}(x, y) = \int_0^\infty \operatorname{Ai}(x+\lambda)\operatorname{Ai}(y+\lambda)\,d\lambda = \frac{\operatorname{Ai}(x)\operatorname{Ai}'(y) - \operatorname{Ai}'(x)\operatorname{Ai}(y)}{x-y}. \tag{A.4}$$

Example A.8 (Bessel kernel)**.** Given $\alpha > 0$, take $Lu = -4(xu')' + \alpha^2 x^{-1} u$ on $(0, \infty)$. Since a fundamental system of solutions of $Lu = 0$ is given by $u(x) = x^{\pm\alpha/2}$, the endpoint $x = 0$ is LP for $\alpha \geqslant 1$ and LC otherwise; the endpoint $x = \infty$ is LP in both cases. Fixing the LP case at $x = 0$, we restrict ourselves to the case $\alpha \geqslant 1$.

For a given $\operatorname{Im} z > 0$ the solutions u_a (u_b) of $(L - z)u = 0$ being L^2 at 0 (∞) are spanned by [1, equations (9.1.7-9) and (9.2.5-6)]

$$u_a(x) = J_\alpha\big(\sqrt{xz}\big), \qquad u_b(x) = J_\alpha\big(\sqrt{xz}\big) + iY_\alpha\big(\sqrt{xz}\big).$$

Thus, Theorem A.2 applies: L is essentially selfadjoint on $C_0^\infty(0, \infty)$, the resolvent of its unique selfadjoint extension is represented, for $\operatorname{Im} z > 0$, by the Green's kernel

$$G_z(x, y) = \frac{i\pi}{4}\begin{cases} J_\alpha(\sqrt{xz})\big(J_\alpha\big(\sqrt{yz}\big) + iY_\alpha\big(\sqrt{yz}\big)\big), & x > y, \\ J_\alpha(\sqrt{yz})\big(J_\alpha\big(\sqrt{xz}\big) + iY_\alpha\big(\sqrt{xz}\big)\big), & \text{otherwise.} \end{cases}$$

For $\lambda > 0$ there is the limit

$$\pi^{-1}\operatorname{Im} G_{\lambda+i0}(x, y) = K_\lambda(x, y) = \frac{1}{4} J_\alpha\big(\sqrt{x\lambda}\big) J_\alpha\big(\sqrt{y\lambda}\big),$$

for $\lambda \leqslant 0$ the limit is zero; both limits are locally uniform in $x, y \in \mathbb{R}$. Hence, Theorem A.3 applies: $\sigma(L) = \sigma_{\text{ac}}(L) = [0, \infty)$ and (A.2) holds for each Borel set $\Delta \subset \mathbb{R}$. Given $0 \leqslant s < \infty$, we may estimate for the specific choice $\Delta = (-\infty, 1)$ that

$$\left(\int_0^s \int_0^s \left(\int_\Delta |K_\lambda(x, y)|\,d\lambda\right)^2 dx\,dy\right)^{1/2} \leqslant \frac{1}{4}\int_0^s \int_0^1 J_\alpha\big(\sqrt{x\lambda}\big)^2 d\lambda\,dx = \tau_\alpha(s).$$

Therefore, Theorem A.3 yields that $\chi_{(-\infty,1)}(L)\chi_{(0,s)}$ is Hilbert–Schmidt with the *Bessel kernel*

$$\begin{aligned}\int_{-\infty}^1 K_\lambda(x, y)\,d\lambda &= \frac{1}{4}\int_0^1 J_\alpha\big(\sqrt{x\lambda}\big) J_\alpha\big(\sqrt{y\lambda}\big)\,d\lambda \\ &= \frac{J_\alpha(\sqrt{x})\sqrt{y}J_\alpha'(\sqrt{y}) - \sqrt{x}J_\alpha'(\sqrt{x})J_\alpha(\sqrt{y})}{2(x-y)},\end{aligned}$$

restricted to $x, y \in (0, s)$. Here, the last equality is obtained from a Christoffel–Darboux type of argument: First, we use the underlying differential equation,

$$xJ_\alpha\big(\sqrt{x\lambda}\big) = -4\frac{d}{d\lambda}\left(\lambda \frac{d}{d\lambda} J_\alpha\big(\sqrt{x\lambda}\big)\right) + \alpha^2\lambda^{-1} J_\alpha\big(\sqrt{x\lambda}\big),$$

and partial integration to obtain

$$\frac{x}{4}\int_0^1 J_\alpha\big(\sqrt{x\lambda}\big) J_\alpha\big(\sqrt{y\lambda}\big)\,d\lambda$$

$$
\begin{aligned}
&= -\int_0^1 \frac{d}{d\lambda}\left(\lambda \frac{d}{d\lambda} J_\alpha(\sqrt{x\lambda})\right) J_\alpha(\sqrt{y\lambda})\, d\lambda + \frac{\alpha^2}{4}\int_0^1 \lambda^{-1} J_\alpha(\sqrt{x\lambda}) J_\alpha(\sqrt{y\lambda})\, d\lambda \\
&= -\frac{1}{2}\sqrt{x} J_\alpha'(\sqrt{x}) J_\alpha(\sqrt{y}) \\
&\quad + \int_0^1 \lambda \left(\frac{d}{d\lambda} J_\alpha(\sqrt{x\lambda})\right)\left(\frac{d}{d\lambda} J_\alpha(\sqrt{y\lambda})\right) d\lambda + \frac{\alpha^2}{4}\int_0^1 \lambda^{-1} J_\alpha(\sqrt{x\lambda}) J_\alpha(\sqrt{y\lambda})\, d\lambda.
\end{aligned}
$$

Next, we exchange the roles of x and y and substract to get the assertion. Since the resulting kernel is continuous, Theorem A.3 gives that $\chi_{(-\infty,1)}(L)\chi_{(0,s)}$ is a trace class operator with trace

$$
\operatorname{tr} \chi_{(-\infty,1)}(L)\chi_{(0,s)} = \tau_\alpha(s) \to \infty, \qquad s \to \infty.
$$

To summarize, we have thus obtained the following lemma.

Lemma A.9. *Given $\alpha \geqslant 1$, the differential operator $Lu = -4(xu')' + \alpha^2 x^{-1} u$ is essentially selfadjoint on $C_0^\infty(0,\infty)$. The spectrum of its unique selfadjoint extension is*

$$
\sigma(L) = \sigma_{\mathrm{ac}}(L) = [0,\infty).
$$

Given $0 \leqslant s < \infty$, the operator $\chi_{(-\infty,1)}(L)\chi_{(0,s)}$ is trace class with kernel

$$
K_{\mathrm{Bessel}}^{(\alpha)}(x,y) = \frac{1}{4}\int_0^1 J_\alpha(\sqrt{x\lambda}) J_\alpha(\sqrt{y\lambda})\, d\lambda = \frac{J_\alpha(\sqrt{x})\sqrt{y} J_\alpha'(\sqrt{y}) - \sqrt{x} J_\alpha'(\sqrt{x}) J_\alpha(\sqrt{y})}{2(x-y)}. \tag{A.5}
$$

Remark A.10. Lemma A.9 extends to $0 \leqslant \alpha < 1$ if we choose the particular selfadjoint realization of L that is defined by the boundary condition (1.10), cf. [10, Example 10.5.12].

References

[1] Abramowitz M., Stegun I.A., Handbook of mathematical functions with formulas, graphs, and mathematical tables, *National Bureau of Standards Applied Mathematics Series*, Vol. 55, U.S. Government Printing Office, Washington, D.C., 1964.

[2] Anderson G.W., Guionnet A., Zeitouni O., An introduction to random matrices, *Cambridge Studies in Advanced Mathematics*, Vol. 118, Cambridge University Press, Cambridge, 2010.

[3] Andrews G.E., Askey R., Roy R., Special functions, *Encyclopedia of Mathematics and its Applications*, Vol. 71, Cambridge University Press, Cambridge, 1999.

[4] Borodin A., Olshanski G., Asymptotics of Plancherel-type random partitions, *J. Algebra* **313** (2007), 40–60, math.PR/0610240.

[5] Collins B., Product of random projections, Jacobi ensembles and universality problems arising from free probability, *Probab. Theory Related Fields* **133** (2005), 315–344, math.PR/0406560.

[6] Deift P.A., 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, Amer. Math. Soc., Providence, RI, 1999.

[7] Erdélyi A., Magnus W., Oberhettinger F., Tricomi F.G., Higher transcendental functions. Vol. II, McGraw-Hill Book Company, New York, 1953.

[8] Forrester P.J., Log-gases and random matrices, *London Mathematical Society Monographs Series*, Vol. 34, Princeton University Press, Princeton, NJ, 2010.

[9] Gohberg I., Goldberg S., Krupnik N., Traces and determinants of linear operators, *Operator Theory: Advances and Applications*, Vol. 116, Birkhäuser Verlag, Basel, 2000.

[10] Hutson V., Pym J.S., Cloud M.J., Applications of functional analysis and operator theory, *Mathematics in Science and Engineering*, Vol. 200, 2nd ed., Elsevier B.V., Amsterdam, 2005.

[11] Johansson K., Shape fluctuations and random matrices, *Comm. Math. Phys.* **209** (2000), 437–476, math.CO/9903134.

[12] Johnstone I.M., On the distribution of the largest eigenvalue in principal components analysis, *Ann. Statist.* **29** (2001), 295–327.

[13] Johnstone I.M., Multivariate analysis and Jacobi ensembles: largest eigenvalue, Tracy–Widom limits and rates of convergence, *Ann. Statist.* **36** (2008), 2638–2716, arXiv:0803.3408.

[14] Kuijlaars A.B.J., Universality, in The Oxford Handbook of Random Matrix Theory, Oxford University Press, Oxford, 2011, 103–134, arXiv:1103.5922.

[15] Lax P.D., Functional analysis, Pure and Applied Mathematics, Wiley-Interscience, New York, 2002.

[16] Lubinsky D.S., A new approach to universality limits involving orthogonal polynomials, *Ann. of Math.* **170** (2009), 915–939, math.CA/0701307.

[17] Reed M., Simon B., Methods of modern mathematical physics. I. Functional analysis, Academic Press, New York – London, 1972.

[18] Simon B., Trace ideals and their applications, *Mathematical Surveys and Monographs*, Vol. 120, 2nd ed., Amer. Math. Soc., Providence, RI, 2005.

[19] Stolz G., Weidmann J., Approximation of isolated eigenvalues of ordinary differential operators, *J. Reine Angew. Math.* **445** (1993), 31–44.

[20] Tao T., Topics in random matrix theory, *Graduate Studies in Mathematics*, Vol. 132, Amer. Math. Soc., Providence, RI, 2012.

[21] Tracy C.A., Widom H., Level-spacing distributions and the Airy kernel, *Comm. Math. Phys.* **159** (1994), 151–174, hep-th/9211141.

[22] Tracy C.A., Widom H., Level spacing distributions and the Bessel kernel, *Comm. Math. Phys.* **161** (1994), 289–309, hep-th/9304063.

[23] Wachter K.W., The limiting empirical measure of multiple discriminant ratios, *Ann. Statist.* **8** (1980), 937–957.

[24] Weidmann J., Linear operators in Hilbert spaces, *Graduate Texts in Mathematics*, Vol. 68, Springer-Verlag, New York – Berlin, 1980.

[25] Weidmann J., Spectral theory of ordinary differential operators, *Lecture Notes in Mathematics*, Vol. 1258, Springer-Verlag, Berlin, 1987.