ORTHOGONAL POLYNOMIALS

1. Prelude: The van der Monde Determinant

The link between random matrix theory and the classical theory of orthogonal polynomials is *van der Monde's determinant*:

$$\Delta_n := \prod_{1\le i<j\le n}(x_j - x_i) = \det\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \dots & & & \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1}\end{vmatrix} \tag{1}$$

The product on the left side occurs to some power ($\beta = 1,2,4$) in the joint density of the eigenvalues for each of the basic Gaussian random matrix ensembles (orthogonal, unitary, and symplectic, respectively), and it also occurs in a number of other important models studied in statistical physics, called generically *determinantal point processes*. Van der Monde's formula implies that this joint density can be rewritten as a determinant, or product of determinants. This is more useful than it might first appear, because the determinant (1), can be re-expressed using other polynomial bases: Suppose that for each index $k = 0,1,2,\dots,$

$$\pi_k(x) = x^k + \text{lower order terms}$$

is a monic polynomial of degree k; then by elementary column operations on the Van der Monde determinant (1),

$$\prod_{1\le i<j\le n}(x_j - x_i) = \det\begin{vmatrix} \pi_0(x_1) & \pi_0(x_2) & \dots & \pi_0(x_n) \\ \pi_1(x_1) & \pi_1(x_2) & \dots & \pi_1(x_n) \\ \dots & & & \\ \pi_{n-1}(x_1) & \pi_{n-1}(x_2) & \dots & \pi_{n-1}(x_n)\end{vmatrix} \tag{2}$$

Which polynomial basis to use? For studying the Gaussian ensembles, it will turn out that the monic orthogonal polynomials with respect to the Gaussian measure on $\mathbb{R}$, the *Hermite polynomials* are especially useful. In other models of statistical physics and stochastic processes (Examples: exclusion processes, the XYZ model) the van der Monde determinant occurs in conjunction with measures other than the Gaussian, and in such problems other systems of orthogonal polynomials come into play.

Proof of van der Monde's Formula (1). There are various proofs. The most commonly recited is algebraic, using the fact that both sides vanish when $x_i = x_j$ for some pair of distinct indices $i \neq j$. Here is a more combinatorial proof that does not rely on the unique factorization theorem for polynomials: (A) Check that both sides are homogeneous polynomials of degree $\binom{n}{2}$ (*homogeneous* means that all terms have the same degree). (B) Check that both sides are *antisymmetric*, that is, if for some pair $i \neq j$ the variables x_i and x_j are switched, then both sides change sign. For this it suffices to check nearest-neighbor pairs $i, i+1$, because nearest-neighbor transpositions generate the group of permutations. (C) Use antisymmetry to deduce that there are no terms (on either side) of the form

$$C\prod_{i=1}^{n} x^{r_i}$$

with $r_i = r_j$ for some pair $i \neq j$. (D) Conclude that it suffices to check that the coefficients of $x_2x_3^2x_4^3\cdots x_n^{n-1}$ are the same on both sides. □

2. Orthogonal Polynomials: Basic Theory

2.1. **Determinantal Representation.** Let μ be a probability measure on the real line with finite moment generating function in some neighborhood of 0. Assume that μ is not fully supported by any finite subset of $\mathbb{R}$; then $L^2(\mu)$ is infinite-dimensional, and the monomials $1, x, x^2, \dots$ are linearly independent as elements of $L^2(\mu)$. For any such measure, the monomial basis $1, x, x^2, \dots$ can be orthogonalized by the Gram-Schmidt procedure to produce an orthonormal set $\{p_n(x)\}_{n\geq 0}$ of $L^2(\mu)$ consisting of polynomials $p_n(x)$ of degree n with real coefficients:

$$\langle p_n, p_m\rangle_\mu := \int_{\mathbb{R}} p_n(x)p_m(x)\,d\mu(x) = \delta_{n,m} \tag{3}$$

Up to signs, this condition uniquely specifies the polynomials $p_n(x)$. It is conventional to choose the signs so that the leading coefficients γ_n are *positive*: the resulting polynomials are the normalized *orthogonal polynomials* associated with the measure μ. Examples for various familiar probability measures will be computed below. The corresponding *monic* orthogonal polynomials will be denoted by $\pi_n(x)$: thus,

$$p_n(x) = \gamma_n\pi_n(x) \quad \text{where} \quad \pi_n(x) = x^n + \text{lower order terms.} \tag{4}$$

Exercise 1. (A) Prove that if μ is not supported by a finite subset of $\mathbb{R}$ then the monomials are linearly independent in $L^2(\mu)$. (B) Show that if support(μ) has cardinality n then $1, x, x^2, \dots, x^{n-1}$ are linearly independent, and therefore form a basis of $L^2(\mu)$. (C) Prove that if μ has a finite moment generating function in a neighborhood of 0 then the polynomials are dense in $L^2(\mu)$. Hint: Use the fact that the Laplace transform of a measure is analytic in its domain of definition. (D)** What if μ merely has all moments finite?

Proposition 1. *The orthogonal polynomials $p_n(x)$ and $\pi_n(x)$ are given by the determinantal formulas (vaid for $n \geq 1$)*

$$\pi_n(x) = D_n(x)/D_{n-1}, \tag{5}$$

$$p_n(x) = D_n(x)/\sqrt{D_nD_{n-1}} \tag{6}$$

$$\gamma_n = \sqrt{D_{n-1}/D_n}. \tag{7}$$

where

$$D_n(x) := \det\begin{vmatrix} m_0 & m_1 & m_2 & \dots & m_n \\ m_1 & m_2 & m_3 & \dots & m_{n+1} \\ \dots & & & & \\ m_{n-1} & m_n & m_{n+1} & \dots & m_{2n-1} \\ 1 & x & x^2 & \dots & x^n \end{vmatrix}, \tag{8}$$

(9) $$D_n := \det \begin{vmatrix} m_0 & m_1 & m_2 & \dots & m_n \\ m_1 & m_2 & m_3 & \dots & m_{n+1} \\ \dots \\ m_n & m_{n+1} & m_{n+2} & \dots & m_{2n} \end{vmatrix},$$

and m_n is the nth moment of the measure μ:

(10) $$m_n := \int x^n \, d\mu(x).$$

Consequently, the polynomials $p_n(x)$ and $\pi_n(x)$ have real *coefficients. Moreover,*

(11) $$D_n = \prod_{k=1}^{n} \gamma_k^{-2} = \prod_{k=0}^{n} \gamma_k^{-2}.$$

Note: (1) The fact that the orthogonal polynomials $p_n(x)$ have real coefficients implies that in inner products involving these polynomials, no complex conjugation is necessary. (2) Matrices $(a_{i,j})$ such as occur in formula (9), that is, such that $a_{i,j} = a_{i+j}$, are called *Hankel matrices*. Matrices that are constant down diagonals, that is, such that $a_{i,j} = a_{j-i}$, are called *Toeplitz matrices*.

Proof. The function $D_n(x)$ is obviously a polynomial of degree n, and it is clear from the determinantal formula (8) that the coefficient of x^n is the minor determinant D_{n-1}. Multiplying $D_n(x)$ by x^k can be accomplished in the determinant (8) by replacing each x^i by x^{i+k}. Integrating $x^k D_n(x)$ against $d\mu(x)$ can de done by first integrating each factor x^{i+k} in the bottom row, then taking the determinant: see Lemma 2 below. Thus, for $k \leq n$,

$$\langle D_n(x), x^k \rangle = \det \begin{vmatrix} m_0 & m_1 & m_2 & \dots & m_n \\ m_1 & m_2 & m_3 & \dots & m_{n+1} \\ \dots \\ m_{n-1} & m_n & m_{n+1} & \dots & m_{2n-1} \\ m_k & m_{k+1} & m_{k+2} & \dots & m_{k+n} \end{vmatrix} = \delta_{n,k} D_n,$$

because if $k < n$ two rows of the matrix agree. This shows that the polynomials $D_n(x)$ are orthogonal in $L^2(\mu)$, and equation (5) follows. Furthermore, since the lead coefficient of $D_n(x)$ is D_{n-1}, it follows that

$$\langle D_n(x), D_n(x) \rangle = \langle D_n(x), x^n D_{n-1} \rangle = D_n D_{n-1},$$

and this together with (5) implies the normalization equations (6)–(7). The final assertion, equation (11), follows from (7), since $D_0 = 1$. □

Lemma 2.

(12) $$\int \det \begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \dots \\ a_{m-1,1} & a_{m-1,2} & \dots & a_{m-1,m} \\ f_1(x) & f_2(x) & \dots & f_m(x) \end{vmatrix} d\mu(x) = \det \begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ a_{2,1} & a_{2,2} & \dots & a_{2,m} \\ \dots \\ a_{m-1,1} & a_{m-1,2} & \dots & a_{m-1,m} \\ \int f_1 \, d\mu & \int f_2 \, d\mu & \dots & \int f_m \, d\mu \end{vmatrix}$$

Proof. The left side is $\int (\sum_{j=1}^m c_j f_j) \, d\mu$ for constants c_j gotten by taking the appropriate minors of the matrix $(a_{i,j})$. Reversing the order of integration and summation and usiing the usual cofactor representation of a determinant yields the result. □

2.2. **Three-Term Recurrence Formula.** The monic orthogonal polynomials $\pi_n(x)$ satisfy a simple, recursive system of linear equations called the three-term recurrence. Multiply the monic polynomial $\pi_n(x)$ by x: The resulting polynomial $x\pi_n(x)$ is again monic, and of degree $n+1$, so it must equal $\pi_{n+1}(x)+$ lower order polynomials. What is noteworthy about this is that the lower order terms only involve $\pi_n(x)$ and $\pi_{n-1}(x)$.

Proposition 3. *There exist real α_n and positive β_n so that*

$$x\pi_n(x) = \pi_{n+1}(x) + \alpha_n\pi_n(x) + \beta_n\pi_{n-1}(x). \tag{13}$$

Moreover, the coefficients β_n obey

$$\beta_n = \frac{\gamma_{n-1}^2}{\gamma_n^2} \implies \gamma_n^2 = \prod_{i=1}^{n} \beta_n^{-1}. \tag{14}$$

Proof. Since $x\pi_n(x)$ is monic of degree $n+1$, the difference $x\pi_n(x) - \pi_{n+1}(x)$ is a polynomial of degree $\leq n$, and so it can be expressed as a linear combination of the orthogonal polynomials $p_k(x)$. Since the polynomials p_k form an orthonormal basis of $L^2(\mu)$, the coefficients of $p_k(x)$ in this expansion are the inner products $\langle x\pi_n(x) - \pi_{n+1}(x), p_k(x)\rangle_\mu$. But π_{n+1} is orthogonal to p_k for each $k \leq n$. Consequently,

$$x\pi_n(x) = \pi_{n+1}(x) + \sum_{k=0}^{n} \langle x\pi_n(x), p_k(x)\rangle p_k(x).$$

Now clearly $\langle x\pi_n(x), p_k(x)\rangle = \langle \pi_n(x), xp_k(x)\rangle$; but $xp_k(x)$ is a polynomial of degree $k+1$, and hence is orthogonal to $\pi_n(x)$ if $k+1 < n$. Thus, all but the top two terms in the sum vanish, leaving an identity equivalent to (13). Since the inner products are all real, so are the coefficients α_n and β_n. Finally, since $p_n(x) = \gamma_n\pi_n(x)$,

$$\begin{aligned}
\beta_n &= \langle x\pi_n(x), p_{n-1}(x)\rangle\gamma_{n-1} \\
&= \langle x\pi_n(x), \pi_{n-1}(x)\rangle\gamma_{n-1}^2 \\
&= \langle \pi_n(x), x\pi_{n-1}(x)\rangle\gamma_{n-1}^2 \\
&= \langle \pi_n(x), \pi_n(x)\rangle\gamma_{n-1}^2 \\
&= \gamma_{n-1}^2/\gamma_n^2.
\end{aligned}$$

(The second last equality follows because $x\pi_{n-1}(x) = \pi_n(x)$+lower order terms.) □

2.3. **Van der Monde Determinant and the Christoffel-Darboux Formula.** As we will see, the Hermite polynomials are both monic and of L^2 norm 1 relative to the Gaussian measure on $\mathbb{R}$. Few other orthogonal polynomial systems have this property, and for these systems it is usually more convenient to work with the *normalized* orthogonal polynomials $p_n(x)$ than the *monic* orthogonal polynomials $\pi_n(x)$. The formula (2), which exhibits the van der Monde factor $\prod_{1\leq i<j\leq n}(x_j - x_i)$ as a determinant of *monic* polynomials, is easily converted to a formula involving the *normalized* orthogonal polynomials $\pi_n(x)$, since replacing $\pi_k(x)$ by $p_k(x)$ in the determinant just multiplies the kth row by

γ_k^{-1}. Thus,

$$\Delta_n = \left(\prod_{k=0}^{n-1} \gamma_k^{-1}\right) \det \begin{vmatrix} p_0(x_1) & p_0(x_2) & \dots & p_0(x_n) \\ p_1(x_1) & p_1(x_2) & \dots & p_1(x_n) \\ \dots \\ p_{n-1}(x_1) & p_{n-1}(x_2) & \dots & p_{n-1}(x_n) \end{vmatrix} \tag{15}$$

$$= \sqrt{D_{n-1}} \det \begin{vmatrix} p_0(x_1) & p_0(x_2) & \dots & p_0(x_n) \\ p_1(x_1) & p_1(x_2) & \dots & p_1(x_n) \\ \dots \\ p_{n-1}(x_1) & p_{n-1}(x_2) & \dots & p_{n-1}(x_n) \end{vmatrix}$$

This leads to the following useful formula for the *square* Δ_n^2 of the van der Monde determinant. Recall that this appears in the eigenvalue density of the Gaussian unitary ensemble and the complex Wishart ensemble.

Corollary 4.

$$\Delta_n^2 = D_{n-1} \det\left(K_n(x_i, x_j)\right)_{1\le i,j\le n} \quad \textit{where} \tag{16}$$

$$K_n(x, y) := \sum_{k=0}^{n} p_k(x) p_k(y). \tag{17}$$

Proof. By equation (2), the factor Δ_n^2 can be expressed as

$$\begin{aligned} \Delta_n^2 &= D_{n-1} \det(p_k(x_i)) \det(p_k(x_j)) \\ &= D_{n-1} \det\left(\sum_{k=0}^{n-1} p_k(x_i) p_k(x_j)\right)_{1\le i,j\le n} \\ &= D_{n-1} \det\left(K_n(x_i, x_j)\right)_{1\le i,j\le n}. \end{aligned}$$

□

Proposition 5. *(Christoffel-Darboux Formula)*

$$K_n(x, y) = \frac{\pi_n(x)\pi_{n-1}(y) - \pi_n(y)\pi_{n-1}(x)}{x - y} \tag{18}$$

Proof. This follows by a routine calculation from the 3−term recurrence. □

Proposition 6. *The Christoffel-Darboux kernels are self-reproducing, that is, for each* $n = 0, 1, 2, \dots,$

$$\int K_n(x, y) K_n(y, z)\, d\mu(y) = K_n(x, z). \tag{19}$$

Proof.

$$\begin{aligned}\int K_n(x,y)K_n(y,z)\,d\mu(y) &= \int \sum_{k=0}^{n}\sum_{l=0}^{n} p_k(x)p_k(y)p_l(y)p_l(z)\,d\mu(y)\\ &= \sum_{k=0}^{n}\sum_{l=0}^{n} p_k(x)p_l(z)\int p_k(y)p_l(y)\,d\mu(y)\\ &= \sum_{k=0}^{n}\sum_{l=0}^{n} p_k(x)p_l(z)\delta_{k,l}\\ &= \sum_{k=0}^{n} p_k(x)p_k(z) = K_n(x,z).\end{aligned}$$

□

3. Jacobi Matrices

3.1. **Jacobi matrices and orthogonal polynomials.** A *Jacobi matrix* is a symmetric, tridiagonal matrix, with positive off-diagonal entries. The spectral theory of Jacobi matrices and their infinite counterparts is intimately tied up with the theory of orthogonal polynomials. The point of connection is the fundamental 3−term recurrence formula.

Proposition 7. *For each $n \geq 0$, the monic orthogonal polynomial $\pi_{n+1}(x)$ is the characteristic polynomial of the Jacobi matrix*

$$J_n = \begin{pmatrix} \alpha_0 & \sqrt{\beta_1} & 0 & 0 & \cdots & 0 & 0\\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & 0 & \cdots & 0 & 0\\ 0 & \sqrt{\beta_2} & \alpha_2 & \sqrt{\beta_3} & \cdots & 0 & 0\\ & & & \cdots & & & \\ 0 & 0 & 0 & 0 & \cdots & \alpha_{n-1} & \sqrt{\beta_n}\\ 0 & 0 & 0 & 0 & \cdots & \sqrt{\beta_n} & \alpha_n \end{pmatrix} \tag{20}$$

Proof. The most direct proof is by induction on n, using the 3-term recurrence formula: expansion of the determinant $\det(xI - J_n)$ along the bottom row, together with the induction hypothesis, gives

$$\det(xI - J_n) = (x - \alpha_n)\pi_n(x) - \beta_n\pi_{n-1}(x).$$

□

Remark 1. The same argument shows that $\pi_n(x)$ is also the characteristic polynomial of the asymmetric tridiagonal matrix

$$A_n := \begin{pmatrix} \alpha_0 & \beta_1 & 0 & 0 & \cdots & 0 & 0\\ 1 & \alpha_1 & \beta_2 & 0 & \cdots & 0 & 0\\ 0 & 1 & \alpha_2 & \beta_3 & \cdots & 0 & 0\\ & & & \cdots & & & \\ 0 & 0 & 0 & 0 & \cdots & \alpha_{n-1} & \beta_n\\ 0 & 0 & 0 & 0 & \cdots & 1 & \alpha_n \end{pmatrix}. \tag{21}$$

(Alternatively, one can check that the matrices A_n and J_n are similar, and therefore have the same characteristic polynomial and the same eigenvalues.)

Lemma 8. *The eigenvalues of a Jacobi matrix are all real. If $v = (v_0, v_1, \dots, v_n)$ is a nonzero eigenvector of the Jacobi matrix J_{n+1}, then $v_0 \neq 0$ and $v_n \neq 0$. Consequently, the spectrum of J_{n+1} is* simple*, that is, the eigenspace of any eigenvalue is one-dimensional.*

Proof. The first statement is true of any symmetric real matrix, hence in particular for a Jacobi matrix. Consider the eigenvalue equation(s) $(J-\lambda)v = 0$: Since J is tridiagonal, these equations are 2nd order difference equations (i.e., 3-term recurrences), and since the off-diagonal entries $\sqrt{\beta_j} > 0$, each component equation determines v_{k+1} as a linear combination of v_k and v_{k-1} (and v_1 as a nonzero scalar multiple of v_0):

$$\begin{aligned} v_1 &= \beta_1^{-1/2}(\lambda - a_0)v_0; \\ v_2 &= \beta_2^{-1/2}((\lambda - a_1)v_1 - \beta_1^{1/2}v_0); \\ &\text{etc.} \end{aligned}$$

Hence, if $v_0 = 0$ then $v_1 = 0$; then $v_2 = 0$; and so by induction all entries $v_k = 0$. A similar argument shows that if $v_n = 0$ then all $v_k = 0$.

Now suppose that for some eigenvalue λ there were two nonzero eigenvectors v, w. By the previous paragraph, both v_0 and w_0 are nonzero. By scaling, we may arrange that $v_0 = w_0 = 1$. But then $v - w$ is an eigenvector with eigenvalue λ whose 0th entry is 0. It folows that $v - w$ is identically zero, that is, $v = w$. □

Corollary 9. *The nth orthogonal polynomial $\pi_n(x)$ has n distinct real zeros. Furthermore, the zeros of $\pi_n(x)$ and $\pi_{n+1}(x)$ are strictly interlaced, that is, if ζ_j^k, for $1 \leq j \leq k$, are the zeros of $\pi_k(x)$ then*

$$\zeta_1^{n+1} < \zeta_1^n < \zeta_2^{n+1} < \zeta_2^n < \cdots < \zeta_n^{n+1} < \zeta_n^n < \zeta_{n+1}^{n+1}. \tag{22}$$

Proof. The zeros of $\pi_k(x)$ are the eigenvalues of $J_{k=1}$, which by the preceding lemma are real and simple; thus, $\pi_k(x)$ has k distinct zeros. Next, since J_n is a submatrix of J_{n+1}, Cauchy's interlacing principle for Hermitian matrices implies that the eigenvalues of J_n and J_{n+1} are at least *weakly* interlaced, so what must be shown is that the inequalities in (22) are strict. Suppose not; then $\pi_n(x)$ and $\pi_{n+1}(x)$ would have at least one zero ζ in common. But then, by the 3−term recurrence, ζ would be a zero of $\pi_{n-1}(x)$; and by induction, ζ would be a zero of every $\pi_k(x)$, for $0 \leq k \leq n+1$. But this is impossible, because $\pi_0(x) = 1$ has no zeros! □

Proposition 10. *For each zero $\lambda = \zeta_j^{n+1}$ of $\pi_{n+1}(x)$, the eigenspace of J_n corresponding to the eigenvalue λ is spanned by the eigenvector*

$$v(\lambda) := (\pi_0(\lambda), \pi_1(\lambda), \dots, \pi_n(\lambda))^T. \tag{23}$$

Since distinct eigenspaces of J_{n+1} are orthogonal, it follows that for distinct zeros λ, λ' of $\pi_{n+1}(x)$,

$$\sum_{j=0}^{n} \pi_j(\lambda)\pi_j(\lambda') = 0. \tag{24}$$

Proof. It suffices to check that $v(\lambda)$ is an eigenvector of A_{n+1}, because the matrices J_{n+1} and A_{n+1} are similar. But this is an immediate consequence of the 3−term recurrence relations (13). □

3.2. **Spectral resolution of a Jacobi operator.** Proposition 7 shows that for each probability measure μ on the real line with finite moment generating function there is an associated sequence of Jacobi matrices whose characteristic polynomials are the monic orthogonal polynomials for μ. The reverse is also true: for each infinite Jacobi matrix J satisfying certain mild hypotheses there is a probability measure μ whose orthogonal polynomials satisfy 3-term recurrence relations with coefficients taken from J. This fact won't be needed for applications to random matrices, but it is of some importance in other areas of probability, notably in the study of birth-and-death processes.

Let J be an infinite Jacobi matrix:

$$J = \begin{pmatrix} \alpha_0 & \sqrt{\beta_1} & 0 & 0 & \cdots \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & 0 & \cdots \\ 0 & \sqrt{\beta_2} & \alpha_2 & \sqrt{\beta_3} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix} \tag{25}$$

Assume that

(1) the entries of J are all bounded in absolute value by a constant $C < \infty$, and
(2) every $\beta_i > 0$.

Then J can be viewed as a linear operator

$$J : \ell^2 \longrightarrow \ell^2$$

where ℓ^2 is the space of square-summable infinite sequences. Because the matrix J is real and symmetric, the operator J is *self-adjoint*, that is, for any two elements $x, y \in \ell^2$,

$$\langle x, Jy \rangle = \langle Jx, y \rangle.$$

Here $\langle x, y \rangle$ is the standard inner product on ℓ^2. Also, since the entries of J are bounded by C, the operator J is *bounded* in the following sense: for any $x \in \ell^2$ the ℓ^2 norm of Jx is bounded by $\sqrt{3}C\|x\|_2$. In other words, the *operator norm* of J is bounded by $\sqrt{3}C$.

Definition 1. For any linear operator $T : \ell^2 \to \ell^2$ the *operator norm* of T is defined as follows:

$$\|T\| = \sup_{x \in \ell^2 : \|x\|_2 = 1} \|T\|_2$$

where $\|x\|_2$ is the usual norm on ℓ^2. An operator with finite operator norm is said to be *bounded.*

Proposition 11. *Let $T : \ell^2 \to \ell^2$ be a bounded, self-adjoint linear operator. Then for each unit vector $u \in \ell^2$ there is a Borel probability measure μ_u on the interval $[-\|T\|, \|T\|]$ whose moments are given by*

$$\langle u, T^n u \rangle = \int x^n \, d\mu_u(x) \quad \textit{for all } n \geq 0. \tag{26}$$

The proof would take us too far into functional analysis, so it is omitted. (See REED & SIMON, *Functional Analysis*, sec. VII.2.) The measures μ_u are called the *spectral measures* for the operator T. Note that the identity (26) extends from monomials to polynomials by linearity: in particular, if $p(x), q(x)$ are any polynomials with real coefficients, then

$$\langle p(T)u, q(T)u \rangle = \int p(x) q(x) \, d\mu_u(x). \tag{27}$$

Suppose now that $T = J$ is a Jacobi operator (25) with bounded entries. Let $e_0, e_1, \dots$ be the standard unit vectors in ℓ^2

$$e_0 = (1,0,0,0,\dots),$$
$$e_1 = (0,1,0,0,\dots), \quad \text{etc.},$$

and let μ_n be the spectral probability measures of the vectors e_n.

Proposition 12. *Let $\pi_n(x)$ and $p_n(x)$ be the monic and normalized orthogonal polynomials, respectively, associated with the spectral probability measure μ_0. Then*

(a) *the coefficients α_n and β_n in the* 3*−term recurrence relations* (13) *are the same as the coefficients α_n and β_n that appear in the entries of the Jacobi matrix J; and*
(b) *for each $n \geq 1$, the vector $p_n(T)e_1$ is the nth standard unit vector $e_n = (0,0,\dots,0,1,0,\dots)$, that is,*

$$p_n(T)e_0 = e_n.$$

Proof. Let's begin with statement (b). Since J is tri-diagonal and p_n is a polynomial of degree n, the vector $p_n(T)e_0$ must be a linear combination of $e_0, e_1, \dots, e_n$. Hence, to prove (b) it suffices to show that for $n \neq m$ the vectors $p_n(T)e_0$ and $p_m(T)e_0$ are orthogonal and both have norm 1. But this follows immediately from the inner product formula (27) and the definition of the orthogonal polynomials $p_n(x)$.

Next, let's prove statement (a). Using the formula (25) and (b), multiply the vector e_n by J to obtain

$$Je_n = \sqrt{\beta_n}e_{n-1} + \alpha_n e_n + \sqrt{\beta_{n+1}}e_{n+1} \Longrightarrow$$
$$Jp_n(J)e_0 = \sqrt{\beta_n}p_{n-1}(J) + \alpha_n p_n(J)e_0 + \sqrt{\beta_{n+1}}p_{n+1}(J)e_0.$$

It now follows from the inner product formula (27) that

$$xp_n(x) = \sqrt{\beta_n}p_{n-1}(x) + \alpha_n p_n(x) + \sqrt{\beta_{n+1}}p_{n+1}(x).$$

This is equivalent to (13), by (14). $\square$

4. Hermite Polynomials

4.1. **Definition.** The most important orthogonal polynomials in probability theory are the *Hermite polynomials*. These are the monic orthogonal polynomials associated with the unit Gaussian distribution. Alternatively, the Hermite polynomials can be defined by the *Rodrigues formula*

$$H_n(x) = \exp\{x^2/2\}\left(-\frac{d}{dx}\right)^n \exp\{-x^2/2\}. \tag{28}$$

For each $n \geq 0$, the function $H_n(x)$ so defined is a monic polynomial of degree n. The first few Hermite polynomials are

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

4.2. **Orthogonality.** *The Hermite polynomials $H_n(x)$ are orthogonal in $L^2(\mu)$, where μ is the unit Gaussian distribution:*

$$\int H_n(x) H_m(x) \exp\{-x^2/2\}\, dx/\sqrt{2\pi} = \delta_{n,m} m! \tag{29}$$

Proof. Assume that $n \geq m$. Use the definition (28) to rewrite the integral (29) as

$$\int H_m(x) \left(-\frac{d}{dx}\right)^n \exp\{-x^2/2\}\, dx/\sqrt{2\pi}.$$

Now integrate by parts repeatedly to move the derivatives d/dx onto the factor $H_m(x)$; note that the boundary terms in all of these integrations by parts vanish, because the polynomial factors are killed by the $\exp\{-x^2/2\}$. The result is

$$\int \left(\left(\frac{d}{dx}\right)^n H_m(x)\right) \exp\{-x^2/2\}\, dx/\sqrt{2\pi}.$$

If $n > m$, the first factor will be identically 0, because $H_m(x)$ is a polynomial of degree m. If $n = m$, then the first factor will be $m!$, because $H_m(x)$ is *monic*, and so the integral will be $m!$. □

4.3. 3−**Term Recurrence.** *The Hermite polynomials satisfy the* 3− *term recurrence relation*

$$xH_n(x) = H_{n+1}(x) + nH_{n-1}(x). \tag{30}$$

Proof. Using the shorthand notation $D = d/dx$, $E_\pm = \exp\{\pm x^2/2\}$,

$$\begin{aligned} H_{n+1}(x) &= E_+(-D)^n(-DE_-) \\ &= E_+(-D)^n(xE_-) \\ &= xE_+(-D)^n E_- - nE_+(-D)^{n-1}E_- \\ &= xH_n(x) - nH_{n-1}(x). \end{aligned}$$

The third equation is gotten by noting that in taking $(-D)^n$ there are n chances to let the $-D$ act on the factor x. □

4.4. **Generating Function.** *The exponential generating function of the Hermite polynomials is given by the identity*

$$\mathscr{H}(t,x) := \sum_{n=0}^{\infty} \frac{t^n}{n!} H_n(x) = \exp\left\{tx - t^2/2\right\}. \tag{31}$$

Proof. Taking the exponential generating function of both sides of the 3−term recurrence relation (30) yields

$$x\mathscr{H}(t,x) = \frac{\partial}{\partial t}\mathscr{H}(t,x) + t\mathscr{H}(t,x).$$

For each fixed x, this is a first-order ODE whose solution is determined by the initial condition $\mathscr{H}(0,x) = H_0(x) = 1$. This implies that $\log \mathscr{H}(0,x) = 0$, and so the solution to the ODE is

$$\mathscr{H}(t,x) = \exp\{xt - t^2/2\}.$$

□

4.5. Distribution of Zeros.

Proposition 13. *Let $\{\zeta_k^n\}_{1\le k\le n}$ be the zeros of the nth Hermite polynomial $H_n(x)$, and let F_n be the uniform distribution on the set $\{\zeta_k^n/(2\sqrt{n})\}$. Then as $n\to\infty$,*

$$F_n \Longrightarrow \text{semicircle law.} \tag{32}$$

Proof. The zeros of $H_n(x)$ are the eigenvalues of the Jacobi matrix J_n defined by (20) with $\alpha_j = 0$ and $\beta_j = j$ (by the 3−term recurrence relations (30)). Thus, the moments of the distribution F_n are

$$\int x^k F_n(dx) = \operatorname{tr}(J_n^k)\Big/ 2^k n^{k/2}.$$

Because a Jacobi matrix is tridiagonal, the trace of its kth power, for any nonnegative integer k, is the sum of the *weights* $w(\gamma)$ of all nearest-neighbor loops

$$\gamma = x_0 x_1 x_2 \cdots x_k \quad \text{where } x_k = x_0 \text{ and } 0 \le x_j \le n$$

of length k (here nearest-neighbor means that $|x_j - x_{j+1}| = 1$), with weights $w(\gamma)$ defined by

$$w(\gamma) = \prod_{j=1}^{k} \sqrt{x_j}.$$

For each integer $0 \le m \le n$, denote by $\mathscr{P}_k(m)$ the set of all length-k nearest-neighbor loops that begin and end at m. Observe that $\mathscr{P}_k(m) = \emptyset$ if k is odd; hence, all odd moments of F_n vanish. Thus, it suffices to show that the even moments of F_n converge as $n \to \infty$ to those of the semicircle law.

Fix $2k \ge 0$ even, and consider loops γ of length $2k$ that begin and end at sites $m \gg k$. For any site x_j on such a path, $x_j/x_0 \sim 1$ (because $|x_j - x_0| \le k$), and so

$$w(\gamma) \sim m^k.$$

This approximation holds uniformly for paths whose endpoints m are large compared to k. Since paths with endpoints m that are not large compared to k contribute negligibly to $\operatorname{tr}(J_n^k)$, it follows that

$$\operatorname{tr}(J_n^{2k})/n^k \sim \sum_{m=1}^{n} (m/n)^k \#\mathscr{P}_{2k}(m).$$

Except for endpoints m that are within distance k of 0 or m, the set $\mathscr{P}_{2k}(m)$ of loops rooted at m that lie entirely in $[0,n]$ coincides with the set of all loopsin the full integer lattice $\mathbb{Z}$ of length $2k$ that are rooted at m; hence, for all such m,

$$\#\mathscr{P}_{2k}(m) = \binom{2k}{k}.$$

Consequently,

$$\operatorname{tr}(J_n^{2k})/(2^{2k}n^k) \sim 2^{-2k}\sum_{m=1}^{n}(m/n)^k\binom{2k}{k} \sim 2^{-2k}\binom{2k}{k}\int_0^1 y^k\,dy = 2^{-2k}\binom{2k}{k}\Big/(k+1),$$

which we recognize as the $2k$th moment of the semicircle law. □

Note: Deift, ch. 6, shows by a much more roundabout argument that the semicircle law governs the asymptotic distribution of zeros not only for the Hermite ensemble but for a large class of orthogonal polynomial ensembles.

Exercise 2. The Hermite polynomials have an interesting and sometimes useful combinatorial interpretation. This involves the notion of a *dimer arrangement* (or *matching*) on the set of vertices $[n]$ (more precisely, on the complete graph K_n on the vertex set $[n]$). A *dimer arrangement* on $[n]$ is a set $\mathscr{D}$ of unordered pairs $d = \{j, j'\}$ of distinct elements such that no two pairs $d, d' \in \mathscr{D}$ have a vertex in common. The number of dimers $d \in \mathscr{D}$ is the *weight* $w(\mathscr{D})$ of the dimer arrangement.

(A) Show that $H_n(x/2) = \sum_{\mathscr{D}} (-1)^{w(\mathscr{D})} x^{n-2w(\mathscr{D})}$.

(B) Show that if n is even then the number of *dimer coverings* of $[n]$ (that is, dimer arrangements of weight $n/2$) is the nth moment of the unit Gaussian distribution:

$$\int x^n e^{-x^2/2}\, dx/\sqrt{2\pi}$$

4.6. **Plancherel-Rotach Asymptotics.** By Corollary 4, the factor Δ_n^2 in the GUE eigenvalues density can be expressed as a determinant involving the Christoffel-Darboux kernel $K_n(x, y)$, which in turn, by the Christoffel-Darboux formula, can be expressed as a function of the Hermite polynomials $H_n(x)$ and $H_{n-1}(x)$. Now the bulk of the spectrum of a GUE random matrix lies between $-2\sqrt{n}$ and $+2\sqrt{n}$, so the interesting values of the argument x are in the range $-2 \le x/\sqrt{n} \le 2$, and in particular at the edge $x \approx 2\sqrt{n}$, where the maximum eigenvalue will occur.

Proposition 14. *For each $u \in \mathbb{R}$, and uniformly on compact subsets of $\mathbb{R}$,*

$$H_n(2\sqrt{n} + n^{-1/6}u) \sim \frac{n! \exp\{\frac{3}{2}n + n^{1/3}u\}}{2\pi n^{1/3} n^{n/2}} Ai(u) \tag{33}$$

where $Ai(u)$ is the Airy function

$$Ai(u) = \frac{1}{\pi} \int_0^\infty \cos\left(ut + t^3/3\right)\, dt. \tag{34}$$

Proof. The generating function (31) for the Hermite polynomials can be used, in conjunction with Fourier inversion, to obtain an exact integral formula for $H_n(x)$:

$$H_n(x) = \frac{n!}{2\pi r^n} \int_{-\pi}^{\pi} \exp\{re^{i\theta}x - r^2 e^{2i\theta}/2\} e^{-in\theta}\, d\theta.$$

This is valid for any real $r > 0$. The usual approach to deducing asymptotic formulas from integrals with a free parameter r such as this is to look for a *saddle-point*, that is, a point $r > 0$ where the integrand has a local maximum as a function of θ (and thus a local *minimum* in r for $\theta = 0$). The integrand for $\theta = 0$ takes the form $r^{-n} e^{rx - r^2/2}$. To find a saddle-point, set the derivative equal to 0: this leads to the quadratic equation

$$-n + rx - r^2/2 = 0 \quad \Longrightarrow \quad r = \frac{-x \pm \sqrt{x^2 - 4n}}{-2}.$$

Note that when $x = 2\sqrt{n}$ the discriminant is 0, so the quadratic has a double root: this is ultimately why we see see cube root asymptotics in (33). When $x = 2\sqrt{n} + n^{-1/6}u$, the double root splits for $u \neq 0$, but the two roots are close, and so it is just as easy to use the approximate saddle-point

$$r = \sqrt{n}$$

in the Fourier integral. Thus,

$$H_n(2\sqrt{n} + n^{-1/6}u) = \frac{n!}{2\pi \sqrt{n}^n} \int_{-\pi}^{\pi} \exp\left\{(2n + n^{1/3}u)e^{i\theta} - ne^{2i\theta}/2 - in\theta\right\} d\theta$$

Now collect terms grouped in powers of θ in the exponential:

$$\begin{aligned}\exp\left\{(2n+n^{1/3}u)e^{i\theta}-ne^{2i\theta}/2-in\theta\right\} &= \exp\left\{2n+n^{1/3}u-n/2\right\}\\ &\quad \exp\left\{(2n+n^{1/3}u)i\theta-(n/2)2i\theta-in\theta\right\}\\ &\quad \exp\left\{-(2n+n^{1/3}u+4n^2/2)\theta^2/2\right\}\\ &\quad \exp\left\{-((2n+n^{1/3}u)+8n/2)(i\theta^3/6)+\cdots\right\}\\ &= \exp\left\{\frac{3}{2}n+n^{1/3}u\right\}\\ &\quad \exp\left\{in^{1/3}u\theta-\theta^2 n^{1/3}u/2+in\theta^3/3-n^{1/3}u\theta^3/6+\cdots\right\}\end{aligned}$$

Finally, make the substitution $t=n^{1/3}\theta$ to obtain

$$H_n(2\sqrt{n}+n^{-1/6}u)=\frac{n!}{2\pi\sqrt{n}^n}\exp\left\{\frac{3}{2}n+n^{1/3}u\right\}n^{-1/3}\int_{-\infty}^{\infty}\exp\left\{iut+it^3/3+\cdots\right\}dt$$

$\square$

5. Exercises: Laguerre Polynomials

The *Laguerre polynomials* $L_n^{(\alpha)}(x)$ are the orthogonal polynomials for the Gamma distribution with density proportional to $x^\alpha e^{-x}$ on the positive reals $(0,\infty)$. Recall that the Gamma density occurs in the eigenvalue density of the Wishart ensemble in the same way that the Gaussian density occurs in the GUE and GOE eigenvalue densities. Thus, the Laguerre polynomials play the same role in the theory of the WIshart distribution that the Hermite polynomials play for the GUE and GOE ensembles. As is the case for the Hermite polynomials, the Laguerre polynomials can be defined by a *Rodrigues formula*

$$L_n^{(\alpha)}(x)=\frac{1}{n!}x^{-\alpha}e^x\left(\frac{d}{dx}\right)^n(x^{n+\alpha}e^{-x}). \tag{35}$$

Exercise 3. Check that the formula (35) defines a polynomial of degree n whose leading term is $(-1)^n/n!$:

$$L_n^{(\alpha)}(x)=(-1)^n x^n/n!+\cdots$$

Thus, the *monic* orthogonal polynomials for the Gamma weight function are $n!L^{(\alpha)_n(x)}$.

Exercise 4. Check that the Laguerre polynomials are orthogonal with respect to the Gamma density:

$$\int_0^\infty L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)x^\alpha e^{-x}\,dx=\frac{\Gamma(\alpha+n+1)}{\Gamma(n+1)}\delta_{m,n}$$

Note: Recall that $n!=\Gamma(n+1)$.

Exercise 5. Show that the generating function of the Laguerre polynomials is

$$\sum_{n=0}^\infty t^n L_n^{(\alpha)}(x)=\exp\left\{-\frac{xt}{1-t}\right\}\Big/(1-t)^{\alpha+1}:=\mathscr{L}(t,x).$$

Hint: For this it is easiest (I think) to use the orthogonality relations of Exercise 4. Start by evaluating

$$\int_0^\infty \mathscr{L}(t,x)\mathscr{L}(s,x)x^\alpha e^{-x}\,dx.$$

Exercise 6. Use the generating function to check that the Laguerre polynomials satisfy the 3−term recurrence relations

$$nL_n^{(\alpha)}(x) = (-x+2n+\alpha-1)L_{n-1}^{(\alpha)}(x) - (n+\alpha-1)L_{n-2}^{(\alpha)}(x).$$

Exercise 7. Formulate and prove a theorem about the distribution of zeros of $L_n^{(\alpha)}(x)$ for large n.