CHAPTER 3

GOE and GUE

We quickly recall that a GUE matrix can be defined in the following three equivalent ways. We leave it to the reader to make the three analogous statements for GOE.

In the previous chapters, GOE and GUE matrices appeared merely as special cases of Wigner matrices for which computations were easier. However they have a great many neat properties not shared by other Wigner matrices. The main fact is that the exact density of eigenvalues of GOE and GUE can be found explicitly! And even more surprisingly, these exact densities have a nice structure that make them amenable to computations. Many results that are true for general Wigner matrices are much harder to prove in general but fairly easy for these two cases. Crucial to the "integrability" properties of GOE and GUE are their invariance under orthogonal and unitary conjugations respectively.

Exercise 46. (a) Let X and Y be $n \times n$ GUE and GOE matrices respectively. Then, for any fixed $U \in \mathcal{U}(n)$ and $P \in \mathcal{O}(n)$, we have $U^*XU \stackrel{d}{=} X$ and $P^tYP \stackrel{d}{=} Y$.

(b) If X is a random matrix such that $X_{i,j}$, $i \le j$ are independent real valued entries and suppose that $PXP^t \stackrel{d}{=} X$ for all $P \in \mathcal{O}(n)$, then show that X has the same distribution as $c\tilde{X}$ where c is a constant and $\tilde{X}$ is a GOE matrix. The analogous statement for unitary invariance is also true.

Remark 47. This is analogous to the following well known fact. Let X be a random vector in $\mathbb{R}^n$. Then the following are equivalent.

(1) $X \sim N_n(0, \sigma^2 I)$ for some σ^2.

(2) X_k are independent and $PX \stackrel{d}{=} X$ for any $P \in \mathcal{O}(n)$.

To see that the second implies the first, take for P an orthogonal matrix whose first column is $(1/\sqrt{2}, 1/\sqrt{2}, 0, \ldots, 0)$ to get $X_1 \stackrel{d}{=} (X_1 + X_2)/\sqrt{2}$. Further, X_1, X_2 are i.i.d - independence is given, and choosing P to be a permutation matrix we get identical distributions. It is well known that the only solutions to this distributional equation are the $N(0, \sigma^2)$ distributions. If not convinced, use characteristic functions or otherwise show this fact.

What is the use of unitary or orthogonal invariance? Write the spectral decomposition of a GUE matrix $X = VDV^*$. For any fixed $U \in \mathcal{U}(n)$, then $UXU^* = (UV)D(UV)^*$. By the unitary invariance, we see that VDV^* has the same distribution as $(UV)D(UV)^*$. This suggests that V and D are independent. The only hitch in this reasoning is that the spectral decomposition is not exactly unique, but it can be taken care of[1]

[1]The eigenspace for a given eigenvalue is well-defined. This is the source of non-uniqueness. The set S of Hermitian matrices having distinct eigenvalues is a dense open set in the space of all Hermitian matrices. Therefore, almost surely, a GUE matrix has no eigenvalues of multiplicity more than one (explain why). However, even when the eigenspace is one dimensional, we can multiply the eigenvector by $e^{i\theta}$ for some $\theta \in \mathbb{R}$ and that leads to non-uniqueness. To fix this, let $\mathcal{D}(n)$ be the group of $n \times n$ diagonal unitary matrices and consider the quotient space $Q = \mathcal{U}(n)/\mathcal{D}(n)$ consisting of right cosets. Then, the mapping $X \to ([V], D)$ is one to one and onto on S. Now observe that for any U, $([UV], D) \stackrel{d}{=} ([V], D)$ and hence $[V]$ and D are independent.

1. Tridiagonalization

Let A be an $n \times n$ GOE. Write it as

$$A = \begin{bmatrix} a & \mathbf{u}^t \\ \mathbf{u} & B \end{bmatrix}$$

so that $a \sim N(0,2)$, $\mathbf{u} \sim N_{n-1}(0,I)$, $B \sim \mathrm{GOE}_{n-1}$, and all three are independent. Condition on $\mathbf{u}$. Then pick any orthogonal matrix $P \in O(n-1)$ such that $P\mathbf{u} = \|u\|\mathbf{e}_1$. To be specific, we can take the transformation defined by

$$P\mathbf{v} = \mathbf{v} - 2\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle}\mathbf{w}, \quad \text{with } \mathbf{w} = \mathbf{u} - \mathbf{e}_1.$$

For any $\mathbf{w} \neq 0$, the transformation defined on the left is the reflection across the hyperplane perpendicular to $\mathbf{w}$. These are also referred to as *Householder reflections*. Check that P is indeed unitary and that $P\mathbf{u} = \mathbf{e}_1$.

Since P depends on $\mathbf{u}$ and B is independent of U, the orthogonal invariance of GOE shows that $A_1 := P^tBP \stackrel{d}{=} B$, that is A_1 is a GOE matrix. Also A_1 is independent of $\mathbf{u}$ and a. Thus,

$$C := \begin{bmatrix} 1 & \mathbf{0}^t \\ \mathbf{0} & P^t \end{bmatrix} \begin{bmatrix} a & \mathbf{u}^t \\ \mathbf{u} & B \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^t \\ \mathbf{0} & P \end{bmatrix} = \begin{bmatrix} a & r_1\mathbf{e}_1^t \\ r_1\mathbf{e}_1 & A_1 \end{bmatrix}$$

where $A_1 \sim GOE_{n-1}$, $a \sim N(0,1)$ and $r_1 = \|u\|$ are all independent. Since C is an orthogonal conjugation of A, the eigenvalues of A and C are exactly the same. Observe that $C_{j,1} = C_{1,j} = 0$ for $2 \leq j \leq n$. Note that $r_1^2 = \|u\|^2$ has χ^2_{n-1} distribution.

Now A_1 is a GOE matrix of one less order. We can play the same game with A_1 and get a matrix D which is conjugate to A_1 but has $D_{1,j} = D_{j,1} = 0$ for $2 \leq j \leq n-1$. Combining with the previous one, we get

$$C_2 := \begin{bmatrix} a & r_1 & \mathbf{0}^t \\ r_1 & a' & r_2\mathbf{e}_1^t \\ \mathbf{0} & r_2\mathbf{e}_1 & D \end{bmatrix}$$

with the following properties. C_2 is conjugate to A and hence has the same eigenvalues. $D \sim \mathrm{GOE}_{n-2}$, $a, a' \sim N(0,2)$, $r_1^2 \sim \chi^2_{n-1}$, $r_2^2 \sim \chi^2_{n-2}$, and all these are independent.

It is clear that this procedure can be continued and we end up with a tridiagonal matrix that is orthogonally conjugate to A and such that

$$T_n = \begin{bmatrix} a_1 & b_1 & 0 & 0 & \dots & 0 \\ b_1 & a_2 & b_2 & 0 & \dots & 0 \\ 0 & b_2 & a_3 & b_3 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & b_{n-2} & a_{n-1} & b_{n-1} \\ 0 & \dots & 0 & 0 & b_{n-1} & a_n \end{bmatrix} \tag{20}$$

where $a_k \sim N(0,2)$, $b_k^2 \sim \chi^2_{n-k}$, and all these are independent.

Exercise 48. If A is an $n \times n$ GUE matrix, show that A is conjugate to a tridiagonal matrix T as in (20) where a_k, b_k are all independent, $a_k \sim N(0,1)$ and $b_k^2 \sim \mathrm{Gamma}(n-k,1)$.

Recall that χ^2_p is the same as $\mathrm{Gamma}(\frac{p}{2}, \frac{1}{2})$ or equivalently, the distribution of $2Y$ where $Y \sim \mathrm{Gamma}(\frac{p}{2})$. Thus, we arrive at the following theorem[2].

[2]The idea of tridiagonalizing the GOE and GUE matrices was originally due to Hale Trotter **?**. Part of his original motivation was to give a simple proof of the semicircle law for GOE and GUE matrices.

Theorem 49. *Let T be a tridiagonal matrix as in* (20) *where a_k, b_k are all independent.*

(1) If $a_k \sim N(0,2)$ and $b_k^2 \sim \chi^2_{n-k}$, then the vector of eigenvalues of T has the same distribution as the vector of eigenvalues of a GOE_n matrix.

(2) If $a_k \sim N(0,2)$ and $b_k^2 \sim \chi^2_{2(n-k)}$, then the vector of eigenvalues of T has the same distribution as the eigenvalues of a GUE_n matrix scaled by a factor of $\sqrt{2}$.

2. Tridiagonal matrices and probability measures on the line

Our objective is to find eigenvalue density for certain random matrices, and hence we must find $n-1$ auxiliary parameters in addition to the n eigenvalues (since there are $2n-1$ parameters in the tridiagonal matrix) to carry out the Jacobian computation. The short answer is that if UDU^* is the spectral decomposition of the tridiagonal matrix, then $p_j = |U_{1,j}|^2$, $1 \le j \le n-1$ are the right parameters to choose. However, there are many conceptual reasons behind this choice and we shall spend the rest of this section on these concepts.

Fix $n \ge 1$ and write $T = T(a,b)$ for the real symmetric $n \times n$ tridiagonal matrix with diagonal entries $T_{k,k} = a_k$ for $1 \le k \le n$ and $T_{k,k+1} = T_{k+1,k} = b_k$ for $1 \le k \le n-1$.

Let $\mathcal{T}_n$ be the space of all $n \times n$ real symmetric tridiagonal matrices and let $\mathcal{T}_n^0$ be those $T(a,b)$ in $\mathcal{T}_n$ with n distinct eigenvalues. Let $\mathcal{P}_n$ be the space of all probability measures on $\mathbb{R}$ whose support consists of at most n distinct points and let $\mathcal{P}_n^0$ be those elements of $\mathcal{P}_n$ whose support has exactly n distinct points.

Tridiagonal matrix to probability measure: Recall that the *spectral measure* of a Hermitian operator T at a vector $\mathbf{v}$ is the unique measure ν on $\mathbb{R}$ such that $\langle T^p \mathbf{v}, \mathbf{v}\rangle = \int x^p \nu(dx)$ for all $p \ge 0$. For example, if T is a real symmetric matrix, write its spectral decomposition as $T = \sum_{k=1}^n \lambda_k \mathbf{u}_k \mathbf{u}_k^*$. Then $\{\mathbf{u}_k\}$ is an ONB of $\mathbb{R}^n$ and λ_k are real. In this case, the spectral decomposition of T at any $\mathbf{v} \in \mathbb{R}^n$ is just $\nu = \sum_{k=1}^n |\langle \mathbf{v}, \mathbf{u}_k\rangle|^2 \delta_{\lambda_k}$. Thus $\nu \in \mathcal{P}_n$ (observe that the support may have less than n points as eigenvalues may coincide). In particular, the spectral measure of T at $\mathbf{e}_1$ is $\sum p_j \delta_{\lambda_j}$ where $p_j = |U_{1,j}|^2$ (here $U_{1,j}$ is the first co-ordinate of $\mathbf{u}_j$).

Given a real symmetric tridiagonal matrix T, let ν_T be the spectral measure of T at the standard unit vector $\mathbf{e}_1$. This gives a mapping from $\mathcal{T}_n$ into $\mathcal{P}_n$ which maps T_n^0 into P_n^0.

Probability measure to Tridiagonal matrix: Now suppose a measure $\mu \in \mathcal{P}_n^0$ is given. Write $\mu = p_1\delta_{\lambda_1} + \ldots + p_n\delta_{\lambda_n}$ where λ_j are distinct real numbers and $p_j > 0$. Its moments are given by $\alpha_k = \sum p_j \lambda_j^k$. Let $h_k(x) = x^k$, so that $\{h_0, h_1, \ldots, h_{n-1}\}$ is a basis for $L^2(\mu)$ (how do you express h_n as a linear combination of $h_0, \ldots, h_{n-1}$?).

Apply Gram-Schmidt to the sequence $h_0, h_1, \ldots$ by setting $\varphi_0 = \psi_0 = h_0$, and for $k \ge 1$ inductively by

$$\psi_k = h_k - \sum_{j=0}^{k-1} \langle h_k, \varphi_j\rangle \varphi_j, \qquad \varphi_k = \frac{\psi_k}{\|\psi_k\|_{L^2(\mu)}}.$$

This process is stopped when $\|\psi_k\| = 0$. Here are some elementary observations.

(a) Since $\{h_0, \ldots, h_{n-1}\}$ is a linear basis for $L^2(\mu)$, it follows that $\{\varphi_0, \ldots, \varphi_{n-1}\}$ are well-defined and form an ONB for $L^2(\mu)$.

(b) For $0 \le k \le n-1$, φ_k is a polynomial of degree k and is orthogonal to all polynomials of degree less than k.

(c) As h_n is a linear combination of $h_0, \ldots, h_{n-1}$ (in $L^2(\mu)$), we see that ψ_n is well-defined but $\|\psi_n\| = 0$ and hence φ_n is not defined. Note that $\|\psi_n\| = 0$ means that $\psi_n(\lambda_k) = 0$

for all $k \le n$, not that ψ_n is the zero polynomial. In fact, ψ_n is monic, has degree n and vanishes at λ_k, $k \le n$, which implies that $\psi_n(\lambda) = \prod_{j=1}^n (\lambda - \lambda_j)$.

Fix $0 \le k \le n-1$ and expand $x\varphi_k(x)$ as

$$x\varphi_k(x) \overset{L^2(\mu)}{=} \sum_{j=0}^n c_{k,j}\varphi_j(x), \qquad c_{k,j} = \int x\varphi_k(x)\varphi_j(x)d\mu(x).$$

Now, $x\varphi_j(x)$ has degree less than k if $j<k$ and $x\varphi_k(x)$ has degree less than j if $k<j$. Hence, $c_{k,j} = 0$ if $j \le k-2$ or if $j \ge k+2$. Further, $c_{k,k+1} = c_{k+1,k}$ as both are equal to $\int x\varphi_k(x)\varphi_{k+1}(x)d\mu(x)$. Thus, we get the *three term recurrences*

$$\text{(21)} \qquad \begin{aligned} x\varphi_k(x) &\overset{L^2(\mu)}{=} b_{k-1}\varphi_{k-1}(x) + a_k\varphi_k(x) + b_k\varphi_{k+1}(x), \qquad 0 \le k \le n \\ \text{where} \quad & a_k = \int x\varphi_k(x)^2 d\mu(x),\ b_k = \int x\varphi_k(x)\varphi_{k+1}(x)d\mu(x). \end{aligned}$$

We adopt the convention that φ_{-1}, φ_n, b_{-1} and b_{n-1} are all zero, so that these recurrences also hold for $k=0$ and $k=n$. Since φ_k all have positive leading co-efficients, it is not hard to see that b_k is nonnegative.

From $\mu \in \mathcal{P}_n^0$ we have thus constructed a tridiagonal matrix $T_\mu := T(a,b) \in \mathcal{T}_n$ (caution: here we have indexed a_k, b_k starting from $k=0$). If $\mu \in \mathcal{P}_m^0$ for some $m<n$, the T_μ constructed as before will have size $m \times m$. Extend this by padding $n-m$ columns and rows of zeros to get a real symmetric tridiagonal matrix (we abuse notation and denote it as T_μ again) in $\mathcal{T}_n$. Thus we get a mapping $\mu \to T_\mu$ from $\mathcal{P}_n$ into $\mathcal{T}_n$.

The following lemma shows that $T \to \nu_T$ is a bijection, and relates objects defined on one side (matrix entries, characteristic polynomials, eigenvalues) to objects defined on the other side (the support $\{\lambda_j\}$, the weights p_j, associated orthogonal polnomials).

Lemma 50. *Fix $n \ge 1$.*

(a) The mapping $T \to \nu_T$ is a bijection from $\mathcal{T}_n^0$ into $\mathcal{P}_n^0$ whose inverse is $\mu \to T_\mu$.

(b) Let $T = T(a,b)$ and let $\mu = \nu_T$. For $0 \le k \le n-1$ P_k be the characteristic polynomial of the top $k \times k$ principal submatrix of T and let ψ_k, φ_k be as constructed earlier. Then $\psi_k = P_k$ for $k \le n$ and hence there exist constants d_k such that $\varphi_k = d_k P_k$ (for $k \le n-1$).

(c) The zeros of φ_n are precisely the eigenvalues of T.

(d) If $T = T(a,b)$ and $\nu_T = \sum_{k=1}^n p_j \delta_{\lambda_j}$, then

$$\prod_{k=1}^n b_k^{2(n-k+1)} = \prod_{k=1}^n p_k \prod_{i<j} |\lambda_i - \lambda_j|^2. \tag{22}$$

In particular, $\mathcal{T}_n^0$ gets mapped into $\mathcal{P}_n^0$ (but not onto).

PROOF. (a) Let $\mu = \sum_{j=1}^n p_j \delta_{\lambda_j} \in \mathcal{P}_n$ and let $T = T_\mu$. For $0 \le k \le n-1$, let

$$\mathbf{u}_k = (\sqrt{p_1}\varphi_0(\lambda_k), \dots, \sqrt{p_n}\varphi_{n-1}(\lambda_k))^t.$$

The three-term recurrences can be written in terms of T as $T\mathbf{u}_k = \lambda_k \mathbf{u}_k$. Thus, $\mathbf{u}_k$ is an eigenvector of T with eigenvalue λ_k. If U is the matrix with columns $\mathbf{u}_k$, then the rows of U are orthonormal because φ_k are orthogonal polynomials of μ. Thus $UU^* = I$ and hence also $U^*U = I$, that is $\{\mathbf{u}_k\}$ is an ONB of $\mathbb{R}^n$.

Consequently, $T = \sum_{k=1}^n \lambda_k \mathbf{u}_k \mathbf{u}_k^*$ is the spectral decomposition of T. In particular,

$$T^p \mathbf{e}_1 = \sum_{k=1}^n |u_{k,1}|^2 \lambda_k^p = \sum_{k=1}^n p_k \lambda_k^p$$

because $u_{k,1} = \sqrt{p_k}\varphi_0(\lambda_k) = \sqrt{p_k}$ (as $h_0 = 1$ is already of unit norm in $L^2(\mu)$ and hence after Gram-Schmidt $\varphi_0 = h_0$). Thus, $\langle T^p \mathbf{e}_1, \mathbf{e}_1 \rangle = \int x^p \mu(dx)$ which shows that $\nu_T = \mu$. This proves the first part of the lemma.

(b) We saw earlier that φ_n is zero in $L^2(\mu)$. Hence $\varphi_n(\lambda_j) = 0$ for $1 \le j \le n$. Thus, φ_n and P_n are non-zero polynomials of degree n both of which vanish at the same n points. Hence, $\varphi_n = d_n P_n$ for some constant P_n.

If S is the top $k \times k$ principal submatrix of T, then it is easy to see that the first k orthogonal polynomials of ν_S are the same as $\varphi_0, \ldots, \varphi_k$ (which were obtained as orthogonal polynomials of ν_T). This is easy to see from the three-term recurrences. Thus the above fact shows $\varphi_k = d_k P_k$ for some constant d_k.

(c) By the first part, $\nu_T = \sum p_j \delta_{\lambda_j}$ where λ_j are the eigenvalues of T and $p_j > 0$. The footnote on the previous page also shows that φ_n vanishes at λ_j, $j \le n$. Since it has degree n, φ_n has only these zeros.

(d) This proof is taken from Forrester's book. Let T_k denote the bottom $(n-k) \times (n-k)$ principal submatrix of T. Let Q_k be its characteristic polynomial and let $\lambda_j^{(k)}$, $1 \le j \le n-k$ be its eigenvalues. In particular, $T_0 = T$.

If $T = \sum \lambda_k \mathbf{u}_k \mathbf{u}_k^*$ is the spectral decomposition of T and λ is not an eigenvalue of T, then $(\lambda I - T)^{-1} = \sum (\lambda - \lambda_k)^{-1} \mathbf{u}_k \mathbf{u}_k^*$. Hence, $(\lambda I - T)^{1,1} = \langle (\lambda I - T)^{-1} \mathbf{e}_1, \mathbf{e}_1 \rangle = \sum_j p_j (\lambda - \lambda_j)^{-1}$ for $\lambda \notin \{\lambda_j\}$. But we also know that $(\lambda I - T)^{1,1}$ is equal to $\det(\lambda I - T_1)/\det(\lambda I - T) = Q_1(\lambda)/Q_0(\lambda)$. Let λ approach λ_k to see that

$$p_k = \lim_{\lambda \to \lambda_k} (\lambda - \lambda_k)(\lambda I - T)^{1,1} = \lim_{\lambda \to \lambda_k} (\lambda - \lambda_k) \frac{Q_1(\lambda)}{Q_0(\lambda)} = \frac{Q_1(\lambda_k)}{\prod\limits_{j=1 \, j \neq k}^{n} (\lambda_k - \lambda_j)}.$$

Take product over k to get (the left side is positive, hence absolute values on the right)

$$\prod_{k=1}^{n} p_k \prod_{i<j} (\lambda_i - \lambda_j)^2 = \prod_{k=1}^{n} |Q_1(\lambda_k^{(0)})|. \tag{23}$$

Let A be any $n \times n$ matrix with characteristic polynomial χ_A and eigenvalues λ_i. Let B be an $m \times m$ matrix with characteristic polynomial χ_B and eigenvalues μ_j. Then we have the obvious identity

$$\prod_{i=1}^{n} |\chi_B(\lambda_i)| = \prod_{i=1}^{n} \prod_{j=1}^{m} |\mu_j - \lambda_i| = \prod_{j=1}^{m} |\chi_A(\mu_j)|.$$

Apply to T_0 and T_1 to get $\prod\limits_{k=1}^{n} |Q_1(\lambda_k^{(0)})| = \prod\limits_{k=1}^{n-1} |Q_0(\lambda_k^{(1)})|$. But by expanding $\det(\lambda I - T)$ by the first row, we also have the identity

$$Q_0(\lambda) = (\lambda - a_1) Q_1(\lambda) - b_1^2 Q_2(\lambda).$$

Therefore $Q_0(\lambda_k^{(1)}) = b_1^2 Q_2(\lambda_k^{(1)})$ for $k \le n-1$. Thus $\prod\limits_{k=1}^{n} |Q_1(\lambda_k^{(0)})| = b_1^{2n-2} \prod\limits_{k=1}^{n-1} Q_2(\lambda_k^{(1)})$. The right side is of a similar form to the left side, with matrix size reduced by one. Thus, inductively we get $\prod_{k=1}^{n} |Q_1(\lambda_k^{(0)})| = \prod_{k=1}^{n-1} b_k^{2n-2k}$. Plugging into (23) we get the statement of the lemma. ■

3. Tridiagonal matrix generalities

Fix $n \ge 1$ and write $T = T(a,b)$ for the real symmetric $n \times n$ tridiagonal matrix with diagonal entries $T_{k,k} = a_k$ for $1 \le k \le n$ and $T_{k,k+1} = T_{k+1,k} = b_k$ for $1 \le k \le n-1$. Let $\mathcal{T}_n$

be the space of all $n\times n$ real symmetric tridiagonal matrices and let $\mathcal{T}_n^0$ be those $T(a,b)$ in $\mathcal{T}_n$ with b_k strictly positive. Let $\mathcal{P}_n$ be the space of all probability measures on $\mathbb{R}$ whose support consists of at most n distinct points and let $\mathcal{P}_n^0$ be those elements of $\mathcal{P}_n$ whose support has exactly n distinct points.

Given a real symmetric tridiagonal matrix T, let ν_T be the spectral measure of T at the standard unit vector $\mathbf{e}_1$[3]. This gives a mapping from $\mathcal{T}_n$ into $\mathcal{P}_n$. For future purpose, we also give the following idea to find eigenvalues of T.

Fix some $\lambda\in\mathbb{R}$ and suppose we want to find a vector $\mathbf{v}$ such that $T\mathbf{v}=\lambda\mathbf{v}$. This means

$$b_{k-1}v_{k-1}+a_kv_k+b_kv_{k+1}=\lambda v_k \implies v_{k+1}=\frac{\lambda v_k-b_{k-1}v_{k-1}-a_kv_k}{b_k}.$$

where we adopt the convention that $b_0=0$. We have also assumed that $b_k\neq 0$ for all k (if $b_k=0$, the matrix splits into a direct sum of two matrices). Thus, we set $\mathbf{v}_1=x$ to be arbitrary (non-zero)and solve for $v_1,v_2,\ldots$ successively. Denote these as $v_1(x),v_2(x),\ldots$. Therefore,

Now suppose a measure $\mu\in\mathcal{P}_n^0$ is given. We can construct a tridiagonal matrix T as follows. Write $\mu=p_1\delta_{\lambda_1}+\ldots+p_n\delta_{\lambda_n}$ where λ_j are distinct real numbers and $p_j>0$. The moments are given by $\alpha_k=\sum p_j\lambda_j^k$. Let $h_k(x)=x^k$, so that $\{h_0,h_1,\ldots,h_{n-1}\}$ is a basis for $L^2(\mu)$. [Q: How do you express h_n as a linear combination of $h_0,\ldots,h_{n-1}$?].

Apply Gram-Schmidt to the sequence $h_0,h_1,\ldots$ to get an orthonormal basis $\{\varphi_k:0\le k\le n-1\}$ of $L^2(\mu)$. It is easy to see that φ_k is a polynomial of degree exactly k, and is orthogonal to all polynomials of degree less than k. Fix any k and write

$$x\varphi_k(x)\overset{L^2(\mu)}{=}\sum_{j=0}^{n}c_{k,j}\varphi_j(x),\qquad c_{k,j}=\int x\varphi_k(x)\varphi_j(x)d\mu(x).$$

Now, $x\varphi_j(x)$ has degree less than k if $j<k$ and $x\varphi_k(x)$ has degree less than j if $k<j$. Hence, $c_{k,j}=0$ if $j\le k-2$ or if $j\ge k+2$. Further, $c_{k,k+1}=c_{k+1,k}$ as both are equal to $\int x\varphi_k(x)\varphi_{k+1}(x)d\mu(x)$. Thus, we get the *three term recurrences*

$$\begin{aligned}(24)\qquad x\varphi_k(x)&\overset{L^2(\mu)}{=} b_{k-1}\varphi_{k-1}(x)+a_k\varphi_k(x)+b_k\varphi_{k+1}(x),\qquad 0\le k\le n\\ \text{where}\qquad & a_k=\int x\varphi_k(x)^2d\mu(x),\ b_k=\int x\varphi_k(x)\varphi_{k+1}(x)d\mu(x).\end{aligned}$$

We adopt the convention that φ_{-1}, φ_n, b_{-1} and b_{n-1} are all zero, so that these recurrences also hold for $k=0$ and $k=n$.

From $\mu\in\mathcal{P}_n^0$ we have thus constructed a tridiagonal matrix $T_\mu:=T(a,b)\in\mathcal{T}_n$ (caution: here we have indexed a_k,b_k starting from $k=0$). If $\mu\in\mathcal{P}_m^0$ for some $m<n$, the T_μ constructed as before will have size $m\times m$. Extend this by padding $n-m$ columns and rows of zeros to get a real symmetric tridiagonal matrix (we abuse notation and denote it as T_μ again) in $\mathcal{T}_n$. Thus we get a mapping $\mu\to T_\mu$ from $\mathcal{P}_n$ into $\mathcal{T}_n$.

Lemma 51. *Fix $n\ge 1$.*

(a) The mapping $T\to\nu_T$ is a bijection from $\mathcal{T}_n$ into $\mathcal{P}_n$ whose inverse is $\mu\to T_\mu$.

[3]The *spectral measure* of a Hermitian operator T at a vector $\mathbf{v}$ is the unique measure ν on $\mathbb{R}$ such that $\langle T^p\mathbf{v},\mathbf{v}\rangle=\int x^p\nu(dx)$ for all $p\ge 0$. For example, if T is a real symmetric matrix, write its spectral decomposition as $T=\sum_{k=1}^n\lambda_k\mathbf{u}_k\mathbf{u}_k^*$. Then $\{\mathbf{u}_k\}$ is an ONB of $\mathbb{R}^n$ and λ_k are real. In this case, the spectral decomposition of T at any $\mathbf{v}\in\mathbb{R}^n$ is just $\nu=\sum_{k=1}^n|\langle\mathbf{v},\mathbf{u}_k\rangle|^2\delta_{\lambda_k}$. Thus $\nu\in\mathcal{P}_n$ (observe that the support may have less than n points as eigenvalues may coincide). In particular, if $T=UDU^*$ the spectral measure of T at $\mathbf{e}_1$ is $\nu_T=\sum p_i\delta_{\lambda_i}$, where $p_i=|U_{1,i}|^2$.

(b) *Let $\mu = \nu_T$. Write P_k for the characteristic polynomial of the top $k \times k$ submatrix of T for $k \le n$ and let φ_k be the orthogonal polynomials for μ as defined earlier. Then $\varphi_k = d_k P_k$ for for constants d_k. In particular, zeros of φ_n are precisely the eigenvalues of T.*

(c) *If $T = T(a,b)$ and $\mu = \sum_{k=1}^n p_j \delta_{\lambda_j}$ correspond to each other in this bijection, then*

$$\prod_{k=1}^{n} b_k^{2(n-k+1)} = \prod_{k=1}^{n} p_k \prod_{i<j} |\lambda_i - \lambda_j|^2. \tag{25}$$

In particular, $\mathcal{T}_n^0$ gets mapped into $\mathcal{P}_n^0$ (but not onto).

PROOF. (a) Let $\mu = \sum_{j=1}^n p_j \delta_{\lambda_j} \in \mathcal{P}_n$ and let $T = T_\mu$. For $0 \le k \le n-1$, let

$$\mathbf{u}_k = (\sqrt{p_1}\varphi_0(\lambda_k), \ldots, \sqrt{p_n}\varphi_{n-1}(\lambda_k))^t.$$

The three-term recurrences can be written in terms of T as $T\mathbf{u}_k = \lambda_k \mathbf{u}_k$. Thus, $\mathbf{u}_k$ is an eigenvector of T with eigenvalue λ_k. If U is the matrix with columns $\mathbf{u}_k$, then the rows of U are orthonormal because φ_k are orthogonal polynomials of μ. Thus $UU^* = I$ and hence also $U^*U = I$, that is $\{\mathbf{u}_k\}$ is an ONB of $\mathbb{R}^n$.

Consequently, $T = \sum_{k=1}^n \lambda_k \mathbf{u}_k \mathbf{u}_k^*$ is the spectral decomposition of T. In particular,

$$T^p \mathbf{e}_1 = \sum_{k=1}^{n} |u_{k,1}|^2 \lambda_k^p = \sum_{k=1}^{n} p_k \lambda_k^p$$

because $u_{k,1} = \sqrt{p_k}\varphi_0(\lambda_k) = \sqrt{p_k}$ ($h_0 = 1$ is already of unit norm in $L^2(\mu)$ and hence after Gram-Schmidt $\varphi_0 = h_0$). Thus, $\langle T^p \mathbf{e}_1, \mathbf{e}_1 \rangle = \int x^p \mu(dx)$ which shows that $\nu_T = \mu$. This proves the first part of the lemma.

(b) By part (a), the coefficients in the three term recurrence (24) are precisely the entries of T. Note that the equality in (24) is in $L^2(\mu)$, which means the same as saying that equality holds for $x = \lambda_k$, $1 \le k \le n$.

Here is a way to find

$$T$$

(c) Let A be any $n \times n$ matrix with characteristic polynomial χ_A and eigenvalues λ_i. Let B be an $m \times m$ matrix with characteristic polynomial χ_B and eigenvalues μ_j. Then we have the obvious identity

$$\prod_{i=1}^{n} \chi_B(\lambda_i) = \prod_{i=1}^{n}\prod_{j=1}^{m} (\mu_j - \lambda_i) = (-1)^{mn} \prod_{j=1}^{m} \chi_A(\mu_j)$$

If b_k are all positive, the right hand side of (**??**) is non-zero and hence λ_k must be distinct. This shows that $\mathcal{T}_n^0$ gets mapped into $\mathcal{P}_n^0$. It is obviously not onto (why?). ∎

Lemma 52. *For $T = T(a,b)$ having the spectral measure $\sum_{j=1}^{n+1} p_i \delta_{\lambda_j}$ at $\mathbf{e}_0$, we have the identity*

$$\prod_{k=0}^{n-1} b_k^{2(n+1-k)} = \prod_{i=1}^{n+1} p_i \prod_{i<j\le n+1} (\lambda_i - \lambda_j)^2.$$

4. More on tridiagonal operators*

This section may be omitted as we shall not use the contents in this course. However, as we are this close to a very rich part of classical analysis, we state a few interesting facts. The following four objects are shown to be intimately connected.

(1) Positive measures $\mu \in \mathcal{P}(\mathbb{R})$ having all moments.

(2) Positive definite sequences $\alpha = (\alpha_k)_{k\geq 0}$ such that $(\alpha_{i+j})_{i,j\geq 0}$ is non-negative definite (that is every principal submatrix has non-negative determinant).
(3) Orthogonal polynomials. Given an inner product on the vector space of all polynomials, one can obtain an orthonormal basis $\{\varphi_k\}$ by applying Gram-Schmidt process to the basis $\{h_k\}_{k\geq 0}$ where $h_k(x) = x^k$. The sequence $\{\varphi_k\}$ (which may be finite) is called an orthogonal polynomial sequence.
(4) Real symmetric tridiagonal matrices. We now consider semi-infinite matrices, that is $T_{k,k} = a_k$, $T_{k,k+1} = T_{k+1,k} = b_k$, for $k \geq 0$. Finite matrices are a subset of these, by padding them with zeros at the end.

Measure to Positive definite sequences: If μ is a measure that has all moments, define the moment sequence $\alpha_k = \int x^k d\mu(x)$. Then for any for any $m \geq 1$ and any $\mathbf{u} \in \mathbb{R}^{m+1}$, we have

$$\mathbf{u}^t \left(\alpha_{i+j}\right)_{0\leq i,j\leq m} \mathbf{u} = \sum_{i,j\leq m} \alpha_{i+j} u_i u_j = \int \Big| \sum_{i=0}^{n} u_i x^i \Big|^2 \mu(dx) \geq 0.$$

Hence α is a positive definite sequence. It is easy to see that μ is finitely supported if and only $L^2(\mu)$ is finite dimnesional if and only if $(\alpha_{i+j})_{i,j\geq 0}$ has finite rank.
Positive definite sequence to orthogonal polynomials: Let α be a positive definite sequence For simplicity we assume that $(\alpha_{i+j})_{i,j\geq 0}$ is strictly positive definite. Then the formulas $\langle h_i, h_j\rangle = \alpha_{i+j}$ define a valid inner product on the vector space $\mathcal{P}$ of all polynomials. Complete $\mathcal{P}$ under this inner product to get a Hilbert space H.

In H, h_k are linearly independent and their span (which is $\mathcal{P}$) is dense. Hence, applying Gram-Schmidt procedure to the sequence $h_0, h_1, \ldots$ give a sequence of polynomials $\varphi_0, \varphi_2, \ldots$ which form an orthonormal basis for H. Clearly φ_k has degree k.
Orthogonal polynomials to tridiagonal matrices: Let φ_k be an infinite sequence of polynomials such that φ_k has degree exactly k. Then it is clear that φ_k are linearly independent, that h_k is a linear combination of $\varphi_0, \ldots, \varphi_k$.

Consider an inner product on $\mathcal{P}$ such that $\langle \varphi_k, \varphi_\ell\rangle = \delta_{k,\ell}$. The same reasoning as before gives the three term recurrences (24) for φ_k. Thus we get $a_k \in \mathbb{R}$ and $b_k > 0$, $k \geq 0$. Form the infinite real symmetric tridiagonal matrix $T = T(a,b)$.
Symmetric tridiagonal operators to measures: Let T be a semi-infinite real symmetric tridiagonal matrix. Let $\mathbf{e}_k$ be the co-ordinate vectors in $\ell^2(\mathbb{N})$. Let $D = \{\sum x_k \mathbf{e}_k : x_k \neq 0 \text{ finitely often}\}$. This is a dense subspace of $\ell^2(\mathbb{N})$. T is clearly well-defined and linear on D. It is symmetric in the sense that $\langle Tu, v\rangle = \langle u, Tv\rangle$ for all $u, v \in D$ and the inner product is in $\ell^2(\mathbb{N})$.

Suppose T' is a self-adjoing extension of T. That is, there is a subspace D' containing D and a linear operator $T' : D \to \mathbb{R}$ such that $T'|_D = T$ and such that T' is self-adjoint (we are talking about unbounded operators on Hilbert spaces, hence self-adjointness and symmetry are two distinct things, and this is not the place to go into the definitions. Consult for example, chapter 13 of Rudin's *Functional Analysis*). Then it is a fact that T' has a spectral decomposition. The spectral measure of T' at $\mathbf{e}_0$ is a measure. In general there can be more than one extension. If the extension is unique, then μ is uniquely defined.

This cycle of connections is quite deep. For example, if we start with any positive definite sequence α_k and go through this cycle, we get an OP sequence and a tridiagonal symmetric operator. The spectral measure of any self-adjoint extension of this operator has the moment sequence α_k. Further, there is a unique measure with moments α_k if and only if T has a unique self-adjoint extension!

Remark 53. In the above discussion we assumed that α is a strictly positive definite sequence, which is the same as saying that the measure does not have finite support or that the orthogonal polynomial sequence is finite or that the tridigonal matrix is essentially finite. If we start with a finitely supported measure, we can still go through this cycle, except that the Gram-Schmidt process stops at some finite n etc.