Circular Jacobi Ensembles and Deformed Verblunsky Coefficients

Paul Bourgade[1], Ashkan Nikeghbali[2], and Alain Rouault[3]

[1]ENST, 46 rue Barrault, 75634 Paris Cedex 13, Université Paris 6, LPMA, 175, rue du Chevaleret F-75013 Paris, France, [2]Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland, and [3]Université Versailles-Saint Quentin, LMV, Bâtiment Fermat, 45 avenue des Etats-Unis, 78035 Versailles Cedex, France

Correspondence to be sent to: ashkan.nikeghbali@math.uzh.ch, ashkan.nikeghbali@gmail.com

Using the spectral theory of unitary operators and the theory of orthogonal polynomials on the unit circle, we propose a simple matrix model for the following circular analog of the Jacobi ensemble:

$$c_{\delta,\beta}^{(n)} \prod_{1 \le k < l \le n} |e^{i\theta_k} - e^{i\theta_l}|^\beta \prod_{j=1}^{n} (1 - e^{-i\theta_j})^\delta (1 - e^{i\theta_j})^{\overline{\delta}}$$

with $\Re\mathfrak{e}\,\delta > -1/2$. If e is a cyclic vector for a unitary $n \times n$ matrix U, the spectral measure of the pair (U, e) is well parameterized by its Verblunsky coefficients $(\alpha_0, \ldots, \alpha_{n-1})$. We introduce here a deformation $(\gamma_0, \ldots, \gamma_{n-1})$ of these coefficients so that the associated Hessenberg matrix (called GGT) can be decomposed into a product $r(\gamma_0) \cdots r(\gamma_{n-1})$ of elementary reflections parameterized by these coefficients. If $\gamma_0, \ldots, \gamma_{n-1}$ are independent random variables with some remarkable distributions, then the eigenvalues of the GGT matrix follow the circular Jacobi distribution above.

These deformed Verblunsky coefficients also allow us to prove that, in the regime $\delta = \delta(n)$ with $\delta(n)/n \to \beta\mathrm{d}/2$, the spectral measure and the empirical spectral distribution

weakly converge to an explicit nontrivial probability measure supported by an arc of the unit circle. We also prove the large deviations for the empirical spectral distribution.

1   Introduction

1.1   The circular Jacobi ensemble

The theory of random unitary matrices was developed using the existence of a natural probability uniform measure on compact Lie groups, namely the Haar measure. The statistical properties of the eigenvalues as well as the characteristic polynomial of these random matrices have played a crucial role both in physics (see [25] for an historical account) and in analytic number theory to model L-functions (see [19] and [20] where Keating and Snaith predict moments of L-functions on the critical line using knowledge on the moments of the characteristic polynomial of random unitary matrices).

The circular unitary ensemble (CUE) is $U(n)$, the unitary group over $\mathbb{C}^n$, equipped with its Haar probability measure $\mu_{U(n)}$. Weyl's integration formula allows one to average any (bounded measurable) function on $U(n)$ which is conjugation-invariant

$$\int f d\mu_{U(n)} = \frac{1}{n!} \int \cdots \int |\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^2 f(\mathrm{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})) \frac{d\theta_1}{2\pi} \cdots \frac{d\theta_n}{2\pi}, \qquad (1.1)$$

where $\Delta(e^{i\theta_1}, \dots, e^{i\theta_n}) = \prod_{1 \le j < k \le n}(e^{i\theta_k} - e^{i\theta_j})$ denotes the Vandermonde determinant.

The circular orthogonal ensemble (COE) is the subset of $U(n)$ consisting of symmetric matrices, i.e. $U(n)/O(n) = \{VV^T; V \in U(n)\}$ equipped with the measure obtained by pushing forward $\mu_{U(n)}$ by the mapping $V \mapsto VV^T$. The integration formula is similar to (1.1) but with $|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^2$ replaced by $|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|$ and with the normalizing constant changed accordingly.

For the circular symplectic ensemble (CSE), which will not be recalled here, the integration formula uses $|\Delta(e^{i\theta_1}, \dots, e^{i\theta_n})|^4$.

Dyson observed that the induced eigenvalue distributions correspond to the Gibbs distribution for the classical Coulomb gas on the circle at three different temperatures. More generally, n identically charged particles confined to move on the unit circle, each interacting with the others through the usual Coulomb potential $-\log|z_i - z_j|$, give rise to the Gibbs measure at temperature $1/\beta$ (see the discussion and references in [23]

and [12], Chapter 2):

$$\mathbb{E}_n^\beta(f) = c_{0,\beta}^{(n)} \int f(e^{i\theta_1}, \ldots, e^{i\theta_n}) |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^\beta d\theta_1 \ldots d\theta_n, \tag{1.2}$$

where $c_{0,\beta}^{(n)}$ is a normalizing constant chosen so that

$$h_{0,\beta}^{(n)}(\theta_1, \ldots, \theta_n) = c_{0,\beta}^{(n)} |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^\beta \tag{1.3}$$

is a probability density on $(0, 2\pi)^n$ and where f is any symmetric function. The unitary, orthogonal, and symplectic circular ensembles correspond to matrix models for the Coulomb gas at $\beta = 1, 2, 4$ respectively, but are there matrix models for general $\beta > 0$ for Dyson's circular eigenvalue statistics?

Killip and Nenciu [23] provided matrix models for Dyson's circular ensemble, using the theory of orthogonal polynomials on the unit circle. In particular, they obtained a sparse matrix model which is five-diagonal, called CMV (after the names of the authors Cantero, Moral, and Velásquez [8]). In this framework, there is not a natural underlying measure such as the Haar measure; the matrix ensemble is characterized by the laws of its elements.

There is an analog of Dyson's circular ensembles on the real line: the probability density function of the eigenvalues $(x_1, \ldots, x_n)$ for such ensembles with temperature $1/\beta$ is proportional to

$$|\Delta(x_1, \ldots, x_n)|^\beta \prod_{j=1}^n e^{-x_j^2/2} \tag{1.4}$$

For $\beta = 1, 2$, or 4, this corresponds to the classical Gaussian ensembles. Dimitriu and Edelman [10] gave a simple tridiagonal matrix model for (1.4). Killip and Nenciu [23] gave an analog matrix model for the Jacobi measure on $[-2, 2]^n$, which is up to a normalizing constant,

$$|\Delta(x_1, \ldots, x_n)|^\beta \prod_{j=1}^n (2 - x_j)^a (2 + x_j)^b dx_1 \ldots dx_n, \tag{1.5}$$

where $a, b > 0$, relying on the theory of orthogonal polynomials on the unit circle and its links with orthogonal polynomials on the segment. When a and b are strictly positive integers, the Jacobi measure (1.5) can be interpreted as the potential $|\Delta(x_1, \ldots, x_{n+a+b})|^\beta$ on $[-2, 2]^{n+a+b}$ conditioned to have a elements located at 2 and b elements located at -2. Consequently, the Jacobi measure on the unit circle should be a two parameters extension of (1.3), corresponding to conditioning to have specific given eigenvalues.

Such an analog was defined as the *circular Jacobi ensemble* in [12] and [14]. If $\delta \in \mathbb{R}$, we recover the cJUE as in Witte and Forrester [35].

Definition 1.1. Throughout this paper, we note $h_{\delta,\beta}^{(n)}$ the probability density function on $(0, 2\pi)$ given by

$$h_{\delta,\beta}^{(n)}(\theta_1, \ldots, \theta_n) = c_{\delta,\beta}^{(n)} |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^\beta \prod_{j=1}^{n} (1 - e^{-i\theta_j})^\delta (1 - e^{i\theta_j})^{\overline{\delta}} \tag{1.6}$$

with $\delta \in \mathbb{C}$, $\mathfrak{Re}(\delta) > -\frac{1}{2}$.

If $\delta \in \frac{\beta}{2}\mathbb{N}$, this measure coincides with (1.3) conditioned to have eigenvalues at 1. For $\beta = 2$, such measures were first considered by Hua [18] and Pickrell [28, 29]. This case was also widely studied in [27] and [5] for its connections with the theory of representations and in [7] for its analogies with the Ewens measures on permutation groups.

One of our goals in this paper is to provide matrix models for the circular Jacobi ensemble, i.e. a distribution on $U(n)$ such that the arguments of the eigenvalues $(e^{i\theta_1}, \ldots, e^{i\theta_n})$ are distributed as in (1.6). One can guess that additional problems may appear because the distribution of the eigenvalues is not rotation-invariant anymore. Nevertheless, some statistical information for the circular Jacobi ensemble can be obtained from Dyson's circular ensemble by a sampling (or a change of probability measure) with the help of the determinant. More precisely, let us first define the notion of sampling.

Definition 1.2. Let $(X, \mathcal{F}, \mu)$ be a probability space, and $h : X \mapsto \mathbb{R}^+$ be a measurable and integrable function with $\mathbb{E}_\mu(h) > 0$. Then a measure μ' is said to be the *h-sampling* of μ, if for all bounded measurable functions f,

$$\mathbb{E}_{\mu'}(f) = \frac{\mathbb{E}_\mu(fh)}{\mathbb{E}_\mu(h)}.$$

If we consider a matrix model for $h_{0,\beta}^{(n)}$, we can define (for $\mathfrak{Re}(\delta) > -1/2$, due to an integrability constraint) a matrix model for $h_{\delta,\beta}^{(n)}$ by the means of a sampling, noticing that when the charges are actually the eigenvalues of a matrix U, then (1.3) differs from (1.6) by a factor which is a function of $\det(\mathrm{Id} - U)$. We define $\det_\delta$ for a unitary matrix U as

$$\det_\delta(U) = \det(\mathrm{Id} - U)^{\overline{\delta}} \det(\mathrm{Id} - \overline{U})^\delta,$$

and we will use this $\det_\delta$-sampling.

Actually we look for an effective construction of a random matrix, for instance starting from a reduced number of independent random variables with known distributions. Notice that in the particular case $\beta = 2$, the density $h_{0,2}$ corresponds to eigenvalues of a matrix under the Haar measure on $U(n)$ and the $\det_\delta$-sampling of this measure is the Hua–Pickrell measure studied in our previous paper [7].

1.2 Orthogonal polynomials on the unit circle

We now wish to outline the main ingredients which are needed from the theory of orthogonal polynomials on the unit circle to construct matrix models for the general Dyson's circular ensemble. The reader can refer to [33] and [34] for more results and references; in particular, all the results about orthogonal polynomials on the unit circle (hereafter OPUC) can be found in these volumes.

Let us explain why OPUC play a prominent role in these constructions. In this paper, $\mathbb{D}$ denotes the open unit disk $\{z \in \mathbb{C} : |z| < 1\}$ and $\mathbb{T}$ denotes the unit circle $\{z \in \mathbb{C} : |z| = 1\}$. Let $(\mathcal{H}, u, e)$ be a triple where $\mathcal{H}$ is a Hilbert space, u be a unitary operator, and e be a cyclic unit vector, i.e. $\{u^j e\}_{j=-\infty}^{\infty}$ is total in $\mathcal{H}$. We say that two triples $(\mathcal{H}, u, e)$ and $(\mathcal{K}, v, e')$ are equivalent if and only if there exits an isometry $k : \mathcal{H} \to \mathcal{K}$ such that $v = kuk^{-1}$ and $e' = ke$. The spectral theorem says that for each equivalence class, there exists a unique probability measure μ on $\mathbb{T}$ such that

$$\langle e, u^k e \rangle_{\mathcal{H}} = \int_{\mathbb{T}} z^k d\mu(z), \quad k = 0, \pm 1, \dots .$$

Conversely, such a probability measure μ gives rise to a triple consisting of the Hilbert space $L^2(\mathbb{T}, \mu)$, the operator of multiplication by z, i.e. $h \mapsto (z \mapsto zh(z))$ and the vector $\mathbf{1}$, i.e. the constant function 1. When the space $\mathcal{H}$ is fixed, the probability measure μ associated with the triple $(\mathcal{H}, u, e)$ is called the *spectral measure of the pair* (u, e).

Let us consider the finite n-dimensional case. Assume that u is unitary and e is cyclic. It is classical that u has n different eigenvalues $(e^{i\theta_j}, j = 1, \dots, n)$ with $\theta_j \in [0, 2\pi)$. In any orthonormal basis whose first vector is e, say $(e_1 = e, e_2, \dots, e_n)$, u is represented by a matrix U and there is a unitary matrix Π diagonalizing U. It is then straightforward that the spectral measure is

$$\mu = \sum_{j=1}^{n} \pi_j \, \delta_{e^{i\theta_j}}, \tag{1.7}$$

where the weights are defined as $\pi_j = |\langle e_1, \Pi e_j \rangle|^2$. Note that $\pi_j > 0$ because a cyclic vector cannot be orthogonal to any eigenvector (and we also have $\sum_{j=1}^{n} \pi_j = 1$ because Π is

unitary). The eigenvalues $(e^{i\theta_j}, j = 1, \ldots n)$ and the vector $(\pi_1, \ldots, \pi_n)$ can then be used as coordinates for the probability measure μ.

Keeping in mind our purpose, we see that the construction of a matrix model from a vector $(e^{i\theta_j}, j = 1, \ldots, n)$ may be achieved in two steps: first, give a vector of weights $(\pi_1, \ldots, \pi_n)$, then find a matricial representative of the equivalence class with a rather simple form. The key tool for the second task is the sequence of orthogonal polynomials associated with the measure μ. In $L^2(\mathbb{T}, \mu)$ equipped with the natural basis $\{1, z, z^2, \ldots, z^{n-1}\}$, the Gram–Schmidt procedure provides the family of monic orthogonal polynomials $\Phi_0, \ldots, \Phi_{n-1}$. We can still define Φ_n as the unique monic polynomial of degree n with $\| \Phi_n \|_{L^2(\mathbb{T}, \mu)} = 0$, namely

$$\Phi_n(z) = \prod_{j=1}^{n} (z - e^{i\theta_j}). \tag{1.8}$$

The Φ_k's ($k = 0, \ldots, n$) obey the Szegö recursion relation

$$\Phi_{j+1}(z) = z\Phi_j(z) - \bar{\alpha}_j \Phi_j^*(z), \tag{1.9}$$

where

$$\Phi_j^*(z) = z^j \overline{\Phi_j(\bar{z}^{-1})}. \tag{1.10}$$

The coefficients α_j's ($0 \le j \le n - 1$) are called Verblunsky coefficients and satisfy the condition $\alpha_0, \ldots, \alpha_{n-2} \in \mathbb{D}$ and $\alpha_{n-1} \in \mathbb{T}$.

When the measure μ has infinite support, one can define the family of orthogonal polynomials $(\Phi_n)_{n \ge 0}$ associated with μ for all n. Then there are infinitely many Verblunsky coefficients (α_n) which all lie in $\mathbb{D}$.

Verblunsky's theorem (see for example [33, 34]) states that there is a bijection between probability measures on the unit circle and sequences of Verblunsky coefficients. The matrix of the multiplication by z in $L^2(\mathbb{T}, \mu)$, in the basis of orthonormal polynomials, has received much attention. This unitary matrix, noted $\mathcal{G}(\alpha_0, \ldots, \alpha_{n-1})$, called GGT by Simon (see Chapter 4.1 in [33] for more details and an historical account), is in the Hessenberg form: all entries below the subdiagonal are zero, whereas the entries above the subdiagonal are nonzero and the subdiagonal is nonnegative (see formulae (4.1.5) and (4.1.6) in [33] for an explicit expression for the entries in terms of the Verblunsky coefficients, or formula (1.11) in Lemma 1.3 below).

For H an $n \times n$ complex matrix, the subscript H_{ij} stands for $\langle e_i, H(e_j) \rangle$, where $\langle x, y \rangle = \sum_{k=1}^{n} \bar{x}_k y_k$. Killip and Nenciu state that any unitary matrix in the Hessenberg form with nonnegative subdiagonal is the matrix of multiplication by z in $L^2(\mathbb{T}, \mu)$ for some

measure μ. More precisely, from Killip-Nenciu [23], Lemma 3.2, we have the following lemma.

Lemma 1.3. Let μ be a probability measure on $\mathbb{T}$ supported at n points. Then the matrix H of $f(z) \mapsto zf(z)$, in the basis of orthonormal polynomials of $L^2(\mathbb{T}, \mu)$, is in the Hessenberg form.

More precisely

$$H_{i+1,j+1} = \begin{cases} -\alpha_{i-1}\overline{\alpha}_j \prod_{p=i}^{j-1} \rho_p, & \text{if } i < j+1, \\ \rho_{j-1}, & \text{if } i = j+1, \\ 0, & \text{if } i > j+1 \end{cases} \tag{1.11}$$

with $\rho_j = \sqrt{1 - |\alpha_j|^2}$ and $\alpha_{-1} = -1$, the α_k's being the Verblunsky coefficients associated with μ. Conversely, if $\alpha_0, \ldots, \alpha_{n-1}$ are given in $\mathbb{D}^{n-1} \times \mathbb{T}$, and if we define the matrix H by (1.11), then the spectral measure of the pair (H, e_1) is the measure μ whose Verblunsky coefficients are precisely $\alpha_0, \ldots, \alpha_{n-1}$.

Besides, there is a very useful decomposition of these matrices into product of block matrices, called the AGR decomposition by Simon [31], after the paper [2]. For $0 \le k \le n-2$, let

$$\Theta^{(k)}(\alpha) = \text{Id}_k \oplus \begin{pmatrix} \overline{\alpha}_k & \rho_k \\ \rho_k & -\alpha_k \end{pmatrix} \oplus \text{Id}_{n-k-2},$$

and set $\Theta^{(n-1)}(\alpha_{n-1}) = \text{Id}_{n-1} \oplus (\overline{\alpha}_{n-1})$, with $|\alpha_{n-1}| = 1$. Then the AGR decomposition states that ([31], Theorem 10.1)

$$\mathcal{G}(\alpha_0, \ldots, \alpha_{n-1}) = \Theta^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1) \cdots \Theta^{(n-1)}(\alpha_{n-1}).$$

Now, we state a crucial result of Killip and Nenciu which enabled them to obtain a matrix model in the Hessenberg form for Dyson's circular ensemble. The challenge consists in randomizing $(\alpha_0, \ldots, \alpha_{n-1})$ in $\mathbb{D}^{n-1} \times \mathbb{T}$ in such a way that the angles $(\theta_1, \ldots, \theta_n)$ of the spectral measure

$$\mu = \sum_{j=1}^{n} \pi_j \delta_{e^{i\theta_j}}$$

have the density $h_{0,\beta}^{(n)}$ (see (1.3)). To make the statement precise, let us introduce three definitions (for the properties of the Beta and Dirichlet distributions, see [30]).

Definition 1.3.   For $a_1, a_2 > 0$, let $\mathrm{Beta}(a_1, a_2)$ be the distribution on $[0, 1]$ with density

$$\frac{\Gamma(a_1 + a_2)}{\Gamma(a_1)\Gamma(a_2)} x^{a_1 - 1}(1 - x)^{a_2 - 1}.$$

Its generalization is the following definition.

Definition 1.4.   For $n \geq 2$ and $a_1, \ldots, a_n > 0$, let $\mathrm{Dir}(a_1, \ldots, a_n)$ be the distribution on the simplex $\{(x_1, \ldots, x_n) \in [0, 1]^n : \sum_{i=1}^n x_i = 1\}$ with density

$$\frac{\Gamma(a_1 + \cdots + a_n)}{\Gamma(a_1) \cdots \Gamma(a_n)} \prod_{k=1}^n x_k^{a_k - 1}.$$

If $a_1 = \cdots = a_n = a$, it is called the Dirichlet distribution of order $n \geq 2$ with parameter $a > 0$ and denoted by $\mathrm{Dir}_n(a)$. (For $a = 1$, this is the uniform distribution.)

Definition 1.5.   For $s > 1$ let ν_s be the probability measure on $\mathbb{D}$ with density

$$\frac{s - 1}{2\pi}(1 - |z|^2)^{(s-3)/2}.$$

It is the law of $re^{i\psi}$ where r and ψ are independent, ψ is uniformly distributed on $(0, 2\pi)$, and r^2 has the $\mathrm{Beta}(1, (s - 1)/2)$ distribution. We adopt the convention that ν_1 is the uniform distribution on the unit circle. We denote by $\eta_{0,\beta}^{(n)}$ the distribution on $\mathbb{D}^{n-1} \times \mathbb{T}$ given by

$$\eta_{0,\beta}^{(n)} = \otimes_{k=0}^{n-1} \nu_{\beta(n-k-1)+1} . \tag{1.12}$$

Proposition 1.7 (Killip–Nenciu [23], Proposition 4.2).   The following formulae express the same measure on the manifold of probability distributions on $\mathbb{T}$ supported at n points:

$$\frac{2^{1-n}}{n!}|\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^\beta \prod_{j=1}^n \pi_j^{\beta/2-1} d\theta_1 \cdots d\theta_n d\pi_1 \cdots d\pi_{n-1}$$

in the (θ, π) coordinates and

$$\prod_{k=0}^{n-2}(1 - |\alpha_k|^2)^{(\beta/2)(n-k-1)-1} d^2\alpha_0 \cdots d^2\alpha_{n-2} \frac{d\phi}{2\pi} \tag{1.13}$$

in terms of the Verblunsky coefficients.

Proposition 1.7 may be restated as follows: to pick at random a measure μ such that $(\alpha_0, \ldots, \alpha_{n-1})$ is $\eta_{0,\beta}^{(n)}$ distributed is equivalent to pick the support $(\theta_1, \ldots, \theta_n)$ according to $h_{0,\beta}^{(n)}$ (see (1.3)) and to pick the weights $(\pi_1, \ldots, \pi_n)$ independently according to $\mathrm{Dir}_n(\beta/2)$.

As a consequence, if one takes independent coefficients $(\alpha_0, \ldots, \alpha_{n-1})$ such that α_k is $\nu_{\beta(n-k-1)+1}$ distributed for $0 \le k \le n-1$, then the GGT matrix $\mathcal{G}(\alpha_0, \ldots, \alpha_{n-1})$ will be a matrix model for Dyson's circular ensemble with temperature $1/\beta$ (see also Proposition 2.11 in [12]). Actually in [23], Killip and Nenciu provide a matrix model which is much sparser (five-diagonal), as explained in Section 4.

Let us now define the laws on $U(n)$ which we will consider in the sequel.

Definition 1.8. We denote by $\mathrm{CJ}_{0,\beta}^{(n)}$ the probability distribution supported by the set of $n \times n$ GGT matrices of the form (1.11), corresponding to the law of $\mathcal{G}(\alpha_0, \ldots, \alpha_{n-1})$ defined above. We denote by $\mathrm{CJ}_{\delta,\beta}^{(n)}$ the probability distribution on $U(n)$ which is the $\det_\delta$-sampling of $\mathrm{CJ}_{0,\beta}^{(n)}$.

The standard GGT approach is not sufficient to produce matrix models for the circular Jacobi ensemble because, as we shall see in Section 3, under the measure $\mathrm{CJ}_{\delta,\beta}^{(n)}$, the Verblunsky coefficients are not independent anymore. To overcome this difficulty, we associate with a measure on the unit circle, or equivalently with its Verblunsky coefficients, a new sequence of coefficients $(\gamma_k)_{0 \le k \le n-1}$, which we call deformed Verblunsky coefficients. There is a simple bijection between the original sequence $(\alpha_k)_{0 \le k \le n-1}$ and the new one $(\gamma_k)_{0 \le k \le n-1}$. These coefficients satisfy among other nice properties that $|\alpha_k| = |\gamma_k|$, and that they are independent under $\mathrm{CJ}_{\delta,\beta}^{(n)}$ (and for $\delta = 0$ the α_k's and the γ_k's have the same distribution). They have a geometric interpretation in terms of reflections: this leads to a decomposition of the GGT matrix $\mathcal{G}(\alpha_0, \ldots \alpha_{n-1})$ as a product of independent elementary reflections (constructed from the γ_k's). The explicit expression of the densities allows an asymptotic study (as $n \to \infty$) of the γ_k's, and consequently of the spectral measure, and finally of the empirical spectral distribution.

1.3 Organization of the paper

In Section 2, after recalling basic facts about the reflections introduced in [7], we define the deformed Verblunsky coefficients $(\gamma_k)_{0 \le k \le n-1}$ and give some of their basic properties. In particular, we prove that the GGT matrix $\mathcal{G}(\alpha_0, \ldots \alpha_{n-1})$ can be decomposed into a product of elementary complex reflections (Theorem 2.8).

In Section 3, we derive the law of the γ_k's under $\mathrm{CJ}_{\delta,\beta,}^{(n)}$ (Theorem 3.2); in particular we show that they are independent and that the actual Verblunsky coefficients are dependent if $\delta \neq 0$. We then prove an analog of Proposition 1.7 on the (θ,π) coordinates of μ (Theorem 3.3).

In Section 4, we propose our matrix model (Theorem 4.1). It is a modification of the AGR factorization, where we transform the Θ_k's so that they become reflections:

$$\Xi^{(k)}(\alpha) = \mathrm{Id}_k \oplus \begin{pmatrix} \overline{\alpha} & e^{\mathrm{i}\phi}\rho \\ \rho & -e^{\mathrm{i}\phi}\alpha \end{pmatrix} \oplus \mathrm{Id}_{n-k-2}$$

with $e^{\mathrm{i}\phi} = \frac{1-\overline{\alpha}}{1-\alpha}$. Of course, the CMV representation [8], which is five-diagonal, is also available, but this time the α_k's are not independent. Using the following elementary fact proven in Section 2:

$$\Phi_n(1) = \det(\mathrm{Id} - U) = \prod_{k=0}^{n-1}(1 - \gamma_k),$$

we are able to generalize our previous results in [6] and [7] about the decomposition of the characteristic polynomial evaluated at 1 as a product of independent complex variables (Proposition 4.3).

In Section 5, we study asymptotic properties of our model as $n \to \infty$, when $\delta = \beta n \mathrm{d}/2$, with $\mathfrak{Re}\,\mathrm{d} \geq 0$. We first prove that the Verblunsky coefficients have deterministic limits in probability. This entails that the spectral measure converges weakly in probability to a deterministic measure (denoted by μ_d) which is supported by an arc of the unit circle (Theorem 5.1). Besides, we consider the empirical spectral distribution (ESD), where the Dirac masses have the same weight $1/n$. Bounding the distances between both random measures, we prove that the ESD has the same limit (Theorem 5.4). Moreover, starting from the explicit joint distribution (1.6), we also prove that the ESD satisfies the large deviation principle (LDP) at scale $(\beta/2)n^2$ whose rate function reaches its minimum at μ_d (Theorem 5.5).

2 Deformed Verblunsky Coefficients and Reflections

In this section, we introduce the deformed Verblunsky coefficients and establish some of their relevant properties, in particular a geometric interpretation in terms of reflections. One remarkable property of the Verblunsky coefficients, as it appears in Proposition 1.7, is that they are independent under $\mathrm{CJ}_{0,\beta}^{(n)}$. As we shall see in Section 3, this does not hold anymore under $\mathrm{CJ}_{\delta,\beta}^{(n)}$. This motivated us to introduce a new set of coefficients

$(\gamma_0, \ldots, \gamma_{n-2}, \gamma_{n-1})$, called deformed Verblunsky coefficients, which are uniquely associated with a set of Verblunsky coefficients. In particular, $\gamma_k \in \mathbb{D}$ for $0 \le k \le n-2$, $\gamma_{n-1} \in \mathbb{T}$ and the map $(\gamma_0, \ldots, \gamma_{n-1}) \mapsto (\alpha_0, \ldots, \alpha_{n-1})$ is a bijection. Moreover, the characteristic polynomial at 1 can be expressed simply in terms of $(\gamma_0, \ldots, \gamma_{n-1})$.

2.1 Analytical properties

Let μ be a probability measure on the unit circle supported at n points. Keeping the notations of the introduction, we let $(\Phi_k(z))_{0 \le k \le n}$ denote the monic orthogonal polynomials associated with μ and $(\alpha_k)_{0 \le k \le n-1}$ its corresponding set of Verblunsky coefficients through Szegö's recursion formula (1.9). The functions

$$b_k(z) = \frac{\Phi_k(z)}{\Phi_k^*(z)}, \quad k \le n-1 \tag{2.1}$$

are known as the inverse Schur iterates ([34], p. 476, after Khrushchev [22], p. 273). They are analytic in a neighborhood of $\bar{\mathbb{D}}$ and meromorphic in $\mathbb{C}$. Each b_k is a finite Blashke product

$$b_k(z) = \prod_{j=1}^{k} \left(\frac{z - z_j}{1 - \bar{z}_j z} \right),$$

where $z_1, \ldots, z_k$ are the zeros of Φ_k. Let us now explain the term "inverse Schur iterate."

The Schur function is a fundamental object in the study of the orthogonal polynomials on the unit circle. Let us briefly recall its definition (see [33] or [32] for more details and proofs): if μ is a probability measure on the unit circle (supported at finitely many points or not), its Schur function $f : \mathbb{D} \to \mathbb{D}$ is defined as

$$f(z) = \frac{1}{z} \frac{F(z) - 1}{F(z) + 1}, \quad \text{where } F(z) = \int \frac{e^{i\theta} + z}{e^{i\theta} - z} \, d\mu(e^{i\theta}). \tag{2.2}$$

It is a bijection between the set of probability measures on the unit circle and analytic functions mapping $\mathbb{D}$ to $\bar{\mathbb{D}}$. The Schur algorithm (which is described in [33] or [32], p. 438) allows us to parametrize the Schur function f by a sequence of so-called Schur parameters, which are actually the Verblunsky coefficients associated with μ (Geronimus theorem). In particular, there are finitely many Verblunsky coefficients (or equivalently the measure μ is finitely supported) if and only if f is a finite Blaschke product. The name "inverse Schur iterate" [21] for b_k comes from the result (1.12) of the latter paper where b_k is identified as the Schur function corresponding to the "reversed sequence" $(-\bar{\alpha}_{k-1}, \ldots, -\bar{\alpha}_0, 1)$ (see also [34], Proposition 9.2.3).

Let us define our sequence of functions, which shall lead us to the deformed coefficients.

Definition 2.1. If μ is supported at n points and with the notation above, define $\gamma_k(z)$ for $0 \leq k \leq n - 1$, as

$$\gamma_k(z) = z - \frac{\Phi_{k+1}(z)}{\Phi_k(z)} \,. \tag{2.3}$$

From the Szegö's recursion formula (1.9) and notation (2.1), this is equivalent to

$$\gamma_k(z) = \frac{\bar{\alpha}_k}{b_k(z)} \,, \tag{2.4}$$

so that γ_k is meromorphic, with poles in $\mathbb{D}$ and zeros lying outside $\overline{\mathbb{D}}$.

The next proposition shows how the functions $\gamma_k(z)$ can be defined recursively with the help of the coefficients α_k. As a consequence, the $\gamma_k(z)$ are very closely related to a fundamental object in the theory of random matrices: the characteristic polynomial.

Proposition 2.2. For any $z \in \mathbb{C}$, $\gamma_0(z) = \bar{\alpha}_0$ and the following decomposition for $\Phi_k(z)$ holds:

$$\Phi_k(z) = \prod_{j=0}^{k-1} (z - \gamma_j(z)), \quad k = 1, \dots, n \,. \tag{2.5}$$

The $\gamma_k(z)$'s may be also defined by means of the α's through the recursion

$$\gamma_k(z) = \bar{\alpha}_k \prod_{j=0}^{k-1} \frac{1 - z\widetilde{\gamma}_j(z)}{z - \gamma_j(z)} \,, \tag{2.6}$$

$$\widetilde{\gamma}_k(z) = \overline{\gamma_k(\bar{z}^{-1})} \,. \tag{2.7}$$

Proof. The first claim is an immediate consequence of (2.3). Now, using $\Phi_k(z) = \prod_{j=0}^{k-1}(z - \gamma_j(z))$, we obtain

$$\Phi_k^*(z) = \prod_{j=0}^{k-1} (1 - z\widetilde{\gamma}_j(z)),$$

and hence (we use (2.4))

$$\gamma_k(z) = \bar{\alpha}_k \prod_{j=0}^{k-1} \frac{1 - z\widetilde{\gamma}_j(z)}{z - \gamma_j(z)} \,. \qquad \blacksquare$$

Note that when $|z| = 1$, $|\gamma_k(z)| = |\alpha_k|$. Combined with the above proposition, this leads us to introduce the following set of coefficients.

Definition 2.3. Define the coefficients $(\gamma_k)_{0 \le k \le n-1}$ by

$$\gamma_k := \gamma_k(1), \quad k = 0, \ldots, n-1. \tag{2.8}$$

We shall refer to the γ_k's as the *deformed Verblunsky coefficients*.

Proposition 2.4. The following properties hold for the deformed Verblunsky coefficients:

(a) For all $0 \le k \le n-1$, $|\gamma_k| = |\alpha_k|$, and in particular $\gamma_{n-1} \in \mathbb{T}$;

(b) $\gamma_0 = \bar{\alpha}_0$, and

$$\gamma_k = \bar{\alpha}_k e^{i\varphi_{k-1}}, \quad e^{i\varphi_{k-1}} = \prod_{j=0}^{k-1} \frac{1 - \bar{\gamma}_j}{1 - \gamma_j}, \quad (k = 1, \ldots, n-1). \tag{2.9}$$

The last term is special. Since $|\alpha_{n-1}| = 1$, we set $\alpha_{n-1} = e^{i\psi_{n-1}}$, so that

$$\gamma_{n-1} = e^{i(-\psi_{n-1} + \varphi_{n-2})} := e^{i\theta_{n-1}}; \tag{2.10}$$

(c) Let μ be the spectral measure associated with (U, e_1), $U \in U(n)$. Then $\Phi_n(z)$ is the characteristic polynomial of U, in particular,

$$\Phi_n(1) = \det(\mathrm{Id} - U) = \prod_{k=0}^{n-1} (1 - \gamma_k). \tag{2.11}$$

Proof. All the results are direct consequences of the Definition 2.3 and the formulae in Proposition 2.2 evaluated at 1. ∎

Remark 1. In [24], Killip and Stoiciu have considered variables which are the complex conjugate of our deformed Verblunsky coefficients as auxiliary variables in the study of the Prüfer phase (Lemma 2.1 in [24]). Nevertheless, the way we define them as well as the use we make of them are different.

Remark 2. The formula (2.9) shows that the γ_k's can be obtained from the α_k's recursively. Hence, starting from a spectral measure associated with a unitary matrix, one can associate with it the Verblunsky coefficients and then the deformed Verblunsky coefficients. Conversely, one can translate any property of the deformed ones into properties for the spectral measure associated with it by inverting the transformations (2.9).

Remark 3. The distribution of the characteristic polynomial of random unitary matrices evaluated at 1, through its Mellin–Fourier transform, plays a key role in the theory of random matrices, especially through its links with analytic number theory (see [26] for an account). In [6] it is proven that it can be decomposed in law into a product of independent random variables when working on the unitary and orthogonal groups endowed with the Haar measure; since we will prove in Section 3 that the γ_k's are independent under $\mathrm{CJ}_{\delta,\beta}^{(n)}$, then we can conclude that this latter result holds for any circular Jacobi ensemble.

2.2 Geometric interpretation

We give a connection between the coefficients $(\gamma_k)_{0 \le k \le n-1}$ and reflections defined below. This allows us to obtain a new decomposition of the GGT matrix associated with a measure μ supported at n points on the unit circle as a product of n elementary reflections.

Many distinct definitions of reflections on the unitary group exist, the most well known may be the Householder reflections. The transformations which will be relevant to us are the following ones.

Definition 2.5. An element r in $U(n)$ will be referred to as a reflection if $r -$ Id has rank 0 or 1.

If $v \in \mathbb{C}^n$, we denote by $\langle v|$ the linear form $w \mapsto \langle v, w \rangle$. The reflections can also be described in the following way. If e and $m \ne e$ are unit vectors of $\mathbb{C}^n$, there is a unique reflection r such that $r(e) = m$, and

$$r = \mathrm{Id} - \frac{1}{1 - \langle m, e \rangle}(m - e)\langle(m - e)|. \tag{2.12}$$

Let $F := \mathrm{span}\{e, m\}$ be the two-dimensional vector space which is spanned by the vectors e and m. It is clear that the reflection given by formula (2.12) leaves $F^\perp$ invariant. Now set

$$\gamma = \langle e, m \rangle, \quad \rho = \sqrt{1 - |\gamma|^2}, \quad e^{\mathrm{i}\varphi} = \frac{1 - \gamma}{1 - \bar{\gamma}}, \tag{2.13}$$

and let $g \in F$ be the unit vector orthogonal to e obtained by the Gram–Schmidt procedure. Then in the basis (e, g) of F, the matrix of the restriction of r is

$$\Xi(\gamma) := \begin{pmatrix} \gamma & \rho e^{\mathrm{i}\varphi} \\ \rho & -\bar{\gamma}e^{\mathrm{i}\varphi} \end{pmatrix}. \tag{2.14}$$

Conversely, for $\gamma \in \mathbb{D}$, such a matrix represents the unique reflection in $\mathbb{C}^2$ provided with its canonical basis, mapping e_1 onto $\gamma e_1 + \sqrt{1 - |\gamma|^2} e_2$. The eigenvalues of r are 1 and $-e^{i\varphi}$.

Let u be a unitary operator in $\mathbb{C}^n$ and e be a unit cyclic vector for u. We define n reflections $r_1, \ldots, r_n$ recursively as follows. Let $(\varepsilon_1, \ldots, \varepsilon_n)$ be the orthonormal basis obtained from the Gram–Schmidt procedure applied to $(e, ue, \ldots, u^{n-1}e)$.

Let r_1 be the reflection, mapping $e = \varepsilon_1$ onto $ue = u\varepsilon_1$. More generally, for $k \geq 2$ let r_k be the reflection mapping ε_k onto $r_{k-1}^{-1} r_{k-2}^{-1} \ldots r_1^{-1} u\varepsilon_k$. We will identify these reflections and establish the decomposition of u. Following the basics recalled about GGT matrices in the introduction, we note that the matrix of u in the basis $(\varepsilon_1, \ldots, \varepsilon_n)$ is the GGT matrix associated with the measure μ, i.e. the matrix $\mathcal{G}(\alpha_0, \cdots, \alpha_{n-2}, \alpha_{n-1})$, where $(\alpha_0, \cdots, \alpha_{n-2}, \alpha_{n-1})$ are the Verblunsky coefficients associated with the measure μ. We will use formula (4.1.6) of [33] or formula (1.11) of (our) Lemma 1.3 for the identification of scalar products.

Proposition 2.6. (1) For every $1 \leq k \leq n-1$, the reflection r_k leaves invariant the $n-2$-dimensional space $\mathrm{Span}\{\varepsilon_1, \ldots, \varepsilon_{k-1}, \varepsilon_{k+2}, \ldots, \varepsilon_n\}$. The reflection r_n leaves invariant $\mathrm{Span}\{\varepsilon_1, \ldots, \varepsilon_{n-1}\}$.

(2) The following decomposition holds:

$$u = r_1 \cdots r_n. \tag{2.15}$$

Proof. (1) In view of Section 2.1, it is enough to prove that for $j \notin \{k, k+1\}$, the vectors ε_j and $r_k \varepsilon_k$ are orthogonal.

For $k = 1$, $\langle \varepsilon_j, r_1 \varepsilon_1 \rangle = \langle \varepsilon_j, u\varepsilon_1 \rangle = 0$ as soon as $j \geq 3$ from (1.11).

Assume that for every $\ell \leq k-1$, the reflection r_ℓ leaves invariant $\mathrm{Span}\{\varepsilon_1, \ldots, \varepsilon_{\ell-1}, \varepsilon_{\ell+2}, \ldots, \varepsilon_n\}$. For every $j = 1, \ldots, n$, we have

$$\langle \varepsilon_j, r_k \varepsilon_k \rangle = \left\langle \varepsilon_j, r_{k-1}^{-1} r_{k-2}^{-1} \ldots r_1^{-1} u\varepsilon_k \right\rangle = \langle r_1 \cdots r_{k-1} \varepsilon_j, u\varepsilon_k \rangle. \tag{2.16}$$

For $j \geq k+2$, by assumption, the reflections $r_1, \ldots, r_{k-1}$ leave invariant ε_j, so that the above expression reduces to $\langle \varepsilon_j, u\varepsilon_k \rangle$, which is 0 again by (1.11).

For $j = k-1$, we have $r_1 \cdots r_{k-1} \varepsilon_{k-1} = u\varepsilon_{k-1}$ by definition of r_{k-1}, so that (2.16) gives $\langle \varepsilon_{k-1}, r_k \varepsilon_k \rangle = \langle u\varepsilon_{k-1}, u\varepsilon_k \rangle$, which is 0 since u is unitary.

For $j < k-1$, by assumption, the reflections $r_{j+1}, \ldots, r_{k-1}$ leave invariant ε_j, so that the right-hand side of (2.16) reduces to $\langle r_1 \cdots r_j \varepsilon_j, u\varepsilon_k \rangle$. By definition of r_j, it is $\langle u\varepsilon_j, u\varepsilon_k \rangle$ which is 0.

(2) For k fixed, it is clear from (1) that $r_1 \cdots r_n \varepsilon_k = r_1 \cdots r_k \varepsilon_k$, which is $u \varepsilon_k$ by definition of r_k. ∎

Proposition 2.7. For $k = 1, \ldots, n-1$, the matrix of the restriction of r_k to the basis $(\varepsilon_k, \varepsilon_{k+1})$ is $\Xi(\gamma_{k-1})$ as defined in (2.14). In particular

$$\langle \varepsilon_k, r_k \varepsilon_k \rangle = \gamma_{k-1} \, . \tag{2.17}$$

The restriction of r_n to $\mathbb{C}\varepsilon_n$ is the multiplication by γ_{n-1}.

Proof. Note that for every $k \leq n-1$

$$\langle \varepsilon_{k+1}, r_k \varepsilon_k \rangle = \langle r_1 \cdots r_{k-1} \varepsilon_{k+1}, u \varepsilon_k \rangle = \langle \varepsilon_{k+1}, u \varepsilon_k \rangle = \rho_{k-1} \, . \tag{2.18}$$

Since r_k is a reflection acting on the subspace $\mathrm{Span}\{\varepsilon_k, \varepsilon_{k+1}\}$, identities (2.18) and (2.17) entail that the matrix representing r_k in the basis $(\varepsilon_1, \ldots, \varepsilon_n)$ is precisely $\Xi(\gamma_{k-1})$ (see (2.14)). It is then enough to prove (2.17).

For $k = 1$ it is immediate that

$$\langle \varepsilon_1, r_1 \varepsilon_1 \rangle = \langle \varepsilon_1, u \varepsilon_1 \rangle = \bar{\alpha}_0 = \gamma_0 \, .$$

Let us proceed by induction. For $j \geq 1$ set $q_j := \langle \varepsilon_j, r_j \varepsilon_j \rangle$. Assume that $q_j = \gamma_{j-1}$ for $j \leq k$. We have $q_{k+1} = \langle \varepsilon_{k+1}, r_{k+1} \varepsilon_{k+1} \rangle = \langle r_1 \ldots r_k \varepsilon_{k+1}, u \varepsilon_{k+1} \rangle$. Equation (2.12) implies

$$r_k \varepsilon_{k+1} = \varepsilon_{k+1} - \frac{1}{1 - \bar{\gamma}_{k-1}} (r_k \varepsilon_k - \varepsilon_k) \langle r_k \varepsilon_k, \varepsilon_{k+1} \rangle, \tag{2.19}$$

and since $r_j \varepsilon_\ell = \varepsilon_\ell$ for $\ell \geq j+2$, we get

$$r_1 \ldots r_k \varepsilon_{k+1} = \varepsilon_{k+1} - \frac{1}{1 - \bar{\gamma}_{k-1}} (r_1 \ldots r_k \varepsilon_k - r_1 \ldots r_{k-1} \varepsilon_k) \langle u \varepsilon_k, \varepsilon_{k+1} \rangle.$$

Now, it is known that $\langle u \varepsilon_k, \varepsilon_{k+1} \rangle = \overline{\langle \varepsilon_{k+1}, u \varepsilon_k \rangle} = \rho_k$. If we set $v_1 = \varepsilon_1$,

$$v_j = r_1 \ldots r_{j-1} \varepsilon_j, \quad a_j = \frac{\rho_{j-1}}{1 - \bar{\gamma}_{j-1}} \, , \quad w_{j+1} = \varepsilon_{j+1} - a_j u \varepsilon_j,$$

we get the recursion

$$v_{j+1} = a_j v_j + w_{j+1} \, , \ (j \leq k) \, , \tag{2.20}$$

which we solve in

$$v_{k+1} = \left(\prod_{j=1}^{k} a_j \right) \varepsilon_1 + \sum_{\ell=2}^{k+1} \left(\prod_{j=\ell}^{k} a_j \right) w_\ell. \tag{2.21}$$

Taking the scalar product with $u\varepsilon_{k+1}$ yields

$$q_{k+1} = \left(\prod_{j=1}^{k} \bar{a}_j\right) \langle \varepsilon_1, u\varepsilon_{k+1}\rangle + \sum_{\ell=2}^{k+1} \left(\prod_{j=\ell}^{k} \bar{a}_j\right) \langle w_\ell, u\varepsilon_{k+1}\rangle.$$

But $\langle w_\ell, u\varepsilon_{k+1}\rangle = \langle \varepsilon_\ell, u\varepsilon_{k+1}\rangle - \bar{a}_{\ell-1}\langle u\varepsilon_{\ell-1}, u\varepsilon_{k+1}\rangle$, and since $\ell \leq k+1$, we have

$$\langle w_\ell, u\varepsilon_{k+1}\rangle = \langle \varepsilon_\ell, u\varepsilon_{k+1}\rangle = -\bar{\alpha}_k \alpha_{\ell-2} \prod_{m=\ell-1}^{k-1} \rho_m,$$

which yields (with $\alpha_{-1} = -1$)

$$-\frac{q_{k+1}}{\bar{\alpha}_k} = \sum_{\ell=1}^{k+1} \left(\prod_{j=\ell}^{k} \bar{a}_j\right) \alpha_{\ell-2} \prod_{m=\ell-1}^{k-1} \rho_m$$

$$= \sum_{\ell=1}^{k+1} \prod_{m=\ell-1}^{k-1} \rho_m^2 \prod_{j=0}^{\ell-3}(1-\bar{\gamma}_j)\frac{\bar{\gamma}_{\ell-2}}{\prod_{s=0}^{k-1}(1-\gamma_s)}(1-\gamma_{\ell-2})$$

$$= \frac{1}{\prod_{s=0}^{k-1}(1-\gamma_s)} \sum_{\ell=1}^{k+1} \left[\prod_{m=\ell-2}^{k-1} \rho_m^2 \prod_{j=0}^{\ell-3}(1-\bar{\gamma}_j) - \prod_{m=\ell-1}^{k-1} \rho_m^2 \prod_{j=0}^{\ell-2}(1-\bar{\gamma}_j)\right]$$

$$= -\prod_{s=0}^{k-1} \frac{(1-\bar{\gamma}_s)}{(1-\gamma_s)},$$

and eventually $q_{k+1} = \gamma_k$. ∎

Now, we can summarize the above results in the following theorem.

Theorem 2.8. Let $u \in U(n)$ and e a cyclic vector for u. Let μ be the spectral measure of the pair (u, e), and $(\alpha_0, \ldots, \alpha_{n-1})$ its Verblunsky coefficients. Let $(\varepsilon_1, \ldots, \varepsilon_n)$ be the orthonormal basis obtained from the Gram–Schmidt procedure applied to $(e, ue, \ldots, u^{n-1}e)$. Then, u can be decomposed as a product of n reflections $(r_k)_{1 \leq k \leq n}$:

$$u = r_1 \ldots r_n \tag{2.22}$$

where r_1 is the reflection mapping ε_1 onto $u\varepsilon_1$ and by induction for each $2 \leq k \leq n$, r_k maps ε_k onto $r_{k-1}^{-1}r_{k-2}^{-1}\ldots r_1^{-1}u\varepsilon_k$.

This decomposition can also be restated in terms of the GGT matrix:

$$\mathcal{G}(\alpha_0, \cdots, \alpha_{n-1}) = \Xi^{(0)}(\gamma_0)\Xi^{(1)}(\gamma_1)\ldots\Xi^{(n-1)}(\gamma_{n-1}), \tag{2.23}$$

where for $0 \leq k \leq n-2$, the matrix $\Xi^{(k)}$ is given by

$$\Xi^{(k-1)}(\gamma_{k-1}) = \mathrm{Id}_{k-1} \oplus \Xi(\gamma_{k-1}) \oplus \mathrm{Id}_{n-k-1}, \tag{2.24}$$

with $\Xi(\gamma)$ defined in (2.14). For $k = n - 1$,

$$\Xi^{(n-1)}(\gamma_{n-1}) = \mathrm{Id}_{n-1} \oplus (\gamma_{n-1}). \tag{2.25}$$

3   Deformed Verblunsky Coefficients and independence

We now use the point of view of sampling (or change of probability measure) to compute the distribution of the deformed Verblunsky coefficients under $\mathrm{CJ}_{\delta,\beta}^{(n)}$. Let us first remark that if the α_k's are independent and with rotational invariant distribution, then from (2.9)

$$(\alpha_0, \ldots, \alpha_{n-1}) \overset{\text{law}}{=} (\gamma_0, \ldots, \gamma_{n-1}). \tag{3.1}$$

This is the case under $\mathrm{CJ}_{0,\beta}^{(n)}$.

We first prove that when $\delta \neq 0$ the Verblunsky coefficients are not independent anymore by studying the simple case $n = 2, \beta = 2$, and then we compute the distribution of $(\gamma_0, \ldots, \gamma_{n-1})$ under $\mathrm{CJ}_{\delta,\beta}^{(n)}$. We then show that under this distribution, the weights of the measure associated with the Verblunsky coefficients $(\alpha_0, \ldots, \alpha_{n-1})$ are independent from the points at which the measure is supported and follow a Dirichlet distribution.

Let $\delta \in \mathbb{C}$ such that $\mathfrak{Re}\,\delta > -1/2$. The formula

$$\lambda^{(\delta)}(\zeta) = \frac{\Gamma(1 + \delta)\Gamma(1 + \bar{\delta})}{\Gamma(1 + \delta + \bar{\delta})}(1 - \zeta)^{\bar{\delta}}(1 - \bar{\zeta})^{\delta} , \quad \zeta \in \mathbb{T} \tag{3.2}$$

defines a probability density with respect to the Haar measure on $\mathbb{T}$, which is discontinuous at 1 when $\mathfrak{Im}\,\delta \neq 0$ (see [7]).

When $\beta = 2$ and $\delta \neq 0$, the Verblunsky coefficients are dependent. Indeed, let $M \in U(2)$ with Verblunsky coefficients α_0 and α_1. Then

$$\det(\mathrm{Id} - M) = [1 - \bar{\alpha}_0 - \bar{\alpha}_1(1 - \alpha_0)],$$

with $|\alpha_0| < 1$ and $|\alpha_1| = 1$. Under $\mathrm{CJ}_{0,2}^{(2)}$, the variables α_0 and α_1 are independent and uniformly distributed on $\mathbb{D}$ and $\mathbb{T}$ respectively (see [23] or Proposition 1.7). The $\mathrm{CJ}_{\delta,2}^{(2)}$ is a $\det_\delta$ sampling of $\mathrm{CJ}_{0,2}^{(n)}$ (see the Introduction for the definition and notation for the $\det_\delta$-sampling). So, the joint density of (α, φ) on $\mathbb{D} \times \mathbb{T}$ is proportional to

$$\begin{aligned} f(\alpha_0, \alpha_1) &= [1 - \bar{\alpha}_0 - \bar{\alpha}_1(1 - \alpha_0)]^{\bar{\delta}}[1 - \alpha_0 - \alpha_1(1 - \bar{\alpha}_0)]^{\delta} \\ &= (1 - \bar{\alpha}_0)^{\bar{\delta}}(1 - \alpha_0)^{\delta}[1 - \gamma\bar{\alpha}_1]^{\bar{\delta}}[1 - \bar{\gamma}\alpha_1]^{\delta} , \end{aligned}$$

where

$$\gamma = \frac{1 - \alpha}{1 - \bar{\alpha}}.$$

It is then clear that the conditional density of α_1 given α_0 is

$$\alpha_1 \mapsto \lambda^{(\delta)}(\gamma \bar{\alpha_1}), \quad \alpha_1 \in \mathbb{T},$$

and this last quantity does not depend on α_0 (i.e. on γ) if and only if $\delta = 0$. Otherwise, the Verblunsky coefficients α_0 and α_1 are dependent.

The next theorem illustrates our interest in the deformed Verblunsky coefficients: under $\mathrm{CJ}_{\delta,\beta}^{(n)}$, they are independent. For the proof of this theorem, we need the following lemma which will also be useful when we study limit theorems.

Lemma 3.1.   Let $s, t, \ell \in \mathbb{C}$ such that $\mathfrak{Re}(s + \ell + 1) > 0, \mathfrak{Re}(t + \ell + 1) > 0$. Then, the following identity holds:

$$\int_{\mathbb{D}} (1 - |z|^2)^{\ell-1}(1 - z)^s(1 - \bar{z})^t d^2z = \frac{\pi \Gamma(\ell)\Gamma(\ell + 1 + s + t)}{\Gamma(\ell + 1 + s)\Gamma(\ell + 1 + t)}. \tag{3.3}$$

Proof.   A Taylor expansion yields

$$(1 - z)^s(1 - \bar{z})^t = \sum_{m,n \geq 0} \rho^{m+n} \frac{(-s)_n(-t)_m}{n!m!} e^{i(m-n)\theta},$$

with $z = \rho e^{i\theta}$ and $0 \leq \rho < 1$. We obtain by integrating

$$\int_{\mathbb{D}} (1 - |z|^2)^{\ell-1}(1 - z)^s(1 - \bar{z})^t d^2z = 2\pi \sum_{n \geq 0} \frac{(-s)_n(-t)_n}{n!n!} \int_0^1 (1 - \rho^2)^{\ell-1} \rho^{2n+1} d\rho$$

$$= \pi \sum_{n \geq 0} \frac{(-s)_n(-t)_n}{n!} \frac{(\ell - 1)!}{(n + \ell)!} = \frac{\pi}{\ell} {}_2F_1(-s, -t; \ell + 1; 1),$$

where ${}_2F_1$ is the classical hypergeometric function (see [3]) and an application of Gauss formula (see [3]) shows that the last expression is exactly the right-hand side of (3.3).   ∎

Theorem 3.2.   Let $\delta \in \mathbb{C}$ with $\mathfrak{Re}\,\delta > -1/2$ and $\beta > 0$. Set $\beta' = \beta/2$. Under $\mathrm{CJ}_{\delta,\beta}^{(n)}$, the distribution of $(\gamma_0, \dots, \gamma_{n-1})$, denoted hereafter $\eta_{\delta,\beta}^{(n)}$, is the following:

   (1) the variables $\gamma_0, \dots, \gamma_{n-2}, \gamma_{n-1}$ are independent;
   (2) for $k = 0, \dots, n - 2$ the density of γ_k with respect to the Lebesgue measure d^2z on $\mathbb{C}$ is

$$c_{k,n}(\delta)(1 - |z|^2)^{\beta'(n-k-1)-1}(1 - z)^{\bar{\delta}}(1 - \bar{z})^{\delta} 1_{\mathbb{D}}(z),$$

where

$$c_{k,n}(\delta) = \frac{\Gamma(\beta'(n-k-1)+1+\delta)\Gamma(\beta'(n-k-1)+1+\bar{\delta})}{\pi\,\Gamma(\beta'(n-k-1))\Gamma(\beta'(n-k-1)+1+\delta+\bar{\delta})} \; ; \qquad (3.4)$$

(3) the density of γ_{n-1} with respect to the Haar measure on $\mathbb{T}$ is $\lambda^{(\delta)}$ (given by (3.2)).

Proof. The distribution of the α's in the β-circular unitary ensemble is $\eta_{0,\beta}^{(n)}$. More precisely, as seen in Definition 1.4 they are independent and if

$$\alpha_k = r_k e^{i\psi_k} \ \ (0 \le k \le n-2), \quad \alpha_{n-1} = e^{i\psi_{n-1}} \,,$$

then r_k and ψ_k are independent, ψ_k is uniformly distributed and r_k^2 has the Beta$(1, \beta'(n-k-1))$ distribution. Moreover, α_{n-1} is uniformly distributed on $\mathbb{T}$. From (2.11), the sampling factor is

$$\det(\mathrm{Id} - U)^{\bar{\delta}}\det(\mathrm{Id} - \bar{U})^{\delta} = (1 - \gamma_{n-1})^{\bar{\delta}}(1 - \bar{\gamma}_{n-1})^{\delta} \prod_{k=0}^{n-2}(1 - \gamma_k)^{\bar{\delta}}(1 - \bar{\gamma}_k)^{\delta} \,,$$

so that, under $\mathrm{CJ}_{\delta,\beta}^{(n)}$, the density of $(r_0, \ldots, r_{n-2}, \psi_0, \ldots, \psi_{n-1})$ is proportional to

$$\lambda^{(\delta)}(\gamma_{n-1}) \prod_{k=0}^{n-2} \left(1 - r_k^2\right)^{\beta'(n-1-k)-1} r_k (1 - \gamma_k)^{\bar{\delta}}(1 - \bar{\gamma}_k)^{\delta} \mathbf{1}_{(0,1)}(r_k) \,,$$

with

$$\gamma_k = r_k e^{i\theta_k} \ \ (0 \le k \le n-2), \quad \gamma_{n-1} = e^{i\theta_{n-1}} \,.$$

Thanks to the relations (2.9) and (2.10), the Jacobian matrix of the mapping

$$(r_0, \ldots, r_{n-2}, \psi_0, \ldots, \psi_{n-1}) \to (r_0, \ldots, r_{n-2}, \theta_0, \ldots, \theta_{n-1})$$

is lower triangular with diagonal elements ± 1, so that under $\mathrm{CJ}_{\delta,\beta}^{(n)}$ the density of $(r_0, \ldots, r_{n-2}, \theta_0, \ldots, \theta_{n-1})$ is proportional to

$$\lambda^{(\delta)}(\gamma_{n-1}) \prod_{k=0}^{n-2} \left(1 - r_k^2\right)^{\beta'(n-1-k)-1} r_k (1 - \gamma_k)^{\bar{\delta}}(1 - \bar{\gamma}_k)^{\delta} \mathbf{1}_{(0,1)}(r_k) \,, \qquad (3.5)$$

which proves the required independence and the expression of the distributions, up to the determination of $c_{k,n}(\delta)$. This quantity is obtained by taking $\ell = \beta'(n-k-1), s = \bar{\delta}, t = \delta$ in (3.3), which gives (3.4) and completes the proof of the theorem. ∎

Starting with a set of deformed Verblunsky coefficients, with distribution $\eta_{\delta,\beta}^{(n)}$, we obtain the coefficients $(\alpha_0, \ldots, \alpha_{n-1})$ by inverting formula (2.11). These are the coordinates

of some probability measure μ supported at n points on the unit circle

$$\mu = \sum_{k=1}^{n} \pi_k \delta_{e^{i\theta_k}},$$

with $\pi_k > 0$ and $\sum_{k=1}^{n} \pi_k = 1$. The next theorem gives the distribution induced on the vector $(\pi_1, \ldots, \pi_n, \theta_1, \ldots, \theta_n)$ by $(\gamma_0, \ldots, \gamma_{n-1})$.

Theorem 3.3. The following formulae express the same measure on the manifold of probability distribution on $\mathbb{T}$ supported at n points:

$$K_{\delta,\beta}^{(n)} |\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^\beta \prod_{k=1}^{n} (1 - e^{-i\theta_k})^\delta (1 - e^{i\theta_k})^{\bar\delta} \prod_{k=1}^{n} \pi_k^{\beta'-1} d\theta_1 \cdots d\theta_n d\pi_1 \cdots d\pi_{n-1}$$

in the (θ, π) coordinates and

$$K_{\delta,\beta}^{(n)} \prod_{k=0}^{n-2} (1 - |\gamma_k|^2)^{\beta'(n-k-1)-1} \prod_{k=0}^{n-1} (1 - \gamma_k)^{\bar\delta} (1 - \bar\gamma_k)^\delta d^2\gamma_0 \cdots d^2\gamma_{n-2} d\phi$$

in terms of the deformed Verblunsky coefficients, with $\gamma_{n-1} = e^{i\phi}$. Here, $K_{\delta,\beta}^{(n)}$ is a constant:

$$K_{\delta,\beta}^{(n)} = \frac{\Gamma(1+\delta)\Gamma(1+\bar\delta)}{2^{n-1}\pi\,\Gamma(1+\delta+\bar\delta)} \prod_{k=0}^{n-2} c_{k,n}(\delta),$$

with $c_{k,n}(\delta)$ given in Theorem 3.2. Consequently, if $(\gamma_0, \ldots, \gamma_{n-1})$ is $\eta_{\delta,\beta}^{(n)}$ distributed, then $(\pi_1, \ldots, \pi_n)$ and $(\theta_1, \ldots, \theta_n)$ are independent; the vector of weights $(\pi_1, \ldots, \pi_n)$ follows the $\mathrm{Dir}_n(\beta')$ distribution and the vector $(\theta_1, \ldots, \theta_n)$ has the density $h_{\delta,\beta}^{(n)}$.

Proof. In the course of this proof, we shall adopt the following point of view. Starting with a measure supported at n points on the unit circle, we associate with it its Verblunsky coefficients $(\alpha_0, \ldots, \alpha_{n-1})$ and then the corresponding GGT matrix which we note G for simplicity. Then e_1 is a cyclic vector for G and μ is the spectral measure of (G, e_1). Conversely, starting with the set of deformed Verblunsky coefficients with $\eta_{\delta,\beta}$ distribution, we construct the coefficients $(\alpha_0, \ldots, \alpha_{n-1})$ with the transformations (2.9), then the GGT matrix associated with it and finally μ the spectral measure associated with this matrix and e_1.

We use the following well-known identity (see [33] or [23] Lemma 4.1):

$$|\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^2 \prod_{k=1}^{n} \pi_k = \prod_{k=0}^{n-2} (1 - |\alpha_k|^2)^{n-k-1}. \tag{3.6}$$

Since $|\gamma_k| = |\alpha_k|$, we can also write

$$|\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^2 \prod_{k=1}^{n} \pi_k = \prod_{k=0}^{n-2} (1 - |\gamma_k|^2)^{n-k-1}. \tag{3.7}$$

Moreover, from (2.11),

$$\det(\mathrm{Id} - G) = \prod_{k=1}^{n} (1 - e^{i\theta_k}) = \prod_{k=0}^{n-1} (1 - \gamma_k). \tag{3.8}$$

In our setting, π_i is modulus squared of the first component of the ith eigenvector of the matrix G. Now, define

$$q_k^2 = \pi_k, \quad \text{for } k = 1, \ldots, n.$$

It is known (see for example Forrester [12], Chapter 2 and [13] Theorem 2) that the Jacobian of the map $(\alpha_0, \ldots, \alpha_{n-1}) \mapsto (\theta_1, \ldots, \theta_n, q_1, \ldots, q_{n-1})$ is given by

$$\frac{\prod_{k=0}^{n-2} (1 - |\alpha_k|^2)}{q_n \prod_{k=1}^{n} q_k}.$$

Moreover, the map $(\gamma_0, \ldots, \gamma_{n-1}) \mapsto (\alpha_0, \ldots, \alpha_{n-1})$ is invertible and its Jacobian is 1, as already seen. The result now follows from simple integral manipulations combined with the identities (3.7) and (3.8).   ∎

4   Matrix Models for the Circular Jacobi Ensemble

The results of the previous sections can now be used to propose some simple matrix models for the circular Jacobi ensemble. There are mainly two ways to generate matrix models for a given spectral measure encoded by its Verblunsky coefficients.

   The AGR decomposition. If $U = \Theta^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1) \cdots \Theta^{(n-1)}(\alpha_{n-1})$ (the Θ_k's are defined in the introduction), the Verblunsky coefficients for the spectral measure associated with (U, e_1) are precisely $(\alpha_0, \ldots, \alpha_{n-1})$ (see [2] or [31], Section 10). Therefore, taking independent α_k's with law $\eta_{0,\beta}^{(n)}$, the density of the eigenvalues of

$$U = \Theta^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1) \cdots \Theta^{(n-1)}(\alpha_{n-1})$$

is proportional to $|\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^\beta$. The matrix U obtained above is the GGT matrix associated with the α_k's. It is in the Hessenberg form.

   The CMV form. Set

$$\begin{cases} \mathcal{L} = \Theta^{(0)}(\alpha_0)\Theta^{(2)}(\alpha_2) \ldots, \\ \mathcal{M} = \Theta^{(1)}(\alpha_1)\Theta^{(3)}(\alpha_3) \ldots. \end{cases}$$

Cantero, Moral, and Velazquez [8] proved that the Verblunsky coefficients associated with $(\mathcal{L}\mathcal{M}, e_1)$ are precisely $(\alpha_0, \ldots, \alpha_{n-1})$. Therefore, taking as previously independent α_k's with distribution $\eta_{0,\beta}^{(n)}$, the density of the eigenvalues of the spectral law of $\mathcal{L}\mathcal{M}$ is proportional to $|\Delta(e^{i\theta_1}, \ldots, e^{i\theta_n})|^\beta$ [23]. This matrix model is very sparse: it is pentadiagonal.

We now propose a matrix model for the circular Jacobi ensemble: it is reminiscent of the AGR factorization with the noticeable difference that it is based on the deformed Verblunsky coefficients and actual reflections as defined in Section 2.

Theorem 4.1. If $(\gamma_0, \ldots, \gamma_{n-1})$ is $\eta_{\delta,\beta}^{(n)}$ distributed, then with the notation of (2.24) and (2.25),

$$\Xi^{(0)}(\gamma_0)\,\Xi^{(1)}(\gamma_1)\cdots\Xi^{(n-1)}(\gamma_{n-1})$$

is a matrix model for the circular Jacobi ensemble, i.e. the density of the eigenvalues is $h_{\delta,\beta}^{(n)}$ (see (1.6)).

Proof. We know from Theorem 2.8 that

$$\mathcal{G}(\alpha_0, \ldots, \alpha_{n-1}) = \Xi^{(0)}(\gamma_0)\,\Xi^{(1)}(\gamma_1)\cdots\Xi^{(n-1)}(\gamma_{n-1}). \tag{4.1}$$

We also proved in Theorem 3.3 that the set of deformed Verblunsky coefficients with probability distribution $\eta_{\delta,\beta}^{(n)}$ induces a distribution on the eigenvalues of the GGT matrix $\mathcal{G}(\alpha_0, \ldots, \alpha_{n-1})$, which has exactly the density $h_{\delta,\beta}$. This completes the proof of the theorem. $\blacksquare$

Remark 3. We now explain the CMV form obtained by Killip and Nenciu in [23]. Cantero, Moral, and Velazquez [8] introduced the basis $\chi_0, \ldots, \chi_{n-1}$ obtained by orthogonalizing the sequence $1, z, z^{-1}, \ldots$. They prove that in this basis the matrix is pentadiagonal. We name this matrix $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1})$. It turns out that there exists a unitary P such that

$$P\mathcal{G}(\alpha_0, \ldots, \alpha_{n-1})P^\star = \mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}) \,, \ \ P\varphi_0 = \chi_0.$$

The two pairs $(\mathcal{G}(\alpha_0, \ldots, \alpha_{n-1}), \varphi_0)$ and $(\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1}), \chi_0)$ are equivalent, they admit the α_k's as Verblunsky coefficients and have the same spectral measure. We conclude that if we start with the γ_k's distributed as $\eta_{\delta,\beta}^{(n)}$ and build the α_k's by inverting the transformation (2.9), then $\mathcal{C}(\alpha_0, \ldots, \alpha_{n-1})$ will be a matrix model for the circular Jacobi ensemble. But, we do not know how to construct the CMV matrix from the γ_k's directly. We saw at the beginning of this section that Cantero et al. introduced the matrices $\mathcal{L}$ and $\mathcal{M}$ as direct

product of small blocks $\Theta^{(k)}(\alpha_k)$ and obtained $\mathcal{C}$ as $\mathcal{C} = \mathcal{L}\mathcal{M}$. It would be interesting to have an analog construction based on the independent γ_k's.

Theorem 2.8, which is a deterministic result, has also the following consequence.

Proposition 4.2. Let $(\alpha_0, \ldots, \alpha_{n-2}, \alpha_{n-1}) \in \mathbb{D}^{n-1} \times \mathbb{T}$ be independent random variables with rotationally invariant distribution. Then

$$\Theta^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1)\cdots\Theta^{(n-1)}(\alpha_{n-1}) \overset{\text{law}}{=} \Xi^{(0)}(\alpha_0)\Xi^{(1)}(\alpha_1)\cdots\Xi^{(n-1)}(\alpha_{n-1}).$$

Proof. We give two proofs of this result. The first one is a consequence of Theorem 2.8 from which we know that

$$\Theta^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1)\cdots\Theta^{(n-1)}(\alpha_{n-1}) = \Xi^{(0)}(\gamma_0)\Xi^{(1)}(\gamma_1)\cdots\Xi^{(n-1)}(\gamma_{n-1}),$$

and the remark at the beginning of Section 3.

For the second proof, we proceed by induction on n. For $n = 1$ the result is obvious. Suppose the result holds at rank $n-1$: thanks to the recurrence hypothesis

$$\Xi^{(0)}(\alpha_0)\Xi^{(1)}(\alpha_1)\cdots\Xi^{(n-1)}(\alpha_{n-1}) \overset{\text{law}}{=} \Xi^{(0)}(\alpha_0)\Theta^{(1)}(\alpha_1)\cdots\Theta^{(n-1)}(\alpha_{n-1}).$$

Let $e^{i\phi_0} = \frac{1-\overline{\alpha_0}}{1-\alpha_0}$. An elementary calculation gives

$$\begin{aligned}
\Xi^{(0)}(\alpha_0)&\Theta^{(1)}(\alpha_1)\cdots\Theta^{(n-2)}(\alpha_{n-2})\Theta^{(n-1)}(\alpha_{n-1}) \\
&= \Theta^{(0)}(\alpha_0)\Theta^{(1)}(e^{-i\phi_0}\alpha_1)\cdots\Theta^{(n-2)}(e^{-i\phi_0}\alpha_{n-1})\Theta^{(n-1)}(e^{-i\phi_0}\alpha_{n-1}).
\end{aligned}$$

As the α_k's are independent with law invariant by rotation:

$$(\alpha_0, e^{-i\phi_0}\alpha_1, \ldots, e^{-i\phi_0}\alpha_{n-2}, e^{i\phi_0}\alpha_{n-1}) \overset{\text{law}}{=} (\alpha_0, \alpha_1, \ldots, \alpha_{n-2}, \alpha_{n-1}),$$

which completes the proof. ∎

Now that we have a matrix model for the circular Jacobi ensemble, we can study the characteristic polynomial for such matrices; the key formula will be (2.11).

Proposition 4.3. Let U be a unitary matrix of size n and let $Z_n = \det(\mathrm{Id} - U)$ be its characteristic polynomial evaluated at 1. Then, in the circular Jacobi ensemble, Z_n can

be written as a product of n independent complex random variables:

$$Z_n = \prod_{k=0}^{n-1} (1 - \gamma_k),$$

where the laws of the γ_k's are given in Theorem 3.2. Consequently, for any $s, t \in \mathbb{C}$, with $\mathfrak{Re}(t) > -\frac{1}{2}$, the Mellin–Fourier transform of Z_n is

$$\mathbb{E}[|Z_n|^t e^{\mathrm{is\,arg}\,Z_n}] = \prod_{k=0}^{n-1} \frac{\Gamma(\beta'k + 1 + \delta)\Gamma(\beta'k + 1 + \bar{\delta})\Gamma(\beta'k + 1 + \delta + \bar{\delta} + t)}{\Gamma(\beta'k + 1 + \delta + \bar{\delta})\Gamma\left(\beta'k + 1 + \delta + \frac{t-s}{2}\right)\Gamma\left(\beta'k + 1 + \bar{\delta} + \frac{t+s}{2}\right)}. \quad (4.2)$$

Proof. The first part is an easy consequence of (2.11) and Theorem 3.2. To prove the second part, we note that if $X_k = (1 - \gamma_k)$, then $|X_k|^t e^{\mathrm{is\,arg}\,X_k} = (1 - \gamma_k)^a (1 - \bar{\gamma}_k)^b$, where $a = (t+s)/2$ and $b = (t-s)/2$. Consequently, by independence of the γ_k's, we obtain

$$\mathbb{E}[|Z_n|^t e^{\mathrm{is\,arg}\,Z_n}] = \prod_{k=0}^{n-1} \mathbb{E}[(1 - \gamma_k)^a (1 - \bar{\gamma}_k)^b],$$

and formula (4.2) then easily follows from Lemma 3.1.   ∎

5 Limiting Spectral Measure and Large Deviations

In (1.7) we defined the spectral measure, which is a central tool for the study of our circular ensembles. Let us rewrite this measure on $\mathbb{T}$ as

$$\mu_{\mathrm{sp}}^{(n)} := \sum_{k=1}^{n} \pi_k^{(n)} \delta_{e^{\mathrm{i}\theta_k^{(n)}}}, \quad (5.1)$$

where we put a superscript $^{(n)}$ to stress on the dependency on n. Besides, in classical random matrix theory, many authors are mainly interested in the empirical spectral distribution (ESD) defined by

$$\mu_{\mathrm{esd}}^{(n)} = \frac{1}{n} \sum_{k=1}^{n} \delta_{e^{\mathrm{i}\theta_k^{(n)}}}. \quad (5.2)$$

We are concerned with their asymptotics under $\mathrm{CJ}_{\delta,\beta}^{(n)}$ when $n \to \infty$ with

$$\delta = \delta(n) = \beta' n \mathrm{d},$$

where $\mathfrak{Re}\,\mathrm{d} \geq 0$ (and as usual $\beta' = \beta/2$). In this framework the variables $(\theta_1^{(n)}, \ldots, \theta_n^{(n)})$ and $(\pi_1^{(n)}, \ldots, \pi_n^{(n)})$ are distributed as in Theorem 3.3. It is well known that in the CUE (i.e. for $\mathrm{d} = 0$ and $\beta' = 1$), the sequence $(\mu_{\mathrm{esd}}^{(n)})$ converges weakly in probability to the

uniform measure on $\mathbb{T}$ and that it satisfies the large deviation principle at scale n^2 ([16], Chapter 5).

In this section, we prove that in our model, both sequences $(\mu_{\mathrm{sp}}^{(n)})$ and $(\mu_{\mathrm{esd}}^{(n)})$ converge weakly in probability to the same limit measure supported by an arc of the unit circle, and that $(\mu_{\mathrm{esd}}^{(n)})$ satisfies the large deviation principle at scale n^2. In the next two subsections, we first recall some definitions (which may be skipped by probabilist readers) and then state the main results, which are Theorems 5.1, 5.4, and 5.5.

5.1   Weak convergence of measures

Let $\mathcal{M}_1(\mathbb{T})$ be the space of probability measures on $\mathbb{T}$. The weak topology on $\mathcal{M}_1(\mathbb{T})$ is defined from the duality with the space of continuous functions on $\mathbb{T}$; that is, ν_n weakly converges to ν means $\int f d\nu_n \to \int f d\nu$ for every continuous function f. Since $\mathbb{T}$ is compact, it is equivalent to the convergence of moments. If we define the distribution function of $\nu \in \mathcal{M}_1(\mathbb{T})$ as $F_\nu(t) = \nu(\{e^{\mathrm{i}\theta}; \theta \in [0, t)\})$, then $\nu_n \to \nu$ weakly if and only if

$$F_{\nu_n}(t) \to F_\nu(t)$$

for all $t \in [0, 2\pi)$ at which $F_\nu(t)$ is continuous. The Lévy distance d_L (whose precise definition is not needed here, see [9], Theorem D8) is compatible with this topology and makes $\mathcal{M}_1(\mathbb{T})$ a compact metric space. It should be noted that for all pair of elements of $\mathcal{M}_1(\mathbb{T})$

$$d_L(\mu, \nu) \leq \sup_t |F_\mu(t) - F_\nu(t)| . \tag{5.3}$$

In the following, we consider random probability measures (ν_n) and the weak convergence in probability of (ν_n) to a deterministic $\nu \in \mathcal{M}_1(\mathbb{T})$. This means that for every $n \geq 1$ we have a probability space $(\Omega_n, \mathcal{F}_n, Q_n)$, an $\mathcal{M}_1(\mathbb{T})$-valued random variable ν_n, and that $d(\nu_n, \nu) \to 0$ in probability, where d is any distance compatible with the weak topology (e.g. the Lévy distance d_L). In other words,

$$\lim_n \nu_n = \nu \ (\text{in probability}) \Leftrightarrow \lim_n d(\nu_n, \nu) = 0 \ (\text{in probability})$$
$$\Leftrightarrow \forall \varepsilon > 0 \quad \lim_n Q_n(d(\nu_n, \nu) > \varepsilon) = 0. \tag{5.4}$$

5.1.1   *The spectral measure*

The following theorem states the convergence of the sequence $(\mu_{\mathrm{sp}}^{(n)})$ to an explicit limit, which we identify now, with the help of some more notation. We assume that $\mathfrak{Re}\, d \geq 0$.

Let

$$\alpha_{\mathrm{d}} = -\frac{\overline{\mathrm{d}}}{1+\overline{\mathrm{d}}} \tag{5.5}$$

and let $\theta_{\mathrm{d}} \in [0, \pi)$ and $\xi_{\mathrm{d}} \in [-|\theta_{\mathrm{d}}|, |\theta_{\mathrm{d}}|]$ be such that

$$\sin \frac{\theta_{\mathrm{d}}}{2} = \left| \frac{\mathrm{d}}{1+\mathrm{d}} \right|, \quad e^{\mathrm{i}\xi_{\mathrm{d}}} = \frac{1+\mathrm{d}}{1+\overline{\mathrm{d}}}.$$

If $\mathfrak{Re}\, \mathrm{d} \geq 0, \mathrm{d} \neq 0$, let w_{d} be defined by

$$w_{\mathrm{d}}(\theta) = \begin{cases} \dfrac{\sqrt{\sin^2((\theta - \xi_{\mathrm{d}})/2) - \sin^2(\theta_{\mathrm{d}}/2)}}{|1+\alpha_{\mathrm{d}}|\sin(\theta/2)}, & \text{if } \theta \in (\theta_{\mathrm{d}} + \xi_{\mathrm{d}}, 2\pi - \theta_{\mathrm{d}} + \xi_{\mathrm{d}}), \\ 0, & \text{otherwise.} \end{cases} \tag{5.6}$$

In this case, let

$$d\mu_{\mathrm{d}}(\zeta) := w_{\mathrm{d}}(\theta) \frac{d\theta}{2\pi} \quad (\zeta = e^{\mathrm{i}\theta}), \tag{5.7}$$

and for $\mathrm{d} = 0$, let

$$d\mu_0(\zeta) := \frac{d\theta}{2\pi} \quad (\zeta = e^{\mathrm{i}\theta})$$

the Haar measure on $\mathbb{T}$.

Theorem 5.1.   Assume $\mathfrak{Re}\, \mathrm{d} \geq 0$. As $n \to \infty$,

$$\lim \mu_{\mathrm{sp}}^{(n)} = \mu_{\mathrm{d}} \quad \text{(in probability).}$$

To prove this convergence, we use the parameterization of measures by their modified Verblunsky coefficients and the following two lemmas, whose proofs are postponed until the end of the section.

Lemma 5.2.   For every fixed $k \geq 0$, as $n \to \infty$,

$$\lim \gamma_k^{(n)} = -\frac{\mathrm{d}}{1+\overline{\mathrm{d}}} \quad \text{(in probability),} \tag{5.8}$$

and, consequently,

$$\lim \alpha_k^{(n)} = \alpha_{\mathrm{d}} e^{-\mathrm{i}(k+1)\xi_{\mathrm{d}}} \quad \text{(in probability).} \tag{5.9}$$

Lemma 5.3.   The sequence of Verblunsky coefficients of μ_{d} is precisely $(\alpha_{\mathrm{d}} e^{-\mathrm{i}(k+1)\xi_{\mathrm{d}}})_{k \geq 0}$.

Proof of Theorem 5.1.   For $v \in \mathcal{M}_1(\mathbb{T})$, let $m_k(v) = \int e^{\mathrm{i}k\theta} dv(\theta)$ be its kth moment and let $\alpha_k(v)$ be its kth Verblunsky, in short V-coefficient. Moments are related to V-coefficients in a continuous way: for every $j \geq 1$, $m_j(v)$ is a continuous function of $(\alpha_0, \ldots, \alpha_{(j \wedge N)-1})$, where N is the cardinal of the support of v. We know for Lemma 5.2 that for every fixed k, $\lim \alpha_j(\mu_{\mathrm{sp}}^{(n)}) = \alpha_j(\mu_{\mathrm{d}})$ for $j \leq k$, in probability. It entails the convergence of $m_k(\mu_{\mathrm{sp}}^{(n)})$ to $m_k(\mu_{\mathrm{d}})$, in probability, hence the weak convergence of the measures. ∎

Remark 3.   The above method can be easily adapted to show the trivial asymptotics of two different scaling regimes:

- if $\delta(n) = o(n)$, the limit is the uniform measure on $\mathbb{T}$;
- if $\delta(n)$ is real and $n = o(\delta(n))$, the limit is the Dirac measure at -1.

Proof of Lemma 5.2.   For $\gamma_k^{(n)}$ we use the Mellin transform

$$\mathbb{E}\big((1 - \gamma_k^{(n)})^s\big) = \frac{\Gamma(\beta'(n-k-1) + \delta + \overline{\delta} + s + 1)\Gamma(\beta'(n-k-1) + \overline{\delta} + 1)}{\Gamma(\beta'(n-k-1) + \delta + \overline{\delta} + 1)\Gamma(\beta'(n-k-1) + \overline{\delta} + s + 1)}$$

(this comes immediately from (3.3)).

Since for fixed $z \in \mathbb{C}$

$$\lim_{n \to \infty} \frac{\Gamma(n+z)}{\Gamma(n)n^z} = 1,$$

we get, for fixed s (and k),

$$\lim_{n \to \infty} \mathbb{E}\big((1 - \gamma_k^{(n)})^s\big) = \left(\frac{1 + \overline{\mathrm{d}} + \mathrm{d}}{1 + \overline{\mathrm{d}}}\right)^s,$$

which implies that

$$\lim_{n \to \infty} (1 - \gamma_k^{(n)}) = \frac{1 + \overline{\mathrm{d}} + \mathrm{d}}{1 + \overline{\mathrm{d}}} \quad \text{in probability},$$

which is equivalent to (5.8). The statement (5.9) is a direct consequence of (5.8) and (2.9). ∎

Remark 3.   The convergences in probability in the above lemma actually hold in L^p, for all $p > 0$ because all variables are bounded by 1.

Proof of Lemma 5.3.   If d is real, the Verblunsky coefficients are all equal. In fact, measures satisfying this property are known to have an absolute continuous part supported by an arc of the unit circle and a possible additional Dirac mass ([33], p. 87). More precisely, let $\alpha \in \mathbb{D}$ and let v be the measure on $\mathbb{T}$ such that $\alpha_k(v) = \alpha$ for every $k \geq 0$. If $\theta(\alpha)$

and $\xi(\alpha)$ are defined by

$$\theta(\alpha) = 2\arcsin|\alpha|, \quad e^{i\xi(\alpha)} = \frac{1+\alpha}{1+\overline{\alpha}},$$

then

$$d\nu(\zeta) = w(\theta)\frac{d\theta}{2\pi} \quad (\zeta = e^{i\theta}),$$

where

$$w(\theta) = \begin{cases} \dfrac{\sqrt{\sin^2(\theta/2) - \sin^2(\theta(\alpha)/2)}}{|1+\alpha|\,\sin((\theta+\xi(\alpha))/2)}, & \text{if } \theta \in (\theta(\alpha), 2\pi - \theta(\alpha)), \\ 0, & \text{otherwise}, \end{cases} \tag{5.10}$$

as soon as

$$\left|\alpha + \frac{1}{2}\right| \le \frac{1}{2} \tag{5.11}$$

(otherwise, there is an additional Dirac mass). The orthogonal polynomials with respect to this measure are known as the "Geronimus polynomials."

When $\alpha = \alpha_d$, we set $\theta(\alpha) = \theta_d$, $\xi(\alpha) = \xi_d$ and $\nu = \nu_d$, $w = w_d$. We see that

$$\alpha_d + \frac{1}{2} = \frac{1 - \overline{d}}{2(1 + \overline{d})},$$

so that condition (5.11) is fulfilled if and only if $\mathfrak{Re}\,d \ge 0$, which we assumed.

Moreover, it is known (see [34], p. 960) that if $(\alpha_k)_{k\ge 0}$ is the sequence of Verblunsky coefficients of a measure μ, then the coefficients $(e^{-i(k+1)\xi_d}\alpha_k)_{k\ge 0}$ are associated with μ rotated by ξ_d. Consequently,

$$d\mu_d(\zeta) = d\nu_d(\zeta e^{-i\xi_d}),$$

which is precisely (5.7). ∎

5.1.2   *The ESD*

In matrix models, the convergence of the ESD is often tackled directly via the convergence of moments or the convergence of the Cauchy transform. Here, we follow a different way: we use Theorem 5.1 and prove that the two sequences $(\mu_{\text{sp}}^{(n)})_n$ and $(\mu_{\text{esd}}^{(n)})_n$ are "contiguous."

Theorem 5.4.   Assume $\mathfrak{Re}\,d \ge 0$. As $n \to \infty$,

$$\lim_n \mu_{\text{esd}}^{(n)} = \mu_d \quad \text{(in probability)},$$

where μ_d is given in (5.7).

Proof.   From Theorem 5.1, we know, with the notation from the beginning of Section 5.1, that

$$\lim_{n \to \infty} d_L\big(\mu_{\mathrm{sp}}^{(n)}, \mu_{\mathrm{d}}\big) = 0 \ \text{ in probability.}$$

By the triangle inequality

$$d_L\big(\mu_{\mathrm{esd}}^{(n)}, \mu_{\mathrm{d}}\big) \le d_L\big(\mu_{\mathrm{sp}}^{(n)}, \mu_{\mathrm{d}}\big) + d_L\big(\mu_{\mathrm{sp}}^{(n)}, \mu_{\mathrm{esd}}^{(n)}\big),$$

so it is enough to prove that $d_L(\mu_{\mathrm{sp}}^{(n)}, \mu_{\mathrm{esd}}^{(n)})$ converges to 0 in probability. Thanks to (5.3), it is enough to prove that

$$\lim_{n \to \infty} \sup_t \big|F_{\mu_{\mathrm{sp}}^{(n)}}(t) - F_{\mu_{\mathrm{esd}}^{(n)}}(t)\big| = 0 \ \text{ in probability.} \tag{5.12}$$

In fact,

$$\sup_t \big|F_{\mu_{\mathrm{sp}}^{(n)}}(t) - F_{\mu_{\mathrm{esd}}^{(n)}}(t)\big| = \max_k \left|S_k^{(n)} - \frac{k}{n}\right|, \tag{5.13}$$

where $S_k^{(n)} = \sum_{j=1}^k \pi_j^{(n)}$. Using the union bound and the Markov inequality, we may write

$$\mathbb{P}\left(\max_k \left|S_k^{(n)} - \frac{k}{n}\right| > \varepsilon\right) \le \sum_{k=1}^n \mathbb{P}\left(\left|S_k^{(n)} - \frac{k}{n}\right| > \varepsilon\right) \le \varepsilon^{-4} \sum_{k=1}^n \mathbb{E}\left(\left(S_k^{(n)} - \frac{k}{n}\right)^4\right), \tag{5.14}$$

and then use explicit distributions. Indeed, we showed in Theorem 3.3 that the vector $(\pi_1^{(n)}, \dots, \pi_n^{(n)})$ follows the $\mathrm{Dir}_n(\beta')$ distribution. It entails that for $k = 1, \dots, n-1$, the variable $S_k^{(n)}$ is $\mathrm{Beta}(\beta' k, \beta'(n-k))$ distributed.

Recall that the Mellin transform of a beta variable $\mathrm{Beta}(a, b)$ with positive parameters a and b is

$$\mathbb{E}(\mathrm{Beta}(a, b)^s) = \frac{\Gamma(a+s)\Gamma(a+b)}{\Gamma(a)\Gamma(a+b+s)},$$

for $s > -a$. As a consequence, $\mathbb{E}\mathrm{Beta}(a, b) = a/(a+b)$ and a straightforward calculation gives

$$\mathbb{E}((\mathrm{Beta}(a, b) - \mathbb{E}\mathrm{Beta}(a, b)))^4) = \frac{3ab(2a^2 + 2b^2 - 2ab + a^2 b + ab^2)}{(a+b)^4 (a+b+1)(a+b+2)(a+b+3)} = O\left(\frac{ab}{(a+b)^4}\right).$$

When $a = \beta' k$ and $b = \beta'(n-k)$, this shows that the kth term in the sum of (5.14) is $O\left(\frac{k(n-k)}{n^4}\right)$, so that

$$\mathbb{P}\left(\max_k \left|S_k^{(n)} - \frac{k}{n}\right| > \varepsilon\right) = O\left(\sum_{k=1}^n \frac{k(n-k)}{n^4}\right) = O\left(\frac{1}{n}\right).$$

Thanks to (5.13), this yields (5.12) and completes the proof.   ∎

5.2 Large deviations for the ESD

It turns out that the convergence in Theorem 5.4 is exponentially fast, and to make this statement more precise we need to recall the definition of some notions of large deviations, for which the reference is the book of Dembo and Zeitouni [9]. We say that a sequence (P_n) of probability measures on a measurable Hausdorff space $(\mathcal{X}, B(\mathcal{X}))$ satisfies the LDP at scale u_n (with $u_n \to \infty$), if there exists a lower semicontinous function $I : \mathcal{X} \to [0, \infty]$ such that

$$\limsup \frac{1}{u_n} \log P_n(F) \leq -\inf\{I(x); x \in F\}, \tag{5.15}$$

$$\liminf \frac{1}{u_n} \log P_n(G) \geq -\inf\{I(x); x \in G\}, \tag{5.16}$$

for every closed set $F \subset \mathcal{X}$ and every open set $G \subset \mathcal{X}$. The rate function I is called good if its level sets are compact. More generally, a sequence of $\mathcal{X}$-valued random variables is said to satisfy the LDP if their distributions satisfy the LDP. From now on, we work with $\mathcal{X} = \mathcal{M}_1(\mathbb{T})$, whose compacity makes our task simpler. In fact, according to [9] Theorem 4.1.11, in this case the LDP is equivalent to the following property: for every $\mu \in \mathcal{M}_1(\mathbb{T})$

$$-I(\mu) = \lim_{\varepsilon \downarrow 0} \limsup_n \frac{1}{u_n} \log P_n(B(\mu, \varepsilon)), \tag{5.17}$$

$$= \lim_{\varepsilon \downarrow 0} \liminf_n \frac{1}{u_n} \log P_n(B(\mu, \varepsilon)), \tag{5.18}$$

where $B(\mu, \varepsilon)$ is the open ball of radius ε, centered at μ

$$B(\mu, \varepsilon) = \{\nu : d_L(\mu, \nu) < \varepsilon\}.$$

It is easy to see that the latter is equivalent to the pair of inequalities

$$\lim_{\varepsilon \downarrow 0} \limsup_n \frac{1}{u_n} \log P_n(B(\mu, \varepsilon)) \leq -I(\mu), \tag{5.19}$$

$$\lim_{\varepsilon \downarrow 0} \limsup_n \frac{1}{u_n} \log P_n(B(\mu, \varepsilon)) \geq -I(\mu). \tag{5.20}$$

For the sake of completeness, let us explain shortly why (5.19) and (5.20) lead to (5.15) and (5.16), respectively. On the one hand, every closed (hence compact) set in $\mathcal{M}_1(\mathbb{T})$ may be recovered by a finite number of balls and applying (5.19) for all these balls leads to (5.15). On the other hand, every open set contains open balls and it remains to optimize with respect to the centers of the balls.

Our large deviation result follows the way initiated by the pioneer paper of Ben Arous and Guionnet [4] and continued by Hiai and Petz (see [15, 17]).

We work with the set $\mathcal{M}_1(\mathbb{T})$ of probability measures on the unit circle. For $\mu \in \mathcal{M}_1(\mathbb{T})$, the Voiculescu entropy is

$$\Sigma(\mu) = \iint \log|\zeta - \zeta'| d\mu(\zeta) d\mu(\zeta') . \tag{5.21}$$

With a different sign it is the logarithmic energy of μ (see Chapter 5.3 in [16]). We also define the potential

$$Q_d(\zeta) = \begin{cases} -2(\mathfrak{Re}\,d) \log\left(2\sin\frac{\theta}{2}\right) - (\mathfrak{Im}\,d)(\theta - \pi), & \text{if } \zeta = e^{i\theta}, \theta \in (0, 2\pi), \\ \infty, & \text{if } \zeta = 1 \text{ and } \mathfrak{Re}\,d > 0, \\ -|\mathfrak{Im}\,d|\pi, & \text{if } \zeta = 1 \text{ and } \mathfrak{Re}\,d = 0. \end{cases} \tag{5.22}$$

It should be noted that Q_d is a lower semicontinuous function. The main result of this section is an extension of the theorem of Hiai and Petz (Theorem 5.4.10 in [16]), which corresponds to the case $d = 0$ + a *continuous* potential.

Theorem 5.5.   Let $d \in \mathbb{C}$ with $\mathfrak{Re}\,d \geq 0$. For $n \in \mathbb{N}$ and $(\zeta_1, \ldots, \zeta_n) \in \mathbb{T}^n$ let

$$h(\zeta_1, \ldots, \zeta_n) := \prod_{k=1}^n (1 - \zeta_k)^{\overline{d}\beta'n} (1 - \overline{\zeta_k})^{d\beta'n} \prod_{j<k} |\zeta_j - \zeta_k|^{2\beta'}$$

and let $\mathbb{P}_d^{(n)}$ be the distribution on $\mathbb{T}^n$ having the density

$$\frac{1}{\mathcal{Z}_d(n)} h(\zeta_1, \ldots, \zeta_n) , \tag{5.23}$$

where $\mathcal{Z}_d(n)$ is the normalization constant.

(1) We have

$$\lim_{n\to\infty} \frac{1}{n^2\beta'} \log \mathcal{Z}_d(n) = B(d), \tag{5.24}$$

where

$$B(d) = \int_0^1 x \log \frac{x(x + \mathfrak{Re}\,d)}{|x + d|^2} \, dx.$$

(2) The sequence of distributions of

$$\mu_{\text{esd}}^{(n)} = \frac{\delta_{\zeta_1} + \cdots + \delta_{\zeta_n}}{n}$$

under $\mathbb{P}_d^{(n)}$ satisfies the LDP in $\mathcal{M}_1(\mathbb{T})$ at scale $\beta'n^2$ with (good) rate function

$$I_d(\mu) = -\Sigma(\mu) + \int_{\mathbb{T}} Q_d(\zeta) d\mu(\zeta) + B(d) . \tag{5.25}$$

(3) The rate function vanishes only at $\mu = \mu_d$.

Proof.   (1) An exact expression of $\mathcal{Z}(n)$ is obtained using the following lemma, whose proof is postponed to the end of this subsection.

Lemma 5.6.   If we define

$$\mathcal{Z}_{s,t}(n) = \int_{[0,2\pi)^n} \prod_{k=1}^n (1 - e^{i\theta_k})^s (1 - e^{-i\theta_k})^t \prod_{j<k} |e^{i\theta_j} - e^{i\theta_k}|^\beta \, d\theta_1 \cdots d\theta_n,$$

then we have

$$\mathcal{Z}_{s,t}(n) = \frac{\Gamma(\beta'n + 1)}{(\Gamma(\beta' + 1))^n} \prod_0^{n-1} \frac{\Gamma(\beta'j + 1)\Gamma(\beta'j + 1 + s + t)}{\Gamma(\beta'j + 1 + s)\Gamma(\beta'j + 1 + t)}. \tag{5.26}$$

We have $\mathcal{Z}_{\mathrm{d}}(n) = \mathcal{Z}_{\bar{\mathrm{d}}\beta'n, \mathrm{d}\beta'n}(n)$ and then,

$$\log \mathcal{Z}_{\mathrm{d}}(n) = \log \Gamma(\beta'n + 1) - n \log \Gamma(\beta' + 1) + \sum_{j=0}^{n-1} \log \Gamma(\beta'j + 1)$$

$$+ \sum_{j=0}^{n-1} [\log \Gamma(\beta'j + 1 + 2\Re \mathrm{d}\beta'n) - 2\Re \log \Gamma(\beta'j + 1 + \mathrm{d}\beta'n)].$$

From the Binet formula (Abramowitz and Stegun [1] or Erdélyi *et al.* [11], p. 21), we have for $\Re x > 0$

$$\log \Gamma(x) = \left(x - \frac{1}{2}\right) \log x - x + \frac{1}{2} \log(2\pi) + \int_0^\infty f(s) e^{-sx} \, ds, \tag{5.27}$$

where the function f is defined by

$$f(s) = \left[\frac{1}{2} - \frac{1}{s} + \frac{1}{e^s - 1}\right] \frac{1}{s} = 2 \sum_{k=1}^\infty \frac{1}{s^2 + 4\pi^2 k^2},$$

and satisfies for every $s \geq 0$

$$0 < f(s) \leq f(0) = 1/12, \quad 0 < \left(sf(s) + \frac{1}{2}\right) < 1.$$

Using (5.27), a straightforward study of Riemann sums gives

$$\frac{1}{\beta'n^2} \log \mathcal{Z}_{\mathrm{d}}(n) \to \int_0^1 x \log \frac{x(x + \Re \mathrm{d})}{|x + \mathrm{d}|^2} \, dx = B(\mathrm{d}). \tag{5.28}$$

(2) The proof is based on the explicit form of the joint eigenvalue density. We follow the lines of [4, 15, 17], and [16]. We skip the index esd for notational convenience. Our goal is the proof of the two inequalities (see (5.19) and (5.20)), which hold for every

$\mu \in \mathcal{M}_1(\mathbb{T})$:

$$\lim_{\varepsilon \to 0} \limsup_n \frac{1}{\beta' n^2} \log \mathbb{P}_{\mathrm{d}}^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) \leq -I_{\mathrm{d}}(\mu), \tag{5.29}$$

$$\lim_{\varepsilon \to 0} \liminf_n \frac{1}{\beta' n^2} \log \mathbb{P}_{\mathrm{d}}^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) \geq -I_{\mathrm{d}}(\mu) \tag{5.30}$$

We rest on the LDP known in the case of $\mathrm{d} = 0$, and use the general method which consists in estimating the Radon–Nikodym derivative ruling the change of probability. The key formula is

$$\mathbb{P}_{\mathrm{d}}^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) = \frac{\mathcal{Z}_0(n)}{\mathcal{Z}_{\mathrm{d}}(n)} \mathbb{E}_0^{(n)}\left[1_{\mu^{(n)} \in B(\mu, \varepsilon)} e^{-n^2 \beta' \int Q_{\mathrm{d}} d\mu^{(n)}} \right]. \tag{5.31}$$

The upper bound (5.29). Let us first assume $\mathfrak{Re}\, \mathrm{d} > 0$, so that $Q_{\mathrm{d}}(\zeta) \to \infty$ as $\zeta \to 1$. For $R > 0$ we consider the cutoff $Q^R = \min(Q_{\mathrm{d}}, R)$. Since Q^R is continuous, the mapping $\nu \in \mathcal{M}_1(\mathbb{T}) \mapsto \int Q^R d\nu$ is continuous, so

$$\inf_{B(\mu, \varepsilon)} \int Q^R d\nu \geq \int Q^R d\mu - r_1(\varepsilon, R) \tag{5.32}$$

with $\lim_\varepsilon r_1(\varepsilon, R) = 0$. Since $Q_{\mathrm{d}} \geq Q^R$, we get

$$\frac{1}{\beta' n^2} \log \mathbb{E}_0^{(n)}\left[1_{\mu^{(n)} \in B(\mu, \varepsilon)} e^{-n^2 \beta' \int Q_{\mathrm{d}} d\mu^{(n)}} \right] \leq \frac{1}{\beta' n^2} \log \mathbb{P}_0^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) - \int Q^R d\mu + r_1(\varepsilon, R). \tag{5.33}$$

Thanks to the LDP known for $\mathrm{d} = 0$, we can take limsup in n and then limit in ε and obtain

$$\lim_\varepsilon \limsup_n \frac{1}{\beta' n^2} \log \mathbb{E}_0^{(n)}\left[1_{\mu^{(n)} \in B(\mu, \varepsilon)} e^{-n^2 \beta' \int Q_{\mathrm{d}} d\mu^{(n)}} \right] \leq \Sigma(\mu) - \int Q^R d\mu,$$

and thanks to (5.31) and (5.24)

$$\lim_\varepsilon \limsup_n \frac{1}{\beta' n^2} \log \mathbb{P}_{\mathrm{d}}^{(n)}(\mu^{(n)} \in B(\mu, \varepsilon)) \leq \Sigma(\mu) - \int Q^R d\mu - B(\mathrm{d}).$$

By the monotone convergence theorem $\lim_{R \to \infty} \int Q^R d\mu = \int Q_{\mathrm{d}} d\mu$, which proves (5.29).

If $\mathfrak{Re}\, \mathrm{d} = 0$, then Q_{d} is lower semicontinuous and bounded, so that the mapping $\nu \mapsto \int Q_{\mathrm{d}} d\nu$ is lower semicontinuous and (5.32) still holds with Q_{d} instead of Q^R and some $r(\varepsilon)$ instead of $r_1(\varepsilon, R)$. The rest of the argument is the same as above.

The lower bound (5.30). If $I_{\mathrm{d}}(\mu) = \infty$, the bound is trivial, so that we can assume that μ has no atom. To overcome the problem of the singularity at 1, we use a classical

approximation of μ by a probability vanishing in a neighborhood of 1, i.e.

$$d\mu_M(\zeta) = \frac{\mathbf{1}_{|1-\zeta|\geq M^{-1}}}{\mu(|1-\zeta|\geq M^{-1})}d\mu(\zeta)\,.$$

The benefit is that the mapping $\nu \mapsto \int Q_d\,d\nu$ is continuous in a neighborhood of μ_M. Choosing M large enough, we ensure that $d_L(\mu_M,\mu) \leq \varepsilon/2$ and then, thanks to the triangle inequality,

$$B(\mu,\varepsilon) \supset B(\mu_M,\varepsilon/2),$$

which leads to

$$\mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu,\varepsilon)) \geq \mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu_M,\varepsilon/2))\,.$$

We come back to (5.31) with μ_M and $\varepsilon/2$. For $\nu \in B(\mu_M,\varepsilon/2)$, we have $\int Q_d\,d\nu \leq \int Q_d\,d\mu_M + r_2(\varepsilon,M))$ where $\lim_\varepsilon r_2(\varepsilon,M) = 0$. We get

$$\frac{1}{\beta'n^2}\log\mathbb{E}_0^{(n)}\big[\mathbf{1}_{\mu^{(n)}\in B(\mu,\varepsilon)}e^{-n^2\beta'\int Q_d\,d\mu^{(n)}}\big] \geq \frac{1}{\beta'n^2}\mathbb{P}_0^{(n)}(\mu^{(n)} \in B(\mu_M,\varepsilon/2)) - \int Q_d\,d\mu_M - r_2(\varepsilon,M).$$

Again, we take liminf in n and lim in ε and use successively the LDP for d $= 0$, (5.31) and (5.24) to obtain

$$\lim_\varepsilon \limsup_n \frac{1}{\beta'n^2}\log\mathbb{P}_d^{(n)}(\mu^{(n)} \in B(\mu,\varepsilon)) \geq \Sigma(\mu_M) - \int Q_d\,d\mu_M - B(d)\,.$$

Now, since $I(\mu) < \infty$ and since Q_d and Σ are bounded below, the monotone convergence theorem yields

$$\lim_{M\to\infty}\Sigma(\mu_M) = \Sigma(\mu), \quad \lim_{M\to\infty}\int Q_d\,d\mu_M = \int Q_d\,d\mu\,.$$

This ends the proof of (5.30).

(3) The uniqueness of the minimizer is a direct consequence of the strict convexity of I which comes from the strict concavity of Σ.

We do not give a self-contained proof of the identity of the minimizer, but rather use a probabilistic argument. On the one hand, in Theorem 5.4 we proved that $\mu_{\text{esd}}^{(n)}$ converges weakly in probability to μ_d and, on the other hand, the LDP combined with the uniqueness of the minimizer imply that $\mu_{\text{esd}}^{(n)}$ converges weakly in probability to this minimizer. This completes the proof. ∎

Proof of Lemma 5.6.   We have

$$\frac{\mathcal{Z}_{s,t}(n)}{\mathcal{Z}_{0,0}(n)} = \mathbb{E}(\det(\text{Id} - U)^s\det(\text{Id} - \bar{U})^t),$$

where the mean is taken under the $CJ_{0,\beta}^{(n)}$ distribution. Under this distribution, $\det(\mathrm{Id} - U)$ has the same law as the product of independent variables $1 - \bar{\alpha}_k$, where α_k is $\nu_{\beta(n-k-1)+1}$ distributed. We get

$$\mathbb{E}(\det(\mathrm{Id} - U)^s \det(\mathrm{Id} - \bar{U})^t) = \prod_{j=0}^{n-1} \mathbb{E}(1 - \bar{\alpha}_j)^s (1 - \alpha_j))^t.$$

From (3.3) we get

$$\frac{\mathscr{Z}_{s,t}(n)}{\mathscr{Z}_{0,0}(n)} = \prod_{0}^{n-1} \frac{\Gamma(\beta' j + 1)\Gamma(\beta' j + 1 + s + t)}{\Gamma(\beta' j + 1 + s)\Gamma(\beta' j + 1 + t)}.$$

Besides, Lemma 4.4 in [23] gives

$$\mathscr{Z}_{0,0}(n) = \frac{\Gamma(\beta' n + 1)}{(\Gamma(\beta' + 1))^n}. \qquad \blacksquare$$

References

[1]   Abramowitz, M., and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. New York: Dover, 1972.

[2]   Ammar, G., W. Gragg, and L. Reichel. "Constructing a unitary Hessenberg matrix from spectral data." In *Numerical Linear Algebra, Digital Signal Processing and Parallel Algorithms,* 385–95. NATO Advanced Science Institutes Series F: Computer and Systems Sciences 70. Berlin: Springer, 1988.

[3]   Andrews, George E., Richard A. Askey, and Ranjan Roy. *Special Functions.* Encyclopedia of Mathematics and Its Applications 71. Cambridge, UK: Cambridge University Press, 1999.

[4]   Ben Arous, G., and A. Guionnet. "Large deviations for Wigner's law and Voiculescu's non-commutative entropy." *Probability Theory and Related Fields* 108 (1997): 517–42.

[5]   Borodin, A., and G. Olshanski. "Infinite random matrices and ergodic measures." *Communications in Mathematical Physics* 203 (2001): 87–123.

[6]   Bourgade, P., C. P. Hughes, A. Nikeghbali, and M. Yor. "The characteristic polynomial of a random unitary matrix: A probabilistic approach." *Duke Mathematical Journal* 145, no. 1 (2008) 45–69.

[7]   Bourgade, P., A. Nikeghbali, and A. Rouault. "Hua–Pickrell measures on general compact groups." (2007): preprint arXiv:0712.0848.

[8]   Cantero, M. J., L. Moral, and L. Velazquez. "Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle." *Linear Algebra and Its Applications* 362 (2003): 29–56.

[9]   Dembo, A., and O. Zeitouni. *Large Deviations Techniques and Applications,* 2nd ed. New York: Springer, 1998.

[10] Dumitriu, I., and A. Edelman. "Matrix models for beta ensembles." *Journal of Mathematical Physics* 43 (2002): 5830–47.

[11] Erdelyi, A., W. Magnus, F. Oberhettinger, and F. G. Tricomi. *Higher Transcendental Functions,* vol 1. New York: Krieger, 1981.

[12] Forrester, P. J. *Log-Gases and Random Matrices*. Book available from: http://www.ms.unimelb.edu.au/~matpjf/matpjf.html.

[13] Forrester, P. J., and E. M. Rains. "Jacobians and rank-1 perturbations relating to unitary Hessenberg matrices." *International Mathematics Research Notices* (2006).

[14] Forrester, P. J., and N. S. Witte. "Application of the τ-function theory of Painleve equations to random matrices: Pvi, the JUE, CyUE, cJUE and scaled limits." *Nagoya Mathematical Journal* 174 (2004): 29–114.

[15] Hiai, F., and D. Petz. "A large deviation theorem for the empirical eigenvalue distribution of random unitary matrices." *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques* 36 (2000): 71–85.

[16] Hiai, F., and D. Petz. *The Semicircle Law, Free Random Variables and Entropy*. Mathematical Surveys and Monographs 77. Providence, RI: American Mathematical Society, 2000.

[17] Hiai, F., and D. Petz. "Large deviations for functions of two random projection matrices." *Acta Universitatis Szegediensis. Acta Scientiarum Mathematicarum* 72 (2006): 581–609.

[18] Hua, L. K. *Harmonic Analysis of Functions of Several Complex Variables in the Classical Domains*. Translations of Mathematical Monographs 6. Providence, RI: American Mathematical Society, 1963.

[19] Keating, J. P., and N. C. Snaith. "Random matrix theory and $\zeta(1/2 + it)$." *Communications in Mathematical Physics* 214 (2000): 57–89.

[20] Keating, J. P., and N. C. Snaith. "Random matrix theory and L-functions at $s = 1/2$." *Communications in Mathematical Physics* 214 (2000): 91–110.

[21] Khrushchev, S. V. "Schur's algorithm, orthogonal polynomials, and convergence of Wall's continued fractions in $L^2(\mathbb{T})$." *Journal of Approximation Theory* 108, no. 2 (2001): 161–248.

[22] Khrushchev, S. V. "Classification theorems for general orthogonal polynomials on the unit circle." *Journal of Approximation Theory* 116, no. 2 (2002): 268–342.

[23] Killip, R., and I. Nenciu. "Matrix models for circular ensembles." *International Mathematics Research Notices* 50 (2004): 2665–701.

[24] Killip, R., and M. Stoiciu. "Eigenvalue statistics for CMV matrices: From Poisson to clock via CβE." *Duke Mathematical Journal* 146, no. 3 (2009): 361–99.

[25] Mehta, M. L. *Random Matrices*, 2nd ed. Boston, MA: Academic Press, 1991.

[26] Mezzadri, F., and N. C. Snaith (eds). *Recent Perspectives in Random Matrix Theory and Number Theory*. London Mathematical Society Lecture Note Series 322. Cambridge: Cambridge University Press, 2005.

[27] Neretin, Y. A. "Hua-type integrals over unitary groups and over projective limits of unitary groups." *Duke Mathematical Journal* 114 (2002): 239–66.

[28] Pickrell, D. "Measures on infinite-dimensional Grassmann manifolds." *Journal of Functional Analysis* 70, no. 2 (1987): 323–56.

[29] Pickrell, D. "Mackey analysis of infinite classical motion groups." *Pacific Journal of Mathematics* 150 (1991): 139–66.

[30] Port, Sidney C. *Theoretical Probability for Applications*. New York: Wiley, 1994.

[31]   Simon, B. "CMV matrices: Five years later." *Journal of Computational and Applied Mathematics* 208 (2007): 120–54.

[32]   Simon, B. "OPUC on one foot." *Bulletin of the American Mathematical Society* 42, no. 4 (2005): 431–60.

[33]   Simon, B. *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*. AMS Colloquium Publications 54. Providence, RI: American Mathematical Society, 2005.

[34]   Simon, B. *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*. AMS Colloquium Publications 54. Providence, RI: American Mathematical Society, 2005.

[35]   Witte, N. S., and P. J. Forrester. "Gap probabilities in the finite and scaled Cauchy random matrix ensembles." *Nonlinearity* 13 (2000): 1965–86.