ON THE SCALING LIMITS OF DETERMINANTAL POINT PROCESSES
WITH KERNELS INDUCED BY STURM–LIOUVILLE OPERATORS

FOLKMAR BORNEMANN

ABSTRACT. By applying an idea of Borodin and Olshanski (2007), 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
unitary random matrix ensembles (GUE, LUE/Wishart, JUE/MANOVA).

1. INTRODUCTION

We consider determinantal point processes on an interval \( \Lambda = (a, b) \) with trace
class projection kernel
\[
K_n(x, y) = \sum_{j=0}^{n-1} \phi_j(x) \phi_j(y),
\]
where \( \phi_0, \phi_1, \ldots, \phi_{n-1} \) are orthonormal in \( L^2(\Lambda) \); each \( \phi_j \) may have some depen-
dence on \( n \) that we suppress from the notation. We recall (see, e.g., Anderson,
Guionnet and Zeitouni 2010, §4.2) that for such processes the joint probability
density of the \( n \) points is given by
\[
p_n(x_1, \ldots, x_n) = \frac{1}{n!} \det K_n(x_i, x_j),
\]
the mean counting probability is given by the density
\[
\rho_n(x) = n^{-1} K_n(x, x),
\]
and the gap probabilities are given, by the inclusion-exclusion principle, in terms
of a Fredholm determinant, namely
\[
E_n(J) = \mathbb{P}(\{x_1, \ldots, x_n\} \cap J = \emptyset) = \det(I - 1_J K_n 1_J).
\]
The various scaling limits are usually derived from a pointwise convergence of the
kernel function \( K_n(x, y) \) obtained from considering the large \( n \) asymptotic of the
eigenfunctions \( \phi_j \), which can be technically very involved. Borodin and Olshan-
ski (2007) suggested, for discrete point processes, a different, conceptually and
technically much simpler approach based on selfadjoint difference operators. We

\[\text{2010 Mathematics Subject Classification.} \quad \text{15B52, 34B24, 33C45.} \]

\[\text{1\textsuperscript{1} Based on the two-scale Plancherel–Rotach asymptotic of classical orthogonal polynomials or,}
\text{methodically more general, on the asymptotic of Riemann–Hilbert problems; see, e.g., Deift (1999),}
\text{Johnstone (2001), Collins (2005), Johnstone (2008), Anderson et al. (2010), and Kuijlaars (2011).} \]
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 unitary random matrix ensembles (GUE, LUE/Wishart, JUE/MANOVA) 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

$$K_n = 1_{(-\infty,0)}(L_n)$$

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 scaled objects

$$\tilde{E}_n(j) = \det(I - \mathbb{1}_j \tilde{K}_n \mathbb{1}_j), \quad \tilde{K}_n = 1_{(-\infty,0)}(\tilde{L}_n),$$

where $L_n$ 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

$$\tilde{L}_n u \to \tilde{L} u \quad (u \in C). \quad (1.2)$$

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 $L_n$—a convergence that is, typically, an easy calculus exercise. Now, given (1.2), the concept of *strong resolvent convergence* (see Theorem 2) immediately yields, if $0 \notin \sigma_{pp}(\tilde{L})$,

$$K_n \mathbb{1}_{\tilde{\Lambda}_n} \to 1_{(-\infty,0)}(L_n) \mathbb{1}_{\tilde{\Lambda}} \xrightarrow{s} 1_{(-\infty,0)}(L).$$

Third, we take an interval $J \subset \tilde{\Lambda}$, eventually satisfying $J \subset \tilde{\Lambda}_n$, such that the operator $1_{(-\infty,0)}(\tilde{L}) \mathbb{1}_J$ is trace class with kernel $\tilde{K}(x,y)$ (which can be obtained from the generalized eigenfunction expansion of $\tilde{L}$, see §A.2). Then, we immediately get the strong convergence

$$\tilde{K}_n \mathbb{1}_J \xrightarrow{s} \tilde{K} \mathbb{1}_J.$$

**Remark 1.1.** Tao (2012, §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 $1_{(-\infty,0)}(\tilde{L})$, he puts his sketch under the headline “The Dyson and Airy kernels of GUE via semiclassical analysis”.

**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 $\tilde{K}_n \mathbb{1}_J \to \tilde{K} \mathbb{1}_J$, 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 pointwise convergence of the kernel functions to the convergence of gap probabilities, that is,

$$\tilde{E}_n(j) = \det(I - \mathbb{1}_j \tilde{K}_n \mathbb{1}_j) \to \det(I - \mathbb{1}_j \tilde{K} \mathbb{1}_j) = \tilde{E}(j).$$

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

---

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

---
(1) By Hadamard’s inequality, convergence of the determinants follows directly from the locally uniform convergence of the kernels \( K_n \) (Anderson et al. 2010, Lemma 3.4.5) and, for unbounded \( J \), from further large deviation estimates (Anderson et al. 2010, 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. 2010, §§3.5 and 3.7).

(2) Since \( A \mapsto \det(I - A) \) is continuous with respect to the trace class norm (Simon 2005, Thm. 3.4), \( \tilde{K}_n 1_J \to \tilde{K} 1_J \) 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 on the scaling limits of the LUE/Wishart ensembles (2001) and that on the limits of the JUE/MANOVA ensembles (2008).

(3) Since \( 1_J \tilde{K}_n 1_J \) and \( 1_J \tilde{K} 1_J \) are selfadjoint and positive semi-definite, yet another way is by observing that the convergence \( \tilde{K}_n 1_J \to \tilde{K} 1_J \) in trace class norm is, for continuous kernels, equivalent (Simon 2005, Thm. 2.20) to the combination of both, the convergence \( \tilde{K}_n 1_J \to \tilde{K} 1_J \) in the weak operator topology and the convergence of the traces

\[
\int_J \tilde{K}_n(\xi, \zeta) d\xi \to \int_J \tilde{K}(\xi, \zeta) d\xi. \tag{T}
\]

Once again, these convergences follow from locally uniform convergence of the kernels; see Deift (1999, §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 (T) by some additional, similarly short and simply argument. Note that, by the ideal property of the trace class, condition (T) implies the same condition for all \( J' \subset J \). We fall, however, short of conceiving a proof strategy for condition (T) that would be independent of all the laborious proofs of locally uniform convergence of the kernels.

Remark 1.2. Contrary to the discrete case originally 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, \( \tilde{K}_n(\xi, \eta) \to \tilde{K}(\xi, \eta) \).

For instance, in §2, we will need just a single such instance, namely

\[
\tilde{K}_n(0, 0) \to \tilde{K}(0, 0), \quad \text{(K0)}
\]

to prove a limit law \( \tilde{\rho}_n(t) dt \xrightarrow{\text{w}} \tilde{\rho}(t) dt \) for the mean counting probability. Using mollified Dirac deltas, pointwise convergence would generally follow, for continuously differentiable \( \tilde{K}_n(\xi, \eta) \), if were able to bound, locally uniform, the gradient of \( \tilde{K}_n(\xi, \eta) \). Then, by dominated convergence, criterion (T) 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. Using the Borodin–Olshanski method, we will prove the following general result on selfadjoint Sturm–Liouville operators; a result that adds a further class of problems to the universality (Kuijlaars 2011) of the Dyson, Airy, and Bessel kernel in the bulk, soft-edge, and hard-edge scaling limits.

**Theorem 1.** Consider a selfadjoint realization $L_n$ of the formally selfadjoint Sturm–Liouville operator

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

on $\Lambda = (-\infty, \infty)$, $\Lambda = (0, \infty)$, or $\Lambda = (0, 1)$ with $p, q_n \in C^\infty(\Lambda)$ and $p(x) > 0$ for all $x \in \Lambda$ such that, for $t \in \Lambda$ and $n \to \infty$,

$$n^{-2\alpha} \lambda_n \sim \omega, \quad n^{-2\alpha} q_n(n^k t) \sim \tilde{q}(t), \quad n^{2\alpha''} p(n^k t) \sim \tilde{p}(t) > 0,$$

asymptotically up to an error $O(n^{-1})$; the scaling exponents should satisfy

$$\kappa + \kappa' + \kappa'' = 1.$$  \hfill (1.4)

We assume that these expansions can be differentiated\(^4\) at least twice, that the roots of $\tilde{q}(t) - \omega$ are simple and that the spectral projection $K_n = \mathbb{1}_{(-\infty,0)}(L_n)$ satisfies $\text{tr } K_n = n$. Let $x = \sigma_n x + \mu_n$ induce the transformed projection $\tilde{K}_n$ and let $x = n^t$ induce the transformed mean counting probability density $\tilde{\rho}_n$. Then the following scaling limits hold.

- **Bulk scaling limit:** given $t \in \Lambda$ with $\tilde{q}(t) < \omega$, the scaling

$$\mu_n = n^k t, \quad \sigma_n = n^{\kappa - 1} \tilde{\rho}(t), \quad \tilde{\rho}(t) = \frac{1}{\pi} \sqrt{\frac{1}{\tilde{q}(t)} - \frac{\omega - \tilde{q}(t)}{\tilde{q}(t)}},$$

yields, for any bounded interval $J$, the strong operator limit $\tilde{K}_n \mathbb{1}_J \xrightarrow{s} K_{\text{Dyson}} \mathbb{1}_J$.

Under condition $(K_0)$ and if $\tilde{\rho}$ has unit mass on $\Lambda$, there is the limit law

$$\tilde{\rho}_n(t) dt \xrightarrow{w} \tilde{\rho}(t) dt.$$  \hfill (1.3)

- **Soft-edge scaling limit:** given $t_+ \in \Lambda$ with $\tilde{q}(t_+) = \omega$, the scaling

$$\mu_n x = n^k t_+, \quad \sigma_n = n^{\kappa - 3/2} \left( \frac{\tilde{\rho}(t_+)}{\tilde{q}(t_+)} \right)^{1/3},$$

yields, for $s > -\infty$ and any interval $J \subset (s, \infty)$, the limit $\tilde{K}_n \mathbb{1}_J \xrightarrow{s} K_{\text{Airy}} \mathbb{1}_J$.

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

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

the scaling

$$\mu_n = 0, \quad \sigma_n = \frac{p'(0)}{4 \omega_n n^{2\alpha'}}, \quad \alpha = \frac{2\gamma}{\sqrt{p'(0)}},$$

yields,\(^5\) for any bounded interval $J \subset (0, \infty)$, the limit $\tilde{K}_n \mathbb{1}_J \xrightarrow{s} K_{\text{Bessel}}^{(n)} \mathbb{1}_J$.

---

\(^3\)Since, in this paper, we always consider a particular selfadjoint realization of a formal differential operator, we will use the same letter to denote both.

\(^4\)We say that an expansion $f_n(t) - f(t) = O(1/n)$ can be differentiated if $f_n'(t) - f'(t) = O(1/n)$.

\(^5\)Here, if $0 \leq \alpha < 1$, the selfadjoint realization $L_n$ is defined by means of the boundary condition

$$2xu'(x) - au(x) = o(x^{-\alpha/2}) \ (x \to 0).$$

\hfill (1.6)
Remark 1.3. Whether the intervals $J$ can be chosen unbounded or not depends on whether the limit operator $K \mathbb{1}_J$ would then be trace class or not, see the explicit formulæ for the traces in the appendix: only in the former case we get the representation of the scaling limit in terms of a particular integral kernel. Note that we can never use $J = \Lambda$ since $\text{tr} K_n = n \to \infty$.

Outline of the paper. The proof of Theorem 1 is subject of §2. In §3 we apply it to the classical orthogonal polynomials, which yields a short and unified derivation of the known formulæ for the scaling limits for the classical unitary random matrix ensembles (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 $p_n$ and $q_n$ as well as to the eigenvalues $\lambda_n$ of the Sturm–Liouville operator $L_n$ and next, by applying Theorem 1, to the particular scalings limits.

To emphasize that our main result and its application is largely independent of concretely identifying the kernel $\tilde{K}$ of the spectral projection $1_{(-\infty,0)}(L)$, we postpone this identification to the appendix: there, using generalized eigenfunction expansions, we calculate the Dyson, Airy, and Bessel kernels directly from the limit differential operators $\tilde{L}$.

2. Proof of the Main Result for Sturm–Liouville Operators

To begin with, we first state the expansions of the coefficients, which also motivate the particular regularity assumptions made in Theorem 1.

Expansions. Since $L_n$ is a selfadjoint realization of

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

with $p, q_n \in C^\infty(\Lambda)$ and $p(x) > 0$ for $x \in \Lambda$, we have $C^\infty_0(\Lambda) \subset D(L_n)$. The scaling

$$x = \sigma_n \xi + n^\kappa t \quad (\sigma_n \neq 0) \quad (2.1)$$

(with $t \in \Lambda$ a parameter) yields $\tilde{\Lambda} = (-\infty, \infty)$ in each of the following cases:

- $\Lambda = (-\infty, \infty)$;
- $\Lambda = (0, \infty)$, $\kappa \geq 0$, $\sigma_n \to 0$;
- $\Lambda = (0,1)$, $\kappa = 0$, $\sigma_n \to 0$.

By virtue of $K_n(\xi, \eta) \, d\eta = K_n(x,y) \, dy$ (with $y = \sigma_n \eta + n^\kappa t$), such a scaling induces

$$\tilde{K}_n(\xi, \eta) = \sigma_n K_n(\sigma_n \xi + n^\kappa t, \sigma_n \eta + n^\kappa t).$$

Because of $\sigma_n^2 > 0$ and $p(n^\kappa t) > 0$, we see that $\tilde{K}_n = 1_{(-\infty,0)}(\tilde{L}_n)$ with

$$\tilde{L}_n = -\frac{1}{p(n^\kappa t)} \frac{d}{d\xi} \left( p(\sigma_n \xi + n^\kappa t) \frac{d}{d\xi} \right) + \frac{\sigma_n^2}{p(n^\kappa t)} \left( q_n(\sigma_n \xi + n^\kappa t) - \lambda_n \right).$$
Now, if $\sigma_n = o(n^{k-1/2})$, we get from (1.3) and (1.4) by Taylor expansion that
\[
\tilde{L}_n u(\xi) = -u''(\xi) + \sigma_n n^{2-2k} \left( \tilde{\varrho}'(t) - \omega + \sigma_n n^{-k} \tilde{q}'(t) \cdot \xi \right) u(\xi) + o(1).
\]

**Bulk Scaling Limit.** If $\tilde{q}(t) \neq \omega$, by choosing
\[
\sigma_n = \sigma_n(t) = \pi n^{k-1} \sqrt{\frac{\tilde{p}(t)}{|\omega - \tilde{q}(t)|}},
\]
we get $\tilde{L}_n u(\xi) = -u''(\xi) - s\pi^2 u(\xi) + o(1)$; that is, $\tilde{L}_n u \to \tilde{L} u$ with the limit
\[
\tilde{L} = -\frac{d^2}{d\xi^2} - s\pi^2, \quad s = \text{sign}(\omega - \tilde{q}(t)).
\]

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

**Limit Law.** The result for the bulk scaling limit allows, in passing, to calculate a limit law of the mean counting probability $\rho_n(x) dx$. We observe that, by virtue of $\rho_n(x) dx = \tilde{\varrho}_n(t) dt$, the scaling $x = n^k t$ transforms $\rho_n(x)$ into $\tilde{\rho}_n(t) = n^k \varrho_n(n^k t) = n^{k-1} K_n(n^k t, n^k t) = \frac{n^{k-1}}{\sigma_n(t)} \tilde{K}_n(0, 0) = \frac{1}{\pi} \sqrt{\frac{|\omega - \tilde{q}(t)|}{\tilde{p}(t)}} \tilde{K}_n(0, 0).

Thus, to get to a limit, we have to *assume* condition (K0), that is,$^8
\tilde{K}_n(0, 0) \to [\tilde{q}(t) < \omega] K_{\text{Dyson}}(0, 0) = [\tilde{q}(t) < \omega]$.

We then get, for $\tilde{q}(t) \neq \omega$ (that is, almost everywhere in $\Lambda$),
\[
\tilde{\rho}_n(t) \to \tilde{\varrho}(t) = \frac{1}{\pi} \sqrt{\frac{(\omega - \tilde{q}(t))^+}{\tilde{p}(t)}}.
\]

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

**Soft-Edge Scaling Limit.** If $\tilde{q}(t_*) = \omega$, by choosing$^9$
\[
\sigma_n = \sigma_n(t_*) = n^{k-2/3} \left( \frac{\tilde{\varrho}(t_*)}{\tilde{q}'(t_*)} \right)^{1/3},
\]
we get $\tilde{L}_n u(\xi) = -u''(\xi) + \xi u(\xi) + o(1)$; that is, $\tilde{L}_n u \to \tilde{L} u$ with the limit
\[
\tilde{L} = -\frac{d^2}{d\xi^2} + \xi.
\]

---

$^6$If $\sigma_n = o(n^{k-1/2})$, the error term is $O(\max(\sigma_n n^{-k}, \sigma_n^2 n^{1-2k}))$, which amounts for $O(n^{-1})$ in the bulk scaling limit and for $O(n^{-1/3})$ in the soft-edge scaling limit.

$^7$It is the cancellation of the powers of $n$ in the last equality, which called for assumption (1.4).

$^8$Following Knuth, the bracketed notation $[S]$ stands for 1 if the statement $S$ is true, 0 otherwise.

$^9$Note that, by the assumption made on the simplicity of the roots of $\tilde{q}(t) - \omega$, we have $\tilde{q}'(t_*) \neq 0$. 

We recall that \( \tilde{L} \) is essentially selfadjoint on \( C_0^\infty(\tilde{\Lambda}) \) and that its unique selfadjoint extension has absolutely continuous spectrum: \( \sigma(\tilde{L}) = \sigma_{ac}(\tilde{L}) = (-\infty, \infty) \). The spectral projection \( 1_{(-\infty,0)}(\tilde{L}) \) can straightforwardly be calculated by a generalized eigenfunction expansion, yielding the \textit{Airy kernel} (A.4); see Lemma A.2.

**Hard-Edge Scaling Limit.** In the case \( a = 0 \), we take the scaling

\[
x = \sigma_n \xi \quad (\sigma_n > 0),
\]

with \( \sigma_n = o(1) \) appropriately chosen, to explore the vicinity of this “hard edge”; note that such a scaling yields \( \tilde{\Lambda} = (0, \infty) \). To simplify we make the assumptions stated in (1.5). We see that \( \tilde{K}_n = 1_{(-\infty,0)}(\tilde{L}_n) \) with

\[
\tilde{L}_n = -\frac{4}{p'(0)} \frac{d}{d\xi} \left( p(\sigma_n \xi) \frac{d}{d\xi} \right) + \frac{4\sigma_n}{p'(0)} (q(\sigma_n \xi) - \lambda_n).
\]

If we choose

\[
\sigma_n = n^{-2\alpha'} p'(0) 4\omega,
\]

we get\(^\text{10}\) \( \tilde{L}_n u(\xi) = -4(\xi u'(\xi))' + (\alpha^2 \xi^{-1} - 1) u(\xi) + o(1) \); that is, \( \tilde{L}_n u \rightarrow \tilde{L} u \),

\[
\tilde{L} = -4 \frac{d}{d\xi} \left( \tilde{\xi} \frac{d}{d\xi} \right) + \alpha^2 \xi^{-2} - 1, \quad \alpha = \frac{2\gamma}{\sqrt{p'(0)}}.
\]

We recall that, if \( \alpha \geq 1 \), 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: \( \sigma(\tilde{L}) = \sigma_{ac}(\tilde{L}) = [-1, \infty) \). The spectral projection can be calculated by a generalized eigenfunction expansion, yielding the \textit{Bessel kernel} (A.5); see Lemma A.3.

**Remark 2.1.** The theorem also holds in the case \( 0 \leq \alpha < 1 \) if the particular selfadjoint realization \( L_n \) is defined by the boundary condition (1.6), see Remark A.1.

3. **Projection Kernels Associated to Classical Orthogonal Polynomials**

In this section we apply Theorem 1 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 Laguerre Unitary Ensemble (LUE) or Wishart ensemble, and the Jacobi Unitary Ensemble (JUE) or MANOVA\(^\text{11}\) ensemble.

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 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 \( \phi_n \) 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 (1999) (§5.7 for the Jacobi polynomials, §6.5 for the Hermite and Laguerre polynomials). By a result of Tricomi (see Erdélyi, Magnus, Oberhettinger and Tricomi 1953, §10.7), the \( P_n(x) \) satisfy the eigenvalue problem

\[
- \frac{1}{w(x)} \frac{d}{dx} \left( p(x) w(x) \frac{d}{dx} P_n(x) \right) = \lambda_n P_n(x), \quad \lambda_n = -n(n' + \frac{1}{2}(n + 1)p''),
\]

\(^\text{10}\)The error term is \( O(n^{-\min(1,2\alpha')}) \), which amounts for \( O(n^{-1}) \) if \( \alpha' \geq \frac{1}{2} \) (as throughout §3).

\(^\text{11}\)MANOVA = multivariate analysis of variance
where $p(x)$ is a quadratic polynomial and $r(x)$ a linear polynomial such that
\[
\frac{w'(x)}{w(x)} = \frac{r(x)}{p(x)}.
\]
In terms of $\phi_n$, a simple calculation shows that
\[
-\frac{d}{dx} \left( p(x) \frac{d}{dx} \phi_n(x) \right) + q(x) \phi_n(x) = \lambda_n \phi_n(x), \quad q(x) = \frac{r(x)^2}{4p(x)} + \frac{r'(x)}{2}.
\]
Therefore, by the completeness of the $\phi_n$, the formally selfadjoint Sturm–Liouville operator
\[
L = -\frac{d}{dx} p(x) \frac{d}{dx} + q(x)
\]
has a particular selfadjoint realization (which we continue to denote by the letter $L$) with spectrum
\[
\sigma(L) = \{\lambda_0, \lambda_1, \lambda_2, \ldots\}
\]
and corresponding eigenfunctions $\phi_n$. Hence, the projection kernel (1.1) induces an integral operator $K_n$ with \(\text{tr} K_n = n\) that satisfies
\[
K_n = 1_{(-\infty,0)}(L_n), \quad L_n = L - \lambda_n.
\]
Note that this relation remains true if we choose to make some parameters of the weight $w$ (and, therefore, of the functions $\phi_j$) to depend on $n$. For the scaling limits of $K_n$, we are now in the realm of Theorem 1: given the weight $w(x)$ as the only input, all the other quantities can now simply be obtained by routine calculations.

**Hermite Polynomials.** The weight is $w(x) = e^{-x^2}$ on $\Lambda = (-\infty, \infty)$; hence
\[
p(x) = 1, \quad r(x) = -2x, \quad q(x) = x^2 - 1, \quad \lambda_n = 2n,
\]
and, therefore,
\[
\kappa = \kappa' = \frac{1}{2}, \quad \kappa'' = 0, \quad \tilde{p}(t) = 1, \quad \tilde{q}(t) = t^2, \quad \omega = 2.
\]
Theorem 1 is applicable and we directly read off the following well-known scaling limits of the GUE (see, e.g., Anderson et al. 2010, Chap. 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 $\tilde{\rho}_n$ with a weak limit given by the Wigner semicircle law
  \[
  \tilde{\rho}(t) = \frac{1}{\pi t} \sqrt{(2 - t^2)_+};
  \]
- soft-edge scaling limit: the transformation
  \[
x = \pm(2^{-1/2} n^{-1/6} \xi + \sqrt{2n})
  \]
  induces $\tilde{K}_n$ with a strong limit given by the Airy kernel.

\[\text{12}^\text{With the sign chosen that } p(x) > 0 \text{ for } x \in (a,b).\]
Laguerre Polynomials. The weight is \( w(x) = x^a e^{-x} \) on \( \Lambda = (0, \infty) \); hence
\[
p(x) = x, \quad r(x) = a - x, \quad q(x) = \frac{(a-x)^2}{4x} - \frac{1}{2}, \quad \lambda_n = n.
\]
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 \( m \geq n \); the Laguerre parameter \( a \) is then given by \( a = m - n \geq 0 \). Of particular interest in statistics (Johnstone 2001) is the simultaneous limit \( m, n \to \infty \) with
\[
\frac{m}{n} \to \theta \geq 1,
\]
for which we get
\[
k = 1, \quad k' = \frac{1}{2}, \quad k'' = -\frac{1}{2}, \quad \tilde{p}(t) = t, \quad \tilde{q}(t) = \frac{(\theta - 1 - t)^2}{4t}, \quad \omega = 1.
\]
Note that
\[
\omega - \tilde{q}(t) = \frac{(t_+ - t)(t - t_-)}{4t}, \quad t_\pm = (\sqrt{\theta} \pm 1)^2.
\]
Theorem 1 is applicable and we directly read off the following well-known scaling limits of the Wishart ensemble (Johnstone 2001):

- **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 \( \tilde{\rho}_n \) with a weak limit given by the *Marčenko–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 −,
  \[
  x = \pm n^{1/3} \theta^{-1/6} t^{2/3} \xi + nt_\pm
  \tag{3.1}
  \]
  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
\[
  x = \sigma \xi + \mu, \quad \mu = (\sqrt{m} \pm \sqrt{n})^2, \quad \sigma = (\sqrt{m} \pm \sqrt{n}) \left( \frac{1}{\sqrt{m}} \pm \frac{1}{\sqrt{n}} \right)^{1/3},
\]
which is obtained from (3.1) by replacing \( \theta \) with \( m/n \) (see Johnstone 2001, 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 \( a = m - n \) is a constant,\(^{13}\) \( x = \xi/(4n) \) induces
  \( \tilde{K}_n \) with a strong limit given by the *Bessel kernel* \( \tilde{K}_{\text{Bessel}}^{(x)} \).

\(^{13}\)By Remark 2.1, there is no need to restrict ourselves to \( a \geq 1 \); since \( \phi_n(x) = x^a \chi_n(x) \) with \( \chi_n(x) \) extending smoothly to \( x = 0 \), we have, for \( a \geq 0 \),
\[
x^{a/2}(2x\phi_n' - a\phi_n) = 2x^{1+a}\chi_n'(x) = O(x) \quad (x \to 0).
\]
Hence, the selfadjoint realization \( L_0 \) is compatible with the boundary condition (1.6).
Jacobi Polynomials. The weight is \( w(x) = x^\alpha (1 - x)^\beta \) on \( \Lambda = (0, 1) \); hence

\[
p(x) = x(1 - x), \quad r(x) = a - (a + \beta)x, \quad q(x) = \frac{(a - (a + \beta)x)^2}{4a(1 - x)} - \frac{a + \beta}{2},
\]

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 \( m_1, m_2 \geq n \); the Jacobi parameters \( a, \beta \) are then given by \( a = m_1 - n \geq 0 \) and \( \beta = m_2 - n \geq 0 \). Of particular interest in statistics (Johnstone 2008) is the simultaneous limit \( m_1, m_2, n \to \infty \) with

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

for which we get

\[
k = \kappa'' = 0, \quad \kappa' = 1, \quad \bar{p}(t) = t(1 - t), \quad \bar{q}(t) = \frac{(\theta - \tau - (1 - 2\tau)t)^2}{4\tau^2(1 - t)}, \quad \omega = 1 - \frac{\tau}{\theta}.
\]

Note that

\[
\omega - \bar{q}(t) = \frac{(t_+ - t)(t - t_-)}{4\tau^2(1 - t)}, \quad t_\pm = \left( \sqrt{\theta(1 - \tau)} \pm \sqrt{\tau(1 - \theta)} \right)^2.
\]

Theorem 1 is applicable and we directly read off the following (less well-known) scaling limits of the MANOVA ensemble (Collins 2005, Johnstone 2008):

- **bulk scaling limit:** if \( t_- < t < t_+ \),

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

induces \( \bar{K}_n \) with a strong limit given by the *Dyson kernel*;

- **limit law:** (because of \( \kappa = 0 \) there is no transformation here) \( \rho_n \) has a weak limit given by the law (Wachter 1980)

\[
\rho(t) = \frac{1}{2\pi \tau t(1 - t)} \sqrt{(t_+ - t)(t - t_-)},
\]

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

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

induces \( \bar{K}_n \) with a strong limit given by the *Airy kernel*.

**Remark 3.2.** Johnstone (2008, p. 2651) gives the soft-edge scaling in terms of a trigonometric parametrization of \( \theta \) and \( \tau \). By putting

\[
\theta = \sin^2 \frac{\phi}{2}, \quad \tau = \sin^2 \frac{\Psi}{2},
\]

we immediately get

\[
t_\pm = \sin^2 \frac{\phi \pm \Psi}{2}
\]
and (3.2) becomes
\[ x = \pm \sigma \pm \xi + t, \quad \sigma = n^{-2/3} \left( \frac{\tau^2 \sin^4 (\phi \pm \psi)}{4 \sin \phi \sin \psi} \right)^{1/3}. \]

In the case \( \theta = \tau = 1/2 \), which is equivalent to \( m_1/n, m_2/n \rightarrow 1 \), we have \( t_0 = 0 \) and \( t_\pm = 1 \). 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,\(^{14}\) \( x = \xi/(4n^2) \) 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 (Weidmann 1980, §9.3) links the convergence of differential operators, pointwise for an appropriate class of smooth test 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 \( a_n \rightarrow a \) and \( b_n \rightarrow b \), selfadjoint operators
\[ L : D(L) \subset L^2(a, b) \rightarrow L^2(a, b), \quad L_n : D(L_n) \subset L^2(a_n, b_n) \rightarrow L^2(a_n, b_n). \]

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 \( 1_{(a_n, b_n)} \), which we continue to denote by \( 1_{(a_n, b_n)} \), constitutes the orthogonal projection of \( L^2(a, b) \) onto \( L^2(a_n, b_n) \). Following Stolz and Weidmann (1993, §2), we say that \( L_n \) converges to \( L \) in the sense of generalized strong convergence \((\text{gsc})\), if for some \( z \in \mathbb{C} \setminus \mathbb{R} \), and hence, a fortiori, for all such \( z \),
\[ R_z(L_n)1_{(a_n, b_n)} \xrightarrow{s} R_z(L) \quad (n \rightarrow \infty) \]
in the strong operator topology of \( L^2(a, b) \).\(^{15}\)

**Theorem 2** (Stolz and Weidmann (1993, Thm. 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, \( C \subset D(L_n) \).

1. If \( L_n u \rightarrow Lu \) for all \( u \in C \), then \( L_n \xrightarrow{\text{gsc}} L \).
2. If \( L_n \xrightarrow{\text{gsc}} L \) and 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 \( \Delta \) converge as
\[ 1_{\Delta}(L_n)1_{(a_n, b_n)} \xrightarrow{s} 1_{\Delta}(L). \]

\(^{14}\)For the cases \( 0 \leq \alpha < 1 \) and \( 0 \leq \beta < 1 \), see the justification of the limit given in Footnote 13.

\(^{15}\)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 (Weidmann 1980, Thm. 8.27) this is then true for all such $c$ and $z$ and, moreover, in the LP case there exists, for every $z \in \mathbb{C} \setminus \mathbb{R}$, a one-dimensional space of solutions $u$ of the equation $(L - z)u = 0$ for which $u \in L^2(a, c)$. The same structure and notion holds for the boundary point $b$.

**Theorem 3.** 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

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

(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 (and a proof) that includes the LC case can be found, e.g., in Weidmann (1980, Thm. 8.22/8.29). In the following, we write (A.1) briefly in the form

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

If the imaginary part of the thus defined Green’s kernel $G_z(x, y)$ 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 4.** (i) Assume that there exits, as $\epsilon \downarrow 0$, the limit

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

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

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

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

$$\langle \mathbb{1}_\Delta(L)\phi, \psi \rangle = \int_\Delta (K_\lambda \phi, \psi)\,d\lambda \quad (\phi, \psi \in C_0^\infty(a, b)).$$

(A.2)

16See Weidmann (1987, pp. 41–42) for a proof that $C_0^\infty(a, b)$ is a core if the coefficients are smooth.
(ii) Assume further, for some \((a', b') \subset (a, b)\), that
\[
\int_{a'}^{b'} \int_{a'}^{b'} \left( \int_{\Delta} |K_\lambda(x, y)| \, d\lambda \right)^2 \, dx \, dy < \infty.
\]
Then \(\mathbbm{1}_{\Delta}(L) \mathbbm{1}_{(a', b')}\) is a Hilbert-Schmidt operator on \(L^2(a, b)\) with kernel
\[
\mathbbm{1}_{(a', b')} (y) \int_{\Delta} K_\lambda(x, y) \, d\lambda.
\]
If \(\int_{\Delta} K_\lambda(x, y) \, d\lambda\) is a continuous function of \(x, y \in (a', b')\), \(\mathbbm{1}_{\Delta}(L) \mathbbm{1}_{(a', b')}\) is a trace class operator with trace
\[
\text{tr} \mathbbm{1}_{\Delta}(L) \mathbbm{1}_{(a', b')} = \int_{a'}^{b'} \int_{a'}^{b'} K_\lambda(x, x) \, d\lambda \, dx.
\]

Proof. With \(E\) denoting the spectral resolution of the selfadjoint operator \(L\), we observe that, for a given \(\phi \in \mathcal{C}_0^\infty(a, b)\), 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 Lax 2002, §32.1):
\[
\int_{-\infty}^{\infty} \frac{d\mu_\phi(\lambda)}{\lambda - z} = (R_z(L) \phi, \phi).
\]
If we take \(z = \lambda + ie\) and let \(e \downarrow 0\), we obtain by the locally uniform convergence of the integral kernel of \(R_z\) that there exits either the limit
\[
\pi^{-1} \text{Im} \langle R_{\lambda+ie}(L) \phi, \phi \rangle \rightarrow \langle K_\lambda \phi, \phi \rangle
\]
or, at isolated points \(\lambda\),
\[
e \text{Im} \langle R_{\lambda+ie}(L) \phi, \phi \rangle \rightarrow 0.
\]
By a theorem of de la Vallée-Poussin (see Simon 2005, Thm. 11.6(ii/iii)), the singular part of \(\mu_\phi\) vanishes, \(\mu_{\phi, \text{sing}} = 0\); by Plemelj’s reconstruction the absolutely continuous part satisfies (see Simon 2005, Thm. 11.6(iv))
\[
d\mu_{\phi, \text{ac}}(\lambda) = \langle K_\lambda \phi, \phi \rangle \, d\lambda.
\]
Since \(\mathcal{C}_0^\infty(a, b)\) is dense in \(L^2(a, b)\), approximation shows that \(E_{\text{sing}} = 0\), that is, \(\sigma'(L) = \sigma_{\text{ac}}(L)\). Since \(\langle \mathbbm{1}_{\Delta}(L) \phi, \phi \rangle = \int_{\Delta} d\mu_{\phi}(\lambda)\), 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 Gohberg, Goldberg and Kruupnik (2000, Thm. IV.8.3) since \(\mathbbm{1}_{(a', b')} \mathbbm{1}_{\Delta}(L) \mathbbm{1}_{(a', b')}\) is a selfadjoint, positive-semidefinite operator.

We apply this theorem to the spectral projections used in the proof of Theorem 1. The first two examples could have been dealt with by Fourier techniques (Tao 2012, §3.3); applying, however, the same method in all the examples renders the approach more systematic.

Example 1 (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 \(\text{Im } z > 0\) the solutions \(u_a\) (\(u_b\)) of \((L - z)u = 0\) being \(L^2\) at \(-\infty\) \((\infty)\) are spanned by
\[
u_a(x) = e^{-izx}, \quad u_b(x) = e^{izx}.
\]
Thus, Theorem 3 applies: $L$ is essentially selfadjoint on $C_0^\infty(-\infty, \infty)$, the resolvent of its unique selfadjoint extension is represented, for $\text{Im} \, z > 0$, by the Green’s kernel
\[ G_z(x, y) = \frac{i}{2\sqrt{z}} \begin{cases} e^{(x-y)\sqrt{z}} & x > y \\ e^{-i(x-y)\sqrt{z}} & \text{otherwise.} \end{cases} \]

For $\lambda > 0$ there is the limit
\[ \pi^{-1} \text{Im} \, G_{\lambda+i0}(x, y) = K_\lambda(x, y) = \frac{\cos((x-y)\sqrt{\lambda})}{2\pi\sqrt{\lambda}}, \]

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
\[ \epsilon \text{ Im} \, G_{ie}(x, y) \to 0 \quad (\epsilon \downarrow 0), \]
locally uniform in $x, y \in \mathbb{R}$. Hence, Theorem 4 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
\[
\int_a^b \int_a^b \left( \int_{-\infty}^{\pi^2} |K_\lambda(x, y)| \, d\lambda \right)^2 \, dx \, dy \\
= \int_a^b \int_a^b \left( \int_0^{\pi^2} \left| \frac{\cos((x-y)\sqrt{\lambda})}{2\pi\sqrt{\lambda}} \right| \, d\lambda \right)^2 \, dx \, dy \\
\leq \left( \int_a^b \int_0^{\pi^2} \frac{d\lambda}{2\pi\sqrt{\lambda}} \right)^2 = (b-a)^2.
\]

Therefore, Theorem 4 yields that $1_{(-\infty, \pi^2)}(L)1_{(a,b)}$ is Hilbert–Schmidt with the Dyson kernel
\[ \int_{-\infty}^{\pi^2} K_\lambda(x, y) \, d\lambda = \int_0^{\pi^2} \frac{\cos((x-y)\sqrt{\lambda})}{2\pi\sqrt{\lambda}} \, d\lambda = \frac{\sin(\pi(x-y))}{\pi(x-y)}, \]

restricted to $x, y \in (a,b)$. Here, the last equality is simply obtained from
\[ (x-y) \int_0^{\pi^2} \frac{\cos((x-y)\sqrt{\lambda})}{2\sqrt{\lambda}} \, d\lambda = \int_0^{\pi^2} \frac{d\lambda}{2\pi\sqrt{\lambda}} \sin((x-y)\sqrt{\lambda}) \, d\lambda = \sin(\pi(x-y)). \]

Since the resulting kernel is continuous for $x, y \in (a,b)$, Theorem 4 gives that $1_{(-\infty, \pi^2)}(L)1_{(a,b)}$ is a trace class operator with trace
\[ \text{tr} \, 1_{(-\infty, \pi^2)}(L)1_{(a,b)} = b-a. \]

To summarize, we have thus obtained the following lemma.

**Lemma A.1.** 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, $1_{(-\infty, \pi^2)}(L)1_{(a,b)}$ is trace class with trace $b-a$ and kernel
\[ K_{\text{Dyson}}(x, y) = \int_0^{\pi^2} \frac{\cos((x-y)\sqrt{\lambda})}{2\pi\sqrt{\lambda}} \, d\lambda = \frac{\sin(\pi(x-y))}{\pi(x-y)} \quad (A.3) \]
Example 2 (Airy kernel). Consider the differential operator $Lu = -u'' + xu$ on $(-\infty, \infty)$. Since the specific solution $u(x) = Bi(x)$ of $Lu = 0$ is not locally $L^2$ at each of the endpoints, both endpoints are LP. For a given $\Im z > 0$ the solutions $u_a$ ($u_b$) of $(L-z)u = 0$ being $L^2$ at $-\infty$ ($\infty$) are spanned by (see Abramowitz and Stegun 1964, Eq. (10.4.59–64))

$$u_a(x) = Ai(x - z) - i Bi(x - z), \quad u_b(x) = Ai(x - z).$$

Thus, Theorem 3 applies: $L$ is essentially selfadjoint on $C^\infty_0(-\infty, \infty)$, the resolvent of its unique selfadjoint extension is represented, for $\Im z > 0$, by the Green's kernel

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

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

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

locally uniform in $x, y \in \mathbb{R}$. Hence, Theorem 4 applies: $\sigma(L) = \mathcal{S}_\infty(L) = \mathbb{R}$ 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

$$\left(\int_0^{s} \int_{s}^{s} \left(\int_{\Delta} |K_{\lambda}(x, y)| d\lambda \right)^2 dx dy\right)^{1/2} \leq \int_0^{s} \int_{0}^{s} \lambda Ai(x + \lambda)^2 d\lambda dx = \tau(s)$$

with

$$\tau(s) = \frac{1}{3} \left(2s^2 \Im(s)^2 - 2s \Im(s)^2 - \Im(s) \Re(s)\right).$$

Therefore, Theorem 4 yields that $1_{(-\infty, 0)}(L)1_{(s, \infty)}$ is Hilbert–Schmidt with the Airy kernel

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

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,

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

and partial integration to obtain

$$x \int_{0}^{s} Ai(x + \lambda) Ai(y + \lambda) d\lambda$$

$$= \int_{0}^{s} Ai''(x + \lambda) Ai(y + \lambda) d\lambda - \int_{0}^{s} \lambda Ai(x + \lambda) Ai(y + \lambda) d\lambda$$

$$= -Ai'(x) Ai(y) - \int_{0}^{s} Ai(x + \lambda) Ai'(y + \lambda) d\lambda - \int_{0}^{s} \lambda Ai(x + \lambda) Ai(y + \lambda) d\lambda.$$

Next, we exchange the roles of $x$ and $y$ and substract to get the assertion. Since the resulting kernel is continuous, Theorem 4 gives that $1_{(-\infty, 0)}(L)1_{(s, \infty)}$ is a trace class operator with trace

$$\text{tr} 1_{(-\infty, 0)}(L)1_{(s, \infty)} = \tau(s) \rightarrow \infty \quad (s \rightarrow -\infty).$$

To summarize, we have thus obtained the following lemma.
Lemma A.2. The differential operator $Lu = -u'' + xu$ is essentially selfadjoint on $C^0_0(-\infty, \infty)$. The spectrum of its unique selfadjoint extension is
\[ \sigma(L) = \sigma_{ac}(L) = (-\infty, \infty). \]

Given $s > -\infty$, the operator $\mathbb{1}_{(-\infty,0)}(L)\mathbb{1}_{(s,\infty)}$ is trace class with kernel
\[ K_{\text{Airy}}(x,y) = \int_0^\infty \text{Ai}(x + \lambda)\text{Ai}(y + \lambda) d\lambda = \frac{\text{Ai}(x)\text{Ai}'(y) - \text{Ai}'(x)\text{Ai}(y)}{x - y}. \quad (A.4) \]

Example 3 (Bessel kernel). Given $\alpha > 0$, take $Lu = -4(xu)' + \alpha^2x^{-1}u$ 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 $\alpha \geq 1$ 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 $\alpha \geq 1$.

For a given $\text{Im}z > 0$ the solutions $u_a(u_b)$ of $(L - z)u = 0$ being $L^2$ at $0 (\alpha)$ are spanned by (see Abramowitz and Stegun 1964, Eq. (9.1.7–9) and (9.2.5–6))
\[ u_a(x) = J_\alpha(\sqrt{\lambda x}), \quad u_b(x) = J_\alpha(\sqrt{\lambda x}) + i Y_\alpha(\sqrt{\lambda x}). \]

Thus, Theorem 3 applies: $L$ is essentially selfadjoint on $C^0_0(0, \infty)$, the resolvent of its unique selfadjoint extension is represented, for $\text{Im}z > 0$, by the Green’s kernel
\[ G_z(x,y) = \frac{i\pi}{4} \begin{cases} \int_0^\infty (J_\alpha(\sqrt{\lambda x}) + i Y_\alpha(\sqrt{\lambda x})) & x > y \\ \int_0^\infty (J_\alpha(\sqrt{\lambda x}) + i Y_\alpha(\sqrt{\lambda x})) & \text{otherwise}. \end{cases} \]

For $\lambda > 0$ there is the limit
\[ \pi^{-1} \text{Im} G_{\lambda+i0}(x,y) = K_\lambda(x,y) = \frac{1}{4} J_\alpha(\sqrt{\lambda^2}) J_\alpha(\sqrt{\lambda^2}), \]
for $\lambda \leq 0$ the limit is zero; both limits are locally uniform in $x, y \in \mathbb{R}$. Hence, Theorem 4 applies: $\sigma(L) = \sigma_{ac}(L) = [0, \infty)$ and (A.2) holds for each Borel set $\Delta \subset \mathbb{R}$. Given $0 \leq s < \infty$, we may estimate for the specific choice $\Delta = (-\infty, 1)$ that
\[ \left( \int_0^s \int_0^s \left( \int_\Delta |K_\lambda(x,y)| d\lambda \right)^2 dx dy \right)^{1/2} \leq \frac{1}{4} \int_0^s \int_0^1 J_\alpha(\sqrt{\lambda^2})^2 d\lambda dx = \tau_\alpha(s). \]

Therefore, Theorem 4 yields that $\mathbb{1}_{(-\infty,1)}(L)\mathbb{1}_{(0,s)}$ is Hilbert–Schmidt with the Bessel kernel
\[ \int_{-\infty}^1 K_\lambda(x,y) d\lambda = \frac{1}{4} \int_0^1 J_\alpha(\sqrt{\lambda^2}) J_\alpha(\sqrt{\lambda^2}) d\lambda \]
\[ = \frac{J_\alpha(\sqrt{\lambda}) \sqrt{\lambda} J'_\alpha(\sqrt{\lambda}) - \sqrt{\lambda} J'_\alpha(\sqrt{\lambda}) J_\alpha(\sqrt{\lambda})}{2(x - y)}, \]
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,
\[ x J_\alpha(\sqrt{\lambda^2}) = -4 \frac{d}{d\lambda} \left( \lambda \frac{d}{d\lambda} J_\alpha(\sqrt{\lambda^2}) \right) + \alpha^2 \lambda^{-1} J_\alpha(\sqrt{\lambda^2}), \]
and partial integration to obtain
\[
\frac{x}{4} \int_0^1 J_a(\sqrt{x\lambda})J_a(\sqrt{y\lambda}) \, d\lambda
\]
\[
= - \int_0^1 \frac{d}{d\lambda} \left( \lambda \frac{d}{d\lambda} J_a(\sqrt{x\lambda}) \right) J_a(\sqrt{y\lambda}) \, d\lambda + \frac{\alpha^2}{4} \int_0^1 \lambda^{-1} J_a(\sqrt{x\lambda})J_a(\sqrt{y\lambda}) \, d\lambda
\]
\[
= - \frac{1}{2} \sqrt{x} J'_a(\sqrt{x}) J_a(\sqrt{y}) + \int_0^1 \lambda \left( \frac{d}{d\lambda} J_a(\sqrt{x\lambda}) \right) \left( \frac{d}{d\lambda} J_a(\sqrt{y\lambda}) \right) \, d\lambda + \frac{\alpha^2}{4} \int_0^1 \lambda^{-1} J_a(\sqrt{x\lambda})J_a(\sqrt{y\lambda}) \, d\lambda.
\]
Next, we exchange the roles of \(x\) and \(y\) and substract to get the assertion. Since the resulting kernel is continuous, Theorem 4 gives that \(1_{(-\infty,1)}(L)1_{(0,\infty)}\) is a trace class operator with trace
\[
\text{tr} 1_{(-\infty,1)}(L)1_{(0,\infty)} = \tau_a(s) \to \infty \quad (s \to \infty).
\]
To summarize, we have thus obtained the following lemma.

**Lemma A.3.** Given \(\alpha \geq 1\), 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
\[
\sigma(L) = \sigma_{ac}(L) = [0, \infty).
\]
Given \(0 \leq s < \infty\), the operator \(1_{(-\infty,1)}(L)1_{(0,s)}\) is trace class with kernel
\[
K^{(\alpha)}_{\text{Bessel}}(x,y) = \frac{1}{4} \int_0^1 J_a(\sqrt{x\lambda})J_a(\sqrt{y\lambda}) \, d\lambda
\]
\[
= \frac{J_a(\sqrt{x})\sqrt{y} J'_a(\sqrt{y}) - \sqrt{x} J'_a(\sqrt{x}) J_a(\sqrt{y})}{2(x - y)}.
\] (A.5)

**Remark A.1.** Lemma A.3 extends to \(0 \leq \alpha < 1\) if we choose the particular selfadjoint realization of \(L\) that is defined by the boundary condition (1.6), cf. Hutson, Pym and Cloud (2005, Example 10.5.12).

**References**

Abramowitz, M. and Stegun, I. A.: 1964, Handbook of mathematical functions with formulas, graphs, and mathematical tables, National Bureau of Standards, Washington.

Anderson, G. W., Guionnet, A. and Zeitouni, O.: 2010, An Introduction to Random Matrices, Cambridge University Press.

Andrews, G. E., Askey, R. and Roy, R.: 1999, Special Functions, Cambridge University Press, Cambridge.

Borodin, A. and Olshanski, G.: 2007, Asymptotics of Plancherel-type random partitions, J. Algebra 313, 40–60.

Collins, B.: 2005, Product of random projections, Jacobi ensembles and universality problems arising from free probability, Probab. Theory Related Fields 133, 315–344.

Deift, P.: 1999, Orthogonal polynomials and random matrices: a Riemann-Hilbert approach, AMS, Providence.

Erdélyi, A., Magnus, W., Oberhettinger, F. and Tricomi, F. G.: 1953, Higher transcendental functions. Vol. II, McGraw-Hill, New York.

Gohberg, I., Goldberg, S. and Krupnik, N.: 2000, Traces and determinants of linear operators, Birkhäuser Verlag, Basel.

Hutson, V., Pym, J. S. and Cloud, M. J.: 2005, Applications of functional analysis and operator theory, second edn, Elsevier B. V., Amsterdam.
Johnstone, I. M.: 2001, On the distribution of the largest eigenvalue in principal components analysis, *Ann. Statist.* 29, 295–327.

Johnstone, I. M.: 2008, Multivariate analysis and Jacobi ensembles: largest eigenvalue, Tracy-Widom limits and rates of convergence, *Ann. Statist.* 36, 2638–2716.

Kuijlaars, A. B. J.: 2011, Universality, in G. Akemann, J. Baik and P. D. Francesco (eds), *Oxford Handbook of Random Matrix Theory*, Oxford University Press. arXiv:1103.5922.

Lax, P. D.: 2002, *Functional analysis*, Wiley-Interscience, New York.

Simon, B.: 2005, *Trace ideals and their applications*, second edn, American Mathematical Society, Providence, RI.

Stolz, G. and Weidmann, J.: 1993, Approximation of isolated eigenvalues of ordinary differential operators, *J. Reine Angew. Math.* 445, 31–44.

Tao, T.: 2012, *Topics in random matrix theory*, American Mathematical Society.

Wachter, K. W.: 1980, The limiting empirical measure of multiple discriminant ratios, *Ann. Statist.* 8, 937–957.

Weidmann, J.: 1980, *Linear operators in Hilbert spaces*, Springer-Verlag, New York.

Weidmann, J.: 1987, *Spectral theory of ordinary differential operators*, Springer-Verlag, Berlin.

Zentrum Mathematik – M3, Technische Universität München, 80290 München, Germany

E-mail address: bornemann@ma.tum.de