Notes with Diane Holcomb

1 Tridiagonal matrices

Definition 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 2. We can 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.

Example 3. 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.

Why do we introduce this notion of 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 4. 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.

1.1 Graph Limits

Definition 5. 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 6. 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 7. *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 8. 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 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.

Exercise 9. Suppose you have the random d-regular graph on n vertices in the configuration model (look this up). Then in the limit this converges to the d-regular infinite tree in the limit.

Definition 10. 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 11. *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

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

$\square$

Examples of Benjamini-Schram convergence:

- Cycle to $\mathbb{Z}$
- Path of length n to $\mathbb{Z}$.
- Large box of Z^d to Z^d.
- (Exercise) 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, 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 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.

1.2 Equivalence of spectral measure

Let's say two symmetric matrices are equivalent if they have the same eigenvalues with multiplicity. We understand this equivalence relation well: two matrices are equivalen if and only if they are conjugates by an orthongoal matrix (in group theory language, the equivalence classes are the orbits of the conjugation action of the orthongal group). There is a canonical representative in each class, a diagonl matrix with nonincreasing diagonals.

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

Definition 12. 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 13. *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 $O_{11} = 0, O_{1i} = 0$.*

Definition 14. A Jacobi matrix is a real symmetric tridiagonal matrix with positive off-diagonals.

Theorem 15. *For all A there exists a unique Jacobi matrix J such that $\sigma_J = \sigma_A$*

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 $n \times n$ 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. $\square$

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

Proposition 16. *(Trotter) 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, $a_i \sim N(0,2)$ and $b_i \sim \chi_{n-i}$. In particular, $OAO^t =$ has the same spectral measure as A.

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. □

Example 17 (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.

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

1.3 Wigner semicircle law

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

Exercise 18. Check that $\chi_n - \sqrt{n} \to \mathcal{N}(0, 1/2)$.

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. 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.

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

The top eigenvalue

The eigenvalue distribution of the GOE converges after scaling by $\sqrt{n}$ to the Wigner semicircle 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 19 (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 20. *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

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

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

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

We used sublinearity of λ_1, which follows from the Rayleigh quotient representation. □

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 19.

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.

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 21 (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 22 (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 23.

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

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 $a + \frac{1}{a}$. $\square$

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 24. *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. $\square$

3 β-ensembles

Exercise 25. 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

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

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

Theorem 26. *If $\beta > 0$ then the joint density of the eigenvalue of M_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 M_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 27 (Dumitriu, Edelman, Krishnapur, Rider, Virág)**.** *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 $\{(a_i, b_i), i \geq 1\}$ 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 The Stochastic Airy operator

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 18, for $k \ll n$ we have

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

Now we consider the matrix

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

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

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 28. 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).

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.

Statement of 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.

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

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

The asymptotics are classical.

The bilinear form and eigenvalues

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 29. 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 30. 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 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 31. There exists a random constant C so that we have the following inequality of functions:

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

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

Lemma 32. *We have*

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

for some random constant C.

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.

Exercise 33. Prove the bounds in the previous statement, and the following

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

Using these bound we get the following bound:

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

Corollary 34. *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

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

where the λ_k are the Airy eigenvalues (1). But this follows from the operator inequality of Lemma 32 and Exercise 30. $\square$

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.

To be continued.