Operator Limits of Random Matrices

Diane Holcomb
with contributions from Bálint Virág

These notes were written for a short course taught by myself, Diane Holcomb, at KTH in November-December of 2017. They incorporate notes written for a short course taught by Bálint Virág at PCMI summer school in July 2017 which were written as a collaboration between myself and Báling Virág. Portions of these notes will also be based on a random matrices course given by Benedek Valkó at University of Wisconsin circa Spring 2012.

The intention of the lectures here at KTH is to provide an introduction to the type of operator convergence work that is becoming better known in the random matrix community, but is still not widely studied or understood. Major contributions to this area include the work of Alan Edelman, Ioana Dumitriu, Brian Sutton, Jose Ramírez, Brian Rider, Bálint Virág, and Benedek Valkó. There will naturally be portions of the area that I will not be able to cover in full detail in the lecture. These notes will fill in some, though probably not all of these details.

I also would like to make a few notes about the style and organization of the notes. The goal here is the introduction to the ideas underlying the field. Because of this I will not be doing things in full generality, nor proving common statements from random matrices. I hope that when I finish the notes each chapter will have a first section that gives in some sense the narrative of the chapter. The lectures will largely be drawn from these sections. Remaining sections will discuss the results from the chapter in greater generality and/or prove statements left unproven in the first section. The exception to this is chapter 1 which includes a variety of preliminaries that are necessary for the remainder of the notes. The current version is still very much a work in progress.

Contents

CHAPTER 1

The Gaussian Ensembles

1. The Gaussian Orthogonal and Unitary Ensembles

One of the earliest random matrix models was introduced by Eugene Wigner in the 1950's. The model can be characterized in many ways, but we will start with the following construction: let M_n be a symmetric matrix with

$$m_{ij} \sim \mathcal{N}(0,1) \quad \text{for } i < j, \qquad \text{and} \qquad m_{ii} \sim \mathcal{N}(0,2).$$

The symmetry condition completes the description of the matrix. A matrix constructed in this way is said to be a member of the Gaussian Orthogonal Ensemble. If we instead take $m_{ij} \sim \mathbb{C}\mathcal{N}(0,1) \sim X + iY$ a complex standard normal where X and Y are independent with $\mathcal{N}(0,1)$ distribution and the constraint $M_n = M_n^*$ we get the Gaussian Unitary ensemble.

THEOREM 1.1. *Suppose M_n has GOE or GUE distribution then M_n has eigenvalue density*

$$f(\lambda_1, ..., \lambda_n) = \frac{1}{Z_n} \prod_{k=1}^{n} e^{-\frac{\beta}{4}\lambda_k^2} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \tag{1}$$

with $\beta = 1$ for the GOE and $\beta = 2$ for the GUE.

For convenience we will take $\Lambda^{(n)} = \{\lambda_i\}_{i=1}^n$ to denote the set of eigenvalues of the GOE or GUE. This notation will be used later to denote the eigenvalues or points in whatever random matrix model is being discussed at the time.

Observe that for the joint density above we can move all the the terms into the exponent giving us that

$$f(\lambda_1, ..., \lambda_n) = \frac{1}{Z_n} \exp\left[-\frac{\beta}{4}\sum_{k=1}^{n}\lambda_k^2 + \beta\sum_{i<j}\log|\lambda_i - \lambda_j|\right].$$

From this we can see that this is a model for n particles interacting through a Hamiltonian with a potential $V(x) = x^2$ and an interaction term $\log|x-y|$. These two terms work against each other, the potential term pushing the particles together while the interaction term pushes them apart. Now consider the what happens as n increases. The first sum has n terms while the second has $\binom{n}{2}$, therefore as n tends to infinity the repulsion will be the dominant term. This means that the points will be found further and further out and will not remain in a compact interval. To put these two terms on the same scale we do the mapping

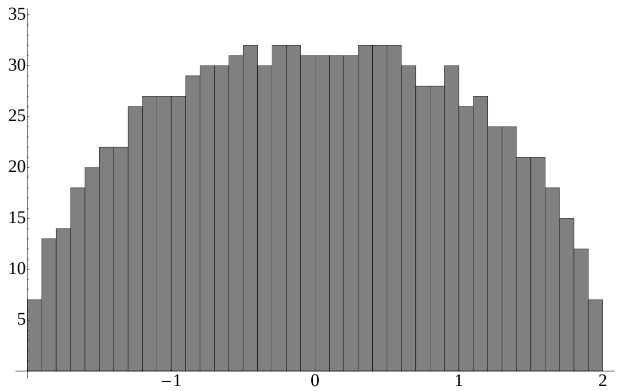

FIGURE 1. Rescaled eigenvalues of a 1000×1000 GOE matrix

$\lambda_i \mapsto \sqrt{n}x_i$. Then the joint density for the x_i's will be given by

$$\begin{aligned} f_x(x_1, ..., x_n) &= \frac{e^{\binom{n}{2}\log\sqrt{n}}}{Z_n n^{n/2}} \exp\left[-n\frac{\beta}{4}\sum_{k=1}^n x_k^2 + \beta\sum_{i<j}\log|x_i - x_j|\right] \\ &= \frac{e^{\binom{n}{2}\log\sqrt{n}}}{Z_n n^{n/2}} \exp\left[-n^2\frac{\beta}{4}\int x^2 d\nu_n(x) + n^2\beta\iint \log(x-y)d\nu_n(x)d\nu_n(y)\right], \end{aligned}$$

where here $\nu_n = \frac{1}{n}\sum_{k=1}^n \delta_{x_k}$ is the empirical measure of the x_i's. Notice that the original mapping was equivalent to scaling the eigenvalues down by a factor of $\sqrt{n}$ and so we get that

$$\nu_n = \frac{1}{n}\sum_{k=1}^n \delta_{\lambda_i/\sqrt{n}}. \tag{2}$$

If these measures converge as $n \to \infty$ it must be to the measure which minimizes the energy of the system with Hamiltonian

$$H_\infty(\sigma) = \frac{1}{2}\int x^2 d\sigma(x) - \iint \log(x-y)d\sigma(x)d\sigma(y).$$

This Hamiltonian is minimized by the semicircle measure σ_{sc} with density

$$\frac{d\sigma_{sc}}{dx} = \frac{1}{2\pi}\sqrt{4-x^2}\ \mathbf{1}_{x\in[-2,2]} = \rho_{sc}(x).$$

EXERCISE 1.2. Show that σ_{sc} minimizes H_∞.

THEOREM 1.3 (Wigner's semi-circle law). *Let ν_n be the sequence of measures defined in (2). Then as $n \to \infty$*

$$\nu_n \Rightarrow \sigma_{sc}.$$

One way of thinking of this statement is to consider the histogram of the rescaled eigenvalues of a large GOE matrix. This histogram will be approximately semi-circular in shape. See figure 1. For a proof of the semicircle law involving graph convergence see chapter 2.

1.1. Properties of the GOE. One of the most important properties of the GOE after the fact that its eigenvalue density has an explicit form is that the distribution of the GOE is invariant under orthogonal conjugation.

THEOREM 1.4. *Let M_n have GOE distribution and O be an independent orthogonal matrix. Then $M_n \stackrel{d}{=} OM_nO^t$.*

Before proving the above theorem we need to introduce at least one other description of the GOE.

PROPOSITION 1.5. *The following characterizations of the Gaussian Orthogonal Ensemble are equivalent:*

(1) *$M_n = M_n^t$ with $m_{ii} \sim \mathcal{N}(0,2)$ and $m_{ij} \sim \mathcal{N}(0,1)$ for $i < j$ all independent.*
(2) *$M_n = \frac{X_n + X_n^t}{\sqrt{2}}$ whre $x_{i,j} \sim \mathcal{N}(0,1)$ all independent.*
(3) *M_n has density $\frac{1}{Z}\exp(-\frac{1}{4}Tr\ MM^T)$ on the space of symmetric matrices.*

PROOF. If $\underline{v}$ is a standard normal vector and O is orthogonal then $O\underline{v} \stackrel{d}{=} \underline{v}$. Therefore, for $X = [v_1, v_2, ..., v_n]$ i.i.d vectors we get $OX \stackrel{d}{=} X$. Then using the second description of the GOE we get that $OX_nO^T \stackrel{d}{=} X_n$.

Alternatively, consider the third description of the GOE. We look at the transformation $X \mapsto OXO^T$. The density will be invariant if we can show that $\text{Tr}\ XX^T = \text{Tr}\ OXO^T(OXO^T)^T$ and the Jacobian has a magnitude of 1. Indeed we do get the equality of the traces because conjugation with an orthogonal matrix fixes the eigenvalues. For the second part we need the following lemma:

LEMMA 1.6. *The Jacobian of $X \mapsto OXO^T$ is 1.*

This completes the proof of the previous lemma

□

PROOF OF LEMMA 1.6. Let $v_X = (x_{11}, x_{22}, ..., x_{nn}, x_{12}, x_{13}, ..., x_{n-1,n})^T$. Then there is a matrix A with $v_{O^TXO} = Av_X$. We need $|\det A|$. Let $D = diag(1, 1, ..., 1, 2, ..., 2)$ (n 1's and $(n^2 - n)/2$ 2's). Then

$$v_X^T D v_X = \text{Tr}\ XX^T$$

moreover we get that

$$v_X^T A^T D A v_X = v_{O^TXO} D v_{O^TXO} = \text{Tr}\ XX^T$$

Therefore $A^T DA = D$and so $|\det A|^2 \det D = \det D$ which implies $|\det A| = 1$. □

REMARK 1.7. Similar results hold for the GUE, but the orthogonal matrix O is replaced by a unitary matrix U.

2. Global and local convergence

The Wigner semi-circle law is an example of a global or macroscopic convergence result. In this setting the rescaled eigenvalues $\{\lambda_i/\sqrt{n}\}_{i=1}^n$ accumulate on a compact interval and so in the limit become indistinguishable from each other. If we are instead interested in the local interactions between eigenvalues (the interactions of eigenvalues with their neighbors) we need to be able to see the behavior of individual points in the limit.

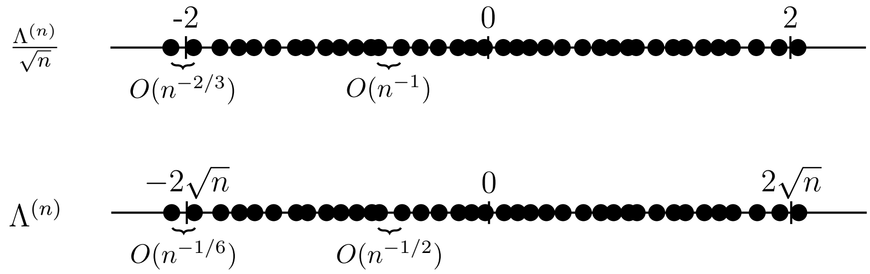

FIGURE 2. Spectrum of a GOE scaled down to $[-2, 2]$ and in natural scale.

Figure 2 shows the spectrum of a GOE on two different scales along with the order of magnitude of the spacing. Notice that at this point if we took $n \to \infty$ for either $\Lambda^{(n)}/\sqrt{n}$ or $\Lambda^{(n)}$ all of the points would become indistinguishable even at the edge. To see where the spacing comes from consider the Wigner semi-circle law. When n is large we get that for $a < b \in [-2, 2]$

$$\#\{x_i = \lambda_i/\sqrt{n} \in [a, b]\} \approx n \int_a^b d\sigma_{sc}(x) = n \int_a^b \frac{1}{2\pi}\sqrt{4 - x^2}dx$$

When $[a, b] = [a, a+\delta]$ with $a \in (-2, 2-\delta)$ this become approximately the length of the interval multiplied by ρ_{sc} evaluated at a. Therefore $\#\{x_i \in [a, b]\} \approx n\delta\rho_{sc}(a)$ (recall $x_i = \lambda_i/\sqrt{n}$), meaning that there are order n points in any such interval and the approximate spacing of the points is $O(1/n)$.

EXERCISE 2.1. Check that the number of points x_i in an interval of the form $[-2, -2 + \delta]$ is on the order of $n^{2/3}$ meaning that the spacing is $O(n^{-2/3})$.

At this point we can see that the scale at which the points of a GOE will be order 1 about will change depending on the location. We call the set $(-2\sqrt{n}, 2\sqrt{n})$ the bulk of the spectrum and the points $\pm 2\sqrt{n}$ the edge. In the bulk of the spectrum we need to scale up by $f(a)\sqrt{n}$ where $f(a)$ is some function of a. At the edge we need to scale up by $n^{1/6}$. See Figure 2 for a picture of what is happening. The limiting spectral measure (in this case the semi-circle) should be taken as a guide for the correct scale at which to see local interaction but is not guaranteed to give the right answer. In this case it does and we get the following results:

THEOREM 2.2 ([**7**]). *Let $\lambda_1 \geq \lambda_2 \geq \cdots$ denote the ordered eigenvalues of the GOE then*

$$\left\{n^{1/6}(2\sqrt{n} - \lambda_i)\right\}_{i=1,\dots,k} \Rightarrow \{\Lambda_1, \Lambda_2, \dots, \Lambda_k\}$$

where Λ_i are the ordered eigenvalues of a certain random differential operator.

The differential operator and the theorem will be discussed in further detail in chapter 4.

THEOREM 2.3 ([**9**],[**10**]). *Let $\Lambda^{(n)}$ be the set of eigenvalues of a GOE and $a \in (-2, 2)$ then*

$$\rho_{sc}(a)\sqrt{n}(\Lambda^{(n)} - a\sqrt{n}) \Rightarrow \mathrm{Sine}_1$$

where Sine_1 *is a point process that may be characterized as the eigenvalues of a certain random Dirac operator.*

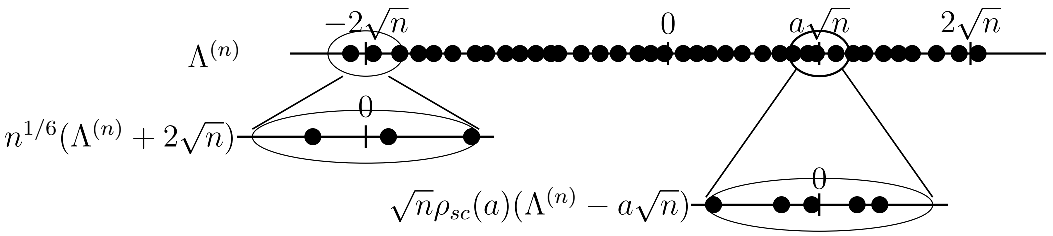

FIGURE 3. Illustration of the correct scale to see local interactions.

REMARK 2.4. These theorems were originally proved and stated using the integrable structure of the GOE. The GOE eigenvalues form a Pfaffian point process with a kernel constructed from Hermite polynomials. The limiting processes may be identified to looking at the limit of the kernel in the appropriate scale.

3. Tridiagonalization and spectral measure

Up to this point we have discussing the empirical measure associated to the eigenvalues, but there is another measure that we can discuss which will turn out to be useful. That is the spectral measure of a matrix A at a vector v. For this section we will always take $v = \mathbf{e}_1$ the first coordinate vector.

DEFINITION 3.1. Suppose you have a symmetric matrix A, we can define its spectral measure (at the first coordinate vector) σ_A. The spectral measure is the measure for which

$$\int x^k d\sigma_A = A^k_{11}.$$

EXERCISE 3.2. Check that if $\lambda_1, ..., \lambda_n$ are the eigenvalues of A then

$$\sigma_A = \sum_i \delta_{\lambda_i} \varphi_i(1)^2$$

where φ_i is the ith normalized eigenvector of A.

We say two symmetric matrices are equivalent if they have the same eigenvalues with multiplicity. This equivalence is well understood: two matrices are equivalent if and only if they are conjugates by an orthogonal matrix (in group theory language, the equivalence classes are the orbits of the conjugation action of the orthogonal group). There is a canonical representative in each class, a diagonal matrix with non-increasing diagonals.

Can we have a similar characterization for matrices with the same same spectral measure?

DEFINITION 3.3. A vector v is cyclic for an $n \times n$ matrix A if $v, Av, ..., A^{n-1}v$ is a basis for the vector space $\mathbb{R}^n$.

THEOREM 3.4. *Let A and B be two matrices, then $\sigma_A = \sigma_B$ if an only if $O^{-1}AO = B$ where O is orthogonal of the form $O_{11} = 1, O_{1i} = 0$ for $i \neq 1$.*

DEFINITION 3.5. A Jacobi matrix is a real symmetric tridiagonal matrix with positive off-diagonals.

THEOREM 3.6. *For all A there exists a unique Jacobi matrix J such that $\sigma_J = \sigma_A$*

PROOF OF EXISTENCE. We will use the Householder transformation to take an $n \times n$ symmetric matrix to a similar tridiagonal matrix. Write our matrix in a block form,

$$A = \left[\begin{array}{c|c} a & b^t \\ \hline b & C \end{array}\right]$$

Now let O be an $(n-1) \times (n-1)$ orthogonal matrix, and let

$$Q = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & O \end{array}\right]$$

Then Q is orthogonal and

$$QAQ^t = \left[\begin{array}{c|c} a & (Ob)^t \\ \hline Ob & OCO^t \end{array}\right]$$

Now we can choose the orthogonal matrix O so that Ou is in the direction of the first coordinate vector, namely $Ou = |u|\mathbf{e}_1$.

An expliction option for O is the following Householder reflection:

$$Ov = v - 2\frac{\langle v, w\rangle}{\langle w, w\rangle} w \qquad \text{where} \qquad w = b - |b|\mathbf{e}_1$$

Check that $OO^t = I$, $Ob = |b|\mathbf{e}_1$.

Therefore

$$QAQ^t = \left[\begin{array}{c|cccc} a & |b| & 0 & \dots & 0 \\ \hline |b| & & & & \\ 0 & & & & \\ \vdots & & & OCO^t & \\ 0 & & & & \end{array}\right].$$

We now repeat the previous step, but this time choosing the first two rows and columns to be 0 except having 1's in the diagonal entries, and than again until the matrix becomes tridiagonal. ☐

EXERCISE 3.7. Suppose e_1 is cyclic. Apply Gram-Schmidt to the vectors $(e_1, Ae_1, ..., A^{n-1})$ to get a new orthonormal basis. Show that A written in this basis will be a Jacobi matrix.

EXERCISE 3.8. A tridiagonal matrix J is called a Jacobi matrix if the off diagonal entries are all positive. We define the spectrum of J to be the measure that satisfies $\int p(x) d\sigma(x) = \langle e_1, p(A)e_1\rangle$ for all polynomials $p(x)$. Show that if J and J' are two Jacobi matrices with $\sigma_J = \sigma_{J'}$ then $J = J'$.
Hint: Look at the moments of the measure.

3.1. Tridiagonalization and the GOE. When we do this tridiagonalization procedure to the GOE we get a very clean result because of the invariance of the distribution under independent orthogonal transformation.

PROPOSITION 3.9 ([**8**]). *Let A be GOE_n. There exists a random orthogonal matrix fixing the first coordinate vector e so that*

$$OAO^t = \begin{bmatrix} a_1 & b_1 & 0 & \dots & 0 \\ b_1 & a_2 & \ddots & & \\ 0 & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & a_{n-1} & b_{n-1} \\ 0 & & & b_{n-1} & a_n \end{bmatrix}$$

with all entries indepdent, $a_i \sim \mathcal{N}(0,2)$ and $b_i \sim \chi_{n-i}$. In particular, $OAO^t =$ has the same spectral measure as A.

For those unfamiliar with the chi distribution we make the following observations. Suppose that k is an integer and v is a vector of independent $\mathcal{N}(0,1)$ random variables of length k, then $\chi_k \stackrel{d}{=} \|v\|$. The density of a χ random for $k > 0$ is given by

$$f_{\chi_k}(x) = \frac{1}{2^{\frac{k}{2}-1}\Gamma(k/2)} x^{k-1} e^{-x^2/2},$$

where $\Gamma(x)$ is the Gamma function.

PROOF. The argument above can be applied to the random matrix. After the first step, OCO^t will be independent of a, b and have a GOE distribution. This is because GOE is invariant by conjugation with a fixed O, and O is only a function of b. So conditionally on a,b, both C and OCO^t have GOE distribution. $\square$

EXERCISE 3.10. Let $X_{n\times m}$ be an $n \times m$ matrix with $x_{i,j} \sim \mathcal{N}(0,1)$ (not symmetric nor Hermitian). The distribution of this matrix is invariant under left and right multiplication by independent Unitary matrices. Show that such a matrix X may be lower bidiagonalized such that the distribution of the singular values is the same for both matrices. Note that the singular values of a matrix are unchanged by multiplication by a unitary matrix.

(1) Start by coming up with a matrix that right multiplied with A gives you a matrix where the first row is 0 except the 11 entry.
(2) What can you say about the distribution of the rest of the matrix after this transformation to the first row?
(3) Next apply a left multiplication. Continue using right and left multiplication to finish the bidiagonalization.

4. β-ensembles

EXERCISE 4.1. For every spectral measure σ there exists a symmetric matrix with that spectral measure. This implies that there exists a Jacobi matrix with this spectral measure.

Let

$$A_n = \frac{1}{\sqrt{\beta}} \begin{bmatrix} a_1 & b_1 & 0 & \dots & 0 \\ b_1 & a_2 & \ddots & & \\ 0 & \ddots & \ddots & \ddots & \\ \vdots & & \ddots & a_{n-1} & b_{n-1} \\ 0 & & & b_{n-1} & a_n \end{bmatrix}$$

That is a tridiagonal matrix with $a_1, a_2, ..., a_n \sim N(0,2)$ on the diagonal and $b_1, ..., b_{n-1}$ with $b_k \sim \chi_{\beta(n-k)}$ and everything independent. Recall that if $z_1, z_2, ...$ are iid $N(0,1)$ then $z_1^2 + \cdots + z_k^2 \sim \chi_k^2$.

If $\beta = 1$ then A_n is similar to a GOE matrix (the joint density of the eigenvalues is the same). If $\beta = 2$ then A_n is similar to a GUE matrix.

Theorem 4.2 ([**2**]). *If $\beta > 0$ then the joint density of the eigenvalue of A_n is given by*

$$f(\lambda_1, ..., \lambda_n) = \frac{1}{Z_{n,\beta}} e^{-\frac{\beta}{4}\sum_{i=1}^n \lambda_i^2} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^\beta.$$

Before we start the proof we begin with a few observations. In the matrix A_n the off-diagonal entries are $b_i > 0$, and there are $2n-1$ variable in the matrix.

We wish to compute the Jacobian of the map $(\overline{\lambda}, \overline{q}) \to (\overline{a}, \overline{b})$. To do this we work by going through the moments.

$$m_k = \int x^k d\sigma = \sum \lambda_i^k q_i^2$$

We look at maps from both sets to $(m_1, ..., m_{2n-1})$. These are simple transformations. We will write down the appropriate matrices and hope we can find their determinants. F

Theorem 4.3 (Dumitriu, Edelman, Krishnapur, Rider, Virág, [**2, 5**]). *Let V be a potential (think convex) and $\overline{a}, \overline{b}$ are chosen from then density proportional to*

$$\exp\left(-TrV(J)\right) \prod_{k=1}^{n-1} b_{n-k}^{k\beta - 1}$$

then the eigenvalues have distribution

$$f(\lambda_1, .., \lambda_n) = \frac{1}{Z} \exp\left(-\sum_i V(\lambda_i)\right) \prod_{i<j} |\lambda_i - \lambda_j|^\beta$$

and the q_i are independent of the λ with $(q_1, ..., q_n) = (\varphi_1(1)^2, ..., \varphi_n(1)^2)$ have dirichlet$(\frac{\beta}{2}, ..., \frac{\beta}{2})$ distribution.

Suppose we have the sequence $\{(a_i, b_i), i \ge 1\}$ with the distribution from the theorem. This is a Markov Chain. This holds no matter which polynomial you choose, though you might need to take bigger blocks of (a_i, b_i).

4.1. The β-Laguerre ensemble. We finish this section by introducing a few more classical random matrix models and their associated β-ensembles. Suppose that you have a rectangular matrix $n \times p$ matrix X_n with $p \geq n$ and $x_{i,j} \sim \mathcal{N}(0,1)$ all independent. The matrix

$$M_n = X_n X_n^t$$

is a symmetric matrix which may be thought of as a sample covariance matrix for a population with independent normally distributed traits. As in the case of the Gaussian ensembles we could have started with complex entries and looked instead at XX^* to form a Hermitian matrix. The eigenvalues of this matrix have distribution

$$f_{L,\beta}(\lambda_1, ..., \lambda_n) = \frac{1}{Z_{\beta,n,p}} \prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(p-n+1)-1} e^{-\frac{\beta}{2}\lambda_i} \prod_{j<k} |\lambda_j - \lambda_k|^\beta, \tag{3}$$

with $\beta = 1$. This generalizes to the β-Laguerre ensemble which is a set of points with density $f_{L,\beta}$. The matrix model M_n is part of a wider class of random matrix models called Wishart matrices. This class of models was originally introduced by Eugene Wigner in the 1920's.

As in the case of the Gaussian ensembles there is a limiting spectral measure when the eigenvalues are in the correct scale.

Theorem 4.4 (Marchenko-Pastur law). *Let $\lambda_1, ..., \lambda_n$ have β-Laguerre distribution,*

$$\nu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i/n},$$

and suppose that $\frac{n}{p} \to \gamma \in (0,1]$. Then as $n \to \infty$

$$\nu_n \Rightarrow \sigma_{mp}, \qquad \textit{where} \qquad \frac{d\sigma_{mp}}{dx} = \rho_{mp}(x) = \frac{\sqrt{(\gamma_+ - x)(x - \gamma_-)}}{2\pi\gamma x} \mathbf{1}_{[\gamma_-,\gamma_+]}, \tag{4}$$

and $\gamma_\pm = (1 \pm \sqrt{\gamma})^2$.

Notice that this density can display different behavior at the lower end point depending on the value of γ. In particular if $\gamma = 1$ then $\gamma_- = 0$ and so the lower end of the spectrum is 0 and the density simplifies to.

$$\rho_{mp}(x) = \frac{1}{2\pi}\sqrt{\frac{4-x}{x}}\mathbf{1}_{[0,4]}.$$

In this case the lower edge has an asymptote at 0, however for any $\gamma < 1$ we get that the lower edge has the same $\sqrt{x}$ type behavior that we see at the edge of the semi-circle distribution.

As in the case of the β-Hermite ensemble there exists a tridiagonal matrix model which has joint eigenvalue distribution $f_{L,\beta}$.

Theorem 4.5. *Let*

$$L_{\beta,p} = \frac{1}{\sqrt{\beta}} \begin{bmatrix} \chi_{\beta p} & \tilde{\chi}_{\beta(n-1)} & & & \\ & \chi_{\beta(p-1)} & \tilde{\chi}_{\beta(n-2)} & & \\ & & \ddots & \ddots & \\ & & & \chi_{\beta(p-n+2)} & \tilde{\chi}_{\beta} \\ & & & & \chi_{\beta(p-n+1)} \end{bmatrix}$$

with the χ and $\tilde{\chi}$ random variables independent χ-distributed random variables with the parameter given by the subscript. Then the eigenvalues of $L_{\beta,p}L_{\beta,p}^t$ have β-Laguerre distribution give in equation (3).

4.2. Circular β-ensemble. We introduce one final classical β-ensemble. The circular orthogonal and unitary ensembles are defined by choosing an orthogonal or unitary matrix according to Haar measure on the spaces of orthogonal or unitary matrices respectively. When a matrix is chosen according to this distribution its eigenvalues fall on the circle and their joint density is given by

$$f_C(e^{i\theta_1}, ..., e^{i\theta_n}) = \frac{1}{Z} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^\beta \tag{5}$$

with $\beta = 1$ or 2 for the orthogonal or unitary ensembles respectively. A set of points on the circle that has joint density (5) is call a circular β-ensemble. This matrix model does not have a corresponding tridiagonal representation. Instead there is a corresponding matrix model called a CMV matrix which is a 5-diagonal matrix with a particular structure. This CMV matrix may be defined by looking at a sequence of numbers called Verblunsky coefficients. These coefficients will be discussed further in Chapter 6.

CHAPTER 2

The Wigner semi-circle law

1. Graph convergence

The proof of the Wigner semi-circle law given here will rest on a graph convergence argument, so we begin by introducing the notions of convergence needed for the proof and the connection between graph and spectral convergence. We will try to give examples throughout this section so the precise nature of the convergence is easier to understand.

Example 1.1. Suppose that A is the adjacency matrix of a graph G. We can label the vertex corresponding to the first coordinate vector. Then A^k_{11} is the sum over weighted paths that are length k and return to the starting vertex.

What is the connection between graph convergence and spectral measure? We will be interested in taking limits of rooted graphs. The notion of spectral measure will be useful in this context. In particular we can define the spectral measure of a rooted graph in the following way:

Definition 1.2. Let (G, ρ) be a rooted graph and A_G its adjacency matrix where ρ corresponds to the first coordinate vector. We define the spectral measure of G rooted at ρ to be the spectral measure of A_G.

Definition 1.3. Suppose you have a sequence of graphs G_n with a distinguished vertex we call the root. We say that a sequence of graphs converges if any ball around the root stabilizes at some point.

Examples 1.4. We give two examples

(1) Suppose we take the n cycle with any vertex chosen as the root. This converges to $\mathbb{Z}$ in the limit
(2) Suppose we have the k by k grid with vertices at the intersection. If we choose a vertex at the center of the grid as the root, we will get convergence to $\mathbb{Z}^2$.

Proposition 1.5. *If (G_n, ρ_n) converges to (G, ρ) is the sense of rooted local convergence, then the spectral measure of the graphs converge.*

Exercise 1.6. Consider paths of length n rooted at the left end point. This sequence converges to $\mathbb{Z}_+$ in the limit. What is the limit of the spectrum? It is the

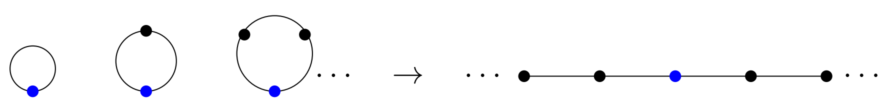

Figure 1. Rooted n-cycles to $\mathbb{Z}$

Wigner semi-circle law, since the moments are Dyck paths. But one can prove this directly, since the paths on length n are easy to diagonalize. This is an example where the spectral measure has a different limit than the eigenvalue distribution.

DEFINITION 1.7. Suppose you wish to construct a graph on n vertices with a given degree sequence $(d_1, ..., d_n)$. The configuration model may be constructed by attaching the appropriate number of half edges to the vertices and choosing a uniformly at random perfect matching of the half edges.

EXERCISE 1.8. Suppose you have the random d-regular graph on n vertices in the configuration model. Then in the limit this converges to the d-regular infinite tree in the limit.

DEFINITION 1.9. Benjamini-Schramm convergence: We define this convergence for unrooted graphs. Choose a vertex uniformly at random to be the root. If this converges in distribution in with respect to rooted convergence to a random rooted graph, we say the graphs converge in this sense.

This is essentially asking for convergence of the local statistics.

PROPOSITION 1.10. *Suppose we have a finite graph G and we choose a vertex uniformly at random. This defines a random graph and its associated random spectral measure σ. Then $E\sigma = \mu$ is the eigenvalue distribution.*

PROOF. Recall that for the spectral measure of a matrix (and so a graph) we have

$$\sigma_{(G,\rho)} = \sum_{i=1}^{n} \delta_{\lambda_i} \varphi_i^2(\rho)$$

Since φ_i is of length one, we have

$$\sum_{\rho \in V(G)} \varphi_i(\rho)^2 = 1$$

hence

$$E\sigma_{G,\rho} = \frac{1}{n} \sum_{\rho \in V(G)} \sigma_{G,\rho} = \mu_G.$$

□

EXAMPLES 1.11. The following are examples of Banjamini-Schram convergence:

(1) A cycle graph converges to the graph of $\mathbb{Z}$.
(2) A path of length n converges to the graph of $\mathbb{Z}$.
(3) Large box of Z^d converges to the full Z^d lattice.

Notice that for the last two the probability of being in a neighborhood of the edge goes to 0 and so the limiting graph doesn't see the edge effects.

EXERCISE 1.12. Say a sequence of d-regular graphs G_n with n vertices is of essentially large girth if for every k the number of k-cycles in G_n is $o(n)$. Show that G_n is essentially large girth if and only if it Benjamini-Schramm converges to the d-regular tree.

EXERCISE 1.13. (harder) Show that for $d \geq 3$ the d-regular tree is not the Benjamini-Schramm limit of finite trees.

How does this help? We will get that $\sigma_n \to \sigma_\infty$ in distribution and also in expectation. By bounded convergence, the eigenvalue distributions will converge as well: $\mu_n = E\sigma_n \to E\sigma_\infty$.

We can also consider the case where G are weighted graphs. This corresponds to general symmetric matrices A. In this case we require that the neighborhoods stabilize and the weights also converge. Everything goes through.

EXAMPLE 1.14 (The spectral measure of $\mathbb{Z}$). We use Bejamini-Schram convergence of the cycle graph to $\mathbb{Z}$. We begin by computing the spectral measure of the n-cycle G_n. We get that $A = T + T^t$ where

$$T = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 1 & & & & 0 \end{bmatrix}$$

The eigenvalues of $2T$ will be the roots of unity. We can think of this as $A = (\frac{T+T-1}{2})2$. So think of this as having the eigenvalues uniformly spaced on a circle of radius 2. The spectrum will be given by the projection of these points onto the real line. In the limit this will become the projection of the uniform measure on the circle. This gives you the arcsine distribution.

$$\sigma_{\mathbb{Z}} = \frac{1}{2\pi\sqrt{4-x^2}} \mathbf{1}_{x\in[-2,2]}\, dx$$

EXERCISE 1.15. Let B_n be the unweighted finite binary tree with n levels. Suppose a vertex is chosen uniformly at random from the set of vertices. Give the distribution of the limiting graph.

EXERCISE 1.16. Let (G, ρ) be a rooted graph, we define the spectrum of (G, ρ) to be the measure that satisfies that for all polynomials $p(x)$ we get that $\int p(x) d\mu(x) = \langle e_1, p(A)e_1 \rangle$ where we take the root to correspond to the first row and column of the matrix.

(1) Suppose a graph G_n has the $n \times n$ adjacency matrix

$$A = \begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & 0 & 1 \\ & & & 1 & 0 \end{bmatrix}$$

What is the graph with adjacency matrix A?

(2) Suppose we root G_n at the vertex corresponding to the first row and column. What is limit of (G_n, ρ_n)?

(3) Suppose we choose the root uniformly at random. What is the limit of the graph? What is the spectrum of the limiting graph? This is a distribution you have seen before.

EXERCISE 1.17. Let A be a rescaled $n \times n$ Dumitriu-Edelman tridiagonal matrix

$$A = \frac{1}{\sqrt{\beta n}} \begin{bmatrix} b_1 & a_1 & & & \\ a_1 & b_2 & a_2 & & \\ & a_2 & \ddots & \ddots & \\ & & \ddots & b_{n-1} & a_{n-1} \\ & & & a_{n-1} & b_n \end{bmatrix}, \quad a_i \sim \chi_{\beta(n-i)}, \quad b_i \sim \mathcal{N}(0,1)$$

all independent, and suppose that A is the adjacency matrix of a weighted graph.

(1) Draw the graph with adjacency matrix A. (There can be loops)
(2) Suppose a root for your graph is chosen uniformly at random, what is the limiting distribution of your graph?
(3) What is the limiting spectral measure of the graph rooted at the vertex corresponding to the first row and column?
(4) What is the limiting spectral measure of the unweighted graph?

2. Wigner semicircle law

Note that a Jacobi matrix can be thought of the adjacency matrix of a weighted path with loops.

EXERCISE 2.1. Check that $\chi_n - \sqrt{n} \xrightarrow[n\to\infty]{P} \mathcal{N}(0, 1/2)$.

PROOF 1. Take the previous graph, divide all the labels by $\sqrt{n}$ and then take a Benjamini-Schramm limit.

What is the limit? $\mathbb{Z}$, but then we need labels. The labels in a randomly rooted neighborhood will now be the square root of a single uniform random variable in $[0, 1]$. Call this U. Then the edge weights are $\sqrt{U}$. Recall that $\mu_n \to E\sigma$. In the case where U was fixed we would just get a scaled arcsine measure.

Choose a point uniformly on the circle of radius $\sqrt{U}$ and project it down to the real line. But this point is in fact a uniform random point in the disk. This gives us the semicircle law.

$$\mu_{sc} = \frac{1}{2\pi}\sqrt{4 - x^2}\,1_{x\in[-2,2]}\,dx.$$

□

PROOF 2. We take the rooted limit where we choose the root to be the matrix $J/\sqrt{n}$ corresponding to the first coordinate vector.

In the limit this graph converges to $\mathbb{Z}_+$. Therefore $\sigma_n \to \sigma_{\mathbb{Z}_+} = \rho_{sc}$. This convergence is weak convergence in probability.

So going back to the full matrix model of GOE, we see that the spectral measure at an arbitrary root converges weakly in probability to μ_{sc}. But then this must hold also if we average the spectral measures over the choice of root (but not the randomness in the matrix).

Thus we get $\mu_n \to \mu_{sc}$ in probability. □

Dilemma: The limit of the spectral measure should have nothing to do with the limit of the eigenvalue distribution in the general case. This tells you that the Jacobi matrices that we get in the case of the GOE are very special.

CHAPTER 3

The top eigenvalue and the Baik-Ben Arous-Pechet transition

1. The top eigenvalue

The eigenvalue distribution of the GOE converges after scaling by $\sqrt{n}$ to the Wigner semi-circle law. From this, it follows that the top eigenvalue, $\lambda_1(n)$ satisfies for every $\varepsilon > 0$

$$P(\lambda_1(n)/\sqrt{n} > 2 - \varepsilon) \to 1$$

the 2 here is the top of the support of the semicircle law. However, the upper bound does not follow and needs more work. This is the content of the following theorem.

Theorem 1.1 (Füredi-Komlós).

$$\frac{\lambda_1(n)}{\sqrt{n}} \to 2 \textit{ in probability.}$$

This holds for more general Wigner matrices; we have a simple proof for the GOE case.

Lemma 1.2. *If J is a Jacobi matrix (a's diagonal, b's off-diagonal) then*

$$\lambda_1(J) \le \max_i (a_i + b_i + b_{i-1})$$

Here we take the convention $b_0 = b_n = 0$.

Proof. Observe that J may be written as

$$J = -AA^T + \operatorname{diag}(a_i + b_i + b_{i-1})$$

where

$$A = \begin{bmatrix} 0 & \sqrt{b_1} & & & \\ & -\sqrt{b_1} & \sqrt{b_2} & & \\ & & -\sqrt{b_2} & \sqrt{b_3} & \\ & & & \ddots & \ddots \end{bmatrix}$$

and AA^t is nonnegative definite. So for the top eigenvalues we have

$$\lambda_1(J) \le -\lambda_1(AA^T) + \lambda_1(\operatorname{diag}(a_i + b_i + b_{i-1})) \le \max_i (a_i + b_i + b_{i-1}).$$

We used sublinearity of λ_1, which follows from the Rayleigh quotient representation. $\square$

If we apply this to our setting we get that

$$\lambda_1(GOE) \le \max_i (N_i, \chi_{n-i} + \chi_{n-i+1}) \le 2\sqrt{n} + c\sqrt{\log n}$$

the right inequality is an exercise (using the Gaussian tails in χ) and holds with probability tending to 1 if c is large enough. This completes the proof of Theorem 1.1.

This gives you that the top eigenvalue cannot go further than an extra $\log n$ outside of the spectrum. Indeed we have that

$$\lambda_1(GOE) = 2\sqrt{n} + TW_1 n^{-1/6} + o(n^{-1/6})$$

so this result from before is not optimal.

2. Baik-Ben Arous-Pechet transition

Historically one of the areas that random matrices have been used is to study correlations. When considering principle component analysis we use the Wishart model to study what would happen if we just had noise and nothing else? How likely is it that the correlations that we see are more than chance?

Wishart consider matrices $X_{n\times m}$ with independent entries and studies the eigenvalues of XX^t. We will take the noise matrix X to have i.i.d. normal entries. We study rank one perturbations and consider the top eigenvalue. Consider the case $n = m$.

THEOREM 2.1 (BBP transition).

$$\frac{1}{n}\lambda_1\Big(X\, diag(1+a^2, 1, 1, ..., 1)X^t\Big) \to \varphi(a)^2$$

where

$$\varphi(a) = \begin{cases} 2 & a \le 1 \\ a + \frac{1}{a} & a \ge 1 \end{cases}$$

Heuristically, correlation in the populations appears in the asymptotics in the top eigenvalue of the data only if it is sufficiently large, $a > 1$. Otherwise, it gets washed out by the fake correlations coming from noise.

We will prove the GOE analouge of this theorem, and leave the Wishart case as an exercise.

One can also study the fluctuations of the eigenvalues. In the case $a < 1$ you get Tracy-Widom. In the case $a > 1$ you get Gaussian fluctuations. Very close to th point $a = 1$ you get a deformed Tracy-Widom.

The GOE analogue is the following.

THEOREM 2.2 (Top eigenvalue of GOE with nontrivial mean).

$$\lambda_1(\mathrm{GOE}_n + \frac{a}{\sqrt{n}} 11^t) \to \varphi(a)$$

where 1 *is the all-1 vector, and* 11^t *is the all-1 matrix.*

It may be suprising how little change in the mean in fact changes the top eigenvalue!

We will not use the following theorem, but will include it only to show where the function ϕ comes from. It will also motivate the proof for the GOE case.

THEOREM 2.3.

$$\lambda_1(\mathbb{Z}^+ + \text{ loop of weight } a \text{ on } 0) = \varphi(a)$$

We will leave this as an exercise for the reader.

HEURISTICS. The spectrum of $\mathbb{Z}^+$ is absolutely continuous. This doesn't change unless a is big enough. Why? there is an eigenvector $(1, a^{-1}, a^{-2}, ...)$ which is only in ℓ^2 when $a > 1$. Moreover this has eigenvalue $a + \frac{1}{a}$. □

PROOF OF GOE CASE. The first observation is that because the GOE is an invariant ensemble, we can replace 11^t by vv^t for any vector v having the same length as the vector 1. We can replace the perturbation with $\sqrt{n} a e_1 e_1^t$. Such perturbations commute with tridiagonalization.

Therefore we can consider Jacobi matrices of the form

$$J(a) = \frac{1}{\sqrt{n}} \begin{bmatrix} a\sqrt{n} + N_1 & \chi_{n-1} & & \\ \chi_{n-1} & N_2 & \chi_{n-2} & \\ & \ddots & \ddots & \ddots \end{bmatrix}$$

Case 1: $a \leq 1$. Since the perturbation is positive, we only need an upper bound. We use the maximum bound from before. For $i = 1$, the first entry, there was a room of size $\sqrt{n}$. For $i = 1$ the max bound still holds.

Case 2: $a > 1$

Now fix k and let $v = (1, 1/a, 1/a^2, ..., 1/a^k, 0, ..., 0)$. We get that the error from the noise will be of order $1/\sqrt{n}$ so that

$$\left\| \frac{J(a)}{v} - v(a + \frac{1}{a}) \right\| \leq ca^{-k}$$

with probability tending to 1.

Now if for a symmetric matrix A and a vector v of length at least 1 we have $\|Av - xv\| < \varepsilon$, then A has an eigenvalue ε-close to x.

Thus $J(a)$ has an eigenvalue λ^* that is ca^{-k}-close to $a + 1/a$.

We now need to check that this eigenvalue will actually be the maximum.

LEMMA 2.4. *Consider adding a positive rank 1 perturbation to a symmetric matrix. Then the eigenvalues of the two matrices will interlace and the shift under perturbation will be to the right.*

By interlacing,

$$\lambda_2(J(a)) \leq \lambda_1(J) = 2 + o(1) < a + 1/a - ca^k$$

if we chose k large enough. Thus the eigenvalue λ^* we identified must be λ_1. □

EXERCISE 2.5. Suppose you have $\mathbb{Z}^+$ with a loop of weight a on the end and let $\lambda_1(a)$ be the largest eigenvalue the graph. We can think of this as a rank 1 positive definite pertubation of $\mathbb{Z}^+$. By interlacing this means that the largest eigenvalue will shift to the right under the perturbation. Prove the following lower bound:

$$\lambda_1(a) \geq \begin{cases} 2 & a \leq 1 \\ a + \frac{1}{a} & a > 1 \end{cases} \tag{6}$$

CHAPTER 4

The soft edge

1. The heuristic convergence argument

The distribution of top eigenvalue of the β-Hermite ensemble has nice closed formulas in the case of $\beta = 1, 2, 4$. The goal here is to give the characterization of the top eigenvalue for general β. To do this we look at the geometric structure of the tridiagonal matrix.

Simulations show that the eigenvectors corresponding to the top eigenvalues of the matrix tend to be supported in the first $o(n)$ coordinates. This suggests that we in some sense need to look at the top corner of the matrix in order to see the behavior of the top eigenvalue.

Now suppose that you have the matrix

$$A = \begin{bmatrix} 0 & 1 & & \\ 1 & 0 & 1 & \\ & 1 & 0 & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

We look at the matrix $A - 2I$, and we look at the action of this matrix on the first order m entries of a vector, where we guess that $m = n^\alpha$, with α to be determined later. Let $f : \mathbb{R}^+ \to \mathbb{R}$ and $v_f = (f(0), f(1/m), f(2/m), ..., f(n/m))^t$. Then $B = m^2(A - 2I)$ acts as a discrete second derivative on f, in the sense that $Bv_f \approx v_{f''}$.

Returning to the β-Hermite case, by Exercise 2.1, for $k \ll n$ we have

$$\chi_{n-k} \approx \sqrt{\beta(n-k)} + \mathcal{N}(0, 1/2) \approx \sqrt{\beta}(\sqrt{n} - \frac{k}{2\sqrt{n}}) + \mathcal{N}(0, 1/2)$$

Now we consider the matrix

$$m^\gamma(2\sqrt{n}I - J) \approx$$

$$m^\gamma\sqrt{n}\begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 0 & \ddots \\ & & \ddots & \ddots \end{bmatrix} + \frac{m^\gamma}{2\sqrt{n}}\begin{bmatrix} 0 & 1 & & & \\ 1 & 0 & 2 & & \\ & 2 & 0 & 3 & \\ & & 3 & 0 & \ddots \\ & & & \ddots & \ddots \end{bmatrix} + \frac{m^\gamma}{\sqrt{\beta}}\begin{bmatrix} N_1 & \tilde{N}_1 & & \\ \tilde{N}_1 & N_2 & \tilde{N}_2 & \\ & \tilde{N}_2 & N_3 & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

Now assume that we have $m = n^\alpha$ for some α. What choice of α should we make? For the first term we want

$$m^\gamma \sqrt{n} \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 0 & \ddots \\ & & \ddots & \ddots \end{bmatrix}$$

to behave like a second derivative. This means that $m^\gamma \sqrt{n} = m^2$ which gives $2\alpha = \alpha\gamma + 1/2$. We do a similar analysis on the second term. We want this term to behave like multiplication by t. for this we want $\frac{m^\gamma}{\sqrt{n}} = \frac{1}{m}$ which gives $\alpha\gamma - 1/2 = -\alpha$. Solving this system we get $\alpha = 1/3$ and $\gamma = 1/2$. For the noise term, multiplication by it should yield a distribution (in the Schwarz sense), which means that its integral over intervals should be of order 1. In other words, the average of m noise terms times m^γ should be of order 1. This gives $\gamma = 1/2$, consistent with the previous computations.

This means that we need to look at the section of the matrix that is $m = n^{1/3}$ and we rescale by $n^{1/6}$. That is we look at the matrix

$$n^{1/6}(2\sqrt{n} - J_n)$$

acting on functions with mesh size $n^{-1/3}$.

Exercise 1.1. Show that in this scaling, the second matrix in the expansion above has the same limit as the diagonal matrix with $0, 2, 4, 5, 6, ...$ on the diagonal (scaled the same way).

The final piece that we need to check is what happens to the final matrix with noise terms. We may use a similar trick to the previous exercise in order to consider the matrix given by

$$W_n = \begin{bmatrix} N_1 & & & \\ & N_2 + 2\tilde{N}_1 & & \\ & & N_3 + 2\tilde{N}_2 & \\ & & & \ddots \end{bmatrix}$$

With the exception of the first entry all the of the diagonal entries are now independent $2\mathcal{N}(0,1)$. In order to consider the action of W_n on a vector v_f we need to look at the 'integrated version' via Riemann sums. In particular we consider

$$\frac{m^{1/2}}{\sqrt{\beta}} \sum_{k=1}^{\lfloor xm \rfloor} \frac{1}{m} (W_n v_f)_k = \frac{m^{-1/2}}{\sqrt{\beta}} \sum_{k=1}^{\lfloor xm \rfloor} W_{kk} f(k/m).$$

and we consider the entries of W_n to be the increments of some process $Y(x)$ with $Y(\frac{k}{m}) - Y(\frac{k-1}{m}) = \frac{m^{-1/2}}{2} W_k k$. Then we may write

$$\begin{aligned}
\frac{m^{-1/2}}{\sqrt{\beta}} \sum_{k=1}^{\lfloor xm \rfloor} (W_n v_f)_k &= \frac{2}{\sqrt{\beta}} \sum_{k=1}^{\lfloor xm \rfloor} f(\tfrac{k}{m})\big(Y(\tfrac{k}{m}) - Y(\tfrac{k-1}{m})\big) \\
&\approx \frac{2}{\sqrt{\beta}} \sum_{k=1}^{\lfloor xm \rfloor} \big(f(\tfrac{k}{m}) - f(\tfrac{k-1}{m})\big) Y(\tfrac{k}{m}) \\
&\approx \frac{2}{\sqrt{\beta}} \frac{1}{m} \sum_{k=1}^{\lfloor xm \rfloor} f'(\tfrac{k}{m}) Y(\tfrac{k}{m}) \Rightarrow \frac{2}{\sqrt{\beta}} \int_0^x f'(s) b_s ds
\end{aligned}$$

We've integrated and done an integration by parts here, so the noise term should actually give a contribution of $\frac{2}{\sqrt{\beta}} b_x'$ in the limit.

Conclusion. This matrix acting on functions with this mesh size behaves like a differential operator. That is

$$H_n = n^{1/6}(2\sqrt{n} - J_n) \approx -\partial_x^2 + x + \frac{2}{\sqrt{\beta}} b_x' = \mathrm{SAO}_\beta$$

here b_x' is a white noise. This operator will be called the Stochastic Airy operator (SAO_β). We also set the boundary condition to be Dirichlet. This conclusion can be made precise.

There are two problems at this point that must be overcome in order to make this convergence rigorous. The first is that we need to actually be able to make sense of that limiting operator. The second is that the matrix even embedded an an operator on step functions acts on a different space that the SAO_β so we need to make sense of what the convergence statement should actually be. The following are the main ideas behind the operator convergence.

Ideas on the operator convergence.

(1) Embed $\mathbb{R}^n$ into $L^2(\mathbb{R})$ via

$$e_i \mapsto \sqrt{m} \mathbf{1}_{[\frac{i-1}{m}, \frac{i}{m})}.$$

This gives an embedding of the matrix J acting on a subspace of $L^2(\mathbb{R}^+)$.

(2) It is not clear what functions the Stochastic Airy Operator acts on at this point. Certainly nice functions multiplied by the derivative of Brownian motion will not be functions, but distributions. The only way we get nice functions as results if this is cancelled out by the second derivative. Nevertheless, the domain of SAO_β can be defined.

In any case, these operators act on two completely different sets of functions. The matrix acts on piecewise constant functions, while SAO_β acts on some exotic functions.

(3) The nice thing is that if there are no zero eigenvalues, both H_n^{-1} and J^{-1} can be defined in their own domains, and the resulting operators have compact extensions to the entire L^2.

The sense of convergence we have is

$$\|H_n^{-1} - A_\beta^{-1}\|_{2\to 2} \to 0.$$

This is called norm resolvent convergence, and it implies convergence of eigenvalues and eigenvectors if the limit has discrete simple spectrum.

(4) The simplest way to deal with the limiting operator and the issues of white noise is to think of it as a bilinear form. This is the approach we follow in the next section. The kth eigenvalue can be identified using the Courant-Fisher characterization.

Exercise 1.2. We will consider cases where a matrix $A_{n\times n}$ can be embedded as an operator acting on the space of step function with mesh size $1/m_n$. In particular we can encode these step functions in to vectors $v_f = [f(\frac{1}{m_n}), f(\frac{2}{m_n}), ..., f(\frac{n}{m_n})]^t$.

Let A be the matrix

$$A = \begin{bmatrix} -1 & 1 & & \\ & -1 & 1 & \\ & & \ddots & \ddots \\ & & & -1 \end{bmatrix}$$

For which k_n to we get $k_n A v_f \to f'$?

Exercise 1.3. Let A be the diagonal matrix with diagonal entries $(1, 4..., n^2)$. Find a k_n such that $k_n A v_f$ converges to something nontrivial in the limit. What is k_n and what does the limit converge to?

Exercise 1.4. Let J be a Jacobi matrix (tridiagonal with positive off-diagonal entries) and v be an eigenvector with eigenvalue λ. The number of times that v changes sign is equal to the number of eigenvalues above λ. More generally the equation $Jv = \lambda v$ determines a recurrence for the entries of v. If we run this recurrence for an arbitrary λ (not necessarily an eigenvalue) and count the number of times that v changes sign this still gives the number of eigenvalues greater than λ.

(1) Based on this gives a description of the number of eigenvalues in the interval $[a, b]$.
(2) Suppose that $v^t = (v_1, ..., v_n)$ solves the recurrence defined by $Jv = \lambda v$. What is the recurrence for $r_k = v_{k+1}/v_k$? What are the boundary conditions for r that would make v and eigenvector?

2. The bilinear form and making sense of the SAO$_\beta$

Recall the Airy operator

$$Af = -\partial_x^2 + xf$$

acting on $L^2(\mathbb{R}^+)$ with boundary condition $f(0) = 0$. The equation $Af = 0$ has two solutions $\mathrm{Ai}(x)$ and $\mathrm{Bi}(x)$, called Airy functions. Note that the solution of $(A - \lambda)f = 0$ is just a shift of these functions by λ.

Since only Ai^2 is integrable, the eigenfunctions A are the shifts of Ai with the eigenvalues the amount of the shift. We know that the kth zero of the Ai function is at $z_k = -\left(\frac{3}{2}\pi k\right)^{2/3} + o(1)$, therefore to satisfy the boundary conditions the shift must place a 0 at 0, so the kth eigenvalue is given by

$$\lambda_k = -z_k = \left(\frac{3}{2}\pi k\right)^{2/3} + o(1) \tag{7}$$

The asymptotics are classical.

For the Airy operator A and a.e. differentiable, continuous functions f with $f(0) = 0$ we can define

$$\|f\|_*^2 := \langle Af, f\rangle = \int_0^\infty f^2(x)x + f'(x)^2\, dx.$$

Let L^* be the space of such functions with $\|f\|_* < \infty$

EXERCISE 2.1. Show that there is $c > 0$ so that

$$\|f\|_2 \le c\|f\|_*$$

for every $f \in L^*$. In particular, $L^* \subset L^2$.

Recall the Rayleigh quotient characterization of the eigenvalues λ_1 of A.

$$\lambda_1 = \inf_{f\in L^*, \|f\|_2=1} \langle Af, f\rangle.$$

More generally, the Courant-Fisher characterization is

$$\lambda_k = \inf_{B\subset L^*, \dim B=k}\ \sup_{f\in B, \|f\|_2=1} \langle Af, f\rangle$$

where the infimum is over subspaces B.

For two operators we say $A \le B$ if any $f \in L^*$

$$\langle f, Af\rangle \le \langle f, Bf\rangle.$$

EXERCISE 2.2. If $A \le B$, then $\lambda_k(A) \le \lambda_k(B)$.

Our next goal is to define the bilinear form associated with the Stochastic Airy operator on funcitons in L^*. Clearly, the only missing part is to define

$$\int_0^\infty f^2(x)b'(x)\, dx.$$

At this point you could say that this is defined in terms of stochastic integration, but the standard L^2 theory is not strong enough – we need it to be defined in the almost sure sense for all functions in L^*. We could define it in the following way:

$$\langle f, b'f\rangle \text{``}=\text{''} \int_0^\infty f^2(x)b'(x)dx = -\int_0^\infty 2f'(x)f(x)b(x)dx$$

This is now a perfectly fine integral, but it may not converge. The main idea will be to write b as its average together with an extra term.

$$b(x) = \int_x^{x+1} b(s)ds + \tilde{b}(x) = \bar{b}(x) + \tilde{b}(x)$$

In this decomposition we will get that $\bar{b}$ is differentiable and $\tilde{b}$ is small. The averaging term decouples quickly (at time intervals of length 1), so this term is analogous to a sequence of i.i.d. random variables. We define the inner product in terms of this decomposition

$$\langle f, b'f\rangle := \int_0^\infty f^2(x)\bar{b}'(x)dx - 2\int_0^\infty f'(x)f(x)\tilde{b}(x)dx.$$

This is well defined because we have the following bounds:

EXERCISE 2.3. There exists a random constant C so that we have the following inequality of functions:

$$|\bar{b}'|, |\tilde{b}| \leq C\sqrt{\log(2+x)} \tag{8}$$

Now we return to the Stochastic Airy operator, the following lemma with give us that the operator is bounded from below.

LEMMA 2.4. *We have*

$$-CI + (1-\varepsilon)A \leq \mathrm{SAO}_\beta \leq (1+\varepsilon)A + CI$$

for some random constant C. This may also be written as

$$-C\|f\|_2^2 + (1-\varepsilon)\|f\|_*^2 \leq \langle f, \mathrm{SAO}_\beta f\rangle \leq (1+\varepsilon)\|f\|_*^2 + C\|f\|_2^2 \tag{9}$$

The upper bound here implies that the bilinear form is defined for all functions $f \in L^*$. The lemma will follow from the fact that

$$b' \leq \varepsilon(1+\varepsilon)A + cI,$$

in the sense of operators.

PROPOSITION 2.5. *The eigenvalues of* SAO_β *admit a variational characterization.*

PROOF FOR Λ_0. Let $\tilde{\Lambda}_0 = \inf\{\langle f, \mathrm{SAO}_\beta\rangle : f \in L^*, \|f\|_2 = 1\}$, then we may choose a minimizing sequence $f_n \in L^*, \|f_n\|_2 = 1$ with $\langle f_n, \mathrm{SAO}_\beta\rangle \to \tilde{\Lambda}_0$. From equation (9) we get that the f_n may be chosen so that $\|f_n\|_* \leq B$ for some constant B (a function of the brownian path). Therefore we can choose a subsequence which converges to some function f_0 uniformly on compact subsets, in L^2, and weakly in H^1. Using bounds on $\tilde{b}$ and $\bar{b}$ from equation (8) we get that for $\varepsilon > 0$ we may choose X so that

$$\langle f_n, \mathrm{SAO}_\beta f_n\rangle = \|f\|_*^2 - \|f\|_2^2 + \frac{4}{\sqrt{\beta}}\int_0^X f(x)f'(x)b'(x)dx + \mathcal{E}$$

where $|\mathcal{E}| \leq \varepsilon\|f\|_*^2$. Taking the limit as $n \to \infty$ we see that the middle two terms converge and $\|f_0\|_*^2 \leq \liminf \|f_n\|_*^2$. Taking $\varepsilon \to 0$ we get that $\langle f_0, \mathrm{SAO}_\beta f\rangle \leq \tilde{\Lambda}_0$. The definition of $\tilde{\Lambda}_0$ gives the opposite inequality. Therefore we have that $\langle f_0, \mathrm{SAO}_\beta f\rangle = \tilde{\Lambda}_0$ meaning that the infimum is achieved at some function f_0. It remains to be shown that f_0 is an eigenfunction, that is, for any test function φ we have $\langle \varphi, \mathrm{SAO}_\beta f_0\rangle = \Lambda\langle \varphi, f_0\rangle$.

Look at

$$f_\epsilon = \frac{f_0 + \epsilon\varphi}{\|f_0 + \epsilon\varphi\|_2}$$

differentiating in ϵ we find

$$\frac{d}{d\epsilon}\langle f, \mathrm{SAO}_\beta f_\epsilon\rangle|_{\epsilon=0}$$

This means that we have a minimum at $\epsilon = 0$, and so the directional derivative in the direction of φ is 0. This shows that (f_0, Λ_0) is an eigenfunction/eigenvalue pair for SAO_β.

□

EXERCISE 2.6. Prove the bounds in the previous statement, and the following

$$|\bar{b}'| \le C_\varepsilon + \varepsilon x, \qquad |\tilde{b}|^2 \le C_\varepsilon + \varepsilon x.$$

Using these bound we get the following bound:

$$\begin{aligned}|\langle f, b'f\rangle| &\le C_\varepsilon \|f\|_2^2 + \varepsilon\langle f, Af\rangle + \varepsilon \int (f')^2 dx + \frac{1}{\varepsilon}\int f^2(C_{\varepsilon^2} + \varepsilon^2 x)dx \\ &\le C'_\varepsilon \|f\|_2^2 + 2\varepsilon\langle f, Af\rangle\end{aligned}$$

COROLLARY 2.7. *The eigenvalues of SAO_β satisfy*

$$\frac{\lambda_k^\beta}{k^{2/3}} \to \left(\frac{2\pi}{3}\right)^{2/3} \qquad a.s.$$

PROOF. It suffices to show that a.s. for every rational $\varepsilon > 0$ there exists $C_\varepsilon > 0$ so that

$$(1-\varepsilon)\lambda_k - C_\varepsilon \le \lambda_k^\beta \le (1+\varepsilon)\lambda_k + C_\varepsilon$$

where the λ_k are the Airy eigenvalues (7). But this follows from the operator inequality of Lemma 2.4 and Exercise 2.2. □

If you look at the empirical distribution of the eigenvalues as $k \to \infty$ then the "density" behaves like $\sqrt{\lambda}$. More precisely, the number of eigenvalues less then λ is of order $\lambda^{3/2}$. This is the Airy-β version of the Wigner semicircle law. We only see the edge of the semicircle here.

3. The discrete bilinear form and limits

We have characterized the eigenvalues of the limiting distribution by a variational problem. In order to handle the convergence question we use the same formulation for the finite n problem. We will discuss the discrete to continuous convergence in a slightly simplified case. Assume that H_n has the form

$$(H_n v)_k = n^{2/3}(-v_{k-1} + 2v_k - v_{k+1}) + \frac{k}{n^{1/3}} v_k + n^{1/6}\frac{2}{\sqrt{\beta}} N_k v_k,$$

where $N_k \sim \mathcal{N}(0,1)$ are independent and we take $v_0 = v_{n+1} = 0$. Then the bilinear form in $\mathbb{R}^n$ is given by

$$\begin{aligned}\langle v, H_n v\rangle &\approx n^{2/3}\sum_{k=1}^n v_k(-v_{k+1} + 2v_k - v_{k-1}) + \frac{1}{n^{1/3}}\sum_{k=1}^n k v_k^2 + \frac{2n^{1/6}}{\sqrt{\beta}}\sum_{k=1}^n v_k^2 N_k \\ &\approx n^{2/3}\sum_{k=1}^n (v_k - v_{k-1})^2 + \frac{1}{n^{1/3}}\sum_{k=1}^n k v_k^2 + \frac{2n^{1/6}}{\sqrt{\beta}}\sum_{k=1}^n v_k^2 N_k.\end{aligned}$$

This is the traditional inner product on $\mathbb{R}^n$. In order to consider the inner product with respect to step functions in L^2 we use the embedding

$$e_i \mapsto n^{1/6}\mathbf{1}_{[\frac{i-1}{n^{1/3}}, \frac{i}{n^{1/3}})},$$

and so we need to multiply the existing bilinear form by $n^{-1/3}$. We define the new bilinear form on step functions in L^2 given by

$$\begin{aligned}\langle f, H_n f\rangle_n &= \frac{1}{n^{1/3}}\langle v_f, H_n v_f\rangle \\ &= n^{1/3}\sum_{k=1}^{n}(f(\tfrac{k-1}{n^{1/3}}) - f(\tfrac{k-2}{n^{1/3}}))^2 + \frac{1}{n^{2/3}}\sum_{k=1}^{n} k f^2(\tfrac{k-1}{n^{1/3}}) + \frac{2n^{-1/6}}{\sqrt{\beta}}\sum_{k=1}^{n} f^2(\tfrac{k-1}{n^{1/3}})N_k.\end{aligned}$$

and the norm defined by

$$\|f\|_{*,n}^2 = n^{1/3}\sum(f(\tfrac{k-1}{n^{1/3}}) - f(\tfrac{k-2}{n^{1/3}}))^2 + n^{-1/3}\sum f^2(\tfrac{k-1}{n^{1/3}}).$$

Using this norm and the standard L^2 we can get bounds on the bilinear form similar to those in the limiting case.

LEMMA 3.1. *For some random constant C almost surely we get*

$$-C\|f\|_2^2 + (1-\varepsilon)\|f\|_{*,n}^2 \le \langle f, H_n f\rangle_n \le (1+\varepsilon)\|f\|_{*,n}^2 + C\|f\|_2^2. \tag{10}$$

Lastly we introduce the projection operator P_n which projects any function f onto the step function defined by the v_f introduced earlier.

Now that we have the L^2 framework for the finite n case we can discuss the limits. We start by showing the following convergence statements:

$$n^{-1/3}\sum_{k=1}^{\lfloor n^{1/3}x\rfloor}\frac{k}{n^{1/3}} \Rightarrow \frac{x^2}{2}$$

$$n^{-1/6}\sum_{k=1}^{\lfloor n^{1/3}x\rfloor} N_k \Rightarrow b(x).$$

These together with some tightness conditions will be enough to show that for $\varphi, f \in L^*$ we get

$$\langle P_n\varphi, H_n P_n f\rangle_n \to \langle \varphi, \mathrm{SAO}_\beta f\rangle.$$

LEMMA 3.2. *Suppose we have a sequence $f_n \in L^{*,n}$ with $\|f_n\|_{*,n} < K$ and $\|f_n\|_2 = 1$, then there exists $f \in L^*$ and subsequence n_k such that*

1. *$f_n \to f$ weakly in L^2 and $n^{1/3}\left(f_n(x+n^{1/3}) - f_n(x)\right) \to f'$ weakly in L^2.*
2. *$f_{n_k} \to f$ in L^2.*
3. *For any $\varphi \in C_0^\infty$, $\langle P_n\varphi, H_n f_{n_k}\rangle \to \langle \varphi, \mathrm{SAO}_\beta f\rangle$*

Once this lemma is established we can turn to the actual eigenvalue convergence.

LEMMA 3.3. *Let $\lambda_{k,n}$ denote the kth smallest eigenvalue of H_n, then for $k \ge 1$ we have $\underline{\lambda}_k = \liminf \lambda_{k,n} \ge \Lambda_k$.*

PROOF FOR $k = 1$. Assume $\underline{\lambda}_1 < \infty$. Since the eigenvalues of H_n are uniformly bounded from below we can find a subsequence for which $\lambda_{1,n_j} \to \underline{\lambda}_1$. The corresponding eigenfunctions have L_n^* uniformly bounded and so there is a further subsequence along which the L^2 limit of the eigenfunctions exists. Call this function f_1. This gives us that

$$\underline{\lambda}_1\langle\varphi, f_1\rangle = \lim_{j\to\infty}\lambda_{1,n_j}\langle P_{n_j}\varphi, f_{1,n_j}\rangle = \lim_{j\to\infty}\langle P_{n_j}\varphi, H_{n_j} f_{1,n_j}\rangle = \langle\varphi, \mathrm{SAO}_\beta f_1\rangle$$

which gives us that f_1 is an eigenfunction with eigenvalue $\underline{\lambda}_1$. By definition this gives us that $\underline{\lambda}_1 \geq \Lambda_1$. □

LEMMA 3.4. *For $k \geq 1$ we have $\lambda_{k,n} \to \Lambda_k$ and $f_{n,k} \to_{L^2} f_k$.*

PROOF FOR $k = 1$. Let f_1^ε be in an ε neighborhood of f_1 in the L^* norm, where f_1 is the eigenfunction of SAO_β with eigenvalue Λ_1. Then define $g_{1,n} = P_n f_1$ by using the variational characterization of $\lambda_{k,n}$ we get that

$$\limsup \lambda_{k,n} \leq \limsup_{n\to\infty} \frac{\langle g_{1,n}, H_n g_{1,n}\rangle}{\langle g_{1,n}, g_{1,n}\rangle} = \Lambda_1 + O(\varepsilon).$$

Now recall that lemma 3.2 implies that any subsequence of $f_{1,n}$ has a further subsequence converging in L^2 to some function f. We now know that g must satisfy $\mathrm{SAO}_\beta g = \lambda_1 g$ therefore $g = f_1$ and so $f_{1,n} \to_{L^2} g$.

□

4. Tails of the Tracy Widom$_\beta$ distribution

DEFINITION 4.1. We define the Tracy-Widom-β distribution

$$TW_\beta = -\lambda_1(A_\beta)$$

In the case $\beta = 1, 2$ this is consistent with the classical distribution.

If we look at the distribution of the Tracy-Widom$_\beta$ distribution it will look at the distribution of the tails we will get the the right tail is approximately $\exp(-\frac{2}{3}\beta a^{3/2})$ and on the left tail we get $\exp(-\frac{\beta+o(1)}{24}a^3)$. We will prove the left tail.

THEOREM 4.2.

$$P(\mathrm{TW}_\beta < -a) = \exp(-\frac{\beta + o(1)}{24}a^3) \qquad as \quad a \to \infty$$

PROOF OF THE UPPER BOUND. Suppose we have $\lambda_1 > a$, this implies that for all f we get the bound

$$\langle f, A_\beta f\rangle \geq a\|f\|_2^2.$$

Therefore we are interested in the probability

$$P\left(\|f'\|_2^2 + \|\sqrt{x}f\|_2^2 + \frac{2}{\sqrt{\beta}}\int f^2 b' dx \geq a\|f\|_2^2\right) =?$$

Notice that the first two terms are deterministic, but we can study the distribution of the third term. In particular understood in the sense of Itô we get that

$$\frac{2}{\sqrt{\beta}}\int f^2 b' dx \stackrel{d}{=} \frac{2}{\sqrt{\beta}}\int f^2(x) db_x$$

The variance of this term will be given by the quadratic variation which is $\frac{4}{\beta}(\int f^4 dx) = \frac{4}{\beta}\|f\|_4^4$. This leads us to computing

$$P\left(\|f'\|_2^2 + \|\sqrt{x}f\|_2^2 + N\|f\|_4^2 \geq a\|f\|_2^2\right) =?$$

where N is a normal random variable with variance $4/\beta$. Using the standard tail bound for a normal random variable we get

(11)

$$P\left(\|f'\|_2^2 + \|\sqrt{x}f\|_2^2 + N\|f\|_4^2 \geq a\|f\|_2^2\right) \leq 2\exp\left(-\frac{\beta(a\|f\|_2^2 - \|f'\|_2^2 - \|f\sqrt{x}\|_2^2)^2}{8\|f\|_4^4}\right)$$

We want to optimize over possible choices of f. It turns out the optimal f will have small derivative, so we will drop the derivative term and then optimize the remaining terms. That is we wish to maximize.

$$\frac{(a\|f\|_2^2 - \|f\sqrt{x}\|_2^2)^2}{\|f\|_4^4}$$

With some work we can show that the optimal function will be approximately $f(x) \approx \sqrt{(a-x)^+}$. This needs to be modified a bit in order to keep the derivative small, so we cut this function off and replace is with a linear piece

$$f(x) = \sqrt{(a-x)^+} \wedge (a-x)^+ \wedge x\sqrt{a}$$

We can check that

$$a\|f\|_2^2 \sim \frac{a^3}{2} \qquad \|f\| \sim O(a) \qquad \|\sqrt{x}f\|^2 \sim \frac{a^3}{6} \qquad \|f\|_4^4 \sim \frac{a^3}{3}.$$

Using these values in equation (11) give us the correct upper bound.

□

Proof of the lower bound. We begin by introducing the Riccati transform: Suppose we have an operator

$$L = -\partial_{xx} + V(x)$$

then the eigenvalue equation is

$$\lambda f = (-\partial_{xx} + V(x))f$$

We can pick a λ and attempt to solve this equation. The left boundary condition is given, so you can check if the solution satisfies $f \in L^*$, which would give you an eigenfunction. Most of the time this won't be true, but we can still gain information by studying these solutions. To study this problem we first make the transformation

$$p = \frac{f'}{f}, \qquad \text{which gives} \qquad p' = V(x) - \lambda - p^2, \qquad p(0) = \infty.$$

Under this transformation we have the following:

Proposition 4.3. *Choose λ, we will have $\lambda < \lambda_1$ if an only if the solution to the Ricatti equation does not blow up.*

We can draw the slope field: If we take the case $V(x) = 0$ and $\lambda = 0$ we will get that there is aright facing parabola $p^2 = x$ where the upper branch is attracting and the lower branch is repelling. The drift will be negative outside the parabola and positive inside. Shifting the initial condition to the left is equivalent to shifting the λ to the right, so this picture may be used to consider the problem for all λ.

Now suppose we put back in the b'. The solution of the Ricatti equation is now a diffusion. In this case we have that there is some positive chance of the diffusion moving against the drift, including crossing the parabola. If we use $P_{-\lambda,y}$ to denote the probability measure associated with starting our diffusion with initial condition $p(-\lambda) = y$, then we get

$$P(\lambda_1 > a) = P_{-a,+\infty}\left(p \text{ does not blow up}\right)$$

We can bound this below by starting our particle at 1.

$$P_{-a,+\infty}\left(p \text{ does not blow up}\right) \geq P_{-a,1}\left(p \text{ does not blow up}\right)$$

We now bound this below by requiring that our diffusion stays in $p(x) \in [0, 2]$ on the interval $x \in [-a, 0)$ and then choosing convergence to the upper edge of the parabola after 0. This gives

$$\begin{aligned} &P_{-a,1}\left(p \text{ does not blow up}\right) \\ &\qquad\qquad \geq P_{-a,1}\left(p \text{ stays in } [0,2] \text{ for } x < 0\right) \cdot P_{0,0}\left(p \text{ does not blow up}\right). \end{aligned}$$

Notice that the second event does not depend on a and so will just be some constant. We focus on the first event.

We can do a Girsanov change of measure to determine the likelihood. This change of measure moves us to working on the space where p is replaces by a standard brownian motion (started at 1). The Radon-Nikodym derivative of this change of measure may be computed explicitly. We compute

$$\begin{aligned} &P_{-a,1}\left(p \text{ stays in } [0,2] \text{ for } x < 0\right) = E_{-a,1}\left[\mathbf{1}(p_x \in [0,2], x \in (-a,0))\right] \\ &\quad = E_{-a,1}\left[\exp\left(\frac{\beta}{4}\int_{-a}^{0}(x - b^2)db - \frac{\beta}{8}\int_{-a}^{0}(x-b^2)^2dx\right)\mathbf{1}(b_x \in [0,2], x \in (-a,0))\right]. \end{aligned}$$

Notice that

$$\frac{\beta}{4}\int_{-a}^{0}(x - b^2)db \sim O(a), \qquad \text{and} \qquad \frac{\beta}{8}\int_{-a}^{0}(x-b^2)^2dx \approx -\frac{\beta}{24}a^3,$$

which gives us the desired lower bound.

$\square$

CHAPTER 5

The hard edge

1. The heuristic convergence argument

We change from working with the β-Hermite ensemble to working with the β-Laguerre ensemble introduced in subsection 4.1. Recall that for $\Lambda^{(n)}$ with β-Laguerre distribution and

$$\nu_n = \frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k/n}$$

we get that ν_n converges weekly to the Marchenko-Pastur distribution with parameter $\frac{n}{p} \to \gamma$ 4. We made the observation that if $\gamma = 1$ than the Marchenko Pastur density has an asymptote at 0. This suggests that the local behavior in this case should be different than the soft edge limit discussed in the previous section. This turns out to be true some of the time. In particular if p and n differ by a constant a, or $p - n = a_n \to a$ then the local behavior will be something called hard-edge behavior. In the intermediate regime where $a_n \to \infty, a_n/n \to 0$ it is expected that the behavior is soft edge and so the limiting process will be Airy_β. For this regime there is a partial result in the case $\beta = 2$ for $a_n \sim c\sqrt{n}$ by Deift, Menon, and Trogdon (see [**1**]), but otherwise the problem remains open.

The process is called a hard-edge process because it happens when the spectrum of a random matrix is forced against some hard constraint. Recalling the full matrix models for $\beta = 1, 2$ we observe that the matrices are positive definite. This gives a hard lower constraint of 0 for the eigenvalues. If p is close to n than this hard constraint on the lower edge will be felt and so result in different local behavior.

For the remainder of this section we replace the parameter p used in the original description of the β-Laguerre ensemble with $n + a$ giving us a as the extra parameter. This gives us joint density

$$f_{L,\beta}(\lambda_1, ..., \lambda_n) = \frac{1}{Z_{\beta,n,a}} \prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(a+1)-1} e^{-\frac{\beta}{2}\lambda_i} \prod_{j<k} |\lambda_j - \lambda_k|^\beta.$$

We can also rewrite the associated bidiagonal matrice $L_{\beta,a}$ which has singular values distributed according to $f_{L,\beta}$. This now has the form.

$$L_{\beta,a} = \frac{1}{\sqrt{\beta}} \begin{bmatrix} \chi_{\beta(n+a)} & \tilde{\chi}_{\beta(n-1)} & & & \\ & \chi_{\beta(n+a-1)} & \tilde{\chi}_{\beta(n-2)} & & \\ & & \ddots & \ddots & \\ & & & \chi_{\beta(a+2)} & \tilde{\chi}_{\beta} \\ & & & & \chi_{\beta(a+1)} \end{bmatrix}.$$

Using the same heuristic arguments that were used in the soft-edge case Edelman and Sutton conjectured that the singular values of $L_{\beta,a}$ scaled up by $\sqrt{n}$

should converge to the the singular values of the operator

$$\mathcal{L}_{\beta,a} = -\sqrt{x}\frac{d}{dx} + \frac{a}{2\sqrt{n}} + \frac{1}{\sqrt{\beta}}b'(x)$$

where b' is understood as white noise. This is the case, but rather than following the same method of proof we instead work with the inverses. There are several reasons for working with the inverses instead of the original matrices. First the inverse of $\mathcal{L}_{\beta,a}\mathcal{L}^*_{\beta,a}$ will be an integral operator and we can also treat the matrix inverses as integral operators on the same space. Showing convergence of these integral operators particularly in this case turns out to be easier that showing convergence to a differential operator.

There is another reason we choose to use the integral operators here but did not in the case of the soft edge and that is because these matrices are all positive definite. The shifted matrix we considered in the case of the soft edge could have negative eigenvalues. We we show convergence of the eigenvalues we do this by showing convergence of the operator norm. This gives convergence of the eigenvalue with largest absolute value. This is only enough information if we also know these eigenvalues to be positive.

Recall that we want to show that the eigenvalues near 0 converge in distribution to the eigenvalues of some operator. Our final statement should be something like

$$m_n^2 \Lambda^{(n)} \Rightarrow \{\Lambda_1, \Lambda_2, ...\}$$

where the Λ_i are the eigenvalues of some differential operator. If that limiting operator also has strictly positive eigenvalues this is equivalent to showing the convergence of the top eigenvalues of the inverses. We choose to work with m_n^2 instead of m_n because we break the problem in half to start with by working with the bidiagonal matrices which both need to be scaled up by m_n.

In this setting we can take advantage of the natural embedding of matrices as integral operators into $L^2[0,1]$ that works as follows: let $A = a_{i,j} \in \mathbb{R}^{n\times n}$ and $x_i = i/n$ then there exists a natural operator embedding into $L^2[0,1]$ that does not chance the spectrum given by

$$(Af)(x) = \sum_{j=1}^{n} a_{i,j} n \int_{x_{j-1}}^{x_j} f(x)dx \quad \text{for} \quad x \in [x_{i-1}, x_i).$$

In order to work with this embedding we first need to have an inverse to our original matrices.

Lemma 1.1. *For any lower bidiagonal matrix $B = [b_{i,j}]$ the inverse if it exists is lower triangular given by*

$$B^{-1}_{i,j} = \frac{(-1)^{i+j}}{b_{i,i}} \prod_{k=j}^{i-1} \frac{b_{k+1,k}}{b_{k,k}} \quad \textit{for} \quad j \le i.$$

In order to make this inverse as nice as possible we will conjugate our bidiagonal matrix by an orthogonal matrix in order to get

$$M_{a,\beta} = \frac{1}{\sqrt{\beta}} \begin{bmatrix} \chi_{\beta(a+1)} & & & & \\ -\tilde{\chi}_{\beta} & \chi_{\beta(a+2)} & & & \\ & -\tilde{\chi}_{\beta 2} & \ddots & & \\ & & \ddots & \chi_{\beta(n+a-1)} & \\ & & & -\tilde{\chi}_{\beta(n-1)} & \chi_{\beta(n+a)} \end{bmatrix}.$$

If we consider the matrix $(m_n M_{a,\beta})^{-1}$ embedded as an operator we get that

$$\left(m_n^{-1} M_{a,\beta}^{-1} f\right)(x) = \sum_{j=1}^{\lfloor nx \rfloor} \frac{n\sqrt{\beta}}{m_n \chi_{\beta(\lfloor nx \rfloor + a)}} \prod_{k=j}^{\lfloor nx \rfloor - 1} \frac{\tilde{\chi}_{\beta k}}{\chi_{\beta(k+a)}} \int_{x_{j-1}}^{x_i} f dx.$$

This is an integral operator $K_{\beta,a}^n$ with discrete kernel

(12)
$$k_{\beta,a}^n(x,y) = \frac{n\sqrt{\beta}}{m_n \chi_{\beta(\lfloor nx \rfloor + a)}} \exp\left[\sum_{k=j}^{i-1} \log \tilde{\chi}_{\beta k} - \log \chi_{\beta(k+a)}\right] \mathbf{1}_{x \in [x_{i-1}, x_i)} \mathbf{1}_{y \in [x_{j-1}, x_j)}.$$

Our goal at this point is to determine the proper choice of m as well as the limit of $k_{\beta,a}^n$ in the $n \to \infty$ limit. There are two separate pieces, the exponential piece and the leading term. Recall the interpretation of χ^2 as the sum of normal random variables. Therefore the law of large number implies

$$\frac{\chi_{\beta(\lfloor nx \rfloor + a)}}{\sqrt{n}} \xrightarrow[n\to\infty]{\text{a.s.}} \sqrt{x}.$$

Therefore we should choose $m_n = \sqrt{n}$ and we get

$$\frac{\sqrt{n\beta}}{\chi_{\beta(\lfloor nx \rfloor + a)}} \Rightarrow \frac{1}{\sqrt{x}} \tag{13}$$

for $x \in (0,1]$.

For the exponential term begin by recording the following information about χ random variables.

Proposition 1.2. *Let χ_r be a chi random variable with index $r > 0$, then as $r \to \infty$*

$$E[\log \chi_r] = \frac{1}{2}\log r - \frac{3}{2r} + O(1/r^2) \qquad \mathrm{Var}[\log \chi_r] = \frac{1}{2r} + O(1/r^2),$$

while $E[(\log \chi_r - E \log \chi_r)^{2m}] = O(r^{-m})$ for positive integer m.

Exercise 1.3. Use the expected value expansion from Proposition 1.2 to show that

$$\lim_{n\to\infty} \sum_{k=\lfloor ny \rfloor}^{\lfloor nx \rfloor} \left(E \log \tilde{\chi}_{\beta k} - E \log \chi_{\beta(k+a)}\right) = \frac{a}{2} \log(y/x). \tag{14}$$

for $y < x \in (0,1]$.

This takes care of the mean of the exponential term, but we still need to determine what happens to the noise. For that we use the following theorem.

THEOREM 1.4 (After Theorem 1.3, Chapter 7 of [**3**]). *Let $y_{n,k}$ be a sequence of mean-zero processes starting at* 0 *with independent increments $\Delta y_{n,k}$. Assume*

$$nE(\Delta y_{n,k})^2 = f(k/n) + o(1), \quad nE(\Delta y_{n,k})^4 = o(1)$$

uniformly for k/n in compact sets of $[0,T)$ with continuous $f \in L^1_{loc}[0,T)$. Then $y_n(t) = y_{n,\lfloor nt \rfloor} \Rightarrow \int_0^t f^{1/2}(s)db(s)$. with a standard brownian motion b (in the Skorohod topology).

This may be used to show that

$$\sum_{k=\lfloor nx \rfloor}^{n} (\log \chi_{\beta(k+c)} - E \log \chi_{\beta(k+c)}) \Rightarrow \int_x^1 \frac{1}{\sqrt{2\beta z}} db(z). \tag{15}$$

for a standard brownian motion b.

Conclusion. The limits in (13) (14), and (15) suggest that $K^n_{\beta,a}$ should converge to some limiting kernel operator with kernel

$$k_{\beta,a}(x,y) = x^{-\frac{1+a}{2}} \exp\left[\int_y^x \frac{db_z}{\sqrt{\beta z}}\right] y^{a/2} \mathbf{1}_{y<x}. \tag{16}$$

As in the case of the soft-edge this statement may be made precise.

As a full statement about the inverse operators we should have the following:

$$n((M^T_{a,\beta})^{-1} M^{-1}_{a,b} f)(x) \Rightarrow \int_x^1 \int_0^y x^{\frac{a}{2}} e^{\int_x^y \frac{db_s}{\sqrt{\beta s}}} y^{-a+1} e^{\int_z^y \frac{db_s}{\sqrt{\beta s}}} z^{\frac{a}{2}} f(z)dzdy$$

To get to a differential operator characterization of the point process we consider the eigenvalue equation which gives us

$$\begin{aligned} f(x) &= \lambda \int_x^1 \int_0^y x^{\frac{a}{2}} e^{\int_x^y \frac{db_s}{\sqrt{\beta s}}} y^{-a+1} e^{\int_z^y \frac{db_s}{\sqrt{\beta s}}} z^{\frac{a}{2}} f(z)dzdy \\ &= \int_0^1 (xz)^{a/2} \left(\int_{x \vee y}^1 e^{-2\int_y^1 \frac{db_s}{\sqrt{\beta s}}} y^{-(a+1)} dy \right) e^{\int_x^1 \frac{db_s}{\sqrt{\beta s}}} e^{\int_z^1 \frac{db_s}{\sqrt{\beta s}}} f(z)dz \end{aligned}$$

where in the final line we switched the order of integration. We now make the substitution $g(x)x^{-a/2}e^{-\int x^1 \frac{db_s}{\sqrt{\beta s}}} f(x)$ and use the time change $\int_x^1 \frac{1}{\sqrt{x}} db_s = \hat{b}(\log(1/x))$ for some new Brownian motion $\hat{b}$. Then

$$g(x) = \lambda \int_0^1 \left(\int_{x \vee z}^1 e^{-\frac{2}{\sqrt{\beta}}\hat{b}(\log 1/y)} y^{-(a+1)} dy \right) g(z) z^a e^{\frac{2}{\sqrt{\beta}}\hat{b}(\log 1/z)} dz.$$

we make a final change of variables $x, y, z \mapsto e^{-x}, e^{-y}, e^{-z}$ to get

$$h(x) = \lambda \int_0^\infty \left(\int_0^{x \wedge y} e^{-\frac{2}{\sqrt{\beta}}\hat{b}(z)} e^{az} dz \right) h(y) e^{-(a+1)y} e^{\frac{2}{\sqrt{\beta}}\hat{b}(y)} dy.$$

From the speed and scale characterization of a diffusion we can say that the generator of a diffusion with speed measure $m(dx)$ and scale measure $s(dx)$ is given by $\frac{dm}{dx}\frac{d}{dx}[\frac{ds}{dx}\frac{d}{dx}]$ and its inverse operator acting on a function ψ is given by $-\int_0^\infty (\int_0^{x \wedge y} s(dz)\psi(y)m(dy)$. This gives us that for $\mathfrak{G}_{\beta,a} = (K^*_{\beta,a} K_{\beta,a})^{-1}$ we get

$$\mathfrak{G}_{\beta,a} = -\exp[(a+1)x + \frac{2}{\sqrt{\beta}} b(x)] \frac{d}{dx} \left(\exp[-ax - \frac{2}{\sqrt{\beta}} b(x)] \frac{d}{dx} \right). \tag{17}$$

2. The operator and spectrum convergence

Now that we have identified the limiting operator we will state the actual operator convergence statement and then give the proof of convergence of the spectrum.

LEMMA 2.1. *Let $K_{\beta,a}$ be the integral operator with kernel $k_{\beta,a}$ defined in (16), and $K_{\beta,a}$ the integral operator with kernel defined in (12). Both operators are almost surely Hilbert-Schmidt and for any sequence of $K^n_{\beta,a}$ there exists a subsequence n' along which $K^{n'}_{\beta,a}$ converges to $K_{\beta,a}$ in Hilbert-Schmidt norm with probability 1. That is*

$$\lim_{n'\to\infty}\int_0^1\int_0^1 |k^{n'}_{\beta,a}(x,y)-k_{\beta,a}(x,y)|^2dxdy=0, \qquad \textit{almost surely.}$$

Now that we have this we need to show that this implies convergence of the eigenvalues. To do this as in the case of the soft edge we use the Rayleigh characterization of the eigenvalues so that if we take $\lambda_1(n)<\lambda_2(n)<...$ to be the ordered eigenvalue of the β-Laguerre ensemble then we get

$$n\lambda_1(n)=\inf_{\|v\|_{\ell_1}=1}\langle v,nM^T_{a,\beta}M_{a,\beta}v\rangle=\left(\sup_{\|f\|_{L^2}=1}\langle f,(K^n_{\beta,a})^TK^n_{\beta,a}f\rangle\right)^{-1}=\|(K^n_{\beta,a})^TK^n_{\beta,a}\|^{-1}$$

where the final norm is the $L^2\to L^2$ operator norm. Convergence in Hilbert-Schmidt norm of the subsequence $K^{n'}_{\beta,a}\to K_{\beta,a}$ gives us that for $f\in L^2$ we have $(K^{n'}_{\beta,a})^TK^{n'}_{\beta,a}f\to K^T_{\beta,a}K_{\beta,a}f$ in L^2. This is strong convergence of the operators and will imply convergence of the operator norms $\|(K^{n'}_{\beta,a})^TK^{n'}_{\beta,a}\|\to\|K^T_{\beta,a}K_{\beta,a}\|$. Since all of the operators involved are Hilbert-Schmidt these norms are exactly the largest eigenvalue of each operator giving us convergence of the eigenvalues.

From this we may show convergence of the eigenfunctions. Define $\{f_n\}$ and f to be the eigenvalues of $J_n=(K^n_{\beta,a})^TK^n_{\beta,a}$ and $J=K^T_{\beta,a}K_{\beta,a}$ respectively. Since $\{f_n\}$ is uniformly bounded in L^2 it has a weakly convergent subsequence with limit function f_∞. Then for all $\varphi\in L^2$ we get that

$$\langle\varphi,Jf_\infty\rangle=\lim_{n'\to\infty}\langle\varphi J_{n'}f_{n'}\rangle=\lim_{n'\to\infty}\langle\varphi,\lambda_{1,n'}f_{n'}\rangle=\Lambda_1\langle\varphi,f_\infty\rangle$$

since this holds for all φ we get that f_∞ is the eigenfunction of J associated to Λ_1. Since every subsequence has a further subsequence that converges to f this gives us that $\lim_{n\to\infty}f_n=f$.

In order to get convergence of the next eigenvalue we consider a sequence of functions $|f_n-f|\to 0$ in L^2 and for a function $g\in L^2$ we denote by P_g the projection onto the orthogonal compliment of g. Then we will get that there exists a subsequence along which $P_{f_n}(K^n_{\beta,a})^TK_{\beta,a}\to P_f(K_{\beta,a})^TK_{\beta,a}$.

Now using norms of these operators we get convergence of the next next largest eigenvalue. We can induct using projections to get convergence of the largest k eigenvalues for any finite k.

3. Stochastic differential equations and the hard edge

As in the case of the soft edge operator we may consider the eigenvalue equation in order to get a related differential equation. If we consider again the eigenvalue equation for the inverse operator we see

$$h(x) = \lambda \int_0^\infty \left(\int_0^{x \wedge y} e^{-\frac{2}{\sqrt{\beta}}\hat{b}(z)} e^{az} dz \right) h(y) e^{-(a+1)y} e^{\frac{2}{\sqrt{\beta}}\hat{b}(y)} dy.$$

We may compute that the derivative is

$$h'(x) = -\lambda e^{-\frac{2}{\sqrt{\beta}}\hat{b}(x)} e^{ax} \int_x^\infty h(y) e^{-(a+1)y} e^{\frac{2}{\sqrt{\beta}}\hat{b}(y)} dy$$

For the next derivative we need to define it in the sense of Itô. Making use of this we get

$$\begin{aligned} dh'(x) &= \frac{2}{\sqrt{\beta}} h'(x) db(x) + \left((a + \tfrac{2}{\beta}) h'(x) - \lambda e^{-x} h(x) \right) dx \\ dh(x) &= h'(x) dx \end{aligned}$$

We use the Riccati map $p(x) = h'(x)/h(x)$ which is valid away from the zeros of h. We get

$$dp(x) = \frac{2}{\sqrt{\beta}} p(x) db(x) + \left((a + \tfrac{2}{\beta}) p(x) - p^2(x) - \lambda e^{-x} \right) dx$$

We can use this Ricatti transformation in order to describe the counting function of the eigenvalue process.

THEOREM 3.1 ([**6**]). *Let $\Lambda_0(\beta, a) < \Lambda_1(\beta, a) < \ldots$ be the ordered eigenvalues of $\mathfrak{G}_{\beta,a}$, and let $P_{\infty,t}$ denote the law induced by $p(\cdot : \beta, a, \lambda)$ started at $+\infty$ at time t, and restarted at $+\infty$ and time $\mathfrak{m}$ upon any $\mathfrak{m} < \infty, p(\mathfrak{m}) = -\infty$. Then,*

$$P(\Lambda_0(\beta, a) > \lambda) = P_{\infty,0}(p \text{ never hits } 0), \tag{18}$$

$$P(\Lambda_k(\beta, a) < \lambda) = P_{\infty,0}(p \text{ hits 0 at least } k+1 \text{ times}). \tag{19}$$

In other words the counting function of the eigenvalue process is the number of times that $p_\lambda(t)$ hits 0.

CHAPTER 6

The Bulk Limit

The original description of the bulk limit process was through a process called the Browning Carousel first introduced by Valkó and Virág in [**9**]. The limiting process introduced there could also be described in terms of a system of coupled stochastic differential equations which gave the counting function of the process. In particular let α_λ satisfy

$$d\alpha_\lambda = \lambda \frac{\beta}{4} e^{-\frac{\beta}{4}t} dt + \operatorname{Re}\left[(e^{-i\alpha_\lambda} - 1)dZ\right], \tag{20}$$

where $Z_t = X_t + iY_t$ with X and Y standard Brownian motions and $\alpha_\lambda(0) = 0$. The α_λ are coupled through the noise term. Define $N_\beta(\lambda) = \frac{1}{2\pi} \lim_{t\to\infty} \alpha_\lambda(t)$, then $N_\beta(\lambda)$ is the counting function for Sine_β.

More recent work characterizes this same limit process as the eigenvalues of a certain Dirac operator. In order to work with this material we need some background in orthogonal polynomials and Dirac operators, we give this background first and then move on to the proof of convergence.

1. The Szegö Recursion

Suppose we have an $n \times n$ unitary matrix U. To this matrix there is associated a spectral measure μ_n on the unit circle which satisfies

$$\int_{\partial\mathbb{D}} p \, d\mu_n = \langle p(U)e_1, e_1\rangle = \sum_{i=1}^{n} p(\lambda_i)\langle v_i, e_1\rangle^2$$

for polynomials p. We use the notation $q_i = \langle v_i, e_1\rangle$ and refer to the q_i as the spectral weights. To this spectral measure there is associated a sequence of n Verblunsky coefficients $\alpha_0, \alpha_1, ..., \alpha_{n-1} \in \mathbb{C}$, which give the recurrence relation for the polynomials orthogonal with respect to the measure μ_n.

More generally there is a bijection between measures μ on $\partial\mathbb{D}$ and sequences of Verblunsky coefficients $\{\alpha_k\}_{k=0}^\infty$ where $\alpha_k \in \mathbb{D}$. In the case of μ_n mentioned above we have $|\alpha_{n-1}| = 1$ and $\alpha_k = 0$ for $k \geq n$. We define the family of orthogonal polynomials $\Phi_0(z), \Phi_1(z), ...$ on the unit circle to be the polynomials with respect to the spectral measure μ. We simultaneously define the conjugate polynomials $\Phi_k^*(z) = z^k\overline{\Phi}_k(1/z)$. The families Φ and Φ^* obey the recurrence

$$\begin{aligned}\Phi_{k+1} &= z\Phi_k(z) - \bar{\alpha}_k\Phi_k^*(z)\\ \Phi_{k+1}^* &= \Phi_k^*(z) - \alpha_k z\Phi_k(z)\end{aligned}$$

with initial condition $\Phi_0(z) = \Phi_0^*(z) = 1$. This recurrence can be encoded into a matrix equation by

$$\begin{bmatrix} \Phi_{k+1}(z) \\ \Phi_{k+1}^*(z) \end{bmatrix} = \begin{bmatrix} z & -\bar{\alpha}_k \\ -\alpha_k z & 1 \end{bmatrix} \begin{bmatrix} \Phi_k(z) \\ \Phi_k^*(z) \end{bmatrix} \tag{21}$$

$$= \begin{bmatrix} 1 & -\bar{\alpha}_k \\ -\alpha_k & 1 \end{bmatrix} \begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Phi_k(z) \\ \Phi_k^*(z) \end{bmatrix} = A_k Z \begin{bmatrix} \Phi_k(z) \\ \Phi_k^*(z) \end{bmatrix} \tag{22}$$

Where we use A_k and Z to denote the first and second matrices in that row respectively. Making use of the recurrence we have that

$$\begin{bmatrix} \Phi_{k+1}(z) \\ \Phi_{k+1}^*(z) \end{bmatrix} = A_k Z A_{k-1} Z \cdots A_0 Z \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

We will get that z is an eigenvalue of the original matrix U, or in the support of the spectral measure μ_n if and only if $[\Phi_n(z), \Phi_n^*(z)]^t$ is parallel to $[0,1]^t$. This condition may be transformed to

$$\begin{bmatrix} \bar{\alpha}_{n-1} \\ 1 \end{bmatrix} \parallel Z \begin{bmatrix} \Phi_{n-1}(z) \\ \Phi_{n-1}^*(z) \end{bmatrix} = Z A_{n-2} Z \cdots A_0 Z \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{23}$$

1.1. From products to conjugations. We "change the basis" of the functions that we study. Instead of tracking the vectors with the orthogonal polynomials we define the functions $f_{k+1}(\lambda)$ given by by

(24)

$$f_{k+1}(\lambda) = e^{-i\lambda(k+1)/2}(A_{k-1}A_{k-2}\cdots A_0)^{-1} \begin{bmatrix} \Phi_k(e^{i\lambda}) \\ \Phi_k^*(e^{i\lambda}) \end{bmatrix} = e^{-i\lambda(k+1)/2} M_k^{-1} \begin{bmatrix} \Phi_k(e^{i\lambda}) \\ \Phi_k^*(e^{i\lambda}) \end{bmatrix}$$

where $M_k = A_{k-1}A_{k-2}\cdots A_0$. From this definition we can compute

$$\begin{aligned} f_{k+1}(\lambda) &= e^{-i\lambda(k+1)/2} M_k^{-1} A_{k-1} Z \begin{bmatrix} \Phi_{k-1}(z) \\ \Phi_{k-1}^*(z) \end{bmatrix} \\ &= e^{-i\lambda(k+1)/2} M_k^{-1} A_{k-1} Z M_{k-1} e^{i\lambda k/2} \left(e^{-i\lambda k/2} M_{k-1}^{-1} \begin{bmatrix} \Phi_{k-1}(z) \\ \Phi_{k-1}^*(z) \end{bmatrix} \right) \\ &= e^{-i\lambda/2} M_{k-1}^{-1} Z M_{k-1} f_k(\lambda) \\ &= \begin{bmatrix} e^{-i\lambda/2} & 0 \\ 0 & e^{i\lambda/2} \end{bmatrix}^{M_{k-1}} f_k(\lambda). \end{aligned}$$

Where we use the notation $Z^A = A^{-1}ZA$. Now recall the condition in (23). Written in terms of the f_ks we get that λ is an eigenvalue

$$\begin{bmatrix} \bar{\alpha}_{n-1} \\ 1 \end{bmatrix} \parallel M_{n-1} f_n(\lambda), \qquad \text{which can be rewritten as} \qquad M_{n-1}^{-1} \begin{bmatrix} \bar{\alpha}_{n-1} \\ 1 \end{bmatrix} \parallel f_n(\lambda).$$

Notice that the parallel condition means that it is enough to consider the ratio of the two entries of the vector. To this end we introduce the notation

$$\mathcal{P} \begin{bmatrix} x \\ y \end{bmatrix} = x/y,$$

Which gives that z is an eigenvalue if and only if

$$\mathcal{P}(A_{n-2}\cdots A_0)^{-1} \begin{bmatrix} \bar{\alpha}_{n-1} \\ 1 \end{bmatrix} = \mathcal{P} Z^{M_{n-1}} Z^{M_{n-2}} \cdots Z^{M_0} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

1.2. The geometric picture. We make use of a correspondence between linear fractional transformations and 2×2 matrices. In particular we have

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto m_A(w) = \frac{aw+b}{cw+d}.$$

A few observations

(1) This mapping preserves structure in that

$$m_{AB}(w) = m_A(m_B(w)) = m_A \circ m_B(w).$$

(2) Z maps to rotation by z (recall that $z \in \mathbb{D}$).

$$\begin{bmatrix} z & 0 \\ 0 & 1 \end{bmatrix} \mapsto m_Z(w) = zw$$

(3) Assume we have $A^{-1}ZA$ where A is a bijection on $\mathbb{D}$. Then $m_{Z^A}(w)$ will be conformal rotation about the point $m_{A^{-1}}(0)$.

(4) Let $\mathcal{P}$ be defined as before then

$$m_A(0) = \mathcal{P}A \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \text{and} \qquad m_A(1) = \mathcal{P}A \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Now the evolution of the Γ_k may be studied as follows: We define the points

$$b_k = m_{M_k^{-1}}(0) = \mathcal{P}(A_{k-1} \cdots A_0)^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{25}$$

These points b_k will form a walk in the disk. Now in the geometric picture we get that $m_{Z^{M_k}}(w)$ is rotation by z around the point b_k. We also define the points

$$X_k(z) = m_{Z^{M_k} \cdots Z^{M_0}}(1) = \mathcal{P} Z^{M_k} Z^{M_{k-1}} \cdots Z^{M_0} \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{26}$$

Notice that the circle $\partial\mathbb{D}$ is fixed under these transformations so $X_k(z) \in \partial\mathbb{D}$ for all z.

In the framework we should think of the sequence of points $\{X_k(z)\}$ as being obtained in the following way: X_k is obtained from the point X_{k-1} by rotating about the point b_k by z where the angle of ration is the hyperbolic angle. In this set up we get that z is an eigenvalue if and only if $X_{k-1}(z) = b_* = \mathcal{P}(A_{n-2} \cdots A_0)^{-1} \begin{bmatrix} \bar{\alpha}_{n-1} \\ 1 \end{bmatrix}$. Notice that both of these points lie on the unit circle because $\bar{\alpha}_{n-1} \in \partial\mathbb{D}$. So we can think of b_* as a "target" point.

2. Dirac operators

One way of understanding a Dirac operator is as the square root of a second order differential operator. That is, for some second order differential operator L (for example the Laplacian) we are looking for a first order differential operator D such that $D^2 = L$. Let us begin with the classical examples.

Example 2.1 (1-dimensional Laplacian). Suppose L is the 1-dimensional Laplacian $L = -\frac{d^2}{dt^2}$. The natural choice is $D = i\frac{d}{dt}$. This D is formally self-adjoint.

That is suppose we have $f, g :\in C^\infty(\mathbb{R}, \mathbb{C})$ then

$$\langle Df, g\rangle_{L^2} = i\int_{-\infty}^{\infty} f'(t)g(t)dt = -i\int_{-\infty}^{\infty} f(t)g'(t)dt = \langle f, Dg\rangle.$$

By considering f decomposed into its real and imaginary parts $f(t) = u(t) + iv(t)$ we get $Df(t) = iu'(t) - v'(t)$

$$Df(t) = D\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\frac{d}{dt}\begin{bmatrix} u(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\frac{d}{dt}f(t).$$

This is the type of system that we will be considering.

Now consider this same system acting on functions $f : \mathbb{R}_+ \to \mathbb{R}^2$ with boundary condition $f(0) = (1, 1)^t$. Let us look at the eigenvalue equation for this system, this gives us

$$\lambda f(t) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} f'(t) \tag{27}$$

this may be rewritten in the classical differential equation form to give us

$$f'(t) = \lambda\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} f(t)$$

This is now a standard ODE system so we may solve by diagonalizing and taking the exponential. This gives us

$$f(t) = S\begin{bmatrix} e^{-i\lambda} & 0 \\ 0 & e^{i\lambda} \end{bmatrix} S^{-1}f(0).$$

2.1. A Dirac operator from the Szegö recurrsion. Recall from the previous section we had that

$$f_{k+1}(\lambda) = \begin{bmatrix} e^{i\lambda/2} & 0 \\ 0 & e^{-i\lambda/2} \end{bmatrix}^{M_k} f_k(\lambda).$$

Now define $M(t) = M_{\lfloor nt\rfloor}$ for $t \in [0, 1)$ and consider the differential operator τ acting on functions $g : [0, 1) \to \mathbb{C}^2$ given by

$$\tau g = 2\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}^{M_t} g'(t)$$

with boundary conditions

$$g(0) \parallel u_0 := \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \text{and} \qquad g(1) \parallel u_1 := M_{n-1}^{-1}\begin{bmatrix} \bar{\alpha}_{n-1} \\ 1 \end{bmatrix}.$$

Then the solution of the eigenvalue equation $\tau g = \mu g$ satisfies

$$g'(t) = \begin{bmatrix} i\mu/2 & 0 \\ 0 & -i\mu/2 \end{bmatrix}^{M(t)} g(t)$$

which we can solve explicitly on the intervals $[\frac{k}{n}, \frac{k+1}{n})$ giving us

$$g(\tfrac{k+1}{n}) = \begin{bmatrix} e^{i\frac{\mu}{2n}} & 0 \\ 0 & e^{-i\frac{\mu}{2n}} \end{bmatrix}^{M(k/n)} g(\tfrac{k}{n}).$$

At this point we can see that $g_\mu(k/n) = f_k(\mu/n)$ and so the eigenvalues of τ are given by

$$\{\mu \in \mathbb{R} : e^{i\mu/n} \text{ is an eigenvalue of the Szegö recursion}\}.$$

Notice that M_k has the form $\begin{bmatrix} p_k & q_k \\ \bar{q}_k & \bar{p}_k \end{bmatrix}$ with $|p_k|^2 - |q_k|^2$ since the A_k have this form and the property is inherited. Recall our definition of $b_k = -\frac{q_k}{p_k} = \mathcal{P}M_k^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then we can compute

$$\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}^{M_k} = \frac{1}{1-|b_k|^2}\begin{bmatrix} 1+|b_k|^2 & 2b_k \\ 2\bar{b}_k & 1+|b_k|^2 \end{bmatrix}\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix}.$$

If we take $b(t) = b_{\lfloor nt \rfloor}$ for $t \in [0,1]$ we can write τ in a new form

$$\tau g = \frac{2}{1-|b|^2}\begin{bmatrix} 1+|b|^2 & 2b \\ 2\bar{b} & 1+|b|^2 \end{bmatrix}\begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} g', \qquad t \in [0,1) \tag{28}$$

acting on functions $g : [0,1) \to \mathbb{C}^2$ with initial and end conditions $[1,1]^t$ and $[b(1),1]^t$.

2.2. Circular β-ensembles. We look at the n-point measure on the circle introduced in section 4.2 with joint intensity

$$f(\theta_1, ..., \theta_n) = \frac{1}{Z_{n,\beta}} \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^\beta.$$

We further equip a spectral measure with $(q_1^2, ..., q_n^2)$ ∼Dirichlet$(\frac{\beta}{2}, ..., \frac{\beta}{2})$. For the cases $\beta = 1, 2$ and 4 this is the distribution of the spectral measure induced by the Haar measure on Orthogonal, Unitary or Symplectic matrices. With this choice of spectral measure the associated Verblunsky coefficients will be independent with rotationally invariant distribution and

$$|\alpha_k|^2 \sim \begin{cases} \text{Beta } (1, \frac{\beta}{2}(n-k-1)) & k < n-1 \\ 1 & k = n-1 \end{cases}$$

Recall that if $X \sim \text{Beta}(a,b)$ then

$$EX = \frac{a}{a+b}, \qquad \operatorname{Var} X = \frac{ab}{(a+b)^2(a+b+1)}.$$

From this sequence of Verblunsky coefficient we can define the points b_k by

$$b_k = \mathcal{P}M_k^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{29}$$

In order to get some sort of limit we will need to show that $\{b_k\}$ is a random walk path that converges in some sense to a hyberbolic Brownian motion. If we get this that the Dirac operator defined with the piecewise constant R will in some sense converge to a Dirac operator driven by the hyperbolic Brownian motion. This operator will have the same form at (28) but the random walk path b will be replaced by a hyperbolic Brownian motion.

In order to make this convergence statement precess we actually begin by moving from the hyperbolic disk model to the half place. We then show that the inverse operators are Hilbert-Schmidt integral operators that converge in Hilbert-Schmidt norm at some rate. This gives convergence of the eigenvalues.

3. The Canonical system setup

The Dirac operator associated to the Laplacian is the nicest case of a more general framework. In particular we will consider operators of the form

$$Jv'(t) = \lambda R(t)v(t), \qquad J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{30}$$

Notice that this is the same framework as in equation (27) except that there is an extra matrix $R(t)$. We will take $R(t)$ to be an integrable positive definite real 2 by 2 matrix valued function. Notice that such a system may be defined for R semi-definite, but with the choice strictly positive definite we there is an associated differential operator τ given by

$$\tau v(t) = R^{-1}(t)Jv'(t).$$

since $R(t)$ is positive semi-definite it has a unique representation

$$R = \frac{f}{y}\begin{bmatrix} 1 & -x \\ -x & x^2+y^2 \end{bmatrix} = \frac{f}{\det X}X^tX, \quad \text{where} \quad X = \begin{bmatrix} 1 & -x \\ 0 & y \end{bmatrix}$$

where $f > 0, y > 0$ and $x \in \mathbb{R}$. The map $X \mapsto R$ is a bijection between the affine group and all positive definite 2×2 matrices of determinant 1. We specify boundary conditions and study solutions to the equation

$$\lambda v(t) = R^{-1}(t)v'(t), \qquad v(0) = u_0,\ v(T) = u_T.$$

These solutions will correspond to rotating a point on the boundary of $\mathbb{H}$ the upper half plane around the point $x(t) + iy(y)$ at rate λ (in the hyperbolic angle).

3.1. An associated Hilbert space and inverses. For any $R(t)$ positive definite for $t \in [0,T)$ and that there exists some nonzero vector $v_* \in \mathbb{R}^2$ with $\int_0^T u_*^t R(t)u_* dt < \infty$, then we can define the Hilbert space $L^2_R[0,T)$ equipped with the inner product

$$\langle v, w\rangle_R = \int_0^T v(t)R(t)w(t)dt$$

We now have that τ acts on functions $f : [0,T) \to \mathbb{R}^2$ with $\|f\|_R^2 = \langle f, f\rangle_R < \infty$. This means that for different R our differential operators will act on different spaces. In particular if we try to do the discrete to continuous convergence argument our operators do not all act on the same space. In order to get around this we may work instead with the operators $\hat{\tau} = X\tau X^{-1}$. This $\hat{\tau}$ is a self-adjoint operator acting on v such that $Xv \in \mathrm{dom}(\tau) \subset L^2$. Moreover the spectrum of $\hat{\tau}$ is the same as that of τ.

In order to prove the final convergence statement we actually show convergence the inverses of these $\hat{\tau}$ operators. The inverse operator $\hat{\tau}^{-1}$ is a Hilbert-Schmidt integral operator with kernel

$$K_{\hat{\tau}^{-1}}(x,y) = \frac{1}{2}\left(X(x)u_0(X(y)u_T)^t\mathbf{1}(x<y) + X(x)u_T(X(y)u_0)^t\mathbf{1}(y<x).\right)$$

3.2. From $\mathbb{H}$ to $\mathbb{D}$. For any operator τ that is defined as in (30) there is a corresponding operator $\tilde{\tau}$ with the same spectrum where instead of tracking the path $x+iy$ in the half-plane model we instead work in the Poincaré disk model.

Let $\tilde{U}$ be the linear transformation corresponding to the Cayley map $U(z)$ (conformal with $(\infty, 1, -1) \mapsto (1, -i, i)$),

$$\tilde{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}$$

and consider the operator

$$\tilde{\tau} u(t) = \tilde{U} \tau (\tilde{U}^{-1} u(t)) = \tilde{U} R^{-1}(t) J \tilde{U}^{-1} u'(t)$$

defined on functions $u : [0,1] \to \mathbb{C}^2$. Direct computation shows that

$$\tilde{\tau} u(t) = \frac{1}{f(t)(1 - |\gamma(t)|^2)} \begin{bmatrix} 1 + |\gamma(t)|^2 & 2\gamma(t) \\ 2\bar{\gamma}(t) & 1 + |\gamma(t)|^2 \end{bmatrix} \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} u'(t),$$

where $\gamma(t) = U(x(t) + iy(t))$ is the image of the driving path in the disk model. Notice that this form is identical the one we found by considering the operator associated to the Szegö recursion. Therefore we can move from that operator to the half plane and use the structure developed there.

CHAPTER 7

Operator limits via Stochastic Differential Equations

1. Characterizing the limits by differential equations

Recall that in chapters 4 and 5 we gave a characterization of the eigenvalue processes for the two operators by looking at the eigenvalue equation and considering the Ricatti transformation of the associated functions that solve the equation (possibly not eigenfunctions if they do not satisfy the boundary conditions). We also had a description of the bulk process Sine_β in terms of its counting function. For the bulk process we began by defining the diffusion α_λ with

$$d\alpha_\lambda = \lambda \frac{\beta}{4} e^{-\frac{\beta}{4}t} dt + \mathrm{Re}\left[(e^{-i\alpha_\lambda} - 1)dZ\right], \tag{31}$$

where $Z_t = X_t + iY_t$ with X and Y standard Brownian motions and $\alpha_\lambda(0) = 0$. The α_λ are coupled through the noise term. Define $N_\beta(\lambda) = \frac{1}{2\pi} \lim_{t\to\infty} \alpha_\lambda(t)$, then $N_\beta(\lambda)$ is the counting function for Sine_β. This diffusion equation simplifies further in the case where λ is fixed. In particular we may rewrite the martingale part in the following way:

$$\begin{aligned}
\mathrm{Re}\left[\int_0^t (e^{-i\alpha_\lambda} - 1)dZ\right] &= \mathrm{Re}\left[\int_0^t (e^{-i\alpha_\lambda/2} - e^{i\alpha_\lambda/2})e^{-i\alpha_\lambda/2} dZ\right] \\
&= \mathrm{Re}\left[\int_0^t (e^{-i\alpha_\lambda/2} - e^{i\alpha_\lambda/2})(dX^{(\lambda)} + idY^{(\lambda)}\right] \\
&\stackrel{d}{=} \int_0^t 2\sin\left(\frac{\alpha_\lambda}{2}\right) dY^{(\lambda)}
\end{aligned}$$

where $X^{(\lambda)}$ and $Y^{(\lambda)}$ are standard Brownian motions (depending on λ). This gives us that for a fixed λ the diffusion α_λ satisfies the SDE

$$d\alpha_\lambda = \lambda \frac{\beta}{4} e^{-\frac{\beta}{4}t} dt + 2\sin\left(\frac{\alpha_\lambda}{2}\right) dB_t, \qquad \alpha_\lambda(0) = 0.$$

In this section we will focus on proving results about the counting function $N_\beta(\lambda)$. Many of these results can also be proved in a similar way for the hard edge. This is because we can get a similar characterization for the hard edge process. Recall from chapter 5 we have that the Ricatti diffusion p_λ satisfying

$$dp(x) = \frac{2}{\sqrt{\beta}} p(x)db(x) + \left((a + \tfrac{2}{\beta})p(x) - p^2(x) - \lambda e^{-x}\right) dx$$

may be used to define the counting function of the eigenvalue process. We can perform a change of variables to get this into a form where the counting function may be more easily identified. In particular if we take $\varphi_{a,\lambda}$ to satisfy the diffusion

$$d\varphi_{a,\lambda} = \frac{\beta}{2}(a + \tfrac{1}{2}) \sin\left(\frac{\varphi_{a,\lambda}}{2}\right) dt + \beta\lambda e^{-\beta t/8} dt + \frac{\sin \varphi_{a,\lambda}}{2} dt + 2\sin\left(\tfrac{\varphi_{a,\lambda}}{2}\right) dB_t \tag{32}$$

with initial condition $\varphi_{a,\lambda}(0) = 2\pi$. Then we can define

$$M_{a,\beta}(\lambda) \stackrel{d}{=} \lim_{t\to\infty} \left\lfloor \frac{1}{4\pi}\varphi_{a,\lambda}(t) \right\rfloor.$$

This will be the counting function for the singular values of $\mathfrak{G}_{\beta,a}$. The counting function of the eigenvalue process is obtained by looking at $M_{a,\beta}(\sqrt{\lambda})$. For convenience for the rest of the chapter we will use $\mathrm{Bess}_{a,\beta}$ to denote this singular value process.

2. Transitions

Before computing probabilities of various rare events we will start by showing a transition from the singular value process $\mathrm{Bess}_{a,\beta}$ to the bulk process Sine_β. To do this we need to show that the finite dimensional marginals of the counting function near λ converge to those of $N_\beta(x)$ as λ goes to infinity. We can get the counting function of intervals near λ by considering $M_{a,\beta}(\lambda + x) - M_{a,\beta}(\lambda)$. In order to get a diffusion that describes this counting function we need to consider what is the relationship between two diffusions that satisfy stochastic differential equations of the form (32) for λ and $\lambda + x$ which are coupled through their noise terms. Let $\psi_{a,\lambda,x} = \varphi_{a,\lambda+x} - \varphi_{a,\lambda}$, then the SDE for $\psi_{a,\lambda,x}$ is

$$\begin{aligned} d\psi_{a,\lambda,x} = {}& \frac{\beta}{2}(a+1/2)\mathrm{Im}\left[e^{i\frac{\varphi_{a,\lambda}}{2}}\left(e^{-i\frac{\psi_{a,\lambda,x}}{2}}-1\right)\right]dt + \frac{1}{2}\mathrm{Im}\left[e^{i\varphi_{a,\lambda}}\left(e^{-i\psi_{a,\lambda,x}}-1\right)\right]dt \\ & + \beta x e^{-\beta t/8}dt + \mathrm{Im}\left[e^{i\frac{\varphi_{a,\lambda}}{2}}\left(e^{-i\frac{\psi_{a,\lambda,x}}{2}}-1\right)\right]dB_t, \end{aligned} \tag{33}$$

with initial condition $\psi_{a,\lambda,x}(0) = 0$. This follows from standard Itô techniques together with the application of angle addition formulas. The idea here is that the oscillatory integrals will be “small.”

To start we want to determine what is the behavior of $\psi_{a,\lambda,x}$ at finite time as $\lambda \to \infty$. If we can determine this and then show that the impact of the large time behavior of the diffusion is irrelevant this will be enough to determine the behavior of the counting function in the limit.

Proposition 2.1. *For any fixed T and c we get that*

$$\int_0^T e^{ic\varphi_{a,\lambda}}\left(e^{-ic\psi_{a,\lambda,x}}-1\right)dt \xrightarrow{P} 0.$$

This immediately gives us that the first line of (33) will vanish in the $\lambda \to \infty$ limit. Now consider what happens to the noise term. Define

$$X_\lambda(t) = \int_0^t \cos\left(\frac{\varphi_{a,\lambda}}{2}\right)dB_s \quad \text{and} \quad Y_\lambda(t) = \int_0^t \sin\left(\frac{\varphi_{a,\lambda}}{2}\right)ds.$$

In the limit we can compute the quadratic variations

$$[X_\lambda]_t = \int_0^t \cos^2\left(\frac{\varphi_{a,\lambda}}{2}\right)ds = \int_0^2 \cos^2\left(\frac{\varphi_{a,\lambda}}{2}\right) = \frac{t}{2} + \int_0^t \cos\left(\varphi_{a,\lambda}\right)ds \xrightarrow{P} \frac{t}{2}$$

Similarly we may compute

$$[Y_\lambda]_t \xrightarrow{P} \frac{t}{2} \quad \text{and} \quad [X_\lambda, Y_\lambda]_t \xrightarrow{P} 0.$$

At this point the martingale central limit theorem gives us that $(X_\lambda, Y_\lambda) \Rightarrow \frac{1}{\sqrt{2}}(W^{(1)}, W^{(2)})$ where $W^{(1)}$ and $W^{(2)}$ are independent standard Brownian motions. Therefore

$$\begin{aligned}\operatorname{Im}\int_0^T \left[e^{i\frac{\varphi_{a,\lambda}}{2}}\left(e^{-i\frac{\psi_{a,\lambda,x}}{2}}-1\right)\right]dB_t &\Rightarrow \operatorname{Im}\frac{1}{\sqrt{2}}\int_0^T\left(e^{-i\frac{\hat{\psi}_{a,x}}{2}}-1\right)(dW^{(1)}+idW^{(2)})\\ &\stackrel{d}{=} \operatorname{Re}\frac{1}{\sqrt{2}}\int_0^T\left(e^{-i\frac{\hat{\psi}_{x}}{2}}-1\right)(dW^{(2)}+idW^{(1)}).\end{aligned}$$

All together if $\psi_{a,\lambda,x}$ has a limit $\hat{\psi}_x$ this limit should satisfy the SDE

$$d\hat{\psi}_x = \beta x e^{-\beta t/8} + \frac{1}{\sqrt{2}}\left[\left(e^{-i\frac{\hat{\psi}_x}{2}}-1\right)dZ\right].$$

We can check that this is just a time and space changed version of the α_x diffusion. In particular we get that

THEOREM 2.2 ([**4**]). *Let $a > 0$ and $\beta > 0$ fixed, then*

$$\frac{1}{4}(\mathrm{Bess}_{a,\beta} - \lambda) \Rightarrow \mathrm{Sine}_\beta \tag{34}$$

as $\lambda \to \infty$.

List of symbols

$\Rightarrow$	convergence in distribution/measure
$\xrightarrow{P}$	convergence in probability
$\xrightarrow{\text{a.s.}}$	convergence almost surely
$\stackrel{d}{=}$	equality in distribution

Bibliography

[1] P. Deift, T. Trogdon, and G. Menon. On the condition number of the critically-scaled Laguerre unitary ensemble. *Discrete and Cont. Dyn. Sys.*, 36(8):4287–4347, 2016.

[2] I. Dumitriu and A. Edelman. Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.

[3] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[4] D. Holcomb. The random matrix hard edge: rare events and a transition. *https://arxiv.org/abs/1610.05367*, Oct 2016.

[5] M. Krishnapur, B. Rider, and B. Virág. Universality of the stochastic Airy operator. *Comm. Pure Appl. Math.*, 69:145–199, 2016.

[6] J. A. Ramírez and B. Rider. Diffusion at the random matrix hard edge. *Comm. Math. Phys.*, 288(3):887–906, 2009.

[7] J. A. Ramírez, B. Rider, and B. Virág. Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc.*, 24(4):919–944, 2011.

[8] H. 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.

[9] B. Valkó and B. Virág. Continuum limits of random matrices and the Brownian carousel. *Inventiones Math.*, 177:463–508, 2009.

[10] B. Valkó and B. Virág. The Sine_β operator. *http://arxiv.org/abs/1604.04381*, April 2016.