The Stochastic Operator Approach to Random Matrix Theory

by

Brian D. Sutton

B.S., Virginia Polytechnic Institute and State University, 2001

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2005

Author .....
Department of Mathematics
April 29, 2005

Certified by .....
Alan Edelman
Professor of Applied Mathematics
Thesis Supervisor

Accepted by .....
Rodolfo R. Rosales
Chairman, Applied Mathematics Committee

Accepted by .....
Pavel I. Etingof
Chairman, Department Committee on Graduate Students

The Stochastic Operator Approach to Random Matrix Theory

by

Brian D. Sutton

Submitted to the Department of Mathematics
on April 29, 2005, in partial fulfillment of the
requirements for the degree of
DOCTOR OF PHILOSOPHY

Abstract

Classical random matrix models are formed from dense matrices with Gaussian entries. Their eigenvalues have features that have been observed in combinatorics, statistical mechanics, quantum mechanics, and even the zeros of the Riemann zeta function. However, their eigenvectors are Haar-distributed—completely random. Therefore, these classical random matrices are rarely considered as operators.

The stochastic operator approach to random matrix theory, introduced here, shows that it is actually quite natural and quite useful to view random matrices as random operators. The first step is to perform a change of basis, replacing the traditional Gaussian random matrix models by carefully chosen distributions on structured, e.g., tridiagonal, matrices. These structured random matrix models were introduced by Dumitriu and Edelman, and of course have the same eigenvalue distributions as the classical models, since they are equivalent up to similarity transformation.

This dissertation shows that these structured random matrix models, appropriately rescaled, are finite difference approximations to stochastic differential operators. Specifically, as the size of one of these matrices approaches infinity, it looks more and more like an operator constructed from either the Airy operator,

$$\mathcal{A} = \frac{d^2}{dx^2} - x,$$

or one of the Bessel operators,

$$\mathcal{J}_a = -2\sqrt{x}\frac{d}{dx} + \frac{a}{\sqrt{x}},$$

plus noise.

One of the major advantages to the stochastic operator approach is a new method for working in "general β" random matrix theory. In the stochastic operator approach, there is always a parameter β which is inversely proportional to the variance of the noise. In contrast, the traditional Gaussian random matrix models identify

the parameter β with the real dimension of the division algebra of elements, limiting much study to the cases $\beta = 1$ (real entries), $\beta = 2$ (complex entries), and $\beta = 4$ (quaternion entries).

An application to general β random matrix theory is presented, specifically regarding the universal largest eigenvalue distributions. In the cases $\beta = 1, 2, 4$, Tracy and Widom derived exact formulas for these distributions. However, little is known about the general β case. In this dissertation, the stochastic operator approach is used to derive a new asymptotic expansion for the mean, valid near $\beta = \infty$. The expression is built from the eigendecomposition of the Airy operator, suggesting the intrinsic role of differential operators.

This dissertation also introduces a new matrix model for the Jacobi ensemble, solving a problem posed by Dumitriu and Edelman, and enabling the extension of the stochastic operator approach to the Jacobi case.

Thesis Supervisor: Alan Edelman
Title: Professor of Applied Mathematics

For Mom and Dad.

Acknowledgments

Greatest thanks go to my family: to Mom for the flash cards, to Dad for the computer, to Michael for showing me how to think big, and to Brittany for teaching me that it's possible to be both smart and cool. Also, thanks to Grandma, Granddaddy, and Granny for the encouragement, and to Mandy for all the pictures of Hannah.

To my adviser, Alan Edelman, I say, thanks for teaching me how to think. You took a guy who said the only two things he would never study were numerical analysis and differential equations, and you turned him into a Matlab addict.

To Charles Johnson, thank you for looking after me. You taught me that I could do real mathematics, and your continuing generosity over the years will not be forgotten.

To my "big sister" Ioana Dumitriu, thanks for all your knowledge. You always had a quick answer to my toughest questions. As for my "brothers," I thank Raj Rao for many interesting—and hard—problems and Per-Olof Persson for humoring my monthly emails about his software.

At Virginia Tech, Marge Murray instilled a love for mathematics, and Dean Riess instilled a love for the Mathematics Department. Go Hokies!

Finally, I wish the best for all of the great friends I have made at MIT, including Andrew, Bianca, Damiano, Fumei, Ilya, Karen, Michael, and Shelby.

Contents

List of Figures

Chapter 1

A hint of things to come

Random matrix theory can be divided into *finite random matrix theory* and *infinite random matrix theory*. Finite random matrix theory is primarily concerned with the eigenvalues of random matrices. Infinite random matrix theory considers what happens to the eigenvalues as the size of a random matrix approaches infinity.

This dissertation proposes that not only do the eigenvalues have $n \to \infty$ limiting distributions, but that the random matrices *themselves* have $n \to \infty$ limits. By this we mean that certain commonly studied random matrices are finite difference approximations to *stochastic differential operators.*

The random matrix distributions that we consider model the classical ensembles of random matrix theory and are displayed in Figures 1.0.1–1.0.3. The stochastic differential operators are built from the operators in Figures 1.0.4–1.0.6.[1] These six figures contain the primary objects of study in this dissertation.

[1]The reader may find the list of notation on pp. 151–153 helpful when viewing these figures.

JACOBI MATRIX MODEL

$$J^{\beta}_{a,b} \sim \begin{bmatrix} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) \\ B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) \end{bmatrix}$$

$$= \left[\begin{array}{ccccc|ccccc} c_n & -s_n c'_{n-1} & & & & s_n s'_{n-1} & & & & \\ & c_{n-1}s'_{n-1} & -s_{n-1}c'_{n-2} & & & c_{n-1}c'_{n-1} & s_{n-1}s'_{n-2} & & & \\ & & c_{n-2}s'_{n-2} & \ddots & & & c_{n-2}c'_{n-2} & s_{n-2}s'_{n-3} & & \\ & & & \ddots & -s_2 c'_1 & & & \ddots & \ddots & \\ & & & & c_1 s'_1 & & & & c_1 c'_1 & s_1 \\ \hline -s_n & -c_n c'_{n-1} & & & & c_n s'_{n-1} & & & & \\ & -s_{n-1}s'_{n-1} & -c_{n-1}c'_{n-2} & & & -s_{n-1}c'_{n-1} & c_{n-1}s'_{n-2} & & & \\ & & -s_{n-2}s'_{n-2} & \ddots & & & -s_{n-2}c'_{n-2} & c_{n-2}s'_{n-3} & & \\ & & & \ddots & -c_2 c'_1 & & & \ddots & \ddots & \\ & & & & -s_1 s'_1 & & & & -s_1 c'_1 & c_1 \end{array}\right]$$

$$\beta > 0,\ a,\ b > -1$$

$$\Theta = (\theta_n, \ldots, \theta_1) \in [0, \tfrac{\pi}{2}]^n \qquad \Phi = (\phi_{n-1}, \ldots, \phi_1) \in [0, \tfrac{\pi}{2}]^{n-1}$$

$$c_k = \cos\theta_k \qquad c'_k = \cos\phi_k$$

$$s_k = \sin\theta_k \qquad s'_k = \sin\phi_k$$

$$c_k \sim \sqrt{\text{beta}\left(\tfrac{\beta}{2}(a+k), \tfrac{\beta}{2}(b+k)\right)} \qquad c'_k \sim \sqrt{\text{beta}\left(\tfrac{\beta}{2}k, \tfrac{\beta}{2}(a+b+1+k)\right)}$$

Figure 1.0.1: The Jacobi matrix model. beta$(\cdot,\cdot)$ denotes a beta-distributed random variable. The angles $\theta_1, \ldots, \theta_n, \phi_1, \ldots, \phi_{n-1}$ are independent.

LAGUERRE MATRIX MODEL (square)

$$L_a^\beta \sim \frac{1}{\sqrt{\beta}} \begin{bmatrix} \chi_{(a+n)\beta} & & & & & \\ -\chi_{(n-1)\beta} & \chi_{(a+n-1)\beta} & & & & \\ & -\chi_{(n-2)\beta} & \chi_{(a+n-2)\beta} & & & \\ & & \ddots & \ddots & & \\ & & & -\chi_{\beta} & \chi_{(a+1)\beta} \end{bmatrix} \quad \begin{array}{l} \beta > 0 \\ a > -1 \\ n\text{-by-}n \end{array}$$

LAGUERRE MATRIX MODEL (rectangular)

$$M_a^\beta \sim \frac{1}{\sqrt{\beta}} \begin{bmatrix} -\chi_{n\beta} & \chi_{(a+n-1)\beta} & & & & \\ & -\chi_{(n-1)\beta} & \chi_{(a+n-2)\beta} & & & \\ & & -\chi_{(n-2)\beta} & \chi_{(a+n-3)\beta} & & \\ & & & \ddots & \ddots & \\ & & & & -\chi_{\beta} & \chi_{a\beta} \end{bmatrix} \quad \begin{array}{l} \beta > 0 \\ a > 0 \\ n\text{-by-}(n+1) \end{array}$$

Figure 1.0.2: The Laguerre matrix models. χ_r denotes a chi-distributed random variable with r degrees of freedom. Both models have independent entries. The square model was introduced in [6].

HERMITE MATRIX MODEL

$$H^\beta \sim \frac{1}{\sqrt{2\beta}} \begin{bmatrix} \sqrt{2}G & \chi_{(n-1)\beta} & & & \\ \chi_{(n-1)\beta} & \sqrt{2}G & \chi_{(n-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \chi_{2\beta} & \sqrt{2}G & \chi_{\beta} \\ & & & \chi_{\beta} & \sqrt{2}G \end{bmatrix}, \quad \beta > 0$$

Figure 1.0.3: The Hermite matrix model. χ_r denotes a chi-distributed random variable with r degrees of freedom, and G denotes a standard Gaussian random variable. The matrix is symmetric with independent entries in the upper triangular part. The model was introduced in [6].

AIRY OPERATOR

Operator:

$$\mathcal{A} = \frac{d^2}{dx^2} - x \text{ on } [0, \infty)$$

Boundary conditions:

$$f(0) = 0, \lim_{x \to \infty} f(x) = 0$$

Eigenvalue decomposition:

$$\mathcal{A}\left[\mathrm{Ai}(x + \zeta_k)\right] = \zeta_k \left[\mathrm{Ai}(x + \zeta_k)\right], \; 0 > \zeta_1 > \zeta_2 > \ldots \text{ zeros of Ai}$$

Figure 1.0.4: The Airy operator.

BESSEL OPERATOR

Operator:

$$\mathcal{J}_a = -2\sqrt{x}\frac{d}{dx} + \frac{a}{\sqrt{x}},\ \mathcal{J}_a^* = 2\sqrt{x}\frac{d}{dx} + \frac{a+1}{\sqrt{x}} \text{ on } (0,1]$$

Boundary conditions:

$\mathcal{J}_a : f \mapsto g$

(i) $f(1) = 0,\ g(0) = 0$

(ii) $g(0) = g(1) = 0$

Singular value decomposition:

(i) $\begin{cases} \mathcal{J}_a[j_a(\zeta_k\sqrt{x})] = \zeta_k[j_{a+1}(\zeta_k\sqrt{x})],\ \mathcal{J}_a^*[j_{a+1}(\zeta_k\sqrt{x})] = \zeta_k[j_a(\zeta_k\sqrt{x})] \\ 0 < \zeta_1 < \zeta_2 < \dots \text{ zeros of } j_a \end{cases}$

(ii) $\begin{cases} \mathcal{J}_a[j_a(\zeta_k\sqrt{x})] = \zeta_k[j_{a+1}(\zeta_k\sqrt{x})],\ \mathcal{J}_a^*[j_{a+1}(\zeta_k\sqrt{x})] = \zeta_k[j_a(\zeta_k\sqrt{x})] \\ 0 < \zeta_1 < \zeta_2 < \dots \text{ zeros of } j_{a+1} \\ \mathcal{J}_a[x^{a/2}] = 0 \end{cases}$

BESSEL OPERATOR (Liouville form)

Operator:

$$\tilde{\mathcal{J}}_a = -\frac{d}{dy} + \left(a+\tfrac{1}{2}\right)\frac{1}{y},\ \tilde{\mathcal{J}}_a^* = \frac{d}{dy} + \left(a+\tfrac{1}{2}\right)\frac{1}{y} \text{ on } (0,1]$$

Boundary conditions:

$\tilde{\mathcal{J}}_a : f \mapsto g$

(i) $f(1) = 0,\ g(0) = 0$

(ii) $g(0) = g(1) = 0$

Singular value decomposition:

(i) $\begin{cases} \tilde{\mathcal{J}}_a[\sqrt{y}j_a(\zeta_k y)] = \zeta_k[\sqrt{y}j_{a+1}(\zeta_k y)],\ \tilde{\mathcal{J}}_a^*[\sqrt{y}j_{a+1}(\zeta_k y)] = \zeta_k[\sqrt{y}j_a(\zeta_k y)] \\ 0 < \zeta_1 < \zeta_2 < \dots \text{ zeros of } j_a \end{cases}$

(ii) $\begin{cases} \tilde{\mathcal{J}}_a[\sqrt{y}j_a(\zeta_k y)] = \zeta_k[\sqrt{y}j_{a+1}(\zeta_k y)],\ \tilde{\mathcal{J}}_a^*[\sqrt{y}j_{a+1}(\zeta_k y)] = \zeta_k[\sqrt{y}j_a(\zeta_k y)] \\ 0 < \zeta_1 < \zeta_2 < \dots \text{ zeros of } j_{a+1} \\ \tilde{\mathcal{J}}_a[x^{a+1/2}] = 0 \end{cases}$

Figure 1.0.5: Two families of Bessel operators, related by the change of variables $x = y^2$. The families are parameterized by $a > -1$.

SINE OPERATOR

Operator:

$$\mathcal{J}_{-1/2} = -2\sqrt{x}\frac{d}{dx} - \frac{1}{2\sqrt{x}},\ \mathcal{J}^*_{-1/2} = 2\sqrt{x}\frac{d}{dx} + \frac{1}{2\sqrt{x}} \text{ on } (0,1]$$

Boundary conditions:

$$\mathcal{J}_{-1/2} : f \mapsto g$$
(i) $f(1) = 0,\ g(0) = 0$
(ii) $g(0) = g(1) = 0$

Singular value decomposition:

$$\text{(i)} \begin{cases} \mathcal{J}_{-1/2}\left[\frac{\cos(\zeta_k\sqrt{x})}{x^{1/4}}\right] = \zeta_k\left[\frac{\sin(\zeta_k\sqrt{x})}{x^{1/4}}\right],\ \mathcal{J}^*_{-1/2}\left[\frac{\sin(\zeta_k\sqrt{x})}{x^{1/4}}\right] = \zeta_k\left[\frac{\cos(\zeta_k\sqrt{x})}{x^{1/4}}\right] \\ \zeta_k = \pi(k-\frac{1}{2}),\ k = 1,2,\ldots \end{cases}$$

$$\text{(ii)} \begin{cases} \mathcal{J}_{-1/2}\left[\frac{\cos(\zeta_k\sqrt{x})}{x^{1/4}}\right] = \zeta_k\left[\frac{\sin(\zeta_k\sqrt{x})}{x^{1/4}}\right],\ \mathcal{J}^*_{-1/2}\left[\frac{\sin(\zeta_k\sqrt{x})}{x^{1/4}}\right] = \zeta_k\left[\frac{\cos(\zeta_k\sqrt{x})}{x^{1/4}}\right] \\ \zeta_k = \pi k,\ k = 1,2,\ldots \\ \mathcal{J}_{-1/2}[x^{-1/4}] = 0 \end{cases}$$

SINE OPERATOR (Liouville form)

Operator:

$$\tilde{\mathcal{J}}_{-1/2} = -\frac{d}{dy},\ \tilde{\mathcal{J}}^*_{-1/2} = \frac{d}{dy} \text{ on } (0,1]$$

Boundary conditions:

$$\tilde{\mathcal{J}}_{-1/2} : f \mapsto g$$
(i) $f(1) = 0,\ g(0) = 0$
(ii) $g(0) = g(1) = 0$

Singular value decomposition:

$$\text{(i)} \begin{cases} \tilde{\mathcal{J}}_{-1/2}[\cos(\zeta_k y)] = \zeta_k[\sin(\zeta_k y)],\ \tilde{\mathcal{J}}^*_{-1/2}[\sin(\zeta_k y)] = \zeta_k[\cos(\zeta_k y)] \\ \zeta_k = \pi(k-\frac{1}{2}),\ k = 1,2,\ldots \end{cases}$$

$$\text{(ii)} \begin{cases} \tilde{\mathcal{J}}_{-1/2}[\cos(\zeta_k y)] = \zeta_k[\sin(\zeta_k y)],\ \tilde{\mathcal{J}}^*_{-1/2}[\sin(\zeta_k y)] = \zeta_k[\cos(\zeta_k y)] \\ \zeta_k = \pi k,\ k = 1,2,\ldots \\ \tilde{\mathcal{J}}_{-1/2}[1] = 0 \end{cases}$$

Figure 1.0.6: Two sine operators, related by the change of variables $x = y^2$. These operators are special cases of the Bessel operators.

Chapter 2

Introduction

In the physics literature, random matrix theory was developed to model nuclear energy levels. Spurred by the work of Porter and Rosenzweig [29] and of Dyson [8], the assumption of orthogonal, unitary, or symplectic invariance ($A \stackrel{d}{=} U^*AU$) quickly became canonized. As Bronk described the situation, "It is hard to imagine that the basis we choose for our ensemble should affect the eigenvalue distribution" [3]. Even the names of the most studied random matrix distributions reflect the emphasis on invariance, e.g., the Gaussian Orthogonal, Unitary, and Symplectic Ensembles (GOE, GUE, GSE).

The starting point for this dissertation is the following notion.

> Although physically appealing, unitary invariance is not the most useful choice mathematically.

In all of the books and journal articles that have been written on random matrix theory in the past fifty years, random matrices are rarely, if ever, treated as operators. The eigen*values* of random matrices have become a field unto themselves, encompassing parts of mathematics, physics, statistics, and other areas, but the eigen*vectors* are usually cast aside, because they are Haar-distributed—quite uninteresting. Haar-distributed eigenvectors stand in the way of an operator-theoretic approach to random

matrix theory.

Fortunately, new random matrix models have been introduced in recent years, pushed by Dumitriu and Edelman [4, 6, 19, 21]. In contrast to the dense matrix models from earlier years, the new random matrix models have structure. They are bidiagonal and symmetric tridiagonal, and hence are *not* unitarily invariant. Indeed, the eigenvectors of these matrix models are far from Haar-distributed. As will be seen, they often resemble special functions from applied mathematics, specifically, the Airy, Bessel, and sine functions. This eigenvector structure is a hint that perhaps random matrices should be considered as random operators.

What develops is the *stochastic operator approach* to random matrix theory:

> Rescaled random matrix models are finite difference approximations to stochastic differential operators.

The stochastic differential operators are built from the Airy, Bessel, and sine operators displayed in the previous chapter.

The concrete advantage of the stochastic operator approach is a new method for working in "general β" random matrix theory. Although classical random matrices come in three flavors—real, complex, and quaternion—their eigenvalue distributions generalize naturally to a much larger family, parameterized by a real number $\beta > 0$. (The classical cases correspond to $\beta = 1, 2, 4$, based on the real dimension of the division algebra of elements.) These generalized eigenvalue distributions are certainly interesting in their own right. For one, they are the Boltzmann factors for certain log gases, in which the parameter β plays the role of inverse temperature, $\beta = \frac{1}{kT}$.

This dissertation is particularly concerned with extending *local statistics* of random eigenvalues to general $\beta > 0$. Local statistics and their applications to physics, statistics, combinatorics, and number theory have long been a major motivation for studying random matrix theory. In particular, the spacings between consecutive eigenvalues of a random matrix appear to resemble the spacings between consecu-

tive energy levels of slow neutron resonances [22] as well as the spacings between the critical zeros of the Riemann zeta function [24]. Also, the largest eigenvalue of a random matrix resembles the length of the longest increasing subsequence of a random permutation [2]. The past two decades have seen an explosion of analytical results concerning local statistics for $\beta = 1, 2, 4$. In 1980, Jimbo et al discovered a connection between random eigenvalues and Painlevé transcendents [16], which Tracy and Widom used over a decade later to develop a method for computing exact formulas for local eigenvalue statistics [34, 35, 36, 37, 38]. With their method, Tracy and Widom discovered exact formulas for the distributions of (1) the largest eigenvalue, (2) the smallest singular value, and (3) the spacing between consecutive eigenvalues,[1] for commonly studied random matrices, notably the Hermite, or Gaussian, ensembles and the Laguerre, or Wishart, ensembles. With contributions from Forrester [12, 13], Dyson's threefold way [8] was extended to local eigenvalue statistics: exact formulas were found for the three cases $\beta = 1, 2, 4$. These were extraordinary results which breathed new life into the field of random matrix theory. Unfortunately, these techniques have not produced results for general $\beta > 0$.

The stochastic operator approach is a new avenue into the general β realm. Each of the stochastic differential operators considered in this dissertation involves an additive term of the form σW, in which σ is a scalar and W is a diagonal operator that injects white noise. In each case, σ is proportional to $\frac{1}{\sqrt{\beta}}$, so that the variance of the noise is proportional to $kT = \frac{1}{\beta}$. This connection between β and variance is more explicit and more natural than in classical random matrix models, where β is the dimension of the elements' division algebra.

In the final chapter, the stochastic operator approach is successfully employed to derive a new quantity regarding general β local eigenvalue statistics. Specifically, an asymptotic expansion, valid near $\beta = \infty$, for the mean of the universal largest eigenvalue distribution is derived.

[1] These universal statistics arise from the Airy, Bessel, and sine kernels, respectively, in the work of Tracy and Widom.

2.1 Random matrix ensembles and scaling limits

The three *classical ensembles* of random matrix theory are Jacobi, Laguerre, and Hermite, with the following densities on $\mathbb{R}^n$:

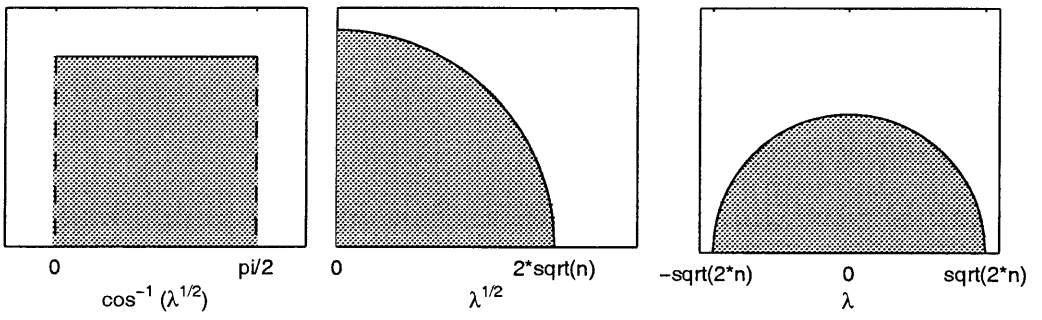
Ensemble	Joint density (up to a constant factor)	Domain
Jacobi	$\prod_{i=1}^n \lambda_i^{\frac{\beta}{2}(a+1)-1}(1-\lambda_i)^{\frac{\beta}{2}(b+1)-1}\prod_{1\le i<j\le n}\lvert\lambda_i-\lambda_j\rvert^\beta$	$(0,1)$
Laguerre	$e^{-\frac{\beta}{2}\sum_{i=1}^n\lambda_i}\prod_{i=1}^n\lambda_i^{\frac{\beta}{2}(a+1)-1}\prod_{1\le i<j\le n}\lvert\lambda_i-\lambda_j\rvert^\beta$	$(0,\infty)$
Hermite	$e^{-\frac{\beta}{2}\sum_{i=1}^n\lambda_i^2}\prod_{1\le i<j\le n}\lvert\lambda_i-\lambda_j\rvert^\beta$	$(-\infty,\infty)$

In each case, β may take any positive value, and a and b must be greater than -1.

The name *random matrix ensemble* at times seems like a misnomer; these densities have a life of their own outside of linear algebra, describing, for example, the stationary distributions of three log gases—systems of repelling particles subject to Brownian-like fluctuations. In this statistical mechanics interpretation, the densities are usually known as Boltzmann factors, and β plays the role of inverse temperature, $\beta = \frac{1}{kT}$.

Curiously, though, these densities also describe the eigenvalues, singular values, and CS values of very natural random matrices. (See Definition 3.1.4 for the CS decomposition.) For example, the eigenvalues of $\frac{1}{2}(N+N^T)$, in which N is a matrix with independent real standard Gaussian entries, follow the Hermite law with $\beta = 1$, while the eigenvalues of $\frac{1}{2}(N+N^*)$, in which N has complex entries, follow the Hermite law with $\beta = 2$. Similarly, the singular values of a matrix with independent real (resp., complex) standard Gaussian entries follow the Laguerre law with $\beta = 1$ (resp., $\beta = 2$) after a change of variables. (The parameter a is the number of rows minus the number of columns.) If the eigenvalues or singular values or CS values of a random matrix follow one of the ensemble laws, then the matrix distribution is often called a *random matrix model*.

This dissertation is particularly concerned with large n asymptotics of the ensembles. One of the first experiments that comes to mind is (1) sample an n-particle



(a) Jacobi (b) Laguerre (c) Hermite

Figure 2.1.1: Level densities, $n \to \infty$, β, a, b fixed. The level density is the density of a randomly chosen particle.

configuration of an ensemble, and (2) draw a histogram of the positions of all n particles. (The *particles* of an ensemble are simply points on the real line whose positions are given by the entries of a sample vector from the ensemble.) Actually, the plots look a little nicer if $\{\cos^{-1}(\sqrt{\lambda_i})\}_{i=1,\ldots,n}$ is counted in the Jacobi case and $\{\sqrt{\lambda_i}\}_{i=1,\ldots,n}$ is counted in the Laguerre case, for reasons that will be apparent later. As $n \to \infty$, the histograms look more and more like the plots in Figure 2.1.1, that is, a step function, a quarter-ellipse, and a semi-ellipse, respectively. In the large n limit, the experiment just described is equivalent to finding the marginal distribution of a single particle, chosen uniformly at random from the n particles. (This is related to *ergodicity*.) This marginal distribution is known as the *level density* of an ensemble, for its applications to the energy levels of heavy atoms.

For our purposes, the level density is a guide, because we are most interested in local behavior. In the level density, all n particles melt into one continuous mass. In order to see individual particles as $n \to \infty$, we must "zoom in" on a particular location. The hypothesis of *universality* in random matrix theory states that the local behavior then observed falls into one of three categories, which can be predicted in the Jacobi, Laguerre, and Hermite cases by the form of the level density at that location. Before explaining this statement, let us look at a number of examples. Each example describes an experiment involving one of the classical ensembles and the limiting

behavior of the experiment as the number of particles approaches infinity. Bear with us; seven examples may appear daunting, but these seven cases play a pivotal role throughout the dissertation.

Example 2.1.1. Let $\lambda_{\max}$ denote the position of the rightmost particle of the Hermite ensemble. As $n \to \infty$, $\sqrt{2}n^{1/6}(\lambda_{\max} - \sqrt{2n})$ converges in distribution to one of the *universal largest eigenvalue distributions*.[2] Cases $\beta = 1, 2, 4$ are plotted in Figure 2.1.2(a).[3]

Example 2.1.2. Let $\lambda_{\max}$ denote the position of the rightmost particle of the Laguerre ensemble. As $n \to \infty$, $2^{-4/3}n^{-1/3}(\lambda_{\max} - 4n)$ converges in distribution to one of the universal largest eigenvalue distributions, just as in the previous example.

Example 2.1.3. Let $\lambda_{\min}$ denote the position of the leftmost particle of the Laguerre ensemble. As $n \to \infty$, $2\sqrt{n}\sqrt{\lambda_{\min}}$ converges in distribution to one of the *universal smallest singular value distributions*.[4] The cases $\beta = 1, 2, 4$ for $a = 0$ are plotted in Figure 2.1.2(b).

Example 2.1.4. Let $\lambda_{\min}$ denote the position of the leftmost particle of the Jacobi ensemble. As $n \to \infty$, $2n\sqrt{\lambda_{\min}}$ converges in distribution to one of the universal smallest singular value distributions, just as in the previous example. The precise distribution depends on a and β, but b is irrelevant in this scaling limit. However, if we look at $2n\sqrt{1 - \lambda_{\max}}$ instead, then we see the universal smallest singular value distribution with parameters b and β.

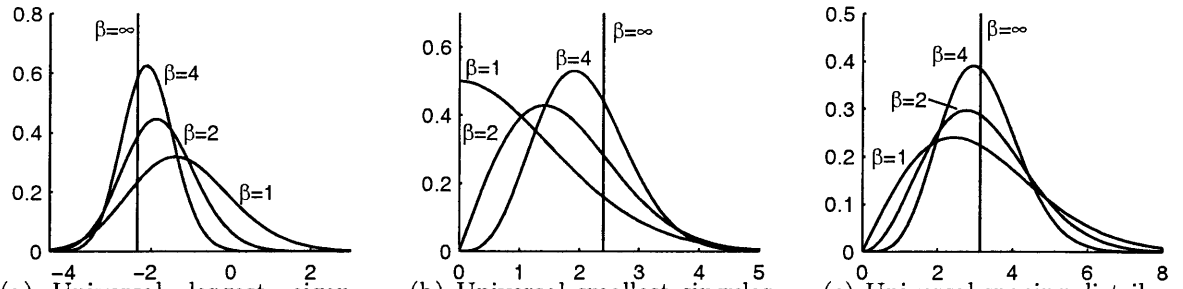
Example 2.1.5. Fix some $x \in (-1, 1)$, and find the particle of the Hermite ensemble just to the right of $x\sqrt{2n}$. Then compute the distance to the next particle to the right, and rescale this random variable by multiplying by $\sqrt{2n(1 - x^2)}$. As $n \to \infty$, the distribution converges to one of the *universal spacing distributions*.[5] The cases

[2]These are the distributions studied in [35]. See Subsection 3.7.1.

[3]In Examples 2.1.1–2.1.7, the behavior is not completely understood when $\beta \neq 1, 2, 4$. (This is, in fact, a prime motivation for the stochastic operator approach.) See Chapter 9 for a discussion of what is known about the general β case.

[4]These are the distributions studied in [36]. See Subsection 3.7.2.

[5]These are the distributions studied in [34]. See Subsection 3.7.3.



(a) Universal largest eigenvalue distributions (soft edge).

(b) Universal smallest singular value distributions for $a = 0$ (hard edge).

(c) Universal spacing distributions (bulk).

Figure 2.1.2: Local statistics of random matrix theory. See Section 3.7. Exact formulas are not known for other values of β.

$\beta = 1, 2, 4$ are plotted in Figure 2.1.2(c).

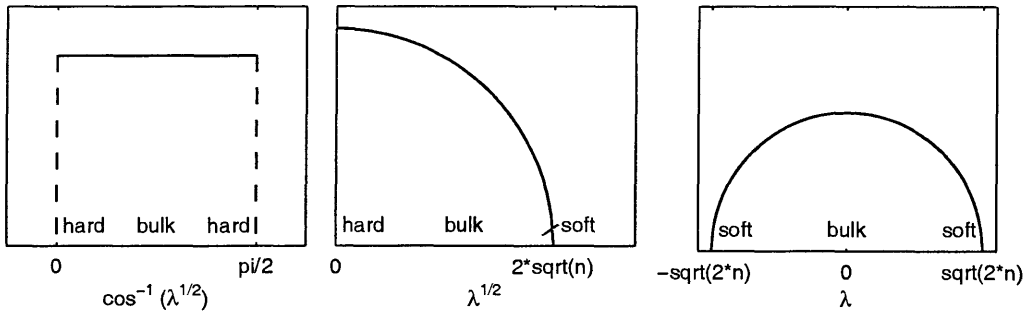
Example 2.1.6. Fix some $x \in (0, 1)$, and find the particle of the Laguerre ensemble just to the right of $4nx^2$, and then the particle just to the right of that one. Compute the difference $\sqrt{\lambda_+} - \sqrt{-\lambda_-}$, where λ_+, λ_- are the positions of the two particles, and rescale this random variable by multiplying by $2\sqrt{n(1-x^2)}$. As $n \to \infty$, the distribution converges to one of the universal spacing distributions, just as in the previous example.

Example 2.1.7. Fix some $x \in (0, \frac{\pi}{2})$, and find the particle of the Jacobi ensemble just to the right of $(\cos x)^2$, and then the particle just to the right of that one. Compute the difference $\cos^{-1}\sqrt{\lambda_-} - \cos^{-1}\sqrt{\lambda_+}$, where λ_+, λ_- are the positions of the two particles, and rescale this random variable by multiplying by $2n$. As $n \to \infty$, the distribution converges to one of the universal spacing distributions, just as in the previous two examples.

The universal distributions in Figure 2.1.2 are defined explicitly in Section 3.7.

The technique in each example of recentering and/or rescaling an ensemble as $n \to \infty$ to focus on one or two particles is known as taking a *scaling limit*.

Although there are three different ensembles, and we considered seven different scaling limits, we only saw three families of limiting distributions. This is univer-



(a) Jacobi (b) Laguerre (c) Hermite

Figure 2.1.3: Local behavior depends on the location in the global ensemble.

sality in action. Arguably, the behavior was predictable from the level densities. In the first two examples, the scaling limits focused on the right edges of the Hermite and Laguerre ensembles, where the level densities have square root branch points. In both cases, the universal largest eigenvalue distributions were observed. In the next two examples, the scaling limit focused on the left edges of the Laguerre and Jacobi ensembles, where the level densities have jump discontinuities. In both cases, the universal smallest singular value distributions were observed. In the final three examples, the scaling limit focused on the middle of the Hermite, Laguerre, and Jacobi ensembles, where the level densities are differentiable. In these three cases, the universal spacing distributions were observed.

Hence, scaling limits are grouped into three families, by the names of *soft edge*, *hard edge*, and *bulk*. The left and right edges of the Hermite ensemble, as well as the right edge of the Laguerre ensemble, are soft edges. The left and right edges of the Jacobi ensemble, as well as the left edge of the Laguerre ensemble, are hard edges. The remaining scaling limits, in the middle as opposed to an edge, are bulk scaling limits. Scaling at a soft edge, one sees a universal largest eigenvalue distribution; scaling at a hard edge, one sees a universal smallest singular value distribution; and scaling in the bulk, one sees a universal spacing distribution. The scaling regimes for the classical ensembles are indicated in Figure 2.1.3.

The central thesis of this dissertation is that the local behavior of the ensembles explored in Examples 2.1.1–2.1.7 can be modeled by the eigenvalues of stochastic differential operators. The stochastic differential operators are discovered by interpreting scaling limits of structured random matrix models as finite difference approximations.

2.2 Results

Matrix models for the three ensembles are displayed in Figures 1.0.1–1.0.3. The square Laguerre model and the Hermite model appeared in [6], but the Jacobi model is an original contribution of this thesis. (But see [19] for related work.)

Contribution 1 (The Jacobi matrix model). *When a CS decomposition of the Jacobi matrix model is taken,*

$$J_{a,b}^{\beta} = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T,$$

the diagonal entries of C, squared, follow the law of the β-Jacobi ensemble. Hence, $J_{a,b}^{\beta}$ "models" the Jacobi ensemble.

See Definition 3.1.4 for a description of the CS decomposition. The Jacobi matrix model is developed in Chapter 5. The need for a Jacobi matrix model and the history of the problem are discussed there.

The differential operators $\mathcal{A}$, $\mathcal{J}_a$, and $\tilde{\mathcal{J}}_a$ in Figures 1.0.4–1.0.6 have not been central objects of study in random matrix theory, but we propose that they should be.

Contribution 2 (Universal eigenvalue statistics from stochastic differential operators). *Let W denote a diagonal operator that injects white noise. We argue that*

- *The rightmost eigenvalue of $\mathcal{A} + \frac{2}{\sqrt{\beta}}W$ follows a universal largest eigenvalue distribution (Figure 2.1.2(a)).*

- *The smallest singular value of* $\mathcal{J}_a + \frac{2}{\sqrt{\beta}}W$ *follows a universal smallest singular value distribution (Figure 2.1.2(b)), as does the smallest singular value of* $\tilde{\mathcal{J}}_a + \sqrt{\frac{2}{\beta}}\frac{1}{\sqrt{y}}W$.

- *The spacings between consecutive eigenvalues of*

$$\begin{bmatrix} & \mathcal{J}_{-1/2} \\ \mathcal{J}_{-1/2}^* & \end{bmatrix} + \frac{1}{\sqrt{\beta}} \begin{bmatrix} 2W_{11} & \sqrt{2}W_{12} \\ \sqrt{2}W_{12} & 2W_{22} \end{bmatrix}$$

  follow a universal spacing distribution (Figure 2.1.2(c)). Here, W_{11}, W_{12}, W_{22} *denote independent noise operators.*

See Chapter 8 for more details, in particular, the role of boundary conditions. These claims are supported as follows.

- Scaling limits of the matrix models are shown to be finite difference approximations to the stochastic differential operators. Therefore, the eigenvalues of the stochastic operators should follow the $n \to \infty$ limiting distributions of the eigenvalues of the matrix models.

- In the zero temperature case ($\beta = \infty$), this behavior is confirmed: The eigenvalues and eigenvectors of the zero temperature matrix models are shown to approximate the eigenvalues and eigenvectors of the zero temperature stochastic differential operators.

- Numerical experiments are presented which support the claims.

Above, we should write "eigenvalues/singular values/CS values" in place of "eigenvalues."

Finding the correct interpretation of the noise operator W, e.g., Itô or Stratonovich, is an open problem. Without an interpretation, numerical experiments can only go so far. Our experiments are based on finite difference approximations to the stochastic

operators and show that eigenvalue behavior is fairly robust: different finite difference approximations to the same stochastic operator appear to have the same eigenvalue behavior as $n \to \infty$.

The method of studying random matrix eigenvalues through stochastic differential operators will be called the *stochastic operator approach to random matrix theory*. The approach is introduced in this dissertation, although a hint of the stochastic Airy operator appeared in [5]. The stochastic operator approach is developed in Chapters 6–8.

We close with an application of the stochastic operator approach. This application concerns large β asymptotics for the universal largest eigenvalue distributions.

Contribution 3 (Large β asymptotics for the universal largest eigenvalue distributions). *We argue that the mean of the universal largest eigenvalue distribution of parameter β is*

$$\zeta_1 + \frac{1}{\beta}\left(-4\int_0^\infty G_1(t,t)(v_1(t))^2 dt\right) + O\left(\frac{1}{\beta^2}\right) \qquad (\beta \to \infty),$$

in which ζ_1 is the rightmost zero of Ai, $v_1(t) = \frac{1}{\mathrm{Ai}'(\zeta_1)}\mathrm{Ai}(t+\zeta_1)$, *and $G_1(s,t)$ is a Green's function for the translated Airy operator $\frac{d^2}{dx^2} - x - \zeta_1$.*

These large β asymptotics are developed in Chapter 9, where a closed form expression for the Green's function can be found. The development is based on Contribution 2, and the result is verified independently by comparing with known means for $\beta = \infty, 4, 2, 1$.

2.3 Organization

The dissertation is organized into three parts.

1. Finite random matrix theory.

2. Finite-to-infinite transition.

3. Infinite random matrix theory.

Chapters 4 and 5 deal with finite random matrix theory. They unveil the Jacobi matrix model for the first time and prove identities involving the matrix models. Chapters 6 through 8 make the finite-to-infinite transition. They show that scaling limits of structured random matrix models converge to stochastic differential operators, in the sense of finite difference approximations. Chapter 9 applies the stochastic operator approach to derive asymptotics for the universal largest eigenvalue distribution. This chapter is unusual in the random matrix literature, because there are no matrices! The work begins and ends in the infinite realm, avoiding taking an $n \to \infty$ limit by working with stochastic differential operators instead of finite matrices.

Chapter 3

Background

This chapter covers known results that will be used in the sequel.

3.1 Matrix factorizations

There are three closely related matrix factorizations that will play crucial roles. These are eigenvalue decomposition, singular value decomposition, and CS decomposition. For our purposes, these factorizations should be unique, so we define them carefully.

Definition 3.1.1. Let A be an n-by-n Hermitian matrix, and suppose that A has n distinct eigenvalues. Then the *eigenvalue decomposition* of A is uniquely defined as $A = Q\Lambda Q^*$, in which Λ is real diagonal with increasing entries, and Q is a unitary matrix in which the last nonzero entry of each column is real positive.

When A is real symmetric, Q is real orthogonal.

Our definition of the singular value decomposition, soon to be presented, only applies to full rank m-by-n matrices for which either $n = m$ or $n = m + 1$, which is certainly an odd definition to make. However, these two cases fill our needs, and in these two cases, the SVD can easily be made unique. In other cases, choosing a basis for the null space becomes an issue.

Definition 3.1.2. Let A be an m-by-n complex matrix, with either $n = m$ or $n = m + 1$, and suppose that A has m distinct singular values. Then the *singular value decomposition* (SVD) of A is uniquely defined as $A = U\Sigma V^*$, in which Σ is an m-by-n nonnegative diagonal matrix with increasing entries on the main diagonal, U and V are unitary, and the last nonzero entry in each column of V is real positive.

When A is real, U and V are real orthogonal.

CS decomposition is perhaps less familiar than eigenvalue and singular value decompositions, but it has the same flavor. In fact, the CS decomposition provides SVD's for various submatrices of a unitary matrix. A proof of the following proposition can be found in [27].

Proposition 3.1.3. *Let X be an m-by-m unitary matrix, and let p, q be nonnegative integers such that $p \geq q$ and $p+q \leq m$. Then there exist unitary matrices U_1 (p-by-p), U_2 ($(m-p)$-by-$(m-p)$), V_1 (q-by-q), and V_2 ($(m-q)$-by-$(m-q)$) such that*

$$X = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \left[\begin{array}{c|c|c} C & S & \\ & I_{p-q} & \\ \hline -S & C & \\ \hline & & I_{m-p-q} \end{array}\right] \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^*, \tag{3.1.1}$$

with C and S q-by-q nonnegative diagonal. The relationship $C^2+S^2 = I$ is guaranteed.

Definition 3.1.4. Assume that in the factorization (3.1.1), the diagonal entries of C are distinct. Then the factorization is made unique by imposing that the diagonal entries of C are increasing and that the last nonzero entry in each column of $V_1 \oplus V_2$ is real positive. This factorization is known as the *CS decomposition* of X (with partition size p-by-q), and the entries of C will be called the (p-by-q) *CS values* of X.

This form of CSD is similar to the "Davis-Kahan-Stewart direct rotation form" of [27].

The random matrices defined later have distinct eigenvalues, singular values, or CS values, as appropriate, with probability 1, so the fact that the decompositions are not defined in degenerate cases can safely be ignored.

3.2 Airy and Bessel functions

Airy's Ai function satisfies the differential equation $\mathrm{Ai}''(x) = x\,\mathrm{Ai}(x)$, and is the unique solution to this equation that decays at $+\infty$. Bi is a second solution to the differential equation. See [1] or [25] for its definition.

The Bessel function of the first kind of order a, for $a > -1$, is the unique solution to

$$x^2\frac{d^2f}{dx^2} + x\frac{df}{dx} + (x^2 - a^2)f = 0$$

that is on the order of x^a as $x \to 0$. It is denoted j_a.

3.3 Orthogonal polynomial systems

3.3.1 Definitions and identities

Orthogonal polynomials are obtained by orthonormalizing the sequence of monomials $1, x, x^2, \ldots$ with respect to an inner product defined by a weight, $\langle f, g\rangle = \int_c^d fgw$. This dissertation concentrates on the three classical cases, Jacobi, Laguerre, and Hermite. For Jacobi, the weight function $w^J(a, b; x) = x^a(1-x)^b$ is defined over the interval $(0, 1)$, and a and b may take any real values greater than -1. (Note that many authors work over the interval $(-1, 1)$, modifying the weight function accordingly.) For Laguerre, the weight function $w^L(a; x) = e^{-x}x^a$ is defined over $(0, \infty)$, and a may take any real value greater than -1. For Hermite, the weight function $w^H(x) = e^{-x^2}$ is defined over the entire real line. The associated orthogonal polynomials will be

denoted by $\pi_n^J(a,b;\cdot)$, $\pi_n^L(a;\cdot)$, and $\pi_n^H(\cdot)$, respectively, so that

$$\int_0^1 \pi_m^J(a,b;x)\pi_n^J(a,b;x)x^a(1-x)^b dx = \delta_{mn}, \tag{3.3.1}$$

$$\int_0^\infty \pi_m^L(a;x)\pi_n^L(a;x)e^{-x}x^a dx = \delta_{mn}, \tag{3.3.2}$$

$$\int_{-\infty}^\infty \pi_m^H(x)\pi_n^H(x)e^{-x^2} dx = \delta_{mn}. \tag{3.3.3}$$

Signs are chosen so that the polynomials increase as $x \to \infty$. The notation $\psi_n^J(a,b;x) = \pi_n^J(a,b;x)w^J(a,b;x)^{1/2}$, $\psi_n^L(a;x) = \pi_n^L(a;x)w^L(a;x)^{1/2}$, $\psi_n^H(x) = \pi_n^H(x)w^H(x)^{1/2}$ will also be used for convenience. These functions will be known as the Jacobi, Laguerre, and Hermite *functions*, respectively.

All three families of orthogonal polynomials satisfy recurrence relations. For Jacobi, [1, (22.7.15–16,18–19)]

$$\begin{aligned}\sqrt{x}\cdot\psi_n^J(a,b;x) &= -\sqrt{\tfrac{b+n}{a+b+2n}}\sqrt{\tfrac{n}{1+a+b+2n}}\psi_{n-1}^J(a+1,b;x)\\ &\quad+\sqrt{\tfrac{a+(n+1)}{a+b+2(n+1)}}\sqrt{\tfrac{1+a+b+n}{1+a+b+2n}}\psi_n^J(a+1,b;x),\end{aligned} \tag{3.3.4}$$

$$\begin{aligned}\sqrt{x}\cdot\psi_n^J(a+1,b;x) &= \sqrt{\tfrac{a+(n+1)}{a+b+2(n+1)}}\sqrt{\tfrac{1+a+b+n}{1+a+b+2n}}\psi_n^J(a,b;x)\\ &\quad-\sqrt{\tfrac{b+(n+1)}{a+b+2(n+1)}}\sqrt{\tfrac{(n+1)}{1+a+b+2(n+1)}}\psi_{n+1}^J(a,b;x),\end{aligned} \tag{3.3.5}$$

$$\begin{aligned}\sqrt{1-x}\cdot\psi_n^J(a,b;x) &= \sqrt{\tfrac{a+n}{a+b+2n}}\sqrt{\tfrac{n}{1+a+b+2n}}\psi_{n-1}^J(a,b+1;x)\\ &\quad+\sqrt{\tfrac{b+(n+1)}{a+b+2(n+1)}}\sqrt{\tfrac{1+a+b+n}{1+a+b+2n}}\psi_n^J(a,b+1;x),\end{aligned} \tag{3.3.6}$$

$$\begin{aligned}\sqrt{1-x}\cdot\psi_n^J(a,b+1;x) &= \sqrt{\tfrac{b+(n+1)}{a+b+2(n+1)}}\sqrt{\tfrac{1+a+b+n}{1+a+b+2n}}\psi_n^J(a,b;x)\\ &\quad+\sqrt{\tfrac{a+(n+1)}{a+b+2(n+1)}}\sqrt{\tfrac{(n+1)}{1+a+b+2(n+1)}}\psi_{n+1}^J(a,b;x).\end{aligned} \tag{3.3.7}$$

When $n = 0$, equations (3.3.4) and (3.3.6) hold after dropping the term involving $\psi_{n-1}^J(a+1,b;x)$.

For Laguerre, [1, (22.7.30–31)]

$$\sqrt{x}\psi_n^L(a;x) = -\sqrt{n}\psi_{n-1}^L(a+1;x) + \sqrt{a+n+1}\psi_n^L(a+1;x), \tag{3.3.8}$$

$$\sqrt{x}\psi_n^L(a+1;x) = \sqrt{a+n+1}\psi_n^L(a;x) - \sqrt{n+1}\psi_{n+1}^L(a;x). \tag{3.3.9}$$

When $n = 0$, the first equation holds after dropping the term involving $\psi_{n-1}^L(a+1;x)$.

For Hermite, [33, (5.5.8)]

$$x\psi_n^H(x) = \sqrt{\tfrac{n}{2}}\psi_{n-1}^H(x) + \sqrt{\tfrac{n+1}{2}}\psi_{n+1}^H(x). \tag{3.3.10}$$

When $n = 0$, the equation holds after dropping the term involving $\psi_{n-1}^H(x)$.

In Section 4.2, we show that some familiar identities involving orthogonal polynomials have random matrix model analogues, not previously observed. The orthogonal polynomial identities [33, (4.1.3), (5.3.4), (5.6.1)], [43, 05.01.09.0001.01] are

$$\pi_n^J(a,b;x) = (-1)^n\pi_n^J(b,a;1-x), \tag{3.3.11}$$

$$\pi_n^L(a;x) = \lim_{b\to\infty} b^{-\frac{1}{2}(a+1)}\pi_n^J(a,b;\tfrac{1}{b}x), \tag{3.3.12}$$

$$\pi_{2n}^H(x) = (-1)^n 2^n \pi^{-1/4}\sqrt{\tfrac{\Gamma(n+\frac{1}{2})\Gamma(n+1)}{\Gamma(2n+1)}}\pi_n^L(-\tfrac{1}{2};x^2), \tag{3.3.13}$$

$$\pi_{2n+1}^H(x) = (-1)^n 2^{n+1/2}\pi^{-1/4}\sqrt{\tfrac{\Gamma(n+1)\Gamma(n+\frac{3}{2})}{\Gamma(2n+2)}}x\pi_n^L(\tfrac{1}{2};x^2), \tag{3.3.14}$$

$$\pi_n^H(x) = \lim_{a\to\infty} a^{-n/2}\pi^{-1/4}\sqrt{\Gamma(a+n+1)}\cdot\pi_n^L(a;a-\sqrt{2a}x). \tag{3.3.15}$$

The Jacobi, Laguerre, and Hermite *kernels* are defined by

$$K_n^J(a,b;x,y) = \sum_{k=0}^{n}\psi_k^J(a,b;x)\psi_k^J(a,b;y),$$

$$K_n^L(a;x,y) = \sum_{k=0}^{n}\psi_k^L(a;x)\psi_k^L(a;y),$$

$$K_n^H(x,y) = \sum_{k=0}^{n}\psi_k^H(x)\psi_k^H(y).$$

3.3.2 Orthogonal polynomial asymptotics

The following large n asymptotics follow immediately from Theorems 8.21.8, 8.21.12, 8.22.4, 8.22.6, 8.22.8, and 8.22.9 of [33].

$$\frac{1}{\sqrt{2n}}\psi_n^J(a,b;x) = j_a(\sqrt{t}) + O(n^{-1}), \qquad x = \tfrac{1}{4n^2}t, \tag{3.3.16}$$

$$\frac{1}{\sqrt{2n}}\psi_n^J(a,b;x) = j_b(\sqrt{t}) + O(n^{-1}), \qquad x = 1 - \tfrac{1}{4n^2}t, \tag{3.3.17}$$

$$\frac{(-1)^m\sqrt{\pi}}{2}\psi_{2m}^J(a,b;x) = \cos\left(\tfrac{(-a+b)\pi}{4} + \pi t\right) + O(m^{-1}), \qquad x = \tfrac{1}{2} + \tfrac{\pi}{2(2m)}t, \tag{3.3.18}$$

$$\frac{(-1)^{m+1}\sqrt{\pi}}{2}\psi_{2m+1}^J(a,b;x) = \sin\left(\tfrac{(-a+b)\pi}{4} + \pi t\right) + O(m^{-1}), \qquad x = \tfrac{1}{2} + \tfrac{\pi}{2(2m+1)}t, \tag{3.3.19}$$

$$\psi_n^L(a;x) = j_a(\sqrt{t}) + O(n^{-2}), \qquad x = \tfrac{1}{4n}t, \tag{3.3.20}$$

$$(-1)^n(2n)^{1/3}\psi_n^L(a;x) = \mathrm{Ai}(t) + O(n^{-2/3}), \qquad x = 4n + 2a + 2 + 2^{4/3}n^{1/3}t, \tag{3.3.21}$$

$$2^{n/2}\pi^{1/4}\tfrac{\Gamma(n/2+1)}{\Gamma(n+1)^{1/2}}\psi_n^H(x) = \cos\left(t - \frac{n\pi}{2}\right) + O(n^{-1}), \qquad x = \tfrac{1}{\sqrt{2n+1}}t, \tag{3.3.22}$$

$$2^{-1/4}n^{1/12}\psi_n^H(x) = \mathrm{Ai}(t) + O(n^{-2/3}), \qquad x = \sqrt{2n+1} + 2^{-1/2}n^{-1/6}t. \tag{3.3.23}$$

In (3.3.16,3.3.17,3.3.20), the error term is uniform for t in compact subsets of the positive half-line. In (3.3.18,3.3.19,3.3.21,3.3.22,3.3.23), the error term is uniform for t in compact subsets of the real line.

3.3.3 Zero asymptotics

Asymptotics for zeros of orthogonal polynomials follow. $z_{n1} < z_{n2} < \cdots < z_{nn}$ denote the zeros of the indicated polynomial.

Polynomial	Zero asymptotics ($n \to \infty$)	Eq. in [33]
$\pi_n^J(a,b;\cdot)$	$z_{nk} \sim \frac{1}{4}n^{-2}\zeta_k^2$, $0 < \zeta_1 < \zeta_2 < \ldots$ zeros of j_a	(6.3.15)
$\pi_n^J(a,b;\cdot)$	$z_{n,n+1-k} \sim 1 - \frac{1}{4}n^{-2}\zeta_k^2$, $0 < \zeta_1 < \zeta_2 < \ldots$ zeros of j_b	(6.3.15)
$\pi_n^J(a,b;\cdot)$, n even	$z_{n,K_n+k} \sim \frac{1}{2} + \frac{\pi}{2n}(\frac{a-b}{4} + \frac{1}{2} + k)$	(8.9.8)
$\pi_n^J(a,b;\cdot)$, n odd	$z_{n,K_n+k} \sim \frac{1}{2} + \frac{\pi}{2n}(\frac{a-b}{4} + k)$	(8.9.8)
$\pi_n^L(a;\cdot)$	$z_{nk} \sim \frac{1}{4}n^{-1}\zeta_k^2$, $0 < \zeta_1 < \zeta_2 < \ldots$ zeros of j_a	(6.31.6)
$\pi_n^L(a;\cdot)$	$\sqrt{z_{n,n+1-k}} \sim 2\sqrt{n} + 2^{-2/3}n^{-1/6}\zeta_k$, $0 > \zeta_1 > \zeta_2 > \ldots$ zeros of Ai	(6.32.4)
$\pi_n^H(\cdot)$, n even	$z_{n,n/2+k} \sim \frac{\pi}{\sqrt{2n}}(k - \frac{1}{2})$	(6.31.16)
$\pi_n^H(\cdot)$, n odd	$z_{n,\lceil n/2\rceil+k} \sim \frac{\pi}{\sqrt{2n}}k$	(6.31.17)
$\pi_n^H(\cdot)$	$z_{n,k} \sim -\sqrt{2n} - \frac{1}{\sqrt{2}}n^{-1/6}\zeta_k$, $0 > \zeta_1 > \zeta_2 > \ldots$ zeros of Ai	(6.32.5)
$\pi_n^H(\cdot)$	$z_{n,n+1-k} \sim \sqrt{2n} + \frac{1}{\sqrt{2}}n^{-1/6}\zeta_k$, $0 > \zeta_1 > \zeta_2 > \ldots$ zeros of Ai	(6.32.5)

In the third and fourth rows of the table, $\{K_n\}_{n=1,2,\ldots}$ is a sequence of integers independent of k. K_n should be asymptotic to $\lfloor \frac{n}{2} \rfloor$, but determining K_n precisely requires uniform asymptotics that are not available in Szegő's book.

3.3.4 Kernel asymptotics

The universal largest eigenvalue distributions, the universal smallest singular value distributions, and the universal bulk spacing distributions can be expressed in terms of Fredholm determinants of integral operators. The kernels of these operators are the *Airy kernel*, the *Bessel kernel*, and the *sine kernel*, defined on $\mathbb{R}^2$, $(0,\infty)\times(0,\infty)$, and $\mathbb{R}^2$, respectively.

$$
\begin{aligned}
K^{\text{Airy}}(s,t) &= \begin{cases} \frac{\text{Ai}(s)\,\text{Ai}'(t)-\text{Ai}'(s)\,\text{Ai}(t)}{s-t} & (s\neq t) \\ -t\,\text{Ai}(t)^2+\text{Ai}'(t)^2 & (s=t) \end{cases} \\
&= \int_0^\infty \text{Ai}(s+r)\,\text{Ai}(t+r)\,dr, \\
K^{\text{Bessel}}(a;s,t) &= \begin{cases} \frac{j_a(\sqrt{s})\sqrt{t}j_a'(\sqrt{t})-\sqrt{s}j_a'(\sqrt{s})j_a(\sqrt{t})}{2(s-t)} & (s\neq t) \\ \frac{1}{4}((j_a(\sqrt{t}))^2-j_{a-1}(\sqrt{t})j_{a+1}(\sqrt{t})) & (s=t) \end{cases} \\
&= \frac{1}{4}\int_0^1 j_a(\sqrt{sr})j_a(\sqrt{tr})\,dr, \\
K^{\text{sine}}(s,t) &= \begin{cases} \frac{\sin\pi(s-t)}{\pi(s-t)} & (s\neq t) \\ 1 & (s=t) \end{cases} \\
&= \int_0^1 \sin(\pi sr)\sin(\pi tr)dr+\int_0^1 \cos(\pi sr)\cos(\pi tr)dt.
\end{aligned}
$$

The original and best known development of the Painlevé theory of universal local statistics, which leads to Figure 2.1.2, starts from the Fredholm determinants. There is a rich theory here, covered in part by [12, 22, 34, 35, 36].

Our application of the Airy, Bessel, and sine kernels is somewhat more direct. It is well known that orthogonal polynomial kernels provide the discrete weights used in Gaussian quadrature schemes. These discrete weights also play a role in the CS/singular value/eigenvalue decompositions of zero temperature random matrix models; see Chapter 6. As $n\to\infty$, the Jacobi, Laguerre, and Hermite kernels,

appropriately rescaled, converge to the Airy, Bessel, and sine kernels. We have, as $n \to \infty$,

$$K_n^J(a,b;x,y)|\tfrac{dy}{dt}| = K^{\text{Bessel}}(a;s,t) + O(n^{-1}), \qquad x = \tfrac{1}{4n^2}s, y = \tfrac{1}{4n^2}t, \tag{3.3.24}$$

$$K_n^J(a,b;x,y)|\tfrac{dy}{dt}| = K^{\text{Bessel}}(b;s,t) + O(n^{-1}), \qquad x = 1 - \tfrac{1}{4n^2}s, y = 1 - \tfrac{1}{4n^2}t, \tag{3.3.25}$$

$$K_n^J(a,b;x,y)|\tfrac{dy}{dt}| = K^{\text{sine}}(s,t) + O(n^{-1}), \qquad x = \tfrac{1}{2} + \tfrac{\pi}{2n}s, y = \tfrac{1}{2} + \tfrac{\pi}{2n}t, \tag{3.3.26}$$

$$K_n^L(a;x,y)|\tfrac{dy}{dt}| = K^{\text{Bessel}}(a;s,t) + O(n^{-2}), \qquad x = \tfrac{1}{4n}s, y = \tfrac{1}{4n}t, \tag{3.3.27}$$

$$K_n^L(a;x,y)|\tfrac{dy}{dt}| = K^{\text{Airy}}(s,t) + O(n^{-2/3}), \qquad \begin{cases} x = 4n + 2a + 2^{4/3}n^{1/3}s \\ y = 4n + 2a + 2^{4/3}n^{1/3}t \end{cases}, \tag{3.3.28}$$

$$K_n^H(x,y)|\tfrac{dy}{dt}| = K^{\text{sine}}(s,t) + O(n^{-1/2}), \qquad x = \tfrac{1}{\sqrt{2n+1}}s, y = \tfrac{1}{\sqrt{2n+1}}t, \tag{3.3.29}$$

$$K_n^H(x,y)|\tfrac{dy}{dt}| = K^{\text{Airy}}(s,t) + O(n^{-2/3}), \qquad \begin{cases} x = \sqrt{2n+1} + 2^{-1/2}n^{-1/6}s \\ y = \sqrt{2n+1} + 2^{-1/2}n^{-1/6}t \end{cases}. \tag{3.3.30}$$

In (3.3.24,3.3.25,3.3.27), the error term is uniform for s,t in compact intervals of $(0,\infty)$ [20, (1.10)],[14, (7.8)]. In (3.3.26,3.3.28,3.3.29,3.3.30), the error term is uniform for s,t in compact intervals [20, (1.9)],[14, (7.6–7)].

(3.3.28) is not stated explicitly in [14], but the proof for (3.3.30) works for the Laguerre kernel as well as the Hermite kernel. (Note that [14] actually proves a stronger uniformity result. The weaker result stated here is especially straightforward to prove.)

3.4 The Selberg integral

This subsection follows Chapter 3 of [14].

Let $\beta > 0$ and $a, b > -1$. Selberg's result [31] is

$$\int_0^1 \cdots \int_0^1 \prod_{i=1}^n \lambda_i^{\frac{\beta}{2}(a+1)-1}(1-\lambda_i)^{\frac{\beta}{2}(b+1)-1} \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|^\beta d\lambda_1 \cdots d\lambda_n$$
$$= \prod_{k=1}^n \frac{\Gamma(\frac{\beta}{2}(a+k))\Gamma(\frac{\beta}{2}(b+k))\Gamma(\frac{\beta}{2}k+1)}{\Gamma(\frac{\beta}{2}(a+b+n+k))\Gamma(\frac{\beta}{2}+1)}. \quad (3.4.1)$$

Performing the change of variables $\lambda_i = \frac{1}{b}\lambda_i'$, $i = 1, \ldots, n$, and taking the limit $b \to \infty$ gives

$$\int_0^1 \cdots \int_0^1 \prod_{i=1}^n e^{-\frac{\beta}{2}\lambda_i}\lambda_i^{\frac{\beta}{2}(a+1)-1} \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|^\beta d\lambda_1 \cdots d\lambda_n$$
$$= \left(\frac{2}{\beta}\right)^{\frac{\beta}{2}n(a+n)} \prod_{k=1}^n \frac{\Gamma(\frac{\beta}{2}(a+k))\Gamma(\frac{\beta}{2}k+1)}{\Gamma(\frac{\beta}{2}+1)}. \quad (3.4.2)$$

Performing an alternative change of variables, setting $b = a$, and letting $a \to \infty$ gives

$$\int_0^1 \cdots \int_0^1 \prod_{i=1}^n e^{-\frac{\beta}{2}\lambda_i^2} \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|^\beta d\lambda_1 \cdots d\lambda_n$$
$$= \beta^{-(\beta/2)\binom{n}{2}-n/2}(2\pi)^{n/2} \prod_{k=1}^n \frac{\Gamma(\frac{\beta}{2}k+1)}{\Gamma(\frac{\beta}{2}+1)}. \quad (3.4.3)$$

(See [14, Proposition 3.18].)

3.5 Random variable distributions

The real standard Gaussian distribution has p.d.f. $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. A complex standard Gaussian is distributed as $\frac{1}{\sqrt{2}}(G_1 + \sqrt{-1}G_2)$, in which G_1 and G_2 are independent real standard Gaussians. The chi distribution with r degrees of freedom has

p.d.f. $2^{1-r/2}\Gamma(\frac{r}{2})^{-1}x^{r-1}e^{-r^2/2}$. A random variable following this distribution is often denoted χ_r. The beta distribution with parameters c, d has p.d.f. $\frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)}x^{c-1}(1-x)^{d-1}$. A random variable following this distribution is often denoted beta(c, d).

Remark 3.5.1. If $p, q \to \infty$ in such a way that $\frac{q}{p} \to c$, $0 < c < 1$, then

$$2\sqrt{1+c}\sqrt{p}\left(\cos^{-1}\sqrt{\text{beta}(p,q)} - \cos^{-1}((1+c)^{-1/2})\right)$$

appears to approach a standard Gaussian in distribution.

It is known that the beta distribution itself approaches a Gaussian as $p, q \to \infty$ in the prescribed fashion [11]. The remark applies the change of variables $\theta = \cos^{-1}\sqrt{x}$. The Gaussian asymptotics should hold because the change of variables is nearly linear, locally. This remark is mentioned again in Section 8.2, but it is not crucial to any part of the dissertation.

3.6 Finite differences

The stochastic operator approach to random matrix theory is developed by viewing certain random matrices as finite difference approximations to differential operators. These finite difference approximations are mostly built from a few square matrices, whose sizes will usually be evident from context.

$$D_1 = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & \ddots & \\ & & & \ddots & 1 \\ & & & & -1 \end{bmatrix}, \tag{3.6.1}$$

$$D_2 = \begin{bmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{bmatrix}, \tag{3.6.2}$$

$$F = F_n = \begin{bmatrix} & & & & 1 \\ & & & 1 & \\ & & 1 & & \\ & \iddots & & & \\ 1 & & & & \end{bmatrix}, \tag{3.6.3}$$

$$\Omega = \Omega_n = \begin{bmatrix} -1 & & & & \\ & 1 & & & \\ & & -1 & & \\ & & & 1 & \\ & & & & \ddots \end{bmatrix}. \tag{3.6.4}$$

Also, let P denote a "perfect shuffle" permutation matrix, whose size will be evident from context,

$$P = P_n = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & & 1 \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \end{bmatrix} \ (n \text{ even}), \qquad P = P_n = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \end{bmatrix} \ (n \text{ odd}). \tag{3.6.5}$$

The stochastic eigenvalue problems that appear later are related to classical Sturm-Liouville problems. A reference for both the theory and numerical solution of Sturm-Liouville problems is [30].

3.7 Universal local statistics

The universal distributions in Figure 2.1.2 are expressible in terms of Painlevé transcendents. This is work pioneered by Jimbo, Miwa, Môri, and Sato in [16] and perfected by Tracy and Widom in [34, 35, 36]. The densities plotted in Figure 2.1.2 are called f_β^{soft}, $f_{\beta,a}^{\text{hard}}$, and f_β^{bulk} below. The plots were created with Per-Olof Persson's software [10], after extensions by the present author.

3.7.1 Universal largest eigenvalue distributions

Let $(\lambda_1, \ldots, \lambda_n)$ be distributed according to the β-Hermite ensemble. Consider the probability that an interval (s, ∞) contains no rescaled particle $\sqrt{2}n^{1/6}(\lambda_i - \sqrt{2n})$, i.e., the c.d.f. of the rescaled rightmost particle. As $n \to \infty$, this probability approaches a limit, which will be denoted $F_\beta^{\text{soft}}(s)$. (The same limit can be obtained by scaling at the right edge of the Laguerre ensemble [12, 17, 18].) Also define $f_\beta^{\text{soft}}(s) = \frac{d}{ds}F_\beta^{\text{soft}}(s)$, which is, of course, the density of the rightmost particle. In this subsection, exact formulas for $F_1^{\text{soft}}(s)$, $F_2^{\text{soft}}(s)$, and $F_4^{\text{soft}}(s)$ are presented. The $\beta = 2$ theory is found in [35]. The $\beta = 1, 4$ theory was originally published in [37], and a shorter exposition can be found in [38].

Let q be the unique solution to the following differential equation, which is a special case of the Painlevé II equation,

$$q'' = sq + 2q^3,$$

with boundary condition

$$q(s) \sim \text{Ai}(s) \quad (s \to \infty).$$

Then

$$\begin{aligned} F_2^{\text{soft}}(s) &= \exp\left[-\int_s^\infty (x-s)(q(x))^2 dx\right], \\ F_1^{\text{soft}}(s) &= \sqrt{F_2^{\text{soft}}(s)}\exp\left[-\frac{1}{2}\int_s^\infty q(x)dx\right], \\ F_4^{\text{soft}}(2^{-2/3}s) &= \sqrt{F_2^{\text{soft}}(s)}\cosh\left[\frac{1}{2}\int_s^\infty q(x)dx\right]. \end{aligned}$$

The plots in Figure 2.1.2(a) can be produced by differentiating these three equations to find $f_\beta^{\text{soft}}(s)$.

3.7.2 Universal smallest singular value distributions

Let $(\lambda_1, \ldots, \lambda_n)$ be distributed according to the β-Laguerre ensemble with parameter a. Consider the probability that an interval $(0, s)$ contains no rescaled particle $4n\lambda_i$. As $n \to \infty$, this probability approaches a limit, which will be denoted $E^{\text{hard}}_{\beta,a}(s)$. Of course, $1 - E^{\text{hard}}_{\beta,a}(s)$ is the c.d.f. of the rescaled leftmost particle $4n\lambda_1$. In this subsection, exact formulas for $E^{\text{hard}}_{1,a}(s)$, $E^{\text{hard}}_{2,a}(s)$, and $E^{\text{hard}}_{4,a}(s)$ are presented. This material is somewhat scattered in the literature. The $\beta = 2$ formulas are derived in [36]. The derivation starts with an expression for $E^{\text{hard}}_{\beta,a}(s)$ in terms of the Fredholm determinant of an integral operator whose kernel is K^{Bessel}, due to [12]. The extension of the formulas to $\beta = 1, 4$ is found in [13]. The limiting distributions $E^{\text{hard}}_{\beta,a}(s)$ can also be obtained by scaling the Jacobi ensemble at either edge. This fact, restricted to $\beta = 2$, can be found in [20], where more references are also cited.

The probability $E^{\text{hard}}_{\beta,a}(s)$ of a particle-free interval will be expressed in terms of the Painlevé transcendent p, satisfying the differential equation

$$s(p^2 - 1)(sp')' = p(sp')^2 + \frac{1}{4}(s - a^2)p + \frac{1}{4}sp^3(p^2 - 2)$$

and the boundary condition

$$p(s) \sim j_a(\sqrt{s}) \quad (s \to 0).$$

The differential equation is a special case of the Painlevé V equation. (Painlevé III also appears in the hard edge theory. See [13].)

The formulas below are taken from [13]. We have

$$E_{2,a}^{\text{hard}}(s) = \exp\left(-\frac{1}{4}\int_0^s (\log s/x)(p(x))^2 dx\right),$$
$$E_{1,a}^{\text{hard}}(s) = \sqrt{E_{2,a}^{\text{hard}}(s)}\exp\left(-\frac{1}{4}\int_0^s \frac{p(x)}{\sqrt{x}}dx\right),$$
$$E_{4,a}^{\text{hard}}(s/2) = \sqrt{E_{2,2a}^{\text{hard}}(s)}\cosh\left(\frac{1}{4}\int_0^s \frac{p(x)}{\sqrt{x}}dx\right).$$

In order to create the plot in Figure 2.1.2(b), a change of variables is required. Set $\sigma_1 = \sqrt{\lambda_1}$, so that $P[2\sqrt{n}\sigma_1 < t] = P[4n\lambda_1 < t^2] = 1 - E_{\beta,a}^{\text{hard}}(t^2)$. Then the density of $2\sqrt{n}\sigma_1$ at t converges to $f_{\beta,a}^{\text{hard}}(t) := -2t\left(\frac{d}{ds}E_{\beta,a}^{\text{hard}}(s)\big|_{s=t^2}\right)$ as $n \to \infty$.

3.7.3 Universal spacing distributions

Let $(\lambda_1, \ldots, \lambda_n)$ be distributed according to the β-Hermite ensemble. Consider the probability that an interval (c, d) contains no rescaled particle $\sqrt{2n}\lambda_i$. (Note that our rescaling differs from most authors' by a factor of π.) As $n \to \infty$, this probability approaches a limit that depends only on the difference $s := d - c$. The limiting probability will be denoted $E_\beta^{\text{bulk}}(s)$. Also, define $F_\beta^{\text{bulk}}(s) = -\pi\frac{d}{ds}E_\beta^{\text{bulk}}(s)$ and $f_\beta^{\text{bulk}}(s) = -\frac{d}{ds}F_\beta^{\text{bulk}}(s)$, so that $1 - F_\beta^{\text{bulk}}(s)$ is the c.d.f. of the distance from a randomly chosen particle to the particle immediately to the right, and $f_\beta^{\text{bulk}}(s)$ is the p.d.f. of this random variable. In this subsection, exact formulas for $E_1^{\text{bulk}}(s)$, $E_2^{\text{bulk}}(s)$, and $E_4^{\text{bulk}}(s)$ are presented. See [16, 22, 34] for more details. Also see [20] and references therein for proof that the same limit can be obtained by scaling in the bulk of the Laguerre and Jacobi ensembles.

Let $\sigma_V(x)$ be the unique solution to the following equation, which is the "σ representation" of the Painlevé V equation,

$$(x\sigma_V'')^2 + 4(x\sigma_V' - \sigma_V)(x\sigma_V' - \sigma_V + (\sigma_V')^2) = 0$$

with boundary condition

$$\sigma_V(x) \sim -\frac{1}{\pi}x - \frac{1}{\pi^2}x^2 \quad (x \to 0),$$

and define $D(s), D_{\pm}(s)$ by

$$D(s) = \exp\left(\int_0^s \frac{\sigma_V(x)}{x}dx\right),$$
$$\log D_{\pm}(s) = \frac{1}{2}\log D(s) \mp \frac{1}{2}\int_0^s \sqrt{-\tfrac{d^2}{dy^2}\log D(y)\Big|_{y=x}}\,dx.$$

Then

$$E_1^{\text{bulk}}(s) = D_+(s),$$
$$E_2^{\text{bulk}}(s) = D(s),$$
$$E_4^{\text{bulk}}(s) = \frac{1}{2}(D_+(2s) + D_-(2s)).$$

Once these quantities are computed, Figure 2.1.2(c) can be generated by differentiating $E_\beta^{\text{bulk}}(s)$ twice and multiplying by π to get $f_\beta^{\text{bulk}}(0; s)$.

Chapter 4

Random matrix models

4.1 The matrix models and their spectra

Matrix models for the three classical ensembles—Jacobi, Laguerre, and Hermite—are defined in Figures 1.0.1–1.0.3. The Jacobi matrix model is an original contribution of this thesis. The Laguerre and Hermite matrix models were introduced in [6] and further developed in [4]. Special cases ($\beta = 1, 2$) of the Laguerre and Hermite models appeared earlier, in [9, 32] and [40], respectively. $J_{a,b}^{\beta}$ is a random orthogonal matrix with a special structure we call *bidiagonal block form*. The angles $\theta_1, \ldots, \theta_n, \phi_1, \ldots, \phi_{n-1}$ are independent. L_a^{β} is a random real bidiagonal matrix with independent entries. H^{β} is a random real symmetric tridiagonal matrix with independent entries in the upper triangular part.

The following theorem is proved in Chapter 5.

Theorem 4.1.1. *Compute the n-by-n CS decomposition of the 2n-by-2n Jacobi matrix model,*

$$J_{a,b}^{\beta} = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T.$$

Then the CS values, squared, follow the law of the β-Jacobi ensemble with parameters

a, b. That is, if $C = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n})$, *then*

$$(\lambda_1, \ldots, \lambda_n) \sim \frac{1}{c} \prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(a+1)-1} (1-\lambda_i)^{\frac{\beta}{2}(b+1)-1} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^\beta,$$

in which the constant c is the quantity in (3.4.1).

The next two propositions were proved in [6].

Proposition 4.1.2 (Dumitriu, Edelman). *The singular values* $\sigma_1 < \cdots < \sigma_n$ *of* L_a^β, *squared, follow the law of the* β*-Laguerre ensemble with parameter a. That is, if* $\lambda_i = \sigma_i^2$, $i = 1, \ldots, n$, *then*

$$(\lambda_1, \ldots, \lambda_n) \sim \frac{1}{c} e^{-\frac{\beta}{2}\sum_{i=1}^n \lambda_i} \prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(a+1)-1} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^\beta,$$

in which the constant c is the quantity in (3.4.2).

Proposition 4.1.3 (Dumitriu, Edelman). *The eigenvalues* $\lambda_1 < \cdots < \lambda_n$ *of* H^β *follow the law of the* β*-Hermite ensemble,*

$$(\lambda_1, \ldots, \lambda_n) \sim \frac{1}{c} e^{-\frac{\beta}{2}\sum_{i=1}^n \lambda_i^2} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^\beta,$$

in which the constant c is the quantity in (3.4.3).

The rectangular Laguerre model M_a^β has not appeared before, to the best of our knowledge. It is crucial to understanding the connection between the Laguerre and Hermite ensembles.

Conjecture 4.1.4. *The singular values, squared, of the n-by-*$(n+1)$ *rectangular Laguerre matrix model* M_a^β *follow the law of the* β*-Laguerre ensemble with parameter a.*

The conjecture is straightforward to prove in the cases $\beta = 1, 2$, $a \in \mathbb{Z}$, by applying

Householder reflectors to a dense matrix with independent standard Gaussian entries, in the spirit of [6].

4.2 Identities

In this subsection, we show that several identities from orthogonal polynomial theory have random matrix model analogues.

Analogous to (3.3.11),

Theorem 4.2.1. *We have*

$$J_{a,b}^{\beta} \stackrel{d}{=} \begin{bmatrix} & \Omega \\ \Omega & \end{bmatrix} J_{b,a}^{\beta} \begin{bmatrix} -\Omega & \\ & \Omega \end{bmatrix}. \tag{4.2.1}$$

Analogous to (3.3.12),

Theorem 4.2.2. *We have*

$$\lim_{b\to\infty} \sqrt{b} J_{a,b}^{\beta} \stackrel{d}{=} \left[\begin{array}{c|c} (L_a^{\beta})^T & * \\ \hline * & L_a^{\beta} \end{array}\right].$$

The off-diagonal blocks of $\sqrt{b}J_{a,b}^{\beta}$ *do not converge as* $b \to \infty$. *The top-right block tends toward* $\mathrm{diag}(\infty, \ldots, \infty)$ *and the bottom-left block tends toward* $\mathrm{diag}(-\infty, \ldots, -\infty)$.

Proof. The beta distribution with parameters c, d is equal in distribution to $X_1/(X_1 + X_2)$, for $X_1 \sim \chi^2_{2c}$, $X_2 \sim \chi^2_{2d}$. So, in the Jacobi model,

$$\begin{aligned} c_k &\sim \sqrt{\chi^2_{\beta(a+k)}/(\chi^2_{\beta(a+k)} + \chi^2_{\beta(b+k)})}, \\ s_k &\sim \sqrt{\chi^2_{\beta(b+k)}/(\chi^2_{\beta(a+k)} + \chi^2_{\beta(b+k)})}, \\ c'_k &\sim \sqrt{\chi^2_{\beta k}/(\chi^2_{\beta k} + \chi^2_{\beta(a+b+1+k)})}, \\ s'_k &\sim \sqrt{\chi^2_{\beta(a+b+1+k)}/(\chi^2_{\beta k} + \chi^2_{\beta(a+b+1+k)})}. \end{aligned}$$

As $r \to \infty$, the mean of χ_r^2 becomes asymptotic to r, while the standard deviation is $\sqrt{2r}$. Therefore, as $b \to \infty$, $\sqrt{b}c_k \xrightarrow{d} \frac{1}{\sqrt{\beta}}\chi_{\beta(a+k)}$, $s_k \xrightarrow{d} 1$, $\sqrt{b}c'_k \xrightarrow{d} \frac{1}{\sqrt{\beta}}\chi_{\beta k}$, and $s'_k \xrightarrow{d} 1$. □

Remark 4.2.3. We have no analogous theorem for the rectangular Laguerre model. We suspect that there may be a "rectangular Jacobi model," containing rectangular blocks, from which the rectangular Laguerre model would arise after a $b \to \infty$ limit.

Analogous to (3.3.13–3.3.14),

Theorem 4.2.4. *Let H^β be the n-by-n Hermite matrix model. If n is even, then*

$$PH^\beta P^T \stackrel{d}{=} \left[\begin{array}{c|c} \frac{1}{\sqrt{\beta}}\operatorname{diag}(G_1, G_3 \ldots, G_{n-1}) & \Omega L_{-1/2}^{2\beta}\Omega \\ \hline \Omega(L_{-1/2}^{2\beta})^T\Omega & \frac{1}{\sqrt{\beta}}\operatorname{diag}(G_2, G_4, \ldots, G_n) \end{array}\right],$$

in which $G_1, \ldots, G_n$ are i.i.d. real standard Gaussians, $L_{-1/2}^{2\beta}$ is a Laguerre matrix model, independent from the Gaussians, and the matrix on the right hand side is symmetric, so that only one Laguerre matrix is "sampled." If, on the other hand, n is odd, then

$$PH^\beta P^T \stackrel{d}{=} \left[\begin{array}{c|c} \frac{1}{\sqrt{\beta}}\operatorname{diag}(G_1, G_3 \ldots, G_n) & -\Omega(M_{1/2}^{2\beta})^T\Omega \\ \hline -\Omega M_{1/2}^{2\beta}\Omega & \frac{1}{\sqrt{\beta}}\operatorname{diag}(G_2, G_4, \ldots, G_{n-1}) \end{array}\right].$$

Analogous to (3.3.15),

Theorem 4.2.5. *In the following equation, both occurrences of L_a^β represent the same random matrix sampled from the Laguerre matrix model. We have*

$$H^\beta \stackrel{d}{=} \lim_{a\to\infty} \frac{1}{\sqrt{2}}(\Omega L_a^\beta \Omega + \tilde{\Omega}(L_a^\beta)^T\Omega - 2\sqrt{a}I).$$

Proof. For $k = 1, \ldots, n$, the (k,k) entry of the right hand side, before taking the limit, is $\sqrt{\frac{2}{\beta}}(\chi_{(a+n+1-k)\beta} - \sqrt{a\beta})$. As $r \to \infty$, $\sqrt{2}(\chi_r - \sqrt{r})$ tends toward a standard Gaussian random variable [11], so that the (k,k) entry converges in distribution to

a normal random variable, $\frac{1}{\sqrt{\beta}}G_k$. For $k = 1, \ldots, n-1$, the $(k, k+1)$ and $(k+1, k)$ entries of the right hand side are precisely $\frac{1}{\sqrt{2\beta}}\chi_{(n-k)\beta}$. $\square$

Chapter 5

The Jacobi matrix model

This chapter derives the Jacobi matrix model by designing and running a numerically-inspired algorithm on a Haar-distributed orthogonal/unitary matrix. Then the model is extended from $\beta = 1, 2$ to general $\beta > 0$ in the obvious way.

While the algorithmic approach is inspired by [6], crucial pieces are novel: the connection between the Jacobi ensemble and Haar measure on the orthogonal and unitary groups is rarely, if ever, observed in the literature, and the algorithm, while simple, is certainly not widely known.

The central theme of the entire dissertation is that matrix models for the classical ensembles, appropriately rescaled, are finite difference approximations to stochastic differential operators. The present chapter strengthens support for the stochastic operator approach by providing a new, structured Jacobi matrix model. We shall see later that the new model discretizes the Bessel and sine operators from Chapter 1.

5.1 Introduction

Traditionally, the Hermite ensemble is modeled by the *eigenvalues* of a symmetric matrix with Gaussian entries, and the Laguerre ensemble is modeled by the *singular values* of a matrix with Gaussian entries. This chapter begins by showing that the

Jacobi ensemble arises from a *CS decomposition* problem. Specifically, the $\beta = 1$ Jacobi ensemble arises from the CS decomposition of a Haar-distributed orthogonal matrix, and the $\beta = 2$ Jacobi ensemble arises from the CS decomposition of a Haar-distributed unitary matrix. This observation completes the following table, and enables the development of the Jacobi matrix model seen in Figure 1.0.1.

Ensemble	Random linear algebra problem
Hermite	eigenvalue decomposition
Laguerre	singular value decomposition
Jacobi	CS decomposition

For several decades, random matrix theory concentrated on three values of β in the ensemble densities. The $\beta = 1$ ensembles were shown to arise from real random matrices, the $\beta = 2$ ensembles from complex random matrices, and the $\beta = 4$ ensembles from quaternion random matrices, according to Dyson's "threefold way" [8]. In recent years, the development of a *general* β theory, extending beyond $\beta = 1, 2, 4$ to all $\beta > 0$, has gained momentum. One of the fundamental problems in developing a general β theory is to find a random matrix distribution that "models" the desired ensemble in some fashion. Dumitriu and Edelman solved the matrix model problems for the Hermite and Laguerre ensembles [6]. In the Hermite case, for example, they provided a random symmetric tridiagonal matrix for each β whose eigenvalues follow the law of the Hermite ensemble. Dumitriu and Edelman posed the development of a β-Jacobi matrix model as an open problem, which has been considered in [19, 21].

The major contribution of this chapter is the introduction of the *Jacobi matrix model* in Figure 1.0.1. The matrix model is a distribution on structured orthogonal matrices, parameterized by $\beta > 0$, $a > -1$, and $b > -1$. Its CS decomposition has entries from the Jacobi ensemble with the same parameters. We argue that this model is *the* β-Jacobi matrix model.

The development of the model is in the spirit of [6], utilizing an algorithm inspired by bidiagonalization and tridiagonalization algorithms from numerical linear algebra.

The use of CS decomposition breaks from previous work, which has focused on eigenvalues. Notable among the existing work is that of Killip and Nenciu [19], which provides a random matrix model whose *eigenvalues* follow the law of the Jacobi ensemble. In fact, this eigenvalue model is closely related to our CS decomposition model. However, the approach in the present dissertation has the following advantages:

- Our matrix model is a random orthogonal matrix, generalizing certain features of the orthogonal and unitary groups to general β.
- CS decomposition is used in place of eigenvalue decomposition, which is natural considering that the Jacobi ensemble is a distribution on $[0,1]^n$ rather than all of $\mathbb{R}^n$. (CS values lie in $[0,1]$ by definition.)
- The matrix model has both left and right CS vectors, rather than just eigenvectors.
- The development of the matrix model is illuminating, based on a numerically-inspired algorithm.
- There is an immediate connection to multivariate analysis of variance (MANOVA), based on the connections between CS decomposition and generalized singular value decomposition.

5.1.1 Background

The Jacobi ensemble, proportional to

$$\prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(a+1)-1}(1-\lambda_i)^{\frac{\beta}{2}(b+1)-1} \prod_{i<j} |\lambda_i - \lambda_j|^{\beta},$$

has been studied extensively, motivated by applications in both physics and statistics.

In statistical mechanics, the ensemble arises in the context of log gases. A log gas is a system of charged particles on the real line that are subject to a logarithmic interaction potential as well as Brownian-like fluctuations. If the particles are constrained to the interval $[0, 1]$ and are also subject to the external potential $\sum_{i=1}^{n}(\frac{a+1}{2} - \frac{1}{\beta})\log\lambda_i + \sum_{i=1}^{n}(\frac{b+1}{2} - \frac{1}{\beta})\log(1-\lambda_i)$, then the long term stationary distribution of the system of charges is the Jacobi ensemble [7, 14, 42].

In statistics, the ensemble arises in the context of MANOVA, starting from a pair of independent Gaussian matrices N_1, N_2. If N_1 and N_2 have real entries, then their generalized singular values follow the law of the Jacobi ensemble with $\beta = 1$. If they have complex entries, then their generalized singular values follow the law of the Jacobi ensemble with $\beta = 2$. Now we define GSVD and make these statements precise.

Definition 5.1.1 (Generalized singular value decomposition (GSVD)). Let A be $(n+a)$-by-n and B be $(n+b)$-by-n with complex entries. Then there exist matrices R, U_1, U_2, V, C, and S such that

$$\left[\begin{array}{c} A \\ \hline B \end{array}\right] = \left[\begin{array}{c|c} U_1 & \\ \hline & U_2 \end{array}\right] \left[\begin{array}{c} C \\ 0 \\ \hline -S \\ 0 \end{array}\right] V^*R, \tag{5.1.1}$$

in which R is n-by-n upper triangular, U_1 is $(n+a)$-by-$(n+a)$ unitary, U_2 is $(n+b)$-by-$(n+b)$ unitary, V is n-by-n unitary, and C and S are nonnegative diagonal, satisfying $C^2 + S^2 = I$. The diagonal entries of C are known as the *generalized singular values* of the pair A, B, and the factorization in (5.1.1) is a *generalized singular value decomposition* (GSVD).

There are a few observations worth mentioning. First, this definition does not define the GSVD of a pair of matrices uniquely. Second, if A and B have real entries,

then R, U_1, U_2, and V have real entries. Third, many authors refer to the cotangents $\frac{c_k}{s_k}$, $k = 1, \ldots, n$, instead of the cosines c_k, as generalized singular values.

One way to construct a GSVD, which may not be the most numerically accurate, is to first compute a QR decomposition of $\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right]$, and then to compute SVD's for the top and bottom blocks of Q. See [41] for details.

The Jacobi ensemble can be seen in the generalized singular values of a pair of Gaussian matrices.

Proposition 5.1.2. *Let N_1 and N_2 be independent random matrices. Suppose that N_1 is $(n+a)$-by-n and N_2 is $(n+b)$-by-n, each with i.i.d. real (resp. complex) standard Gaussian entries. Then the generalized singular values, squared, of the pair N_1, N_2 follow the law of the Jacobi ensemble with parameters a, b, for $\beta = 1$ (resp. $\beta = 2$).*

Proof. The generalized singular values, squared, are equal to the eigenvalues of $N_1^* N_1 (N_1^* N_1 + N_2^* N_2)^{-1}$, which behave as the Jacobi ensemble [23]. To see this, note that $N_1^* N_1 (N_1^* N_1 + N_2^* N_2)^{-1} = N_1^* N_1 \left(\left[\begin{smallmatrix} N_1 \\ N_2 \end{smallmatrix}\right]^* \left[\begin{smallmatrix} N_1 \\ N_2 \end{smallmatrix}\right] \right)^{-1}$, so if the CSD of $\left[\begin{smallmatrix} N_1 \\ N_2 \end{smallmatrix}\right]$ is

$$\left[\begin{array}{c|c} U_1 & \\ \hline & U_2 \end{array}\right] \left[\begin{array}{c} C \\ \hline 0 \\ \hline -S \\ \hline 0 \end{array}\right] V^* R,$$

then $N_1^* N_1 (N_1^* N_1 + N_2^* N_2)^{-1} = (R^* V) C^2 (R^* V)^{-1}$. □

The preceding proposition provides matrix models for the Jacobi ensemble in the cases $\beta = 1$ and $\beta = 2$, for integral a and b. The primary contribution of this chapter is a general β matrix model, which also removes the quantization on a and b.

5.1.2 Results

We show that the Jacobi ensemble arises from Haar measure on compact matrix groups, through the CS decomposition. This viewpoint is central to the development

of the Jacobi matrix model.

There is a deep connection between CSD and GSVD. Specifically, if a unitary X is partitioned into $X = \left[\begin{smallmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{smallmatrix}\right]$, with X_{11} of size p-by-q, then the generalized singular values of the pair X_{11}, X_{21} equal the p-by-q CS values of X. This fact is evident from the definitions.

Theorem 5.1.3. *Let n, a, and b be positive integers, and define $m = 2n + a + b$. Let X be an m-by-m Haar-distributed orthogonal matrix, and take the CS decomposition of X with partition size $(n + a)$-by-n. Then the CS values of X, squared, follow the law of the $\beta = 1$ Jacobi ensemble with parameters a, b. If, instead, X is a Haar-distributed unitary matrix, then the CS values, squared, obey the law of the $\beta = 2$ Jacobi ensemble.*

Proof. Let $\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right]$ be an m-by-n matrix of independent standard Gaussian entries, with A $(n+a)$-by-n and B $(n+b)$-by-n. We claim that the CS values of X share the same distribution with the generalized singular values of the pair A, B. Upon showing this, the proof will follow by Proposition 5.1.2.

With probability 1, the generalized singular values are distinct, so we can take a QR decomposition, $\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right] = QR$, with R invertible. Next, randomize signs, $QR = (QD)(D^*R)$, using a diagonal matrix D with i.i.d. entries chosen uniformly from either $\{-1, 1\}$ (if X is real orthogonal) or the unit circle (if X is complex unitary). It is clear that QD shares the same distribution with the first n columns of X. Therefore, the CS values of X share the same distribution with the singular values of the first $n + a$ rows of QD. But these singular values equal the generalized singular values of the pair A, B. (Note that $\left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right] = (QD)(D^*R)$, generalized singular values are invariant under right multiplication by an invertible matrix, and the first $n + a$ and last $n + b$ rows of QD must have the same right singular vector matrix since $(QD)^*(QD) = I$.) $\square$

The Jacobi matrix model introduced in this dissertation extends beyond $\beta = 1, 2$ to general β, and removes the quantization on a and b. The model is a distribution

on orthogonal matrices with a special structure.

Definition 5.1.4. Given $\Theta = (\theta_n, \ldots, \theta_1)$ and $\Phi = (\phi_{n-1}, \ldots, \phi_1)$, we define four n-by-n bidiagonal matrices, $B_{11}(\Theta, \Phi)$, $B_{12}(\Theta, \Phi)$, $B_{21}(\Theta, \Phi)$, and $B_{22}(\Theta, \Phi)$.

$$\left[\begin{array}{c|c} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) \\ \hline B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) \end{array}\right] =$$

$$= \left[\begin{array}{cccc|cccc} c_n & -s_n c'_{n-1} & & & s_n s'_{n-1} & & & \\ & c_{n-1}s'_{n-1} & \ddots & & c_{n-1}c'_{n-1} & s_{n-1}s'_{n-2} & & \\ & & \ddots & -s_2 c'_1 & & \ddots & \ddots & \\ & & & c_1 s'_1 & & & c_1 c'_1 & s_1 \\ \hline -s_n & -c_n c'_{n-1} & & & c_n s'_{n-1} & & & \\ & -s_{n-1}s'_{n-1} & \ddots & & -s_{n-1}c'_{n-1} & c_{n-1}s'_{n-2} & & \\ & & \ddots & -c_2 c'_1 & & \ddots & \ddots & \\ & & & -s_1 s'_1 & & & -s_1 c'_1 & c_1 \end{array}\right],$$

in which $c_i = \cos\theta_i$, $s_i = \sin\theta_i$, $c'_i = \cos\phi_i$, $s'_i = \sin\phi_i$.

To prevent any possible confusion, we clarify that the $(n-1, n-1)$ entry of $B_{12}(\Theta, \Phi)$ is $s_2 s'_1$, and the $(n-1, n-1)$ entry of $B_{22}(\Theta, \Phi)$ is $c_2 s'_1$.

Lemma 5.1.5. *For any real Θ, Φ, the matrix*

$$\left[\begin{array}{cc} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) \\ B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) \end{array}\right]$$

is orthogonal.

The Jacobi matrix model (Figure 1.0.1) is defined by placing a distribution on Θ, Φ. Hence, the Jacobi matrix model is a random orthogonal matrix.

The Jacobi matrix model is first derived in the real and complex cases ($\beta = 1, 2$) by applying unitary transformations to a Haar-distributed matrix from the orthogonal or unitary group. These unitary transformations are structured to preserve CS

values. In fact, they are direct sums of Householder reflectors, chosen by an algorithm reminiscent of familiar algorithms from numerical analysis; see Section 5.2. This algorithmic approach is used in Section 5.3 to prove the following theorem in the special cases $\beta = 1, 2$.

Theorem. *Let β be any positive real number, let n be a positive integer, and let $a, b > -1$. Take the n-by-n CS decomposition of the $2n$-by-$2n$ Jacobi matrix model,*

$$J_{a,b}^{\beta} = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T.$$

The diagonal entries of C, squared, follow the law of the Jacobi ensemble with parameters β, a, b.

The theorem is stated again as Theorem 5.4.1 and proved in full generality in Section 5.4, and the following corollary is presented in Section 5.5 as Corollary 5.5.1.

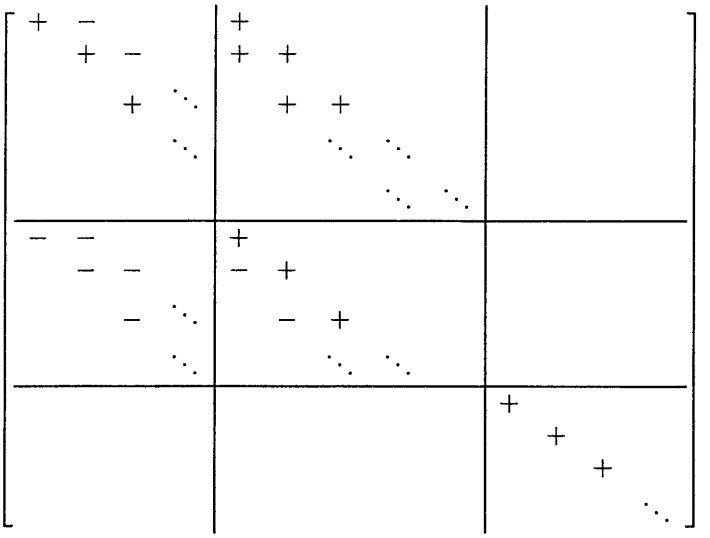
Corollary. *Under the hypotheses above, with $J_{a,b}^{\beta} = \left[\begin{smallmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{smallmatrix}\right]$, the generalized singular values, squared, of the pair B_{11}, B_{21} follow the law of the Jacobi ensemble with the same parameters.*

5.2 Bidiagonalization

5.2.1 Bidiagonal block form

A matrix is in bidiagonal block form if it satisfies a certain sign pattern. Throughout this chapter, $+$ in a sign pattern denotes a nonnegative entry, $-$ denotes a nonpositive entry, $\times$ denotes an unconstrained entry, and blanks denote zero entries.

Definition 5.2.1. Let A be a real m-by-m matrix, and let $p \geq q$ be nonnegative



The rows are partitioned as $p, q, (m-p-q)$ and the columns are partitioned as $q, p, (m-p-q)$.

Figure 5.2.1: Bidiagonal block form.

integers such that $p+q \leq m$. A is in *bidiagonal block form* with partition size p-by-q if A has the sign pattern in Figure 5.2.1.

Bidiagonal block form is most interesting in the context of unitary matrices. We shall see an analogy:

Finite computation	Infinite computation
tridiagonal form	eigenvalue decomposition
bidiagonal form	singular value decomposition
bidiagonal block form	CS decomposition

Lemma 5.2.2. *If Y is an m-by-m orthogonal matrix in bidiagonal block form with partition size p-by-q, then there exist unique $\Theta = (\theta_q, \ldots, \theta_1)$ and $\Phi = (\phi_{q-1}, \ldots, \phi_1)$,*

with entries between 0 *and* $\frac{\pi}{2}$, *such that*

$$Y = \left[\begin{array}{c|c|c} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) & \\ & I_{p-q} & \\ \hline B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) & \\ \hline & & I_{m-p-q} \end{array}\right]. \tag{5.2.1}$$

Proof. Uniqueness: θ_q is determined by the first column. Then ϕ_{q-1} is determined by the first row. Then θ_{q-1} is determined by the second column, ϕ_{q-2} by the second row, and so on.

Existence: Once θ_q is fixed, rows 1 and $p+1$ must have the forms given by the right hand side of (5.2.1) in order for these two rows to be orthogonal. Now ϕ_{q-1} is fixed, and columns 2 and $q+1$ must have the forms given by the right hand side of (5.2.1) in order for these two columns to be orthogonal. The proof continues by induction. □

Lemma 5.2.3. *Suppose that a unitary matrix* $Y = (y_{ij})$ *has the sign pattern in Figure 5.2.2. Then* Y *must be in bidiagonal block form.*

Proof. First we will show that the $(1,2)$ and $(2,1)$ blocks of Y have zeros in the necessary positions. If $y_{i,q+1}$ were nonzero, with $3 \leq i \leq p$, then the inner product of rows 1 and i would be nonzero. Now knowing that column $q+1$ has zeros in the locations just checked, the same argument applies to column $q+2$ in rows 4 through p, and then to column $q+3$, and so on. A similar argument shows that $y_{p+i,j}$ is zero for every $j \geq i+2$. Now we check for zeros in the $(2,2)$ block of Y. If $y_{p+i,q+j}$ were nonzero, with $1 \leq i \leq q$ and $1 \leq j \leq p$ and $j > i$, then the inner product of columns i and $q+j$ would be nonzero. If $y_{p+i,q+j}$ were nonzero, with $1 \leq i \leq q$ and $1 \leq j \leq p$ and $j < i-1$, then the inner product of rows j and $p+i$ would be nonzero. Now we check for zeros in the $(2,3)$ block of Y. If $y_{p+i,p+q+j}$ were nonzero, with $1 \leq i \leq q$ and $1 \leq j \leq m-p-q$, then the inner product of columns 1 and $p+q+j$ would be

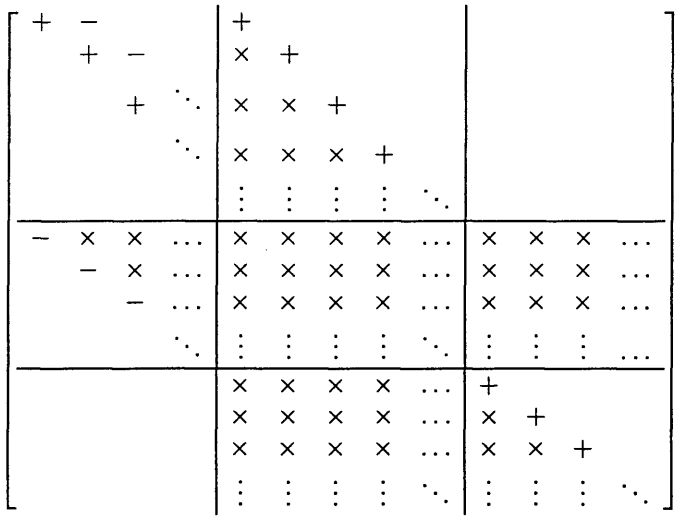
nonzero. Now, knowing that the $p+1$ row of Y has zeros in columns $p+q+1$ through m, we can use the same argument on the next row, taking inner products of column 2 with columns $p+q+1$ through m to show that row $p+2$ has zeros. Continuing, the $(2,3)$ block of Y has all zeros. A similar argument works for the $(3,2)$ block taking inner products of rows 1 through p with rows $p+q+1$ through $p+q+m$ in a judicious order. For the $(3,3)$ block of Y, first check column $p+q+1$. This column must have zeros below the main diagonal by taking the inner product of row $p+q+1$ with row $p+q+i$, $2 \leq i \leq m-p-q$. Then the same technique may be applied to column $p+q+2$, then $p+q+3$, and so on.

It remains to check the signs on the subdiagonal of block $(1,2)$, the superdiagonal of block $(2,1)$, and the main diagonal and subdiagonal of block $(2,2)$. Taking the inner product of columns j and $j+1$ shows that $y_{p+j,j+1}$ must be negative, for $1 \leq j \leq q-1$. Taking the inner product of rows i and $i+1$ shows $y_{i+1,q+i}$ must be positive, for $1 \leq i \leq p-1$. Finally, taking the inner products of columns j and $q+j$, $1 \leq j \leq p$ shows that the main diagonal of the $(2,2)$ block of Y must be positive, and taking the inner product of column $q+j$ with column $q+j+1$, $1 \leq j \leq p-1$ shows that the subdiagonal of the same block must be negative. □

5.2.2 The algorithm

We present an algorithm, hereafter known as "the algorithm," that transforms any matrix into a matrix having the sign pattern in Figure 5.2.2. The transformation is accomplished using unitary matrices with special structure. When the algorithm is run on a unitary matrix, the output is a unitary matrix in bidiagonal block form with the same CS values as the input matrix.

The complete algorithm is presented as Matlab code in Appendix A. Here is a rough sketch.



The rows are partitioned as $p, q, (m-p-q)$ and the columns are partitioned as $q, p, (m-p-q)$.

Figure 5.2.2: A related sign pattern.

The algorithm

Input: X (m-by-m) and $p \geq q \geq 0$ such that $p+q \leq m$
Output: Y

```
for k = 1 : q
  H := blkdiag(householder(...), householder(...));
  Y_{2k-1} := H * Y_{2k-2};
  H := blkdiag(householder(...), householder(...));
  Y_{2k} := Y_{2k-1} * H;
end
Y := postprocess(Y_{2q});
```

In the appendix, Y_i is denoted `Z(:,:,i)`.

The behavior of the algorithm on non-unitary X is suggested in Figure A.0.1, and the behavior on unitary X is suggested in Figure A.0.2. In the first step, Householder reflectors are used to place zeros in several entries in the first column, and then a second pair of Householder reflectors is used to place zeros in several entries in the first row, without modifying the first column. The algorithm continues, designing rows in the second column, then the second row, then the third column, then the

third row, and so on. The direct sums of Householder reflectors are designed to preserve CS values when the input is unitary.

5.2.3 Analysis of the algorithm

Theorem 5.2.4. *When X is unitary, the algorithm produces matrices Y, U, V such that*

1. $U^*XV = Y$.
2. *U is unitary and block diagonal, with blocks of sizes p-by-p and $(m-p)$-by-$(m-p)$.*
3. *V is unitary and block diagonal, with blocks of sizes q-by-q and $(m-q)$-by-$(m-q)$.*
4. *Y is an orthogonal matrix in bidiagonal block form with partition size p-by-q.*
5. *X and Y share the same p-by-q CS values.*

Proof. (1), (2), and (3) are straightforward and hold even if X is not unitary. The Householder reflectors are chosen so that Y satisfies the sign pattern in Figure 5.2.2. Hence, (4) holds by Lemma 5.2.3. (5) is immediate from (1), (2), and (3). □

Eventually, we want to run the algorithm on a random matrix. Analyzing the behavior will require a better understanding of the intermediate matrices Y_1, Y_2, ..., Y_{2q}. The algorithm proceeds by fixing one column, then two rows, then three columns, then four rows, and so on. In the remainder of this section, we develop this idea.

Let

$$P_{2k-1} = \begin{bmatrix} I_{q,k} & 0 \\ 0 & I_{p,k-1} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_{q,k} & 0 \\ 0 & I_{p,k-1} \\ 0 & 0 \end{bmatrix}^T, \tag{5.2.2}$$

$$P_{2k} = \begin{bmatrix} I_{p,k} & 0 \\ 0 & I_{q,k} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_{p,k} & 0 \\ 0 & I_{q,k} \\ 0 & 0 \end{bmatrix}^T . \tag{5.2.3}$$

For given $\hat{\Theta} = (\theta_q, \ldots, \theta_{q-k+1})$ and $\hat{\Phi} = (\phi_{q-1}, \ldots, \phi_{q-k+1})$, let

$$\hat{Y}_{2k-1}(\hat{\Theta}, \hat{\Phi}) = \left[\begin{array}{cc|c} B_{11} & B_{12} & \\ & I_{p-q} & \\ \hline B_{21} & B_{22} & \\ \hline & & I_{m-p-q} \end{array}\right] P_{2k-1}, \tag{5.2.4}$$

and for given $\hat{\Theta} = (\theta_q, \ldots, \theta_{q-k+1})$ and $\hat{\Phi} = (\phi_{q-1}, \ldots, \phi_{q-k})$, let

$$\hat{Y}_{2k}(\hat{\Theta}, \hat{\Phi}) = P_{2k} \left[\begin{array}{cc|c} B_{11} & B_{12} & \\ & I_{p-q} & \\ \hline B_{21} & B_{22} & \\ \hline & & I_{m-p-q} \end{array}\right] . \tag{5.2.5}$$

In each case, $B_{ij} = B_{ij}(\Theta, \Phi)$, where Θ is formed by extending $\hat{\Theta}$ arbitrarily to a q-vector and Φ is formed by extending $\hat{\Phi}$ arbitrarily to a $(q-1)$-vector. The particular extensions chosen do not affect the right hand side of (5.2.4) or (5.2.5).

Lemma 5.2.5. *Suppose that the algorithm produces intermediate matrices* $Y_1, Y_2, \ldots, Y_{2q}$ *and final matrix* Y *defined by* $\Theta = (\theta_q, \ldots, \theta_1)$ *and* $\Phi = (\phi_{q-1}, \ldots, \phi_1)$. *Then*

$$Y_{2k-1}P_{2k-1} = \hat{Y}_{2k-1}(\theta_q, \ldots, \theta_{q-k+1}; \phi_{q-1}, \ldots, \phi_{q-k+1}), \quad k = 1, \ldots, q,$$

and

$$P_{2k}Y_{2k} = \hat{Y}_{2k}(\theta_q, \ldots, \theta_{q-k+1}; \phi_{q-1}, \ldots, \phi_{q-k}), \quad k = 1, \ldots, q.$$

Lemma 5.2.6. *Suppose that the algorithm produces intermediate matrices $Y_1, Y_2, \ldots, Y_{2q}$ and final matrix Y defined by $\Theta = (\theta_q, \ldots, \theta_1)$ and $\Phi = (\phi_{q-1}, \ldots, \phi_1)$. Then*

$$\theta_{q-k} = \tan^{-1} \frac{\|(Y_{2k})_{((p+1+k):m),(1+k)}\|}{\|(Y_{2k})_{((1+k):p),(1:k)}\|}, \quad k = 0, \ldots, q-1$$
$$\phi_{q-1-k} = \tan^{-1} \frac{\|(Y_{2k+1})_{(1+k),((q+1+k):m)}\|}{\|(Y_{2k+1})_{(1+k),((2+k):q)}\|}, \quad k = 0, \ldots, q-2,$$

with all angles between 0 and $\frac{\pi}{2}$.

In the lemma, submatrices of Y_{2k} are specified in Matlab-like notation.

5.3 Real and complex random matrices

Let G be either $O(m)$ or $U(m)$, i.e., either the orthogonal group or the unitary group of m-by-m matrices. Let X be a random matrix from G whose distribution is Haar measure. Running the algorithm on X produces a sequence of intermediate matrices $Y_1, Y_2, \ldots, Y_{2q}$. Each Y_i is itself a random matrix, and we are interested in its distribution. In constructing Y_i, the algorithm observes certain rows and columns when choosing Householder reflectors, but it does not directly observe the remaining entries. Therefore, it is natural to consider the distribution of Y_i conditioned on $Y_i P_i$ if i is odd or $P_i Y_i$ if i is even. By Lemma 5.2.5, this is equivalent to conditioning on specific values for $\theta_n, \ldots, \theta_{n-k+1}, \phi_{n-1}, \ldots, \phi_{n-k+1}$ if $i = 2k-1$ or $\theta_n, \ldots, \theta_{n-k+1}, \phi_{n-1}, \ldots, \phi_{n-k}$ if $i = 2k$.

With this consideration in mind, define

$$\mu_{2k-1}(\hat{\Theta}, \hat{\Phi}) = \text{distribution of } Y_{2k-1}, \text{ conditioned on } Y_{2k-1}P_{2k-1} = \hat{Y}_{2k-1}(\hat{\Theta}, \hat{\Phi}),$$
$$\mu_{2k}(\hat{\Theta}, \hat{\Phi}) = \text{distribution of } Y_{2k}, \text{ conditioned on } P_{2k}Y_{2k} = \hat{Y}_{2k}(\hat{\Theta}, \hat{\Phi}).$$

We shall show that these distributions are defined by invariance properties. Let

$$\mathcal{U}_i(\hat{\Theta}, \hat{\Phi}) = \{U \in G : U^* \hat{Y}_i = \hat{Y}_i\},$$
$$\mathcal{V}_i(\hat{\Theta}, \hat{\Phi}) = \{V \in G : \hat{Y}_i V = \hat{Y}_i\},$$

in which $\hat{Y}_i$ is short for $\hat{Y}_i(\hat{\Theta}, \hat{\Phi})$. We call a matrix distribution $\mathcal{U}_i(\hat{\Theta}, \hat{\Phi})$-*invariant* if it is invariant w.r.t. left-multiplication by U^*, for all $U \in \mathcal{U}_i(\hat{\Theta}, \hat{\Phi})$. Similarly, we call a matrix distribution $\mathcal{V}_i(\hat{\Theta}, \hat{\Phi})$-*invariant* if it is invariant w.r.t. right-multiplication by all $V \in V_i(\hat{\Theta}, \hat{\Phi})$. Notice that $\mathcal{U}_0 \supset \mathcal{U}_1(\theta_0) \supset \mathcal{U}_2(\theta_0; \phi_0) \supset \mathcal{U}_3(\theta_0, \theta_1; \phi_0) \supset \ldots$ and $\mathcal{V}_0 \supset \mathcal{V}_1(\theta_0) \supset \mathcal{V}_2(\theta_0; \phi_0) \supset \mathcal{V}_3(\theta_0, \theta_1; \phi_0) \supset \ldots$.

Lemma 5.3.1. *For any appropriately sized $\hat{\Theta}$ and $\hat{\Phi}$ with entries between 0 and $\frac{\pi}{2}$, $\mu_i(\hat{\Theta}, \hat{\Phi})$ is $\mathcal{U}_i(\hat{\Theta}, \hat{\Phi})$-invariant and $\mathcal{V}_i(\hat{\Theta}, \hat{\Phi})$-invariant.*

Before proving the lemma, we must state and prove another lemma.

Lemma 5.3.2. *Suppose that A follows a random matrix distribution for which*

1. *$AP_i = \hat{Y}_i(\hat{\Theta}, \hat{\Phi})$ if i is odd, or*
2. *$P_i A = \hat{Y}_i(\hat{\Theta}, \hat{\Phi})$ if i is even,*

for some fixed $\hat{\Theta}$, $\hat{\Phi}$. Then the distribution of A is $\mathcal{U}_i$-invariant if and only if it is $\mathcal{V}_i$-invariant.

Proof. We prove only the case when i is odd. The case when i is even is very similar.

A can be broken into two terms, $A = AP_i + A(I - P_i)$. Let $\hat{U}$ and $\hat{V}$ be m-by-$(m-i)$ matrices containing orthonormal bases for the orthogonal complements of the column space and row space, respectively, of $\hat{Y}_i(\hat{\Theta}, \hat{\Phi})$. The following statements are equivalent.

1. $A(I - P_i) \stackrel{d}{=} \hat{U}\,[\text{Haar}]\,\hat{V}^*$.
2. A is $\mathcal{U}_i$-invariant.

3. A is $\mathcal{V}_i$-invariant.

The proofs of (1)$\Leftrightarrow$(2) and (1)$\Leftrightarrow$(3) are straightforward. □

Proof of Lemma 5.3.1. The proof uses induction. We will abuse notation and suppress $\hat{\Theta}$ and $\hat{\Phi}$. For example, instead of writing $\mathcal{U}_{2k}(\hat{\Theta}_{1:k}, \hat{\Phi}_{1:k})$, we will simply write $\mathcal{U}_{2k}$.

Base case: μ_0 is Haar measure by definition, and $\mathcal{U}_0 = \mathcal{V}_0 = G$.

Induction step: Assume that μ_{i-1} is $\mathcal{U}_{i-1}$-invariant and $\mathcal{V}_{i-1}$-invariant. By Lemma 5.3.2, it suffices to prove that μ_i is *either* $\mathcal{U}_i$-invariant or $\mathcal{V}_i$-invariant. When $i = 2k$ is even, we show that μ_{2k} is $\mathcal{U}_{2k}$-invariant; the case when i is odd is left to the reader. Because $\mathcal{U}_{2k-1} \supset \mathcal{U}_{2k}$, μ_{2k-1} is $\mathcal{U}_{2k}$-invariant. The action of any $U \in \mathcal{U}_{2k}$ also has the special property that it preserves rows k and $p+k$, so that it preserves ϕ_{n-k}. Therefore, $\mu_{2k-1}|\phi_{n-k}$, the distribution obtained from μ_{2k-1} by conditioning on ϕ_{n-k}, is also $\mathcal{U}_{2k}$-invariant. Now, μ_{2k} is simply the pushforward distribution of $\mu_{2k-1}|\phi_{n-k}$ along right-multiplication by a suitably chosen Householder reflector H. The $\mathcal{U}_{2k}$-invariance of μ_{2k} follows from associativity: $UY_{2k} = U(Y_{2k-1}H) = (UY_{2k-1})H \stackrel{d}{=} Y_{2k-1}H = Y_{2k}$. (It is worth emphasizing that the $\stackrel{d}{=}$ is conditioned on ϕ_{n-k} and is justified because we checked that μ_{2k-1} is $\mathcal{U}_{2k}$-invariant, even after conditioning on ϕ_{n-k}.) □

Theorem 5.3.3. *Let n, a, and b be positive integers, and define $m = 2n + a + b$. Suppose that X follows Haar measure on G (either $O(m)$ or $U(m)$), and run the algorithm on X using partition size $(n + a)$-by-n. Then the output Y is a random orthogonal matrix in bidiagonal block form,*

$$Y = \left[\begin{array}{c|c|c} B_{11} & B_{12} & \\ & I & \\ \hline B_{21} & B_{22} & \\ \hline & & I \end{array}\right],$$

distributed in such a way that

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \stackrel{d}{=} J^{\beta}_{a,b}.$$

In other words, the distribution of $\left[\begin{smallmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{smallmatrix}\right]$ *is the* $2n$*-by-*$2n$ β*-Jacobi matrix model with parameters* a *and* b*, where* $\beta = 1$ *if* $G = O(m)$*, or* $\beta = 2$ *if* $G = U(m)$*.*

Proof. First assume that $G = O(m)$. Let $\sqrt{P_{2k-1}}$ denote the left factor in (5.2.2), and let $\sqrt{P_{2k}}$ denote the left factor in (5.2.3).

$\mathcal{V}_{2k-1}(\hat{\Theta}, \hat{\Phi})$ does not actually depend on $\hat{\Theta}$ or $\hat{\Phi}$; it contains precisely the orthogonal matrices V such that $P_{2k-1}V = P_{2k-1}$. By Lemma 5.3.1, if $Y_{2k-1} \sim \mu_{2k-1}(\hat{\Theta}, \hat{\Phi})$, then $Y_{2k-1} \stackrel{d}{=} Y_{2k-1}V$. This implies that the *direction* of each row of $Y_{2k-1}\sqrt{P_{2k-1}}$ is uniformly distributed on the real sphere S^{2k-2}. By Lemma 5.2.6, ϕ_{n-k} depends only on the direction of row k of this matrix, and the distribution is given by $\cos\phi_{n-k} \sim \sqrt{\text{beta}\left(\frac{1}{2}(n-k), \frac{1}{2}(a+b+1+n-k)\right)}$. The same distribution results regardless of which $\hat{\Theta} = (\theta_n, \ldots, \theta_{n-k+1}), \hat{\Phi} = (\phi_{n-1}, \ldots, \phi_{n+1-k})$ are assumed by conditioning. Hence, ϕ_{n-k} is actually independent of $\theta_n, \ldots, \theta_{n-k+1}, \phi_{n-1}, \ldots, \phi_{n-k+1}$ and has the indicated distribution.

The preceding discussion applied to $i = 2k-1$ odd. For $i = 2k$ even, note that $\mathcal{U}_{2k}(\hat{\Theta}, \hat{\Phi})$ does not depend on $\hat{\Theta}$ or $\hat{\Phi}$ and contains precisely the orthogonal matrices U such that $U^T P_{2k} = P_{2k}$. For $Y_{2k} \sim \mu_{2k}(\hat{\Theta}, \hat{\Phi})$, the directions of the columns of $\left(\sqrt{P_{2k}}\right)^T Y_{2k}$ are uniformly distributed on the real sphere S^{2k-1}, regardless of which $\hat{\Theta}, \hat{\Phi}$ are used. Therefore, θ_{n-k} is independent from $\theta_n, \ldots, \theta_{n-k+1}, \phi_{n-1}, \ldots, \phi_{n-k}$, and its distribution is defined by $\cos\theta_{n-k} \sim \sqrt{\text{beta}\left(\frac{1}{2}(a+n-k), \frac{1}{2}(b+n-k)\right)}$.

When $G = U(m)$, the proof is exactly the same, except that (1) $\mathcal{V}_{2k-1}(\hat{\Theta}, \hat{\Phi})$ and $\mathcal{U}_{2k}(\hat{\Theta}, \hat{\Phi})$ contain unitary matrices in addition to orthogonal matrices, (2) conjugate transpose replaces transpose, (3) complex spheres replace real spheres, and (4) $\beta = 2$ replaces $\beta = 1$ in $\cos\phi_{n-k} \sim \sqrt{\text{beta}\left(\frac{\beta}{2}(n-k), \frac{\beta}{2}(a+b+1+n-k)\right)}$, $\cos\theta_{n-k} \sim \sqrt{\text{beta}\left(\frac{\beta}{2}(a+n-k), \frac{\beta}{2}(b+n-k)\right)}$. □

5.4 General β matrix models: Beyond real and complex

Theorems 5.1.3 and 5.3.3 imply that the n-by-n CS values of $J^{\beta}_{a,b}$ follow the β-Jacobi law when $\beta = 1$ or $\beta = 2$ and a, b are nonnegative integers. It remains to show that the CS values model the β-Jacobi ensemble for other values of β, a, and b.

Theorem 5.4.1. *Let β be any positive real number, let n be a positive integer, and let $a, b > -1$. Take the n-by-n CS decomposition of the $2n$-by-$2n$ Jacobi matrix model,*

$$J^{\beta}_{a,b} = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T.$$

The diagonal entries of C, squared, follow the law of the β-Jacobi ensemble with parameters a, b. Also, the first row of V_1, up to sign, is distributed as a vector of i.i.d. χ_β random variables, normalized to unit length.

The proof is at the end of this section.

The β-Jacobi matrix model is a distribution on $2n$-by-$2n$ orthogonal matrices, but the CS values are completely determined by the upper-left n-by-n block. In fact, the CS values are precisely the singular values of this matrix. Their distribution will be obtained by changing variables.

Given $\Theta = (\theta_n, \ldots, \theta_1)$ and $\Phi = (\phi_{n-1}, \ldots, \phi_1)$, let $c_i = \cos\theta_i$, $s_i = \sin\theta_i$, $c'_i = \cos\phi_i$, and $s'_i = \sin\phi_i$. Also, let $\sigma_1 > \cdots > \sigma_n$ be the singular values of $B_{11}(\Theta, \Phi)$, and, for $i = 1, \ldots, n-1$, let v_i be the first entry of the right singular vector of $B_{11}(\Theta, \Phi)$ corresponding to σ_i, constrained to be nonnegative.

Lemma 5.4.2. *The $2n-1$ parameters $\sigma_1, \ldots, \sigma_n, v_1, \ldots, v_{n-1}$ defined above uniquely determine a matrix of the form $\begin{bmatrix} B_{11}(\Theta,\Phi) & B_{12}(\Theta,\Phi) \\ B_{21}(\Theta,\Phi) & B_{22}(\Theta,\Phi) \end{bmatrix}$. The Jacobian for the change of variables between $(c_n, \ldots, c_1, c'_{n-1}, \ldots, c'_1)$ and $(\sigma_1, \ldots, \sigma_n, v_1, \ldots, v_{n-1})$ can be ex-*

pressed as

$$\prod_{i=1}^{n}\left(c_i^{\beta(i-1)+1}s_i^{\beta(i-1)}dc_i\right)\prod_{i=1}^{n-1}\left((c_i')^{\beta i-1}(s_i')^{\beta(i-1)+2}dc_i'\right)$$
$$=\left(\prod_{i<j}(\sigma_i^2-\sigma_j^2)^{\beta}\prod_{i=1}^{n}\sigma_i d\sigma_i\right)\left(\prod_{i=1}^{n-1}v_i^{\beta-1}dv_i\right)$$

for any $\beta > 0$.

Proof. Denote the entries of $B_{11}(\Theta,\Phi)$ by

$$B_{11}(\Theta,\Phi)=\begin{bmatrix} x_n & -y_{n-1} & & & \\ & x_{n-1} & -y_{n-2} & & \\ & & x_{n-2} & \ddots & \\ & & & \ddots & -y_1 \\ & & & & x_1 \end{bmatrix}.$$

It follows from Lemmas 2.7, 2.9, and 2.11 of [6] that

$$\left(\prod_{i<j}(\sigma_i^2-\sigma_j^2)^{\beta}\prod_{i=1}^{n}\sigma_i\prod_{i=1}^{n}d\sigma_i\right)\left(\prod_{i=1}^{n}v_i^{\beta-1}\prod_{i=1}^{n-1}dv_i\right)$$
$$=\left(\prod_{i=1}^{n}x_i^{\beta(i-1)+1}\prod_{i=1}^{n}dx_i\right)\left(\prod_{i=1}^{n-1}y_i^{\beta i-1}\prod_{i=1}^{n-1}dy_i\right).$$

The Jacobian matrix $\frac{\partial(x_n,\ldots,x_1,y_{n-1},\ldots,y_1)}{\partial(c_n,\ldots,c_1,c_n',\ldots,c_1')}$ is

$$\left[\begin{array}{c|c|c} \multicolumn{2}{c|}{\multirow{2}{*}{$\operatorname{diag}(1,s_{n-1}',\ldots,s_1')$}} & 0 \quad \cdots \quad 0 \\ \cline{3-3} \multicolumn{2}{c|}{} & \operatorname{diag}(-c_{n-1}\frac{c_{n-1}'}{s_{n-1}'},\ldots,-c_1\frac{c_1'}{s_1'}) \\ \hline \operatorname{diag}(\frac{-c_n}{s_n}c_{n-1}',\ldots,\frac{-c_2}{s_2}c_1') & \begin{matrix}0\\ \vdots \\ 0\end{matrix} & \operatorname{diag}(s_n,\ldots,s_{n-1}) \end{array}\right].$$

Using the Schur complement formula for the determinant [15], we find that the Jaco-

bian is the product of the determinant of the top-left block with the determinant of a lower triangular matrix whose diagonal is given by the bottom-right block. Hence,

$$\prod_{i=1}^{n} dx_i \prod_{i=1}^{n-1} dy_i = \prod_{i=2}^{n} s_i dc_i \prod_{i=1}^{n-1} s_i' dc_i'.$$

Changing variables again using this Jacobian and evaluating x_i, y_i gives

$$\begin{aligned}&\left(\prod_{i=1}^{n} x_i^{\beta(i-1)+1} \prod_{i=1}^{n} dx_i\right)\left(\prod_{i=1}^{n-1} y_i^{\beta i-1} \prod_{i=1}^{n-1} dy_i\right)\\ &\quad = \prod_{i=1}^{n} c_i^{\beta(i-1)+1} \prod_{i=1}^{n} s_i^{\beta(i-1)} \prod_{i=1}^{n-1} (c_i')^{\beta i-1} \prod_{i=1}^{n-1} (s_i')^{\beta(i-1)+2} \prod_{i=1}^{n} dc_i \prod_{i=1}^{n-1} dc_i'.\end{aligned}$$

□

Proof of Theorem 5.4.1. The differential for the β-Jacobi matrix model is

$$dJ = \text{const} \times \prod_{i=1}^{n} \left(c_i^{\beta(a+i)-1} s_i^{\beta(b+i)-2} dc_i\right) \prod_{i=1}^{n-1} \left((c_i')^{\beta i-1} (s_i')^{\beta(a+b+1+i)-2} dc_i'\right).$$

Changing variables using the lemma gives

$$\begin{aligned}dJ &= \text{const} \times \prod_{i=1}^{n} c_i^{\beta(a+1)-2} \prod_{i=1}^{n} s_i^{\beta(b+1)-2} \prod_{i=1}^{n-1} (s_i')^{\beta(a+b+2)-4} \times\\ &\quad \times \left(\prod_{i<j} (\sigma_i^2 - \sigma_j^2)^{\beta} \prod_{i=1}^{n} \sigma_i d\sigma_i\right)\left(\prod_{i=1}^{n-1} v_i^{\beta-1} dv_i\right)\\ &= \text{const} \times \prod_{i=1}^{n} x_i^{\beta(a+1)-2} \prod_{i=1}^{n} w_i^{\beta(b+1)-2} \left(\prod_{i<j} (\sigma_i^2 - \sigma_j^2)^{\beta} \prod_{i=1}^{n} \sigma_i d\sigma_i\right)\left(\prod_{i=1}^{n-1} v_i^{\beta-1} dv_i\right),\end{aligned}$$

in which $x_n, \ldots, x_1$ are the diagonal entries of B_{11} and $w_n, \ldots, w_1$ are the diagonal entries of B_{21}. Now notice that $\prod_{i=1}^{n} x_i$ is the determinant of B_{11}, that $\prod_{i=1}^{n} w_i$ is the

determinant of B_{21}, and that $B_{21}^T B_{21} = I - B_{11}^T B_{11}$, so that

$$\begin{aligned} dJ &= \text{const} \times \det(B_{11}^T B_{11})^{\frac{\beta}{2}(a+1)-1} \det(I - B_{11}^T B_{11})^{\frac{\beta}{2}(b+1)-1} \times \\ &\quad \times \left(\prod_{i<j} (\sigma_i^2 - \sigma_j^2)^\beta \prod_{i=1}^{n} \sigma_i d\sigma_i \right) \left(\prod_{i=1}^{n-1} v_i^{\beta-1} dv_i \right) \\ &= \text{const} \times \left(\prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(a+1)-1} (1-\lambda_i)^{\frac{\beta}{2}(b+1)-1} \prod_{i<j} (\lambda_i - \lambda_j)^\beta \prod d\lambda_i \right) \left(\prod_{i=1}^{n-1} v_i^{\beta-1} \prod_{i=1}^{n-1} dv_i \right), \end{aligned}$$

in which $\lambda_i = \sigma_i^2$. □

5.5 Multivariate analysis of variance

Pairs of Gaussian matrices are often important in multivariate analysis of variance (MANOVA). The following corollary may be useful in this context.

Corollary 5.5.1. *Let n, a, and b be nonnegative integers. Suppose that N_1 ($(n+a)$-by-n) and N_2 ($(n+b)$-by-n) are independent random matrices, each with i.i.d. real (resp., complex) standard Gaussian entries. If*

$$J_{a,b}^\beta = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

with $\beta = 1$ (resp., $\beta = 2$), then

$$\text{gsvd}(N_1, N_2) \stackrel{d}{=} \text{gsvd}(B_{11}, B_{21}).$$

Here, $\text{gsvd}(C, D)$ evaluates to the generalized singular values of C, D in decreasing order.

Proof. Because the columns of $J_{a,b}^\beta$ are orthonormal, the generalized singular values of B_{11}, B_{21} are just the CS values of $J_{a,b}^\beta$ (see [26]), whose squares follow the Jacobi

law. The proof follows by Proposition 5.1.2. □

Chapter 6

Zero temperature matrix models

Each of the matrix models has a natural extension to $\beta = \infty$, displayed in Figures 6.0.1–6.0.3. These "zero temperature" random matrix models are not random at all. In fact, they encode orthogonal polynomial recurrence relations.

6.1 Overview

This chapter provides explicit formulas and large n asymptotics for the CS decomposition (CSD) of $J_{a,b}^{\infty}$, the singular value decomposition (SVD) of L_a^{∞}, and the eigenvalue decomposition of H^{∞}.

It will be shown that the CSD of $J_{a,b}^{\infty}$ can be constructed from Jacobi polynomials and their zeros, the SVD of L_a^{∞} can be constructed from Laguerre polynomials and their zeros, and the eigenvalue decomposition of H^{∞} can be constructed from Hermite polynomials and their zeros.

The large n asymptotics for eigenvalues/singular values/CS values proved later in this chapter are suggested by the following table.

JACOBI MATRIX MODEL, $\beta = \infty$

$$J^{\infty}_{a,b} = \begin{bmatrix} B_{11}(\bar{\Theta},\bar{\Phi}) & B_{12}(\bar{\Theta},\bar{\Phi}) \\ B_{21}(\bar{\Theta},\bar{\Phi}) & B_{22}(\bar{\Theta},\bar{\Phi}) \end{bmatrix}$$

$$= \left[\begin{array}{ccccc|ccccc} \bar{c}_n & -\bar{s}_n\bar{c}'_{n-1} & & & & \bar{s}_n\bar{s}'_{n-1} & & & & \\ & \bar{c}_{n-1}\bar{s}'_{n-1} & -\bar{s}_{n-1}\bar{c}'_{n-2} & & & \bar{c}_{n-1}\bar{c}'_{n-1} & \bar{s}_{n-1}\bar{s}'_{n-2} & & & \\ & & \bar{c}_{n-2}\bar{s}'_{n-2} & \ddots & & & \bar{c}_{n-2}\bar{c}'_{n-2} & \bar{s}_{n-2}\bar{s}'_{n-3} & & \\ & & & \ddots & -\bar{s}_2\bar{c}'_1 & & & \ddots & \ddots & \\ & & & & \bar{c}_1\bar{s}'_1 & & & & \bar{c}_1\bar{c}'_1 & \bar{s}_1 \\ \hline -\bar{s}_n & -\bar{c}_n\bar{c}'_{n-1} & & & & \bar{c}_n\bar{s}'_{n-1} & & & & \\ & -\bar{s}_{n-1}\bar{s}'_{n-1} & -\bar{c}_{n-1}\bar{c}'_{n-2} & & & -\bar{s}_{n-1}\bar{c}'_{n-1} & \bar{c}_{n-1}\bar{s}'_{n-2} & & & \\ & & -\bar{s}_{n-2}\bar{s}'_{n-2} & \ddots & & & -\bar{s}_{n-2}\bar{c}'_{n-2} & \bar{c}_{n-2}\bar{s}'_{n-3} & & \\ & & & \ddots & -\bar{c}_2\bar{c}'_1 & & & \ddots & \ddots & \\ & & & & -\bar{s}_1\bar{s}'_1 & & & & -\bar{s}_1\bar{c}'_1 & \bar{c}_1 \end{array}\right]$$

$$\bar{\Theta} = (\bar{\theta}_n, \ldots, \bar{\theta}_1) \in [0, \tfrac{\pi}{2}]^n \qquad \bar{\Phi} = (\bar{\phi}_{n-1}, \ldots, \bar{\phi}_1) \in [0, \tfrac{\pi}{2}]^{n-1}$$

$$\bar{c}_i = \cos\bar{\theta}_i \qquad \bar{c}'_i = \cos\bar{\phi}_i$$

$$\bar{s}_i = \sin\bar{\theta}_i \qquad \bar{s}'_i = \sin\bar{\phi}_i$$

$$\bar{c}_i = \sqrt{\frac{a+i}{a+b+2i}} \qquad \bar{c}'_i = \sqrt{\frac{i}{a+b+1+2i}}$$

Figure 6.0.1: The $\beta = \infty$ (zero temperature) Jacobi matrix model.

LAGUERRE MATRIX MODEL, $\beta = \infty$ (square)

$$L_a^\infty = \begin{bmatrix} \sqrt{a+n} & & & & \\ -\sqrt{n-1} & \sqrt{a+n-1} & & & \\ & -\sqrt{n-2} & \sqrt{a+n-2} & & \\ & & \ddots & \ddots & \\ & & & -\sqrt{1} & \sqrt{a+1} \end{bmatrix}$$

LAGUERRE MATRIX MODEL, $\beta = \infty$ (rectangular)

$$M_a^\infty = \begin{bmatrix} -\sqrt{n} & \sqrt{a+n-1} & & & \\ & -\sqrt{n-1} & \sqrt{a+n-2} & & \\ & & -\sqrt{n-2} & \sqrt{a+n-3} & \\ & & & \ddots & \ddots \\ & & & & -\sqrt{1} \quad \sqrt{a} \end{bmatrix}$$

Figure 6.0.2: The $\beta = \infty$ (zero temperature) Laguerre matrix models.

HERMITE MATRIX MODEL, $\beta = \infty$

$$H^\infty = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & \sqrt{n-1} & & & \\ \sqrt{n-1} & 0 & \sqrt{n-2} & & \\ & \ddots & \ddots & \ddots & \\ & & \sqrt{2} & 0 & \sqrt{1} \\ & & & \sqrt{1} & 0 \end{bmatrix}$$

Figure 6.0.3: The $\beta = \infty$ (zero temperature) Hermite matrix model.

Matrix model ($\beta = \infty$)	Scaling limit	Scaling limit type	Eigenvalue/singular value/ CS value asymptotics
Jacobi	left edge	hard edge	Bessel zeros
Jacobi	right edge	hard edge	Bessel zeros
Jacobi	center	bulk	linearly spaced
Laguerre (square)	left edge	hard edge	Bessel zeros
Laguerre (square)	right edge	soft edge	Airy zeros
Laguerre (rectangular)	left edge	hard edge	Bessel zeros
Laguerre (rectangular)	right edge	soft edge	Airy zeros
Hermite	center	bulk	linearly spaced
Hermite	right edge	soft edge	Airy zeros

The large n asymptotics for eigenvectors/singular vectors/CS vectors proved later in this chapter are suggested by the following table.

Matrix model ($\beta = \infty$)	Scaling limit	Scaling limit type	Eigenvector/singular vector/CS vector asymptotics
Jacobi	left edge	hard edge	Bessel functions
Jacobi	right edge	hard edge	Bessel functions
Jacobi	center	bulk	sine waves
Laguerre (square)	left edge	hard edge	Bessel functions
Laguerre (square)	right edge	soft edge	Airy functions
Laguerre (rectangular)	left edge	hard edge	Bessel functions
Laguerre (rectangular)	right edge	soft edge	Airy functions
Hermite	center	bulk	sine waves
Hermite	right edge	soft edge	Airy functions

6.2 Eigenvalue, singular value, and CS decompositions

In this subsection, some of the results are new, and some are old. The proposition regarding the Hermite matrix model is common knowledge. It states that the zero temperature Hermite matrix model encodes the Hermite polynomial recurrence, so that the eigenvalues are Hermite polynomial roots, and the eigenvectors are constructed from Hermite eigenvectors. The propositions for the Jacobi and Laguerre

models are similar, but use CS decomposition and singular value decomposition instead of eigenvalue decomposition. The Jacobi case is likely an original contribution of this dissertation, because of the relative obscurity of the CS decomposition, and the Laguerre case is also original to this dissertation as far as we know.

Many works on orthogonal polynomials mention the ubiquity of three-term recurrences, which lead to tridiagonal matrices. It is obvious that these tridiagonal matrices can be generated from the Lanczos iteration, a topic discussed in [39]. However, in the Jacobi and Laguerre cases, two-term recurrences are actually available, if one considers $\pi_n^L(a+1;x)$ in addition to $\pi_n^L(a;x)$, for example. These two-term recurrences are encoded by bidiagonal matrices, as seen below, which can be generated from the conjugate gradient iteration. This idea follows from viewing the conjugate gradient iteration as a "one-sided" Lanczos iteration [28].

Proposition 6.2.1. *Take the n-by-n CS decomposition of the 2n-by-2n Jacobi matrix model* $J_{a,b}^{\infty}$,

$$J_{a,b}^{\infty} = \begin{bmatrix} U_1 & \\ & U_2 \end{bmatrix} \begin{bmatrix} C & S \\ -S & C \end{bmatrix} \begin{bmatrix} V_1 & \\ & V_2 \end{bmatrix}^T.$$

Then

1. *The diagonal entries of* C, *squared, are the zeros* $z_1 < z_2 < \cdots < z_n$ *of* $\pi_n^J(a,b;\cdot)$.

2. *The* $(1,j)$ *entry of* V_1 *is*

$$\frac{\sqrt{\frac{a+n}{a+b+2n}} z_j^{-1/2} \psi_{n-1}^J(a,b+1;z_j)}{\sqrt{K_{n-2}^J(a+1,b+1;z_j,z_j) + \frac{a+n}{a+b+2n} z_j^{-1} \psi_{n-1}^J(a,b+1,z_j)^2}}, \quad j = 1,\ldots,n,$$

and for $i \geq 2$, *the* (i,j) *entry is*

$$\frac{\psi_{n-i}^J(a+1,b+1;z_j)}{\sqrt{K_{n-2}^J(a+1,b+1;z_j,z_j) + \frac{a+n}{a+b+2n} z_j^{-1} \psi_{n-1}^J(a,b+1,z_j)^2}}, \quad j = 1,\ldots,n.$$

3. *For all i, j, the (i,j) entry of V_2 is $\psi^J_{n-i}(a,b;z_j)/\sqrt{K^J_{n-1}(a,b;z_j,z_j)}$.*

4. *For all i, j, the (i,j) entry of U_1 is $\psi^J_{n-i}(a,b+1;z_j)/\sqrt{K^J_{n-1}(a,b+1;z_j,z_j)}$.*

5. *For all i, j, the (i,j) entry of U_2 is $\psi^J_{n-i}(a+1,b;z_j)/\sqrt{K^J_{n-1}(a+1,b;z_j,z_j)}$.*

Proof. For any $x_1, \ldots, x_n$ between 0 and 1, we have

$$\begin{aligned}
\operatorname{diag}(\sqrt{x_1}, \ldots, \sqrt{x_n}) \left[\psi^J_{j-1}(a,b,x_i) \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} &= \left[\psi^J_{j-1}(a+1,b,x_i) \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} B, \\
\operatorname{diag}(\sqrt{x_1}, \ldots, \sqrt{x_n}) \left[\psi^J_{j-1}(a+1,b,x_i) \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} &= \left[\psi^J_{j-1}(a,b,x_i) \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n+1}} C,
\end{aligned}$$

in which the entries of the upper bidiagonal n-by-n matrix B and the lower bidiagonal $(n+1)$-by-n matrix C are determined by the recurrence relations (3.3.4–3.3.5). In fact, the leading n-by-n submatrix of C equals the transpose of B. Now, when $x_1, \ldots, x_n$ are the zeros $z_1, \ldots, z_n$ of $\pi^J_n(a,b;\cdot)$, the last column of

$$\left[\psi^J_{j-1}(a,b,x_i) \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n+1}}$$

contains zeros, and so the equations still hold after dropping the last column of this matrix and the last row of C, reading

$$\begin{aligned}
\operatorname{diag}(\sqrt{z_1}, \ldots, \sqrt{z_n}) \left[\psi^J_{j-1}(a,b,x_i) \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} &= \left[\psi^J_{j-1}(a+1,b,x_i) \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} B, \\
\operatorname{diag}(\sqrt{z_1}, \ldots, \sqrt{z_n}) \left[\psi^J_{j-1}(a+1,b,x_i) \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} &= \left[\psi^J_{j-1}(a,b,x_i) \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} B^T.
\end{aligned}$$

Because B is precisely the bottom-right block of $J^\infty_{a,b}$ after permuting rows and columns, we have explicitly constructed the singular value decomposition of this block, up to normalization of the singular vectors. It is quite clear that the correct normalization constant is given by the kernel, as in the statement of the theorem. This shows that the CS values are the zeros of the polynomial and that the entries of U_2 and V_2 are as advertised.

It remains to check the singular vectors of the top-left block of $J^{\infty}_{a,b}$ are given by U_1 and V_1. The proof is similar, but this time we already know that the singular values are the zeros of $\pi^J_n(a,b;\cdot)$. □

The proof shows that the top-left and bottom-right blocks of $J^{\infty}_{a,b}$ encode recurrence relations (3.3.4–3.3.5). It is worth noting that the top-right and bottom-left blocks encode recurrence relations (3.3.6–3.3.7).

Proposition 6.2.2. *Take the SVD of the n-by-n Laguerre matrix model L^{∞}_a to get $L^{\infty}_a = U\Sigma V^T$. Then*

1. *The singular values, squared, are the zeros $z_1 < z_2 < \cdots < z_n$ of $\pi^L_n(a;\cdot)$.*
2. *For all i,j, the (i,j) entry of V is $\psi^L_{n-i}(a;z_j)/\sqrt{K^L_{n-1}(a;z_j,z_j)}$.*
3. *For all i,j, the (i,j) entry of U is $\psi^L_{n-i}(a+1;z_j)/\sqrt{K^L_{n-1}(a+1;z_j,z_j)}$.*

Proof. The proof is similar to the proof of the previous proposition. For any positive $x_1,\ldots,x_n$, we have

$$\operatorname{diag}(\sqrt{x_1},\ldots,\sqrt{x_n})\left[\ \psi^L_{j-1}(a,x_i)\ \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} = \left[\ \psi^L_{j-1}(a+1,x_i)\ \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} B,$$
$$\operatorname{diag}(\sqrt{x_1},\ldots,\sqrt{x_n})\left[\ \psi^L_{j-1}(a+1,x_i)\ \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n}} = \left[\ \psi^L_{j-1}(a,x_i)\ \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n+1}} C,$$

in which the entries of the n-by-n upper bidiagonal matrix B and the $(n+1)$-by-n lower bidiagonal matrix C are determined by the recurrence relations (3.3.8–3.3.9). When $x_1,\ldots,x_n$ are the zeros $z_1,\ldots,z_n$ of $\pi^L_n(a;\cdot)$, the last column of

$$\left[\ \psi^L_{j-1}(a,x_i)\ \right]_{\substack{i=1,\ldots,n,\\ j=1,\ldots,n+1}}$$

is zero, so this column and the last row of C can be dropped. Observing that $B = C^T_{1:n,1:n} = FL^{\infty}_aF$, we have explicitly constructed the SVD of the $\beta = \infty$ Laguerre

model. It remains to check the normalization constants involving the Laguerre kernel, which is straightforward. □

Proposition 6.2.3. *Take the SVD of the n-by-$(n+1)$ rectangular Laguerre matrix model M_a^∞ to get $M_a^\infty = U\Sigma V^T$. Then*

1. *The singular values, squared, are the zeros $z_1 < z_2 < \cdots < z_n$ of $\pi_n^L(a;\cdot)$.*
2. *For $i = 1, \ldots, n+1$ and $j = 1, \ldots, n$, the (i,j) entry of V is $\psi_{n+1-i}^L(a-1;z_j)/\sqrt{K_n^L(a-1;z_j,z_j)}$. For $i = 1, \ldots, n+1$, the $(i,n+1)$ entry is $\sqrt{\frac{a\Gamma(n+1)\Gamma(a+n+1-i)}{\Gamma(a+n+1)\Gamma(n+2-i)}}$.*
3. *For all i,j, the (i,j) entry of U is $\psi_{n-i}^L(a;z_j)/\sqrt{K_{n-1}^L(a;z_j,z_j)}$.*

The proof is similar to the previous proof and is omitted.

Proposition 6.2.4. *Take the eigenvalue decomposition of the n-by-n Hermite matrix model H^∞ to get $H^\infty = Q\Lambda Q^T$. Then*

1. *The diagonal entries of Λ are the zeros $z_1 < z_2 < \cdots < z_n$ of π_n^H.*
2. *The (i,j) entry of Q is $\psi_{n-i}^H(z_j)/\sqrt{K_{n-1}^H(z_j,z_j)}$.*

Proof. This time, we have an eigenvalue problem. For any real $x_1, \ldots, x_n$,

$$\operatorname{diag}(x_1, \ldots, x_n) \left[\psi_{j-1}^H(x_i) \right]_{i=1,\ldots,n, j=1,\ldots,n} = \left[\psi_{j-1}^H(x_i) \right]_{i=1,\ldots,n, j=1,\ldots,n+1} T,$$

in which the entries of the $(n+1)$-by-n tridiagonal matrix T are determined by the Hermite recurrence (3.3.10). When the x_i are the zeros $z_1, \ldots, z_n$ of π_n^H, the last column of

$$\left[\psi_{j-1}^H(x_i) \right]_{i=1,\ldots,n, j=1,\ldots,n+1}$$

is zero, and so the equation holds after deleting this column and the last row of T. What remains of T is precisely $FH^\infty F$, and the equation gives the eigenvalue decomposition. It remains to check the normalization constants involving the Hermite kernel, which is straightforward. □

Remark 6.2.5. The signs of the entries in the matrix models were chosen to make the statements of these four propositions concise.

6.3 Large n asymptotics

Proofs are omitted in this section, because they are straightforward applications of results in Subsections 3.3.2, 3.3.3, 3.3.4 and Section 6.2.

This section uses a substantial body of notation. The reader will find Section 3.6 and the list of notation at the end of the dissertation particularly helpful.

6.3.1 Jacobi at the left edge

Theorem 6.3.1. *Take the CS decomposition of the 2n-by-2n Jacobi matrix model $J_{a,b}^{\infty}$, as in Proposition 6.2.1. Fix a positive integer k, and let ζ_k denote the kth smallest positive zero of j_a. Then, as $n \to \infty$ with $h = \frac{1}{n+(a+b+1)/2}$, we have*

1. *The kth diagonal entry of C is asymptotic to $\frac{1}{2n}\zeta_k$.*
2. *Let $x_i = 1 - h(n-i)$ and $\frac{1}{\sqrt{h}}FV_1 = (v_{ij}^{(1)})$. Then $\sum_{i=1}^{n-1} v_{ik}^{(1)} \mathbb{1}_{(x_i - h/2, x_i + h/2]}$ converges pointwise to $\sqrt{x} j_{a+1}(\zeta_k x) / \sqrt{2K^{\text{Bessel}}(a+1; \zeta_k^2, \zeta_k^2)}$ on $(0,1)$.*
3. *Let $x_i = (1-h) - h(n-i)$ and $\frac{1}{\sqrt{h}}FV_2 = (v_{ij}^{(2)})$. Then $\sum_{i=2}^{n} v_{ik}^{(2)} \mathbb{1}_{(x_i - h/2, x_i + h/2]}$ converges pointwise to $\sqrt{x} j_{a}(\zeta_k x) / \sqrt{2K^{\text{Bessel}}(a; \zeta_k^2, \zeta_k^2)}$ on $(0,1]$.*
4. *Let $x_i = (1-\frac{h}{2}) - h(n-i)$ and $\frac{1}{\sqrt{h}}FU_1 = (u_{ij}^{(1)})$. Then $\sum_{i=1}^{n} u_{ik}^{(1)} \mathbb{1}_{(x_i - h/2, x_i + h/2]}$ converges pointwise to $\sqrt{x} j_{a}(\zeta_k x) / \sqrt{2K^{\text{Bessel}}(a; \zeta_k^2, \zeta_k^2)}$ on $(0,1]$.*
5. *Let $x_i = (1-\frac{h}{2}) - h(n-i)$ and $\frac{1}{\sqrt{h}}FU_2 = (u_{ij}^{(2)})$. Then $\sum_{i=1}^{n} u_{ik}^{(2)} \mathbb{1}_{(x_i - h/2, x_i + h/2]}$ converges pointwise to $\sqrt{x} j_{a+1}(\zeta_k x) / \sqrt{2K^{\text{Bessel}}(a+1; \zeta_k^2, \zeta_k^2)}$ on $(0,1]$.*

In parts (2)–(5), the convergence is uniform on any interval $[\varepsilon, 1-\varepsilon]$ or $[\varepsilon, 1]$, $\varepsilon > 0$, as appropriate.

6.3.2 Jacobi at the right edge

Theorem 6.3.2. *Take the CS decomposition of the 2n-by-2n Jacobi matrix model $J^{\infty}_{a,b}$, as in Proposition 6.2.1. Fix a positive integer k, and let ζ_k denote the kth smallest positive zero of j_b. Then, as $n \to \infty$ with $h = \frac{1}{n+(a+b+1)/2}$, we have*

1. *The $(n+1-k, n+1-k)$ entry of S is asymptotic to $\frac{1}{2n}\zeta_k$.*
2. *Let $x_i = 1-h(n-i)$ and $-\frac{1}{\sqrt{h}}\Omega F V_1 = (v^{(1)}_{ij})$. Then $\sum_{i=1}^{n-1} v^{(1)}_{i,n+1-k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\sqrt{x}j_{b+1}(\zeta_k x)/\sqrt{2K^{\text{Bessel}}(b+1;\zeta_k^2,\zeta_k^2)}$ on $(0,1)$.*
3. *Let $x_i = (1-h)-h(n-i)$ and $-\frac{1}{\sqrt{h}}\Omega F V_2 = (v^{(2)}_{ij})$. Then $\sum_{i=2}^{n} v^{(2)}_{i,n+1-k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\sqrt{x}j_{b}(\zeta_k x)/\sqrt{2K^{\text{Bessel}}(b;\zeta_k^2,\zeta_k^2)}$ on $(0,1]$.*
4. *Let $x_i = (1-\frac{h}{2})-h(n-i)$ and $-\frac{1}{\sqrt{h}}\Omega F U_1 = (u^{(1)}_{ij})$. Then $\sum_{i=1}^{n} u^{(1)}_{i,n+1-k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\sqrt{x}j_{b+1}(\zeta_k x)/\sqrt{2K^{\text{Bessel}}(b+1;\zeta_k^2,\zeta_k^2)}$ on $(0,1]$.*
5. *Let $x_i = (1-\frac{h}{2})-h(n-i)$ and $-\frac{1}{\sqrt{h}}\Omega F U_2 = (u^{(2)}_{ij})$. Then $\sum_{i=1}^{n} u^{(2)}_{i,n+1-k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\sqrt{x}j_{b}(\zeta_k x)/\sqrt{2K^{\text{Bessel}}(b;\zeta_k^2,\zeta_k^2)}$ on $(0,1]$.*

In parts (2)–(5), the convergence is uniform on any interval $[\varepsilon, 1-\varepsilon]$ or $[\varepsilon, 1]$, $\varepsilon > 0$, as appropriate.

6.3.3 Jacobi near one-half

Theorem 6.3.3. *Suppose that $n = 2m$ is even, take the CSD of the 2n-by-2n Jacobi matrix model $J^{\infty}_{a,b}$ as in Proposition 6.2.1, and consider the SVD of the bottom-right block, $B_{22} = U_2 C V_2^T$. Fix an integer k, and define K_n as in Subsection 3.3.3. Then, as $n \to \infty$ over even values with $h = \frac{2}{n+(a+b+1)/2}$, we have*

1. *The (K_n+k)th diagonal entry of C^2 is asymptotic to $\frac{1}{2} + \frac{\pi}{2n}(\frac{a-b}{4} + \frac{1}{2} + k)$.*
2. *Let $x_i = (1-h)-h(m-i)$, $y_i = (1-\frac{h}{2})-h(m-i)$, and $\frac{1}{\sqrt{h}}(\Omega_m \oplus \Omega_m)PF(\sqrt{2}V_2) = (v^{(2)}_{ij})$. Then $\sum_{i=2}^{m} v^{(2)}_{i,K_n+k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $-\sqrt{2}\cos(\pi(\frac{a-b}{4}+$*

$\frac{1}{2}+k)x+\frac{(-a+b)\pi}{4})$ *on* $(0,1]$ *and* $\sum_{i=1}^{m} v^{(2)}_{m+i,K_n+k}\mathbb{1}_{(y_i-h/2,y_i+h/2]}$ *converges pointwise to* $\sqrt{2}\sin(\pi(\frac{a-b}{4}+\frac{1}{2}+k)y+\frac{(-a+b)\pi}{4})$ *on* $(0,1]$.

3. *Let* $x_i=(1-\frac{3h}{4})-h(m-i)$, $y_i=(1-\frac{h}{4})-h(m-i)$, *and* $\frac{1}{\sqrt{h}}(\Omega_m\oplus\Omega_m)PF(\sqrt{2}U_2)=(u^{(2)}_{ij})$. *Then* $\sum_{i=1}^{m} u^{(2)}_{i,K_n+k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ *converges pointwise to* $-\sqrt{2}\cos(\pi(\frac{a-b}{4}+\frac{1}{2}+k)x+\frac{(-(a+1)+b)\pi}{4})$ *on* $(0,1)$ *and* $\sum_{i=1}^{m} u^{(2)}_{m+i,K_n+k}\mathbb{1}_{(y_i-h/2,y_i+h/2]}$ *converges pointwise to* $\sqrt{2}\sin(\pi(\frac{a-b}{4}+\frac{1}{2}+k)y+\frac{(-(a+1)+b)\pi}{4})$ *on* $(0,1]$.

In parts (2)-(3), the convergence is uniform on any interval $[\varepsilon,1-\varepsilon]$ *or* $[\varepsilon,1]$, $\varepsilon>0$, *as appropriate.*

Theorem 6.3.4. *Suppose that* $n=2m+1$ *is odd, take the CSD of the* $2n$*-by-*$2n$ *Jacobi matrix model* $J^{\infty}_{a,b}$ *as in Proposition 6.2.1, and consider the SVD of the bottom-right block,* $B_{22}=U_2CV_2^T$. *Fix an integer* k, *and define* K_n *as in Subsection 3.3.3. Then, as* $n\to\infty$ *over odd values with* $h=\frac{2}{n+(a+b+1)/2}$, *we have*

1. *The* (K_n+k)*th diagonal entry of* C^2 *is asymptotic to* $\frac{1}{2}+\frac{\pi}{2n}(\frac{a-b}{4}+k)$.

2. *Let* $x_i=(1-\frac{h}{2})-h(m+1-i)$, $y_i=(1-h)-h(m-i)$, *and* $\frac{1}{\sqrt{h}}(\Omega_{m+1}\oplus\Omega_m)PF(\sqrt{2})V_2=(v^{(2)}_{ij})$. *Then* $\sum_{i=2}^{m+1} v^{(2)}_{i,K_n+k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ *converges pointwise to* $-\sqrt{2}\cos(\pi(\frac{a-b}{4}+k)x+\frac{(-a+b)\pi}{4})$ *on* $(0,1]$, *and* $\sum_{i=1}^{m} v^{(2)}_{m+1+i,K_n+k}\mathbb{1}_{(y_i-h/2,y_i+h/2]}$ *converges pointwise to* $\sqrt{2}\sin(\pi(\frac{a-b}{4}+k)y+\frac{(-a+b)\pi}{4})$ *on* $(0,1]$.

3. *Let* $x_i=(1-\frac{h}{4})-h(m+1-i)$, $y_i=(1-\frac{3h}{4})-h(m-i)$, *and* $\frac{1}{\sqrt{h}}(\Omega_{m+1}\oplus\Omega_m)PF(\sqrt{2}U_2)=(u^{(2)}_{ij})$. *Then* $\sum_{i=1}^{m+1} u^{(2)}_{i,K_n+k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ *converges pointwise to* $-\sqrt{2}\cos(\pi(\frac{a-b}{4}+k)x+\frac{(-(a+1)+b)\pi}{4})$ *on* $(0,1]$, *and* $\sum_{i=1}^{m} u^{(2)}_{m+1+i,K_n+k}\mathbb{1}_{(y_i-h/2,y_i+h/2]}$ *converges pointwise to* $\sqrt{2}\sin(\pi(\frac{a-b}{4}+k)y+\frac{(-(a+1)+b)\pi}{4})$ *on* $(0,1)$.

In parts (2)-(3), the convergence is uniform on any interval $[\varepsilon,1-\varepsilon]$ *or* $[\varepsilon,1]$, $\varepsilon>0$, *as appropriate.*

6.3.4 Laguerre at the left edge

Theorem 6.3.5. *Take the SVD of the n-by-n Laguerre matrix model L_a^∞, as in Proposition 6.2.2. Fix a positive integer k, and let ζ_k denote the kth smallest positive zero of j_a. Then, as $n \to \infty$ with $h = \frac{1}{n+(a+1)/2}$, we have*

1. *The kth smallest singular value is asymptotic to $\frac{1}{2\sqrt{n}}\zeta_k$.*

2. *Let $x_i = (1-h) - h(n-i)$ and $\frac{1}{\sqrt{h}}FV = (v_{ij})$. Then $\sum_{i=1}^n v_{ik}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $j_a(\zeta_k\sqrt{x})/\sqrt{4K^{\text{Bessel}}(a;\zeta_k^2,\zeta_k^2)}$ on $(0,1]$.*

3. *Let $x_i = (1-\frac{h}{2}) - h(n-i)$ and $\frac{1}{\sqrt{h}}FU = (u_{ij})$. Then $\sum_{i=1}^n u_{ik}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $j_{a+1}(\zeta_k\sqrt{x})/\sqrt{4K^{\text{Bessel}}(a+1;\zeta_k^2,\zeta_k^2)}$ on $(0,1]$.*

In parts (2)-(3), the convergence is uniform on any interval $[\varepsilon,1]$, $\varepsilon > 0$.

Theorem 6.3.6. *Take the SVD of the n-by-$(n+1)$ Laguerre matrix model M_a^∞, as in Proposition 6.2.3. Fix a positive integer k, and let ζ_k denote the kth smallest positive zero of j_a. Then, as $n \to \infty$ with $h = \frac{1}{n+(a+1)/2}$, we have*

1. *The kth smallest singular value is asymptotic to $\frac{1}{2\sqrt{n}}\zeta_k$.*

2. *Let $x_i = (1-\frac{h}{2}) - h(n+1-i)$ and $\frac{1}{\sqrt{h}}FV = (v_{ij})$. Then $\sum_{i=1}^{n+1} v_{ik}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $j_{a-1}(\zeta_k\sqrt{x})/\sqrt{4K^{\text{Bessel}}(a-1;\zeta_k^2,\zeta_k^2)}$ on $(0,1]$.*

3. *With x_i and v_{ij} as in part (2), $\sum_{i=1}^{n+1} v_{i,n+1}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\sqrt{a}\,x^{(a-1)/2}$ on $(0,1]$.*

4. *Let $x_i = (1-h) - h(n-i)$ and $\frac{1}{\sqrt{h}}FU = (u_{ij})$. Then $\sum_{i=1}^n u_{ik}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $j_a(\zeta_k\sqrt{x})/\sqrt{4K^{\text{Bessel}}(a;\zeta_k^2,\zeta_k^2)}$ on $(0,1]$.*

In parts (2)-(4), the convergence is uniform on any interval $[\varepsilon,1]$, $\varepsilon > 0$.

6.3.5 Laguerre at the right edge

Theorem 6.3.7. *Take the SVD of the n-by-n Laguerre matrix model L_a^∞, as in Proposition 6.2.2. Fix a positive integer k, and let ζ_k denote the kth rightmost zero of* Ai. *Then, as $n \to \infty$ with $h = 2^{2/3}n^{-1/3}$, we have*

1. *The kth largest singular value is asymptotic to $2\sqrt{n + \frac{a+1}{2}} + 2^{-2/3}n^{-1/6}\zeta_k$.*
2. *Let $x_i = hi$ and $(-1)^n \frac{1}{\sqrt{h}}\Omega V = (v_{ij})$. Then $\sum_{i=1}^n v_{i,n+1-k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\mathrm{Ai}(x+\zeta_k)/\sqrt{K^{\mathrm{Airy}}(\zeta_k,\zeta_k)}$ on $[0,\infty)$.*
3. *Let $x_i = \frac{h}{2} + h(i-1)$ and $(-1)^n \frac{1}{\sqrt{h}}\Omega U = (u_{ij})$. Then $\sum_{i=1}^n u_{i,n+1-k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\mathrm{Ai}(x+\zeta_k)/\sqrt{K^{\mathrm{Airy}}(\zeta_k,\zeta_k)}$ on $[0,\infty)$.*

The convergence in parts (2)–(3) is uniform on any interval $[0, M]$, $M > 0$.

Theorem 6.3.8. *Take the SVD of the n-by-$(n+1)$ Laguerre matrix model M_a^∞, as in Proposition 6.2.3. Fix a positive integer k, and let ζ_k denote the kth rightmost zero of* Ai. *Then, as $n \to \infty$ with $h = 2^{2/3}n^{-1/3}$, we have*

1. *The kth largest singular value is asymptotic to $2\sqrt{n + \frac{a+1}{2}} + 2^{-2/3}n^{-1/6}\zeta_k$.*
2. *Let $x_i = \frac{h}{2} + h(i-1)$ and $(-1)^{n+1} \frac{1}{\sqrt{h}}\Omega V = (v_{ij})$. Then $\sum_{i=1}^n v_{i,n+1-k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\mathrm{Ai}(x+\zeta_k)/\sqrt{K^{\mathrm{Airy}}(\zeta_k,\zeta_k)}$ on $[0,\infty)$.*
3. *Let $x_i = hi$ and $(-1)^n \frac{1}{\sqrt{h}}\Omega U = (u_{ij})$. Then $\sum_{i=1}^n u_{i,n+1-k}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\mathrm{Ai}(x+\zeta_k)/\sqrt{K^{\mathrm{Airy}}(\zeta_k,\zeta_k)}$ on $[0,\infty)$.*

The convergence in parts (2)–(3) is uniform on any interval $[0, M]$, $M > 0$.

6.3.6 Hermite near zero

As special cases of Theorem 4.2.4, we have, for $n = 2m$ even,

$$(\Omega_m \oplus \Omega_m) P H^\infty P^T (\Omega_m \oplus \Omega_m) = \left[\begin{array}{c|c} & L_{-1/2}^\infty \\ \hline (L_{-1/2}^\infty)^T & \end{array}\right], \tag{6.3.1}$$

in which H^∞ is $2m$-by-$2m$ and $L^\infty_{-1/2}$ is m-by-m, and for $n = 2m+1$ odd,

$$(\Omega_{m+1} \oplus \Omega_m) P H^\infty P^T (\Omega_{m+1} \oplus \Omega_m) = \left[\begin{array}{c|c} & -(M^\infty_{1/2})^T \\ \hline -M^\infty_{1/2} & \end{array}\right], \tag{6.3.2}$$

in which H^∞ is $(2m+1)$-by-$(2m+1)$ and $M^\infty_{1/2}$ is m-by-$(m+1)$.

Theorem 6.3.9. *Suppose that $n = 2m$ is even, and consider the n-by-n Hermite matrix model H^∞. Compute an eigenvalue decomposition of $(\Omega_m \oplus \Omega_m) P H^\infty P^T (\Omega_m \oplus \Omega_m)$ and permute rows and columns to find*

$$(\Omega_m \oplus \Omega_m) P H^\infty P^T (\Omega_m \oplus \Omega_m) = \begin{bmatrix} U & -U \\ V & V \end{bmatrix} \begin{bmatrix} \Sigma & \\ & -\Sigma \end{bmatrix} \begin{bmatrix} U & -U \\ V & V \end{bmatrix}^T,$$

in which Σ has positive diagonal entries in increasing order and U and V are m-by-m. Fix a positive integer k, and let $\zeta_k = (k - \frac{1}{2})\pi$. As $n \to \infty$ over even values with $h = \frac{4}{2n+1}$, we have

1. *The kth diagonal entry of Σ is asymptotic to $\frac{1}{\sqrt{2n}}\zeta_k$.*

2. *Let $x_i = (1-h) - h(m-1)$ and $\sqrt{\frac{2}{h}} FV = (v_{ij})$. Then $\sum_{i=1}^m v_{ik} \mathbb{1}_{(x_i - h/2, x_i + h/2]}$ converges pointwise to $\sqrt{\frac{2}{\pi}} \zeta_k^{-1/2} x^{-1/4} \cos(\zeta_k \sqrt{x}) / \sqrt{4K^{\text{Bessel}}(-\frac{1}{2}; \zeta_k^2, \zeta_k^2)}$ on $(0, 1]$.*

3. *Let $x_i = (1 - \frac{h}{2}) - h(m-i)$ and $\sqrt{\frac{2}{h}} FU = (u_{ij})$. Then $\sum_{i=1}^m u_{ik} \mathbb{1}_{(x_i - h/2, x_i + h/2]}$ converges pointwise to $\sqrt{\frac{2}{\pi}} \zeta_k^{-1/2} x^{-1/4} \sin(\zeta_k \sqrt{x}) / \sqrt{4K^{\text{Bessel}}(\frac{1}{2}; \zeta_k^2, \zeta_k^2)}$ on $(0, 1]$.*

In parts (2)–(3), the convergence is uniform on any interval $[\varepsilon, 1]$, $\varepsilon > 0$.

Theorem 6.3.10. *Suppose that $n = 2m+1$ is odd, and consider the n-by-n Hermite matrix model H^∞. Compute an eigenvalue decomposition of $-(\Omega_{m+1} \oplus \Omega_m) P H^\infty P^T (\Omega_{m+1} \oplus$*

Ω_m) and permute rows and columns to find

$$-(\Omega_{m+1}\oplus\Omega_m)PH^{\infty}P^T(\Omega_{m+1}\oplus\Omega_m) = \begin{bmatrix} \hat{V} & -\hat{V} & z \\ U & U & 0 \end{bmatrix} \begin{bmatrix} \Sigma & & \\ & -\Sigma & \\ & & 0 \end{bmatrix} \begin{bmatrix} \hat{V} & -\hat{V} & z \\ U & U & 0 \end{bmatrix}^T,$$

in which Σ is m-by-m with positive diagonal entries in increasing order, $\hat{V}$ is $(m+1)$-by-m, U is m-by-m, and z is $(m+1)$-by-1. Fix a positive integer k, and let $\zeta_k = k\pi$. As $n \to \infty$ over odd values with $h = \frac{4}{2n+1}$, we have

1. *The kth diagonal entry of Σ is asymptotic to $\frac{1}{\sqrt{2n}}\zeta_k$.*
2. *Let $x_i = (1-\frac{h}{2})-h(m+1-i)$ and $\sqrt{\frac{2}{h}}F\hat{V} = (v_{ij})$. Then $\sum_{i=1}^{m+1} v_{ik}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\sqrt{\frac{2}{\pi}}\zeta_k^{-1/2}x^{-1/4}\cos(\zeta_k\sqrt{x})/\sqrt{4K^{\text{Bessel}}(-\frac{1}{2};\zeta_k^2,\zeta_k^2)}$ on $(0,1]$.*
3. *Let x_i be defined as in part (2), and let $\frac{1}{\sqrt{h}}Fz = (z_i)$. Then $\sum_{i=1}^{m+1} z_i\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\frac{1}{\sqrt{2}}x^{-1/4}$ on $(0,1]$.*
4. *Let $x_i = (1-h)-h(m-i)$ and $\sqrt{\frac{2}{h}}FU = (u_{ij})$. Then $\sum_{i=1}^{m} u_{ik}\mathbb{1}_{(x_i-h/2,x_i+h/2]}$ converges pointwise to $\sqrt{\frac{2}{\pi}}\zeta_k^{-1/2}x^{-1/4}\sin(\zeta_k\sqrt{x})/\sqrt{4K^{\text{Bessel}}(\frac{1}{2};\zeta_k^2,\zeta_k^2)}$ on $(0,1]$.*

In parts (2)–(4), the convergence is uniform on any interval $[\varepsilon,1]$, $\varepsilon > 0$.

Remark 6.3.11. Theorems 6.3.9 and 6.3.10 hold because of the perfect symmetry of the $\beta = \infty$ Hermite ensemble: the eigenvalues come in positive/negative pairs, whose absolute values are precisely the singular values of a Laguerre matrix model. For finite β, the diagonal of the Hermite matrix model becomes nonzero, and the positive/negative symmetry of the eigenvalues is broken. In more technical language, the Hermite ensemble and the Laguerre chiral ensemble coincide at $\beta = \infty$, and the theorems say as much about the chiral ensemble as they do about the Hermite ensemble. When stepping to $\beta < \infty$, care must be taken to differentiate the Hermite ensemble from the chiral ensemble. The chiral ensemble is discussed in [14].

6.3.7 Hermite at the right edge

Theorem 6.3.12. *Take the eigenvalue decomposition of the n-by-n Hermite matrix model H^∞, as in Proposition 6.2.4. Fix a positive integer k, and let ζ_k denote the kth rightmost zero of* Ai. *Then, as $n \to \infty$ with $h = n^{-1/3}$, we have*

1. *The kth rightmost eigenvalue of H^∞ is asymptotic to $\sqrt{2n+1} + \frac{1}{\sqrt{2}} n^{-1/6} \zeta_k$.*
2. *Let $x_i = hi$ and $\frac{1}{\sqrt{h}} Q = (q_{ij})$. Then $\sum_{i=1}^{n} q_{i,n+1-k} \mathbb{1}_{(x_i - h/2, x_h + h/2]}$ converges pointwise to $\mathrm{Ai}(x + \zeta_k) / \sqrt{K^{\mathrm{Airy}}(\zeta_k, \zeta_k)}$ on $[0, \infty)$. The convergence is uniform on any interval $[0, M]$, $M > 0$.*

Chapter 7

Differential operator limits: the zero temperature case

Take a moment to study the operators in Figures 1.0.4–1.0.6, and compare them with the results in Section 6.3. It appears that the Airy, Bessel, and sine operators play roles in the spectra of the zero temperature matrix models as $n \to \infty$. Recalling the soft edge, hard edge, and bulk from Section 2.1, the operators appear to be associated with scaling limits:

scaling limit	operator	
soft edge	$\mathcal{A}$	(Airy)
hard edge	$\mathcal{J}_a$, $\tilde{\mathcal{J}}_a$	(Bessel)
bulk	$\mathcal{J}_{-1/2}$, $\tilde{\mathcal{J}}_{-1/2}$	(sine)

The present chapter argues that not only do the spectra of the matrix models resemble the spectra of the continuous operators, but actually the matrix models themselves are finite difference approximations to the Airy, Bessel, and sine operators. In order to view the matrix models as finite difference approximations, the models must be recentered and rescaled as $n \to \infty$. For example, a Laguerre model can

be viewed as a finite difference approximation of either an Airy or Bessel operator, depending on the rescaling. The situation is exactly as in Figure 2.1.3.

7.1 Overview

The limiting differential operators are indicated in the following table.

Matrix model ($\beta = \infty$)	Scaling limit	Scaling limit type	Differential operator
Laguerre (square or rectangular)	right edge	soft edge	$\mathcal{A}$
Hermite	right edge	soft edge	$\mathcal{A}$
Jacobi	left edge	hard edge	$\tilde{\mathcal{J}}_a$ type (i) b.c.
Jacobi	right edge	hard edge	$\tilde{\mathcal{J}}_b$ type (i) b.c.
Laguerre (square)	left edge	hard edge	$\mathcal{J}_a$ type (i) b.c.
Laguerre (rectangular)	left edge	hard edge	$\mathcal{J}_{a-1}$ type (ii) b.c.
Jacobi (n even)	center	bulk	$\begin{bmatrix} & -\tilde{\mathcal{J}}^*_{-1/2} \\ -\tilde{\mathcal{J}}_{-1/2} & \end{bmatrix}$ type (i) b.c.
Jacobi (n odd)	center	bulk	$\begin{bmatrix} & -\tilde{\mathcal{J}}^*_{-1/2} \\ -\tilde{\mathcal{J}}_{-1/2} & \end{bmatrix}$ type (ii) b.c.
Hermite (n even)	center	bulk	$\begin{bmatrix} & \mathcal{J}_{-1/2} \\ \mathcal{J}^*_{-1/2} & \end{bmatrix}$ type (i) b.c.
Hermite (n odd)	center	bulk	$\begin{bmatrix} & \mathcal{J}^*_{-1/2} \\ \mathcal{J}_{-1/2} & \end{bmatrix}$ type (ii) b.c.

Remark 7.1.1. The "center" scaling limits in the table refer to the very center of the spectrum: $\frac{1}{2}$ for Jacobi and 0 for Hermite. Applying the stochastic operator approach to other bulk locations in the Jacobi and Hermite ensembles, and anywhere in the bulk of the Laguerre ensemble, is an open problem.

7.2 Soft edge

In this section and the remaining two, we make a slight abuse of notation. Several times, we refer to row $\lfloor \frac{M}{h} \rfloor$ or $\lceil \frac{\varepsilon}{h} \rceil$ of an n-by-n matrix, in which M and ε are fixed constants, and $h = h_n$ is the step size, as in the previous chapter. For small values of n, these indices may be greater than n, referring to rows that do not exist. However, for sufficiently large n, the indices are valid. This is sufficient, since all of the theorems concern large n asymptotics.

7.2.1 Laguerre at the right edge

The first two theorems concern the square Laguerre matrix model L_a^∞.

Theorem 7.2.1. *With step size $h = 2^{2/3}n^{-1/3}$ and mesh $x_k = hk$, $k = 1, \ldots, n$, make the approximation*

$$\frac{h}{4}(\Omega(L_a^\infty)^T L_a^\infty \Omega - 4(n + \tfrac{a+1}{2})I_n) = \frac{1}{h^2}D_2 - \frac{1}{4}\begin{bmatrix} 2x_1 & x_1 & & & \\ x_1 & 2x_2 & x_2 & & \\ & \ddots & \ddots & \ddots & \\ & & x_{n-2} & 2x_{n-1} & x_{n-1} \\ & & & x_{n-1} & 2x_n \end{bmatrix} + E.$$

Then E is symmetric tridiagonal, and its entries are $O(h)$ as $n \to \infty$, uniformly.

Proof. E is clearly symmetric tridiagonal. Its diagonal entries are exactly $-\frac{a+1}{4}h$. For $k = 1, \ldots, n$-1, the $(k, k+1)$ entry is $-h^{-2} + \frac{k}{4}h + \frac{h}{4}\sqrt{(n-k)(a+n-k)}$. Rewriting the last term as $\frac{h}{4}\sqrt{(\frac{a}{2}+n-k)^2 - (\frac{a}{2})^2}$ and taking a series expansion of the square root factor about $\frac{a}{2} + n - k$, we see that the entry equals $\frac{h}{8}(a + R[\frac{a^2}{4\sqrt{a+1}}])$, in which $R[x]$ represents a quantity bounded in magnitude by x. □

Theorem 7.2.2. *With step size $h = 2^{2/3}n^{-1/3}$ and mesh $x_k = h(k - \frac{1}{2})$, $k = 1, \ldots, n$,*

make the approximation

$$\frac{h}{4}(\Omega L_a^\infty (L_a^\infty)^T \Omega - 4(n+\tfrac{a+1}{2})I_n) = \frac{1}{h^2}\tilde{D}_2 - \frac{1}{4}\begin{bmatrix} 2x_1 & x_1 & & & & \\ x_1 & 2x_2 & x_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & x_{n-2} & 2x_{n-1} & x_{n-1} & \\ & & & x_{n-1} & 2x_n & \end{bmatrix} + E,$$

in which $\tilde{D}_2$ is obtained from D_2 by replacing the $(1,1)$ entry with -3. Then E is symmetric tridiagonal, and its entries are $O(h)$ as $n \to \infty$, uniformly.

Proof. E is clearly symmetric tridiagonal, and every diagonal entry equals $-\frac{a+1}{4}h$. For $k = 1, \ldots, n-1$, the $(k, k+1)$ entry is $-\frac{1}{h^2} + \frac{-1+2k}{8}h + \frac{h}{4}\sqrt{(n-k)(a+1+n-k)}$. Rewriting the last term as $\frac{h}{4}\sqrt{(\frac{a+1}{2}+n-k)^2 - (\frac{a+1}{2})^2}$ and taking a series expansion of the square root factor about $\frac{a+1}{2} + n - k$, we see that the entry equals $\frac{h}{8}(a + R[\frac{(a+1)^2}{4\sqrt{a+2}}])$. □

Claim 7.2.3. *$(L_a^\infty)^T L_a^\infty$ and $L_a^\infty (L_a^\infty)^T$, scaled at the soft edge, are finite difference approximations to the Airy operator $\mathcal{A}$.*

The claim is supported by Theorems 6.3.7, 7.2.1, and 7.2.2.

The next two theorems concern the rectangular Laguerre model M_a^∞.

Theorem 7.2.4. *With step size $h = 2^{2/3}n^{-1/3}$ and mesh $x_k = h(k-\frac{1}{2})$, $k = 1, \ldots, n+1$, make the approximation*

$$\frac{h}{4}(\Omega (M_a^\infty)^T M_a^\infty \Omega - 4(n+\tfrac{a+1}{2})I_{n+1}) = \frac{1}{h^2}\tilde{D}_2 - \frac{1}{4}\begin{bmatrix} 2x_1 & x_1 & & & & \\ x_1 & 2x_2 & x_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & x_{n-1} & 2x_n & x_n & \\ & & & x_n & 2x_{n+1} & \end{bmatrix} + E,$$

in which $\tilde{D}_2$ is obtained from D_2 by replacing the $(1,1)$ entry with -3. Then E is symmetric tridiagonal, and its entries are $O(h)$ as $n \to \infty$, uniformly.

Proof. E is clearly symmetric tridiagonal. For $k = 2, \ldots, n$, the (k,k) entry is $-\frac{a+1}{4}h$. The $(1,1)$ entry and the $(n+1, n+1)$ entry both equal $-\frac{2a+1}{4}h$. For $k = 1, \ldots, n$, the $(k, k+1)$ entry is $-\frac{1}{h^2} + \frac{-1+2k}{8}h + \frac{h}{4}\sqrt{(n+1-k)(a+n-k)}$. Rewriting the last term as $\frac{h}{4}\sqrt{(\frac{a+1}{2}+n-k)^2 - (\frac{a-1}{2})^2}$ and taking a series expansion of the square root factor about $\frac{a+1}{2}+n-k$, we see that the entry is $\frac{h}{8}(a + R[\frac{(a-1)^2}{4\sqrt{a}}])$. □

Theorem 7.2.5. *With step size $h = 2^{2/3}n^{-1/3}$ and mesh $x_k = hk$, $k = 1, \ldots, n$, make the approximation*

$$\frac{h}{4}(\Omega M_a^\infty (M_a^\infty)^T \Omega - 4(n+\tfrac{a+1}{2})I_n) = \frac{1}{h^2}D_2 - \frac{1}{4}\begin{bmatrix} 2x_1 & x_1 & & & \\ x_1 & 2x_2 & x_2 & & \\ & \ddots & \ddots & \ddots & \\ & & x_{n-2} & 2x_{n-1} & x_{n-1} \\ & & & x_{n-1} & 2x_n \end{bmatrix} + E.$$

Then E is symmetric tridiagonal, and its entries are $O(h)$ as $n \to \infty$, uniformly.

Proof. E is clearly symmetric tridiagonal. The diagonal entries are exactly $-\frac{a+1}{4}h$. For $k = 1, \ldots, n-1$, the $(k, k+1)$ entry is $-\frac{1}{h^2} + \frac{k}{4}h + \frac{h}{4}\sqrt{(n-k)(a+n-k)}$. Rewriting the last term as $\frac{h}{4}\sqrt{(\frac{a}{2}+n-k)^2 - (\frac{a}{2})^2}$ and taking a series expansion of the square root factor about $\frac{a}{2}+n-k$, we see that the entry is $\frac{h}{8}(a + R[\frac{a^2}{4\sqrt{a+1}}])$. □

Claim 7.2.6. *$(M_a^\infty)^T M_a^\infty$ and $M_a^\infty (M_a^\infty)^T$, scaled at the soft edge, are finite difference approximations to the Airy operator $\mathcal{A}$.*

The claim is supported by Theorems 6.3.8, 7.2.4, and 7.2.5.

7.2.2 Hermite at the right edge

Claim 7.2.7. *Let H^∞ be the $(2m)$-by-$(2m)$ zero temperature Hermite matrix model, and let $h = 2^{2/3}m^{-1/3}$. We claim that $\frac{h}{4}(P(H^\infty)^2 P^T - 4(m+\frac{1}{4})I_{2m})$ is a finite*

difference approximation to $[\mathcal{A}\ _{\mathcal{A}}]$. *Similarly, when the size of the matrix model is odd,* $n = 2m+1$, *we claim that* $\frac{h}{4}(P(H^\infty)^2P^T - 4(m+\frac{3}{4})I_{2m+1})$ *is a finite difference approximation to* $[\mathcal{A}\ _{\mathcal{A}}]$, *in which* $h = 2^{2/3}m^{-1/3}$.

The claim is supported by Theorem 6.3.12 as well as Subsection 7.2.1. To apply Theorem 6.3.12, note that the eigenvalues of H^∞ come in positive/negative pairs, so that squaring H^∞ forms 2-dimensional eigenspaces. The λ_k eigenspace is spanned by a vector of the form $\left[\begin{smallmatrix} u \\ 0 \end{smallmatrix}\right]$ and a vector of the form $\left[\begin{smallmatrix} 0 \\ v \end{smallmatrix}\right]$, with u an $\lceil \frac{n}{2} \rceil$ vector and v an $\lfloor \frac{n}{2} \rfloor$ vector. To apply Subsection 7.2.1, note that when $n = 2m$ is even, $(H^\infty)^2$ can be expressed in terms of the m-by-m Laguerre model,

$$P(H^\infty)^2P^T = \begin{bmatrix} \Omega L^\infty_{-1/2}(L^\infty_{-1/2})^T\Omega & \\ & \Omega(L^\infty_{-1/2})^T L^\infty_{-1/2}\Omega \end{bmatrix},$$

so that Theorems 7.2.1 and 7.2.2 can be used, and when $n = 2m+1$ is odd, $(H^\infty)^2$ can be expressed in terms of the m-by-$(m+1)$ rectangular Laguerre model,

$$P(H^\infty)^2P^T = \begin{bmatrix} \Omega(M^\infty_{1/2})^T M^\infty_{1/2}\Omega & \\ & \Omega M^\infty_{1/2}(M^\infty_{1/2})^T\Omega \end{bmatrix},$$

so that Theorems 7.2.4 and 7.2.5 can be used.

The Hermite matrix model can also be analyzed directly, without appealing to the Laguerre matrix model.

Theorem 7.2.8. *With step size* $h = n^{-1/3}$ *and mesh* $x_k = hk$, $k = 1, \ldots, n$, *make the approximation*

$$\sqrt{\frac{2}{h}}(H^\infty - \sqrt{2}h^{-3/2}I_n) = \frac{1}{h^2}D_2 - \frac{1}{2}\begin{bmatrix} 0 & x_1 & & & \\ x_1 & 0 & x_2 & & \\ & \ddots & \ddots & \ddots & \\ & & x_{n-2} & 0 & x_{n-1} \\ & & & x_{n-1} & 0 \end{bmatrix} + E.$$

Then E is symmetric tridiagonal, and the entries in rows $1, \ldots, \lfloor \frac{M}{h} \rfloor$ of E are uniformly $O(h^2)$, for any fixed $M > 0$.

Proof. E is clearly symmetric tridiagonal, and its diagonal entries are exactly zero. For $k = 1, \ldots, n-1$, the $(k, k+1)$ entry is $\sqrt{\frac{n-k}{h}} + \frac{hk}{2} - \frac{1}{h^2}$. Taking a series expansion of $\sqrt{n-k}$ about $\sqrt{n}$, we see that the entry equals $\frac{1}{\sqrt{h}}\sqrt{n} + \frac{1}{\sqrt{h}}\frac{1}{2}n^{-1/2}(-k) + R[-\frac{1}{4}\frac{1}{\sqrt{h}}(n-k)^{-3/2}(-k)^2] + \frac{hk}{2} - \frac{1}{h^2}$. Substituting $h = n^{-1/3}$, the entry becomes $n^{2/3} - \frac{1}{2}n^{-1/3}k + \frac{1}{2}n^{-1/3}k - n^{2/3} + R[\frac{1}{4}n^{1/6}(n-k)^{-3/2}k^2]$, which is bounded in magnitude by $\frac{1}{4}n^{1/6}(n-k)^{-3/2}k^2$. For all rows of interest, $k \leq \frac{M}{h}$, so the entry is bounded in magnitude by $\frac{1}{4}(1 - Mh)^{-3/2}M^2h^2$. □

The theorem only considers the first $O(n^{1/3})$ rows of the Hermite matrix model. However, from Theorem 6.3.12, the k "rightmost" eigenvectors have almost all of their support within this regime, for any k. Under the $n \to \infty$ scaling limit, the bottom-right portion of the matrix is often irrelevant.

Claim 7.2.9. *$\sqrt{\frac{2}{h}}(H^\infty - \sqrt{2}h^{-3/2}I_n)$ is a finite difference approximation to $\mathcal{A}$.*

The claim is supported by Theorems 6.3.12 and 7.2.8.

7.3 Hard edge

7.3.1 Jacobi at the left edge

Theorem 7.3.1. *Let B_{22} denote the bottom-right n-by-n block of the $2n$-by-$2n$ Jacobi matrix model $J^\infty_{a,b}$. With step size $h = \frac{1}{n+(a+b+1)/2}$ and mesh $x_k = (1 - \frac{h}{2}) - h(n-k)$,*

$k = 1, \ldots, n$, make the approximation

$$\frac{2}{h} F B_{22} F = -\left(\frac{1}{h} D_1\right) + \left(a + \frac{1}{2}\right) \cdot \frac{1}{2} \begin{bmatrix} \frac{1}{x_1} & \frac{1}{x_1} & & & \\ & \frac{1}{x_2} & \frac{1}{x_2} & & \\ & & \frac{1}{x_3} & \ddots & \\ & & & \ddots & \frac{1}{x_{n-1}} \\ & & & & \frac{1}{x_n} \end{bmatrix} + E.$$

Then E is upper bidiagonal, and, for fixed $\varepsilon \in (0, 1]$, the entries in rows $\lceil \frac{\varepsilon}{h} \rceil, \ldots, n$ are $O(h)$, uniformly.

Proof. E is clearly upper bidiagonal. Check that entry $(n+1-k, n+1-k)$ of the left hand side equals

$$\frac{1}{h} \left(\tfrac{x_{n+1-k}}{h} + \tfrac{a-b}{2}\right)^{1/2} \left(\tfrac{x_{n+1-k}}{h} + \tfrac{a+b}{2}\right)^{1/2} \left(\tfrac{x_{n+1-k}}{h}\right)^{-1/2} \left(\tfrac{x_{n+1-k}}{h} - \tfrac{1}{2}\right)^{-1/2}.$$

After rewriting this expression as

$$\frac{1}{h} \left(\left(\tfrac{x_{n+1-k}}{h} + \tfrac{a}{2}\right)^2 - \tfrac{b^2}{4}\right)^{1/2} \left(\left(\tfrac{x_{n+1-k}}{h} - \tfrac{1}{4}\right)^2 - \tfrac{1}{16}\right)^{-1/2},$$

it is straightforward to check that the entry is

$$\tfrac{1}{h} + (a + \tfrac{1}{2}) \cdot \tfrac{1}{2} \tfrac{1}{x_{n+1-k}} + O(h),$$

with the $O(h)$ term bounded uniformly over all k such that $n + 1 - k \geq \lceil \frac{\varepsilon}{h} \rceil$.

The superdiagonal error terms can be bounded similarly. □

Theorem 7.3.2. *Let B_{11} denote the top-left n-by-n block of the $2n$-by-$2n$ Jacobi matrix model $J^{\infty}_{a,b}$. With step size $h = \frac{1}{n+(a+b+1)/2}$ and mesh $x_k = 1 - h(n-k)$,*

$k = 1, \ldots, n$, make the approximation

$$\frac{2}{h} F B_{11} F = \frac{1}{h}\left(-D_1^T\right) + \left(a + \frac{1}{2}\right) \cdot \frac{1}{2} \begin{bmatrix} \frac{1}{x_1} & & & & \\ \frac{1}{x_1} & \frac{1}{x_2} & & & \\ & \frac{1}{x_2} & \frac{1}{x_3} & & \\ & & \ddots & \ddots & \\ & & & \frac{1}{x_{n-1}} & \frac{1}{x_n} \end{bmatrix} + E.$$

Then E is lower bidiagonal, and, for fixed $\varepsilon \in (0, 1]$, the entries in rows $\lceil \frac{\varepsilon}{h} \rceil, \ldots, n-1$ are $O(h)$, uniformly.

The proof is similar to the proof of the previous theorem. Note that the theorem says nothing about the last row of $\frac{2}{h} F B_{11} F$. In fact, the (n, n) entry is qualitatively different from nearby entries, enforcing the right boundary condition evident in the previous chapter. We omit details for brevity.

Claim 7.3.3. *The Jacobi matrix model scaled at the hard edge on the left, $\frac{2}{h}(F_n \oplus F_n) J_{a,b}^{\infty} (F_n \oplus F_n)$, is a finite difference approximation to*

$$\begin{bmatrix} \tilde{J}_a^* & * \\ * & \tilde{J}_a \end{bmatrix},$$

with type (i) boundary conditions.

The claim is supported by Theorems 6.3.1, 7.3.1, and 7.3.2.

Remark 7.3.4. Notice that type (ii) boundary conditions are not seen in this subsection. Perhaps the rectangular Jacobi model, conjectured in Remark 4.2.3, would reveal type (ii) boundary conditions. This idea is suggested by Subsection 7.3.3.

7.3.2 Jacobi at the right edge

Claim 7.3.5. *The Jacobi matrix model scaled at the hard edge on the right,* $\frac{2}{h}(\Omega_n \oplus \Omega_n)(F_n \oplus F_n)J_{a,b}^{\infty}(F_n \oplus F_n)(\Omega_n \oplus \Omega_n)$*, is a finite difference approximation to*

$$\begin{bmatrix} * & \tilde{\mathcal{J}}_b \\ \tilde{\mathcal{J}}_b^* & * \end{bmatrix},$$

with type (i) boundary conditions.

This claim is equivalent to the claim of the previous subsection, considering (4.2.1).

7.3.3 Laguerre at the left edge

Theorem 7.3.6. *With step size* $h = \frac{1}{n+(a+1)/2}$ *and mesh* $x_k = (1 - \frac{h}{2}) - h(n-k)$, $k = 1, \ldots, n$*, make the approximation*

$$\frac{2}{\sqrt{h}} F L_a^{\infty} F$$

$$= -2\,\mathrm{diag}(\sqrt{x_1}, \ldots, \sqrt{x_n}) \left(\frac{1}{h} D_1\right) + a \cdot \frac{1}{2} \begin{bmatrix} \frac{1}{\sqrt{x_1}} & \frac{1}{\sqrt{x_1}} & & & \\ & \frac{1}{\sqrt{x_2}} & \frac{1}{\sqrt{x_2}} & & \\ & & \frac{1}{\sqrt{x_3}} & \ddots & \\ & & & \ddots & \frac{1}{\sqrt{x_{n-1}}} \\ & & & & \frac{1}{\sqrt{x_n}} \end{bmatrix} + E.$$

Then E *is upper bidiagonal, and the entries in rows* $\lceil \frac{\varepsilon}{h} \rceil, \ldots, n$ *of* E *are uniformly* $O(h)$*, for any fixed* $0 < \varepsilon < 1$.

Proof. For $k = 1, \ldots, n$, the (k,k) entry of E is $(2\sqrt{k+a} - 2\sqrt{k+a/2} - \frac{a}{2\sqrt{k+a/2}})\frac{1}{\sqrt{h}}$. Taking a series expansion of $\sqrt{k+a}$ about $\sqrt{k+a/2}$, the entry can be bounded in magnitude by $\frac{a^2}{16(\varepsilon-h)^{3/2}}h$ for all $k \geq \frac{\varepsilon}{h}$. The superdiagonal of E can be bounded in a similar fashion, at one point taking a series expansion of $\sqrt{k}$ about $\sqrt{k+a/2}$. $\square$

Claim 7.3.7. *The Laguerre matrix model L_a^∞, scaled at the hard edge, is a finite difference approximation to the Bessel operator $\mathcal{J}_a$ with type (i) boundary conditions.*

The claim is supported by Theorems 6.3.5 and 7.3.6.

Theorem 7.3.8. *With step size $h = \frac{1}{n+(a+1)/2}$ and mesh $x_k = (1-h) - h(n-k)$, $k = 1, \ldots, n$, make the approximation*

$$\frac{2}{\sqrt{h}} F M_a^\infty F = -2\,\mathrm{diag}(\sqrt{x_1}, \ldots, \sqrt{x_n}) \left(\frac{1}{h} (D_1)_{1:n,1:n+1} \right) + (a-1) \cdot \frac{1}{2} \begin{bmatrix} \frac{1}{\sqrt{x_1}} & \frac{1}{\sqrt{x_1}} & & & & \\ & \frac{1}{\sqrt{x_2}} & \frac{1}{\sqrt{x_2}} & & & \\ & & \ddots & \ddots & & \\ & & & \frac{1}{\sqrt{x_{n-1}}} & \frac{1}{\sqrt{x_{n-1}}} & \\ & & & & \frac{1}{\sqrt{x_n}} & \frac{1}{\sqrt{x_n}} \end{bmatrix} + E.$$

Then all nonzero entries of E lie on the main diagonal and superdiagonal, and the entries in rows $\lceil \frac{\varepsilon}{h} \rceil, \ldots, n$ of E are uniformly $O(h)$, for any fixed $0 < \varepsilon < 1$.

Proof. For $k = 1, \ldots, n$, the (k,k) entry of E is $(2\sqrt{k+a-1} - 2\sqrt{k+(a-1)/2} - \frac{a-1}{2\sqrt{k+(a-1)/2}}) \frac{1}{\sqrt{h}}$. Taking a series expansion of $\sqrt{k+a-1}$ about $\sqrt{k+(a-1)/2}$, the entry can be bounded in magnitude by $\frac{(a-1)^2}{16(\varepsilon-h)^{3/2}} h$ for all $k \geq \frac{\varepsilon}{h}$. The superdiagonal of E can be bounded in a similar fashion, at one point taking a series expansion of $\sqrt{k}$ about $\sqrt{k+(a-1)/2}$. □

Claim 7.3.9. *The rectangular Laguerre matrix model M_a^∞, scaled at the hard edge, is a finite difference approximation to the Bessel operator $\mathcal{J}_{a-1}$ with type (ii) boundary conditions.*

The claim is supported by Theorems 6.3.6 and 7.3.8.

7.4 Bulk

7.4.1 Jacobi near one-half

Theorem 7.4.1. *Suppose that $a = b$ and that $n = 2m$ is even, and let B_{22} denote the bottom-right block of the $2n$-by-$2n$ Jacobi matrix model $J_{a,b}^{\infty}$. With step size $h = \frac{2}{n+(a+b+1)/2}$, make the approximation*

$$\frac{4}{h}\left((\Omega_m \oplus \Omega_m)PFB_{22}^T B_{22} F P^T (\Omega_m \oplus \Omega_m) - \frac{1}{2}I\right) = \begin{bmatrix} & \frac{1}{h}D_1^T \\ \frac{1}{h}D_1 & \end{bmatrix} + E.$$

Then E is symmetric and has the same sparsity pattern as $\begin{bmatrix} & \frac{1}{h}D_1^T \\ \frac{1}{h}D_1 & \end{bmatrix}$. Also, for any fixed $\varepsilon \in (0,1]$, the entries in rows $\lceil \frac{\varepsilon}{h} \rceil, \ldots, m$ of the bottom-left block of E are $O(h)$, uniformly.

Proof. First, consider the symmetry and sparsity pattern claims. The $(1,1)$ entry of the left hand side is $\frac{4}{h}\frac{a-b}{2(a+b+2)} = 0$. For $k = 2, \ldots, m$, the (k,k) entry of the left hand side is

$$\frac{4}{h}\left(\frac{2(k-1)(b+2(k-1))}{(a+b+2(2(k-1)))(1+a+b+2(2(k-1)))} + \frac{(1+a+b+2(k-1))(a+(2k-1))}{(1+a+b+2(k-1))(a+b+2(2k-1))} - \frac{1}{2}\right) = 0.$$

For $k = 1, \ldots, m$, the $(m+k, m+k)$ entry of the left hand side is

$$\frac{4}{h}\left(\frac{(2k-1)(b+(2k-1))}{(a+b+2(2k-1))(1+a+b+2(2k-1)))} + \frac{(1+a+b+(2k-1))(a+2k)}{(1+a+b+2(2k-1))(a+b+2(2k))} - \frac{1}{2}\right) = 0.$$

Now E is clearly symmetric with the indicated sparsity pattern.

For $2 \leq k \leq m$, the (k,k) entry of the bottom-left block of E is

$$-\frac{4}{h}\sqrt{\frac{a+(2k-1)}{a+b+2(2k-1)}}\sqrt{\frac{b+(2k-1)}{a+b+2(2k-1)}}\sqrt{\frac{2k-1}{a+b+1+2(2k-1)}}\sqrt{\frac{a+b+1+2(k-1)}{a+b+1+2(2(k-1))}} - (-\frac{1}{h}),$$

which can be rewritten

$$-\tfrac{4}{h}\left((2k+\tfrac{a+b}{2}-1)^2-(\tfrac{a-b}{2})^2\right)^{1/2}\left((2k+\tfrac{a+b}{2}-1)^2-(\tfrac{a+b}{2})^2\right)^{1/2}$$
$$\cdot((4k+a+b-2)^2-1)^{-1/2}((4k+a+b-2)^2)^{-1/2}+\tfrac{1}{h}.$$

Asymptotically, then, the entry is

$$\begin{aligned}
&-\tfrac{4}{h}(2k+\tfrac{a+b}{2}-1+O(k_{\min}^{-1}))(2k+\tfrac{a+b}{2}-1+O(k_{\min}^{-1}))\\
&\qquad\cdot((4k+a+b-2)^{-1}+O(k_{\min}^{-3}))(4k+a+b-2)^{-1}+\tfrac{1}{h}\\
&=-\tfrac{4}{h}(\tfrac{1}{4}+O(k_{\min}^{-2}))+\tfrac{1}{h}\\
&=O(h^{-1}k_{\min}^{-2})
\end{aligned}$$

for all $k \geq k_{\min} \geq 1$ as $n \to \infty$. (The implicit constant is uniform for all $k \geq k_{\min}$.) If the sequence of $k_{\min}$'s is bounded below by $\frac{\varepsilon}{h}$, then the entry is $O(h)$.

The superdiagonal can be bounded similarly. □

Theorem 7.4.2. *Suppose that $a = b$ and that $n = 2m+1$ is odd, and let B_{22} denote the bottom-right block of the $2n$-by-$2n$ zero temperature Jacobi matrix model $J_{a,b}^{\infty}$. With step size $h = \frac{2}{n+(a+b+1)/2}$, make the approximation*

$$\frac{4}{h}\left((\Omega_{m+1}\oplus\Omega_m)PFB_{22}^TB_{22}FP^T(\Omega_{m+1}\oplus\Omega_m)-\frac{1}{2}I\right)$$
$$=\begin{bmatrix} & \frac{1}{h}(D_1)_{1:m,1:m+1}^T \\ \frac{1}{h}(D_1)_{1:m,1:m+1} & \end{bmatrix}+E.$$

Then E is symmetric and has the same sparsity pattern as the other term on the right hand side. Also, for any fixed $\varepsilon \in (0,1]$, the entries in rows $\lceil\frac{\varepsilon}{h}\rceil,\ldots,m$ of the bottom-left block of E are $O(h)$, uniformly.

The proof is similar to the proof of the previous theorem.

Claim 7.4.3. *When $a = b$ and $n = 2m$ is even, the Jacobi matrix model scaled in the center of the spectrum, $-\frac{4}{h}((\Omega_m \oplus \Omega_m)PFB_{22}^T B_{22} F P^T (\Omega_m \oplus \Omega_m) - \frac{1}{2}I)$, is a finite difference approximation to*

$$\begin{bmatrix} & -\tilde{\mathcal{J}}_{-1/2}^* \\ -\tilde{\mathcal{J}}_{-1/2} & \end{bmatrix},$$

with type (i) boundary conditions. When $a = b$ and $n = 2m+1$ is odd, the Jacobi matrix model scaled in the center of the spectrum, $-\frac{4}{h}((\Omega_{m+1} \oplus \Omega_m)PFB_{22}^T B_{22} F P^T (\Omega_{m+1} \oplus \Omega_m) - \frac{1}{2}I)$, is a finite difference approximation to

$$\begin{bmatrix} & -\tilde{\mathcal{J}}_{-1/2}^* \\ -\tilde{\mathcal{J}}_{-1/2} & \end{bmatrix},$$

with type (ii) boundary conditions.

The claim is supported by Theorems 6.3.3, 6.3.4, 7.4.1, and 7.4.2.

The above discussion applies only to the ultraspherical case $a = b$. When $a \neq b$, Theorems 6.3.3 and 6.3.4 suggest that similar finite difference interpretations exist, but with different boundary conditions.

7.4.2 Hermite near zero

Claim 7.4.4. *If H^∞ is the zero temperature $2m$-by-$2m$ Hermite matrix model, then, with $h = \frac{4}{2n+1}$, $\frac{2}{\sqrt{h}}(\Omega_m \oplus \Omega_m)(F_m \oplus F_m)PH^\infty P^T (F_m \oplus F_m)(\Omega_m \oplus \Omega_m)$ is a finite difference approximation to*

$$\begin{bmatrix} & \mathcal{J}_{-1/2} \\ \mathcal{J}_{-1/2}^* & \end{bmatrix},$$

with type (i) boundary conditions. If, instead, H^∞ is $(2m+1)$-by-$(2m+1)$, then, with $h = \frac{4}{2n+1}$, $\frac{2}{\sqrt{h}}(\Omega_{m+1} \oplus \Omega_m)(F_{m+1} \oplus F_m)PH^\infty P^T (F_{m+1} \oplus F_m)(\Omega_{m+1} \oplus \Omega_m)$ is a finite

difference approximation to

$$\begin{bmatrix} & \mathcal{J}^*_{-1/2} \\ \mathcal{J}_{-1/2} & \end{bmatrix},$$

with type (ii) boundary conditions.

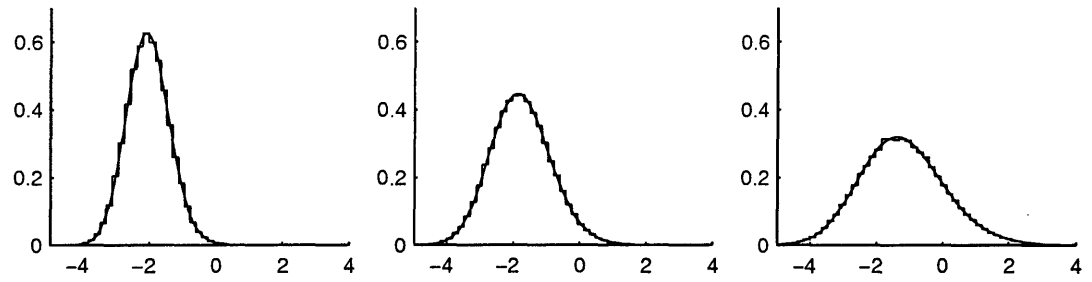
The claim is a natural consequence of (6.3.1) and Claim 7.3.7, in the even case, and (6.3.2) and Claim 7.3.9, in the odd case.

Chapter 8

Stochastic differential operator limits

The previous two chapters considered zero temperature ($\beta = \infty$) random matrix models. The present chapter moves to the finite β case, in which randomness actually plays a role. We argue that the random matrix models, rescaled appropriately, are finite difference approximations to stochastic differential operators. In other words, we add noise to the previous chapter.

The arguments work in two steps. First, we replace the matrix models with Gaussian approximations, a step that appears to be benign as $n \to \infty$. Second, we claim that the $n \to \infty$ limit of a diagonal or bidiagonal matrix of Gaussian entries is a diagonal operator that injects white noise, usually denoted W. However, we do not propose an interpretation, in the Itô or Stratonovich sense, for example, for white noise. The numerical experiments in this chapter justify the use of Gaussian approximations and may someday be useful for finding the correct interpretation of white noise. They alone, however, cannot justify the step from finite matrix to stochastic operator in our arguments.



(a) $\beta = 4$ (b) $\beta = 2$ (c) $\beta = 1$

Figure 8.1.1: Largest eigenvalue of a finite difference approximation to $\mathcal{A} + \frac{2}{\sqrt{\beta}}W$. The smooth curves are $n \to \infty$ theoretical densities (see Figure 2.1.2(a)), while the square waves are finite n histograms.

8.1 Overview

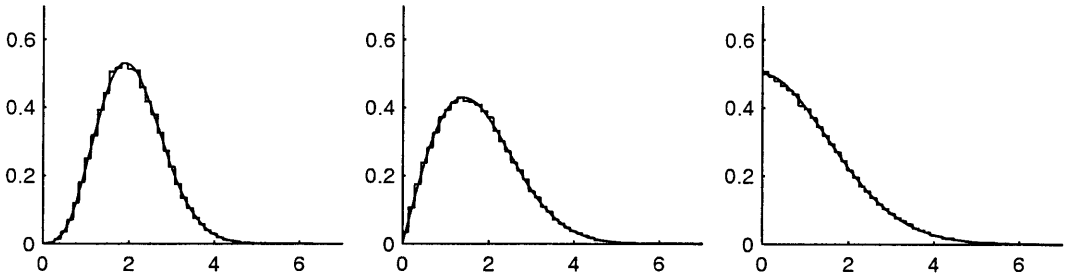
This chapter presents several conjectures regarding the eigenvalues and singular values of stochastic differential operators. These conjectures are motivated by the observation that scaling limits of random matrix models are finite difference approximations to the stochastic differential operators, and the conjectures are partially supported by numerical experiments. The conjectures follow. Each should be prefixed by, "Under the appropriate SDE interpretation for the diagonal noise operator W..."

Conjecture 8.1.1. *The largest eigenvalue of* $\mathcal{A} + \frac{2}{\sqrt{\beta}}W$ *follows the universal largest eigenvalue distribution with parameter* $\beta > 0$.

Results from numerical experiments are presented in Figure 8.1.1. More details can be found in Section 8.4.

Conjecture 8.1.2. *The smallest singular value of* $\tilde{\mathcal{J}}_a + \sqrt{\frac{2}{\beta}}\frac{1}{\sqrt{y}}W$ *with type (i) boundary conditions follows the universal smallest singular value distribution with parameters* β, a.

Results from numerical experiments are presented in Figure 8.1.2. More details can be found in Section 8.5.



(a) $a = 0, \beta = 4$ (b) $a = 0, \beta = 2$ (c) $a = 0, \beta = 1$

Figure 8.1.2: Smallest singular value of a finite difference approximation to $\tilde{\mathcal{J}}_a + \sqrt{\frac{2}{\beta}}\frac{1}{\sqrt{y}}W$ with type (i) boundary conditions. The smooth curves are $n \to \infty$ theoretical densities (see Figure 2.1.2(b)), while the square waves are finite n histograms.

Conjecture 8.1.3. *The smallest singular value of $\mathcal{J}_a + \frac{2}{\sqrt{\beta}}W$ with type (i) boundary conditions follows the universal smallest singular value distribution with parameters β, a. With type (ii) boundary conditions, the smallest singular value follows the universal smallest singular value distribution with parameters β, $a + 1$.*

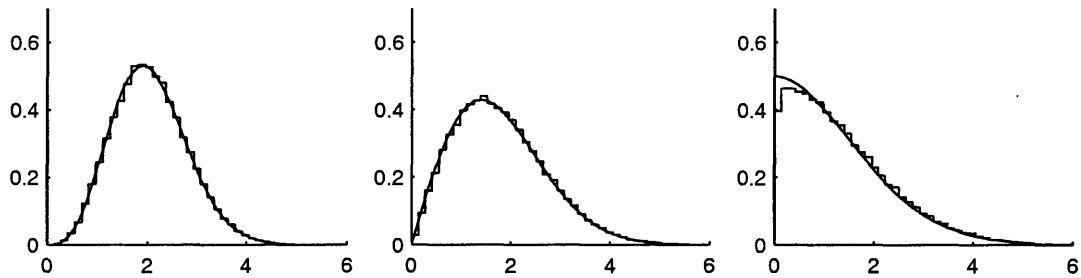
Results from numerical experiments are displayed in Figure 8.1.3. More details can be found in Section 8.5.

Conjecture 8.1.4. *With W_{11}, W_{12}, W_{22} denoting independent diagonal noise operators, the eigenvalue gap about zero of*

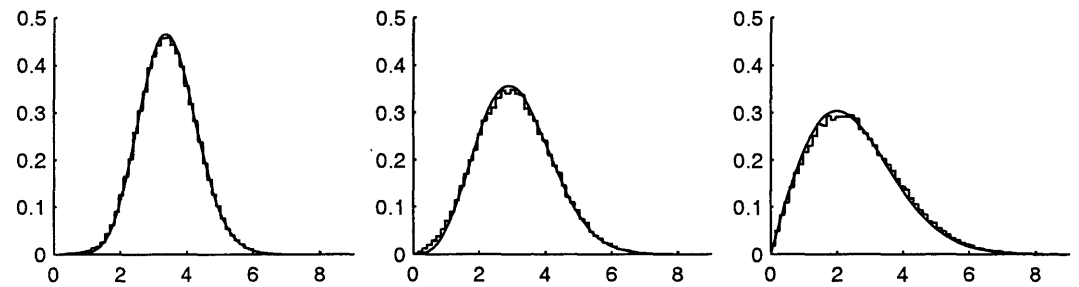
$$\begin{bmatrix} & \mathcal{J}_{-1/2} \\ \mathcal{J}^*_{-1/2} & \end{bmatrix} + \frac{1}{\sqrt{\beta}} \begin{bmatrix} 2W_{11} & \sqrt{2}W_{12} \\ \sqrt{2}W_{12} & 2W_{22} \end{bmatrix} \tag{8.1.1}$$

follows the universal spacing distribution with parameter $\beta > 0$. We conjecture that this holds whether type (i) or type (ii) boundary conditions are imposed on the sine operators.

Results from numerical experiments are displayed in Figure 8.1.4. More details can be found in Section 8.6.

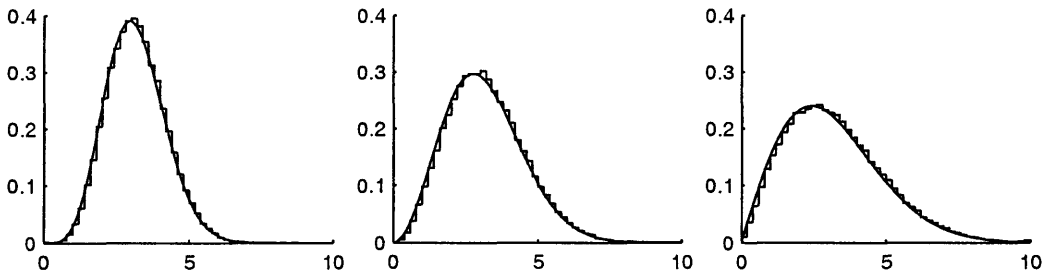


(a) $a = 0$, $\beta = 4$, type (i) b.c. (b) $a = 0$, $\beta = 2$, type (i) b.c. (c) $a = 0$, $\beta = 1$, type (i) b.c.

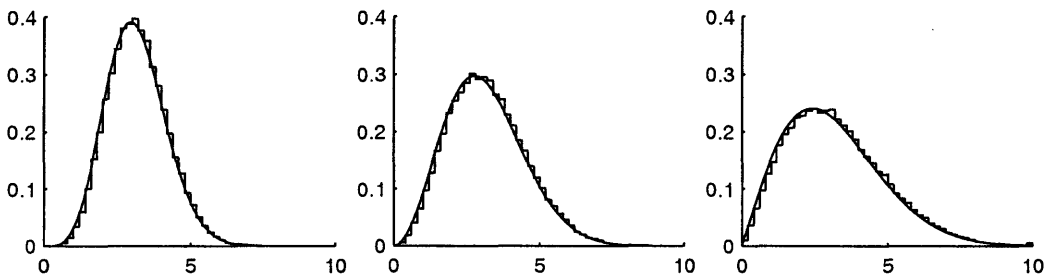


(d) $a = 0$, $\beta = 4$, type (ii) b.c. (e) $a = 0$, $\beta = 2$, type (ii) b.c. (f) $a = 0$, $\beta = 1$, type (ii) b.c.

Figure 8.1.3: Smallest singular value of a finite difference approximation to $\mathcal{J}_a + \frac{2}{\sqrt{\beta}}W$. The smooth curves are $n \to \infty$ theoretical densities (see Figure 2.1.2(b)), while the square waves are finite n histograms. The problem area near zero in figure (c) appears to disappear as $n \to \infty$.



(a) $\beta = 4$, type (i) b.c. (b) $\beta = 2$, type (i) b.c. (c) $\beta = 1$, type (i) b.c.



(d) $\beta = 4$, type (ii) b.c. (e) $\beta = 2$, type (ii) b.c. (f) $\beta = 1$, type (ii) b.c.

Figure 8.1.4: Bulk spacings for finite difference approximations to (8.1.1). The smooth curves are $n \to \infty$ theoretical densities (see Figure 2.1.2(c)), while the square waves are finite n histograms.

Notably absent in these conjectures is the sine operator in Liouville form, $\tilde{\mathcal{J}}_{-1/2}$, which played a vital role in understanding the Jacobi ensemble scaled in the center of its spectrum in the previous chapter. Work concerning this operator is presented in Subsection 8.6.1, where a more measured approach is taken because of apparent sensitivity to the interpretation of white noise.

The above conjectures are motivated by scaling limits of random matrix models, suggested by the following table.

Matrix model	Scaling limit	Scaling limit type	Stochastic differential operator
Laguerre (square or rectangular)	right edge	soft edge	$\mathcal{A}$ plus noise
Hermite	right edge	soft edge	$\mathcal{A}$ plus noise
Jacobi	left edge	hard edge	$\tilde{\mathcal{J}}_a$ plus noise, type (i) b.c.
Jacobi	right edge	hard edge	$\tilde{\mathcal{J}}_b$ plus noise, type (i) b.c.
Laguerre (square)	left edge	hard edge	$\mathcal{J}_a$ plus noise, type (i) b.c.
Laguerre (rectangular)	left edge	hard edge	$\mathcal{J}_{a-1}$ plus noise, type (ii) b.c.
Hermite (n even)	center	bulk	$\begin{bmatrix} & \mathcal{J}_{-1/2} \\ \mathcal{J}^*_{-1/2} & \end{bmatrix}$ plus noise, type (i) b.c.
Hermite (n odd)	center	bulk	$\begin{bmatrix} & \mathcal{J}^*_{-1/2} \\ \mathcal{J}_{-1/2} & \end{bmatrix}$ plus noise, type (ii) b.c.

8.2 Gaussian approximations

The approximations presented in this section are based on the observation that, for large n, the Jacobi, Laguerre, and Hermite matrix models contain many independent random variables, and these random variables do not have heavy tails. These are

the sorts of hypotheses necessary for central limit theorems. We conjecture that for our purposes, the only relevant features of the entry distributions are (1) mean, (2) variance, and (3) no heavy tails. We can replace an entry distribution with a different distribution without affecting the eigenvalue statistics of interest as $n \to \infty$, as long as we preserve these three features. In practice, we use Gaussian approximations, for ease of notation and sampling.

Although we doubt that there is anything special about Gaussians, we note that Gaussians can give especially good approximations to the matrix models, based on the observation that $\sqrt{2}(\chi_r - \sqrt{r})$ approaches a standard Gaussian as $r \to \infty$ [11] and based on Remark 3.5.1, concerning asymptotics of the beta distribution. (Note that, for example, nearly all χ_r entries in the matrix models have $r \to \infty$ as $n \to \infty$.)

The approximate Jacobi matrix model is obtained by first replacing angles $\theta_1, \ldots, \theta_n$, $\phi_1, \ldots, \phi_{n-1}$ by Gaussians (Figure 8.2.1), replacing certain instances of $\bar{c}_i$, $\bar{s}_i$, $\bar{c}'_i$, and $\bar{s}'_i$ by $\frac{1}{\sqrt{2}}$, multiplying out the matrix entries, and throwing away "second-order terms." For example, a diagonal entry in the top-left block of the preliminary approximation is of the form $\tilde{c}_i\tilde{s}'_i$, which equals $\bar{c}_i\bar{s}'_i - \bar{s}'_i\frac{1}{2\sqrt{\beta}}\frac{\sqrt{2}\bar{s}_i}{\sqrt{a+b+2i}}G_i + \bar{c}_i\frac{1}{2\sqrt{\beta}}\frac{\sqrt{2}\bar{c}'_i}{\sqrt{a+b+1+2i}}G'_i - \frac{1}{4\beta}\frac{2\bar{s}_i\bar{c}'_i}{\sqrt{a+b+2i}\sqrt{a+b+1+2i}}G_iG'_i$. For moderately large i, this expression is well approximated by $\bar{c}_i\bar{s}'_i - \frac{1}{2\sqrt{2\beta}}\frac{1}{\sqrt{a+b+2i}}G_i + \frac{1}{2\sqrt{2\beta}}\frac{1}{\sqrt{a+b+1+2i}}G'_i$. Using this approximation for all i, based on the assumption that a not-so-good approximation for a finite number of entries does not significantly change the CS values of interest as $n \to \infty$, we have the final Gaussian approximation to the Jacobi model, displayed in Figure 8.2.2. Gaussian approximations to the Laguerre and Hermite models are also displayed in this figure. They are obtained by replacing chi-distributed random variables with Gaussian approximations.

Whether or not the Gaussian approximations are appropriate to use depends on the eigenvalue statistic of interest.

JACOBI MATRIX MODEL, PRELIMINARY APPROXIMATION

$$
\left[\begin{array}{ccccc|ccccc}
\tilde{c}_n & -\tilde{s}_n\tilde{c}'_{n-1} & & & & \tilde{s}_n\tilde{s}'_{n-1} & & & & \\
 & \tilde{c}_{n-1}\tilde{s}'_{n-1} & -\tilde{s}_{n-1}\tilde{c}'_{n-2} & & & \tilde{c}_{n-1}\tilde{c}'_{n-1} & \tilde{s}_{n-1}\tilde{s}'_{n-2} & & & \\
 & & \tilde{c}_{n-2}\tilde{s}'_{n-2} & \ddots & & & \tilde{c}_{n-2}\tilde{c}'_{n-2} & \tilde{s}_{n-2}\tilde{s}'_{n-3} & & \\
 & & & \ddots & -\tilde{s}_2\tilde{c}'_1 & & & \ddots & \ddots & \\
 & & & & \tilde{c}_1\tilde{s}'_1 & & & & \tilde{c}_1\tilde{c}'_1 & \tilde{s}_1 \\
\hline
-\tilde{s}_n & -\tilde{c}_n\tilde{c}'_{n-1} & & & & \tilde{c}_n\tilde{s}'_{n-1} & & & & \\
 & -\tilde{s}_{n-1}\tilde{s}'_{n-1} & -\tilde{c}_{n-1}\tilde{c}'_{n-2} & & & -\tilde{s}_{n-1}\tilde{c}'_{n-1} & \tilde{c}_{n-1}\tilde{s}'_{n-2} & & & \\
 & & -\tilde{s}_{n-2}\tilde{s}'_{n-2} & \ddots & & & -\tilde{s}_{n-2}\tilde{c}'_{n-2} & \tilde{c}_{n-2}\tilde{s}'_{n-3} & & \\
 & & & \ddots & -\tilde{c}_2\tilde{c}'_1 & & & \ddots & \ddots & \\
 & & & & -\tilde{s}_1\tilde{s}'_1 & & & & -\tilde{s}_1\tilde{c}'_1 & \tilde{c}_1
\end{array}\right]
$$

$$\beta > 0,\ a, b > -1$$

$$\tilde{c}_i = \bar{c}_i - \frac{1}{2\sqrt{\beta}}\frac{\sqrt{2}\bar{s}_i}{\sqrt{a+b+2i}}G_i \qquad \tilde{c}'_i = \bar{c}'_i - \frac{1}{2\sqrt{\beta}}\frac{\sqrt{2}\bar{s}'_i}{\sqrt{a+b+1+2i}}G'_i$$

$$\tilde{s}_i = \bar{s}_i + \frac{1}{2\sqrt{\beta}}\frac{\sqrt{2}\bar{c}_i}{\sqrt{a+b+2i}}G_i \qquad \tilde{s}'_i = \bar{s}'_i + \frac{1}{2\sqrt{\beta}}\frac{\sqrt{2}\bar{c}'_i}{\sqrt{a+b+1+2i}}G'_i$$

Figure 8.2.1: A preliminary Gaussian approximation to the Jacobi matrix model. $\bar{c}_i$, $\bar{s}_i$, $\bar{c}'_i$, and $\bar{s}'_i$ are defined in Figure 6.0.1.

JACOBI GAUSSIAN APPROXIMATION

$$\tilde{J}^{\beta}_{a,b} \sim J^{\infty}_{a,b} + \frac{1}{2\sqrt{2\beta}} \begin{bmatrix} X & \\ & X \end{bmatrix} \begin{bmatrix} -(-\Omega D_1 \Omega) & -D_1^T \\ D_1 & -(-\Omega D_1 \Omega)^T \end{bmatrix} + \frac{1}{2\sqrt{2\beta}} \begin{bmatrix} -\Omega D_1 \Omega & -D_1^T \\ D_1 & (-\Omega D_1 \Omega)^T \end{bmatrix} \begin{bmatrix} Y_1 & \\ & Y_2 \end{bmatrix}$$

$$X = \mathrm{diag}(a+b+2(n:-1:1))^{-1/2}\,\mathrm{diag}(G_n,\ldots,G_1)$$
$$Y_1 = \mathrm{diag}(1, a+b+1+2(n-1:-1:1))^{-1/2}\,\mathrm{diag}(0, G'_{n-1},\ldots,G'_1)$$
$$Y_2 = \mathrm{diag}(a+b+1+2(n-1:-1:1), 1)^{-1/2}\,\mathrm{diag}(G'_{n-1},\ldots,G'_1, 0)$$

LAGUERRE GAUSSIAN APPROXIMATION (square)

$$\tilde{L}^{\beta}_{a} \sim L^{\infty}_{a} + \frac{1}{\sqrt{2\beta}} \begin{bmatrix} G_n & & & & \\ G'_{n-1} & G_{n-1} & & & \\ & G'_{n-2} & G_{n-2} & & \\ & & \ddots & \ddots & \\ & & & G'_1 & G_1 \end{bmatrix}$$

LAGUERRE GAUSSIAN APPROXIMATION (rectangular)

$$\tilde{M}^{\beta}_{a} \sim M^{\infty}_{a} + \frac{1}{\sqrt{2\beta}} \begin{bmatrix} G_n & G'_n & & & \\ & G_{n-1} & G'_{n-1} & & \\ & & \ddots & \ddots & \\ & & & G_2 & G'_2 \\ & & & & G_1 & G'_1 \end{bmatrix}$$

HERMITE GAUSSIAN APPROXIMATION

$$\tilde{H}^{\beta} \sim H^{\infty} + \frac{1}{2\sqrt{\beta}} \begin{bmatrix} 2G_n & G'_{n-1} & & & \\ G'_{n-1} & 2G_{n-1} & G'_{n-2} & & \\ & \ddots & \ddots & \ddots & \\ & & G'_2 & 2G_2 & G'_1 \\ & & & G'_1 & 2G_1 \end{bmatrix}$$

Figure 8.2.2: Gaussian approximations to the matrix models. In each case, $G_1,\ldots,G_n,G'_1,\ldots,G'_{n-1}$ (or, in the rectangular Laguerre case, $G_1,\ldots,G_n$, $G'_1,\ldots,G'_n$) are i.i.d. standard Gaussian random variables. Also, $a+b+2(n:-1:1)$ is Matlab notation for the vector $(a+b+2n, a+b+2(n-1),\ldots,a+b+2)$.

8.3 White noise operator

Given a uniform grid on $[a, b]$ with mesh size h, i.e., $x_k = a + hk$, $k = 1, \ldots, n$, where $n = \lfloor \frac{b-a}{h} \rfloor$, our intuitive idea of a diagonal white noise operator is $\frac{1}{\sqrt{h}} \operatorname{diag}(G_1, \ldots, G_n)$, in which $G_1, \ldots, G_n$ are i.i.d. real standard Gaussians. The diagonal of this matrix approximates white noise in the sense that its values at distinct locations are uncorrelated, and the variance per unit interval of the real line is 1. To illustrate this last statement, suppose that $a = 0$ and $b = 1$, so that $\int_a^b W_t dt$ equals Brownian motion at time 1, i.e., a standard Gaussian:

$$\begin{aligned} G &\stackrel{d}{=} \int_0^1 W_t dt \\ &\approx \sum_{k=1}^{\lfloor \frac{1}{h} \rfloor} \left(\frac{1}{\sqrt{h}} G_k \right) \cdot h \\ &\stackrel{d}{=} \sqrt{h}\, N\left(0, \left\lfloor \frac{1}{h} \right\rfloor\right) \\ &\to N(0,1) \quad (h \to 0). \end{aligned}$$

A diagonal matrix with i.i.d. Gaussian entries is not the only reasonable approximation to a white noise operator. In fact, many textbooks construct Brownian motion from a sum of uniform random variables on $\{-1, 1\}$ instead of Gaussians. As long as the construction has 0 mean, unit variance per interval, independent increments, and no heavy tails, everything is fine. The remainder of this chapter identifies various bidiagonal and tridiagonal random matrices as discretizations of the white noise operator.

8.4 Soft edge

8.4.1 Laguerre at the right edge

We would like to extend Claims 7.2.3 and 7.2.6 to $\beta < \infty$.

Claim 8.4.1. *The Laguerre matrix models, scaled at the right edge, are finite difference approximations to the Airy operator plus white noise. Specifically, if* $h = 2^{2/3}n^{-1/3}$, *then* $\frac{h}{4}(\Omega(L_a^\beta)^T L_a^\beta \Omega - 4(n + \frac{a+1}{2})I_n)$, $\frac{h}{4}(\Omega L_a^\beta (L_a^\beta)^T \Omega - 4(n + \frac{a+1}{2})I_n)$, $\frac{h}{4}(\Omega(M_a^\beta)^T M_a^\beta \Omega - 4(n + \frac{a+1}{2})I_{n+1})$, *and* $\frac{h}{4}(\Omega M_a^\infty (M_a^\infty)^T \Omega - 4(n + \frac{a+1}{2})I_n)$ *are finite difference approximations to* $\mathcal{A} + \frac{2}{\sqrt{\beta}}W$.

W should represent a diagonal operator that injects white noise, but the correct interpretation of white noise is not clear. Our argument gives some guidance for future work.

The support for the claim is heuristic and numerical.

For $\frac{h}{4}(\Omega(L_a^\beta)^T L_a^\beta \Omega - 4(n + \frac{a+1}{2})I_n)$, start by replacing L_a^β with its Gaussian approximation $\tilde{L}_a^\beta$. The resulting approximation can be expanded to $\frac{h}{4}(\Omega(L_a^\infty)^T L_a^\infty \Omega - 4(n + \frac{a+1}{2})I_n) + \frac{h}{4\sqrt{2\beta}}\Omega(L_a^\infty)^T B\Omega + \frac{h}{4\sqrt{2\beta}}\Omega B^T L_a^\infty \Omega + \frac{h}{8\beta}\Omega B^T B\Omega$, in which B is a lower bidiagonal matrix with i.i.d. standard Gaussian entries. The first of these terms is a finite difference approximation to the Airy operator, by a claim of the previous chapter. The entries of the fourth term are negligible in magnitude compared to the entries of the other terms, except in the bottom-right portion of the matrix. Fortunately, this bottom-right portion is insignificant, because the dominant eigenvectors of the nonrandom term have almost all of their mass in their first $O(n^{1/3})$ entries (Theorem 6.3.7). After dropping the fourth term, the remaining random bit is $\frac{h}{4\sqrt{2\beta}}\Omega((L_a^\infty)^T B + B^T L_a^\infty)\Omega$. This is a symmetric tridiagonal matrix with Gaussian entries, which we claim discretizes a diagonal white noise operator. It remains to carefully consider the covariance matrix of the entries. Each of the entries is a sum of two Gaussians, each of whose standard deviation is determined by an entry of L_a^∞. Because only the first $O(n^{1/3})$ rows of the matrix are really significant, the entries in row k of the noise matrix will be approximated by $\frac{h\sqrt{n}}{4\sqrt{2\beta}}(G_{n+1-k} - G'_{n-(k-1)})$, $\frac{h\sqrt{n}}{4\sqrt{2\beta}}(2G_{n+1-k} - 2G'_{n-k})$, and $\frac{h\sqrt{n}}{4\sqrt{2\beta}}(G_{n+1-(k+1)} - G'_{n-k})$. Taking into account correlations, the "average standard deviation per row" is $\frac{2}{\sqrt{\beta}}\frac{1}{\sqrt{h}}$, which provides the constant in the white noise term $\frac{2}{\sqrt{\beta}}W$. We should also note that the noise at grid point x is

independent of the noise at grid point $y \neq x$, i.e., rows $\lfloor \frac{x}{h} \rfloor$ and $\lfloor \frac{y}{h} \rfloor$ are independent for sufficiently large n, so that the noise is "white."

The arguments for the three other rescaled matrix models mentioned in the claim are similar.

8.4.2 Hermite at the right edge

Claim 8.4.2. *The Hermite matrix model, scaled at the right edge, is a finite difference approximation to the Airy operator plus white noise. Specifically, if* $h = n^{-1/3}$, *then* $\sqrt{\frac{2}{h}}(H^\beta - \sqrt{2}h^{-3/2}I_n)$ *is a finite difference approximation to* $\mathcal{A} + \frac{2}{\sqrt{\beta}}W$.

To justify this claim, replace H^β by its Gaussian approximation $\tilde{H}^\beta$. This gives a nonrandom term $\sqrt{2}n^{1/6}(H^\infty - \sqrt{2n}I)$, which we previously argued is a finite difference approximation to $\mathcal{A}$, and a random term which is a symmetric tridiagonal matrix with Gaussian entries. Taking into account correlations between the entries, the "average standard deviation per row" is $\frac{2}{\sqrt{\beta}}\frac{1}{\sqrt{h}}$, so that the $h \to 0$ limit should give $\frac{2}{\sqrt{\beta}}W$.

8.4.3 Numerical experiments

The claims presented earlier in this section give rise to Conjecture 8.1.1. The best way we know to investigate this conjecture is numerically, through a finite difference approximation to the stochastic Airy operator $\mathcal{A} + \frac{2}{\sqrt{\beta}}W$. Of course, the point of the conjecture is to jump from finite matrices to continuous operators, so jumping back to finite matrices to test the conjecture is not ideal. Nevertheless, the technique does have its merits. We have already seen two random matrices that can reasonably be called discretizations of the stochastic Airy operator—the Laguerre and Hermite matrix models. If a number of different discretizations to the stochastic operator all show similar eigenvalue behavior, then support for the conjecture is strengthened.

Figure 8.1.1 was created by approximating $\mathcal{A}$ by $\frac{1}{h^2}D_2 - \mathrm{diag}(x_1, \ldots, x_n)$ and $\frac{2}{\sqrt{\beta}}W$

by $\frac{2}{\sqrt{\beta}}\frac{1}{\sqrt{h}}\,\mathrm{diag}(G_1,\ldots,G_n)$, with $n = 200$, $h = n^{-1/3}$, $x_k = hk$, and $G_1,\ldots,G_n$ i.i.d. standard Gaussians. Over 10^5 trials, the largest eigenvalue of this random matrix was computed, for the three cases $\beta = 1, 2, 4$.

8.5 Hard edge

8.5.1 Jacobi at the left edge

Claim 8.5.1. *The Jacobi matrix model, scaled at the hard edge on the left, is a finite difference approximation to the Bessel operator in Liouville form plus white noise. Specifically, if* $h = \frac{1}{n+(a+b+1)/2}$ *and* B_{22} *denotes the bottom-right* n*-by-*n *block of the* $2n$*-by-*$2n$ *Jacobi matrix model* $J^\beta_{a,b}$*, then* $\frac{2}{h}FB_{22}F$ *is a finite difference approximation to* $\tilde{\mathcal{J}}_a + \sqrt{\frac{2}{\beta}}\frac{1}{\sqrt{y}}W$.

As in the soft edge case, the correct interpretation of the diagonal noise operator W is not clear.

To justify the claim, start by replacing $J^\beta_{a,b}$ by its Gaussian approximation $\tilde{J}^\beta_{a,b}$. The bottom-right block can be expressed as the sum of a nonrandom term and a random term. The nonrandom term comes from the $\beta = \infty$ matrix model, and we have already argued that this matrix is a finite difference approximation to the Bessel operator. The random term is $\frac{2}{h}(\frac{1}{2\sqrt{2\beta}}FXF(-(-\Omega D_1\Omega)) + \frac{1}{2\sqrt{2\beta}}(-\Omega D_1\Omega)FY_2F)$. This matrix is upper bidiagonal, and its (k,k) entry is $\frac{1}{h}\frac{1}{\sqrt{2\beta}}(-(a+b+2k)^{-1/2}G_k + (a+b+1+2(k-1))^{-1/2}G'_{k-1}) \approx \frac{1}{h}\frac{1}{2\sqrt{\beta}}\frac{1}{\sqrt{k}}(-G_k + G'_{k-1})$, for $k > 1$, while its $(k,k+1)$ entry is $\frac{1}{h}\frac{1}{\sqrt{2\beta}}(-(a+b+2k)^{-1/2}G_k + (a+b+1+2k)^{-1/2}G'_k) \approx \frac{1}{h}\frac{1}{2\sqrt{\beta}}\frac{1}{\sqrt{k}}(-G_k + G'_k)$, for $k < n$. Taking into account correlations, the standard deviation per row is approximately $\frac{1}{h}\frac{\sqrt{2}}{\sqrt{\beta}}\frac{1}{\sqrt{k}} = \frac{\sqrt{2}}{\sqrt{\beta}}\frac{\sqrt{h}}{\sqrt{y_k}}\frac{1}{h} = \sqrt{\frac{2}{\beta}}\frac{1}{\sqrt{y_k}}\frac{1}{\sqrt{h}}$ in the vicinity of row k, in which $y_k = hk$.

There is an obvious analogue for the top-left block of the Jacobi matrix model, in which $\tilde{\mathcal{J}}_a$ should be replaced by its adjoint.

8.5.2 Jacobi at the right edge

Claim 8.5.2. *The Jacobi matrix model, scaled at the hard edge on the right, is a finite difference approximation to a Bessel operator in Liouville form plus white noise. Specifically, if* $h = \frac{1}{n+(a+b+1)/2}$ *and* B_{12} *denotes the top-right n-by-n block of the Jacobi matrix model* $J_{a,b}^{\beta}$*, then* $\frac{2}{h}\Omega F B_{12} F \Omega$ *is a finite difference approximation to* $\tilde{\mathcal{J}}_b + \sqrt{\frac{2}{\beta}}\frac{1}{\sqrt{y}}W$.

This claim is equivalent to the claim of the previous subsection, considering (4.2.1).

8.5.3 Laguerre at the left edge

Claim 8.5.3. *The square Laguerre matrix model, scaled at the hard edge, is a finite difference approximation to a Bessel operator plus white noise. Specifically, if* $h = \frac{1}{n+(a+1)/2}$*, then* $\frac{2}{\sqrt{h}} F L_a^{\beta} F$ *is a finite difference approximation to* $\mathcal{J}_a + \frac{2}{\sqrt{\beta}}W$*, in which the Bessel operator has type (i) boundary conditions.*

To justify the claim, start by replacing L_a^{β} with its Gaussian approximation $\tilde{L}_a^{\beta}$. Then $\frac{2}{\sqrt{h}} F \tilde{L}_a^{\beta} F$ can be expressed as the sum of a nonrandom term $\frac{2}{\sqrt{h}} F L_a^{\infty} F$ and a random term whose (k,k) entry is $\sqrt{\frac{2}{\beta}}\frac{1}{\sqrt{h}}G_k$ and whose $(k,k+1)$ entry is $\sqrt{\frac{2}{\beta}}\frac{1}{\sqrt{h}}G'_k$, for all possible k. We have already argued that the nonrandom term is a finite difference approximation to $\mathcal{J}_a$ with type (i) boundary conditions, and the total standard deviation over row k of the random term is $\frac{2}{\sqrt{\beta}}\frac{1}{\sqrt{h}}$, for $k < n$.

We make an analogous claim for the rectangular Laguerre model.

Claim 8.5.4. *The rectangular Laguerre matrix model, scaled at the hard edge, is a finite difference approximation to a Bessel operator plus white noise. Specifically, if* $h = \frac{1}{n+(a+1)/2}$*, then* $\frac{2}{\sqrt{h}} F M_a^{\beta} F$ *is a finite difference approximation to* $\mathcal{J}_{a-1} + \frac{2}{\sqrt{\beta}}W$*, in which the Bessel operator has type (ii) boundary conditions.*

8.5.4 Numerical experiments

The claims presented earlier in this section give rise to Conjectures 8.1.2 and 8.1.3.

Results from numerical experiments regarding $\tilde{\mathcal{J}}_a + \sqrt{\frac{2}{\beta}}\frac{1}{\sqrt{y}}W$ are displayed in Figure 8.1.2. The Bessel operator was approximated by $-\frac{1}{h}D_1 + (a+\frac{1}{2})\operatorname{diag}(\frac{1}{y_1},\ldots,\frac{1}{y_n})$, with $n = 2000$, $h = \frac{1}{n}$, and $y_k = hk$. The random term was approximated by $\sqrt{\frac{2}{\beta}}\operatorname{diag}(\frac{1}{\sqrt{y_1}},\ldots,\frac{1}{\sqrt{y_n}}) \cdot \left(\frac{1}{\sqrt{h}}\operatorname{diag}(G_1,\ldots,G_n)\right) \cdot \left(\frac{1}{2}(-\Omega D_1 \Omega)\right)$, in which $G_1,\ldots,G_n$ are i.i.d. standard Gaussians. (The averaging matrix $\frac{1}{2}(-\Omega D_1 \Omega)$ splits the noise over two diagonals instead of one.) Over 10^5 trials, the smallest singular value was computed in the cases $\beta = 1, 2, 4$.

Results from numerical experiments regarding $\mathcal{J}_a + \frac{2}{\sqrt{\beta}}W$ are displayed in Figure 8.1.3. First, we consider type (i) boundary conditions. In this case, the Bessel operator $\mathcal{J}_a$ was approximated by $-2\operatorname{diag}(\sqrt{x_1},\ldots,\sqrt{x_n})(\frac{1}{h}D_1) + a\operatorname{diag}(\frac{1}{\sqrt{x_1}},\ldots,\frac{1}{\sqrt{x_n}})$, with $n = 2000$, $h = \frac{1}{n}$, and $x_k = hk$. The random term was approximated by $\frac{2}{\sqrt{\beta}}\left(\frac{1}{2}(-\Omega D_1 \Omega)\right)\left(\frac{1}{\sqrt{h}}\operatorname{diag}(G_1,\ldots,G_n)\right)$, in which $G_1,\ldots,G_n$ are independent standard Gaussians. Over 10^5 trials, the smallest singular value was computed in the cases $\beta = 1, 2, 4$. In Figures 8.1.3(a)–(c), the experimental agreement with the universal smallest singular value distributions is quite good for $\beta = 2, 4$, but not terribly good for $\beta = 1$. Based on our experience with several values of n, it appears that the problem area near zero shrinks as n grows larger, and likely disappears as $n \to \infty$. Next, we consider type (ii) boundary conditions. In this case, the Bessel operator $\mathcal{J}_a$ was approximated by $-2\operatorname{diag}(\sqrt{x_1},\ldots,\sqrt{x_n})(\frac{1}{h}(D_1)_{1:n,:}) + a\operatorname{diag}(\frac{1}{\sqrt{x_1}},\ldots,\frac{1}{\sqrt{x_n}},0)_{1:n,:}$, with $n = 2000$, $h = \frac{1}{n}$, and $x_k = hk$. The random term was approximated by $\frac{2}{\sqrt{\beta}}$ times an n-by-$(n+1)$ matrix with i.i.d. Gaussian entries on the main diagonal and superdiagonal, each with mean 0 and variance $\frac{1}{2}$. (This approximation to the white noise operator, using two independent diagonals of noise, is different from the other approximations, which multiply a diagonal noise matrix by an averaging operator. An $n \to \infty$, we believe that either scheme works, but for the Bessel operator in Liouville form with type (ii) boundary conditions, the convergence is significantly faster with

two independent diagonals of noise.) Over 10^5 trials, the smallest singular value was computed in the cases $\beta = 1, 2, 4$. The resulting histograms are plotted in Figures 8.1.3(d)–(f).

Remark 8.5.5. Numerical experiments tend to produce histograms that better match the universal distributions when the noise is centered.

For example, in the Airy case, with a centered second difference scheme involving D_2, either a diagonal or tridiagonal noise matrix produces good results. Each places the noise in any row symmetrically about the corresponding mesh point (in terms of variance).

The Bessel cases are more interesting. The experiments involving both $\tilde{\mathcal{J}}_a$ and $\mathcal{J}_a$ use forward differences. Therefore, there are two meshes—one mesh for the domain, and a different mesh for the codomain—and the meshes interlace each other. Experimentally, putting all of the noise on either the main diagonal or superdiagonal produces histograms that do not fit the universal distributions. However, splitting the noise evenly across the two diagonals often produces good fits, e.g., Figures 8.1.2 and 8.1.3. Splitting the noise can be accomplished in at least two ways: (1) by using two independent diagonals of noise, adjusting the variance appropriately, as in the experiment that produced Figures 8.1.3(d)–(f), or (2) by multiplying a diagonal noise matrix by an averaging operator, e.g., $\frac{1}{2}(-\Omega D_1 \Omega)$, as in the experiment that produced Figures 8.1.2 and 8.1.3(a)–(c).

8.6 Bulk

8.6.1 Jacobi near one-half

Suppose that $a = b$ and that n is even, and let B_{22} denote the bottom-right block of the $2n$-by-$2n$ Jacobi matrix model $J^{\beta}_{a,b}$. Then, with $h = \frac{2}{n+(a+b+1)/2}$, $\frac{4}{h}((\Omega \oplus$

$\Omega)PFB_{22}^T B_{22} F P^T(\Omega \oplus \Omega) - \frac{1}{2}I)$ is well approximated by

$$\begin{bmatrix} & \frac{1}{h}D_1^T \\ \frac{1}{h}D_1 & \end{bmatrix} + \frac{1}{\sqrt{\beta}} \begin{bmatrix} E_{11} & E_{21}^T \\ E_{21} & E_{22} \end{bmatrix}, \tag{8.6.1}$$

in which

$$\begin{aligned} E_{11} &= \mathrm{diag}(\tfrac{1}{\sqrt{x_1}}, \ldots, \tfrac{1}{\sqrt{x_{n/2}}})(-\tfrac{1}{\sqrt{2}}\hat{W}_1 + \tfrac{1}{\sqrt{2}}(D_1 + I)^T \hat{W}_2 (D_1 + I)), \\ E_{22} &= \mathrm{diag}(\tfrac{1}{\sqrt{x_1}}, \ldots, \tfrac{1}{\sqrt{x_{n/2}}})(\tfrac{1}{\sqrt{2}}\hat{W}_1 - \tfrac{1}{\sqrt{2}}\hat{W}_2), \\ E_{21} &= \mathrm{diag}(\tfrac{1}{\sqrt{x_1}}, \ldots, \tfrac{1}{\sqrt{x_{n/2}}})(\tfrac{1}{\sqrt{2}}\hat{W}_3 \tfrac{1}{2}(-\Omega D_1 \Omega) - \tfrac{1}{\sqrt{2}}\tfrac{1}{2}(-\Omega D_1 \Omega)\hat{W}_4), \end{aligned}$$

with $\hat{W}_1 = \frac{1}{\sqrt{h}}\,\mathrm{diag}(G_1, G_3, \ldots, G_{n-1})$, $\hat{W}_2 = \frac{1}{\sqrt{h}}\,\mathrm{diag}(G_2, G_4, \ldots, G_n)$, $\hat{W}_3 = \frac{1}{\sqrt{h}}\,\mathrm{diag}(G'_1, G'_3, \ldots, G'_{n-1})$, and $\hat{W}_4 = \frac{1}{\sqrt{h}}\,\mathrm{diag}(0, G'_2, G'_4, \ldots, G'_{n-2})$.

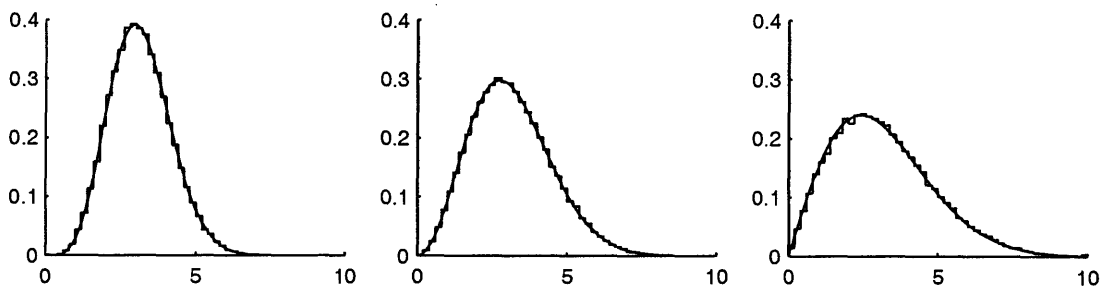
This random matrix appears to be a finite difference approximation to

$$\begin{bmatrix} & -\tilde{\mathcal{J}}_{-1/2}^* \\ -\tilde{\mathcal{J}}_{-1/2} & \end{bmatrix} + \frac{1}{\sqrt{\beta}} \left[\begin{array}{c|c} \frac{1}{\sqrt{y}}W_1 & \frac{1}{\sqrt{y}}W_2 \\ \hline \frac{1}{\sqrt{y}}W_2 & -\frac{1}{\sqrt{y}}W_1 \end{array}\right] \tag{8.6.2}$$

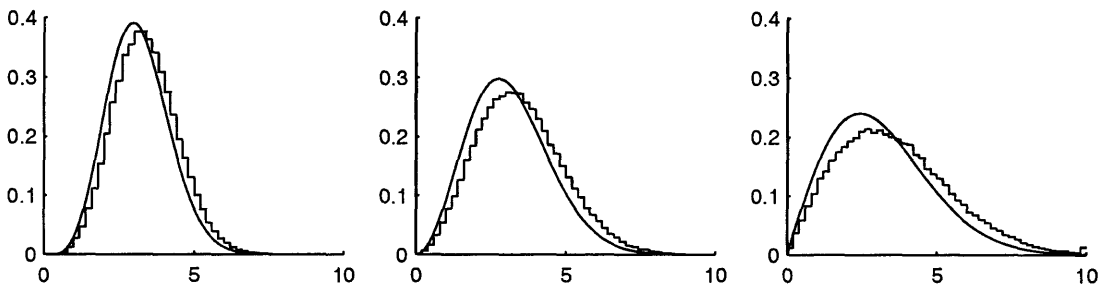
with type (i) boundary conditions (see Subsection 7.4.1), and with W_1 and W_2 denoting independent diagonal noise operators. Numerical experiments indicate that this stochastic operator is especially sensitive to the interpretation of white noise, however. Figures 8.6.1(a–c) contain histograms of the spacings between eigenvalues m and $m+1$ of the $2m$-by-$2m$ random matrix

$$\begin{bmatrix} & \frac{1}{h}D_1^T \\ \frac{1}{h}D_1 & \end{bmatrix} + \frac{1}{\sqrt{\beta}} \left[\begin{array}{c|c} E_{11} & (\frac{1}{2}(-\Omega D_1 \Omega))^T \hat{W}_5\,\mathrm{diag}(\frac{1}{\sqrt{x_1}}, \ldots, \frac{1}{\sqrt{x_m}}) \\ \hline \mathrm{diag}(\frac{1}{\sqrt{x_1}}, \ldots, \frac{1}{\sqrt{x_m}})\hat{W}_5(\frac{1}{2}(-\Omega D_1 \Omega)) & E_{22} \end{array}\right],$$

in which E_{11} and E_{22} are defined as above, and $\hat{W}_5$ is $\frac{1}{\sqrt{h}}$ times a diagonal matrix



(a) $\beta = 4$, discretization I (b) $\beta = 2$, discretization I (c) $\beta = 1$, discretization I



(d) $\beta = 4$, discretization II (e) $\beta = 2$, discretization II (f) $\beta = 1$, discretization II

Figure 8.6.1: Bulk spacings for two different finite difference approximations to (8.6.2). The differences seem to arise from different interpretations of white noise, somewhat analogously to Itô versus Stratonovich.

with independent standard Gaussian entries. With $m = 200$, the agreement with the universal spacing distributions for $\beta = 1, 2, 4$ is quite good. However, a slightly different experiment gives very different results. Figures 8.6.1(d–f) contain histograms of the spacings between eigenvalues m and $m+1$ of the $2m$-by-$2m$ random matrix

$$\begin{bmatrix} & \frac{1}{h}D_1^T \\ \frac{1}{h}D_1 & \end{bmatrix} + \frac{1}{\sqrt{\beta}} \left[\begin{array}{c|c} \operatorname{diag}(\frac{1}{\sqrt{x_1}},\ldots,\frac{1}{\sqrt{x_m}})(\frac{1}{\sqrt{2}}(-\hat{W}_1+\hat{W}_2)) & \frac{1}{2}(-\Omega D_1\Omega)^T\hat{W}_5 \operatorname{diag}(\frac{1}{\sqrt{x_1}},\ldots,\frac{1}{\sqrt{x_m}}) \\ \hline \operatorname{diag}(\frac{1}{\sqrt{x_1}},\ldots,\frac{1}{\sqrt{x_m}})\hat{W}_5(\frac{1}{2}(-\Omega D_1\Omega)) & \operatorname{diag}(\frac{1}{\sqrt{x_1}},\ldots,\frac{1}{\sqrt{x_m}})(\frac{1}{\sqrt{2}}(\hat{W}_1-\hat{W}_2)) \end{array}\right].$$

Again we take $m = 200$, but this time the histograms do not appear to fit the universal spacing distributions. The only difference between the two experiments is in the top-left block of the noise matrix. In the first experiment, $\hat{W}_2$ is shifted up and to the left by one row and one column; in the second experiment, $\hat{W}_2$ is not shifted. The first experiment conforms more closely to the Jacobi matrix model and its approximation in (8.6.1), while the second experiment appears to be a perfectly reasonable approximation to (8.6.2).

The experiments suggest that the spacing statistic is sensitive to the interpretation of white noise. In the first experiment, the noise is placed "to the left" in the mesh corresponding to the top-left block, while in the second experiment, the noise is placed "to the right." This issue is very reminiscent of Itô versus Stratonovich. Because one has to be particularly careful when discretizing (8.6.2), we stop short of making a conjecture concerning the eigenvalues. Further work on the interpretation of white noise is warranted.

8.6.2 Hermite near zero

Claim 8.6.1. *First, the even case. Let H^β be the $2m$-by-$2m$ Hermite matrix model. We claim that if $h = \frac{1}{m+1/4}$, then $\frac{2}{\sqrt{h}}(\Omega \oplus \Omega)(F \oplus F)PH^\beta P^T(F \oplus F)(\Omega \oplus \Omega)$ is a*

finite difference approximation to

$$\begin{bmatrix} & \mathcal{J}_{-1/2} \\ \mathcal{J}_{-1/2}^* & \end{bmatrix} + \frac{1}{\sqrt{\beta}} \begin{bmatrix} 2W_{11} & \sqrt{2}W_{12} \\ \sqrt{2}W_{12} & 2W_{22} \end{bmatrix},$$

in which the sine operators have type (i) boundary conditions, and W_{11}, W_{12}, W_{22} *denote independent diagonal noise operators.*

Now for the odd case. Suppose that H^β *is the* $(2m+1)$*-by-*$(2m+1)$ *Hermite matrix model. We claim that if* $h = \frac{1}{m+3/4}$, *then* $\frac{2}{\sqrt{h}}(\Omega \oplus \Omega)(F \oplus F)PH^\beta P^T(F \oplus F)(\Omega \oplus \Omega)$ *is a finite difference approximation to*

$$\begin{bmatrix} & \mathcal{J}_{-1/2}^* \\ \mathcal{J}_{-1/2} & \end{bmatrix} + \frac{1}{\sqrt{\beta}} \begin{bmatrix} 2W_{11} & \sqrt{2}W_{12} \\ \sqrt{2}W_{12} & 2W_{22} \end{bmatrix},$$

in which the Bessel operators have type (ii) boundary conditions, and W_{11}, W_{12}, W_{22} *denote independent diagonal noise operators.*

As with the other stochastic eigenvalue/singular value problems, the correct interpretation of the noise operators is not clear.

To justify the claim, first replace H^β with its Gaussian approximation $\tilde{H}^\beta$. When the size of the matrix model is $2m$-by-$2m$, $\frac{2}{\sqrt{h}}(\Omega \oplus \Omega)(F \oplus F)P\tilde{H}^\beta P^T(F \oplus F)(\Omega \oplus \Omega)$ follows the same distribution as

$$\begin{bmatrix} & \frac{2}{\sqrt{h}}F\tilde{L}_{-1/2}^{2\beta}F \\ \frac{2}{\sqrt{h}}F(\tilde{L}_{-1/2}^{2\beta})^T F & \end{bmatrix} + \frac{2}{\sqrt{\beta}}\frac{1}{\sqrt{h}}\operatorname{diag}(G, \ldots, G),$$

in which the Gaussians in $\tilde{L}_{-1/2}^{2\beta}$ are $G_1', \ldots, G_{n-1}'$ rearranged appropriately, and the Gaussians in the diagonal matrix are $G_1, \ldots, G_n$, rearranged appropriately. We argued earlier that $\frac{2}{\sqrt{h}}F\tilde{L}_{-1/2}^{2\beta}F$ is a finite difference approximation to $\mathcal{J}_{-1/2} + \frac{2}{\sqrt{2\beta}}W$.

The argument for odd-sized Hermite matrix models is similar.

8.6.3 Numerical experiments

Claim 8.6.1 gives rise to Conjecture 8.1.4. We would like to make an analogous conjecture for the sine operator in Liouville form, $\tilde{\mathcal{J}}_{-1/2}$, but as discussed in Subsection 8.6.1, numerical experiments do not fully support such a conjecture.

The results of numerical experiments regarding Conjecture 8.1.4 are displayed in Figure 8.1.4. For type (i) boundary conditions, $\mathcal{J}_{-1/2}$ was approximated by

$$-2\,\mathrm{diag}(\sqrt{x_1},\ldots,\sqrt{x_m})(\tfrac{1}{h}D_1) - \tfrac{1}{2}\,\mathrm{diag}(\tfrac{1}{\sqrt{x_1}},\ldots,\tfrac{1}{\sqrt{x_m}}),$$

with $m = 200$, $h = \frac{1}{m}$, and $x_k = hk$. Each of W_{11} and W_{22} was approximated by $\frac{1}{\sqrt{h}}$ times a diagonal of independent standard Gaussians, and W_{12} was approximated by $\frac{1}{\sqrt{h}}\,\mathrm{diag}(G,\ldots,G)(\frac{1}{2})(-\Omega D_1\Omega)$, in which the diagonal matrix contains independent standard Gaussians. Over 10^5 trials, the gap between eigenvalues m and $m+1$ was computed for $\beta = 1, 2, 4$, and the resulting histograms are displayed in Figures 8.1.4(a)–(c). For type (ii) boundary conditions, $\mathcal{J}_{-1/2}$ was approximated by the m-by-$(m+1)$ matrix

$$-2\,\mathrm{diag}(\sqrt{x_1},\ldots,\sqrt{x_m})(\tfrac{1}{h}D_1)_{1:m,:} - \tfrac{1}{2}\,\mathrm{diag}(\tfrac{1}{\sqrt{x_1}},\ldots,\tfrac{1}{\sqrt{x_m}},0)_{1:m,:},$$

with $m = 200$, $h = \frac{1}{m}$, and $x_k = hk$. Each of W_{11} and W_{22} was approximated by $\frac{1}{\sqrt{h}}$ times a diagonal of independent standard Gaussians, and W_{12} was approximated by $\frac{1}{\sqrt{h}}\,\mathrm{diag}(G,\ldots,G)(\frac{1}{2})(-\Omega(D_1)_{1:m,:}\Omega)$, in which the diagonal matrix contains independent standard Gaussians. Over 10^5 trials, the gap between eigenvalues m and $m+1$ was computed for $\beta = 1, 2, 4$, and the resulting histograms are displayed in Figures 8.1.4(d)–(f).

Chapter 9

Application: Large β asymptotics

This chapter presents an asymptotic expression for the mean of the largest eigenvalue distribution, valid near $\beta = \infty$. The expression involves quantities intimately related to the Airy operator, suggesting the intrinsic role of stochastic differential operators in random matrix theory.

9.1 A technical note

Scaling limits of the ensembles are not well understood when $\beta \neq 1, 2, 4$. In Examples 2.1.1–2.1.7, a discussion of the general β case was postponed. Here, we address the difficulties presented when $\beta \neq 1, 2, 4$.

The examples in the introduction describe experiments that can be conducted on the classical ensembles: Sample an ensemble, select a particular particle, rescale the position of this particle in some judicious fashion, and draw a histogram of the rescaled positions over a large number of trials. For each experiment, the histogram approaches one of the plots in Figure 2.1.2 as the number of particles n approaches infinity, when $\beta = 1, 2, 4$. However, when $\beta \neq 1, 2, 4$, it is not known whether the histograms approach $n \to \infty$ limits; it is not known if universality still holds; and there are no known analytic expressions for the universal distributions (if the

distributions even exist). Despite these theoretical difficulties, physical arguments and numerical experiments strongly indicate that universal distributions exist for general β. In fact, this chapter gives some idea of how the universal largest eigenvalue distributions should look for large β. The following conjecture provides our definition of the universal largest eigenvalue distribution for general β.

Conjecture 9.1.1. *Let $\beta > 0$. Let λ_k^H denote the kth rightmost particle of the n particle β-Hermite ensemble, and let λ_k^L denote the kth rightmost particle of the n particle β-Laguerre ensemble with parameter a, for some fixed $a > -1$. We conjecture that as $n \to \infty$, $\sqrt{2}n^{1/6}(\lambda_k^H - \sqrt{2n})$ and $2^{-4/3}n^{-1/3}(\lambda_k^L - 4n)$ converge in distribution to the same distribution, which we shall denote X_k^β.*

The conjecture is known to be true in the cases $\beta = 1, 2$ by [12, 17, 18]. In particular, X_1^1, X_1^2, X_1^4 are the universal largest eigenvalue distributions plotted in Figure 2.1.2(a) using the Painlevé theory of Tracy and Widom. We conjecture that these known results can be extended to general $\beta > 0$.

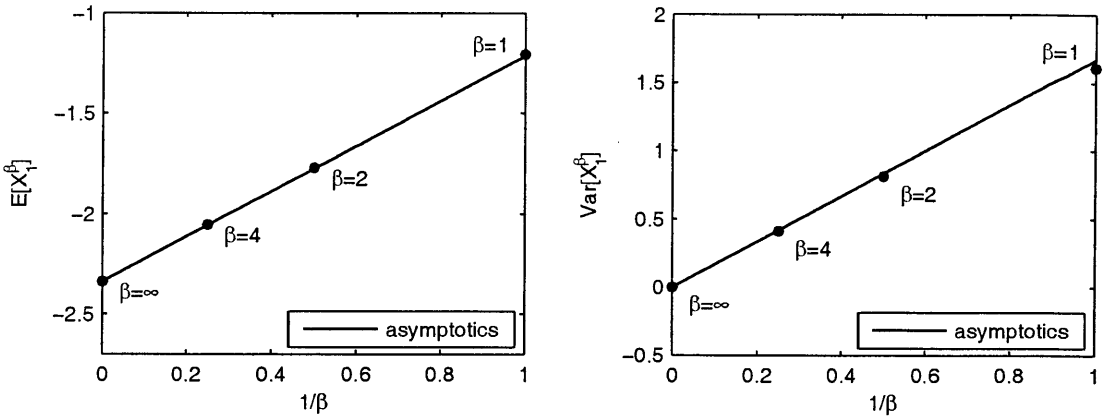
9.2 The asymptotics

In the following conjecture, ζ_k is the kth rightmost zero of Ai, $v_k(t) = \frac{1}{\mathrm{Ai}'(\zeta_k)} \mathrm{Ai}(t+\zeta_k)$, and

$$G_k(t,t) = -v_k(t)v_k'(t) - v_k'(t)v_k(t) + \frac{\pi \,\mathrm{Bi}'(\zeta_k)}{\mathrm{Ai}'(\zeta_k)} \mathrm{Ai}(t+\zeta_k)^2 - \pi \,\mathrm{Ai}(t+\zeta_k)\,\mathrm{Bi}(t+\zeta_k).$$

Conjecture 9.2.1. *We conjecture a large β asymptotic expansion for the mean of X_k^β,*

$$\mathrm{E}[X_k^\beta] = \zeta_k + \frac{1}{\beta}\left(-4\int_0^\infty G_k(t,t)(v_k(t))^2 dt\right) + O\left(\frac{1}{\beta^2}\right), \qquad (9.2.1)$$



(a) Mean (b) Variance

Figure 9.2.1: Large β asymptotics for the universal largest eigenvalue distribution, and four known data points.

and a large β asymptotic expansion for the variance,

$$\mathrm{Var}[X_k^\beta] = \frac{1}{\beta}\left(4\int_0^\infty (v_k(t))^4 dt\right) + O\left(\frac{1}{\beta^2}\right). \tag{9.2.2}$$

The case $k = 1$ for the variance asymptotics appeared previously in [5]. Our argument is somewhat different.

All of the expressions in the conjecture are related to the Airy operator $\mathcal{A}$. The kth eigenvalue of the operator is ζ_k, the kth normalized eigenvector is v_k, and the pseudoinverse of $\mathcal{A} - \zeta_k$ is the integral operator with kernel $G_k(s,t)$, to be defined below. $G_k(t,t)$ is the diagonal of this kernel.

The Green's function $G_k(s,t)$ is known analytically, so each of the coefficients in (9.2.1) and (9.2.2) can be evaluated with a simple numerical quadrature to give

$$\mathrm{E}[X_1^\beta] \approx -2.3381 + 1.1248\frac{1}{\beta},$$
$$\mathrm{Var}[X_1^\beta] \approx 1.6697\frac{1}{\beta}.$$

These approximations are plotted against known data points in Figure 9.2.1.

We give what we feel to be a convincing argument for the second conjecture, using the stochastic operator approach. We state the result as a conjecture, though, since we have not given a concrete interpretation of the diagonal noise operators in the previous chapter. We suspect that the interpretation of white noise will make little difference in this chapter, however, because our argument relies only on the following benign properties of white noise.

1. $\mathrm{E}[W_t] = 0$,

2. $\mathrm{E}[W_s W_t] = \delta(s-t)$,

3. $\mathrm{E}[W_r W_s W_t] = 0$.

All three equations refer to a single white noise "path," and state that the "values" of white noise at distinct time values are uncorrelated, and that the first three moments of white noise at any single time are essentially $0, 1, 0$, just like a Gaussian.

9.3 Justification

Under Conjecture 8.1.1, X_k^β is the distribution of the kth largest eigenvalue of $\mathcal{A} + \frac{2}{\sqrt{\beta}}W$, in which W is a diagonal operator that injects white noise. The boundary conditions are the same as those on the unperturbed operator $\mathcal{A}$: $f(0) = 0$ and $\lim_{t\to\infty} f(t) = 0$.

Using eigenvalue perturbation theory, if λ is the kth largest eigenvalue of $\mathcal{A} + \varepsilon W$ and v is the corresponding eigenvector, then we should have

$$\lambda = \lambda^{(0)} + \varepsilon\lambda^{(1)} + \varepsilon^2\lambda^{(2)} + \varepsilon^3\lambda^{(3)} + O(\varepsilon^4),$$
$$v = v^{(0)} + \varepsilon v^{(1)} + \varepsilon^2 v^{(2)} + O(\varepsilon^3),$$

with

$$\begin{aligned}
\lambda^{(0)} &= \zeta_k, \\
v^{(0)}(t) &= v_k(t) = \frac{1}{\mathrm{Ai}'(\zeta_k)} \mathrm{Ai}(t+\zeta_k), \\
\lambda^{(1)} &= \langle W v^{(0)}, v^{(0)} \rangle, \\
v^{(1)} &= -(\mathcal{A} - \lambda^{(0)})^{+} W v^{(0)}, \\
\lambda^{(2)} &= \langle W v^{(0)}, v^{(1)} \rangle, \\
v^{(2)} &= (\mathcal{A} - \lambda^{(0)})^{+} (-W v^{(1)} + \lambda^{(1)} v^{(1)}), \\
\lambda^{(3)} &= \langle W v^{(0)}, v^{(2)} \rangle.
\end{aligned}$$

(In the asymptotic expansion for v, $O(\varepsilon^3)$ refers to the elements of v.)

Note that $\lambda^{(0)}, v^{(0)}$ is an eigenvalue, eigenvector pair of the unperturbed Airy operator: $(\frac{d^2}{dt^2} - t)\,\mathrm{Ai}(t+\zeta_k) = \zeta_k \,\mathrm{Ai}(t+\zeta_k)$ and $\left[\int_0^\infty \left(\frac{1}{\mathrm{Ai}'(\zeta_k)} \mathrm{Ai}(t+\zeta_k)\right)^2 dt\right]^{1/2} = 1$. In the expression for $v^{(1)}$, $(\mathcal{A} - \lambda^{(0)})^{+}$ denotes the Moore-Penrose pseudoinverse of the operator $\mathcal{A} - \lambda^{(0)}$, that is, the pseudoinverse that sends $v^{(0)}$ to 0. The pseudoinverse is an integral operator whose kernel is a Green's function from ODE theory, $\left[(\mathcal{A} - \lambda^{(0)})^{+} u\right](t) = \int_0^\infty G_k(t,s) u(s) ds$. Computing the Green's function G_k is somewhat tricky since the operator is singular, but verifying it is straightforward.

Lemma 9.3.1. *Let ζ_k be the kth rightmost zero of* Ai, *and let v_k denote the eigenvector of $\mathcal{A}$ belonging to ζ_k, $v_k(t) = \frac{1}{\mathrm{Ai}'(\zeta_k)} \mathrm{Ai}(t+\zeta_k)$. The Green's function $G(s,t)$ for $\mathcal{A} - \zeta_k$ that sends v_k to zero and that satisfies the boundary conditions $G(0,t) = 0$, $\lim_{s\to\infty} G(s,t) = 0$, is*

$$\begin{aligned}
G_k(s,t) := -v_k(s) v_k'(t) - v_k'(s) v_k(t) + \frac{\pi \,\mathrm{Bi}'(\zeta_k)}{\mathrm{Ai}'(\zeta_k)} \mathrm{Ai}(s+\zeta_k)\,\mathrm{Ai}(t+\zeta_k) \\
- \pi \begin{cases} \mathrm{Ai}(t+\zeta_k)\,\mathrm{Bi}(s+\zeta_k) & s \le t \\ \mathrm{Ai}(s+\zeta_k)\,\mathrm{Bi}(t+\zeta_k) & s > t \end{cases}. \qquad (9.3.1)
\end{aligned}$$

Proof. Noting that $\mathcal{A} - \zeta_k$ is Hermitian and $G_k(s,t)$ is symmetric, three conditions must be checked.

First, the boundary conditions must be checked. This step is omitted.

Next, it must be checked that $[(\mathcal{A} - \zeta_k)G_k(\cdot,t)](s) = \delta(s-t) - v_k(s)v_k(t)$. For convenience, denote the terms of the right hand side of (9.3.1) by $C(s) = -v_k(s)v_k'(t)$, $D(s) = -v_k'(s)v_k(t)$, $E(s) = \frac{\pi \operatorname{Bi}'(\zeta_k)}{\zeta_k} \operatorname{Ai}(s+\zeta_k) \operatorname{Ai}(t+\zeta_k)$, and $F(s)$ equaling the final, piecewise term, regarding t as fixed. That $[(\mathcal{A}-\zeta_k)C](s) = [(\mathcal{A}-\zeta_k)D](s) = 0$ is quick to check. It is also straightforward to see that $[(\mathcal{A}-\zeta_k)B](s) = \frac{d}{ds}(-v_k''(s)v_k(t)) - (s+\zeta_k)(-v_k'(s)v_k(t)) = -v_k(s)v_k(t)$. Computing $[(\mathcal{A}-\zeta_k)E](s) = \delta(s-t)$ is the final and most difficult part. First, notice that $[(\mathcal{A}-\zeta_k)]E(s)$ is zero when $s \neq t$, since both Ai and Bi are solutions to Airy's equation. To see what happens at the diagonal, compute

$$\begin{aligned}
\lim_{\delta\to 0}\int_{t-\delta}^{t+\delta}[(\mathcal{A}-\zeta_k)G_k(\cdot,t)](s)ds &= \lim_{\delta\to 0}\int_{t-\delta}^{t+\delta}\left[(\frac{d^2}{du^2} - (u+\zeta_k))G_k(u,t)|_{u=s}\right]ds \\
&= \lim_{\delta\to 0}\int_{t-\delta}^{t+\delta}\left[(\frac{d}{du}(\frac{d}{du}G_k(u,t)))|_{u=s}\right]ds \\
&= \lim_{\delta\to 0}(\frac{d}{ds}G_k(s,t)|_{s=t+\delta} - \frac{d}{ds}G_k(s,t)|_{s=t-\delta}) \\
&= -\pi(\operatorname{Ai}'(t+\zeta_k)\operatorname{Bi}(t+\zeta_k) - \operatorname{Ai}(t+\zeta_k)\operatorname{Bi}'(t+\zeta_k)) \\
&= 1,
\end{aligned}$$

where the final step involves evaluation of the Wronskian of Ai and Bi [25]. This shows that $[(\mathcal{A}-\zeta_k)E](s) = \delta(s-t)$, which completes the proof that $[(\mathcal{A}-\zeta_k)G_k(\cdot,t)](s) = \delta(s-t) - v_k(s)v_k(t)$.

Finally, we must check that the integral operator with kernel G_k sends v_k to zero. Using identities found in [25], notably the Wronskian again, it is straightforward to show that $\int_0^\infty G_k(s,t)v_k(t)dt = 0$. □

Now, we have a stochastic expression for the largest eigenvalue of $\mathcal{A}+\varepsilon W$, specif-

ically, $\lambda = \lambda^{(0)} + \varepsilon\lambda^{(1)} + \varepsilon^2\lambda^{(2)} + \varepsilon^3\lambda^{(3)} + O(\varepsilon^4)$, in which

$$\begin{aligned}
\lambda^{(0)} &= \zeta_k, \\
\lambda^{(1)} &= \int_0^\infty (v_k(t))^2 W_t dt, \\
\lambda^{(2)} &= \int_0^\infty v^{(0)}(t)v^{(1)}(t)W_t dt \\
&= \int_0^\infty v_k(t)\left[-\int_0^\infty G_k(t,s)v_k(s)W_s ds\right] W_t dt \\
&= -\int_0^\infty \int_0^\infty v_k(t)G_k(t,s)v_k(s)W_t W_s ds dt, \\
\lambda^{(3)} &= \int_0^\infty v^{(0)}(t)v^{(2)}(t)W_t dt \\
&= \int_0^\infty v_k(t)\left[\int_0^\infty G_k(t,s)(-W_s v^{(1)}(s) + \lambda^{(1)}v^{(1)}(s))ds\right] W_t dt \\
&= \int_0^\infty \int_0^\infty \int_0^\infty v_k(t)G_k(t,s)G_k(s,r)v_k(r)W_r W_s W_t dr ds dt \\
&\quad - \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty v_k(t)G_k(t,s)G_k(s,q)v_k(q)(v_k(r))^2 W_q W_r W_t dq dr ds dt.
\end{aligned}$$

Taking expectations reveals

$$\begin{aligned}
\mathrm{E}[\lambda^{(0)}] &= \zeta_k, \\
\mathrm{E}[\lambda^{(1)}] &= 0, \\
\mathrm{E}[\lambda^{(2)}] &= -\int_0^\infty \int_0^\infty v_k(t)G_k(t,s)v_k(s)\delta(s-t)ds dt \\
&= -\int_0^\infty G_k(t,t)(v_k(t))^2 dt, \\
\mathrm{E}[\lambda^{(3)}] &= 0.
\end{aligned}$$

Finally, substituting $\varepsilon = \frac{2}{\sqrt{\beta}}$ gives the asymptotic expansion for the mean of the

eigenvalue,

$$\begin{aligned}\mathrm{E}[\lambda] &= \zeta_k + \frac{2}{\sqrt{\beta}} \cdot 0 + \frac{4}{\beta}\left(-\int_0^\infty G_k(t,t)(v_k(t))^2 dt\right) + \frac{8}{\beta^{3/2}} \cdot 0 + O\left(\frac{1}{\beta^2}\right) \\ &= \zeta_k - \frac{4}{\beta}\int_0^\infty G_k(t,t)(v_k(t))^2 dt + O\left(\frac{1}{\beta^2}\right).\end{aligned}$$

For the variance, we have $(\lambda - \mathrm{E}[\lambda])^2 = (\varepsilon\lambda^{(1)} + \varepsilon^2\lambda^{(2)} + O(\varepsilon^3))^2 = \varepsilon^2(\lambda^{(1)})^2 + 2\varepsilon^3\lambda^{(1)}\lambda^{(2)} + O(\varepsilon^4)$, so we compute

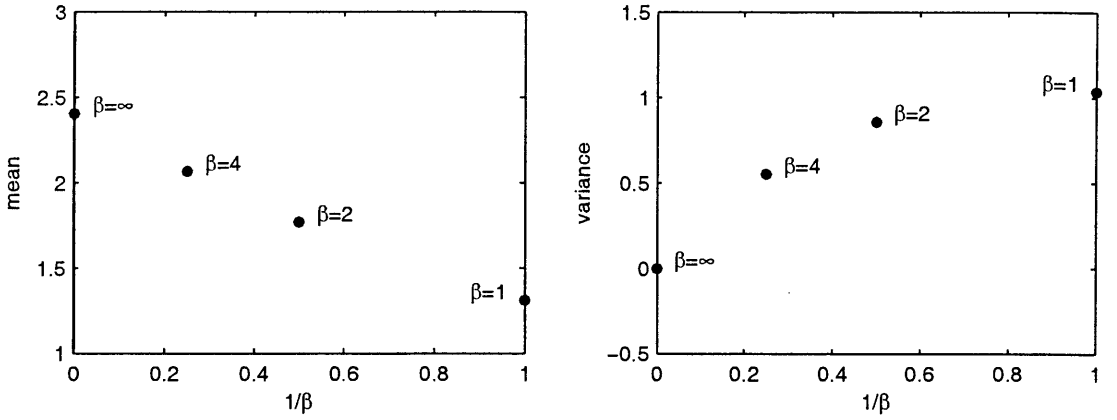
$$\begin{aligned}\mathrm{E}[(\lambda^{(1)})^2] &= \mathrm{E}\left[\int_0^\infty \int_0^\infty (v_k(t))^2(v_k(s))^2 W_s W_t ds dt\right] \\ &= \int_0^\infty (v_k(t))^4 dt, \\ \mathrm{E}[\lambda^{(1)}\lambda^{(2)}] &= \mathrm{E}\left[\left(\int_0^\infty (v_k(t))^2 W_t dt\right)\left(-\int_0^\infty \int_0^\infty v_k(t)G_k(t,s)v_k(s)W_t W_s ds dt\right)\right] \\ &= 0.\end{aligned}$$

Substituting $\varepsilon = \frac{2}{\sqrt{\beta}}$, this gives the asymptotic expansion for the variance,

$$\mathrm{Var}[\lambda] = \mathrm{E}[(\lambda - \mathrm{E}[\lambda])^2] = \frac{4}{\beta}\int_0^\infty (v_k(t))^4 dt + O\left(\frac{1}{\beta^2}\right).$$

9.4 Hard edge and bulk

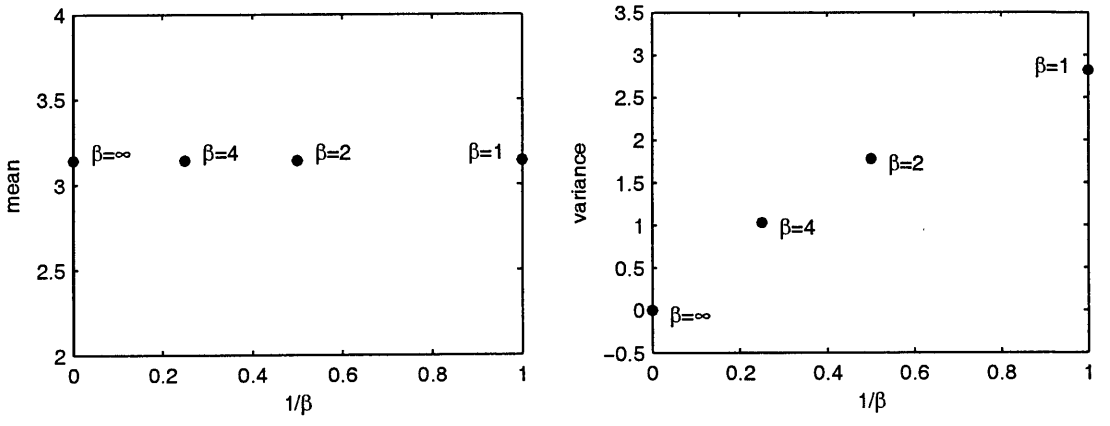
Ideas from this chapter can likely be applied to the hard edge and bulk. Four known data points for the universal smallest singular value distributions and the universal spacing distributions are plotted in Figures 9.4.1 and 9.4.2. In general, these figures do not show as strong of a linear trend as the universal largest eigenvalue distributions, for β as small as 1. In the largest eigenvalue case, $\beta = 1$ may be considered "moderately large," because the linear approximation gives a somewhat accurate approximation to the mean and variance. However, in the smallest singular value and



(a) Mean ($a = 0$) (b) Variance ($a = 0$)

Figure 9.4.1: Mean and variance of the universal smallest singular value distribution.

spacing cases, $\beta = 1$ should not be considered large. Based on this observation, the experiments in Chapter 8 appear to support the claim that the stochastic operator approach is valid for all $\beta > 0$, not just "large β."



(a) Mean (b) Variance

Figure 9.4.2: Mean and variance of the universal spacing distribution.

Appendix A

The algorithm to produce the Jacobi matrix model

Matlab code for the algorithm introduced in Section 5.2.2 follows.

```
function [Y,U,V]=blockbidiagonalize(X,p,q)

%blockbidiagonalize     Transform a unitary matrix into
%                       block bidiagonal form.
%
%   [Y,U,V] = blockbidiagonalize(X,p,q), where X is unitary,
%   produces an equivalent Y in block bidiagonal form with
%   partition size p-by-q.
%
%   1. Y is in bidiagonal block form.
%   2. Y = U'*X*V.
%   3. U is of the form [ U1    ]
%                       [    U2 ]
%      with U1 p-by-p unitary and U2 (m-p)-by-(m-p) unitary.
%   4. V is of the form [ V1    ]
%                       [    V2 ]
%      with V1 q-by-q unitary and V2 (m-q)-by-(m-q) unitary.
%
%   If X is not unitary, then Y has the "related sign pattern" in
%   Chapter 5.
%
%   For demonstration purposes only. This code may be numerically
%   unstable.

%%%%%%%%%%  Argument checking  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%% If p<q, work with conjugate transpose.
```

```
if p<q
  [Y2,U2,V2]=blockbidiagonalize(X',q,p);
  Y=Y2';U=V2;V=U2;
  return;
end

[m,n]=size(X);
if m~=n
  error('X must be square.')
end
if p<0||q<0||p+q>m
  error('Invalid partition size.')
end

%%%%%%%%%% Initialization %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Y=X;U=eye(m);V=eye(n);Z=zeros(m,m,2*q);

%%%%%%%%%% Main loop %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%% Iterate over rows and columns.
%%% At each iteration, choose Householder reflectors based on
%%% sections of column p and then on sections of row p.
for k=1:q

  angles(q+1-k,1)=atan(norm(Y(p+k:end,k))/norm(Y(k:p,k)));

  %%% Put zeros in Y(k+1:p,k).
  [w,eta]=gallery('house',Y(k:p,k));
  F=blkdiag(eye(k-1),eye(p-k+1)-eta*w*w',eye(m-p));U=U*F';Y=F*Y;
  D=singlephase(m,k,conj(nzsign(Y(k,k))));U=U*D';Y=D*Y;

  %%% Put zeros in Y(p+k+1:end,k).
  [w,eta]=gallery('house',Y(p+k:end,k));
  F=blkdiag(eye(p+k-1),eye(m-p-k+1)-eta*w*w');U=U*F';Y=F*Y;
  D=singlephase(m,p+k,-conj(nzsign(Y(p+k,k))));U=U*D';Y=D*Y;

  Z(:,:,2*k-1)=Y;

  if k<q
    angles(q-k,2)=atan(norm(Y(k,q+k:end))/norm(Y(k,k+1:q)));
  end

  %%% Put zeros in Y(k,k+2:q).
  if k<q
    [w,eta]=gallery('house',Y(k,k+1:q)');
    F=blkdiag(eye(k),eye(q-k)-eta*w*w',eye(m-q));V=V*F;Y=Y*F;
    D=singlephase(m,k+1,-conj(nzsign(Y(k,k+1))));V=V*D;Y=Y*D;
  end

  %%% Put zeros in Y(k,q+k+1:end).
```

```
  [w,eta]=gallery('house',Y(k,q+k:end)');
  F=blkdiag(eye(q+k-1),eye(m-q-k+1)-eta*w*w');V=V*F;Y=Y*F;
  D=singlephase(m,q+k,conj(nzsign(Y(k,q+k))));V=V*D;Y=Y*D;

  Z(:,:,2*k)=Y;

end

%%%%%%%%%%  Post-processing  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%% Transform Y([1+q:p p+q+1:end],2*q+1:end) into lower triangular.
if 2*q<m
  [Q,R]=qr(Y([1+q:p p+q+1:end],2*q+1:end)');
  Q=Q*diag(conj(sign(diag(R))));
  Q=blkdiag(eye(2*q),Q);
  V=V*Q;Y=Y*Q;
end

%%%%%%%%%%  Utility functions  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function D=singlephase(m,i,phase)
D=diag([ones(1,i-1) phase ones(1,m-i)]);

function omega=nzsign(z)
if abs(sign(z))>0.5,omega=sign(z);else,omega=1;end
```

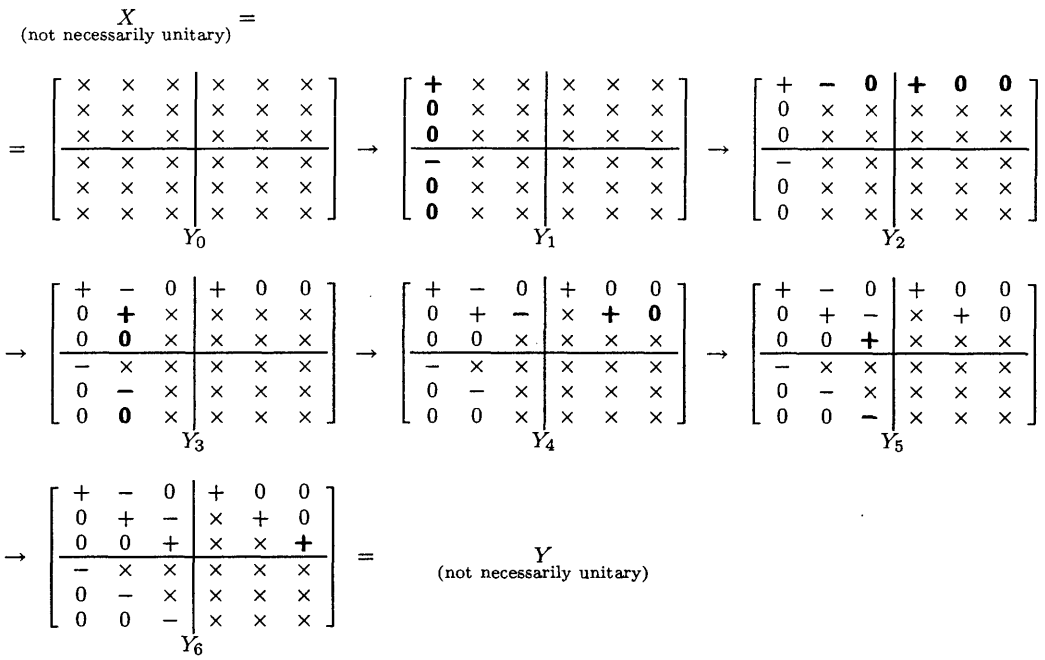


Figure A.0.1: An example run of the algorithm on a non-unitary matrix. In this example, $m = 6$, $p = 3$, and $q = 3$. When $p + q < m$, a little more work is involved.

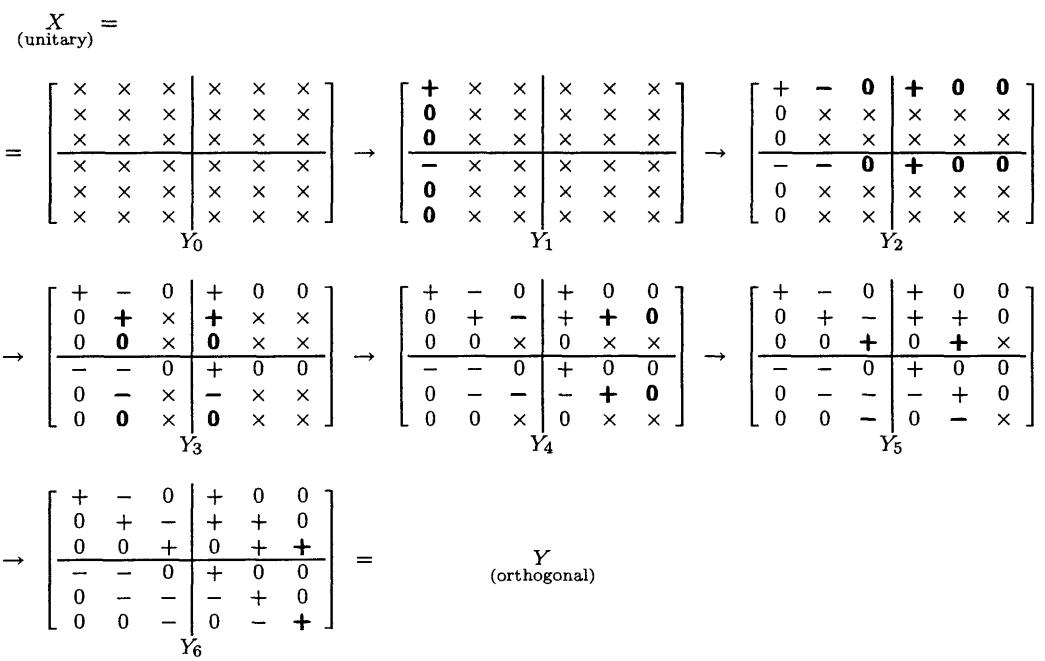


Figure A.0.2: An example run of the algorithm on a unitary matrix. The extra zeros, contrasted with Figure A.0.1, are guaranteed by the orthogonality constraints; no extra work is necessary.

Bibliography

[1] Milton Abramowitz and Irene A. Stegun, editors. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* Dover Publications Inc., New York, 1992. Reprint of the 1972 edition.

[2] Jinho Baik, Percy Deift, and Kurt Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.*, 12(4):1119–1178, 1999.

[3] Burt V. Bronk. Accuracy of the semicircle approximation for the density of eigenvalues of random matrices. *J. Mathematical Phys.*, 5:215–220, 1964.

[4] Ioana Dumitriu. *Eigenvalue statistics for beta-ensembles.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA 02139, June 2003.

[5] Ioana Dumitriu and Alan Edelman. Eigenvalues of Hermite and Laguerre ensembles: large beta asymptotics. *Annales de L'Institut Henri Poincare (B).* To appear.

[6] Ioana Dumitriu and Alan Edelman. Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.

[7] Freeman J. Dyson. Statistical theory of the energy levels of complex systems. I. *J. Mathematical Phys.*, 3:140–156, 1962.

[8] Freeman J. Dyson. The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics. *J. Mathematical Phys.*, 3:1199–1215, 1962.

[9] Alan Edelman. *Eigenvalues and condition numbers of random matrices.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA 02139, May 1989.

[10] Alan Edelman and Per-Olof Persson. Numerical methods for eigenvalue distributions of random matrices. 2005. `http://arxiv.org/abs/math-ph/0501068`.

[11] Merran Evans, Nicholas Hastings, and Brian Peacock. *Statistical distributions.* Wiley Series in Probability and Statistics: Texts and References Section. Wiley-Interscience, New York, third edition, 2000.

[12] P. J. Forrester. The spectrum edge of random matrix ensembles. *Nuclear Phys. B*, 402(3):709–728, 1993.

[13] P. J. Forrester. Painlevé transcendent evaluation of the scaled distribution of the smallest eigenvalue in the laguerre orthogonal and symplectic ensembles. 2000. `http://arxiv.org/abs/nlin.SI/0005064`.

[14] Peter J. Forrester. *Log-Gases and Random Matrices.* `http://www.ms.unimelb.edu.au/~matpjf/matpjf.html`. 14 July 2004.

[15] Roger A. Horn and Charles R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, 1990. Corrected reprint of the 1985 original.

[16] Michio Jimbo, Tetsuji Miwa, Yasuko Môri, and Mikio Sato. Density matrix of an impenetrable Bose gas and the fifth Painlevé transcendent. *Phys. D*, 1(1):80–158, 1980.

[17] Kurt Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.*, 209(2):437–476, 2000.

[18] Iain M. Johnstone. On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.*, 29(2):295–327, 2001.

[19] Rowan Killip and Irina Nenciu. Matrix models for circular ensembles. *Int. Math. Res. Not.*, (50):2665–2701, 2004.

[20] A. B. J. Kuijlaars and M. Vanlessen. Universality for eigenvalue correlations from the modified Jacobi unitary ensemble. *Int. Math. Res. Not.*, (30):1575–1600, 2002.

[21] Ross A. Lippert. A matrix model for the β-Jacobi ensemble. *J. Math. Phys.*, 44(10):4807–4816, 2003.

[22] Madan Lal Mehta. *Random matrices.* Academic Press Inc., Boston, MA, second edition, 1991.

[23] Robb J. Muirhead. *Aspects of multivariate statistical theory.* John Wiley & Sons Inc., New York, 1982. Wiley Series in Probability and Mathematical Statistics.

[24] A. M. Odlyzko. On the distribution of spacings between zeros of the zeta function. *Math. Comp.*, 48(177):273–308, 1987.

[25] Frank W. J. Olver. *Digital library of mathematical functions*, chapter AI: Airy and related functions. National Institute of Standards and Technology, 1998. `http://dlmf.nist.gov/Contents/AI/AI.html`.

[26] C. C. Paige and M. A. Saunders. Towards a generalized singular value decomposition. *SIAM J. Numer. Anal.*, 18(3):398–405, 1981.

[27] C. C. Paige and M. Wei. History and generality of the CS decomposition. *Linear Algebra Appl.*, 208/209:303–326, 1994.

[28] Beresford Parlett. Private communication. 24 Dec 2004.

[29] Charles E. Porter and Norbert Rosenzweig. Statistical properties of atomic and nuclear spectra. *Ann. Acad. Sci. Fenn. Ser. A VI No.*, 44:66, 1960.

[30] John D. Pryce. *Numerical solution of Sturm-Liouville problems.* Monographs on Numerical Analysis. The Clarendon Press Oxford University Press, New York, 1993. Oxford Science Publications.

[31] Atle Selberg. Remarks on a multiple integral. *Norsk Mat. Tidsskr.*, 26:71–78, 1944.

[32] Jack W. Silverstein. The smallest eigenvalue of a large-dimensional Wishart matrix. *Ann. Probab.*, 13(4):1364–1368, 1985.

[33] Gábor Szegő. *Orthogonal polynomials.* American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.

[34] Craig A. Tracy and Harold Widom. Introduction to random matrices. In *Geometric and quantum aspects of integrable systems (Scheveningen, 1992)*, volume 424 of *Lecture Notes in Phys.*, pages 103–130. Springer, Berlin, 1993.

[35] Craig A. Tracy and Harold Widom. Level-spacing distributions and the Airy kernel. *Comm. Math. Phys.*, 159(1):151–174, 1994.

[36] Craig A. Tracy and Harold Widom. Level spacing distributions and the Bessel kernel. *Comm. Math. Phys.*, 161(2):289–309, 1994.

[37] Craig A. Tracy and Harold Widom. On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.*, 177(3):727–754, 1996.

[38] Craig A. Tracy and Harold Widom. The distribution of the largest eigenvalue in the Gaussian ensembles: $\beta = 1, 2, 4$. In *Calogero-Moser-Sutherland models (Montréal, QC, 1997)*, CRM Ser. Math. Phys., pages 461–472. Springer, New York, 2000.

[39] Lloyd N. Trefethen and David Bau, III. *Numerical linear algebra.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[40] Hale F. Trotter. Eigenvalue distributions of large Hermitian matrices; Wigner's semicircle law and a theorem of Kac, Murdock, and Szegő. *Adv. in Math.*, 54(1):67–82, 1984.

[41] Charles F. Van Loan. Generalizing the singular value decomposition. *SIAM J. Numer. Anal.*, 13(1):76–83, 1976.

[42] Ming Chen Wang and G. E. Uhlenbeck. On the theory of the Brownian motion. II. *Rev. Modern Phys.*, 17:323–342, 1945.

[43] Wolfram functions site. `http://functions.wolfram.com/`. April 2005.

Notation

^{T}	transpose
*	Hermitian adjoint
$\oplus$	matrix direct sum: $X \oplus Y = \left[\begin{smallmatrix} X & \\ & Y \end{smallmatrix}\right]$
$\stackrel{d}{=}$	equal in distribution
$\mathbb{1}_S$	indicator function for the set S
$\mathcal{A}$	Airy operator (Figure 1.0.4)
Ai	Airy Ai function (§3.2)
beta(c, d)	beta-distributed random variable (§3.5)
Bi	Airy Bi function (§3.2)
χ_r	chi-distributed random variable (§3.5)
D_1	first difference matrix (eq. (3.6.1))
D_2	second difference matrix (eq. (3.6.2))
δ_{ij}	Kronecker delta
F	flip permutation matrix (eq. (3.6.3))
G	real standard Gaussian-distributed random variable (except in Chapter 5)
H^β	Hermite matrix model (Figures 1.0.3, 6.0.3)
$\tilde{H}^\beta$	Gaussian approximation to the Hermite matrix model (Figure 8.2.2)
j_a	Bessel function of the first kind (§3.2)
$\mathcal{J}_a$	Bessel operator (Figure 1.0.5)
$\tilde{\mathcal{J}}_a$	Bessel operator (Liouville form) (Figure 1.0.5)
$\mathcal{J}_{-1/2}$	sine operator (Figure 1.0.6)
$\tilde{\mathcal{J}}_{-1/2}$	sine operator (Liouville form) (Figure 1.0.6)

$J_{a,b}^{\beta}$	Jacobi matrix model (Figures 1.0.1, 6.0.1)
$\tilde{J}_{a,b}^{\beta}$	Gaussian approximation to the Jacobi matrix model (Figure 8.2.2)
$K^{\text{Airy}}(x,y)$	Airy kernel (§3.3.4)
$K^{\text{Bessel}}(a;x,y)$	Bessel kernel (§3.3.4)
$K_n^H(x,y)$	Hermite kernel (§3.3.1)
$K_n^J(a,b;x,y)$	Jacobi kernel (§3.3.1)
$K_n^L(a;x,y)$	Laguerre kernel (§3.3.1)
$K^{\text{sine}}(x,y)$	sine kernel (§3.3.4)
L_a^{β}	square Laguerre matrix model (Figures 1.0.2, 6.0.2)
$\tilde{L}_a^{\beta}$	Gaussian approximation to the square Laguerre matrix model (Figure 8.2.2)
M_a^{β}	rectangular Laguerre matrix model (Figures 1.0.2, 6.0.2)
$\tilde{M}_a^{\beta}$	Gaussian approximation to the rectangular Laguerre matrix model (Figure 8.2.2)
N	matrix with i.i.d. standard Gaussian entries, either real or complex
$O(m)$	group of m-by-m orthogonal matrices
Ω	alternating signs permutation matrix (eq. (3.6.4))
P	perfect shuffle permutation matrix (eq. (3.6.5))
$\pi_n^H(x)$	Hermite polynomial (§3.3.1)
$\pi_n^J(a,b;x)$	Jacobi polynomial (§3.3.1)
$\pi_n^L(a;x)$	Laguerre polynomial (§3.3.1)
$\psi_n^H(x)$	Hermite function (§3.3.1)
$\psi_n^J(a,b;x)$	Jacobi function (§3.3.1)
$\psi_n^L(a;x)$	Laguerre function (§3.3.1)
$\mathbb{R}$	the real numbers
$R[x]$	remainder term bounded in magnitude by $\lvert x\rvert$ (proof of Theorem 7.2.1)
$U(m)$	group of m-by-m unitary matrices
W	diagonal white noise operator (Chapter 8)
W_t	white noise generalized stochastic process (§9.2)
$w^H(x)$	Hermite weight e^{-x^2} (§3.3.1)

$w^J(a,b;x)$	Jacobi weight $x^a(1-x)^b$ (§3.3.1)
$w^L(a;x)$	Laguerre weight $e^{-x}x^a$ (§3.3.1)
$\mathbb{Z}$	the integers

Index