Operators and Special Functions in Random Matrix Theory

Andrew McCafferty, MSci

Submitted for the degree of Doctor of Philosophy
at Lancaster University
March 2008

Operators and Special Functions in Random Matrix Theory

Andrew McCafferty, MSci

Submitted for the degree of Doctor of Philosophy

at Lancaster University, March 2008

Abstract

The Fredholm determinants of integral operators with kernel of the form

$$\frac{A(x)B(y) - A(y)B(x)}{x - y}$$

arise in probabilistic calculations in Random Matrix Theory. These were extensively studied by Tracy and Widom, so we refer to them as Tracy–Widom operators. We prove that the integral operator with Jacobi kernel converges in trace norm to the integral operator with Bessel kernel under a hard edge scaling, using limits derived from convergence of differential equation coefficients. The eigenvectors of an operator with kernel of Tracy–Widom type can sometimes be deduced via a commuting differential operator. We show that no such operator exists for TW integral operators acting on $L^2(\mathbb{R})$. There are analogous operators for discrete random matrix ensembles, and we give sufficient conditions for these to be expressed as the square of a Hankel operator: writing an operator in this way aids calculation of Fredholm determinants. We also give a new example of discrete TW operator which can be expressed as the sum of a Hankel square and a Toeplitz operator.

Previously unsolvable equations are dealt with by threats of reprisals . . .

Woody Allen

Acknowledgements

I would like to thank many people for helping me through what has sometimes been a difficult three years. First must come my supervisor Gordon Blower, who put up with my many questions, and made many helpful suggestions. Next, my family, in particular my parents, and the many friends I have made during my studies, particularly at the Chaplaincy Centre. Lastly, to Andrea for her “remote” support, and for chivvying me along in the final stages: *vielen dank!*.

Financial support for my studies was provided by an EPSRC doctoral training grant.

Declaration

The work in this thesis is my own, except where otherwise indicated, and has not been submitted elsewhere for the award of a higher degree. There are reworked proofs of standard results included for completeness, but where this is the case, this will be indicated in the text. Part of Chapter 5 has been accepted for publication in the the Proceedings of the Edinburgh Mathematical Society [8].

Contents

1 Introduction and background results

1.1 Introduction

Random matrix theory began in 1928, when the concept of a random matrix was introduced in a paper by Wishart [49]. Wigner [48], who worked on the energy levels in atomic nuclei, observed that such systems could be described well by the eigenvalues of a random matrix, and calculated asymptotic eigenvalue distributions. In this thesis, we consider a special type of operator which is associated with probabilistic calculations for eigenvalue distributions. The basic form of the operator kernel we consider is

$$K(x,y) = \frac{A(x)B(y) - A(y)B(x)}{x-y}, \tag{1}$$

where the functions A and B are varied depending on the context. The prototypical example, as considered by Tracy and Widom in [44], and also Blower, in [6] is when A and B are bounded functions on $\mathbb{R}$, and $K(x,y)$ is the kernel of an integral operator on $L^2(\mathbb{R})$. Tracy and Widom observed in [44] that many of the important examples in random matrix theory involve pairs of functions A and B in the kernel $K(x,y)$ which satisfy differential equations of the form

$$m(x)\frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} A(x) \\ B(x) \end{bmatrix} = \begin{bmatrix} \alpha(x) & \beta(x) \\ -\gamma(x) & -\alpha(x) \end{bmatrix}\begin{bmatrix} A(x) \\ B(x) \end{bmatrix}, \tag{2}$$

where $m(x)$, $\alpha(x)$, $\beta(x)$ and $\gamma(x)$ are polynomials. We shall call these *Tracy–Widom systems*. The Fredholm determinants $\det(I-K)$ of such operators can be used to calculate for instance the probability that a given interval contains n eigenvalues of a random matrix. The Airy kernel

$$\frac{\mathrm{Ai}\,(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}\,(y)\,\mathrm{Ai}'(x)}{x-y}, \tag{3}$$

and Bessel kernel

$$\frac{\mathrm{J}_\alpha(\sqrt{x})\sqrt{y}\,\mathrm{J}'_\alpha(\sqrt{y}) - \mathrm{J}_\alpha(\sqrt{y})\sqrt{x}\,\mathrm{J}'_\alpha(\sqrt{x})}{(x-y)}. \tag{4}$$

are particular examples of kernels satisfying equations of the form (2).

In Chapter 2, we prove that the integral operator which describes the eigenvalues at the edge of the spectrum of a random matrix from the Jacobi unitary ensemble converges in trace class norm to the Bessel integral operator, with kernel (4). This result is important because trace class convergence implies convergence of determinants, and hence of probabilities. Others, including [5], [15], [31], [45] make reference to this result in terms of convergence of kernels, but as far as we can see, they do not make clear the mode of convergence of the operators. Central to the proof are some limits of scaled Jacobi polynomials, which we establish using convergence of differential equation coefficients. Since many other orthogonal polynomial systems satisfy simple differential equations, this approach could be used in other cases. Thus, our methods provide a more elementary alternative to the asymptotic expansions based on Riemann Hilbert theory used by other authors to establish convergence theorems for Tracy–Widom kernels.

Several authors (including [27, pp.98-101] and [45, §III B]) calculate the eigenfunctions of a Tracy–Widom integral operator by finding a differential operator which commutes with it. A particular example of this is the operator

$$-\frac{\mathrm{d}}{\mathrm{d}x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right) + x^2$$

which commutes with the integral operator with Airy kernel (3), operating on $L^2[0,\infty)$. In Chapter 3, we prove that no non-zero Tracy–Widom operator operating on $L^2(\mathbb{R})$ can commute with a self-adjoint differential operator. The proof relies on the fact that the Hilbert transform commutes with differentiation on

$L^2(\mathbb{R})$.

By analogy with the continuous case, we consider operators whose matrix entries are $K(m,n)$, in which A and B are now functions $\mathbb{Z}_+ \to \mathbb{R}$ which satisfy a one-step difference equation of the form

$$\begin{bmatrix} A(n+1) \\ B(n+1) \end{bmatrix} = T(n) \begin{bmatrix} A(n) \\ B(n) \end{bmatrix},$$

where $T(x)$ is a 2×2 matrix of rational functions of x with $\det T(x) = 1$. Such operators arise in the theory of discrete random matrix ensembles, as in [4] and [21], where they play the same role as in the continuous case. One-step difference equations of this form are also important in the theory of discrete Schrödinger operators, and are investigated in the context of Anderson localisation (see [8]). Tracy and Widom, in [43] and [45] proved that the Airy and Bessel integral kernels could be expressed as the squares of Hankel integral operators, enabling them to calculate their eigenfunctions and eigenvalues. Seeing the utility of expressions of this form (in particular, the fact that $\det(I-K) = \det(I+\Gamma)\det(I-\Gamma)$), Blower [7] gave sufficient conditions for a Tracy–Widom operator to be expressible as Γ^2 or $\Gamma^*\Gamma$, where Γ is a Hankel integral operator. In Chapter 5, we prove a new result which gives sufficient conditions for the matrix K to be expressible as Γ^2, where Γ is a Hankel matrix. We also consider a weaker condition, in which $K = W + \Gamma^2$, where W is a Toeplitz operator. Although exact calculations of the eigenvalues of K from those of Γ is not possible in the latter case, expressions of this kind may still be useful, since the spectrum of Toeplitz operators can often be calculated. In Proposition 5.7, we give what appears to be a new example of this type of operator, which we present as an analogue to the Laguerre integral operator. Theorem 5.1 and Proposition 5.7 appear in a paper by the author and G. Blower in the Proceedings of the Edinburgh Mathematical Society [8].

The final part of Chapter 5 considers briefly Tracy–Widom operators which are compact perturbations of a Hankel square. We use Weyl's theorem to deduce the essential spectrum of a particular discrete Tracy–Widom operator, which we show is a Hilbert–Schmidt perturbation of the square of the operator induced by Hilbert's Hankel matrix.

In Chapter 4, we consider the Tracy–Widom operators which describe the eigenvalue distributions for circular ensembles. We use standard results on Hankel and Toeplitz operators to give sufficient conditions for these operators to be written as $\Gamma^*\Gamma$, where Γ is a Hankel operator on $H^2(\mathbb{T})$. In the particular case where the functions in the kernel are Blaschke products, we show that the range of the operator can be calculated explicitly.

The remainder of this first chapter introduces most of the background in Random Matrix Theory and operator theory needed for later chapters, and (in §1.9) relevant special functions. In §1.8, we show how to define orthogonal polynomials, and demonstrate (in §1.9 and §1.10) how they can be used to express the joint probability density for the eigenvalues of a random matrix as a determinant of the form

$$\det \left[K(\lambda_j, \lambda_k)\right]_{j,k},$$

where $0 \leq K \leq I$, and K is trace class, as in Soshnikov's theory of determinantal point processes [40]. The material in this chapter is standard, but its presentation has been adapted to fit our purposes, and to point out how it applies to the problems we tackle later in the thesis.

1.2 Hardy spaces and the unilateral shift operator

Here we shall introduce notation and definitions for the important Hilbert spaces which will be used throughout the thesis. To begin with, for ease of reference, we state the following well-known and basic theorem.

Theorem 1.1 (Riesz–Fischer) *Let (e_k) be an orthonormal basis for a Hilbert space H. Then, for every $x \in H$ we have*

$$x = \sum_k \langle x, e_k \rangle e_k, \tag{5}$$

and

$$\|x\|^2 = \sum_k |\langle x, e_k \rangle|^2. \tag{6}$$

Remark

Although we refer to this as the Riesz–Fischer theorem, it should be noted that these authors were actually considering the specific case of trigonometric expansions, as in Theorem 1.2 below.

Proof. Let $y = x - \sum_k \langle x, e_k \rangle e_k$. Then, for any j

$$\langle y, e_j \rangle = \langle x, e_j \rangle - \sum_k \langle x, e_k \rangle \langle e_k, e_j \rangle = 0.$$

Since (e_k) is an orthonormal basis, we know that the only vector orthogonal to all the e_j for all j is the zero vector, so we deduce that $y = 0$, which gives (5). The second equation (6) then follows by Pythagoras theorem and the continuity of the norm. ∎

The general outline of our treatment of Hardy spaces follows Martinez-Rosenthal [26, Chapter 1]. We write $\mathbb{Z}_+ = \{0, 1, 2, 3, \ldots\}$, and for the Hilbert space of square-summable sequences indexed by $\mathbb{Z}_+$, we write

$$\ell^2 = \{(a_n)_{n=0}^{\infty} : \sum_{n=0}^{\infty} |a_n|^2 < \infty\},$$

where the inner product of two vectors $a, b \in \ell^2$ is

$$\langle a, b\rangle = \sum_{n=0}^{\infty} a_n \bar{b_n}.$$

As usual, the Kronecker delta function is defined as

$$\delta_{j,k} = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases} \tag{7}$$

Let the open unit disk be denoted by $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$, and the unit circle by $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. The Hardy Hilbert space $H^2(\mathbb{D})$ of analytic functions whose power series representations have square-summable coefficients is

$$H^2(\mathbb{D}) = \left\{ f : f(z) = \sum_{n=0}^{\infty} a_n z^n, \text{ with } \sum_{n=0}^{\infty} |a_n|^2 < \infty \right\},$$

and the inner product of two functions $f, g \in H^2(\mathbb{D})$ expressed as $f(z) = \sum_{n\geq 0} a_n z^n$ and $g(z) = \sum_{n\geq 0} b_n z^n$ is

$$\langle f, g\rangle = \sum_{n=0}^{\infty} a_n \overline{b_n}.$$

We can show that any function in $H^2(\mathbb{D})$ is analytic on $\mathbb{D}$ as follows. Take $f \in H^2(\mathbb{D})$ with $f(z) = \sum_{n=0}^{\infty} a_n z^n$. Then, by definition, (a_n) is in ℓ^2, so there exists $M > 0$ such that $|a_n| < M$ for all $n \geq 0$. Take $z_0 \in \mathbb{D}$. Then

$$\sum_{n=0}^{\infty} |a_n z_0^n| \leq M \sum_{n=0}^{\infty} |z_0|^n,$$

where the right-hand side is a convergent geometric series, so the series $\sum_{n=0}^{\infty} a_n z_0^n$ is absolutely convergent. Thus $f(z)$ is an analytic function on $\mathbb{D}$. The spaces $H^2(\mathbb{D})$ and ℓ^2 may be identified by the natural isometric isomorphism $(a_n) \mapsto \sum_{n=0}^{\infty} a_n z^n$. Since ℓ^2 is a Hilbert space, this shows that $H^2(\mathbb{D})$ is also a Hilbert space.

$L^2(\mathbb{T})$ is the Hilbert space of square integrable functions on the unit circle $\mathbb{T}$, with inner product

$$\langle f, g\rangle = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})\overline{g(e^{i\theta})}\,\mathrm{d}\theta. \tag{8}$$

It is a well-known result (see, for example [10, p.21]) that $(e^{in\theta})_{n=-\infty}^{\infty}$ forms an orthonormal basis for $L^2(\mathbb{T})$, so that any function $f \in L^2(\mathbb{T})$ can be written as $f(e^{i\theta}) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{in\theta}$, where we define the n^{th} Fourier coefficient to be

$$\hat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})e^{-in\theta}\,\mathrm{d}\theta.$$

Theorem 1.2 (Parseval) *Let $f \in L^2(\mathbb{T})$, with Fourier coefficients $\hat{f}(n)$. Then*

$$\frac{1}{2\pi}\int_0^{2\pi} \left|f(e^{i\theta})\right|^2 \mathrm{d}\theta = \sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2.$$

Proof. This is simply a consequence of Theorem 1.1, applied to the orthonormal basis $(e^{in\theta})_{n=-\infty}^{\infty}$. ■

For convenience, we shall usually write $z = e^{i\theta}$, so that the standard orthonormal basis for $L^2(\mathbb{T})$ is $(z^n)_{n=-\infty}^{\infty}$. We can define a subspace $H^2(\mathbb{T})$ of $L^2(\mathbb{T})$ by selecting functions for which the negative Fourier coefficients are zero:

$$H^2(\mathbb{T}) = \{f \in L^2(\mathbb{T}) : \hat{f}(n) = 0 \text{ for } n < 0\}.$$

A typical $H^2(\mathbb{T})$ function has the form $f(e^{i\theta}) = \sum_{n=0}^{\infty} a_n e^{in\theta}$, so we see that this space is naturally isomorphic to $H^2(\mathbb{D})$ via the correspondence $\sum_{n=0}^{\infty} a_n e^{in\theta} \mapsto \sum_{n=0}^{\infty} a_n z^n$. We write P_+ and P_- for the Riesz projection operators of $L^2(\mathbb{T})$ onto (respectively) $H^2(\mathbb{T})$ and onto its orthogonal complement, the space

$$H^2_-(\mathbb{T}) = \{f \in L^2(\mathbb{T}) : \hat{f}(n) = 0 \text{ for } n \geq 0\}.$$

To be more explicit, we state the effect on the standard orthonormal basis $(z^n)_{n=-\infty}^{\infty}$ of the projections P_+ and P_-:

$$P_+ z^n = \begin{cases} z^n & \text{if } n \geq 0 \\ 0 & \text{if } n < 0 \end{cases} \tag{9}$$

and

$$P_- z^n = \begin{cases} z^n & \text{if } n < 0 \\ 0 & \text{if } n \geq 0. \end{cases} \tag{10}$$

A measurable function ϕ on $\mathbb{T}$ is said to be *essentially bounded* if there exists a number M_0 such that

$$m\left(\{z \in \mathbb{T} : |\phi(z)| > M_0\}\right) = 0,$$

where m is the normal Lebesgue measure on the unit circle. $L^\infty(\mathbb{T})$ is a Banach space which is a subspace of $L^2(\mathbb{T})$, consisting of essentially bounded functions on the unit circle, with norm

$$\|f\|_{L^\infty} = \inf\{M : m\{z \in \mathbb{T} : |\phi(z)| > M\} = 0\}.$$

The unilateral shift operator S is of fundamental importance in functional analysis. For instance, its invariant subspaces have been closely studied, and later we shall see that it provides a characterisation of Hankel operators. To define S, we take a sequence $(a_0, a_1, a_2, \ldots) \in \ell^2$, and then $S(a_0, a_1, a_2, \ldots) = (0, a_0, a_1, \ldots)$. To calculate the adjoint S^*, take $a, b \in \ell^2$ with $a = (a_0, a_1, a_2, \ldots)$ and $b = (b_0, b_1, b_2, \ldots)$, and set $c = S^* b$. Then since $\langle Sa, b\rangle = \langle a, S^* b\rangle$, we have

$$\sum_{n=1}^{\infty} a_{n-1}\overline{b_n} = \sum_{n=0}^{\infty} a_n \overline{c_n},$$

so that $c_n = b_{n+1}$. Hence $S^*b = (b_1, b_2, \ldots)$. For this reason, S^* is sometimes called the backwards shift operator. Since ℓ^2 is naturally isomorphic to $H^2(\mathbb{T})$, we can also define the shift S on the latter space, via the correspondence $(\hat{f}(n)) \mapsto f$. We have $S(\hat{f}(0), \hat{f}(1), \ldots) = (0, \hat{f}(0), \hat{f}(1), \ldots)$, so it is clear that the operation of S on $H^2(\mathbb{T})$ is multiplication by the independent variable z.

Throughout this thesis, whenever we refer to a subspace, it is assumed without comment that it is a closed linear subspace. Also, all Hilbert spaces are assumed to be complex and seperable. Let E be a subspace of a Hilbert space H, and let A be an operator on H. If the set $AE = \{Ax : x \in E\}$ is a subspace of E, we say that E is an *invariant subspace* for A. If the inclusion $AE \subset E$ is proper, E is said to be *simply invariant*, while if in fact $AE = E$, then it is *doubly invariant*. We now present the famous result of Beurling on the invariant subspaces in $H^2(\mathbb{T})$ of the shift operator S, which we shall use later on in Chapter 4. The proof follows [2, p.46], but is included here for completeness. Since the operation of the shift S on $H^2(\mathbb{T})$ is multiplication by the independent variable z, a subspace E is doubly invariant with respect to the shift if and only if $zE = E$, where $zE = \{ze : e \in E\}$. Because $z\bar{z} = 1$ on $\mathbb{T}$, this is equivalent to the two inclusions $zE \subset E$ and $\bar{z}E \subset E$, a fact that will be useful in the proof.

Theorem 1.3 (Beurling's Theorem) *If $E \neq \{0\}$ is a subspace of $H^2(\mathbb{T})$ such that SE is a subspace of E, then there exists a measurable function $\Theta \in H^2(\mathbb{T})$ with $|\Theta| = 1$ a.e. on $\mathbb{T}$ such that $E = \Theta H^2(\mathbb{T})$.*

Proof. First, we need to show that E is simply invariant, i.e. that $SE \neq E$. As noted above, this is equivalent to showing that $\bar{z}E$ is not a subspace of E. Assume the contrary, and take $f \in E$ with $f \neq 0$. Then there exists $n \in \mathbb{Z}_+$ such that $\hat{f}(n) \neq 0$, so we have $\widehat{g}(-1) = \hat{f}(n) \neq 0$, where $g(z) = \bar{z}^{n+1} f(z)$. But this implies that $\bar{z}^{n+1} f \notin H^2$, so that $\bar{z}^{n+1} f \notin E$, which by our assumption implies that $f \notin E$. This contradiction shows that SE is a proper subspace of E. Hence $E \ominus SE$ is non-trivial, so we can select $\Theta \in E \ominus SE$ with $\|\Theta\| = 1$. Note that, by invariance, $S^n\Theta \in E$ for all $n > 0$, and also that $S^n\Theta \in SE$. Hence

$$0 = \langle S^n\Theta, \Theta\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{in\theta}\left|\Theta(e^{i\theta})\right|^2 \,\mathrm{d}\theta \quad \text{for all } n > 0.$$

If we take complex conjugates in the integral above, we get the same condition for $n < 0$ as well, and hence $|\Theta|$ is constant almost everywhere. But $\|\Theta\| = 1$, so this gives $|\Theta| = 1$ almost everywhere on $\mathbb{T}$. We see therefore that the operation of multiplication by Θ is isometric, and so, using $\vee$ to stand for "closed linear span", we have

$$\bigvee_{n=0}^{\infty}\{S^n\Theta\} = \Theta\bigvee_{n=0}^{\infty}\{S^n 1\} = \Theta H^2.$$

This shows that $\Theta H^2 \subseteq E$. To prove equality, choose $f \in E \ominus \Theta H^2$. By the above, we have $S^n\Theta \in \Theta H^2$ for all $n \geq 0$, so

$$\langle f, S^n\Theta\rangle = 0 \quad \text{for all } n \geq 0.$$

Also, it is clear that $S^n f \in SE$ and $S^n f \in E$, so that

$$\langle \Theta, S^n f\rangle = 0 \quad \text{for all } n > 0,$$

where we recall that $\Theta \in E \ominus SE$. Putting these two facts together, we get $f\overline{\Theta} = 0$ almost everywhere on $\mathbb{T}$, so that $f = 0$ and hence $E = \Theta H^2$. ■

Remark

A function $\Theta \in H^2(\mathbb{T})$ which satisfies $|\Theta| = 1$ a.e. on $\mathbb{T}$ is called an *inner function.*

1.3 The spectrum and essential spectrum of an operator

The spectrum of an operator can be seen as the generalisation of its set of eigenvalues. It is perfectly possible to define the spectrum of an operator in Banach space, but for our purposes we shall only need to define it for operators on a Hilbert space.

Definition 1.4 *Let $A : H \to H$ be a bounded linear operator on a Hilbert space H. Then the* spectrum *of A is the set*

$$\sigma(A) = \{\lambda \in \mathbb{C} : A - \lambda I \textit{ is not invertible}\}.$$

We list a few well-known facts about the spectrum. First, note that $\sigma(A)$ is a non-empty set. To see why this is so, suppose for a contradiction that $\sigma(A)$ is empty. Then (as [50, p.80] notes) for all $\lambda \in \mathbb{C}$ we would have $(A - \lambda I)^{-1}$ bounded, and the map

$$\lambda \mapsto (A - \lambda I)^{-1}$$

from $\mathbb{C}$ into the space $B(H)$ of bounded operators on H would be bounded and entire, and hence constant, by an operator-valued version of Liouville's theorem. Since $(A - \lambda I)^{-1}$ cannot be constant, we conclude that $\sigma(A)$ is non-empty. It can also be shown that $\sigma(A)$ is a compact subset of $\mathbb{C}$ (see e.g. [17, p. 226]). The complement of $\sigma(A)$ in $\mathbb{C}$ is called the *resolvent set* of A. If an operator has eigenvalues, these are clearly contained in the spectrum: the subset of the spectrum consisting of eigenvalues of the operator is called the *point spectrum*, and written as $\sigma_p(A)$. In particular, the spectrum of a compact operator consists of 0 together with a sequence of eigenvalues of finite multiplicity. This sequence, if infinite, converges to zero. The essential spectrum does not include such points, and may be defined as the following subset of the spectrum:

$$\sigma_{ess}(A) = \{\lambda \in \sigma(A) : \lambda \text{ is not an isolated eigenvalue of finite multiplicity}\}.$$

1.4 Hilbert–Schmidt and Trace class operators, and operator convergence

Throughout this section, H is a complex, infinite-dimensional and seperable Hilbert space, and all operators we consider are assumed to be bounded. An operator A on H is said to be *positive* if

$$\langle Ax, x\rangle \geq 0 \text{ for all } x \in H.$$

If A is positive and self-adjoint, then we can define a square root, that is, a positive self-adjoint operator B that satisfies $B^2 = A$ (see [25, p.157] for a proof of this). We write this B as $A^{1/2}$. The theory of trace class and Hilbert–Schmidt operators can be tackled from different directions: our approach here follows broadly [34]. Although all the results are well known, we include their proofs in the interest of completeness. We begin with a proof of the result that every compact operator has an expansion in terms of its generalised eigenvectors and eigenvalues.

Proposition 1.5 *A linear operator A on H is compact if and only if there exists a finite or countably infinite sequence of scalars $\sigma_j(A)$ decreasing to zero, and orthonormal sequences (v_i) and (w_i) in H, such that*

$$Ax = \sum_{j=0}^{\infty} \sigma_j(A)\langle x, v_j\rangle w_j \tag{11}$$

for all $x \in H$.

Proof. First, suppose that the expansion (11) holds, and let

$$A_m x = \sum_{j=0}^{m} \sigma_j(A)\langle x, v_j\rangle w_j.$$

Then A_m is a sequence of finite rank (at most m) operators, and

$$\|A - A_m\| = \sigma_{m+1} \to 0$$

as $m \to \infty$. Hence A is compact, since it is the norm limit of a sequence of finite rank operators. Conversely, suppose that A is compact. Note that A^*A is a positive and self-adjoint operator, since

$$\langle A^*Ax, x\rangle = \langle Ax, Ax\rangle \geq 0,$$

and

$$\langle A^*Ax, y\rangle = \langle x, (A^*A)^*y\rangle = \langle x, A^*Ay\rangle.$$

Hence if λ_j are the non-zero eigenvalues arranged in decreasing order of size, we have $\lambda_j \geq 0$ for all j, and we may define $\sigma_j(A) := \sqrt{\lambda_j}$. Let v_j be the corresponding normalised eigenvectors, which form an orthonormal sequence. Also, let $w_j = \frac{1}{\sigma_j(A)}Av_j$. Then (w_j) is an orthonormal sequence:

$$\begin{aligned}
\langle w_j, w_k\rangle &= \frac{1}{\sigma_j(A)\sigma_k(A)}\langle Av_j, Av_k\rangle \\
&= \frac{1}{\sigma_j(A)\sigma_k(A)}\langle A^*Av_j, v_k\rangle \\
&= \frac{\sigma_j(A)^2}{\sigma_j(A)\sigma_k(A)}\langle v_j, v_k\rangle \\
&= \delta_{j,k}.
\end{aligned}$$

By the spectral theorem for compact self-adjoint operators (as in [50, p.99]), we now have the required expansion (11). ■

Definition 1.6 *The real sequence* $(\sigma_j(A))$ *is called the sequence of* singular numbers *or* generalised eigenvalues *for the operator* A*. For* $1 \leq p < \infty$*, we say that a compact operator* A *is in the* Schatten–von Neumann class C_p *if its singular* *numbers satisfy*

$$\sum_{j=0}^{\infty} \sigma_j(A)^p < \infty. \tag{12}$$

Recall that C_p is a Banach space with norm

$$\|A\|_{C_p} = \left(\sum_{j=0}^{\infty} \sigma_j(A)^p \right)^{1/p}.$$

The two classes of interest in this thesis (and indeed in many applications) are C_1, which we call the *trace class* operators, and C_2, which are the *Hilbert–Schmidt operators*. The norm of Hilbert–Schmidt operators is particularly easy to calculate, as the next result shows.

Proposition 1.7 *If $A \in C_2$, then for any orthonormal basis (e_j) of H, we have*

$$\|A\|_{C_2}^2 = \sum_{j=0}^{\infty} \|Ae_j\|^2 = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |\langle Ae_j, e_k \rangle|^2. \tag{13}$$

Remark

If we represent an operator A by a matrix $[A_{m,n}]_{m,n \geq 0}$ with respect to some orthonormal basis, then the second sum in (13) is simply $\sum_{m,n \geq 0} |A_{m,n}|^2$.

Proof. Choose any orthornormal bases (e_j) and (f_j) of H. Then

$$Ae_i = \sum_{j=0}^{\infty} \langle Ae_i, f_j \rangle f_j,$$

so by the Riesz–Fischer theorem we have

$$\begin{aligned} \sum_{i=0}^{\infty} \|Ae_i\|^2 &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} |\langle Ae_i, f_j \rangle|^2 \\ &= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} |\langle e_i, A^* f_j \rangle|^2 \\ &= \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} |\langle A^* f_j, e_i \rangle|^2 \end{aligned}$$

$$= \sum_{j=0}^{\infty} \| A^* f_j \|^2 ,$$

and we can interchange the rôles of A and A^* to see that the value of the first sum in (13) is independent of the basis chosen. For any $x \in H$, we have

$$Ax = \sum_{j=0}^{\infty} \sigma_j(A) \langle x, v_j \rangle w_j$$

where v_j are the normalised eigenvectors corresponding to the eigenvalues $\sigma_j(A)^2$ of (A^*A). We extend (v_j) to an orthonormal basis of H, by including vectors from the kernel of A, and then

$$\sum_{j=0}^{\infty} \sigma_j(A)^2 = \sum_{j=0}^{\infty} \| Av_i \|^2$$

as required. Now note that since (e_j) is an orthonormal basis, we can write $Ae_j = \sum_{k=0}^{\infty} \langle Ae_j, e_k \rangle e_k$, and hence by the Riesz-Fischer theorem we have

$$\sum_{j=0}^{\infty} \| Ae_j \|^2 = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |\langle Ae_j, e_k \rangle|^2 .$$

■

Lemma 1.8 *If C is trace class, then it can be written as AB for some Hilbert–Schmidt operators A and B.*

Remark

The converse of this result is also true, that is, AB is trace class whenever A and B are Hilbert–Schmidt. See [34, p.11] for the proof.

Proof. If C is trace class, then, as before,

$$Cx = \sum_{j=0}^{\infty} \sigma_j(C) \langle x, v_j \rangle w_j \quad \text{for all } x \in H$$

where $\sum_{j=0}^{\infty} \sigma_j(C) < \infty$. We define

$$Ax = \sum_{j=0}^{\infty} \sigma_j(C)^{1/2} \langle x, v_j \rangle w_j \quad \text{for all } x \in H$$

and

$$Bx = \sum_{j=0}^{\infty} \sigma_j(C)^{1/2} \langle x, v_j \rangle v_j \quad \text{for all } x \in H,$$

where both series are convergent in the norm of H. Then it is clear that A and B are Hilbert–Schmidt, since they have singular values $(\sigma_j(C)^{1/2})_{j=0}^{\infty}$, which is clearly a square-summable sequence because C is trace class. Also, for any $x \in H$, we have

$$\begin{aligned}
ABx &= A\left(\sum_{j=0}^{\infty} \sigma_j(C)^{1/2} \langle x, v_j \rangle v_j\right) \\
&= \sum_{i=0}^{\infty} \sigma_i(C)^{1/2} \left\langle \sum_{j=0}^{\infty} \sigma_j(C)^{1/2} \langle x, v_j \rangle v_j, v_i \right\rangle w_i \\
&= \sum_{i=0}^{\infty} \sigma_i(C)^{1/2} \sum_{j=0}^{\infty} \sigma_j(C)^{1/2} \langle x, v_j \rangle \langle v_j, v_i \rangle w_i \\
&= \sum_{j=0}^{\infty} \sigma_j(C) \langle x, v_j \rangle w_j = Cx
\end{aligned}$$

■

Recall that the *trace* of a matrix is the sum of the entries along its main diagonal. We wish to make a similar definition for operators. The obvious starting point for a definition of trace(A) is to define it as the sum of the diagonal entries of the matrix of A with respect to some orthonormal basis. There are two potential problems with this: (i) we do not know whether our choice of orthonormal basis will affect the value of this "trace", and (ii) the sum will in general be infinite, and there is no obvious guarantee of convergence. Fortunately, it turns out that we can overcome these problems when A is trace class, as the next result shows.

Proposition 1.9 *Let A be a trace class operator. Then, for any orthonormal basis (e_j) of H, the sum*

$$\sum_{j=0}^{\infty} \langle Ae_j, e_j \rangle$$

is convergent, and the value of the sum is independent of the choice of basis.

Proof. Pick an orthonormal basis (e_j). Note that Lemma 1.8 allows us to write $A = BC$, for some Hilbert–Schmidt operators B and C, and so

$$\sum_{j=0}^{\infty} \langle Ae_j, e_j \rangle = \sum_{j=0}^{\infty} \langle BCe_j, e_j \rangle = \langle Ce_j, B^* e_j \rangle. \tag{14}$$

We need to show that this sum is convergent, and that its value is independent of the basis chosen. To this end, we define an inner product on C_2, the space of Hilbert–Schmidt operators by

$$\langle T_1, T_2 \rangle_{C_2} = \sum_{j=0}^{\infty} \langle T_1 e_j, T_2 e_j \rangle.$$

This sum is convergent (and so (14) is), since, by the Cauchy–Schwarz inequality applied twice,

$$\begin{aligned} \sum_{j=0}^{\infty} |\langle T_1 e_j, T_2 e_j \rangle| &\leq \sum_{j=0}^{\infty} \|T_1 e_j\| \, \|T_2 e_j\| \\ &\leq \left(\sum_{j=0}^{\infty} \|T_1 e_j\|^2 \right)^{1/2} \left(\sum_{j=0}^{\infty} \|T_2 e_j\|^2 \right)^{1/2}, \end{aligned}$$

and Proposition 1.7 tells us that the sums on the last line are the Hilbert–Schmidt norms of T_1 and T_2, which are by definition finite. Observe that $\langle , \rangle_{C_2}$ inherits the properties of the ordinary inner product in the sum, so that it is linear in the first argument, conjugate linear in the second argument, and $\langle T, T \rangle_{C_2} \geq 0$ for all T, and hence is a genuine inner product. We still need to show, however, that the value of $\langle T_1, T_2 \rangle_{C_2}$ does not depend on the orthonormal basis chosen. We already know that

$$\langle T, T\rangle_{C_2} = \sum_{j=0}^{\infty}\langle Te_j, Te_j\rangle = \sum_{j=0}^{\infty}\|Te_j\|^2$$

does not depend on the choice of basis (e_j), so by the polarisation identity,

$$4\langle x, y\rangle = \langle x+y, x+y\rangle - \langle x-y, x-y\rangle + i\langle x+iy, x+iy\rangle - i\langle x-iy, x-iy\rangle, \quad (15)$$

valid for any inner product, we see that

$$\langle C, B^*\rangle_{C_2} = \langle Ce_j, B^*e_j\rangle$$

does not depend on the choice of (e_j) either. ■

We can now unambiguously make the following definition.

Definition 1.10 *For a trace class operator A, and any orthonormal basis (e_j), the* trace *is defined by*

$$\text{trace}\,(A) = \sum_{j=0}^{\infty}\langle Ae_j, e_j\rangle.$$

Notice that this definition agrees with the "sum down the main diagonal" value for finite rank operators (which correspond to finite-dimensional matrices). In the matrix definition of trace, the diagonal sum coincides with the sum of the eigenvalues. This is true for trace class operators as well: a proof of the following famous theorem can be found in [24, p.334].

Theorem 1.11 (Lidskii) *Let A be a trace class operator, and $\lambda_j(A)$ be its eigenvalues, counted according to geometric multiplicity. Then*

$$\text{trace}\,(A) = \sum_j \lambda_j(A).$$

Remark

If A is trace class, and in addition positive and self-adjoint, then $\sigma_j(A)$ are just the eigenvalues of A, counted according to geometric multiplicity and

$$\|A\|_{C_1} = \operatorname{trace}(A).$$

Let K be an operator on $L^2[a,b]$ which operates on a function $f \in L^2[a,b]$ as follows:

$$Kf(x) = \int_a^b K(x,y) f(y)\,\mathrm{d}y.$$

Then we say that K is an *integral operator* with *kernel* $K(x,y)$. The adjoint operator K^* is also an integral operator, with kernel $\overline{K(y,x)}$ (see [50, p.77]) so that K is self-adjoint if and only if $K(x,y) = \overline{K(y,x)}$. It can be shown that integral operators with continuous kernel on $[a,b]^2$ are compact. The corollary to the following result, due to Mercer (see [24, p.343] for a proof), will allow us to convert statements about convergence of integrals into trace convergence results for integral operators.

Theorem 1.12 (Mercer) *Let $K(x,y)$ be a real-valued, symmetric, continuous function of x and y, and suppose that the integral operator $K : L^2[0,1] \to L^2[0,1]$ with kernel $K(x,y)$ is non-negative:*

$$\langle Ku, u\rangle \geq 0 \textit{ for all } u \in L^2[0,1].$$

Then $K(x,y)$ can be expanded in a uniformly convergent series

$$K(x,y) = \sum_{j=0}^{\infty} \lambda_j \phi_j(x) \phi_j(y), \tag{16}$$

where λ_j are the eigenvalues, and ϕ_j the normalized eigenfunctions of the operator K.

Corollary 1.13 *Suppose that $K : L^2[0,1] \to L^2[0,1]$ is an integral operator whose* *kernel $K(x,y)$ satisfies the conditions of Mercer's theorem. Then K is trace class, and*

$$\operatorname{trace} K = \int_0^1 K(x,x)\,\mathrm{d}x.$$

Proof. Set $x = y$ in (16) and integrate. Since the sum is uniformly convergent, we can exchange the order of summation and integration to get

$$\begin{aligned}\int_0^1 K(x,x)\,\mathrm{d}x &= \sum_{j=0}^{\infty} \lambda_j \int_0^1 \phi_j(x)^2\,\mathrm{d}x \\ &= \sum_{j=0}^{\infty} \lambda_j \\ &= \operatorname{trace}(K),\end{aligned}$$

by Lidskii's theorem (Theorem 1.11). ■

At this point, we say some words about convergence for operators, which will mostly be relevant in the context of Chapter 2. Let A_n be a sequence of bounded operators on H. Then we identify three types of convergence:

$$\begin{array}{ll} A_n \to A \text{ in norm} & \text{if } \|A_n - A\| \to 0 \\ A_n \to A \text{ strongly} & \text{if } A_n x \to Ax \text{ for all } x \in H \\ A_n \to A \text{ weakly} & \text{if } \langle A_n x, y\rangle \to \langle Ax, y\rangle \text{ for all } x, y \in H \\ & \text{as } n \to \infty. \end{array}$$

The sequence of inequalities

$$|\langle (A - A_n)x, y\rangle| \le \|y\|\,\|Ax - A_n x\| \le \|y\|\,\|x\|\,\|A - A_n\|$$

shows that there is the following hierarchy in these definitions: norm convergence implies strong convergence, which implies weak convergence. Note that, because of the polarisation identity (15), the definition of weak convergence above is equivalent to the condition

$$\langle A_n x, x\rangle \to \langle Ax, x\rangle \text{ for all } x \in H \text{ as } n \to \infty.$$

The following result is a special case of one proved in [39] (see Theorem 2.20 there). We shall use it to get the required trace norm convergence in Theorem 2.7. The notation $|A|$ means $(A^*A)^{1/2}$, so that $|A| = A$ if A is positive.

Theorem 1.14 *Let A_n and A be trace class operators, and suppose that $A_n \to A$, $|A_n| \to |A|$, and $|A_n^*| \to |A^*|$ all weakly, and that $\|A_n\|_{C_1} \to \|A\|_{C_1}$ as $n \to \infty$. Then $\|A_n - A\|_{C_1} \to 0$ as $n \to \infty$.*

1.5 Fredholm determinants

We shall make a definition of operator determinants which agrees with the usual definition when we specialise to finite rank operators (or in other words, finite-dimensional matrices). Our approach follows [39, pp.33-36], and we refer the reader to this source for some of the proofs. To begin with, we need some new notation. Let $H_1, \ldots H_n$ be Hilbert spaces, and let $\hom(H_1, \ldots, H_n)$ be the space of multilinear maps $\ell : H_1 \times \ldots H_n \to \mathbb{C}$. Let $\phi_i \in H_i$. We introduce the notation $\phi_1 \otimes \ldots \otimes \phi_n$ for the multilinear function

$$\phi_1 \otimes \ldots \otimes \phi_n : (\psi_1, \ldots, \psi_n) \mapsto \prod_{i=1}^{n} \langle \phi_i, \psi_i \rangle.$$

Now write $\hom_f(H_1, \ldots, H_n)$ for the algebraic span of the $\phi_1 \otimes \ldots \otimes \phi_n$ in $\hom(H_1, \ldots, H_n)$. We may define an inner product on $\hom_f(H_1, \ldots, H_n)$ which acts on basis elements as follows:

$$\langle \ell, \phi_1 \otimes \ldots \otimes \phi_n \rangle = \ell(\phi_1, \ldots, \phi_n).$$

Using this inner product, we can realise the completion of $\hom_f(H_1, \ldots, H_n)$ as a subset of $\hom(H_1, \ldots, H_n)$: we call this $H_1 \otimes \ldots \otimes H_n$. Given maps $A_i : H_i \to H_i$, there is a map of $\hom(H_1, \ldots, H_n)$ into itself defined by

$$((A_1 \otimes \ldots \otimes A_n)(\ell))(\psi_1, \ldots, \psi_n) = \ell(A_1^*\psi_1, \ldots, A_n^*\psi_n).$$

Note that this map also takes $H_1 \otimes \ldots \otimes H_n$ into itself, and satisfies

$$\begin{aligned}
& (A_1 \otimes \ldots \otimes A_n)(\phi_1 \otimes \ldots \otimes \phi_n)(\psi_1, \ldots, \psi_n) \\
= \; & (\psi_1 \otimes \psi_n)(A_1^*\psi_1, \ldots, A_n^*\psi_n) \\
= \; & \prod_{i=1}^{n} \langle \psi_i, A_i\psi_i \rangle \\
= \; & \prod_{i=1}^{n} \langle A_i\phi_i, \psi_i \rangle \\
= \; & (A_1\phi_1) \otimes \ldots \otimes (A_n\phi_n)(\psi_1, \ldots, \psi_n).
\end{aligned}$$

Given $\psi_1, \ldots, \psi_n \in H$, we define a new object

$$\psi_1 \wedge \ldots \wedge \psi_n = \frac{1}{\sqrt{n!}} \sum_{\pi \in \sigma_n} (-1)^\pi \psi_{\pi(1)} \otimes \ldots \otimes \psi_{\pi(n)},$$

where σ_n is the group of all permutations on $\{1, \ldots, n\}$, and $(-1)^\pi$ is the sign of the permutation π. We write $\Lambda^n H$ for the Hilbert-span of the $\psi_1 \wedge \ldots \wedge \psi_n$, and introduce the notation $\Lambda^k(A)$ for the operator $A \otimes \ldots \otimes A$ on $\Lambda^n H$. We now proceed to define a determinant on the space of trace class operators C_1.

Lemma 1.15 *Let A be a trace class operator on a Hilbert space H. Then $\Lambda^k(A)$ is also trace class on $\Lambda^k H$ with*

$$\left\| \Lambda^k(A) \right\|_{C_1} \leq \frac{1}{k!} \left\| A \right\|_{C_1}^k. \tag{17}$$

In particular, the series

$$\det(I + zA) := \sum_{k=0}^{\infty} z^k \operatorname{trace}\left(\Lambda^k(A)\right)$$

defines an entire function which satisfies the estimate

$$|\det(I + zA)| \leq \exp\left(|z| \, \|A\|_{C_1}\right). \tag{18}$$

Moreover, for any fixed $\epsilon > 0$, we have that

$$|\det(I + zA)| \leq C_\epsilon \exp(\epsilon \, |z|). \tag{19}$$

. *Further, if A is a finite rank operator, then this definition of the determinant agrees with that previously defined for matrices.*

Proof. Let $(\mu_j(A))_{j\geq 0}$ be the singular values of A. Notice that $\left|\Lambda^k(A)\right| = \Lambda^k(|A|)$, so that the singular values of $\Lambda^k(A)$ are $\{\mu_{i_1}(A) \dots \mu_{i_k}\}_{i_1 < \dots < i_k}$. Then

$$\left\|\Lambda^k(A)\right\|_{C_1} = \sum_{i_1 < \dots < i_k} \mu_{i_1}(A) \dots \mu_{i_k}(A).$$

Thus, the inequality (17) is clear. It is then easy to see that (18) is true, since

$$\begin{aligned}
& \left|\sum_{k=0}^{\infty} z^k \operatorname{trace}\left(\Lambda^k(A)\right)\right| \\
\leq \; & \sum_{k=0}^{\infty} |z|^k \left\|\Lambda^k(A)\right\|_{C_1} \\
\leq \; & \sum_{k=0}^{\infty} \frac{|z|^k}{k!} \|A\|_{C_1}^k \\
= \; & \exp\left(|z| \, \|A\|_{C_1}\right).
\end{aligned}$$

The series $\sum_{n=0}^{\infty} \mu_n(A)$ is convergent, so for fixed $\epsilon > 0$, we can pick N so that

$$\sum_{n=N+1}^{\infty} \mu_n(A) \leq \epsilon/2.$$

Also,

$$\begin{aligned}
|\det(I + zA)| &\leq \sum_{k=0}^{\infty} |z|^k \left\|\Lambda^k(A)\right\|_{C_1} \\
&= \sum_{k=0}^{\infty} |z|^k \sum_{i_1 < \ldots < i_k} \mu_{i_1}(A) \ldots \mu_{i_k}(A) \\
&= \prod_{k=1}^{\infty} (1 + |z| \, \mu_k(A)) \\
&= \prod_{k=1}^{N} (1 + |z| \, \mu_k(A)) \prod_{k=N+1}^{\infty} (1 + |z| \, \mu_k(A)) \\
&\leq \prod_{k=1}^{N} (1 + |z| \, \mu_k(A)) \prod_{k=N+1}^{\infty} \exp\left(|z| \, \mu_k(A)\right) \\
&= \prod_{k=1}^{N} (1 + |z| \, \mu_k(A)) \exp(|z| \sum_{k=N+1}^{\infty} \mu_k(A)) \\
&= \prod_{k=1}^{N} (1 + |z| \, \mu_k(A)) \exp(|z| \, \epsilon/2) \\
&\leq C_\epsilon \exp(\epsilon \, |z|),
\end{aligned}$$

for some constant C_ϵ. The final part of the theorem is discussed in [39, p.7]. ■

Theorem 1.16 *The mapping $A \mapsto \det(I + A)$ is a continuous function on C_1, the set of trace class operators.*

Lemma 1.17 *If A is a trace class operator, then there exists a sequence of finite-rank operators A_n such that $\|A_n - A\|_{C_1} \to 0$ as $n \to \infty$.*

Theorem 1.18 *Let A and B be trace class operators. Then:*

(i) $\det(I + A + B + AB) = \det(I + A) \det(I + B)$

(ii) $\det(I + A) \neq 0$ *if and only if* $I + A$ *is invertible.*

(iii) If λ is an eigenvalue of A of multiplicity n, and $z_0 = -1/\lambda$, then $\det(I + zA)$ has a zero of order n at z_0.

Proof. *(i)* By the continuity of the determinant proved in Theorem 1.16, and the convergence result Lemma 1.17, we only need to prove this result for finite-rank operators. But this result is clear for finite rank operators, as it follows from the fact that $\det(CD) = \det(C)\det(D)$. For parts *(ii)* and *(iii)*, we refer the reader to [39, p.34], to avoid verbatim repitition. ■

The following result is due to Hadamard, and appears in Titchmarsh [42, p. 250, §8.24].

Lemma 1.19 *Let $f(z)$ be an entire function with zeros at $z_1, z_2, \ldots$ (counting multiplicity). Suppose that $f(0) = 1$, $\sum_{n=0}^{\infty} |z_n|^{-1} < \infty$, and that for any $\epsilon > 0$*

$$|f(z)| \leq C_\epsilon \exp(\epsilon |z|).$$

Then

$$f(z) = \prod_{n=1}^{\infty} (1 - z_n^{-1} z).$$

The following result is a direct application of Hurwitz's theorem to the determinant function. It essentially asserts that zeros of both sides of (20) match up, and no new zeros arise from the limiting process. For a proof of Hurwitz's theorem, see [42, p.119].

Lemma 1.20 *Suppose that (A_n) is a sequence of finite-rank operators which converge to a trace-class operator A in trace norm as $n \to \infty$. Let Γ be a contour such that no zero of $\det(I + zA)$ lies on Γ, and let $\chi(f; \Gamma)$ be the number of zeros of a function f which lie inside Γ. Then*

$$\chi(\det(I + zA_n); \Gamma) \to \chi(\det(I + zA); \Gamma) \text{ as } n \to \infty.$$

Theorem 1.21 *Let A be a trace class operator, with non-zero eigenvalues $\lambda_k(A)$. Then, for the determinant we defined in Lemma 1.15, we have*

$$\det(I + zA) = \prod_{k=1}^{\infty}(I + z\lambda_k(A)). \tag{20}$$

Proof. Let $f(z) = \det(I + zA)$, and let $(\lambda_n(A))$ be the non-zero eigenvalues of A arranged in descending order. By Theorem 1.18, f has zeros at $z_n = -1/\lambda_n(A)$. By the Lalesco inequality (see [23])

$$\sum_{i=1}^{N} |\lambda_i(A)| \leq \sum_{i=0}^{\infty} |\mu_i(A)|,$$

and hence $\sum_{n=0}^{\infty} |z_n|^{-1} < \infty$. It is clear that $f(0) = 1$, and by (18), $f(z) \leq C_\epsilon \exp(\epsilon |z|)$. Thus we can use Lemma 1.19 to get the required expansion. ∎

Finally in this section, we illustrate the Fredholm determinant of a integral operator, which will be important later on, when we consider random matrix theory. We refer the reader to [39, p.36] for the proof

Theorem 1.22 *Let K be a trace-class integral operator on $L^2(a, b)$ with kernel $K(x, y)$, and K continuous. Then*

$$\det(I - K) = \sum_{n=0}^{\infty} (-1)^n \frac{\alpha_n}{n!},$$

where

$$\alpha_n = \int_a^b \dots \int_a^b \det\left(K(x_i, y_j)\right)_{1 \leq i,j \leq n} \mathrm{d}x_1 \dots \mathrm{d}x_n.$$

1.6 Hankel and Toeplitz operators

Hankel operators will be considered in several contexts in this thesis, mainly because they are an aid to spectral calculation for Tracy–Widom operators. To begin with, we think of them as matrices. If an operator $\Gamma : \ell^2 \to \ell^2$ has matrix of the form

$$\begin{bmatrix} a_0 & a_1 & a_2 & \cdots \\ a_1 & a_2 & a_3 & \ddots \\ a_2 & a_3 & a_4 & \ddots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

or in other words

$$[\Gamma]_{j,k} = a_{j+k}, \quad j, k \geq 0$$

for some sequence $(a_n)_{n=0}^{\infty} \in \ell^2$, then we say that Γ is a *Hankel operator* on ℓ^2, and refer to its matrix $[\Gamma]$ as a *Hankel matrix*. Since the value of each entry depends only on the sum of its indices, another way to state this is to say that $[\Gamma]$ is constant on the diagonals perpendicular to the main diagonal. Following on from this idea, we can use the shift operator S defined in §1.2 to give a characterisation of Hankel operators.

Proposition 1.23 *Γ has a Hankel matrix in the standard basis of ℓ^2 if and only if*

$$\Gamma S = S^* \Gamma. \tag{21}$$

Proof. Let $(e_n)_{n=0}^{\infty}$ be the standard basis for ℓ^2, and suppose that Γ is an operator on ℓ^2. Then

$$\langle \Gamma S e_m, e_n \rangle = \langle \Gamma e_{m+1}, e_n \rangle$$

while

$$\langle S^* \Gamma e_m, e_n \rangle = \langle \Gamma e_n, S e_n \rangle = \langle \Gamma e_m, e_{n+1} \rangle.$$

Hence $\langle \Gamma S e_m, e_n \rangle = \langle S^* \Gamma e_m, e_n \rangle$ if and only if $\langle \Gamma e_{m+1}, e_n \rangle = \langle \Gamma e_m, e_{n+1} \rangle$, or in other words, Γ is a Hankel matrix. ■

The following special property of the kernel of a Hankel operator will be important later on, when we consider operators on the circle in Chapter 4. It is a direct consequence of the shift characterisation we just proved. Recall that the *kernel* of a linear operator on H is the set

$$\mathrm{Ker} A = \{x \in H : Ax = 0\}.$$

Corollary 1.24 *The kernel* $\mathrm{Ker}\,\Gamma_\phi$ *of a Hankel operator* Γ_ϕ *on* $H^2(\mathbb{T})$ *is shift-invariant, i.e.* $S\,\mathrm{Ker}\,\Gamma_\phi$ *is a subspace of* $\mathrm{Ker}\,\Gamma_\phi$*.*

Proof. Take $x \in S\mathrm{Ker}(\Gamma)$. Then by definition $x = Sy$, where $\Gamma y = 0$, so by Proposition 1.23,

$$\Gamma x = \Gamma S y = S^* \Gamma y = 0,$$

and hence $x \in \mathrm{Ker}(\Gamma)$ as required. ■

We would like to define an operator on $H^2(\mathbb{T})$ which has a Hankel matrix, since these arise naturally in problems that we shall study in later parts of the thesis. Furthermore, expressing a Hankel matrix in this way can make spectral calculation much easier, because of results like Theorem 1.28 below. Let $(z^n)_{n=-\infty}^{\infty}$ be the standard orthonormal basis for $L^2(\mathbb{T})$. Let M_ϕ denote the operator of multiplication by a function $\phi \in L^2(\mathbb{T})$. Define a flip operator J on the basis of $L^2(\mathbb{T})$ by $Jz^n = z^{-n}$. Now let $\Gamma_\phi : H^2 \to H^2$ be defined by $\Gamma_\phi = P_+ M_\phi J$. Then the matrix of Γ_ϕ with respect to $(z^n)_{n=-\infty}^{\infty}$ has $(m,n)^{th}$ entry

$$\begin{aligned} \langle P_+ M_\phi J z^m, z^n \rangle &= \langle P_+ \sum_{k=-\infty}^{\infty} \hat{\phi}(k) z^{k-m}, z^n \rangle \\ &= \langle P_+ \sum_{l=-\infty}^{\infty} \hat{\phi}(m+l) z^l, z^n \rangle \\ &= \langle \sum_{l=0}^{\infty} \hat{\phi}(m+l) z^l, z^n \rangle \\ &= \hat{\phi}(m+n) \quad (m, n \in \mathbb{Z}_+), \end{aligned} \tag{22}$$

which shows that Γ_ϕ has a Hankel matrix with respect to the standard basis. We call ϕ a *symbol* of the Hankel operator Γ_ϕ. Notice that ϕ is not unique: Γ_ϕ

determines only the positive Fourier coefficients $(\hat{\phi}(n))_{n=0}^{\infty}$ of a symbol function. Nonetheless, there is in some sense a "best choice" of symbol, which appears in the next result, where the bounded Hankel operators are characterised. The result was first given by Nehari in 1957, and states that a Hankel operator Γ is bounded if and only if it has a bounded symbol. There are many proofs in the literature, see for example [32, p.26], [37, p.3], or for a gentler approach, [26, p.131]. It is worth pointing out at this stage that there is an alternative way of defining a Hankel operator, often found in the literature, in which matrix entries determine instead the negative Fourier coefficients. There is no advantage of one approach over another, but it is important to be aware of which definition is being used when applying results from a particular source.

Theorem 1.25 (Nehari) *A Hankel operator* Γ *with matrix* $[a_{j+k}]_{j,k\geq 0}$ *is bounded if and only if there exists a function* $\phi \in L^{\infty}(\mathbb{T})$ *such that* $\hat{\phi}(n) = a_n$ *for* $n \geq 0$. *Furthermore,*

$$\|\Gamma\| = \inf\{\|\phi\| : \phi \in L^{\infty}(\mathbb{T}) \textit{ with } \hat{\phi}(n) = a_n \textit{ for } n \geq 0\}.$$

We now list some more basic properties of Hankel operators. The adjoint of the Hankel operator Γ_ϕ has matrix

$$\left[\overline{\hat{\phi}(m+n)}\right]_{m,n\geq 0},$$

and a symbol function for this matrix is

$$\sum_{j=-\infty}^{\infty} \overline{\hat{\phi}(j)} z^j = \overline{\phi(\bar{z})},$$

so $\Gamma_\phi^* = \Gamma_{\phi^*}$, where we define $\phi^*(z) = \overline{\phi(\bar{z})}$. Thus, a Hankel operator Γ_ϕ is self-adjoint if and only if $\hat{\phi}(n) \in \mathbb{R}$ for all $n \geq 0$.

Proposition 1.26 *A Hankel operator Γ_ϕ on H^2 is Hilbert-Schmidt if and only if*

$$\sum_{n=1}^{\infty} n|\hat{\phi}(n)|^2 < \infty.$$

Proof. Any Hankel operator has a matrix which is constant on diagonals perpendicular to the main diagonal, so the sum of its squared matrix entries is $\sum_{n=1}^{\infty} n|\hat{\phi}(n)|^2$. By Proposition 1.7, this sum is equal to $\|\Gamma_\phi\|^2$. ∎

There is a simple compactness criterion due to Hartman which we state without proof (see [32, p.214]), and will use later on. We write $C(\mathbb{T})$ for the space of continuous functions on the unit circle.

Theorem 1.27 (Hartman) *Let Γ be a Hankel operator on $H^2(\mathbb{T})$. Then Γ is compact if and only there exists a function $\phi \in C(\mathbb{T})$ such that $\Gamma = \Gamma_\phi$.*

We mentioned earlier that if we define operators on H^2 which have a Hankel matrix with respect to the standard basis, then spectral results about these matrices can be deduced. We now make this idea clear, by presenting a result due to Power [36] on the essential spectrum of a Hankel operator with piecewise continuous symbol. The latter term means that at each point of $\mathbb{T}$, ϕ is right-continuous, and has left and right limits there. We shall need some notation: if ϕ is piecewise-continuous, and $\lambda \in \mathbb{T}$ is a point of discontinuity, then we define the jump (*saltus*) at λ as

$$s_\lambda = \phi(\lambda_+) - \phi(\lambda_-),$$

where $\phi(\lambda_+) = \lim_{\theta \to 0_+} \phi(\lambda e^{i\theta})$ and $\phi(\lambda_-) = \lim_{\theta \to 0_-} \phi(\lambda e^{i\theta})$.

Theorem 1.28 (Power) *Let $\phi \in L^\infty(\mathbb{T})$ be piecewise-continuous. Then*

$$\sigma_{ess}(\Gamma_\phi) = \frac{i}{2}[0, s_1] \cup \frac{i}{2}[0, s_{-1}] \bigcup_{\lambda \in \mathbb{T} \setminus \{\pm 1\}} \frac{i}{2}\left[-(s_\lambda s_{\bar{\lambda}})^{1/2}, (s_\lambda s_{\bar{\lambda}})^{1/2}\right].$$

Thus the essential spectrum of a Hankel operator with piecewise-continuous symbol is a union of intervals in the complex plane, which all pass through or start from the origin. Although this makes calculation of the essential spectrum fairly straightforward in this case, the full spectrum remains more elusive, since it could contain isolated eigenvalues. As an illustration, and because the information will be useful later, we show how Power's result can be used to recover the following famous result on the spectrum of the Hilbert matrix.

Proposition 1.29 *Hilbert's Hankel matrix* $\Gamma = [1/(m+n+1)]_{m,n\geq 0}$ *has* $\sigma_{ess}(\Gamma) = [0, \pi]$.

Remark

In fact, more is true: a famous result of Magnus shows that the full spectrum of Γ is $[0, \pi]$ (see [32, p.287]).

Proof. We have first to find a symbol function ϕ for Γ. The positive Fourier coefficients are already determined by the Hilbert matrix, so we must have $\hat{\phi}(n) = 1/(n+1)$ for $n \geq 0$, and we choose the negative coefficients to mirror these. Thus we have

$$\begin{aligned}
\phi(e^{i\theta}) &= \sum_{n=0}^{\infty} \frac{e^{in\theta}}{n+1} - \sum_{n=2}^{\infty} \frac{e^{-in\theta}}{n-1} \\
&= -e^{-i\theta} \log(1 - e^{i\theta}) + e^{-i\theta} \log(1 - e^{-i\theta}) \\
&= e^{-i\theta} \left(\log \left| \frac{1 - e^{-i\theta}}{1 - e^{i\theta}} \right| + i \arg \left(\frac{1 - e^{-i\theta}}{1 - e^{i\theta}} \right) \right) \\
&= ie^{-i\theta} \arg \left(\frac{1 - e^{-i\theta}}{1 - e^{i\theta}} \right),
\end{aligned}$$

in which we take the logarithm to be defined by the principal value of the argument. An easy calculation shows that

$$\frac{1 - e^{-i\theta}}{1 - e^{i\theta}} = \frac{1 - 2\cos\theta + \cos 2\theta}{2(1 - \cos\theta)} + i \frac{2\sin\theta - \sin 2\theta}{2 - 2\cos\theta},$$

for $\theta \in [0, 2\pi)$, and so

$$\begin{aligned}
\arg\left(\frac{1-e^{-i\theta}}{1-e^{i\theta}}\right) &= \tan^{-1}\left(\frac{2\sin\theta - \sin 2\theta}{1 - 2\cos\theta + \cos 2\theta}\right) \\
&= \tan^{-1}\left(\frac{2\sin\theta\,(1-\cos\theta)}{2\cos\theta(\cos\theta - 1)}\right) \\
&= \tan^{-1}\left(-\frac{\sin\theta}{\cos\theta}\right) \\
&= \pi - \theta.
\end{aligned}$$

Thus

$$\phi(e^{i\theta}) = ie^{-i\theta}(\pi - \theta) \quad \text{for } \theta \in [0, 2\pi).$$

Clearly, the only point of discontinuity of this function on the unit circle is at 1. We have $\phi(1_+) = i\pi$ and $\phi(1_-) = -i\pi$, so $s_1 = 2\pi$. Hence, by Theorem 1.28

$$\sigma_{ess}(\Gamma) = \frac{i}{2}[0, 2\pi] = [0, \pi].$$

■

Another class of operators having a special matrix form are the *Toeplitz operators*. A Toeplitz matrix is characterised by being constant on the diagonals parallel to the main diagonal

$$\begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots \\ a_1 & a_0 & a_{-1} & \ddots \\ a_2 & a_1 & a_0 & \ddots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix};$$

thus it has the form $[a_{j-k}]_{j,k\geq 0}$ for some sequence $(a_n)_{n\geq 0}$. As with the Hankel case, it is useful to define operators on $H^2(\mathbb{T})$ which have a Toeplitz matrix with respect to the standard basis for $H^2(\mathbb{T})$. We set $T_\phi = P_+ M_\phi$, call ϕ the symbol of the Toeplitz operator, and then the $(m, n)^{th}$ element of the matrix of T_ϕ with respect to the standard basis $(z^n)_{n=0}^{\infty}$ is

$$\begin{aligned}
\langle P_+ M_\phi z^m, z^n\rangle &= \langle P_+ \sum_{k=-\infty}^{\infty} \hat{\phi}(k) z^{m+k}, z^n\rangle \\
&= \langle P_+ \sum_{l=-\infty}^{\infty} \hat{\phi}(l-m) z^l, z^n\rangle \\
&= \langle \sum_{l=0}^{\infty} \hat{\phi}(l-m) z^l, z^n\rangle \\
&= \hat{\phi}(n-m) \quad (m,n \in \mathbb{Z}_+).
\end{aligned}$$

There is a result analogous to Nehari's theorem, which states that the bounded Toeplitz operators are those with bounded symbols (see [33, p.312, Theorem 26]). The *essential range* of a function $\phi \in L^\infty(\mathbb{T})$ is defined to be the set

$$\{\lambda : m\left\{e^{i\theta} : \left|\phi(e^{i\theta}) - \lambda\right| < \epsilon\right\} > 0 \text{ for all } \epsilon > 0\}.$$

When the symbol ϕ is real-valued, Hartman and Wintner (see [32, p.248]) proved that

$$\sigma(T_\phi) = \sigma_{ess}(T_\phi) = [\text{ess inf } \phi, \text{ess sup } \phi],$$

where ess inf ϕ and ess sup ϕ are, respectively, the greatest lower bound and least upper bound of the essential range of ϕ. The resemblance between Hankel and Toeplitz operators on H^2 leads to the following useful identity, which can be found in [32, p.253].

Proposition 1.30 *Let $\phi, \psi \in L^\infty(\mathbb{T})$. Then*

$$T_\phi T_\psi - T_{\phi\psi} = -\Gamma_\phi \Gamma_{J\psi} = -\Gamma_\phi \Gamma_{\bar{\psi}}^*. \tag{23}$$

Proof. Following the definition of T_ϕ as above, and using the fact that $P_+ + P_- = I$ and $J^2 = I$, we have

$$T_\phi T_\psi - T_{\phi\psi} = P_+ M_\phi P_+ M_\psi - P_+ M_\phi M_\psi$$

$$
\begin{aligned}
&= P_+M_\phi P_+M_\psi - P_+M_\phi(P_+ + P_-)M_\psi \\
&= -P_+M_\phi P_-M_\psi \\
&= -(P_+M_\phi J)(JP_-M_\psi).
\end{aligned}
$$

Recall that $P_+M_\phi J = \Gamma_\phi$, and observe that

$$
\begin{aligned}
\langle JP_-M_\psi z^m, z^n\rangle &= \langle JP_- \sum_{j=-\infty}^{\infty} \hat{\psi}(j)z^{m+j}, z^n\rangle \\
&= \langle JP_- \sum_{l=\infty}^{\infty} \hat{\psi}(l-m)z^l, z^n\rangle \\
&= \langle J \sum_{l=-\infty}^{-1} \hat{\psi}(l-m)z^l, z^n\rangle \\
&= \langle J \sum_{l=1}^{\infty} \hat{\psi}(-l-m)z^{-l}, z^n\rangle \\
&= \langle \sum_{l=1}^{\infty} \hat{\psi}(-l-m)z^l, z^n\rangle \\
&= \hat{\psi}(-n-m).
\end{aligned}
$$

Now, since

$$
J\psi(z) = \psi(\bar{z}) = \sum_{k=-\infty}^{\infty} \hat{\psi}(k)z^{-k} = \sum_{k=-\infty}^{\infty} \hat{\psi}(-k)z^k
$$

and also

$$
\overline{\psi}(z) = \sum_{j=-\infty}^{\infty} \overline{\hat{\psi}(j)}z^{-j} = \sum_{j=-\infty}^{\infty} \overline{\hat{\psi}(-j)}z^j,
$$

it is then clear that

$$
JP_-M\psi = \Gamma_{J\psi} = \Gamma^*_{\bar{\psi}}
$$

and hence we have the result. ■

1.7 Spectral multiplicity

The following version of the spectral theorem is proved in [13, p.47].

Theorem 1.31 *Let N be a normal operator on a seperable Hilbert space, with spectrum X. Then N is unitarily equivalent to a multiplication operator M_f on some $L^2(\mu)$ space, where the measure μ is a positive Radon measure defined on the space $X_\infty = X \times \mathbb{N}$ of countably many distinct copies of X.*

The measure μ in the above theorem is called the *scalar spectral measure.*

Definition 1.32 *The spectral multiplicity measure ν_W of a normal operator W is given by $\nu_W(A) = \mu(A \times N)$ for all Borel subsets A of X. In particular, for a compact and self-adjoint operator C, the spectrum consists of 0 together with a sequence of real eigenvalues λ and $\nu_C(\{\lambda\}) = \dim E_\lambda$, where E_λ is the eigenspace that corresponds to the eigenvalue λ.*

As a further motivation for the results of Chapters 4 and 5, we see how expressing an operator as a Hankel square can yield information about its spectral multiplicity function. Megretskiĭ, Peller and Treil [30] have proved the following important result about the operators unitarily equivalent to a Hankel operator. We state it, and then give some consequences in our context.

Theorem 1.33 *Let Γ be a bounded and self-adjoint operator on a Hilbert space H with a scalar spectral measure μ and spectral multiplicity function ν. Then Γ is unitarily equivalent to a Hankel operator if and only if the following conditions are satisfied:*

(i) Either $\dim \operatorname{Ker} \Gamma = 0$ *or* $\dim \operatorname{Ker} \Gamma = \infty$*;*

(ii) Γ *is non-invertible;*

(iii) $|\nu(\lambda) - \nu(-\lambda)| \leq 2 \quad \mu_a$*-a.e. and* $|\nu(\lambda) - \nu(-\lambda)| \leq 1 \quad \mu_s$*-a.e., where* μ_s *and* μ_a *are the singular and absolutely continuous components of* μ*.*

The following elementary fact about the kernel of an operator is trivial, but will be useful in the proof below, and in one of the results of Chapter 4:

Lemma 1.34 *Let A be a bounded linear operator on H. Then*

$$\operatorname{Ker}(A^*A) = \operatorname{Ker}(A).$$

Proof. It is immediate that $\operatorname{Ker} A \subset \operatorname{Ker}(A^*A)$, so suppose that $x \in \operatorname{Ker}(A^*A)$. Then $A^*Ax = 0$, so

$$0 = \langle A^*Ax, x\rangle = \langle Ax, Ax\rangle,$$

and hence $Ax = 0$, i.e. $x \in \operatorname{Ker} A$. ■

Proposition 1.35 *Suppose that an operator K satisfies $K = \Gamma^2$, where Γ is a compact self-adjoint Hankel operator. Then*

(i) $\nu_K(0) = 0$ *or* $\nu_K(0) = \infty$.

(ii) $\nu_K(\lambda) < \infty$ *and* $\nu_K(\lambda) = \nu_\Gamma(\sqrt{\lambda}) + \nu_\Gamma(-\sqrt{\lambda})$ *for all* $\lambda > 0$.

(iii) If $\nu_K(\lambda)$ *is even, then* $\nu_\Gamma(\sqrt{\lambda}) = \nu_\Gamma(-\sqrt{\lambda})$.

(iv) If $\nu_K(\lambda)$ *is odd, then* $\left|\nu_\Gamma(\sqrt{\lambda}) - \nu_\Gamma(-\sqrt{\lambda})\right| = 1$

Proof. *(i)* Note that $\nu_K(0)$ is the dimension of $\operatorname{Ker} K$. Since Γ is self-adjoint, by Lemma 1.34 we have

$$\operatorname{Ker}(K) = \operatorname{Ker}(\Gamma^2) = \operatorname{Ker}(\Gamma),$$

where $\operatorname{Ker}(\Gamma) = \{0\}$ or $\Theta H^2(\mathbb{T})$, for some inner Θ, by Beurling's theorem, since the kernel of any Hankel operator is a shift-invariant subspace of $H^2(\mathbb{T})$ (Corollary 1.24).

(ii) K is compact because Γ is, so the spectrum of K consists of eigenvalues of finite multiplicity, i.e. $\nu_K(\lambda) < \infty$ for all λ. Note that if $\lambda > 0$ is an eigenvalue of K, then

$$E_\lambda(K) = E_\lambda(\Gamma^2) = \{x \in H : \ \Gamma x = \sqrt{\lambda}x \text{ or } \Gamma x = -\sqrt{\lambda}x\},$$

which gives the required statement on spectral multiplicity.

(iii) If $\nu_K(\lambda)$ is even, then by *(ii)*, $\nu_\Gamma(\sqrt{\lambda})+\nu_\Gamma(-\sqrt{\lambda})$ is even, and hence $\nu_\Gamma(\sqrt{\lambda})-\nu_\Gamma(-\sqrt{\lambda})$ is even. Since Theorem 1.33 tells us that $|\nu_\Gamma(\sqrt{\lambda})-\nu_\Gamma(-\sqrt{\lambda})| \leq 1$ (Γ is compact, so there is no absolutely continuous component to the spectrum), we therefore have $\nu_\Gamma(\sqrt{\lambda})=\nu_\Gamma(-\sqrt{\lambda})$.

(*iv*) This follows by a very similar argument to *(iii)*. ∎

1.8 Orthogonal polynomials

Let w be a non-negative function defined on a (possibly infinite) interval $[a,b]$. Assuming all the moments $\int_a^b w(x)x^n\,\mathrm{d}x$ of w exist, we can construct, via the Gram–Schmidt process applied to the set $\{1,x,x^2,\ldots\}$, a sequence of orthonormal polynomials $p_n(x)$ such that

$$\int_a^b p_j(x)p_k(x)w(x)\,\mathrm{d}x = \delta_{j,k}, \tag{24}$$

where $\delta_{j,k}$ is the Kronecker delta function as defined in (7), and the leading coefficient of p_n is positive. Note that the assumptions we have so far made about the weight w do not imply that the polynomials (p_n) form an orthonormal basis for $L^2[a,b]$. To get this, we need an extra condition. The following result is proved in [29, p.333].

Proposition 1.36 *Let $w(x)$ be a non-negative weight function on a (possibly infinite) interval $[a,b]$. Suppose that*

$$\int_a^b e^{r|x|}w(x)\,\mathrm{d}x < \infty.$$

Then all the moments of w exist, and further the sequence $(p_n(x))$ of polynomials *orthonormal on $[a,b]$ with respect to w which arises from the Gram-Schmidt process has the property that if any function $f \in L^2[a,b]$ is orthogonal to all the p_n, then $f = 0$ almost everywhere, i.e. $(p_n(x))$ is complete in $L^2[a,b]$.*

An important property of all orthogonal polynomial sequences is that they satisfy a recurrence relation: we shall exploit this later to simplify expressions. This property is well-known, but we include a proof for completeness.

Lemma 1.37 *Let $w(x)$ be a non-negative weight function on an interval $[a,b]$ such that all the moments $\int_a^b x^n w(x)\,\mathrm{d}x$ exist, and let $(p_n(x))$ be the sequence of polynomials arising from the Gram–Schmidt process which are orthonormal with respect to w on the interval $[a,b]$. Then there is a three-term recurrence relation*

$$xp_{n-1}(x) = A_n p_{n-2}(x) + B_n p_{n-1}(x) + C_n p_n(x) \quad (n = 1, 2, 3, \ldots), \tag{25}$$

in which

$$A_n = \int_a^b xp_{n-1}(x)p_{n-2}(x)w(x)\,\mathrm{d}x,$$

$$B_n = \int_a^b xp_{n-1}(x)^2 w(x)\,\mathrm{d}x,$$

$$C_n = \int_a^b xp_n(x)p_{n-1}(x)w(x)\,\mathrm{d}x,$$

and we define $p_{-1} = 0$.

Remark

Notice that $C_n = A_{n+1}$, simply from the definition of the integrals.

Proof. The polynomial $xp_{n-1}(x)$ has degree n, so that

$$xp_{n-1}(x) = \sum_{k=0}^{n} c_k p_k(x),$$

for some constants c_k. We use the fact that (p_k) is an orthonormal sequence to show that

$$c_k = \langle xp_{n-1}(x), p_k(x) \rangle = \int_a^b xp_{n-1}(x)p_k(x)w(x)\,\mathrm{d}x. \tag{26}$$

In fact, since $xp_k(x)$ is a polynomial of degree $k+1$, and p_{n-1} has degree $n-1$, $c_k = 0$ for $k = 0, \ldots, n-3$, and so we have the required expression for $xp_{n-1}(x)$, where the values of A_n, B_n and C_n follow from (26). ■

Remark

The three-term recurrence relation can also be stated in the following matrix form

$$\begin{bmatrix} p_n(x) \\ p_{n+1}(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{A_{n+1}}{C_{n+1}} & \frac{x-B_{n+1}}{C_{n+1}} \end{bmatrix} \begin{bmatrix} p_{n-1}(x) \\ p_n(x) \end{bmatrix},$$

in which the entries of the one step transition matrix are rational functions (in fact, linear polynomials) in x. Clearly, if $A_{n+1} = C_{n+1}$, then this matrix has determinant 1, and we are in the territory of the operators considered in Chapter 5. Since the form of the matrix system is similar to the continuous Tracy–Widom matrix systems considered in [44] and here in §1.11, writing the three-term recurrence relation in this way thus helps to unify the discrete and continuous cases.

The following formula is well-known in the theory of orthogonal polynomials (see, for example [41, p.43]). We include the easy proof for completeness, and also add in a simple consequence for an important operator which appears in the determinantal expressions for random matrix eigenvalue distributions in §1.9 and §1.10.

Proposition 1.38 (Christoffel-Darboux formula) *Let a weight w be as above, and let $(p_n(x))$ be the sequence of polynomials arising from the Gram–Schmidt process which are orthonormal with respect to w on the interval $[a, b]$. Then we have*

$$\sum_{k=0}^{n-1} p_k(x)p_k(y) = C_n \frac{p_n(x)p_{n-1}(y) - p_{n-1}(x)p_n(y)}{x-y}, \tag{27}$$

where the value of C_n comes from the three-term recurrence relation (25). Moreover, the integral operator K_n on $L^2(w,[a,b])$ with kernel

$$K_n(x,y) = \sum_{k=0}^{n-1} p_k(x)p_k(y)$$

satisfies $K_n^2 = K_n$ and $K_n^* = K_n$.

Proof. We use the three-term recurrence relation (25) to rewrite the numerator on the right hand side. We find

$$\begin{aligned}
& C_n\left(p_n(x)p_{n-1}(y) - p_{n-1}(x)p_n(y)\right) \\
= \;& p_{n-1}(y)\left(xp_{n-1}(x) - A_n p_{n-2}(x) - B_n p_{n-1}(x)\right) \\
& -p_{n-1}(x)\left(yp_{n-1}(y) - A_n p_{n-2}(y) - B_n p_{n-1}(y)\right) \\
= \;& (x-y)p_{n-1}(x)p_{n-1}(y) + A_n\left(p_{n-1}(x)p_{n-2}(y) - p_{n-1}(y)p_{n-2}(x)\right) \\
= \;& (x-y)p_{n-1}(x)p_{n-1}(y) + C_{n-1}\left(p_{n-1}(x)p_{n-2}(y) - p_{n-1}(y)p_{n-2}(x)\right),
\end{aligned}$$

where the last line follows from the remark after Lemma 1.37, and then we repeat the argument a further $n-3$ times to get the required formula (and recall the convention $p_{-1}(x) = 0$). Now note that K_n is the projection onto the subspace

$$\overline{\text{span}}\left\{p_k(x) : k = 0, \ldots, n-1\right\}$$

so it satisfies $K_n^2 = K_n$ and $K_n^* = K_n$ as an operator on $L^2(w,[a,b])$. ∎

Corollary 1.39 *For $[c,d] \subseteq [a,b]$, the operator $\mathbb{I}_{[c,d]}K\mathbb{I}_{[c,d]}$ that has kernel*

$$C_n\mathbb{I}_{[c,d]}(x)\frac{p_n(x)p_{n-1}(y) - p_{n-1}(x)p_n(y)}{x-y}\mathbb{I}_{[c,d]}(y)$$

satisfies

$$0 \leq \mathbb{I}_{[c,d]}K\mathbb{I}_{[c,d]} \leq I$$

and is of trace class.

1.9 The Gaussian Unitary Ensemble

Let $x_{j,k}$ and $y_{j,k}$ for $1 \leq j < k \leq N$ be $N(0, 1/2)$ random variables and $x_{j,j}$ for $1 \leq j \leq N$ be $N(0, 1/4)$ random variables, where $x_{j,j}$, $x_{j,k}$ and $y_{j,k}$ are all mutually independent. Now define a space of $N \times N$ Hermitian (self-adjoint) matrices Z with entries

$$Z_{j,k} = \begin{cases} (x_{j,k} + iy_{j,k}) & \text{for } j < k \\ x_{j,j} & \text{for } j = k \\ (x_{j,k} - iy_{j,k}) & \text{for } k < j. \end{cases}$$

The p.d.f. for a $N(0, \sigma^2)$ distribution is

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-x^2/(2\sigma^2)\right),$$

and hence the joint p.d.f. of the matrix elements is

$$\begin{aligned} &\prod_{1\leq j<k\leq N} \left(\frac{1}{\sqrt{\pi/2}}\right)^2 \exp\left(-2(x_{j,k}^2 + y_{j,k}^2)\right) \prod_{j=1}^{N} \frac{1}{\sqrt{\pi}} \exp\left(-x_{j,j}^2\right) \\ = \; &\frac{2^{N(N-1)/2}}{\pi^{N^2/2}} \exp\left(-2 \sum_{1\leq j<k\leq N} (x_{j,k}^2 + y_{j,k}^2) - \sum_{1\leq j\leq N} x_{j,j}^2\right). \end{aligned}$$

It is easy to see that

$$\operatorname{trace}(Z^* Z) = \sum_{1\leq j\leq N} x_{j,j}^2 + 2 \sum_{1\leq j<k\leq N} (x_{j,k}^2 + y_{j,k}^2),$$

so we define a probability measure ν at the level of matrices by

$$\nu(\,\mathrm{d}Z) = \frac{2^{N(N-1)/2}}{\pi^{N^2/2}} \exp(-\operatorname{trace}(Z^* Z))\,\mathrm{d}Z,$$

where

$$\mathrm{d}Z = \prod_{1\le j\le k\le N} \mathrm{d}x_{j,k} \prod_{1\le j<k\le N} \mathrm{d}y_{j,k}.$$

This space of matrices, with the given probability measure, is known as the Gaussian Unitary Ensemble (GUE). The probability measure is invariant under unitary transformations $Z \mapsto UZU^*$, where U is an $N \times N$ unitary matrix, in the sense that, for any continuous and bounded function f,

$$\int f(UZU^*)\nu(\mathrm{d}Z) = \int f(Z)\nu(\mathrm{d}Z).$$

To define instead a probability measure on eigenvalues of Z, we need the Jacobian which arises from the change of variables from the space of Hermitian matrices to the simplex of ordered eigenvalues. This is derived in [27, p.62], and the result is that the joint p.d.f. for the N eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ is

$$p_N(\lambda_1, \lambda_2, \ldots, \lambda_N) = \frac{1}{Z_N} \exp\left(-\sum_{j=1}^{N} \lambda_j^2\right) \prod_{1\le j<k\le N} (\lambda_k - \lambda_j)^2, \tag{28}$$

in which Z_N is a constant which ensures that the p.d.f. integrates to 1. Observe that the exponential term arises from the general fact that $\operatorname{trace}(Z^*Z) = \sum_{j=1}^{N} \lambda_j^2$. To prove the next result, we need the following "integrating out" result. For a proof, see [27, Theorem 5.2.1].

Lemma 1.40 *Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, where the x_j all lie in a (possibly infinite) interval $[a, b]$, and let $A_n(x)$ be an $n \times n$ matrix with entries $A_n(\mathbf{x})_{i,j} = K(x_i, x_j)$, where $K \in L^2[a, b] \times L^2[a, b]$ is a real-valued, symmetric, and continuous function which satisfies the "reproducing kernel" property*

$$\int_a^b K(x, y)K(y, z)\,\mathrm{d}y = K(x, z),$$

and is also non-negative:

$$\int_a^b \int_a^b K(x,y) f(y) f(x)\,\mathrm{d}y\,\mathrm{d}x \geq 0 \quad \textit{for all } f \in L^2[a,b].$$

Then

$$\int \det\left(A_n\left(x_1, x_2, \ldots, x_n\right)\right) \mathrm{d}x_n = \left(q - (n-1)\right) \det\left(A_{n-1}(x_1, x_2, \ldots, x_{n-1})\right),$$

where

$$q = \int_a^b K(x,x)\,\mathrm{d}x.$$

Remark

The conditions that $K(x,y)$ be real-valued, symmetric, continuous and positive are as in Mercer's theorem (see Theorem 1.12), and ensure that the number q exists. The result below is due to Mehta and Gaudin [28], and is discussed by the former in [27, p.91].

Proposition 1.41 *The joint p.d.f. of the eigenvalues for the Gaussian Unitary Ensemble is given by the determinantal formula*

$$p_N(\lambda_1, \lambda_2, \ldots, \lambda_N) = \frac{1}{N!} \det\left(K_N(\lambda_j, \lambda_k)\right)_{j,k=1}^N, \tag{29}$$

in which

$$K_N(x,y) = \sum_{j=0}^{N-1} \phi_j(x)\phi_j(y) \tag{30}$$

and $\phi_j(x)$ are the normalised Hermite polynomials $\phi_j(x) = \pi^{-1/4}\exp(-x^2/2)h_n(x)$, orthonormal on $\mathbb{R}$, in the sense that

$$\int_{-\infty}^{\infty} \phi_j(x)\phi_k(x)\,\mathrm{d}x = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} h_j(x)h_k(x)\exp(-x^2)\,\mathrm{d}x = \delta_{j,k}.$$

Further, the normalisation constant in (28) is

$$Z_N = \frac{\pi^{N/2}\prod_{j=0}^{N} j!}{2^{N(N-1)/2}}.$$

Proof. First, note that the product term in (28) is the famous van der Monde determinant

$$\prod_{1\le j<k\le N}(\lambda_k-\lambda_j) = \det\begin{bmatrix} 1 & \dots & 1 \\ \lambda_1 & \dots & \lambda_N \\ \vdots & \vdots & \vdots \\ \lambda_1^{N-1} & \dots & \lambda_N^{N-1}\end{bmatrix},$$

so we define the matrix

$$A := \left[\lambda_k^{j-1}\frac{e^{-\lambda_k^2/2}}{\pi^{1/4}}\right]_{j,k=1}^{N},$$

which satisfies

$$\det A = \pi^{-N/4}\exp\left(-1/2\sum_{j=1}^{N}\lambda_j^2\right)\prod_{1\le j<k\le N}(\lambda_k-\lambda_j). \tag{31}$$

This can be seen by simply multiplying the columns $k=1,\dots,N$ of the van der Monde determinant by exponential factors $\exp(-\lambda_k^2/2)$. Now let

$$\phi_j(x) = \frac{1}{\pi^{1/4}}\exp\left(-x^2/2\right)h_j(x),$$

where $h_j(x) = a_j x^j + \dots$ is the Hermite polynomial of degree j, with coefficients chosen so that $(\phi_j(x))$ is an orthonormal sequence in $L^2(\mathbb{R})$, and define

$$B = [\phi_{j-1}(\lambda_k)]_{j,k=1}^{N}.$$

By linearly combining rows in (31), we have

$$a_0\dots a_{N-1}\det A = \det B,$$

and hence

$$\begin{aligned}(\det B)^2 &= (a_0 \ldots a_{N-1})^2 (\det A)^2 \\ &= \pi^{-N/2} (a_0 \ldots a_{N-1})^2 \exp\left(-\sum_{j=1}^{N} \lambda_j^2\right) \prod_{1 \le j < k \le N} (\lambda_k - \lambda_j)^2.\end{aligned}$$

Recall that $(\det X)^2 = \det(X^t)\det(X) = \det(X^t X)$ for any matrix X, so we can write

$$\begin{aligned}(\det B)^2 &= \det(B^t B) \\ &= \det\left[\sum_{i=1}^{N} B^t_{ji} B_{ik}\right]_{j,k=1}^{N} \\ &= \det\left[\sum_{i=1}^{N} \phi_{i-1}(\lambda_j)\phi_{i-1}(\lambda_k)\right]_{j,k=1}^{N} \\ &= \det\left[K_N(\lambda_j, \lambda_k)\right]_{j,k=1}^{N},\end{aligned}$$

where $K_N(x, y)$ is as defined in (30). To summarise our findings so far:

$$\det\left[K_N(\lambda_j, \lambda_k)\right]_{j,k=1}^{N} = \pi^{-N/2}(a_0 \ldots a_{N-1})^2 \exp\left(-\sum_{j=1}^{N} \lambda_j^2\right) \prod_{1 \le j < k \le N} (\lambda_k - \lambda_j)^2, \tag{32}$$

in which the right hand side is the joint p.d.f. for the eigenvalues in GUE as in (28). The kernel $K_N(x, y)$ satisfies the conditions of Lemma 1.40, since

$$\begin{aligned}\int_{-\infty}^{\infty} K_N(x,y)K_N(y,z)\,\mathrm{d}y &= \int_{-\infty}^{\infty} \sum_{k=0}^{N-1}\sum_{j=0}^{N-1} \phi_k(x)\phi_k(y)\phi_j(y)\phi_j(z)\,\mathrm{d}y \\ &= \sum_{k=0}^{N-1}\sum_{j=0}^{N-1} \phi_k(x)\phi_j(z) \int_{-\infty}^{\infty} \phi_k(y)\phi_j(y)\,\mathrm{d}y \\ &= \sum_{k=0}^{N-1} \phi_k(x)\phi_k(z) \\ &= K_N(x, z),\end{aligned}$$

and we can calculate

$$\int_{-\infty}^{\infty} K_N(x,x)\,\mathrm{d}x = \sum_{j=0}^{N-1}\int_{-\infty}^{\infty} \phi_j(x)\phi_j(x)\,\mathrm{d}x = N.$$

Hence we can integrate the left hand side of (32) over the variable λ_1 to get

$$\begin{aligned}\int_{-\infty}^{\infty} \det\left[K_N(\lambda_j,\lambda_k)\right]_{j,k=1}^{N}\,\mathrm{d}\lambda_1 &= (N-(N-1))\det\left[K_N(\lambda_j,\lambda_k)\right]_{j,k=1}^{N-1}\\ &= \det\left[K_N(\lambda_j,\lambda_k)\right]_{j,k=1}^{N-1},\end{aligned}$$

and then again over λ_2 to get

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \det\left[K_N(\lambda_j,\lambda_k)\right]_{j,k=1}^{N}\lambda_1\,\mathrm{d}\lambda_2 = (N-(N-2))\det\left[K_N(\lambda_j,\lambda_k)\right]_{j,k=1}^{N-2}.$$

Continuing in this way, we can integrate over $\lambda_1,\ldots,\lambda_N$, and we obtain

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \det\left[K_N(\lambda_j,\lambda_k)\right]_{j,k=1}^{N}\,\mathrm{d}\lambda_1\ldots\mathrm{d}\lambda_N = N!.$$

Thus, if we divide (32) by $N!$, we get an expression which integrates to 1, and so we have the formula (29). We can now see that the normalisation constant in (28) must satisfy

$$Z_N^{-1} = \pi^{-N/2}(N!)^{-1}(a_0a_1\ldots a_{N-1})^2,$$

so our task is to find a_n, the coefficient of the highest power of x in the n^{th} Hermite polynomial $h_n(x) = a_nx^n + \ldots$. We have the Rodrigues's formula (see [41, p.105])

$$h_n(x) = c_ne^{x^2}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(e^{-x^2}\right).$$

It is clear from this $a_n = c_n(-2)^n$. By the orthonormality condition on $h_n(x)$, and integrating by parts n times, we have

$$1 = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} h_n(x)^2e^{-x^2}\,\mathrm{d}x$$

$$
\begin{aligned}
&= \frac{c_n}{\sqrt{\pi}} \int_{-\infty}^{\infty} h_n(x) \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left(e^{-x^2}\right) \mathrm{d}x \\
&= \frac{c_n}{\sqrt{\pi}} (-1)^n \int_{-\infty}^{\infty} e^{-x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \left(h_n(x)\right) \mathrm{d}x \\
&= \frac{c_n}{\sqrt{\pi}} (-1)^n \int_{-\infty}^{\infty} e^{-x^2} (n! a_n) \, \mathrm{d}x \\
&= (-1)^n c_n a_n n! \\
&= \frac{(-1)^n a_n^2 n!}{(-2)^n}
\end{aligned}
$$

and hence

$$a_n = \frac{2^{n/2}}{\sqrt{n!}}.$$

Thus

$$
\begin{aligned}
Z_N &= \frac{\pi^{N/2} 0! 1! \ldots N!}{2^0 2^1 \ldots 2^{N-1}} \\
&= \frac{\pi^{N/2} \prod_{j=0}^{N} j!}{2^{N(N-1)/2}}.
\end{aligned}
$$

■

The n-point correlation function is given by the formula

$$R_n(\lambda_1, \lambda_2, \ldots, \lambda_n) = \frac{N!}{(N-n)!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_N(\lambda_1, \lambda_2, \ldots, \lambda_N) \, \mathrm{d}\lambda_{n+1} \, \mathrm{d}\lambda_{n+2} \ldots \mathrm{d}\lambda_N,$$

in which the constant factor is the number of ways of arranging n items selected from N. Ignoring the fact that it integrates to $N!/(N-n)!$ rather than 1, and so cannot be considered as a proper p.d.f., $R_n(\lambda_1, \ldots, \lambda_n)$ can be seen as the probability density that we find eigenvalues at $\lambda_1, \ldots, \lambda_n$, with the remaining $N-n$ eigenvalues unobserved. In particular, the mean eigenvalue density around the point z is given by $R_1(z)$. It is a very convenient fact that R_n can, like p_N, be expressed in terms of the kernel $K_N(x, y)$. As in the proof of Proposition 1.41, we can integrate the determinantal expression in (29) over the variables $\lambda_{n+1}, \ldots, \lambda_N$

to get

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \det\left[K_N(\lambda_j, \lambda_k)\right]_{j,k=1}^{N} \mathrm{d}\lambda_{n+1} \ldots \mathrm{d}\lambda_N = (N-n)!,$$

and hence

$$\begin{aligned}
R_n(\lambda_1, \ldots, \lambda_n) &= \frac{N!}{(N-n)!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p_N(\lambda_1, \ldots, \lambda_N)\, \mathrm{d}\lambda_{n+1}\, \mathrm{d}\lambda_{n+2} \ldots \mathrm{d}\lambda_N \\
&= \frac{N!}{(N-n)!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{1}{N!} \det\left(K_N(\lambda_j, \lambda_k)\right]_{j,k=1}^{N} \mathrm{d}\lambda_{n+1}\, \mathrm{d}\lambda_{n+2} \ldots \mathrm{d}\lambda_N \\
&= \frac{1}{(N-n)!} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \det\left[K_N(\lambda_j, \lambda_k)\right]_{j,k=1}^{N} \mathrm{d}\lambda_{n+1} \ldots \mathrm{d}\lambda_N \\
&= \det\left[K_N(\lambda_j, \lambda_k)\right]_{j,k=1}^{n}.
\end{aligned}$$

What happens to the eigenvalue distribution when we let $N \to \infty$? The answer is not as simple as it might appear, since it turns out that a scaling operation is needed to keep all the quantities finite. This is described well by Tracy and Widom, in [43]. Following these authors' approach, we choose a point z in the spectrum, make this the new origin, and scale the eigenvalues so that the eigenvalue density at this point is equal to 1 in the limit. Scaling eigenvalues in this manner corresponds to replacing the kernel $K_N(x, y)$ by

$$\frac{1}{R_1(z)} K_N\left(z + \frac{x}{R_1(z)}, z + \frac{y}{R_1(z)}\right).$$

The asymptotic eigenvalue density can be found, since $R_1(z) = K_N(z, z)$, and it can be shown that

$$R_1(z) = K_N(z, z) \sim \frac{1}{\pi}\sqrt{2N} \text{ as } N \to \infty$$

If we choose the scaling

$$\lambda_j \mapsto z + \frac{\pi \lambda_j}{\sqrt{2N}},$$

and consider the asymptotics of the Hermite polynomials under this scaling, it turns out that the kernel $K_N(x,y)$ converges to what is known as the sine kernel:

$$\lim_{N\to\infty} \frac{\pi}{\sqrt{2N}} K_N\left(z + \frac{\pi x}{\sqrt{2N}}, z + \frac{\pi y}{\sqrt{2N}}\right) = \frac{\sin(x-y)}{\pi(x-y)}.$$

Thus the eigenvalues in the bulk of the spectrum can be described asymptotically by the sine kernel. This situation is called "bulk" scaling, because it describes the distribution of eigenvalues in the middle of the spectrum. The edge of the spectrum is located at approximately $\sqrt{2N}$, and substituting into $K_N(x,x)$, it can be shown that the eigenvalue density here is asymptotically $\sqrt{2}N^{1/6}$. Using some more asymptotic formulae for the Hermite polynomials, the "soft edge" limit

$$\lim_{N\to\infty} \frac{1}{\sqrt{2}N^{1/6}} K_N\left(\sqrt{2N} + \frac{x}{\sqrt{2}N^{1/6}}, \sqrt{2N} + \frac{y}{\sqrt{2}N^{1/6}}\right) = \frac{\operatorname{Ai}(x)\operatorname{Ai}'(y) - \operatorname{Ai}(y)\operatorname{Ai}'(x)}{x-y},$$

can be obtained, and so the eigenvalue distribution at the soft edge of the spectrum is described by the Airy kernel. The Airy function $\operatorname{Ai}(x)$ is defined in §1.11 below.

1.10 Random Matrix Theory

Random matrix theory is the study of eigenvalue distributions for matrices chosen according to some probability measure on the matrix space in question. We refer to the space of matrices with an associated probability measure as a *Random Matrix Ensemble*, and the protoptypical example is the Gaussian Unitary Ensemble discussed in the previous §1.9. A variety of ensembles are studied in the literature, but in this thesis, we will be dealing with $N \times N$ Hermitian or symmetric matrices, with a probability measure which is invariant under a certain class of matrix transformations. The main three types of invariance are orthogonal, unitary and symplectic, named after the type of matrix performing the transformation. Once we have defined the space of matrices, and the probability measure, it is possible to find the probability distribution of the eigenvalues. The calculations are fairly involved, but the interested reader can consult the famous book by Mehta [27, Chapter 3] for details of how to do this for Gaussian ensembles. Since the matrices are Hermitian, the eigenvalues are all found in a (possibly infinite) interval $[a, b]$ of the real line. It turns out that the probability density for the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ in a random matrix ensemble is

$$p_{N,\beta}(\lambda_1, \lambda_2, \ldots, \lambda_N) = \frac{1}{Z_{N,\beta}} \prod_{j=1}^{N} w(\lambda_j) \prod_{1 \le j < k \le N} |\lambda_j - \lambda_k|^\beta \, \mathrm{d}\lambda_1 \ldots \lambda_N, \tag{33}$$

(this is covered briefly in [27, §19.3]) where w is a non-negative weight function, integrable on $[a, b]$, $Z_{N,\beta}$ is a normalisation constant which ensures that the right-hand side integrates to 1, and β is a parameter which reflects the invariance properties of the underlying probability measure on matrices. The weight for the GUE is $w(x) = \pi^{-1/2} \exp(-x^2)$. The values $\beta = 1$, 2, and 4 correspond, respectively, to orthogonal, unitary, and symplectic ensembles. Notice that the last product term in (33) reflects the fact that the eigenvalues are not independent. In the case of unitary ensembles ($\beta = 2$), with weight $w(x)$, it is possible to give a determinantal form for the eigenvalue probability distribution (33), which makes it possible to calculate some statistics explicitly. The proof is almost identical to that of the special case of Proposition 1.41, and can be obtained by replacing $\pi^{-1/4} \exp(-x^2)$ by $w(x)$, and the real line by the interval $[a, b]$.

Proposition 1.42 *The joint p.d.f. of the eigenvalues for a unitary ensemble is given by*

$$p_N(\lambda_1, \lambda_2, \ldots, \lambda_N) = \frac{1}{N!} \det \left(K_N(\lambda_j, \lambda_k)\right)_{j,k=1}^{N}, \tag{34}$$

in which

$$K_N(x, y) = (w(x)w(y))^{1/2} \sum_{j=0}^{N-1} \pi_j(x)\pi_j(y), \tag{35}$$

and $(\pi_j(x))$ is the sequence of polynomials arising from the Gram–Schmidt process which are orthonormal with respect to w on the interval $[a, b]$.

K_N is often referred to as a "kernel". We can now see why Tracy–Widom operators are so fundamental to Random Matrix Theory, since the Christoffel-Darboux formula (Lemma 1.38) allows us to write the kernel as

$$K_N(x,y) = C_N(w(x)w(y))^{1/2}\frac{P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y)}{x-y},$$

where the constant C_N comes from the three-term recurrence relation for the orthogonal polynomials $P_j(x)$. This is in the form of a Tracy–Widom kernel (1).

The n-point correlation function $R_n(x_1, x_2, \dots, x_n)$ (in the context of general ensembles) is given by the formula

$$R_n(\lambda_1, \lambda_2, \dots, \lambda_n) = \frac{N!}{(N-n)!}\int_a^b \dots \int_a^b p_{N,\beta}(\lambda_1, \lambda_2, \dots, \lambda_N)\,\mathrm{d}\lambda_{n+1}\,\mathrm{d}\lambda_{n+2}\dots\,\mathrm{d}\lambda_N.$$

For unitary ensembles, as in the GUE case, we have the formula

$$R_n(\lambda_1, \lambda_2, \dots, \lambda_n) = \det(K_N(\lambda_j, \lambda_k))_{j,k=1}^n, \tag{36}$$

which is obtained exactly as before.

Another quantity of interest in random matrix theory is the so-called "gap" or "hole" probability $A(L)$, which is defined as the probability that an interval $[-\theta/2, \theta/2]$ will contain no eigenvalues. The evaluation of quantities such as this is one of the main reasons for studying integral operators with Tracy–Widom kernels. A derivation of the well-known relation

$$A(\theta) = \sum_{j=0}^{N}\frac{(-1)^j}{j!}\int_{-\theta/2}^{\theta/2}\cdots\int_{-\theta/2}^{\theta/2} R_j(\lambda_1, \dots, \lambda_j)\,\mathrm{d}\lambda_1\dots\,\mathrm{d}\lambda_j$$

can be found in [16, p.44-45]. With the expression (36) for the correlation function R_n given above, we see that this can also be written as

$$A(\theta) = \sum_{j=0}^{N} \frac{(-1)^j}{j!} \int_{-\theta/2}^{\theta/2} \cdots \int_{-\theta/2}^{\theta/2} \det \left(K_N(\lambda_i, \lambda_k)\right)_{1\le i,k\le j} \, \mathrm{d}\lambda_1 \ldots \mathrm{d}\lambda_j,$$

which is the Fredholm determinant

$$\det\left(I - \mathbb{I}_{[-\frac{\theta}{2},\frac{\theta}{2}]} K_N \mathbb{I}_{[-\frac{\theta}{2},\frac{\theta}{2}]}\right)$$

of the integral operator K_N with kernel $K_N(x,y)$ compressed to act on $L^2[\frac{-\theta}{2}, \frac{\theta}{2}]$. Using the theory of Fredholm determinants (see §1.5), we see that the eigenvalues $x_0, \ldots, x_{N-1}$ of the integral operator $\mathbb{I}_{[-\frac{\theta}{2},\frac{\theta}{2}]} K_N \mathbb{I}_{[-\frac{\theta}{2},\frac{\theta}{2}]}$ give the value of the determinant: this is

$$\det(I - \mathbb{I}_{[-\frac{\theta}{2},\frac{\theta}{2}]} K_N \mathbb{I}_{[-\frac{\theta}{2},\frac{\theta}{2}]}) = \prod_{i=0}^{N-1} (1 - x_i).$$

1.11 Special functions

Here we summarise the definitions and properties of the special functions we will be using throughout this thesis, and state, where relevant, the associated Tracy–Widom system.

The Airy function

The function

$$\mathrm{Ai}\,(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(zt+t^3/3)} \, \mathrm{d}t$$

satisfies the differential equation

$$u''(z) - zu(z) = 0,$$

which we can write as the Tracy–Widom system

$$\frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} u(x) \\ u'(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ x & 0 \end{bmatrix}\begin{bmatrix} u(x) \\ u'(x) \end{bmatrix}, \tag{37}$$

and the associated Tracy–Widom kernel is the *Airy kernel*

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

The Airy function has the following asymptotic formula (see [41, p.18]):

$$\mathrm{Ai}\,(x) = \frac{1}{2x^{1/4}\sqrt{\pi}}\left(1 + O(x^{-3/2})\right)\exp\left(-\frac{2}{3}x^{3/2}\right),$$

so, in particular, $\mathrm{Ai}\,(x) \to 0$ as $x \to \infty$.

Bessel functions

We shall always deal with the Bessel function of the first kind, which is defined for all values of $\alpha \in \mathbb{R}$ by the series

$$\mathrm{J}_\alpha(z) = \sum_{r=0}^{\infty} \frac{(-1)^r(\frac{1}{2}z)^{\alpha+2r}}{r!\Gamma(\alpha+r+1)}.$$

It is easily verified from this series expansion that $\mathrm{J}_\alpha(z)$ satisfies the Bessel differential equation, that is

$$z^2\,\mathrm{J}''_\alpha(z) + z\,\mathrm{J}'_\alpha(z) + (z^2-\alpha^2)\,\mathrm{J}_\alpha(z) = 0. \tag{38}$$

If $\alpha + r + 1 \leq 0$, we replace $(\Gamma(\alpha+r+1))^{-1}$ by zero. With this in mind, we also have the relations

$$\mathrm{J}_{-\alpha}(z) = (-1)^\alpha\,\mathrm{J}_\alpha(z),$$

and

$$\mathrm{J}_\alpha(z) + \mathrm{J}_{\alpha+2}(z) = \frac{2(\alpha+1)}{z}\,\mathrm{J}_{\alpha+1}(z),$$

for all $\alpha \in \mathbb{R}$, which can also be obtained from the series expansion (see [41, p.14]).

We have the Tracy–Widom system

$$x\frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} \mathrm{J}_\alpha(\sqrt{x}) \\ \sqrt{x}\,\mathrm{J}'_\alpha(\sqrt{x}) \end{bmatrix} = \begin{bmatrix} 0 & \sqrt{x} \\ -1/4\,(\alpha^2 - x) & 1/4 \end{bmatrix}\begin{bmatrix} \mathrm{J}_\alpha(\sqrt{x}) \\ \sqrt{x}\,\mathrm{J}'_\alpha(\sqrt{x}) \end{bmatrix}, \tag{39}$$

and the associated Bessel kernel

$$\frac{\mathrm{J}_\alpha(\sqrt{x})\sqrt{y}\,\mathrm{J}'_\alpha(\sqrt{y}) - \mathrm{J}_\alpha(\sqrt{y})\sqrt{x}\,\mathrm{J}'_\alpha(\sqrt{x})}{x-y},$$

which will form part of the calculations in Chapter 2.

Laguerre functions

Let $(L_n^{(\alpha)}(x))$ be the sequence of polynomials orthogonal with respect to the weight $w(x) = x^\alpha \exp(-x/2)$ on the interval $[0,\infty)$, where we take $\alpha > -1$ to ensure that the weight is integrable. These are the Laguerre polynomials, and they satisfy the differential equation (see [41, p.99])

$$x\frac{\mathrm{d}^2}{\mathrm{d}x^2}L_n^{(\alpha)}(x) + (\alpha+1-x)\frac{\mathrm{d}}{\mathrm{d}x}L_n^{(\alpha)}(x) + nL_n^{(\alpha)}(x) = 0. \tag{40}$$

Take $u(x) = xe^{-x/2}L_n^{(1)}(x)$. Then an easy calculation shows that

$$u''(x) + \left(-1/4 + \frac{n+1}{x}\right)u(x) = 0 \quad (x>0),$$

which gives rise to the Tracy–Widom system

$$\frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} u(x) \\ u'(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1/4 - (n+1)/x & 0 \end{bmatrix}\begin{bmatrix} u(x) \\ u'(x) \end{bmatrix}, \tag{41}$$

and the associated kernel

$$\frac{u(x)u'(y) - u'(x)u(y)}{x-y}.$$

A generalisation of the function u above, the Laguerre functions may be defined as

$$u_n^{(\alpha)}(x) = x^\alpha e^{-x/2} L_n^{(\alpha)}(x), \tag{42}$$

and when $\alpha = 0$ they form the basis of $L^2[0,\infty)$ under which Hankel integral operators and Hankel operators on ℓ^2 may be identified: see §1.12. The Laplace transform of the Laguerre functions will be useful in some later applications in the thesis, and it can be calculated as follows. We use the following alternative formulation, often called the Rodrigues's formula for the Laguerre polynomials (see [41, p.100]):

$$L_n^{(\alpha)}(x) = \frac{e^x}{n! x^\alpha} \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(x^{n+\alpha} e^{-x}\right), \tag{43}$$

and so

$$u_n^{(\alpha)}(x) = \frac{e^{x/2}}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(x^{n+\alpha} e^{-x}\right).$$

We have, on integrating by parts,

$$\begin{aligned}
n!\mathfrak{L}(u_n^{(\alpha)}(x);\lambda) &= \int_0^\infty e^{t/2} \frac{\mathrm{d}^n}{\mathrm{d}t^n}\left(t^{n+\alpha} e^{-t}\right) e^{-\lambda t}\,\mathrm{d}t \\
&= \int_0^\infty e^{-t(\lambda-1/2)} \frac{\mathrm{d}^n}{\mathrm{d}t^n}\left(t^{n+\alpha} e^{-t}\right)\,\mathrm{d}t \\
&= \left[e^{-t(\lambda-1/2)} \frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}}\left(t^{n+\alpha} e^{-t}\right)\right]_0^\infty \\
&\quad +(\lambda - 1/2)\int_0^\infty e^{-t(\lambda-1/2)} \frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}}\left(t^{n+\alpha} e^{-t}\right)\,\mathrm{d}t.
\end{aligned}$$

The boundary terms are clearly zero, and if we repeat the above calculation a further $n-1$ times, the resulting boundary terms will always be zero. Thus we have

$$\begin{aligned}
n!\mathfrak{L}(u_n^{(\alpha)}(x);\lambda) &= (\lambda - 1/2)^n \int_0^\infty e^{-t(\lambda-1/2)} t^{n+\alpha} e^{-t}\,\mathrm{d}t \\
&= (\lambda - 1/2)^n \int_0^\infty e^{-t(\lambda+1/2)} t^{n+\alpha}\,\mathrm{d}t \\
&= (\lambda - 1/2)^n \int_0^\infty \left(\frac{x}{\lambda+1/2}\right)^{n+\alpha} e^{-x} \frac{\mathrm{d}x}{\lambda+1/2}
\end{aligned}$$

$$= \frac{(\lambda - 1/2)^n}{(\lambda + 1/2)^{n+\alpha+1}} \int_0^\infty x^{n+\alpha} e^{-x}\, \mathrm{d}x$$
$$= \frac{(\lambda - 1/2)^n}{(\lambda + 1/2)^{n+\alpha+1}} \Gamma(n+\alpha),$$

where the Gamma function is ([42, p.55])

$$\Gamma(z) = \int_0^\infty y^{z-1} e^{-y}\, \mathrm{d}y.$$

Hence we have

$$\mathfrak{L}(u_n^{(\alpha)}(x); \lambda) = \frac{\Gamma(n+\alpha)}{n!} \frac{(\lambda - 1/2)^n}{(\lambda + 1/2)^{n+\alpha+1}},$$

and when α is a positive integer, this simplifies, via the relation $\Gamma(n) = (n-1)!$, to

$$\mathfrak{L}(u_n^{(\alpha)}(x); \lambda) = (n+\alpha)(n+\alpha-1)\dots(n+1)\frac{(\lambda - 1/2)^n}{(\lambda + 1/2)^{n+\alpha+1}}.$$

1.12 Hankel operators on $L^2[0, \infty)$

There is a continuous analogue of the Hankel operators discussed in §1.6, namely, integral operators on the half line $[0, \infty)$, with kernel of the form $\phi(x+y)$:

$$\Gamma_\phi f(x) = \int_0^\infty \phi(x+y) f(y)\, \mathrm{d}y.$$

These are called *Hankel integral operators*, and it is a remarkable fact that they are unitarily equivalent to Hankel operators on ℓ^2 (as defined in §1.6). The calculations needed to show this are somewhat involved (they are detailed in [37, p.46-53]) but we can give some indication of how this correspondence comes about as follows. Define $H^2(\mathbb{C}_+)$ as the space of functions holomorphic on the right half plane and such that there exists $M < \infty$ with

$$\int_{-\infty}^\infty |f(x+iy)|^2\, \mathrm{d}y < M \text{ for all } x > 0.$$

We have already observed that the sequence space ℓ^2 can be identified with $H^2(\mathbb{D})$ by the obvious correspondence $(a_n)_{n=0}^\infty \mapsto \sum_{n=0}^\infty a_n z^n$. The Möbius transformation

$$\lambda \mapsto \frac{\lambda - 1/2}{\lambda + 1/2}$$

suggests the map

$$\sum_{n=0}^\infty a_n z^n \mapsto \sum_{n=0}^\infty a_n \frac{(\lambda - 1/2)^n}{(\lambda + 1/2)^{n+1}},$$

which maps $H^2(\mathbb{D})$ into $H^2(\mathbb{C}^+)$, and we can show that this map is unitary. Finally, the Laplace transform is, by the Paley-Wiener theorem (see [22, p.132]), a unitary map from $L^2[0,\infty)$ to $H^2(\mathbb{C}^+)$, and it maps the basis

$$\{u_n^{(0)}(4t)\}_{n\geq 0}$$

of the space $L^2[0,\infty)$ to the basis of $H^2(\mathbb{C}^+)$ described above, where u_n^0 are the Laguerre functions as defined in (42). Peller [37, p.46-53] shows that a Hankel integral operator on $L^2[0,\infty)$ has a Hankel matrix with respect to this basis of Laguerre functions.

1.13 Hankel operator squares in discrete and continuous contexts

We wish to find operators K such that $0 \leq K \leq I$ and K is trace class. The kernels of operators of this kind lead to determinantal point processes as described by Soshnikov [40], which are processes whose probability density functions have the form

$$\det[K(x_i, x_j)]_{i,j}.$$

Particular examples of this are the eigenvalue systems described in §1.9 and §1.10. If $K = \Gamma^*\Gamma$, where Γ is a Hilbert–Schmidt Hankel operator satisfying $\|\Gamma\| \leq 1$, then clearly $0 \leq K \leq I$ and K is trace class. If further we can write $K = \Gamma^2$, then we have a way to calculate the Fredholm determinant, since

$$\det(I-K)=\det(I-\Gamma)\det(I+\Gamma).$$

Fredholm determinants are involved in the calculation of probabilistic quantities for Random Matrix Ensembles (see the end of §1.8), so writing a Tracy–Widom operator as $\Gamma^*\Gamma$ or Γ^2 is a useful aim. In the continuous case, Tracy and Widom have observed that many important systems arising in random matrix theory can be analysed by kernels arising from differential equations of the form

$$\frac{\mathrm{d}}{\mathrm{d}x}a(x)=\begin{bmatrix}\alpha(x) & \beta(x)\\ -\gamma(x) & -\alpha(x)\end{bmatrix}a(x),$$

where the matrix entries α, β and γ are rational functions, and $a(x)$ is a 2×1 vector of functions. Blower [7] gives sufficient conditions for such kernels to be expressible as the square of a Hankel operator, and we quote and prove this result in a simplified form. We define the involution matrix, which will be used at several points here and in Chapter 5:

$$J=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}.$$

It is obvious that this matrix satisfies $J^t=-J$ and $J^2=-I$.

Theorem 1.43 *Let $a(x)$ be a 2×1 vector of differentiable $L^2(0,\infty)$ functions satisfying the differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}x}a(x)=\Omega(x)a(x),$$

for some 2×2 matrix $\Omega(x)$ of rational functions. Suppose that $a(x)\to 0$ as $x\to\infty$, and also that

$$\frac{J\Omega(x)+\Omega(y)^tJ}{x-y}=C$$

where C is a constant and rank one matrix with non-zero eigenvector $\lambda < 0$ and corresponding eigenvector v_λ. Then

$$K(x,y) := \frac{\langle Ja(x), a(y)\rangle}{x-y} = \int_0^\infty \phi(x+t)\phi(y+t)\,\mathrm{d}t \tag{44}$$

where

$$\phi(z) = |\lambda|^{1/2} \langle v_\lambda, a(z)\rangle.$$

Remark

Note that Γ will in general be different from the positive square root $K^{1/2}$ of K.

Proof. The trick is to consider the effect of a special differential operator, namely $\frac{\partial}{\partial x} + \frac{\partial}{\partial y}$, on both sides of the equation (44). This introduces two "un-differentiated" terms which cancel, and we get

$$\begin{aligned}
\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) K(x,y) &= \frac{1}{x-y}\left(\langle J\frac{\mathrm{d}}{\mathrm{d}x}a(x), a(y)\rangle + \langle Ja(x), \frac{\mathrm{d}}{\mathrm{d}y}a(y)\rangle\right) \\
&= \frac{1}{x-y}\left(\langle J\Omega(x)a(x), a(y)\rangle + \langle Ja(x), \Omega(y)a(y)\rangle\right) \\
&= \frac{1}{x-y}\langle\left(J\Omega(x) + \Omega(y)^t J\right) a(x), a(y)\rangle \\
&= \langle Ca(x), a(y)\rangle \\
&= -\phi(x)\phi(y).
\end{aligned}$$

The other side of the equation yields

$$\begin{aligned}
&\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) \int_0^R \phi(x+t)\phi(y+t)\,\mathrm{d}t \\
= &\int_0^R \left(\phi'(x+t)\phi(y+t) + \phi(x+t)\phi'(y+t)\right)\,\mathrm{d}t \\
= &\left([\phi(x+t)\phi(y+t)]_0^R - \int_0^R \phi(x+t)\phi'(y+t)\,\mathrm{d}t\right) \\
&+ \int_0^R \phi(x+t)\phi'(y+t)\,\mathrm{d}t \\
\to &-\phi(x)\phi(y) \quad \text{as} \quad R \to \infty,
\end{aligned}$$

since $\phi(x) \to 0$ as $x \to \infty$, by the vanishing condition on $a(x)$. Hence

$$K(x,y) - \int_0^\infty \phi(x+t)\phi(y+t)\,\mathrm{d}t = g(x-y) \quad \text{for all} \quad x, y \in \mathbb{R}$$

for some function g. It is clear that $K(x,y) \to 0$ as x or $y \to \infty$, and the same is true of the integral expression:

$$\begin{aligned} \left| \int_0^\infty \phi(x+t)\phi(y+t)\,\mathrm{d}t \right| &\leq \left(\int_0^\infty \phi(x+t)^2\,\mathrm{d}t \right)^{1/2} \left(\int_0^\infty \phi(y+t)^2\,\mathrm{d}t \right)^{1/2} \\ &\to 0 \quad \text{as} \quad x \text{ or } y \to \infty, \end{aligned}$$

since $\phi \in L^2(0,\infty)$. We deduce that $g(x-y) \to 0$ as x or $y \to \infty$, which implies that g is in fact identically zero for all x and y, and we have the required identity (44). ■

Example 1.44

We consider the Airy kernel, which arises when we scale the eigenvalues at the soft spectral edge of the Gaussian Unitary Ensemble (see [43], or §1.9 above). The differential equation satisfied by the Airy function can be written as

$$\frac{\mathrm{d}}{\mathrm{d}x} \begin{bmatrix} \mathrm{Ai}\,(x) \\ \mathrm{Ai}'(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ x & 0 \end{bmatrix} \begin{bmatrix} \mathrm{Ai}\,(x) \\ \mathrm{Ai}'(x) \end{bmatrix}.$$

If we write the 2×2 matrix of coefficients as $\Omega(x)$, observe that

$$\frac{J\Omega(x) + \Omega(y)^t J}{x-y} = \mathrm{diag}\,(-1, 0)$$

and also that the Airy function vanishes at infinity, so the Airy Kernel can be written as the square of a Hankel integral operator

$$\frac{\mathrm{Ai}\,(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}\,(y)}{x-y} = \int_0^\infty \mathrm{Ai}\,(x+z)\,\mathrm{Ai}\,(y+z)\,\mathrm{d}z.$$

Discrete analogues

The new results in Chapter 5 of this thesis are about how to express discrete operators as Hankel squares, i.e. to find sufficient conditions on K such that there exists a Hankel operator Γ such that $\Gamma = \Gamma^*$ and $K = \Gamma^2$. This amounts to finding a function $\phi : \mathbb{Z}_+ \to \mathbb{R}$ such that the matrix entries satisfy

$$K(m,n) = \sum_{k=0}^{\infty} \phi(m+k)\phi(n+k) \quad \text{for } m \neq n,$$

and the diagonal entries (which are in general unspecified) are then defined by the Hankel operator itself. The uninitiated reader might think that that the results already obtained for continuous operators would transfer across to the discrete case, particularly given the unitary equivalence between Hankel operators on $L^2[0,\infty)$ and on ℓ^2 which we discussed in §1.12. We have found that this is not the case, and that the conditions are fundamentally different. Nonetheless, it is interesting to draw parallels, and notice the formal similarity between the two cases, which we do in the table on the following page.

Discrete operators	**Continuous operators**
Matrix entries: $K(m,n) = \dfrac{\langle Ja(m), a(n)\rangle}{m-n},$ for a real sequence (a(n)) of 2×1 vectors.	Kernel: $Kf(x) = \displaystyle\int_0^\infty K(x,y)f(y)\,\mathrm{d}y,$ where $K(x,y) = \dfrac{\langle Ja(x), a(y)\rangle}{x-y}$
Recurrence relation: $a(j+1) = T(j)a(j)$ for some 2×2 matrix $T(j)$ satisfying $\det(T(j)) = 1$	Differential equation: $\dfrac{\mathrm{d}}{\mathrm{d}x}a(x) = \Omega(x)a(x),$ for some 2×2 matrix of rational functions satisfying $\det(\Omega(x)) = 0$
Discrete Lyapunov condition: $\dfrac{J - T(n)^t JT(m)}{m-n} = B(n)^t CB(m),$ where C is a rank 1 real symmetric matrix with non-zero eigenvalue $\lambda < 0$ and corresponding eigenvector v_λ and $B(n)$ is another 2×2 matrix.	Continuous Lyapunov condition: $\dfrac{J\Omega(x) + \Omega(y)^t J}{x-y} = C,$ where C is a rank one real symmetric matrix with non-zero eigenvalue $\lambda < 0$ and corresponding eigenvector v_λ.
Square of Hankel matrix $K(m,n) = \displaystyle\sum_{k=0}^{\infty} \phi(m+k)\phi(n+k),$ where $\phi(j) = \langle v_\lambda, B(j)a(j)\rangle$	Square of Hankel integral operator $K(x,y) = \displaystyle\int_0^\infty \phi(x+t)\phi(y+t)\,\mathrm{d}t,$ where $\phi(z) = \langle v_\lambda, a(z)\rangle$.

2 Eigenvalue scalings and the Jacobi Kernel

2.1 Introduction

The limiting eigenvalue distribution of a random matrix ensemble as the matrix dimension N tends to infinity is of interest both in random matrix theory, and in its applications to nuclear physics (see [27, §1.1] for some discussion of this). Since an $N \times N$ matrix will typically have a largest eigenvalue which grows with N, we can never simply let N tend to infinity and hope to get a sensible asymptotic distribution without some sort of scaling operation. We can carry out scaling in different parts of the spectrum, resulting in different kernels which describe the asymptotic eigenvalue distribution. The scaling and limit taking operation can be carried out on the kernel $K_N(x, y)$: here we shall do this for the Jacobi kernel, and we use a "hard edge" scaling, which describes the eigenvalue distribution at the right-hand end of the interval $[-1, 1]$. The terminology "hard" refers to the fact that no eigenvalues can be found outside this interval, in contrast to "soft edge" scaling in, for example, the Laguerre ensemble (see [15, §2]), where the eigenvalues lie in the interval $[0, \infty)$. The first step of the argument involves taking limits of the coefficients of the scaled Jacobi differential equation to prove convergence of the scaled Jacobi polynomials. We then apply this convergence result to show that the scaled Jacobi kernel converges to another Tracy-Widom kernel, which we call the Bessel kernel. A result from operator theory can then be applied to show that we have trace class convergence of the integral operator arising from the Jacobi kernel to that of the Bessel kernel. The eigenvalues of the Jacobi integral operator can be used to calculate the gap probability for the Jacobi ensemble, and so this trace class convergence is important, since it implies convergence of determinants. The result was suggested by Forrester [15], then proved by Borodin [5] in a special case. Others, including Nagao and Wadati [31] have obtained the general case, but we use different techniques, and also make clear that the operators converge in trace class norm. The Bessel kernel, and the system of partial differential equations satisfied by its Fredholm determinant, are considered in [45], but again these authors do note make clear the mode of convergence of the Jacobi integral operator to the Bessel operator. The technique of finding limits via convergence of differential equation coefficients would be easily applicable to other orthogonal polynomial systems (and their associated random matrix ensembles), since the polynomials will often satisfy a simple second order differential equation.

2.2 The Jacobi ensemble and electrostatic systems

We introduce the Jacobi weight

$$w_{\alpha,\beta}(x) = 2^{-(\alpha+\beta+1)}(1-x)^{\alpha}(1+x)^{\beta}, \quad x \in [-1,1], \quad (\alpha > -1, \beta > -1)$$

in which the constraints on α and β are to ensure integrability in the interval $[-1,1]$. Following Szegö [41, p.68], we define the Jacobi polynomials $P_n^{\alpha,\beta}(x)$ to be the sequence of orthogonal polynomials arising from the Gram–Schmidt process for which P_n has degree n and positive leading coefficient and

$$\int_{-1}^{1} P_j(x)P_k(x)w(x)\,\mathrm{d}x = 0$$

for $j \neq k$, while

$$\int_{-1}^{1} P_n(x)^2 w(x)\,\mathrm{d}x = \frac{1}{2n+\alpha+\beta+1}\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}.$$

The Jacobi unitary ensemble of order N has joint eigenvalue probability density function

$$p_N^{\alpha,\beta}(\lambda_1,\dots,\lambda_N) = \frac{1}{Z_N}\prod_{j=1}^{N} w_{\alpha,\beta}(\lambda_j) \prod_{1<j<k\leq N}(\lambda_j-\lambda_k)^2, \tag{45}$$

for a normalisation constant Z_N, where the eigenvalues are constrained to lie in the interval $[-1,1]$. Following the theory outlined in §1.10, the joint p.d.f can be written as

$$p_N^{\alpha,\beta}(\lambda_1, \ldots, \lambda_N) = \frac{1}{N!} \det(K_N(\lambda_i, \lambda_j))_{i,j=1,\ldots N},$$

in which

$$K_N(x, y) = \sum_{j=0}^{N-1} \pi_j(x)\pi_j(y)(w(x)w(y))^{1/2} \tag{46}$$

and $\pi_j(x)$ are the polynomials which are orthonormal with respect to the weight $w(x)$ on $[-1, 1]$. Thus

$$\pi_j(x) = Q_j P_j^{\alpha,\beta}(x),$$

where

$$Q_j = \left((2j + \alpha + \beta + 1)\frac{\Gamma(j + 1)\Gamma(j + \alpha + \beta + 1)}{\Gamma(j + \alpha + 1)\Gamma(j + \beta + 1)}\right)^{1/2}.$$

We wish to write this in the form of a Tracy–Widom kernel, using the Christoffel-Darboux formula. For this we need the constant C_n from the three-term recurrence relation for the π_j

$$x\pi_{n-1}(x) = A_n\pi_{n-2}(x) + B_n\pi_{n-1}(x) + C_n\pi_n(x).$$

The three-term recurrence relation for the Jacobi polynomials $P_j(x)$ (see [41, p.71]) can be written as

$$\begin{aligned} xP_{n-1}(x) &= \frac{2(n + \alpha - 1)(n + \beta - 1)}{(2n + \alpha + \beta - 1)(2n + \alpha + \beta - 2)} P_{n-2}(x) \\ &\quad + \frac{(\alpha^2 - \beta^2)}{(2n + \alpha + \beta)(2n + \alpha + \beta - 2)} P_{n-1}(x) \\ &\quad + \frac{2n(n + \alpha + \beta)}{(2n + \alpha + \beta - 1)(2n + \alpha + \beta)} P_n(x), \end{aligned}$$

and hence, using Lemma 1.37,

$$\begin{aligned} C_n &= \int_{-1}^{1} x\pi_{n-1}(x)\pi_n(x)w(x)\,\mathrm{d}x \\ &= \int_{-1}^{1} xQ_{n-1}P_{n-1}(x)Q_nP_n(x)w(x)\,\mathrm{d}x \end{aligned}$$

$$
\begin{aligned}
&= Q_{n-1}Q_n \int_{-1}^{1} \frac{2n(n+\alpha+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)} P_n(x)^2 w(x)\,\mathrm{d}x \\
&= Q_{n-1}Q_n \frac{2n(n+\alpha+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)} Q_n^{-2} \\
&= \frac{Q_{n-1}}{Q_n} \frac{2n(n+\alpha+\beta)}{(2n+\alpha+\beta-1)(2n+\alpha+\beta)}.
\end{aligned}
$$

We can now use Proposition 1.38 to write

$$
\begin{aligned}
&\sum_{j=0}^{N-1} \pi_j(x)\pi_j(y) \\
= &\ \frac{Q_{N-1}}{Q_N} \frac{2N(N+\alpha+\beta)}{(2N+\alpha+\beta-1)(2N+\alpha+\beta)} \frac{\pi_N(x)\pi_{N-1}(y) - \pi_{N-1}(x)\pi_N(y)}{x-y} \\
= &\ Q_{N-1}^2 \frac{2N(N+\alpha+\beta)}{(2N+\alpha+\beta-1)(2N+\alpha+\beta)} \frac{P_N(x)P_{N-1}(y) - P_{N-1}(x)P_N(y)}{x-y} \\
= &\ \frac{2N(N+\alpha+\beta)\Gamma(N)\Gamma(N+\alpha+\beta)}{(2N+\alpha+\beta)\Gamma(N+\alpha)\Gamma(N+\beta+1)} \frac{P_N(x)P_{N-1}(y) - P_{N-1}(x)P_N(y)}{x-y}.
\end{aligned}
$$

Later we shall let $N \to \infty$, so to save notational messiness, observe that the constant term in the latter expression is like $(N+\alpha+\beta)$ as $N \to \infty$, and define

$$
K_N(x,y) = (N+\alpha+\beta) \frac{P_N(x)P_{N-1}(y) - P_{N-1}(x)P_N(y)}{x-y} (w(x)w(y))^{1/2}. \tag{47}
$$

The zeros of the Jacobi orthogonal polynomials correspond to a certain minimisation problem in electrostatics. Suppose that we place fixed positively charged particles of charge p and q respectively at the endpoints of the interval $[-1,1]$, and n negatively charged particles x_n on the interior of the interval. The negative charges repel each other, and are attracted towards the positively charged particles at either end. We look for the arrangement of charges which minimises the expression

$$
-\log\left(\prod_{i=1}^{n} (1-x_i)^p (1+x_i)^q \prod_{1 \le i < j \le n} |x_i - x_j|\right),
$$

which is the potential energy of the system. Szegö [41, p.140] shows that the minimising set for this system is precisely the zeros of the Jacobi polynomial $P_n^{\alpha,\beta}$, where $\alpha = 2p-1$ and $\beta = 2q-1$. Note that the expression inside the logarithm is the joint p.d.f (without normalisation constant) of the eigenvalues from the Jacobi Orthogonal ensemble, since the power on the van der Monde determinant term is one. Similar results hold for other orthogonal polynomials, and the optimal distribution of particles in each case is the modal eigenvalue distribution for the corresponding random matrix ensemble.

2.3 Scaling in the Jacobi Ensemble

Let the eigenvalues of a matrix from the Jacobi ensemble be the sequence $(\lambda_j)_{j\leq n}$. We shall show that at the right-hand edge of the interval $[-1,1]$, the asymptotic distribution of eigenvalues can be described, under the scaling

$$\lambda_n \mapsto \cos\left(\frac{\sqrt{\lambda_n}}{n}\right),$$

by the Bessel kernel

$$K(x,y) = \frac{\mathrm{J}_\alpha(\sqrt{x})\sqrt{y}\,\mathrm{J}'_\alpha(\sqrt{y}) - \mathrm{J}_\alpha(\sqrt{y})\sqrt{x}\,\mathrm{J}'_\alpha(\sqrt{x})}{(x-y)}. \tag{48}$$

on the space $L^2[0,1]$. Specifically, we consider the subinterval in which there are $O(1)$ eigenvalues. Nagao and Wadati [31] show that the level density in the Jacobi ensemble is like

$$\frac{n}{\pi\sqrt{1-x^2}},$$

so that the interval $[1-a_n/n^2, 1]$, with

$$a_n = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{(2k)!n^{2k-2}} = n^2\left(1-\cos\left(\frac{1}{n}\right)\right),$$

contains the required $O(1)$ eigenvalues. The scaled eigenvalues are equal to

$$n^2\left(\cos^{-1}\lambda_j\right)^2,$$

so the endpoint 1 of the subinterval $[1 - a_n/n^2, 1]$ is mapped to 0, and $1 - a_n/n^2$ is mapped to

$$\begin{aligned} n^2 \cos^{-1}\left(1 - \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{(2k)!n^{2k}}\right)^2 &= n^2 \cos^{-1}\left(\cos\left(\frac{1}{n}\right)\right)^2 \\ &= 1. \end{aligned}$$

Having established that the rescaled eigenvalues lie in the interval $[0, 1]$, we define an integral operator on $L^2[0, 1]$ with kernel

$$\tilde{K}_n(x, y) = \frac{2^{-\alpha}}{n^2} K_n\left(\cos\frac{\sqrt{x}}{n}, \cos\frac{\sqrt{y}}{n}\right), \tag{49}$$

which describes the distribution of the rescaled eigenvalues at the edge of the interval in the Jacobi ensemble of order n.

2.4 Asymptotic formulae for the Jacobi polynomials

We now prove the two asymptotic limits of the Jacobi polynomials which are needed to establish the convergence of the scaled Jacobi kernel to the Bessel kernel. These limits were known to Mehler and Heine for the special case of the Legendre polynomials ($\alpha = \beta = 0$), and Szegö describes how to derive them from other asymptotic formulae (see [41, p.190]), but the proofs here depend on elementary results about differential equations in the complex domain. We state the following result, which is known as Grönwall's inequality (see [18, p.15] for discussion and a proof).

Lemma 2.1 *Let K be a continuous, non-negative and integrable function on an interval $[a, b]$, and let f and g be continuous non-negative functions on $[a, b]$. Suppose that for all t in $[a, b]$*

$$f(t) \leq g(t) + \int_a^t K(s)f(s)\,\mathrm{d}s. \tag{50}$$

Then

$$f(t) \leq g(t) + \int_a^t K(s) \exp\left(\int_s^t K(u)\,\mathrm{d}u\right) g(s)\,\mathrm{d}s. \tag{51}$$

We now use Grönwall's inequality in a special case to prove a convergence result for differential equations, which we will ultimately use to deduce limits for the Jacobi polynomials.

Lemma 2.2 *Suppose $(g_n(z))$ is a sequence of $k \times 1$ vectors of functions satisfying the differential equation*

$$\frac{\mathrm{d}}{\mathrm{d}z} g_n(z) = \Omega_n(z) g_n(z), \tag{52}$$

where $\Omega_n(z)$ is a $k \times k$ matrix of analytic functions which converges uniformly on compact subsets of $\mathbb{C}$ to a limiting matrix $\Omega(z)$ as $n \to \infty$. Let $g(z)$ be a $k \times 1$ solution vector of

$$\frac{\mathrm{d}}{\mathrm{d}z} g(z) = \Omega(z) g(z),$$

and suppose that

$$g_n(0) \to g(0) \qquad \text{as } n \to \infty.$$

Then the sequence of vectors $(g_n(z))$ converges to $g(z)$ as $n \to \infty$, uniformly on compact subsets of $\mathbb{C}$.

Proof. Integrating the differential equation satisfied by g_n gives

$$g_n(z) - g_n(0) = \int_{[0,z]} \Omega_n(\zeta) g_n(\zeta)\,\mathrm{d}\zeta,$$

and similarly

$$g(z) - g(0) = \int_{[0,z]} \Omega(\zeta) g(\zeta)\,\mathrm{d}\zeta.$$

Subtracting these two equations, and rewriting, we obtain

$$g(z) - g_n(z) = g(0) - g_n(0) + \int_{[0,z]} \Omega_n(\zeta)\left(g(\zeta) - g_n(\zeta)\right)\mathrm{d}\zeta + \int_{[0,z]} \left(\Omega(\zeta) - \Omega_n(\zeta)\right) g(\zeta)\,\mathrm{d}\zeta.$$

Now let t be the real parameter for the interval $[0,z]$, and let $a_n(t) := \|g(z) - g_n(z)\|$, $\omega_n(t) := \|\Omega_n(z)\|$, $b_n(t) := \|\Omega(z) - \Omega_n(z)\|$, and $c(t) := \|g(z)\|$. Then, by the triangle inequality,

$$a_n(t) \leq a_n(0) + \int_0^t \omega_n(s) a_n(s)\,\mathrm{d}s + \int_0^t b_n(s)c(s)\,\mathrm{d}s.$$

Since the convergence of Ω_n to Ω is uniform on $[0,z]$, we have a bound $\omega_n(s) \leq W(s)$, say, for all n, so that

$$a_n(t) \leq a_n(0) + \int_0^t W(s) a_n(s)\,\mathrm{d}s + \int_0^t b_n(s)c(s)\,\mathrm{d}s. \tag{53}$$

By definition, a_n, b_n, c and W are all positive and continuous on $[0,z]$, so we can apply Grönwall's inequality (Lemma 2.1) to (53), and we get

$$\begin{aligned} a_n(t) \;&\leq\; a_n(0) + \int_0^t b_n(s)c(s)\,\mathrm{d}s \\ &\quad + \int_0^t W(s) \exp\left(\int_s^t W(u)\,\mathrm{d}u\right)\left(a_n(0) + \int_0^s b_n(x)c(x)\,\mathrm{d}x\right)\,\mathrm{d}s. \end{aligned}$$

It is clear that $a_n(0) \to 0$ as $n \to \infty$ and that $b_n(t) \to 0$ as $n \to \infty$, uniformly on $[0,z]$, so the right hand side tends to zero as $n \to \infty$, and we get the result. ■

Lemma 2.3 *Let $f_n(z) = P_n\left(\cos\frac{\sqrt{z}}{n}\right)$. Then $f_n(z)$ is entire, and*

$$z\frac{\mathrm{d}}{\mathrm{d}z}\begin{bmatrix} f_n(z) \\ f_n'(z) \end{bmatrix} = \Omega_n \begin{bmatrix} f_n(z) \\ f_n'(z) \end{bmatrix}, \tag{54}$$

where

$$\Omega_n = \begin{bmatrix} 0 & z \\ -\frac{1}{4n}(n+\alpha+\beta+1) & \frac{\sqrt{z}}{2n}\left((\beta-\alpha)\operatorname{cosec}\frac{\sqrt{z}}{n} - (\alpha+\beta+1)\cot\frac{\sqrt{z}}{n}\right) - \frac{1}{2} \end{bmatrix}.$$

Further, as $n \to \infty$,

$$\Omega_n(z) \to \begin{bmatrix} 0 & z \\ -1/4 & -(\alpha+1) \end{bmatrix}$$

uniformly on compact subsets of $\mathbb{C} \setminus (-\infty, 0)$.

Proof. The Jacobi polynomials satisfy the following differential equation (see [41, p.60])

$$(1-t^2)P_n''(t) + (\beta - \alpha - (\alpha+\beta+2)t)\, P_n'(t) + n(n+\alpha+\beta+1)P_n(t) = 0. \quad (55)$$

We make the substitution $t = \cos\frac{\sqrt{z}}{n}$, and write $f_n(z) = P_n\left(\cos\frac{\sqrt{z}}{n}\right)$. We then have

$$f_n'(z) = -\frac{\sin(\frac{\sqrt{z}}{n})}{2n\sqrt{z}} P_n'\left(\cos\frac{\sqrt{z}}{n}\right), \quad (56)$$

and

$$f_n''(z) = \frac{\sin^2\left(\frac{\sqrt{z}}{n}\right)}{4n^2 z} P_n''\left(\cos\frac{\sqrt{z}}{n}\right) - \left(\frac{\cos\frac{\sqrt{z}}{n}}{4n^2 z} - \frac{\sin\frac{\sqrt{z}}{n}}{4nz^{3/2}}\right) P_n'\left(\cos\frac{\sqrt{z}}{n}\right), \quad (57)$$

so that

$$P_n'\left(\cos\frac{\sqrt{z}}{n}\right) = -2n\sqrt{z}\,\mathrm{cosec}\left(\frac{\sqrt{z}}{n}\right) f_n'(z) \quad (58)$$

and

$$\begin{aligned} P_n''\left(\cos\frac{\sqrt{z}}{n}\right) &= 4n^2 z\,\mathrm{cosec}^2\left(\frac{\sqrt{z}}{n}\right)\left(f_n''(z) + \left(\frac{\cos\frac{\sqrt{z}}{n}}{4n^2 z} - \frac{\sin\frac{\sqrt{z}}{n}}{4nz^{3/2}}\right) P_n'\left(\cos\frac{\sqrt{z}}{n}\right)\right) \\ &= 4n^2 z\,\mathrm{cosec}^2\left(\frac{\sqrt{z}}{n}\right)\left(f_n''(z) - \frac{\cot\frac{\sqrt{z}}{n}}{2n\sqrt{z}} f_n'(z) + \frac{1}{2z} f_n'(z)\right). \quad (59) \end{aligned}$$

Hence the f_n satisfy

$$\begin{aligned} & 4n^2 z\left(f_n''(z) - \frac{\cot\frac{\sqrt{z}}{n}}{2n\sqrt{z}} f_n'(z) + \frac{1}{2z} f_n'(z)\right) \\ - \; & 2n(\beta-\alpha)\sqrt{z}\,\mathrm{cosec}\left(\frac{\sqrt{z}}{n}\right) f_n'(z) + 2n(\alpha+\beta+2)\sqrt{z}\cot\left(\frac{\sqrt{z}}{n}\right) f_n'(x) \\ + \; & n(n+\alpha+\beta+1)f_n(z) = 0, \quad (60) \end{aligned}$$

or

$$
\begin{aligned}
& 4n^2 z f_n''(z) - 2n\sqrt{z}\cot\left(\frac{\sqrt{z}}{n}\right) f_n'(z) + 2n^2 f_n'(z) \\
- \quad & 2n(\beta-\alpha)\sqrt{z}\,\mathrm{cosec}\left(\frac{\sqrt{z}}{n}\right) f_n'(z) + 2n\sqrt{z}(\alpha+\beta+2)\cot\left(\frac{\sqrt{z}}{n}\right) f_n'(z) \\
+ \quad & n(n+\alpha+\beta+1) f_n(z) = 0. \qquad (61)
\end{aligned}
$$

Dividing by n^2, this becomes

$$
\begin{aligned}
4z f_n''(z) + \Bigg(2 - \frac{2}{n}(\beta-\alpha)\sqrt{z}\,\mathrm{cosec}\,\frac{\sqrt{z}}{n} & \\
+\frac{2}{n}(\alpha+\beta+1)\sqrt{z}\cot\frac{\sqrt{z}}{n}\Bigg) f_n'(z) & \\
+ \quad \frac{1}{n}(n+\alpha+\beta+1) f_n(z) = 0, & \qquad (62)
\end{aligned}
$$

and hence the differential equation satisfied by the rescaled polynomials is

$$
z\frac{\mathrm{d}}{\mathrm{d}z}\begin{bmatrix} f_n(z) \\ f_n'(z) \end{bmatrix} = \Omega_n \begin{bmatrix} f_n(z) \\ f_n'(z) \end{bmatrix},
$$

where

$$
\Omega_n = \begin{bmatrix} 0 & z \\ -\frac{1}{4n}(n+\alpha+\beta+1) & \frac{\sqrt{z}}{2n}\left((\beta-\alpha)\,\mathrm{cosec}\,\frac{\sqrt{z}}{n} - (\alpha+\beta+1)\cot\frac{\sqrt{z}}{n}\right) - \frac{1}{2} \end{bmatrix}.
$$

We note the limits

$$
\frac{1}{n}\,\mathrm{cosec}\,\frac{\sqrt{z}}{n} = \frac{1}{n\left(\frac{\sqrt{z}}{n} + O\left(\frac{1}{n^3}\right)\right)} \to \frac{1}{\sqrt{z}} \quad \text{as} \quad n \to \infty
$$

and

$$
\frac{1}{n}\cot\frac{\sqrt{z}}{n} = \frac{1 + O\left(\frac{1}{n^2}\right)}{n\left(\frac{\sqrt{z}}{n} + O\left(\frac{1}{n^3}\right)\right)} \to \frac{1}{\sqrt{z}} \quad \text{as} \quad n \to \infty,
$$

and deduce that

$$
\Omega_n(z) \to \begin{bmatrix} 0 & z \\ -1/4 & -(\alpha+1) \end{bmatrix}
$$

as $n \to \infty$, uniformly on compact subsets of $\mathbb{C} \setminus (-\infty, 0]$. ■

Lemma 2.4 *Let* J *be any solution of Bessel's equation*

$$z^2 \, \mathrm{J}''(z) + z \, \mathrm{J}'(z) + (z^2 - \alpha^2) \, \mathrm{J}(z) = 0. \tag{63}$$

Then

$$f(z) = z^{-\alpha/2} \, \mathrm{J}\left(\sqrt{z}\right),$$

where we define

$$z^{-\alpha/2} = \exp\left(\frac{-\alpha}{2} \log z\right),$$

satisfies

$$z \frac{\mathrm{d}}{\mathrm{d}z} \begin{bmatrix} f(z) \\ f'(z) \end{bmatrix} = \begin{bmatrix} 0 & z \\ -\frac{1}{4} & -(\alpha+1) \end{bmatrix} \begin{bmatrix} f(z) \\ f'(z) \end{bmatrix}, \tag{64}$$

for $z \in \mathbb{C} \setminus (-\infty, 0]$.

Proof. We substitute $\sqrt{z}$ for z in equation (63), to get

$$z \, \mathrm{J}''(\sqrt{z}) + \sqrt{z} \, \mathrm{J}'(\sqrt{z}) + (z - \alpha^2) \, \mathrm{J}(\sqrt{z}) = 0, \tag{65}$$

and show that equation (64) can be obtained by making the substitution $f(z) = z^{-\alpha/2} \, \mathrm{J}(\sqrt{z})$ in this modified Bessel equation. We have

$$f'(z) = -\frac{\alpha}{2} z^{-\alpha/2-1} \, \mathrm{J}(\sqrt{z}) + \frac{1}{2} z^{-\alpha/2-1/2} \, \mathrm{J}'(\sqrt{z})$$

and

$$\begin{aligned} f''(z) &= \frac{\alpha}{2}\left(\frac{\alpha}{2}+1\right) z^{-\alpha/2-2} \, \mathrm{J}(\sqrt{z}) - \frac{1}{2} \alpha z^{-\alpha/2-3/2} \, \mathrm{J}'(\sqrt{z}) \\ &\quad + \frac{1}{4} z^{-\alpha/2-1} \left(\mathrm{J}''(\sqrt{z}) - z^{-1/2} \, \mathrm{J}'(\sqrt{z}) \right), \end{aligned} \tag{66}$$

so

$$\mathrm{J}(\sqrt{z}) = z^{\alpha/2} f(z),$$

$$
\begin{aligned}
J'(\sqrt{z}) &= 2z^{\alpha/2+1/2}(f'(z) + \frac{\alpha}{2} z^{-\alpha/2-1}\, J(\sqrt{z})) \\
&= 2z^{\alpha/2+1/2}\left(f'(z) + \frac{1}{2}\alpha z^{-1} f(z)\right), \qquad (67)
\end{aligned}
$$

and

$$
\begin{aligned}
J''(\sqrt{z}) &= 4z^{\alpha/2+1}\left(f''(z) - \frac{\alpha}{2}\left(\frac{\alpha}{2}+1\right) z^{-\alpha/2-2}\, J(\sqrt{z}) + \frac{1}{2}\left(\alpha + \frac{1}{2}\right) z^{-\alpha/2-3/2}\, J'(\sqrt{z})\right) \\
&= 4z^{\alpha/2+1}\left(f''(z) - \frac{\alpha}{2}\left(\frac{\alpha}{2}+1\right) z^{-2} f(z) + \left(\alpha + \frac{1}{2}\right) z^{-1}\left(f'(z) + \frac{1}{2}\alpha z^{-1} f(z)\right)\right) \\
&= 4z^{\alpha/2+1}\left(f''(z) + \frac{1}{4}\alpha(\alpha-1) z^{-2} f(z) + \left(\alpha + \frac{1}{2}\right) z^{-1} f(z)\right) \qquad (68)
\end{aligned}
$$

Hence $f(z)$ satisfies (substituting into (65), and dividing throughout by $z^{\alpha/2}$)

$$
4z^2\left(y'' + \frac{1}{4}\alpha(\alpha-1)z^{-2}y + \left(\alpha+\frac{1}{2}\right)z^{-1}y\right) + 2z\left(y' + \frac{1}{2}\alpha z^{-1}y\right) + (z-\alpha^2)y = 0. \qquad (69)
$$

Cancellation, and a further division by z^2 gives

$$
4zf''(z) + 4(\alpha+1)f'(z) + f(z) = 0
$$

which is equivalent to the system (64). ■

Remark (on uniqueness of solutions)

Note that, for $\alpha > -1$,

$$
\psi(z) = 2^{\alpha} z^{-\alpha/2}\, J_{\alpha}\left(\sqrt{z}\right) = \sum_{j=0}^{\infty} \frac{(-1)^j \left(\frac{1}{4}z\right)^{\alpha+2j}}{j!\ \Gamma(j+\alpha+1)} \qquad (70)
$$

is an entire function, and that $y = \psi$ satisfies the differential equation

$$
4zy'' + 4(\alpha+1)y' + y = 0. \qquad (71)
$$

Further, when α is not an integer there is another independent solution

$$\begin{aligned}\phi(z) &= 2^{-\alpha} z^{-\alpha/2}\, \mathrm{J}_{-\alpha}(\sqrt{z}) \\ &= z^{-\alpha} \sum_{j=0}^{\infty} \frac{(-1)^j (\frac{1}{4}z)^j}{j!\, \Gamma(j-\alpha+1)}. \end{aligned} \tag{72}$$

Hence for $0 < \alpha < 1$, $\psi(z)$ gives the unique solution of (71) such that $y(0) = \psi(0) = 1/\Gamma(\alpha+1)$, since $\phi(z)$ is singular at $z = 0$. Further, when $\alpha = 0$, $\mathrm{J}_0(z)$ gives the unique solution of (71) such that $y(0) = 1$ since the Bessel function of the second kind Y_0 has a logarithmic singularity at $z = 0$ (see [19, p.171])

Proposition 2.5 *For the Jacobi polynomials $P_n = P_n^{\alpha,\beta}$, we have the limits*

$$(i) \quad n^{-\alpha} P_n \left(\cos \frac{\sqrt{z}}{n} \right) \to \frac{2^\alpha}{z^{\alpha/2}}\, \mathrm{J}_\alpha \left(\sqrt{z} \right) \tag{73}$$

and

$$(ii) \quad \frac{\sin \frac{\sqrt{z}}{n}}{2n\sqrt{z}} n^{-\alpha} P_n' \left(\cos \frac{\sqrt{z}}{n} \right) \to 2^{\alpha-1} z^{-\alpha/2} \left(\alpha z^{-1}\, \mathrm{J}_\alpha(\sqrt{z}) - z^{-1/2}\, \mathrm{J}_\alpha'(\sqrt{z}) \right), \tag{74}$$

as $n \to \infty$, which are uniform for z in compact subsets of $\mathbb{C} \setminus (-\infty, 0]$.

Proof. Note that *(ii)* follows from *(i)* by uniform convergence: we simply differentiate the expressions on both sides of the limit. Let

$$f_n(z) = n^{-\alpha} P_n \left(\cos \frac{\sqrt{z}}{n} \right).$$

Then f_n satisfies the differential equation (54). Lemma 2.3 tells us that

$$f(z) = 2^\alpha z^{-\alpha/2}\, \mathrm{J}_\alpha(\sqrt{z}) = \sum_{m=0}^{\infty} \frac{(-1)^m (\frac{1}{2})^{2m} z^m}{m! \Gamma(m+\alpha+1)}$$

is a solution of the system

$$z\frac{\mathrm{d}}{\mathrm{d}z}\begin{bmatrix} f(z) \\ f'(z) \end{bmatrix} = \Omega(z)\begin{bmatrix} f(z) \\ f'(z) \end{bmatrix},$$

where

$$\Omega(z) = \begin{bmatrix} 0 & z \\ -1/4 & -(\alpha+1) \end{bmatrix}.$$

We use the following fact about the normalisation of the Jacobi polynomials (see [41, p.58])

$$P_n^{\alpha,\beta}(1) = \frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)},$$

to get

$$f_n(0) = n^{-\alpha}P_n^{\alpha,\beta}(1) = n^{-\alpha}\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)}.$$

By considering the series expansion of $f(z)$, we get

$$f(0) = \frac{2^\alpha}{z^{\alpha/2}}\,\mathrm{J}\left(\sqrt{z}\right)\Big|_{z=0} = \frac{1}{\Gamma(\alpha+1)}.$$

By a result in [42, p.58], we have

$$\frac{\Gamma(n+\alpha)}{\Gamma(n)} \sim n^\alpha \text{ as } n \to \infty, \tag{75}$$

so

$$f_n(0) \to \frac{1}{\Gamma(\alpha+1)} = f(0) \text{ as } n \to \infty.$$

By differentiating the series expansion of $f(z)$, we get

$$f'(0) = \frac{-1}{4\Gamma(\alpha+1)}.$$

The following formula is a special case of one in [14, p.170]

$$\frac{\mathrm{d}}{\mathrm{d}x}P_n^{\alpha,\beta}(x) = \frac{1}{2}(n+\alpha+\beta+1)P_{n-1}^{\alpha+1,\beta+1}(x),$$

and we use it, and the limit (75), to get

$$\begin{aligned}
f_n'(0) &= -\frac{n^{-\alpha}}{2n^2}P_n'(1) \\
&= n^{-\alpha}\left(\frac{-1}{2n^2}\right)\frac{1}{2}(n+\alpha+\beta+1)\frac{\Gamma(n+\alpha+1)}{\Gamma(n)\Gamma(\alpha+2)} \\
&\sim -\frac{1}{4}n^{-(\alpha+1)}\frac{\Gamma(n+\alpha+1)}{\Gamma(n)\Gamma(\alpha+2)} \\
&\to \frac{-1}{4\Gamma(\alpha+2)} = f'(0).
\end{aligned}$$

We now have $f_n(z) \to f(z)$ as $n \to \infty$, on compact subsets of $\mathbb{C} \setminus (-\infty, 0]$, by Lemma 2.2. ■

2.5 Convergence of the Jacobi operator

For brevity, we shall write $x_n = \cos\left(\frac{\sqrt{x}}{n}\right)$.

Proposition 2.6 *The scaled operator $\tilde{K}_n : L^2[0,1] \to L^2[0,1]$ converges to the Bessel operator $K : L^2[0,1] \to L^2[0,1]$ weakly as $n \to \infty$, that is*

$$\langle \tilde{K}_n f, f\rangle \to \langle K f, f\rangle \ \text{ as } n \to \infty, \text{ for all } f \in L^2[0,1]$$

Proof. We use the relation (from [41, p.72])

$$2(n+\alpha)(n+\beta)P_{n-1}(z) = (2n+\alpha+\beta)(1-z^2)P_n'(z) + n\left((2n+\alpha+\beta)z+\beta-\alpha\right)P_n(z) \tag{76}$$

to get

$$\begin{aligned}
& P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y) \\
= \ & \frac{P_n(x)}{2(n+\alpha)(n+\beta)}\left((2n+\alpha+\beta)(1-y^2)P_n'(y) + n\left((2n+\alpha+\beta)y+\beta-\alpha\right)P_n(y)\right) \\
& -\frac{P_n(y)}{2(n+\alpha)(n+\beta)}\left((2n+\alpha+\beta)(1-x^2)P_n'(x) + n\left((2n+\alpha+\beta)x+\beta-\alpha\right)P_n(x)\right) \\
= \ & \frac{2n+\alpha+\beta}{2(n+\alpha)(n+\beta)}\left((1-y^2)P_n(x)P_n'(y) - (1-x^2)P_n(y)P_n'(x) - nP_n(x)P_n(y)\right),
\end{aligned} \tag{77}$$

and hence we have

$$
\begin{aligned}
&\frac{P_n(x_n)\,P_{n-1}(y_n) - P_n(y_n)\,P_{n-1}(x_n)}{x_n - y_n} \\
= &\left(\frac{2n+\alpha+\beta}{2(n+\alpha)(n+\beta)}\right)\frac{\sin^2\left(\frac{\sqrt{y}}{n}\right)P_n(x_n)\,P_n'(y_n) - \sin^2\left(\frac{\sqrt{y}}{n}\right)P_n(y_n)\,P_n'(x_n)}{x_n - y_n} \\
&-\frac{n(2n+\alpha+\beta)}{2(n+\alpha)(n+\beta)}P_n(x_n)\,P_n(y_n)\,. \qquad (78)
\end{aligned}
$$

Expanding some of the terms as Taylor series, and writing the constant terms as they appear in the limit, (78) becomes

$$
\begin{aligned}
&n\frac{\left(\frac{\sqrt{y}}{n}+O\left(\frac{1}{n^3}\right)\right)\sin\left(\frac{\sqrt{y}}{n}\right)P_n'(y_n)P_n(x_n) - \left(\frac{\sqrt{x}}{n}+O\left(\frac{1}{n^3}\right)\right)\sin\left(\frac{\sqrt{x}}{n}\right)P_n'(x_n)P_n(y_n)}{\frac{1}{2}(y-x)+O\left(\frac{1}{n^4}\right)} \\
&-P_n(x_n)\,P_n(y_n)\,.
\end{aligned}
$$

We do the same for the weight function:

$$
\begin{aligned}
w(x_n) &= 2^{-(\alpha+\beta+1)}\left(1-\cos\frac{\sqrt{x}}{n}\right)^{\alpha}\left(1+\cos\frac{\sqrt{x}}{n}\right)^{\beta} \\
&= \left(\frac{x}{n^2}+O\left(\frac{1}{n^4}\right)\right)^{\alpha}\left(2-\frac{x}{n^2}+O\left(\frac{1}{n^4}\right)\right)^{\beta}2^{-(\alpha+\beta+1)}.
\end{aligned}
$$

Using this information, and Proposition 2.5, we now have the following limit on the off-diagonal ($x \neq y$) for the rescaled kernel defined in (49)

$$
\begin{aligned}
&\tilde{K}_n(x,y) \\
\to\ &\frac{1}{y-x}\Big(y^{1-\alpha/2}x^{-\alpha/2}\left(\alpha y^{-1}\,\mathrm{J}_\alpha(\sqrt{y}) - y^{-1/2}\,\mathrm{J}_\alpha'(\sqrt{y})\right)\mathrm{J}_\alpha(\sqrt{x}) \\
&\qquad -x^{1-\alpha/2}y^{-\alpha/2}\left(\alpha x^{-1}\,\mathrm{J}_\alpha(\sqrt{x}) - x^{-1/2}\,\mathrm{J}_\alpha'(\sqrt{x})\right)\mathrm{J}_\alpha(\sqrt{y})\Big)(xy)^{\alpha/2} \\
=\ &\frac{\mathrm{J}_\alpha(\sqrt{x})\sqrt{y}\,\mathrm{J}_\alpha'(\sqrt{y}) - \mathrm{J}_\alpha(\sqrt{y})\sqrt{x}\,\mathrm{J}_\alpha'(\sqrt{x})}{(x-y)}
\end{aligned}
$$

as $n \to \infty$, uniformly on $[0,1]^2 \setminus \{(x,y) \in [0,1]^2 : x = y\}$. Notice that the contribution from the last term in (78) tends to zero under the n^{-2} scaling. For the diagonal ($x = y$), note that if A and B are once differentiable functions, and

$T_{A,B}(x,y) = \frac{1}{x-y}\left(A(x)B(y) - A(y)B(x)\right)$, then, by L'Hôpital's rule, $T_{A,B}(x,x) = A'(x)B(x) - A(x)B'(x)$. Hence

$$K_n(x,x) = (n+\alpha+\beta)\left(P_n'(x)P_{n-1}(x) - P_n(x)P_{n-1}'(x)\right)w(x). \tag{79}$$

We differentiate (76) to get an expression for P_{n-1}' in terms of P_n and its derivatives:

$$\begin{aligned} P_{n-1}'(x) &= \frac{1}{2(n+\alpha)(n+\beta)}\Big((2n+\alpha+\beta)\left((1-x^2)P_n''(x) - 2xP_n'(x)\right) \\ &\qquad +n((2n+\alpha+\beta)x+\beta-\alpha)P_n'(x) + n(2n+\alpha+\beta)P_n(x)\Big). \end{aligned}$$

Substituting this into (79), also using (76), and neglecting constants which will not appear in the limit, we have

$$\begin{aligned} K_n(x,x) &= w(x)\Big((1-x^2)P_n'(x)^2 - (1-x^2)P_n(x)P_n''(x) + 2xP_n(x)P_n'(x) - nP_n(x)^2\Big) \\ &= w(x)\Big((1-x^2)P_n'(x)^2 + \left(\beta(1-x) - \alpha(1+x)\right)P_n(x)P_n'(x) \\ &\qquad +n(n+\alpha+\beta)P_n(x)^2\Big). \end{aligned}$$

The second equality follows by (55), the differential equation satisfied by the Jacobi polynomials. Thus the diagonal of the rescaled kernel is given by

$$\tilde{K}_n(x,x) = \frac{2^{-\alpha}}{n^2}\left(f_1(x;n) + f_2(x;n,\alpha,\beta) + f_3(x;n,\alpha,\beta)\right)w\left(x_n\right) \tag{80}$$

where

$$\begin{aligned} f_1(x;n) &= \sin^2\left(\frac{\sqrt{x}}{n}\right)P_n'\left(x_n\right)^2, \\ f_2(x;n,\alpha,\beta) &= \left(\beta\left(1-x_n\right) - \alpha\left(1+x_n\right)\right)P_n\left(x_n\right)P_n'\left(x_n\right), \\ f_3(x;n,\alpha,\beta) &= n(n+\alpha+\beta)P_n\left(x_n\right)^2. \end{aligned}$$

As before, the scaled weight $w(x_n)$ contributes $n^{-2\alpha}$, which effects the convergence of all the terms, and also x^α and the constant factor $2^{-\alpha-1}$. We use Proposition

2.5 *(i)* to get

$$\frac{1}{n^2} f_3(x;n,\alpha,\beta) w(x_n) \to 2^{\alpha-1} \mathrm{J}_\alpha(\sqrt{x})^2,$$

and Proposition 2.5 *(ii)* gives

$$\begin{aligned}
\frac{1}{n^2} f_1(x;n) w(x_n) &\to 2^{\alpha-1} x^{-1} \left(\alpha x^{-1} \mathrm{J}_\alpha(\sqrt{x}) - x^{-1/2} \mathrm{J}'_\alpha(\sqrt{x}) \right)^2 \\
&= 2^{\alpha-1} \left(\alpha^2 x^{-1} \mathrm{J}_\alpha(\sqrt{x})^2 - 2\alpha x^{-1/2} \mathrm{J}_\alpha(\sqrt{x}) \mathrm{J}'_\alpha(\sqrt{x}) + \mathrm{J}'_\alpha(\sqrt{x})^2 \right).
\end{aligned}$$

For f_2, note first that

$$\frac{\sin \frac{\sqrt{x}}{n}}{n\sqrt{x}} n^{-\alpha} P'_n\left(\cos\frac{\sqrt{x}}{n}\right) = \frac{\sin \frac{\sqrt{x}}{n}}{\frac{\sqrt{x}}{n}} \frac{1}{n^2} n^{-\alpha} P'_n\left(\cos\frac{\sqrt{x}}{n}\right),$$

where the left-hand side is known to converge by Proposition 2.5 *(ii)*, and the first factor on the right-hand side tends to 1 as $n \to \infty$. Since

$$\begin{aligned}
f_2(x;n,\alpha,\beta) &= \left(\beta \left(\frac{x}{n^2} + O\left(\frac{1}{n^4}\right) \right) - \alpha \left(2 - \frac{x}{n^2} + O\left(\frac{1}{n^4}\right) \right) \right) \\
&\quad \times P_n\left(\cos\frac{\sqrt{x}}{n}\right) P'_n\left(\cos\frac{\sqrt{x}}{n}\right),
\end{aligned}$$

we have

$$\frac{1}{n^2} f_2(x;n,\alpha,\beta) w(x_n) \to -2\alpha 2^{\alpha-1} \left(\alpha x^{-1} \mathrm{J}_\alpha(\sqrt{x})^2 - x^{-1/2} \mathrm{J}_\alpha(\sqrt{x}) \mathrm{J}'_\alpha(\sqrt{x}) \right) \text{ as } n \to \infty,$$

and adding the above limits together we get

$$\tilde{K}_n(x,x) \to \frac{1}{2} \left(\left(1 - \frac{\alpha^2}{x} \right) \mathrm{J}_\alpha(\sqrt{x})^2 + \mathrm{J}'_\alpha(\sqrt{x})^2 \right),$$

uniformly for $x \in [0,1]$. Applying L'Hôpital's rule again to the Bessel kernel, we find

$$\begin{aligned}
K(x,x) &= \frac{1}{2} \left(\mathrm{J}'_\alpha(\sqrt{x})^2 - \mathrm{J}_\alpha(\sqrt{x}) \left(\mathrm{J}''_\alpha(\sqrt{x}) + \frac{1}{\sqrt{x}} \mathrm{J}'_\alpha(\sqrt{x}) \right) \right) \\
&= \frac{1}{2} \left(\left(1 - \frac{\alpha^2}{x} \right) \mathrm{J}_\alpha(\sqrt{x})^2 + \mathrm{J}'_\alpha(\sqrt{x})^2 \right),
\end{aligned}$$

where we use equation (65) to write the second derivative of $\mathrm{J}_\alpha(\sqrt{x})$ as a sum of lower order terms. Hence we have shown that $\tilde{K}_n(x,x) \to K(x,x)$, uniformly for $x \in [0,1]$. It is now clear that

$$\begin{aligned} \langle \tilde{K}_n f, f\rangle = \int_0^1 \int_0^1 \tilde{K}_n(x,y) f(x) f(y)\,\mathrm{d}y\,\mathrm{d}x \quad &\to \quad \int_0^1 \int_0^1 K(x,y) f(x) f(y)\,\mathrm{d}y\,\mathrm{d}x \\ &= \quad \langle K f, f\rangle. \end{aligned}$$

■

Theorem 2.7 *The scaled operator $\tilde{K}_n : L^2[0,1] \to L^2[0,1]$ converges to the Bessel operator $K : L^2[0,1] \to L^2[0,1]$ in trace class norm as $n \to \infty$.*

Proof. We have already shown that $\tilde{K}_n(x,x) \to K(x,x)$, uniformly for $x \in [0,1]$, and hence

$$\int_0^1 K_n(x,x)\,\mathrm{d}x \to \int_0^1 K(x,x)\,\mathrm{d}x \quad \text{as } n \to \infty.$$

We must show that this limit implies convergence of traces. $\tilde{K}_n$ is clearly non-negative, since it is the orthogonal projection onto the subspace

$$\overline{\text{span}} \left\{ P_k\left(\cos\frac{\sqrt{x}}{n}\right), k = 0, \ldots, n \right\},$$

of $L^2[0,1]$, and by weak convergence (Proposition 2.6), K is also non-negative. Further, both kernels $\tilde{K}_n$ and K are real symmetric and continuous, and hence Mercer's theorem (Theorem 1.12) gives

$$\operatorname{trace} \tilde{K}_n = \int_0^1 \tilde{K}_n(x,x)\,\mathrm{d}x$$

and

$$\operatorname{trace} K = \int_0^1 K(x,x)\,\mathrm{d}x,$$

and so $\operatorname{trace} \tilde{K}_n \to \operatorname{trace} K$ as $n \to \infty$. Since $\tilde{K}_n$ and K are positive and self-adjoint, this is equivalent to

$$\|\tilde{K}_n\|_{C_1} \to \|K\|_{C_1}$$

and hence, together with the weak convergence proved in Proposition 2.6, Theorem 1.14 gives the required trace norm convergence. ■

The previous result is useful when proving convergence for level spacing distributions in the Jacobi kernel. Determinants carry the probabilistic information about eigenvalue distributions of random matrices. We use the following result to deduce convergence of determinants from trace norm convergence. For a proof, see [24, p.342].

Lemma 2.8 *Let T be a trace class operator, and T_N a sequence of operators converging to T in trace class norm as $N \to \infty$. Then $\det(I - T_N) \to \det(I - T)$ as $N \to \infty$.*

Corollary 2.9 *For the operators $\tilde{K}_n$ and K above, we have*

$$\det(I - \mathbb{I}_{(a,b)}\tilde{K}_n\mathbb{I}_{(a,b)}) \to \det(I - \mathbb{I}_{(a,b)}K\mathbb{I}_{(a,b)}) \textit{ as } n \to \infty$$

Proof. Since multiplication by $\mathbb{I}_{(a,b)}$ is bounded, we have, by Theorem 2.7

$$\mathbb{I}_{(a,b)}\tilde{K}_n\mathbb{I}_{(a,b)} \to \mathbb{I}_{(a,b)}K\mathbb{I}_{(a,b)}$$

in trace norm, as $n \to \infty$, and we use Lemma 2.8 to deduce the result. ■

3 Tracy–Widom operators not commuting with a differential operator

3.1 Introduction

In several important examples (see, e.g. [27, pp.98-101] and [45, §III B]), the eigenvectors of a Tracy–Widom integral operator K can be found by using the fact (a proof of which is given below) that a differential operator L satisfying $KL = LK$ has the same eigenvectors as K. In this section, we consider the possibility of finding a self-adjoint differential operator which commutes with a Tracy–Widom integral operator K (with kernel $K(x, y)$) on $L^2(\mathbb{R})$. We observe that a TW integral operator can be written using commutators, via the multiplication and Hilbert transform operators, and show that the Hilbert transform commutes with differentiation by expressing both these operators in their Fourier transform state. Using these facts, and some commutator formulae, we expand the commutator of K and a typical self-adjoint differential operator, and find that for this to be zero, K must also be zero. We conclude that a commuting self-adjoint differential operator cannot be found for any non-zero TW integral operator acting on the whole real line.

3.2 Self-adjointness for differential operators

Differential operators are unbounded, so we need to be careful when discussing what is meant by self-adjointness, so that spectral theory results can be applied (see [12, p.7,§1.2]). To begin with though, we need to define clearly what we mean by an operator in the unbounded context. A densely defined linear operator on a Hilbert space H is a pair comprising a dense linear subspace, which we call $\mathrm{Dom}(A)$ (the domain) and a linear map $A : \mathrm{Dom}(A) \to H$. If E is a linear subspace of H which contains $\mathrm{Dom}(A)$ and $\tilde{A}$ is a map $E :\to H$ which satisfies $\tilde{A}f = Af$ for all $f \in$Dom(A), then we say that $\tilde{A}$ is an *extension* of A. The adjoint A^* of an operator A satisfies the condition

$$\langle Ax, y\rangle = \langle x, A^*y\rangle$$

for all $x \in$Dom(A) and $y \in$ Dom(A^*), where we define Dom(A^*) to be the set of all $y \in H$ for which there exists $z \in H$ with

$$\langle Ax, y\rangle = \langle x, z\rangle$$

for all $x \in$Dom(A). Thus, A is self-adjoint if and only if it is symmetric and

$$\text{Dom}(A) = \text{Dom}(A^*).$$

In practice, we usually define a differential operator on some dense subspace of the Hilbert space H on which it acts, and use a result known as Friedrich's theorem to show that there is a self-adjoint extension. We state this in the form in which we need it (for a proof, see [24, p.402-404], which also shows how to construct the extension).

Theorem 3.1 (Friedrich's theorem) *Let L be an operator defined on a dense subspace $\mathcal{D}$ of a Hilbert space H. If the following conditions are satisfied:*
(i) $L(\mathcal{D}) \subseteq \mathcal{D}$
(ii) $\langle Lu, v\rangle = \langle u, Lv\rangle$ *for all* $u, v \in \mathcal{D}$
(iii) $\langle Lu, u\rangle \geq 0$ *for all* $u \in \mathcal{D}$,
then L has a self-adjoint extension.

3.3 Definitions and motivation

Let I be a (possibly infinite) interval of the real line. We consider integral operators $K_{A,B}$ with kernel $K_{A,B}(x,y)$ of Tracy-Widom type

$$K_{A,B}(x,y) = \frac{A(x)B(y) - A(y)B(x)}{x-y},$$

which act on a function $f \in L^2(I)$ in the usual way:

$$K_{A,B}f(x) = \int_I K_{A,B}(x,y)f(y)\,\mathrm{d}y.$$

When the dependence on the functions A and B is clear, we shall omit the subscripts. In several important examples (e.g. for the Bessel kernel [45, §III B], the sine kernel [27, pp.96-101] and the Airy kernel [43, §IV]) the eigenvectors and eigenvalues of an operator of this form can be found via a commuting self-adjoint differential operator. By this, we mean a differential operator L on some suitable space of functions which satisfies the condition

$$L_x \int_I K(x,y)f(y)\,\mathrm{d}y = \int_I K(x,y)L_y f(y)\,\mathrm{d}y,$$

for all x, in which L_x means that L acts on the x variable and so on. The following general theorem on commuting operators means that if we can find such a differential operator, its eigenvectors will be the same as those of the Tracy–Widom operator K.

Proposition 3.2 *Let A and B be compact self-adjoint operators on a Hilbert space H, and suppose that $AB = BA$. Then there exists an orthonormal basis (ϕ_j) of H such that $A\phi_j = \lambda_j\phi_j$ and $B\phi_j = \mu_j\phi_j$, for some scalars λ_j and μ_j.*

Proof. Take $0 \neq \lambda \in \sigma_p(A)$, let $E_\lambda = \{x \in H : Ax = \lambda x\}$, so that E_λ is finite-dimensional, and let P_λ be the projection from H onto E_λ. By Cauchy's theorem, we have

$$P_\lambda = \frac{1}{2\pi i}\int_{C(\lambda,\delta)} \frac{\mathrm{d}z}{z-A}.$$

Note that, since A is compact, and hence bounded, we have, for $|z| > \|A\|$

$$\begin{aligned}
B(z-A)^{-1} &= \frac{B}{z}\sum_{n=0}^{\infty}\left(\frac{A}{z}\right) \\
&= \frac{1}{z}\sum_{n=0}^{\infty}\left(\frac{A}{z}\right)^n B \\
&= (z-A)^{-1}B
\end{aligned}$$

where we also use the fact that $AB = BA$, so that B commutes with A^n. This identity can be extended to hold for all z in the complement of $\sigma(A)$ (its resolvent set) by analytic continuation. It is now clear that

$$BP_\lambda = \frac{1}{2\pi i}\int_{C(\lambda,\delta)} B\frac{\mathrm{d}z}{z-A} = \frac{1}{2\pi i}\int_{C(\lambda,\delta)} \frac{\mathrm{d}z}{z-A}B = P_\lambda B.$$

Hence BP_λ is self-adjoint, since $(BP_\lambda)^* = P_\lambda B = BP_\lambda$. Let (ϕ_j) be an orthonormal basis for E_λ consisting of eigenvectors of BP_λ. Then clearly the ϕ_j are also eigenvectors of A. We can repeat the argument with A taking the place of B, and thus we see that A and B have a common basis of eigenvectors. ■

Example 3.3

The Airy kernel arises when we scale the eigenvalues of a random matrix at the soft edge of the spectrum in the Gaussian Unitary Ensemble (see [43], or §1.9 in this thesis). It can be written as the square of the Hankel operator with kernel $\mathrm{Ai}\,(x+y)$:

$$\frac{\mathrm{Ai}\,(x)\,\mathrm{Ai}'(y) - \mathrm{Ai}'(x)\,\mathrm{Ai}\,(y)}{x-y} = \int_0^\infty \mathrm{Ai}\,(x+t)\,\mathrm{Ai}\,(y+t)\,\mathrm{d}t, \qquad (81)$$

Let L be a differential operator defined on the space $\mathcal{D} = C_c^\infty(\mathbb{R}_+)$ of smooth functions on $\mathbb{R}_+$ with compact support by

$$L = -\frac{\mathrm{d}}{\mathrm{d}x}\left(x\frac{\mathrm{d}}{\mathrm{d}x}\right) + x^2$$

Note that L is symmetric. To see this, take $f, g \in \mathcal{D}$. Then integration by parts, and the fact that $f(x), g(x) \to 0$ as $x \to \infty$ together give

$$\begin{aligned}
\langle Lf, g\rangle &= \int_0^\infty \left(-(xf'(x))' + x^2 f(x)\right) g(x)\,\mathrm{d}x \\
&= [-xf'(x)g(x)]_0^\infty + \int_0^\infty xf'(x)g'(x)\,\mathrm{d}x + \int_0^\infty x^2 f(x)g(x)\,\mathrm{d}x \\
&= \int_0^\infty xf'(x)g'(x)\,\mathrm{d}x + \int_0^\infty x^2 f(x)g(x)\,\mathrm{d}x,
\end{aligned}$$

while

$$\begin{aligned}
\langle f, Lg\rangle &= \int_0^\infty f(x)\left(-(xg'(x))' + x^2 g(x)\right)\,\mathrm{d}x \\
&= [-xf(x)g'(x)]_0^\infty + \int_0^\infty xf'(x)g'(x)\,\mathrm{d}x + \int_0^\infty x^2 f(x)g(x)\,\mathrm{d}x \\
&= \int_0^\infty xf'(x)g'(x)\,\mathrm{d}x + \int_0^\infty x^2 f(x)g(x)\,\mathrm{d}x \\
&= \langle Lf, g\rangle.
\end{aligned}$$

A similar argument shows that L is a positive operator, and hence L has a self-adjoint extension, by Friedrich's theorem (Theorem 3.1) Also, L commutes with the Hankel integral operator Γ with kernel $\mathrm{Ai}\,(x+y)$. Recall that $\mathrm{Ai}\,''(x) = x\,\mathrm{Ai}\,(x)$, and then

$$\begin{aligned}
L\Gamma f(x) &= L\int_0^\infty \mathrm{Ai}\,(x+y)f(y)\,\mathrm{d}y \\
&= \int_0^\infty \left(-\left(x\,\mathrm{Ai}\,'(x+y)\right)' + x^2\,\mathrm{Ai}\,(x+y)\right) f(y)\,\mathrm{d}y \\
&= \int_0^\infty \left(-x(x+y)\,\mathrm{Ai}\,(x+y) - \mathrm{Ai}\,'(x+y) + x^2\,\mathrm{Ai}\,(x+y)\right) f(y)\,\mathrm{d}y \\
&= \int_0^\infty (-xy\,\mathrm{Ai}\,(x+y) - \mathrm{Ai}\,'(x+y))\, f(y)\,\mathrm{d}y,
\end{aligned}$$

while integration by parts gives

$$\begin{aligned}
\Gamma L f(x) &= \int_0^\infty \mathrm{Ai}\,(x+y)\left(-(yf'(y))' + y^2 f(y)\right)\,\mathrm{d}y \\
&= [\,\mathrm{Ai}\,(x+y) y f'(y)]_0^\infty + \int_0^\infty \mathrm{Ai}\,'(x+y) y f'(y)\,\mathrm{d}y \\
&\quad + \int_0^\infty \mathrm{Ai}\,(x+y) y^2 f(x)\,\mathrm{d}y \\
&= \int_0^\infty \mathrm{Ai}\,'(x+y) y f'(y)\,\mathrm{d}y + \int_0^\infty \mathrm{Ai}\,(x+y) y^2 f(y)\,\mathrm{d}y \\
&= [\,\mathrm{Ai}\,'(x+y) y f(y)]_0^\infty - \int_0^\infty (\,\mathrm{Ai}\,'(x+y) + y\,\mathrm{Ai}\,''(x+y))\, f(y)\,\mathrm{d}y \\
&\quad + \int_0^\infty \mathrm{Ai}\,(x+y) y^2 f(y)\,\mathrm{d}y \\
&= -\int_0^\infty \mathrm{Ai}\,'(x+y) f(y)\,\mathrm{d}y - \int_0^\infty y(x+y)\,\mathrm{Ai}\,(x+y) f(y)\,\mathrm{d}y \\
&\quad + \int_0^\infty y^2\,\mathrm{Ai}\,(x+y) f(y)\,\mathrm{d}y \\
&= L\Gamma f(x)
\end{aligned}$$

where we used the fact that $\mathrm{Ai}\,(x) \to 0$ as $x \to \infty$ to show that the boundary terms are zero. In the light of (81), it is then clear that L commutes with the Airy Tracy–Widom operator on the half line

In contrast to the above example, the main result in this chapter tells us that there are no non-zero Tracy–Widom operators which commute with a self-adjoint differential operator on the whole real line. The first step is to express the TW operator as the sum of two terms involving commutators with the Hilbert transform. We show that the Hilbert transform commutes with differentiation on $L^2(\mathbb{R})$, and then, using a number of commutator identities, expand the commutator of the TW operator and differential operator.

3.4 Results on commutators

Definition 3.4 *Let R and S be operators, at least one of which is bounded. The* commutator *of R and S is*

$$[R, S] = RS - SR.$$

Clearly, the commutator of two operators is zero if and only if they commute.

We shall need the well-known identities in the next result to calculate the commutator of $K_{A,B}$ with a differential operator.

Lemma 3.5 *For operators R, S and T, of which at least two are bounded, the following identities hold:*

$$[R, [S, T]] + [S, [T, R]] + [T, [R, S]] = 0 \quad \textit{(Jacobi's identity)} \tag{82}$$

$$[RS, T] = [R, T]\,S + R\,[S, T] \tag{83}$$

$$[R, ST] = [R, S]\,T + S\,[R, T]\,. \tag{84}$$

Proof. Clear on calculation of the commutators in question. ∎

Remark

The Jacobi identity is important in some branches of algebra, particularly in the theory of Lie groups. See, for example [38, p.10].

We shall sometimes consider integrals in which the integrand diverges at one point. In such cases, we invoke the Cauchy Principal Value (PV) (see [3, p.238]). Suppose that the integrand $f(x)$ diverges at $x = x_0$. Then we interpret the integral as follows:

$$PV \int_a^b f(x)\,\mathrm{d}x = \lim_{\epsilon \to 0} \left(\int_a^{x_0 - \epsilon} + \int_{x_0 + \epsilon}^b f(x)\,\mathrm{d}x \right).$$

Definition 3.6 *The Hilbert transform is the integral operator on $L^2(\mathbb{R})$ with kernel $1/(x-y)$. Thus*

$$Hf(x) = PV \int_{-\infty}^{\infty} \frac{f(y)}{x-y}\,\mathrm{d}y \quad (f \in L^2(\mathbb{R})).$$

Note that, for later convenience, this definition lacks the usual factor of $1/\pi$. The calculations in the following result are standard (see, for instance [42, p.103]), but we include them here for completeness. We define the Fourier Transform on the real line to be

$$\mathcal{F}(f;\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ix\xi}\,\mathrm{d}x.$$

Lemma 3.7 *For the Hilbert Transform H, and Fourier transform $\mathcal{F}$, we have*

$$H = \mathcal{F}^* M_\phi \mathcal{F},$$

where $\phi(\xi) = -i\pi\,\mathrm{sgn}\,(\xi)$. Moreover, H is a bounded operator, and it commutes with differentiation on the space $C_c^\infty(\mathbb{R})$ of smooth functions with compact support in $\mathbb{R}$.

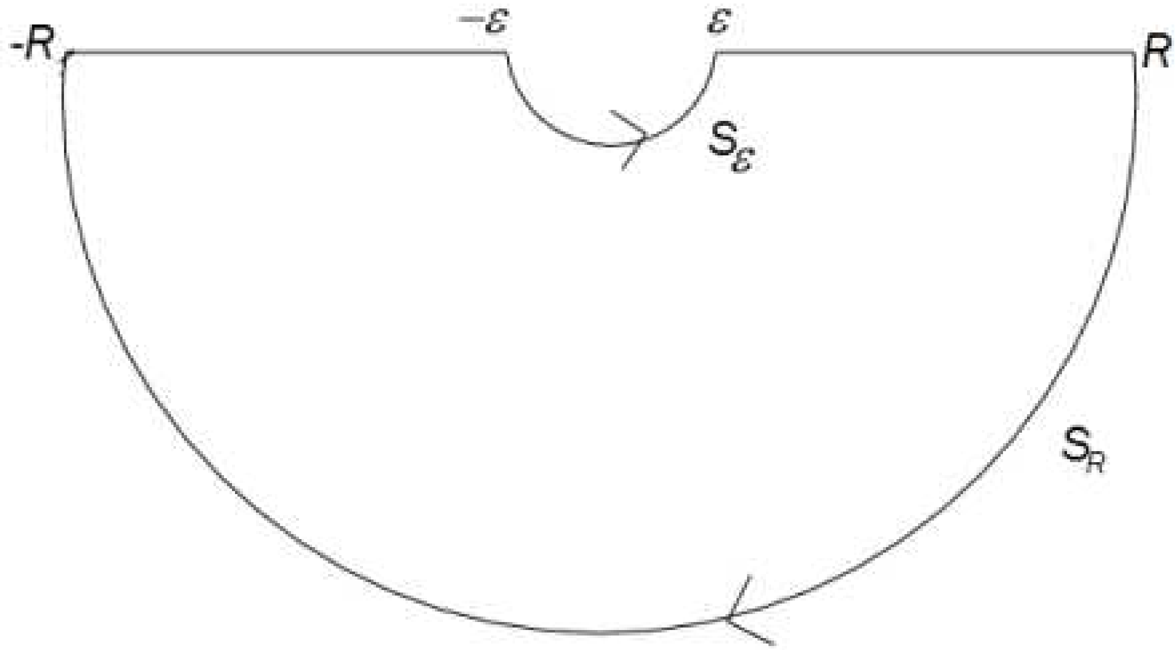
Remark

Note that the Hilbert transform does not commute with differentiation on the half line $[0,\infty)$, since it can be shown that H does not map $C_c^\infty([0,\infty))$ to itself.

Proof. Note that $Hf(x)$ is the convolution of the functions $\frac{1}{x}$ and $f(x)$, so that

$$\begin{aligned} \mathcal{F}(Hf;\xi) &= \sqrt{2\pi}\,\mathcal{F}\left(\frac{1}{x};\xi\right)\mathcal{F}(f;\xi) \\ &= \mathcal{F}(f;\xi) PV \int_{-\infty}^{\infty} \frac{e^{-ix\xi}}{x}\,\mathrm{d}x. \end{aligned}$$

We calculate the principal value integral by contour integration as follows. Assume to begin with that $\xi > 0$. Let Γ be the contour in the lower half-plane, shown in the diagram below and taken once in the negative sense.



Since $e^{-ix\xi}/x$ is holomorphic inside and on this contour, we have

$$
\begin{aligned}
0 &= \int_\Gamma \frac{e^{-ix\xi}}{x}\,\mathrm{d}x \\
&= \int_{-R}^{-\epsilon} + \int_{\epsilon}^{R} \frac{e^{-ix\xi}}{x}\,\mathrm{d}x + \int_{S_\epsilon} \frac{e^{-ix\xi}}{x}\,\mathrm{d}x + \int_0^\pi \exp\left(-i\xi(R\cos\theta - iR\sin\theta)\right) i\,\mathrm{d}\theta.
\end{aligned}
$$

The integral of $e^{-ix\xi}/x$ around a small circle centred at 0 is $2\pi i$, so the integral around the semicircle S_ϵ is πi. For the last integral, note that

$$|\exp\left(-i\xi(R\cos\theta - iR\sin\theta) + i\theta\right) i| = \exp\left(-\xi R\sin\theta\right)$$

and, by symmetry

$$\int_0^\pi e^{-\xi R\sin\theta}\,\mathrm{d}\theta = 2\int_0^{\frac{\pi}{2}} e^{-\xi R\sin\theta}\,\mathrm{d}\theta.$$

We use the bound $\sin\theta \geq \frac{2}{\pi}\theta \quad (0 \leq \theta \leq \frac{\pi}{2})$ to get

$$
\begin{aligned}
\left|\int_{S_R} \frac{e^{i\xi x}}{x}\,\mathrm{d}x\right| &\leq 2\int_0^{\frac{\pi}{2}} e^{-\xi R\frac{2}{\pi}\theta}\,\mathrm{d}\theta \\
&= \frac{2}{\xi\,\frac{\pi}{2}R}(1 - e^{-\xi R}) \to 0 \quad \text{as} \quad R \to \infty.
\end{aligned}
$$

We are therefore left with

$$\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty} \frac{e^{-ix\xi}}{x}\,\mathrm{d}x = -i\pi \text{ (for } \xi > 0).$$

The case of $\xi < 0$ can be dealt with by considering the reflection of the contour Γ in the upper half-plane. This leads to

$$\int_{-\infty}^{-\epsilon} + \int_{\epsilon}^{\infty} \frac{e^{-ix\xi}}{x}\,\mathrm{d}x = i\pi \text{ (for } \xi < 0).$$

We can combine these two statements together to write

$$\mathcal{F}(Hf;\xi) = -i\pi\,\mathrm{sgn}\,(\xi)\mathcal{F}(f;\xi),$$

or in other words,

$$H = \mathcal{F}^* M_\phi \mathcal{F},$$

where $\phi(x) = -i\pi\,\mathrm{sgn}\,(\xi)$, since the Fourier transform is a unitary operator. This shows that H is bounded. Take $f \in C_c^\infty(\mathbb{R})$. We observe that

$$\begin{aligned} \int_{-\infty}^{\infty} f'(x)e^{-ix\xi}\,\mathrm{d}x &= \left[f(x)e^{-ix\xi}\right]_{-\infty}^{\infty} + i\xi \int_{-\infty}^{\infty} f(x)e^{-ix\xi}\,\mathrm{d}x \\ &= i\xi\mathcal{F}(f(\xi)), \end{aligned}$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}x} = \mathcal{F}^* M_\psi \mathcal{F},$$

where $\psi(\xi) = i\xi$. We have shown that the two operators H and $\frac{\mathrm{d}}{\mathrm{d}x}$ are unitarily equivalent to multiplication operators on the space $C_c^\infty(\mathbb{R})$, so they commute there. ■

Lemma 3.8 *Let A and B be bounded and measurable functions $\mathbb{R} \to \mathbb{C}$. The*

Tracy-Widom integral operator $K_{A,B}$ on $L^2(\mathbb{R})$ is bounded, and satisfies

$$K_{A,B} = M_A\left[H, M_B\right] + M_B\left[M_A, H\right].$$

Proof. M_A, M_B and H are all bounded, so for a function $f \in L^2(\mathbb{R})$, the expression

$$\left(M_A\left[H, M_B\right] + M_B\left[M_A, H\right]\right) f(x)$$

makes sense. Also, we may write

$$\begin{aligned}
[H, M_B]f(x) &= \int_{-\infty}^{\infty} \frac{B(y) - B(x)}{x - y} f(y)\, \mathrm{d}y \\
&= \int_{-\infty}^{\infty} \frac{B(y)}{x - y} f(y)\, \mathrm{d}y - B(x) \int_{-\infty}^{\infty} \frac{1}{x - y} f(y)\, \mathrm{d}y.
\end{aligned}$$

Note that

$$\begin{aligned}
\left(M_A\left[H, M_B\right] + M_B\left[M_A, H\right]\right) f(x) &= A(x) \int_{-\infty}^{\infty} \frac{B(y) - B(x)}{x - y} f(y)\, \mathrm{d}y \\
&\quad + B(x) \int_{-\infty}^{\infty} \frac{A(x) - A(y)}{x - y} f(y)\, \mathrm{d}y. \\
&= \int_{-\infty}^{\infty} \frac{A(x)B(y) - A(y)B(x)}{x - y} f(y)\, \mathrm{d}y
\end{aligned}$$

■

Lemma 3.9 *For a differentiable function ϕ, we have*

$$\left[\frac{\mathrm{d}}{\mathrm{d}x}, M_\phi\right] = M_{\phi'}. \tag{85}$$

Proof. Observe that, by Leibniz's rule,

$$\begin{aligned}
\left(\frac{\mathrm{d}}{\mathrm{d}x} M_\phi - M_\phi \frac{\mathrm{d}}{\mathrm{d}x}\right) f(x) &= \frac{\mathrm{d}}{\mathrm{d}x} \phi(x) f(x) - \phi(x) f'(x) \\
&= \phi'(x) f(x) + \phi(x) f'(x) - \phi(x) f'(x) \\
&= \phi'(x) f(x).
\end{aligned}$$

■

3.5 Tracy–Widom operators do not commute with self-adjoint differential operators on the real line

Theorem 3.10 *Suppose that α is a strictly increasing and differentiable real function, v is a continuous real function, and that A and B are continuously differentiable functions that are bounded $\mathbb{R} \to \mathbb{C}$. Then $K_{A,B}$ is bounded, and is self-adjoint when A and B are real-valued. Further, if the operator*

$$L = -\alpha \frac{\mathrm{d}^2}{\mathrm{d}x^2} - \alpha' \frac{\mathrm{d}}{\mathrm{d}x} + v$$

commutes with $K_{A,B}$ on $C_c^\infty(\mathbb{R})$, then A and B are proportional, and hence $K_{A,B} = 0$.

Proof. Throughout this proof, we define all operators on the space $C_c^\infty(\mathbb{R})$. Suppose A and B are real-valued. Then it is clear that $K(x, y) = \overline{K(y, x)}$ so that K is self-adjoint in this case. By Lemma 3.8,

$$K_{A,B} = M_A \left[H, M_B\right] + M_B \left[M_A, H\right],$$

where all the operators are bounded. To prove the last part of the result, we calculate part of the commutator of L and $K_{A,B}$. In fact, it turns out that considering only the second order terms will be sufficient. Using Lemma 3.8, we find:

$$\left[-M_\alpha \frac{\mathrm{d}}{\mathrm{d}x} \frac{\mathrm{d}}{\mathrm{d}x}, K_{A,B}\right] = \left[-M_\alpha \frac{\mathrm{d}}{\mathrm{d}x} \frac{\mathrm{d}}{\mathrm{d}x}, M_A \left[H, M_B\right]\right] - \left[M_\alpha \frac{\mathrm{d}}{\mathrm{d}x} \frac{\mathrm{d}}{\mathrm{d}x}, M_B \left[M_A, H\right]\right]$$

We examine both the terms on the right-hand side in turn:

$$\left[-M_\alpha \frac{\mathrm{d}}{\mathrm{d}x} \frac{\mathrm{d}}{\mathrm{d}x}, M_A \left[H, M_B\right]\right]$$

$$
\begin{aligned}
&= \left[\left(M_\alpha \frac{\mathrm{d}}{\mathrm{d}x}\right) \frac{\mathrm{d}}{\mathrm{d}x}, M_A\left[M_B, H\right]\right] \\
&= \left[M_\alpha \frac{\mathrm{d}}{\mathrm{d}x}, M_A\left[M_B, H\right]\right] \frac{\mathrm{d}}{\mathrm{d}x} + M_\alpha \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\mathrm{d}}{\mathrm{d}x}, M_A\left[M_B, H\right]\right] \\
&= \left(\left[M_\alpha, M_A\left[M_B, H\right]\right] \frac{\mathrm{d}}{\mathrm{d}x} + M_\alpha\left[\frac{\mathrm{d}}{\mathrm{d}x}, M_A\left[M_B, H\right]\right]\right) \frac{\mathrm{d}}{\mathrm{d}x} \\
&\quad + M_\alpha \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{\mathrm{d}}{\mathrm{d}x}, M_A\left[M_B, H\right]\right],
\end{aligned}
$$

where the third and fourth lines follow from Lemma 3.5, formula (83). We find

$$
\begin{aligned}
&\left[\frac{\mathrm{d}}{\mathrm{d}x}, M_A\left[M_B, H\right]\right] \\
= \;& \left[\frac{\mathrm{d}}{\mathrm{d}x}, M_A\right]\left[M_B, H\right] + M_A\left[\frac{\mathrm{d}}{\mathrm{d}x}, \left[M_B, H\right]\right] \quad \text{(by (84))} \\
= \;& M_{A'}\left[M_B, H\right] + M_A\left(\left[M_B, \left[\frac{\mathrm{d}}{\mathrm{d}x}, H\right]\right] + \left[M_{B'}, H\right]\right) \quad \text{(by (82) and Lemma 3.9)} \\
= \;& M_{A'}\left[M_B, H\right] + M_A\left[M_{B'}, H\right] \qquad (86)
\end{aligned}
$$

in which we have used the fact that H and d/dx commute (Lemma 3.7). We now have

$$
\begin{aligned}
&\left[-M_\alpha \frac{\mathrm{d}}{\mathrm{d}x} \frac{\mathrm{d}}{\mathrm{d}x}, M_A\left[H, M_B\right]\right] \\
= \;& \left(\left[M_\alpha, M_A\left[M_B, H\right]\right] \frac{\mathrm{d}}{\mathrm{d}x} + M_\alpha\left(M_{A'}\left[M_B, H\right] + M_A\left[M_{B'}, H\right]\right)\right) \frac{\mathrm{d}}{\mathrm{d}x} \\
& + M_\alpha \frac{\mathrm{d}}{\mathrm{d}x}\left(M_{A'}\left[M_B, H\right] + M_A\left[M_{B'}, H\right]\right).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&\left[-M_\alpha \frac{\mathrm{d}}{\mathrm{d}x} \frac{\mathrm{d}}{\mathrm{d}x}, M_B\left[M_A, H\right]\right] \\
= \;& -\left(\left[M_\alpha, M_B\left[M_A, H\right]\right] \frac{\mathrm{d}}{\mathrm{d}x} + M_\alpha\left(M_{B'}\left[M_A, H\right] + M_B\left[M_{A'}, H\right]\right)\right) \frac{\mathrm{d}}{\mathrm{d}x} \\
& - M_\alpha \frac{\mathrm{d}}{\mathrm{d}x}\left(M_{B'}\left[M_A, H\right] + M_B\left[M_{A'}, H\right]\right).
\end{aligned}
$$

Second order terms in the commutator result only from the action of second order terms in L, so we see that the second order terms in $[L, K_{A,B}]$ must be

$$\begin{aligned}
& [M_\alpha, M_A [M_B, H]] - [M_\alpha, M_B [M_A, H]] \\
= \;& [M_\alpha, M_A] [M_B, H] + M_A [M_\alpha, [M_B, H]] \\
- \;& [M_\alpha, M_B] [M_A, H] - M_B [M_\alpha, [M_A, H]] \text{ (by (84))} \\
= \;& M_A [M_\alpha, [M_B, H]] - M_B [M_\alpha, [M_A, H]] .
\end{aligned}$$

Hence, if L and $K_{A,B}$ commute, then

$$M_A [M_\alpha, [M_B, H]] - M_B [M_\alpha, [M_A, H]] = 0,$$

which is true if and only if the kernels of the above operators satisfy

$$\begin{aligned}
& A(x) \left(\alpha(x) \frac{B(x) - B(y)}{x - y} - \frac{B(x) - B(y)}{x - y} \alpha(y) \right) \\
= \;& B(x) \left(\alpha(x) \frac{A(x) - A(y)}{x - y} - \frac{A(x) - A(y)}{x - y} \alpha(y) \right) .
\end{aligned}$$

We rearrange this equation to get

$$(\alpha(x) - \alpha(y)) (A(y)B(x) - A(x)B(y)) = 0,$$

and since α is strictly increasing, the first factor is never zero, so we have the statement

$$A(y)B(x) - A(x)B(y) = 0 \quad \forall x, y \in \mathbb{R};$$

that is, A and B are proportional, and hence $K_{A,B}(x, y) = 0$ for all $x, y \in \mathbb{R}$. ■

Remark

We point out to the reader that this result does not contradict the Airy example we discussed in §3.3. If we were to try to run this argument on the half line, then equation (86) would not hold, by the remark after Lemma 3.7, the commutator would not have the simple form given here, and we would not be able to conclude that $K_{A,B}$ is zero.

4 Tracy–Widom operators on the circle

4.1 Introduction

Circular random matrix ensembles were introduced by Dyson to remove the assumption of statistical independence of matrix elements, which is somewhat artificial in physical applications (see [27, p.52 and p.181]). In this chapter, we consider the corresponding Tracy–Widom operators on the circle. A known identity linking Toeplitz and Hankel operators leads to a simple formula for the projection onto $H^2(\mathbb{T})$ of a TW operator on $L^2(\mathbb{T})$. The expression we obtain is a difference of the form $\Gamma_\phi^*\Gamma_\phi - \Gamma_\psi^*\Gamma_\psi$. If we then select one of the functions in (1) to be anti-analytic, we have automatically expressed the TW operator as $\Gamma_\phi^*\Gamma_\phi$. In the particular case where the functions A and B in (1) are a Blaschke product and its complex conjugate, we find a formula for the range of the projection onto $H^2(\mathbb{T})$ of the integral operator on $L^2(\mathbb{T})$. We consider Tracy–Widom kernels on the unit circle, of the form

$$K(\theta, \phi) = \frac{f(e^{i\theta})g(e^{i\phi}) - g(e^{i\theta})f(e^{i\phi})}{1 - e^{i(\theta-\phi)}}, \tag{87}$$

in which f and g are functions in $L^\infty(\mathbb{T})$. To see how these relate to the kernels studied elsewhere in the thesis, consider first the kernel

$$K(z, w) = \frac{f(z)g(w) - f(w)g(z)}{z - w}$$

which acts on a function $h \in L^2(\mathbb{T})$ in the usual way:

$$Kh(z) = \frac{1}{2\pi i}\int_{\mathbb{T}} \frac{f(z)g(w) - f(w)g(z)}{z - w} h(w)\,\mathrm{d}w.$$

Now we make the change of variables $z = e^{i\theta}$ and $w = e^{i\phi}$, and we find

$$Kh(e^{i\theta}) \quad = \quad \frac{1}{2\pi i}\int_0^{2\pi} \frac{f(e^{i\theta})g(e^{i\phi}) - f(e^{i\phi})g(e^{i\theta})}{e^{i\theta} - e^{i\phi}} ie^{i\phi}h(e^{i\phi})\,\mathrm{d}\phi$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \frac{f(e^{i\theta})g(e^{i\phi}) - f(e^{i\phi})g(e^{i\theta})}{e^{i(\theta-\phi)} - 1} h(e^{i\phi})\,\mathrm{d}\phi.$$

In fact, for later convenience in calculations, we shall consider a related kernel (multiplied by -1)

$$K(\theta, \phi) = \frac{f(e^{i\theta})g(e^{i\phi}) - f(e^{i\phi})g(e^{i\theta})}{1 - e^{i(\theta-\phi)}}.$$

The operation of K on a function $h \in L^2(\mathbb{T})$ is then

$$Kh(e^{i\theta}) = \frac{1}{2\pi}\int_0^{2\pi} K(\theta, \phi)h(e^{i\phi})\,\mathrm{d}\phi.$$

4.2 Sufficient conditions for Hankel factorisation

Lemma 4.1 *Suppose that $f, g \in L^\infty$ have $\bar{f} = g$. Then K defines a bounded and self-adjoint operator on $L^2(\mathbb{T})$. Further, the projection of K onto $H^2(\mathbb{T})$ satisfies*

$$P_+K = \Gamma_g^*\Gamma_g - \Gamma_f^*\Gamma_f. \tag{88}$$

Moreover, when f is continuous, P_+K is compact.

Proof. The condition $\bar{f} = g$ gives immediately $K(e^{i\theta}, e^{i\phi}) = \overline{K(e^{i\phi}, e^{i\theta})}$, and so K is self-adjoint. We shall write the Riesz projection P_+ as an integral operator on $L^2(\mathbb{T})$. Let $f \in L^2(\mathbb{T})$, choose $0 < r < 1$, and note that

$$\begin{aligned}
P_+f(re^{i\theta}) &= \sum_{n=0}^{\infty} \hat{f}(n)r^n e^{in\theta} \\
&= \sum_{n=0}^{\infty} \left(\frac{1}{2\pi}\int_0^{2\pi} f(e^{i\phi})e^{-in\phi}\,\mathrm{d}\phi\right) r^n e^{in\theta} \\
&= \frac{1}{2\pi}\sum_{n=0}^{\infty}\int_0^{2\pi} r^n e^{in(\theta-\phi)} f(e^{i\phi})\,\mathrm{d}\phi \\
&= \frac{1}{2\pi}\int_0^{2\pi}\sum_{n=0}^{\infty} r^n e^{in(\theta-\phi)} f(e^{i\phi})\,\mathrm{d}\phi
\end{aligned}$$

$$= \quad \frac{1}{2\pi}\int_0^{2\pi} \frac{f(e^{i\phi})}{1-re^{i(\theta-\phi)}}\,\mathrm{d}\phi,$$

where reversing the order of integration and summation is justified by the uniform convergence of the series, by a simple application of the $M-$test:

$$\left|\sum_{n=0}^{\infty} r^n e^{in(\theta-\phi)}\right| \le \sum_{n=0}^{\infty} r^n.$$

Now note that

$$\lim_{r\to 1^-} \left\| P_+ f(e^{i\theta}) - P_+ f(re^{i\theta})\right\|_{L^2} = 0$$

(a simple proof of this is in [26, p.6]) and so we have

$$P_+ f(e^{i\theta}) = \frac{1}{2\pi} PV \int_0^{2\pi} \frac{f(e^{i\phi})}{1-e^{i(\theta-\phi)}}\,\mathrm{d}\phi.$$

We can decompose the operator K as follows

$$Kh(e^{i\theta}) = g(e^{i\theta})\frac{1}{2\pi}\int_0^{2\pi} \frac{f(e^{i\theta})-f(e^{i\phi})}{1-e^{i(\theta-\phi)}} h(e^{i\phi})\,\mathrm{d}\phi - f(e^{i\theta})\frac{1}{2\pi}\int_0^{2\pi} \frac{g(e^{i\theta})-g(e^{i\phi})}{1-e^{i(\theta-\phi)}} h(e^{i\phi}))\,\mathrm{d}\phi$$

which can, using the expression for P_+ obtained above, be written in commutator notation as

$$K = M_g\left[M_f, P_+\right] - M_f\left[M_g, P_+\right], \tag{89}$$

where all the operators are bounded because f and g are in $L^\infty(\mathbb{T})$. To show that (88) holds, we expand the commutators and use Proposition 1.30, as follows:

$$\begin{aligned}
P_+K &= P_+\left(M_g[M_f,P_+] - M_f[M_g,P_+]\right) \\
&= P_+\left(M_g\left(M_fP_+ - P_+M_f\right) - M_f\left(M_gP_+ - P_+M_g\right)\right) \\
&= P_+(M_{gf} - M_gP_+M_f - M_{fg}P_+ + M_fP_+M_g) \\
&= (T_{gf} - T_gT_f) - (T_{fg} - T_fT_g) \\
&= \Gamma_g\Gamma^*_{\bar{f}} - \Gamma_f\Gamma^*_{\bar{g}} \\
&= \Gamma_g\Gamma^*_{g} - \Gamma_f\Gamma^*_{f}.
\end{aligned}$$

The last statement of the result follows by Hartman's theorem (Theorem 1.27): the Hankel operators on the right-hand side of (88) are compact when f is continuous. ■

Of course, we are looking for expressions of the form $\Gamma^*\Gamma$. We can arrange for this to happen by choosing the function to have all its nonnegative Fourier coefficients zero, i.e. $f \in H^2_-(\mathbb{T})$.

Proposition 4.2 *Suppose $w \in \mathbb{D}$, and define the function*

$$k_{w,l}(z) = \frac{l!\ z^l}{(1-\bar{w}z)^{l+1}} \quad (l \in \mathbb{Z}_+).$$

Then $k_{w,l}(z) \in H^2(\mathbb{D})$, and, for any $f \in H^2(\mathbb{D})$, we have

$$\langle f, k_{w,l} \rangle_{H^2} = f^{(l)}(w).$$

Remark The function $k_{w,l}$ is called the *reproducing kernel* for the l^{th} derivative.

Proof. Note that the change of variables $z = e^{i\theta}$ transforms the inner product integral as follows. Take f and g in $H^2(\mathbb{D})$. Then

$$\begin{aligned}\langle f, g\rangle &= \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta})\overline{g(e^{i\theta})}\,\mathrm{d}\theta \\ &= \frac{1}{2\pi i}\int_{\mathbb{T}} f(z)\overline{g(z)}\frac{\mathrm{d}z}{z},\end{aligned}$$

and hence, by Cauchy's formula for derivatives (see [42, p.82]),

$$\begin{aligned}\langle f, k_{w,l}\rangle &= \frac{1}{2\pi i}\int_{\mathbb{T}} f(z)\frac{l! z^{-l}}{(1-w\bar{z})^{l+1}}\frac{\mathrm{d}z}{z} \\ &= \frac{l!}{2\pi i}\int_{\mathbb{T}} \frac{f(z)}{(z-w)^{l+1}}\,\mathrm{d}z \\ &= f^{(l)}(w).\end{aligned}$$

■

Definition 4.3 *For (α_j) a sequence of distinct points in $\mathbb{D}$ with multiplicities $m(\alpha_j)$ such that $\sum(1-|\alpha_j|)m(\alpha_j)<\infty$, the Blashke product, with zeros at each α_j having multiplicity $m(\alpha_j)$ and at 0 with multiplicity $m(0)$ is*

$$b(z) = z^{m(0)} \prod_{j=0}^{\infty} \left(\frac{|\alpha_j|}{\alpha_j} \frac{z-\alpha_j}{1-\overline{\alpha_j} z} \right)^{m(\alpha_j)}.$$

For convenience in the next result, we use one of the alternative definitions of Hankel operator, namely, $\Gamma_\phi = JP_- M_\phi$. Using calculations which are very similar to those in 1.30, we can adapt the identity of Proposition 4.1 to fit this new definition. It now reads

$$P_+ K = \Gamma_f^* \Gamma_f - \Gamma_g^* \Gamma_g$$

where now $\Gamma_\phi = 0$ when $\phi \in H^2(\mathbb{T})$ instead of when $\phi \in H^2_-(\mathbb{T})$. The result is essentially contained in a remark in [37, p.22], but we include it here to show how it fits in with our work linking Hankel and Tracy–Widom operators.

Proposition 4.4 *Let $f=\bar{b}$ and $g=b$. Then $P_+K = \Gamma_{\bar{b}}^* \Gamma_{\bar{b}}$ and*

$$\overline{\text{Range}}(P_+K) = \overline{\text{span}}\{k_{\alpha_j,l} : l = 0,\ldots,m(\alpha_j)-1, j=0,1,\ldots\}.$$

In particular, if b is a finite Blashke product, then P_+K has finite rank.

Proof. Applying Lemma 4.1 gives $P_+K = \Gamma_{\bar{b}}^* \Gamma_{\bar{b}}$, so that we can find the closure of the range of P_+K by calculating $(\text{Ker } \Gamma_{\bar{b}}^* \Gamma_{\bar{b}})^\perp$, which by Lemma 1.34 is equal to $(\text{Ker } \Gamma_{\bar{b}})^\perp$. The kernel of any Hankel operator is shift-invariant (Proposition 1.24), so by Beurling's theorem (Corollary 1.3) we have $\text{Ker } \Gamma_{\bar{b}} = uH^2$, for an inner function u which is unique up to a unimodular constant. Clearly $b \in \text{Ker } \Gamma_{\bar{b}}$, so we have $b = uv$, where v is inner. But the function v is constant, since

$$0 = \Gamma_{\bar{b}} u(z) = \Gamma_{\bar{b}} b(z) \bar{v}(z) = \bar{z}(\bar{v}(\bar{z}) - \bar{v}(0))$$

for all z, and so we have

$$\text{Ker } \Gamma_{\bar{b}} = bH^2.$$

Thus

$$\text{Range}(K_+W) = (bH^2)^\perp.$$

Now suppose $h \in bH^2$, or equivalently $h/b \in H^2$. This can only happen if each zero α_j of b is also a zero of h, with the same multiplicity. This, in turn, is true if and only if $\langle h, k_{\alpha_j,l} \rangle = 0$ for $l = 0, \ldots, m(\alpha_j) - 1$ and all j. We have shown that

$$\text{Range}(K_+W) = (bH^2)^\perp = \overline{\text{span}}\{k_{\alpha_j,l} : l = 0, \ldots, m(\alpha_j) - 1, j = 0, 1, \ldots\}.$$

■

5 Discrete Tracy–Widom Operators expressible as Hankel Squares

5.1 Introduction

Blower [6] has found sufficient conditions for a Tracy–Widom integral operator to be expressible as the square of a Hankel integral operator. Here we find an analogous set of sufficient conditions for the discrete case, that is, operators on $\ell^2(\mathbb{Z}_+)$ whose matrices with respect to the standard basis have entries

$$K(m,n) = \frac{\langle Ja(m), a(n)\rangle}{m-n} \ (m \neq n; m,n \in \mathbb{Z}_+), \tag{90}$$

where $(a(j))_{j=0}^{\infty}$ is a sequence of 2×1 vectors, and

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

These are discrete analogues of the continuous operators considered in Chapter 2 and 3. The comparison is not simply one of formal similarity: the discrete operators have genuine applications to discrete random matrix ensembles. For instance, Johannson [21] and Borodin *et al* [4] use the discrete Bessel kernel to describe the distribution of points at the edge of a growing Young diagram. We look for conditions under which we can construct a function $\phi : \mathbb{Z}_+ \to \mathbb{R}$ with $(\phi(j)) \in \ell^2$ such that

$$K(m,n) = \sum_{k=0}^{\infty} \phi(m+k)\phi(n+k), \quad (m \neq n; m,n \in \mathbb{Z}_+) \tag{91}$$

that is, the Tracy–Widom operator K equals the square of the Hankel matrix with entries $\phi(m+n)$. We recover a result of Johansson [21] and also Borodin *et al* [4], that the discrete Bessel kernel can be expressed in the form (91). In the continuous case, some examples fall 'just short' of being exact Hankel squares, in that they they are of the form $\Gamma^*\Gamma$ or $\Gamma_\phi\Gamma_\psi + \Gamma_\psi\Gamma_\phi$, where Γ, Γ_ϕ and Γ_ψ are Hankel operators (for example, see [1]). In the discrete context, we add another case, in which the Tracy–Widom operator can be written as the sum of a Hankel square and Toeplitz operator. The spectrum of Toeplitz operators can in particular cases be determined, as we noted in §1.6, so this expression may also be useful in calculating the spectrum of the operator K. We give what appears to be a new example, which we view as the discrete analogue of the Laguerre kernel, and is expressible as the sum of a Hankel square and Toeplitz operator. We warn the reader about a change in notation in this chapter. Throughout, a Hankel matrix Γ_ϕ will have matrix entries given by the function $\phi : \mathbb{Z}_+ \to \mathbb{R}$ in the following way:

$$\Gamma_\phi(m, n) = \phi(m + n) \quad m, n \in \mathbb{Z}_+.$$

This differs from the definition we gave in §1.6, in that ϕ is no longer an $L^\infty(\mathbb{T})$ function whose Fourier coefficients give the entries of the Hankel matrix, so it cannot be described as a symbol function. We make this change to avoid having to use hat notation when dealing with the entries of Γ_ϕ, and we hope that it will not cause confusion. Note also that

$$\Gamma_\phi^2(m, n) = \sum_{k=0}^{\infty} \phi(m + k)\phi(n + k).$$

Following the pattern of the continuous case, we analyse operators arising from recurrence relations of the form

$$a(j + 1) = T(j)a(j)$$

where $T(x)$ is a 2×2 matrix whose entries are rational functions of x, and show that an analogous set of conditions gives Hankel factorisation.

5.2 Sufficient conditions for Hankel factorisation in the discrete case

The first result considers the more general problem of expressing K as the sum of a Hankel and Toeplitz operator. It then turns out to be an easy matter to narrow the conditions to make K an exact Hankel square; this is the content of Corollary 5.2. We make use again of the involution matrix

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},$$

which satisfies $J^2 = -I$ and $J^t = -J$.

Theorem 5.1 *Let $T(j)$ and $B(j)$ be 2×2 real matrices and $(a(j))$ a sequence of real 2×1 vectors such that*

$$a(j+1) = T(j)a(j) \quad (j \in \mathbb{Z}_+) \tag{92}$$

$$\sum_{j=1}^{\infty} \| B(j)a(j) \|^2 < \infty. \tag{93}$$

Suppose further that there exists a matrix C with eigenvalues $\lambda > 0$ and 0, with (respectively) eigenvectors v_λ and v_0 and for which

$$\frac{J - T(n)^t J T(m)}{m-n} = B(n)^t C B(m) \quad (m \neq n; m, n \in \mathbb{Z}_+). \tag{94}$$

Let

$$\phi(j) = \lambda^{1/2} \langle v_\lambda, B(j)a(j) \rangle$$

Then

$$\Gamma_\phi^2 = K - W, \tag{95}$$

where K has entries as in (90), and W is a Toeplitz matrix with $W(0) = 0$, and *entries*

$$W(m-n) = \lim_{k\to\infty} \frac{\langle Ja(m+k), a(n+k)\rangle}{m-n} \quad (m \neq n; m, n \in \mathbb{Z}_+). \tag{96}$$

Remark

Notice that we let the (previously undefined) diagonal entries of K be defined by the Hankel operator Γ_ϕ in (95).

Proof. Let K have entries as in (90), and let U be the orthogonal matrix with v_λ in the first column, and the eigenvalue corresponding to the eigenvalue 0 in the second column. If $m \neq n$, then we have

$$\begin{aligned}
& K(m+k, n+k) - K(m+k+1, n+k+1) \\
= \; & \frac{1}{m-n} \left(\langle Ja(m+k), a(n+k)\rangle - \langle JT(m+k)a(m+k), T(n+k)a(n+k)\rangle\right) \\
= \; & \frac{1}{m-n} \langle \big(J - T(n+k)^t JT(m+k)\big)\, a(m+k), a(n+k)\rangle \\
= \; & \langle B(n+k)^t CB(m+k)a(m+k), a(n+k)\rangle \\
= \; & \langle B(n+k)^t U \operatorname{diag}(\lambda, 0) U^t B(m+k)a(m+k), a(n+k)\rangle \\
= \; & \lambda\langle \operatorname{diag}(1,0) U^t B(m+k)a(m+k), U^t B(n+k)a(n+k)\rangle \\
= \; & \lambda\langle \operatorname{diag}(1,0) U^t B(m+k)a(m+k), \operatorname{diag}(1,0) U^t B(n+k)a(n+k)\rangle \\
= \; & \phi(m+k)\phi(n+k),
\end{aligned}$$

Summing over $k = 0, 1, 2, \ldots, N$, we get

$$K(m,n) - K(m+N+1, n+N+1) = \sum_{k=0}^{N} \phi(m+k)\phi(n+k). \tag{97}$$

By the Cauchy-Schwarz inequality

$$\left| \sum_{k=0}^{N} \phi(m+k)\phi(n+k) \right| \leq \left(\sum_{k=0}^{N} |\phi(m+k)|^2 \right)^{1/2} \left(\sum_{k=0}^{N} |\phi(n+k)|^2 \right)^{1/2}, \tag{98}$$

and $(\phi(j))$ is an ℓ^2 sequence, so we can let $N \to \infty$ in (97), to get

$$K(m,n) = \sum_{k=0}^{\infty} \phi(m+k)\phi(n+k) + W(m-n) \quad (m \neq n; m,n \in \mathbb{Z}_+). \tag{99}$$

The first term on the right-hand side of (99) is the $(m,n)^{th}$ $(m \neq n)$ entry of the square of the Hankel operator Γ_ϕ. Since K previously had no definition for $m = n$, we now let $K(m,m) = \sum_{k=0}^{\infty} \phi(m+k)^2$, the diagonal $(m = n)$ entries of Γ_ϕ^2. ■

The case of most interest, and which gives the most spectral information, is when the operator factors exactly as a product of Hankels: this is a special case of the result above, in which the Toeplitz operator W is zero.

Corollary 5.2 *Let $T(j)$, $B(j)$ and $(a(j))$ are as in Theorem 5.1, and suppose further that*

$$a(j) \to 0 \text{ as } j \to \infty. \tag{100}$$

Then

$$\Gamma_\phi^2 = K \tag{101}$$

where K has entries as in (90). Further, K is compact.

Proof. By Theorem 5.1, we have, for $m \neq n$,

$$K(m,n) = \Gamma_\phi^2(m,n) + W(m-n),$$

where by hypothesis

$$W(m-n) = \lim_{k\to\infty} \frac{\langle Ja(m+k), a(n+k)\rangle}{m-n} = 0 \quad (m \neq n; m,n \in \mathbb{Z}_+).$$

Note that K can be viewed as the composition of the discrete Hilbert transform

$$a_n \mapsto \sum_{m,n;m\neq n} \frac{1}{m-n} a_n$$

which is bounded, by Theorem 294 of [20], with the compact operators $\ell^2(\mathbb{Z}_+, \mathbb{C}) \to \ell^2(\mathbb{Z}_+, \mathbb{C}^2)$, given by $(x_n) \mapsto (Ja(n)x_n)$ and the adjoint of $(x(n)) \mapsto (a(n)x_n)$; so K is compact. Since the sequence $a(n)$ is bounded, it is clear that the two compact operators introduced above are also bounded, so K is also bounded. ■

Proposition 5.3 *Let $a(j)$, $T(j)$ and $B(j)$ satisfy conditions (92), (94), (100), and also*

$$\sum_{j=0}^{\infty} j \, \| B(j) \|^2 < \infty. \tag{102}$$

Now let $K = \Gamma_\phi^2$ as in Corollary 5.2. Then K is a trace-class operator.

Proof. We show that Γ_ϕ is Hilbert–Schmidt, by considering the sum of squares of its matrix entries, as in Proposition 1.7. Since the matrix of Γ_ϕ is constant on diagonals perpendicular to the main diagonal, we have

$$\begin{aligned} \| \Gamma_\phi \|_{C_2} &= \sum_{k=1}^{\infty} k \, |\phi(k)|^2 \\ &= \sum_{k=1}^{\infty} k \left| \lambda^{1/2} \langle v_\lambda, B(k) a(k) \rangle \right|^2 \\ &\le M_0 \sum_{k=1}^{\infty} k \, |a(k)|^2 \, \| B(k) \|^2 < \infty, \end{aligned}$$

where we used the condition (102), the Cauchy–Schwarz inequality, and the fact that the sequence $(a(j))$ is by hypothesis bounded. ■

5.3 Lyapunov equations

There is an alternative way of viewing the results in the previous sections, using the concept of Lyapunov equations. These arise in linear system theory (see [9, p.135]), where they can be used to determine whether or not a linear system is stable.

Theorem 5.4 *Suppose that K is bounded, $K = K^*$ and $K - S^*KS \geq 0$, so that*

$$K - S^*KS = MM^*, \tag{103}$$

for some bounded operator M. Suppose also that

$$\sum_{k=0}^{\infty} S^{*k} MM^* S^k \tag{104}$$

converges in the weak operator topology. Then

$$W = \lim_{k\to\infty} S^{*k} K S^k$$

exists in the weak operator topology, and satisfies

$$K = W + \sum_{k=0}^{\infty} S^{*k} MM^* S^k.$$

Remark The sum (104) is the controllability operator as used by Peller *et al* in [30, p.281] (see also [37, p.457]).

Proof. We have the sequence of equalities

$$(S^*)^k K S^k - (S^*)^{k+1} K S^{k+1} = (S^*)^k MM^* S^k \quad \text{for } k \in \mathbb{Z}_+, \tag{105}$$

and adding these for $k = 0, \ldots, n$ gives

$$K - (S^*)^{n+1} K S^{n+1} = \sum_{k=0}^{n} (S^*)^k MM^* S^k.$$

The sum on the right is convergent in the weak operator topology as $n \to \infty$, so we get

$$K = W + \sum_{k=0}^{\infty} (S^*)^k MM^* S^k, \tag{106}$$

where $W = \lim_{n\to\infty} (S^*)^n K S^n$. ■

The following corollary shows how the above Theorem relates to the results of §5.2.

Corollary 5.5 *Suppose further that in Theorem 5.4, $K - S^*KS \geq 0$ has rank one, so that the matrix of MM^* is $[\phi(j)\overline{\phi(k)}]_{j,k\geq 0}$, for some $\phi : \mathbb{Z}_+ \to \mathbb{C}$. Then $K = W + \Gamma\Gamma^*$ where Γ is the Hankel operator that has matrix $[\phi(j+k)]_{j,k\geq 0}$.*

Proof. We have

$$K = W + \sum_{k=0}^{\infty} (S^*)^k MM^* S^k,$$

where

$$MM^* = [\phi(j)\overline{\phi(k)}]_{j,k\geq 0}.$$

Let (e_m) be the standard basis for ℓ^2. Then the $(m,n)^{th}$ entry of the matrix of $\sum_{k=0}^{\infty}(S^*)^k MM^* S^k)$

$$\begin{aligned}
&\left\langle (\sum_{k=0}^{\infty} (S^*)^k MM^* S^k) e_m, e_n \right\rangle \\
= &\sum_{k=0}^{\infty} \langle (S^*)^k MM^* S^k e_m, e_n \rangle \\
= &\sum_{k=0}^{\infty} \langle MM^* S^k e_m, S^k e_n \rangle \\
= &\sum_{k=0}^{\infty} \langle MM^* e_{m+k}, e_{n+k} \rangle \\
= &\sum_{k=0}^{\infty} \phi(m+k)\overline{\phi(n+k)},
\end{aligned}$$

which shows that this sum has the same matrix as $\Gamma\Gamma^*$. ■

5.4 The discrete Bessel kernel

We show how Corollary 5.2 can be applied to the discrete Bessel kernel to recover a result which appears in [4] and [21], without their use of asymptotic formulae for the Bessel functions. The definition of the Bessel function, and two of the identities we use here, appear elsewhere in the thesis (§1.11).

Proposition 5.6 *Let* $\mathrm{J}_n(z)$ *be the Bessel function of the first kind of order* n*, let* $\mathrm{J}_n = \mathrm{J}_n(2\sqrt{\theta})$*, where* $\theta > 0$*; let* $\phi(n) = \mathrm{J}_{n+1}$ *and* $a(n) = [\sqrt{\theta}\,\mathrm{J}_n,\ \mathrm{J}_{n+1}]^t$*. Then the Hankel operator* Γ_ϕ *is Hilbert–Schmidt, and* $B = \Gamma_\phi^2$ *has entries*

$$B(m,n) = \frac{\langle Ja(m), a(n)\rangle}{m-n} \quad (m \neq n; m, n \in \mathbb{Z}_+)$$

Proof. It is clear that (2.1) holds, since we have the recurrence relation

$$\mathrm{J}_{n+2}(2z) = \frac{n+1}{z}\,\mathrm{J}_{n+1}(2z) - \mathrm{J}_n(2z),$$

giving $a(n+1) = T(n)a(n)$, where

$$T(n) = \begin{bmatrix} 0 & \sqrt{\theta} \\ \frac{-1}{\sqrt{\theta}} & \frac{n+1}{\sqrt{\theta}} \end{bmatrix}.$$

Note that

$$\frac{T(n)^t JT(m) - J}{m-n} = C,$$

where $C = \mathrm{diag}(0,-1)$, which is clearly of rank one. The non-zero eigenvalue of C is $\lambda = -1$, and a corresponding unit eigenvector is $v_\lambda = [0,1]^t$, so

$$|\lambda|^{1/2}\,\langle v_\lambda, a(n)\rangle = \mathrm{J}_{n+1} = \phi(n).$$

We now verify condition (102), and thus (100). Note that

$$\frac{1}{\theta}\sum_{n=1}^{\infty} n\,\mathrm{J}_{n+1}^2 < \frac{1}{\theta}\sum_{n=1}^{\infty}(n+1)^2\,\mathrm{J}_{n+1}^2 = \sum_{n=1}^{\infty}(\mathrm{J}_{n+2} + \mathrm{J}_n)^2 \leq 4\sum_{n=1}^{\infty}\mathrm{J}_n^2. \tag{107}$$

We have, by a standard formula given in [46, p.379]

$$e^{i2\sqrt{\theta}\sin\psi} = \mathrm{J}_0(2\sqrt{\theta}) + 2\sum_{m=1}^{\infty}\mathrm{J}_{2m}(2\sqrt{\theta})\cos 2m\psi + 2i\sum_{m=1}^{\infty}\mathrm{J}_{2m-1}(2\sqrt{\theta})\sin(2m-1)\psi,$$

and since $\mathrm{J}_k(z) = (-1)^k\, \mathrm{J}_k(z)$, we can write this as

$$\begin{aligned} e^{i2\sqrt{\theta}\sin\psi} &= \mathrm{J}_0 + \sum_{k=1}^{\infty} (\,\mathrm{J}_k + \,\mathrm{J}_{-k}) \cos k\psi + i \sum_{k=1}^{\infty} (\,\mathrm{J}_k - \,\mathrm{J}_{-k}) \sin k\psi \\ &= J_0(2\sqrt{\theta}) + \sum_{k=1}^{\infty} \mathrm{J}_k\, (\cos k\psi + i \sin k\psi) + \,\mathrm{J}_{-k}\, (\cos k\psi - i \sin k\psi) \\ &= \sum_{k=-\infty}^{\infty} \mathrm{J}_k e^{ik\psi}. \end{aligned}$$

By Parseval's identity (Theorem 1.2) we get

$$\sum_{k=-\infty}^{\infty} |\,\mathrm{J}_k|^2 = \frac{1}{2\pi} \int_0^{2\pi} \left| e^{i2\sqrt{\theta}\sin\psi} \right|^2 \mathrm{d}\psi = 1,$$

and hence $\mathrm{J}_0(2\sqrt{\theta})^2 + 2\sum_{m=1}^{\infty} \mathrm{J}_m(2\sqrt{\theta})^2 = 1$ for all $\theta > 0$, so that the sum on the right hand side of (107) is finite. ■

5.5 A discrete analogue of the Laguerre kernel

The Laguerre kernel, as considered in [7], arises from the solutions of the differential equation,

$$u''(x) + (-1/4 + (n+1)/x)\, u(x) = 0,$$

which we can write as

$$\frac{\mathrm{d}}{\mathrm{d}x} \begin{bmatrix} u(x) \\ u'(x) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1/4 - (n+1)/x & 0 \end{bmatrix} \begin{bmatrix} u(x) \\ u'(x) \end{bmatrix}.$$

We pick the solution $u(x) = u(x) = xe^{-x/2} L_n^{(1)}(x)$, as in §1.11, and then Blower's result (Theorem 1.43 in this thesis) can then be applied directly to show that

$$\frac{u(x)u'(y) - u(y)u'(y)}{x-y} = \frac{\langle Ja(x), a(y)\rangle}{x-y} = (n+1) \int_0^{\infty} \frac{u(x+t)u(y+t)}{(x+t)(y+t)}\, \mathrm{d}t. \quad (108)$$

We consider an analogous recurrence relation (see Proposition 5.8 below) and the Hankel operator we obtain has a similar shape. Furthermore we shall see that the generating function of the discrete analogue of u is similar in form to the Laplace transform of u. Unlike the continuous case, however, we do not obtain K as an exact Hankel square: an extra Toeplitz operator is involved, as in Theorem 5.1.

The following easy result will be used in the course of the proof of Proposition 5.8.

Lemma 5.7 *Suppose a_k is a sequence of complex numbers such that*

$$|a_{k+1} - a_k| = O\left(\frac{1}{k^2}\right).$$

for all sufficiently large k. Then there exists $a \in \mathbb{C}$ such that $a_k \to a$ as $k \to \infty$.

Proof. We have

$$|a_{n+1} - a_1| = \left|\sum_{k=1}^{n}(a_{k+1} - a_k)\right| \leq \sum_{k=1}^{n}|a_{k+1} - a_k| \leq \sum_{k=1}^{n}\frac{M}{n^2}$$

for some $M > 0$. The sum on the right is convergent as $n \to \infty$, so the sequence $(a(n))$ has a limit as required. ■

Proposition 5.8 *For $\theta \in \mathbb{R}$, let $(a(j))$ satisfy the recurrence relation $a(j+1) = T(j)a(j)$, where*

$$T(j) = \begin{bmatrix} \theta/(j+1) & -1 \\ 1 & 0 \end{bmatrix}$$

and $a(1) = [\theta, 1]^t$. Then there exist polynomials $p_j(\theta)$ of degree j, with real coefficients, such that:

(i) $a(j) = [p_j(\theta), p_{j-1}(\theta)]^t$;

(ii) The self adjoint Hankel matrix $\Gamma_\phi = [\phi(j+k-1)]_{j,k\geq 1}$ with entries

$$\phi(j) = \frac{\sqrt{\theta}p_j(\theta)}{j+1}$$

is a bounded linear operator such that $\Gamma^2 = K - W$*, where* K *has entries as in (90), and* W *is the bounded Toeplitz operator with matrix*

$$W(m-n) = \frac{\langle J^{\bar{m}-\bar{n}+1} T_\infty a(1), T_\infty a(1)\rangle}{m-n} \qquad (m \neq n; m, n \in \mathbb{Z}_+)$$

for some constant 2×2 *matrix* T_∞*, where* $\bar{m}$ *is the conjugacy class of* m *mod 4.*

Proof. Firstly, note that

$$\frac{J - T(n)^t J T(m)}{m-n} = \begin{bmatrix} \theta/(m+1)(n+1) & 0 \\ 0 & 0 \end{bmatrix} = B(n)^t C B(m),$$

where $B(j) = \operatorname{diag}(1/j, 0)$ and $C = \operatorname{diag}(\theta, 0)$. We show that $(a(j))$ is bounded, and deduce that condition (93) of Theorem 5.1 holds. Clearly $a(n) = T(n)\dots T(1)a(1)$, and the matrix $T(n)$ is like J (which has $J^4 = I$) for large n, so we consider the product of T-matrices in bunches of 4. Note that

$$\begin{aligned} \Phi(j) &:= T(4j)T(4j-1)T(4j-2)T(4j-3) \\ &= I + O\left(\frac{1}{j^2}\right) - \frac{\theta}{2}\begin{bmatrix} 0 & \frac{-1}{(1-\frac{1}{4j})(j+\frac{1}{4})} \\ \frac{(1+\frac{1}{4j})}{j-\frac{1}{2}} & 0 \end{bmatrix} \\ &= I + O\left(\frac{1}{j^2}\right) - \frac{\theta}{2j}J \end{aligned}$$

for $j \in \mathbb{Z}_+$. We deduce that

$$\begin{aligned} \|\Phi(j)\|^2 &= \|\Phi(j)^*\Phi(j)\| \\ &= \left\| \left(I + O\left(\frac{1}{j^2}\right) + \frac{\theta}{2j}J\right)\left(I + O\left(\frac{1}{j^2}\right) - \frac{\theta}{2j}J\right)\right\| \\ &= \left\| I + \frac{\theta^2}{4j^2} I \right\| \\ &= 1 + O\left(\frac{1}{j^2}\right) \end{aligned}$$

and similarly for $\Phi(j)^{-1}$. Hence there exists $C(\theta)$, independent of n, with

$$\|T(n)T(n-1)\dots T(1)\| \leq C(\theta)$$

and likewise

$$\left\|T(1)^{-1}T(2)^{-1}\dots T(n)^{-1}\right\| \leq C(\theta)$$

for $n \in \mathbb{Z}_+$. Hence there exists $\kappa(\theta) > 0$ such that $\kappa(\theta) < \|a(n)\| < \kappa(\theta)^{-1}$ for all $n \in \mathbb{Z}_+$, that is, $(a(n))$ is bounded as required. Now let

$$C_k = \prod_{j=1}^{k} \exp\left(\frac{\theta J}{2j}\right)\Phi(j),$$

and note that

$$\begin{aligned}
C_{k+1} - C_k &= \left(\exp\left(\frac{\theta J}{2(k+1)}\right)\Phi(k+1) - I\right)\prod_{j=1}^{k}\exp\left(\frac{\theta J}{2j}\right)\Phi(j) \\
&= \left(\left(I + \frac{\theta J}{2(k+1)} + O\left(\frac{1}{k^2}\right)\right)\left(I - \frac{\theta J}{2(k+1)} + O\left(\frac{1}{k^2}\right)\right) - I\right) \\
&\quad \times \prod_{j=1}^{k}\left(\left(I + \frac{\theta J}{2j} + O\left(\frac{1}{j^2}\right)\right)\left(I - \frac{\theta J}{2j} + O\left(\frac{1}{j^2}\right)\right)\right) \\
&= O\left(\frac{1}{k^2}\right)\prod_{j=1}^{k}\left(I + O\left(\frac{1}{j^2}\right)\right) \\
&= O\left(\frac{1}{k^2}\right),
\end{aligned}$$

so that the limit

$$T_\infty = \lim_{k\to\infty} C_k$$

exists by Lemma 5.7. We have

$$T(m+4k)\dots T(1)$$

$$
= \begin{cases} \Phi(k+\frac{m}{4})\Phi(k-1+\frac{m}{4})\dots\Phi(1) & \text{if } m \equiv 0 \bmod 4 \\ T(m+4k)\Phi(k+\frac{m-1}{4})\dots\Phi(1) & \text{if } m \equiv 1 \bmod 4 \\ T(m+4k)T(m+4k-1)\Phi(k+\frac{m-2}{4})\dots\Phi(1) & \text{if } m \equiv 2 \bmod 4 \\ T(m+4k)T(m+4k-1)T(m+4k-2)\Phi(k+\frac{m-3}{4})\dots\Phi(1) & \text{if } m \equiv 3 \bmod 4. \end{cases}
$$

Now suppose $m, n \equiv 0 \bmod 4$. Since

$$
\begin{aligned}
& \left(\prod_{j=1}^{k+n/4} \exp\left(\frac{\theta J}{2j}\right)\right)^* \prod_{j=1}^{k+m/4} \exp\left(\frac{\theta J}{2j}\right) \\
= \;& \exp\left(\frac{\theta J}{2}\left(\sum_{j=1}^{k+m/4}\frac{1}{j} - \sum_{j=1}^{k+n/4}\frac{1}{j}\right)\right) \\
\to \;& \exp(0) = I,
\end{aligned}
$$

we have

$$
\begin{aligned}
& \lim_{k\to\infty} \left\langle JC_{k+m/4}a(1), C_{k+n/4}a(1)\right\rangle \\
= \;& \lim_{k\to\infty} \left\langle J \prod_{j=1}^{k+m/4} \exp\left(\frac{\theta J}{2j}\right) a(m+4k), \prod_{j=1}^{k+n/4} \exp\left(\frac{\theta J}{2j}\right) a(n+4k)\right\rangle \\
= \;& \lim_{k\to\infty} \left\langle J \left(\prod_{j=1}^{k+n/4} \exp\left(\frac{\theta J}{2j}\right)\right)^* \prod_{j=1}^{k+m/4} \exp\left(\frac{\theta J}{2j}\right) a(m+4k), a(n+4k)\right\rangle \\
= \;& \lim_{k\to\infty} \left\langle Ja(m+4k), a(n+4k)\right\rangle .
\end{aligned}
$$

If m or n is in any of the other conjugacy classes modulo 4, then since $T(j) \to J$ as $j \to \infty$, the extra T-matrices contribute factors of J as in the case analysis above. The other parts of the calculation are unaffected, and so we have

$$
\begin{aligned}
& \Big\langle Ja(m+4k), a(n+4k)\Big\rangle \\
= \;& \Big\langle JT(m+4k)T(m+4k-1)\dots T(1)a(1), T(n+4k)T(n+4k-1)\dots a(1)\Big\rangle \\
\to \;& \langle J^{\bar{m}+1}T_\infty a(1), J^{\bar{n}}a(1)\rangle
\end{aligned}
$$

as $k \to \infty$, where $\bar{m}$ and $\bar{n}$ are the conjugacy classes of m and n modulo 4. ∎

It is of interest to consider a further analogy with the continuous case, by finding the generating function of the polynomials p_j in Proposition 5.8. This turns out to have a similar shape to the Laplace transform of the function $u(x)$ in the Laguerre kernel, which we calculated in §1.11:

$$\mathcal{L}(u;\lambda) = (n+1)\frac{(\lambda-\frac{1}{2})^n}{(\lambda+\frac{1}{2})^{n+2}} \quad (\mathrm{Re}\,\lambda > -1/2).$$

Proposition 5.9 *Suppose that $p_0(\theta) = 1$ and $p_1(\theta) = \theta$. Then the generating function $f(z) = \sum_{n=0}^{\infty} p_j(\theta)z^n$ satisfies*

$$f(z) = \left(\frac{1-iz}{1+iz}\right)^{i\theta/2} \frac{1}{1+z^2} \quad (|z| < 1).$$

Proof. Write $p_n = p_n(\theta)$, and then

$$f(z) = \sum_{n=0}^{\infty} p_n z^n.$$

The recurrence relation for p_n is

$$(n+1)p_{n+1} = \theta p_n - (n+1)p_{n-1},$$

and we multiply through by the complex variable z^n and sum from $n = 1$ to ∞ to get

$$\begin{aligned}\sum_{n=1}^{\infty}(n+1)p_{n+1}z^n &= \theta\sum_{n=1}^{\infty} p_n z^n - \sum_{n=1}^{\infty}(n+1)p_{n-1}z^n \\ &= \theta\sum_{n=1}^{\infty} p_n z^n - \sum_{n=1}^{\infty}((n-1)+2)p_{n-1}z^n,\end{aligned}$$

which is equivalent to

$$f'(z) - p_1 = \theta(f(z) - p_0) - z^2 f'(z) - 2zf(z). \tag{109}$$

Now recall that $p_0 = 1$ and $p_1 = \theta$ to get

$$(1+z^2)f'(z) + (2z-\theta)f(z) = 0. \tag{110}$$

This is a separable differential equation:

$$\int \frac{1}{f(z)}\,\mathrm{d}f = \int \frac{\theta - 2z}{1+z^2}\,\mathrm{d}z \tag{111}$$

and the general solution is

$$\log(f(z)) = \theta\arctan(z) - \log(1+z^2) + C. \tag{112}$$

We have $f(o) = p_0 = 1$, so $C = 0$, and hence

$$f(z) = \frac{\exp(\theta\arctan(z))}{1+z^2}. \tag{113}$$

The complex arctangent function can be written as

$$\arctan(z) = \frac{-i}{2}\log\left(\frac{1+iz}{1-iz}\right),$$

and hence $f(z)$ is as stated. ■

5.6 The essential spectrum of Tracy–Widom operators

We now consider the case where of a Tracy–Widom operator which is "nearly" equal to a squared Hankel operator. To be more precise, we look for operators which are compact perturbations of Γ^2. A well-known result in operator theory then tells us that their essential spectra must be equal. We show that a discrete analogue of the Carleman operator is a compact perturbation of the operator induced by the Hilbert matrix.

Definition 5.10 *Let A and B be bounded operators on a Hilbert space H. We say A is a* compact perturbation *of B if there exists a compact operator K such that $A - B = K$.*

The following famous result is due to Weyl [47]: we state it here without proof.

Proposition 5.11 *Let A and B be bounded self-adjoint operators on a Hilbert space H, such that B is compact perturbation of A. Then $\sigma_{ess}(A) = \sigma_{ess}(B)$.*

Example 5.12

The Carleman operator is an example of a bounded Hankel integral operator, remarkable in having continuous spectrum of multiplicity two, as shown by Power in [35]. It is defined on $L^2(0,\infty)$ by

$$\Gamma h(x) = \int_0^\infty \frac{h(y)\,\mathrm{d}y}{x+y}.$$

As another example of a Tracy–Widom operator expressible as a Hankel square, we consider the system

$$\frac{\mathrm{d}}{\mathrm{d}x}\begin{bmatrix} \log x \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 1/x \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \log x \\ 1 \end{bmatrix}.$$

Writing $\Omega(x)$ for the 2×2 matrix above, we calculate

$$\frac{J\Omega(x) + \Omega(y)^t J}{x-y} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{-1}{xy} \end{bmatrix},$$

and then by Theorem 1.43 we have

$$K(x,y) := \frac{\log x - \log y}{x-y} = \int_0^\infty \frac{1}{(x+t)(y+t)}\,\mathrm{d}t,$$

where the right hand side is clearly the kernel of the squared Carleman operator Γ^2. By Power [35], $\sigma(\Gamma) = [0,\pi]$ of multiplicity two, so we can deduce by the spectral mapping theorem that $\sigma(K) = [0, \pi^2]$, also of multiplicity two.

The Carleman operator is the continuous analogue of the Hilbert matrix, and we now present a result showing that the discrete analogue of the kernel K above, while not equal to the square of the Hilbert matrix, is at least a Hilbert–Schmidt perturbation of it, so that we can use Proposition 5.11 to find its essential spectrum.

Theorem 5.13 *Let K be the matrix with elements*

$$K(m,n) = \begin{cases} \frac{1}{m-n}(\log m - \log n) & \text{for } m \neq n \\ \frac{1}{n} & \text{for } m = n \end{cases},$$

Then $\sigma_{ess}(K) = \sigma_{ess}(\Gamma^2) = [0, \pi^2]$, where Γ is the Hilbert matrix as defined in Proposition 1.29.

Proof. We shall show that the difference $K - \Gamma^2$ is Hilbert-Schmidt, and hence compact. Fix a column n, and sum the squared entries along the rows, excluding the term on the diagonal. We use the estimate

$$\sum_{k=1}^{N} \frac{1}{k} = \log N + \gamma + O\left(\frac{1}{N}\right),$$

where γ is Euler's constant, to get the following bound on the terms:

$$\begin{aligned}
&\sum_{m,n\geq 1; m\neq n} \left(K(m,n) - \Gamma^2(m,n)\right)^2 \\
= &\sum_{m,n\geq 1; m\neq n} \left(\frac{\log m - \log n}{m-n} - \sum_{k=1}^{\infty} \frac{1}{(m+k-1)(n+k-1)}\right)^2 \\
= &\sum_{m,n\geq 1; m\neq n} \frac{1}{(m-n)^2}\left(\log m - \log n - \sum_{k=1}^{\infty}\left(\frac{1}{n+k-1} - \frac{1}{m+k-1}\right)\right)^2 \\
= &\sum_{m,n\geq 1; m\neq n} \frac{1}{(m-n)^2}\left(\log m - \log n - \log(m-1) + O\left(\frac{1}{m}\right) + \log(n-1) + O\left(\frac{1}{n}\right)\right)^2 \\
= &\sum_{m,n\geq 1; m\neq n} \frac{1}{(m-n)^2}\left(\log\left(\frac{m}{m-1}\right) + \log\left(\frac{n-1}{n}\right) + O\left(\frac{1}{m}\right) + O\left(\frac{1}{n}\right)\right)^2.
\end{aligned}$$

Note that $\log(m/(m-1)) \to 0$ as $m \to \infty$, and similarly for $\log((n-1)/n)$, so the parts of the sum arising from these terms are finite, because $\sum_{j=1}^{\infty} 1/j^2 < \infty$. Also,

$$\begin{aligned}
\sum_{m,n\geq 1; m\neq n} \frac{1}{(m-n)^2}\frac{1}{mn} &\leq 2\sum_{m,n\geq 1, m\neq n} \frac{1}{(m-n)^2}\frac{1}{m^2} \\
&= 2\sum_{m=1}^{\infty}\left(\sum_{n=1}^{\infty}\frac{1}{(m-n)^2}\right)\frac{1}{m^2} \\
&\leq \sum_{m=1}^{\infty}\left(4\sum_{n=1}^{\infty}\frac{1}{n^2}\right)\frac{1}{m^2} \\
&= \frac{2\pi^2}{3}\sum_{m=1}^{\infty}\frac{1}{m^2} \\
&= \frac{\pi^4}{9}
\end{aligned}$$

so that the sum of off-diagonal terms is finite. For the diagonal terms, note that

$$0 \leq \sum_{k=1}^{\infty}\frac{1}{(k+n-1)^2} \leq \int_0^{\infty}\frac{1}{(t+n-1)^2}\,\mathrm{d}t = \frac{1}{n-1},$$

and hence

$$\begin{aligned}
\sum_{n=1}^{\infty}\left(K(n,n)-\Gamma^2(n,n)\right)^2 &= \sum_{n=1}^{\infty}\left(\frac{1}{n}-\sum_{k=1}^{\infty}\frac{1}{(k+n-1)^2}\right)^2 \\
&\leq \sum_{n=1}^{\infty}\left(\frac{1}{n-1}-\sum_{k=1}^{\infty}\frac{1}{(k+n-1)^2}\right)^2 \\
&\leq \sum_{n=1}^{\infty}\frac{1}{(n-1)^2} \\
&< \infty
\end{aligned} \tag{114}$$

We have shown that $K-\Gamma^2$ is Hilbert-Schmidt (and hence compact), so we can apply Proposition 5.11. The essential spectrum of the Hilbert matrix Γ is the interval $[0,\pi]$ (by Proposition 1.29), so, by the spectral mapping theorem, $\sigma_{ess}(K) = \sigma_{ess}(\Gamma^2) = [0,\pi^2]$. ■

References

[1] G. Aubrun, A sharp small deviation inequality for the largest eigenvalue of a random matrix, Springer Lecture Notes in Mathematics, 1857, Springer, Berlin, 2005.

[2] J. Agler, J.E. McCarthy, *Pick interpolation and Hilbert function spaces*, AMS Rhode Island, 2002

[3] M.J. Ablowitz, A.S. Fokas, *Complex Variables: Introduction and Applications*, 2nd edition, CUP, 2003.

[4] A. Borodin, A. Okounkov, G. Olshanski, Asymptotics of Plancherel measures for symmetric groups, *J. Amer. Math. Soc. 13 (2000) 481-515.*

[5] A. Borodin, Biorthogonal ensembles, *Nuclear Phys. B 536 (1999), no. 3, 704-732.*

[6] G Blower, Operators associated with Soft and Hard Spectral Edges from Unitary Ensembles, *J. Math. Anal. Appl 337 (1) (2007) pp. 239-265.*

[7] G. Blower, Integrable operators and the squares of Hankel operators, *J. Math. Anal. Appl 337 (2007), pp 239-265 (doi:10.1016/j.jmaa.2007.03.084).*

[8] G. Blower, A. McCafferty, Discrete Tracy–Widom Operators, *Proc. Edin. Math. Soc.* accepted May 2008, publication pending.

[9] Chi-Tsong Chen, *Linear System Theory and Design*, 3rd edition, OUP New York, 1999.

[10] J. Conway, *A course in functional analysis*, Springer-Verlag New York, 1985.

[11] P. Deift, *Orthogonal Polynomials and Random Matrices: a Riemann-Hilbert approach*, Courant Inst. Lecture Notes, AMS Rhode Island, 2000.

[12] E.B. Davies, *Spectral theory and differential operators*, Cambridge University Press, New York, 1995.

[13] K. R. Davidson, *C^*-Algebras by Example*, American Mathematical Society, Providence, RI, 1996.

[14] A Erdélyi, *Higher Transcendental functions*, Vol II, McGraw-Hill, 1953.

[15] P.J Forrester, The spectrum edge of random matrix ensembles, *Nucl. phys, B, 402 (1993), no3, pp. 709-728.*

[16] Y.V. Fyodorov, *Introduction to the random matrix theory: Gaussian Unitary Ensemble and beyond*, in *Recent Perspectives in Random Matrix Theory and Number Theory*, ed. F Mezzadri and N.C. Snaith, CUP 2005.

[17] I Gohberg, S Goldberg, *Basic Operator Theory*, Birkhuser, Boston, 1981.

[18] E. Hille, *Differential equations in the complex domain*, Wiley, New York, 1976.

[19] E. L Ince, *Ordinary differential equations*, Dover Publications, New York, 1956.

[20] G.H. Hardy, J.E. Littlewood, G. Pôlya, *Inequalities*, 2nd edition, CUP, 1988.

[21] K. Johansson, Discrete Orthogonal Polynomial ensembles and the Plancherel measure, Ann. Math. (2) 153 (2001), no. 2, 259-296.

[22] P. Koosis, *Introduction to H_p spaces*, 2nd edition, CUP, 1998.

[23] T. Lalesco, Une théorème sur les noyaux composes, *Bull. Acad. Sci. Roumaine 3 (1915), 271-272*.

[24] P. Lax, *Functional Analysis*, Wiley, New York, 2002.

[25] Lisuternik, Sobolev, *Elements of functional analysis*, Frederick Ungar, New York, 1961.

[26] R.A. Martînez and P. Rosenthal, *An Introduction to Operators on the Hardy-Hilbert Space*, Springer, New York, 2007.

[27] M.L. Mehta, *Random Matrices*, 2nd edition, Academic Press, Boston, 1991.

[28] M. L. Mehta and M. Gaudin, On the density of eigenvalues of a random matrix, *Nucl. Phys. B 18, 420-427.*

[29] B. Sz.-Nagy, *Introduction to Real Functions and Orthogonal Expansions*, OUP New York, 1965

[30] A.N. Megretskiĭ, V.V. Peller and S.R. Treil, The inverse spectral problem for self-adjoint Hankel operators, *Acta Math.* **174** (1995), 241–309.

[31] T. Nagao, M Wadati, Eigenvalue distribution of random matrices at the spectrum edge, *J. Phys. Soc. Japan 62 (1993), 3845-3856.*

[32] N.K. Nikolski, *Operators, Functions and Systems: an easy reading*, vol. 2, AMS, Providence, R.I., 2002.

[33] N.K. Nikolski, *Treatise on the shift operator*, Springer-Verlag, Berlin, New York, 1986.

[34] J. Partington, *An Introduction to Hankel Operators*, LMS student texts; 13, CUP, New York, 1988.

[35] S.C. Power, *Hankel Operators on Hilbert Space*, Research Notes in Math., 64, Pitman Adv. Publ. Progr., Boston-London-Melbourne, 1982.

[36] S. Power, The essential spectrum of a Hankel operator with piecewise continuous symbol, *Michigan Math. J. Volume 25, Issue 1 (1978), 117-121.*

[37] V.V. Peller, *Hankel Operators and Their Applications*, Springer, New York, 2003.

[38] R. Gilmore, *Lie groups, Lie algebras, and some of their applications*, Wiley Interscience, 1974.

[39] B. Simon, *Trace ideals and their applications*, AMS, Providence, R.I., 2005.

[40] A. Soshnikov, Determinantal Random Point Fields, 2000 *arXiv.org:math/0002099.*

[41] G. Szegö, *Orthogonal Polynomials*, AMS, 1959.

[42] E.C. Titchmarsh, *The Theory of Functions*, second edition, OUP, 1939.

[43] C.A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Comm. Math. Phys. 159 (1994), 151-174.*

[44] C.A. Tracy, H. Widom, Fredholm Determinants, Differential Equations and Matrix Models, *Comm. Math. Phys.* **163** *(1994), no. 1, 33–72.*

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

[46] E.T Whittaker and G.N. Watson, *A Course of Modern Analysis*, Cambridge University Press, 4th ed. 1963.

[47] H. Weyl, Über beschrnkte quadratische Formen, deren Differenz vollstetig ist, *Rend. Circ. Mat. Palermo* **27** *(1909)*, included in *Hermann Weyl: Gesammelte Abhanlungen, Band I*, Springer-Verlag, Berlin 1968

[48] E.P. Wigner, On the statistical distribution of the widths and spacings of nuclear resonance levels, *Proc. Cambridge Philos. Soc.* **47** *(1951) 790-798.*

[49] J. Wishart, Generalized product moment distribution in samples, *Biometrika* A **20** 32, 1928

[50] N. Young, *An Introduction to Hilbert Space*, CUP, 1988.