

# Lectures on Random Matrices (Spring 2025)

## Lecture 4: Semicircle law via tridiagonalization.

### Orthogonal polynomial ensembles

Leonid Petrov

Wednesday, January 29, 2025\*

## Contents

<b>1</b>	<b>Recap</b>	<b>2</b>
1.1	Gaussian ensembles . . . . .	2
1.2	Tridiagonalization . . . . .	2
<b>2</b>	<b>Tridiagonal random matrices</b>	<b>3</b>
2.1	Distribution of the tridiagonal form of the GOE . . . . .	3
2.2	Dumitriu–Edelman $G\beta E$ tridiagonal random matrices . . . . .	4
2.3	The case $\beta = 2$ . . . . .	4
<b>3</b>	<b>Wigner semicircle law via tridiagonalization</b>	<b>6</b>
3.1	Moments for tridiagonal matrices . . . . .	6
3.2	Asymptotics of chi random variables . . . . .	7
3.3	Bounds on products in the bulk . . . . .	8
<b>4</b>	<b>Conclusion</b>	<b>9</b>
4.1	Completing the proof: global semicircle behavior . . . . .	10
<b>5</b>	<b>Wigner semicircle law via Stieltjes transform</b>	<b>10</b>
<b>6</b>	<b>Characteristic Polynomial and Three-Term Recurrence</b>	<b>10</b>
<b>7</b>	<b>Semicircle Law via the Tridiagonal Form</b>	<b>11</b>
7.1	The Scaling and Law of Large Numbers in $\alpha_j$ . . . . .	11
7.2	Diagonal vs. Subdiagonal Entries . . . . .	11
7.3	Characteristic Polynomial and Recurrence Analysis . . . . .	12
7.4	Conclusion: Wigner’s Semicircle Law . . . . .	12
<b>8</b>	<b>Exercises (Due 2025-02-28)</b>	<b>13</b>

---

\*Course webpage • TeX Source • Updated at 12:35, Wednesday 29<sup>th</sup> January, 2025

<b>9 Further Reading and Next Steps</b>	<b>14</b>
<b>D Problems (due 2025-02-28)</b>	<b>15</b>
D.1 Eigenvalue density of $G\beta E$ . . . . .	15
D.2 Chi-square mean and variance . . . . .	15
D.3 Characteristic Polynomial and Three-Term Recurrence . . . . .	15
D.4 Sketch of the Semicircle Limit Proof . . . . .	16

## 1 Recap

Note: I did some live random matrix simulations [here](#) and [here](#) — check them out. More simulations to come.

### 1.1 Gaussian ensembles

We introduced Gaussian ensembles, and for GOE ( $\beta = 1$ ) we computed the joint eigenvalue density. The normalization is so that the off-diagonal elements have variance  $\frac{1}{2}$  and the diagonal elements have variance 1. Then the joint eigenvalue density is

$$p(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_n} \prod_{i=1}^n e^{-\frac{1}{2}\lambda_i^2} \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j), \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n.$$

### 1.2 Tridiagonalization

We showed that any real symmetric matrix  $A$  can be tridiagonalized by an orthogonal transformation  $Q$ :

$$Q^\top A Q = T,$$

where  $T$  is real symmetric tridiagonal, having nonzero entries only on the main diagonal and the first super-/subdiagonals:

$$T = \begin{pmatrix} d_1 & \alpha_1 & 0 & \cdots & 0 \\ \alpha_1 & d_2 & \alpha_2 & \cdots & 0 \\ 0 & \alpha_2 & d_3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \alpha_{n-1} \\ 0 & 0 & \cdots & \alpha_{n-1} & d_n \end{pmatrix}.$$

In the proof, each time we need to act in the orthogonal complement to the subspace  $e_1, \dots, e_{k-1}$  (starting from  $e_1$ ), and apply a Householder reflection to zero out everything strictly below the subdiagonal. (We apply the transformations like  $A \mapsto HAH^\top$ , so that the first row transforms in the same way as the first column of  $A$ ).

## 2 Tridiagonal random matrices

### 2.1 Distribution of the tridiagonal form of the GOE

Applying the tridiagonalization to GOE, we obtain the following random matrix model.

**Theorem 2.1.** *Let  $W$  be an  $n \times n$  GOE matrix (real symmetric) with variances chosen so that each off-diagonal entry has variance  $1/2$  and each diagonal entry has variance 1. Then there exists an orthogonal matrix  $Q$  such that*

$$W = Q^\top T Q,$$

where  $T$  is a real symmetric tridiagonal matrix

$$T = \begin{pmatrix} d_1 & \alpha_1 & 0 & \cdots \\ \alpha_1 & d_2 & \alpha_2 & \ddots \\ 0 & \alpha_2 & d_3 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \quad (2.1)$$

and the random variables  $\{d_i, \alpha_j\}_{1 \leq i \leq n, 1 \leq j \leq n-1}$  are mutually independent, with

$$d_i \sim \mathcal{N}(0, 1), \quad \alpha_j = \sqrt{\frac{\chi_{n-j}^2}{2}},$$

where  $\chi_\nu^2$  is a chi-square distribution with  $\nu$  degrees of freedom.

**Remark 2.2** (Chi-square distributions). The *chi-square distribution* with  $\nu$  degrees of freedom, denoted by  $\chi_\nu^2$ , is a fundamental distribution in statistics and probability theory. It arises naturally as the distribution of the sum of the squares of  $\nu$  independent standard normal random variables. Formally, if  $Z_1, Z_2, \dots, Z_\nu$  are independent random variables with  $Z_i \sim \mathcal{N}(0, 1)$ , then the random variable

$$Q = \sum_{i=1}^{\nu} Z_i^2$$

follows a chi-square distribution with  $\nu$  degrees of freedom, i.e.,  $Q \sim \chi_\nu^2$ . In the context of Theorem 2.1, the  $\alpha_j$ 's can be called *chi random variables*.

The parameter  $\nu$  does not need to be an integer, and the chi-square distribution is well defined for any positive real  $\nu$ , for example, by continuation of the density formula. The probability density is

$$f(x) = \frac{1}{2^{\nu/2} \Gamma(\nu/2)} x^{\nu/2-1} e^{-x/2}, \quad x \geq 0.$$

*Proof of Theorem 2.1.* In the process of tridiagonalization, we apply Householder reflections. Note that the diagonal entries stay fixed, and we only change the off-diagonal entries. Let us consider these off-diagonal entries.

In the first step, we apply the reflection in  $\mathbb{R}^{n-1}$  to turn the column vector  $(a_{2,1}, a_{3,1}, \dots, a_{n,1})$  into a vector parallel to  $(1, 0, \dots, 0) \in \mathbb{R}^{n-1}$ . Since the Householder reflection is orthogonal, it preserves lengths. So,

$$\alpha_1 = \sqrt{a_{21}^2 + a_{31}^2 + \dots + a_{n1}^2}, \quad a_{i1} \sim \mathcal{N}(0, \frac{1}{2}).$$

This implies that  $\alpha_1$  has the desired chi distribution. The distribution of the other entries is obtained similarly by the recursive application of the Householder reflections.

Note that  $\alpha_j$ 's and  $d_i$ 's depend on nonintersecting subsets of the matrix entries, so they are independent. This completes the proof.  $\square$

## 2.2 Dumitriu–Edelman $G\beta E$ tridiagonal random matrices

Let us define a general  $\beta$  extension of the tridiagonal model for the GOE.

**Definition 2.3.** Let  $\beta > 0$  be a parameter. The tridiagonal  $G\beta E$  is a random  $n \times n$  tridiagonal real symmetric matrix  $T$  as in (2.1), where  $d_i \sim \mathcal{N}(0, 1)$  are independent standard Gaussians, and

$$\alpha_j \sim \frac{1}{\sqrt{2}} \chi_{\beta(n-j)}, \quad 1 \leq j \leq n-1,$$

are chi-distributed random variables.

We showed that for  $\beta = 1$ , the  $G\beta E$  is the tridiagonal form of the GOE random matrix model. The same holds for the two other classical betas:

**Proposition 2.4** (Without proof). *For  $\beta = 2$ , the  $G\beta E$  is the tridiagonal form of the GUE random matrix model, which is the random complex Hermitian matrix with Gaussian entries and maximal independence. Similarly, for  $\beta = 4$ , the  $G\beta E$  is the tridiagonal form of the GSE random matrix model.*

Moreover, for all  $\beta$ , the joint eigenvalue density of  $G\beta E$  is explicit:

**Theorem 2.5** ([DE02]). *Let  $T$  be a  $G\beta E$  matrix as in Definition 2.3. Then the joint eigenvalue density is given by*

$$p(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_{n,\beta}} e^{-\frac{1}{2} \sum_{i=1}^n \lambda_i^2} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^\beta, \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n.$$

This theorem is also given without proof. The proof involves linear algebra and computation of the Jacobians of the change of variables from the matrix entries to the eigenvalues in the tridiagonal setting. It can be found in the original paper [DE02].

## 2.3 The case $\beta = 2$

For many questions involving *local eigenvalue statistics*, the case  $\beta = 2$  (the GUE, Gaussian Unitary Ensemble) is the most tractable. This is because the joint density of the eigenvalues admits a determinantal structure coming from a *square* Vandermonde factor  $\prod_{i < j} (\lambda_i - \lambda_j)^2$  and the Gaussian exponential  $\exp(-\frac{1}{2} \sum \lambda_j^2)$ . Moreover, for  $\beta = 2$ , the random matrix model and its correlation functions can be expressed explicitly through determinants involving *orthogonal polynomials*, namely, the *Hermite polynomials*.

**Proposition 2.6** (Joint density for GUE and orthogonal polynomials). *Consider the GUE (Gaussian Unitary Ensemble) random matrix model, i.e. an  $n \times n$  complex Hermitian matrix whose entries are i.i.d. up to the Hermitian condition, with each off-diagonal entry distributed as  $\mathcal{N}(0, \frac{1}{2}) + i\mathcal{N}(0, \frac{1}{2})$  and each diagonal entry  $\mathcal{N}(0, 1)$ . The ordered eigenvalues  $\lambda_1 \geq \dots \geq \lambda_n$  (or, without ordering, thought of as an unordered set) satisfy the joint probability density*

$$p(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_{n,2}} \prod_{j=1}^n e^{-\frac{1}{2}\lambda_j^2} \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j)^2, \quad (2.2)$$

where  $Z_{n,2}$  is a normalization constant.

Moreover, if  $\{\psi_k(\lambda)\}_{k=0}^\infty$  is the family of Hermite polynomials, orthonormal with respect to the measure  $w(\lambda) d\lambda = e^{-\lambda^2/2} d\lambda$  on  $\mathbb{R}$  (i.e.,  $\int_{-\infty}^\infty \psi_k(\lambda) \psi_\ell(\lambda) w(\lambda) d\lambda = \mathbf{1}_{k=\ell}$ ), then one can also write

$$p(\lambda_1, \dots, \lambda_n) = \text{const} \cdot \det \left[ \psi_{j-1}(\lambda_k) e^{-\frac{\lambda_k^2}{4}} \right]_{j,k=1}^n \det \left[ \psi_{j-1}(\lambda_k) e^{-\frac{\lambda_k^2}{4}} \right]_{j,k=1}^n \quad (2.3)$$

(the two determinants are identical, but let us keep this notation for future convenience).

The square determinant structure is extremely useful. It is precisely the  $\beta = 2$  counterpart of the squared Vandermonde factor  $\prod_{i < j} (\lambda_i - \lambda_j)^2$ .

**Remark 2.7** (Hermite polynomials). There are various normalizations of Hermite polynomials. In random matrix theory for the Gaussian ensembles, we often use the *probabilists' Hermite polynomials* (sometimes called  $\text{He}_k$ , but we use the notation  $H_k$ ). There are various normalizations due to the factor in the exponent of  $x^2$ .

A convenient definition for use with the weight  $e^{-x^2/2}$  is:

$$H_k(x) = (-1)^k e^{\frac{x^2}{2}} \frac{d^k}{dx^k} \left( e^{-\frac{x^2}{2}} \right), \quad k = 0, 1, \dots,$$

whose leading term is  $x^k$ . Polynomials with the leading coefficient 1 are called *monic*. The first few monic Hermite polynomials are

$$H_0(x) = 1, \quad H_1(x) = x, \quad H_2(x) = x^2 - 1, \quad H_3(x) = x^3 - 3x, \quad H_4(x) = x^4 - 6x^2 + 3.$$

The difference between  $H_k$  and  $\psi_k$  entering Proposition 2.6 is in a constant normalization, since  $H_k$  are monic but not orthonormal, while  $\psi_k$  are orthonormal but not monic.

*Sketch of the determinantal representation.* In brief, one observes that the factor  $\prod_{i < j} (\lambda_i - \lambda_j)$  is exactly the Vandermonde determinant  $\Delta(\lambda_1, \dots, \lambda_n) = \det [\lambda_k^{j-1}]_{j,k=1}^n$ . Next, the Vandermonde determinant is also equal to the determinant built out of any monic family of polynomials of the corresponding degrees (by linear transformations), and so we get the desired representation.  $\square$

We will work with Hermite polynomials and the determinantal structure in Proposition 2.6 in the next [Lecture 5](#)).

### 3 Wigner semicircle law via tridiagonalization

If  $W$  is an  $n \times n$  real Wigner matrix with entries of mean zero and variance 1 on the off-diagonal, then as  $n \rightarrow \infty$ , the empirical spectral distribution (ESD) of  $W/\sqrt{n}$  converges weakly almost surely to the Wigner semicircle distribution:

$$\mu_{\text{sc}}(dx) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{|x| \leq 2} dx.$$

We already derived this in [Lecture 2](#) by a direct combinatorial argument on the trace. Now we present another proof by using the tridiagonal form of  $W$ . The argument is conceptually simpler in some steps, because the matrix is sparser (only tridiagonal). At the same time, we will establish the Wigner semicircle law for the general  $\text{G}\beta\text{E}$  case (but only Gaussian), and thus it will apply to GUE and GSE.

#### 3.1 Moments for tridiagonal matrices

Consider the rescaled  $\text{G}\beta\text{E}$  matrix  $T/\sqrt{n}$ :

$$\frac{T}{\sqrt{n}} = \begin{pmatrix} d_1/\sqrt{n} & \alpha_1/\sqrt{n} & 0 & \cdots \\ \alpha_1/\sqrt{n} & d_2/\sqrt{n} & \alpha_2/\sqrt{n} & \ddots \\ 0 & \alpha_2/\sqrt{n} & d_3/\sqrt{n} & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix},$$

where  $d_i \sim \mathcal{N}(0, 1)$  and  $\alpha_j \sim \frac{1}{\sqrt{2}} \chi_{\beta(n-j)}$ . We want to show that the ESD of  $T/\sqrt{n}$  converges to the semicircle law. We will mostly consider expected traces of powers, and leave the analytic parts of the argument to the reader.

The  $k$ -th (random) moment of the ESD  $\frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i/\sqrt{n}}$  is

$$\frac{1}{n} \text{Tr} \left( \frac{T}{\sqrt{n}} \right)^k = \frac{1}{n^{1+\frac{k}{2}}} \sum_{i_1, \dots, i_k=1}^n t_{i_1, i_2} \cdots t_{i_k, i_1}, \quad (3.1)$$

where  $t_{ij}$  are the non-rescaled entries of  $T$ . But now  $t_{ij}$  is nonzero only if  $|i - j| \leq 1$ , i.e. the  $(i, j)$  entry is on the main or first super-/subdiagonal. In a closed product  $t_{i_1 i_2} \cdots t_{i_k i_1}$ , we thus get a *closed walk* in a linear graph on the vertex set  $\{1, 2, \dots, n\}$  with edges only between consecutive indices.

The relevant combinatorial objects encoding these walks are *Motzkin paths*: lattice walks in  $\mathbb{Z}_{\geq 0}^2$  starting at  $(0, 0)$ , ending at  $(k, 0)$ , and consisting of steps  $(1, 0)$ ,  $(1, 1)$ , and  $(1, -1)$ . The steps  $(1, 0)$  correspond to picking the diagonal element; steps  $(1, 1)$  correspond to picking  $i_{\ell+1} = i_\ell + 1$ , and steps  $(1, -1)$  correspond to  $i_{\ell+1} = i_\ell - 1$ . See [Figure 1](#) for an illustration of a Motzkin path.

Now, each term in the sum in (3.1) corresponds to a Motzkin path. Moreover, for each Motzkin path, there are  $O(n)$  summands corresponding to it. The number of Motzkin paths of length  $k$  is finite (independent of  $n$ ), so we need to look more closely at the asymptotics of the product in (3.1). This product involves chi random variables which depend on  $n$ , too.

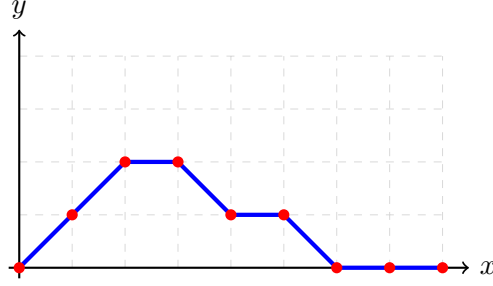


Figure 1: Example of a Motzkin path.

### 3.2 Asymptotics of chi random variables

One additional technical point in analyzing  $T/\sqrt{n}$  is to note that  $\alpha_j$  is roughly  $\sqrt{\beta(n-j)/2}$  for large  $n$ . Indeed, we have

$$\chi_\nu^2 = \sum_{i=1}^{\nu} Z_i^2, \quad \mathbb{E}[\chi_\nu^2] = \nu, \quad \text{Var}[\chi_\nu^2] = 2\nu.$$

Now, since we are dividing by  $\sqrt{n}$ , we have

$$\frac{\alpha_j}{\sqrt{n}} \sim \sqrt{\frac{\beta}{2}} \sqrt{1 - \theta}, \quad \theta = \frac{j}{n} \in [0, 1].$$

This estimate is valid in the “bulk” region, that is, when  $\theta$  is strictly between 0 and 1.

Let us make these estimates more precise. We have:

**Proposition 3.1** (Pointwise asymptotics in the bulk). *Fix any  $\delta \in (0, 1)$  and let  $j$  range so that  $\theta_j := j/n \in [\delta, 1 - \delta]$ . Then for each such  $j$ , we have<sup>1</sup>*

$$\frac{\alpha_j}{\sqrt{n}} = \sqrt{\beta \left(1 - \frac{j}{n}\right)} + O_p\left(\frac{1}{\sqrt{n}}\right),$$

In particular,

$$\lim_{n \rightarrow \infty} \frac{\alpha_j}{\sqrt{n}} = \sqrt{\beta(1 - \theta_j)} \quad \text{in probability.}$$

**Remark 3.2.** Outside the bulk region (i.e. very close to  $j = 0$  or  $j = n$ ), one would need a different statement to handle the case  $\beta(n - j)$  is not large. In our application, we only need the bulk behavior.

Meanwhile, on the diagonal,  $d_i/\sqrt{n}$  almost surely vanishes in the limit as  $n \rightarrow \infty$ , because  $d_i$  is standard Gaussian and does not depend on  $n$ .

<sup>1</sup>Here and below,  $O_p(\cdot)$  denotes a term that is stochastically bounded at the indicated order as  $n \rightarrow \infty$ . That is,  $X_n = O_p(a_n)$  means that for any  $\epsilon > 0$ , there exists  $M > 0$  such that  $\Pr(|X_n/a_n| > M) < \epsilon$  for all sufficiently large  $n$ .

**Remark 3.3** (Diagonal entries). For completeness, note that in many random matrix models with real/complex/Quaternion Gaussian entries, one also has diagonal terms  $d_i \sim N(0, 1)$  (independent of  $n$ ). Then

$$\frac{d_i}{\sqrt{n}} \rightarrow 0 \quad \text{almost surely,}$$

because  $d_i$  does not scale with  $n$ . Hence on the diagonal,  $\alpha_j$ -type behavior is irrelevant (or replaced by  $d_i$  that vanish in  $\sqrt{n}$  scaling).

### 3.3 Bounds on products in the bulk

We now give a couple of corollaries/lemmas that control products of the form

$$\prod_{j=a}^b \frac{\alpha_j}{\sqrt{n}}.$$

Whether you need deterministic bounds or stochastically dominating bounds, these simple expansions (and standard large-deviation arguments around the  $O_p(\cdot)$  notation) usually suffice.

**Corollary 3.4** (Finite product bound). *Fix a finite integer  $m$ , and let  $1 \leq j \leq m$  with  $j/n \in [0, 1]$ . Then*

$$\prod_{j=1}^m \frac{\alpha_j}{\sqrt{n}} = \prod_{j=1}^m \left( \sqrt{\beta(1 - \frac{j}{n})} + O_p\left(\frac{1}{\sqrt{n}}\right) \right) = \prod_{j=1}^m \sqrt{\beta(1 - \frac{j}{n})} \left( 1 + O_p\left(\frac{1}{\sqrt{n}}\right) \right).$$

*In particular, as  $n \rightarrow \infty$  (with  $m$  fixed),*

$$\prod_{j=1}^m \frac{\alpha_j}{\sqrt{n}} = \beta^{m/2} \prod_{j=1}^m \sqrt{1 - \frac{j}{n}} \left( 1 + O_p\left(\frac{1}{\sqrt{n}}\right) \right).$$

*Proof.* By Proposition 3.1 (even applied naively for small  $j$ ),

$$\frac{\alpha_j}{\sqrt{n}} = \sqrt{\beta(1 - \frac{j}{n})} + O_p\left(\frac{1}{\sqrt{n}}\right).$$

Hence each factor in the product equals  $\sqrt{\beta(1 - j/n)} \times [1 + O_p(\frac{1}{\sqrt{n}})]$ . Since  $m$  is fixed, the product of  $m$  such terms introduces only another  $1 + O_p(\frac{1}{\sqrt{n}})$  factor overall.  $\square$

**Lemma 3.5** (Product bounds on a bulk interval). *Let  $\delta \in (0, 1)$  be fixed, and consider indices  $j$  such that  $j/n \in [\delta, 1 - \delta]$ . For any integer interval  $[a, b] \subset \{\lceil \delta n \rceil, \dots, \lfloor (1 - \delta)n \rfloor\}$ , we have*

$$\prod_{j=a}^b \frac{\alpha_j}{\sqrt{n}} = \exp\left(\sum_{j=a}^b \ln\left[\sqrt{\beta(1 - \frac{j}{n})} + O_p\left(\frac{1}{\sqrt{n}}\right)\right]\right).$$

*Moreover, using the expansion*

$$\ln\left[\sqrt{\beta(1 - \frac{j}{n})} + \epsilon_j\right] = \ln\sqrt{\beta(1 - \frac{j}{n})} + O_p\left(\frac{|\epsilon_j|}{\sqrt{\beta(1 - j/n)}}\right),$$



one finds that with high probability,

$$\prod_{j=a}^b \frac{\alpha_j}{\sqrt{n}} = \prod_{j=a}^b \sqrt{\beta \left(1 - \frac{j}{n}\right)} \exp\left(O_p\left(\frac{b-a+1}{\sqrt{n}}\right)\right).$$

*Proof.* Applying Proposition 3.1 in the bulk region, each factor satisfies

$$\frac{\alpha_j}{\sqrt{n}} = \sqrt{\beta \left(1 - \frac{j}{n}\right)} + O_p\left(\frac{1}{\sqrt{n}}\right).$$

Denote  $\epsilon_j = O_p\left(\frac{1}{\sqrt{n}}\right)$ . Then

$$\ln\left[\sqrt{\beta \left(1 - \frac{j}{n}\right)} + \epsilon_j\right] = \ln \sqrt{\beta \left(1 - \frac{j}{n}\right)} + \ln\left[1 + \frac{\epsilon_j}{\sqrt{\beta(1-j/n)}}\right].$$

As long as  $j/n \in [\delta, 1 - \delta]$ , the denominator  $\sqrt{\beta(1-j/n)}$  is bounded away from zero, so

$$\ln\left[1 + \frac{\epsilon_j}{\sqrt{\beta(1-j/n)}}\right] = O_p\left(\frac{|\epsilon_j|}{\sqrt{\beta(1-j/n)}}\right) = O_p\left(\frac{1}{\sqrt{n}}\right).$$

Hence summing from  $j = a$  to  $j = b$  (and noting that  $b - a + 1 \leq n$  but is often smaller or on the order of  $n$ ) gives the claimed product expansion.  $\square$

**Remark 3.6.** Lemma 3.5 can be refined to a local law or large-deviation style bound where one shows that with overwhelming probability (exponentially high), the deviation from  $\prod_{j=a}^b \sqrt{\beta(1-j/n)}$  is at most  $\exp(C(b-a+1)/\sqrt{n})$  for some constant  $C$ . Such results require standard concentration inequalities for  $\chi^2$  random variables (or normal approximations) rather than just the  $O_p(\cdot)$  expansions.

## 4 Conclusion

In summary:

- For large  $n$ , each  $\alpha_j = \sqrt{\chi_{\beta(n-j)}^2}$  is on the order of  $\sqrt{\beta(n-j)}$  with a Gaussian fluctuation of size  $O_p(1)$ .
- Consequently,  $\alpha_j/\sqrt{n}$  concentrates near  $\sqrt{\beta(1-j/n)}$ , and the diagonal entries (if they are standard Gaussians independent of  $n$ ) vanish at order  $n^{-1/2}$ .
- Products  $\prod_{j=a}^b (\alpha_j/\sqrt{n})$  in the bulk regime are therefore very close (in probability or even with high probability) to the corresponding product of deterministic factors  $\prod_{j=a}^b \sqrt{\beta(1-j/n)}$ , up to multiplicative corrections of order  $\exp(O_p((b-a)/\sqrt{n}))$ .

## 4.1 Completing the proof: global semicircle behavior

Putting the above pieces together, we see that  $T/\sqrt{n}$  is asymptotic to

$$\sum_{\text{Motzkin paths of length } k} \prod_{\ell=1}^k \frac{t_{i_\ell, i_{\ell+1}}}{\sqrt{n}}$$

behaves—globally—like a “Jacobi-type matrix” whose main diagonal is small and whose sub-diagonal is roughly  $\sqrt{(n-j)/2n}$ . One can then do a continuum limit of the characteristic polynomial or use the combinatorial expansion (Motzkin path argument) or the resolvent method (Stieltjes transform approach) to find that the limiting eigenvalue distribution is exactly Wigner’s semicircle law. Thus:

**Theorem 4.1** (Semicircle Law via Tridiagonalization). *Let  $W$  be an  $n \times n$  real Wigner matrix with off-diagonal variance 1. Then  $W$  can be unitarily conjugated into a real symmetric tridiagonal matrix  $T$  whose diagonal entries are i.i.d.  $\mathcal{N}(0, 1)$  and subdiagonal entries are chi-type with typical value  $\sqrt{\frac{n-j}{2}}$ . By examining  $T/\sqrt{n}$  and counting or bounding its trace moments (or via its resolvent), one concludes that the ESD converges almost surely to the semicircle distribution on  $[-2, 2]$ .*

This completes the second proof of Wigner’s semicircle law.

## 5 Wigner semicircle law via Stieltjes transform

BEGIN DRAFT

## 6 Characteristic Polynomial and Three-Term Recurrence

Once we have  $T = (t_{ij})$  tridiagonal, the characteristic polynomial of  $T$  takes a well-known form governed by a three-term recurrence. Denote

$$T - \lambda I = \begin{pmatrix} d_1 - \lambda & \alpha_1 & 0 & \cdots & 0 \\ \alpha_1 & d_2 - \lambda & \alpha_2 & \cdots & 0 \\ 0 & \alpha_2 & d_3 - \lambda & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \alpha_{n-1} \\ 0 & 0 & \cdots & \alpha_{n-1} & d_n - \lambda \end{pmatrix}.$$

**Definition 6.1** (Characteristic Polynomials  $p_k(\lambda)$ ). For  $1 \leq k \leq n$ , let  $T_k$  be the top-left  $k \times k$  submatrix of  $T$ . Define

$$p_k(\lambda) = \det(T_k - \lambda I_k).$$

Moreover, set  $p_0(\lambda) := 1$  by convention.

**Lemma 6.2** (Three-Term Recurrence). *Let  $p_k(\lambda)$  be as above, with  $p_1(\lambda) = d_1 - \lambda$ . Then for  $k \geq 1$ ,*

$$p_{k+1}(\lambda) = (d_{k+1} - \lambda)p_k(\lambda) - \alpha_k^2 p_{k-1}(\lambda).$$

*Idea of proof.* One checks the base case  $k = 1, 2$  directly by computing determinants of  $1 \times 1$  and  $2 \times 2$  blocks. For the general step, expand  $\det(T_{k+1} - \lambda I_{k+1})$  by minors along the last row or column. The block structure ensures exactly the claimed recurrence.  $\square$

**Remark 6.3.** This is analogous to recurrences in orthogonal polynomials, e.g. the three-term recursion for polynomials orthonormal with respect to certain measures. In fact, the polynomial  $p_n(\lambda)$  can be viewed as an orthogonal polynomial if  $\alpha_j > 0$ . This interplay is vital in random matrix theory.

## 7 Semicircle Law via the Tridiagonal Form

We now present a fairly detailed sketch (or roadmap) for how the tridiagonal form yields the Wigner semicircle law. Recall we aim to show:

$$\frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i/\sqrt{n}} \longrightarrow \mu_{\text{sc}}(dx) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{|x| \leq 2} \quad \text{almost surely.}$$

### 7.1 The Scaling and Law of Large Numbers in $\alpha_j$

Write

$$T = \begin{pmatrix} d_1 & \alpha_1 & & \\ \alpha_1 & d_2 & \alpha_2 & \\ & \alpha_2 & d_3 & \ddots \\ & & \ddots & \ddots \end{pmatrix}.$$

In the Dumitriu–Edelman setting,  $\alpha_j^2 = \frac{1}{2} \chi_{(n-j)}^2$ , so  $E[\alpha_j^2] = (n-j)/2$  and  $\alpha_j^2$  is tightly concentrated around its mean for large  $n$ . Precisely, for any  $\varepsilon > 0$ ,

$$\Pr\left(\left|\alpha_j^2 - \frac{n-j}{2}\right| > \varepsilon(n-j)\right) \leq e^{-c(n-j)}$$

for some  $c > 0$ , or a similar bound from large deviation estimates. Hence, with overwhelming probability,

$$\alpha_j \approx \sqrt{\frac{n-j}{2}}.$$

### 7.2 Diagonal vs. Subdiagonal Entries

When we examine  $\frac{1}{\sqrt{n}} T$ , the diagonal entries become  $\frac{d_i}{\sqrt{n}}$ . Since  $d_i \sim \mathcal{N}(0, 1)$  i.i.d., with probability going to 1 these entries lie within  $O(n^{-1/2})$ , so they vanish in the large- $n$  limit. Meanwhile,

$$\frac{\alpha_j}{\sqrt{n}} \approx \sqrt{\frac{n-j}{2n}} \approx \sqrt{\frac{1-j/n}{2}}.$$

In the bulk region (i.e.  $j \approx \theta n$  for  $\theta \in (0, 1)$ ), we have  $\alpha_j/\sqrt{n} \approx \sqrt{\frac{1-\theta}{2}}$ . In short, the subdiagonal elements (scaled by  $1/\sqrt{n}$ ) remain of order 1, while the diagonal elements vanish.

### 7.3 Characteristic Polynomial and Recurrence Analysis

Denote

$$p_n(\lambda) = \det\left(\frac{1}{\sqrt{n}}T - \lambda I\right).$$

Equivalently,  $\lambda$  is an eigenvalue of  $\frac{1}{\sqrt{n}}T$  if and only if  $\mu = \sqrt{n}\lambda$  is an eigenvalue of  $T$ . We want to understand the distribution of the roots  $\lambda_i(\frac{1}{\sqrt{n}}T)$  as  $n \rightarrow \infty$ . The three-term recurrence for  $p_n(\lambda)$  can be viewed in the limit as  $n \rightarrow \infty$ , turning into an integral equation for the Stieltjes transform of the limiting measure.

**Stieltjes Transform Argument (Sketch).** Set

$$G_n(z) = \frac{1}{n} \text{Tr}\left(\left(\frac{1}{\sqrt{n}}T - z\right)^{-1}\right),$$

the Stieltjes transform. As  $n$  grows, the main input is that  $\alpha_j \approx \sqrt{\frac{1-j/n}{2}}$ ; substituting this approximate profile into the recursion for Green's functions (a linear difference equation akin to orthogonal polynomials) yields a limiting functional equation. Solving that equation for  $G(z)$  leads to

$$G(z) = \frac{z \pm \sqrt{z^2 - 4}}{2},$$

with the appropriate branch cut. This is the well-known Stieltjes transform for the semicircle law on  $[-2, 2]$ . (See advanced texts, e.g. *Deift* or *Tao's* books on RMT for a full rigorous derivation.)

**Moment Argument (Sketch).** One can also proceed by computing or bounding the moments  $E\left[\frac{1}{n} \text{Tr}\left(\frac{1}{\sqrt{n}}T\right)^k\right]$  and showing that they match the semicircle moments. Indeed, as  $n \rightarrow \infty$ , the diagonal part becomes negligible, while the subdiagonal structure essentially forces closed loops in the sum expansions, reproducing the Catalan numbers that appear in the Wigner moment method (Lectures 1–2). The advantage of tridiagonalization is a simpler combinatorial interpretation of the entries used in each closed loop.

### 7.4 Conclusion: Wigner's Semicircle Law

Putting these ingredients together, we conclude that with probability 1,

$$\frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i(\frac{1}{\sqrt{n}}W)} \longrightarrow \mu_{\text{sc}}(dx) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{|x| \leq 2}$$

This completes the proof by tridiagonal methods.

**Remark 7.1** (Universality). The argument here specifically used the Gaussian Wigner distribution for off-diagonal entries to get an explicit  $\chi^2$  structure. However, the final result (semicircle law) remains true under vastly weaker assumptions on the entries. The *universality principle* states that the global spectral behavior is insensitive to fine details of the distribution (e.g. it depends primarily on the first two moments).

## 8 Exercises (Due 2025-02-28)

Below are problems elaborating on the main concepts. They range from verifying computations to exploring deeper aspects of tridiagonalization and simulations.

### 1. Detailed Householder Steps and Reflection Properties

- (a) **Constructing a Reflection that Maps One Vector to Another.** Let  $x, y \in \mathbb{R}^n$  be nonzero. Show how to pick  $v$  so that  $Hx = y$ , where  $H = I - 2\frac{vv^\top}{\|v\|^2}$  is a Householder reflection. Why does  $H$  remain orthogonal?
- (b) **Eliminating Below-Diagonal Entries.** In the first step of the tridiagonalization algorithm, pick a reflection  $H_1$  that modifies the subspace spanned by  $a_{21}, \dots, a_{n1}$ . Show explicitly how  $H_1$  zeroes out all these subdiagonal entries while keeping  $a_{11}$  unchanged (aside from possibly a sign).
- (c) **Symmetry and Superdiagonal.** Why does  $H_1(A)H_1^\top$  also have the corresponding superdiagonal entries zeroed out in row 1? (Hint: Use  $A$  is symmetric:  $A_{ij} = A_{ji}$ .)

### 2. Three-Term Recurrence Warmup

- (a) **Base Cases.** Compute  $p_1(\lambda)$  and  $p_2(\lambda)$  for a  $2 \times 2$  matrix

$$\begin{pmatrix} d_1 - \lambda & \alpha_1 \\ \alpha_1 & d_2 - \lambda \end{pmatrix}.$$

Verify that  $p_2(\lambda) = (d_2 - \lambda)p_1(\lambda) - \alpha_1^2 p_0(\lambda)$ .

- (b) **Inductive Step.** Using determinant expansion along the  $(k+1)$ -st row or column, outline why  $p_{k+1}(\lambda) = (d_{k+1} - \lambda)p_k(\lambda) - \alpha_k^2 p_{k-1}(\lambda)$  for  $k \geq 1$ .

### 3. Tridiagonal Model for GOE and GUE

- (a) **Real Case (GOE).** Starting with a real Wigner matrix  $W$  ( $X_{ij} \sim \mathcal{N}(0, 1)$  for  $i < j$ ,  $X_{ii} \sim \mathcal{N}(0, 2)$ ), show that the Householder steps produce a tridiagonal  $T$  whose diagonal entries are  $d_i \sim \mathcal{N}(0, 1)$  and subdiagonals  $\alpha_j^2 = \frac{1}{2}\chi_{(n-j)}^2$ .
- (b) **Complex Case (GUE).** For a complex Hermitian  $W$ , the off-diagonal entries are  $\mathcal{N}(0, \frac{1}{2}) + i\mathcal{N}(0, \frac{1}{2})$ . Sketch how the same approach yields a complex Hermitian tridiagonal form with  $d_i$  real normal and  $\alpha_j$  drawn from appropriate  $\chi$  distributions. (You do not need a fully rigorous proof; just highlight the changes in dimension counting for real vs. complex parts.)

### 4. Semicircle Law via Stieltjes Transform

- (a) **Defining the Green's Function.** For  $z \in \mathbb{C} \setminus \mathbb{R}$ , let

$$G_n(z) = \frac{1}{n} \text{Tr} \left( \left( \frac{1}{\sqrt{n}} T - zI \right)^{-1} \right).$$

Argue (informally) that  $G_n(z)$  converges to a limit  $G(z)$  which must satisfy an algebraic equation derived from the tridiagonal structure and the typical size of  $\alpha_j$ .

- (b) **Solving for  $G(z)$ .** Show that  $G(z)$  satisfies

$$G(z)^2 + z G(z) + 1 = 0,$$

and deduce that

$$G(z) = \frac{-z + \sqrt{z^2 - 4}}{2}.$$

Hence identify the imaginary part of  $G(z)$  on the real interval  $(-2, 2)$ , concluding that the limiting distribution is the semicircle law.

## 5. Simulation of Tridiagonal vs. Dense Wigner

Write a small program (in Python, MATLAB, or another language):

- Generate a dense GOE matrix  $W$  of size  $n = 1000$  and scale by  $\frac{1}{\sqrt{n}}$ . Compute eigenvalues and plot a histogram.
- Generate the corresponding tridiagonal matrix  $T$  from the Dumitriu–Edelman approach (diagonal  $d_i \sim \mathcal{N}(0, 1)$ , subdiag  $\alpha_j = \sqrt{\frac{1}{2} \chi_{n-j}^2}$ ). Compute eigenvalues of  $\frac{1}{\sqrt{n}}T$  and compare histograms. They should match well with the semicircle shape for sufficiently large  $n$ .
- (Optional) Investigate how large  $n$  must be before the histogram looks convincingly semicircular.

## 6. Wishart and MANOVA Exercises

- Wishart from Data Matrix.** Generate an  $n \times m$  data matrix  $X$  with iid  $\mathcal{N}(0, 1)$ . Form  $W = X X^\top$ . Plot the normalized eigenvalues  $\lambda_i$  vs. the Marchenko–Pastur distribution  $\mu_{MP}$  (with aspect ratio  $m/n$ ). Discuss approximate agreement for moderate  $n, m$ .
- Jacobi (MANOVA).** Let  $X$  be  $n \times t$  and  $Y$  be  $k \times t$  independent  $\mathcal{N}(0, 1)$ . Form the matrix  $M = (X X^\top + Y Y^\top)^{-1}(X X^\top)$  and find its eigenvalues in  $[0, 1]$ . Plot their histogram and compare with the Jacobi Beta distribution.

## 9 Further Reading and Next Steps

- **Local Laws and Universality.** After establishing the global semicircle distribution, one may delve into local spectral laws (e.g. the sine kernel or GOE Tracy–Widom distribution at the edge). The *local semicircle law* refines the analysis of the Green’s function in small intervals.
- **Dyson Brownian Motion.** Another approach interprets the eigenvalues as particles with logarithmic repulsion and uses stochastic differential equations to show that equilibrium distributions converge to the  $\beta$ -ensembles.

- **Other Ensembles.** Beyond Wigner, Wishart, and Jacobi, one finds many integrable and combinatorial random matrices (e.g. discrete random partitions, polynomial ensembles, random tilings). Orthogonal polynomial techniques remain central in these broader contexts.
- **References.** - I. Dumitriu and A. Edelman, *Matrix models for beta ensembles*, J. Math. Phys., 2002. - T. Tao, *Topics in Random Matrix Theory*, 2012. - P. Deift, *Orthogonal Polynomials and Random Matrices*, 2000. - M. Mehta, *Random Matrices*, 3rd ed., Elsevier, 2004. - T. Anderson, *An Introduction to Multivariate Statistical Analysis*, for Wishart and MANOVA.

**End of Lecture 4.** In this expanded lecture, we detailed how any real symmetric matrix is tridiagonalized, derived the three-term recurrence for the characteristic polynomial, and showed how this structure underlies a clean proof of Wigner’s semicircle law for random Wigner matrices. We also saw how these ideas extend to Wishart and Jacobi ensembles, bridging us toward the broader  $\beta$ -ensemble world. Upcoming lectures will further explore local eigenvalue statistics, universality results, and connections with integrable probability.

## D Problems (due 2025-02-28)

### D.1 Eigenvalue density of $G\beta E$

Read and understand the main principles of the proof of Theorem 2.5 in [DE02].

### D.2 Chi-square mean and variance

Let  $X$  be a random variable with  $\chi_\nu^2$  distribution. Compute the mean and variance of  $X$ . (If  $\nu$  is an integer, you can use the fact that  $\chi_\nu^2$  is a sum of  $\nu$  independent squares of standard normal random variables. How to extend this to non-integer  $\nu$ ?)

## Notes for the lecturer

Add / finish up the discussion about tridiagonalization. The notes for L3 are updated, but mention it here.

The reflection is mapping any vector into any other vector (unit vectors); Also, we apply it in  $(n-1)$ -dimensional space in the first pass.

maybe make a simulation for Wishart and MANOVA in L4

### D.3 Characteristic Polynomial and Three-Term Recurrence

Consider  $p_n(\lambda) = \det(T - \lambda I)$ . Because  $T$  is tridiagonal, we have the classical three-term recurrence for these characteristic polynomials:

$$p_0(\lambda) := 1, \quad p_1(\lambda) := d_1 - \lambda,$$

$$p_{k+1}(\lambda) = (d_{k+1} - \lambda)p_k(\lambda) - \alpha_k^2 p_{k-1}(\lambda), \quad (k = 1, \dots, n-1).$$

The eigenvalues of  $T$  are precisely the roots of  $p_n(\lambda)$ .

## D.4 Sketch of the Semicircle Limit Proof

We want to show that the empirical distribution

$$L_n = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}$$

(where  $\lambda_1, \dots, \lambda_n$  are the eigenvalues of  $T$ ) converges weakly to the semicircle law

$$\mu_{\text{sc}}(dx) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{|x| \leq 2} dx$$

as  $n \rightarrow \infty$ . A typical outline:

1. **Law of Large Numbers for  $\alpha_j$ .** Since  $\alpha_j^2 = \frac{1}{2} \chi_{n-j}^2$  has mean  $\frac{n-j}{2}$ , it is typically of order  $n/2$ . More precisely, for large  $n$ ,  $\alpha_j \approx \sqrt{\frac{n-j}{2}}$  with high probability.
2. **Scaling by  $\sqrt{n}$ .** One rescales  $T$  by  $\frac{1}{\sqrt{n}}$ . This gives subdiagonal entries

$$\frac{\alpha_j}{\sqrt{n}} \approx \sqrt{\frac{n-j}{2n}} \approx \sqrt{\frac{1-j/n}{2}},$$

while the diagonal entries become  $\frac{d_i}{\sqrt{n}}$ , which vanish in the large- $n$  limit. So effectively, the subdiagonal structure drives the main spectral behavior in the bulk, producing the semicircle shape in the limit.

3. **Orthogonal Polynomial / Recurrence Analysis.** The polynomial  $p_n(\lambda)$  satisfies a discrete three-term recurrence whose “continuum limit” yields a certain integral equation (specifically the Stieltjes transform for the measure) whose solution is precisely the semicircle distribution. In more detailed treatments, one shows that the moments or the Cauchy transform of  $L_n$  converge to that of  $\mu_{\text{sc}}$ . The relevant PDE or integral equation is exactly solvable, producing the semicircle.

Hence, with probability 1, as  $n \rightarrow \infty$ , the empirical spectrum of  $\frac{1}{\sqrt{n}} W$  converges to the semicircle distribution on  $[-2, 2]$ . This precisely recovers *Wigner’s semicircle law*.

**Remark D.1** (Extensions). A very similar approach works for the Gaussian Unitary Ensemble ( $\beta = 2$ ), leading to a random *complex Hermitian* tridiagonal matrix. For  $\beta = 4$ , there is a quaternionic block-tridiagonal model. All of these point toward the same semicircle law for the global spectral distribution.

## References

- [DE02] I. Dumitriu and A. Edelman, *Matrix models for beta ensembles*, Journal of Mathematical Physics **43** (2002), no. 11, 5830–5847. arXiv:math-ph/0206043. [↑4](#), [15](#)

L. PETROV, UNIVERSITY OF VIRGINIA, DEPARTMENT OF MATHEMATICS, 141 CABELL DRIVE, KERCHOF HALL, P.O. BOX 400137, CHARLOTTESVILLE, VA 22904, USA  
E-mail: lenia.petrov@gmail.com