

HUA-PICKRELL MEASURES ON GENERAL COMPACT GROUPS

P. BOURGADE, A. NIKEGHBALI, AND A. ROUAULT

ABSTRACT. Take a generic subgroup \mathcal{G} , endowed with its Haar measure, from $U(n, K)$, the unitary group of dimension n over the field K of real, complex or quaternion numbers. We give some equalities in law for $Z := \det(\text{Id} - G)$, $G \in \mathcal{G}$: under some general conditions, Z can be decomposed as a product of independent random variables, whose laws are explicitly known (Section 2). Consequently \mathcal{G} , endowed with a generalization of its Haar measure (the Hua-Pickrell measure), can be generated as a product of independent reflections. This constitutes a generalization of the well known Ewens sampling formula, corresponding to $\mathcal{G} = \mathcal{S}_n$, the n -dimensional symmetric group (Section 3). Finally, explicit determinantal point processes can be associated to the spectrum induced by the Hua-Pickrell measures, implying asymptotics on correlation functions (Section 4).

CONTENTS

1. Introduction	1
2. A decomposition of the Haar measure and its consequences	3
2.1. The general equality in law.	3
2.2. Decomposition of determinants as products of independent random variables.	5
2.3. Unitary groups.	6
2.4. Symplectic groups.	7
2.5. Symmetric groups.	9
3. Hua-Pickrell measures on general compact groups	10
3.1. Ewens sampling formula on general compact groups.	11
3.2. Examples.	12
4. Determinantal point processes for the Hua-Pickrell measures	13
4.1. Determinantal form for the spectral density.	14
4.2. Implied asymptotics.	18
References	20

2000 *Mathematics Subject Classification.* 15A52, 60F05, 60F15.

Key words and phrases. Decomposition of Haar Measure, Random Matrices, Characteristic Polynomials, Ewens's sampling formula, Hua-Pickrell measures, determinantal processes.

1. INTRODUCTION

Let G be distributed with the Haar measure on the n -dimensional complex unitary group. Recently the study of $\det(\text{Id}_n - G)$ has been highly motivated by the Keating-Snaith paradigm (see [10]) : its distribution is conjecturally closely related to the repartition of the values of the zeta function on the critical line.

This characteristic polynomial can surprisingly be decomposed as a product of n independent random variables (see [4]) :

$$\det(\text{Id}_n - G) \stackrel{\text{law}}{=} \prod_{k=1}^n \left(1 - e^{i\theta_k} \sqrt{\beta_{1,k-1}}\right), \quad (1.1)$$

where $\theta_1, \dots, \theta_n, \beta_{1,0}, \dots, \beta_{1,n-1}$ are independent random variables, the θ_k 's being uniformly distributed on $(-\pi, \pi)$ and the $\beta_{1,j}$'s ($0 \leq j \leq n-1$) being beta distributed with parameters 1 and j (with the convention that $\beta_{1,0} = 1$).

The present paper aims first to extend (1.1) to other compact groups, such as unitary groups over other fields and the symplectic group (Section 2). Another example is given by the group of permutations of dimension n , \mathcal{S}_n , or more precisely the symmetrized group $\tilde{\mathcal{S}}_n = \{(e^{i\theta_j} \delta_{\sigma(i)}^j)_{1 \leq i, j \leq n} \mid \sigma \in \mathcal{S}_n, (\theta_1, \dots, \theta_n) \in (-\pi, \pi)^n\}$: the corresponding decomposition is

$$\det(\text{Id}_n - G) \stackrel{\text{law}}{=} \prod_{k=1}^n \left(1 - e^{i\theta_k} X_k\right), \quad (1.2)$$

where $\theta_1, \dots, \theta_n, X_1, \dots, X_n$ are independent random variables, the θ_k 's being uniformly distributed on $(-\pi, \pi)$ and the X_k 's being Bernoulli variables : $\mathbb{P}(X_k = 1) = 1/k$, $\mathbb{P}(X_k = 0) = 1 - 1/k$.

The analogy between formulae (1.1) and (1.2) suggests that the well-known Ewens sampling formula on \mathcal{S}_n should admit an equivalent on the unitary group $U(n, \mathbb{C})$. Let us recall this formula for the symmetric group. Define :

- (a) $\sigma_1 := \tau_n \circ \dots \circ \tau_1$ where the τ_k 's are independent transpositions in \mathcal{S}_n , $\tau_k = [1, j]$, ($1 \leq j \leq k$), with

$$\mathbb{P}(\tau_k(1) = j) = \begin{cases} \frac{\theta}{\theta+k-1} & \text{if } j = 1 \\ \frac{1}{\theta+k-1} & \text{if } 2 \leq j \leq k \end{cases} ;$$

- (b) σ_2 with law $\mu^{(\theta)}$, the sampling of the Haar measure μ on \mathcal{S}_n by a factor θ^{k_σ} (k_σ : the number of cycles of a permutation σ) :

$$\mathbb{E}_{\mu^{(\theta)}}(f(\sigma_2)) = \frac{\mathbb{E}_\mu(f(\sigma_2)\theta^{k_{\sigma_2}})}{\mathbb{E}_\mu(\theta^{k_{\sigma_2}})}$$

for any bounded measurable function f .

Then the Ewens sampling formula can be expressed as the simple equality

$$\sigma_1 \stackrel{\text{law}}{=} \sigma_2. \quad (1.3)$$

More details about this can be found in [2].

Section 3 generalizes (1.3) to unitary groups and a particular class of their subgroups. The analogues of transpositions in decomposition (a) are reflections and the sampling (b) is defined relatively to the factor $\det(\text{Id} - G)^{\bar{\delta}} \det(\text{Id} - \bar{G})^{\delta}$, $\delta \in \mathbb{C}$.

Such samplings with $\delta \in \mathbb{R}$ have already been studied on the finite-dimensional unitary group by Hua ([6]), and results about the infinite dimensional case (on complex Grassmannians) were given by Pickrell ([14] and [15]). More recently, Neretin also studied these sampled measures, introducing the possibility $\delta \in \mathbb{C}$ ([12]). Borodin and Olshanski ([3]) have used the analogue of this sampled Haar measure on the infinite dimensional unitary group and studied resulting ergodic properties.

Following their work about the unitary group, we will refer to these sampled Haar measures as the Hua-Pickrell probability measures, although they might also have been called Ewens probability measures, as a generalization of the special case of the symmetric group. Forrester and Witte ([5]) also studied these measures, referring to them as the cJUE distribution.

Finally, inspired by Borodin and Olshanski's work and questions on $U(n, \mathbb{C})$, Section 4 explains which determinantal processes get involved in the sampled spectral measure, and the resulting limit theorems.

Notations. In this paper, as mentioned in the abstract, $U(n, K)$ denotes the unitary group over K^n with $K = \mathbb{R}, \mathbb{C}$ or \mathbb{H} , that is to say the set of elements U in $M(n, K)$ with $U\bar{U}^{\top} = \text{Id}$. An element R in $U(n, K)$ will be referred to as a reflection if $R - \text{Id}$ has rank 0 or 1.

The reflections can also be described in the following way. Let $\mathcal{M}(n, K)$ be the set of matrices M in $M(n, K)$ that can be written

$$M = \left(M_1 \| e_2 - k \frac{\bar{M}_{12}}{1 - \bar{M}_{11}} \| \dots \| e_n - k \frac{\bar{M}_{1n}}{1 - \bar{M}_{11}} \right). \quad (1.4)$$

with (e_1, \dots, e_n) an orthonormal basis, the vector $M_1 = (M_{11}, \dots, M_{1,n})^{\top} \neq e_1$ on the unit n -dimensional sphere and $k = M_1 - e_1$. Then the reflections are exactly the elements R in $M(n, K)$ having the following form :

$$R = \begin{pmatrix} \text{Id}_j & 0 \\ 0 & M \end{pmatrix}$$

with $M \in \mathcal{M}(n - j, K)$, $0 \leq j \leq n$. Note that in (1.4) the vector k must come before the coefficients because the quaternion field is not commutative.

2. A DECOMPOSITION OF THE HAAR MEASURE AND ITS CONSEQUENCES

In this section, we show how an element of any classical compact group (under the Haar measure) can be generated as a product of independent elementary transformations. This will lead to remarkable identities for the characteristic polynomial.

2.1. The general equality in law. Let \mathcal{G} be a subgroup of $U(n, K)$, the group of unitary matrices of size n over K . Let (e_1, \dots, e_n) be an orthonormal basis of K^n and $\mathcal{H} := \{H \in \mathcal{G} \mid H(e_1) = e_1\}$, the subgroup of \mathcal{G} which stabilizes e_1 . For a generic compact group \mathcal{A} , we write $\mu_{\mathcal{A}}$ for the unique Haar probability measure on \mathcal{A} . Then the following result holds (it is a generalization of Proposition 2.1 in [4]):

Proposition 2.1. *Let M and H be independent random matrices, $M \in \mathcal{G}$ and $H \in \mathcal{H}$ with distribution $\mu_{\mathcal{H}}$. Then $MH \sim \mu_{\mathcal{G}}$ if and only if $M(e_1) \sim p_1(\mu_{\mathcal{G}})$, where p_1 is the map $p_1 : G \mapsto G(e_1)$.*

Proof. If $MH \sim \mu_{\mathcal{G}}$, then $M(e_1) = MH(e_1) \sim p_1(\mu_{\mathcal{G}})$.

Suppose now that $M(e_1) \sim p_1(\mu_{\mathcal{G}})$. Thanks to the uniqueness of the Haar probability measure, to prove $MH \sim \mu_{\mathcal{G}}$, it suffices to show

$$GMH \stackrel{\text{law}}{=} MH$$

for any fixed $G \in \mathcal{G}$. Since $M(e_1) \sim p_1(\mu_{\mathcal{G}})$, $GM(e_1) \sim p_1(\mu_{\mathcal{G}})$. Therefore in an orthonormal basis with first element e_1 , the matrix GM can be written $(P(e_1) \parallel \tilde{P})$ with $P(e_1) \stackrel{\text{law}}{=} M(e_1)$. Consequently, by conditioning on the value $M(e_1) = P(e_1) = v$, it is sufficient to show that

$$(v \parallel P')H \stackrel{\text{law}}{=} (v \parallel M')H,$$

for some distributions on P' and M' , still assumed to be independent of H . As $v = M(e_1) \sim p_1(\mu_{\mathcal{G}})$, there exists almost surely an element $G_v \in \mathcal{G}$ with $G_v(e_1) = v$. By multiplication of the above equality by G_v^{-1} , we only need to show that

$$P''H \stackrel{\text{law}}{=} M''H$$

for some elements P'' and M'' in \mathcal{H} , again assumed to be independent of H . By conditioning on P'' (resp M''), we know that $P''H \stackrel{\text{law}}{=} H$ (resp $M''H \stackrel{\text{law}}{=} H$) by definition of the Haar measure $\mu_{\mathcal{H}}$. This gives the desired result. \square

This proposition will enable us to give a simple way to generate the Haar measure on the group \mathcal{G} . Before stating the corresponding theorem, we need to introduce the following designation.

Definition 2.2. Let \mathcal{G} be a subgroup of $U(n, K)$ and, for all $1 \leq k \leq n-1$, let μ_k be the Haar measure on this subgroup of $\mathcal{G} : \mathcal{H}_k := \{G \in \mathcal{G} \mid G(e_j) = e_j, 1 \leq j \leq k\}$. We also set $\mathcal{H}_0 = \mathcal{G}$, $\mu_0 = \mu_{\mathcal{G}}$. Moreover, for all $1 \leq k \leq n$ we define p_k as the map $p_k : G \mapsto G(e_k)$.

A sequence $(\nu_0, \dots, \nu_{n-1})$ of probability measures on \mathcal{G} is said to be coherent with $\mu_{\mathcal{G}}$ if for all $0 \leq k \leq n-1$, $\nu_k(\mathcal{H}_k) = 1$ and the probability measures $p_{k+1}(\nu_k)$ and $p_{k+1}(\mu_k)$ are the same.

In the following, $\nu_0 \times \nu_1 \times \dots \times \nu_{n-1}$ stands for the law of a random variable $H_0 H_1 \dots H_{n-1}$ where all H_i 's are independent and $H_i \sim \nu_i$. Now

we can provide a general method to generate an element of \mathcal{G} endowed with its Haar measure.

Theorem 2.3. *Let \mathcal{G} be a subgroup of $U(n, K)$. Let $(\nu_0, \dots, \nu_{n-1})$ be a sequence of coherent measures with $\mu_{\mathcal{G}}$. Then $\mu_{\mathcal{G}}$ and $\nu_0 \times \nu_1 \times \dots \times \nu_{n-1}$ are the same.*

Proof. It is sufficient to prove by induction on $1 \leq k \leq n$ that

$$\nu_{n-k} \times \nu_{n-k+1} \times \dots \times \nu_{n-1} = \mu_{\mathcal{H}_{n-k}},$$

which gives the desired result for $k = n$. If $k = 1$ this is obvious. If the result is true at rank k , it remains true at rank $k + 1$ by a direct application of Proposition 2.1 to the groups \mathcal{H}_{n-k-1} and its subgroup \mathcal{H}_{n-k} . \square

As an example, take the orthogonal group $O(n)$. Let $\mathcal{S}_{\mathbb{R}}^{(k)} = \{x \in \mathbb{R}^k \mid |x| = 1\}$ and, for $x_k \in \mathcal{S}_{\mathbb{R}}^{(k)}$, $r_k(x_k)$ the matrix representing the reflection which transforms x_k in the first element of the basis. If the x_k 's are uniformly distributed on the $\mathcal{S}_{\mathbb{R}}^{(k)}$'s and independent, then Theorem 2.3 implies that

$$r_n(x_n) \begin{pmatrix} \text{Id}_1 & 0 \\ 0 & r_{n-1}(x_{n-1}) \end{pmatrix} \cdots \begin{pmatrix} \text{Id}_{n-2} & 0 \\ 0 & r_2(x_2) \end{pmatrix} \begin{pmatrix} \text{Id}_{n-1} & 0 \\ 0 & r_1(x_1) \end{pmatrix} \sim \mu_{O(n)}.$$

2.2. Decomposition of determinants as products of independent random variables. Note that from now on in this article, the determinants are supposed to be over commutative fields : the study of determinants over the quaternionic field is not our purpose here.

Let \mathcal{G} and \mathcal{H} be as in the previous subsection and \mathcal{R} be the set of elements of \mathcal{G} which are reflections. Define also

$$g : \begin{cases} \mathcal{H} & \rightarrow U(n-1, K) \\ H & \mapsto H_{\text{span}(e_2, \dots, e_n)} \end{cases},$$

where $H_{\text{span}(e_2, \dots, e_n)}$ is the restriction of H to $\text{span}(e_2, \dots, e_n)$. Now suppose that

$$\{G(e_1) \mid G \in \mathcal{G}\} = \{R(e_1) \mid R \in \mathcal{R}\}. \quad (2.1)$$

Under this additional condition the following Proposition allows to represent the characteristic polynomial of \mathcal{G} as a product of two independent variables.

Proposition 2.4. *Let G ($\sim \mu_{\mathcal{G}}$), G' ($\sim \mu_{\mathcal{G}}$) and H ($\sim \mu_{\mathcal{H}}$) be independent. Suppose that condition (2.1) holds. Then*

$$\det(\text{Id}_n - G) \stackrel{\text{law}}{=} (1 - \langle e_1, G'(e_1) \rangle) \det(\text{Id}_{n-1} - g(H)).$$

Proof. Note that in Proposition 2.1, we can choose any matrix $M \in U(n, K)$ with $M(e_1)$ distributed as $G(e_1)$. Let us choose the simplest suitable transformation R : the reflection transforming e_1 in $R(e_1)$ if $R(e_1) \neq e_1$ (Id if $R(e_1) = e_1$) with $R(e_1) \stackrel{\text{law}}{=} G(e_1)$ independent of H . Thanks to condition

(2.1), $R \in \mathcal{G}$. Let the vector k be $R(e_1) - e_1$. From (1.4) there exists $(\lambda_2, \dots, \lambda_n) \in K^{n-1}$ such that

$$R = (e_1 + k \| e_2 + k\lambda_2 \| \dots \| e_n + k\lambda_n).$$

Hence from Proposition 2.1, one can write

$$\det(\text{Id} - G) \stackrel{\text{law}}{=} \det(\text{Id} - RH) = \det(\overline{H}^T - R) \det H.$$

If we call $(u_1, \dots, u_{n-1}) := \overline{g(H)}^T$ then using the multi-linearity of the determinant we get

$$\begin{aligned} \det(\overline{H}^T - R) &= \det\left(-k, \begin{pmatrix} 0 \\ u_1 \end{pmatrix} - e_2 - k\lambda_2, \dots, \begin{pmatrix} 0 \\ u_{n-1} \end{pmatrix} - e_n - k\lambda_n\right) \\ &= \det\left(-k, \begin{pmatrix} 0 \\ u_1 \end{pmatrix} - e_2, \dots, \begin{pmatrix} 0 \\ u_{n-1} \end{pmatrix} - e_n\right) \\ &= \det\left(\begin{array}{c|c} -k_1 & 0 \\ \dots & \overline{g(H)}^T - I_{n-1} \end{array}\right) \\ \det(\overline{H}^T - R) &= -k_1 \det(\overline{g(H)}^T - I_{n-1}). \end{aligned}$$

Finally, $\det(\text{Id} - G) \stackrel{\text{law}}{=} -k_1 \det(\text{Id} - g(H))$, with $-k_1 = 1 - \langle e_1, R(e_1) \rangle \stackrel{\text{law}}{=} 1 - \langle e_1, G'(e_1) \rangle$ and H independent. \square

This decomposition can be iterated to write the determinant as a product of independent random variables. We first need the equivalent of condition (2.1) for every dimension.

Definition 2.5. Note \mathcal{R}_k the set of elements in \mathcal{H}_k which are reflections. If for all $0 \leq k \leq n-1$

$$\{R(e_{k+1}) \mid R \in \mathcal{R}_k\} = \{H(e_{k+1}) \mid H \in \mathcal{H}_k\},$$

the group \mathcal{G} will be said to satisfy condition (R) (R standing for reflection).

The following result now follows immediately from Proposition 2.4 combined with an induction on n :

Theorem 2.6. *Let \mathcal{G} be a subgroup of $U(n, K)$ satisfying condition (R), and $(\nu_0, \dots, \nu_{n-1})$ be coherent with $\mu_{\mathcal{G}}$. Take $H_k \sim \nu_k$, $0 \leq k \leq n-1$, and $G \sim \mu_{\mathcal{G}}$, all being assumed independent. Then*

$$\det(\text{Id} - G) \stackrel{\text{law}}{=} \prod_{k=0}^{n-1} (1 - \langle H_k(e_{k+1}), e_{k+1} \rangle).$$

2.3. Unitary groups. Take $\mathcal{G} = U(n, \mathbb{C})$. Then $\mu_{\mathcal{H}_k} = f_k(\mu_{U(n-k, \mathbb{C})})$ where $f_k : A \in U(n-k, \mathbb{C}) \mapsto \text{Id}_k \oplus A$. As all reflections with respect to a hyperplane of \mathbb{C}^{n-k} are elements of $U(n-k, \mathbb{C})$, one can apply Theorem 2.6. The Hermitian products $\langle e_k, H_k(e_k) \rangle$ are distributed as the first coordinate of the first vector of an element of $U(n-k, \mathbb{C})$, that is to say the first coordinate of the $(n-k)$ -dimensional unit complex sphere with uniform measure : $\langle H_k(e_{k+1}), e_{k+1} \rangle \stackrel{\text{law}}{=} e^{i\theta_n} \sqrt{\beta_{1, n-k-1}}$ with θ_n uniform on $(-\pi, \pi)$ and independent of $\beta_{1, n-k-1}$, a beta variable with parameters 1 and $n-k-1$.

Therefore we get the following Corollary of Theorem 2.6.

Corollary 2.7. Unitary group over complex numbers. ([4]) *Let $G \in U(n, \mathbb{C})$ be $\mu_{U(n, \mathbb{C})}$ distributed. Then*

$$\det(\text{Id} - G) \stackrel{\text{law}}{=} \prod_{k=1}^n \left(1 - e^{i\theta_k} \sqrt{\beta_{1, k-1}}\right),$$

with $\theta_1, \dots, \theta_n, \beta_{1,0}, \dots, \beta_{1, n-1}$ independent random variables, the θ_k 's uniformly distributed on $(-\pi, \pi)$ and the $\beta_{1, j}$'s ($0 \leq j \leq n-1$) being beta distributed with parameters 1 and j (by convention, $\beta_{1,0} = 1$).

Let $SO(2n) = \{X \in U(2n, \mathbb{R}) \mid \det X = 1\}$. A similar reasoning (with the complex unit spheres replaced by the real ones) gives this analogous result.

Corollary 2.8. Special orthogonal group. *Let $G \in SO(2n)$ be $\mu_{SO(2n)}$ distributed. Then*

$$\det(\text{Id} - G) \stackrel{\text{law}}{=} 2 \prod_{k=2}^{2n} \left(1 - \epsilon_k \sqrt{\beta_{\frac{1}{2}, \frac{k-1}{2}}}\right),$$

with $\epsilon_1, \dots, \epsilon_{2n}, \beta_{1/2, 1/2}, \dots, \beta_{1/2, (2n-1)/2}$ independent random variables, $\mathbb{P}(\epsilon_k = 1) = \mathbb{P}(\epsilon_k = -1) = 1/2$, and the β 's being beta distributed with the indicated parameters.

2.4. Symplectic groups. Following the Katz-Sarnak philosophy ([8] and [9]), the study of moments of families of L -functions gives great importance to $\det(\text{Id} - G)$ for $G \in U(n, \mathbb{C})$, $SO(2n)$ but also $USp(2n, \mathbb{C}) = \{U \in U(2n, \mathbb{C}) \mid UJU^\top = J\}$, with

$$J = \begin{pmatrix} 0 & \text{Id}_n \\ -\text{Id}_n & 0 \end{pmatrix}. \quad (2.2)$$

Getting a decomposition as a product of independent random variables for $\mathcal{G} = USp(2n, \mathbb{C})$ requires some additional work, detailed below.

Let K be a subfield of \mathbb{R} , \mathbb{C} or \mathbb{H} and K' be a subfields of \mathbb{R} or \mathbb{C} . We write $M(m, K')$ for the ring of linear applications on K'^m . Let $\phi : K \rightarrow M(m, K')$ be a continuous injective ring morphism such that $\phi(\bar{x}) = \overline{\phi(x)}$. This morphism trivially induces the ring morphism

$$\Phi : \begin{cases} M_n(K) & \rightarrow & M_{nm}(K') \\ (a_{ij})_{1 \leq i, j \leq n} & \mapsto & (\phi(a_{ij}))_{1 \leq i, j \leq n} \end{cases}.$$

Let \mathcal{G} be a subgroup of $U(n, K)$; then $\Phi(\mathcal{G})$ is a subgroup of $U(nm, K')$. The action of Φ can be applied to Theorem 2.3 and implies, with the notation of the Theorem, $\Phi(\mu_{\mathcal{G}}) = \Phi(\nu_0) \times \Phi(\nu_1) \times \cdots \times \Phi(\nu_{n-1})$. As invariance in law by left translation is conserved by a multiplicative morphism, $\Phi(\mu_{\mathcal{G}}) = \mu_{\Phi(\mathcal{G})}$, so

$$\mu_{\Phi(\mathcal{G})} = \Phi(\nu_0) \times \Phi(\nu_1) \times \cdots \times \Phi(\nu_{n-1}). \quad (2.3)$$

This constitutes an analogue of Theorem 2.3 about the decomposition of the Haar measure. What would be the counterpart of Theorem 2.6 about the decomposition of the determinant? A straightforward although lengthy computation, similar to that in the proof of Proposition 2.4, answers this question.

Theorem 2.9. *Let \mathcal{G} be a subgroup of $U(n, K)$ checking condition (R), and $(\nu_0, \dots, \nu_{n-1})$ coherent with $\mu_{\mathcal{G}}$. Take $H_k (\sim \Phi(\nu_k))$, $0 \leq k \leq n-1$, and $G (\sim \mu_{\Phi(\mathcal{G})})$, all being assumed independent. Then*

$$\det(\text{Id}_{nm} - G) \stackrel{\text{law}}{=} \prod_{i=0}^{n-1} \det(\text{Id}_m - \Phi(\langle H_k(e_{k+1}), e_{k+1} \rangle)).$$

Consequently, $\det(\text{Id} - G)$ for $G \sim \mu_{USp(2n, \mathbb{C})}$, can be split in a product of n independent random variables. This is an easy application of Theorem 2.9 with

$$\phi : \begin{cases} \mathbb{H} & \rightarrow & M_2(\mathbb{C}) \\ a + ib + jc + kd & \mapsto & \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix} \end{cases},$$

the usual representation of quaternions. Indeed, for such a choice of ϕ , $\Phi(U(n, \mathbb{C}))$ is precisely the set of elements in $G \in U(2n, \mathbb{C})$ satisfying $G\tilde{J}G^{\top} = \tilde{J}$. Here, $J_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\tilde{J} = J_2 \oplus \cdots \oplus J_2$ is conjugate to J , defined by (2.2). The set $\Phi(U(n, \mathbb{C}))$ is therefore conjugate to $USp(2n, \mathbb{C})$, so the law of $\det(\text{Id} - G)$ is the same in both sets endowed with their respective Haar measure. As

$$\det \left(\text{Id} - \begin{pmatrix} a + ib & c + id \\ -c + id & a - ib \end{pmatrix} \right) = (a - 1)^2 + b^2 + c^2 + d^2,$$

the expected decomposition results from Theorem 2.9.

Corollary 2.10. Symplectic group. *Let $G \in USp(2n, \mathbb{C})$ be $\mu_{USp(2n, \mathbb{C})}$ distributed. Then*

$$\det(\text{Id} - G) \stackrel{\text{law}}{=} \prod_{k=1}^n ((a_k - 1)^2 + b_k^2 + c_k^2 + d_k^2),$$

with the vectors (a_k, b_k, c_k, d_k) , $1 \leq k \leq n$, being independent and (a_k, b_k, c_k, d_k) $4k$ coordinates of the $4k$ -dimensional real unit sphere endowed with the uniform measure; hence $(a_k, b_k, c_k, d_k) \stackrel{\text{law}}{=} \frac{1}{\sqrt{\mathcal{N}_1^2 + \cdots + \mathcal{N}_{4k}^2}} (\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3, \mathcal{N}_4)$, with the \mathcal{N}_i 's independent standard normal variables.

Remark. If $\mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_k, \dots, \mathcal{N}_n$ are independent standard normal variables, then

$$\frac{\mathcal{N}_1^2 + \dots + \mathcal{N}_k^2}{\mathcal{N}_1^2 + \dots + \mathcal{N}_n^2} \stackrel{\text{law}}{=} \beta_{\frac{k}{2}, \frac{n-k}{2}}.$$

Consequently, with the notation of Corollary 2.10,

$$(a_k^2, b_k^2 + c_k^2 + d_k^2) \stackrel{\text{law}}{=} \left(\beta_{\frac{1}{2}, 2k-\frac{1}{2}}, \left(1 - \beta_{\frac{1}{2}, 2k-\frac{1}{2}}\right) \beta'_{\frac{3}{2}, 2k-2} \right),$$

with β and β' independent beta variables with the specified parameters. This gives the somehow more tractable identity in law

$$\det(\text{Id} - G) \stackrel{\text{law}}{=} \prod_{k=1}^n \left(\left(1 + \epsilon_k \sqrt{\beta_{\frac{1}{2}, 2k-\frac{1}{2}}}\right)^2 + \left(1 - \beta_{\frac{1}{2}, 2k-\frac{1}{2}}\right) \beta'_{\frac{3}{2}, 2k-2} \right),$$

with all variables independent, $\mathbb{P}(\epsilon_k = 1) = \mathbb{P}(\epsilon_k = -1) = 1/2$.

Moreover, note that our method can be applied to other interesting groups such as $USp(2n, \mathbb{R}) = \{U \in U(2n, \mathbb{R}) \mid UJU^\top = J\}$ thanks to the morphism

$$\phi : \begin{cases} \mathbb{C} & \rightarrow & M_2(\mathbb{R}) \\ a + ib & \mapsto & \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \end{cases}.$$

The traditional representation of the quaternions in $M(4, \mathbb{R})$

$$\phi : \begin{cases} \mathbb{C} & \rightarrow & M_4(\mathbb{R}) \\ a + ib + jc + kd & \mapsto & \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & -c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \end{cases}$$

gives another identity in law for a compact exotic subgroup of $U(4n, \mathbb{R})$.

2.5. Symmetric groups. Take now \mathcal{S}_n the group of permutations of size n . An element $\sigma \in \mathcal{S}_n$ can be identified with the matrix $(\delta_{\sigma(i)}^j)_{1 \leq i, j \leq n}$ (δ is Kronecker's symbol).

Let H be a subgroup of $\{x \in \mathbb{H} \mid |x|^2 = 1\}$, with Haar probability measure μ_H , and H_n the group of diagonal matrices of size n with diagonal elements in H . Then the semidirect product $\mathcal{G} = H_n \cdot \mathcal{S}_n$ gives another example of determinant-splitting. More explicitly,

$$\mathcal{G} = \{(h_j \delta_{\sigma(i)}^j)_{1 \leq i, j \leq n} \mid \sigma \in \mathcal{S}_n, (h_1, \dots, h_n) \in H^n\}.$$

As the reflections now correspond to the transpositions, condition (R) holds. Moreover, with the notation of Theorem 2.6 $H_k(e_{k+1})$ is uniformly distributed on the unit sphere $\{he_{k+1} \mid h \in H\} \cup \dots \cup \{he_n \mid h \in H\}$. Therefore, following Theorem 2.6, we can state the following decomposition.

Corollary 2.11. *Let $G \in \mathcal{G} = H_n \cdot \mathcal{S}_n$ be $\mu_{\mathcal{G}}$ distributed. Then*

$$\det(\text{Id} - G) \stackrel{\text{law}}{=} \prod_{k=1}^n (1 - h_k X_k),$$

with $h_1, \dots, h_n, X_1, \dots, X_n$ independent random variables, the h_k 's μ_H distributed, $\mathbb{P}(X_k = 1) = 1/k$, $\mathbb{P}(X_k = 0) = 1 - 1/k$.

Remark. Let k_σ be the number of cycles of a random permutation of size n , with respect to the (probability) Haar measure. Corollary 2.11 allows us to recover the well-known law of k_σ :

$$k_\sigma \stackrel{\text{law}}{=} X_1 + \dots + X_n,$$

with the X_k 's Bernoulli variables as previously. Indeed, take for example $H = \{-1, 1\}$ in the Corollary. If a permutation $\sigma \in \mathcal{S}_n$ has k_σ cycles with lengths l_1, \dots, l_{k_σ} ($\sum_k l_k = n$), then it is easy to see that

$$\det(x\text{Id} - G) = \prod_{k=1}^{k_\sigma} (x^{l_k} - \eta_k)$$

with the η_k 's independent and uniform on $\{-1, 1\}$. Using this relation and the result of Corollary 2.11 we get

$$\prod_{k=1}^n (1 - \epsilon_k X_k) \stackrel{\text{law}}{=} \prod_{k=1}^{k_\sigma} (1 - \eta_k).$$

The equality of the Mellin transforms of the modulus of the above members easily implies the expected result : $k_\sigma \stackrel{\text{law}}{=} X_1 + \dots + X_n$.

In this section we have shown that for $G \in \mathcal{G}$, a general compact group endowed with its Haar measure, $\det(\text{Id} - G)$ can be decomposed as a product of independent random variables. This can be generalized to some *h-sampling* of the Haar measure. This will lead us to a generalization of the Ewens sampling formula, well known for the symmetric group.

3. HUA-PICKRELL MEASURES ON GENERAL COMPACT GROUPS

As previously, a permutation $\sigma \in \mathcal{S}_n$ is identified with the matrix $(\delta_{\sigma(i)}^j)_{1 \leq i, j \leq n}$. The Haar measure on \mathcal{S}_n can be generated by induction. Indeed, let τ_1, \dots, τ_n be independent transpositions respectively in $\mathcal{S}_1, \dots, \mathcal{S}_n$, with $\mathbb{P}(\tau_k(1) = j) = 1/k$ for any $1 \leq j \leq k$. Theorem 2.3 shows that if $\sigma \sim \mu_{\mathcal{S}_n}$ then

$$\sigma \stackrel{\text{law}}{=} \tau_n \begin{pmatrix} \text{Id}_1 & 0 \\ 0 & \tau_{n-1} \end{pmatrix} \cdots \begin{pmatrix} \text{Id}_{n-2} & 0 \\ 0 & \tau_2 \end{pmatrix} \begin{pmatrix} \text{Id}_{n-1} & 0 \\ 0 & \tau_1 \end{pmatrix}. \quad (3.1)$$

Read from right to left, the RHS of (3.1) corresponds to the so-called Chinese restaurant process, while from left to right this is the Feller decomposition of the symmetric group (see e.g. [2]).

What if the independent distributions of the $\tau_k(1)$'s are not uniform anymore ? Let $\theta > 0$. If for all $k \geq 1$

$$\mathbb{P}(\tau_k(1) = j) = \begin{cases} \frac{\theta}{\theta+k-1} & \text{if } j = 1 \\ \frac{1}{\theta+k-1} & \text{if } j \neq 1 \end{cases}, \quad (3.2)$$

then the distribution $\mu_{\mathcal{S}_n}^{(\theta)}$ of

$$\sigma := \tau_n \begin{pmatrix} \text{Id}_1 & 0 \\ 0 & \tau_{n-1} \end{pmatrix} \cdots \begin{pmatrix} \text{Id}_{n-2} & 0 \\ 0 & \tau_2 \end{pmatrix} \begin{pmatrix} \text{Id}_{n-1} & 0 \\ 0 & \tau_1 \end{pmatrix}.$$

can be expressed easily with respect to the Haar measure $\mu_{\mathcal{S}_n} = \mu_{\mathcal{S}_n}^{(1)}$: for a fixed $\Sigma \in \mathcal{S}_n$

$$\mathbb{P}_{\mu_{\mathcal{S}_n}^{(\theta)}}(\sigma = \Sigma) = \frac{\theta^{k_\Sigma}}{\mathbb{E}_{\mu_{\mathcal{S}_n}}(\theta^{k_\sigma})} \mathbb{P}_{\mu_{\mathcal{S}_n}}(\sigma = \Sigma),$$

with k_Σ the number of cycles of the permutation Σ . This is the Ewens sampling formula (for a direct proof, see e.g. [2]), which can also be formulated this way: for any function f from \mathcal{S}_n to \mathbb{R}

$$\mathbb{E}_{\mu_{\mathcal{S}_n}^{(\theta)}}(f(\sigma)) = \frac{\mathbb{E}_{\mu_{\mathcal{S}_n}}(f(\sigma)\theta^{k_\sigma})}{\mathbb{E}_{\mu_{\mathcal{S}_n}}(\theta^{k_\sigma})}, \quad (3.3)$$

which means that $\mu_{\mathcal{S}_n}^{(\theta)}$ is the θ^{k_σ} -sampling of $\mu_{\mathcal{S}_n}$. Our purpose here is to generalize the non-uniform measure (3.2) to any compact group, and to derive the corresponding equivalent to Ewens sampling formula (3.3).

3.1. Ewens sampling formula on general compact groups. We first give a general definition of the h -sampling.

Definition 3.1. Let (X, \mathcal{F}, μ) be a probability space, and $h : X \mapsto \mathbb{R}^+$ a measurable function with $\mathbb{E}_\mu(h(x)) > 0$. Then a measure μ' is said to be the h -sampling of μ if for all bounded measurable functions f

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

As usual, in the following, \mathcal{G} is any subgroup of a unitary group $U(n, K)$. Take $\delta \in \mathbb{C}$ such that

$$0 < \mathbb{E}_{\mu_{\mathcal{G}}} \left(\det(\text{Id} - G)^{\bar{\delta}} \det(\text{Id} - \bar{G})^{\delta} \right) < \infty. \quad (3.4)$$

For $0 \leq k \leq n - 1$ we note

$$\exp_{\delta}^{(k)} : \begin{cases} \mathcal{G} & \rightarrow \mathbb{R}^+ \\ G & \mapsto (1 - \langle G(e_{k+1}), e_{k+1} \rangle)^{\bar{\delta}} (1 - \langle \bar{G}(e_{k+1}), e_{k+1} \rangle)^{\delta} \end{cases}.$$

Moreover, define \det_{δ} as the function

$$\det_{\delta} : \begin{cases} \mathcal{G} & \rightarrow \mathbb{R}^+ \\ G & \mapsto \det(\text{Id} - G)^{\bar{\delta}} \det(\text{Id} - \bar{G})^{\delta} \end{cases}.$$

Then the following generalization of Theorem 2.3 (which corresponds to the case $\delta = 0$) holds. However, note that, contrary to Theorem 2.3, in the following result we need that the coherent measures ν_0, \dots, ν_{n-1} be supported by the set of reflections.

Theorem 3.2. Generalized Ewens sampling formula. *Let \mathcal{G} be a subgroup of $U(n, K)$ checking condition (R) and (3.4). Let $(\nu_0, \dots, \nu_{n-1})$ be a sequence of measures coherent with $\mu_{\mathcal{G}}$, with $\nu_k(\mathcal{R}_k) = 1$. We note $\mu_{\mathcal{G}}^{(\delta)}$ the \det_{δ} -sampling of $\mu_{\mathcal{G}}$ and $\nu_k^{(\delta)}$ the $\exp_{\delta}^{(k)}$ -sampling of ν_k . Then*

$$\nu_0^{(\delta)} \times \nu_1^{(\delta)} \times \dots \times \nu_n^{(\delta)} = \mu_{\mathcal{G}}^{(\delta)},$$

that is to say, for all bounded measurable functions f on \mathcal{G} ,

$$\mathbb{E}_{\nu_0^{(\delta)} \times \dots \times \nu_{n-1}^{(\delta)}} (f(R_0 R_1 \dots R_{n-1})) = \frac{\mathbb{E}_{\mu_{\mathcal{G}}} \left(f(G) \det(\text{Id} - G)^{\bar{\delta}} \det(\text{Id} - \bar{G})^{\delta} \right)}{\mathbb{E}_{\mu_{\mathcal{G}}} \left(\det(\text{Id} - G)^{\bar{\delta}} \det(\text{Id} - \bar{G})^{\delta} \right)}.$$

Proof. From Theorem 2.3,

$$\begin{aligned} & \frac{\mathbb{E}_{\mu_{\mathcal{G}}} \left(f(G) \det(\text{Id} - G)^{\bar{\delta}} \det(\text{Id} - \bar{G})^{\delta} \right)}{\mathbb{E}_{\mu_{\mathcal{G}}} \left(\det(\text{Id} - G)^{\bar{\delta}} \det(\text{Id} - \bar{G})^{\delta} \right)} \\ &= \frac{\mathbb{E}_{\nu_0 \times \dots \times \nu_{n-1}} \left(f(R_0 \dots R_{n-1}) \det(\text{Id} - R_0 \dots R_{n-1})^{\bar{\delta}} \det(\text{Id} - \overline{R_0 \dots R_{n-1}})^{\delta} \right)}{\mathbb{E}_{\nu_0 \times \dots \times \nu_{n-1}} \left(\det(\text{Id} - R_0 \dots R_{n-1})^{\bar{\delta}} \det(\text{Id} - \overline{R_0 \dots R_{n-1}})^{\delta} \right)}. \end{aligned}$$

As R_k is almost surely a reflection, we know from the proof of Theorem 2.6 that $\det(\text{Id} - R_0 \dots R_{n-1}) = \prod_{k=0}^{n-1} (1 - r_k)$ a.s. where $r_k = \langle R_k(e_{k+1}), e_{k+1} \rangle$. So thanks to the independence of the R_k 's

$$\begin{aligned} & \frac{\mathbb{E}_{\mu_{\mathcal{G}}} \left(f(G) \det(\text{Id} - G)^{\bar{\delta}} \det(\text{Id} - \bar{G})^{\delta} \right)}{\mathbb{E}_{\mu_{\mathcal{G}}} \left(\det(\text{Id} - G)^{\bar{\delta}} \det(\text{Id} - \bar{G})^{\delta} \right)} \\ &= \mathbb{E}_{\nu_0 \times \dots \times \nu_{n-1}} \left(f(R_0 \dots R_{n-1}) \prod_{k=0}^{n-1} \frac{(1 - r_k)^{\bar{\delta}} (1 - \bar{r}_k)^{\delta}}{\mathbb{E}_{\nu_k} \left((1 - r_k)^{\bar{\delta}} (1 - \bar{r}_k)^{\delta} \right)} \right). \end{aligned}$$

By the definition of the measures $\nu_k^{(\delta)}$, this is the expected result. \square

Remark. Of course a generalized Ewens sampling formula could also be given for $\Phi(\mathcal{G})$, with \mathcal{G} checking condition (R) and Φ the morphism from subsection 2.4. For simplicity it is stated in the restricted case when \mathcal{G} directly checks condition (R).

For $\mathcal{G} = U(n, \mathbb{C})$, Borodin and Olshanski [3] call $\mu_{\mathcal{G}}^{(\delta)}$ a Hua-Pickrell measure (see the introduction for an explanation). We keep this name for all groups checking condition (R).

Definition 3.3. The measures $\mu_{\mathcal{G}}^{(\delta)}$ are called the Hua-Pickrell measures on the group \mathcal{G} (which must satisfy the conditions of Theorem 3.2).

3.2. Examples. The general Theorem 3.2 gives a proof of formula (3.3), although $\det(\text{Id} - G) = 0$ in the case of the symmetric group. Indeed, we can consider the semidirect product $\mathcal{G} := \{-1, 1\} \cdot \mathcal{S}_n$, consisting of all matrices $(\epsilon_j \delta_{\sigma(j)}^i)$ with $\epsilon_j = \pm 1$, $\sigma \in \mathcal{S}_n$. The group \mathcal{G} checks all conditions of Theorem 3.2 for $\delta \in \mathbb{R}^+$. Moreover a sampling by the function $\exp_\delta^{(k)}$ corresponds to a sampling by a parameter $\theta := 2^{2\delta-1}$ in (2.11). Consequently (the first equality follows from Theorem 3.2),

$$\mathbb{E}_{\mu_{\mathcal{S}_n}^{(\theta)}}(f(G)) = \mathbb{E}_{\mu_{\mathcal{G}}^\delta} f(|G|) = \frac{\mathbb{E}_{\mu_{\mathcal{G}}} (f(|G|) |\det(\text{Id} - G)|^{2\delta})}{\mathbb{E}_{\mu_{\mathcal{G}}} (|\det(\text{Id} - G)|^{2\delta})}.$$

By conditioning on the permutation and integrating on the ϵ_j 's, we have

$$\begin{aligned} \mathbb{E}_{\mu_{\mathcal{G}}} \left(f(|G|) |\det(\text{Id} - G)|^{2\delta} \right) &= \mathbb{E}_{\mu_{\mathcal{G}}} \left(\mathbb{E} \left(f(|G|) |\det(\text{Id} - G)|^{2\delta} \mid \sigma \right) \right) \\ &= \mathbb{E}_{\mu_{\mathcal{G}}} \left(f(|G|) 2^{(2\delta-1)k_\sigma} \right) = \mathbb{E}_{\mu_{\mathcal{S}_n}} \left(f(G) 2^{(2\delta-1)k_\sigma} \right). \end{aligned}$$

We thus get the expected result :

$$\mathbb{E}_{\mu_{\mathcal{S}_n}^{(\theta)}}(f(G)) = \frac{\mathbb{E}_{\mu_{\mathcal{S}_n}} (f(G) \theta^{k_\sigma})}{\text{cst}}.$$

Another consequence of Theorem 3.2 consists in applying it to the unitary group for $\delta = 1$. Let f be a symmetric function of $(\theta_1, \dots, \theta_n)$, the eigenangles of a unitary matrix. The Weyl integration formula and Theorem 3.2 together give

$$\begin{aligned} &\mathbb{E}_{\nu_0^{(1)} \times \dots \times \nu_{n-1}^{(1)}} (f(\theta_1, \dots, \theta_n)) \\ &= \text{cst} \int_{(-\pi, \pi)^n} f(\theta_1, \dots, \theta_n) \prod_{j < k} |e^{i\theta_j} - e^{i\theta_k}|^2 \prod_{l=1}^n |1 - e^{i\theta_l}|^2 d\theta_1 \dots d\theta_n. \end{aligned}$$

This means that the distribution of $(\theta_1, \dots, \theta_n)$ under $\nu_0^{(1)} \times \dots \times \nu_{n-1}^{(1)}$ is the same as the distribution of the n first eigenangles of $G \sim \mu_{U(n+1, \mathbb{C})}$, conditionally to $\theta_{n+1} = 0$.

Of course, the same fact holds for the orthogonal group in the case $\delta = 1/2$.

4. DETERMINANTAL POINT PROCESSES FOR THE HUA-PICKRELL MEASURES

Let $(e^{i\theta_1}, \dots, e^{i\theta_n})$ be the eigenvalues of a generic element in $U(n, \mathbb{C})$. This section explains that Hua-Pickrell measures on $U(n, \mathbb{C})$ induce determinantal point processes for this spectrum. Of course, the following discussion could be held for Hua-Pickrell measures on orthogonal and symplectic groups.

Among many others, two motivations for the study of this determinantal form are the following :

- (a) Olshanski has shown that the Hua-Pickrell measures are closely linked to the natural analogue on the infinite unitary group of the biregular representations. Such representations can be described by the spectral measure of their characters. Theorem 4.6 gives information about this spectral measure. More details about these links between Hua-Pickrell measures and representations of infinite dimensional groups can be found in [13].
- (b) Many limit theorems about the Hua-Pickrell measure can be derived from Theorem 4.6. For instance, the number of eigenangles on any compact set of $(-\pi, \pi)$ satisfies a central limit theorem. Such results are easy applications of the general theory of determinantal point processes (see [16]).

4.1. Determinantal form for the spectral density. Consider a point process ζ on the unit circle C , with successive correlation functions $\rho_1, \rho_2 \dots$ (more precisions about correlation functions and determinantal point processes can be found in [7]).

Definition 4.1. If there exists a function $K : C \times C \rightarrow \mathbb{C}$ such that for all $k \geq 1$ and $(z_1, \dots, z_k) \in C^k$

$$\rho_k(z_1, \dots, z_k) = \det \left(K(z_i, z_j)_{i,j=1}^k \right)$$

then ζ is said to be a determinantal point process with correlation kernel K .

An example of determinantal point process related to the Hua-Pickrell measure is explained below. Let $H(n)$ be the set of $n \times n$ complex Hermitian matrices. Consider the Cayley transform

$$\begin{cases} H(n, \mathbb{C}) & \rightarrow U(n, \mathbb{C}) \\ X & \mapsto \frac{i-X}{i+X} \end{cases} .$$

Its reciprocal is defined almost everywhere and transforms the Hua-Pickrell measure $\mu_{U(n, \mathbb{C})}^{(\delta)}$ in a measure $\mu_{H(n, \mathbb{C})}^{(\delta)}$. A. Borodin and G. Olshanski ([3]) studied $\mu_{H(n, \mathbb{C})}^{(\delta)}$: they exhibit a determinantal form for the eigenvalues correlation functions, involving hypergeometric functions. Moreover, they suggest that such a form may exist for $\mu_{U(n, \mathbb{C})}^{(\delta)}$ itself. Theorem 4.4 gives this determinantal kernel.

Before stating this theorem, we need the following proposition, which exhibits the sequence of orthogonal polynomials on the unit circle for the measure

$$\begin{aligned} \lambda^{(\delta)}(\theta) d\theta &= c(\delta)(1 - e^{i\theta})^{\bar{\delta}}(1 - e^{-i\theta})^{\delta} d\theta \\ &= c(\delta)(2 - 2 \cos \theta)^a e^{-b(\pi \operatorname{sgn} \theta - \theta)} d\theta \end{aligned} \quad (4.1)$$

with

$$\delta = a + ib \quad , \quad c(\delta) = \frac{\Gamma(1 + \delta)\Gamma(1 + \bar{\delta})}{\Gamma(1 + \delta + \bar{\delta})}$$

$d\theta$ the Lebesgue measure on $(-\pi, \pi)$. Orthogonality and norm here are with respect to the Hermitian product

$$\langle \phi, \psi \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{\phi(e^{i\theta})} \psi(e^{i\theta}) \lambda^{(\delta)}(\theta) d\theta. \quad (4.2)$$

Proposition 4.2. *Let $\delta \in \mathbb{C}$ with $\Re(\delta) > -1/2$. Then $\lambda^{(\delta)}(\theta) \frac{d\theta}{2\pi}$ is a probability measure with successive monic orthogonal polynomials*

$$P_n(X) = X^n {}_2F_1(\delta, -n; -n - \bar{\delta}; X^{-1}), \quad n \geq 0.$$

Moreover,

$$\|P_n\|^2 = \frac{(\bar{\delta} + \delta + 1)_n n!}{(\bar{\delta} + 1)_n (\delta + 1)_n}.$$

Remark. In this proposition and in the following, $(x)_n$ stands for the Pochhammer symbol : if $n \geq 0$, $(x)_n = x(x+1)\dots(x+n-1)$, and if $n \leq 0$ $(x)_n = 1/(x+n)_{-n}$.

The hypergeometric function ${}_2F_1(\delta, -n; -n - \bar{\delta}; X^{-1})$ is strictly speaking not well defined : ${}_2F_1(a, b; c; z)$ is generally not convergent for $|z| = 1$ if $\Re(c - a - b) < 0$. However, in our case, as $-b \in \mathbb{N}$, the hypergeometric series contains a finite number of terms and therefore converges. Actually

$$P_n(z) = \sum_{k=0}^n \binom{n}{k} \frac{\Gamma(n-k+\delta)\Gamma(k+1+\bar{\delta})}{\Gamma(\delta)\Gamma(n+1+\bar{\delta})} z^k.$$

Proof. Let $n \geq 0$. We first calculate the moment of order n

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\theta} \lambda^{(\delta)}(\theta) d\theta = \frac{c(\delta)}{2\pi} \int_{-\pi}^{\pi} e^{-in\theta} (1 - e^{i\theta})^{\bar{\delta}} (1 - e^{-i\theta})^{\delta} d\theta.$$

As the Taylor series

$$(1 - e^{i\theta})^{\bar{\delta}} (1 - e^{-i\theta})^{\delta} = \left(\sum_{k \geq 0} \frac{(-\bar{\delta})_k}{k!} e^{ik\theta} \right) \left(\sum_{l \geq 0} \frac{(-\delta)_l}{l!} e^{-il\theta} \right)$$

agrees with formula (4.1), after an expansion of the sum all terms with $n+l \neq k$ cancel (the exchange of order between integral and sum requires some attention) and we get

$$c_n = c(\delta) \sum_{l \geq 0} \frac{(-\delta)_l}{l!} \frac{(-\bar{\delta})_{l+n}}{(l+n)!} = c(\delta) \frac{(-\bar{\delta})_n}{n!} {}_2F_1(-\delta, -\bar{\delta} + n; n+1; 1).$$

Combining our choice for $c(\delta)$ and Gauss formula (see [1])

$${}_2F_1(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \quad \Re(c-a-b) > 0,$$

leads to

$$c_n = \frac{(-\bar{\delta})_n}{(\delta+1)_n}. \quad (4.3)$$

The same method shows that formula (4.3) stands also for $n \leq 0$ ($(x)_n$ is defined in the preceding remark for a negative n). Note that $c_0 = 1$ so $\lambda^{(\delta)}(\theta) \frac{d\theta}{2\pi}$ is a probability measure.

The polynomials P_n ($n \geq 0$) are clearly monic, and they are orthogonal if and only if

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} P_n(e^{i\theta}) e^{-il\theta} \lambda^{(\delta)}(\theta) d\theta = 0, \quad 0 \leq l \leq n-1,$$

for all $n \geq 1$. Note that

$$\begin{aligned} & \frac{1}{2\pi} \int_{-\pi}^{\pi} P_n(e^{i\theta}) e^{-il\theta} \lambda^{(\delta)}(\theta) d\theta \\ &= \sum_{k=0}^n \frac{(\delta)_k (-n)_k}{(-n-\bar{\delta})_k k!} c_{l+k-n} = \sum_{k=0}^n \frac{(\delta)_k (-n)_k}{(-n-\bar{\delta})_k k!} \frac{(-\bar{\delta})_{l+k-n}}{(\delta+1)_{l+k-n}} \\ &= \frac{(-\bar{\delta})_{l-n}}{(\delta+1)_{l-n}} {}_3F_2(\delta, -\bar{\delta}+l-n, -n; 1+\delta+l-n, -n-\bar{\delta}; 1). \end{aligned}$$

The Pfaff-Saalschutz identity (see [1]) states that if $-c \in \mathbb{N}$ and $d+e = a+b+c+1$ then

$${}_3F_2(a, b, c; d, e; 1) = \frac{(d-a)_{-c} (d-b)_{-c}}{(d)_{-c} (d-a-b)_{-c}}.$$

Consequently, as $l < n$, $\langle P_n, X^l \rangle = 0$ and so the P_n 's are orthogonal. Moreover, as they are monic,

$$\|P_n\|^2 = \langle P_n, X^n \rangle = {}_3F_2(\delta, -\bar{\delta}, -n; -n-\bar{\delta}, 1+\delta; 1),$$

which is $\frac{(\bar{\delta}+\delta+1)_n n!}{(\bar{\delta}+1)_n (\delta+1)_n}$ once again thanks to the Pfaff-Saalschutz identity. \square

In fact, the second claim of Proposition 4.2 gives a new proof of the Keating-Snaith formula for the average of the characteristic polynomial $Z_n := \det(\text{Id} - U)$ of a random unitary matrix, originally derived with Selberg's integrals :

Corollary 4.3. *Let $Z_n := \det(\text{Id} - U)$, with $U \sim \mu_{U(n, \mathbb{C})}$; then:*

$$\mathbb{E}_{\mu_{U(n, \mathbb{C})}} (|Z_n|^t e^{s \arg Z_n}) = \prod_{k=1}^n \frac{\Gamma(k) \Gamma(k+t)}{\Gamma(k + \frac{t+is}{2}) \Gamma(k + \frac{t-is}{2})}. \quad (4.4)$$

Proof. An application of Heine's formula yields:

$$\mathbb{E}_{\mu_{U(n, \mathbb{C})}} (|Z_n|^t e^{s \arg Z_n}) = \frac{1}{c(\delta)^n} \det \left(\langle X^k, X^l \rangle_{k,l=0}^{n-1} \right),$$

where $t > -1$, $s \in \mathbb{R}$ and where the Hermitian product is defined in (4.2) with $\delta = \frac{t+is}{2}$. It is well known that this determinant of a Gram matrix is the square of the volume of the parallelepiped determined by the vectors

$1, X, \dots, X^{n-1}$. The "base times height" formula implies that this volume is the product of the norms of the successive monic orthogonal polynomials:

$$\mathbb{E}_{\mu_{U(n, \mathbb{C})}} (|Z_n|^t e^{s \arg Z_n}) = \frac{1}{c(\delta)^n} \prod_{k=0}^{n-1} \|P_n\|^2 = \prod_{k=1}^n \frac{\Gamma(k) \Gamma(k+t)}{\Gamma(k + \frac{t+is}{2}) \Gamma(k + \frac{t-is}{2})}.$$

This agrees with (4.4). \square

Note that [4] contains still another proof of (4.4), relying on the decomposition (2.7).

Theorem 4.4. *Let $\delta \in \mathbb{C}$, $\Re(\delta) > -\frac{1}{2}$. For $U \sim \mu_{U(n, \mathbb{C})}^{(\delta)}$, consider $\zeta = \{e^{i\theta_1}, \dots, e^{i\theta_n}\}$ the eigenvalues of U . Then ζ is a determinantal point process with correlation kernel*

$$K_{\delta}^{(n)}(e^{i\alpha}, e^{i\beta}) = d_n(\delta) \sqrt{\lambda^{(\delta)}(\alpha) \lambda^{(\delta)}(\beta)} \frac{e^{i\frac{n(\alpha-\beta)}{2}} Q_n^{(\delta)}(e^{-i\alpha}) Q_n^{(\bar{\delta})}(e^{i\beta}) - e^{-i\frac{n(\alpha-\beta)}{2}} Q_n^{(\bar{\delta})}(e^{i\alpha}) Q_n^{(\delta)}(e^{-i\beta})}{e^{i\frac{\alpha-\beta}{2}} - e^{-i\frac{\alpha-\beta}{2}}}.$$

Here $d_n(\delta) = \frac{1}{2\pi} \frac{(\bar{\delta}+1)_n (\delta+1)_n}{(\bar{\delta}+\delta+1)_n n!}$, $Q_n^{(\delta)}(x) = {}_2F_1(\delta, -n; -n - \bar{\delta}; x)$ and $\lambda^{(\delta)}(\alpha)$ is defined by (4.1).

Remark. Substituting $\delta = 0$ leads to the kernel

$$K_0^{(n)}(e^{i\alpha}, e^{i\beta}) = \frac{\sin \frac{n(\alpha-\beta)}{2}}{\sin \frac{\alpha-\beta}{2}}. \quad (4.5)$$

Proof. This is straightforward once we know the orthogonal polynomials from Proposition 4.2 : the following arguments are standard in Random Matrix Theory, and more details can be found in [11]. Let $f(e^{i\theta})d\theta$ be a probability measure on $(-\pi, \pi)$. Consider the probability distribution

$$F(e^{i\theta_1}, \dots, e^{i\theta_n}) d\theta_1 \dots d\theta_n = c(n, f) \prod_j f(e^{i\theta_j}) \prod_{k < l} |e^{i\theta_l} - e^{i\theta_k}|^2 d\theta_1 \dots d\theta_n$$

on $(-\pi, \pi)^n$, with $c(n, f)$ the normalization constant. Let P_k ($0 \leq k \leq n-1$) be monic polynomials with degree k . Thanks to Vandermonde's formula and multilinearity of the determinant

$$\prod_j \sqrt{f(e^{i\theta_j})} \prod_{k < l} (e^{i\theta_l} - e^{i\theta_k}) = \sqrt{\prod_{k=0}^{n-1} \|P_k\|_{L^2(f)}^2} \det \left(\sqrt{f(e^{i\theta_j})} \frac{P_k(e^{i\theta_j})}{\|P_k\|_{L^2(f)}} \right)_{k,j=1}^n.$$

Multiplying this identity with its conjugate and using $\det(AB) = \det A \det B$ gives

$$F(e^{i\theta_1}, \dots, e^{i\theta_n}) = \det \left(K^{(n)}(e^{i\theta_j}, e^{i\theta_k})_{j,k=1}^n \right)$$

with $K^{(n)}(x, y) = c\sqrt{f(x)f(y)}\sum_{k=0}^{n-1}\frac{P_k(x)\overline{P_k(y)}}{\|P_k\|_{L^2(f)}^2}$, the constant c depending on f , n and the P_i 's. This shows that the correlation ρ_n has the desired determinantal form. Gaudin's lemma (see [11]) implies that if the polynomials P_k 's are orthogonal in $L^2(f)$, then

$$\rho_l(e^{i\theta_1}, \dots, e^{i\theta_l}) = \det \left(K^{(n)}(e^{i\theta_j}, e^{i\theta_k})_{j,k=1}^l \right)$$

for all $1 \leq l \leq n$. As $\int_{-\pi}^{\pi} \rho_1(e^{i\theta})d\theta = n$, then $c = 1$, so the Christoffel-Darboux formula gives

$$\begin{aligned} K^{(n)}(x, y) &= \sqrt{f(x)f(y)} \sum_{k=0}^{n-1} \frac{P_k(x)\overline{P_k(y)}}{\|P_k\|_{L^2(f)}^2} \\ &= \frac{\sqrt{f(x)f(y)}}{\|P_n\|_{L^2(f)}^2} \frac{P_n^*(x)\overline{P_n^*(y)} - P_n(x)\overline{P_n(y)}}{x - y}. \end{aligned}$$

where $P_n^*(x) = x^n \overline{P_n(1/\bar{x})}$.

Concerning the Hua-Pickrell measure, taking in the above discussion $\lambda^{(\delta)}$ for f , replacing the kernel $K^{(n)}(e^{i\alpha}, e^{i\beta})$ by $e^{i\frac{n\alpha}{2}} K^{(n)}(e^{i\alpha}, e^{i\beta}) e^{-i\frac{n\beta}{2}}$ (this doesn't change the determinant), we get directly the result of Theorem 4.4. \square

4.2. Implied asymptotics. The asymptotic of the kernel (4.5) is given by

$$K_0^{(\infty)}(e^{i\alpha}, e^{i\beta}) = \lim_{n \rightarrow \infty} \frac{1}{n} K_0^{(n)}(e^{i\frac{\alpha}{n}}, e^{i\frac{\beta}{n}}) = \frac{\sin\left(\frac{\alpha-\beta}{2}\right)}{\frac{\alpha-\beta}{2}}.$$

A similar limit holds for the Hua-Pickrell determinantal kernel, given in Theorem 4.5. To this end, we shall need the following asymptotics:

Proposition 4.5. *Let $\delta \in \mathbb{C}$. Then*

$$\lim_{n \rightarrow \infty} n^{-\bar{\delta}} {}_2F_1(\bar{\delta}, -n; -n - \delta; e^{i\frac{\theta}{n}}) = \frac{\Gamma(\delta + 1)}{\Gamma(\delta + \bar{\delta} + 1)} {}_1F_1(\bar{\delta}, \delta + \bar{\delta} + 1, i\theta),$$

uniformly on $\{\theta \in \mathcal{K}\}$, with \mathcal{K} any compact set of \mathbb{R} .

Proof. The function $g^{(n)} : \theta \mapsto n^{-\bar{\delta}} {}_2F_1(\bar{\delta}, -n; -n - \delta; e^{i\frac{\theta}{n}})$ satisfies the ordinary differential equation

$$\begin{aligned} &\left[n(1 - e^{-i\frac{\theta}{n}}) \right] \partial_{\theta\theta} g^{(n)} \\ &+ \left[-i(1 - e^{-i\frac{\theta}{n}}) + ie^{-i\frac{\theta}{n}}(n + \delta) - i(n - \bar{\delta} - 1) \right] \partial_{\theta} g^{(n)} + \bar{\delta} g^{(n)} = 0 \quad (4.6) \end{aligned}$$

with initial conditions (here we use the Chu-Vandermonde identity)

$$\begin{cases} g^{(n)}(0) &= \frac{(-n - \delta - \bar{\delta})_n}{n^{\bar{\delta}}(-n - \delta)_n} &= \frac{(\delta + \bar{\delta} + 1)_n}{n^{\bar{\delta}}(\delta + 1)_n} \\ g^{(n)'}(0) &= \frac{i\bar{\delta}(-n - \delta - \bar{\delta})_{n-1}}{n^{\bar{\delta}}(n + \delta)(-n - \delta + 1)_{n-1}} &= \frac{i\bar{\delta}(\delta + \bar{\delta} + 2)_{n-1}}{n^{\bar{\delta}}(n + \delta)(\delta + 1)_{n-1}} \end{cases}$$

Taking $n \rightarrow \infty$ in (4.6) and using classical theory of differential equations, we can conclude that $g^{(n)}$ converges uniformly on any compact set to the solution g of the differential equation

$$(i\theta)\partial_{\theta\theta}g + (i(\delta + \bar{\delta} + 1) + \theta)\partial_{\theta}g + \bar{\delta}g = 0 \quad (4.7)$$

with initial values $g(0) = \lim_n g^{(n)}(0)$ and $g'(0) = \lim_n g^{(n)'}(0)$ i.e.

$$g(0) = \frac{\Gamma(\delta + 1)}{\Gamma(\delta + \bar{\delta} + 1)}, \quad g'(0) = \frac{i\bar{\delta}\Gamma(\delta + 1)}{\Gamma(\delta + \bar{\delta} + 2)}.$$

The unique solution of (4.7) is

$$g(\theta) = \frac{\Gamma(\delta + 1)}{\Gamma(\delta + \bar{\delta} + 1)} {}_1F_1(\bar{\delta}, \delta + \bar{\delta} + 1, i\theta).$$

□

Consequently, using this proposition and Theorem 4.4 the re-scaled correlation function associated to the Hua-Pickrell measure can be written in a determinantal form.

Theorem 4.6. *Let $\delta \in \mathbb{C}$, $\Re(\delta) > -\frac{1}{2}$. For $U \sim \mu_{U(n, \mathbb{C})}^{(\delta)}$, consider $\zeta = \{e^{i\theta_1}, \dots, e^{i\theta_n}\}$ the eigenvalues of U and write $\rho_k^{(n)}(e^{i\theta_1}, \dots, e^{i\theta_k})$ ($k \geq 0$) for the associate correlation functions of the eigenvalues. Then, uniformly on any compact set of \mathbb{R}^k ,*

$$\frac{1}{n^k} \rho_k^{(n)}(e^{i\frac{\theta_1}{n}}, \dots, e^{i\frac{\theta_k}{n}}) \xrightarrow{n \rightarrow \infty} \det \left(K_{\delta}^{(\infty)}(e^{i\theta_i}, e^{i\theta_j})_{i,j=1}^k \right).$$

with

$$\begin{aligned} & K_{\delta}^{(\infty)}(e^{i\alpha}, e^{i\beta}) \\ &= e(\delta) |\alpha\beta|^{\Re\delta} e^{-\frac{\pi}{2}(\Im\delta)(\operatorname{sgn}\alpha + \operatorname{sgn}\beta)} \frac{e^{i\frac{\alpha-\beta}{2}} Q^{(\delta)}(-i\alpha) Q^{(\bar{\delta})}(i\beta) - e^{-i\frac{\alpha-\beta}{2}} Q^{(\bar{\delta})}(i\alpha) Q^{(\delta)}(-i\beta)}{\alpha - \beta}. \end{aligned}$$

Here $e(\delta) = \frac{1}{2i\pi} \frac{\Gamma(\delta+1)\Gamma(\bar{\delta}+1)}{\Gamma(\delta+\bar{\delta}+1)^2}$ and $Q^{(\delta)}(x) = {}_1F_1(\delta, \delta + \bar{\delta} + 1; x)$.

Remark. As expected, the kernel $K_0^{(\infty)}$ coincides with the sine kernel for $\delta = 0$:

$$K_0(e^{i\alpha}, e^{i\beta}) = \frac{\sin\left(\frac{\alpha-\beta}{2}\right)}{\frac{\alpha-\beta}{2}}.$$

For $\Re(\delta) > -1/2$, ${}_1F_1(\delta, \delta + \bar{\delta} + 1, x)$ (and then $K_{\delta}^{(\infty)}$) can be expressed in terms of Whittaker functions, and for $\delta \in \mathbb{R}$ as Bessel functions (see [1]).

Remark. In [3] the determinantal kernel on the real line associated to Hua-Pickrell measures on the Hermitian matrices $H(n, \mathbb{C})$ is given. Its asymptotics, after the scaling $x \mapsto nx$ on the eigenvalues, is noted $K_{\delta, H}^{(\infty)}(x, y)$. With no surprise, the expression given by Borodin and Olshanski coincides with ours after a suitable change of variables :

$$K_{\delta, H}^{(\infty)}(x, y) = f(\delta) K_{\delta, U}^{(\infty)}(e^{\frac{2i}{x}}, e^{\frac{2i}{y}}) \quad (4.8)$$

for a constant $f(\delta)$. This was observed by Borodin and Olshanski thanks to the following argument, linking the unitary and Hermitian ensembles via the Cayley transform

$$\begin{cases} H(n, \mathbb{C}) & \rightarrow U(n, \mathbb{C}) \\ X & \mapsto \frac{i-X}{i+X} \end{cases} .$$

Indeed, this function transforms the Hua-Pickrell measure on $H(n, \mathbb{C})$ into the Hua-Pickrell measure on $U(n, \mathbb{C})$. Therefore, a scaling $x \mapsto nx$ of the eigenvalues on $H(n, \mathbb{C})$ corresponds to a scaling $\alpha \mapsto \frac{\alpha}{n}$ for the eigenangles on $U(n, \mathbb{C})$:

$$\frac{i - nx}{i + nx} = -e^{-\frac{2i}{nx}} + O\left(\frac{1}{n^2}\right), \quad (4.9)$$

leading to the correspondence $\alpha = \frac{2}{x}$. Theorem 4.6 gives an alternative proof of (4.8), by direct calculation. Note that for fixed n we do not have an identity like (4.8), because of the error term in (4.9).

Acknowledgements. P. B. thanks his advisors A. S. Ustunel for suggesting the references [14] and [15] and M. Yor for many precious discussions and a very careful reading. The authors are grateful to P.-O. Dehaye for suggesting the proof of Proposition 4.5.

REFERENCES

- [1] G.E. Andrews, R.A. Askey, R. Roy, Special functions, Encyclopedia of Mathematics and its Applications, 71, Cambridge University Press, Cambridge, 1999.
- [2] R. Arratia , A.D. Barbour, S. Tavaré, Logarithmic Combinatorial Structures: A Probabilistic Approach. 352 pp, 2003. EMS Monographs in Mathematics, 1. European Mathematical Society Publishing House, Zürich.
- [3] A. Borodin, G. Olshanski, Infinite Random Matrices and Ergodic Measures, Comm. Math. Phys. 203 (2001), 87-123.
- [4] P. Bourgade, C.P. Hughes, A. Nikeghbali, M. Yor, The characteristic polynomial of a random unitary matrix: a probabilistic approach, to appear in Duke Mathematical Journal.
- [5] N. S. Witte and P. J. Forrester, Gap probabilities in the finite and scaled Cauchy random matrix ensembles, Nonl., 13 (2000), 1965-1986.
- [6] L. K. Hua, Harmonic analysis of functions of several complex variables in the classical domains, Chinese edition: Science Press, Peking, 1958; Russian edition: IL, Moscow, 1959; English edition: Transl. Math. Monographs 6, Amer. Math. Soc., 1963.
- [7] K. Johansson, Random matrices and determinantal processes, math-ph/0510038, 2005.

- [8] N. M. Katz and P. Sarnak, Random Matrices, Frobenius Eigenvalues and monodromy, American Mathematical Society Colloquium Publications, 45. American Mathematical Society, Providence, Rhode island, 1999.
- [9] N. M. Katz and P. Sarnak, Zeros of zeta functions and symmetry, Bull. Amer. Soc. 36, 1-26, 1999. Colloquium Publications, 45. American Mathematical Society, Providence, Rhode island, 1999.
- [10] J.P. Keating and N.C. Snaith, Random Matrix Theory and $\zeta(1/2+it)$, Comm. Math. Phys. 214, p 57-89, 2000.
- [11] M. L. Mehta, Random matrices, Pure and Applied Mathematics Series 142, Elsevier (London - 2004).
- [12] Yu. A. Neretin, Hua type integrals over unitary groups and over projective limits of unitary groups, Duke Math. J. 114 (2002), 239-266.
- [13] G. Olshanski, Journal of Functional Analysis, Volume 205, Number 2, 20 December 2003 , pp. 464-524(61).
- [14] D. Pickrell, Measures on infinite-dimensional Grassmann manifolds, J. Func. Anal. 70 (1987), no. 2, 323-356.
- [15] D. Pickrell, Mackey analysis of infinite classical motion groups, Pacific J. Math. 150 (1991), 139-166.
- [16] A Soshnikov, Determinantal random point fields, Russ. Math. Surv., 2000, 55 (5), 923-975.

ENST, 46 RUE BARRAULT, 75634 PARIS CEDEX 13. UNIVERSITÉ PARIS 6, LPMA,
175, RUE DU CHEVALERET F-75013 PARIS.

E-mail address: `bourgade@enst.fr`

INSTITUT FÜR MATHEMATIK, UNIVERSITÄT ZÜRICH, WINTERTHURERSTRASSE 190,
CH-8057 ZÜRICH, SWITZERLAND

E-mail address: `ashkan.nikeghbali@math.unizh.ch`

UNIVERSITÉ VERSAILLES-SAINT QUENTIN, LMV, BÂTIMENT FERMAT, 45 AVENUE DES
ÉTATS-UNIS, 78035 VERSAILLES CEDEX

E-mail address: `rouault@fermat.math.uvsq.fr`