MATRIX MODELS FOR CIRCULAR ENSEMBLES

ROWAN KILLIP AND IRINA NENCIU

Abstract. The Gibbs distribution for n particles of the Coulomb gas on the unit circle at inverse temperature β is given by

$$\mathbb{E}_n^\beta(f) = \frac{1}{Z_{n,\beta}} \int \cdots \int f(e^{i\theta_1}, \ldots, e^{i\theta_n}) |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^\beta \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_n}{2\pi}$$

for any symmetric function f, where Δ denotes the Vandermonde determinant and $Z_{n,\beta}$ the normalization constant.

We will describe an ensemble of (sparse) random matrices whose eigenvalues follow this distribution. Our approach combines elements from the theory of orthogonal polynomials on the unit circle with ideas from recent work of Dumitriu and Edelman. In particular, we resolve a question left open by them: find a tri-diagonal model for the Jacobi ensemble.

1. Introduction

In 1962, Dyson [6] introduced three ensembles of random unitary matrices with a view to simplifying the study of energy level behavior in complex quantum systems. Earlier work in this direction, pioneered by Wigner, focused on ensembles of Hermitian matrices.

The simplest of these three models is the unitary ensemble, which is just the group $\mathbb{U}(n)$ of $n \times n$ unitary matrices together with its Haar measure. The induced probability measure on the eigenvalues is given by the Weyl integration formula (cf. [17, §VII.4]): for any symmetric function of the eigenvalues,

$$\mathbb{E}(f) = \frac{1}{n!} \int_0^{2\pi} \cdots \int_0^{2\pi} f(e^{i\theta_1}, \ldots, e^{i\theta_n}) \big|\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})\big|^2 \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_n}{2\pi} \tag{1.1}$$

where Δ denotes the Vandermonde determinant,

$$\Delta(z_1, \ldots, z_n) = \prod_{1 \le j < k \le n} (z_k - z_j) = \begin{vmatrix} 1 & \ldots & 1 \\ \vdots & & \vdots \\ z_1^{n-1} & \ldots & z_n^{n-1} \end{vmatrix}. \tag{1.2}$$

The orthogonal ensemble consists of symmetric $n \times n$ unitary matrices together with the unique measure that is invariant under $U \mapsto W^T U W$ for all $W \in \mathbb{U}(n)$. Alternatively, if U is chosen according to the unitary ensemble, then $U^T U$ is distributed as a random element from the orthogonal ensemble. The distribution of eigenvalues is given by (1.1) but with $|\Delta|^2$ replaced by $|\Delta|$ and a new normalization constant.

Date: May 12, 2004.

The symplectic ensemble is a little more complicated. Let Z denote the $2n \times 2n$ block diagonal matrix

$$\begin{bmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{bmatrix} \tag{1.3}$$

and define the dual of a matrix by $U^R = Z^T U^T Z$. The symplectic ensemble consists of self-dual unitary $2n \times 2n$ matrices; the measure is that induced from $\mathbb{U}(2n)$ by the map $U \mapsto U^R U$. (This is the unique measure invariant under $V \mapsto W^R V W$ for all $W \in \mathbb{U}(2n)$.) The eigenvalues of such matrices are doubly degenerate and the pairs are distributed on the circle as in (1.1) but now with $|\Delta|^2$ replaced by $|\Delta|^4$. Again, the normalization constant needs to be changed.

Dyson also observed that these eigenvalue distributions correspond to the Gibbs distribution for the classical Coulomb gas on the circle at three different temperatures. Let us elaborate.

Consider n identically charged particles confined to move on the unit circle in the complex plane. Each interacts with the others through the usual Coulomb potential, $-\log|z_i - z_j|$, which gives rise to the Hamiltonian

$$H(z_1, \dots, z_n) = \sum_{1 \le j < k \le n} -\log|z_j - z_k|.$$

(One may add a kinetic energy term; however, as we are interested only in the distribution of the particle positions, it has no effect.) This gives rise to the Gibbs measure (with parameters n, the number of particles, and β, the inverse temperature)

$$\mathbb{E}_n^\beta(f) = \frac{1}{(2\pi)^n Z_{n,\beta}} \int \cdots \int f(e^{i\theta_1}, \dots, e^{i\theta_n}) e^{-\beta H(e^{i\theta_1}, \dots, e^{i\theta_n})}\, d\theta_1 \cdots d\theta_n \tag{1.4}$$

$$= \frac{1}{(2\pi)^n Z_{n,\beta}} \int \cdots \int f(e^{i\theta_1}, \dots, e^{i\theta_n}) \big|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})\big|^\beta \, d\theta_1 \cdots d\theta_n \tag{1.5}$$

for any symmetric function f. The partition function is given by

$$Z_{n,\beta} = \frac{\Gamma(\frac{1}{2}\beta n + 1)}{\big[\Gamma(\frac{1}{2}\beta + 1)\big]^n} \tag{1.6}$$

as conjectured by Dyson. This was proved by Gunson [11] and Wilson [18], though the Good proof [10] is even better. We give another proof at the end of Section 4.

From the discussion above, we see that the orthogonal, unitary, and symplectic ensembles correspond to the Coulomb gas at inverse temperatures $\beta = 1$, 2, and 4.

Viewed from the opposite perspective, one may say that Dyson provided matrix models for the Coulomb gas at three different temperatures. Our first goal here is to present a family of matrix models for all temperatures. These matrices will be sparse—approximately $4n$ non-zero entries—which suggests certain computational advantages. To state the theorem, we need the following

Definition 1.1. We say that a complex random variable, X, with values in the unit disk, $\mathbb{D}$, is Θ_ν-distributed (for $\nu > 1$) if

$$\mathbb{E}\{f(X)\} = \tfrac{\nu-1}{2\pi} \iint_{\mathbb{D}} f(z)(1-|z|^2)^{(\nu-3)/2}\, d^2z. \tag{1.7}$$

For $\nu \geq 2$ an integer, this has the following geometric interpretation: If v is chosen from the unit sphere S^ν in $\mathbb{R}^{\nu+1}$ at random according to the usual surface measure, then $v_1 + iv_2$ is Θ_ν-distributed. (See Corollary A.3.)

As a continuation of this geometric picture, we shall say that X is Θ_1-distributed if it is uniformly distributed on the unit circle in $\mathbb{C}$.

Let us now describe the family of matrix models.

Theorem 1. *Given $\beta > 0$, let $\alpha_k \sim \Theta_{\beta(n-k-1)+1}$ be independent random variables for $0 \leq k \leq n-1$, $\rho_k = \sqrt{1-|\alpha_k|^2}$, and define*

$$\Xi_k = \begin{bmatrix} \bar{\alpha}_k & \rho_k \\ \rho_k & -\alpha_k \end{bmatrix}$$

for $0 \leq k \leq n-2$, while $\Xi_{-1} = [1]$ and $\Xi_{n-1} = [\bar{\alpha}_{n-1}]$ are 1×1 matrices. From these, form the $n \times n$ block-diagonal matrices

$$L = \operatorname{diag}\big(\Xi_0, \Xi_2, \Xi_4, \dots\big) \quad \textit{and} \quad M = \operatorname{diag}\big(\Xi_{-1}, \Xi_1, \Xi_3, \dots\big).$$

Both LM and ML give (sparse) matrix models for the Coulomb gas at inverse temperature β. That is, their eigenvalues are distributed according to (1.5).

Remark. As each of the Ξ_k is unitary, so are L and M. (In the case of Ξ_{n-1}, we should reiterate that $\alpha_{n-1} \sim \Theta_1$ is uniformly distributed on the unit circle.) As a result, the eigenvalues of LM and ML lie on the unit circle. Note also that, since M conjugates one to the other, LM and ML have the same eigenvalues.

In proving this theorem, we will be following the recent paper of Dumitriu and Edelman [5] rather closely, while incorporating the nuances of the theory of polynomials orthogonal on the unit circle. The matrices L and M that appear in the theorem have their origin in the work of Cantero, Moral, and Velásquez [3]; this is discussed in Section 2.

Dumitriu and Edelman constructed tri-diagonal matrix models for two of the three standard examples of the Coulomb gas on the real line. A model for the third will be constructed below.

The simplest way to obtain a normalizable Gibbs measure on the real line is to add an external harmonic potential $V(x) = \frac{1}{2}x^2$. This gives rise to the probability measure

$$\mathbb{E}(f) \propto \int\!\cdots\!\int f(x_1,\dots,x_n)\,\big|\Delta(x_1,\dots,x_n)\big|^\beta \prod_j e^{-V(x_j)}\, dx_1\cdots dx_n \tag{1.8}$$

on $\mathbb{R}^n$. This is known as the Hermite ensemble, because of its intimate connection to the orthogonal polynomials of the same name, and when $\beta = 1$, 2, or 4, arises as the eigenvalue distribution in the classical Gaussian ensembles of random matrix theory. Dumitriu and Edelman showed that (1.8) is the distribution of eigenvalues for a symmetric tri-diagonal matrix with independent entries (modulo symmetry). The diagonal entries have standard Gaussian distribution and the lower diagonal entries are $2^{-1/2}$ times a χ-distributed random variable with the number of degrees of freedom equal to β times the number of the row.

The second example treated by Dumitriu and Edelman is the Laguerre ensemble. In statistical circles, this is known as the Wishart ensemble, special cases of which arise in the empirical determination of the covariance matrix of a multivariate Gaussian distribution. For this ensemble, one needs to modify the distribution given in (1.8) in two ways: each particle x_j is confined to lie in $[0,\infty)$ and is subject to the external potential $V(x) = -a\log(x) + x$, where $a > -1$ is a parameter. In [5], it is shown that if B is a certain $n\times n$ matrix with independent χ-distributed entries on the main and sub-diagonal (the number of degrees of freedom depends on a, β, and the element in question) and zeros everywhere else, then the eigenvalues of $L = BB^T$ follow this distribution.

The third canonical form of the Coulomb gas on $\mathbb{R}$ is the Jacobi ensemble. The distribution is as in (1.8), but now the particles are confined to lie within $[-2,2]$ and are subject to the external potential $V(x) = -a\log(2-x) - b\log(2+x)$, where $a, b > -1$ are parameters. This corresponds to the probability measure on $[-2,2]^n$ that is proportional to

$$\big|\Delta(x_1,\ldots,x_n)\big|^\beta \prod_j (2-x_j)^a(2+x_j)^b\,dx_1\cdots dx_n. \tag{1.9}$$

The partition function (or normalization coefficient) was determined by Selberg [14]. This will be discussed in Section 6.

Dumitriu and Edelman did not give a matrix model for this ensemble, listing it as an open problem. We present a tri-diagonal matrix model in Theorem 2 below. The independent parameters follow a beta distribution:

Definition 1.2. A real-valued random variable X is said to be beta-distributed with parameters $s, t > 0$, which we denote by $X \sim B(s,t)$, if

$$\mathbb{E}\{f(X)\} = \frac{2^{1-s-t}\Gamma(s+t)}{\Gamma(s)\Gamma(t)}\int_{-1}^{1} f(x)(1-x)^{s-1}(1+x)^{t-1}\,dx. \tag{1.10}$$

Note that $B(\frac{\nu}{2},\frac{\nu}{2})$ is the distribution of the first component of a random vector from the ν-sphere. (See Corollary A.3.)

Theorem 2. *Given $\beta > 0$, let α_k, $0 \le k \le 2n-2$, be distributed as follows*

$$\alpha_k \sim \begin{cases} B(\frac{2n-k-2}{4}\beta + a + 1, \frac{2n-k-2}{4}\beta + b + 1) & k \text{ even},\\ B(\frac{2n-k-3}{4}\beta + a + b + 2, \frac{2n-k-1}{4}\beta) & k \text{ odd}.\end{cases} \tag{1.11}$$

Let $\alpha_{2n-1} = \alpha_{-1} = -1$ and define

$$b_{k+1} = (1-\alpha_{2k-1})\alpha_{2k} - (1+\alpha_{2k-1})\alpha_{2k-2} \tag{1.12}$$

$$a_{k+1} = \big\{(1-\alpha_{2k-1})(1-\alpha_{2k}^2)(1+\alpha_{2k+1})\big\}^{1/2} \tag{1.13}$$

for $0 \le k \le n-1$; then the eigenvalues of the tri-diagonal matrix

$$J = \begin{bmatrix} b_1 & a_1 & & \\ a_1 & b_2 & \ddots & \\ & \ddots & \ddots & a_{n-1} \\ & & a_{n-1} & b_n \end{bmatrix}$$

are distributed according to the the Jacobi ensemble (1.9).

We know of two other papers which discuss the Jacobi ensemble in a manner inspired by the work of Dumitriu–Edelman: [7, §4.2] and [13]. These results are however of a rather different character; in particular, we contend that Theorem 2 is the true Jacobi ensemble analogue of the results of [5].

In Section 6, we show how the ideas developed in the earlier parts of this paper lead to new derivations of the classical integrals of Aomoto [2] and Selberg [14]. The main novelty of these proofs is their directness: They treat all values of β and n on an equal footing. In particular, we do not prove the result for β an integer and then make recourse to Carlson's Theorem. These remarks are also applicable to the proof of (1.6) given at the end of Section 4.

Acknowledgements: We would like to thank Barry Simon for encouragement and for access to preliminary drafts of his forthcoming book [15].

2. Overview of the Proofs and Background Material

We begin by examining the $\beta = 2$ case of Theorem 1, that is, Haar measure on the unitary group.

Rather than study the eigenvalues as the fundamental statistical object, we will consider the spectral measure associated to U and the vector $e_1 = (1, 0, \dots, 0)^T$. It will be denoted by $d\mu$. As Haar measure is invariant under conjugation, any choice of unit vector e_1 leads to the same probability distribution on $d\mu$.

The most natural coordinates for $d\mu$ are the eigenvalues $e^{i\theta_1}, \dots, e^{i\theta_n}$ and the mass that $d\mu$ gives to them: $\mu_1 = \mu(\{e^{i\theta_1}\}), \dots, \mu_{n-1} = \mu(\{e^{i\theta_{n-1}}\})$. As $\int d\mu = 1$, we omit $\mu_n = \mu(\{e^{i\theta_n}\})$. Note that we have chosen not to order the eigenvalues, which means that the natural parameter space gives an $n!$-fold cover of the set of measures. We have already used this way of thinking a number of times, beginning with (1.1).

The above system of coordinates does not cover the possibility that U has degenerate eigenvalues. However, as the Weyl integration formula shows, the set of such U has zero Haar measure; in fact, the density vanishes quadratically at these points. The reason for this is worth repeating (cf. [17, §VII.4]): The submanifold where two eigenvalues coincide has co-dimension three in $\mathbb{U}(n)$; one degree of freedom is lost from the reduction of the number of eigenvalues and two more are lost in the reduction from two orthogonal one-dimensional eigenspaces to a single two-dimensional eigenspace. One should compare this to spherical polar coordinates in $\mathbb{R}^3$ where $r = 0$ is a submanifold of co-dimension three and consequently, the density also vanishes to second order.

In Section 3, we will determine the probability distribution on $d\mu$ induced from Haar measure on $\mathbb{U}(n)$, in the θ, μ coordinates. Conjugation invariance of Haar measure implies that the eigenvalues and masses are statistically independent; it is then easy to see that the former are distributed as in (1.1) and $(\mu_1, \dots, \mu_n)$ is uniformly distributed on the simplex $\sum \mu_j = 1$. See Proposition 3.1. This implies that $d\mu$ gives non-zero weight to each of the eigenvalues with probability one. As a consequence, we can always recover the eigenvalues from $d\mu$.

We will now introduce different coordinates, $(\alpha_0, \dots, \alpha_{n-1})$, for $d\mu$ that arise in the study of orthogonal polynomials on the unit circle.

The monomials $1, z, \dots, z^{n-1}$ form a basis for $L^2(d\mu)$ and so, applying the Gram–Schmidt procedure, we can construct an orthogonal basis of monic polynomials: Φ_j,

$0 \le j < n$, with Φ_j monic of degree j. We also define $\phi_j = \Phi_j/\|\Phi_j\|$, which gives an orthonormal basis.

There is a well-developed theory of such orthogonal polynomials, parts of which will be important in what follows. For a proper discussion of this theory, see [9], [15], or [16, Ch. XI].

The first important fact about the orthogonal polynomials is that they obey recurrence relations:

$$\Phi_{k+1}(z) = z\Phi_k(z) - \bar{\alpha}_k \Phi_k^*(z) \tag{2.1}$$

$$\Phi_{k+1}^*(z) = \Phi_k^*(z) - \alpha_k z \Phi_k(z) \tag{2.2}$$

where the α_k are the recurrence coefficients and Φ_k^* denotes the reversed polynomial:

$$\Phi_k(z) = \sum_{l=0}^{k} c_l z^l \quad \Rightarrow \quad \Phi_k^*(z) = \sum_{l=0}^{k} \bar{c}_{k-l} z^l. \tag{2.3}$$

Equivalently, $\Phi_k^*(z) = z^k \overline{\Phi_k(\bar{z}^{-1})}$. These recurrence equations imply

$$\big\|\Phi_k\big\|_{L^2(d\mu)} = \prod_{l=0}^{k-1} \rho_l \quad \text{where} \quad \rho_l = \sqrt{1 - |\alpha_l|^2}, \tag{2.4}$$

from which the recurrence relations for the orthonormal polynomials are easily derived.

The recurrence coefficients α_k have been called by many names; we will follow [15], where they were recently dubbed 'Verblunsky parameters'. Each of $\alpha_0, \dots, \alpha_{n-2}$ lies inside the unit disk, $\mathbb{D}$, while α_{n-1} lies on its boundary, S^1.

There is an alternate way of relating measures to their Verblunsky parameters, namely the Schur algorithm: If $d\mu$ is a probability measure, then we define its Schur function $f : \mathbb{D} \to \mathbb{D}$ by

$$f(z) = \frac{1}{z} \frac{F(z) - 1}{F(z) + 1} \quad \text{where} \quad F(z) = \int \frac{e^{i\theta} + z}{e^{i\theta} - z} d\mu(e^{i\theta}). \tag{2.5}$$

The Schur algorithm parameterizes analytic maps $f : \mathbb{D} \to \bar{\mathbb{D}}$ by finitely or infinitely many parameters α_k.

There are finitely many parameters if and only if f is a finite Blaschke product, or, equivalently, iff $d\mu$ has finite support. More precisely, the support of $d\mu$ consists of n points iff f is a Blaschke product of degree n. This is the case when there are n Verblunsky parameters: $\alpha_0, \dots, \alpha_{n-2} \in \mathbb{D}$ and $\alpha_{n-1} \in S^1$. When there are finitely many parameters, the last must always be unimodular. In fact, the final parameter is essentially equal to the product of the locations of the mass points of $d\mu$; see (B.4).

When $d\mu$ has infinite support, there are infinitely many Verblunsky parameters, all of which lie in the unit disk.

Just as the Schur algorithm gives a bijection, so there is a bijection between measures $d\mu$ on S^1 supported at n points and sequences of parameters $\alpha_0, \dots, \alpha_{n-2} \in \mathbb{D}$, $\alpha_{n-1} \in S^1$. This justifies their use as coordinates for the measure $d\mu$.

In Proposition 3.3, we determine the probability distribution on $d\mu$ (induced by Haar measure on $\mathbb{U}(n)$) in these new coordinates. Interestingly, the αs turn out to be statistically independent, with $\alpha_k \sim \Theta_{2n-2k-1}$.

It is now but a few short steps to the $\beta = 2$ case of Theorem 1.

Consider the operator $f(z) \mapsto zf(z)$ in $L^2(d\mu)$. The spectral measure associated to the vector $f(z) \equiv 1$ is simply $d\mu$. To obtain a matrix model, we need only choose a basis in which to represent this operator. The most obvious choice is the basis of orthonormal polynomials, $\{\phi_k\}$. This leads to a matrix whose entries can be expressed simply in terms of the αs. However, this matrix is not sparse: all entries above and including the sub-diagonal are non-zero (with probability one). Such matrices are typically known as being in Hessenberg form. In deference to this, we will denote the matrix by H. It plays an important role in the determination of the distribution of the Verblunsky parameters, but does not appear in Theorem 1.

The matrix LM described in Theorem 1 is $f(z) \mapsto zf(z)$ in $L^2(d\mu)$ in the orthonormal basis formed from $1, z, z^{-1}, \dots$ by applying the Gram–Schmidt procedure. That this matrix can be expressed so simply in terms of the Verblunsky coefficients is a discovery of Cantero, Moral, and Velázquez, [3]. Related matters are discussed in Appendix B. (The matrix ML is the same operator in the basis formed by applying Gram–Schmidt to $1, z^{-1}, z, \dots$.)

Thus far, we have only spoken about the unitary group, that is, about $\beta = 2$. In this case, we have found a random ensemble of measures whose mass points are distributed as the particles in the Coulomb gas at inverse temperature $\beta = 2$. The key discovery, however, was that the corresponding Verblunsky parameters turn out to be independent.

For more general β, we wish to find an ensemble of measures so that the mass points are distributed appropriately; we have complete freedom in choosing how the weights are distributed. By the same token, we want the induced probability distribution on the Verblunsky parameters to retain independence. We can then form the matrix set out in Theorem 1 and its eigenvalues are guaranteed to follow the proper distribution.

The key to satisfying these desires is Lemma 4.1. It expresses the value of the Toeplitz determinant associated to $d\mu$ in terms of the θ, μ coordinates and in terms of the Verblunsky parameters. Multiplying the probability distribution from the $\beta = 2$ case by the appropriate power of the Toeplitz determinant gives Proposition 4.2, which is exactly the resolution of the goals set forth in the previous paragraph.

As an off-shoot of proving Theorem 1, we are able to determine the Jacobian for the map from the (θ, μ)-coordinates to the Verblunsky parameters α_k. That this is possible is a delightful idea of Dumitriu and Edelman [5]. (See Lemmas 4.3 and 4.4.)

Were we granted the Jacobian for this map, the paper could be much shorter—though we contend that the scenic route followed below is not without merit. It is a natural quantity to calculate and the answer takes a rather simple form. This behooves us to find a simple, direct derivation. Thus far, we have failed. We would be much obliged to any reader who can resolve this matter.

The proof of Theorem 2 is very similar. Again we begin by studying the problem for $\beta = 2$. The relevant group in this instance is not $\mathbb{U}(n)$, but rather $\mathbb{SO}(2n)$. Such matrices have eigenvalues in complex conjugate pairs and the corresponding eigenvectors are complex conjugates of one another. Consequently, the spectral measure associated to e_1 is symmetric with respect to complex conjugation. The most natural coordinates are $\theta_j \in (0, \pi)$ and $\mu_j \in [0, 1]$ where

$$\int f d\mu = \sum_{j=1}^{n} \tfrac{1}{2}\mu_j [f(e^{i\theta_j}) + f(e^{-i\theta_j})] \tag{2.6}$$

and $\sum \mu_j = 1$.

Once again, we use the Verblunsky coefficients as a second set of coordinates. These are now real as a consequence of the complex conjugation symmetry of the measure. Indeed, a measure has this symmetry if and only if its Verblunsky coefficients are real. From this and the foregoing discussion of the general case, we see that the last Verblunsky coefficient, α_{2n-1}, must be real and unimodular. In fact it must be -1 because the product of the eigenvalues of a matrix from $\mathbb{SO}(2n)$ is equal to one; see (B.4). The remaining Verblunsky coefficients α_k, $0 \le k \le 2n-2$, are free to range over $(-1, 1)$.

By proceeding very much as before, we can construct certain ensembles of orthogonal matrices for which the spectral measure is distributed in a desirable fashion. When the Verblunsky coefficients are real, both $\det(1 - U)$ and $\det(1 + U)$ have simple expressions in terms of these coefficients. As a result, we are able to add two new parameters a and b to our family of distributions. This line of reasoning leads us to Proposition 5.3.

Given a measure, $d\mu$, on S^1 that is symmetric with respect to complex conjugation, one may define a measure on $[-2, 2]$ by

$$\int_{S^1} f(z + z^{-1})\, d\mu(z) = \int_{-2}^{2} f(x)\, d\nu(x). \tag{2.7}$$

In particular, if $d\mu$ is of the form (2.6), then we find

$$\int f d\nu = \sum f(x_j)\mu_j \qquad \text{where} \qquad x_j = 2\cos(\theta_j).$$

In this way, we find that Proposition 5.3 relates an ensemble of probability measures on $[-2, 2]$ to a certain ensemble of Verblunsky coefficients. In fact, the locations of the masses of $d\nu$ are distributed as the points in the Jacobi ensemble, (1.9), and are independent of the masses.

Theorem 2 follows immediately from the fact that the matrix J represents $f(x) \mapsto xf(x)$ in $L^2(d\nu)$ with respect to the basis of orthonormal polynomials. The remainder of this section is devoted to explaining the origin of this fact.

Let $P_k(x)$ denote the monic polynomials orthogonal with respect to $d\nu$ and $p_k(x)$, the corresponding orthonormal polynomials. These obey a three-term recurrence relation [16, §3.2]:

$$xp_k(x) = a_{k+1}p_{k+1}(x) + b_{k+1}p_k(x) + a_k p_{k-1}(x), \tag{2.8}$$

which explains the structure of the matrix J.

It is a famous observation of Szegő (see [16, §11.5]) that the polynomials orthogonal with respect to $d\mu$ are intimately related to those orthogonal with respect to $d\nu$. Specifically,

$$P_k(z + z^{-1}) = \frac{z^{-k}\Phi_{2k}(z) + z^k\Phi_{2k}(z^{-1})}{1 - \alpha_{2k-1}} \tag{2.9}$$

or equivalently,

$$p_k(z+z^{-1}) = \frac{z^{-k}\phi_{2k}(z) + z^k\phi_{2k}(z^{-1})}{\sqrt{2(1-\alpha_{2k-1})}}.$$

Attendant to this relation between the orthogonal polynomials is a relation between their recurrence coefficients:

$$\begin{aligned} b_{k+1} &= (1-\alpha_{2k-1})\alpha_{2k} - (1+\alpha_{2k-1})\alpha_{2k-2} \\ a_{k+1} &= \big\{(1-\alpha_{2k-1})(1-\alpha_{2k}^2)(1+\alpha_{2k+1})\big\}^{1/2}. \end{aligned}$$

These equations are known as the Geronimus relations. They are exactly equations (1.12) and (1.13) from the statement of Theorem 2. This shows that the matrix J defined in that theorem really does represent multiplication by x in $L^2(d\nu)$.

The connection between the orthogonal polynomials for $d\mu$ and $d\nu$ suggests a relation between the matrices LM of Theorem 1 and J of Theorem 2. This is investigated in Appendix B; in particular we are able to give a short proof of the Geronimus relations.

3. Distribution of $d\mu$ for $\mathbb{U}(n)$ and $\mathbb{SO}(2n)$

Let e_1 denote the standard unit vector $(1,0,\dots,0)^T$. As described in the Introduction, for each $U \in \mathbb{U}(n)$ we consider the spectral measure associated to the pair (U, e_1), that is, the unique measure on $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ that obeys

$$\langle e_1 | U^n e_1 \rangle = \int z^n \, d\mu(z)$$

for all $n \in \mathbb{Z}$. The first goal of this section is to determine the probability distribution of $d\mu$ when $U \in \mathbb{U}(n)$ is chosen according to Haar measure. We will then prove the analogous result for $\mathbb{SO}(2n)$. In each case, we give the distribution both in terms of the natural parameters (the eigenvalues and associated masses) and in terms of the Verblunsky parameters.

Proposition 3.1. *If U is chosen according to Haar measure on $\mathbb{U}(n)$, then the probability measure on $d\mu$ is given by*

$$\frac{(n-1)!}{n!(2\pi)^n}\big|\Delta(e^{i\theta_1},\dots,e^{i\theta_n})\big|^2 d\theta_1 \cdots d\theta_n \, d\mu_1 \dots d\mu_{n-1} \tag{3.1}$$

where $0 \le \theta_j \le 2\pi$ and $0 \le \mu_j \le 1$ with $\sum \mu_j \le 1$.

Proof. Conjugation invariance of Haar measure on $\mathbb{U}(n)$ implies that the distribution of the eigenvectors is independent of that of the eigenvalues. The distribution of the eigenvalues is given by the Weyl integration formula (1.1), while the masses are simply the square moduli of the top entries of the normalized eigenvectors.

Of course, the normalized eigenvectors are only determined up to a phase factor. If we choose this phase factor at random from the unit circle, then by conjugation invariance, the top entries of the eigenvectors are distributed as a random unit vector from $\mathbb{C}^n$. Therefore, by Corollary A.3, $(\mu_1,\dots,\mu_n)$ is uniformly distributed on the $(n-1)$-simplex $\sum \mu_j = 1$. □

The key to writing the probability measure on $d\mu$ in the Verblunsky coordinates is the Householder algorithm [12, §6.4], which converts any matrix to one in Hessenberg form (i.e., with zeros below the sub-diagonal) via unitary conjugation. The algorithm proceeds iteratively, killing the undesirable elements in each column in turn, working left to right.

Consider a matrix A whose first $k-1$ columns are zero below the sub-diagonal. Let us write $A_k = [a_{1,k}, \ldots, a_{n,k}]^T$ for the kth column of A and then define

$$v = [0, \ldots, 0, \alpha, a_{k+2,k}, \ldots, a_{n,k}]^T \tag{3.2}$$

$$\alpha = a_{k+1,k} - \frac{a_{k+1,k}}{|a_{k+1,k}|} \sqrt{|a_{k+1,k}|^2 + \cdots + |a_{n,k}|^2}. \tag{3.3}$$

The reflection through the plane perpendicular to v is given by $R = I - 2\frac{vv^\dagger}{\|v\|^2}$. It maps A_k to $A_k - v$, which has zeros in the desired places. Moreover, the first $k-1$ columns of A are unchanged by left multiplication by R because they are orthogonal to v—indeed, their only non-zero entries coincide with zeros in v. Similarly, right multiplication of any matrix by R leaves its first k columns unchanged.

As R is a reflection, $R^{-1} = R = R^\dagger$ and so we find that conjugating A by R gives a new matrix whose first k columns agree with Hessenberg form (i.e., with zeros below the sub-diagonal). In this way, we have described one step of the usual Householder algorithm. However, we wish to add one further conjugation so as to make the entries on the sub-diagonal non-negative. To do this, we form $DRARD^\dagger$ where D differs from the identity matrix by having $(k+1, k+1)$ entry $e^{-i\phi}$ with ϕ chosen appropriately.

Starting with a unitary matrix U, we can apply the above algorithm to obtain a unitarily equivalent matrix, H, in Hessenberg form with non-negative sub-diagonal. Moreover, the spectral measure for (H, e_1) is the same as that for (U, e_1)—namely $d\mu$—because the vector e_1 is fixed by all the unitary matrices by which U is conjugated.

In the Introduction, we used H to denote the matrix representation of $f(z) \mapsto zf(z)$ on $L^2(d\mu)$ in the basis of orthonormal polynomials. It is easily seen to be in Hessenberg form and to have a non-negative (indeed positive) sub-diagonal. As one might hope from the notation, these two matrices are the same (see also [4, Corollary 3.3]):

Lemma 3.2. *Suppose H is an $n \times n$ unitary matrix in Hessenberg form with non-negative sub-diagonal and let $d\mu$ denote the spectral measure associated to the vector e_1. If the support of $d\mu$ consists of n points, then H represents $f(z) \mapsto zf(z)$ in the basis of orthonormal polynomials. Consequently,*

$$H_{i+1,j+1} = \langle \phi_i | z\phi_j \rangle = \begin{cases} -\alpha_{i-1}\bar{\alpha}_j \prod_{l=i}^{j-1} \rho_l & i < j+1 \\ \rho_{j-1} & i = j+1 \\ 0 & i > j+1 \end{cases} \tag{3.4}$$

where $\rho_j = \sqrt{1 - |\alpha_j|^2}$ and $\alpha_{-1} = -1$.

Proof. As $d\mu$ is the spectral measure for (H, e_1) and as $L^2(d\mu)$ has the same dimension as the space on which the operator H acts, there must be an orthonormal basis $f_1, \ldots, f_n$ for $L^2(d\mu)$ with $f_0 \equiv 1$ such that H represents $f(z) \mapsto zf(z)$ in this basis. This is just the spectral theorem combined with the fact that e_1 must be cyclic (for otherwise, $L^2(d\mu)$ wouldn't have full dimension).

From the cyclicity argument we also learn that H must have a strictly positive sub-diagonal.

To finish the proof of the first claim, we need only show that $f_j(z) = \phi_{j-1}(z)$; that is, that the orthonormal basis in question is precisely that of the orthonormal polynomials. Because H is in Hessenberg form with positive sub-diagonal, the standard basis vectors arise from applying the Gram–Schmidt procedure to $e_1, He_1, \ldots, H^{n-1}e_1$. Consequently, the vectors f_j must be the result of applying the same procedure to $1, z, \ldots, z^{n-1}$; that is, f_j must be ϕ_{j-1}.

The first part of (3.4) merely reexpresses what we have just proved. The second follows from the recursion relations and (2.4); however, the proof is not particularly enlightening and we refer the reader to [8] for details. □

We will now apply the Householder algorithm outlined earlier to a matrix chosen at random from $\mathbb{U}(n)$. By the lemma above, this will allow us to determine the induced distribution on the Verblunsky parameters associated to the spectral measure for (U, e_1).

Proposition 3.3. *Let $d\mu$ be the spectral measure for (U, e_1) with U chosen at random from $\mathbb{U}(n)$ according to Haar measure. In terms of the Verblunsky parameters, $\alpha_0, \ldots, \alpha_{n-2}$, and $\alpha_{n-1} = e^{i\phi}$, this probability distribution is given by*

$$\tfrac{1}{2}\tfrac{(n-1)!}{\pi^n} \prod_{k=0}^{n-2} (1 - |\alpha_k|^2)^{n-2-k}\, d^2\alpha_0 \cdots d^2\alpha_{n-2}\, d\phi. \tag{3.5}$$

That is, the Verblunsky parameters are independent and $\alpha_j \sim \Theta_{2n-2j-1}$.

Proof. The key to applying the Householder algorithm to a random element $U \in \mathbb{U}(n)$ is the following realization of Haar measure: Choose the first column at random from the unit sphere; then choose the second column from the unit sphere of vectors orthogonal to the first; then the third column and so forth. In this way, one could say that the columns of U form a random orthonormal basis for $\mathbb{C}^n$. (That this is indeed Haar measure is a simple consequence of invariance under left multiplication by unitary matrices.)

The first column of U is a random vector from the unit sphere. After left multiplication by the appropriate reflection R, the new first column takes the form $[\bar{\alpha}_0, b, 0, \ldots, 0]^T$ where $\bar{\alpha}_0$ is the the original $(1,1)$ entry of U and so Θ_{2n-1}-distributed, while b has modulus ρ_0 and arbitrary argument. Subsequent left multiplication by D converts the first column to $[\bar{\alpha}_0, \rho_0, 0, \ldots, 0]^T$, as it will remain. The other columns are still orthogonal to the first; indeed, they form a random orthonormal basis for the orthogonal complement of the first column.

Right multiplication by $RD^\dagger$ leaves the first column untouched while orthogonally intermixing the other columns. Of course, this means that they remain a random orthonormal basis for the orthogonal complement of the first column. (Remember, Haar measure is also invariant under right multiplication by a unitary.)

For the subsequent columns, the procedure is similar. Let us skip ahead to dealing with the kth column.

From the unitarity of the matrix H from the previous lemma,

$$X = \begin{bmatrix} \rho_0\rho_1\rho_2\cdots\rho_{k-2} \\ -\alpha_0\rho_1\rho_2\cdots\rho_{k-2} \\ -\alpha_1\rho_2\cdots\rho_{k-2} \\ \vdots \\ -\alpha_{k-3}\rho_{k-2} \\ -\alpha_{k-2} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

is a unit vector orthogonal to the first $k-1$ columns. As the kth column is a random vector orthogonal to the first $k-1$ columns, its inner product with X is distributed as the top entry of a random vector from the $(2n-2k+1)$-sphere and is independent of $\alpha_1,\dots,\alpha_{k-2}$. Let us call this inner product $\bar\alpha_{k-1}$, noting that this implies α_{k-1} is $\Theta_{2n-2k+1}$-distributed as stated in the proposition.

We now multiply the matrix at hand from the left by the appropriate reflection and rotation to bring the kth column into the desired form. Neither of these operations alters the top k rows and so the inner product of the kth column with X is unchanged. But now the kth column is uniquely determined; it must be $\bar\alpha_{k-1}X + \rho_{k-1}e_{k+1}$, just as in (3.4).

Lastly, we should multiply on the right by $RD^\dagger$, but this leaves the first k columns unchanged while orthogonally intermixing the other columns. In this way, we obtain a matrix whose first k columns conform to the structure of H, while the remaining columns form a random basis for the orthogonal complement of the span of those k columns.

In this way, we can proceed inductively until we reach the last column. It is obliged to be a random orthonormal basis for the one-dimensional space orthogonal to the preceding $n-1$ columns and hence a random unimodular multiple, say $\bar\alpha_{n-1}$, of X. This is why the last Verblunsky parameter is Θ_1-distributed.

We have now conjugated U to a matrix in the form of (3.4) with parameters distributed as stated in the proposition. The vector e_1 is unchanged under the action of each of the conjugating matrices and consequently, these are precisely the Verblunsky parameters of $d\mu$. □

We now turn to the study of Haar measure on $\mathbb{SO}(2n)$. The proofs follow those given above pretty closely.

Proposition 3.4. *If U is chosen at random from $\mathbb{SO}(2n)$ according to Haar measure, then the spectral measure $d\mu$ associated to (U, e_1) is distributed as*

$$\frac{(n-1)!}{2^{n-1}\,n!}\big|\Delta\big(2\cos\theta_1,\dots,2\cos\theta_n\big)\big|^2\frac{d\theta_1}{\pi}\cdots\frac{d\theta_n}{\pi}d\mu_1\cdots d\mu_{n-1} \tag{3.6}$$

where θ_j and μ_j are the coordinates given in (2.6).

Proof. By the Weyl integration formula for $\mathbb{SO}(2n)$, the marginal distribution of the eigenvalues is as above. (See [17, §VII.9].)

If the eigenvalues are prescribed, say $e^{\pm i\theta_1},\dots,e^{\pm i\theta_n}$, then the conditional distribution of U is given by taking a fixed matrix with this spectrum and conjugating it by a random element from $\mathbb{SO}(2n)$. The natural choice for this fixed matrix is block diagonal:

$$U_0 = \operatorname{diag}\left(\begin{bmatrix} \cos(\theta_1) & \sin(\theta_1) \\ -\sin(\theta_1) & \cos(\theta_1) \end{bmatrix}, \ldots, \begin{bmatrix} \cos(\theta_n) & \sin(\theta_n) \\ -\sin(\theta_n) & \cos(\theta_n) \end{bmatrix} \right).$$

From this we see that the $d\mu$ is the spectral measure for U_0 and a random vector from the $(2n-1)$-sphere. The proposition then follows by diagonalizing U_0 and applying Corollary A.3. $\square$

Proposition 3.5. *Let U be chosen from $\mathbb{SO}(2n)$ according to Haar measure and let $d\mu$ denote the spectral measure for (U, e_1). In terms of the Verblunsky parameters, the probability distribution on $d\mu$ is*

$$\tfrac{(n-1)!}{\pi^n} \prod_{k=0}^{2n-2} (1-\alpha_k^2)^{\frac{2n-k-3}{2}} \, d\alpha_0 \cdots d\alpha_{2n-2} \tag{3.7}$$

with $\alpha_{2n-1} = -1$. That is, the Verblunsky parameters are independent and $\alpha_k \sim B(\frac{2n-k-1}{2}, \frac{2n-k-1}{2})$.

Proof. While we may use the Householder algorithm as set out above, the fact that we are now dealing with real-valued matrices allows the following simplification: the vector defining the reflection is as in (3.2), but now with

$$\alpha = a_{k+1,k} - \sqrt{a_{k+1,k}^2 + \cdots + a_{2n,k}^2}$$

instead of (3.3). This permits us to forgo the conjugation by D.

Haar measure on $\mathbb{SO}(2n)$ can be realized by choosing the first column as a random vector from the unit sphere in $\mathbb{R}^{2n}$, and then the second as a random vector orthogonal to the first, and so forth. However, the fact that the matrix has determinant one means that the first $2n-1$ columns completely determine the last. One may say that the columns of U form a random positively oriented basis for $\mathbb{R}^{2n}$.

Proceeding as in the proof of Proposition 3.3, we see that α_{k-1} is defined as the inner product of a specific vector X with a random unit vector from the $(2n-k-2)$-sphere of vectors orthogonal to the first $k-1$ columns. It follows from Corollary A.2 that $\alpha_{k-1} \sim B(\frac{2n-k-2}{2}, \frac{2n-k-2}{2})$ as stated above.

The last column of H, and hence α_{2n-1}, is uniquely determined by the fact that $\det(H) = 1$. It is just a matter of using (3.4) to determine which value of α_{2n-1} makes this determinant one; moreover, by continuity of the determinant, it suffices to consider the case where all other αs are zero. This gives

$$1 = \det(H) = -\alpha_{-1}\alpha_{2n-1} \operatorname{sign}(\sigma) = -\alpha_{2n-1}$$

where σ is the cyclic permutation $j \mapsto j+1 \mod 2n$, which is odd.

Lastly, we should justify the normalization coefficient given in (3.7); what appears there is very much simpler than one would expect from (1.10). This simplification is based on the duplication formula for the Γ function: $\sqrt{\pi}\Gamma(2t) = 2^{2t-1}\Gamma(t)\Gamma(t+\frac{1}{2})$. Specifically, beginning with (1.10),

$$\int_{-1}^{1} (1-\alpha^2)^{t-1} \, d\alpha = \frac{2^{1-2t}\Gamma(2t)}{\Gamma(t)^2} = \frac{\Gamma(t+\frac{1}{2})}{\sqrt{\pi}\,\Gamma(t)},$$

which causes the product of normalization coefficients to telescope:

$$\prod_{k=0}^{2n-2} \frac{\Gamma(\frac{2n-k}{2})}{\sqrt{\pi}\,\Gamma(\frac{2n-k-1}{2})} = \pi^{\frac{1}{2}-n}\frac{\Gamma(n)}{\Gamma(\frac{1}{2})} = \frac{(n-1)!}{\pi^n},$$

as given in (3.7). ☐

4. The Proof of Theorem 1

Let $d\mu$ be the measure on S^1 given by

$$\int f\,d\mu = \sum_{j=1}^{n} \mu_j f(e^{i\theta_j}) \tag{4.1}$$

with $\theta_j \in [0, 2\pi)$ distinct and $\sum \mu_j = 1$. As discussed in the Introduction, this measure is uniquely determined by its Verblunsky parameters $\alpha_0, \ldots, \alpha_{n-2} \in \mathbb{D}$ and $\alpha_{n-1} = e^{i\phi} \in S^1$.

It is difficult to find functions of $d\mu$ that admit simple expressions in terms of both θ_j, μ_j and the Verblunsky parameters. One such quantity is the determinant of the associated Toeplitz matrix; this is the subject of the next lemma.

Lemma 4.1. *If $d\mu$ is a probability measure of the form given in* (4.1) *and $\alpha_0, \ldots, \alpha_{n-1}$ its Verblunsky coefficients, then*

$$|\Delta(z_1, \ldots, z_n)|^2 \prod_{j=1}^{n} \mu_j = \prod_{k=0}^{n-2} (1 - |\alpha_k|^2)^{n-k-1}. \tag{4.2}$$

Proof. Let $c_k = \sum_{j=1}^{n} \mu_j z_j^k$ denote the moments of $d\mu$. We will prove that both sides of (4.2) are equal to the determinant of the $n \times n$ Toeplitz matrix associated to $d\mu$:

$$T = \begin{bmatrix} c_0 & c_{-1} & \cdots & c_{1-n} \\ c_1 & c_0 & \cdots & c_{2-n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & \cdots & c_0 \end{bmatrix}.$$

If we define A and M by

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_n \\ \vdots & \vdots & \ddots & \vdots \\ z_1^{n-1} & z_2^{n-1} & \cdots & z_n^{n-1} \end{bmatrix} \qquad M = \begin{bmatrix} \mu_1 & 0 & \cdots & 0 \\ 0 & \mu_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_n \end{bmatrix},$$

then $T = AMA^\dagger$. Consequently,

$$\det(T) = \left|\det(A)\right|^2 \det(M) = |\Delta(z_1, \ldots, z_n)|^2 \prod_{j=1}^{n} \mu_j.$$

We will now show that the right-hand side of (4.2) is equal to $\det(T)$. To this end, let

$$B = \begin{bmatrix} \Phi_0(z_1) & \Phi_0(z_2) & \cdots & \Phi_0(z_n) \\ \Phi_1(z_1) & \Phi_1(z_2) & \cdots & \Phi_1(z_n) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{n-1}(z_1) & \Phi_{n-1}(z_2) & \cdots & \Phi_{n-1}(z_n) \end{bmatrix},$$

which has the same determinant as A because each can be reduced to the other by elementary row operations. From the orthogonality property of the $\{\Phi_j\}$, it follows that $BMB^\dagger$ is the diagonal matrix whose entries are the squares of the $L^2(d\mu)$-norms of $\Phi_0, \Phi_1, \ldots, \Phi_{n-1}$. Therefore by (2.4),

$$\det(T) = \det(BMB^\dagger) = \prod_{k=0}^{n-1} \|\Phi_k\|^2_{L^2(d\mu)} = \prod_{k=0}^{n-2} (1 - |\alpha_k|^2)^{n-k-1},$$

just as was required. □

Both expressions for the Toeplitz determinant are well known; indeed, the argument presented above and its Hankel-matrix analog play a central role in random matrix theory.

We are now ready to state and prove the main result of this section. Please note that neither measure given below is normalized; however, they do have the same normalization coefficient. It is calculated in Lemma 4.4 where it is used to give an independent proof of (1.6).

Proposition 4.2. *The following formulae express the same measure on the manifold of probability distributions on S^1 supported at n points:*

$$\frac{2^{1-n}}{n!} |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^\beta \prod_{j=1}^{n} \mu_j^{\frac{\beta}{2}-1}\, d\theta_1 \cdots d\theta_n \, d\mu_1 \cdots d\mu_{n-1} \tag{4.3}$$

in the (θ, μ)-coordinates of (4.1)*, and*

$$\prod_{k=0}^{n-2} (1 - |\alpha_k|^2)^{\frac{\beta}{2}(n-k-1)-1}\, d^2\alpha_0 \cdots d^2\alpha_{n-2}\, d\phi \tag{4.4}$$

in terms of the Verblunsky parameters.

Proof. When $\beta = 2$, this follows immediately from Propositions 3.1 and 3.3. To obtain the general-β version of (4.3) from the $\beta = 2$ version, one has to multiply by

$$|\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^{\beta-2} \prod_{j=1}^{n} \mu_j^{\frac{\beta}{2}-1}, \tag{4.5}$$

while the same transformation of β in (4.4) is effected by multiplying by

$$\prod_{k=0}^{n-2} (1 - |\alpha_k|^2)^{(\frac{\beta}{2}-1)(n-k-1)}. \tag{4.6}$$

But, (4.5) and (4.6) are equal; they are either side of (4.2) raised to the power $\frac{\beta}{2} - 1$. □

Proof of Theorem 1. Theorem 1 is an immediate corollary of Proposition 4.2 and results from [3]: The Verblunsky parameters of the spectral measure for (LM, e_1) are precisely the αs that appear in the definition of L and M. Consequently, if the Verblunsky parameters are distributed according to (4.4), then the eigenvalues are distributed as in (1.5). □

It is fair to suggest that studying the unitary group is a rather roundabout way of proving the above proposition. We simply do not know a better way. The natural suggestion is to first calculate the Jacobian for the map from the (θ, μ)-coordinates to the Verblunsky parameters and to proceed from there. While we can determine this Jacobian *a posteriori* by employing a cunning idea of Dumitriu and Edelman, we do not have a direct derivation.

The idea of Dumitriu and Edelman can be summarized as follows:

Lemma 4.3. *Suppose $\phi : \mathcal{O}_1 \to \mathcal{O}_2$ is an N-fold cover of $\mathcal{O}_2$ by $\mathcal{O}_1$, both of which are open subsets of $\mathbb{R}^n$. If the measure $f(x)\, d^n x$ is the symmetric pull-back of the measure $g(y)\, d^n y$, with both f and g positive, then the Jacobian of ϕ is given by*

$$|\phi'(x)| = \frac{N f(x)}{g \circ \phi\,(x)}$$

for any $x \in \mathcal{O}_1$.

Proof. For every $x \in \mathcal{O}_1$,

$$f(x)\, d^n x = \phi^*\big(\tfrac{1}{N} g(y)\, d^n y\big) = \tfrac{1}{N}[g \circ \phi](x)|\phi'(x)|\, d^n x,$$

which proves the lemma. □

Proposition 4.2 allows us to apply this lemma to the current situation. The map of θ, μ to the Verblunsky parameters is an $n!$-fold cover and so we obtain

$$\left|\frac{\partial(\alpha, \phi)}{\partial(\theta, \mu)}\right| = 2^{1-n} \frac{|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^2}{\prod_{k=0}^{n-2}(1 - |\alpha_k|^2)^{n-k-2}} \tag{4.7}$$

$$= 2^{1-n} \frac{\prod_{k=0}^{n-2}(1 - |\alpha_k|^2)}{\prod_{j=0}^{n} \mu_j} \tag{4.8}$$

where one should regard the Verblunsky parameters as functions of the θs and μs. The formulae above correspond to applying Proposition 4.2 with $\beta = 2$ and $\beta = 0$, respectively. Of course, one can also use other values of β, but the resulting formulae are related to one another by Lemma 4.1.

Earlier we promised to determine the (common) normalization coefficient for the measures (4.3) and (4.4). We also promised to give a new derivation of the partition function, (1.6), for the Coulomb gas. We will now settle these obligations.

Lemma 4.4. *The integral of* (4.4) *is*

$$\int \!\cdots\! \int \prod_{k=0}^{n-2} (1 - |\alpha_k|^2)^{\frac{\beta}{2}(n-k-1)-1}\, d^2\alpha_0 \cdots d^2\alpha_{n-2}\, d\phi = \frac{(2\pi)^n}{\beta^{n-1}(n-1)!} \tag{4.9}$$

while

$$\int \!\cdots\! \int \big|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})\big|^\beta \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_n}{2\pi} = \frac{\Gamma(\frac{1}{2}\beta n + 1)}{\big[\Gamma(\frac{1}{2}\beta + 1)\big]^n}, \tag{4.10}$$

in agreement with (1.5) *and* (1.6).

Proof. Each of the integrals in (4.9) is rendered trivial by switching to polar coordinates:

$$\int (1 - |z|^2)^{(t/2)-1}\, d^2 z = 2\pi t^{-1}.$$

It is this integral that gives the normalization coefficient in the definition of the Θ distributions; cf. (1.7).

The proof of (4.10) begins with the evaluation of the Dirichlet integral

$$\int_{\triangle} \prod_{j=1}^{n} \mu_j^{\frac{\beta}{2}-1} \, d\mu_1 \cdots d\mu_{n-1} = \frac{\Gamma(\frac{\beta}{2})^n}{\Gamma(\frac{n\beta}{2})},$$

which is derived in the proof of Lemma A.4; see (A.4). As the two measures in Proposition 4.2 have the same total mass, the integral of (4.3) is given by (4.9). Combining this with the Dirichlet integral above leads us to

$$\text{LHS (4.10)} = \frac{n!}{2^{1-n}} \frac{1}{\beta^{n-1}(n-1)!} \frac{\Gamma(\frac{n\beta}{2})}{\Gamma(\frac{\beta}{2})^n} = \frac{\Gamma(\frac{n\beta}{2}+1)}{\Gamma(\frac{\beta}{2}+1)^n},$$

which is exactly (1.6). □

5. The Proof of Theorem 2

As explained in Section 2, Theorem 2 is an immediate corollary of Proposition 5.3 and the Geronimus relations. As a result, the primary purpose of this section is to prove this proposition.

Throughout this section, $d\mu$ will denote a probability measure of the form given in (2.6). In particular, it is symmetric with respect to complex conjugation and the last Verblunsky parameter, α_{2n-1}, is equal to -1. We will also use the notation $x_j = 2\cos(\theta_j)$ repeatedly.

In addition to Lemma 4.1 from the previous section, two further lemmas are required. They are the following:

Lemma 5.1. *If $x_j = 2\cos(\theta_j)$, $1 \le j \le n$, then*

$$|\Delta(e^{\pm i\theta_1}, e^{\pm i\theta_2}, \ldots, e^{\pm i\theta_n})| = |\Delta(x_1, x_2, \ldots, x_n)|^2 \prod_{l=1}^{n} |2\sin(\theta_l)| \tag{5.1}$$

where the left-hand side is shorthand for the Vandermonde of the $2n$ quantities $e^{i\theta_1}, e^{-i\theta_1}, \ldots, e^{i\theta_n}, e^{-i\theta_n}$.

Proof. By expanding the Vandermonde as in (1.2),

$$\begin{aligned} &|\Delta(e^{\pm i\theta_1}, e^{\pm i\theta_2}, \ldots, e^{\pm i\theta_n})| \\ &= \prod_{l=1}^{n} |e^{i\theta_l} - e^{-i\theta_l}| \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^2 |e^{i\theta_j} - e^{-i\theta_k}|^2 \\ &= \prod_{l=1}^{n} |2\sin(\theta_l)| \prod_{j<k} \big[2\cos(\theta_j) - 2\cos(\theta_k)\big]^2, \end{aligned}$$

as required. In the last step we used that

$$|(z-w)(z-\bar{w})| = |(z-w)(1-\bar{z}\bar{w})| = |(z+\bar{z}) - (w+\bar{w})|$$

for any pair of points z, w on the unit circle. □

Lemma 5.2. *Let $d\mu$ be a measure on the unit circle of the form given in* (2.6) *and let Φ_k denote the corresponding monic orthogonal polynomials. Then*

$$\prod_j (2-x_j) = \Phi_{2n}(1) = \prod_{k=0}^{2n-1}(1-\alpha_k) = 2\prod_{k=0}^{2n-2}(1-\alpha_k) \tag{5.2}$$

$$\prod_j (2+x_j) = \Phi_{2n}(-1) = \prod_{k=0}^{2n-1}\big(1+(-1)^k\alpha_k\big) = 2\prod_{k=0}^{2n-2}\big(1+(-1)^k\alpha_k\big) \tag{5.3}$$

where $x_j = 2\cos(\theta_j)$ and α_k denote the Verblunsky parameters of $d\mu$.

Proof. As Φ_{2n} is orthogonal to each of $1, z, \dots, z^{2n-1}$, which forms a basis for $L^2(d\mu)$, its zeros must be $e^{\pm i\theta_1}, \dots, e^{\pm i\theta_n}$. Consequently, for any $x \in \mathbb{R}$,

$$\Phi_{2n}(x) = \prod_j |x - e^{i\theta_j}|^2 = \prod_j \big[x^2 - 2x\cos(\theta_j) + 1\big],$$

which gives the first equality in each of (5.2) and (5.3).

Because all Verblunsky parameters are real, the coefficients of the orthogonal polynomials are also real. This implies $\Phi_k^*(z) = z^k\Phi_k(z^{-1})$ and so, for $z = \pm 1$, the recurrence equation (2.1) becomes $\Phi_{k+1}(z) = (z - \alpha_k z^k)\Phi_k(z)$. Each of the second equalities stated in the lemma now follows by the obvious induction. The third equalities simply express $\alpha_{2n-1} = -1$. ☐

The following proposition is the $\mathbb{SO}(2n)$ analogue of Proposition 4.2 and so the main ingredient of the proof of Theorem 2.

Proposition 5.3. *Consider the following measure on $[-2,2]^n \times \triangle$:*

$$\frac{2^{-\kappa}}{n!}|\Delta(x_1,\dots,x_n)|^\beta \prod_{j=1}^n \mu_j^{\frac{\beta}{2}-1} \prod_{j=1}^n \big[(2-x_j)^a(2+x_j)^b\big]\, dx_1\cdots dx_n\, d\mu_1\cdots d\mu_{n-1}$$

where $\kappa = (n-1)\frac{\beta}{2}+a+b+1$. Under (2.6) *and the change of variables $x_j = 2\cos\theta_j$, this gives a measure on $d\mu$ (which is not normalized). Transferring this measure to the Verbunsky parameters gives*

$$\prod_{k=0}^{2n-2}(1-\alpha_k^2)^{\frac{\beta(2n-k-1)}{4}-1}\prod_{k=0}^{2n-2}\big(1-\alpha_k\big)^{a+1-\frac{\beta}{4}}\big(1+(-1)^k\alpha_k\big)^{b+1-\frac{\beta}{4}}\, d\alpha_0\cdots d\alpha_{2n-2}$$

and $\alpha_{2n-1} \equiv 1$. After normalization, this measure corresponds to choosing the Verbunsky parameters independently with distribution given by

$$\alpha_k \sim \begin{cases} B(\frac{2n-k-2}{4}\beta + a + 1, \frac{2n-k-2}{4}\beta + b + 1) & k \text{ even,} \\ B(\frac{2n-k-3}{4}\beta + a + b + 2, \frac{2n-k-1}{4}\beta) & k \text{ odd.}\end{cases} \tag{5.4}$$

The definition of $B(s,t)$ is given in (1.10).

Proof. From Lemma 5.1 and Lemma 4.1, we may deduce that

$$\begin{aligned} |\Delta(x_1,x_2,\dots,x_n)|^2\prod_{l=1}^n|2\sin(\theta_l)| &= |\Delta(e^{\pm i\theta_1}, e^{\pm i\theta_2},\dots,e^{\pm i\theta_n})| \\ &= 2^n\prod_{k=0}^{2n-2}(1-\alpha_k^2)^{(2n-k-1)/2}\prod_{j=1}^n\mu_j^{-1}. \end{aligned} \tag{5.5}$$

By forming the square-root of the product of (5.2) and (5.3), we can rewrite the product of $|2\sin\theta_l|$ in terms of the Verblunsky parameters:

$$\prod_{l=1}^{n} |2\sin(\theta_l)| = 2\prod_{k=0}^{2n-2} \big(1-\alpha_k\big)^{\frac{1}{2}} \big(1+(-1)^k\alpha_k\big)^{\frac{1}{2}}. \tag{5.6}$$

Substituting this into (5.5) above and doing a little rearranging of terms leads us to

$$2^{1-n}|\Delta(x_1,x_2,\dots,x_n)|^2 \prod_{j=1}^{n}\mu_j = \frac{\prod_{k=0}^{2n-2}\big(1-\alpha_k^2\big)^{\frac{2n-k-1}{2}}}{\prod_{k=0}^{2n-2}\big(1-\alpha_k\big)^{\frac{1}{2}}\big(1+(-1)^k\alpha_k\big)^{\frac{1}{2}}}. \tag{5.7}$$

We will return to this equation in a moment.

For $\beta = 2$ and $a = b = -\frac{1}{2}$, the proposition is an immediate corollary of Propositions 3.4 and 3.5 together with (5.6). The latter arises as the Jacobian of the change of variables from θ_j to $x_j = 2\cos(\theta_j)$.

Changing a and b amounts to multiplying this result by the appropriate powers of (5.2) and (5.3), respectively. To see that the two measures are equivalent for $\beta \neq 2$, it suffices to multiply by (5.7) raised to the $\frac{\beta}{2} - 1$ power. □

Combining this proposition with Lemma 4.3 permits us to determine the Jacobian of the map from the (θ,μ)-coordinates to the Verblunsky parameters. From $\beta = 2$, $a = b = -\frac{1}{2}$ we obtain

$$\left|\frac{\partial(\alpha)}{\partial(\theta,\mu)}\right| = \frac{2^{1-n}|\Delta(x_1,\dots,x_n)|^2}{\prod_{k=0}^{2n-2}(1-\alpha_k^2)^{\frac{2n-k-3}{2}}}. \tag{5.8}$$

We do not have a direct derivation of this fact.

6. The Selberg and Aomoto Integrals

In [14], Selberg evaluated the following integral:

$$\int_0^1\!\cdots\!\int_0^1 |\Delta(u_1,\dots,u_n)|^{2z}\prod_{j=1}^{n} u_j^{x-1}(1-u_j)^{y-1}\,du_1\cdots du_n, \tag{6.1}$$

which subsequently turned out to be important in random matrix theory. We will present a new derivation of his result in a manner analogous to the proof of Lemma 4.4 above.

We begin with the evaluation of the (common) integral of the measures given in Proposition 5.3. In the case of the second measure, this gives rise to a product of beta integrals (cf. (1.10)) from which we obtain the answer

$$\begin{aligned}&\prod_{\substack{k=0\\k\text{ even}}}^{2n-2}\frac{\Gamma(\frac{2n-k-2}{4}\beta+a+1)\Gamma(\frac{2n-k-2}{4}\beta+b+1)}{\Gamma(\frac{2n-k-2}{2}\beta+a+b+2)}2^{(2n-k-2)\frac{\beta}{2}+a+b+1}\\ &\times\prod_{\substack{k=1\\k\text{ odd}}}^{2n-3}\frac{\Gamma(\frac{2n-k-3}{4}\beta+a+b+2)\Gamma(\frac{2n-k-1}{4}\beta)}{\Gamma(\frac{2n-k-2}{2}\beta+a+b+2)}2^{(2n-k-2)\frac{\beta}{2}+a+b+1}.\end{aligned} \tag{6.2}$$

To aid in the eventual comparison with Selberg [14], we will switch to his parameters:

$$x = a+1,\quad y = b+1,\quad\text{and}\quad z = \tfrac{1}{2}\beta.$$

We also wish to make the following simplifications: The products of the powers of 2 can be combined since they are the same for odd and even k, where we can easily sum the resulting arithmetic progression in the exponents. Similarly, we combine the products of the denominators into a single product and make the substitution $p = 2n-k-2$. In the even-k numerators, we will make the substitution $r = \frac{1}{2}(2n-k-2)$ and in the odd-k numerators, the substitution $s = \frac{1}{2}(2n-k-3)$. Combining these we find that (6.2) is equal to

$$2^{\sigma} \times \frac{\prod_{r=0}^{n-1} \Gamma(rz+x)\Gamma(rz+y) \ \prod_{s=0}^{n-2} \Gamma(sz+x+y)\Gamma((s+1)z)}{\prod_{p=0}^{2n-2} \Gamma(pz+x+y)} \tag{6.3}$$

where $\sigma = [(n-1)z+x+y-1](2n-1)$. Note that there is a cancellation between the $\Gamma(sz+x+y)$ terms in the numerator and the corresponding terms in the denominator.

This essentially completes the determination of the total mass of the measures in Proposition 5.3. After the transformation $x_j = 4u_j - 2$, the former of these is the tensor product of the measure in (6.1) with a measure in the μ_j coordinates. This leads us to evaluate

$$\frac{2^{-\kappa}}{n!} \int \cdots \int \prod_{j=1}^{n} \mu_j^{z-1} \, d\mu_1 \cdots d\mu_{n-1} = \frac{2^{-\kappa}}{n!} \frac{\Gamma(z)^n}{\Gamma(nz)} = \frac{2^{-\kappa} z^{1-n} \Gamma(z+1)^n}{(n-1)! \, \Gamma(nz+1)} \tag{6.4}$$

where $\kappa = (n-1)z+x+y-1$. Moreover, taking the ratio (6.3)/(6.4) and making the cancellation mentioned above, we obtain

$$\begin{aligned} &\int \cdots \int |\Delta(x_1, \ldots, x_n)|^{2z} \prod_{j=1}^{n} \big[(2-x_j)^{x-1}(2+x_j)^{y-1}\big] \, dx_1 \cdots dx_n \\ &\quad = 2^{\sigma+\kappa} \times \frac{\prod_{r=0}^{n-1} \Gamma(rz+x)\Gamma(rz+y) \ \prod_{s=0}^{n-2} (s+1)z\Gamma((s+1)z)}{\Gamma(z+1)^n \prod_{p=n-1}^{2n-2} \Gamma(pz+x+y)} \Gamma(nz+1). \end{aligned} \tag{6.5}$$

Notice that the term $z^{n-1}(n-1)!$ from (6.4) was split up inside the product over s. To simplify, we use $\xi\Gamma(\xi) = \Gamma(\xi+1)$ inside the product over s and notice that the final factor, $\Gamma(nz+1)$, just corresponds to $s = n-1$. In this way, all products run over the same number of terms and can be combined. Therefore, we reach the final conclusion

$$\text{LHS (6.5)} = 2^{\tau} \prod_{r=0}^{n-1} \frac{\Gamma\big(rz+x\big)\Gamma\big(rz+y\big)\Gamma\big((r+1)z\big)}{\Gamma\big(z+1\big)\Gamma\big((n+r-1)z+x+y\big)}$$

where $\tau = \sigma + \kappa = 2n[(n-1)z+x+y-1]$. This is in perfect agreement with Selberg's paper: the values of (6.1) and (6.5) differ by a factor of 2^{τ} as a result of the change of variables $x_j = 4u_j - 2$; the measure is homogeneous of order $\tau/2$.

We now turn to our second topic: the Aomoto integral.

In [2], Aomoto determined the average value of $\prod(x - x_j)$ when the points x_j are distributed according to the Jacobi ensemble, (1.9). Theorem 2 shows that this is equivalent to evaluating the average of the characteristic polynomial for a certain ensemble of Jacobi matrices. In this way, Proposition 6.1 below reproduces Aomoto's result.

The answer is given in terms of the classical Jacobi polynomials: In the notation of [1],

$$\frac{4^n n!}{(a+b+n+1)_n} P_n^{(a,b)}(\tfrac{1}{2}x)$$

are the monic polynomials that are orthogonal with respect to the measure

$$\int f\,d\mu = \frac{\Gamma(a+b+2)}{4^{a+b+1}\Gamma(a+1)\Gamma(b+1)} \int_{-2}^{2} f(x)\,(2-x)^a(2+x)^b\,dx.$$

Here $(z)_n = \Gamma(z+n)/\Gamma(z)$ is the Pochhammer symbol. The recurrence coefficients for the corresponding system of orthonormal polynomials are (cf. (2.8))

$$\begin{aligned} b_{n+1} &= \frac{2(b^2-a^2)}{(2n+a+b)(2n+a+b+2)} \\ a_{n+1}^2 &= \frac{16(n+1)(n+a+b+1)(n+a+1)(n+b+1)}{(2n+a+b+1)(2n+a+b+2)^2(2n+a+b+3)}. \end{aligned} \tag{6.6}$$

Proposition 6.1. *Let J be the Jacobi matrix whose entries $b_1, a_1, \ldots, b_n$ are distributed as in Theorem 2. Then*

$$\mathbb{E}\big[\det(x-J)\big] = \frac{4^n n!}{(\tilde{a}+\tilde{b}+n+1)_n} P_n^{(\tilde{a},\tilde{b})}(\tfrac{1}{2}x) \tag{6.7}$$

where $\tilde{a} = \frac{2(a+1)}{\beta} - 1$ and $\tilde{b} = \frac{2(b+1)}{\beta} - 1$.

Proof. We will show that both sides of (6.7) are related to the same monic orthogonal polynomial on the unit circle. We begin with the left-hand side.

Let $\alpha_0, \ldots, \alpha_{2n-2}$ be distributed as in (1.11), let $d\mu$ denote the corresponding measure on S^1, and let $d\nu$ denote the measure on $[-2,2]$ induced by (B.5).

The characteristic polynomial of J is equal to $P_n(x)$, the nth monic orthogonal polynomial associated to the measure $d\nu$; indeed the principal $k \times k$ minor of $x - J$ is equal to P_k. (By expanding along the last row, one can see that these minor determinants obey the same recurrence relation as the orthogonal polynomials.) Combining this with (2.9) gives

$$\det((z+z^{-1}) - J) = \frac{z^{-n}\Phi_{2n}(z) + z^n\Phi_{2n}(z^{-1})}{2}. \tag{6.8}$$

As Φ_k is independent of α_k (it depends only on $\alpha_0, \ldots, \alpha_{k-1}$), the recurrence relations (2.1) and (2.2) yield

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

Hence, $\mathbb{E}\Phi_k$ are the monic orthogonal polynomials associated to the averaged Verblunsky parameters:

$$\mathbb{E}(\alpha_k) = \begin{cases} \frac{2b-2a}{(2n-k-2)\beta+2a+2b+4} & k \text{ even}, \\ \frac{\beta-2a-2b-4}{(2n-k-2)\beta+2a+2b+4} & k \text{ odd}. \end{cases}$$

(Note that if $X \sim B(s,t)$ then $\mathbb{E}(X) = \frac{t-s}{t+s}$.)

Applying the Geronimus relations to these averaged Verblunsky coefficients does not produce the recursion coefficients associated to the Jacobi polynomials. However, by Proposition B.2, $\mathbb{E}\Phi_{2n}$ is also the monic orthogonal polynomial of degree

$2n$ associated to the 'reversed' coefficients: $\tilde{\alpha}_k = \mathbb{E}\alpha_{2n-2-k}$, $0 \le k \le 2n-2$, and $\tilde{\alpha}_{2n-1} = -1$. Under the Geronimus relations, these coefficients give rise to

$$\tilde{b}_{k+1} = \frac{2(\tilde{b}^2 - \tilde{a}^2)}{(2k+\tilde{a}+\tilde{b})(2k+\tilde{a}+\tilde{b}+2)}$$
$$\tilde{a}_{k+1}^2 = \frac{16(k+1)(k+\tilde{a}+\tilde{b}+1)(k+\tilde{a}+1)(k+\tilde{b}+1)}{(2k+\tilde{a}+\tilde{b}+1)(2k+\tilde{a}+\tilde{b}+2)^2(2k+\tilde{a}+\tilde{b}+3)}$$

where $\tilde{a} = \frac{2(a+1)}{\beta} - 1$ and $\tilde{b} = \frac{2(b+1)}{\beta} - 1$. By comparison with (6.6), we find

$$\tfrac{1}{2}\mathbb{E}\{z^{-n}\Phi_{2n}(z) + z^n\Phi_{2n}(z^{-1})\} = \tfrac{4^n n!}{(\tilde{a}+\tilde{b}+n+1)_n} P_n^{(\tilde{a},\tilde{b})}(\tfrac{1}{2}x).$$

Equation (6.7) now follows from (6.8). □

Appendix A. The Surface Measure on S^n

This Appendix presents some elementary results used in the text and is provided solely for the reader's convenience.

Let $d\tilde{\sigma}$ denote the usual surface measure on S^n and let $d\sigma$ denote the corresponding normalized probability measure on S^n. We will write $\mathbb{D}^n$ for the n-disk: $\mathbb{D}^n = \{x \in \mathbb{R}^n : |x| < 1\}$.

Lemma A.1. *If $f : \mathbb{R}^{n+1} \to \mathbb{C}$, then*

$$\int_{S^n} f(x)\, d\tilde{\sigma}(x) = \sum_{\pm} \int_{\mathbb{D}^n} f\big(x_1, \dots, x_n, \pm\sqrt{1-|x|^2}\big) \frac{d^n x}{\sqrt{1-|x|^2}}.$$

Proof. The map $\phi : (x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, \sqrt{1-|x|^2})$ is a diffeomorphism onto the open upper hemisphere. The corresponding Gram matrix is

$$G_{i,j} = \Big\langle \frac{\partial\phi}{\partial x_i} \Big| \frac{\partial\phi}{\partial x_j} \Big\rangle = \delta_{i,j} + \frac{|x|^2}{1-|x|^2} \frac{x_i}{|x|} \frac{x_j}{|x|},$$

which is a rank one perturbation of the identity matrix. Therefore,

$$\det[G] = 1 + \frac{|x|^2}{1-|x|^2} = \frac{1}{1-|x|^2},$$

from which the lemma follows immediately. □

We wish to determine the distributions of certain probability measures induced from the normalized surface measure on S^n. These follow easily from this lemma.

Corollary A.2. *For the projection of the n-sphere into its first coordinate, we have*

$$\int_{S^n} f(x_1)\, d\sigma(x) = \frac{2^{1-n}\Gamma(n)}{\Gamma(\frac{n}{2})^2} \int_{-1}^{1} f(t)\,(1-t^2)^{(n-2)/2}\, dt, \tag{A.1}$$

that is, x_1 is $B(\frac{n}{2}, \frac{n}{2})$-distributed. Projection onto the first two coordinates gives

$$\int_{S^n} f(x_1 + ix_2)\, d\sigma(x) = \frac{n-1}{2\pi} \int_{\mathbb{D}} f(z)\,(1-|z|^2)^{(n-3)/2}\, d^2 z, \tag{A.2}$$

which implies $x_1 + ix_2$ is Θ_n-distributed.

Proof. Both formulae follow from the more general statement that for $1 \leq k < n$,

$$\int_{S^n} f(x_1, \dots, x_k)\, d\sigma(x) \propto \int_{\mathbb{D}^k} f(x_1, \dots, x_k)(1 - |x|^2)^{(n-k-1)/2}\, dx_1 \cdots dx_k$$

where the proportionality constant depends on k, but not the function f. The value of the normalization constant can then be determined by substituting $f \equiv 1$.

An inductive proof of the more general statement follows easily from

$$\int_{-s}^{s} \big(s^2 - x_{k+1}^2\big)^{(n-k-2)/2}\, dx_{k+1} \propto s^{n-k-1}$$

where one takes $s = (1 - x_1^2 - \cdots - x_k^2)^{1/2}$. $\square$

Corollary A.3. *If the vector $(x_1, y_1, \dots, x_n, y_n)$ is chosen at random from the $(2n-1)$-sphere according to normalized surface measure, then*

$$(\mu_1, \dots, \mu_n) = \big(x_1^2 + y_1^2, \dots, x_n^2 + y_n^2\big)$$

is uniformly distributed on the $(n-1)$-simplex $\sum \mu_j = 1$. That is,

$$\mathbb{E} f(\mu_1, \dots, \mu_n) = (n-1)! \int_{\triangle} f(s_1, \dots, s_{n-1}, 1 - s_1 - \cdots - s_{n-1})\, ds_1 \cdots ds_{n-1}$$

where $\triangle = \{(s_1, \dots, s_{n-1}) \,:\, 0 \leq s_1 + \cdots + s_{n-1} \leq 1\}$.

Proof. By Lemma A.1, we need only compute

$$\sum_{\pm} \int_{\mathbb{D}^{2n-1}} f(x_1, y_1, \dots, x_n, \pm y_n)\, \frac{dx_1 dy_1 \cdots dx_n}{y_n}$$

where $y_n = \sqrt{1 - (x_1^2 + y_1^2 + \cdots + x_n^2)}$.

Let us change variables by

$$\left.\begin{aligned} x_j &= \cos(\phi_j)\sqrt{s_j} \\ y_j &= \sin(\phi_j)\sqrt{s_j} \end{aligned}\right\} 1 \leq j \leq n-1$$

$$x_n = \cos(\phi_n)\bigg(1 - \sum_{k=1}^{n-1} s_k\bigg)^{1/2},$$

for which the Jacobian is $y_n^{-1} dx_1 dy_1 \cdots dx_n = ds_1 ds_2 \cdots ds_{n-1} d\phi_1 \dots d\phi_n$. Hence, up to proportionality constants that depend only on n,

$$\mathbb{E} f(\mu_1, \dots, \mu_n) \propto \int_{\triangle} f(s_1, \dots, s_{n-1}, 1 - s_1 - \cdots - s_{n-1})\, ds_1 \cdots ds_{n-1}.$$

The determination of the normalization constant is an easy calculation. A more general result will appear in the proof of Lemma A.4. $\square$

Lemma A.4. *For $(\mu_1, \dots, \mu_n)$ uniformly distributed on the $(n-1)$-simplex $\sum \mu_j = 1$ and $p_1, \dots, p_n$ real numbers greater than -1, we have*

$$\mathbb{E}\big\{\textstyle\prod \mu_j^{p_j}\big\} = \frac{(n-1)!\, \Gamma(p_1 + 1) \cdots \Gamma(p_n + 1)}{\Gamma(p_1 + \cdots + p_n + n)}. \tag{A.3}$$

Proof. This is sometimes known as Dirichlet's integral; our method of proof is a standard one. We begin with the product integral

$$\prod_{j=1}^{n} \int_0^\infty s_j^{p_j} e^{-s_j}\, ds_j = \prod_{j=1}^{n} \Gamma(p_j + 1)$$

and change variables to $r = \sum s_j$ and $\mu_j = s_j/r$. The Jacobian is given by $r^{n-1}\,dr\,d\mu_1\cdots d\mu_{n-1} = ds_1\cdots ds_n$ which leads us to

$$\begin{aligned}\prod_{j=1}^{n}\Gamma(p_j+1) &= \int_{\triangle}\int_0^\infty \prod (r\mu_j)^{p_j}\,e^{-r}r^{n-1}\,dr\,d\mu_1\cdots d\mu_{n-1}\\ &= \Gamma(p_1+\cdots+p_n+n)\int_{\triangle}\prod \mu_j^{p_j}\,d\mu_1\cdots d\mu_{n-1}.\end{aligned}$$

By taking $p_1 = \cdots = p_n = 0$ we recover the normalization constant of the μ_j-integral given in Corollary A.3. Moreover,

$$\begin{aligned}\mathbb{E}\big\{\prod \mu_j^{p_j}\big\} &= (n-1)!\int_{\triangle}\prod \mu_j^{p_j}\,d\mu_1\cdots d\mu_{n-1}\\ &= \frac{(n-1)!\,\Gamma(p_1+1)\cdots\Gamma(p_n+1)}{\Gamma(p_1+\cdots+p_n+n)}\end{aligned} \tag{A.4}$$

as promised. □

Appendix B. The CMV Matrix and Geronimus Relations

The purpose of this appendix is to describe some results in the theory of orthogonal polynomials on the unit circle that were used in the main text; we also wish to show how easily they can be derived from the perspective of Cantero, Moral, and Velázquez, [3]. We begin with an outline of that work.

Applying the Gram–Schmidt procedure to $1, z, z^{-1}, z^2, z^{-2}, \ldots$ in $L^2(d\mu)$ produces the orthonormal basis

$$\chi_k(z) = \begin{cases} z^{-k/2}\phi_k^*(z) & k \text{ even}\\ z^{(1-k)/2}\phi_k(z) & k \text{ odd}\end{cases} \tag{B.1}$$

where $k \geq 0$. As in the Introduction, ϕ_k denotes the kth orthonormal polynomial and ϕ_k^*, its reversal (cf. (2.3)). If we apply the procedure to $1, z^{-1}, z, z^{-2}, z^2, \ldots$ instead, then we obtain a second orthonormal basis:

$$X_k(z) = \overline{\chi_k(1/\bar{z})} = \begin{cases} z^{-k/2}\phi_k(z) & k \text{ even}\\ z^{(-1-k)/2}\phi_k^*(z) & k \text{ odd.}\end{cases} \tag{B.2}$$

It is natural to compute the matrix representation of $f(z) \mapsto zf(z)$ in $L^2(d\mu)$ with respect to these bases. This was done in a rather cunning way. The matrices with entries

$$L_{i+1,j+1} = \langle \chi_i(z) | zX_j(z)\rangle \quad\text{and}\quad M_{i+1,j+1} = \langle X_i(z)|\chi_j(z)\rangle$$

are block-diagonal; indeed,

$$L = \operatorname{diag}\big(\Xi_0, \Xi_2, \Xi_4, \ldots\big) \quad\text{and}\quad M = \operatorname{diag}\big(\Xi_{-1}, \Xi_1, \Xi_3, \ldots\big) \tag{B.3}$$

where $\Xi_{-1} = [1]$ and

$$\Xi_k = \begin{bmatrix} \bar{\alpha}_k & \rho_k \\ \rho_k & -\alpha_k\end{bmatrix}.$$

The representation of $f(z) \mapsto zf(z)$ in the χ_j basis is just LM, which is the CMV matrix of the title of this appendix, and in the X_j basis, it is ML.

If $d\mu$ is supported at finitely many points, say m, then α_{m-1} will be unimodular and hence Ξ_{m-1} is diagonal. We replace Ξ_{m-1} by the 1×1 matrix that is its top left entry, $\bar{\alpha}_{m-1}$, and discard all Ξ_k with $k \geq m$. In this way, we find L and M are naturally $m \times m$ block-diagonal matrices.

In a couple of places we noted that the spectral measure for any $U \in \mathbb{SO}(2n)$ has last Verblunsky parameter equal to -1. The CMV matrix allows us to give a particularly short proof of this fact:

Lemma B.1. *Let $d\mu$ be a probability measure of the form*

$$\int f\, d\mu = \sum_j \mu_j f(z_j)$$

where $z_1, \ldots, z_m$ are distinct points on the unit circle. If α_{m-1} denotes the final Verblunsky parameter associated to this measure, then

$$\prod_{j=1}^{m} z_j = (-1)^{m-1} \bar{\alpha}_{m-1}. \tag{B.4}$$

Proof. As one might guess, (B.4) just represents two ways of calculating the determinant of the CMV matrix: the left-hand side is the product of the eigenvalues; the right-hand side is the product of the determinants of L and M. Note that $\det(\Xi_k) = -1$ for $0 \leq k < m$ while $\det(\Xi_{m-1}) = \bar{\alpha}_{m-1}$. □

Next we prove a result used in the derivation of the Aomoto integral:

Proposition B.2. *Given a finite system of Veblunsky parameters $\alpha_k \in \mathbb{D}$, $0 \leq k \leq m-2$, and $\alpha_{m-1} = e^{i\phi}$, define a second system by $\tilde{\alpha}_k = -e^{i\phi}\bar{\alpha}_{m-2-k}$, $0 \leq k \leq m-2$, and $\tilde{\alpha}_{m-1} = e^{i\phi}$. Then the monic orthogonal polynomials of degree m associated to these two systems are the same.*

Proof. If L and M are the matrices associated to the αs and Φ_m the monic orthogonal polynomial of degree m, then $\det(z - LM) = \Phi_m(z)$. This is because both are monic polynomials vanishing at the eigenvalues of LM.

Similarly, we have $\det(z - \tilde{L}\tilde{M}) = \tilde{\Phi}_m(z)$ for the $\tilde{\alpha}$s. Hence the proposition will follow once we show that LM and $\tilde{L}\tilde{M}$ are conjugate. We will give full details when m is even and a few remarks on the changes necessary when m is odd.

Conjugating LM by

$$\begin{bmatrix} 0 & \ldots & 0 & 1 \\ 0 & \ldots & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \ldots & 0 & 0 \end{bmatrix}$$

is equivalent to reversing the order of the rows and columns in each factor. As m is even, both L and M retain their structure. The transformation on L amounts to replacing α_k by $-\bar{\alpha}_{m-2-k}$; for M, we have the additional change that the $(1,1)$ entry becomes $e^{-i\phi}$ and the (m,m) entry, 1.

To convert M to its proper structure (with 1 in the first position), we perform the transformations $L \mapsto U^\dagger L V$ and $M \mapsto V^\dagger M U$ where

$$U = \operatorname{diag}(e^{i\phi}, 1, e^{i\phi}, 1, \ldots), \quad V = \operatorname{diag}(1, e^{i\phi}, 1, e^{i\phi}, \ldots).$$

In this way, we obtain the matrices $\tilde{L}$ and $\tilde{M}$, which shows that $\tilde{L}\tilde{M}$ is conjugate to the original LM.

When m is odd, reversing the order of the rows and columns converts L to a matrix whose structure resembles that of M, while M is converted to an L-like matrix. Proceeding as above shows that LM is conjugate to $\tilde{M}\tilde{L}$, and hence to $\tilde{L}\tilde{M}$. □

Let us now consider the case where the measure $d\mu$ is symmetric with respect to complex conjugation, or what is equivalent, where all Verblunsky parameters are real. It is a famous observation of Szegő (see [16, §11.5]) that the polynomials orthogonal with respect to this measure are intimately related to the polynomials orthogonal with respect to the measure $d\nu$ on $[-2, 2]$ defined by

$$\int_{S^1} f(z + z^{-1})\, d\mu(z) = \int_{-2}^{2} f(x)\, d\nu(x). \tag{B.5}$$

As described in Section 2, the recurrence coefficients for these systems of orthogonal polynomials are related by Geronimus relations:

$$\begin{aligned} b_{k+1} &= (1 - \alpha_{2k-1})\alpha_{2k} - (1 + \alpha_{2k-1})\alpha_{2k-2} \\ a_{k+1} &= \big\{(1 - \alpha_{2k-1})(1 - \alpha_{2k}^2)(1 + \alpha_{2k+1})\big\}^{1/2}. \end{aligned} \tag{B.6}$$

We will now present a short proof of these formulae. As an off-shoot of our method, we also recover relations to the recurrence coefficients for $(4 - x^2)\, d\nu(x)$ and $(2 \pm x)d\nu$. The former appears in the proposition below, the latter in the remark that follows it.

Proposition B.3. *Let α_k be the system of real Verblunsky parameters associated to a symmetric measure $d\mu$ and let L and M denote the matrices of* (B.3)*. Then $LM + ML$ is unitarily equivalent to the direct sum of two Jacobi matrices:*

$$J = \begin{bmatrix} b_1 & a_1 & 0 & \\ a_1 & b_2 & \ddots & \\ 0 & \ddots & \ddots & \end{bmatrix} \qquad \tilde{J} = \begin{bmatrix} \tilde{b}_1 & \tilde{a}_1 & 0 & \\ \tilde{a}_1 & \tilde{b}_2 & \ddots & \\ 0 & \ddots & \ddots & \end{bmatrix}$$

where a_k and b_k are as in (B.6) *and*

$$\begin{aligned} \tilde{b}_{k+1} &= (1 - \alpha_{2k+1})\alpha_{2k} - (1 + \alpha_{2k+1})\alpha_{2k+2} \\ \tilde{a}_{k+1} &= \big\{(1 + \alpha_{2k+1})(1 - \alpha_{2k+2}^2)(1 - \alpha_{2k+3})\big\}^{1/2}. \end{aligned}$$

Moreover, the spectral measure for (J, e_1) is precisely the $d\nu$ of (B.5)*. The spectral measure for $(\tilde{J}, e_1)$ is $\frac{1}{2(1-\alpha_0^2)(1-\alpha_1)}(4 - x^2)\, d\nu(x)$.*

Proof. Let S denote the following unitary block matrix

$$S = \operatorname{diag}([1], S_1, S_3, \dots) \quad \text{where} \quad S_k = \frac{1}{\sqrt{2}} \begin{bmatrix} -\sqrt{1 - \alpha_k} & \sqrt{1 + \alpha_k} \\ \sqrt{1 + \alpha_k} & \sqrt{1 - \alpha_k} \end{bmatrix},$$

which is easily seen to diagonalize M. Indeed, $S^\dagger M S = \operatorname{diag}(+1, -1, +1, -1, \dots)$. We will denote this matrix by R.

The matrix $LM+ML$ is unitarily equivalent to $A = S^\dagger(LM+ML)S = S^\dagger LSR + RS^\dagger LS$, which we will show is the direct sum of two Jacobi matrices. We begin by showing that even-odd and odd-even entries of A vanish, from which it follows that A is the direct sum of its even-even and odd-odd submatrices.

Left multiplication by R changes the sign of the entries in each even-numbered row, while right multiplication by R reverses the sign of each even-numbered column. In this way, $RB + BR$ has the stated direct sum structure for any matrix B and hence, in particular, for $B = S^\dagger LS$.

It remains only to calculate the non-zero entries of A. As S and L are both tri-diagonal, A must be hepta-diagonal and so the direct sum of tri-diagonal matrices. Moreover, A is symmetric (because L is) so there are only four categories of entries to calculate: the odd/even diagonals and the odd/even off-diagonals. We begin with the diagonals:

$$\begin{aligned} A_{2k+1,2k+1} &= \begin{bmatrix}\sqrt{1+\alpha_{2k-1}} & \sqrt{1-\alpha_{2k-1}}\end{bmatrix}\begin{bmatrix}-\alpha_{2k-2} & 0\\ 0 & \alpha_{2k}\end{bmatrix}\begin{bmatrix}\sqrt{1+\alpha_{2k-1}}\\ \sqrt{1-\alpha_{2k-1}}\end{bmatrix}\\ &= (1-\alpha_{2k-1})\alpha_{2k} - (1+\alpha_{2k-1})\alpha_{2k-2}\\ A_{2k,2k} &= -\begin{bmatrix}-\sqrt{1-\alpha_{2k-1}} & \sqrt{1+\alpha_{2k-1}}\end{bmatrix}\begin{bmatrix}-\alpha_{2k-2} & 0\\ 0 & \alpha_{2k}\end{bmatrix}\begin{bmatrix}-\sqrt{1-\alpha_{2k-1}}\\ \sqrt{1+\alpha_{2k-1}}\end{bmatrix}\\ &= (1-\alpha_{2k-1})\alpha_{2k-2} - (1+\alpha_{2k-1})\alpha_{2k}. \end{aligned}$$

Note that the factor of 2 resulting from A being the sum of two terms is cancelled by the factors of $2^{-1/2}$ coming from S and $S^\dagger$. The calculation of the off-diagonal terms proceeds in a similar fashion:

$$\begin{aligned} A_{2k+1,2k+3} &= \begin{bmatrix}\sqrt{1+\alpha_{2k-1}} & \sqrt{1-\alpha_{2k-1}}\end{bmatrix}\begin{bmatrix}0 & 0\\ \rho_{2k} & 0\end{bmatrix}\begin{bmatrix}\sqrt{1+\alpha_{2k+1}}\\ \sqrt{1-\alpha_{2k+1}}\end{bmatrix}\\ &= \sqrt{(1-\alpha_{2k-1})(1-\alpha_{2k}^2)(1+\alpha_{2k+1})}\\ A_{2k,2k+2} &= -\begin{bmatrix}-\sqrt{1-\alpha_{2k-1}} & \sqrt{1+\alpha_{2k-1}}\end{bmatrix}\begin{bmatrix}0 & 0\\ \rho_{2k} & 0\end{bmatrix}\begin{bmatrix}-\sqrt{1-\alpha_{2k+1}}\\ \sqrt{1+\alpha_{2k+1}}\end{bmatrix}\\ &= \sqrt{(1+\alpha_{2k-1})(1-\alpha_{2k}^2)(1-\alpha_{2k+1})}. \end{aligned}$$

That $d\nu$ is the spectral measure for (J, e_1) is an immediate consequence of the spectral theorem, $LM+ML = LM+(LM)^{-1}$, and the fact that S leaves the vector $[1, 0, \dots, 0]$ invariant.

Tracing back through the definitions, we find that the spectral measure for $(\tilde{J}, e_1)$ is equal to that for the operator $f(z) \mapsto (z + z^{-1})f(z)$ in $L^2(d\mu)$ and the vector

$$f(z) = \left(\tfrac{1+\alpha_1}{2}\right)^{\frac12}\chi_2(z) - \left(\tfrac{1-\alpha_1}{2}\right)^{\frac12}\chi_1(z) = \left(\tfrac{1+\alpha_1}{2}\right)^{\frac12} z^{-1}\phi_2^*(z) - \left(\tfrac{1-\alpha_1}{2}\right)^{\frac12}\phi_1(z).$$

From the relations (2.1), (2.2), and (2.4), we find $\rho_1\phi_2^*(z) = \phi_1^*(z) - \alpha_1 z\phi_1(z)$, $\rho_0\phi_1(z) = z - \alpha_0$, and $\rho_0\phi_1^*(z) = 1 - \alpha_0 z$. These simplify the formula considerably:

$$f(z) = \frac{z^{-1} - z}{\rho_0\sqrt{2(1-\alpha_1)}}.$$

The expression for the spectral measure for $(\tilde{J}, e_1)$ now follows from the simple calculation $|z^{-1} - z|^2 = 4 - (z + z^{-1})^2$. □

Remark: In the above proof, we conjugated $LM+ML$ by the unitary matrix which diagonalizes M. One may use the matrix $\operatorname{diag}(S_0, S_2, \dots)$, which diagonalizes L, instead. This also conjugates $LM+ML$ to the direct sum of two Jacobi matrices. In this way, we learn that the recurrence coefficients for $\frac{1}{2(1\pm\alpha_0)}(2\pm x)d\nu(x)$ are given by

$$\begin{aligned} b_{k+1} &= \pm(1\mp\alpha_{2k})\alpha_{2k+1} \mp (1\pm\alpha_{2k})\alpha_{2k-1}\\ a_{k+1} &= \left\{(1\mp\alpha_{2k})(1-\alpha_{2k+1}^2)(1\pm\alpha_{2k+2})\right\}^{1/2}. \end{aligned}$$

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions, with Formulas, Graphs and Mathematical Tables.* National Bureau of Standards, Washington, D.C., 1966.

[2] K. Aomoto, Jacobi polynomials associated with Selberg integrals. *SIAM J. Math. Anal.* **18** (1987), 545–549.

[3] M. J. Cantero, L. Moral, and L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle. *Linear Algebra Appl.* **362** (2003), 29–56.

[4] M. J. Cantero, L. Moral, and L. Velázquez, Minimal representations of unitary operators and orthogonal polynomials on the unit circle. Preprint.

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

[6] F. Dyson, Statistical theory of the energy levels of complex systems. I, II, and III. *J. Math. Phys.* **3** (1962), 140–156, 157–165, and 166–175.

[7] P. J. Forrester and E. M. Rains, Interpretations of some parameter dependent generalizations of classical matrix ensembles. math-ph/0211042.

[8] Ya. L. Geronimus, On polynomials orthogonal on the circle, on trigonometric moment-problem and on allied Carathodory and Schur functions. (Russian.) *Mat. Sbornik* **15(57)**, (1944), 99–130.

[9] Ya. L. Geronimus, *Orthogonal Polynomials.* Consultants Bureau, New York 1961.

[10] I. J. Good, Short proof of a conjecture by Dyson. *J. Math. Phys.* **11** (1970), 1884.

[11] J. Gunson, Proof of a conjecture by Dyson in the statistical theory of energy levels. *J. Math. Phys.* **3** (1962), 752–753.

[12] A. S. Householder, *The Theory of Matrices in Numerical Analysis.* Blaisdell Publishing Company, New York, 1965.

[13] R. A. Lippert, A matrix model for the β-Jacobi ensemble. *J. Math. Phys.* **44** (2002), 4807–4816.

[14] A. Selberg, Bemerkninger om et multipelt integral. *Norsk Mat. Tidsskr.* **26** (1944), 71–78.

[15] B. Simon, *Orthogonal Polynomials on the Unit Circle.* American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, Rhode Island, expected 2004.

[16] G. Szegő, *Orthogonal Polynomials.* American Mathematical Society Colloquium Publications, Vol. XXIII. American Mathematical Society, Providence, Rhode Island, 1975.

[17] H. Weyl, *The Classical Groups: Their Invariants and Representations,* second edition. Princeton University Press, Princeton, New Jersey, 1946.

[18] K. Wilson, Proof of a conjecture by Dyson. *J. Math. Phys.* **3** (1962), 1040–1043.

Rowan Killip, UCLA Mathematics Department, Box 951555, Los Angeles, CA 90095
E-mail address: killip@math.ucla.edu

Irina Nenciu, Mathematics 253-37, Caltech, Pasadena, CA 91125
E-mail address: nenciu@caltech.edu